ALZHEIMER'S DISEASE CLASSIFICATION VIA SCORE-DERIVED SPECTRAL COUPLING FROM LONGITUDINAL CORTICAL DATA

Anonymous authors
Paper under double-blind review

ABSTRACT

Longitudinal neuroimaging provides direct observations of within-subject disease-related change, but such data are typically sparse and irregularly sampled. We propose a spherical score-based framework for learning class-conditioned coupling representations from pairwise longitudinal cortical observations drawn from trajectories with two or more timepoints. Given a baseline–follow-up pair, we represent the interval-normalized cortical change in the spherical harmonic (SH) domain and train a class-conditioned score model using a spherical Ornstein–Uhlenbeck diffusion process. Rather than using the learned score for generation, we exploit its local differential structure: the negative score Jacobian defines a state-dependent local dependency operator over SH modes, reducing to the precision matrix in the Gaussian special case. We convert this score-derived operator into a normalized guidance operator and use it to modulate intermediate spectral features of a spherical cortical network under alternative diagnostic hypotheses. The proposed framework is evaluated on longitudinal cortical thickness, FDG-PET, and AV45-PET data from the ADNI dataset, achieving consistently strong classification performance across modalities. Additional cross-cohort and phase-transfer evaluations further support the generalizability of the learned coupling representation.

1 INTRODUCTION

Motivation and Background. Alzheimer's disease (AD) is a progressive biological process that unfolds over many years, such that a single clinical visit provides only a snapshot of an underlying disease trajectory. Longitudinal neuroimaging instead provides direct observations of within-subject change, while prior biomarker and network studies suggest that AD-related abnormalities evolve in a temporally structured manner and involve coordinated vulnerability across spatially distributed brain systems rather than isolated local changes (Jack et al., 2013; Raj et al., 2012). Neuropathological staging studies further indicate that AD-related neurofibrillary changes progress through ordered anatomical stages (Braak & Braak, 1991; 1995; Ohm et al., 1995), motivating representations of cortical change that are conditioned on diagnostic state. Exploiting such information, however, remains challenging because longitudinal imaging trajectories are often sparse, irregularly sampled, and incomplete even in large cohorts such as ADNI (Mueller et al., 2005) and OASIS (Marcus et al., 2010). A baseline-follow-up pair therefore provides a practical minimal unit of longitudinal supervision: subjects with only two valid observations can contribute directly, while longer trajectories can be decomposed into multiple pairwise transitions. Prior approaches instead infer progression from population-level variation (Young et al., 2014; 2018), fit trajectories to predefined low-dimensional biomarkers (Jedynak et al., 2012; Donohue et al., 2014), aggregate per-visit features across multiple timepoints (Ouyang et al., 2021), or synthesize future disease states (Jung et al., 2023). These either benefit from richer temporal sequences or operate on summarized biomarkers rather than on within-subject anatomical change at surface resolution. Yet the magnitude of a pairwise change alone does not reveal how disease-related alterations are organized across the cortex. Brain network studies suggest that cortical regions linked within distributed systems may exhibit coordinated vulnerability during neurodegeneration, motivating us to model dependencies among longitudinal cortical changes rather than treating them as independent regional alterations. Because cortical measure-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

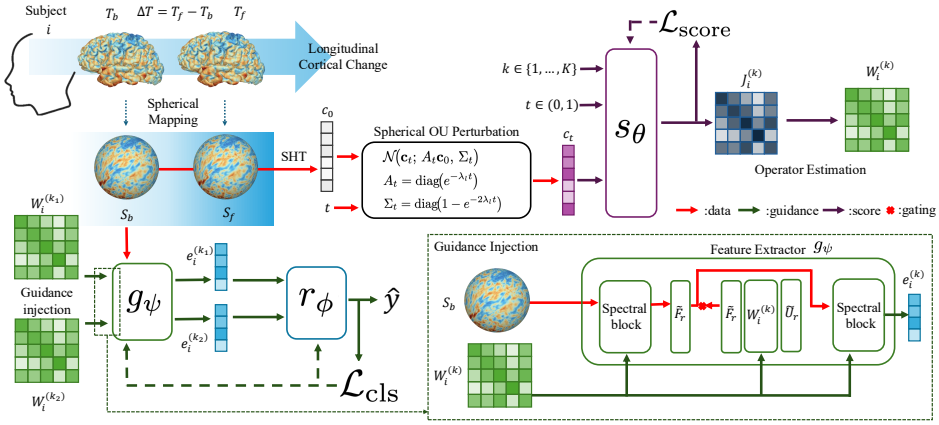


Figure 1: A longitudinal cortical pair is converted into a baseline-to-follow-up deformation, from which a diagnosis-conditioned score model derives class-specific coupling guidance. The guidance modulates a spherical feature extractor, and a shared readout compares the hypothesis-conditioned embeddings for diagnosis.

ments are naturally defined on a continuous surface, we model these dependencies directly in the cortical domain while preserving spatial detail without requiring a predefined regional parcellation. The central question of this work is how subject-specific coupling structure in within-subject cortical change can be learned from sparse longitudinal observations and subsequently used for representation learning.

Proposed method. Motivated by evidence that brain connectome organization and large-scale network structure are linked to patterns of neurodegenerative vulnerability and disease progression (Seeley et al., 2009; Raj et al., 2012), we propose a spherical score-based framework for learning coordinated coupling structure from pairwise longitudinal cortical observations. Given a baseline–follow-up pair, we represent interval-normalized cortical change in the spherical harmonic domain and learn its class-conditional distribution through score matching. We then exploit the local differential structure of the learned score to derive a subject- and hypothesis-conditioned dependency representation among spectral deformation modes, capturing how longitudinal cortical changes are statistically coordinated under alternative diagnostic hypotheses. This score-derived coupling is transformed into structural guidance for a spherical cortical feature extractor, encouraging intermediate representations to incorporate relationships among longitudinal changes in addition to their individual magnitudes. The resulting framework converts sparse pairwise longitudinal observations into data-driven coupling guidance for diagnostic classification without requiring a predefined anatomical connectome. We evaluate the proposed method across longitudinal cortical thickness, FDG-PET, and AV45-PET, together with cross-cohort and phase-transfer distribution shifts. Figure 1 illustrates the overall framework. Our contributions are summarized as follows:

- We introduce a spherical score-based framework for learning class-conditioned coupling representations from sparse pairwise longitudinal cortical observations, without requiring densely sampled trajectories or predefined anatomical connectomes.
- We convert the local differential structure of a class-conditioned score model into an explicit subject- and hypothesis-conditioned dependency operator, and use the resulting guidance to modulate mode-to-mode spectral feature interactions for discriminative representation learning.
- We demonstrate consistent generalization across longitudinal cortical modalities and distribution shifts, and validate the role of the proposed coupling guidance through targeted component ablations and analysis of its emergent spatial structure on the cortical surface.

2 METHODS

2.1 PROBLEM STATEMENT

Let $\mathcal{X}$ denote the space of longitudinal observations $\mathcal{D} = (S_b, S_f, \Delta T)$, where the spherical cortical fields $S_b, S_f : \mathbb{S}^2 \to \mathbb{R}$ are acquired with temporal interval $\Delta T > 0$, and S_b and S_f denote the baseline and follow-up observations, respectively. Our objective is to learn a classifier $f : \mathcal{X} \to \mathcal{Z}$, where $\mathcal{Z} = \{1, \ldots, K\}$ denotes the diagnostic class space.

2.2 SCORE-DERIVED COUPLING REPRESENTATION

Definition. Motivated by spatially coordinated patterns of neurodegenerative change, we seek to characterize not only the magnitude of longitudinal cortical change but also the dependencies among changes across cortical locations. For a longitudinal pair $\mathcal{D} = (S_b, S_f, \Delta T)$ with acquisition interval ΔT, we consider the deformation $\delta(x) = \{S_f(x) - S_b(x)\}/\Delta T$ at $x \in \mathbb{S}^2$ expanded in the spherical harmonics (SH) coefficient space. In the SH encoding, we exclude the degree-zero coefficient to remove the spatially constant component and retain the band-limited coefficients $\mathbf{c} = (c_a) \in \mathbb{R}^M$, where $M = (L+1)^2 - 1$, $a \equiv (l, m)$, $l \geq 1$ and L is a spectral band limit. Across subjects, we regard these SH coefficients $\mathbf{c}$ as realizations of a random vector $\mathbf{C} = (C_1, \ldots, C_M) \in \mathbb{R}^M$. For diagnostic class $y \in \mathcal{Z}$, let $p(\mathbf{c}|y)$ denote the class-conditional density of $\mathbf{c}$. The corresponding class-conditional score is defined as $s_y(\mathbf{c}) = \nabla_{\mathbf{c}} \log p(\mathbf{c}|y)$. This score characterizes the local geometry of the class-conditional density in the SH coefficient space.

For smooth and strictly positive densities, mixed derivatives of the log-density provide a standard characterization of conditional dependence (Baptista et al., 2024; Spantini et al., 2018). This motivates examining the Jacobian of the score or equivalently the Hessian of the log-density, to characterize dependencies among SH modes. More formally, by letting $H_y(\mathbf{c}) = -\nabla_{\mathbf{c}} s_y(\mathbf{c}) = -\nabla^2_{\mathbf{c}} \log p(\mathbf{c} \mid y)$, we refer to $H_y(\mathbf{c})$ as a local dependency operator that characterizes the local second-order interaction structure of the class-conditional log-density around the state $\mathbf{c}$. We instantiate this characterization in the SH spectral domain of longitudinal cortical change.

