LNODE: Uncovering the Latent Dynamics of Aβ in Alzheimer's Disease

Zheyu Wen[1 [0000−0001−9628−1449]] and George Biros[1 [0000−0002−0033−3994]]

Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 E. 24th Street, Austin, Texas 78712, USA
zheyw@utexas.edu, biros@oden.utexas.edu

Abstract. Aβ Positron Emission Tomography (PET) is often used to manage Alzheimer's disease (AD). To better understand Aβ progression, we introduce and evaluate a mathematical model that couples Aβ at parcellated gray matter regions. We term this model LNODE for *"latent network ordinary differential equations"*. At each region, we track normal Aβ, abnormal Aβ, and m latent states that intend to capture unobservable mechanisms coupled to Aβ progression. LNODE is parameterized by subject-specific parameters and cohort parameters. We jointly invert for these parameters by fitting the model to Aβ-PET data from 585 subjects from the ADNI dataset. Although underparameterized, our model achieves population $R^2 \geq 98\%$ compared to $R^2 \leq 60\%$ when fitting without latent states. Furthermore, these preliminary results suggest the existence of different subtypes of Aβ progression.

Keywords: Aβ pathology · PET imaging · Alzheimer's disease · network model · biophysical modeling.

1 Introduction

The deposition of abnormal (misfolded) amyloid beta (Aβ) is widely believed to be linked to Alzheimer's disease [12, 18]. Longitudinal Aβ-PET imaging is one of the best ways to study abnormal Aβ [19]. Here we introduce LNODE as a framework to construct cohort-shared mathematical models of Aβ progression, and eventually extend it to misfolded tau. LNODE is constructed by calibrating it on PET scans from the ADNI dataset [13]. The model's subject-specific parameters aim to capture individual variations in disease progression, while the cohort-shared parameters reflect unobserved dynamics of AD disease. The general methodology is summarized in Figure 1.

Contributions: ❶ We introduce a novel ODE model for the Aβ evolution. ❷ We present a cohort inversion algorithm to simultaneously estimate subject-specific and cohort-shared parameters. ❸ We demonstrate the effectiveness of the model using synthetic data. ❹ We evaluate LNODE on n =585 subjects from the ADNI dataset and study its sensitivity to m, the number of latent states. As we discuss in Section 3, LNODE significantly improves the state of the art while using ∼2.4 extra parameters per subject compared to a model without latent states. The

model without latent states is among the most widely used for describing Aβ spreading [8]. LNODE accurately reconstructs the observations even for left-out subjects whose data was not used to calibrate the cohort parameters.

Related Work: Several approaches have been proposed to predict the progression of AD using deep learning and machine learning techniques. For example, [27] introduced a deep convolutional network model to predict the disease state of AD from 3D-PET images, while [9] developed a multimodal deep learning model that uses both MRI and PET images for AD classification. In addition, [3] used Aβ-PET images for Aβ positivity prediction. These methods generally involve a large number of parameters and do not offer a mechanism-based perspective. Simpler and more interpretable models often target the aggregate Aβ-PET signal. For example, [5] used logistic functions to forecast the progression of scalar biomarkers—including global Aβ-PET and Tau-PET—from cohort data; and [10] utilized ROI-based multimodal data (including Aβ-PET) to predict the transition from mild cognitive impairment (MCI) to AD. More recently, [24] introduced a conceptual framework for aligning personalized disease (biomarker) trajectories with a global progression timeline, highlighting the need for a model that can fit cohort datasets while preserving interpretability.

Complementary to the methods above, ODE/mechanism-based models have been used to fit observational Aβ-PET or Tau-PET data using $O(1)$ subject-specific parameters. Among these, the Fisher-Kolmogorov model is one of the most widely used [8,20]. Recent alternatives have been proposed to improve performance fitting PET data [16, 22]; furthermore, [25] combines physics-informed neural networks and symbolic regression to discover an appropriate partial differential equation for modeling Tau-PET. These works analyze each patient independently and do not leverage cohort-shared disease dynamics. Our approach builds on these methods by integrating both subject-specific and cohort-shared parameters to model the cohort dynamics of disease progression.

2 Methodology

Let N denote the number of gray matter regions (ROIs). We use $\mathbf{b}_a(t) \in \mathbb{R}^N$ to denote the abnormal Aβ and $\mathbf{b}_n(t)$ the normal Aβ. Here t is a disease age that we discuss further below. Following standard modeling strategies for Aβ spread, we represent ROI interactions using a graph Laplacian [1, 20, 23]. Using tractography data from an atlas [2], we define a reference sparse ROI connectivity matrix $\mathbf{W}$ and use it to define the (negative) graph Laplacian by $\mathbf{L} := \mathbf{W} - \mathsf{diag}(\sum_{s,s\neq r} [\mathbf{W}]_{rs})$; $\mathbf{L}$ is the same for all subjects.

