Published in final edited form as:

Med Image Comput Comput Assist Interv. 2020 October ; 12267: 418–427.

doi:10.1007/978-3-030-59728-3_41.

A physics-informed geometric learning model for pathological tau spread in Alzheimer's disease

Tzu-An Song^{**},^{1,2}, Samadrita Roy Chowdhury^{**},^{1,2}, Fan Yang^{1,2}, Heidi I. L. Jacobs², Jorge Sepulcre², Van J. Wedeen², Keith A. Johnson², Joyita Dutta^{1,2}

¹University of Massachusetts Lowell, Lowell, MA, USA

²Massachusetts General Hospital & Harvard Medical School, Boston, MA, USA

Abstract

Tau tangles are a pathophysiological hallmark of Alzheimer's disease (AD) and exhibit a stereotypical pattern of spatiotemporal spread which has strong links to disease progression and cognitive decline. Preclinical evidence suggests that tau spread depends on neuronal connectivity rather than physical proximity between different brain regions. Here, we present a novel physics-informed geometric learning model for predicting tau buildup and spread that learns patterns directly from longitudinal tau imaging data while receiving guidance from governing physical principles. Implemented as a graph neural network with physics-based regularization in latent space, the model enables effective training with smaller data sizes. For training and validation of the model, we used longitudinal tau measures from positron emission tomography (PET) and structural connectivity graphs from diffusion tensor imaging (DTI) from the Harvard Aging Brain Study. The model led to higher peak signal-to-noise ratio and lower mean squared error levels than both an unregularized graph neural network and a differential equation solver. The method was validated using both two-timepoint and three-timepoint tau PET measures. The effectiveness of the approach was further confirmed by a cross-validation study.

Keywords

Alzheimer's disease; tau spread; graph neural networks; PET; DTI

1 Introduction

Alzheimer's disease (AD) is a debilitating neurodegenerative disorder and a looming public health challenge. Extracellular amyloid- β (A β) plaques and intracellular tau tangles, the two hallmark pathologies of AD, are believed to play key mechanistic roles in AD [1]. Studies show that tau pathology in the medial temporal lobe (MTL) is a key driver of memory impairment in AD and is an important marker for neurodegeneration [6,14,16]. It is known that the spatiotemporal spread of tau tangles follows a stereotypical trajectory starting in the locus coeruleus and the transentorhinal cortex and then extending to the entorhinal cortex, the hippocampus, and finally the neocortex [3, 5]. A growing body of evidence indicates that

dutta.joyita@mgh.harvard.edu.

^{**}these authors contributed equally

tau spreads through the brain in a prion-like fashion with neurons carrying pathological tau species transmitting them to their connected neighboring neurons via anatomical or synaptic connections [8, 18]. Comprehension of neurodegenerative pathogenesis requires the understanding of the mechanisms of proliferation and accumulation of tau. Network diffusion [19] and epidemic spread [22] models have had reasonable success at explaining the relationship between structural and functional variables in the human brain. These methods have enabled group-level inference of the sources of atrophy [10, 21, 23]. Yet quantitative, patient-tailored, predictive methods based on these models that use learned population-level information remain unexplored.

In recent years, deep neural networks have been successfully used to solve real-world problems relating to physics-based natural phenomena [9, 17]. A novel subcategory of this research uses physics-informed neural networks to explain naturally occurring dynamic processes [15, 20]. While the principles of physics may correctly model some facets of real-life datasets, data-driven learning can fill the gaps between the known physics and real observations. For tau propagation, source modeling is an open problem, and there tends to be a high degree of subject-to-subject variability when group-level fitting of physics-based models is performed. Here, we present a novel physics-informed geometric learning model for predicting tau spread that (i) is capable of learning patterns of tau buildup not explained by passive diffusion alone, (ii) allows us to incorporate additional variables (e.g. scalar connectivity and diffusivity measures and potentially $A\beta$) which may influence the complex seeding processes for tau, and (iii) could offer robustness against local inaccuracies in the structural connectivity measures (e.g. imperfections of tractography) and local departures from physics guidance. Regularization would also help prevent overfitting, reduce generalization error, and thus ensure effective training with smaller datasets compared to a purely data-driven approach.

