

A Stage-Aware Mixture of Experts Framework for Neurodegenerative Disease Progression Modelling

Tiantian He^{1 2*}, Keyue Jiang^{4*}, An Zhao¹, Anna Schroder¹, Elinor Thompson¹,
Sonja Soskic², Frederik Barkhof^{2 3 5}, Daniel C. Alexander¹

¹ UCL Hawkes Institute and Department of Computer Science, University College London

² UCL Hawkes Institute and Department of Medical Physics and Biomedical Engineering, University College London

³ Queen Square Institute of Neurology, University College London

⁴ Department of Electronic and Electrical Engineering & AI Centre, University College London

⁵ Department of Radiology, Amsterdam University Medical Center, Vrije Universiteit
tiantian.he.20@ucl.ac.uk, keyue.jiang.18@ucl.ac.uk

Abstract

The long-term progression of pathology in neurodegenerative diseases is commonly conceptualized as a spatiotemporal diffusion process that consists of a graph diffusion process across the structural brain connectome and a localized reaction process within brain regions. However, modeling this progression remains challenging due to 1) the scarcity of longitudinal data obtained through irregular and infrequent subject visits and 2) the complex interplay of pathological mechanisms across brain regions and disease stages, where traditional models assume fixed mechanisms throughout disease progression. To address these limitations, we propose a novel stage-aware Mixture of Experts (MoE) framework that explicitly models how different contributing mechanisms dominate at different disease stages through time-dependent expert weighting. This architecture is a key innovation designed to maximize the utility of small datasets and provide interpretable insights into disease etiology. Data-wise, we utilize an iterative dual optimization method to properly estimate the temporal position of individual observations, constructing a cohort-level progression trajectory from irregular snapshots. Model-wise, we enhance the spatial component with an inhomogeneous graph neural diffusion model (IGND) that allows diffusivity to vary based on node states and time, providing more flexible representations of brain networks. We also introduce a localized neural reaction module to capture complex dynamics beyond standard processes. The resulting IGND-MoE model dynamically integrates these components across temporal states, offering a principled way to understand how stage-specific pathological mechanisms contribute to progression. When used to model tau pathology propagation in human brains, IGND-MoE outperforms purely pathophysiological and purely neural baselines in long-term prediction accuracy. Moreover, its stage-wise weights yield novel clinical insights that align with literature, suggesting that graph-related processes are more influential at early stages, while other unknown physical processes become dominant later on. Our findings highlight the necessity of designing hybrid and expert-constrained models that account for the evolving nature of neurodegenerative processes.

*These authors contributed equally.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Code — <https://github.com/Tiantian-H/stage-aware-MoE>

Extended version — <https://arxiv.org/abs/2508.07032>

Introduction

Neurodegenerative diseases exhibit a progressive propagation of pathology through the brain (Busche and Hyman 2020). Understanding their long-term progression from the early to advanced stages is a key challenge for developing disease-modifying treatments. However, constraints of real-world patient data acquisition often hamper such efforts. Since medical scans can be expensive or pose potential health risks, data is often collected irregularly and over a narrow time frame. Accordingly, a set of modern computational approaches, known as data-driven disease progression models (Fonteyn et al. 2012; Young et al. 2014), has emerged to address the challenge of estimating population-level trajectories of change from such sparse and irregularly sampled patient data sets.

Pathophysiological disease progression models (Zhou et al. 2012a; Raj, Kuceyeski, and Weiner 2012b; Seguin, Sporns, and Zalesky 2023; Garbarino et al. 2019; Young et al. 2014) simulate the spread of pathology over the brain using hypothetical mechanisms. These models capture spatiotemporal dynamics through two components: i) a graph that approximates the ability of each region’s pathology occupancy to cause pathology appearance in each other region and ii) a mechanism of propagation between regions given that set of graph links. Network diffusion models (NDMs) (Raj, Kuceyeski, and Weiner 2012b; Weickenmeier, Kuhl, and Goriely 2018a) assume pathology spreads by diffusing along structural brain connections from MRI. Current approaches use brain connectivity measures as proxies for graph link strength and offer interpretability. However, they face three limitations: 1) **Homogeneous graph diffusion** assumes uniform diffusion rates and fixed graphs, oversimplifying brain networks and failing to model complex neurodegeneration that evolves over time (He et al. 2023; Zhou et al. 2012b; Avena-Koenigsberger, Misic, and Sporns 2018); 2) **Constrained localized propagation** inadequately captures alternative processes beyond reaction-type

mechanisms, such as clearance and interactions (Garbarino, Lorenzi, and Initiative 2019; Abi Nader et al. 2022); and 3) The assumption of **fixed mechanisms throughout disease**, whereas contributing mechanisms likely shift during disease progression (Meisl et al. 2021b).

Deep learning-based time series models offer opportunities for enhancing disease progression models, leveraging their inherent flexibility and data-driven characteristics. However, existing approaches have distinct limitations when applied to disease progression. Typical models for time-series analysis, such as: Recurrent Neural Networks (RNNs) (Rumelhart et al. 1985), Long Short-Term Memory Networks (LSTMs) (Graves and Graves 2012), and Gated Recurrent Units (GRUs) (Chung et al. 2014), discretize time series and thus exhibit limitations in handling continuous dynamical systems, irregularly sampled data, and long-term dependencies. Neural ordinary differential equations (ODEs) (Chen et al. 2018) parameterise the derivative of the trajectory using black-box neural networks, enabling them to approximate complex dynamics with great flexibility. However, they cannot decompose learned patterns into mechanistic components that align with biological knowledge. Graph Neural ODEs (Poli et al. 2019) couple Neural ODEs with graph networks to model the complex dynamics evolving on the graph. However, they are rigidly bound to a predefined input graph structure that may not capture all relevant biological connections. Moreover, they cannot easily distinguish between graph-related and non-graph-related pathological processes. Thus, this type of model alone offers limited insights into how different mechanisms contribute at different disease stages or brain regions. The proper combination of them with the pathophysiological model will maximize each of their expertise, and more importantly, allow us to understand the complex mechanism of the disease progression pattern better.