We obtain the *observed data* $\mathbf{d} := \mathbf{b}_a^{\text{obs}}$ using the preprocessing steps summarized in Figure 1**(A)**. We use FSL [15], to register the brain MUSE atlas [2] and each subject's Aβ-PET images to the subject's first T1 MRI. The MUSE atlas segments the brain into $N = 114$ gray matter ROIs. After that the scan values are normalized by the median value of the cerebellum (reference ROI). Finally, for each ROI ℓ at time t, we define $\mathbf{d}[\ell](t)$ as the maximum mean discrepancy (MMD) distributional distance [6] between ROI ℓ and the cerebellum.

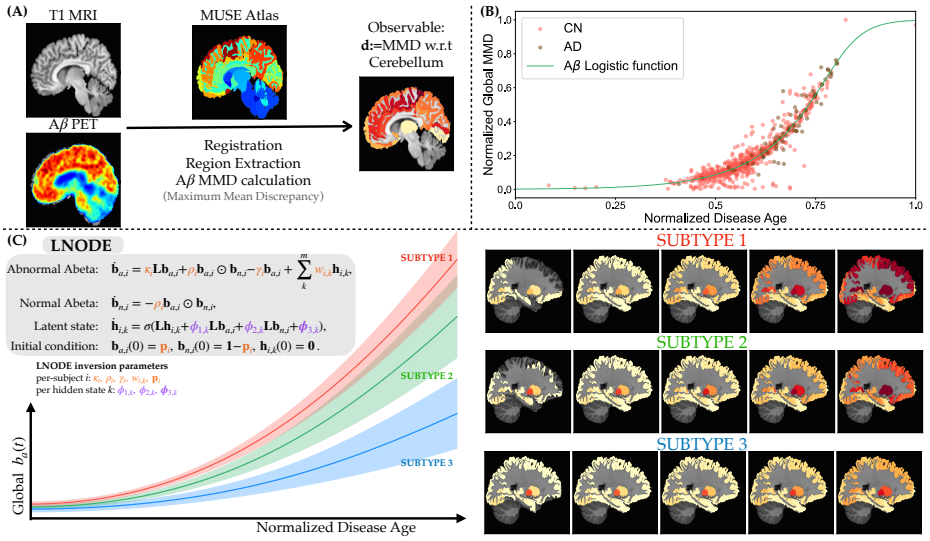

FIG. 1: *Method summary:* **(A)** *First, we register the parcellated MUSE atlas and the subject's Aβ-PET to the subject's MRI to define ROIs. For each ROI, we compute the Aβ Maximum Mean Discrepancy (MMD) between the ROI and the cerebellum resulting in a regional representation of degree of abnormal Aβ. These values collect in a* $\boldsymbol{d} \in \mathbb{R}^N$ *comprise the per-subject observations that drive the inverse problem for the LNODE parameters.* **(B)** *A well known issue is cohort age normalization. Following [5], we normalize chronological time to a "disease age", which we use as the time variable in LNODE. The plots show the fit (line) and all the patient scans. In this way, we can align individual subject data onto a common timeline, thereby enabling effective cohort inversion.* **(C)** *The LNODE formulation and inversion parameters: subject-specific parameters (highlighted in orange) and cohort-shared parameters (highlighted in purple); i indicates subject; k indicates latent state. Disease age is a normalized common time scale that is necessary for our cohort analysis. In the plot, we depict average trajectories and their half variance (for visualization purpose) for three distinct subtypes of Aβ progression found with k-means clustering of* w_{ik}*. The sagittal images show trajectories for the subtypes. White indicates low Aβ abnormality; red indicates high-degree of Aβ abnormality.*

Following [5], we define the unnormalized disease age as $t' = a\mathtt{t} + b$, where $\mathtt{t}$ is the chronological age. We fit a logistic function between t' and the global Aβ MMD score across all subjects in ADNI, as shown in Figure 1**(B)**. The disease age is then normalized to the range $[0, 1]$ using $t = \frac{t' - t'_{\min} + \eta}{t'_{\max} - t'_{\min} + \eta}$, where η maximizes the classification accuracy in distinguishing cognitively normal (CN) from Alzheimer's disease (AD) subjects using t as the sole feature.

Figure 1**(C)** summarizes our ODE model. Let N denote the number of ROIs and n the number of subjects. We denote the abnormal and normal Aβ concentrations for subject i as $\mathbf{b}_{a,i} \in \mathbb{R}^N_+$ and $\mathbf{b}_{n,i} \in \mathbb{R}^N_+$, respectively, for $i = 1, \ldots, n$. To explore the subtypes of Aβ trajectories, we introduce m latent states $\mathbf{h}_{i,k} \in \mathbb{R}^N$ for $k = 1, \ldots, m$, where $m \ll n$.

