Title

Applying Joint Graph Embedding to Study Alzheimer's Neurodegeneration Patterns in Volumetric Data

Permalink

https://escholarship.org/uc/item/9pr3x75z

Journal

Neuroinformatics, 21(3)

ISSN

1539-2791

Authors

He, Rosemary

Tward, Daniel

Publication Date

2023-07-01

DOI

10.1007/s12021-023-09634-6

Peer reviewed

RESEARCH

Applying Joint Graph Embedding to Study Alzheimer's Neurodegeneration Patterns in Volumetric Data

Rosemary He[1] · Daniel Tward[2,3] · Alzheimer's Disease Neuroimaging Initiative

Abstract

Neurodegeneration measured through volumetry in MRI is recognized as a potential Alzheimer's Disease (AD) biomarker, but its utility is limited by lack of specificity. Quantifying spatial patterns of neurodegeneration on a whole brain scale rather than locally may help address this. In this work, we turn to network based analyses and extend a graph embedding algorithm to study morphometric connectivity from volume-change correlations measured with structural MRI on the timescale of years. We model our data with the multiple random eigengraphs framework, as well as modify and implement a multigraph embedding algorithm proposed earlier to estimate a low dimensional embedding of the networks. Our version of the algorithm guarantees meaningful finite-sample results and estimates maximum likelihood edge probabilities from population-specific network modes and subject-specific loadings. Furthermore, we propose and implement a novel statistical testing procedure to analyze group differences after accounting for confounders and locate significant structures during AD neurodegeneration. Family-wise error rate is controlled at 5% using permutation testing on the maximum statistic. We show that results from our analysis reveal networks dominated by known structures associated to AD neurodegeneration, indicating the framework has promise for studying AD. Furthermore, we find network-structure tuples that are not found with traditional methods in the field.

Keywords Alzheimer's disease · Structural MRI · Graph embedding · Network analysis · Familywise error rate control

Alzheimer's Disease Neuroimaging Initiative (ADNI) is a Group/Institutional Author.

✉ Daniel Tward
dtward@mednet.ucla.edu

[1] Departments of Computer Science and Computational Medicine, University of California, Los Angeles, USA

[2] Departments of Computational Medicine and Neurology, University of California, Los Angeles, USA

[3] Neuroscience Research Building (NRB) 635 Charles E Young Drive South, Rm 225J, Los Angeles, CA 90095, USA

Introduction

Alzheimer's disease (AD) is a progressive mental disorder associated with neurodegeneration that generally occurs in old ages. It is one of the most common diseases in seniors, killing more than breast cancer and prostate cancer combined (Association, 2019). However, AD can only be formally diagnosed through an autopsy after a patient is deceased, encouraging the research of alternative proxies for AD diagnosis. There are currently three classes of potential biomarkers that could offer useful alternatives to assess AD diagnosis: $\beta-$amyloid(A), tau(T) and biomarkers for neurodegeneration or neuronal injury(N) (Jack et al., 2016), with the last class being the focus of this paper. Neurodegeneration can be measured noninvasively, such as through structural MRI, but is not specific to AD as it can reflect other diseases (Jack & Holtzman, 2013). For example, atrophy is a biomarker that is often associated with AD, but also occurs in a variety of disorders such as epilepsy and anoxia (Jack & Holtzman, 2013). The lack of specificity of these neurodegeneration biomarkers poses a major limitation in their utility for early-stage clinical diagnosis of AD. A

promising approach to improving specificity is to consider patterns of volumetric changes across the whole brain, rather than focusing on a small number of regions. For example, in discussing the diagnosis of Mild Cognitive Imparment due to AD, Albert et al. suggests the possibility of biomarkers describing "complex patterns of tissue loss" through "data driven statistical approaches in which many different brain regions are evaluated simultaneously" (Albert et al., 2011).

While rare in volumetric analysis, such patterns have been studied extensively in functional brain imaging. One popular approach to studying brain connectomics is to extract information about interactions between volumetric pixels (voxels) from time-series functional magnetic resonance imaging (fMRI) data collected over a timescale of weeks, months or years (Cohen et al., 2017). Functional connectivity, in particular, studies temporal dependencies among anatomically separated regions (Van Den Heuvel & Pol, 2010). There are several approaches to study functional connectivity in fMRI data. As our interests lie within the whole brain instead of a single voxel, we discuss only multivoxel pattern analysis methods that study networks as a whole (Lewis-Peacock & Norman, 2014). A traditional method in the field is seed-based analysis, in which a region of interest (ROI) is selected, and all voxels correlated to the ROI is identified (Cole et al., 2010). For example, in an earlier work Biswal et al. studied the motor cortex to identify the sensorimotor network (Biswal et al., 1995). Unsupervised clustering methods including k-means, hierarchical and graph-based methods do not require a priori ROI and group voxels together by their similarities in time series data (Khosla et al., 2019). In their paper, Lee et al. uses fuzzy-c-means clustering to identify resting state networks (Lee et al., 2012).

Over the years, graph-based approaches have gained popularity in studying functional connectivity. To convert a brain into a graph, the regions are modeled as nodes and connections between regions as edges. Under this model, one can construct a matrix of all pairs of connections in the brain, known as the functional connectome (Fornito et al., 2016), of which decomposition or embedding methods can be applied to uncover latent variables. Independent component analysis decomposes data into linearly independent components, grouping brain regions into networks based on their voxel activation correlations (Calhoun et al., 2001). Non-negative matrix factorization is a dimensionality reduction method that forces non-negativity constraints on the components (Khosla et al., 2019). Some popular embedding methods include Adjacency Spectral Embedding (ASE) (Sussman et al., 2012), which embeds a single symmetric adjacency matrix using eigenvectors corresponding to the largest eigenvalues, and Laplacian Eigenmap (LE) (Belkin & Niyogi, 2003), which embeds a single graph-Laplacian matrix using its eigenvectors corresponding to the smallest nonzero eigenvalues.

However, several limitations lie within these graph-based approaches. First, they embed one graph at a time, and combining individual embeddings across multiple graphs is not a straighforward task. Second, the results are difficult to interpret, and further analysis is required (Yang et al., 2020).

One technique for embedding multiple graphs at once is omnibus embedding, in which the matrices of multiple graphs are combined into one and embedding is done on the big matrix (Levin et al., 2017). However, the combined matrix is usually very large and require lots of computational power. Dictionary learning is another framework for uncovering low-dimensional embeddings across multiple graphs, allowing for group comparison. Drawing upon clustering and linear decomposition methods, this method allows for additional constraints to achieve better formed solutions (Abraham et al., 2013). In the work of D'Souza et al. (D'Souza et al., 2019), they use a dictionary learning method to model interactions between resting state functional MRI and behavioral data in Autism Spectrum Disorder. Their method finds shared dictionary elements across multiple graphs and a subject specific loading onto the elements, which are then used as inputs to a neural network for disease prediction (D'Souza et al., 2019). In the work of Wang et al. (Wang et al., 2019), they propose a joint graph embedding method to estimate a low dimensional embedding across multiple graphs and each graph's projection onto that embedding, which we will discuss more in detail in Section 2.2.

In this work, we shift our attention away from functional connectivity and propose a graph-based approach to study neurodegeneration using correlations in volumetric data over time. In Alzheimer's disease, tau tangle accumulation is known to follow a stereotyped pattern, beginning in the transentorhinal region (stage I-II), spreading to the limbic regions (stage III-IV), and eventually moving throughout the isocortex (stage V-VI) Braak and Braak (1991). Evidence is accumulating from digital pathology and brain morphology that patterns of neurodegeneration follow this tau deposition Tward et al. (2020); Stouffer et al. (2022, 2023); Sadaghiani et al. (2022); Xie et al. (2022); Lyu et al. (2023), and therefore we hypothesize that correlations in neurodegeneration among these temporal lobe structures may provide a signal that is specific to early Alzheimer's. Our previous work Miller et al. (2015a) examined the timing of neurodegeneration throughout this medial temporal lobe network, but this spread of pathology throughout a characteristic network has been observed in other diseases as well such as Huntington's Ross et al. (2014), Parkinson's Visanji et al. (2013); Kordower (2014), depression Small et al. (2011) as well as other work in Alzheimer's disease AD Yin et al. (2014). The analysis we propose here provides an opportunity to

identify these networks from whole brain data in Alzheimer's and other disorders.

Other authors have previously proposed graph-based analysis of volumetric brain imaging data applied to neurological disease. For example Zugman et al. (2015) describes the use of "structural covariance" and Yin et al. (2023) describes the use of a "morphological connectivity network" to quantify characteristic brain networks involved in schizophrenia. Earlier work Bullmore and Bassett (2011) has proposed similar approaches termed "anatomical connectivity". However, each of these methods creates a single graph that describes a population, whereas our approach leverages time series data to produce a connectivity graph for each individual. This approach allows us to model variability between different individuals, and perform statistical testing on a well-posed joint model for graph valued random variables Chung et al. (2021). To our knowledge, this is the first time joint graph embeddings have been used to study volumetric brain data.
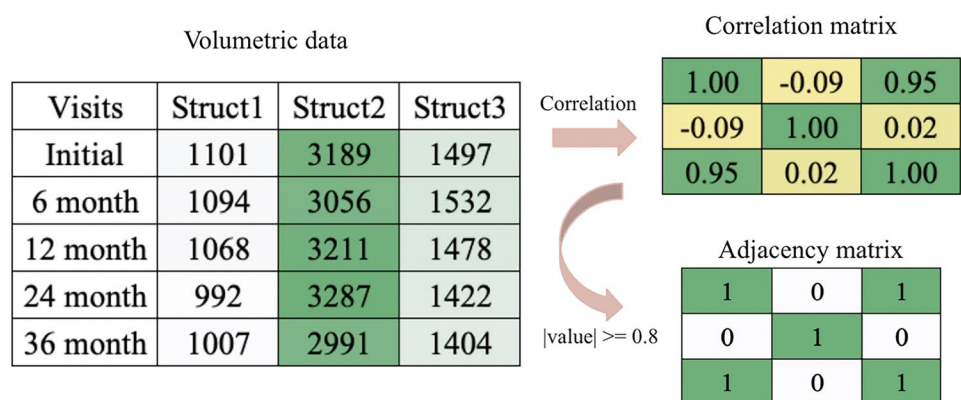
Similar to methods reviewed above, we model each brain as a graph where a node represents a structure of interest and an edge represents a correlation in atrophy patterns between two structures (Xu, 2021). We then use a multigraph embedding technique to try and understand these patterns and uncover potential biomarkers by applying multigraph embedding to study neurodegeneration over a long timescale (relative to fMRI measures). In addition, we hope to increase the specificity of neurodegeneration biomarkers by modeling the dataset with a more complex pattern than existing approaches. Rather than looking at volume changes in each structure individually in the traditional mass univariate method (Pengas et al., 2012), we add complexity by characterizing pair-wise relationships between structures in the context of uncovered networks. Furthermore, we illustrate how our embedding coefficients can be fed into machine learning algorithms for potential diagnostic applications, and how samples can be drawn from our model to visualize the brain's typicality and variability, interpolating between (or extrapolating beyond) healthy and diseased states.

Material and Methods

Data Preprocessing

We obtain our data from Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner,MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Specifically, we took the ADNI1 3Y1.5T Longitudinal FreeSurfer dataset (Wyman et al., 2013) prepared by University of California, San Francisco, comprised of 699 individuals in total. We selected 108 regions of interest common in studying neurodegeneration, excluding non-brain and whole-brain structures as we are interested in structures on a smaller scale. We selected a cohort in which each individual has at least 3 visits during the span of 3 years. Note here that the number of time point required for each subject is not fixed, as long as it is greater than 2, as 2 time points does not give a meaningful correlation matrix. For subjects with missing volumes, we forward filled in time by taking the measure from the most recent previous visit, note that the initial visit had no missing values for all subjects. We model each individual's brain as a graph, where each anatomical structure is a node and an edge exists between two nodes if they are highly correlated during neurodegeneration. For each individual in the cohort, we converted volumetric data into a correlation matrix of size 108 by 108 and then an adjacency matrix based on a threshold of 0.8 by absolute value, as shown in Fig. 1. We note here as there is no "gold standard" for choosing a threshold, previous work have used 0.1 in Kiar et al. (2017), 0.8 in Zhuo et al. (2018), or based on graph density and statistical significance in Bullmore and Bassett (2011). To investigate the potential for sensitivity to

Fig. 1 An example of transformation for one individual's selected structures

Volumetric data

Visits	Struct1	Struct2	Struct3
Initial	1101	3189	1497
6 month	1094	3056	1532
12 month	1068	3211	1478
24 month	992	3287	1422
36 month	1007	2991	1404

Correlation

Correlation matrix

1.00	-0.09	0.95
-0.09	1.00	0.02
0.95	0.02	1.00

|value| >= 0.8

Adjacency matrix

1	0	1
0	1	0
1	0	1

this choice, we ran additional experiments with thresholds of 0.7 and 0.9, and produced similar results. In particular, the structure with the top 10 highest loadings in each significant network were the same. For defining disease groups, the Clinical Dementia Rating (CDR) is referenced, which consists of 5 levels: 0 (None), 0.5 (Questionable), 1 (Mild), 2 (Moderate), and 3 (Severe) (Morris, 1991). We decided a threshold of <= 1 to separate the group into none/mild and severe cognitive impairment. After filtering and pre-processing, the cohort contains 494 individuals, specifically the none/mild group with 322 and the severe group with 172.

Multiple Random Eigengraphs and Joint Graph Embedding

We first review the mathematical model and original joint graph embedding algorithm proposed by Wang et al. (Wang et al., 2019), of which our algorithm is based on. In this work, we refer to a random graph as a graph in which the edges are generated under a probability distribution. The Multiple Random Eigengraphs (MREG) is a mathematical framework modeling multiple random graphs (Wang et al., 2019). Consider a set of m unweighted and undirected graphs with the same n vertices denoted by $\{G_s = (V_s, E_s)\}_{s=1}^m$. Let $h_1, ..., h_d$ be normalized vectors in $\mathbb{R}^n$ that span a subspace of dimension d, contributing to a large amount of variability in the set of graphs, and $\lambda_1, ..., \lambda_d$ be vectors in $\mathbb{R}^m$ such that $\sum_{k=1}^d \lambda_k[s]h_k h_k^T \in [0,1]^{n \times n}$ for all λ (Wang et al., 2019) for subject s. With these parameters known, we could generate a sample graph described here and illustrated in Fig. 2:

1. Identify each group specific network, and subject specific loadings
2. Carry out the product to form a score for each edge

3. Apply the softmax function to form a probability for each edge
4. Sample independent Bernoulli random variables for each possible edge

Under this procedure, the adjacency matrix A_s for each graph G_s should be modeled as follows (Wang et al., 2019):

$$A_s[i,j] \sim Bernoulli\left(\sum_{k=1}^d \lambda_k[s]h_k[i]h_k[j]\right), \text{ for } i \geq j \quad (1)$$

$$A_s[i,j] = A_s[j,i], \text{ for } i < j \quad (2)$$

Note here each A_s is symmetric, as opposed to sampling independently above and below the diagonal. The h vectors span the latent subspace shared by the set of multiple graphs, and the λ vectors represent graph-specific loadings onto the subspace (Wang et al., 2019).

The original joint graph embedding algorithm by Wang et al. (Wang et al., 2019) estimates a low dimensional embedding of the latent space across multiple graphs and each graph's projection onto that embedding under the MREG model. It estimates the subspace by minimizing the sum of squared errors (SSE) between the subspace and adjacency matrices. (Wang et al., 2019). In this work, we implemented a modified version of the Wang et al. algorithm as discussed below.

Limitations of the Original Framework

The original framework assumes a large enough number of vertices to provide accurate estimates. In the original paper's experiment analyzing brain data, there were 1105
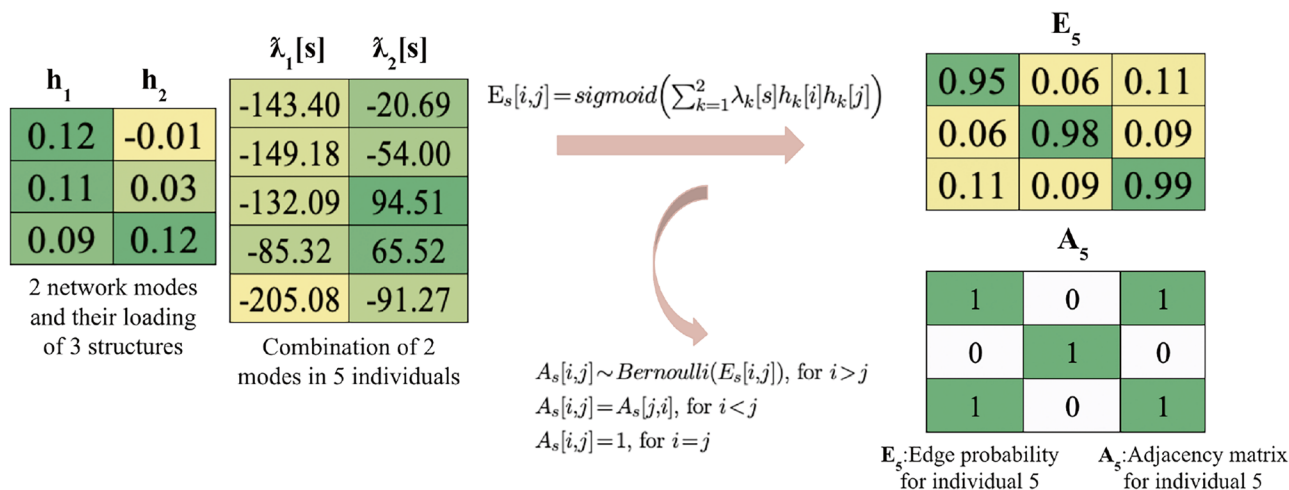

Fig. 2 An example of MREG for the same individual and structure in Fig. 1

vertices, but here we work with only 108 vertices. As such, we identify several limitations when applying the original framework and algorithm above to neurodegeneration data, which we will address and modify in our version:

1. The constraint that probabilities lie in [0, 1] is difficult to enforce in practice. The original method assumes a large enough sample space that gives desirable results, but does not hold up in smaller sample spaces. In fact, when we reanalyzed our data using the least squares estimator described in the original paper, the resulting matrix was not a probability matrix, with a maximum value of 1.53 and a minimum of −0.17. While embeddings from this model still have useful applications, it cannot be sampled from, for example, to produce visualizations as in our Fig. 6.

2. Incorrect estimation of diagonal entries contributes a negligible amount of error when the number of vertices is large, but contributes significantly in our case. We ran experiments by applying our algorithm with and without considering diagonal entries and comparing the binary cross entropy (BCE) loss between the two optimizations for various numbers of structures. If working with fewer than 40 structures, our method reduces the BCE loss by 29%. For our dataset, the BCE loss was reduced by 0.011%.

3. The SSE loss function does not correspond to a log likelihood under the proposed Bernoulli model, and therefore resulting parameter estimates do not have desirable properties of maximum likelihood estimators. For example, they are not guaranteed to be asymptotically unbiased or efficient, whereas our estimators are. On the other hand, the SSE loss can be optimized more efficiently than our method.

Now we state our main contribution and novelty in this work. First, we address the three limitations listed above by (1) introducing a sigmoid function to the model to guarantee edge probabilities are in [0, 1], (2) adding constraints on the diagonal such that parameter estimates are not forced to fit diagonal entries that carry no meaning (since diagonal entries of a correlation matrix are always 1), and (3) extending the original algorithm (Wang et al., 2019) to identify maximum likelihood estimators by gradient descent. With these modifications, our embedding algorithm generates probabilities suitable for a Bernoulli model, which we will describe more in detail below. Secondly, we develop and implement a novel statistical testing framework to detect complex patterns including network-structure pairs and triples, rather than a machine learning classifier. To our knowledge, our approach to testing patterns has not been performed to analyze joint graph embedding results on brain imaging data before.

Modifications to MREG

In this section, we precisely state our modifications to the MREG model in Wang et al. (Wang et al., 2019). Consider the set of m unweighted and undirected graphs with the same n vertices denoted by $\{G_s = (V_s, E_s)\}_{s=1}^{m}$, where a vertex represents a brain structure of our interest and an edge represents a strong correlation between structures. We modify the interpretation of h and state that the h vectors now span a space of parameters that encode the probability when acted on by a sigmoid function. This sigmoid function guarantees $sigmoid(\sum_{k=1}^{d} \lambda_k[s]h_k h_k^T) \in [0, 1]^{n \times n}$ even in the case of small samples, addressing limitation 1. Secondly, we force diagonal entries to be 1 since a structure's relation with itself is not of interest in this work, addressing limitation 2. Our modifications to Eqs. 1 and 2 are as follows:

$$A_s[i,j] \sim Bernoulli\left(sigmoid\left(\sum_{k=1}^{d} \lambda_k[s]h_k[i]h_k[j] \right) \right), \text{ for } i > j \tag{3}$$

$$A_s[i,j] = A_s[j,i], \text{ for } i < j \tag{4}$$

$$A_s[i,j] = 1, \text{ for } i = j \tag{5}$$

$$sigmoid(x) = \frac{1}{1 + \exp(-x)} \tag{6}$$

An example transformation from the latent subspace and subject specific loading to adjacency matrices observed in our data is illustrated in Fig. 2.

Our Maximum Likelihood Joint Graph Embedding Algorithm

We change the algorithm to better fit the constraints of our study and estimate the subspace by minimizing the binary cross entropy (BCE) loss, which is the log likelihood under the model. Let $E_s[i,j]$ be a symmetric matrix representing edge probabilities between structures i and j for subject s, then the BCE loss is minimized as in Eq. (7), where A_s are the observed adjacency matrices:

$$\underset{\lambda,h}{\arg\min} -\frac{1}{m} \sum_{s=1}^{m} \sum_{i=1}^{n} \sum_{j=i+1}^{n} A_s[i,j] \log(E_s[i,j]) + \tag{7}$$
$$(1 - A_s[i,j]) \log(1 - E_s[i,j])$$

Note here that the diagonal terms do not contribute to our loss function. By replacing the original loss function with BCE, we will have a well characterized maximum likelihood estimator even in the small sample case, addressing limitation 3. We take a greedy approach in finding the optimal

representation of the latent space used in the original method Wang et al. (2019), where we start with a 1D optimization problem, and then expand to the second dimension while keeping the first dimension fixed to find the optimal representation. The algorithm is implemented as follows:

```
λ ← random.normal((m, number of networks)), h ← random.normal((n, number of networks)),
normalize h
for d = 1; d <=number of networks; d + + do
    while loss not converged in dimension d do
        E = sigmoid(∑_{k=1}^{d} multiply(multiply(λ[:, : k], h[:, : k]), transpose(h[:, : k])))
        Compute BCE loss as in equation (7)
        Compute gradient of the BCE with respect to λ and h
        Update λ[:, d] by gradient descent, and h[:, d] by projected gradient descent (i.e. update
        and normalize)
    end while
end for
```

We coded in python and PyTorch, and used automatically calculated gradients for gradient descent. The algorithm took 30 minutes to run on a computer with 10-core CPU. Convergence in all 4 dimensions is shown in Fig. 3a, where at every 10,000 iterations we observe the loss dropping quickly after adding another dimension, and converging before the next dimension is added.

Hyperparameter Tuning

For hyperparameter tuning, we first estimate an optimal dimension of the latent space to account for the majority of the variation in observed correlation matrices. To do this, we computed an eigendecomposition for each subject's correlation matrix, and found that 4 components provided a reasonable reconstruction accuracy. A scree plot for one typical subject is shown in Fig. 3b. Next, we performed a grid search over several orders of magnitude for the gradient descent step sizes corresponding to h and λ, and selected the largest parameters that gave convergence without oscillation.

We decided to set the step size for h to 2 and the step size for λ to 20000.

Significance Testing

We develop and implement a novel test statistic similar to F-type statistics (i.e. comparing sum of square error under two different models) to test for differences in networks and structures between the two groups, and use permutation testing on a maximum statistic to control for Familywise Error Rate (FWER) at 5% (Nichols & Hayasaka, 2003). In addition, we compare the difference between groups after accounting for confounders: age, intracranial volume and APOE gene status by least squares regression. APOE gene status was modeled as a categorical (as opposed to cardinal) variable, where an individual may have 0, 1, or 2 copies. Other studies of neurodegeneration have used similar covariates in analysis. For example studying morphometry in early Alzheimer's, Miller et al. (2013) and Miller et al. (2015b) accounted for intracranial volume and sex, in studying morphometry in schizophrenia Yin et al. (2023) accounted for age and sex. Our choices of confounders were inspired by Judea Pearl's backdoor criteria Neuberg (2003), regressing out a sufficient set of variables we believe have a causal effect on both disease status and morphology. In particular, we chose not to adjust for sex, as we believe its largest impact on morphology is mediated by brain size, which we have already adjusted for. We included APOE status as a covariate to understand what additional information morphology can tell us about disease beyond what is already known from genetics. For a cohort with m subjects, d networks, and n structures, we define the test score as $\text{Score}_{sij} = \sum_k \lambda_k[s]h_k[i]h_k[j]$ for subject s and structures i and j. We pass the score through a sigmoid to give probabilities E, and estimate these variables by maximum likelihood. We
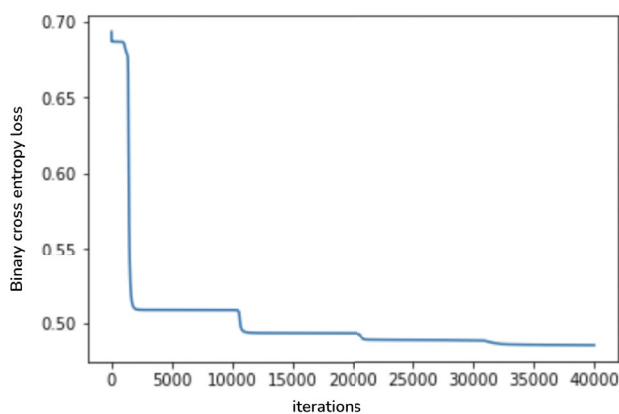
(a)
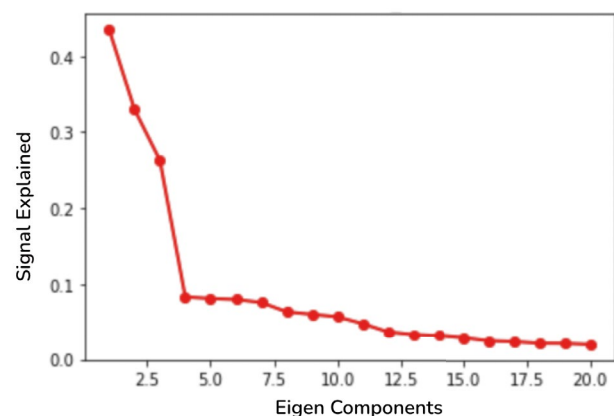
(b)

Fig. 3 a. Convergence in all 4 dimensions with the algorithm **b**. Scree plot with elbow at 4 dimensions

form the tensor $T_{skij} = \lambda_k[s]h_k[i]h_k[j]$ to get a three-tuple for each subject s, of which we will perform statistical tests by comparing how close it is to its group-dependent or group-combined average, after accounting for confounders. By summing over various combinations of indices in T before statistical testing, we are able to test for:

1. networks only
2. structures only
3. network-structure pairs
4. structure-structure pairs
5. network-structure-structure triples

While network only and structure only are fairly straightforward to explain as they represent the networks' and structures' association with neurodegeneration in AD, the rest are more difficult to interpret. We can view network-structure pair as studying the structure's association with disease status through its role in the network. We can view structure-structure pairs as discovering structures that, when correlated with another structure, have significant effect on disease status. Lastly for the triple, we can interpret it as the effect on disease status as the pair of structures, through their combined role in the network. We note here that tests for 2 include the standard mass univariate approach of Hayasaka et al. (2004), and test for 4 include permutation testing on the absolute value of correlation coefficients, described for example in Bullmore and Bassett (2011). As standard tests for items 2 and 4 exist, we will focus our work on 1, 3, and 5. To our knowledge, our framework for testing patterns involving networks, pairs, or triples has not been performed to analyze brain imaging data before.

Confounder Regression Analysis

We perform least square regression analysis to test for true signals not caused by common confounders for AD and neurodegeneration. We start with a design matrix D containing a column of 1 s (for mean), and columns for age, intracranial volume and APOE gene status, and estimate a coefficient matrix $\hat{C}$ for confounders. From $\hat{C}$ the SSE for one group is calculated as follows:

$$\hat{C}^{combined} = (D^T D)^{-1} D^T T \tag{8}$$

$$\hat{T}^{combined} = \hat{C}^{combined} D \tag{9}$$

Here T is reshaped into a matrix from a tensor for calculation. Next, we calculate the SSE for splitting the cohort into two groups: none/mild and severe AD. We first form D', which has an additional column to D indicating disease status. We then perform the same regression, with D' replacing

D in Eqs. 8 and 9 to find $\hat{T}^{twogroups}$. In the next sections, we will use $\hat{T}^{combined}$ and $\hat{T}^{twogroups}$ to calculate the sum of square errors (SSE) when considering one vs. two groups.

Networks

We start by testing for significant networks and obtain the test statistic X_k for network k by reducing (i.e. summing over) additional dimensions. Let $\hat{T}^g_{k,i,j}$ be the expected value matrix under our linear model for dimension k, structures i and j, and subjects in group g. We remove extra dimensions by calculating the SSE between one group vs. two groups and X_k as follows:

$$SSE^g_k = \sum_{s \in g, i, j} (T_{skij} - \hat{T}^g_{kij})^2 , \, g \in \{combined, twogroups\} \tag{10}$$

$$X_k = SSE^{combined}_k - SSE^{twogroup}_k \tag{11}$$

To control the FWER at 5%, we use permutation testing and take the max over k at each iteration and define the threshold as the 95 percentile of 10,000 simulation results (Nichols & Hayasaka, 2003).

Network Structure Pairs

Similar to networks only, we calculate the SSE in the two settings but reduce one fewer level as follows:

$$SSE^g_{k,i} = \sum_{s \in g, j} (T_{skij} - \hat{T}^g_{kij})^2 , \, g \in \{combined, twogroups\} \tag{12}$$

$$X_{k,i} = SSE^{combined}_{k,i} - SSE^{twogroup}_{k,i} \tag{13}$$

We follow the same permutation testing procedure and take a max over k, i instead of just k at each iteration.

Network Structure Structure Triples

Lastly for triples, we form the test statistic:

$$SSE^g_{k,i,j} = \sum_{s \in g} (T_{skij} - \hat{T}^g_{kij})^2 , \, g \in \{combined, twogroups\} \tag{14}$$

$$X_{k,i,j} = SSE^{combined}_{k,i,j} - SSE^{twogroup}_{k,i,j} \tag{15}$$

We follow the same permutation testing procedure and take a max over k, i, j at each iteration.

In our experiments, we performed the analysis twice: with and without adjusting for confounders. Note that not adjusting for confounders is a special case of T estimation, where $\hat{T}^g$ is simply the mean over all subjects in group g. In the

results section, we will present the former, and make brief comparisons to the latter.

Code Availability

Our source code and documentation are available on GitHub at https://github.com/twardlab/joint_graph_embedding_AD. To replicate our work, first agree to the user agreement by ADNI and download the ADNI1 3Y1.5T Longitudinal FreeSurfer dataset by University of California, San Francisco (Wyman et al., 2013). Put all files under a directory named dataset, and first run the Jupyter notebook *preprocess*, then the notebook *joint_graph_embedding_analysis*. We document our code with Sphinx (Brandl, 2021), and save documentations in *docs*. More details on how to replicate our work can be found in our GitHub repository.

Results

Significant Networks Associated with AD Neurodegeneration

Out of the 4 networks identified from joint graph embedding, we found the first 2 extremely significant, after accounting for confounders. Upon examination in Table 1, both networks are dominated by structures believed to be associated with AD neurodegeneration. For example, our work has demonstrated the involvement of amygdala (uncovered in network 1) in early Alzheimer's Miller et al. (2015b). In defining criteria for diagnosing Alzheimer's disease McKhann et al. (2011) describes "disproportionate atrophy on structural magnetic resonance imaging in me-dial, basal, and lateral temporal lobe" (uncovered in network 2). We show the results from permutation testing and the statistics for each network in Fig. 4 and visualize

the structures in network 1 and 2 in Fig. 5. In Table 1, we include the top 10 structure by absolute value of loading in each network.

Significant Network Structure Pairs Associated with AD Neurodegeneration

For network-structure pairs, we rejected 170 out of 432 network structure pairs. While most rejected pairs occur in networks previously found significant, the pairs are not exclusive to network 1 and 2 only. Three structure pairs in network 3 were found significant. In Table 2, we show the top 5 significant pairs ranked by p-value.
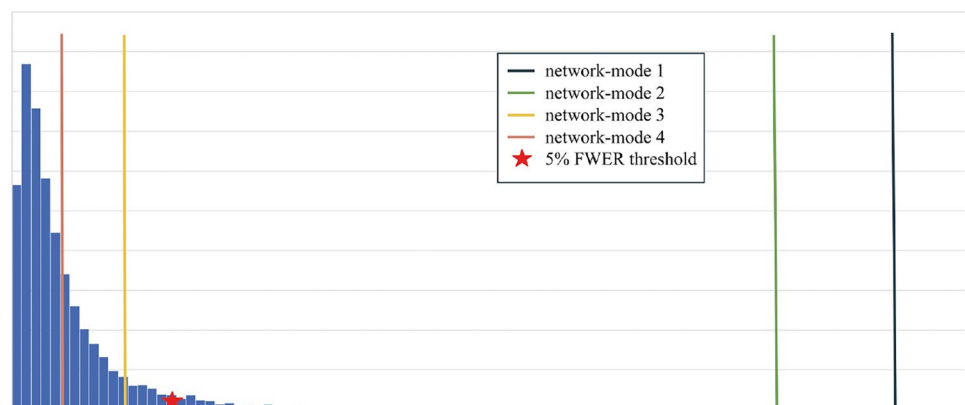
Significant Network Structures Triples Associated with AD Neurodegeneration

For network-structure-structure triples, we rejected 753 out of 46656 possible triples. The network and structures found significant are again not a subset of those found significant in the network-structure pair. In Table 3, we show the top 5 significant triples ranked by p-value.

Results Comparing Confounder Regression and No Regression

While the results were similar to those shown above, the analysis without accounting for confounders found more structures and networks significant than analysis with confounder regression. For networks only, network 3 became slightly significant, and networks 1 and 2 remained highly significant. The top 10 structures were the same for networks 1 and 2 in both analysis, but those in structure 3 and 4 were different. For tuples analysis, we found 205 instead of 170 pairs and 1171 instead of 753 triples significant. The networks and structures found significant in triples are not a subset of those found in pairs and vice versa.

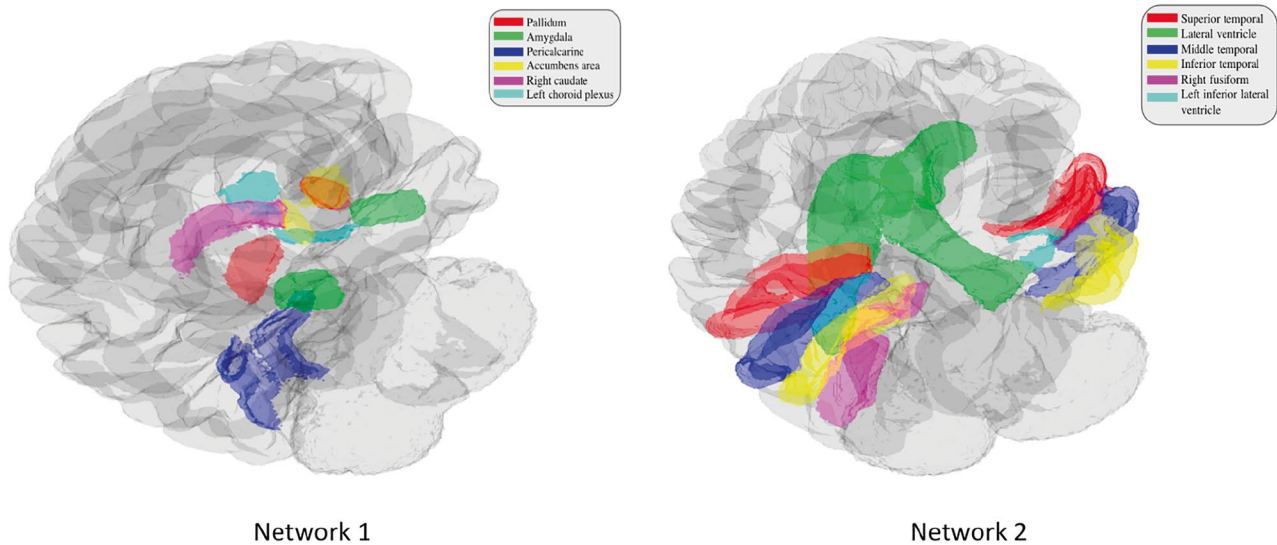**Fig. 4** Histogram of permutation testing after accounting for confounders

Fig. 5 Visualization of significant networks

Results Comparing Significant Structures and Structure-Structure Pairs with Existing Methods

Here we compare the significant structures and structure-structure pairs found by our method and existing methods. For structures only, we compare results to mass univariate analysis Hayasaka et al. (2004). Here we calculate atrophy rate per patient for each structure and find the group

SSE by subtracting from the group mean. The test statistic is calculated as $SSE_{combined} - (SSE_{AD} + SSE_{healthy})$ and the 95 percentile threshold is found using permutation testing. The mass univariate method rejected 36 structures out of 108, which is a subset of the structures our method found significant.

For structure-structure pairs, we compare to a method described in Bernal-Rusiel et al. (2013). For each pair, we fit a linear line between considering the combined group only or AD/normal groups, and calculate the SSE difference. Again, we run permutation testing to find the 95 percentile threshold. The method rejected the null hypothesis for 127 pairs, 79 of which are commonly shared with our finding (62% overlap).

Results Comparing Greedy and Non-Greedy Optimization

We considered the effect of greedy optimization (as described in the original implementation in Wang et al. (2019) versus joint optimization over all dimensions simultaneously. The former has the advantage that network modes are ordered in terms of the variance they explain

Table 1 Top 10 structures ranked by loading of each network

network 1	network 2	network 3	network 4
Left Pericalcarine	Right Middle Temporal	Right Caudal Middle Frontal	Fifth Ventricle
Left Pallidum	Left Middle Temporal	Right Superior Frontal	Right Lingual
Right Pericalcarine	Left Inferior Temporal	Middle Frontal	Left Precuneus
Right Pallidum	Left Superior Temporal	Left Superior Frontal	Right Lateral Occipital
Right Accumbens Area	Right Inferior Temporal	Left Caudal Middle Frontal	Left Superior Parietal
Right Caudate	Right Lateral Ventricle	Left Rostral Middle Frontal	Right Precuneus
Left Accumbens Area	Left Lateral Ventricle	Right Inferior Parietal	Right Superior Parietal
Left Choroid Plexus	Right Superior Temporal	Left Precentral	Left Inferior Parietal
Left Amygdala	Left Inferior Lateral Ventricle	Right Precentral	Left Isthmus Cingulate
Right Amygdala	Right Fusiform	Right Supramarginal	Left Cerebral Cortex

Table 2 Top 5 significant pairs ranked by FWER-corrected p value

network-structure pairs	p-value
1-Right Pallidum	<1e-4
2-Left Inferior Lateral Ventricle	<1e-4
2-Left Fusiform	<1e-4
2-Left Entorhinal	<1e-4
2-Left Cerebral Cortex	1e-4

Table 3 Top 5 significant triples ranked by FWER-corrected p value

network-structure-structure triples	p-value
2-Left Inferior Temporal-Right Middle Temporal	<1e-4
2-Left Middle Temporal-Right Middle Temporal	<1e-4
2-Left Superior Temporal-Right Middle Temporal	<1e-4
2-Left Inferior Temporal-Left Middle Temporal	<1e-4
2-Left Middle Temporal-Left Superior Temporal	1e-4

(similar to principal component analysis). We repeated our experiments using a non-greedy approach, and saw that each network recovered had the same 10 structures with the highest loadings. On the other hand, when using the non-greedy approach our statistical results were different, with our null hypothesis rejected for only one network, but more network-structure pairs and network-structure-structure triples rejected.

Subject Specific Analysis

To illustrate a possibility for further subject-specific analysis using the results from our method, we built a logistic regression model using subject-specific loadings from our method along with age, APOE status and ICV. The classification model predicts each subject's disease status, 1 for AD and 0 for normal. We compare two model results with 10-fold cross validation: one with just age, APOE status and ICV, and the other with added loadings from our method. The first model had an ROC-AUC (area under the receiver operating characteristic curve) of 0.72 while the second had an improved ROC-AUC of 0.78, indicating some added utility by including our results in subject-specific analysis.

Discussion

In this work, we applied a joint graph embedding method (similar to a multivoxel dictionary embedding method), which would typically be used to study functional connectivity, to volumetric data in neurodegeneration. We extended the original algorithm (Wang et al., 2019) and implemented our maximum likelihood joint graph embedding algorithm to identify significant structural networks from volumetric data in Alzheimer's disease cohorts (Wyman et al., 2013). We showed that our version of the algorithm has promise in uncovering latent dimensions that are easy to interpret and visualize. In addition, we developed and implemented a novel testing procedure and tested for significant networks, network-structure pairs and network-structure-structure triples. We performed analysis to regress out common confounders in AD in hope of gaining more discovery power

and make a few comments here between our results. We found fewer significant pairs and triples when taking confounders into account than not. This is expected, as some structures previously found significant may be caused by confounders. While networks 1 and 2 remained unchanged and highly significant, indicating that changes in these networks are due to disease progression, network 3 was no longer a significant network. The structures' change in networks 3 and 4 also indicate these structures may have been significant due to common confounders such as age.

Our framework shows promise in that it discovered structures commonly believed to be associated with AD neurodegeneration. We point out several strengths in our framework. The method offers a new way to view neurodegeneration in AD, where we can not only study networks and structures by themselves, but also their interaction with one another in terms of pairs and triples. Each of the three groups of findings gives us more information on structural correlations than traditional methods (Pengas et al., 2012), and we also find structures not found by previous methods, such as mass univariate analysis (Bernal-Rusiel et al., 2013). Since these pairs and triples are a novel description of neurodegeneration patterns, we will briefly state their interpretation. For example, our discovered pairs can be interpreted as "the right pallidum displays a significant association with disease status, through its role in network 1". As another example, our discovered triples can be interpreted as "the interaction between the left inferior temporal lobe and right middle temporal lobe displays a significant association with disease status, through its role in network 2". One future direction of this study is to examine more closely the interpretation and meaning of the networks and their loadings. Furthermore, the results can be used as an additional source of information in clinical studies. We gave an example of how the results may be used in further subject-specific analysis by implementing a classifier that predicts disease status, and showed that it outperforms the classifier with the same architecture but the loadings removed. Another example is to sample from a model with group average loadings, and interpolate and extrapolate connectivity to generate networks for subjects along a continuum between healthy and diseased. We include a visualization in Fig. 6 using 5 different values of λ (from top to bottom) for a fixed h, and generating networks by sampling from our probability matrix 3 times (from left to right) for each λ. The λ values we used interpolate and extrapolate between the mean for the AD and control group: $p\bar{\lambda}_{control} + (1-p)\bar{\lambda}_{AD}$ for $p \in \{1.5, 1.0, 0.5, 0.0, -0.5\}$. Values outside $[0,1]$ represent extrapolation, which is possible because our model is guaranteed to result in probability matrices in the range $[0,1]$.

Secondly, the dataset used in our study, ADNI (Jack et al., 2008), is very well characterized, it is more robust and spans a longer period than most clinical data. Thirdly,
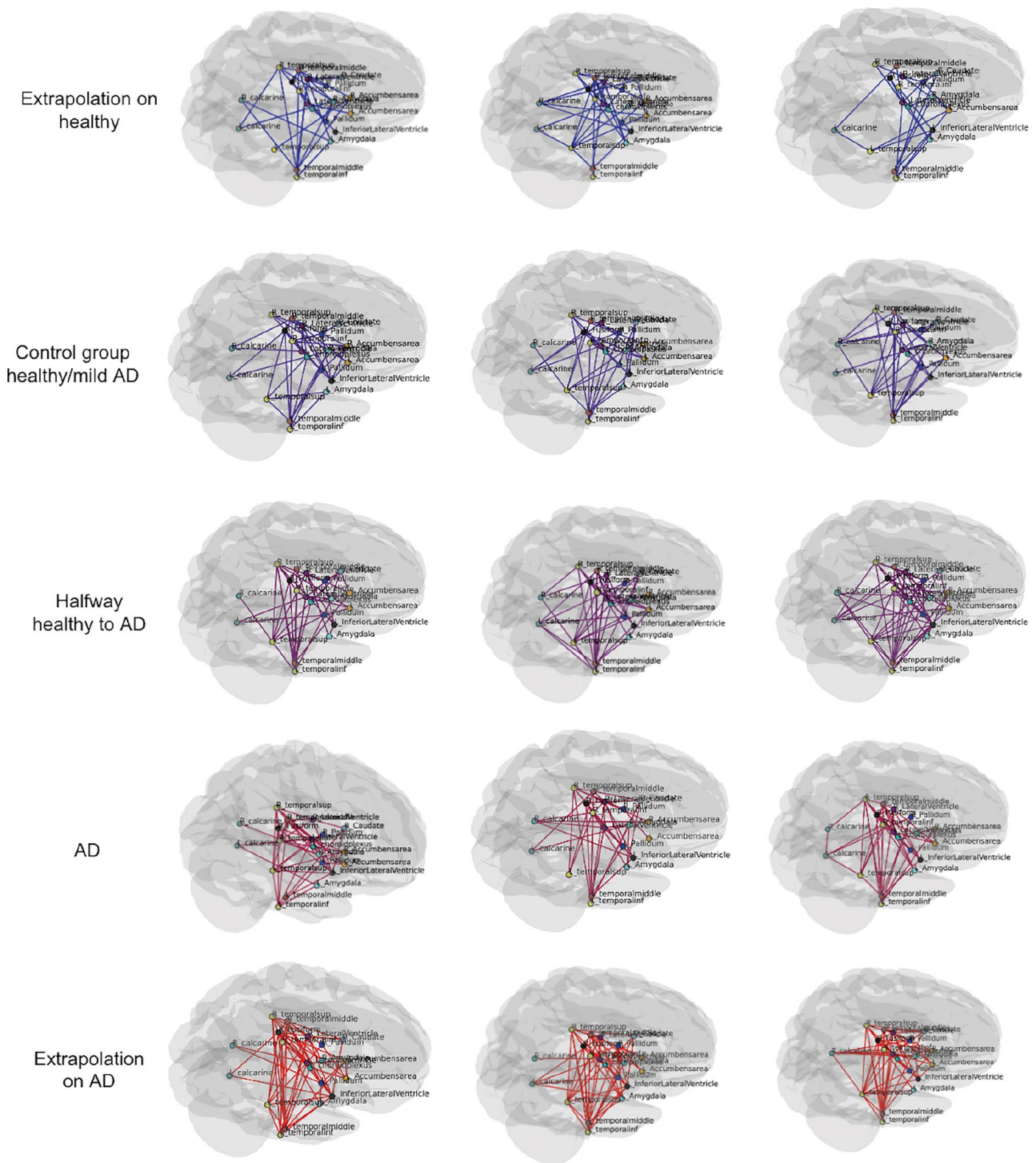
Fig. 6 Visualization of networks sampled from our model using a fixed *h* (measured from data), and values of *λ* along a continuum between the mean for control and the mean for AD. We include extrapolation (top and bottom row) as well as interpolation (middle row), and draw 3 independent samples to understand variability in each case (columns)

though increased complexity in algorithm and testing procedure often require more computational time, our algorithm embeds 494 patients with 108 structures in just 30 minutes and runs significance testing in 4 hours. We experimented with cohort and structure sizes and show runtime results in Fig. 7. We note that while the algorithm's runtime scales
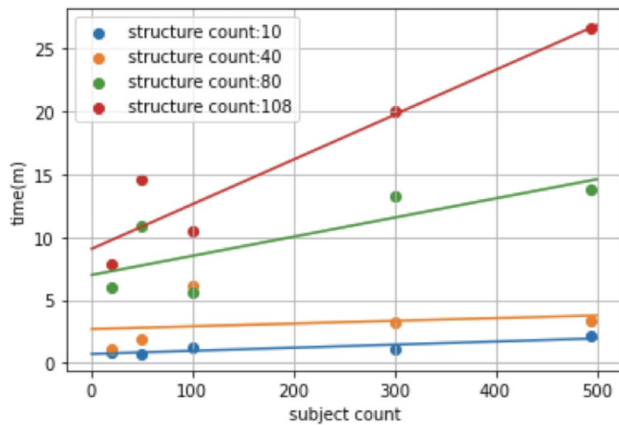
(a) Runtime with increasing subjects
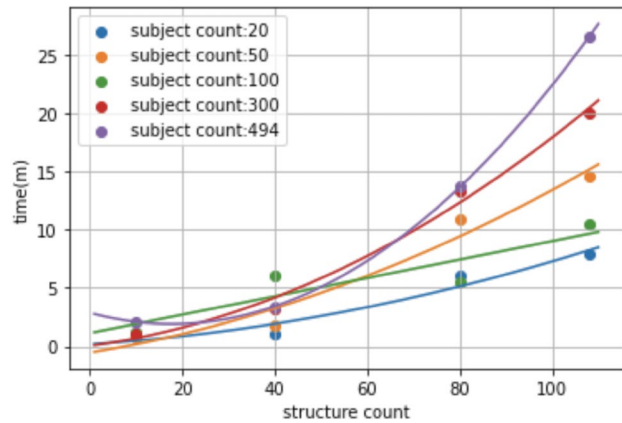
(b) Runtime with increasing structures

Fig. 7 Runtime with increasing parameters

linearly in subjects, it is exponential in structure counts. However, as studies often do not include a large set of structures, it should not be of big concern to users. On the other hand, our algorithm shows good time complexity with increased cohort sizes.

By offering more flexibility in terms of patterns that we can identify, we hope that our framework will reveal biomarker patterns more sensitive to AD.

Next we discuss a few limitations in this work. First, due to our high standard of cohort selection, this dataset may not be representative of clinical MRI subjects. As such, a natural future direction for the study is to apply this framework to clinical datasets, which represent a more diverse population. Secondly, our findings are on a population scale instead of an individual scale, which may lead to additional biases when applying to individual clinical diagnosis. We believe that using a more fine-grained separation of the cohort, as opposed to modeling disease status as only two groups, may help address this issue. Thirdly, we note that since there has been little work done to apply graph embedding methods to study volumetric data in Alzheimer's Disease cohorts, we do not have result comparisons with "state-of-the-art" methods. As with most unsupervised methods, there is no ground truth for evaluation to draw a conclusion on which method is best. Rather, we offer an additional method in analyzing group differences between healthy and diseased individuals based on neurodegeneration. As mentioned above, more work needs to be done to clearly and fully interpret the meaning of the network-structure pairs and triples.

In the fMRI community, network models enabled the transition from focusing on individual hotspots involved in specific tasks, to modeling brain wide activity in resting states Lee et al. (2013). These models helped to launch massive undertakings such as the human connectome project Elam et al. (2021). We believe that the incorporation of

network models such as the one presented here in analysis of structural data may have a similar impact on the field of brain morphometry: transitioning from analysis of single regions in a mass univariate approach, to analyzing brain-wide patterns of tissue loss. With the recent development of potential drugs to treat Alzheimer's disease, methods like ours that can quantify complex patterns of neurodegeneration will be essential for noninvasively identifying patients who would benefit the most. We believe that the development and dissemination of algorithms such as this one, through open source code and well documented examples, will play an important role in helping to reduce the burden of Alzheimer's disease on our aging population.

Information Sharing Statement

Our code is publicly available on Github, details on access in Sect. 2.8. The sample data can be downloaded from ADNI after accepting their data use agreement.

Declarations

Conflict of Interest The authors have nothing to disclose.

References

Abraham, A., Dohmatob, E., Thirion, B., Samaras, D., & Varoquaux, G. (2013). Extracting brain regions from rest fmri with total-variation constrained dictionary learning. In *International Conference on Medical Image Computing and Computer-assisted Intervention*, pages 607–615. Springer.

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., et al. (2011). The diagnosis of mild cognitive impairment due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & Dementia, 7*(3), 270–279.

Association, A. (2019). 2019 alzheimer's disease facts and figures. *Alzheimer's & Dementia, 15*(3), 321–387.

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation, 15*(6), 1373–1396.

Bernal-Rusiel, J. L., Reuter, M., Greve, D. N., Fischl, B., Sabuncu, M. R., Initiative, A. D. N., et al. (2013). Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. *Neuroimage, 81*, 358–370.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine, 34*(4), 537–541.

Braak, H., & Braak, E. (1991). Neuropathological stageing of alzheimer-related changes. *Acta Neuropathologica, 82*(4), 239–259.

Brandl, G. (2021). Sphinx documentation. http://sphinx-doc.org/sphinx.pdf

Bullmore, E. T., & Bassett, D. S. (2011). Brain graphs: Graphical models of the human brain connectome. *Annual Review of Clinical Psychology, 7*, 113–140.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional mri data using independent component analysis. *Human Brain Mapping, 14*(3), 140–151.

Chung, J., Bridgeford, E., Arroyo, J., Pedigo, B. D., Saad-Eldin, A., Gopalakrishnan, V., Xiang, L., Priebe, C. E., & Vogelstein, J. T. (2021). Statistical connectomics. *Annual Review of Statistics and its Application, 8*, 463–492.

Cohen, J. D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., Norman, K. A., Pillow, J., Ramadge, P. J., Turk-Browne, N. B., et al. (2017). Computational approaches to fmri analysis. *Nature Neuroscience, 20*(3), 304–313.

Cole, D. M., Smith, S. M., & Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state fmri data. *Frontiers in Systems Neuroscience*, page 8.

D'Souza, N. S., Nebel, M. B., Wymbs, N., Mostofsky, S., & Venkataraman, A. (2019). Integrating neural networks and dictionary learning for multidimensional clinical characterizations from functional connectomics data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 709–717. Springer.

Elam, J. S., Glasser, M. F., Harms, M. P., Sotiropoulos, S. N., Andersson, J. L., Burgess, G. C., Curtiss, S. W., Oostenveld, R., Larson-Prior, L. J., Schoffelen, J.-M., et al. (2021). The human connectome project: a retrospective. *NeuroImage, 244*, 118543.

Fornito, A., Zalesky, A., & Bullmore, E. (2016). *Fundamentals of Brain Network Analysis*. Academic Press.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage, 22*(2), 676–687.

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Feldman, H. H., Frisoni, G. B., Hampel, H., Jagust, W. J., Johnson, K. A., Knopman, D. S., et al. (2016). A/t/n: an unbiased descriptive classification scheme for alzheimer disease biomarkers. *Neurology, 87*(5), 539–547.

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., Whitwell, L. J., Ward, C., et al. (2008). The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 27*(4), 685–691.

Jack, C. R., Jr., & Holtzman, D. M. (2013). Biomarker modeling of alzheimer's disease. *Neuron, 80*(6), 1347–1358.

Khosla, M., Jamison, K., Ngo, G. H., Kuceyeski, A., & Sabuncu, M. R. (2019). Machine learning in resting-state fmri analysis. *Magnetic Resonance Imaging, 64*, 101–121.

Kiar, G., Bridgeford, E. W., Roncai, W. R. G., for Reliability, C., (CoRR), R., Chandrashekhar, V., Mhembere, D., Ryman, S., Zuo, X.-N., Margulies, D. S., Craddock, R. C., et al. (2017). A high-throughput pipeline identifies robust connectomes but troublesome variability. *bioRxiv*, page 188706.

Kordower, J. H. (2014). The prion hypothesis of parkinson's disease: this hot topic just got hotter.

Lee, M. H., Hacker, C. D., Snyder, A. Z., Corbetta, M., Zhang, D., Leuthardt, E. C., & Shimony, J. S. (2012). Clustering of resting state networks. *PLoS one, 7*(7), e40370.

Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting-state fmri: a review of methods and clinical applications. *American Journal of Neuroradiology, 34*(10), 1866–1872.

Levin, K., Athreya, A., Tang, M., Lyzinski, V., Park, Y., & Priebe, C. E. (2017). A central limit theorem for an omnibus embedding of multiple random graphs and implications for multiscale network inference. arXiv preprint arXiv:1705.09355

Lewis-Peacock, J. A., & Norman, K. A. (2014). Multi-voxel pattern analysis of fmri data. *The Cognitive Neurosciences, 512*, 911–920.

Lyu, X., Duong, M. T., Xie, L., de Flores, R., Richardson, H., Hwang, G., Wisse, L. E., DiCalogero, M., McMillan, C. T., Robinson, J. L., et al. (2023). Tau-neurodegeneration mismatch reveals vulnerability and resilience to comorbidities in alzheimer's continuum. *medRxiv*, pages 2023–02.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., et al. (2011). The diagnosis of dementia due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & Dementia, 7*(3), 263–269.

Miller, M. I., Ratnanather, J. T., Tward, D. J., Brown, T., Lee, D. S., Ketcha, M., Mori, K., Wang, M.-C., Mori, S., Albert, M. S., et al. (2015). Network neurodegeneration in alzheimer's disease via mri based shape diffeomorphometry and high-field atlasing. *Frontiers in Bioengineering and Biotechnology, 3*, 54.

Miller, M. I., Younes, L., Ratnanather, J. T., Brown, T., Trinh, H., Lee, D. S., Tward, D., Mahon, P. B., Mori, S., Albert, M., et al. (2015). Amygdalar atrophy in symptomatic alzheimer's disease based on diffeomorphometry: the biocard cohort. *Neurobiology of Aging, 36*, S3–S10.

Miller, M. I., Younes, L., Ratnanather, J. T., Brown, T., Trinh, H., Postell, E., Lee, D. S., Wang, M.-C., Mori, S., O'Brien, R., et al. (2013). The diffeomorphometry of temporal lobe structures in preclinical alzheimer's disease. *NeuroImage: Clinical*, 3:352–360.

Morris, J. C. (1991). The clinical dementia rating (cdr): Current version and. *Young, 41*, 1588–1592.

Neuberg, L. G. (2003). Causality: models, reasoning, and inference, by judea pearl, cambridge university press, 2000. *Econometric Theory*, 19(4):675–685.

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research, 12*(5), 419–446.

Pengas, G., Williams, G., Acosta-Cabronero, J., Ash, T., Hong, Y., Izquierdo-Garcia, D., Fryer, T., Hodges, J., & Nestor, P. (2012). The relationship of topographical memory performance to regional neurodegeneration in alzheimer's disease. *Frontiers in Aging Neuroscience, 4*, 17.

Ross, C. A., Aylward, E. H., Wild, E. J., Langbehn, D. R., Long, J. D., Warner, J. H., Scahill, R. I., Leavitt, B. R., Stout, J. C., Paulsen, J. S., et al. (2014). Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology, 10*(4), 204–216.

Sadaghiani, S., Trotman, W., Lim, S. A., Chung, E., Ittyerah, R., Ravikumar, S., Khandelwal, P., Prabhakaran, K., Lavery, M. L., Ohm, D. T., et al. (2022). Associations of phosphorylated tau pathology with whole-hemisphere ex vivo morphometry in 7 tesla mri. *Alzheimer's & Dementia.*

Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., & Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience, 12*(10), 585–601.

Stouffer, K. M., Chen, C., Kulason, S., Xu, E., Witter, M. P., Ceritoglu, C., Albert, M. S., Mori, S., Troncoso, J., Tward, D. J., et al. (2023). Early amygdala and erc atrophy linked to 3d reconstruction of rostral neurofibrillary tau tangle pathology in alzheimer's disease. *NeuroImage: Clinical*, 38:103374.

Stouffer, K. M., Witter, M. P., Tward, D. J., & Miller, M. I. (2022). Projective diffeomorphic mapping of molecular digital pathology with tissue mri. *Communications Engineering, 1*(1), 44.

Sussman, D. L., Tang, M., Fishkind, D. E., & Priebe, C. E. (2012). A consistent adjacency spectral embedding for stochastic blockmodel graphs. *Journal of the American Statistical Association, 107*(499), 1119–1128.

Tward, D., Brown, T., Kageyama, Y., Patel, J., Hou, Z., Mori, S., Albert, M., Troncoso, J., & Miller, M. (2020). Diffeomorphic registration with intensity transformation and missing data: Application to 3d digital pathology of alzheimer's disease. *Frontiers in Neuroscience, 14*, 52.

Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fmri functional connectivity. *European Neuropsychopharmacology, 20*(8), 519–534.

Visanji, N. P., Brooks, P. L., Hazrati, L.-N., & Lang, A. E. (2013). The prion hypothesis in parkinson's disease: Braak to the future. *Acta Neuropathologica Communications, 1*(1), 1–12.

Wang, S., Arroyo, J., Vogelstein, J. T., & Priebe, C. E. (2019). Joint embedding of graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(4), 1324–1336.

Wyman, B. T., Harvey, D. J., Crawford, K., Bernstein, M. A., Carmichael, O., Cole, P. E., Crane, P. K., DeCarli, C., Fox, N. C., Gunter, J. L., et al. (2013). Standardization of analysis sets for reporting results from adni mri data. *Alzheimer's & Dementia, 9*(3), 332–337.

Xie, L., Wisse, L. E., Das, S. R., Lyu, X., de Flores, R., Yushkevich, P. A., & Wolk, D. A. (2022). Tau burden is associated with cross-sectional and longitudinal neurodegeneration in the medial temporal lobe in cognitively normal individuals. *Alzheimer's & Dementia, 18*, e067095.

Xu, M. (2021). Understanding graph embedding methods and their applications. *SIAM Review, 63*(4), 825–853.

Yang, J., Gohel, S., & Vachha, B. (2020). Current methods and new directions in resting state fmri. *Clinical Imaging, 65*, 47–53.

Yin, M., Huang, W., Liang, Z., Liu, Q., & Tang, X. (2023). Network analysis on cortical morphometry in first-episode schizophrenia. *IEEE Transactions on Cognitive and Developmental Systems.*

Yin, R.-H., Tan, L., Jiang, T., & Yu, J.-T. (2014). Prion-like mechanisms in alzheimer's disease. *Current Alzheimer Research, 11*(8), 755–764.

Zhuo, C., Wang, C., Wang, L., Guo, X., Xu, Q., Liu, Y., & Zhu, J. (2018). Altered resting-state functional connectivity of the cerebellum in schizophrenia. *Brain Imaging and Behavior, 12*(2), 383–389.

Zugman, A., Assunção, I., Vieira, G., Gadelha, A., White, T. P., Oliveira, P. P. M., Noto, C., Crossley, N., Mcguire, P., Cordeiro, Q., et al. (2015). Structural covariance in schizophrenia and first-episode psychosis: An approach based on graph analysis. *Journal of Psychiatric Research, 71*, 89–96.