To address the challenge of understanding how pathological mechanisms dynamically shift over the disease course, we develop a novel stage-aware Mixture of Experts (MoE) framework that combines an existing pathophysiological model with an inhomogeneous graph neural diffusion model (IGND), where the graph diffusivity varies based on node states and time rather than being constant. Unlike traditional approaches that assume fixed mechanisms throughout disease progression, our IGND-MoE explicitly models how different contributing mechanisms dominate at different disease stages through time-dependent expert weighting. For the spatial component, we introduce an IGND parameterized by a graph auto-encoder (GAE) to address the limitation of homogeneous graph diffusion models (Raj, Kuceyeski, and Weiner 2012b; Weickenmeier, Kuhl, and Goriely 2018a) that assume uniform diffusion rates with fixed connectivity graphs. Our inhomogeneous approach allows diffusivity to vary based on both node states and time, enabling more complex dynamics where pathology spread rates adapt to disease progression. This provides a more flexible representation of network dynamics than prior approaches such as standard Neural ODEs (Chen et al. 2018) and Graph Neural ODEs (Poli et al. 2019), which either lack graph structure or maintain fixed graph properties throughout the dis-

ease course. For the local component at each node, we enhance the traditional logistic growth with neural methods to increase expressiveness. By incorporating temporal attention that modulates each expert’s contribution across disease stages, our model reconstructs cohort-level, long-term spatiotemporal dynamics while revealing stage-specific pathological mechanisms. This comprehensive modelling strategy not only improves predictive accuracy but, more importantly, provides new perspectives on the evolving nature of neurodegenerative processes by distinguishing whether unexplained pathological changes are inherently graph-related, non-graph-related, or emerge from their interactions.

Our contributions are summarized as follows: (1) We propose a stage-aware Mixture of Experts (IGND-MoE) framework to construct a long-term continuous disease progression trajectory from irregular snapshots. It combines pathophysiological models with neural approaches to model how different pathological mechanisms dominate at different disease stages. (2) We introduce an inhomogeneous graph neural diffusion model for spatial propagation and enhanced localized dynamics with neural networks, both integrated through temporal attention that modulates expert contributions across disease stages. (3) Through extensive experiments, we demonstrate our model achieves superior prediction accuracy compared to pure pathophysiological or neural approaches, particularly for long-term prediction, while providing interpretable insights into stage-specific disease mechanisms.

Background: Network Diffusion Models

The propagation of deleterious protein in neurodegenerative disease is commonly formalized as a graph differential equation system based on the Fisher-Kolmogorov equation (Meisl et al. 2021a; Raj, Kuceyeski, and Weiner 2012a; Weickenmeier, Kuhl, and Goriely 2018b), where the graph is constructed based on brain connectivity. Existing literature decomposes the propagation into two principal mechanisms: (i) the spatial diffusion of toxic protein through the brain’s structural network (Raj, Kuceyeski, and Weiner 2012a) named network diffusion model (NDM); and (ii) the localized reaction (production and accumulation) of the protein (Meisl et al. 2021a). The spatial diffusion of the protein is quantified as a *homogeneous* heat diffusion process (Carslaw and Jaeger 1959; Thanou et al. 2017) modulated by the graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and rate k , where \mathbf{A} is the adjacency matrix and $D_{ii} = \sum_j A_{ij}$ is the degree matrix. The local production and aggregation process first exhibits a progressively increasing trend, ultimately stabilizing at a plateau value \mathbf{v} , with a uniformly increasing rate α applicable to all regions. With the above two-component evolution, the overall concentration of the protein \mathbf{c} is governed by the following ordinary differential equation:

$$\frac{d\mathbf{c}(t)}{dt} = \underbrace{-k[\mathbf{L} \cdot \mathbf{c}(t)]}_{\text{Spatial Diffusion (i)}} + \underbrace{\alpha\mathbf{c}(t) \odot [\mathbf{v} - \mathbf{c}(t)]}_{\text{Localized Reaction (ii)}} = f_M, \quad (1)$$

where \odot is the element-wise product. Once the derivative is properly estimated, the long-term prediction is obtained by

integrating the derivatives $\mathbf{c}(t) = \mathbf{c}(0) + \int_0^t (d\mathbf{c}(t)/dt)dt$. Compared to the discrete prediction models, the PDE-based models naturally factor in the continuous dynamics along the timeline, and are thus more suitable for long-term pathology progression modeling.

The limitation of pathophysiological models. We recognise three main drawbacks:

1. *The homogeneity of graph diffusion.* The model in Eq. 1 assumes that the spatial diffusion is *homogeneous* with uniform node and edge characteristics, which suggests that the efficiency at which toxic proteins travel from one region to another is considered constant throughout the entire network and timeline. This hypothesis oversimplifies the actual brain network and fails to model complex neurodegeneration in real-world scenarios (He et al. 2023; Zhou et al. 2012b; Avena-Koenigsberger, Misisic, and Sporns 2018).

2. *The constrained localized propagation.* The constrained physical models inadequately capture localized propagation, typically limited to reaction-type processes. They miss other reported mechanisms like protein clearance and interactions with other biomarkers (Garbarino, Lorenzi, and Initiative 2019; Abi Nader et al. 2022), necessitating a more expressive modeling approach for accurate localized propagation.

3. *The temporal dynamics of expert models.* The model in Eq. 1 only accounts for static contributions of each mechanism, however the contribution of each process to the overall toxic protein propagation varies over time. For instance, previous literature (Meisl et al. 2021b) suggests the local production takes the lead at later disease stages.

Methodology

Problem Formulation

For each subject i , a longitudinal sequence of samples is provided as $\{\mathbf{c}(t_s^i)\}_{s=1}^{S_i}$, where the samples are ordered chronologically for $S_i \geq 1$ distinct time points $t_s, s \in \{1, \dots, S_i\}$. Each sample $\mathbf{c}(t_s^i)$ constitutes an n -dimensional vector representing node states derived from n brain regions. These brain regions define a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ on node set \mathcal{V} representing the brain regions, and edge set \mathcal{E} the connectivity between the regions, where $|\mathcal{V}| = n$ and $|\mathcal{E}| = e$. The adjacency matrix of the graph is represented by $\mathbf{A} \in \mathbb{R}^{n \times n}$. These individual observations are then aligned to a common temporal axis representing the full spectrum of disease development. This alignment facilitates the construction of a cohort-level disease progression trajectory, specified by a sequence of node state vectors $\{\mathbf{c}(t)\}_{t=1}^T$ over T time points. Here, $\mathbf{c}(t) = [c(u, t)]_{u \in \mathcal{V}}^T \in \mathbb{R}^n$ is the vector where $c(u, t)$ denotes the estimated pathological quantity (e.g., protein concentration) in node u at time t on the global timeline. The core problem involves accurately projecting incoming individual samples onto this estimated disease timeline and predicting the subsequent evolution of the node states.