LNODE: With these definitions, our model for subject i reads as follows:

$$\text{Abnormal A}\beta\text{: } \dot{\mathbf{b}}_{a,i} = \kappa_i \mathbf{L}\mathbf{b}_{a,i} + \rho_i \mathbf{b}_{a,i} \odot \mathbf{b}_{n,i} - \gamma_i \mathbf{b}_{a,i} + \sum_k^m w_{i,k}\mathbf{h}_{i,k}, \quad (1a)$$

$$\text{Normal A}\beta\text{: } \dot{\mathbf{b}}_{n,i} = -\rho_i \mathbf{b}_{a,i} \odot \mathbf{b}_{n,i}, \quad (1b)$$

$$\text{Latent states: } \dot{\mathbf{h}}_{i,k} = \sigma(\mathbf{L}\mathbf{h}_{i,k} + \phi_{1,k}\mathbf{L}\mathbf{b}_{a,i} + \phi_{2,k}\mathbf{L}\mathbf{b}_{n,i} + \boldsymbol{\phi}_{3,k}) \ \forall k \quad (1c)$$

$$\text{Initial conditions: } \mathbf{b}_{a,i}(0) = \mathbf{p}_i,\ \mathbf{b}_{n,i}(0) = \mathbf{1} - \mathbf{p}_i,\ \mathbf{h}_{i,k}(0) = \mathbf{0},\ \forall k, \quad (1d)$$

where $\dot{\boldsymbol{b}}$ indicates time derivative, $w_{i,k} \in \mathbb{R}$, and $\odot$ represents the element-wise product. In the latent-state dynamics ROI coupling is done through $\mathbf{L}$ for $\mathbf{h}_{i,k}$, $\mathbf{b}_{a,i}$, $\mathbf{b}_{n,i}$—the last two acting as source terms; $\sigma : \mathbb{R}^N \to \mathbb{R}^N$ is an element-wise function. We have tested $\sigma(h) = h$ and $\sigma(h) = \mathsf{relu}(h)$. In Equation (1), the abnormal Aβ $\mathbf{b}_{a,i}$ is coupled to the latent states through the $w_{i,k}\mathbf{h}_{i,k}$, where $w_{i,k}$ "selects" latent states. The diffusion term $\mathbf{L}\mathbf{b}_{a,i}$ is scaled by the diffusivity constant $\kappa_i \in \mathbb{R}_+$. Abnormal A$\beta$ also causes the transition of normal Aβ $\mathbf{b}_{n,i}$ into abnormal via the term $\rho_i \mathbf{b}_{a,i} \odot \mathbf{b}_{n,i}$, where $\rho_i \in \mathbb{R}_+$. We incorporate clearance of abnormal Aβ $\mathbf{b}_{a,i}$ using $-\gamma_i \mathbf{b}_{a,i}$, with $\gamma_i \in \mathbb{R}_+$. The initial condition (IC) $\mathbf{p}_i$ represents a sparse initial for abnormal Aβ under the assumption that misfolding originates at a small number of regions and then spreads out to the rest of the gray matter [16, 22]. This ODE model is defined for $t \in (0, T]$, where T indicates the time horizon, and t is the disease age. Finally, the inversion parameters are split to subject-specific $\boldsymbol{\theta}_s := \{\kappa_i, \rho_i, \gamma_i, \mathbf{p}_i, w_{i,k}\ \forall i, k\}$, $i = 1, \ldots, n$; and cohort $\boldsymbol{\theta}_c := \phi_{1,k} \in \mathbb{R}$, $\phi_{2,k} \in \mathbb{R}$ and $\boldsymbol{\phi}_{3,k} \in \mathbb{R}^N, k = 1 \ldots m$, shared by all subjects. We define $\boldsymbol{\theta} := \{\boldsymbol{\theta}_s, \boldsymbol{\theta}_c\}$.

Inversion: We define $\mathbf{d}_{ij}$ as that data extracted from the j_{th} scan of subject i, acquired at time t_{ij}. Also, we use s_i to denote the number of scans for subject i. To reconstruct $\boldsymbol{\theta}$, we use the following objective function:

$$\mathcal{J} = \sum_i^n \frac{1}{2s_i} \sum_j^{s_i} \|\mathbf{b}_{a,i}(t_{ij}) - \mathbf{d}_{ij}\|_2^2 + \lambda_1 \sum_i^n \sum_\ell^N \log(1 - \mathbf{p}_i[\ell]) + \lambda_2 \sum_i^n \|\mathbf{w}_i\|_1 . \quad (2)$$

The constraint $\|\mathbf{w}_i\|_1$ reflects a modeling assumption that different disease subtypes have different latent state dynamics. As we will see in Section 3, we will cluster $\mathbf{w}$ to identify the subtype populations.

With these definitions, the overall inverse problem reads as follows:

$$\begin{aligned} &\min_{\boldsymbol{\theta}_s > 0, \boldsymbol{\theta}_c} \mathcal{J}(\boldsymbol{\theta}) \text{ subject to:} \\ &\qquad \text{Equation (1) holds and } \|\mathbf{p}_i\|_0 = s^{\max}\ \forall i. \end{aligned} \quad (3)$$

Note that we set ℓ_0 norm on IC to enforce a sparse initial condition for Aβ. We use gradient method to solve Equation (3). To compute the gradient with respect $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_c$ we use a Lagrangian/adjoint formulation, in which we first solve the backward-in-time adjoint ODEs and then we accumulate the gradients [4, 7]. Let $\boldsymbol{\alpha}_{a,i}$, $\boldsymbol{\alpha}_{n,i}$ and $\boldsymbol{\alpha}_{h,ik}$ be the adjoint variables for $\mathbf{b}_{a,i}$, $\mathbf{b}_{n,i}$, and $\mathbf{h}_{i,k}$. Also define

$\mathbf{o}_{ik} = \mathbf{L}\mathbf{h}_{i,k} + \phi_{1,k}\mathbf{L}\mathbf{b}_{a,i} + \phi_{2,k}\mathbf{L}\mathbf{b}_{n,i} + \boldsymbol{\phi}_{3,k}$. Then, the adjoint equations for the case $\sigma(\cdot) = \mathsf{relu}(\cdot)$ read:

$$\begin{aligned}\dot{\boldsymbol{\alpha}}_{a,i} = &-\kappa_i \mathbf{L}\boldsymbol{\alpha}_{a,i} + \rho_i \mathbf{b}_{n,i} \odot (\boldsymbol{\alpha}_{n,i} - \boldsymbol{\alpha}_{a,i}) + \gamma_i \boldsymbol{\alpha}_{a,i} \\ &+ \xi_{a,i} - \sum_k^m \phi_{1,k} \mathbf{L}\,\mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\boldsymbol{\alpha}_{h,ik}, \end{aligned} \tag{4a}$$

$$\dot{\boldsymbol{\alpha}}_{n,i} = \rho_i \mathbf{b}_{a,i} \odot (\boldsymbol{\alpha}_{n,i} - \boldsymbol{\alpha}_{a,i}) - \sum_k^m \phi_{2,k} \mathbf{L}\,\mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\boldsymbol{\alpha}_{h,ik}, \tag{4b}$$

$$\dot{\boldsymbol{\alpha}}_{h,ik} = -\mathbf{L}\,\mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\boldsymbol{\alpha}_{h,ik} - w_{i,k}\boldsymbol{\alpha}_{a,i}, \tag{4c}$$

$$\boldsymbol{\alpha}_{a,i}(T) = \begin{cases} \mathbf{d}_{ij} - \mathbf{b}_a(t_{ij}) & \text{if } t_{ij} = T \\ \mathbf{0} & \text{otherwise} \end{cases}, \tag{4d}$$

where $\xi_{a,i} = \frac{1}{s_i}\sum_j^{s_i}(\mathbf{b}_a(t_{ij}) - \mathbf{d}_{ij})\delta(t - t_{ij})(1 - \delta(t_{ij} - T))$; $\delta(0) = 1, \delta(x) = 0, \forall x \neq 0$. Once we have solve for the adjoints we can compute the gradient as follows. Let $\mathcal{L}$ be the Lagrangian of Equation (3). $\partial_{\boldsymbol{\theta}_s}\mathcal{L}$ is computed as follows: $\frac{\partial \mathcal{L}}{\partial \kappa_i} = -\int \boldsymbol{\alpha}_{a,i}^\intercal \mathbf{L}\mathbf{b}_{a,i} dt$, $\frac{\partial \mathcal{L}}{\partial \rho_i} = \int (\boldsymbol{\alpha}_{n,i} - \boldsymbol{\alpha}_{a,i})^\intercal (\mathbf{b}_{a,i} \odot \mathbf{b}_{n,i}) dt$, $\frac{\partial \mathcal{L}}{\partial \gamma_i} = \int \boldsymbol{\alpha}_{a,i}^\intercal \mathbf{b}_{a,i} dt$, $\frac{\partial \mathcal{L}}{\partial \mathbf{p}_i} = \boldsymbol{\alpha}_{n,i} - \boldsymbol{\alpha}_{a,i} + \frac{\lambda_1}{\mathbf{p}_i - 1}$, and $\frac{\partial \mathcal{L}}{\partial w_{i,k}} = -\int \boldsymbol{\alpha}_{a,i}^\intercal \mathbf{h}_{i,k} dt + \lambda_2 \mathsf{sign}(w_{i,k})$. The gradient for $\boldsymbol{\theta}_c$ involves a reduction across all subjects in the cohort: $\frac{\partial \mathcal{L}}{\partial \phi_{1,k}} = -\sum_i^n \int \boldsymbol{\alpha}_{h,ik}^\intercal \mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\mathbf{L}\mathbf{b}_a dt$, $\frac{\partial \mathcal{L}}{\partial \phi_{2,k}} = -\sum_i^n \int \boldsymbol{\alpha}_{h,ik}^\intercal \mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\mathbf{L}\mathbf{b}_n dt$, and $\frac{\partial \mathcal{L}}{\partial \boldsymbol{\phi}_{3,k}} = -\sum_i^n \int \mathsf{diag}(\sigma'(\mathbf{o}_{ik}))\boldsymbol{\alpha}_{h,ik} dt$.