In this work, we use longitudinal ^{18}F -flortaucipir tau positron emission tomography (PET) scans for obtaining regional tau measures [12]. White matter fiber bundles generated via diffusion tensor imaging (DTI) are used to map the structural connectome for each subject. In section 2, we describe the geometric learning and physics models. Details on data processing and experiments are provided in section 3. Our main findings are reported in section 4. In section 5, we summarize this work and present our envisioned future directions.

2 Theory

2.1 Fisher-KPP Model for Tau Propagation and Generation

The structural network of the brain can be represented as a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $v_i \in \mathcal{V}$ is the i th node representing a gray-matter parcellation or anatomical region of interest (ROI), $|\mathcal{V}| = N$ is the number of nodes, $e_{ij} \in \mathcal{E}$ is the edge connecting the i th and j th nodes, and $\mathbf{W} \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix. Tau aggregation at a given timepoint can be represented as a graph signal $\mathbf{x}(t) \in \mathbb{R}^N$, the spatiotemporal evolution of which can be modeled as a diffusion process by the inhomogeneous partial differential equation (PDE):

$$\frac{\partial \mathbf{x}(t)}{\partial t} = -\beta \mathbf{L} \mathbf{x}(t) + s(t, \mathbf{x}(t)). \quad (1)$$

Here $s(t, \mathbf{x}(t)) \in \mathbb{R}^N$ is a source term modeling tau generation or clearance processes not explained by node-to-node diffusion, and $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}}$ is the normalized graph Laplacian matrix ($\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix and $\mathbf{D} \in \mathbb{R}^{N \times N}$ is the weighted degree matrix computed from $\mathbf{W}$).

As part of our prior work, we have used impulse-based representations for $s(t, \mathbf{x}(t))$ which led to simple, closed-form, analytic solutions [23]. Such source models do not account for the dependence of tau seeding on the current concentration of tau. In this work, we model tau seeding and spread using the Fisher-Kolmogorov-Petrovski-Puskinovand (Fisher-KPP) equation [7, 13], which is a *reaction-diffusion equation* and has the following source term:

$$s(\mathbf{x}(t)) = \mathbf{x}(t) \odot (1 - \mathbf{x}(t) \oslash \kappa). \quad (2)$$

Here the κ is the carrying capacity, $\odot$ represents the entrywise Hadamard multiplication, and $\oslash$ represents Hadamard division. Unlike source terms that only rely on time t , the semilinear Fisher-KPP source depends on the tau burden $\mathbf{x}(t)$. Although usually not solvable in closed form, such a model is physiologically meaningful since pathological tau species act as seeds or templates promoting further aggregation and spread from cell to cell in a prion-like fashion. Here, we will use the Fisher-KPP model to incorporate physics knowledge into a geometric learning framework. A numerical solution of the Fisher-KPP equation will serve as one of the reference approaches, the other being a standalone geometric learning framework without physics-based guidance.

2.2 Graph Neural Network

Graph Neural Network Architecture—The geometric learning model underlying this study is a graph neural network (GNN) based on combinatorial generalization among nodes, edges, and the entire graph [2]. Such models have been previously utilized to learn complex physical interactions and infer trajectories of dynamical systems from the system's current state. As shown in Fig. 1, the GNN comprises three blocks: an encoder, a core recurrent neural network (RNN), and a decoder. The encoder maps the input graph signal to a latent space, where the physics-based constraints are enforced. The RNN block is a rendition of a nonlocal neural network with update and aggregation functions. The decoder maps the latent space representation back to the observation domain to generate the final output. The GNN is set up to learn residuals in the form of mean annualized differential standardized uptake value ratios (ADSUVRs) for each ROI.