Overview of the Modelling Framework

In general, our method aims to tackle two challenges: 1) The construction of a comprehensive cohort-level disease progression trajectory from snapshots of individual-level obser-

vations (Sec. 3.3), and 2) The development of a mixture of experts system to model how different pathological mechanisms dominate at different disease stages (Sec. 3.4). Figure 1 displays an overview of our stage-aware Mixture of Experts framework (IGND-MoE).

At the start of model training, neither the trajectory shape nor the relative location of each individual on the trajectory is known. We therefore apply a dual optimization strategy that iteratively refines both the disease trajectory and individual subject placements along this trajectory. This optimization process consists of the following steps:

1. **Temporal Initialization Step:** The process begins with prior knowledge about the trajectory, simulated by the pathophysiological model described by equation 1. Given this prior trajectory, each subject i is allocated to the most appropriate location on the temporal axis through optimization of their pseudo time t^i .
2. **Trajectory Construction Step:** Given the estimated subject locations, the trajectory is refined by combining contributions from the three expert components: the pathophysiological model, the inhomogeneous graph neural diffusion, and the localized neural reaction. The temporal attention mechanism modulates each expert's contribution based on disease stage, while regularization ensures interpretability and complementarity between experts. This produces a more flexible trajectory that captures stage-specific pathological mechanisms.
3. **Temporal Alignment Step:** Given the updated stage-aware trajectory, the subject locations along the temporal axis are further refined.
4. **Iterative Refinement:** Steps 2 and 3 are repeated, starting from the previously trained model, until convergence.

Temporal Alignment Step

As displayed by Figure 2, disease progression models aim to construct long-term cohort level trajectories from snapshots of individual data by estimating the location t_0^i of the baseline measurement of subject i on the mutual temporal axis (Young et al. 2024; Lorenzi et al. 2019) so that the measurements can align as closely as possible to the assumed trajectory, by minimizing their sum squared errors:

$$\mathcal{L}_{f_M}(t_0^i) = \sum_{i=1}^N \sum_{s=1}^{S_i} \|c_{obs}(t_0^i + t_s^i) - c(t_0^i + t_s^i)\|^2 \quad (2)$$

where f_M represents the trajectory from pathophysiological model, and t_s^i represents the time gap from the baseline to the s th scan for subject i , which is recorded in the database.

Cohort-level Trajectory Construction Step

Design of the stage-aware Mixture of Expert Model
Our model combines three key components as in Eq. 1: 1) A pathophysiological model that provides interpretable baseline dynamics based on clinical hypotheses; 2) An inhomogeneous graph neural diffusion (IGND) model that enhances spatial propagation by learning flexible, data-driven representations of brain networks; and 3) localized neural component that captures complex local dynamics beyond standard reaction mechanisms. These components are integrated

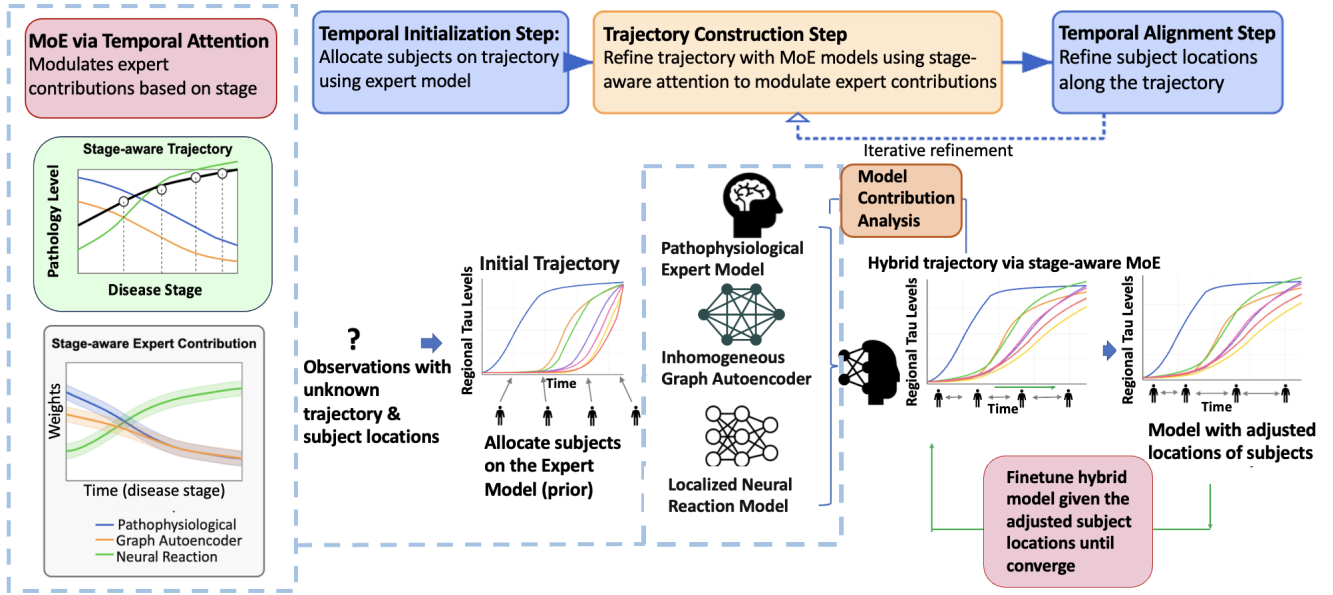


Figure 1: Stage-aware Neurodegenerative Disease Progression Modeling with IGND-MoE. This figure demonstrates the proposed IGND-MoE framework for constructing a full disease progression process from snapshots, by iteratively carrying out the temporal alignment step for mapping each subject to the proper location on the time axis and the trajectory construction step of shaping a better trajectory through the proposed temporal-aware mixture of expert structure.