In summary to compute the gradient, we first solve Equation (1), then solve Equation (4) backwards in time, i.e., in $t \in [T, 0]$; then given forward and adjoint state trajectories we compute $\partial_{\boldsymbol{\theta}}\mathcal{L}$. Parenthetically, the reason we don't use auto-grad is that we have a sparsity constraint $\|\mathbf{p}_i\|_0 = s^{\max}$, which as mentioned is needed to constrain the initial conditions of Aβ.

Numerical Discretization and Inversion Solver: Time discretization of the forward and adjoint ODEs is done using the L-SODA solver [14]. The optimizer is a limited-memory quasi-Newton L-BFGS solver [26]. The regularization parameter λ_1 is chosen to ensure that $\mathbf{p}_i < 1$, and λ_2 is selected to promote sparsity in $\mathbf{w}_i$. For the clinical data experiment, we set $s^{\max} = 10$. The ℓ_0 norm constraint on IC is addressed using the methods described in [17, 23]. We evaluate the model for different values of $m = 0, 1, 2, 3, 4$.

3 Results

In this section, we present results on the performance of LNODE on synthetic and clinical datasets. The purpose of these experiments is to provide answers to the following questions:

(Q1) Can the model accurately reconstruct synthetic observation data?
(Q2) How well does our model fit the ADNI data compared to existing methods?

(Q3) How does the number of latent states affect the performance of LNODE on the ADNI data?
(Q4) How well does LNODE perform on unseen ADNI data?
(Q5) Can we use LNODE-related information to identify population subtypes?

Evaluation metrics: We use two main metrics to evaluate the numerical experiments: relative ℓ_2 and R^2. We define the ℓ_2 relative error for patient i and scan j as $e_{ij} = \frac{\|\mathbf{b}_{a,i}(t_{ij}) - \mathbf{d}_{ij}\|_2}{\|\mathbf{d}_{ij}\|_2}$, with lower values indicating better performance. We use the $R^2 \in [-\infty, 1]$ score to quantify the proportion of variance in the observed data that the model explains. Specifically, we compute two R^2 variants: a per scan-level score R^2_{scan}; and a cohort-level score R^2_{cohort}.

Specifically,

$$R^2_{\text{scan},ij}(\mathbf{d}_{ij}, \mathbf{b}_{a,i}(t_{ij})) = 1 - \frac{\sum_{\ell=1}^{N} \left(\mathbf{d}_{ij}[\ell] - \mathbf{b}_{a,i}[\ell](t_{ij})\right)^2}{\sum_{\ell=1}^{N} \left(\mathbf{d}_{ij}[\ell] - \bar{d}_{ij}\right)^2},$$

where ℓ represents the ℓ^{th} ROI and $\bar{d}_{ij} = \frac{1}{N}\sum_{\ell=1}^{N} \mathbf{d}_{ij}[\ell]$. For R^2_{cohort}, we denote $\mathbf{d}_{\text{avg}} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{s_i}\sum_{j=1}^{s_i} \mathbf{d}_{ij}$, and $\mathbf{b}_{a,\text{avg}} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{s_i}\sum_{j=1}^{s_i} \mathbf{b}_{a,i}(t_{ij})$. Then,

$$R^2_{\text{cohort}}(\mathbf{d}_{\text{avg}}, \mathbf{b}_{a,\text{avg}}) = 1 - \frac{\sum_{\ell=1}^{N} \left(\mathbf{d}_{\text{avg}}[\ell] - \mathbf{b}_{a,\text{avg}}[\ell]\right)^2}{\sum_{\ell=1}^{N} \left(\mathbf{d}_{\text{avg}}[\ell] - \bar{d}_{\text{avg}}\right)^2},$$

where $\bar{d}_{\text{avg}} = \frac{1}{N}\sum_{\ell=1}^{N} \mathbf{d}_{\text{avg}}[\ell]$.

(Q1) Experiment with Synthetic data: We first verify the correctness of our inversion algorithm, its ability of our model to generate distinct subtypes and to reconstruct trajectories, by testing it on synthetic data. To this end, we generate $m = 3$ latent states of n =60 "subjects". For each subject, we choose a single latent state by setting $\mathbf{w}_i$ to be one of $[1, 0, 0]$, $[0, 1, 0]$, or $[0, 0, 1]$. The parameters $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_c$ are randomly set.

Four scans per synthetic subject are used as observation data. We run the forward solver for Equation (1) to generate $\boldsymbol{d}_i$ at times t = [0.8, 0.85, 0.9, 1]. Then we pollute $\boldsymbol{d}_i$ with 10% Gaussian noise. We summarize the results in Figure 2**(A)**. From left to right, the subplots display the true trajectories under *"Ground Truth Trajectories"*, LNODE trajectories by fitting a model without latent states (m=0) under *"No latent states"*, and LNODE trajectories by fitting a model with *"Three latent states"*. The mean relative error $\frac{1}{n}\sum_{i=1}^{n} \frac{1}{s_i}\sum_{j=1}^{s_i} e_{ij}$ for two methods is 11% and 3% respectively. The $m = 3$ reconstruction has errors due to the presence of noise in the data. Near the vertical axis of each subplot we show the average Aβ trajectory for each model (zoom in for full resolution). These results offer a preliminary verification of our formulation and optimization algorithm.

(Q2) ADNI data, evaluation of accuracy: We test our algorithm on the ADNI dataset using Aβ-PET scans from 585 subjects. Among these, 330 are CN, 209 have mild cognitive impairment (MCI), and 46 have AD. The cohort comprises 284 male and 301 female subjects, with a mean age of 71.41 years (SD

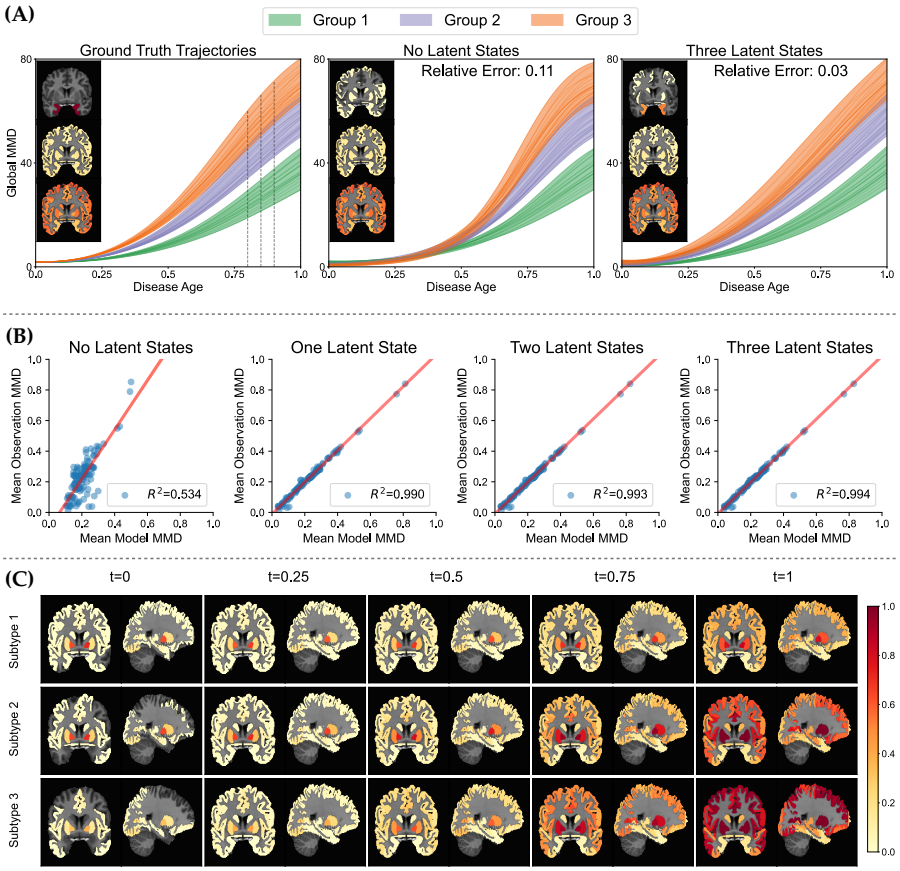

FIG. 2: **(A) Results for synthetically generated data.** *We generate observational data using the proposed forward model and synthesize 60 "subjects" representing three subtypes of disease progression. Two reconstruction methods are applied to the observational data, with the observed disease age indicated by dashed lines. We reconstruct using two LNODE variants, one without latent states ($m = 0$), and one with three latent states ($m=3$). In each subplot, the averaged Aβ trajectory is shown in the upper left corner.* **(B) R^2 Results for ADNI data.** *We evaluate four algorithm variants on the 585 subjects from the ADNI dataset. The group-level R^2_{cohort} score, is displayed in the bottom-right corner of each subplot. In each subplot, every blue dot represents an ROI, while the red line depicts the linear regression fit to these points. Our results demonstrate that our cohort inversion methods ($m = 1, 2, 3$) outperform the $m = 0$ case. We remark that the $m = 0$ represents the current state of the art for mechanism-based ODE models of Aβ progression.* **(C) Observed subtypes in ADNI data.** *We further cluster the inverted $\mathbf{w}_i$ values using the K-means algorithm to identify distinct subtypes of Aβ progression. By examining the elbow in the inertia curve across different numbers of clusters, we determine that the optimal number of clusters is 3. The mean Aβ trajectory for each subtype is then presented.*

$= 6.93$). The data were collected between 2005 and 2024. The Mini-Mental State Examination (MMSE) scores for the CN, MCI, and AD groups are 28.96 ± 0.99, 27.27 ± 2.33, and 22.17 ± 3.64, respectively.

We compare LNODE with an ODE model without latent states [20], originally developed for Tau-PET data. Figure 2**(B)** shows results for the model without latent states, and models with $m = 1, 2, 3$ latent states, respectively. R^2_{cohort} score is displayed in the bottom right corner of each subplot; the average per subject $R^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{s_i}\sum_{j=1}^{s_i} R^2_{\text{scan},ij}$ is 0.369, 0.573, 0.588, and 0.598, respectively.

(Q3) ADNI data, effect of using latent states: We report the mean relative error $\frac{1}{n}\sum_{i=1}^{n}\frac{1}{s_i}\sum_{j=1}^{s_i} e_{ij}$ for the four methods as 0.489, 0.345, 0.332, and 0.320, respectively. In particular, the hidden states significantly increase the performance of the model: In Figure 2**(B)**, we can see that $m = 0$ (No Latent States) case has $R^2_{\text{cohort}} = 0.534$, where the $m = 2$ (Two Latent States) case has $R^2_{\text{cohort}} = 0.993$ a difference of 0.459! And this is not overfitting as the $m = 2$ model uses only 2.4 additional parameters per subject than the $m = 0$ model (having 13 parameters). We also tested the $m = 4$ case but the results remain essentially unchanged.

(Q4) ADNI data, performance on unseen subjects: By unseen subjects, we mean subjects that were not included in the cohort training. For these subjects, we use $\boldsymbol{\theta}_c$ computed by a training cohort and we only invoke the inversion solver for $\boldsymbol{\theta}_s$. We assess the generalizability of the learned cohort-shared parameters, using cross-validation. 80% of the subjects are used to train for $\boldsymbol{\theta}_c$, while the remaining 20% serve as the validation set, where only the subject-specific parameters are updated. We evaluate the model with different numbers of latent states ($m = 1, 2, 3$); the corresponding validation R^2_{cohort} scores are 0.987, 0.988, and 0.991, respectively. Comparing with the results in Figure 2, we only observe a slight reduction in the R^2_{cohort} score.

(Q5) Using LNODE to identify subtypes: Using mechanism-based methodologies to identify subtypes was introduced in [21] for Tau-PET (did not report numbers for Aβ). We attempt something similar with LNODE. First, we can check the sparsity of $\mathbf{w}_i$ for each subject i, to see how many states are active. To this end, we compute the ratio $\frac{\|\mathbf{w}_i\|_\infty}{\|\mathbf{w}_i\|_2}$ (higher means sparser $\mathbf{w}$): for $m = 3$. 94% of the subjects exhibit a ratio greater than 0.7. Second, we cluster $\{\mathbf{w}_i\}_{i=1}^n$ using k-means to identify distinct subtypes of Aβ progression [11]. Figure 2**(C)** displays the mean Aβ trajectory for each subtype.

4 Conclusions

LNODE introduces a novel approach by incorporating ODEs for latent states, dividing model parameters into subject-specific and cohort-shared categories, and using sparse coupling to facilitate subtype detection. The preliminary results are promising, showing a 45.9% improvement in R^2_{cohort} with an average increase of only 2.4 parameters per subject. Our next steps include evaluating LNODE on a larger cohort, testing different models for latent-state dynamics, gaining a better understanding of the clinical characteristics of the identified subtypes, and extending this to Tau-PET and other observables in ADNI subjects.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chung, F.R.: Spectral graph theory, vol. 92. American Mathematical Soc. (1997)
2. Doshi, J., Erus, G., Ou, Y., Resnick, S.M., Gur, R.C., Gur, R.E., Satterthwaite, T.D., Furth, S., Davatzikos, C., Initiative, A.N., et al.: Muse: Multi-atlas region segmentation utilizing ensembles of registration algorithms and parameters, and locally optimal atlas selection. Neuroimage **127**, 186–195 (2016)
3. Fan, S., Ponisio, M.R., Xiao, P., Ha, S.M., Chakrabarty, S., Lee, J.J., Flores, S., LaMontagne, P., Gordon, B., Raji, C.A., et al.: Amyloidpetnet: classification of amyloid positivity in brain pet imaging using end-to-end deep learning. Radiology **311**(3), e231442 (2024)
4. Ghattas, O., Willcox, K.: Learning physics-based models from data: perspectives from inverse problems and model reduction. Acta Numerica **30**, 445–554 (2021)
5. Ghazi, M.M., Nielsen, M., Pai, A., Modat, M., Cardoso, M.J., Ourselin, S., Sørensen, L.: Robust parametric modeling of alzheimer's disease progression. Neuroimage **225**, 117460 (2021)
6. Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B., Smola, A.: A kernel two-sample test. The Journal of Machine Learning Research **13**(1), 723–773 (2012)
7. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints, vol. 23. Springer Science & Business Media (2008)
8. Iturria-Medina, Y., Sotero, R.C., Toussaint, P.J., Evans, A.C., Initiative, A.D.N.: Epidemic spreading model to characterize misfolded proteins propagation in aging and associated neurodegenerative disorders. PLoS computational biology **10**(11), e1003956 (2014)
9. Kim, S.K., Duong, Q.A., Gahm, J.K.: Multimodal 3d deep learning for early diagnosis of alzheimer's disease. IEEE Access (2024)
10. Lee, M.W., Kim, H.W., Choe, Y.S., Yang, H.S., Lee, J., Lee, H., Yong, J.H., Kim, D., Lee, M., Kang, D.W., et al.: A multimodal machine learning model for predicting dementia conversion in alzheimer's disease. Scientific Reports **14**(1), 12276 (2024)
11. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. vol. 5, pp. 281–298. University of California press (1967)
12. Ossenkoppele, R., Pichet Binette, A., Groot, C., Smith, R., Strandberg, O., Palmqvist, S., Stomrud, E., Tideman, P., Ohlsson, T., Jögi, J., et al.: Amyloid and tau pet-positive cognitively unimpaired individuals are at high risk for future cognitive decline. Nature medicine **28**(11), 2381–2387 (2022)
13. Petersen, R.C., Aisen, P.S., Beckett, L.A., Donohue, M.C., Gamst, A.C., Harvey, D.J., Jack Jr, C., Jagust, W.J., Shaw, L.M., Toga, A.W., et al.: Alzheimer's disease neuroimaging initiative (adni) clinical characterization. Neurology **74**(3), 201–209 (2010)
14. Petzold, L.: Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. SIAM journal on scientific and statistical computing **4**(1), 136–148 (1983)
15. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al.:

Advances in functional and structural mr image analysis and implementation as fsl. Neuroimage **23**, S208–S219 (2004)
16. Stockman, C.A., Goriely, A., Kuhl, E., Initiative, A.D.N., et al.: Two for tau: Automated model discovery reveals two-stage tau aggregation dynamics in alzheimer's disease. Brain Multiphysics **7**, 100103 (2024)
17. Tropp, J.A., Wright, S.J.: Computational methods for sparse solution of linear inverse problems. Proceedings of the IEEE **98**(6), 948–958 (2010)
18. Villemagne, V.L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K.A., Salvado, O., Szoeke, C., Macaulay, S.L., Martins, R., Maruff, P., et al.: Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic alzheimer's disease: a prospective cohort study. The Lancet Neurology **12**(4), 357–367 (2013)
19. Villemagne, V.L., Pike, K.E., Chételat, G., Ellis, K.A., Mulligan, R.S., Bourgeat, P., Ackermann, U., Jones, G., Szoeke, C., Salvado, O., et al.: Longitudinal assessment of aβ and cognition in aging and alzheimer disease. Annals of neurology **69**(1), 181–192 (2011)
20. Vogel, J.W., Iturria-Medina, Y., Strandberg, O.T., Smith, R., Levitis, E., Evans, A.C., Hansson, O.: Spread of pathological tau proteins through communicating neurons in human alzheimer's disease. Nature communications **11**(1), 2612 (2020)
21. Vogel, J.W., Young, A.L., Oxtoby, N.P., Smith, R., Ossenkoppele, R., Strandberg, O.T., La Joie, R., Aksman, L.M., Grothe, M.J., Iturria-Medina, Y., et al.: Four distinct trajectories of tau deposition identified in alzheimer's disease. Nature medicine **27**(5), 871–881 (2021)
22. Wen, Z., Ghafouri, A., Biros, G.: Biophysics-based data assimilation of longitudinal tau and amyloid-β pet scans. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 14–24. Springer (2024)
23. Wen, Z., Ghafouri, A., Biros, G.: A single-snapshot inverse solver for two-species graph model of tau pathology spreading in human alzheimer disease. arXiv preprint arXiv:2402.06880 (2024)
24. Young, A.L., Oxtoby, N.P., Garbarino, S., Fox, N.C., Barkhof, F., Schott, J.M., Alexander, D.C.: Data-driven modelling of neurodegenerative disease progression: thinking outside the black box. Nature Reviews Neuroscience **25**(2), 111–130 (2024)
25. Zhang, Z., Zou, Z., Kuhl, E., Karniadakis, G.E.: Discovering a reaction–diffusion model for alzheimer's disease by combining pinns with symbolic regression. Computer Methods in Applied Mechanics and Engineering **419**, 116647 (2024)
26. Zhu, C., Byrd, R.H., Lu, P., Nocedal, J.: Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. ACM Transactions on mathematical software (TOMS) **23**(4), 550–560 (1997)
27. Zou, J., Park, D., Johnson, A., Feng, X., Pardo, M., France, J., Tomljanovic, Z., Brickman, A.M., Devanand, D.P., Luchsinger, J.A., et al.: Deep learning improves utility of tau pet in the study of alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring **13**(1), e12264 (2021)