Physics-Informed Loss Function—A complex phenomenon such as tau spread is not governed by single-variable dynamics. Instead, it involves multi-variable interactions that are not easily explained by known physical models and need to be learned from the data. A latent space representation could lead to a learned composite variable dependent not only on

the tau burden but also on a variety of other inputs, e.g., scalar connectivity and diffusivity measures and potentially $A\beta$. We, therefore, apply the physics-based constraint in the latent space instead of the observation space. This form of physics guidance is effectively a regularizer that stabilizes the underlying inverse problem. The physics-informed loss function uses the Fisher-KPP model in (1) discretized in time and formulated as a difference equation:

$$\frac{\lambda_{n,i} - \lambda_{n-1,i}}{t_n - t_{n-1}} = \sum_j -\beta L_{ij} \lambda_{n,j} + r \lambda_{n-1,i} (1 - \lambda_{n-1,i} / \kappa_i). \quad (3)$$

Here $\lambda_{n,i}$ is a latent space variable, i and j are node indices, n is a time index, t_n is the n th timepoint, L_{ij} is (i, j) th element of the Laplacian, and β and r are hyperparameters. The net loss function, J , comprises an observation-domain L_2 -norm data-fit term and a latent space L_2 -norm penalty term:

$$J = \sum_{n=1}^T \|\mathbf{x}_n - \hat{\mathbf{x}}_n\|^2 + \alpha \sum_{n=1}^{T-(M-1)} \|\lambda_n - \lambda_{n-1} + \beta L \lambda_{n-1} - r \lambda_{n-1} \odot (1 - \lambda_{n-1} \oslash \kappa)\|^2. \quad (4)$$

Here $\mathbf{x}_n$ and $\hat{\mathbf{x}}_n$ are the observed and predicted graph signal vectors representing tau burden in the observation domain, λ_n is the latent space vector, and α is a regularization parameter. The data-fit loss is updated T times with T at least as large as the order of temporal derivative. The physics-based loss is updated $T - (M - 1)$ times, where $M - 1$ is the number of required predictive steps.

3 Methods

3.1 Data Description and Preprocessing

All experiments relied on data from the Harvard Aging Brain Study (HABS) [4], which is an ongoing longitudinal study aimed at revealing the differences between normal aging and preclinical AD. PET and magnetic resonance (MR) imaging data from $N = 70$ human subjects (75.66 ± 6.11 years, 39 females) were used in this study.

Tau PET—Serial tau PET images at 0.6–3.6 years gap between consecutive scans were used to compute regional tau burden. Following the injection of 10 mCi ^{18}F -flortaucipir, PET data were acquired for 30 min using a 3D list-mode dynamic protocol on a Siemens ECAT HR+ scanner. One subset of the data comprised serial tau PET scans at only two timepoints ($N = 60$), henceforth referred to as the two-timepoint or 2TP cohort. A second subset comprised PET scans at three timepoints ($N = 10$), henceforth referred to as the three-timepoint or 3TP cohort. From the PET images, we calculate mean standardized uptake value ratios (SUVRs) for 85 anatomical ROIs based on the FreeSurfer Desikan-Killiany atlas with cerebellar gray matter as the reference.

Diffusion MR—The diffusion-weighted scans were acquired using a spin-echo echo-planar imaging sequence: echo time (TE) 84 ms, repetition time (TR) 8,040 ms, field-of-view $256 \times 256 \times 128$, and voxel size 2 mm isotropic with 30 isotropically distributed orientations for the diffusion-sensitizing gradients at a b-value of 700 s/mm^2 . Diffusion data were preprocessed using FSL to perform corrections relating to subject motion, eddy current distortion, and susceptibility. Diffusion tensors were reconstructed using DSI Studio. Fractional anisotropy (FA) and median diffusivity (MD) measures were computed from the reconstructed tensors. Deterministic fiber tracking via DSI Studio was performed to obtain the number of streamlines between various brain regions. For streamline tracking, we used an angular threshold of 45° and retained default values for all other parameters. The tracts were normalized by length, and only tracts ending in 85 gray-matter ROIs from the Desikan-Killiany atlas were used. The reconstructed streamlines were counted for each pair of ROIs to compute pairwise inter-ROI connection strengths. This led to a series of 85×85 adjacency matrices capturing individualized structural connectivity profiles for each human subject.

3.2 Experiments

Performance Comparison and Evaluation Metrics—The physics-informed GNN (GNN-P) was compared with an unregularized GNN and a numerical ordinary differential equation (ODE) solver. As quantitative figures-of-merit, we compute the mean squared error (MSE) and peak signal-to-noise ratio (PSNR).