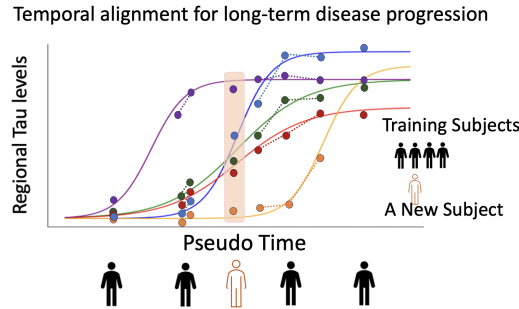


Figure 2: Temporal alignment for long-term progression. This figure visualizes how the framework uses the individual cross-sectional data or short-term longitudinal data to construct the full disease progression trajectory. Each colour represents one brain region. The dots represent real observations. The dots connected with dashed lines represent the longitudinal observations from the same subject, where the real-time gap between the scans is available in the dataset and thus remains. The curves represent the model fitting.

through temporal attention that dynamically modulates their contributions over time.

$$\frac{dc(t)}{dt} = \underbrace{\beta_1(t)f_M}_{\text{Mechanistic Model}} + \underbrace{\beta_2(t)f(\mathbf{A}, \mathbf{c}(t), t, \theta_S)}_{f_S: \text{Graph Neural Diffusion}} + \underbrace{\beta_3(t)f(\mathbf{c}(t), t, \theta_L)}_{f_L: \text{Localized Neural Reaction}} \quad (3)$$

where θ_S and θ_L are parameters for spatial diffusion and localized reaction. The mixture weights $\{\beta_j(t)\}_{j \in \{1,2,3\}}$, which satisfy $\sum_j \beta_j(t) = 1, \forall t$, dynamically modulate the contribution of each expert model throughout the disease stages.

Spatial Modelling with Inhomogeneous Graph Neural Diffusion. Traditional network diffusion models operate on homogeneous graphs, where diffusivity remains constant across all edges and throughout the entire disease progression process, corresponding to a graph diffusion process with uniform diffusion coefficients that do not change with time or node states.

We extend the homogeneous spatial diffusion induced by brain connectivity to an inhomogeneous graph neural diffusion model that accounts for changes in diffusivity over states and time. In contrast to the pathophysiological model in Sec. 2, it accounts for the impact of time and node states on spatial diffusion, thus enabling the modeling of more complex dynamic systems. To do so, we consider the PDE:

$$\frac{\partial c(u, t)}{\partial t} = \text{div}[g(\mathbf{c}(t), t) \nabla c(u, t)] \quad (4)$$

where $\text{div}[\cdot]$ is the divergence operator and ∇ is the gradient. $g(\mathbf{c}(t), t)$ is the diffusivity that governs the rate at which proteins spread between the nodes. In inhomogeneous graph diffusion, the diffusivity defined over each edge varies over time t and state $\mathbf{c}(t)$. Stacking $g(\mathbf{c}(t), t)$ will arrive in a matrix $\mathbf{G}(\mathbf{c}(t), t) \in \mathbb{R}^{e \times e}$. Note that the process described in Eq. 1 corresponds to the homogeneous scenario when $g(\mathbf{c}(t), t) = \text{Const}$.

To derive a discrete version of this PDE, we introduce the incidence matrix $\mathbf{K} \in \mathbb{R}^{n \times e}$, which indicates the connectivity between vertices and edges in the graph (Godsil and Royle 2001). Substituting the expressions for div and ∇ leads to,

$$\begin{aligned} \frac{\partial \mathbf{c}(t)}{\partial t} &= \text{div}[\mathbf{G}(\mathbf{c}(t), t) \mathbf{K}^\top \mathbf{c}(t)] \\ &= -\mathbf{K} \cdot \mathbf{G}(\mathbf{c}(t), t) \cdot \mathbf{K}^\top \mathbf{c}(t) \end{aligned} \quad (5)$$

However, modeling an inhomogeneous diffusion that refines the graph requires considering all possible edges, resulting in a quadratic computational complexity with $e = n^2$. Since \mathbf{G} is symmetric, $\mathbf{K} \cdot \mathbf{G}(\mathbf{c}(t), t) \cdot \mathbf{K}^\top$ is positive semi-definite and can be approximated by a low-rank factorization. This motivates us to learn two matrices $\mathbf{M}_{\text{Enc}}, \mathbf{M}_{\text{Dec}} \in \mathbb{R}^{n \times e'}$, $e' \ll e$ depending on states and time such that,

$$\mathbf{M}_{\text{Dec}} \cdot \mathbf{M}_{\text{Enc}}^\top \approx \mathbf{K} \cdot \mathbf{G}(\mathbf{c}(t), t) \cdot \mathbf{K}^\top. \quad (6)$$

The problem is equivalent to finding the low-rank latent representation $\mathbf{h}(t)$ that correctly predicts $d\mathbf{c}(t)/dt$, which inspires us to use an auto-encoder-like model. To this end, GAE naturally excels in modeling inhomogeneous graph diffusion given its ability to generate latent representations and refine graph structure. The encoder maps the input graph \mathcal{G} with node features $\mathbf{c}(t)$ into a latent representation $\mathbf{h}(t)$ via graph convolutional layers. Specifically, for node u , the encoder calculates the latent representation as,

$$\mathbf{h}(u, t) = \text{Enc}(c(u, t), \bigoplus_{v \in [N]} \text{Prop}(c(u, t), c(v, t), \mathbf{A}_{uv})), \quad (7)$$

where $\text{Enc}(\cdot)$ and $\text{Prop}(\cdot)$ are encoder and propagation functions respectively, \bigoplus is a permutation-invariant aggregation function and $[N] = \{1, \dots, N\}$. Moreover, the decoder refines the graph structure and predicts the node signal by,

$$\frac{d\mathbf{c}(t)}{dt} = \text{Dec}(\mathbf{h}(t), \hat{A}), \quad \text{with } \hat{A} = \sigma(\mathbf{h}(t)\mathbf{h}(t)^\top). \quad (8)$$

$\text{Dec}(\cdot, \cdot)$ is the decoder function that takes in the latent representation of refined graph structure, and $\sigma(\cdot)$ is the normalization function to make sure the refined adjacency matrix is valid. With such a design, the graph auto-encoder provides an estimation of the derivative of the pathology propagation trajectory while refining the graph structure.

