Learning Common Harmonic Waves on Stiefel Manifold – A New Mathematical Approach for Brain Network Analyses

Jiazhou Chen, Guoqiang Han, Hongmin Cai, *Member, IEEE*, Defu Yang, Paul J. Laurienti, Martin Styner, and Guorong Wu, *Member, IEEE*

***Abstract*— Converging evidence shows that disease-relevant brain alterations do not appear in random brain locations, instead, their spatial patterns follow large-scale brain networks. In this context, a powerful network analysis approach with a mathematical foundation is indispensable to understand the mechanisms of neuropathological events as they spread through the brain. Indeed, the topology of each brain network is governed by its native harmonic waves, which are a set of orthogonal bases derived from the Eigen-system of the underlying Laplacian matrix. To that end, we propose a novel connectome harmonic analysis framework that provides enhanced mathematical insights by detecting frequency-based alterations relevant to brain disorders. The backbone of our framework is a novel manifold algebra appropriate for inference across harmonic waves. This algebra overcomes the limitations of using classic Euclidean operations on irregular data structures. The individual harmonic differences are measured by a set of common harmonic waves learned from a population of individual Eigen-systems, where each native Eigen-system is regarded as a sample drawn from the Stiefel manifold. Specifically, a manifold optimization scheme is tailored to find the common harmonic waves, which reside at the center of the Stiefel manifold. To that end, the common harmonic waves constitute a new set of neurobiological bases to understand disease progression. Each harmonic wave exhibits a unique propagation pattern of neuropathological burden spreading across brain networks. The statistical power of our novel connectome harmonic analysis approach is evaluated by identifying frequency-based alterations relevant to Alzheimer's disease, where our learning-based manifold approach discovers more significant and reproducible network dysfunction patterns than Euclidean methods.**

***Index Terms*— Brain network, manifold optimization, harmonic waves, computer-assisted diagnosis.**

Manuscript received September 8, 2020; accepted September 29, 2020. Date of publication October 6, 2020; date of current version December 29, 2020. This work was supported in part by the NIH under Grant AG068399, Grant AG059065, and Grant AG049089; in part by the Key-Area Research and Development of Guangdong Province under Grant 2020B010166002 and Grant 2020B1111190001; in part by the National Natural Science Foundation of China under Grant 61771007; in part by the Health & Medical Collaborative Innovation Project of Guangzhou City under Grant 201803010021 and Grant 202002020049; in part by the China Scholarship Council under Grant 201906150061; and in part by the Alzheimer's Disease Neuroimaging Initiative (ADNI). *(Corresponding authors: Hongmin Cai; Guorong Wu.)*

Jiazhou Chen is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China, and also with the Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA (e-mail: csjzchen@mail.scut.edu.cn).

Guoqiang Han and Hongmin Cai are with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China (e-mail: csgqhan@scut.edu.cn; hmcai@scut.edu.cn).

Defu Yang is with the Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA (e-mail: defu@email.unc.edu).

Paul J. Laurienti is with the Department of Radiology, Wake Forest School of Medicine, Winston Salem, NC 27157 USA (e-mail: plaurien@wakehealth.edu).

Martin Styner and Guorong Wu are with the Department of Psychiatry and Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA (e-mail: styner@email.unc.edu; grwu@med.unc.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2020.3029063

I. Introduction

RECENT advances in neuroimaging offer an *in-vivo* and non-invasive window for investigating connectivity between brain regions [1]–[3]. For example, the combination of diffusion-weighted magnetic resonance imaging (DW-MRI) and tractography technology can be used to reconstruct major fiber bundles in the brain allowing for the visualization of the structural pathways that connect distant brain regions [4]. The ensemble of macroscopic brain connections can then be described as a complex network - the 'connectome'. Various computational and statistical inference methods have been developed to characterize diverse properties of complex networks and then identify network differences in terms of nodes, links, or even subgraphs that are associated with neurological disorders [1], [3].

Due to the high dimensionality of brain connectome data, it is a common practice to analyze node-wise graph variables such as local clustering coefficient and small-worldness [5], instead of using whole-brain connectivity information. By doing so, however, it becomes difficult to discover topological patterns which are an essential aspect of network analyses. On the other hand, there are also a plethora of methods proposed to quantify network changes at the level of individual links rather than nodes [1], [6]–[11]. Like node-wise

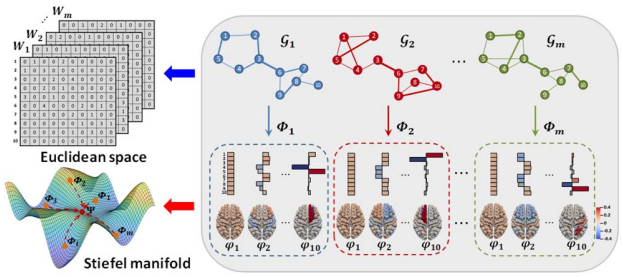


Fig. 1. Conventional methods (top) apply the classic Euclidean operations on the graph structure. Such rigid operations underestimate the irregular data structures and yield poor performances. In comparison, our learning-based approach (bottom) fully respects the irregular graph data structure and discovers brain network harmonics on the Stiefel manifold.

analyses, link-wise analyses are univariate in nature and disregard the multivariate network structure. In addition, due to high dimensionality, link-wise significance tests necessitate strict multiple-comparison correction to alleviate the issue of false positives, which potentially discards scientifically meaningful links [12].

Many neuroimaging studies have found that the progression of neuropathology follows the topology of large-scale networks in the brain [1], [3], [13]–[15]. For instance, a network diffusion model was used in [13], [14] to predict the disease progression in dementia, where the diffusion process is governed by the Laplacian matrix of the underlying brain network. Like various natural phenomena, the Eigen-system of the Laplacian also constitutes the basis of self-organizing patterns (shown in the bottom of **Fig. 1**), where each specific harmonic wave is indeed the Eigen-vector associated with a particular frequency (Eigen-value). Harmonic-based analyses have been used to investigate frequency-based alterations in neuropsychiatric diseases [16], [17] and functional neural activity [18]. Since the harmonic waves are orthogonal to each other, encoding brain connectivity via the harmonic domain offers great flexibility for the analysis of group differences.

As mentioned above, current harmonic analysis approaches have two major limitations. (1) *Lack of an unbiased reference to measure the difference between individuals.* In general, an unbiased reference domain is necessary for conducting group comparisons to provide standardized measurements for the statistical analyses. For example, since intrinsic structural differences are often mixed with external differences (such as the size and shape of the brain), an atlas image is used as a standard spatial reference for voxel-based morphometry (VBM) [19]. The morphometry differences of interest (such as gray matter density [20]) among the spatially normalized images are thought to be more relevant to neurobiological processes. Yet, different networks lead to various Eigen-systems, and thus a harmonic reference space for brain networks needs to represent the common set of harmonic waves that can appropriately represent the majority of the individual Eigen-systems. (2) *Lack of the appropriate manifold algebra.* Despite the well-known importance of a reference space in neuroimaging, finding such reference spaces for manifold data, such as brain networks, it is still an open problem as the complexity of data geometry (topology) is beyond regular data arrays [21], [22]. As shown in the top of **Fig. 1**, current approaches treat high-dimensional network data as a regular matrix or vector. Although applying Euclidean algebra to average brain networks [16] or diffuse connectivity information [23] on a link-wise basis is straightforward, the resulting group-mean network may no longer contain the essential network topology disrupting the geometry of Eigen-systems.

A. Relevant Works

The manifold learning technique has gained popularity in recent studies that use network analyses. For instance, a geodesic distance metric on Riemannian manifold was proposed in [24] to measure the similarity between two Laplacian matrices in a twin study. In addition, the kernel trick has been used to integrate the distance metric defined on Grassmannian manifold with a support vector machine to identify Autism subjects based on functional brain network data in [25]. Recently, Dai *et al.* proposed manifold based approaches to detect changing points [26] and compare functional brain networks [27], with the focus on symmetric positive definite (SPD) subspace of Riemannian manifold. Although none of the existing methods address the problem of unifying individual harmonic waves on the Stiefel manifold, manifold learning has been proven its efficiency in analyzing network and graph data that has intrinsic geometry.

B. Our Contributions

We propose a novel manifold learning method to discover the unbiased population mean of individual Eigen-systems. Since each Eigen-vector is orthogonal to all others, it is reasonable to assume that each Eigen-system behind the individual brain network resides on a high-dimensional Stiefel manifold [28]. Since each Eigen-system is uniquely associated with the corresponding underlying propagation patterns in the brain network, the well-studied Stiefel manifold (red arrow in **Fig. 1**) allows us to find a set of common bases that appropriately express the network propagation patterns, as compared to Euclidean algebra which lacks well-defined algebraic operators for manifold data.

Specifically, our method iteratively alternates two steps. (1) *Adjust each native Eigen-system toward the latent manifold mean.* The construction of each Eigen-system is not only influenced by the underlying Laplacian matrix but also attracted by the latent common harmonic waves at the manifold center. (2) *Update the manifold center.* We first project each Eigen-system to the tangent space at the current manifold center. Then, we estimate the mean tangent which points to the new location of the manifold mean. After mapping the mean tangent back to the Stiefel manifold, we can obtain the new estimation of the manifold center that is used to guide the refinement of individual Eigen-systems in Step 1. The outcome of our manifold optimization is a set of orthogonal vectors located at the manifold center, which represent the common harmonic waves learned over the population of brain networks.

As each harmonic wave exhibits a unique self-organized oscillation pattern across the brain network, our learned set of

TABLE I
LIST OF NOTATIONS USED IN THIS PAPER

Notation	Remark
$x, \boldsymbol{x}, \boldsymbol{X}$	Scalar, vector and matrix
$\mathcal{G}(V, \mathcal{E}, \boldsymbol{W})$	A graph $\mathcal{G}$ with nodes V, edges $\mathcal{E}$ and weights $\boldsymbol{W}$
$\boldsymbol{L}$	A Laplacian matrix of graph $\mathcal{G}$
$\boldsymbol{\Phi}_s$	s^{th} individual network harmonic waves
$\boldsymbol{\Psi}$	Common network harmonic waves
$\mathbb{R}^n$	n-dimensional real space
$\mathcal{O}^n$	Orthogonal group consisting of n-by-n orthogonal matrices
$\mathcal{M}_H$	The Stiefel manifold of harmonic waves
$\mathcal{V}(n, p)$	(n, p)-Stiefel manifold
$\mathcal{T}_{\boldsymbol{X}}, \Delta$	Tangent space and tangent vector of manifold at $\boldsymbol{X}$
exp	Exponential map
$F_{\boldsymbol{X}}$	Matrix derivative of some function F with respect to $\boldsymbol{X}$
$\nabla_{\boldsymbol{X}} F$	Gradient of F at point $\boldsymbol{X}$ in manifold space

common harmonic waves offers a new window to investigate the mechanism of neurodevelopment or neurodegeneration in the setting of brain networks using the classic physics concepts such as power and energy [17]. We have evaluated the statistical power of our new network harmonic analysis approach in a brain network study of Alzheimer's disease (AD). Compared to the conventional approach [16] using Euclidean operations, more statistically significant and reliable frequency-based alterations have been discovered using the common harmonic waves learned on Stiefel manifold.

II. METHOD

First, we provide the brief background on spectral graph theory and Stiefel manifold optimization in Section A. The motivation for discovering common harmonic waves for brain network analyses is explained in Section B. Then we present our manifold learning method for common harmonic waves in Section C, followed by the optimization scheme in Section D. The application of the learned common harmonic waves using a neuroimaging dataset is demonstrated in Section E. The notation used in this paper is summarized in **Table I** for ease of exposition.

A. Background

1) Graph Spectrum and Harmonic Waves: Each brain network can be encoded in a graph $\mathcal{G} = (V, \mathcal{E}, \boldsymbol{W})$, where $V = \{v_i | i \in 1, \cdots, n\}$ represents the node set with n nodes and $\mathcal{E} = \{e_{ij} | (v_i, v_j) \in V \times V\}$ is the set of all possible links. Let $\boldsymbol{W} \in \mathbb{R}^{n \times n}$ be a weighted adjacency matrix where each element w_{ij} in $\boldsymbol{W}$ measures the connectivity strength between node v_i and v_j. Then the symmetric graph Laplacian matrix $\boldsymbol{L}$ of the underlying graph can be calculated by:

$$\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W} \tag{1}$$

where $\boldsymbol{D} = diag(d_1, d_2, \ldots, d_n)$ is the degree matrix of the graph. Each diagonal element equals to the total connectivity degree of the underlying node, i.e., $d_i = \sum_{j=1}^{n} w_{ij}$.

A set of harmonic waves $\boldsymbol{\Phi}$ can be obtained by:

$$\min_{\boldsymbol{\Phi} \in \mathbb{R}^{n \times n}} tr\left(\boldsymbol{\Phi}^T \boldsymbol{L} \boldsymbol{\Phi}\right), \quad s.t.\ \boldsymbol{\Phi}^T \boldsymbol{\Phi} = \boldsymbol{I}_p \tag{2}$$

where $tr(\cdot)$ is the trace operator and $\boldsymbol{I}_p \in \mathbb{R}^{p \times p}$ stands for the identity matrix. The optimization problem in Eq. (2) has the deterministic solution $\widehat{\boldsymbol{\Phi}}$, which is the set of Eigen-vectors of the matrix $\boldsymbol{L}$. Without loss of generality, we can sort each Eigen-vector in $\widehat{\boldsymbol{\Phi}}$, column by column, in increasing order of Eigen-values. Given the connected graph $\mathcal{G}$ (i.e., no isolated nodes), the first smallest Eigen-value is always zero, and each element in the associated Eigen-vector is a constant. As the Eigen-value increases, the corresponding Eigen-vector exhibits more and high-frequency patterns (more rapid and localized oscillations) across the brain network, as shown in the bottom of **Fig. 1**.

The Stiefel manifold is a well-studied space and is defined as a set of ordered orthonormal p-frames of vectors in $\mathbb{R}^n$, denoted by $\mathcal{V}(n, p)$. In this context, any matrix $\boldsymbol{X} \in \mathbb{R}^{n \times p}$ can be regarded as a point sitting on Stiefel manifold $\mathcal{V}(n, p)$ as long as $\boldsymbol{X}$ is an orthogonal matrix, i.e., $\boldsymbol{X}^T \boldsymbol{X} = \boldsymbol{I}_p$. A special case is when $p = 1$, the Stiefel manifold reduces to the set of all unit vectors, which forms the unit sphere. Another special case is when $p = n$, the Stiefel manifold is the group of orthogonal $n \times n$ matrices $\mathcal{V}(n, n) \in \mathcal{O}^n$. Since there is no analytical formula for endpoint geodesics on the Stiefel manifold (*i.e.*, locally shortest length curves between two points $\boldsymbol{X} \in \mathcal{V}(n, p)$ and $\boldsymbol{Y} \in \mathcal{V}(n, p)$), it is common to approximate the geodesic between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in the ambient space by the following squared distance [28]:

$$d^2(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{2} tr(\boldsymbol{X} - \boldsymbol{Y})^T (\boldsymbol{X} - \boldsymbol{Y}) = p - tr(\boldsymbol{X}^T \boldsymbol{Y}) \tag{3}$$

2) Gradient Descent Optimization on Stiefel Manifold: For a point $\boldsymbol{X} \in \mathcal{V}(n, p)$, the tangent space $\mathcal{T}_{\boldsymbol{X}}$ at $\boldsymbol{X}$ consists of a set of tangents $\{\Delta\}$ such that $\boldsymbol{X}^T \Delta = \boldsymbol{0}$. Suppose a real-valued function F is smooth on the Stiefel manifold, the gradient of function F at $\boldsymbol{X}$, *i.e.*, $\nabla_{\boldsymbol{X}} F \in \mathcal{T}_{\boldsymbol{X}}$, can be obtained by [29]:

$$\nabla_{\boldsymbol{X}} F = F_{\boldsymbol{X}} - \boldsymbol{X} F_{\boldsymbol{X}}^T \boldsymbol{X} \tag{4}$$

where $F_{\boldsymbol{X}}$ stands for the matrix derivative of function F with respect to $\boldsymbol{X}$. The gradient calculation in Eq. 4 plays an important role in the application of Stiefel manifold optimization such as nonlinear mean shift [30]. The efficient calculation of the manifold gradient on the flattened tangent space offers a reasonable descent direction for optimizing function F on the manifold. After that, an exponential mapping operation is required to map a tangent $\Delta \in \mathcal{T}_X$ back onto the Stiefel manifold by:

$$exp_{\boldsymbol{X}}(\Delta) = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{Q}\boldsymbol{C} \tag{5}$$

where the matrices $\boldsymbol{B}$, $\boldsymbol{Q}$, and $\boldsymbol{C}$ are calculated in two steps. (1) Apply compact QR decomposition of $(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^T)\Delta$ [31] and thus obtain matrices $\boldsymbol{Q} \in \mathbb{R}^{n \times p}$ and $\boldsymbol{R} \in \mathbb{R}^{p \times p}$. (2) Solve $\boldsymbol{B} \in \mathbb{R}^{p \times p}$ and $\boldsymbol{C} \in \mathbb{R}^{p \times p}$ by:

$$\begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix} = \exp\left(\begin{bmatrix} \boldsymbol{A} & -\boldsymbol{R}^T \\ \boldsymbol{R} & \boldsymbol{0}_p \end{bmatrix}\right) \begin{bmatrix} \boldsymbol{I}_p \\ \boldsymbol{0}_p \end{bmatrix} \tag{6}$$

where $\boldsymbol{A} = \boldsymbol{X}^T \Delta$ is a $p \times p$ matrix and $\boldsymbol{0}_p \in \mathbb{R}^{p \times p}$ stands for the zero matrix.

B. Problem Statement

For each brain network $\mathcal{G}_s$ $(s = 1, \ldots, m)$, we can obtain its set of harmonic waves $\boldsymbol{\Phi}_s$ (an orthogonal matrix) by applying Eigen-decomposition on its Laplacian matrix $\boldsymbol{L}_s$. We are interested in finding the mean (denoted by $\boldsymbol{\Psi}$) of m individual harmonic waves $\{\boldsymbol{\Phi}_s | s = 1, \ldots, m\}$, where $\boldsymbol{\Psi}$ is called the common harmonic waves, which are also an orthogonal matrix.

Although it is efficient to calculate $\boldsymbol{\Psi}$ by simple averaging, *i.e.*, $\bar{\boldsymbol{\Psi}} = \frac{1}{m}\sum_{s=1}^{m} \boldsymbol{\Phi}_s$, the column vectors in $\bar{\boldsymbol{\Psi}}$ are not orthogonal to each other any longer, which compromises the applicability of $\bar{\boldsymbol{\Psi}}$ as the meaningful harmonic waves. Considering that each harmonic set $\boldsymbol{\Phi}_s$ resides on the high dimensional Stiefel manifold $\mathcal{M}_H$, the common harmonic waves $\bar{\boldsymbol{\Psi}}$ generated by arithmetic averaging results in a $\bar{\boldsymbol{\Psi}}$ may not be directly located on the same manifold as all the individual harmonic waves.

An alternative way is to average over the adjacency matrices $\{\boldsymbol{W}_s | s = 1, \ldots, m\}$ and then calculate the Eigen-system $\tilde{\boldsymbol{\Psi}}$ based on the Laplacian matrix which is derived from the average of adjacency matrices $\bar{\boldsymbol{W}} = \frac{1}{m}\sum_{s=1}^{m} \boldsymbol{W}_s$. However, such Euclidean operations are highly sensitive to noises and/or outlying data points. In addition, a heuristic assumption that the intrinsic complex geometry of brain network and Eigen-system data can be well expressed in Euclidean space is difficult to satisfy. As pointed in [32], the arithmetic average of adjacency matrices suffers from *swelling effect*, i.e., the determinant of the average matrix is often much larger than any of the original determinants. Such an inflated determinant artificially disrupts the essential topological properties of brain networks.

Given that each harmonic set $\boldsymbol{\Phi}_s$ is an orthogonal matrix, it is reasonable to consider finding the latent common harmonic waves $\boldsymbol{\Psi}$ on the Stiefel manifold. As we will show in Section II.C, there is no closed-form solution for calculating common harmonic waves. Instead, we propose a gradient descent based optimization scheme to iteratively find the common harmonic waves $\boldsymbol{\Psi}$. To that end, manifold optimization is of necessity to maintain the intrinsic data geometry.