Network Training—GNN and GNN-P implementation was done on the TensorFlow suite using the DeepMind Graph Nets library [2]. Without loss of generality, the carrying capacity of the governing Fisher-KPP equation was set to 1 in the latent space of the graph. The GNN and GNN-P networks were provided the following inputs as graph signal vectors: baseline tau signal vector ($\mathbf{x}_1$) and FA and MD values computed from diffusion MR. The graph adjacency matrices (represented as a list of edge weights) were provided as additional inputs to the networks. GNN and GNN-P training was performed using the 2TP cohort: $N_{\text{train}} = 54$. A batch size of 6 and 1000 epochs were used for network training. The regularization parameter was set to $\alpha = 0.00002$. The hyperparameter values were based on Fisher-KPP literature: $\beta = 0.035$ and $r = 0.0005$.

Network Validation—For the 2TP cohort, the goal was to compute the tau signal vector at follow-up ($\mathbf{x}_2$) from the baseline tau ($\mathbf{x}_1$), FA values, and MD values given the individual structural connectivity profiles. The trained GNN and GNN-P models ($N_{\text{train}} = 54$, 2TP cohort) was independently validated using $N_{\text{test-2TP}} = 6$ distinct samples from the same cohort. This was followed by 10-fold cross-validation. The 3TP cohort was used for validation alone: $N_{\text{test-3TP}} = 10$. To ensure longer-term prediction capability of downstream tau aggregation from a baseline scan, the core module of the GNN was implemented as an RNN. Currently, our data is limited to a maximum of three timepoints. As more data points become available in the HABS cohort, the RNN capabilities can be fully exploited. For the 3TP cohort, the model was iterated over three timepoints to predict tau at t_2 and t_3 from tau at t_1 .

4 Results

A comparison of prediction results in two subjects based on the three methods (ODE, GNN, and GNN-P) is shown in Fig. 2. The figure shows coronal slices in MNI152 space with observed and predicted mean ROI tau SUVRs and ADSUVRs overlaid on an anatomical template. The ROIs displayed in this slice are the entorhinal cortex, fusiform gyrus, inferior temporal cortex, middle temporal cortex, superior temporal cortex, amygdala, and posterior cingulate cortex, all of which are considered critical for the assessment of tau burden in early AD [11]. Tau accumulation is a slow process. Typically the regional tau profiles at consecutive timepoints of serial tau PET scans exhibit a high degree of correlation. For both visual and quantitative assessment, we, therefore, rely on the accuracy of predicted ADSUVRs computed from $\hat{x}_2 - x_1$, instead of $\hat{x}_2$. Rows 2 and 4 of Fig. 2 suggest a higher degree of agreement between the observed ADSUVR and the GNN-P prediction relative to both reference approaches. This was true for both subjects with globally increasing (Fig. 2A) and non-increasing (Fig. 2B) cortical tau. Table 1 shows the predictive accuracy as captured by the mean and standard deviation of the PSNR and MSE values in the validation group of the 2TP cohort for all three methods. To assess the robustness and reproducibility of the results, a 10-fold cross-validation study was performed. The mean, standard deviation, maximum, and minimum values for the PSNR are provided in Table 2.

A comparison of predicted and observed ADSUVRs in one subject from the 3TP cohort is depicted in Fig. 3. Validation in the 3TP cohort was based on the model trained using the 2TP cohort data. GNN-P clearly outperforms both GNN and ODE in this even more challenging scenario thereby demonstrating its strength at making longer-term predictions.