Localized Neural Reaction. In addition to spatial diffusion of pathological proteins via brain connections, their localised replication plays an important role in pathology accumulation. To enhance the expressiveness of this process, we use a multi-layer perception (MLP) to approximate such a process:

$$f_L = \frac{dc(t)}{dt} = \text{MLP}(\mathbf{c}(t), t, \theta_L) \quad (9)$$

Training of the Trajectory Construction Step. Beyond its enhanced modeling capacity, the proposed framework provides an interpretable way to assess the contribution of

each component and identify pathological patterns that the pathophysiological model may have overlooked. We achieve this by incorporating norm and orthogonal regularizations to diversify the model outputs and disentangle the impact of each individual component. The overall IGND-MoE model is trained by the following loss:

$$\mathcal{L} = \mathcal{L}_{\text{traj}} + \lambda_1 \mathcal{L}_{\text{norm}} + \lambda_2 \mathcal{L}_{\text{ortho}} \quad (10)$$

where $\mathcal{L}_{\text{traj}}$ is the ODE trajectory loss, $\mathcal{L}_{\text{norm}}$ is the norm loss on the learning-based model output and $\mathcal{L}_{\text{ortho}}$ is the orthogonal loss on each model's output. λ_1 and λ_2 are the hyper-parameters to balance three loss terms. Specifically, $\mathcal{L}_{\text{traj}}(\theta_S, \theta_L, \theta_M) = \sum_{i=1}^N \sum_{s=1}^{S_i} \|\mathbf{c}_{\text{obs}}(t_0^i + t_s^i) - \mathbf{c}(t_0^i + t_s^i)\|^2$ with $\mathbf{c}_{\text{obs}}(t_s^i)$ is the observation at t_s^i , where i ranges from 1 to N and s ranges from 1 to S_i . The norm loss is used to prioritize the contribution from the pathophysiological model and use the learning-based model as a complement to encourage interpretability and improve generalization ability. Similar to (Yin et al. 2021), we regularize the training through the Frobenius norm of model outputs $\mathcal{L}_{\text{norm}} = \|f_S\|_F + \|f_L\|_F$.

To encourage diversity among the models' predictions, we add an orthogonal regularization term that penalizes the correlation between the outputs of individual ensemble members. This forces each model to learn complementary features or focus on different aspects of the data, which can lead to better generalization and improved overall performance once combined.

$$\mathcal{L}_{\text{ortho}} = \sum_t \sum_{p \neq q} \left(\tilde{f}_p(t) \cdot \tilde{f}_q(t) \right)^2, \quad \tilde{f}_p = \frac{f_p - \text{mean}(f_p)}{\text{std}(f_p)},$$

with p and q index the models and the mean is taken across the subjects.

Experiments and Results

In this section, we will evaluate the proposed model in the task of long-term disease progression of tau pathology in human brains in Alzheimer's disease. First, we will benchmark our model against established spatio-temporal approaches, including discrete models and continuous models with both homogeneous and inhomogeneous configurations. Subsequently, we will conduct in-depth clinical analyses by examining how different component contributions evolve throughout disease progression. We will also identify the specific brain regions and disease stages where our model demonstrates improvement by comparing error maps between our proposed model and classical pathophysiological models. The results will provide insights into the synergistic effects of various model components (each representing distinct physical processes) within the brain's dynamic system. We also demonstrate that with proper constraints from the expert, the proposed method is more robust in the out-of-distribution case (OOD). Please refer to Appendix A.1 in the extended version.

Experimental Setups

Data. In this study we model the dynamics of tau protein, which we compare to standardized uptake value

ratios (SUVRs) obtained from tau-PET imaging from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) (Landau et al. 2021). The SUVRs of cortical regions were normalised to [0,1] across all participants and regions. For each cortical region, we implement a two-component Gaussian mixture model to establish a cutoff for tau-positivity as the mean of the negative distribution plus its one standard deviation. The cohort consists of $N = 216$ individuals (378 observations) who had positive amyloid-beta and tau status. Apart from ADNI, we also carry out validation on two external datasets - the Harvard Ageing Brain Study (HABS)(Dagley et al. 2017) and the Anti-Amyloid Treatment in Asymptomatic Alzheimer’s Disease (A4) study(Sperling et al. 2014, 2020). See Appendix A.2 in the extended version for detailed results. Appendix A.4 displays more detailed data information.

The structural connectome is the group average across 50 individuals from the Microstructure-Informed Connectomics Database (Royer et al. 2022) defined by the Desikan-Killiany Atlas (Desikan et al. 2006), to formulate a stable and continuous cohort-level trajectory that captures the average disease progression.

Training Methodology We implement cross-validation by randomly assigning 35 subjects each to validation and test sets, with the remaining subjects forming the training set. All longitudinal scans from the same subject are kept together in their assigned sets, preserving the actual time intervals between measurements. The validation step happens after an epoch of trajectory optimization on the training data, i.e. the subjects from the validation set are allocated on the trajectory from each training epoch through stage optimization. Finally, the relative location of the subjects from the test set is estimated, and the corresponding model performance is recorded as the test metric.

Benchmarks. We compare our IGND-MoE with three types of models: 1) temporal-GNN models, which are usually combinations of traditional discrete time series models (RNN, LSTM and GRU) and classic GNN models for temporal modelling (Li et al. 2018; Kazemi et al. 2019; Seo et al. 2018). Since these models by nature cannot handle smooth and continuous data, we apply kernel interpolation methods to make them comparable with our proposed models, see Appendix A.3 for details. 2) Neural ODE models designed from homogeneous spatial diffusion (Zang and Wang 2020) 3) Hybrid model with MoE structure but we replace the backbone model with other GNN (Kipf and Welling 2017a) models, such as GCN (Kipf and Welling 2017b), g-transformer (Yu et al. 2020), FROND (Kang et al. 2024), NSD (Bodnar et al. 2022). We also consider the case when we remove the stage-aware characteristics of the mixture mechanisms and thus make the model contribution time-independent.

Metrics. We evaluate model performance using two metrics on the test set: sum of squared errors (SSE), and average Pearson correlation coefficient (R). SSE is calculated across all subjects and regions of interest to quantify prediction accuracy. For each observation, we compute the Pearson correlation coefficient between predicted and measured tau SUVR signals across brain regions, which captures the

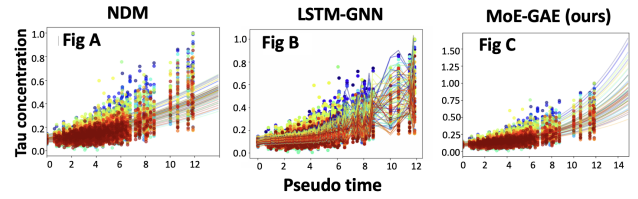


Figure 3: Model predictions. Each curve depicts tau-accumulation trajectories for one of 68 cortical regions, with dots marking individual observations.