Proposition 1. *Let $\mathbf{C} = (C_1, \ldots, C_M) \in \mathbb{R}^M$, for $M \geq 2$, denote a collection of SH-mode coefficients with a strictly positive twice continuously differentiable density $p(\mathbf{c} \mid y)$ and diagnostic class y. Define $s_y(\mathbf{c}) = \nabla_{\mathbf{c}} \log p(\mathbf{c} \mid y)$ and $H_y(\mathbf{c}) = -\nabla_{\mathbf{c}} s_y(\mathbf{c})$. Then, for two distinct SH modes $a \neq b$, $[H_y(\mathbf{c})]_{ab} = -\partial^2 \log p(\mathbf{c} \mid y)/(\partial c_a \, \partial c_b)$. Moreover, $C_a \perp C_b \mid \mathbf{C}_{-(a,b)}, y$ if and only if $[H_y(\mathbf{c})]_{ab} = 0$ for all $\mathbf{c} \in \mathbb{R}^M$ where $\mathbf{C}_{-(a,b)}$ denotes all coefficients except C_a and C_b.*

Proposition 1 provides a probabilistic interpretation of the score Jacobian as a local coupling representation. In particular, $[H_y(c)]_{ab}$ measures how the probability gradient along mode a changes with b, while conditional independence corresponds to the absence of such interaction throughout the class-conditional distribution. Thus, the operator H_y represents state-dependent coupling patterns. For the Gaussian special case, $\mathbf{C} \mid y \sim N(\mu_y, \Sigma_y)$, the operator becomes state-independent and reduces exactly to the precision matrix, i.e., $H_y(\mathbf{c}) = \Sigma_y^{-1}$. Let $Y_a(x)$ denote the real orthonormal SH basis function corresponding to mode $a = (\ell, m)$, and let $Y(x) \in \mathbb{R}^M$ collect the retained basis functions in the same order as $\mathbf{c}$. Then the defined spectral dependency operator $H_y(\mathbf{c})$ induces the band-limited spatial kernel $K_y(x, x'; \mathbf{c}) = Y(x)^\top H_y(\mathbf{c}) Y(x')$ such that for arbitrary band-limited perturbations u, v, $D^2[-\log p](u, v) = \int \int_{\mathbb{S}^2 \times \mathbb{S}^2} u(x) K_y(x, x'; \mathbf{c}) v(x') dx dx'$. Thus, interactions among SH modes encoded by H_y induce a spatially resolved coupling kernel on the cortical surface data.

Proposition 2. *Let T be a linear SO(3)-equivariant operator acting on scalar spherical functions band-limited to degree L. In the SH spectral space, T must take the block scalar form $T = \bigoplus_{l=1}^{L} \gamma_l I_{2l+1}$, where $\gamma_l \in \mathbb{R}$.*

Proposition 2. implies that a linear rotation-equivariant isotropic spectral operator may apply degree-dependent weighting but cannot couple distinct SH modes.

Isotropic reference model. Proposition 2 motivates a restricted isotropic reference family for spectral dependency operators,

$$\mathcal{H}_{\text{iso}} = \left\{ \bigoplus_{l=1}^{L} \gamma_l I_{2l+1} \ : \ \gamma_l \in \mathbb{R} \right\}.$$

For any operator in this family, the spherical-harmonic addition theorem gives the corresponding spatial kernel

$$K_{\text{iso}}(x, x') = \sum_{l=1}^{L} \gamma_l \frac{2l+1}{4\pi} P_l(x \cdot x'),$$

which depends only on the geodesic separation between x and x' where P_l is the Legendre polynomial at degree of l. Thus, the block-scalar family represents degree-wise isotropic spectral coupling without explicit mode-to-mode interactions. In contrast, the spectral dependency operator $H_y(\mathbf{c})$ is not constrained to this family; cross-degree and non-scalar within-degree components allow its induced spatial kernel to represent more general position-dependent coupling patterns.

2.3 Score Estimation via Spherical Diffusion

We now apply the dependency formulation in Sec. 2.2 to a diffused class-conditional density. For diffusion time t, let $p_t(\mathbf{c}_t \mid y)$ denote the class-conditional density of the perturbed SH coefficients, and define $s_{y,t}(\mathbf{c}_t) = \nabla_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid y)$, $H_{y,t}(\mathbf{c}_t) = -\nabla_{\mathbf{c}_t} s_{y,t}(\mathbf{c}_t)$. Thus, $H_{y,t}$ is the finite-time counterpart of the local dependency operator introduced in Sec. 2.2. Direct evaluation of the spectral dependency operator $H_{y,t}$ requires access to the unknown class-conditional density $p_t(\mathbf{c}_t \mid y)$. We therefore estimate its score using a class-conditioned spherical score model, without relying on anatomical connectomes as supervision. The class label y specifies the diagnostic hypothesis, while evaluating the local differential structure of the learned score at an individual coefficient vector $\mathbf{c}_t$ yields a state-dependent coupling representation. The score network constitutes the first stage of our framework: it is optimized solely with a denoising score-matching objective and is subsequently frozen before downstream network training. We next define a spherical diffusion process that provides the perturbed distributions required for score estimation.

Spherical Ornstein–Uhlenbeck forward process. We use a variance-preserving spectral diffusion whose drift is induced by the spherical Laplace-Beltrami operator $\Delta_{\mathbb{S}^2}$ (Lee et al., 2025). Because the spherical harmonic basis functions Y_a are its eigenfunctions, $\Delta_{\mathbb{S}^2} Y_a = -\lambda_l Y_a, \lambda_l = l(l+1)$, the forward process diagonalizes in the SH basis and acts independently on each mode as a mean-reverting Ornstein–Uhlenbeck (OU) process,

$$\mathrm{d}c_a = -\lambda_l \, c_a \, \mathrm{d}t + \sqrt{2\lambda_l} \, \mathrm{d}W_a, \tag{1}$$

where W_a are independent standard Wiener processes. Integrating equation 1 gives the closed-form perturbation kernel for any forward process time $t, c_{a,t} = e^{-\lambda_l t} \, c_{a,0} + \sigma_{l,t} \, \epsilon_a, \sigma_{l,t} = \sqrt{1 - e^{-2\lambda_l t}}, \epsilon_a \sim \mathcal{N}(0, 1)$. Collecting all modes, the kernel is the mode-diagonal Gaussian

$$p_t(\mathbf{c}_t \mid \mathbf{c}_0) = \mathcal{N}\big(\mathbf{c}_t;\, A_t \mathbf{c}_0,\, \Sigma_t\big), \qquad A_t = \operatorname{diag}\big(e^{-\lambda_l t}\big), \qquad \Sigma_t = \operatorname{diag}\big(1 - e^{-2\lambda_l t}\big). \tag{2}$$

The decay rate $\lambda_l = l(l+1)$ grows with degree, so high-degree (fine-scale) modes lose their dependence on the clean deformation $\mathbf{c}_0$ faster than low-degree modes, and the stationary law is the isotropic $\mathcal{N}(0, I)$.

Denoising Score Matching. The network $s_\theta(\mathbf{c}_t, y, t)$ is trained to estimate the score of the perturbation kernel by denoising score matching,

$$\mathcal{L}_{score} = \mathbb{E}_{\mathbf{c}_0,\, y,\, t,\, \boldsymbol{\epsilon}} \Big[\big\| s_\theta(\mathbf{c}_t, y, t) - \nabla_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid \mathbf{c}_0) \big\|_2^2 \Big], \tag{3}$$

whose closed-form target is

$$\nabla_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid \mathbf{c}_0) = -\,\Sigma_t^{-1}\big(\mathbf{c}_t - A_t \mathbf{c}_0\big) = -\,\Sigma_t^{-1/2} \boldsymbol{\epsilon},$$

where $\Sigma_t^{1/2} = \operatorname{diag}(\sigma_{l,t})$, so the target reads $-\epsilon_a / \sigma_{l,t}$ per mode. The regression target is conditioned on $\mathbf{c}_0$, not on y: the forward kernel equation 2 is class-independent, so $p_t(\mathbf{c}_t \mid \mathbf{c}_0, y) = p_t(\mathbf{c}_t \mid \mathbf{c}_0)$. The label y enters only as a network input; taking the expectation over class-y samples drives s_θ to the class-conditional marginal score $s_\theta(\mathbf{c}_t, y, t) \approx \nabla_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid y)$.

Score-derived dependency at finite diffusion time. At the score-matching optimum, the class-conditional score satisfies $s_{y,t}(\mathbf{c}_t) = \nabla_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid y)$, and its negative Jacobian therefore gives the local dependency operator $H_{y,t}(\mathbf{c}_t) := -\nabla_{\mathbf{c}_t} s_{y,t}(\mathbf{c}_t) = -\nabla^2_{\mathbf{c}_t} \log p_t(\mathbf{c}_t \mid y)$. Accordingly, Proposition 1 applies directly to the diffused class-conditional density $p_t(\cdot \mid y)$. For Gaussian $p_t(\mathbf{c}_t \mid y)$, $H_{y,t}$ is state-independent and coincides with its precision matrix. In practice, we evaluate the learned score Jacobian at a small validation-selected guidance scale t_g and use the resulting estimate of H_{y,t_g} as the spectral coupling representation for downstream classification. A finite-time derivation under the spherical OU perturbation is provided in Appendix A.2.

2.4 Dependency-Guided Feature Extraction

We incorporate the score-derived spectral coupling representation into the feature extractor g_ψ through a normalized guidance operator, allowing intermediate features to interact according to class-conditioned dependency structures learned from longitudinal cortical changes. As the backbone, we adopt SPHARM-Net (Ha & Lyu, 2022), a spherical convolutional network that alternates between spectral and spatial representations to extract cortical surface features. This architecture provides a natural interface for the proposed guidance mechanism, since the dependency representation is estimated in the spherical harmonic domain and can therefore directly modulate intermediate spectral features. Given the frozen score model s_{θ^*}, we first estimate a subject- and class-hypothesis-conditioned dependency operator and subsequently transform it into the normalized guidance operator used for downstream feature extraction.

Explicit Score-Derived Dependency. Given the frozen score network, we estimate the spectral dependency operator from the learned score and use it to construct the guidance. Let $M_g = (L_g + 1)^2 - 1$ denote the number of nonzero SH modes retained up to the guidance bandwidth $L_g \leq L$. For subject i and class hypothesis $k \in \{1, ..., K\}$, we evaluate the learned score Jacobian at the fixed guidance scale $t_g \in (0, 1)$, restricted to the retained modes, $J_i^{(k)} = \nabla_{\mathbf{c}_{t_g}} s_{\theta^*}(\mathbf{c}_{t_g}, k, t_g) \in \mathbb{R}^{M_g \times M_g}$. Since the Jacobian of the exact score is a Hessian and is therefore symmetric, small asymmetries in the learned Jacobian are treated as numerical approximation errors. We thus symmetrize the learned score Jacobian and reverse its sign to obtain the estimated spectral dependency operator, $\widehat{H}_i^{(k)} = -\frac{1}{2}\left(J_i^{(k)} + J_i^{(k)\top}\right)$. The restriction to the M_g lowest modes keeps the explicit Jacobian computation tractable and focuses the guidance on coarse spectral scales. We then convert $\widehat{H}_i^{(k)}$ into a normalized guidance operator. Let

$$A_i^{(k)} = |\widehat{H}_i^{(k)}|, \qquad D_i^{(k)} = \text{diag}\left(\sum_j A_{i,aj}^{(k)}\right), W_i^{(k)} = D_i^{(k)-1/2} A_i^{(k)} D_i^{(k)-1/2}.$$

The element-wise absolute value retains the magnitudes of the estimated mode-to-mode dependencies while discarding their signs, and the symmetric normalization stabilizes their use as feature guidance. Accordingly, $W_i^{(k)}$ is used as a normalized guidance operator rather than being interpreted directly as a dependency operator. At each spectral layer r, the normalized guidance operator modulates the intermediate feature $F_i^{(r)} \in \mathbb{R}^{M_g \times d_r}$ through the residual update

$$F_i^{(r)} \leftarrow F_i^{(r)} + \alpha_r \left(W_i^{(k)} \widetilde{F}_i^{(r)}\right) U_r,$$

where $\widetilde{F}_i^{(r)}$ denotes a normalized copy of the intermediate feature, $U_r \in \mathbb{R}^{d_r \times d_r}$ is a learnable channel projection, and α_r is a learnable scalar gate. In this way, the score-derived coupling representation does not replace the original spectral features; instead, it provides structural guidance that biases feature interactions toward the subject- and hypothesis-conditioned dependency structure. Since the score network is frozen after the first training phase, the hypothesis-conditioned guidance operators $\{W_i^{(k)}\}_{k=1}^K$ are computed once for each subject and cached for downstream network.

2.5 Training Objective

Two-stage training. We train the proposed framework in two stages. In the first stage, the score network s_θ is optimized by denoising score matching on the longitudinal SH coefficients using equation 3. After convergence, the score network is frozen. For each subject i and class hypothesis

$k \in \{1, \ldots, K\}$, we obtain the spectral dependency operator $\hat{H}_i^{(k)}$ and its corresponding normalized guidance operator $W_i^{(k)}$ as described in Sec. 2.4. These guidance operators are fixed during the second training stage and are not updated by the classification objective. We then optimize only the feature extractor g_ψ and the shared readout r_ϕ. This two-stage procedure decouples score-based dependency estimation from downstream classification optimization.

Comparative classification by hypothesis scoring. Since the diagnostic class is unknown at inference time, each subject is evaluated under all K class-conditioned coupling hypotheses. For class hypothesis k, the shared feature extractor receives the corresponding guidance operator $W_i^{(k)}$ and produces a hypothesis-conditioned embedding $e_i^{(k)} = g_\psi\left(S_{b,i}; W_i^{(k)}\right)$. A single shared readout $r_\phi : \mathbb{R}^d \to \mathbb{R}$ assigns each embedding a scalar compatibility logit, $g_i^{(k)} = r_\phi\left(e_i^{(k)}\right)$. The logits from all class hypotheses are compared through $p_i^{(k)} = \frac{\exp\left(g_i^{(k)}\right)}{\sum_{j=1}^{K} \exp\left(g_i^{(j)}\right)}$, $\hat{y}_i = \arg\max_{k \in \{1,\ldots,K\}} p_i^{(k)}$. Thus, each branch evaluates the subject under a different class-conditioned coupling structure, while the shared readout learns their relative compatibility. During training, all K hypothesis-conditioned branches are performed jointly, so that the target hypothesis is compared against the competing alternatives. To account for class imbalance, we optimize a class-weighted cross-entropy objective,

$$\mathcal{L}_{\text{cls}} = -\mathbb{E}_i \left[\sum_{k=1}^{K} w_k y_{i,k} \log p_i^{(k)} \right],$$

where $y_{i,k}$ denotes the one-hot ground-truth label for subject i, and w_k is the weight assigned to class k, chosen inversely proportional to its frequency in the training set.

3 Experiments

3.1 Experimental Configuration

Dataset and Preprocessing. We evaluate our proposed framework using the ADNI datasets and OASIS-2 datasets. For each subject, we select baseline–follow-up sMRI or PET scan pairs to minimize group differences in mean acquisition interval ΔT. We reconstruct cortical surfaces using FreeSurfer *recon-all* pipeline and map them to an icosahedral subdivision unit sphere at level 6 with a common grid. We construct three distinct datasets based on their modalities: cortical thickness (CTh), Standardized uptake value ratio (SUVR) of Fluorodeoxyglucose (FDG) and Florbetapir F 18 (AV45) from PET. Further details of the scan pair definition criteria and demographics are provided in Appendix B.

Evaluation. All methods were evaluated under the same subject-level five-fold cross-validation protocol, using disjoint test folds and a separate validation fold for checkpoint selection in each round. For each fold, both training stages are independently initialized. The score network is trained exclusively on the training partition and frozen before downstream classifier training. All selections were based on validation performance. We report accuracy (ACC), area under the ROC curve (AUC), and Matthews Correlation Coefficient (MCC) as performance metrics. All experiments were conducted on a single NVIDIA RTX 3090 GPU with 24GB memory.

Baselines. We compare the proposed method against 20 baseline methods categorized into four groups. (1) **Generic GNNs**: GCN (Kipf & Welling, 2016), GAT (Veličković et al., 2018), GIN (Xu et al., 2019), GraphSAGE (Hamilton et al., 2017), GraphUNet (Gao & Ji, 2019), Exphormer (Shirzad et al., 2023), (2) **Geometric surface learning**: MoNet (Monti et al., 2017), SplineCNN (Fey et al., 2018), FeaStNet (Verma et al., 2018), ChebNet (Defferrard et al., 2016), DiffusionNet (Sharp et al., 2022), (3) **Spherical and mesh**: UGSCNN (Jiang et al., 2019), SUNet (Zhao et al., 2019), S2CNN (Cohen et al., 2018), DeepSphere (Defferrard et al., 2020), MeshCNN (Hanocka et al., 2019), (4) **Cortical surface**: SurfGNN (Li et al., 2025), SiT (Dahan et al., 2022), MS-SiT (Dahan et al., 2024), X-SiT (Bongratz et al., 2025). Since baselines are not tailored to paired longitudinal data, we use $[S_b, S_f]$ as an input.

Table 1: Classification performance on ADNI-CTh, ADNI-FDG, and ADNI-AV45 using 5-fold cross-validation. Mean ± std; ACC and AUC are reported in %. Best in **bold**, second-best underlined.

	ADNI-CTh			ADNI-FDG			ADNI-AV45		
Model	ACC(↑)	AUC(↑)	MCC(↑)	ACC(↑)	AUC(↑)	MCC(↑)	ACC(↑)	AUC(↑)	MCC(↑)
GCN	62.36±5.04	62.13±4.45	0.220±0.076	47.27±9.55	47.90±5.27	0.026±0.130	55.49±8.43	67.79±7.39	0.072±0.113
GAT	70.26±4.22	76.84±3.06	0.394±0.078	66.97±4.43	70.20±4.80	0.301±0.076	84.33±6.44	86.05±6.89	0.448±0.129
GIN	66.84±4.53	69.89±4.60	0.331±0.075	60.91±7.32	63.57±6.00	0.182±0.116	76.05±12.74	80.86±2.09	0.359±0.084
GraphSAGE	67.37±5.04	72.44±3.80	0.344±0.064	62.73±7.94	61.88±4.93	0.207±0.063	78.79±16.46	86.50±7.97	0.432±0.185
GraphUNet	57.72±7.58	59.80±8.57	0.141±0.124	54.24±9.93	50.30±6.55	0.065±0.099	73.58±17.50	83.96±9.77	0.364±0.150
Exphormer	64.44±2.96	69.15±4.59	0.265±0.073	52.42±12.91	61.11±11.46	0.134±0.129	78.20±12.24	80.24±13.58	0.389±0.130
MoNet	67.53±2.06	69.90±3.19	0.302±0.057	67.27±4.35	66.36±5.15	0.251±0.062	73.58±12.29	86.56±6.07	0.353±0.111
SplineCNN	70.97±4.07	79.75±2.85	0.418±0.072	81.82±2.14	89.68±2.30	0.609±0.029	80.04±10.23	85.48±9.65	0.380±0.151
FeaStNet	62.88±10.74	69.07±8.24	0.275±0.164	61.52±7.70	61.95±5.67	0.132±0.073	83.43±5.02	82.60±8.84	0.366±0.097
ChebNet	75.43±4.39	82.95±2.62	0.498±0.085	82.12±2.23	89.55±4.29	0.626±0.060	83.76±7.64	78.25±8.20	0.392±0.174
DiffusionNet	55.48±4.14	57.35±4.73	0.086±0.071	66.06±6.68	67.22±6.31	0.225±0.130	57.27±22.10	69.66±12.41	0.130±0.124
UGSCNN	63.93±4.57	69.49±3.75	0.266±0.085	75.45±2.78	87.55±2.14	0.485±0.051	68.69±15.53	73.09±8.59	0.156±0.145
SUNet	71.13±3.12	76.43±2.07	0.393±0.047	84.24±4.13	90.70±2.98	0.641±0.097	84.65±3.13	84.52±6.19	0.411±0.160
S2CNN	60.47±4.41	65.67±4.52	0.214±0.088	62.12±7.73	68.35±8.50	0.286±0.127	69.02±10.53	64.13±5.19	0.048±0.074
DeepSphere	71.80±4.87	79.83±5.76	0.414±0.120	68.18±3.95	77.23±3.15	0.315±0.062	73.95±11.88	78.97±7.56	0.289±0.146
MeshCNN	60.30±6.51	61.69±7.01	0.189±0.121	59.09±8.46	62.81±3.32	0.149±0.109	76.35±7.63	79.60±6.31	0.317±0.125
SurfGNN	70.28±3.95	74.91±5.40	0.374±0.076	74.85±6.18	87.05±5.37	0.511±0.107	79.13±8.86	71.82±6.57	0.305±0.075
SiT	71.13±3.74	80.21±5.06	0.424±0.085	76.36±5.47	88.43±3.20	0.533±0.084	81.30±7.73	84.48±7.24	0.374±0.138
MS-SiT	73.70±5.02	81.70±2.16	0.478±0.077	73.94±7.00	85.34±3.74	0.501±0.105	85.25±5.05	85.94±5.89	**0.465**±0.127
X-SiT	68.40±5.01	81.84±7.66	0.391±0.095	81.52±4.11	90.46±5.53	0.589±0.101	89.58±1.72	74.64±8.73	0.370±0.122
Ours	**84.19**±2.10	**91.54**±2.28	**0.669**±0.045	**86.06**±3.76	**91.80**±3.94	**0.696**±0.071	**90.78**±1.38	**86.76**±5.00	0.451±0.097

Table 2: Out-of-distribution generalization under two distribution shifts: ADNI-CTh→OASIS-2 and ADNI-1→ADNI-2. Mean ± std; ACC, AUC, and macro-F1 are reported in %. Best in **bold**, second-best underlined.

	ADNI-CTh → OASIS-2				ADNI-1 → ADNI-2			
Model	ACC(↑)	AUC(↑)	MCC(↑)	mF1(↑)	ACC(↑)	AUC(↑)	MCC(↑)	mF1(↑)
GCN	53.53±3.10	56.65±0.61	0.099±0.060	51.93±4.38	66.90±1.05	57.70±1.76	0.195±0.021	57.40±2.74
GAT	55.88±2.23	57.29±1.81	0.102±0.055	52.05±5.63	73.17±2.73	74.62±3.22	0.378±0.058	65.97±4.06
GIN	54.41±1.54	58.46±1.04	0.075±0.042	52.15±3.57	67.62±3.81	68.94±0.89	0.264±0.039	60.78±2.96
GraphSAGE	55.15±1.54	60.51±0.70	0.092±0.043	45.05±4.29	68.57±2.00	68.06±2.56	0.245±0.049	59.35±4.63
GraphUNet	54.56±3.13	57.67±3.33	0.106±0.061	52.73±3.85	65.95±2.72	61.41±2.80	0.191±0.042	57.44±2.29
Exphormer	57.21±3.06	57.36±3.36	0.156±0.061	55.32±4.91	67.30±2.39	63.76±3.80	0.211±0.059	57.22±5.03
MoNet	52.94±0.81	60.23±0.66	−0.016±0.059	36.33±3.12	66.11±1.54	65.46±2.77	0.096±0.062	45.75±5.16
SplineCNN	58.24±2.05	63.00±2.81	0.156±0.043	54.29±4.51	69.92±1.47	69.72±5.01	0.286±0.033	61.99±3.21
FeaStNet	56.47±2.92	60.38±2.01	0.113±0.103	52.56±9.37	66.03±3.49	64.27±3.81	0.196±0.077	56.38±5.22
ChebNet	55.59±3.24	68.31±4.33	0.171±0.065	52.58±4.37	69.37±7.34	80.85±2.51	0.435±0.084	68.45±6.98
DiffusionNet	52.21±0.81	53.69±2.11	0.064±0.025	50.33±1.97	55.32±4.66	52.98±4.87	0.058±0.060	51.42±3.74
UGSCNN	55.88±4.39	63.65±5.62	0.149±0.077	51.83±8.35	57.30±7.22	59.17±3.65	0.143±0.064	52.72±6.23
SUNet	62.21±3.74	66.78±2.22	0.252±0.059	61.60±4.22	66.67±3.18	72.30±3.66	0.318±0.038	63.65±2.69
S2CNN	55.15±5.12	59.20±2.69	0.151±0.085	51.01±8.39	56.59±10.66	58.77±3.23	0.074±0.088	47.73±9.55
DeepSphere	58.09±5.77	65.38±3.48	0.173±0.101	54.85±8.48	55.08±11.22	66.42±5.80	0.192±0.108	51.06±9.99
MeshCNN	51.47±5.24	52.41±5.22	0.024±0.102	50.35±4.59	66.40±3.95	58.48±4.75	0.124±0.065	55.74±3.37
SurfGNN	54.41±3.48	59.11±3.17	0.092±0.080	50.12±8.76	68.02±2.65	71.45±6.16	0.205±0.175	55.98±13.52
SiT	53.24±4.78	66.68±4.45	0.090±0.102	45.27±10.80	62.22±16.81	75.85±6.05	0.296±0.172	56.68±17.46
MS-SiT	64.41±4.15	70.71±3.90	0.311±0.084	62.60±5.31	73.25±4.88	79.81±3.16	0.356±0.147	61.41±12.02
X-SiT	54.71±6.91	68.35±8.05	0.134±0.123	49.08±11.45	66.27±13.58	78.82±4.51	0.322±0.106	58.19±12.17
Ours	**67.06**±2.25	**73.06**±2.97	**0.337**±0.046	**66.71**±2.26	**73.73**±1.05	**82.80**±0.63	**0.488**±0.027	**72.85**±1.13

3.2 Quantitative Results

Classification Performance. Table 1 compares the proposed method with 20 baselines, all of which receive the same paired cortical observations $[S_b, S_f]$. The proposed method achieves the highest AUC across all three modalities and the highest MCC on CTh and FDG. On AV45, MS-SiT achieves a slightly higher MCC by an absolute margin of 0.014; given the strong class imbalance in this cohort, we interpret AV45 primarily using AUC and MCC rather than accuracy. The largest gains are observed on CTh, where the proposed method improves over the runner-up by 8.8%p in accuracy and 8.6%p in AUC. The strongest competing method changes with modality, whereas the proposed framework remains among the top-performing methods throughout. These results suggest that the coupling-guided feature extraction provides a consistent inductive bias across modalities: rather than learning representations solely from the observed cortical measurements, the framework estimates a subject-specific, hypothesis-conditioned dependency structure from longitudinal change and biases intermediate features toward it. Because this structure is defined over relationships among spectral deformation modes rather than over the measured values themselves, it provides a common inductive bias across morphometric, metabolic, and molecular cortical measurements.

Table 3: Ablation study on ADNI-CTh. Checkmarks indicate the observed information used by each variant; for zero-change guidance, the observed follow-up S_f is replaced by S_b. ACC and AUC are reported in % as mean±std across five folds. Best in **bold** and second-best underlined.

Variant	S_b	S_f	ΔT	Guidance	ACC↑	AUC↑	MCC↑
No guidance	✓	✓	×	×	$73.18_{\pm 8.56}$	$87.13_{\pm 3.23}$	$0.502_{\pm 0.127}$
Random guidance	✓	✓	✓	Random	$52.59_{\pm 8.80}$	$51.44_{\pm 6.71}$	$0.022_{\pm 0.167}$
Zero-change guidance	✓	×	×	Full	$65.62_{\pm 7.47}$	$67.13_{\pm 15.12}$	$0.243_{\pm 0.195}$
No interval norm.	✓	✓	×	Full	$\underline{81.62_{\pm 2.11}}$	$87.04_{\pm 2.82}$	$0.616_{\pm 0.044}$
Off-diagonal only	✓	✓	✓	Off-diag.	$78.17_{\pm 2.85}$	$85.86_{\pm 3.15}$	$0.564_{\pm 0.048}$
Diagonal only	✓	✓	✓	Diag.	$81.43_{\pm 4.28}$	$\underline{88.23_{\pm 3.36}}$	$\underline{0.626_{\pm 0.072}}$
Full model (Ours)	✓	✓	✓	**Full**	$\mathbf{84.19_{\pm 2.10}}$	$\mathbf{91.54_{\pm 2.28}}$	$\mathbf{0.669_{\pm 0.045}}$

Out-of-Distribution Generalization. We evaluate two complementary distribution shifts: cross-cohort transfer from ADNI-CTh to OASIS-2 and phase transfer from ADNI-1 to ADNI-2. In both settings, models are trained exclusively on the source cohort and evaluated on the target cohort without fine-tuning, using the same preprocessing pipeline throughout. As shown in Table 2, the proposed method ranks first across all eight shift–metric combinations, maintaining consistent advantages under both cross-cohort and phase-transfer settings. On ADNI-2, the proposed method also exhibits lower fold-level variability in AUC than the strongest competing method (0.63 versus 2.51). This suggests reduced sensitivity to the composition of the source training partition. We attribute this transferability in part to the coupling-guided feature extraction. Rather than learning spectral representations solely from the observed cortical measurements, our framework encourages the feature extractor to respect the subject- and hypothesis-conditioned dependency structure estimated from longitudinal change. We suggest that this structural guidance reduces reliance on cohort-specific feature statistics and promotes representations organized around relationships among deformation modes. This interpretation is also consistent with the in-distribution ablation in Table 3, where removing the guidance decreases AUC by 4.41 %p. Together, these results suggest that the learned coupling structure provides a useful inductive bias for generalization beyond the training distribution.

3.3 Ablation Study

We examine whether the performance gain arises from the backbone, the presence of an additional guidance pathway, the observed longitudinal change and its temporal normalization, or the structure of the learned dependency itself. For the component-wise variants, we first construct the full normalized guidance operator $W_i^{(k)}$ and then mask its diagonal or off-diagonal entries, such that $D_i^{(k)}$ is always computed from the complete magnitude matrix $|\widehat{H}_i^{(k)}|$. Table 3 summarizes these controls.

The unguided backbone remains substantially below the full model, showing that the improvement is not explained by the feature extractor alone. More importantly, replacing the learned operator with normalized random guidance causes performance to collapse despite retaining the same hypothesis-conditioned guidance pathway. The benefit therefore depends on the structure encoded by the learned operator, rather than merely on additional feature mixing or access to longitudinal inputs. The longitudinal controls provide a complementary distinction. Using $S_f - S_b$ only, produces a clear but moderate degradation while preserving much of the advantage over the unguided backbone. Thus, temporal normalization improves the representation of longitudinal change but does not by itself account for the gain. In contrast, replacing the observed follow-up with the baseline ($\delta = 0$) causes a substantially larger degradation, even below the unguided backbone. This separation indicates that the useful guidance is primarily driven by the observed subject-specific longitudinal state, with acquisition-interval normalization providing an additional refinement rather than serving as the sole discriminative signal. The component-wise variants further reveal how the dependency structure is used. The diagonal component provides the stronger standalone contribution, whereas the off-diagonal component alone yields a weaker and metric-dependent effect. However, neither component matches the complete operator: combining diagonal structure with cross-mode interactions consistently gives the strongest performance. This pattern indicates that the off-diagonal dependencies are most useful jointly with the mode-wise structure rather than as an isolated source of prediction.

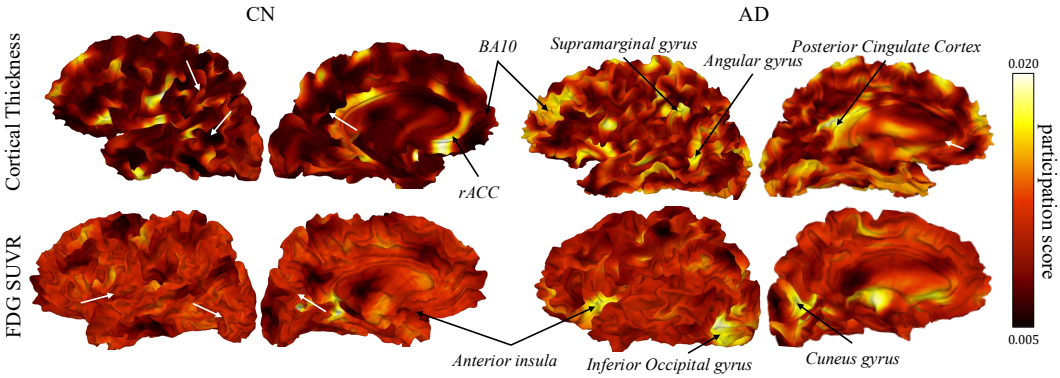


Figure 2: Subject-level top-5% spatial coupling participation maps for randomly selected CN (*left*) and AD (*right*) subjects in CTh (*top*) and FDG-PET (*bottom*). Lateral and medial views are shown for each subject. White arrows indicate the corresponding anatomic locations with black arrows. The shared color scale denotes the participation score.

4 Qualitative Analysis of Emergent Spatial Structure

To visualize the spatial organization of the estimated dependency structure, we construct a top-5% coupling participation map for one randomly selected subject from each of the CN and AD groups. For each cortical location x, we identify the top-5% of locations $x' \neq x$ according to the magnitude of their induced spatial coupling, and assign each location a participation score according to how frequently it appears among the strongest coupling partners of other locations. For CTh, CN subject exhibited high coupling participation in the *rostral Anterior Cingulate Cortex* (rACC), a region associated with age-related cortical thinning (Storsve et al., 2014; Shaw et al., 2016). In contrast, the AD subject emphasizes *Brodmann area 10* (BA10), the *supramarginal* and *angular gyri*, and the *Posterior Cingulate Cortex* (PCC), overlapping with cortical regions implicated in AD-related thinning and the *Default Mode Network* (DMN) (Dickerson et al., 2009; Andrews-Hanna et al., 2010; Greicius et al., 2004). Notably, this large-scale network correspondence emerges without anatomical parcellations, predefined connectomes, or network-level supervision. A complementary pattern is observed in FDG-PET, where high coupling participation appears in *Anterior insula*, *Inferior Occipital gyrus*, and *Cuneus gyrus* associated with AD-related metabolic abnormalities (Mosconi et al., 2004; Herholz et al., 2002; Sorg et al., 2012). Although these maps do not directly measure cortical thinning or regional hypometabolism, the emergence of DMN-related organization in CTh and metabolically relevant posterior territories in FDG-PET provides a consistent cross-modal correspondence with known cortical organization. Together, these observations provide qualitative evidence that the score-derived coupling representation captures modality-relevant cortical structure rather than arbitrary spatial correlations, despite the absence of anatomical, connectomic, or modality-specific spatial supervision.

5 Conclusion

We introduced a spherical score-based framework for classification from pairwise longitudinal cortical observations. By exploiting the local differential structure of a class-conditioned score, the framework estimates a subject- and hypothesis-conditioned spectral dependency operator and converts it into normalized guidance for feature extraction. This augments cortical features with relationships among spectral deformation modes without requiring a predefined anatomical connectome. Across cortical thickness, FDG-PET, and AV45-PET, the framework ranks first in eight of nine modality–metric combinations and remains competitive under cross-cohort and phase-transfer distribution shifts. Ablations show improvements over the unguided and random-guidance controls while supporting complementary diagonal and cross-mode contributions. Limitations include the small AD sample in AV45 and the requirement for a longitudinal pair at inference. Extending the framework to finer spectral bandwidths and multi-stage diagnostic settings is left to future work.

AI USE STATEMENT

In this work, we used generative AI tools for language refinement, text polishing, and improving the readability of the manuscript. We have not used generative AI tools for research conceptualization, literature analysis, data analysis, code generation, experimental design, model development, mathematical derivation, or any other core research process, and no other required-disclosure use cases are applicable to this work. Additionally, we used generative AI tools only for stylistic and grammatical improvements to the manuscript text. We have reviewed all AI-assisted text to ensure that the original scientific meaning, technical content, and claims were preserved without introducing unsupported or altered statements. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work uses de-identified data from the ADNI and OASIS studies, obtained and used under their respective data use agreements. All participants provided informed consent at the acquiring sites, and ethical approval was obtained by the original study investigators. Our analyses were performed entirely on existing imaging and diagnostic records; no new data were collected and no interventions were performed. The proposed framework is intended as a research tool for studying coordinated cortical change. Further validation would be needed before considering potential clinical applications.

REPRODUCIBILITY STATEMENT

We provide the methodological and experimental details necessary to reproduce our results throughout the main paper and appendix. The mathematical assumptions and proofs underlying the score-derived dependency formulation are provided in Appendix A. Appendix B describes the dataset construction and preprocessing procedures, evaluation protocols, and implementation details. Upon acceptance, we will publicly release the software and dataset processing pipeline to facilitate reproduction of the reported results.

REFERENCES

Jessica R. Andrews-Hanna, Jay S. Reidler, Jorge Sepulcre, Renee Poulin, and Randy L. Buckner. Functional-anatomic fractionation of the brain's default network. *Neuron*, 65(4):550–562, 2010. doi: 10.1016/j.neuron.2010.02.005.

Ricardo Baptista, Rebecca Morrison, Olivier Zahm, and Youssef Marzouk. Learning non-gaussian graphical models via hessian scores and triangular transport. *Journal of Machine Learning Research*, 25(85):1–46, 2024.

Fabian Bongratz, Tom Nuno Wolf, Jaume Gual Ramon, and Christian Wachinger. X-sit: Inherently interpretable surface vision transformers for dementia diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 680–690. Springer, 2025.

Heiko Braak and Eva Braak. Neuropathological stageing of alzheimer-related changes. *Acta Neuropathologica*, 82(4):239–259, 1991. doi: 10.1007/BF00308809.

Heiko Braak and Eva Braak. Staging of alzheimer's disease-related neurofibrillary changes. *Neurobiology of Aging*, 16(3):271–278, 1995. doi: 10.1016/0197-4580(95)00021-6.

Taco S. Cohen, Mario Geiger, Jonas Köhler, and Max Welling. Spherical CNNs. In *International Conference on Learning Representations*, 2018.

Simon Dahan, Abdulah Fawaz, Logan Z. J. Williams, Chunhui Yang, Timothy S. Coalson, Matthew F. Glasser, A. David Edwards, Daniel Rueckert, and Emma C. Robinson. Surface vision transformers: Attention-based modelling applied to cortical analysis, 2022. URL https://arxiv.org/abs/2203.16414.

Simon Dahan, Logan Zane John Williams, Daniel Rueckert, and Emma Claire Robinson. The multiscale surface vision transformer. In Ninon Burgos, Caroline Petitjean, Maria Vakalopoulou, Stergios Christodoulidis, Pierrick Coupe, Hervé Delingette, Carole Lartizien, and Diana Mateus (eds.), *Proceedings of The 7nd International Conference on Medical Imaging with Deep Learning*, volume 250 of *Proceedings of Machine Learning Research*, pp. 289–305. PMLR, 03–05 Jul 2024.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 29, pp. 3844–3852, 2016.

Michaël Defferrard, Martino Milani, Frédérick Gusset, and Nathanaël Perraudin. Deepsphere: a graph-based spherical cnn. In *International Conference on Learning Representations (ICLR)*, 2020.

Bradford C. Dickerson, Akram Bakkour, David H. Salat, Eric Feczko, Jenni Pacheco, Douglas N. Greve, Fran Grodstein, Christopher I. Wright, Deborah Blacker, H. Diana Rosas, Reisa A. Sperling, Alireza Atri, John H. Growdon, Bradley T. Hyman, John C. Morris, Bruce Fischl, and Randy L. Buckner. The cortical signature of alzheimer's disease: Regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral Cortex*, 19(3):497–510, 2009. doi: 10.1093/cercor/bhn113.

Michael C. Donohue, Hélène Jacqmin-Gadda, Mélanie Le Goff, Ronald G. Thomas, Rema Raman, Anthony C. Gamst, Laurel A. Beckett, Clifford R. Jack, Michael W. Weiner, Jean-François Dartigues, and Paul S. Aisen. Estimating long-term multivariate progression from short-term data. *Alzheimer's & Dementia*, 10(5 Suppl):S400–S410, 2014. doi: 10.1016/j.jalz.2013.10.003.

Matthias Fey, Jan Eric Lenssen, Frank Weichert, and Heinrich Müller. Splinecnn: Fast geometric deep learning with continuous b-spline kernels. In *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 869–877. IEEE, 2018.

Hongyang Gao and Shuiwang Ji. Graph u-nets. In *International Conference on Machine Learning*, pp. 2083–2092, 2019.

Michael D Greicius, Gaurav Srivastava, Allan L Reiss, and Vinod Menon. Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. *Proceedings of the National Academy of Sciences*, 101(13):4637–4642, 2004.

Seungbo Ha and Ilwoo Lyu. Spharm-net: Spherical harmonics-based convolution for cortical parcellation. *IEEE Transactions on Medical Imaging*, 41(10):2739–2751, 2022. doi: 10.1109/TMI.2022.3168670.

Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

Rana Hanocka, Amir Hertz, Noa Fish, Raja Giryes, Shachar Fleishman, and Daniel Cohen-Or. Meshcnn: A network with an edge. *ACM Transactions on Graphics (TOG)*, 38(4):90:1–90:12, 2019. doi: 10.1145/3306346.3322959.

Karl Herholz, Eric Salmon, Daniela Perani, Jean-Claude Baron, Vjera Holthoff, Lutz Frölich, Peter Schönknecht, K Ito, Rüdiger Mielke, Elke Kalbe, et al. Discrimination between alzheimer dementia and controls by automated analysis of multicenter fdg pet. *Neuroimage*, 17(1):302–316, 2002.

Clifford R. Jr. Jack, David S. Knopman, William J. Jagust, Ronald C. Petersen, Michael W. Weiner, Paul S. Aisen, Leslie M. Shaw, Prashanthi Vemuri, Heather J. Wiste, Stephen D. Weigand, Timothy G. Lesnick, Vernon S. Pankratz, Michael C. Donohue, and John Q. Trojanowski. Tracking pathophysiological processes in alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2):207–216, 2013. doi: 10.1016/S1474-4422(12)70291-0.

Bruno M. Jedynak, Andrew Lang, Bo Liu, Erin Katz, Yuxuan Zhang, Benjamin T. Wyman, David Raunig, Catherine P. Jedynak, Brian Caffo, and Jerry L. Prince. A computational neurodegenerative disease progression score: method and results with the alzheimer's disease neuroimaging initiative cohort. *NeuroImage*, 63(3):1478–1486, 2012. doi: 10.1016/j.neuroimage.2012.07.059.

Chiyu Max Jiang, Jingwei Huang, Karthik Kashinath, Prabhat, Philip Marcus, and Matthias Niessner. Spherical CNNs on unstructured grids. In *International Conference on Learning Representations (ICLR)*, 2019. URL https://openreview.net/forum?id=Bkl-43C9FQ.

Euijin Jung, Miguel Luna, and Sang Hyun Park. Conditional gan with 3d discriminator for mri generation of alzheimer's disease progression. *Pattern Recognition*, 133:109061, 2023. doi: 10.1016/j.patcog.2022.109061.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Seungeun Lee, Sergey Pyatkovskiy, Jaejun Yoo, and Ilwoo Lyu. Spherical Diffusion Process for Score-Guided Cortical Correspondence via Spectral Attention . In *proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*, volume LNCS 15975. Springer Nature Switzerland, September 2025.

Zhuoshuo Li, Jiong Zhang, Youbing Zeng, Jiaying Lin, Dan Zhang, Jianjia Zhang, Duan Xu, Hosung Kim, Bingguang Liu, and Mengting Liu. Surfgnn: A robust surface-based prediction model with interpretability for coactivation maps of spatial and cortical features. *Medical Image Analysis*, pp. 103793, 2025.

Daniel S Marcus, Anthony F Fotenos, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies: longitudinal mri data in nondemented and demented older adults. *Journal of cognitive neuroscience*, 22(12):2677–2684, 2010.

Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 5425–5434. IEEE, 2017.

Lisa Mosconi, Alberto Pupi, M. Teresa R. De Cristofaro, Mozghan Fayyaz, Sandro Sorbi, and Karl Herholz. Functional interactions of the entorhinal cortex: An ^{18}F-FDG PET study on normal aging and alzheimer's disease. *Journal of Nuclear Medicine*, 45(3):382–392, 2004.

Susanne G. Mueller, Michael W. Weiner, Leon J. Thal, Ronald C. Petersen, Clifford R. Jack, William Jagust, John Q. Trojanowski, Arthur W. Toga, and Laurel Beckett. The alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, 15(4):869–877, 2005. doi: 10.1016/j.nic.2005.09.008.

Thomas G. Ohm, Helmut Müller, Heiko Braak, and Johannes Bohl. Close-meshed prevalence rates of different stages as a tool to uncover the rate of alzheimer's disease-related neurofibrillary changes. *Neuroscience*, 64(1):209–217, 1995. doi: 10.1016/0306-4522(95)90397-P.

Jiahong Ouyang, Qingyu Zhao, Edith V. Sullivan, Adolf Pfefferbaum, Susan F. Tapert, Ehsan Adeli, and Kilian M. Pohl. Longitudinal pooling & consistency regularization to model disease progression from mris. *IEEE Journal of Biomedical and Health Informatics*, 25(6):2082–2092, 2021. doi: 10.1109/JBHI.2020.3042447.

Ashish Raj, Amy Kuceyeski, and Michael Weiner. A network diffusion model of disease progression in dementia. *Neuron*, 73(6):1204–1215, 2012. doi: 10.1016/j.neuron.2011.12.040.

William W. Seeley, Richard K. Crawford, Juan Zhou, Bruce L. Miller, and Michael D. Greicius. Neurodegenerative diseases target large-scale human brain networks. *Neuron*, 62(1):42–52, 2009. doi: 10.1016/j.neuron.2009.03.024.

Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. Diffusionnet: Discretization agnostic learning on surfaces. *ACM Transactions on Graphics (TOG)*, 41(3):1–16, 2022. doi: 10.1145/3507905.

Marnie E. Shaw, Walter P. Abhayaratna, Perminder S. Sachdev, Kaarin J. Anstey, and Nicolas Cherbuin. Cortical thinning at midlife: The path through life study. *Brain Topography*, 29(6):875–884, 2016. doi: 10.1007/s10548-016-0509-z.

Hamed Shirzad, Ameya Velingker, Balaji Venkatachalam, Danica J Sutherland, and Ali Kemal Sinop. Exphormer: Sparse transformers for graphs. In *International conference on machine learning*, pp. 31613–31632. PMLR, 2023.

Christian Sorg, Nicholas Myers, Petra Redel, Peter Bublak, Valentin Riedl, Andrei Manoliu, Robert Perneczky, Timo Grimmer, Alexander Kurz, Hans Förstl, Alexander Drzezga, Hermann J. Müller, Afra M. Wohlschläger, and Kathrin Finke. Asymmetric loss of parietal activity causes spatial bias in prodromal and mild Alzheimer's disease. *Biological Psychiatry*, 71:798–804, 2012. doi: 10.1016/j.biopsych.2011.09.027.

Alessio Spantini, Daniele Bigoni, and Youssef Marzouk. Inference via low-dimensional couplings. *Journal of Machine Learning Research*, 19(66):1–71, 2018.

Andreas B. Storsve, Anders M. Fjell, Christian K. Tamnes, Lars T. Westlye, Knut Overbye, Hilde W. Aasland, and Kristine B. Walhovd. Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: Regions of accelerating and decelerating change. *Journal of Neuroscience*, 34(25):8488–8498, 2014. doi: 10.1523/JNEUROSCI.0391-14.2014.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

Nitika Verma, Edmond Boyer, and Jakob Verbeek. Feastnet: Feature-steered graph convolutions for 3d shape analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2598–2606, 2018.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations (ICLR)*, 2019.

Alexandra L. Young, Neil P. Oxtoby, Pankaj Daga, David M. Cash, Nick C. Fox, Sebastien Ourselin, Jonathan M. Schott, and Daniel C. Alexander. A data-driven model of biomarker changes in sporadic alzheimer's disease. *Brain*, 137(9):2564–2577, 2014. doi: 10.1093/brain/awu176.

Alexandra L. Young, Razvan V. Marinescu, Neil P. Oxtoby, Martina Bocchetta, Keir Yong, Nicholas C. Firth, David M. Cash, David L. Thomas, Katrina M. Dick, Jorge Cardoso, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. *Nature Communications*, 9:4273, 2018. doi: 10.1038/s41467-018-05892-0.

Fenqiang Zhao, Shunren Xia, Ziyan Wu, Dingna Duan, Lifang Wang, Weili Lin, John H. Gilmore, Dinggang Shen, and Gang Li. Spherical u-net on cortical surfaces: Methods and applications. In *Information Processing in Medical Imaging (IPMI)*, volume 11492 of *Lecture Notes in Computer Science*, pp. 855–866. Springer, 2019. doi: 10.1007/978-3-030-20351-1_67.

A Theoretical Details

A.1 Proof of Proposition 1

Fix a diagnostic class y. Let

$$C = (C_1, \ldots, C_M)$$

have a strictly positive twice continuously differentiable density $p(c \mid y)$ on $\mathbb{R}^M$. For two distinct coordinates $a \neq b$, write

$$u = c_a, \qquad v = c_b, \qquad z = c_{-(a,b)},$$

and define

$$\ell(u, v, z) := \log p(u, v, z \mid y).$$

By definition,

$$[H_y(c)]_{ab} = -\frac{\partial^2}{\partial u \, \partial v} \ell(u, v, z).$$

We first show the forward implication. Suppose

$$C_a \perp C_b \mid C_{-(a,b)}, y.$$

Since the joint density is strictly positive, conditional independence implies

$$p(u, v, z \mid y) = p(u \mid z, y)\, p(v \mid z, y)\, p(z \mid y).$$

Taking logarithms yields

$$\ell(u, v, z) = \log p(u \mid z, y) + \log p(v \mid z, y) + \log p(z \mid y).$$

The first term is independent of v, while the second is independent of u. Therefore,

$$\frac{\partial^2}{\partial u\, \partial v} \ell(u, v, z) = 0,$$

and hence

$$[H_y(c)]_{ab} = 0 \qquad \text{for all } c \in \mathbb{R}^M.$$

Conversely, suppose

$$[H_y(c)]_{ab} = 0 \qquad \text{for all } c \in \mathbb{R}^M.$$

Equivalently,

$$\frac{\partial^2}{\partial u\, \partial v} \ell(u, v, z) = 0$$

for every (u, v, z). For fixed z, choose arbitrary reference values u_0 and v_0. The vanishing mixed derivative implies the additive decomposition

$$\ell(u, v, z) = \ell(u, v_0, z) + \ell(u_0, v, z) - \ell(u_0, v_0, z).$$

Defining

$$A(u, z) = \ell(u, v_0, z) - \ell(u_0, v_0, z),$$

and

$$B(v, z) = \ell(u_0, v, z),$$

we obtain

$$\ell(u, v, z) = A(u, z) + B(v, z),$$

and therefore

$$p(u, v, z \mid y) = \exp(A(u, z)) \exp(B(v, z)).$$

Conditioning on z gives

$$p(u, v \mid z, y) = \frac{\exp(A(u, z)) \exp(B(v, z))}{\displaystyle\iint \exp(A(u', z)) \exp(B(v', z))\, du'\, dv'}.$$

The denominator factorizes as

$$\iint \exp(A(u', z)) \exp(B(v', z))\, du'\, dv' = \left[\int \exp(A(u', z))\, du'\right] \left[\int \exp(B(v', z))\, dv'\right].$$

Thus,

$$p(u, v \mid z, y) = p(u \mid z, y) p(v \mid z, y),$$

which implies

$$C_a \perp C_b \mid C_{-(a,b)}, y.$$

This proves the equivalence. □

Local versus global interpretation. Proposition 1 distinguishes the local interpretation of an individual entry of $H_y(c)$ from the global conditional-independence statement. A nonzero value $[H_y(c)]_{ab}$ at a particular state c indicates mixed local curvature of the class-conditional log-density around that state. In contrast, conditional independence is characterized by the vanishing of this entry throughout the support. Accordingly, we use $H_y(c)$ as a state-dependent local coupling representation rather than as a state-independent adjacency matrix.

A.2 Score Jacobian under the Spherical OU Perturbation

We next make explicit the probabilistic quantity represented by the score Jacobian at a finite diffusion scale. The spherical OU forward process defined in Sec. 2.3 gives

$$c_t = A_t c_0 + \Sigma_t^{1/2}\epsilon, \qquad \epsilon \sim \mathcal{N}(0, I),$$

where

$$A_t = \operatorname{diag}\left(e^{-\lambda_l t}\right), \qquad \Sigma_t = \operatorname{diag}\left(1 - e^{-2\lambda_l t}\right),$$

and $\lambda_l = l(l+1)$. Hence,

$$p_t(c_t \mid c_0) = \mathcal{N}\left(c_t; A_t c_0, \Sigma_t\right).$$

The class-conditional marginal is

$$p_t(c_t \mid y) = \int p_t(c_t \mid c_0) p_0(c_0 \mid y)\, dc_0.$$

Let

$$s_{y,t}(c_t) = \nabla_{c_t} \log p_t(c_t \mid y).$$

Differentiating the marginal density under the integral gives

$$\begin{aligned} s_{y,t}(c_t) &= \mathbb{E}\left[\nabla_{c_t} \log p_t(c_t \mid c_0) \mid c_t, y\right] \\ &= -\mathbb{E}\left[\Sigma_t^{-1}(c_t - A_t c_0) \mid c_t, y\right] \\ &= \Sigma_t^{-1}\left(A_t \mathbb{E}[c_0 \mid c_t, y] - c_t\right). \end{aligned}$$

Writing

$$m_t(c_t, y) := \mathbb{E}[c_0 \mid c_t, y],$$

we therefore have

$$s_{y,t}(c_t) = \Sigma_t^{-1}\left(A_t m_t(c_t, y) - c_t\right).$$

To differentiate the posterior mean, note that

$$p(c_0 \mid c_t, y) \propto p_0(c_0 \mid y) p_t(c_t \mid c_0).$$

The derivative of its log density with respect to c_t is

$$\begin{aligned} \nabla_{c_t} \log p(c_0 \mid c_t, y) &= \nabla_{c_t} \log p_t(c_t \mid c_0) - s_{y,t}(c_t) \\ &= \Sigma_t^{-1} A_t \left(c_0 - m_t(c_t, y)\right). \end{aligned}$$

Using the standard derivative-of-expectation identity,

$$\begin{aligned} \nabla_{c_t} m_t(c_t, y) &= \operatorname{Cov}\left(c_0, \nabla_{c_t} \log p(c_0 \mid c_t, y) \mid c_t, y\right) \\ &= \operatorname{Cov}(c_0 \mid c_t, y) A_t^\top \Sigma_t^{-1}. \end{aligned}$$

Differentiating the marginal score then yields

$$\nabla_{c_t} s_{y,t}(c_t) = \Sigma_t^{-1} A_t \operatorname{Cov}(c_0 \mid c_t, y) A_t^\top \Sigma_t^{-1} - \Sigma_t^{-1}.$$

Consequently, the finite-time local dependency operator is

$$\boxed{H_{y,t}(c_t) = \Sigma_t^{-1} - \Sigma_t^{-1} A_t \operatorname{Cov}(c_0 \mid c_t, y) A_t^\top \Sigma_t^{-1}} \tag{4}$$

with

$$H_{y,t}(c_t) = -\nabla_{c_t} s_{y,t}(c_t).$$

Equation 4 shows that the operator used by our method is a finite-time property of the diffused class-conditional density itself. Importantly, our interpretation does not require the conditional-independence structure of the clean density p_0 to be preserved by Gaussian smoothing, nor do we require H_{y,t_g} to equal the clean operator at $t = 0$.

Gaussian special case. Suppose

$$c_0 \mid y \sim \mathcal{N}(\mu_y, \Gamma_y).$$

Because the OU perturbation is linear Gaussian,

$$c_t \mid y \sim \mathcal{N}\left(A_t\mu_y,\ A_t\Gamma_y A_t^\top + \Sigma_t\right).$$

The corresponding score is

$$s_{y,t}(c_t) = -\left(A_t\Gamma_y A_t^\top + \Sigma_t\right)^{-1}(c_t - A_t\mu_y),$$

so that

$$\boxed{H_{y,t} = \left(A_t\Gamma_y A_t^\top + \Sigma_t\right)^{-1}.}$$

Thus, for Gaussian class-conditional distributions, $H_{y,t}$ is state-independent and coincides exactly with the finite-time precision matrix. For general non-Gaussian distributions, the same operator remains well defined but can depend on the evaluation state c_t.

A.3 Proof of Proposition 2

Let

$$\mathcal{V}_L = \bigoplus_{l=1}^{L} \mathcal{H}_l$$

denote the space of scalar spherical functions band-limited to degree L, excluding the degree-zero component, where $\mathcal{H}_l$ is the $(2l+1)$-dimensional degree-l spherical harmonic subspace. A rotation $R \in SO(3)$ acts on a spherical function by

$$[\rho(R)f](x) = f(R^{-1}x).$$

Each $\mathcal{H}_l$ is an irreducible representation of $SO(3)$, and distinct degrees correspond to inequivalent irreducible representations.

Let

$$T : \mathcal{V}_L \to \mathcal{V}_L$$

be linear and $SO(3)$-equivariant, i.e.,

$$T\rho(R) = \rho(R)T \qquad \forall R \in SO(3).$$

Let P_l denote the orthogonal projection onto $\mathcal{H}_l$ and define

$$T_{l'l} := P_{l'}TP_l.$$

Equivariance of T implies that each $T_{l'l}$ is an intertwining map between the degree-l and degree-l' representations.

For $l' \neq l$, the two representations are inequivalent. By Schur's lemma,

$$T_{l'l} = 0.$$

For $l' = l$, Schur's lemma gives

$$T_{ll} = \gamma_l I_{2l+1}$$

for some scalar $\gamma_l \in \mathbb{R}$. Therefore,

$$\boxed{T = \bigoplus_{l=1}^{L} \gamma_l I_{2l+1}.}$$

Conversely, any operator of this block-scalar form commutes with the rotation representation and is therefore $SO(3)$-equivariant. This proves Proposition 2. □

Spatial form of the isotropic family. For an operator

$$T = \bigoplus_{l=1}^{L} \gamma_l I_{2l+1},$$

the induced spatial kernel is

$$K_T(x, x') = \sum_{l=1}^{L} \gamma_l \sum_{m=-l}^{l} Y_l^m(x) Y_l^m(x').$$

By the spherical harmonic addition theorem,

$$\sum_{m=-l}^{l} Y_l^m(x) Y_l^m(x') = \frac{2l+1}{4\pi} P_l(x \cdot x'),$$

and hence

$$K_T(x, x') = \sum_{l=1}^{L} \gamma_l \frac{2l+1}{4\pi} P_l(x \cdot x').$$

Thus, the spatial kernel associated with this operator family depends only on geodesic separation. In contrast, the score-derived operator $H_{y,t}(c)$ is not constrained to the block-scalar family and can contain cross-degree and non-scalar within-degree interactions.

Because cortical fields are represented in a common anatomically registered spherical coordinate system, we do not impose global rotational equivariance on the learned coupling operator itself. Position-dependent interactions can therefore retain anatomical meaning in the shared cortical coordinate system.

B Dataset and Experimental Details

B.1 Cohorts Demographics

We construct longitudinal cortical datasets from ADNI for three imaging modalities: cortical thickness (CTh), FDG-PET, and AV45-PET. For each modality, a paired scan is selected according to the pairing criterion used in the main experiments. All pairs belonging to the same subject remain within the same cross-validation partition. Participant identifiers were explicitly checked to ensure that the ADNI-1 source and ADNI-2 target cohorts were subject-disjoint.

Table 4: Demographics for ADNI. Mean ± std or count.

Biomarker	Group	n	Age (mean ± sd)	Male / Female	Scan Interval (yr, mean ± sd)
Thickness	All	582	$75.5_{\pm 6.7}$	309 / 273	$1.52_{\pm 0.66}$
	CN	363	$75.8_{\pm 5.9}$	187 / 176	$1.52_{\pm 0.67}$
	AD	219	$75.0_{\pm 7.8}$	122 / 97	$1.52_{\pm 0.66}$
FDG	All	330	$75.9_{\pm 6.4}$	195 / 135	$1.92_{\pm 1.00}$
	CN	222	$76.0_{\pm 6.1}$	130 / 92	$2.03_{\pm 1.08}$
	AD	108	$75.6_{\pm 7.1}$	65 / 43	$1.71_{\pm 0.79}$
AV45	All	326	$76.3_{\pm 7.0}$	163 / 163	$2.55_{\pm 1.14}$
	CN	296	$76.2_{\pm 7.0}$	145 / 151	$2.58_{\pm 1.15}$
	AD	30	$77.2_{\pm 6.8}$	18 / 12	$2.27_{\pm 0.97}$

B.2 Evaluation Details

All experiments use subject-level five-fold cross-validation with disjoint test folds and a separate validation split for model selection. For each fold, both training stages are independently initialized. The score network is trained only on the training partition and is frozen before downstream classifier training; held-out test subjects are not used for score-network training, checkpoint selection, or hyperparameter selection.

Table 5: Demographics for OASIS-2. Mean ± std or count.

Biomarker	Group	n	Age (mean ± sd)	Male / Female	Scan Interval (yr, mean ± sd)
Thickness	All	136	$75.3_{\pm 7.5}$	58 / 78	$2.83_{\pm 1.44}$
	CN	72	$75.4_{\pm 8.2}$	22 / 50	$3.32_{\pm 1.47}$
	AD	64	$75.1_{\pm 6.7}$	36 / 28	$2.28_{\pm 1.19}$

The score model operates on interval-normalized longitudinal SH coefficients and is trained by denoising score matching under the spherical OU perturbation described in Sec. 2.3. The guidance scale t_g and spectral bandwidth L_g are selected using validation performance. For each subject and diagnostic hypothesis, the score Jacobian is computed on the retained SH modes, symmetrized to obtain $\widehat{H}_i^{(k)}$, and converted into the normalized guidance operator $W_i^{(k)}$ as defined in Sec. 2.4. The resulting operators are cached and kept fixed during downstream classifier training.

All baselines follow the same subject-level data partitions and preprocessing pipeline. Unless otherwise stated, paired baseline and follow-up measurements are provided as separate input channels.

B.3 Implementation Details

To ensure a consistent and fair comparison, the proposed method and all baselines are tuned using a common search space for shared hyperparameters, as summarized in Table 6. Architecture-specific hyperparameters are additionally tuned using model-specific search spaces where applicable (Table 7). All hyperparameter configurations are selected exclusively according to validation ACC, without access to the held-out test fold. This protocol provides a consistent model-selection procedure across methods while accommodating architecture-specific design choices. Hyperparameters not explicitly searched are fixed according to the corresponding reference implementation unless otherwise specified. The score network is implemented as a conditional U-Net and optimized using AdamW.

Table 6: Shared hyperparameter search space used across models.

Hyperparameter	**Search Space**
Learning rate	$\{1\times10^{-5}, 1\times10^{-4}, 3\times10^{-4}, 5\times10^{-4}, 7\times10^{-4}, 1\times10^{-3}\}$
Dropout rate	$\{0, 0.1, 0.3, 0.5, 0.7, 0.9\}$
Weight decay	$\{1\times10^{-5}, 1\times10^{-4}, 1\times10^{-3}\}$
Hidden dimension	$\{8, 16, 32, 64, 128\}$
Number of attention heads	$\{2, 4, 8, 12, 16\}$ (GAT, SiT, MS-SiT, X-SiT, FeaStNet)

Table 7: Model-specific hyperparameter search spaces. Parameters not listed here follow the corresponding reference implementation unless otherwise specified.

Model	**Hyperparameter**	**Search Space**
ChebNet	Chebyshev order K_{Cheb}	$\{3, 4, 5, 6, 7, 8\}$
DiffusionNet	Channel width C_{width}	$\{16, 32, 64\}$
DiffusionNet	Number of diffusion blocks N_{block}	$\{4, 8, 12\}$
SurfGNN	Number of TSL stages	$\{2, 4, 8, 12\}$
Proposed	Guidance scale t_g	$\{4.5, 5.0, 5.5, 6.0\}\times10^{-4}$
Proposed	Guidance bandwidth L_g	$\{9, 10, 11, 12, 13\}$

B.4 Scan Acquisition Interval Fairness Control

Longitudinal neuroimaging cohorts may exhibit diagnosis-associated differences in follow-up availability, resulting in different acquisition-interval distributions across diagnostic groups. Because the proposed method explicitly uses the acquisition interval ΔT through the normalized longitudinal deformation $(S_f - S_b)/\Delta T$, we examine whether ΔT itself provides diagnostic information.

For ADNI-CTh, where longitudinal structural scans permit flexible pair selection, we construct an interval-controlled cohort. For each subject, the latest available scan is used as the follow-up. For

CN subjects, the baseline scan is selected among preceding scans such that the resulting group-level acquisition interval is matched as closely as possible to that of the AD group. The same subject-level partitioning protocol is subsequently used for all models and ablations.

As a direct diagnostic check, we train a univariate logistic-regression classifier using only ΔT under the same subject-level cross-validation partitions used in the main experiments. The resulting AUCs are 48.17% for ADNI-CTh, 56.61% for ADNI-FDG, and 59.96% for ADNI-AV45. Thus, acquisition interval alone is non-discriminative on the interval-controlled CTh cohort, whereas weak residual associations remain in the PET cohorts. Strict interval matching is more constrained for FDG- and AV45-PET because valid longitudinal samples require compatible structural-MRI and PET observations; we therefore report these interval-only results explicitly rather than assuming that interval-related confounding is absent.

C Hyperparameter Sensitivity

We further analyze the sensitivity of the proposed guidance mechanism to the guidance diffusion scale t_g and the spectral guidance bandwidth L_g. As shown in Figure 3, performance remains relatively stable over the evaluated range of t_g, indicating that the extracted local dependency structure is not overly sensitive to small variations in the diffusion scale used for Jacobian evaluation. For L_g, performance generally improves as the spectral bandwidth increases. This trend is consistent with the inclusion of additional SH modes, which allows the guidance operator to capture finer spatial variation and richer mode-to-mode interactions. Overall, the sensitivity analysis shows stable behavior with respect to t_g and a generally increasing performance trend with spectral resolution. For sensitivity analysis only, we additionally evaluate values outside the model-selection search range; these configurations were not used for hyperparameter selection.

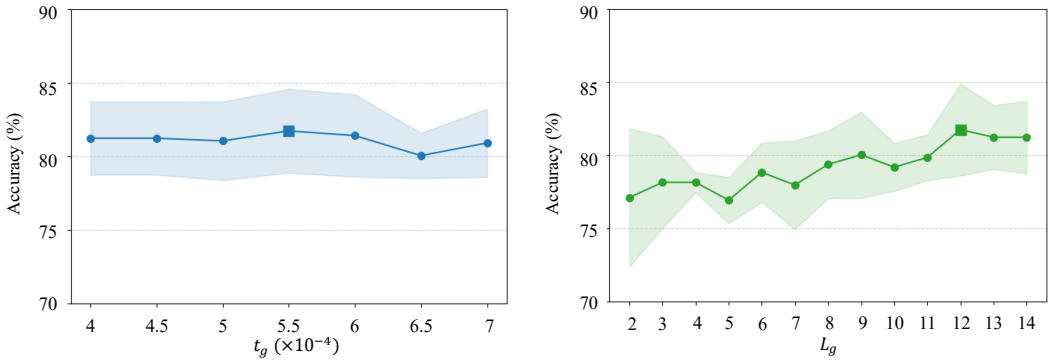


Figure 3: Model sensitivity to t_g (*left*) and L_g (*right*).