5 Conclusion

In this work, we developed and validated a physics-informed geometric learning framework for predicting the spatiotemporal trajectory of misfolded tau protein along the structural network of the brain. We used the Fisher-KPP reaction-diffusion equation as the governing physics model and incorporated physics guidance through a regularization penalty in the latent space generated by an encoder-decoder GNN architecture. To demonstrate proof of concept, we evaluated predictive accuracy of the physics-informed GNN-P network using 2TP and 3TP cohorts. We used MSE and PSNR as figures-of-merit and a numerical ODE solver and an unregularized GNN as benchmarks. We demonstrated that GNN-P is both qualitatively and quantitatively superior to both reference approaches. The improved prediction accuracy with GNN-P extended to both subjects with increasing and decreasing global tau burden. GNN-P was especially robust at the longer-term prediction task of computing the tau signal vector at timepoint t_3 from baseline data at t_1 . Such longitudinal tracking of tau differentials is of great significance in AD prognosis. As part of our future work, we will incorporate regional A β profiles as additional GNN inputs. As more longitudinal data become available, we will increase the training sample size so as to boost the robustness of the GNN and GNN-P models. With the availability of tau PET scans at more timepoints, we will be able to test the prediction capability of the models over even longer observation windows.

Acknowledgments

Supported by the National Institute on Aging grant K01AG050711

References

1. Arriagada PV, Growdon JH, Hedley-Whyte ET, Hyman BT: Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology* 42(3 Pt 1), 631–639 (3 1992) [PubMed: 1549228]
2. Battaglia PW, Hamrick JB, Bapst V, Sanchez-Gonzalez A, Zambaldi V, Malinowski M, Tacchetti A, Raposo D, Santoro A, Faulkner R, Gulcehre C, Song F, Ballard A, Gilmer J, Dahl G, Vaswani A, Allen K, Nash C, Langston V, Dyer C, Heess N, Wierstra D, Kohli P, Botvinick M, Vinyals O, Li Y, Pascanu R: Relational inductive biases, deep learning, and graph networks. *arXiv:1806.01261* (10 2018)
3. Braak H, Braak E: Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica* 82(4), 239–259 (9 1991) [PubMed: 1759558]
4. Dagley A, LaPoint M, Huijbers W, Hedden T, McLaren DG, Chatwal JP, Papp KV, Amariglio RE, Blacker D, Rentz DM, Johnson KA, Sperling RA, Schultz AP: Harvard Aging Brain Study: Dataset and accessibility. *Neuroimage* 144(Pt B), 255–258 (1 2017) [PubMed: 25843019]
5. Delacourte A, Sergeant N, Wattez A, Maurage CA, Lebert F, Pasquier F, David JP: Tau aggregation in the hippocampal formation: an ageing or a pathological process? *Experimental Gerontology* 37(10-11), 1291–1296 (11 2002) [PubMed: 12470843]
6. Dickson DW, Crystal HA, Mattiace LA, Masur DM, Blau AD, Davies P, Yen SH, Aronson MK: Identification of normal and pathological aging in prospectively studied nondemented elderly humans. *Neurobiology of Aging* 13(1), 179–189 (Jan-Feb 1992) [PubMed: 1311804]
7. Fisher RA: The wave of advantageous genes. *Annals of Eugenics* 7(4), 355–369 (6 1937)
8. Frost B, Diamond MI: Prion-like mechanisms in neurodegenerative diseases. *Nature Reviews Neuroscience* 11(3), 155–159 (3 2010) [PubMed: 20029438]
9. Hashemi SMH, Psaltis D: Deep-learning PDEs with unlabeled data and hardwiring physics laws. *arXiv:1904.06578* (4 2019)
10. Hu C, Hua X, Ying J, Thompson PM, Fakhri GE, Li Q: Localizing sources of brain disease progression with network diffusion model. *IEEE Journal of Selected Topics in Signal Processing* 10(7), 1214–1225 (10 2016) [PubMed: 28503250]
11. Jack CR, Wiste HJ, Schwarz CG, Lowe VJ, Senjem ML, Vemuri P, Weigand SD, Therneau TM, Knopman DS, Gunter JL, Jones DT, Graff-Radford J, Kantarci K, Roberts RO, Mielke MM, Machulda MM, Petersen RC: Longitudinal tau PET in ageing and Alzheimer's disease. *Brain* 141(5), 1517–1528 (5 2018) [PubMed: 29538647]
12. Johnson KA, Schultz A, Betensky RA, Becker JA, Sepulcre J, Rentz D, Mormino E, Chhatwal J, Amariglio R, Papp K, Marshall G, Albers M, Mauro S, Pepin L, Alverio J, Judge K, Philiossaint M, Shoup T, Yokell D, Dickerson B, Gomez-Isla T, Hyman B, Vasdev N, Sperling R: Tau positron emission tomographic imaging in aging and early Alzheimer disease. *Ann. Neurol* 79(1), 110–119 (1 2016) [PubMed: 26505746]
13. Kolmogorov AN, Tikhomirov VM, Kolmogorov AN: Mathematics and mechanics. No. v. 1 in *Selected works of A.N. Kolmogorov*, Kluwer Academic Publishers, Dordrecht ; Boston (1991)
14. McKee AC, Kosik KS, Kowall NW: Neuritic pathology and dementia in Alzheimer's disease. *Annals of Neurology* 30(2), 156–165 (8 1991) [PubMed: 1910274]
15. Nabian MA, Meidani H: Physics-driven regularization of deep neural networks for enhanced engineering design and analysis. *Journal of Computing and Information Science in Engineering* 20(1), 011006 (2 2020)
16. Neve RL, Robakis NK: Alzheimer's disease: a re-examination of the amyloid hypothesis. *Trends in Neurosciences* 21(1), 15–19 (1 1998) [PubMed: 9464679]
17. Noé F, Olsson S, Köhler J, Wu H: Boltzmann generators – sampling equilibrium states of many-body systems with deep learning. *arXiv:1812.01729* (7 2019)