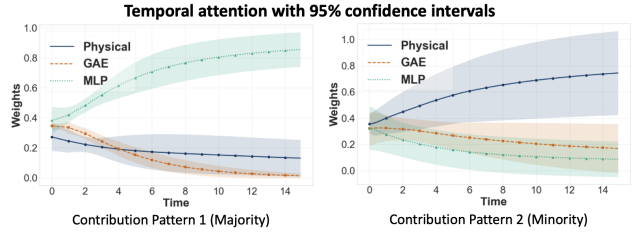


Figure 4: Model contribution versus time. Two patterns found: the majority (70% of runs, high confidence) and the minority (30%, high uncertainty).

similarity of the spatial distribution of pathology. These individual correlations are then averaged to obtain a group-level measure of the model’s ability to reproduce region-specific pathology patterns independent of absolute magnitude.

Quantitative Model Comparison

We conduct the experiments using models mentioned in the Benchmarks section. Table 1 compares models on three metrics described in the Metrics section. With our IGND-MoE, a consistent improvement is found in all metrics on test data, in comparison to the short-term discrete and homogeneous graph diffusion methods. Specifically, 1) When replacing the GAE with other GNNs, the performance drops. The results suggest that GNNs without any graph refinement techniques lack the capacity to fully capture the complex dynamics. 2) When replacing temporal attention and assuming fixed contribution from experts, the model performance also drops. **This shows that the contribution from different disease mechanisms varies at different disease stages.**

Qualitative Results Analysis

With our proposed framework, we construct the long-term trajectory of tau propagation in Alzheimer’s disease using a mixture of different experts. Figure 3 visualises the optimisation trajectories obtained from three different models on training data. Figure 3: A shows that noticeable patterns cannot be described at the later disease stages using the pure pathophysiological model. Figure 3: B shows that without proper constraints, data-driven deep learning models like LSTM-RNN easily overfit. Figure 3: C shows that the pathophysiological component can give a more plausible constraint to the MoE trajectory, and the proposed model

Method	Backbone Model	Metrics	
		Test SSE	Test R Corr
Physical Model (Raj, Kuceyeski, and Weiner 2012a)	/	17.76 ± 0.32	0.657 ± 0.018
Discrete Methods with interpolation	LSTM-GNN (Kazemi et al. 2019)	17.99 ± 1.15	0.649 ± 0.008
	RNN-GNN (Li et al. 2018)	21.34 ± 1.82	0.610 ± 0.020
	GRU-GNN (Seo et al. 2018)	15.66 ± 0.76	0.676 ± 0.020
Homogeneous Graph Diffusion (Zang and Wang 2020)	NeuralODE (localized)	19.65 ± 0.27	0.701 ± 0.002
	NeuralODE (cross-node)	16.73 ± 0.20	0.700 ± 0.002
	NDCN	14.89 ± 0.19	0.677 ± 0.014
Inhomogeneous Graph Diffusion	MoE-GCN	16.50 ± 1.04	0.676 ± 0.019
	MoE-GTrans	16.30 ± 0.83	0.679 ± 0.049
	MoE-FROND	16.61 ± 1.93	0.693 ± 0.030
	MoE-NSD	16.36 ± 3.10	0.671 ± 0.030
	MoE-GAE (w/o Temporal)	15.55 ± 0.44	0.681 ± 0.020
	MoE-GAE (ours)	13.97 ± 0.30	0.717 ± 0.014

* The results are averaged over 3 trials using three random seeds. The mean & standard deviation are reported.

Table 1: Quantitative results for long-term disease trajectory prediction.

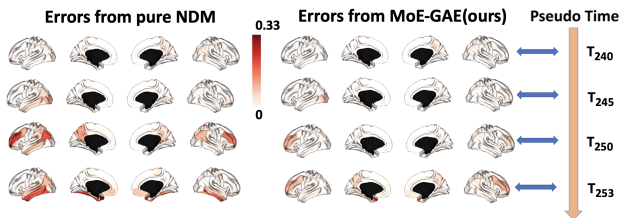


Figure 5: Regional Error mapping on the brain. The plot displays the distribution of the error pattern during disease progression from the real observations using the physical model and the MoE-GAE model. The colour bar, shared by all brain plots, displays the extent of error levels.

structure, with the proper choice of the additional component, can describe more unknown patterns.

A key insight is the dynamic contribution from each expert model, dependent on disease stage. Figure 4 displays each expert model’s contribution versus disease stage. In 10-fold cross-validation, two patterns emerged: a majority pattern (70% of runs, high confidence) and a minority pattern (30%, high uncertainty). The majority pattern shows the physical model and GAE contributions decreasing with disease progression, while the MLP becomes dominant at later stages. This suggests graph-related processes are more influential early in our cohort, while other unknown physical processes dominate later stages. The existing contribution from GAE indicates that network diffusion isn’t the only graph-related mechanism. These findings align with literature showing pathophysiological model pathology diffusion occurs primarily in earlier disease stages with deep learning models complementing later stages where other mechanisms become more prominent (Meisl et al. 2021b). The minority pattern during some of the data splits might hint at the existence of more than one subgroup present in the cohort. Figure 5 demonstrates the regional error of each model with disease progression, focusing on the later stages, from which

we can gain insights about each region and time. The original pathophysiological model provides limited understanding and is complemented by the proposed MoE-GAE model, as shown by the higher regional errors in the left panel.

Discussions and Conclusions

We propose a novel framework for modelling long-term disease progression by combining pathophysiological models with neural networks. This mixture of experts effectively characterises the evolving dynamics of tau propagation across different disease stages in Alzheimer’s disease, providing interpretable insights into disease mechanisms. Our results reveal that graph-related processes are more influential early on, while other unknown physical processes become more prominent later, findings that align with existing literature. The IGND-MoE model achieves superior long-term prediction accuracy and provides robust out-of-sample predictions. The inherent limited number of longitudinal tau-PET data, a well-known constraint in this field (Leuzy et al. 2023; Yang et al. 2021), presents challenges for researchers in this field. However, our framework is specifically designed to address this by constructing cohort-level trajectories from sparse observations and mitigating the risk of overfitting through our expert-constrained architecture with special loss design. The use of a group-averaged structural connectome, a common approach for modelling cohort-level progression, is also implicitly refined by our model’s Graph Auto-Encoder (GAE) component, which refines a data-driven propagation graph from individual features, adapting to individual pathology patterns. Our findings, validated across multiple independent datasets, demonstrate that this complex but carefully constrained approach provides a robust and generalizable proof-of-concept. Our method can be further refined to enhance its clinical utility, including its adaptability to other biomarkers and diseases.

Acknowledgments

T.H, A.S, E.T, S.S and D.C.A are supported by the Wellcome Trust (221915). D.C.A and F.B are supported by the NIHR Biomedical Research Centre at UCLH and UCL. T.H. is supported by the EPSRC funded UCL Centre for Doctoral Training in Intelligent, Integrated Imaging in Healthcare [EP/S021930/1]. K.J. is supported by the UKRI Engineering and Physical Sciences Research Council (EPSRC) [EP/R513143/1].

One dataset used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

One dataset used in the preparation of this article were obtained from the Harvard Aging Brain Study (HABS-P01AG036694). The HABS study was launched in 2010, funded by the National Institute on Aging. and is led by principal investigators Reisa A. Sperling MD and Keith A. Johnson MD at Massachusetts General Hospital/Harvard Medical School in Boston, MA.

The A4 Study was a secondary prevention trial in pre-clinical Alzheimer’s disease targeting cognitive decline in clinically normal older individuals with brain amyloid accumulation. Funded by NIH-National Institute on Aging, Eli Lilly and Company, Alzheimer’s Association, Accelerating Medicines Partnership, GHR Foundation, an anonymous foundation, and private donors, with in-kind support from Avid Radiopharmaceuticals, Cogstate, Albert Einstein College of Medicine, and Foundation for Neurologic Diseases. The companion observational LEARN Study was funded by Alzheimer’s Association and GHR Foundation. Both studies were led by Dr. Reisa Sperling (Brigham and Women’s Hospital, Harvard Medical School) and Dr. Paul Aisen (Alzheimer’s Therapeutic Research Institute, University of Southern California), coordinated by ATRI, with data available through Alzheimer’s Clinical Trial Consortium’s Global Research and Imaging Platform. Complete A4 Study Team list: <https://www.actcinfo.org/a4-study-team-list/>. We acknowledge the dedication of study participants and partners.

References

Abi Nader, C.; Ribaldi, F.; Frisoni, G. B.; Garibotto, V.; Robert, P.; Ayache, N.; and Lorenzi, M. 2022. SimulAD: a dynamical model for personalized simulation and disease staging in Alzheimer’s disease. *Neurobiology of Aging*, 113: 73–83.

Avena-Koenigsberger, A.; Misić, B.; and Sporns, O. 2018. Communication dynamics in complex brain networks. *Nature reviews neuroscience*, 19(1): 17–33.

Bodnar, C.; Di Giovanni, F.; Chamberlain, B.; Lio, P.; and Bronstein, M. 2022. Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns.

Advances in Neural Information Processing Systems, 35: 18527–18541.

Busche, M. A.; and Hyman, B. T. 2020. Synergy between amyloid- β and tau in Alzheimer’s disease. *Nature neuroscience*, 23(10): 1183–1193.

Carslaw, H.; and Jaeger, J. 1959. *Conduction of Heat in Solids*. Oxford science publications. Clarendon Press. ISBN 9780198533689.

Chen, R. T. Q.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. 2018. Neural Ordinary Differential Equations.

Chung, J.; Gulcehre, C.; Cho, K.; and Bengio, Y. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Dagley, A.; LaPoint, M.; Huijbers, W.; Hedden, T.; McLaren, D. G.; Chatwal, J. P.; Papp, K. V.; Amariglio, R. E.; Blacker, D.; Rentz, D. M.; et al. 2017. Harvard aging brain study: dataset and accessibility. *Neuroimage*, 144: 255–258.

Desikan, R. S.; et al. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3): 968–980. Publisher: Academic Press.

Fonteijn, H. M.; Modat, M.; Clarkson, M. J.; Barnes, J.; Lehmann, M.; Hobbs, N. Z.; Scathill, R. I.; Tabrizi, S. J.; Ourselin, S.; Fox, N. C.; et al. 2012. An event-based model for disease progression and its application in familial Alzheimer’s disease and Huntington’s disease. *NeuroImage*, 60(3): 1880–1889.

Garbarino, S.; Lorenzi, M.; and Initiative, A. D. N. 2019. Modeling and inference of spatio-temporal protein dynamics across brain networks. In *International Conference on Information Processing in Medical Imaging*, 57–69. Springer.

Garbarino, S.; Lorenzi, M.; Oxtoby, N. P.; Vinke, E. J.; Marinescu, R. V.; Eshaghi, A.; Ikram, M. A.; Niessen, W. J.; Ciccarelli, O.; Barkhof, F.; et al. 2019. Differences in topological progression profile among neurodegenerative diseases from imaging data. *Elife*, 8: e49298.

Godsil, C. D.; and Royle, G. F. 2001. *Algebraic Graph Theory*. Graduate texts in mathematics. Springer.

Graves, A.; and Graves, A. 2012. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, 37–45.

He, T.; Thompson, E.; Schroder, A.; Oxtoby, N. P.; Abdulla, A.; Barkhof, F.; and Alexander, D. C. 2023. A coupled-mechanisms modelling framework for neurodegeneration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 459–469. Springer.

Kang, Q.; Zhao, K.; Ding, Q.; Ji, F.; Li, X.; Liang, W.; Song, Y.; and Tay, W. P. 2024. Unleashing the Potential of Fractional Calculus in Graph Neural Networks with FROND. In *The Twelfth International Conference on Learning Representations*.

Kazemi, S. M.; Goel, R.; Jain, K.; Kobzyev, I.; Sethi, A.; Forsyth, P.; and Poupart, P. 2019. Relational Representation Learning for Dynamic (Knowledge) Graphs: A Survey. *CoRR*, abs/1905.11485.

- Kipf, T. N.; and Welling, M. 2017a. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR (Poster)*. OpenReview.net.
- Kipf, T. N.; and Welling, M. 2017b. Semi-Supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations (ICLR)*.
- Landau, S.; Ward, T. J.; Murphy, A.; and Jagust, W. 2021. Flortaucipir (AV-1451) processing methods. 8.
- Leuzy, A.; Binette, A. P.; Vogel, J. W.; Klein, G.; Borroni, E.; Toniello, M.; Strandberg, O.; Mattsson-Carlgren, N.; Palmqvist, S.; Pontecorvo, M. J.; et al. 2023. Comparison of group-level and individualized brain regions for measuring change in longitudinal tau positron emission tomography in Alzheimer disease. *JAMA neurology*, 80(6): 614–623.
- Li, Y.; Yu, R.; Shahabi, C.; and Liu, Y. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *ICLR (Poster)*. OpenReview.net.
- Lorenzi, M.; Filippone, M.; Frisoni, G. B.; Alexander, D. C.; Ourselin, S.; Initiative, A. D. N.; et al. 2019. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer’s disease. *NeuroImage*, 190: 56–68.
- Meisl, G.; Hidari, E.; Allinson, K.; Rittman, T.; Devos, S. L.; Sanchez, J. S.; Xu, C. K.; Duff, K. E.; Johnson, K. A.; Rowe, J. B.; Hyman, B. T.; Knowles, T. P. J.; and Klenerman, D. 2021a. In vivo rate-determining steps of tau seed accumulation in Alzheimer’s disease. Technical report.
- Meisl, G.; Hidari, E.; Allinson, K.; Rittman, T.; DeVos, S. L.; Sanchez, J. S.; Xu, C. K.; Duff, K. E.; Johnson, K. A.; Rowe, J. B.; et al. 2021b. In vivo rate-determining steps of tau seed accumulation in Alzheimer’s disease. *Science advances*, 7(44): eabh1448.
- Poli, M.; Massaroli, S.; Park, J.; Yamashita, A.; Asama, H.; and Park, J. 2019. Graph neural ordinary differential equations. *arXiv preprint arXiv:1911.07532*.
- Raj, A.; Kuceyeski, A.; and Weiner, M. 2012a. A Network Diffusion Model of Disease Progression in Dementia. *Neuron*, 73(6): 1204–1215.
- Raj, A.; Kuceyeski, A.; and Weiner, M. 2012b. A network diffusion model of disease progression in dementia. *Neuron*, 73(6): 1204–1215.
- Royer, J.; et al. 2022. An Open MRI Dataset For Multiscale Neuroscience. *Scientific Data*, 9(1).
- Rumelhart, D. E.; Hinton, G. E.; Williams, R. J.; et al. 1985. Learning internal representations by error propagation.
- Seguin, C.; Sporns, O.; and Zalesky, A. 2023. Brain network communication: concepts, models and applications. *Nature reviews neuroscience*, 24(9): 557–574.
- Seo, Y.; Defferrard, M.; Vandergheynst, P.; and Bresson, X. 2018. Structured Sequence Modeling with Graph Convolutional Recurrent Networks. In *ICONIP (1)*, volume 11301 of *Lecture Notes in Computer Science*, 362–373. Springer.
- Sperling, R. A.; Donohue, M. C.; Raman, R.; Sun, C.-K.; Yaari, R.; Holdridge, K.; Siemers, E.; Johnson, K. A.; Aisen, P. S.; Team, A. S.; et al. 2020. Association of factors with elevated amyloid burden in clinically normal older individuals. *JAMA neurology*, 77(6): 735–745.
- Sperling, R. A.; Rentz, D. M.; Johnson, K. A.; Karlawish, J.; Donohue, M.; Salmon, D. P.; and Aisen, P. 2014. The A4 study: stopping AD before symptoms begin? *Science translational medicine*, 6(228): 228fs13–228fs13.
- Thanou, D.; Dong, X.; Kressner, D.; and Frossard, P. 2017. Learning Heat Diffusion Graphs. *IEEE Trans. Signal Inf. Process. over Networks*, 3(3): 484–499.
- Weickenmeier, J.; Kuhl, E.; and Goriely, A. 2018a. Multiphysics of prionlike diseases: Progression and atrophy. *Physical review letters*, 121(15): 158101.
- Weickenmeier, J.; Kuhl, E.; and Goriely, A. 2018b. Multiphysics of Prionlike Diseases: Progression and Atrophy. *Physical Review Letters*, 121(15).
- Yang, F.; Chowdhury, S. R.; Jacobs, H. I.; Sepulcre, J.; Wedeen, V. J.; Johnson, K. A.; and Dutta, J. 2021. Longitudinal predictive modeling of tau progression along the structural connectome. *NeuroImage*, 237: 118126.
- Yin, Y.; Le Guen, V.; Dona, J.; de Bézenac, E.; Ayed, I.; Thome, N.; and Gallinari, P. 2021. Augmenting physical models with deep networks for complex dynamics forecasting. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12): 124012.
- Young, A. L.; Oxtoby, N. P.; Daga, P.; Cash, D. M.; Fox, N. C.; Ourselin, S.; Schott, J. M.; and Alexander, D. C. 2014. A data-driven model of biomarker changes in sporadic Alzheimer’s disease. *Brain*, 137(9): 2564–2577.
- Young, A. L.; Oxtoby, N. P.; Garbarino, S.; Fox, N. C.; Barkhof, F.; Schott, J. M.; and Alexander, D. C. 2024. Data-driven modelling of neurodegenerative disease progression: thinking outside the black box.
- Yu, C.; Ma, X.; Ren, J.; Zhao, H.; and Yi, S. 2020. Spatio-Temporal Graph Transformer Networks for Pedestrian Trajectory Prediction. In *ECCV (12)*, volume 12357 of *Lecture Notes in Computer Science*, 507–523. Springer.
- Zang, C.; and Wang, F. 2020. Neural dynamics on complex networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 892–902.
- Zhou, J.; Gennatas, E. D.; Kramer, J. H.; Miller, B. L.; and Seeley, W. W. 2012a. Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron*, 73(6): 1216–1227.
- Zhou, J.; Gennatas, E. D.; Kramer, J. H.; Miller, B. L.; and Seeley, W. W. 2012b. Predicting Regional Neurodegeneration from the Healthy Brain Functional Connectome. *Neuron*, 73(6): 1216–1227.