It is worth noting that the graph spectrum of each brain network is spanned by its harmonic waves $\boldsymbol{\Phi}_s$, sorted from low frequency to high frequency [21]. Since the harmonic waves associated with high frequency (larger eigenvalues) are more sensitive to possible noise, we only consider the first p $(p \leq n)$ harmonic waves in each $\boldsymbol{\Phi}_s$. In the following, we regard $\boldsymbol{\Phi}_s \in \mathcal{M}_H$ $(\mathcal{M}_H \subset \mathcal{V}(n, p))$ as an $n \times p$ orthogonal matrix unless otherwise stated.

C. Learning Common Harmonic Waves on Stiefel Manifold

Given m Laplacian matrices $\{\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_m\}$, we simultaneously estimate the native harmonic waves $\boldsymbol{\Phi}_s$ for each $\boldsymbol{L}_s$ and optimize the common harmonic waves $\boldsymbol{\Psi}$, which are both optimized on the Stiefel manifold $\mathcal{M}_H$. Specifically, we require the latent common harmonic waves to be located at the manifold center that has the shortest geodesic distances to all individual harmonic waves $\{\boldsymbol{\Phi}_s\}$. To that end, we opt to minimize $\sum_{s=1}^{m} d^2(\boldsymbol{\Phi}_s, \boldsymbol{\Psi})$. By integrating Eq. 2 and Eq. 3, the objective function becomes:

$$\begin{aligned} &\min_{\{\boldsymbol{\Phi}_s\}, \boldsymbol{\Psi}} \sum_{s=1}^{m} \left\{ tr\left(\boldsymbol{\Phi}_s^T \boldsymbol{L}_s \boldsymbol{\Phi}_s\right) + \lambda(p - tr(\boldsymbol{\Phi}_s^T \boldsymbol{\Psi}) \right\} \\ &s.t.\ \forall s : \boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s = \boldsymbol{I}_p \end{aligned} \tag{7}$$

where λ is a scalar balancing two terms in Eq. 7. Specifically, the trace norm (the first term in Eq. 7) is used to require each set of harmonic waves $\boldsymbol{\Phi}_s$ reflect the graph spectrum of its own Laplacian matrix $\boldsymbol{L}_s$. The second term is a manifold distance constraint that ensures the common harmonic waves $\boldsymbol{\Psi}$ is close to all individual harmonic waves $\{\boldsymbol{\Phi}_s\}$. Since the trace norm is not exactly an intrinsic manifold metric, we keep the orthogonal constraint, i.e., $\boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s = \boldsymbol{I}_p$ $(\forall s)$, in the objective function, to ensure the estimated $\{\boldsymbol{\Phi}_s\}$ are located on the Stiefel manifold. Since the optimization of $\boldsymbol{\Psi}$ is completely driven by the intrinsic Stiefel manifold geometry (sub-problem 2 in Section II.D), we do not explicitly add the constraint $\boldsymbol{\Psi}^T \boldsymbol{\Psi} = \boldsymbol{I}_p$ in the objective function.

Given that it is computationally expensive to estimate individual and common harmonic waves jointly, we propose the following gradient descent manifold optimization under the framework of ADMM (Alternating Direction Method of Multipliers) [33], where the augmented Lagrange function becomes:

$$\begin{aligned} &\arg\min_{\{\boldsymbol{\Phi}_s\}, \boldsymbol{\Psi}} \sum_{s=1}^{m} F_{\boldsymbol{\Phi}_s, \boldsymbol{\Psi}} \\ &= \arg\min_{\{\boldsymbol{\Phi}_s\}, \boldsymbol{\Psi}} \sum_{s=1}^{m} \Big\{ tr\left(\boldsymbol{\Phi}_s^T \boldsymbol{L}_s \boldsymbol{\Phi}_s\right) + \lambda\left(p - tr\left(\boldsymbol{\Phi}_s^T \boldsymbol{\Psi}\right)\right) \\ &\quad + tr(\boldsymbol{\Lambda}_s^T(\boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s - \boldsymbol{I}_p)) \Big\}, \end{aligned} \tag{8}$$

where each $\boldsymbol{\Lambda}_s$ $(s = 1, \ldots, m)$ is a $p \times p$ factor matrix of the Lagrange multipliers.

D. Optimization Scheme

We optimize the objective function in Eq. 8 in two alternative steps.

Sub-Problem 1 (Estimating Each Native Harmonic set $\boldsymbol{\Phi}_s$): Since the harmonic waves $\{\boldsymbol{\Phi}_s\}$ are independent, we can estimate each $\boldsymbol{\Phi}_s$ separately by fixing $\boldsymbol{\Psi}$. By dropping the unrelated variables, the objective function becomes:

$$\begin{aligned} \arg\min_{\boldsymbol{\Phi}_s} F_{\Phi_s} = arg\min_{\boldsymbol{\Phi}_s} \Big\{ &tr\left(\boldsymbol{\Phi}_s^T \boldsymbol{L}_s \boldsymbol{\Phi}_s - \lambda \boldsymbol{\Phi}_s^T \boldsymbol{\Psi}\right) \\ &+ tr(\boldsymbol{\Lambda}_s^T(\boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s - \boldsymbol{I}_p)) \Big\} \end{aligned} \tag{9}$$

It is worth noting that the individual harmonic waves $\boldsymbol{\Phi}_s$ are not only determined by their own Laplacian matrix $\boldsymbol{L}_s$, but are also influenced by the latent common harmonic waves $\boldsymbol{\Psi}$. Since Eq. 9 is a typical quadratic problem on the Stiefel manifold, it is often required that $\boldsymbol{L}_s$ is positive definite. Therefore, we first replace $\boldsymbol{L}_s$ with $\tilde{\boldsymbol{L}}_s = \beta \boldsymbol{I} - \boldsymbol{L}_s$, where a relaxation parameter β is used to ensure $\tilde{\boldsymbol{L}}_s$ is a positive definite matrix. We set β as the greatest eigenvalue of $\boldsymbol{L}_s$. By doing so,

the minimization of Eq. 9 becomes the maximizing:

$$\arg\max_{\boldsymbol{\Phi}_s} \tilde{F}_{\boldsymbol{\Phi}_s} = \arg\max_{\boldsymbol{\Phi}_s} \Big\{ tr\left(\boldsymbol{\Phi}_s^T \tilde{\boldsymbol{L}}_s \boldsymbol{\Phi}_s\right) + \lambda tr\left(\boldsymbol{\Phi}_s^T \boldsymbol{\Psi}\right) - tr(\boldsymbol{\Lambda}_s^T(\boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s - \boldsymbol{I}_p)) \Big\} \quad (10)$$

The last term in Eq. 10 is the Lagrange multiplier. We can solve $\boldsymbol{\Phi}_s$ via the KKT condition as:

$$\frac{\partial \tilde{F}_{\boldsymbol{\Phi}_s}}{\partial \boldsymbol{\Phi}_s} = 2\tilde{\boldsymbol{L}}_s \boldsymbol{\Phi}_s + \lambda \boldsymbol{\Psi} - 2\boldsymbol{\Phi}_s \boldsymbol{\Lambda}_s = 0 \quad (11)$$

To overcome the instability as well as reduce the computational cost caused by the matrix inversion involved in $\tilde{\boldsymbol{L}}_s$ and $\boldsymbol{\Lambda}_s$, we adopt the generalized power iteration (GPI) from [34] to iteratively optimize $\boldsymbol{\Phi}_s$ in the following four steps:

(1) Initialize $\boldsymbol{\Phi}_s$ as the Eigen-vector matrix after applying SVD to the underlying Laplacian matrix $\boldsymbol{L}_s$.
(2) Update $\boldsymbol{\Theta} \leftarrow \tilde{\boldsymbol{L}}_s \boldsymbol{\Phi}_s + \lambda \boldsymbol{\Psi}$.
(3) Calculate $\boldsymbol{\Phi}_s$ by maximizing $tr(\boldsymbol{\Phi}_s^T \boldsymbol{\Theta})$ and subject it to the orthogonal constraint $\boldsymbol{\Phi}_s^T \boldsymbol{\Phi}_s = \boldsymbol{I}_p$. We can derive the closed-form solution by $\boldsymbol{\Phi}_s = \boldsymbol{U}\boldsymbol{V}^T$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right Eigen matrix after the full SVD on $\boldsymbol{\Theta}$. (please refer to [34] for detail)
(4) Iteratively perform the steps (2)-(3) until convergence.

Sub-Problem 2 (Estimating the Common Harmonic Set $\boldsymbol{\Psi}$): Given the individual harmonic waves $\{\boldsymbol{\Phi}_s\}$, the objective function of $\boldsymbol{\Psi}$ is:

$$\arg\min_{\boldsymbol{\Psi}} \sum_{s=1}^{m} d^2(\boldsymbol{\Phi}_s, \boldsymbol{\Psi}) = \arg\min_{\boldsymbol{\Psi}} \sum_{s=1}^{m} (p - tr(\boldsymbol{\Phi}_s^T \boldsymbol{\Psi})) \quad (12)$$

The intuition behind in Eq. 12 is to find the latent mean $\boldsymbol{\Psi}$ on the Stiefel manifold which has the shortest geodesic distances to all the observed samples $\{\boldsymbol{\Phi}_s\}$ residing on the Stiefel manifold. Thus, our optimization falls into the classic problem of solving the Fréchet mean on the Stiefel manifold which can be efficiently solved by adopting the Weiszfeld algorithm [35]. Specifically, we alternately perform the following two steps until convergence:

(1) Suppose $\boldsymbol{\Psi}^{(k)}$ is the current estimation of the manifold center. We calculate the gradient $\nabla_{\boldsymbol{\Psi}}$ of the energy function in Eq. 12 with respect to each $\boldsymbol{\Phi}_s$ as: $\nabla_{\boldsymbol{\Psi}} d^2\left(\boldsymbol{\Psi}^{(k)}, \boldsymbol{\Phi}_s\right) = \boldsymbol{\Psi}^{(k)} \boldsymbol{\Phi}_s^T \boldsymbol{\Psi}^{(k)} - \boldsymbol{\Phi}_s$, which are denoted by the black arrows in **Fig. 2**. Then, the mean tangent $\Delta\boldsymbol{\Psi}^{(k+1)} \in \mathcal{T}_{\boldsymbol{\Psi}}$ can be efficiently obtained by:

$$\begin{aligned} \Delta\boldsymbol{\Psi}^{(k+1)} &= -\sum_{s=1}^{m} \nabla_{\boldsymbol{\Psi}} d^2\left(\boldsymbol{\Psi}^{(k)}, \boldsymbol{\Phi}_s\right) \\ &= -\sum_{s=1}^{m} (\boldsymbol{\Psi}^{(k)} \boldsymbol{\Phi}_s^T \boldsymbol{\Psi}^{(k)} - \boldsymbol{\Phi}_s) \end{aligned} \quad (13)$$

As demonstrated in [35], $\Delta\boldsymbol{\Psi}^{(k+1)}$ (red triangle on tangent plane $\mathcal{T}_{\boldsymbol{\Psi}^{(k)}}\mathcal{M}_H$ in **Fig. 2**) is the updated position of the estimated mean and the red arrow specifies the direction from the prior estimation $\boldsymbol{\Psi}^{(k)}$ to the new latent mean on the manifold.

(2) We map the mean tangent $\Delta\boldsymbol{\Psi}^{(k+1)}$ back to the Stiefel manifold to obtain the new estimation of the manifold center $\boldsymbol{\Psi}^{(k+1)} = exp_{\boldsymbol{\Psi}^{(k)}}(\Delta\boldsymbol{\Psi}^{(k+1)})$ (red circle in **Fig. 2**) by Eq. 5-6.

By iteratively calculating the optimal descent direction and mapping it back to the Stiefel manifold, we can obtain the optimal manifold mean $\boldsymbol{\Psi}$, *i.e.,* the common harmonic waves. The entire optimization scheme is summarized in **Table II**.

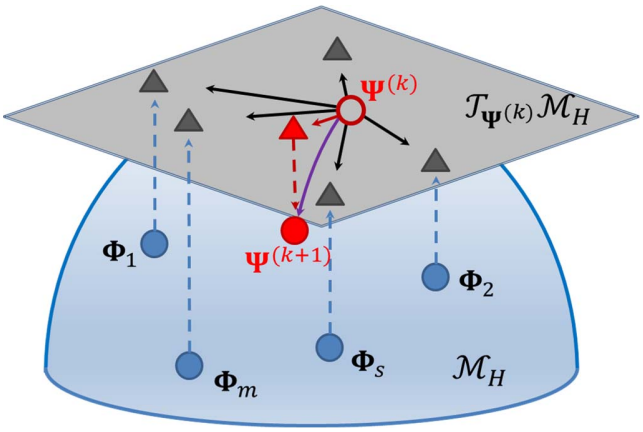


Fig. 2. Illustration for the optimization of common harmonic waves on the Stiefel manifold. Individual harmonic waves (blue solid circle) are located on the Stiefel manifold M_H (blue hemisphere). The individual harmonic waves are projected to corresponding point (grey triangle) in the tangent space $T_{\boldsymbol{\Psi}^{(k)}}M_H$ (grey flat plane) of k^{th} manifold mean $\boldsymbol{\Psi}^{(k)}$ (red hollow circle). The mean tangent (red triangle) is calculated based on all gradient directions (black arrow). Finally, the new manifold mean $\boldsymbol{\Psi}^{(k+1)}$ (red solid circle) is estimated by mapping the mean tangent back to the Stiefel manifold. The manifold trajectory from $\boldsymbol{\Psi}^{(k)}$ to $\boldsymbol{\Psi}^{(k+1)}$ is depicted by purple arrow.

Discussion: Regarding the dimension reduction, we determine p based on the distribution reconstruction error between the observed Laplacian matrix and the reconstructed Laplacian matrix using only the top p smallest Eigen-values and Eigen-vectors. Empirically, we select p around the tipping point that the decrease of reconstruction error is only marginal as p increases. Since our objective function (Eq. 7) is not invariant to any orthogonal matrix $\boldsymbol{Q}$ ($\boldsymbol{Q} \in \mathbb{R}^{n \times p}$, $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$), we opt to optimize the common harmonic waves $\boldsymbol{\Psi}$ on the Stiefel manifold, instead of the Grassmann manifold [36].

It is worth noting that the overarching goal is to find the unbiased common harmonic bases which allow us to understand the cortical spreading pathway of neurodegenerative burdens. In this work, we only focus on unifying harmonic waves, rather than the entire Eigen-system which includes both Eigen-vectors and Eigen-values. Thus, our proposed method is not designed to establish the one-to-one correspondence of harmonic waves across individuals. As shown in **Table II**, our optimization scheme consists of two subproblems. The proof of solving Eq. 9 using GPI is supported by the theorem 1-2 in [34]. The Weiszfeld algorithm [35] has been used in many computer vision applications with proof of convergence.

E. Application in Network Neuroscience

Advanced neuroimaging technology such as MRI and diffusion-weighted MRI allows us to study white matter fiber tracks associated with the progression of cognitive decline. Mounting evidence shows that neurodegenerative diseases such as Alzheimer's disease (AD) can be understood as a disconnection syndrome where the large-scale brain network is progressively disrupted by neuropathological processes [3]. Our proposed harmonic-based network analysis approach

TABLE II
ALGORITHM FOR COMMON HARMONICS DETECTION

Parameters: $\lambda; \varepsilon_1; \varepsilon_2; \gamma;$

Input: **Adjacency matrix** $\boldsymbol{W}_s \in \mathbb{R}^{n\times n},\ s = 1,2,\cdots,m$

Init. Calculate Laplacian matrix $\boldsymbol{L}_s = \boldsymbol{D}_s - \boldsymbol{W}_s$, where $d_i = \sum_{j=1}^{n} \boldsymbol{W}_{ij}$;
Calculate positive definite matrix $\tilde{\boldsymbol{L}}_s = \beta \boldsymbol{I} - \boldsymbol{L}_s$. Set β be the dominant Eigen-value of $\boldsymbol{L}_s$;
Initialize common network harmonic waves $\boldsymbol{\Psi} = eig(\frac{1}{m}\sum_{s=1}^{m} \boldsymbol{L}_s) \in \mathbb{R}^{n\times p}$;
Initialize orthogonal matrix $\boldsymbol{\Phi}_s \in \mathbb{R}^{n\times p}$ through the Eigen-decomposition of Laplacian matrix $\boldsymbol{L}_s$;
Initialize parameter $\lambda = 0.01, \gamma = 0.01$;

do

For $s = 1,2,\cdots,m$ **do**

do

Update $\boldsymbol{\Theta} = \tilde{\boldsymbol{L}}_s \boldsymbol{\Phi}_s^{(k)} + \lambda \boldsymbol{\Psi}$;
Compute $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \boldsymbol{\Theta}$ via the compact SVD method of $\boldsymbol{\Theta}$ where $\mathbf{U} \in \mathbb{R}^{n\times p}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{p\times p}$ and $\mathbf{V} \in \mathbb{R}^{p\times p}$;
Update $\boldsymbol{\Phi}_s^{(k+1)} = \boldsymbol{U}\boldsymbol{V}^T$;

While $\left\| \boldsymbol{\Phi}_s^{(k+1)} - \boldsymbol{\Phi}_s^{(k)} \right\| < \varepsilon_1$

End for

Set start point $\boldsymbol{\Psi}^{(1)} = \boldsymbol{\Phi}_1$;

do

$\Delta\boldsymbol{\Psi}^{(k+1)} = -\lambda \sum_{s=1}^{m} (\boldsymbol{\Psi}^{(k)} \boldsymbol{\Phi}_s^T \boldsymbol{\Psi}^{(k)} - \boldsymbol{\Phi}_s)$;
$\boldsymbol{\Psi}^{(k+1)} = exp_{\boldsymbol{\Psi}^{(k)}}(\gamma \Delta\boldsymbol{\Psi}^{(k+1)})$;

While $\|\Delta\boldsymbol{\Psi}\| < \varepsilon_2$

Update $\boldsymbol{\Psi} = \boldsymbol{\Psi}^{(k+1)}$;
Compute $New_{\text{cost}} = \sum_{s=1}^{m} \{tr(\boldsymbol{\Phi}_s^T \boldsymbol{L}_s \boldsymbol{\Phi}_s) + \lambda d^2(\boldsymbol{\Psi}, \boldsymbol{\Phi}_s)\}$;
$\varepsilon = \text{abs}(New_{\text{cost}} - Old_{\text{cost}})$;
Update $Old_{\text{cost}} = New_{\text{cost}}$;

While ε is less than a pre-defined threshold.

Output Common network harmonic waves $\boldsymbol{\Psi}$.

provides a new methodology to analyze these spatio-temporal neuropathological changes in the progression of AD. Specifically, we assert that the brain is capable of vibrating at a preference frequency (aka. resonance frequency) like any natural object in the universe. In this context, we hypothesize that the resonance between the spatial pattern of pathological burdens and the oscillation pattern in the harmonic waves is the driving force of spreading the neurodegenerative process.

1) Image Processing: As shown in **Fig. 3**, for training data, according to a Desctrieux atlas [37], we parcellate the cortical surface into 148 cortical regions based T1-weighted MR image and then apply surface seed-based probabilistic fiber tractography [37] using the diffusion tensor imaging (DTI) data, thus producing a 148×148 anatomical connectivity matrix. Note, the weight of the anatomical connectivity is defined by the number of fibers linking two brain regions normalized by the total number of fibers in the whole brain. For testing data, we calculate the mean cortical thickness as well as the standard uptake value ratio (SUVR) of the amyloid deposition for each brain region and then assemble them into a column vector, denoted by $\boldsymbol{f}^s$.

2) Harmonic Analyses: Harmonic analysis for network neuroscience was explored in [13], [16], where the well-studied

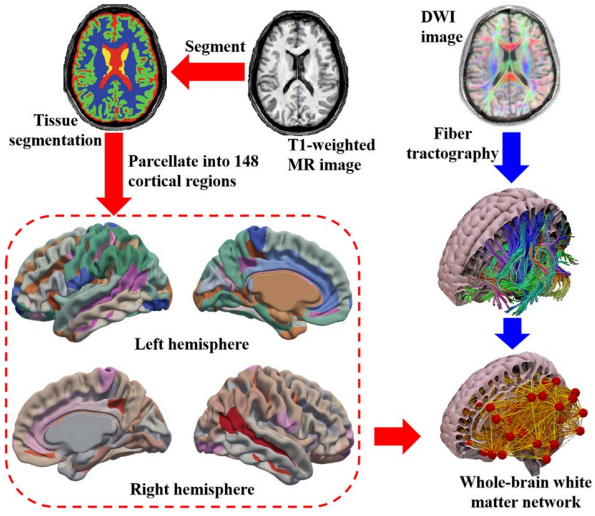


Fig. 3. Image processing pipeline to construct structural brain network from diffusion-weighted MR images.

physics concepts of harmonics such as power and energy have been introduced to quantify the neural activities on brain networks. Specifically, the power of each common harmonic wave $\boldsymbol{\psi}_h$ (h^{th} column vector in $\boldsymbol{\Psi}$) measures the velocity of flow of brain activity for each unique self-organized oscillation pattern across brain networks. The energy of each $\boldsymbol{\psi}_h$ measures the amount of flow of brain activity.

Conventional neuroimaging studies use empirical biomarkers such as region-wise cortical thickness or amyloid level to investigate the pathophysiological mechanism of Alzheimer's disease. In this work, we present a novel harmonic analysis approach to discover the frequency-based alterations that are associated with the cortical spreading pathway of neurodegenerative burdens across brain networks. For each common harmonic wave, we first calculate the harmonic power coefficient of $\boldsymbol{\psi}_h$ to the observed empirical biomarker level (cortical thickness or amyloid deposition) vector $\boldsymbol{f}^s$ of s^{th} individual subject by:

$$\alpha_h^s = \langle \boldsymbol{f}^s, \boldsymbol{\psi}_h \rangle \tag{14}$$

Furthermore, we can calculate the corresponding harmonic-specific energy of $\boldsymbol{\psi}_h$ by:

$$E_h^s = \left| a_h^s \right|^2 \tag{15}$$

The total harmonic energy of brain injury (manifested by cortical thickness or amyloid deposition) with respect to the harmonic waves $\boldsymbol{\Psi}$ is measured by:

$$E_{total}^s = \sum_{h=1}^{p} E_h^s \tag{16}$$

III. RESULTS

To evaluate the power of our new network harmonic analysis approach, we compare the performance of using the common harmonics $\boldsymbol{\Psi}$ optimized by our manifold learning method to three alternative sets of harmonic waves: (1) $\bar{\boldsymbol{\Psi}}$ by simple averaging individual Eigen-systems, (2) $\tilde{\boldsymbol{\Psi}}$ by first averaging

the adjacency matrices and then applying SVD to the average adjacency matrix. and (3) $\check{\boldsymbol{\Psi}}$ by only estimating Fréchet mean from a population of individual Eigen-systems on Stiefel manifold without adjusting each individual Eigen-system (a simplified version of our proposed method). In the following, we call $\bar{\boldsymbol{\Psi}}$, $\tilde{\boldsymbol{\Psi}}$, $\check{\boldsymbol{\Psi}}$ and $\boldsymbol{\Psi}$ as the arithmetic mean (aka. arithmetic harmonic waves), the pseudo manifold mean (aka. pseudo harmonic waves), the Fréchet mean (aka. Fréchet harmonic waves), and Stiefel manifold mean (aka. our common harmonic waves), respectively. In addition, we adopt a grid search strategy to determine the optimal value for parameter λ from 0.0005 to 0.2, where we find $\lambda = 0.01$ gives the highest classification between CN and AD subjects (see Section III.B4 for detail). Thus, we fix $\lambda = 0.01$ in all the following experiments.

A. Experiments on Synthetic Data

First, we use the following low dimension synthetic data to demonstrate the accuracy of the manifold center by our proposed method. Here, we synthesize a set of 3D orthonormal rotation matrices as individual harmonic waves, which are represented as the unit quaternions. A quaternion is denoted as $q = (a, v)$, where a is the real quantity and $v = bi + cj + dk$ with three imaginary quantities (b, c, d). Let $e = (1, 0)$ be the identity quaternion. The transformation between quaternion and rotation matrix can be represented as

$$R = \begin{bmatrix} 1-2c^2-2d^2 & 2bc-2ad & 2ac+2bd \\ 2bc+2ad & 1-2b^2-2d^2 & 2cd-2ad \\ 2bd-2ac & 2ab+2cd & 1-2b^2-2c^2 \end{bmatrix} \tag{17}$$

where $a = \cos\left(\frac{1}{2}\theta\right)$, $b = \sin\left(\frac{1}{2}\theta\right)u_x$, $c = \sin\left(\frac{1}{2}\theta\right)u_y$, $d = \sin(\frac{1}{2}\theta)u_z$, as well as θ and $\boldsymbol{u} = (u_x, u_y, u_z)$ denote rotation angle and rotation axis, respectively.

We generate a random collection of twenty quaternions as follows. First, we set the quaternion with no rotation as the ground truth (starting point), which is displayed in green in Fig. 4. Second, given rotation axis $\boldsymbol{u}$, the rotation angles are sampled from a zero-mean Gaussian distribution with a standard deviation $\sigma = \pi/15$. Third, twenty rotation matrices are obtained through Eq. 17, centered on the identity matrix. Among them, ten rotation matrices are shown in the first two rows of Fig. 4.

Since we do not have the adjacency matrices, we apply naïve averaging and our Stiefel manifold learning method to estimate the common quaternion from the 20 random perturbative quaternions. The arithmetic mean $\bar{\boldsymbol{\Psi}}$ and our Stiefel manifold mean $\boldsymbol{\Psi}$ are shown in Fig. 4(b) and (c), respectively. It is clear that (1) The Stiefel mean is very close to the ground truth on the manifold; (2) The arithmetic mean is not located on the manifold surface (non-orthogonal matrix), as the three rotation axes are not perpendicular to each other; (3) Our iterative manifold optimization can quickly converge to the latent manifold mean, as indicated by the red trajectory in Fig. 4. Although we initialize our optimization from a single individual quaternion (#10) in Fig. 4, no significant difference has been found across the Stiefel mean results initialized with different individual quaternions. We will further examine the replicability in Section III.B2.

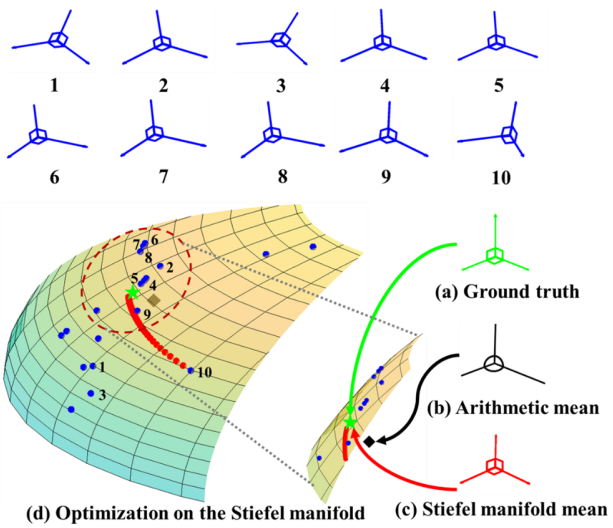


Fig. 4. Comparison of arithmetic mean and our Stiefel manifold mean across individual orthonormal matrices. Top: 10 examples of matrices (first two rows) generated by different rotation of identity matrix (a). Bottom: the ground truth (a), arithmetic mean (b), and our Stiefel manifold mean (c) on the Stiefel manifold (d). It is clear that the manifold center estimated by our manifold optimization method is more reasonable than the arithmetic mean which uses simple averaging operation defined in the Euclidean space.

TABLE III
DEMOGRAPHIC INFORMATION OF TRAINING DATA IN ADNI DATABASE

Gender	Number	Age Range	Mean Age	CN	MCI	AD
Male	47	55.0~90.3	74.3	12	16	19
Female	47	55.6~87.8	73.0	23	8	16
Total	**94**	**55.0~90.3**	**73.7**	**35**	**24**	**35**

B. Experiments on Real Data of Alzheimer's Disease

1) Description of Datasets and Experiment Setup:

a) Training data for learning common harmonic waves: In total, 94 subjects were selected from the ADNI database to learn the common harmonic waves, which consisted of 35 Cognitive Normal (CN), 24 Mild Cognitive Impairment (MCI), and 35 Alzheimer's Disease (AD). Each subject had both T1-weighted MRI and diffusion-weighted MRI scans. The demographic information is shown in Table III. Following the image processing pipeline in Fig. 3, we constructed the structural network for each subject which consisted of 148 nodes.

b) Testing data for identifying frequency-based alterations in AD: In addition, we selected another 50 CN subjects, 44 MCI subjects, and 47 AD subjects from ADNI data as the testing data. As was done with the training data, the cortical surface of each subject was parcellated into 148 regions, and the mean cortical thickness and the standard uptake value ratio (SUVR) of amyloid deposition for each region were computed.

For testing data, we no longer provided the structural network information.

c) Experiment setup on real AD network data: In the following experiments, we compare the performance among the arithmetic mean $\bar{\boldsymbol{\Psi}}$, the pseudo manifold mean $\tilde{\boldsymbol{\Psi}}$, the Fréchet mean $\check{\boldsymbol{\Psi}}$, and the Stiefel manifold mean $\boldsymbol{\Psi}$. The number of harmonic waves p is set to 60 and the parameter λ is set to 0.01. First, we evaluate the replicability of the common harmonic waves using our proposed learning-based method via a resampling test in Section III.B2. Next, we investigate whether the oscillation patterns in the common harmonic waves underline the neurodegenerative process in Section III.B3. Furthermore, we evaluate the diagnostic value of our harmonic feature representations in Section III.B4, with the comparison to empirical AD biomarkers. Finally, we apply the learned common harmonic waves to identify harmonic-based alterations in the cortical thickness (Section III.B5) and amyloid deposition data (Section III.B6).

2) Evaluation of the Replicability: In this experiment, we evaluate the replicability of the learned common harmonic waves via resampling tests. Specifically, we apply the following resample procedure to generate 50 test/retest datasets from the training data: (1) randomly sample 70 networks from the 94 training network data; (2) continue to sample another two sets of networks from the remaining 24 subjects separately, each with 5 networks; (3) form two paired cohorts by combining the networks sampled in step 1 and 2. Then, we deploy our Stiefel manifold learning method on two datasets independently. Because two paired cohorts only have 6.7% (5/75) differences in terms of network data, we can evaluate the replicability of our method by examining whether there exists a significant difference at each element in the harmonic waves via the paired t-test. Fewer elements showing significance indicates better replicability. Since each row in the harmonic matrix is associated with one brain region, we can map the significant findings ($p < 0.01$) onto the cortical surface in **Fig. 5**(a-d). It is apparent that (1) The manifold learning methods, Fréchet mean $\check{\boldsymbol{\Psi}}$ and our manifold mean $\boldsymbol{\Psi}$, yield more consistent (more replicable) common harmonic waves than pseudo manifold mean $\tilde{\boldsymbol{\Psi}}$ across the test/retest datasets in the resampling test, while our common harmonics $\boldsymbol{\Psi}$ achieves even better replicability than Fréchet common harmonics $\check{\boldsymbol{\Psi}}$; (2) Although the arithmetic mean $\bar{\boldsymbol{\Psi}}$ has similar replicability performance as our common harmonic waves $\boldsymbol{\Psi}$, the orthogonality is often not guaranteed in $\bar{\boldsymbol{\Psi}}$ by arithmetic averaging (see the results on synthesized data in **Fig. 4(b)**).

3) Association Between the Oscillation Patterns in Harmonic Waves and Neurodegenerative Process: In this experiment, we are investigating whether the learned common harmonic waves capture information related to the neurodegenerative process in AD. After we compute the common harmonic waves on the training data, we repeat the following steps on the testing data with 100 replicates: (1) randomly select 30 out of 47 AD subjects and 30 out of 50 CN subjects as training data and form the amyloid vector $\boldsymbol{f}$; (2) identify the harmonic power difference α_h between CN and AD for each harmonic wave $\boldsymbol{\psi}_h$; (3) calculate positive power $\alpha_h^+ = \langle \boldsymbol{f}, \boldsymbol{\psi}_h^+ \rangle$ and negative power $\alpha_h^- = \left| \langle \boldsymbol{f}, \boldsymbol{\psi}_h^- \rangle \right|$ of the remaining 37 subjects (testing data), where $\boldsymbol{\psi}_h^+$ and $\boldsymbol{\psi}_h^-$ present the positive-only and negative-only segments in each $\boldsymbol{\psi}_h$; (4) apply the t-test to detect the statistical CN vs AD difference of $\left| \alpha_h^+ - \alpha_h^- \right|$ on each harmonic identified in (2). Since we assume that the resonance between the spatial pattern of AD biomarkers and the oscillation pattern in the harmonic waves is the driving force of the neurodegenerative process in AD, $\left| \alpha_h^+ - \alpha_h^- \right|$ approximately reflects the potential of neurodegenerative burdens between up peaks (positive segment) and bottom peaks (negative segment) within each harmonic wave $\boldsymbol{\psi}_h$.

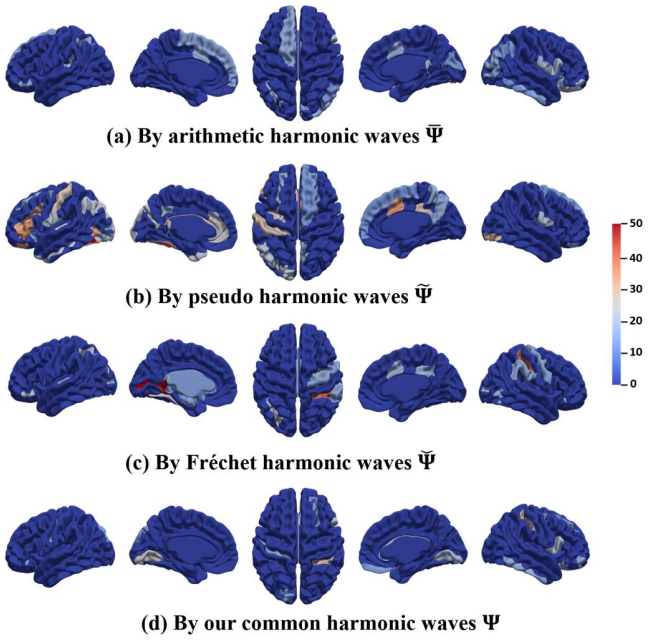


Fig. 5. The replicability test results of the arithmetic harmonic waves $\bar{\boldsymbol{\Psi}}$ (a), pseudo harmonic waves $\tilde{\boldsymbol{\Psi}}$ (b), Fréchet harmonic waves $\check{\boldsymbol{\Psi}}$ (c) and our common harmonic waves $\boldsymbol{\Psi}$ (d), where the color on the cortical surface reflects the number of times with failed replicability tests.

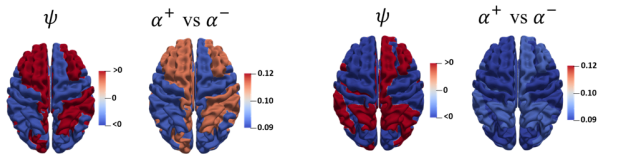
TABLE IV
STATISTICS OF KINETIC POTENTIALS

Methods	Har. Power	+/- Power
Arithmetic harmonic waves $\bar{\boldsymbol{\Psi}}$	7.69 ± 0.94	3.67 ± 0.65
Pseudo harmonic waves $\tilde{\boldsymbol{\Psi}}$	8.01 ± 0.89	1.93 ± 0.99
Fréchet harmonic waves $\check{\boldsymbol{\Psi}}$	8.76 ± 0.97	5.30 ± 0.67
Our common harmonic waves $\boldsymbol{\Psi}$	9.87 ± 1.08	6.16 ± 0.86

First, we find that on average 12.8% (7.69/60) arithmetic harmonic waves, 13.4% (8.01/60) pseudo harmonic waves, 14.6% (8.76/60) Fréchet harmonic waves, and 16.5% (9.87/60) our common harmonic waves show significant harmonic power differences (FDR-adjusted $p < 0.01$) in the amyloid data in step (2) using the Benjamini-Hochberg (BH) procedure [38]. The detailed statistics are shown in **Table IV**. It is clear that the oscillation patterns exhibiting in our common harmonic waves are more aligned with the neurodegenerative process in AD.

TABLE V
CLASSIFICATION RESULTS OF USING EMPIRICAL BIOMARKER FEATURES AND HORMONIC-BASED FEATURE REPRESENTATIONS

Data	Methods	Accuracy	Sensitivity	Specificity	F-score
Cortical thickness	Empirical cortical thickness	0.624 ± 0.032	0.618 ± 0.035	0.645 ± 0.036	0.610 ± 0.034
	Arithmetic harmonic waves $\bar{\mathbf{\Psi}}$	0.609 ± 0.038	0.606 ± 0.048	0.628 ± 0.044	0.591 ± 0.043
	Pseudo harmonic waves $\tilde{\mathbf{\Psi}}$	0.558 ± 0.032	0.546 ± 0.038	0.580 ± 0.036	0.551 ± 0.036
	Fréchet harmonic waves $\check{\mathbf{\Psi}}$	0.601 ± 0.033	0.592 ± 0.038	0.621 ± 0.040	0.589 ± 0.036
	Our common harmonic waves $\mathbf{\Psi}$	$\mathbf{0.661 \pm 0.036}^*$	$\mathbf{0.658 \pm 0.044}^*$	$\mathbf{0.682 \pm 0.042}^*$	$\mathbf{0.651 \pm 0.036}^*$
Amyloid deposition	Empirical Amyloid deposition	0.759 ± 0.029	0.778 ± 0.034	0.762 ± 0.035	0.740 ± 0.032
	Arithmetic harmonic waves $\bar{\mathbf{\Psi}}$	0.664 ± 0.034	0.661 ± 0.042	0.686 ± 0.036	0.650 ± 0.037
	Pseudo harmonic waves $\tilde{\mathbf{\Psi}}$	0.717 ± 0.031	0.733 ± 0.039	0.725 ± 0.035	0.692 ± 0.037
	Fréchet harmonic waves $\check{\mathbf{\Psi}}$	0.697 ± 0.034	0.711 ± 0.047	0.707 ± 0.036	0.675 ± 0.038
	Our common harmonic waves $\mathbf{\Psi}$	$\mathbf{0.773 \pm 0.029}^*$	$\mathbf{0.794 \pm 0.040}^*$	$\mathbf{0.775 \pm 0.031}^*$	$\mathbf{0.751 \pm 0.035}^*$



(a) Oscillation pattern vs kinetic potential pattern on significant harmonic wave

(b) Oscillation pattern vs kinetic potential pattern on non-significant harmonic wave

Fig. 6. Statistics results of the association between the oscillation patterns in harmonic waves and the neurodegeneration process. Please see the above text for more information.

Next, we test the hypothesis that the positive-negative harmonic power difference (kinetic potentials of amyloid level due to the oscillations in harmonic waves) between a_h^+ and a_h^- is the factor leading to such significance. The rationale is that the oscillation patterns in the harmonic waves are correlated with the observations of pathological burdens if not only (1) the harmonic power shows significant differences between CN and AD, but also (2) the positive-negative power differences also manifest significant differences between CN and AD. As such, the observed harmonic waves can serve as biological indicators (factors) for the progression of AD. As a piece of evidence for the above hypothesis, we display the oscillation mapping of both a significant and a non-significant harmonic wave on the cortical surface in **Fig. 6(a)** left and **Fig. 6(b)** left, respectively. This visualization also shows the associated cortical mapping of elementwise vector multiplication between $\boldsymbol{\psi}_h$ and $\boldsymbol{f}$ at the right side of **Fig. 6(a)-(b)**. It is apparent that the two cortical mappings for the significant harmonic wave in **Fig. 6(a)** have a strong resemblance, which is also supported by the statistical significance between α^+ and a^- ($p < 10^{-4}$). On the contrary, such resemblance is not presented in the non-significant harmonic wave (**Fig. 6(b)**), where no significance has been detected between α^+ and a^- ($p = 0.11$).

As shown in the last column in **Table IV**, 62% (6.16/9.87) of the identified significant common harmonic waves in $\mathbf{\Psi}$ support such a hypothesis since average 6.16 common harmonic waves exhibit the statistical significance of $|\alpha_h^+ - \alpha_h^-|$ ($p < 10^{-3}$) in step (4). As a comparison, we find 48% (3.67/7.69), 24% (1.93/8.01), and 61% (5.30/8.76) of the harmonics in $\bar{\mathbf{\Psi}}$, $\tilde{\mathbf{\Psi}}$ and $\check{\mathbf{\Psi}}$ show CN vs AD significance in both harmonic power α_h and positive-negative power difference $|\alpha_h^+ - \alpha_h^-|$. The results in **Table IV** indicate that the oscillation patterns in our learned common harmonic waves have more statistical correlations with pathological neurodegeneration events. In the following experiments, we further demonstrate the potential of our common harmonics in practical clinical applications through classification task, compared with other approaches (B4). Finally, we apply our learned common harmonic waves $\mathbf{\Psi}$ to identify frequency-based harmonic alterations in the context of neurodegeneration biomarker measured by the cortical thickness (B5) and amyloid deposition (B6).

4) Investigating the Diagnostic Value of Harmonic Based Feature Representations: In this experiment, we evaluate the diagnostic value of our harmonic based feature presentations in identifying AD subjects. Specifically, we use the harmonic energy of cortical thickness (or amyloid) as the feature to train the CN vs AD classifier using a linear support vector machine (SVM). To compare the classification performance across harmonic estimation methods, we extract harmonic features based on arithmetic harmonic waves, pseudo harmonic waves, and Fréchet harmonic waves, and train the SVM separately. To show the advantage of harmonic features over the convention AD biomarker, we also train SVM using empirical cortical thickness and amyloid level. The classification performance is quantified using accuracy, specificity, sensitivity, and F-score, based on 5-fold cross-validation.

Table V shows the CN vs AD classification results (mean and standard deviation) of different harmonic-based features and empirical features. It is apparent that harmonic feature based on our common harmonic waves $\mathbf{\Psi}$ achieves the highest score in classification accuracy, sensitivity, specificity, and F-score over all other methods for both cortical thickness or amyloid deposition, where the star '$*$' indicates that our results are significantly better ($p < 0.01$) than the counterpart method ranked at the second place. It is worth noting that our new harmonic-based feature representations outperform empirical biomarker features, which indicate the great potential of applying our harmonic analysis approach in the early diagnosis of AD.

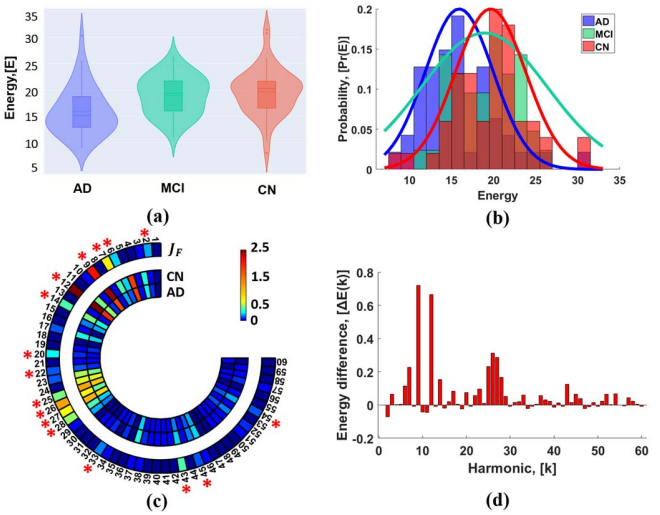


Fig. 7. Harmonic alterations among CN, MCI and AD identified using the learned common harmonics Ψ on cortical thickness. (a)-(b): Significant difference of total harmonic energy has been detected among CN, MCI and AD. (c): Harmonic waves exhibiting significant energy difference between CN and AD cohorts. (d): The plot of CN-to-AD energy difference of each harmonic wave.

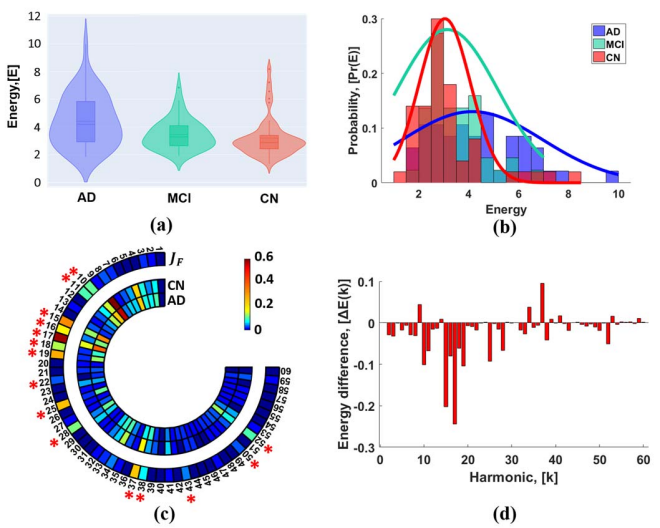


Fig. 8. Harmonic alterations among CN, MCI and AD identified using the learned common harmonics Ψ on amyloid deposition. (a)-(b): Significant difference of total harmonic energy has been detected among CN, MCI and AD. (c): Harmonic waves exhibiting significant energy difference between CN and AD cohorts. (d): The plot of CN-to-AD energy difference of each harmonic wave.

5) Identifying Frequency-Based Harmonic Alterations in Cortical Thickness: A plethora of neuroimaging studies found morphometry differences between CN and AD cohorts. Due to the correlation between our common harmonic waves Ψ and the neurodegeneration process as demonstrated in Section III.B3, as well as its effective classification ability as verified in Section III.B4, we explore the frequency-based alterations of cortical thickness values from the testing data that are relevant to AD progression by using the learned common harmonics from the training data. First, we measure the total harmonic energy of cortical thickness for each subject and plot the statistics (mean and standard deviation) for CN, MCI, and AD groups separately in **Fig. 7(a)**, where the AD group (15.9 ± 4.6) has significantly lower ($p < 10^{-3}$) total energy than both MCI group (18.9 ± 3.7) and CN group (19.6 ± 4.4). There is no significant total energy difference ($p = 0.42$) between MCI group and CN group. However, it is worth noting that empirical feature (cortical thickness) does not show a significant difference ($p = 0.44$) between MCI and CN cohorts either.

To verify the identified significant CN vs AD harmonic energy difference is relevant to AD indeed, we randomly shuffle the diagnostic labels 100 separate times in order to deliberately break down the coherence between network data and the associated diagnostic groups. After that, we fail to find a significant difference by repeating the same statistical test in our experiment (average $p = 0.50$), which indicates the identified harmonic energy differences are relevant to AD.

Furthermore, we plot the distribution of total energy in **Fig. 7(b)**. These results support the evidence that neurodegeneration in AD subjects is associated with reduced neuroanatomical structural integrity. Second, we examine the cross-sectional energy difference for each harmonic, where the mean harmonic-specific energy for CN and AD are shown in the outer and inner rings in **Fig. 7(c)**. In addition, the Fisher score J_F (the ratio between inter-class mean and intra-class variance) of the harmonic-specific energy between CN and AD subjects is shown in the outermost ring in **Fig. 7(c)**, where the harmonic waves exhibiting significant energy differences are tagged with a red star '$*$'. The CN-to-AD difference magnitude at each harmonic wave is displayed in **Fig. 7(d)**. These significant harmonic waves may be critically important in determining the propagation of neuropathological burdens across the brain networks.

6) Identifying Frequency-Based Harmonic Alterations in Amyloid Deposition: Similarly, we calculate the total harmonic energy of the amyloid deposition for each subject and plot the results in **Fig. 8(a)** and **(b)**, where the AD group (4.41 ± 1.85) has significantly higher ($p < 10^{-3}$) total energy than both the MCI group (3.46 ± 1.08) and CN group (3.19 ± 1.40). In addition, we show the statistical significance in energy difference and the CN vs AD energy difference magnitude for each harmonic wave in **Fig. 8(c)** and **(d)**, where there are a total of 15 harmonic waves exhibiting significant difference ($p < 0.01$) between CN and AD, in terms of harmonic energy of amyloid deposition. These results suggest that the aggregation of amyloid peptides is associated with topological features of the brain networks that underlie the network harmonics.

7) Discussions: The deposition of amyloid plaques is one of the hallmarks of AD. Both human and animal data suggest a causal upstream role for amyloid-β in the pathogenesis of AD, which may be sufficient to cause downstream pathologic changes leading to cognitive decline [39]. Our finding of frequency-based harmonic alterations in amyloid deposition complements the current neuroscience and clinical literature, with the AD population having greater amyloid harmonic energy than the CN group (**Fig. 8(d)**). Similarly, reductions

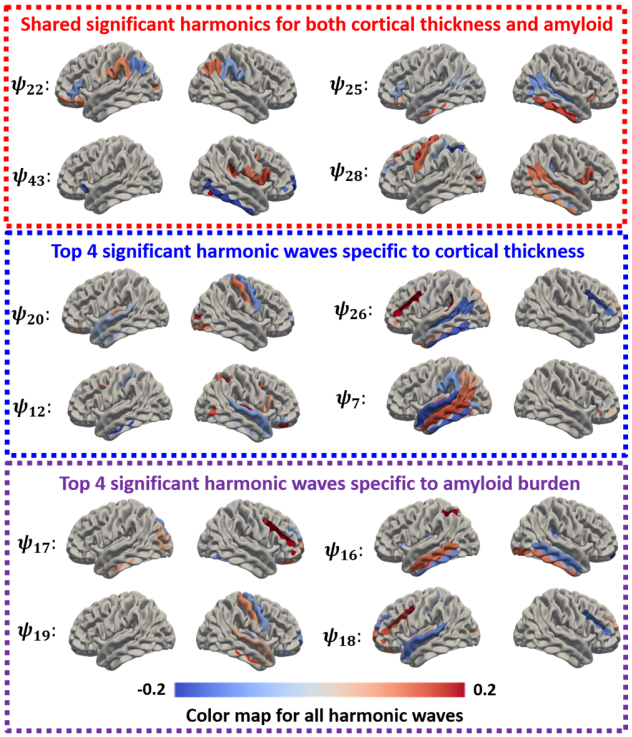


Fig. 9. The spatial pattern of significant harmonic waves in the context of cortical thickness (blue box), and amyloid deposition (purple box). Their shared harmonic waves are displayed in the red box.

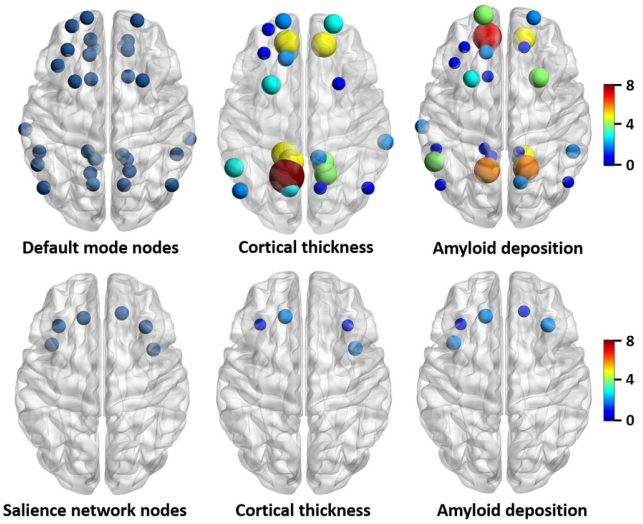


Fig. 10. The spatial alignment of crossing-zeros in the identified significant common harmonic waves with respect to default mode network and salience network. Left: nodes belong to DMN and salience network. Middle: in the context of cortical thickness. Right: in the context of amyloid deposition.

in cortical thickness are thought to reflect neuro-degeneration associated with AD progression. As shown in **Fig 7(d)**, CN subjects have more cortical thickness harmonic energy than AD subjects in most of the harmonic frequency bands, which indicates that degeneration (structural atrophy) is more profound in the AD than the CN cohort.

Furthermore, we found 16 harmonic waves for cortical thickness and 15 out of 60 common harmonic waves for amyloid that were significant differences between CN and AD. We display the oscillation pattern of the identified harmonic waves for neuro-degeneration (cortical thickness) and amyloid deposition in **Fig. 9**, where the shared harmonic waves by the cortical thickness and amyloid deposition are shown at the top. In addition, the top 4 significant harmonic waves with the smallest p-value specific to cortical thickness and amyloid burden are shown in middle and bottom in **Fig. 9**, respectively.

Different neurodegenerative diseases exhibit distinct network alteration patterns [3]. For example, AD is associated with atrophy and hypometabolism in the posterior hippocampal, cingulate, temporal, and parietal regions, which collectively resemble the default mode network (DMN) [40], [41]. In contrast to AD, behavior variant frontotemporal dementia (bvFTD) preferentially affects the salience network (SN) [3], [40]. Here, we examine the association between the oscillation pattern and these large-scale networks. *First*, we mark the location of the top ten crossing-zeros in each harmonic wave which has the largest difference magnitude. In general, 22-24% of the crossing-zeros are found falling in the DMN, compared to only 2% of them are associated with the SN. *Second*, we calculate the frequency of each node being touched by the crossing-zeros across all significant harmonic waves. We show the node frequency maps by cortical thickness and amyloid in the middle and right of **Fig. 10**, respectively. It is clear that much more crossing-zeros are associated with DMN (top) than SN (bottom), which is aligned with the current findings in AD.

IV. Conclusion

In this paper, we present a new network harmonic analysis approach that offers a new window into the investigation of frequency-based alterations in different clinical and research study populations. To achieve this, we propose a manifold optimization method to find the set of common harmonic waves from the native Eigen-systems of individual brain networks. The resulting shared reference space spanned by the common harmonic waves allows us to quantify the individual kinetic differences in terms of propagating neuro-pathological events across brain networks. We have evaluated the power of the common harmonic waves in discovering harmonic-specific alterations between CN and AD. More consistent and reasonable results were achieved by our manifold learning method, than by existing methods based on Euclidean operations on the manifold data.

In the future, we plan to apply our new network harmonic analysis approach to other neurological disorders that manifest network dysfunctions such as frontotemporal dementia and schizophrenia.

References

[1] M. R. Brier, J. B. Thomas, and B. M. Ances, "Network dysfunction in Alzheimer's disease: Refining the disconnection hypothesis," *Brain Connectivity*, vol. 4, no. 5, pp. 299–311, Jun. 2014.

[2] B. C. Dickerson and R. A. Sperling, "Large-scale functional brain network abnormalities in Alzheimer's disease: Insights from functional neuroimaging," *Behav. Neurol.*, vol. 21, nos. 1–2, pp. 63–75, 2009.

[3] J. J. Palop, J. Chin, and L. Mucke, "A network dysfunction perspective on neurodegenerative diseases," *Nature*, vol. 443, no. 7113, pp. 768–773, Oct. 2006.

[4] O. Sporns, "Structure and function of complex brain networks," *Dialogues Clin. Neurosci.*, vol. 15, no. 3, p. 247, 2013.

[5] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[6] H. Aerts, W. Fias, K. Caeyenberghs, and D. Marinazzo, "Brain networks under attack: Robustness properties and the impact of lesions," *Brain*, vol. 139, no. 12, pp. 3063–3083, Dec. 2016.

[7] H. Braak and K. Del Tredici, "Spreading of tau pathology in sporadic Alzheimer's disease along cortico-cortical top-down connections," *Cerebral Cortex*, vol. 28, no. 9, pp. 3372–3384, Sep. 2018.

[8] Y. Chen *et al.*, "Disrupted functional and structural networks in cognitively normal elderly subjects with the APOE ε4 Allele," *Neuropsychopharmacology*, vol. 40, p. 1181, Nov. 2014.

[9] O. Sporns, *Networks of the Brain*. Cambridge, MA, USA: MIT Press, 2011.

[10] O. Sporns, C. J. Honey, and R. Kötter, "Identification and classification of hubs in brain networks," *PLoS ONE*, vol. 2, no. 10, p. e1049, Oct. 2007.

[11] K. Wu, Y. Taki, K. Sato, H. Qi, R. Kawashima, and H. Fukuda, "A longitudinal study of structural brain network changes with normal aging," *Frontiers Hum. Neurosci.*, vol. 7, p. 113, Apr. 2013.

[12] S. J. Hwang *et al.*, "Coupled harmonic bases for longitudinal characterization of brain networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2517–2525.

[13] A. Raj, A. Kuceyeski, and M. Weiner, "A network diffusion model of disease progression in dementia," *Neuron*, vol. 73, no. 6, pp. 1204–1215, Mar. 2012.

[14] A. Raj, E. LoCastro, A. Kuceyeski, D. Tosun, N. Relkin, and M. Weiner, "Network diffusion model of progression predicts longitudinal patterns of atrophy and metabolism in Alzheimer's disease," *Cell Rep.*, vol. 10, no. 3, pp. 359–369, Jan. 2015.

[15] Y. Wang, D. Yang, Q. Li, D. Kaufer, M. Styner, and G. Wu, "Characterizing the propagation pattern of neurodegeneration in Alzheimer's disease by longitudinal network analysis," presented at the IEEE 17th Int. Symp. Biomed. Imag. (ISBI), Iowa City, IA, USA, Apr. 2020.

[16] S. Atasoy, I. Donnelly, and J. Pearson, "Human brain networks function in connectome-specific harmonic waves," *Nature Commun.*, vol. 7, no. 1, p. 10340, Apr. 2016.

[17] S. Atasoy, L. Roseman, M. Kaelen, M. L. Kringelbach, G. Deco, and R. L. Carhart-Harris, "Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under LSD," *Sci. Rep.*, vol. 7, no. 1, Dec. 2017 Art. no. 17661.

[18] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Oct. 2016.

[19] J. Ashburner and K. J. Friston, "Voxel-based morphometry-the methods," *NeuroImage*, vol. 11, pp. 805–821, Jun. 2000.

[20] C. Davatzikos, A. Genc, D. Xu, and S. M. Resnick, "Voxel-based morphometry using the RAVENS maps: Methods and validation using simulated longitudinal atrophy," *NeuroImage*, vol. 14, no. 6, pp. 1361–1369, Dec. 2001.

[21] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.

[22] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[23] B. Wang, J. Zhu, E. Pierson, D. Ramazzotti, and S. Batzoglou, "Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning," *Nature Methods*, vol. 14, p. 414, Mar. 2017.

[24] A. Yamin *et al.*, "Comparison of brain connectomes using geodesic distance on manifold: A twins study," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1797–1800.

[25] L. Dodero, F. Sambataro, V. Murino, and D. Sona, "Kernel-based analysis of functional brain connectivity on Grassmann manifold," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 604–611.

[26] M. Dai, Z. Zhang, and A. Srivastava, "Discovering common change-point patterns in functional connectivity across subjects," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101532.

[27] M. Dai, Z. Zhang, and A. Srivastava, "Analyzing dynamical brain functional connectivity as trajectories on space of covariance matrices," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 611–620, Mar. 2020.

[28] Y. Chikuse, *Statistics on Special Manifolds* (Lecture Notes in Statistics), vol. 174. New York, NY, USA: Springer, 2003.

[29] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, Jan. 1998.

[30] R. Subbarao and P. Meer, "Nonlinear mean shift for clustering over analytic manifolds," presented at the IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2006.

[31] H. E. Cetingul and R. Vidal, "Intrinsic mean shift for clustering on stiefel and Grassmann manifolds," presented at the IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009.

[32] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Geometric means in a novel vector space structure on symmetric positive-definite matrices," *SIAM J. Matrix Anal. Appl.*, vol. 29, no. 1, pp. 328–347, Jan. 2007.

[33] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[34] F. Nie, R. Zhang, and X. Li, "A generalized power iteration method for solving quadratic problem on the stiefel manifold," *Sci. China Inf. Sci.*, vol. 60, no. 11, Nov. 2017, Art. no. 112101.

[35] K. Aftab, R. Hartley, and J. Trumpf, "Generalized Weiszfeld algorithms for Lq optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 4, pp. 728–745, Apr. 2015.

[36] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifold*. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[37] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *NeuroImage*, vol. 53, no. 1, pp. 1–15, Oct. 2010.

[38] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc., B, Methodol.*, vol. 57, no. 1, pp. 289–300, Jan. 1995.

[39] C. R. Jack *et al.*, "NIA-AA research framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's Dementia*, vol. 14, no. 4, pp. 535–562, Apr. 2018.

[40] J. Zhou and W. W. Seeley, "Network dysfunction in Alzheimer's disease and frontotemporal dementia: Implications for psychiatry," *Biol. Psychiatry*, vol. 75, no. 7, pp. 565–573, Apr. 2014.

[41] A. Horn, D. Ostwald, M. Reisert, and F. Blankenburg, "The structural–functional connectome and the default mode network of the human brain," *NeuroImage*, vol. 102, pp. 142–151, Nov. 2014.