18. Nussbaum JM, Seward ME, Bloom GS: Alzheimer disease: a tale of two prions. *Prion* 7(1), 14–19 (2 2013) [PubMed: 22965142]
19. Raj A, Kuceyeski A, Weiner M: A network diffusion model of disease progression in dementia. *Neuron* 73(6), 1204–1215 (3 2012) [PubMed: 22445347]
20. Seo S, Liu Y: Differentiable Physics-informed Graph Networks. *arXiv:1902.02950* (2 2019)
21. Torok J, Maia PD, Powell F, Pandya S, Raj A: A method for inferring regional origins of neurodegeneration. *Brain* 141(3), 863–876 (3 2018) [PubMed: 29409009]
22. Vogel JW, Iturria-Medina Y, Strandberg OT, Smith R, Levitis E, Evans AC, Hansson O, Alzheimer's Disease Neuroimaging Initiative, Swedish BioFinder Study: Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nature Communications* 11(1), 1–15 (2020)
23. Yang F, Roy Chowdhury S, Jacobs HIL, Johnson KA, Dutta J: A longitudinal model for tau aggregation in Alzheimer's disease based on structural connectivity. *Information Processing in Medical Imaging* 11492, 384–393 (5 2019) [PubMed: 31156312]

Fig. 1.
Physics-regularized GNN architecture

Fig. 2.

Prediction results from the 2TP cohort using ODE, GNN, and GNN-P for two archetypal subjects with (A) increasing and (B) non-increasing global tau burden. Mean SUVRs and ADSUVRs in several tau-critical ROIs are plotted on coronal slices in MNI152 space and overlaid on an anatomical MR template. Rows 2 and 4 show a comparison of the ADSUVRs between t_1 and t_2 .

Fig. 3.

Prediction results in one subject from the 3TP cohort using ODE, GNN, and GNN-P. Mean ADSUVRs in several tau-critical ROIs are plotted on coronal slices in MNI152 space and overlaid on an anatomical MR template. The model trained on the 2TP cohort was used to predict tau burden in the 3TP cohort at timepoint t_3 from the baseline scan at t_1 . The MSEs of the residual for this subject are as follows: ODE 0.8174, GNN 0.8758, and GNN-P 0.6216. These prediction results are demonstrative of the longer-term prediction capabilities of GNN-P.

Table 1.

Performance comparison in the 2TP cohort

Metric	ODE	GNN	GNN-P
PSNR (standard deviation)	15.6222 (1.6367)	18.8797 (3.1953)	19.0485 (2.2301)
MSE (standard deviation)	0.0687 (0.0200)	0.0348 (0.0153)	0.0314 (0.0086)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Cross-validation results

Method	Mean PSNR	PSNR standard deviation	Maximum PSNR	Minimum PSNR
GNN	17.2074	0.9646	19.5943	16.2967
GNN-P	18.7978	0.9571	20.3015	17.3277

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript