ENERGY-DRIVEN HIGHER-ORDER GRAPH STATE EVOLUTION FOR BRAIN NETWORK ANALYSIS

Anonymous authors
Paper under double-blind review

ABSTRACT

Characterizing group-wise interactions among brain regions is important for understanding the behaviors of brain disorders. While conventional graph learning methods propagate information through pairwise edges, higher-order graph models introduce group-wise structure but commonly use it as an aggregation pattern under predefined propagation rules. Physics-inspired graph models instead derive state evolution from energy-based dynamics, but existing formulations are built from node-wise or pairwise interactions on fixed graphs. In this regard, we introduce HOPE, a higher-order graph learning framework that discovers instance-specific latent factors with sparse soft memberships and converts their group-wise interactions into factor-wise energies governing node-state evolution. HOPE alternates factor inference with structure-aware gradient-flow updates, allowing higher-order interaction structures and node states to co-evolve, while each fixed-factor state evolution follows dissipative continuous-time dynamics. Experiments on diverse brain networks show that HOPE consistently improves disease prediction across neurodegenerative and neurodevelopmental disorders. Moreover, the learned factors and energies identify disease-relevant ROI groups, revealing structured higher-order interactions associated with heterogeneous brain disorders.

1 INTRODUCTION

Complex systems are often governed by interactions beyond isolated one-to-one links. The stability of many-body physical systems (Della Rossa et al., 2023; Elhatisari et al., 2024), the function of protein complexes (Hu et al., 2021; Feng et al., 2021), and the emergence of collective behavior in social groups (Zhu et al., 2018; Battiston et al., 2020) all arise from couplings among multiple entities. The human brain is a prime example, since cognitive function emerges from the coordinated activity of distributed regions of interest (ROIs) rather than from isolated pairs, and characterizing these group-wise interactions is crucial for the early detection of brain disorders (Han et al., 2025; Deng et al., 2025; Millán et al., 2025). To model such interactions, the brain is commonly represented as a network, with structural connectivity derived from Diffusion Tensor Imaging (DTI) or functional connectivity estimated from functional Magnetic Resonance Imaging (fMRI).

Brain disorders induce distinct alterations in the structure and organization of brain networks. Neurodegenerative disorders such as Alzheimer's disease (AD) and Parkinson's disease (PD) involve progressive deterioration that disrupts white-matter pathways (Matthews et al., 2012; Reid & Evans, 2013; Yang et al., 2023; Festa et al., 2024) and changes regional properties such as cortical thickness and protein deposition measured by Positron Emission Tomography (PET). Neurodevelopmental disorders such as Autism Spectrum Disorder (ASD) instead exhibit atypical functional organization, where coordination among distributed regions deviates from typical development (Heinsfeld et al., 2018). Neuroimaging captures complementary facets of these alterations through structural connectivity from Diffusion-Weighted Imaging (DWI), functional connectivity from fMRI, and regional abnormalities from MRI or PET. This motivates models that jointly leverage network and regional information for early diagnosis and disease biomarker discovery.

Graph neural networks (GNNs) have accordingly become a standard tool for analyzing brain networks with multimodal features (Kan et al., 2022; Peng et al., 2025; Yang et al., 2025). However, stacking pairwise message-passing layers enlarges the receptive field but does not explicitly model group interactions and may exacerbate oversmoothing. Although higher-order models, e.g., hypergraph neural networks, explicitly encode group-wise relations (Feng et al., 2019; Liu et al., 2024;

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.

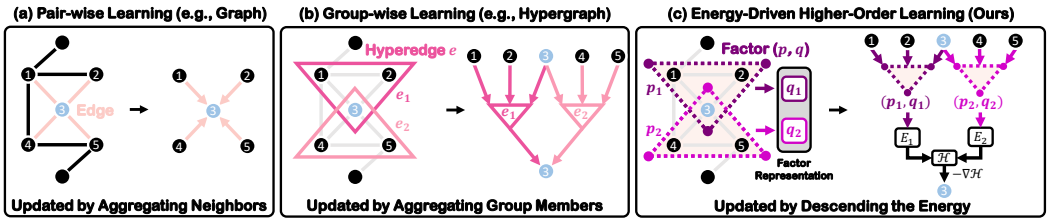


Figure 1: Comparison of how interaction structures govern node-state updates. Pairwise learning (a) and group-wise learning (b) use edges or hyperedges as aggregation operators. In contrast, HOPE (c) represents each learned factor by a soft membership p and latent representation q, whose factor energies form the total energy $\mathcal{H}$. Then, node states evolve along $-\nabla\mathcal{H}$ under the learned higher-order structure.

Mei et al., 2025), and recent brain-specific variants further learn higher-order relations or structures (Wang et al., 2023; Qiu et al., 2024; Sim et al., 2026), the obtained higher-order structure commonly acts as an aggregation operator that routes messages among nodes, while the propagation rule remains predefined and the structure is typically fixed once estimated. Beyond learning how information is aggregated, modeling how regional states evolve under network interactions requires a principled mechanism that explicitly governs state transitions. Physics-inspired GNNs address this perspective by deriving propagation from continuous-time dynamics or energy-based evolution (Chamberlain et al., 2021; Rusch et al., 2022; Di Giovanni et al., 2023), however, their dynamics are primarily built from node-wise or pairwise interactions on a fixed observed graph.

Motivated by these limitations, we develop a framework in which learned higher-order structures directly define the energy governing graph state evolution. As illustrated in Fig. 1, conventional pairwise and group-wise methods use interaction structures as aggregation operators, whereas our approach uses learned higher-order factors to define an energy that drives node-state evolution. Specifically, our framework discovers latent higher-order factors, each represented by a sparse soft node membership and a latent group representation, and converts their group-wise interactions into factor-wise energies. These energies jointly define the total energy $\mathcal{H}$, whose gradient drives node-state evolution through a structure-aware mobility operator, yielding dissipative continuous-time dynamics. By repeatedly re-inferring the factors from evolving states, the higher-order structure and node states co-evolve through alternating updates rather than remaining fixed across evolution steps. Based on these ideas, we propose **H**igher-**O**rder **P**hysics-inspired **E**volution, i.e., **HOPE**.

The **key contributions** of our work are as follows.

i) Introducing an *energy-driven framework* that discovers higher-order factors and converts their interactions into factor-wise energies governing node-state evolution.

ii) Designing *structure-aware dissipative state dynamics* within an alternating scheme that co-evolves higher-order structures and node states.

iii) Establishing *theoretical properties* by proving continuous-time energy dissipation and characterizing factor granularity through temperature-controlled sparse assignments.

Extensive validation was performed on three independent brain-network benchmarks spanning neurodegenerative and neurodevelopmental disorders, using structural DTI networks with multimodal regional biomarkers and fMRI-derived functional networks with region-wise Blood-Oxygen-Level-Dependent (BOLD) signals. HOPE consistently outperforms state-of-the-art methods across these settings, demonstrating its effectiveness across diverse brain-network modalities and disorder types. Moreover, the learned factors and their energies identify disease-relevant ROI groups and group-wise interactions, providing physics-inspired interpretability across heterogeneous disorders.

2 Related Works

Higher-order Graph Learning. Since GNNs aggregate features from neighbors (Veličković et al., 2018; Chen et al., 2020), their interactions remain pairwise and struggle to capture rich group-wise dependencies in complex graphs such as brain connectomes. Several studies have incorporated higher-order structures into graphs to capture complex geometric relationships. For example, HGNN (Feng et al., 2019) captures high-order correlations using hypergraph representations, HNHN (Dong et al., 2020) introduces nonlinear activations on both nodes and hyperedges, UniGCNII (Huang & Yang, 2021) mitigates over-smoothing in hypergraph learning, HyperDrop (Jo et al., 2021) removes

task-irrelevant edges by hypergraph-based pooling, and HyperGT (Liu et al., 2024) encodes global and local structure with a hypergraph transformer. Recently, DHHNN (Mei et al., 2025) dynamically adjusts relationships of nodes and hyperedges in hyperbolic space.

Physics-Inspired Graph Learning. Another line of work derives propagation from continuous-time dynamics rather than from a hand-designed layer. For example, GRAND (Chamberlain et al., 2021) discretizes a diffusion equation on graphs to obtain message passing, PDE-GCN (Eliasof et al., 2021) builds architectures from diffusion and wave equations to control smoothing, and GraphCON (Rusch et al., 2022) builds deep graph networks through ODE dynamics with flexible GNN layers serving as the coupling functions. More broadly, GRAFF (Di Giovanni et al., 2023) characterizes graph convolution as the gradient flow of a parametric energy. Although these formulations offer principled state evolution, their dynamics are primarily built from node-wise or pairwise interactions over the observed graph, without explicitly incorporating learned group-wise couplings.

Brain Network Analysis. Since nodes (i.e., ROIs) within a shared brain template carry distinct locations and unique identities, recent works have adopted or extended GNNs to analyze brain networks. IBGNN (Cui et al., 2022b) identifies disorder-specific ROIs and prominent connections, BrainGB (Cui et al., 2022a) standardizes network construction pipelines and modularizes GNN designs, BrainNetTF (Kan et al., 2022) learns fully pairwise attention with a transformer, and SGCN (Zhou et al., 2024) learns sparse importance probabilities for brain regions and connections. Notably, ALTER (Yu et al., 2024) injects a biased random walk into a transformer for long-range dependencies, BioBGT (Peng et al., 2025) encodes the small-world architecture of brain graphs via network entanglement, and BQN (Yang et al., 2025) adopts quadratic networks for brain network analysis.

Beyond conventional graph-based approaches, recent studies have explored higher-order frameworks in brain networks. For example, dwHGCN (Wang et al., 2023) learns weights over predefined hyperedges to characterize high-order relations among brain regions, and HyBRiD (Qiu et al., 2024) builds discrete hyperedges by sampling binary masks, each carrying a learnable weight. Recently, MuHL (Sim et al., 2026) constructs hierarchical node features and dynamically learns high-order interactions through continuous hyperedge construction. While higher-order methods typically use learned structures as aggregation operators and physics-inspired methods primarily build dynamics from pairwise interactions, HOPE converts discovered higher-order factors into factor-wise energies that jointly define the total energy governing ROI-state evolution.

3 PRELIMINARIES

Graph Dirichlet Energy. An undirected graph is represented as $\mathcal{G}=\{\mathbf{A}, \mathbf{X}\}$ with N nodes, where $\mathbf{A} \in \mathbb{R}^{N\times N}$ is a symmetric adjacency matrix encoding pairwise connectivity and $\mathbf{X} \in \mathbb{R}^{N\times F}$ contains F node features. Let $\mathbf{D}$ be the degree matrix with $D_{i,i}=\sum_j A_{i,j}$ and $\mathbf{L}=\mathbf{D}-\mathbf{A}$ the graph Laplacian. We define the degree-normalized adjacency and Laplacian as $\tilde{\mathbf{A}}=\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ and $\tilde{\mathbf{L}}=\mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$, where $\tilde{\mathbf{L}} \succeq 0$ has eigenvalues in $[0, 2]$. To quantify graph smoothness, for a node representation $\mathbf{Z}=[\mathbf{z}_1, \ldots, \mathbf{z}_N]^\top \in \mathbb{R}^{N\times d}$, the quadratic form of $\tilde{\mathbf{L}}$ defines the Dirichlet energy

$$\mathrm{Tr}(\mathbf{Z}^\top \tilde{\mathbf{L}}\mathbf{Z}) = \frac{1}{2}\sum_{i,j} A_{i,j} \Big\| \frac{\mathbf{z}_i}{\sqrt{D_{i,i}}} - \frac{\mathbf{z}_j}{\sqrt{D_{j,j}}} \Big\|^2, \tag{1}$$

which measures the smoothness of node representations over the graph (Zhou et al., 2021).

Gradient Flow on Graph-Structured States. A growing body of work interprets deep graph learning as the discretization of continuous-time dynamics on graphs (Chamberlain et al., 2021; Eliasof et al., 2021; Rusch et al., 2022). Within this perspective, certain graph neural dynamics can be formulated as gradient flows that evolve node representations by decreasing an associated energy functional (Di Giovanni et al., 2023). Formally, let $\mathbf{Z}(t) \in \mathbb{R}^{N\times d}$ denote the d-dimensional evolving latent states of N nodes at virtual time t, and let $\mathcal{H}(\mathbf{Z}) \in \mathbb{R}$ be a differentiable energy. Generally, a dissipative gradient-flow dynamics is given by

$$\frac{d\mathbf{Z}}{dt} = -\nabla_{\mathbf{Z}}\mathcal{H}(\mathbf{Z}). \tag{2}$$

Thus, the state trajectory evolves toward lower-energy configurations, with the corresponding change in energy given by

$$\frac{d\mathcal{H}}{dt} = \left\langle \nabla_{\mathbf{Z}}\mathcal{H}, \frac{d\mathbf{Z}}{dt} \right\rangle = -\|\nabla_{\mathbf{Z}}\mathcal{H}\|_F^2 \leq 0, \tag{3}$$

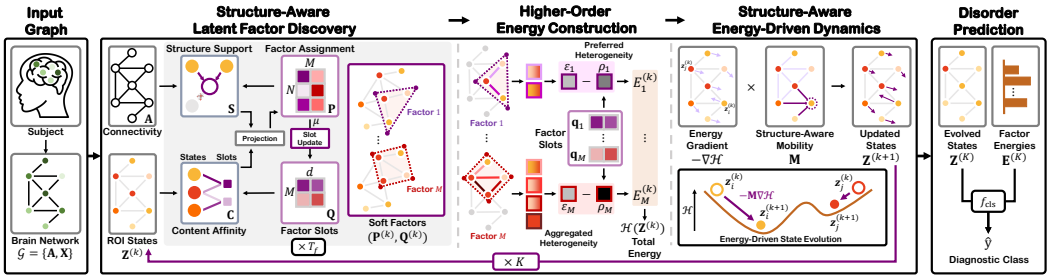


Figure 2: Overview of HOPE. Given a brain network $\mathcal{G}=\{\mathbf{A},\mathbf{X}\}$, a content $\mathbf{C}$ and a structure $\mathbf{S}$ jointly determine the sparse assignment $\mathbf{P}$ of ROIs to the factor slots $\mathbf{Q}$. The discovered factors parameterize a total energy $\mathcal{H}(\mathbf{Z};\mathbf{P},\mathbf{Q})$, and the states descend it under the mobility $\mathbf{M}$. Over K steps, the factors and the states are alternately updated, and the evolved states $\mathbf{Z}^{(K)}$ with the factor energies $\mathbf{E}^{(K)}$ predict the diagnostic class.

where $\langle\cdot,\cdot\rangle$ denotes the Frobenius inner product. Therefore, the energy is non-increasing along the gradient-flow trajectory, providing a principled mechanism for stable state evolution. Representative graph gradient-flow formulations, e.g., GRAFF (Di Giovanni et al., 2023), construct energies from node-wise and pairwise interactions over a fixed graph.

4 Energy-Driven Higher-Order Graph State Evolution

Overview. We introduce HOPE, a framework that (i) discovers instance-specific *latent higher-order factors* coupling multiple nodes, (ii) converts their group-wise interactions into factor-wise energies, and (iii) evolves the states through structure-aware dissipative dynamics driven by the resulting energy. As shown in Fig. 2, from a graph $\mathcal{G}$, initial states $\mathbf{Z}^{(0)}\in\mathbb{R}^{N\times d}$ are encoded from the node features. At each step $k=0,\dots,K-1$, a sparse factor assignment $\mathbf{P}^{(k)}\in\mathbb{R}^{N\times M}$ and M factor slots $\mathbf{Q}^{(k)}\in\mathbb{R}^{M\times d}$ are inferred from the current states $\mathbf{Z}^{(k)}$, which parameterize a total energy $\mathcal{H}^{(k)}=\mathcal{H}(\mathbf{Z}^{(k)};\mathbf{P}^{(k)},\mathbf{Q}^{(k)})\in\mathbb{R}$. Then, the states are updated along the negative energy gradient through the structure-aware mobility $\mathbf{M}$ to obtain $\mathbf{Z}^{(k+1)}$. After K steps, the final factors are inferred from $\mathbf{Z}^{(K)}$, and $\mathbf{Z}^{(K)}$ together with the factor energies $\mathbf{E}^{(K)}$ is used for prediction.

Structure-Aware Latent Factor Discovery. Given a graph $\mathcal{G}=\{\mathbf{A},\mathbf{X}\}$, the node features $\mathbf{X}\in\mathbb{R}^{N\times F}$ (e.g., F imaging measures on N ROIs) are first encoded into initial node states $\mathbf{Z}^{(0)}\in\mathbb{R}^{N\times d}$ using a shallow encoder $f_x:\mathbb{R}^F\to\mathbb{R}^d$, i.e., $\mathbf{Z}^{(0)}=f_x(\mathbf{X})$. To incorporate local graph context, using the normalized adjacency $\tilde{\mathbf{A}}\in\mathbb{R}^{N\times N}$, each state $\mathbf{Z}$ is contextualized as $\mathbf{H}=\mathbf{Z}+\sigma_z(\tilde{\mathbf{A}}\mathbf{Z}\mathbf{W}_z)\in\mathbb{R}^{N\times d}$, where $\mathbf{W}_z\in\mathbb{R}^{d\times d}$ is learnable and σ_z denotes an activation. The residual connection preserves node-specific information while incorporating neighboring states.

To model group-wise interactions without predefined rules, we adopt an iterative slot-based grouping mechanism, inspired by Locatello et al. (2020). For this, we introduce M learnable factor slots $\mathbf{Q}=[\mathbf{q}_1|\dots|\mathbf{q}_M]^\top\in\mathbb{R}^{M\times d}$ and factor assignments $\mathbf{P}=[\mathbf{p}_1|\dots|\mathbf{p}_M]\in\mathbb{R}^{N\times M}$, where each latent higher-order factor is represented by a soft node membership $\mathbf{p}_m$ and a latent group representation $\mathbf{q}_m$. A meaningful factor should group nodes that are compatible not only in their current representations but also in the underlying graph structure. For this, we define a content affinity $\mathbf{C}\in\mathbb{R}^{N\times M}$ and a structural support $\mathbf{S}\in\mathbb{R}^{N\times M}$, whose entries are defined as

$$C_{i,m}=\frac{(\mathbf{h}_i\mathbf{W}_h)(\mathbf{q}_m\mathbf{W}_q)^\top}{\sqrt{d}},\qquad S_{i,m}=\sum_{j=1}^{N}\tilde{A}_{i,j}P_{j,m},\tag{4}$$

where $C_{i,m}$ measures the similarity between the contextualized state $\mathbf{h}_i$ at the i-th node and the m-th factor slot $\mathbf{q}_m$ through learnable projections $\mathbf{W}_h,\mathbf{W}_q\in\mathbb{R}^{d\times d}$, while $S_{i,m}$ measures how strongly the neighbors of node i are associated with factor m based on their previous assignments. At the first iteration, $\mathbf{P}=0$ makes $\mathbf{S}$ vanish, so assignments depend only on $\mathbf{C}$. Subsequently, the structural support encourages connected nodes to share factor memberships.

Selective higher-order participation is encouraged by combining the content affinity and structural support into the assignment scores $\mathbf{O}=\mathbf{C}+\beta\mathbf{S}\in\mathbb{R}^{N\times M}$, where β controls the contribution of the structural term and $\mathbf{o}_i$ denotes the assignment score of node i over the M factors. To avoid dense assignments with excessive factor overlap, the assignment of node i is obtained by projecting the

temperature-scaled scores $\mathbf{o}_i/\tau$ onto the probability simplex Δ^{M-1} as

$$\mathbf{P}_{i,:} = \underset{\mathbf{p}\in\Delta^{M-1}}{\arg\min} \left\| \mathbf{p} - \tfrac{1}{\tau}\mathbf{o}_i \right\|_2^2, \qquad \Delta^{M-1} = \{\mathbf{p} \in \mathbb{R}^M_{\geq 0} : \mathbf{1}^\top \mathbf{p} = 1\}, \tag{5}$$

where temperature $\tau > 0$ controls the sharpness of the assignment. The projection admits the unique closed-form solution $P_{i,m}{=}\max(0, o_{i,m}/\tau - \eta_i)$, where the node-specific threshold η_i is determined by $\sum_m P_{i,m}{=}1$. As a result, sufficiently weak memberships are eliminated, allowing each node to participate in only a subset of factors. Specifically, increasing τ expands the assignment support, whereas decreasing τ yields more selective factor participation.

Proposition 1. *For fixed* $\mathbf{O}$*, let* $\mathbf{P}(\tau)$ *be obtained by projecting each* $\mathbf{o}_i/\tau$ *onto* Δ^{M-1}*. Then, the support of each node assignment expands monotonically with* τ*. As* $\tau \to 0^+$*, the assignment concentrates on the factors with the highest scores, whereas as* $\tau \to \infty$*, it converges to the uniform assignment over all* M *factors.*

Proposition 1 shows that τ controls the granularity of factor participation, ranging from selective assignments to broadly shared memberships across multiple factors.

To construct a representation for each discovered factor, the contextualized states $\mathbf{H}$ are first projected as $\mathbf{V}{=}\mathbf{H}\mathbf{W}_v$, where $\mathbf{W}_v \in \mathbb{R}^{d\times d}$ is learnable. Given the obtained $\mathbf{P}$, each factor aggregates its assigned nodes into the factor representation $\boldsymbol{\mu}_m{=}\sum_i P_{im}\mathbf{v}_i / \sum_i P_{im}$. Using $\boldsymbol{\mu}_m$ together with $\mathbf{q}_m$, the factor slot $\mathbf{q}_m$ is recurrently updated following Chung et al. (2014), and the assignment and slot updates are repeated for T_f refinement iterations.

Higher-Order Energy Construction. Now, the discovered factors are used to construct the total energy that governs the evolution of the node states. Given the current states $\mathbf{Z}$ and factors $(\mathbf{P}, \mathbf{Q})$, each factor induces a factor-wise interaction structure over the observed graph structure through the m-th factor-modulated adjacency defined as

$$\tilde{\mathbf{A}}^{(m)} = \tilde{\mathbf{A}} \odot \left(\mathbf{p}_m \mathbf{p}_m^\top\right), \tag{6}$$

where $\odot$ denotes the Hadamard product, preserving the graph support while reweighting its connections by shared factor membership.

Let $\mathbf{D}^{(m)}$ be the degree matrix of $\tilde{\mathbf{A}}^{(m)}$ and define the factor-guided Laplacian $\mathbf{L}^{(m)}{=}\mathbf{D}^{(m)}{-}\tilde{\mathbf{A}}^{(m)}$. The heterogeneity ε_m within factor m is defined through the corresponding Dirichlet energy as

$$\varepsilon_m = \frac{\mathrm{Tr}\left(\mathbf{V}^\top \mathbf{L}^{(m)} \mathbf{V}\right)}{\mathbf{1}^\top \tilde{\mathbf{A}}^{(m)} \mathbf{1}}, \tag{7}$$

where $\mathbf{1}^\top \tilde{\mathbf{A}}^{(m)} \mathbf{1}$ normalizes ε_m by the total connectivity of factor m. Since regions within a higher-order factor may naturally maintain heterogeneous representations, the factor-level energy E_m is defined relative to a factor-specific preferred heterogeneity ρ_m as

$$E_m = \frac{1}{2}\left(\varepsilon_m - \rho_m\right)^2, \tag{8}$$

where $\rho_m = \sigma(f_q(\mathbf{q}_m))$ denotes the preferred heterogeneity of factor m. Here, $f_q : \mathbb{R}^d \to \mathbb{R}$ maps the factor representation $\mathbf{q}_m$ to this preferred heterogeneity, and σ is the Softplus activation ensuring $\rho_m > 0$. Then, the total energy $\mathcal{H}$ is defined as

$$\mathcal{H}(\mathbf{Z}; \mathbf{P}, \mathbf{Q}) = \frac{1}{M}\sum_{m=1}^{M} E_m. \tag{9}$$

Importantly, although ε_m summarizes pairwise discrepancies within factor m, the nonlinear energy E_m couples these discrepancies jointly, inducing higher-order dependencies among participating nodes. A detailed derivation is provided in Appendix A.

Structure-Aware Energy-Driven Dynamics. At each evolution step k, HOPE computes the energy gradient $\mathbf{G}^{(k)}{=}\nabla_{\mathbf{Z}}\mathcal{H}(\mathbf{Z}^{(k)}; \mathbf{P}^{(k)}, \mathbf{Q}^{(k)})$, which provides the local driving signal for state evolution. To incorporate the input structure into this evolution, we generalize the standard gradient flow in Eq. (2) with a structure-aware mobility operator. Specifically, we parameterize $\mathbf{M}{=}\gamma_0\mathbf{I} + \gamma_1\tilde{\mathbf{L}}$, where $\gamma_0, \gamma_1 > 0$ are learnable coefficients, yielding $d\mathbf{Z}/dt{=}{-}\mathbf{M}\nabla_{\mathbf{Z}}\mathcal{H}$. Here, $\mathbf{I}$ controls the local response, while $\tilde{\mathbf{L}}$ couples the response across nodes according to the input graph structure. This separates the learned higher-order energy from its structure-aware state evolution.

Proposition 2. *Let $\mathcal{H}$ be differentiable in the node states with fixed factors, and let $\mathbf{M} = \gamma_0\mathbf{I}+\gamma_1\tilde{\mathbf{L}}$ with $\gamma_0, \gamma_1 \geq 0$ and $\tilde{\mathbf{L}} \succeq 0$. Then $\mathbf{M} \succeq 0$, and the continuous-time flow $\frac{d\mathbf{Z}}{dt} = -\mathbf{M}\nabla_{\mathbf{Z}}\mathcal{H}$ dissipates the energy as $\frac{d\mathcal{H}}{dt} = -\langle \nabla_{\mathbf{Z}}\mathcal{H}, \mathbf{M}\nabla_{\mathbf{Z}}\mathcal{H}\rangle \leq 0$.*

Proposition 2 follows from the positive semi-definiteness of the mobility operator and does not depend on the particular parameterization of the learned energy. Thus, with fixed factors, the continuous dynamics are dissipative, and the energy is non-increasing. To implement this continuous flow, we discretize $d\mathbf{Z}/dt$ over K steps using the explicit Euler update as

$$\mathbf{Z}^{(k+1)} = \mathbf{Z}^{(k)} - \Delta t\, \mathbf{M}\mathbf{G}^{(k)} \qquad k = 0, \ldots, K-1, \tag{10}$$

where Δt is a fixed step size. Under standard smoothness conditions, a sufficiently small Δt preserves the energy-decreasing behavior of the continuous dynamics. After each state update, the factors are re-inferred from the updated states. Detailed proofs are in Appendix B.

Co-Evolution of States and Factors. Since higher-order interactions depend on the evolving node states, keeping the inferred structure fixed throughout propagation may lead to misalignment. Therefore, HOPE performs K state-evolution steps while re-inferring the factors at every state as $\mathbf{Z}^{(0)} \to (\mathbf{P}^{(0)}, \mathbf{Q}^{(0)}) \to \mathbf{Z}^{(1)} \to \cdots \to (\mathbf{P}^{(K-1)}, \mathbf{Q}^{(K-1)}) \to \mathbf{Z}^{(K)} \to (\mathbf{P}^{(K)}, \mathbf{Q}^{(K)})$, allowing the interaction structure to co-evolve with the latent states rather than remain static. For $k>0$, factor inference is warm-started from $\mathbf{P}^{(k-1)}$ and $\mathbf{Q}^{(k-1)}$, promoting consistent factor identities across the trajectory. During factor inference at k-th step, $\mathbf{Z}^{(k)}$ is treated with stop-gradient, so $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$ are inferred from $\mathbf{Z}^{(k)}$ but treated as fixed when computing the state-update gradient. Additional explanations for the co-evolution process are given in Appendix C.

Training Objective. After K evolutions, $\mathbf{Z}^{(K)}$ represents the final evolved state. Using the factors from the final evolution step, we evaluate the factor energies at $\mathbf{Z}^{(K)}$ as $\mathbf{E}^{(K)}=[E_1^{(K)}, \ldots, E_M^{(K)}]^\top$, which characterize the final factor-wise energy profile. The final states and factor energies are provided to a classifier f_{cls} over C classes (i.e., diagnostic labels) as $\hat{\mathbf{y}}=f_{\text{cls}}(\mathbf{Z}^{(K)}, \mathbf{E}^{(K)})$. The model is trained with the cross-entropy loss and a structural regularizer that encourages the discovered factors to remain coherent with the graph structure throughout the evolution trajectory,

$$\mathcal{L} = -\sum_{c=1}^{C} y_c \log \hat{y}_c + \lambda_{\text{str}} \frac{1}{K+1} \sum_{k=0}^{K} \frac{1}{M} \sum_{m=1}^{M} \Big(1 - \frac{(\mathbf{p}_m^{(k)})^\top \tilde{\mathbf{A}} \mathbf{p}_m^{(k)}}{(\mathbf{p}_m^{(k)})^\top \mathbf{p}_m^{(k)}}\Big), \tag{11}$$

where λ_{str} controls the strength of the structural regularization. The regularizer promotes structurally coherent factor discovery, while the dissipative property of each state update follows from the positive semi-definite mobility in Proposition 2.

5 Experiments

Dataset. We used independent brain network datasets: Alzheimer's Disease Neuroimaging Initiative (ADNI) and Parkinson's Progression Markers Initiative (PPMI) as neurodegenerative benchmarks, alongside Autism Brain Imaging Data Exchange (ABIDE) as a neurodevelopmental benchmark.

For *ADNI*, neuroimaging data from 650 subjects were processed using an in-house pipeline. Each brain was parcellated into 148 cortical and 12 subcortical regions using the Destrieux atlas (Destrieux et al., 2010), and tractography on diffusion-weighted imaging (DWI) was performed to derive white-matter connections among the 160 regions. Region-wise features included Standard Uptake Value Ratio (SUVR) (Thie, 2004) measurements of glucose metabolism (FDG), β-amyloid (AMY), and tau protein (TAU) from PET, together with cortical thickness (CT) from MRI. Subjects were categorized into Control (CN, T=226), Significant Memory Concern (SMC, T=131), Early/Late Mild Cognitive Impairment (EMCI/LMCI, T=217/64), and Alzheimer's disease (AD, T=12) groups.

For the functional brain network benchmarks, we used fMRI data from 181 subjects in *PPMI* (Marek et al., 2011) and 989 subjects in *ABIDE* (Heinsfeld et al., 2018), both of which were preprocessed into brain networks following Xu et al. (2023). Each brain was parcellated into 116 regions using the AAL atlas (Tzourio-Mazoyer et al., 2002), with region-wise Blood-Oxygen-Level-Dependent (BOLD) signals used as node features and pairwise correlations between regional BOLD signals used as edge features. Specifically, subjects from PPMI were divided into Control (CN, T=15), Prodromal (T=53), and Parkinson's disease (PD, T=113) groups, while subjects from ABIDE were

Table 1: In-distribution classification results on the various brain network benchmarks, including ADNI, PPMI, and ABIDE datasets, evaluated using 5-fold cross-validation. The best, second-best, and third-best results are highlighted. (†: Methods designed for brain network analysis, ‡: Physics-inspired methods)

Method	ADNI (CN/SMC/EMCI/LMCI/AD)				PPMI (CN/Prodromal/PD)				ABIDE (CN/ASD)			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
Pair-Wise Graph Method												
GCN	$85.7_{\pm 1.7}$	$88.1_{\pm 3.0}$	$82.7_{\pm 5.9}$	$83.7_{\pm 3.4}$	$66.3_{\pm 2.4}$	$46.2_{\pm 5.4}$	$41.5_{\pm 3.7}$	$38.9_{\pm 4.4}$	$57.5_{\pm 1.8}$	$57.3_{\pm 1.7}$	$56.0_{\pm 2.0}$	$54.3_{\pm 3.3}$
GAT	$89.2_{\pm 2.1}$	$92.7_{\pm 1.8}$	$86.6_{\pm 4.6}$	$88.9_{\pm 3.6}$	$72.9_{\pm 2.5}$	$55.1_{\pm 7.8}$	$52.0_{\pm 3.3}$	$52.3_{\pm 4.3}$	$61.9_{\pm 1.1}$	$61.6_{\pm 1.1}$	$61.4_{\pm 1.3}$	$61.4_{\pm 1.4}$
GCNII	$86.3_{\pm 4.0}$	$91.2_{\pm 1.5}$	$85.8_{\pm 6.6}$	$87.4_{\pm 4.8}$	$68.5_{\pm 2.3}$	$52.5_{\pm 6.0}$	$42.8_{\pm 4.5}$	$41.4_{\pm 6.3}$	$57.1_{\pm 2.3}$	$58.3_{\pm 3.3}$	$55.6_{\pm 3.3}$	$51.6_{\pm 5.7}$
GRAND‡	$90.2_{\pm 2.8}$	$91.8_{\pm 1.6}$	$88.9_{\pm 4.7}$	$89.9_{\pm 3.3}$	$70.7_{\pm 3.9}$	$55.6_{\pm 5.2}$	$46.2_{\pm 6.2}$	$45.9_{\pm 5.0}$	$57.4_{\pm 2.1}$	$57.8_{\pm 2.5}$	$55.1_{\pm 2.9}$	$51.2_{\pm 5.5}$
PDE-GCN‡	$89.6_{\pm 3.2}$	$92.4_{\pm 1.8}$	$88.6_{\pm 4.8}$	$89.9_{\pm 3.8}$	$70.5_{\pm 2.3}$	$60.8_{\pm 5.7}$	$50.8_{\pm 5.6}$	$52.8_{\pm 7.5}$	$60.3_{\pm 1.2}$	$61.0_{\pm 2.1}$	$59.4_{\pm 1.2}$	$58.8_{\pm 1.7}$
GraphCON‡	$90.3_{\pm 2.9}$	$91.8_{\pm 2.7}$	$89.9_{\pm 5.3}$	$90.4_{\pm 4.4}$	$71.8_{\pm 3.9}$	$62.5_{\pm 4.8}$	$50.5_{\pm 5.9}$	$51.0_{\pm 8.8}$	$57.5_{\pm 2.9}$	$58.3_{\pm 2.8}$	$56.7_{\pm 3.9}$	$53.5_{\pm 8.4}$
BrainGNN†	$82.7_{\pm 4.4}$	$85.0_{\pm 4.1}$	$83.1_{\pm 2.4}$	$84.3_{\pm 2.8}$	$68.1_{\pm 2.0}$	$48.8_{\pm 5.3}$	$40.8_{\pm 4.4}$	$38.2_{\pm 5.3}$	$57.9_{\pm 2.0}$	$60.1_{\pm 1.9}$	$56.2_{\pm 2.6}$	$52.2_{\pm 6.0}$
IBGNN†	$81.4_{\pm 2.4}$	$85.8_{\pm 0.7}$	$82.3_{\pm 3.4}$	$83.3_{\pm 2.3}$	$68.1_{\pm 2.4}$	$51.1_{\pm 6.0}$	$42.1_{\pm 5.2}$	$38.7_{\pm 5.4}$	$59.9_{\pm 3.4}$	$59.1_{\pm 3.8}$	$59.0_{\pm 3.7}$	$58.3_{\pm 4.3}$
BrainGB†	$83.5_{\pm 1.6}$	$87.1_{\pm 2.6}$	$80.7_{\pm 3.2}$	$82.2_{\pm 2.3}$	$68.8_{\pm 3.3}$	$48.8_{\pm 6.0}$	$42.8_{\pm 3.2}$	$41.6_{\pm 4.5}$	$57.7_{\pm 3.4}$	$57.9_{\pm 3.9}$	$56.3_{\pm 4.6}$	$53.1_{\pm 7.6}$
BrainNetTF†	$89.4_{\pm 2.5}$	$88.9_{\pm 5.7}$	$84.6_{\pm 7.3}$	$85.8_{\pm 6.0}$	$67.9_{\pm 2.9}$	$55.6_{\pm 8.8}$	$45.7_{\pm 5.1}$	$45.0_{\pm 7.2}$	$60.8_{\pm 2.4}$	$62.3_{\pm 2.7}$	$59.1_{\pm 2.1}$	$57.3_{\pm 3.7}$
GRAFF‡	$90.6_{\pm 4.1}$	$93.2_{\pm 2.8}$	$89.5_{\pm 5.9}$	$90.7_{\pm 4.5}$	$70.2_{\pm 3.4}$	$61.4_{\pm 4.2}$	$53.7_{\pm 5.8}$	$52.9_{\pm 7.9}$	$61.7_{\pm 3.5}$	$61.9_{\pm 3.4}$	$60.3_{\pm 4.7}$	$59.1_{\pm 6.6}$
SGCN†	$84.0_{\pm 2.1}$	$87.6_{\pm 3.3}$	$82.1_{\pm 4.6}$	$83.8_{\pm 3.5}$	$67.6_{\pm 3.2}$	$39.0_{\pm 9.3}$	$40.6_{\pm 4.4}$	$37.6_{\pm 6.8}$	$59.7_{\pm 1.3}$	$61.3_{\pm 2.3}$	$57.7_{\pm 1.9}$	$55.2_{\pm 3.6}$
ALTER†	$90.8_{\pm 2.6}$	$93.3_{\pm 2.4}$	$89.6_{\pm 5.6}$	$90.9_{\pm 4.6}$	$69.6_{\pm 2.3}$	$55.2_{\pm 8.5}$	$55.8_{\pm 5.3}$	$54.5_{\pm 8.3}$	$59.4_{\pm 1.8}$	$59.1_{\pm 2.0}$	$58.6_{\pm 1.7}$	$58.5_{\pm 1.6}$
BioBGT†	$80.3_{\pm 1.7}$	$79.7_{\pm 7.1}$	$77.3_{\pm 6.7}$	$77.2_{\pm 6.1}$	$68.0_{\pm 2.8}$	$54.3_{\pm 8.2}$	$48.7_{\pm 4.2}$	$45.7_{\pm 3.2}$	$62.4_{\pm 1.4}$	$62.1_{\pm 1.4}$	$61.9_{\pm 1.5}$	$61.8_{\pm 1.5}$
BQN†	$81.5_{\pm 4.7}$	$74.3_{\pm 5.6}$	$72.3_{\pm 5.5}$	$72.6_{\pm 5.2}$	$71.9_{\pm 2.9}$	$63.2_{\pm 6.2}$	$51.0_{\pm 8.4}$	$56.4_{\pm 7.3}$	$63.3_{\pm 2.4}$	$63.8_{\pm 2.1}$	$63.4_{\pm 2.0}$	$61.1_{\pm 2.3}$
Higher-Order Graph Method												
HGNN	$83.2_{\pm 2.7}$	$84.8_{\pm 7.1}$	$80.7_{\pm 6.8}$	$82.1_{\pm 6.6}$	$65.2_{\pm 4.8}$	$43.3_{\pm 9.0}$	$41.9_{\pm 6.1}$	$40.3_{\pm 7.1}$	$56.5_{\pm 2.4}$	$56.5_{\pm 2.1}$	$56.4_{\pm 2.1}$	$56.2_{\pm 2.3}$
HNHN	$87.3_{\pm 1.8}$	$89.6_{\pm 1.0}$	$85.6_{\pm 4.2}$	$87.0_{\pm 3.0}$	$69.1_{\pm 2.9}$	$52.6_{\pm 5.1}$	$48.5_{\pm 8.4}$	$48.0_{\pm 9.4}$	$58.7_{\pm 0.8}$	$58.5_{\pm 1.0}$	$58.5_{\pm 1.1}$	$58.4_{\pm 1.1}$
UniGCNII	$89.6_{\pm 1.6}$	$91.1_{\pm 1.8}$	$88.0_{\pm 3.8}$	$89.9_{\pm 2.0}$	$70.8_{\pm 3.3}$	$59.6_{\pm 9.4}$	$54.3_{\pm 7.2}$	$55.2_{\pm 7.3}$	$60.0_{\pm 1.7}$	$60.3_{\pm 2.2}$	$59.0_{\pm 1.8}$	$58.2_{\pm 2.6}$
HyperDrop	$72.3_{\pm 0.8}$	$74.4_{\pm 3.1}$	$65.9_{\pm 0.2}$	$68.6_{\pm 2.2}$	$67.5_{\pm 4.0}$	$41.5_{\pm 8.9}$	$39.2_{\pm 3.8}$	$36.0_{\pm 6.4}$	$52.1_{\pm 4.1}$	$49.0_{\pm 6.7}$	$47.7_{\pm 8.3}$	$46.9_{\pm 6.8}$
dwHGCN†	$90.2_{\pm 1.3}$	$87.2_{\pm 7.1}$	$86.3_{\pm 6.8}$	$86.2_{\pm 6.4}$	$69.0_{\pm 2.9}$	$51.5_{\pm 7.9}$	$44.2_{\pm 6.1}$	$43.3_{\pm 8.0}$	$55.2_{\pm 1.8}$	$59.7_{\pm 9.4}$	$53.1_{\pm 2.4}$	$48.2_{\pm 7.0}$
HyperGT	$81.5_{\pm 1.6}$	$92.1_{\pm 1.8}$	$89.0_{\pm 3.8}$	$89.9_{\pm 2.0}$	$68.5_{\pm 1.5}$	$51.7_{\pm 4.4}$	$42.0_{\pm 4.8}$	$39.6_{\pm 6.1}$	$56.5_{\pm 1.1}$	$60.8_{\pm 5.5}$	$54.8_{\pm 2.5}$	$51.1_{\pm 7.2}$
HyBRiD†	$86.6_{\pm 4.2}$	$87.5_{\pm 3.9}$	$84.3_{\pm 6.4}$	$84.6_{\pm 4.7}$	$67.2_{\pm 6.5}$	$54.8_{\pm 6.8}$	$52.0_{\pm 6.2}$	$52.8_{\pm 6.5}$	$57.3_{\pm 2.6}$	$60.1_{\pm 4.2}$	$55.2_{\pm 3.1}$	$57.5_{\pm 4.5}$
DHHNN	$81.1_{\pm 2.5}$	$89.0_{\pm 4.5}$	$84.5_{\pm 3.2}$	$86.7_{\pm 3.4}$	$60.6_{\pm 3.8}$	$62.3_{\pm 9.8}$	$52.7_{\pm 4.3}$	$54.0_{\pm 6.3}$	$55.2_{\pm 3.5}$	$58.8_{\pm 6.2}$	$53.1_{\pm 4.6}$	$54.6_{\pm 4.1}$
MuHL†	$93.2_{\pm 2.4}$	$95.4_{\pm 1.3}$	$94.2_{\pm 1.6}$	$94.7_{\pm 1.5}$	$76.8_{\pm 3.7}$	$66.6_{\pm 7.6}$	$60.7_{\pm 6.0}$	$62.4_{\pm 6.6}$	$63.7_{\pm 3.6}$	$63.5_{\pm 3.7}$	$63.3_{\pm 3.7}$	$63.3_{\pm 3.7}$
HOPE (Ours)	$96.5_{\pm 2.2}$	$97.5_{\pm 1.6}$	$96.7_{\pm 1.5}$	$97.1_{\pm 1.5}$	$78.2_{\pm 3.4}$	$69.4_{\pm 5.4}$	$63.9_{\pm 4.8}$	$65.7_{\pm 4.9}$	$65.7_{\pm 2.4}$	$65.7_{\pm 2.3}$	$65.0_{\pm 2.9}$	$64.8_{\pm 3.1}$

divided into Control (CN, T=534) and Autism Spectrum Disorder (ASD, T=455) groups. Details of the datasets are explained in Appendix D.

Setup. We evaluated our framework under two main settings: 1) *in-distribution classification* and 2) *zero-shot generalization*. The in-distribution setting performs disease prediction on ADNI using multiple imaging biomarkers, and on PPMI and ABIDE using functional brain networks derived from BOLD signals. In the zero-shot setting, models trained on ADNI-2 are evaluated on unseen ADNI phases (ADNI-1/3/GO), while models trained on PPMI are directly evaluated on two independently collected PD cohorts, TaoWu and Neurocon (Xu et al., 2023). Both settings employ 5-fold cross-validation on the source data, and the mean accuracy (Acc), precision (Pre), recall (Rec), and F1-score (F1s) across folds are reported. Additional details are provided in Appendix E.

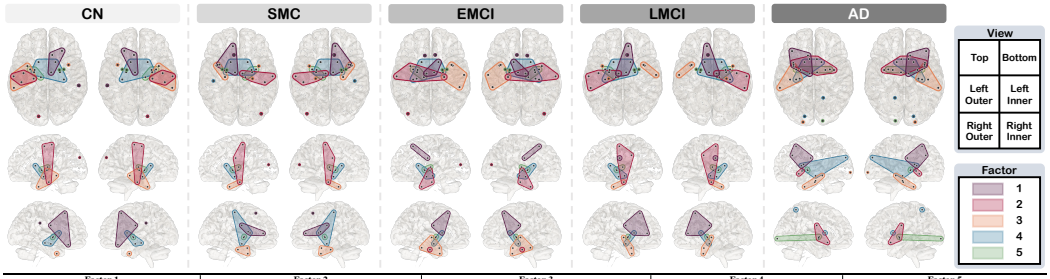
Baselines. To ensure broad comparison, we categorize 24 baselines into pair-wise and higher-order graph methods, including 11 brain-specific and 4 physics-inspired approaches: 1) *Pair-wise graph models*, including GCN (Kipf & Welling, 2016), GAT (Veličković et al., 2018), GCNII (Chen et al., 2020), GRAND (Chamberlain et al., 2021), PDE-GCN (Eliasof et al., 2021), GraphCON (Rusch et al., 2022), BrainGNN (Li et al., 2021), IBGNN (Cui et al., 2022b), BrainGB (Cui et al., 2022a), BrainNetTF (Kan et al., 2022), GRAFF (Di Giovanni et al., 2023), SGCN (Zhou et al., 2024), ALTER (Yu et al., 2024), BioBGT (Peng et al., 2025), and BQN (Yang et al., 2025); and 2) *Higher-order graph models*, including HGNN (Feng et al., 2019), HNHN (Dong et al., 2020), UniGCNII (Huang & Yang, 2021), HyperDrop (Jo et al., 2021), dwHGCN (Wang et al., 2023), HyperGT (Liu et al., 2024), HyBRiD (Qiu et al., 2024), DHHNN (Mei et al., 2025), and MuHL (Sim et al., 2026).

In-Distribution Results. Table 1 reports the in-distribution classification performance of HOPE against 24 baselines on three benchmarks. For example, on ADNI, HOPE achieves the best results on every metric, surpassing the second-best method by 3.3%p in accuracy and 2.4%p in F1-score, with low standard deviations across 5 folds, indicating that the energy-driven evolution yields stable predictions. In addition, on ABIDE, HOPE improves accuracy by 2.0%p and F1-score by 1.5%p, confirming that the framework extends beyond neurodegenerative staging to neurodevelopmental classification. Notably, HOPE surpasses MuHL, the strongest higher-order baseline that learns hyperedge structures but still uses them for aggregation, supporting the benefit of converting discovered higher-order structures into factor-wise energies that directly drive state evolution.

Zero-Shot Generalization. To assess generalizability under heterogeneous distribution shifts, we evaluate whether a model trained on a source phase or dataset can be directly applied to unseen targets without any fine-tuning. Table 2 shows that HOPE achieves strong transfer performance across all three settings, improving upon the second-best method by 2.1%p in accuracy and 1.9%p

Table 2: Zero-shot generalization results between HOPE and brain analysis methods across acquisition phases or datasets. HOPE and baselines are trained on a source phase or dataset and directly evaluated on unseen target phases or datasets without fine-tuning. The best, second-best, and third-best results are highlighted.

Method	ADNI-2→ADNI-1/3/GO				PPMI→TaoWu				PPMI→Neurocon			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
BrainGB	$42.9_{\pm 3.2}$	$48.4_{\pm 8.1}$	$36.5_{\pm 2.4}$	$34.4_{\pm 4.3}$	$53.0_{\pm 1.2}$	$64.1_{\pm 6.8}$	$53.0_{\pm 1.2}$	$43.5_{\pm 4.0}$	$62.9_{\pm 6.4}$	$52.5_{\pm 2.8}$	$51.6_{\pm 7.3}$	$45.1_{\pm 5.3}$
BrainNetTF	$43.5_{\pm 8.4}$	$48.9_{\pm 2.4}$	$50.8_{\pm 2.8}$	$44.8_{\pm 1.6}$	$50.0_{\pm 1.6}$	$46.0_{\pm 3.9}$	$50.0_{\pm 1.6}$	$46.2_{\pm 4.6}$	$62.9_{\pm 4.4}$	$62.7_{\pm 8.1}$	$55.0_{\pm 5.6}$	$50.3_{\pm 8.7}$
ALTER	$43.8_{\pm 3.0}$	$46.8_{\pm 2.0}$	$54.6_{\pm 6.4}$	$44.3_{\pm 2.3}$	$54.5_{\pm 1.9}$	$65.1_{\pm 6.0}$	$52.5_{\pm 1.9}$	$45.3_{\pm 4.4}$	$63.8_{\pm 3.2}$	$63.8_{\pm 7.6}$	$55.3_{\pm 4.4}$	$51.7_{\pm 8.0}$
BQN	$52.3_{\pm 2.9}$	$39.5_{\pm 6.3}$	$32.8_{\pm 7.7}$	$39.0_{\pm 7.6}$	$57.0_{\pm 4.0}$	$56.3_{\pm 4.6}$	$55.0_{\pm 4.0}$	$49.5_{\pm 3.0}$	$55.1_{\pm 3.7}$	$47.8_{\pm 3.2}$	$48.3_{\pm 2.9}$	$46.8_{\pm 2.9}$
dwHGCN	$52.8_{\pm 2.9}$	$48.0_{\pm 2.7}$	$57.0_{\pm 2.8}$	$48.8_{\pm 3.1}$	$50.5_{\pm 2.9}$	$45.1_{\pm 5.2}$	$50.5_{\pm 2.9}$	$36.0_{\pm 4.5}$	$60.9_{\pm 2.7}$	$59.3_{\pm 3.4}$	$53.6_{\pm 3.7}$	$46.1_{\pm 7.0}$
HyBRiD	$47.6_{\pm 5.1}$	$46.7_{\pm 4.6}$	$53.3_{\pm 2.9}$	$44.0_{\pm 5.2}$	$56.2_{\pm 3.1}$	$53.8_{\pm 3.5}$	$49.8_{\pm 4.7}$	$44.7_{\pm 3.8}$	$63.4_{\pm 2.4}$	$51.7_{\pm 2.5}$	$50.0_{\pm 4.2}$	$48.8_{\pm 3.6}$
MuHL	$56.0_{\pm 1.6}$	$52.1_{\pm 6.1}$	$58.4_{\pm 8.3}$	$54.8_{\pm 7.7}$	$60.5_{\pm 1.6}$	$72.8_{\pm 2.4}$	$60.5_{\pm 1.6}$	$58.9_{\pm 3.0}$	$65.9_{\pm 3.4}$	$68.1_{\pm 3.3}$	$57.3_{\pm 3.2}$	$53.7_{\pm 3.6}$
HOPE (Ours)	$58.1_{\pm 2.1}$	$54.2_{\pm 3.4}$	$60.4_{\pm 5.2}$	$56.7_{\pm 6.8}$	$63.5_{\pm 2.4}$	$71.7_{\pm 2.5}$	$63.1_{\pm 1.9}$	$62.9_{\pm 4.1}$	$68.0_{\pm 4.1}$	$70.1_{\pm 3.8}$	$59.4_{\pm 4.0}$	$58.9_{\pm 3.9}$



Factor 1						Factor 2						Factor 3						Factor 4						Factor 5					
ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	SMC	EMCI	LMCI	AD
R.Caudate	0.291	0.262	0.236	0.267	–	L.G.oc.temp.med.Parahip	0.163	0.161	0.173	0.161	–	L.G.oc.temp.med.Parahip	0.155	0.139	–	0.123	0.191	L.Caudate	0.126	0.120	0.117	0.124	0.092	L.Pallidum	0.213	0.185	0.243	0.193	0.151
R.S.pericallosal	0.285	0.273	0.232	0.267	–	L.G.temp.sup.Lateral	0.162	0.172	0.185	0.162	–	L.Pole.temporal	0.150	0.141	–	0.125	0.191	R.Caudate	0.120	0.124	0.107	0.110	0.114	R.Pallidum	0.209	0.181	0.230	0.187	0.129
R.G.front.sup	0.280	0.266	0.239	0.256	–	L.G.precentral	0.155	–	–	0.164	–	R.Pole.temporal	0.141	0.156	0.176	0.137	–	R.Amygdala	0.103	0.109	0.099	0.084	–	R.Putamen	0.192	0.170	0.218	0.167	0.128
L.S.pericallosal	–	–	0.230	0.262	0.188	L.G.occipital.middle	–	0.158	0.158	–	–	R.G.temp.sup.Plan.polar	–	0.138	0.155	0.122	–	L.Amygdala	0.094	0.105	0.103	0.093	–	L.Putamen	0.189	0.168	0.219	0.185	0.145
R.G.S.cingul.Ant	–	0.267	–	0.247	–	L.S.pericallosal	–	0.155	–	0.158	–	R.G.oc.temp.med.Parahip	–	0.146	0.173	–	–	L.Hippocampus	–	–	0.101	0.087	0.086	L.Thalamus.Proper	0.169	0.155	0.194	0.178	–

Figure 3: Top: Visualization of five selected higher-order factors across the ADNI diagnostic stages. Each factor represents a group of representative ROIs. Bottom: Corresponding ROI labels and membership strengths across disease stages. The values show each ROI's association with its factor from CN to AD. The five ROIs are selected based on their overall factor participation across stages.

in F1-score on the cross-phase AD evaluation. On TaoWu, HOPE further improves recall by 2.6%p while achieving competitive precision of 71.7%. Since the factors are re-inferred for every subject at test time, the higher-order factors can adapt to unseen subjects and cohorts, while the same structure-aware energy-driven dynamics are applied without target-specific adaptation. These results demonstrate that HOPE learns transferable disease-related representations that generalize across both acquisition phases and independently collected cohorts.

6 Analyses and Discussions

Higher-Order Structure Analysis. To examine the structure captured by the learned higher-order factors, we analyze five representative factors across diagnostic stages from ADNI, as shown in Fig. 3. Since factor indices are arbitrary across independently trained folds, we first align their ordering based on ROI assignment similarity before aggregating the held-out predictions. Then, we select the five most prominent factors according to their overall membership strengths and characterize their ROI composition using the corresponding profiles. For each selected factor, we compute class-wise ROI membership profiles and select five representative ROIs based on their overall participation across diagnostic stages. Their stage-wise membership strengths are compared to characterize changes in regional participation. The top of Fig. 3 visualizes the resulting spatial organization across stages, while the bottom reports the representative ROIs and their strengths.

In Fig. 3, the learned factors exhibit both persistent anatomical organization and stage-dependent reorganization. For example, Factor 1 (F1), with the largest overall participation mass, is dominated by frontal and pericallosal regions across earlier stages, while its representative ROIs shift toward left pericallosal and frontal regions in AD. F5 consistently captures a subcortical organization involving the pallidum, putamen, and thalamus, with the left pallidum remaining strongly represented from CN (0.213) to AD (0.151). In contrast, F2 shows stronger stage-dependent reorganization, as temporal and parahippocampal ROIs prominent from CN to LMCI no longer appear among the representative ROIs in AD. These patterns suggest that the learned factors capture both anatomically coherent organization and stage-associated changes in regional participation.

Disease-Associated Regional Energy Alterations. To examine how the learned higher-order dynamics differ between disease and control groups, we compare regional factor-energy scores across

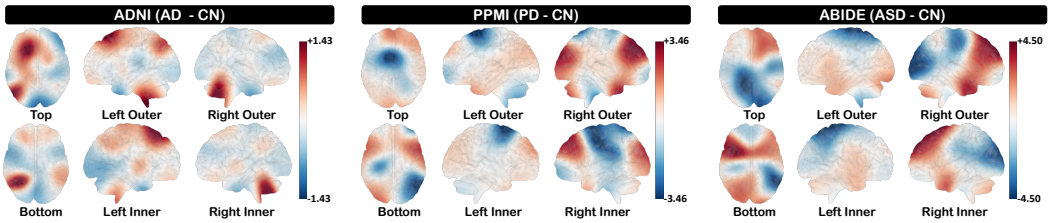


Figure 4: Visualization of higher-order energy differences between disease and control groups across all datasets and six anatomical views. For each ROI, the factor-energy score combines factor membership with the corresponding energy mismatch from preferred heterogeneity. Red indicates a higher score in the disease group, whereas blue indicates a lower score. Colorbar values are scaled by 10^{-2}.

all three datasets. As shown in Fig. 4, for each ROI i, we compute the membership-weighted factor-energy score $\sum_m P_{i,m} E_m$, where E_m measures deviation from the preferred heterogeneity of factor m. The score reflects how strongly each ROI is associated with factors whose heterogeneity deviates from their preferred states. Red regions indicate higher scores in the disease group, whereas blue regions indicate lower scores relative to controls. The resulting maps reveal spatially heterogeneous patterns that differ across all datasets rather than a uniform regional alteration. These findings suggest that disease effects are expressed through region-specific changes in participation in higher-order factors that deviate from their learned preferred states.

Effect of K. To investigate the effect of the number of evolution steps K, we evaluate HOPE on all three brain network benchmarks. In Fig. 5, performance improves rapidly with additional evolution steps, reaching its best performance at K=3 for ADNI and PPMI and at K=4 for ABIDE. Beyond these points, the results remain largely stable with only minor fluctuations, indicating that most of the benefit from iterative factor-state refinement is obtained within a relatively small number of updates. Although the optimal K varies slightly across datasets, the overall trend is consistent. These results show that HOPE efficiently refines both the latent states and their higher-order interaction structure without requiring prolonged evolution or excessive computational depth.

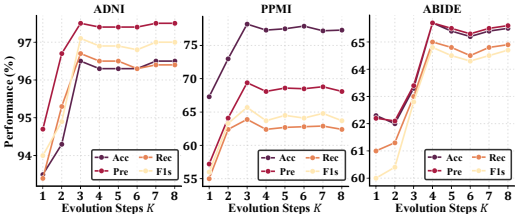


Figure 5: Effect of the number of evolution steps K.

Ablation Studies. To assess each component in HOPE, we conduct ablation studies on ADNI and ABIDE by removing the structural support $\mathbf{S}$, higher-order energy $\mathcal{H}$, connectome constraint $\tilde{\mathbf{L}}$ in $\mathbf{M}$, and factor-state co-evolution. In Table 3, without $\mathcal{H}$, we use conventional Dirichlet energy, while without co-evolution, the initially inferred factors are fixed throughout state evolution. Removing any component consistently degrades performance, demonstrating their complementary roles in learning structure-aware higher-order representations. Notably, replacing the factor-wise energy with pairwise energy causes the largest accuracy drop of 2.6%p on ADNI and 4.0%p on ABIDE, highlighting the importance of group-wise interactions for energy-driven state evolution. Overall, these results show that factor discovery and structure-aware dynamics are most effective when jointly integrated.

Table 3: Ablation study on key components of HOPE.

Method	ADNI			ABIDE		
	Acc ↑	Pre ↑	Rec ↑	Acc ↑	Pre ↑	Rec ↑
w/o $\mathbf{S}$	$94.6_{\pm 2.0}$	$95.8_{\pm 1.4}$	$94.7_{\pm 1.3}$	$64.9_{\pm 2.4}$	$64.8_{\pm 2.5}$	$64.2_{\pm 2.5}$
w/o $\mathcal{H}$	$93.9_{\pm 1.9}$	$95.0_{\pm 1.2}$	$94.1_{\pm 1.3}$	$61.7_{\pm 1.9}$	$61.8_{\pm 2.0}$	$61.3_{\pm 2.1}$
w/o $\tilde{\mathbf{L}}$ in $\mathbf{M}$	$94.5_{\pm 2.3}$	$95.7_{\pm 1.6}$	$94.8_{\pm 1.6}$	$64.6_{\pm 3.1}$	$64.6_{\pm 3.3}$	$64.0_{\pm 2.8}$
w/o co-evolution	$94.8_{\pm 2.0}$	$95.9_{\pm 1.4}$	$95.0_{\pm 1.4}$	$64.7_{\pm 3.5}$	$64.9_{\pm 3.5}$	$64.4_{\pm 3.1}$
HOPE (Full)	$96.5_{\pm 2.2}$	$97.5_{\pm 1.6}$	$96.7_{\pm 1.5}$	$65.7_{\pm 2.4}$	$65.7_{\pm 2.3}$	$65.0_{\pm 2.9}$

7 Conclusion

In this work, we proposed HOPE, an energy-driven framework for higher-order interactions on graphs. HOPE discovers instance-specific latent factors, converts their group-wise interactions into factor-wise energies, and alternates factor inference with structure-aware dissipative state evolution. Experiments on diverse brain networks show that HOPE improves disease classification across neurodegenerative and neurodevelopmental disorders, while generalizing to unseen acquisition phases and independent cohorts without fine-tuning. Furthermore, the learned factors and energies reveal disease-associated patterns of regional organization, demonstrating the potential of physics-inspired higher-order learning as a principled and interpretable tool for heterogeneous brain disorders.

AI USE STATEMENT

Generative AI tools were used to assist with language editing and literature search. All technical content, including the proposed method, experiments, analyses, and theoretical results, was developed and verified by the authors. Any AI-assisted text and suggested references were reviewed for accuracy and relevance. The authors take responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

We provide detailed information to support reproducibility of our results. The main paper describes the proposed framework, training objective, and evaluation protocols. Appendices A and B provide the theoretical derivation and proofs underlying the proposed framework. Appendix C describes the factor-state co-evolution procedure, while Appendix D provides detailed dataset and preprocessing information. Appendix E provides the training configurations, hyperparameter settings, model architectures, and baseline implementation details. Additional experimental results and analyses are provided in Appendices F and G.

REFERENCES

Federico Battiston, Giulia Cencetti, Iacopo Iacopini, et al. Networks beyond pairwise interactions: Structure and dynamics. *Physics reports*, 2020.

Ben Chamberlain, James Rowbottom, Maria I Gorinova, et al. Grand: Graph neural diffusion. In *International conference on machine learning*, 2021.

Ming Chen, Zhewei Wei, Zengfeng Huang, et al. Simple and deep graph convolutional networks. In *International conference on machine learning*. PMLR, 2020.

Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *Advances in neural information processing systems*, 2014.

Hejie Cui, Wei Dai, Yanqiao Zhu, et al. Braingb: a benchmark for brain network analysis with graph neural networks. *IEEE Transactions on Medical Imaging*, 2022a.

Hejie Cui, Wei Dai, Yanqiao Zhu, et al. Interpretable graph neural networks for connectome-based brain disorder analysis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022b.

Fabio Della Rossa, Davide Liuzza, Francesco Lo Iudice, and Pietro De Lellis. Emergence and control of synchronization in networks with directed many-body interactions. *Physical Review Letters*, 2023.

Zhongying Deng, Haoyu Wang, Ziyan Huang, et al. Brain foundation models with hypergraph dynamic adapter for brain disease analysis. *Pattern Recognition*, 2025.

Christophe Destrieux, Bruce Fischl, Anders Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 2010.

Francesco Di Giovanni, James Rowbottom, Benjamin P Chamberlain, et al. Understanding convolution on graphs via energies. *Transactions on Machine Learning Research*, 2023.

Yihe Dong, Will Sawin, and Yoshua Bengio. Hnhn: Hypergraph networks with hyperedge neurons. *ICML Graph Representation Learning and Beyond Workshop*, 2020.

Serdar Elhatisari, Lukas Bovermann, Yuan-Zhuo Ma, et al. Wavefunction matching for solving quantum many-body problems. *Nature*, 2024.

Moshe Eliasof, Eldad Haber, and Eran Treister. Pde-gcn: Novel architectures for graph neural networks motivated by partial differential equations. *Advances in neural information processing systems*, 2021.

Song Feng, Emily Heath, Brett Jefferson, et al. Hypergraph models of biological networks to identify genes critical to pathogenic viral response. *BMC bioinformatics*, 2021.

Yifan Feng, Haoxuan You, Zizhao Zhang, et al. Hypergraph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, 2019.

Lindsay K Festa, Judith B Grinspan, and Kelly L Jordan-Sciutto. White matter injury across neurodegenerative disease. *Trends in neurosciences*, 2024.

Xiangmin Han, Rundong Xue, Jingxi Feng, et al. Hypergraph foundation model for brain disease diagnosis. *IEEE Transactions on Neural Networks and Learning Systems*, 2025.

Anibal Sólon Heinsfeld, Alexandre Rosa Franco, R Cameron Craddock, et al. Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: clinical*, 2018.

Lun Hu, Jun Zhang, Xiangyu Pan, et al. Hiscf: leveraging higher-order structures for clustering analysis in biological networks. *Bioinformatics*, 2021.

Jing Huang and Jie Yang. Unignn: a unified framework for graph and hypergraph neural networks. *International Joint Conference on Artificial Intelligence*, 2021.

Jaehyeong Jo, Jinheon Baek, Seul Lee, et al. Edge representation learning with hypergraphs. *Advances in Neural Information Processing Systems*, 2021.

Xuan Kan, Wei Dai, Hejie Cui, et al. Brain network transformer. *Advances in Neural Information Processing Systems*, 2022.

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

Jacqueline K Kueper, Mark Speechley, and Manuel Montero-Odasso. The alzheimer's disease assessment scale–cognitive subscale (adas-cog): modifications and responsiveness in pre-dementia populations. a narrative review. *Journal of Alzheimer's Disease*, 2018.

Pierre Legendre. Species associations: the kendall coefficient of concordance revisited. *Journal of agricultural, biological, and environmental statistics*, 2005.

Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, et al. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 2021.

Zexi Liu, Bohan Tang, Ziyuan Ye, et al. Hypergraph transformer for semi-supervised classification. In *International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2024.

Francesco Locatello, Dirk Weissenborn, Thomas Unterthiner, et al. Object-centric learning with slot attention. *Advances in neural information processing systems*, 2020.

Kenneth Marek, Danna Jennings, Shirley Lasch, et al. The parkinson progression marker initiative (ppmi). *Progress in neurobiology*, 2011.

Andre Martins and Ramon Astudillo. From softmax to sparsemax: A sparse model of attention and multi-label classification. In *International conference on machine learning*, 2016.

Paul M Matthews, Nicola Filippini, and Gwenaëlle Douaud. Brain structural and functional connectivity and the progression of neuropathology in alzheimer's disease. *Journal of Alzheimer's disease*, 2012.

Zhangyu Mei, Xiao Bi, Dianguo Li, et al. Dhhnn: a dynamic hypergraph hyperbolic neural network based on variational autoencoder for multimodal data integration and node classification. *Information Fusion*, 2025.

Ana P Millán, Hanlin Sun, Lorenzo Giambagli, et al. Topology shapes dynamics of higher-order networks. *Nature Physics*, 2025.

Richard C Mohs and L Cohen. Alzheimer's disease assessment scale (adas). *Psychopharmacol Bull*, 1988.

Susanne G Mueller, Michael W Weiner, Leon J Thal, et al. Ways toward an early diagnosis in alzheimer's disease: the alzheimer's disease neuroimaging initiative (adni). *Alzheimer's & Dementia*, 2005.

Suphakit Niwattanakul, Jatsada Singthongchai, Ekkachai Naenudorn, and Supachanun Wanapu. Using of jaccard coefficient for keywords similarity. In *Proceedings of the international multiconference of engineers and computer scientists*, 2013.

Ciyuan Peng, Yuelong Huang, Qichao Dong, et al. Biologically plausible brain graph transformer. *International Conference on Learning Representations*, 2025.

Weikang Qiu, Huangrui Chu, Selena Wang, et al. Learning high-order relationships of brain regions. In *International conference on machine learning*, 2024.

Andrew T Reid and Alan C Evans. Structural networks in alzheimer's disease. *European neuropsychopharmacology*, 2013.

T Konstantin Rusch, Ben Chamberlain, James Rowbottom, et al. Graph-coupled oscillator networks. In *International conference on machine learning*, 2022.

Pranab Kumar Sen. Estimates of the regression coefficient based on kendall's tau. *Journal of the American statistical association*, 1968.

Jaeyoon Sim, Soojin Hwang, Seunghun Baek, et al. Learning multi-scale hypergraph for high-order brain connectivity analysis. *arXiv preprint arXiv:2606.03310*, 2026.

Joseph A Thie. Understanding the standardized uptake value, its methods, and implications for usage. *Journal of Nuclear Medicine*, 2004.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996.

Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, et al. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 2002.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, et al. Graph attention networks. *International Conference on Learning Representations*, 2018.

Junqi Wang, Hailong Li, Gang Qu, et al. Dynamic weighted hypergraph convolutional network for brain functional connectome analysis. *Medical image analysis*, 2023.

Jiaxing Xu, Yunhan Yang, David Huang, et al. Data-driven network neuroscience: On data collection and benchmark. *Advances in Neural Information Processing Systems*, 2023.

Kai Yang, Zhengqi Wu, Jie Long, et al. White matter changes in parkinson's disease. *npj Parkinson's Disease*, 2023.

Liang Yang, Yuwei Liu, Jiaming Zhuo, et al. Do we really need message passing in brain network modeling? In *International Conference on Machine Learning*, 2025.

Shuo Yu, Shan Jin, Ming Li, et al. Long-range brain graph transformer. *Advances in Neural Information Processing Systems*, 2024.

Houliang Zhou, Lifang He, Brian Y Chen, et al. Multi-modal diagnosis of alzheimer's disease using interpretable graph convolutional networks. *IEEE Transactions on Medical Imaging*, 2024.

Kaixiong Zhou, Xiao Huang, Daochen Zha, et al. Dirichlet energy constrained learning for deep graph neural networks. *Advances in neural information processing systems*, 2021.

Jianming Zhu, Junlei Zhu, Smita Ghosh, et al. Social influence maximization in hypergraph in social networks. *IEEE Transactions on Network Science and Engineering*, 2018.

Appendix Overview

- **A. Theoretical Analysis**
  Derives the factor-wise energy, characterizes its nonseparable higher-order couplings, and visualizes cross-interactions between distinct pairwise connections.
- **B. Proofs of Theoretical Propositions**
  Presents proofs of the theoretical guarantees, including temperature-controlled sparse factor assignments and continuous/discrete energy dissipation.
- **C. Factor-State Co-Evolution Details**
  Describes the alternating factor inference and ROI-state evolution, including stop-gradient, fixed-factor updates, warm-starting, and initialization.
- **D. Dataset Information**
  Provides detailed descriptions of the neuroimaging datasets, preprocessing procedures, and cohort statistics used in our experiments.
- **E. Implementation Details**
  Summarizes training configurations, model architectures, hyperparameter settings, and baseline implementation details.
- **F. Additional Experimental Results**
  Reports additional evaluations on modality reduction, clinical symptom regression, and robustness to random initialization.
- **G. Additional Analysis**
  Analyzes energy dissipation, hyperparameter sensitivity, structural regularization, higher-order factor interpretation, factor-state co-evolution, and computational efficiency.
- **H. Discussions on Our Framework**
  Discusses broader implications, its current limitations, and future extensions and applications of our framework.

A Theoretical Analysis of Higher-Order Energy Coupling

A.1 From Pairwise Dirichlet Energy to Factor-Level Energy

We further clarify how the factor-level energy in Eqs. (6),(7), and (8) differs from a conventional pairwise Dirichlet energy and gives rise to nonseparable interactions within each factor.

For an undirected weighted graph with adjacency $\mathbf{A}$ and Laplacian $\mathbf{L}{=}\mathbf{D}-\mathbf{A}$, the Dirichlet energy of node states $\mathbf{V}{=}[\mathbf{v}_1,\dots,\mathbf{v}_N]^\top$ can be written as

$$\mathrm{Tr}\left(\mathbf{V}^\top\mathbf{L}\mathbf{V}\right)=\frac{1}{2}\sum_{i,j}A_{i,j}\left\|\mathbf{v}_i-\mathbf{v}_j\right\|_2^2, \tag{12}$$

which aggregates pairwise discrepancies independently over the graph. In HOPE, Eq. (6) defines a different interaction structure for each factor m. Its factor-wise edge weight is defined as

$$w_{i,j}^{(m)}:=\tilde{A}_{i,j}^{(m)}=\tilde{A}_{i,j}P_{i,m}P_{j,m}. \tag{13}$$

Hence, an observed connection contributes strongly to factor m only when both incident nodes have high membership in that factor. Now, let

$$Z_m=\mathbf{1}^\top\tilde{\mathbf{A}}^{(m)}\mathbf{1}=\sum_{i,j}w_{i,j}^{(m)} \tag{14}$$

denote the total factor-wise connectivity. Using Eq. (12), the heterogeneity in Eq. (7) can then be expanded as

$$\varepsilon_m=\frac{1}{2Z_m}\sum_{i,j}w_{i,j}^{(m)}\left\|\mathbf{v}_i-\mathbf{v}_j\right\|_2^2=\sum_{i,j}c_{i,j}^{(m)}, \tag{15}$$

where

$$c_{i,j}^{(m)}=\frac{\tilde{A}_{i,j}P_{i,m}P_{j,m}}{2Z_m}\left\|\mathbf{v}_i-\mathbf{v}_j\right\|_2^2 \tag{16}$$

is the contribution of interaction (i,j) to factor m. Therefore, the same pairwise discrepancy can contribute differently across factors according to the joint membership $P_{i,m}P_{j,m}$.

Importantly, the higher-order coupling does not arise from Eq. (7) alone. Substituting Eq. (15) into the nonlinear factor potential in Eq. (8) gives

$$E_m=\frac{1}{2}\left(\sum_{i,j}c_{i,j}^{(m)}-\rho_m\right)^2. \tag{17}$$

For compact notation, let e index the ordered interaction terms (i,j) appearing in Eq. (15). From Eq. (17), expanding this expression yields

$$E_m=\frac{1}{2}\sum_{e}\left(c_e^{(m)}\right)^2+\sum_{e<e'}c_e^{(m)}c_{e'}^{(m)}-\rho_m\sum_{e}c_e^{(m)}+\frac{1}{2}\rho_m^2, \tag{18}$$

where each interaction index e corresponds to a node pair (i,j) with contribution $c_{i,j}^{(m)}$, and e' denotes another distinct interaction within the same factor. Here, $c_e^{(m)}c_{e'}^{(m)}$ couples distinct pairwise discrepancies through the same factor-level energy, including interactions without a shared endpoint. While Eq. (7) summarizes factor-wise pairwise discrepancies, the nonlinear potential in Eq. (8) makes the energy nonseparable across these interactions. Thus, the higher-order behavior of HOPE arises from jointly coupling multiple pairwise interactions through a shared latent factor.

A.2 Physical Interpretation of Factor-Mediated Cross-Interactions

To examine how different energy constructions affect state dynamics, we compare HOPE with GRAFF (Di Giovanni et al., 2023), a representative graph gradient-flow formulation, using the same ADNI subject. GRAFF decomposes its pairwise energy over individual graph edges, whereas HOPE aggregates multiple interactions within a latent factor before applying a nonlinear factor potential.

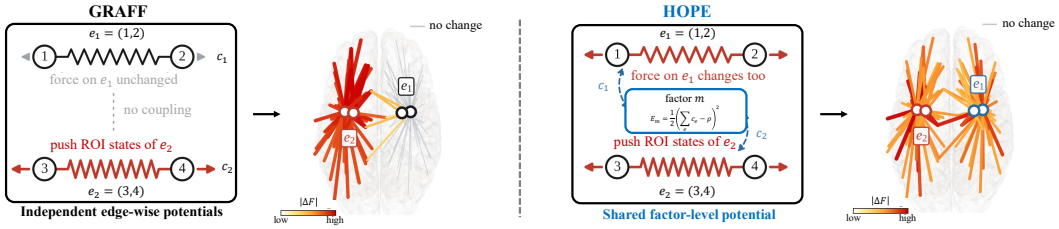


Figure 6: Perturbation-based comparison of pairwise and factor-level energy dynamics. For the same subject, we perturb e_2 and measure ΔF, the resulting change in state-evolution force across ROIs. The interactions e_1 and e_2 are disjoint and not connected. GRAFF yields localized changes with no response at e_1, whereas HOPE propagates the perturbation through shared factor-level energy, producing nonzero responses across ROIs.

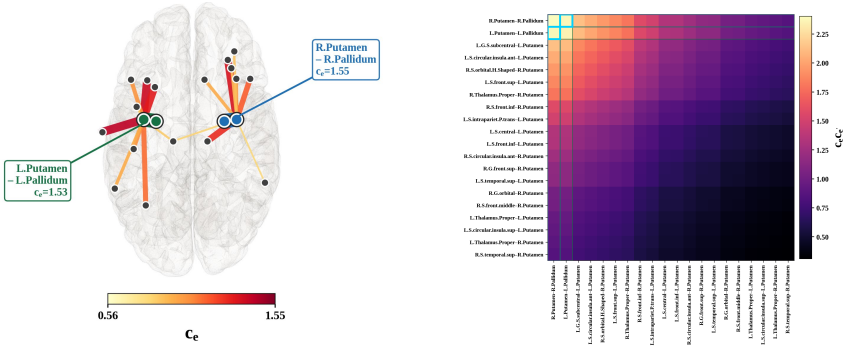


Figure 7: Visualization of higher-order cross-interactions for Factor 14 of subject 116_S_4043. Left: Factor interactions, with edge intensity and thickness indicating c_e. Two strong interactions with disjoint endpoints are highlighted. Right: The cross-interaction matrix $c_e c_{e'}$, whose off-diagonal entries represent coupling between distinct interactions. The highlighted cell corresponds to the two selected disjoint interactions.

This structural difference allows HOPE to introduce cross-interactions among otherwise disjoint connections through a shared factor-level energy.

We select two interactions e_1 and e_2 with disjoint endpoints and no direct connection, perturb the ROI states associated with e_2, and measure the resulting change in the force used for state evolution. In Fig. 6, the perturbation in GRAFF leaves the pairwise force at e_1 unchanged, whereas HOPE changes the force at e_1 through the shared factor-level energy. This result provides empirical evidence of the factor-mediated coupling characterized by Eq. (18), showing that a local perturbation can directly influence the evolution of a disjoint interaction through the shared factor energy.

We further verify that such cross-interactions emerge in the learned factors of HOPE rather than only in the analytical formulation. Fig. 7 visualizes the dominant interactions in Factor 14 for subject 116_S_4043. The highlighted L.Putamen-L.Pallidum and R.Putamen-R.Pallidum interactions share no endpoint but exhibit a nonzero cross-interaction, while the edge-edge interaction matrix shows that similar coupling extends across multiple connections within the same factor. This demonstrates that the learned factors can coordinate anatomically disjoint interactions through a shared energy, enabling HOPE to capture distributed network organization beyond local edge-wise dependencies.

B Proofs of Propositions

Proposition 1. *For fixed* $\mathbf{O}$*, let* $\mathbf{P}(\tau)$ *be obtained by projecting each* $\mathbf{o}_i/\tau$ *onto* Δ^{M-1}*. Then, the support of each node assignment expands monotonically with* τ*. As* $\tau \to 0^+$*, the assignment concentrates on the factors with the highest scores, whereas as* $\tau \to \infty$*, it converges to the uniform assignment over all* M *factors.*

Proof. Fix a row i and omit the index for simplicity. Let the logits be sorted in descending order as $s_{(1)} \geq \cdots \geq s_{(M)}$. To characterize how the active support changes with the temperature, we define

$$D(k) = \sum_{j \leq k} \left(s_{(j)} - s_{(k)} \right) \geq 0, \qquad D(1) = 0. \tag{19}$$

The quantity $D(k)$ measures the cumulative gap between the k largest logits and the k-th largest logit. Its value is non-decreasing with k as

$$D(k{+}1) - D(k) = k\left(s_{(k)} - s_{(k+1)}\right) \geq 0. \tag{20}$$

For simplex projection, the size of the active support is characterized by the threshold condition of sparsemax (Martins & Astudillo, 2016). Applied to the tempered logits, the support size is

$$k^\star = \max\left\{k : 1 + k\frac{s_{(k)}}{\tau} > \sum_{j\leq k}\frac{s_{(j)}}{\tau}\right\}. \tag{21}$$

Multiplying the inequality by $\tau > 0$ gives

$$k^\star(\tau) = \max\left\{k : \tau > D(k)\right\}. \tag{22}$$

Because $D(k)$ is non-decreasing, increasing τ can only make additional values of k satisfy this condition. The support size $k^\star(\tau)$ is therefore non-decreasing with τ.

For the large-temperature limit, $\tau > D(M)$ eventually holds as $\tau \to \infty$, so all M coordinates become active. At the same time, the differences among the tempered logits vanish, and their simplex projection approaches the barycenter of the simplex as

$$\mathbf{p}(\tau) \to \frac{1}{M}\mathbf{1}. \tag{23}$$

For the small-temperature limit, let r denote the number of factors attaining the maximum logit. We have $D(k) = 0$ for $k \leq r$, whereas $D(k) > 0$ for every $k > r$. As $\tau \to 0^+$, the active support is consequently restricted to the r maximal-logit coordinates. Since these coordinates have identical logits, their projected assignments converge to equal weights, giving a uniform assignment over the argmax set. When the maximum logit is unique, this limit reduces to a one-hot assignment. □

Proposition 2. *Let $\mathcal{H}$ be differentiable in the node states with fixed factors, and let $\mathbf{M} = \gamma_0\mathbf{I}+\gamma_1\tilde{\mathbf{L}}$ with $\gamma_0, \gamma_1 \geq 0$ and $\tilde{\mathbf{L}} \succeq 0$. Then $\mathbf{M} \succeq 0$, and the continuous-time flow $\frac{d\mathbf{Z}}{dt} = -\mathbf{M}\nabla_{\mathbf{Z}}\mathcal{H}$ dissipates the energy as $\frac{d\mathcal{H}}{dt} = -\langle\nabla_{\mathbf{Z}}\mathcal{H}, \mathbf{M}\nabla_{\mathbf{Z}}\mathcal{H}\rangle \leq 0$.*

Proof. We first show that the mobility operator is positive semi-definite. For any $\mathbf{x} \in \mathbb{R}^N$, the quadratic form of $\mathbf{M}$ is given by

$$\mathbf{x}^\top\mathbf{M}\mathbf{x} = \mathbf{x}^\top(\gamma_0\mathbf{I} + \gamma_1\tilde{\mathbf{L}})\mathbf{x} = \gamma_0\mathbf{x}^\top\mathbf{x} + \gamma_1\mathbf{x}^\top\tilde{\mathbf{L}}\mathbf{x}. \tag{24}$$

Because $\gamma_0, \gamma_1 \geq 0$, $\mathbf{x}^\top\mathbf{x} \geq 0$, and the normalized Laplacian satisfies $\tilde{\mathbf{L}} \succeq 0$, both terms are non-negative. Therefore,

$$\mathbf{x}^\top\mathbf{M}\mathbf{x} \geq 0, \tag{25}$$

which proves $\mathbf{M} \succeq 0$.

Continuous-time dissipation. Let $\mathbf{G}{=}\nabla_{\mathbf{Z}}\mathcal{H}$ denote the energy gradient. By the chain rule and the mobility-modulated flow $\frac{d\mathbf{Z}}{dt}{=}{-}\mathbf{M}\mathbf{G}$, the energy variation becomes

$$\frac{d\mathcal{H}}{dt} = \left\langle \mathbf{G}, \frac{d\mathbf{Z}}{dt}\right\rangle = -\left\langle \mathbf{G}, \mathbf{M}\mathbf{G}\right\rangle, \tag{26}$$

where $\langle\cdot,\cdot\rangle$ denotes the Frobenius inner product. Because $\mathbf{M} \succeq 0$, we have $\langle\mathbf{G}, \mathbf{M}\mathbf{G}\rangle \geq 0$, and the energy therefore satisfies $\frac{d\mathcal{H}}{dt} \leq 0$.

Discrete-time dissipation. For the Euler discretization, let $\mathbf{G} = \nabla_{\mathbf{Z}}\mathcal{H}(\mathbf{Z})$. The discrete update is

$$\mathbf{Z}^+ = \mathbf{Z} - \Delta t\,\mathbf{M}\mathbf{G}. \tag{27}$$

Unlike the continuous-time flow, the Euler update introduces a finite-step discretization error. Assume that $\nabla_{\mathbf{Z}}\mathcal{H}$ is ℓ-Lipschitz continuous with respect to the Frobenius norm in a convex region containing the iterates. By the descent lemma,

$$\mathcal{H}(\mathbf{Z}^+) \leq \mathcal{H}(\mathbf{Z}) + \left\langle \mathbf{G}, \mathbf{Z}^+ - \mathbf{Z}\right\rangle_F + \frac{\ell}{2}\left\|\mathbf{Z}^+ - \mathbf{Z}\right\|_F^2. \tag{28}$$

Substituting Eq. (27) gives

$$\mathcal{H}(\mathbf{Z}^+) \le \mathcal{H}(\mathbf{Z}) - \Delta t \langle \mathbf{G}, \mathbf{M}\mathbf{G} \rangle + \frac{\ell \Delta t^2}{2} \langle \mathbf{G}, \mathbf{M}^2 \mathbf{G} \rangle . \tag{29}$$

The first term induces energy descent, whereas the second accounts for the Euler discretization error. Since $\mathbf{M} \succeq 0$,

$$\mathbf{M}^2 \preceq \lambda_{\max}(\mathbf{M}) \mathbf{M}, \tag{30}$$

which implies

$$\langle \mathbf{G}, \mathbf{M}^2 \mathbf{G} \rangle \le \lambda_{\max}(\mathbf{M}) \langle \mathbf{G}, \mathbf{M}\mathbf{G} \rangle . \tag{31}$$

Applying this bound to Eq. (29) yields

$$\mathcal{H}(\mathbf{Z}^+) \le \mathcal{H}(\mathbf{Z}) - \Delta t \left(1 - \frac{\ell \Delta t \lambda_{\max}(\mathbf{M})}{2} \right) \langle \mathbf{G}, \mathbf{M}\mathbf{G} \rangle . \tag{32}$$

Therefore, if

$$0 < \Delta t \le \frac{1}{\ell \lambda_{\max}(\mathbf{M})}, \tag{33}$$

then

$$\mathcal{H}(\mathbf{Z}^+) \le \mathcal{H}(\mathbf{Z}) - \frac{\Delta t}{2} \langle \mathbf{G}, \mathbf{M}\mathbf{G} \rangle \le \mathcal{H}(\mathbf{Z}), \tag{34}$$

which establishes discrete-time energy dissipation. For the normalized Laplacian, $\lambda_{\max}(\tilde{\mathbf{L}}) \le 2$. Since $\mathbf{M} = \gamma_0 \mathbf{I} + \gamma_1 \tilde{\mathbf{L}}$, we have

$$\lambda_{\max}(\mathbf{M}) \le \gamma_0 + 2\gamma_1 . \tag{35}$$

Accordingly, a sufficient step-size condition is

$$\Delta t \le \frac{1}{\ell(\gamma_0 + 2\gamma_1)} . \tag{36}$$

If $\gamma_0 > 0$, then $\mathbf{M} \succ 0$, and equality occurs only when $\nabla_{\mathbf{Z}} \mathcal{H} = 0$. □

C Details on Factor-State Co-Evolution

This section details factor-state co-evolution in HOPE, including stop-gradient, fixed-factor state updates, and warm-starting of factor assignments and slots. Starting from $\mathbf{Z}^{(0)}$, K state-evolution steps produce $\mathbf{Z}^{(K)}$, with factors inferred at every state from $k{=}0$ to K.

Algorithm 1 Factor-State Co-Evolution in HOPE

Require: Initial state $\mathbf{Z}^{(0)}$, graph structure $\tilde{\mathbf{A}}$, evolution steps K, factor inference module $\mathcal{F}_\theta$, and energy function $\mathcal{H}$

$\mathbf{P}_{\text{prev}} \leftarrow \mathbf{0}$, $\mathbf{Q}_{\text{prev}} \leftarrow \mathbf{Q}_{\text{init}}$
for $k = 0, \dots, K-1$ **do**
  $(\mathbf{P}^{(k)}, \mathbf{Q}^{(k)}) \leftarrow \mathcal{F}_\theta(\text{sg}(\mathbf{Z}^{(k)}), \tilde{\mathbf{A}}, \mathbf{P}_{\text{prev}}, \mathbf{Q}_{\text{prev}})$
  $\mathcal{H}^{(k)} = \mathcal{H}(\mathbf{Z}^{(k)}; \mathbf{P}^{(k)}, \mathbf{Q}^{(k)})$
  $\mathbf{G}^{(k)} = \nabla_{\mathbf{Z}} \mathcal{H}^{(k)}$
  $\mathbf{Z}^{(k+1)} = \mathbf{Z}^{(k)} - \Delta t \, \mathbf{M}\mathbf{G}^{(k)}$
  $\mathbf{P}_{\text{prev}} \leftarrow \mathbf{P}^{(k)}$, $\mathbf{Q}_{\text{prev}} \leftarrow \mathbf{Q}^{(k)}$
end for
$(\mathbf{P}^{(K)}, \mathbf{Q}^{(K)}) \leftarrow \mathcal{F}_\theta(\text{sg}(\mathbf{Z}^{(K)}), \tilde{\mathbf{A}}, \mathbf{P}_{\text{prev}}, \mathbf{Q}_{\text{prev}})$
return $\mathbf{Z}^{(K)}, \mathbf{P}^{(K)}, \mathbf{Q}^{(K)}$

Stop-Gradient in Factor Inference. At each $k{=}0, \dots, K$, the factor module receives the stop-gradient state $\text{sg}(\mathbf{Z}^{(k)})$. For $k{=}0$, factor inference is initialized with $\mathbf{P}_{\text{prev}}{=}\mathbf{0}$ and $\mathbf{Q}_{\text{prev}}{=}\mathbf{Q}_{\text{init}}$, whereas for $k{>}0$ it is warm-started from $\mathbf{P}^{(k-1)}$ and $\mathbf{Q}^{(k-1)}$. Thus, $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$ are inferred from $\mathbf{Z}^{(k)}$ but treated as fixed when computing the state-update gradient. Without this stop-gradient,

differentiating $\mathcal{H}(\mathbf{Z};\mathbf{P}(\mathbf{Z}),\mathbf{Q}(\mathbf{Z}))$ would introduce chain-rule terms through factor inference, such as $\partial\mathcal{H}/\partial\mathbf{P}\cdot\partial\mathbf{P}/\partial\mathbf{Z}$ and $\partial\mathcal{H}/\partial\mathbf{Q}\cdot\partial\mathbf{Q}/\partial\mathbf{Z}$. Importantly, the stop-gradient is applied only to the pathway used to infer $(\mathbf{P},\mathbf{Q})$. The representation $\mathbf{V}$ used in Eq. (7) remains differentiable with respect to $\mathbf{Z}$, so the energy gradient $\nabla_{\mathbf{Z}}\mathcal{H}$ provides the driving signal for state evolution.

Fixed-Factor State Update. For each state-evolution step $k{=}0,\ldots,K-1$, the inferred factors $(\mathbf{P}^{(k)},\mathbf{Q}^{(k)})$ are held fixed while computing $\mathbf{G}^{(k)}{=}\nabla_{\mathbf{Z}}\mathcal{H}(\mathbf{Z}^{(k)};\mathbf{P}^{(k)},\mathbf{Q}^{(k)})$. The update in Eq. (10) is an Euler discretization of the mobility-modulated fixed-factor flow. After updating $\mathbf{Z}^{(k)}$ to $\mathbf{Z}^{(k+1)}$, the factors are re-inferred from the updated state. The final state $\mathbf{Z}^{(K)}$ similarly produces $\mathbf{P}^{(K)}$ and $\mathbf{Q}^{(K)}$, without an additional state update. Thus, the dissipation guarantee applies to each of the K state-evolution steps with fixed factors, rather than across factor re-inference.

Warm-Starting and Initialization. The initial factor inference at $\mathbf{Z}^{(0)}$ uses $\mathbf{P}_{\text{prev}}{=}\mathbf{0}$ and the learnable slot states $\mathbf{Q}_{\text{prev}}{=}\mathbf{Q}_{\text{init}}$. The zero assignment makes the structural support in Eq. (4) vanish, so $\mathbf{P}^{(0)}$ initially depends only on content affinity. For $k{>}0$, factor inference is warm-started from $(\mathbf{P}^{(k-1)},\mathbf{Q}^{(k-1)})$, which promotes consistent factor identities as the states evolve. This procedure is also applied at the final state $\mathbf{Z}^{(K)}$ to obtain $(\mathbf{P}^{(K)},\mathbf{Q}^{(K)})$ for the final energy readout.

D Dataset Information

To understand our experiments, we describe the datasets and preprocessing steps in detail. We used (i) Alzheimer's Disease Neuroimaging Initiative (ADNI), (ii) Parkinson's Progression Markers Initiative (PPMI), and (iii) Autism Brain Imaging Data Exchange (ABIDE).

(i) ADNI. The ADNI study is a public dataset for Alzheimer's Disease (AD) research (Mueller et al., 2005), providing multimodal imaging and biomarker data across stages of cognitive decline. Structural brain graphs are constructed through a multi-step imaging pipeline: **1)** skull stripping and tissue segmentation on T1-MRI; **2)** brain parcellation using the Destrieux atlas; **3)** obtaining diffusion tensor information from DWI; **4)** cortical surface reconstruction using FreeSurfer; **5)** probabilistic tractography to construct structural brain networks; and **6)** measuring region-wise imaging features, including Standard Uptake Value Ratio (SUVR) of tau protein from Tau-PET (TAU), β-amyloid protein from Amyloid-PET (AMY), metabolism from FDG-PET (FDG), and cortical thickness from MRI (CT), with the cerebellum used as the reference for SUVR normalization.

Table 4: Demographics of the ADNI dataset.

Category	CN	SMC	EMCI	LMCI	AD
# of subjects	226	131	217	64	12
Gender (M / F)	114 / 112	42 / 89	121 / 96	43 / 21	9 / 3
Age (Mean ± Std)	71.27 ± 5.92	71.42 ± 4.76	68.58 ± 7.21	69.66 ± 5.25	73.46 ± 10.34
CDR-SOB (Mean ± Std)	0.02 ± 0.12	0.16 ± 0.40	1.17 ± 0.89	1.27 ± 0.98	3.54 ± 2.14
ADAS11 (Mean ± Std)	4.70 ± 2.51	4.39 ± 2.45	7.24 ± 3.30	8.99 ± 3.79	16.74 ± 4.66
ADAS13 (Mean ± Std)	7.14 ± 3.87	6.52 ± 3.72	11.28 ± 5.16	14.76 ± 5.85	27.03 ± 5.22
MMSE (Mean ± Std)	29.07 ± 1.24	29.08 ± 1.13	28.48 ± 1.58	28.23 ± 1.64	22.75 ± 4.05

Following the clinical outcomes, subjects were categorized as Control (CN), Significant Memory Concern (SMC), Early-stage Mild Cognitive Impairment (EMCI), Late-stage Mild Cognitive Impairment (LMCI), and Alzheimer's Disease (AD). The demographics of the ADNI are summarized in Table 4. It includes a total of 650 subjects, with 226 CN, 131 SMC, 217 EMCI, 64 LMCI, and 12 AD cases. Table 4 reports group-wise gender, age, and AD-related clinical measures, including CDR-SOB, ADAS11, ADAS13, and MMSE. The mean ages range from 68.58 to 73.46 years, with standard deviations between 4.76 and 10.34 years. Gender distribution varies across groups, with CN comprising 114 males and 112 females, while AD includes 9 males and 3 females.

(ii) PPMI. The PPMI aims to identify biological markers associated with the risk, onset, and progression of Parkinson's disease (PD) (Marek et al., 2011). For brain network extraction, resting-state fMRI is parcellated into 116 regions based on the AAL atlas (Tzourio-Mazoyer et al., 2002) to obtain Blood-Oxygen-Level-Dependent (BOLD) signals, which are then used to compute the connectivity matrix. To obtain a non-negative functional graph, we take the absolute correlation values and retain connections with magnitude above 0.5. Node features are extracted from BOLD signals capturing changes in blood flow. Based on clinical outcomes, disease labels were assigned to each subject as Control, Prodromal, or PD. Table 5 summarizes the demographics of the PPMI dataset, which

consists of 181 subjects: 15 Control, 53 Prodromal, and 113 PD. The average ages of these groups range from 62.03 to 64.28 years, with standard deviations between 7.77 and 10.03 years.

Table 5: Demographics of the PPMI dataset.

Category	Control	Prodromal	PD
# of subjects	15	53	113
Gender (M / F)	12 / 3	30 / 23	77 / 36
Age (Mean ± Std)	62.03 ± 10.03	64.00 ± 9.54	64.28 ± 7.77

Table 6: Demographics of the ABIDE dataset.

Category	Control	Autism
# of subjects	534	455
Gender (M / F)	441 / 93	399 / 56
Age (Mean ± Std)	16.89 ± 7.40	16.29 ± 7.58

(iii) ABIDE. The ABIDE initiative aggregated functional brain imaging data collected from laboratories around the world to support research on Autism Spectrum Disorder (ASD), which has stereotyped behaviors such as irritability, hyperactivity, depression, and anxiety (Heinsfeld et al., 2018). For ABIDE, the same preprocessing as PPMI was applied, extracting BOLD signals from 116 AAL regions as node features for functional connectivity estimation. Based on clinical outcomes, disease labels were assigned to each subject as Control or ASD. Table 6 shows the demographics of the ABIDE dataset, which consists of 989 subjects: 534 Control and 455 ASD. The average ages of these groups are 16.89 and 16.29 years, with standard deviations of 7.40 and 7.58 years, respectively.

E Implementation Details

Experimental Settings. Our method and all baseline models were trained on a single NVIDIA RTX 6000 Ada GPU. To ensure reproducibility, all experiments in the main manuscript used a fixed random seed. We optimized all models with Adam for 1000 epochs and evaluated them using 5-fold cross-validation. We report accuracy and macro-averaged precision, recall, and F1-score. A common grid search was performed over hidden units $\{2, 4, 8, 16, 32, 64\}$, dropout rates $\{0.2, 0.5, 0.8\}$, learning rates $\{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$, and weight decay values $\{5 \times 10^{-2}, 5 \times 10^{-3}, 5 \times 10^{-4}\}$. For fair comparison, we used common dataset-specific hyperparameters for all baselines, with hidden dimensions of 16, 8, and 4 for ADNI, PPMI, and ABIDE, respectively, together with a dropout rate of 0.5, a learning rate of 10^{-3}, and a weight decay of 5×10^{-4}.

For the zero-shot setting, the models trained on source folds were evaluated on the entire target cohort without using target data for training, validation, hyperparameter selection, or model selection. Importantly, the source and target cohorts were subject-disjoint, such that no subject appeared in both domains. Hyperparameters were selected using only the training/validation data within each source fold, while the held-out source test fold and all target-cohort data were excluded from model selection. We further used a single scan per subject to avoid repeated measurements across folds.

HOPE Details. For HOPE, the input ROI features were first projected into a latent space with dimension d=16. For higher-order factor discovery, the number of factor slots M was determined separately for each dataset, resulting in M=20, 10, and 6 for ADNI, PPMI, and ABIDE, respectively. The content and structural affinities were combined using the balancing coefficient β=0.25, and the temperature parameter τ=0.5 was used to control the sparsity and granularity of the factor assignments. The number of alternating assignment and slot-update iterations was set to T_f=2 for all datasets, to progressively refine the factor memberships.

For the energy-driven dynamics, we evolved node states for K=3, 3, and 4 discrete steps on ADNI, PPMI, and ABIDE, respectively. The mobility-modulated flow was discretized using explicit Euler with a fixed step size of Δt=0.25. For the connectome-constrained mobility $\mathbf{M}=\gamma_0\mathbf{I}+\gamma_1\tilde{\mathbf{L}}$, we learn the non-negative coefficients γ_0 and γ_1 through softplus parameterization, with both initialized to 1. At every dynamics step, the factors were re-inferred from the updated ROI states and initialized from the assignments and factor slots of the previous step, enabling the factor structure to co-evolve with the latent brain states. Finally, the structural regularization weight was set to λ_{str}=0.1 to balance the classification objective and the connectome coherence of the discovered factors.

Baseline Details. For fair comparisons, we used the official implementations of all baselines and additionally tuned their key model-specific hyperparameters for each dataset.

(i) Pair-wise Graph Method

- **GAT** (Veličković et al., 2018): We tuned the number of transformer heads {2, 4, 8}.

- **GRAND** (Chamberlain et al., 2021): We varied the diffusion integration time $T \in \{1, 2, 4, 8, 16, 32\}$ to control the effective depth of the diffusion process.
- **PDE-GCN** (Eliasof et al., 2021): We explored the number of PDE-GCN layers $L \in \{2, 4, 8, 16, 32, 64\}$ to adjust the depth of PDE-driven feature propagation.
- **GraphCON** (Rusch et al., 2022): We adjusted the damping parameter $\alpha \in [0, 2]$ and the frequency-control parameter $\gamma \in [0, 2]$ to regulate the oscillator dynamics.
- **BrainGNN** (Li et al., 2021): We adjusted the number of ROI communities $\{2, 4, 6, 8\}$ and the pooling ratio for R-pool layer $\{0.1, 0.2, 0.3, 0.4, 0.5\}$.
- **BrainGB** (Cui et al., 2022a): We selected the number of buckets $T \in \{5, 10, 15, 20\}$.
- **BrainNetTF** (Kan et al., 2022): We tuned the number of transformer heads $\{2, 4, 8\}$ and the number of clustering centers $\{2, 3, 4, 5, 10, 50, 100\}$ in the OCREAD readout.
- **GRAFF** (Di Giovanni et al., 2023): We tuned the gradient-flow depth $m \in \{2, 4, 8\}$.
- **SGCN** (Zhou et al., 2024): We tuned hyper-parameters $\lambda_1 \in \{1, 0.1, 0.01\}$, $\lambda_2 \in \{1, 0.1, 0.01\}$, and $\lambda_3 \in \{1, 0.1, 0.01\}$ as penalty coefficients for different regularization.
- **ALTER** (Yu et al., 2024): We tuned the hop counts k ranging from 2 to 16.
- **BioBGT** (Peng et al., 2025): We tuned the number of transformer heads $\{2, 4, 8\}$.

(ii) Higher-Order Graph Method

- **HyperDrop** (Jo et al., 2021): We set the hidden dimension of edges to 16, and randomly searched over the edge drop ratio by increasing the drop ratio from 20% to 80%.
- **HyperGT** (Liu et al., 2024): We tuned the number of transformer heads $\{2, 4, 8\}$, and coefficient for hyperedge structure losses $\{0.1, 0.3, 0.5, 0.7, 0.9\}$.
- **HyBRiD** (Qiu et al., 2024): We mainly tuned the key hyper-parameter such as the number of hyperedges $K \in \{4, 8, 12, 16, 20, 24, 28, 32\}$.
- **DHHNN** (Mei et al., 2025): We adjusted the hyperbolic curvature $c \in \{0.01, 0.1, 1\}$ and the hyperedge construction parameter $k \in \{8, 16, 32, 64, 128\}$.
- **MuHL** (Sim et al., 2026): We tuned the number of spectral scales $J \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ and the number of hyperedges $M \in \{4, 8, 12, 16, 20, 24, 28, 32\}$.

For each baseline, we selected the hyperparameter configuration with the best validation performance. Unlike methods relying solely on structural connectivity, BrainNetTF and BQN construct similarity matrices directly from node features. For higher-order methods such as HGNN, HNHN, and HyperGT, a brain-specific incidence matrix is not inherently available. We therefore constructed input hypergraphs following the strategy introduced in HGNN. In contrast, HOPE directly infers sparse higher-order factors from regional features and connectivity without predefined grouping rules, and converts their interactions into factor-wise energies governing ROI state evolution.

F Additional Experimental Results

Three-Modality In-Distribution Results. To further evaluate the effectiveness of HOPE in multimodal brain network analysis and examine the contribution of individual imaging biomarkers in ADNI, we conducted modality ablation experiments by removing one biomarker at a time from the full set of cortical thickness (CT), β-amyloid (AMY), Fluorodeoxyglucose (FDG), and tau (TAU). As reported in Table 7, we compared HOPE with representative baseline methods across all four three-modality combinations. HOPE consistently achieved the best overall performance, outperforming the second-best method by 1.1%p, 2.9%p, 2.3%p, and 1.5%p in accuracy and by 1.4%p, 1.8%p, 2.8%p, and 2.0%p in F1 score across the four settings, respectively. These results indicate that our method remains effective even when one imaging modality is unavailable, demonstrating the robustness of its higher-order energy-based modeling under reduced multimodal information.

Two-Modality In-Distribution Results. To examine how HOPE performs under more severe modality reduction, we further evaluate all six pairwise combinations of the four biomarkers, where only two imaging modalities are available for each subject. As reported in Table 8, HOPE

Table 7: AD classification performance (CN/SMC/EMCI/LMCI/AD) on the ADNI dataset under various scenarios. Each scenario uses three of the four biomarkers (CT, AMY, FDG, and TAU) by excluding one at a time. The best, second-best, and third-best results are highlighted. (†: Methods designed for brain network analysis, ‡: Physics-inspired methods)

Method	CT + AMY + FDG				CT + AMY + TAU				CT + FDG + TAU				AMY + FDG + TAU			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
Pair-wise Graph Method																
GCN	$84.9_{\pm 1.7}$	$85.3_{\pm 7.6}$	$80.3_{\pm 7.1}$	$81.9_{\pm 7.5}$	$85.5_{\pm 1.6}$	$87.3_{\pm 6.3}$	$81.7_{\pm 3.9}$	$83.6_{\pm 4.5}$	$83.5_{\pm 2.7}$	$87.0_{\pm 4.7}$	$81.0_{\pm 4.8}$	$82.5_{\pm 3.7}$	$86.6_{\pm 2.7}$	$85.9_{\pm 3.8}$	$83.2_{\pm 5.8}$	$83.4_{\pm 4.3}$
GAT	$86.8_{\pm 2.3}$	$89.4_{\pm 1.4}$	$85.6_{\pm 2.5}$	$86.9_{\pm 2.1}$	$88.3_{\pm 1.8}$	$91.3_{\pm 1.8}$	$84.0_{\pm 2.3}$	$86.4_{\pm 1.9}$	$87.4_{\pm 3.0}$	$91.5_{\pm 1.8}$	$86.2_{\pm 5.1}$	$88.1_{\pm 3.9}$	$88.9_{\pm 3.2}$	$91.3_{\pm 3.5}$	$86.6_{\pm 4.7}$	$88.0_{\pm 3.8}$
GCNII	$78.8_{\pm 4.0}$	$86.9_{\pm 2.8}$	$77.1_{\pm 6.3}$	$79.9_{\pm 5.4}$	$81.8_{\pm 4.3}$	$87.1_{\pm 2.5}$	$80.1_{\pm 7.4}$	$85.2_{\pm 5.3}$	$81.1_{\pm 4.8}$	$85.2_{\pm 7.6}$	$77.0_{\pm 7.8}$	$79.3_{\pm 7.5}$	$84.2_{\pm 4.4}$	$85.3_{\pm 7.6}$	$81.2_{\pm 9.3}$	$82.3_{\pm 8.2}$
GRAND‡	$87.0_{\pm 4.2}$	$88.9_{\pm 2.9}$	$85.9_{\pm 5.2}$	$87.7_{\pm 4.1}$	$87.3_{\pm 2.9}$	$90.4_{\pm 1.6}$	$87.4_{\pm 5.5}$	$88.4_{\pm 3.9}$	$87.0_{\pm 1.1}$	$90.3_{\pm 0.9}$	$85.8_{\pm 3.8}$	$87.4_{\pm 2.8}$	$86.8_{\pm 3.6}$	$90.3_{\pm 2.1}$	$86.1_{\pm 4.9}$	$87.5_{\pm 3.6}$
PDE-GCN‡	$87.8_{\pm 2.8}$	$89.3_{\pm 3.7}$	$85.5_{\pm 4.6}$	$86.4_{\pm 4.1}$	$88.5_{\pm 3.0}$	$90.0_{\pm 1.9}$	$87.8_{\pm 4.9}$	$88.4_{\pm 3.5}$	$86.7_{\pm 2.6}$	$88.6_{\pm 2.6}$	$85.4_{\pm 3.9}$	$86.3_{\pm 3.1}$	$87.8_{\pm 2.2}$	$89.8_{\pm 2.1}$	$86.0_{\pm 4.0}$	$87.3_{\pm 2.5}$
GraphCON‡	$87.0_{\pm 3.4}$	$90.5_{\pm 2.0}$	$85.1_{\pm 5.4}$	$87.6_{\pm 4.0}$	$88.4_{\pm 3.1}$	$91.0_{\pm 1.6}$	$87.6_{\pm 5.2}$	$88.8_{\pm 3.8}$	$88.5_{\pm 2.4}$	$91.1_{\pm 1.4}$	$87.6_{\pm 4.9}$	$88.8_{\pm 3.9}$	$88.3_{\pm 2.5}$	$90.9_{\pm 1.6}$	$87.9_{\pm 5.1}$	$88.9_{\pm 3.8}$
BrainGNN†	$76.5_{\pm 4.0}$	$79.2_{\pm 4.5}$	$77.0_{\pm 2.8}$	$78.1_{\pm 3.0}$	$76.8_{\pm 4.8}$	$79.5_{\pm 4.2}$	$77.2_{\pm 3.1}$	$78.3_{\pm 3.2}$	$79.2_{\pm 4.1}$	$81.8_{\pm 3.9}$	$79.8_{\pm 2.2}$	$80.9_{\pm 2.6}$	$80.0_{\pm 4.2}$	$82.8_{\pm 4.0}$	$80.5_{\pm 2.5}$	$81.6_{\pm 2.9}$
IBGNN†	$75.4_{\pm 5.1}$	$81.2_{\pm 3.5}$	$77.6_{\pm 6.1}$	$78.3_{\pm 5.4}$	$74.2_{\pm 4.4}$	$78.7_{\pm 4.3}$	$74.0_{\pm 3.5}$	$75.2_{\pm 1.9}$	$82.9_{\pm 2.5}$	$83.8_{\pm 3.5}$	$79.1_{\pm 2.6}$	$80.5_{\pm 2.7}$	$81.7_{\pm 2.5}$	$84.6_{\pm 5.2}$	$80.4_{\pm 3.7}$	$81.2_{\pm 3.6}$
BrainGB†	$78.8_{\pm 2.2}$	$84.0_{\pm 3.2}$	$77.7_{\pm 3.2}$	$79.5_{\pm 2.7}$	$80.5_{\pm 2.5}$	$86.1_{\pm 2.9}$	$77.3_{\pm 3.5}$	$80.2_{\pm 2.9}$	$81.8_{\pm 4.0}$	$82.5_{\pm 7.9}$	$81.1_{\pm 6.6}$	$80.3_{\pm 6.7}$	$82.8_{\pm 2.5}$	$86.3_{\pm 4.0}$	$82.7_{\pm 5.4}$	$83.6_{\pm 4.1}$
BrainNetTF†	$89.2_{\pm 1.8}$	$91.4_{\pm 2.1}$	$86.6_{\pm 7.1}$	$87.8_{\pm 5.5}$	$86.8_{\pm 2.0}$	$90.1_{\pm 1.7}$	$84.5_{\pm 6.1}$	$86.2_{\pm 4.8}$	$88.6_{\pm 3.5}$	$87.2_{\pm 6.3}$	$86.2_{\pm 7.0}$	$85.9_{\pm 6.2}$	$90.0_{\pm 2.6}$	$90.0_{\pm 6.6}$	$89.9_{\pm 5.2}$	$88.9_{\pm 6.2}$
GRAFF‡	$85.5_{\pm 2.9}$	$88.8_{\pm 1.2}$	$84.8_{\pm 6.1}$	$86.1_{\pm 4.1}$	$87.7_{\pm 3.3}$	$89.9_{\pm 2.6}$	$87.9_{\pm 5.7}$	$88.2_{\pm 4.4}$	$86.3_{\pm 4.6}$	$91.5_{\pm 2.0}$	$84.5_{\pm 7.0}$	$86.6_{\pm 5.5}$	$89.7_{\pm 2.9}$	$92.9_{\pm 1.5}$	$88.6_{\pm 6.0}$	$90.1_{\pm 4.2}$
SGCN†	$81.2_{\pm 1.8}$	$85.2_{\pm 3.2}$	$81.8_{\pm 2.3}$	$82.8_{\pm 1.9}$	$82.2_{\pm 1.1}$	$86.5_{\pm 3.0}$	$83.7_{\pm 2.1}$	$84.3_{\pm 1.8}$	$81.7_{\pm 2.1}$	$87.9_{\pm 2.4}$	$78.2_{\pm 4.7}$	$81.3_{\pm 4.0}$	$84.2_{\pm 2.1}$	$89.2_{\pm 0.8}$	$80.7_{\pm 4.8}$	$83.5_{\pm 3.7}$
ALTER†	$89.1_{\pm 2.1}$	$87.6_{\pm 8.1}$	$85.6_{\pm 8.9}$	$86.0_{\pm 9.4}$	$87.5_{\pm 2.3}$	$90.6_{\pm 1.4}$	$85.1_{\pm 5.9}$	$86.7_{\pm 4.3}$	$88.9_{\pm 3.5}$	$89.4_{\pm 4.6}$	$88.0_{\pm 5.6}$	$88.0_{\pm 5.0}$	$90.9_{\pm 2.8}$	$90.9_{\pm 4.6}$	$90.1_{\pm 5.2}$	$89.9_{\pm 4.4}$
BioBGT†	$79.2_{\pm 2.3}$	$78.5_{\pm 3.2}$	$77.3_{\pm 5.8}$	$77.1_{\pm 4.3}$	$77.8_{\pm 3.1}$	$76.9_{\pm 4.4}$	$76.3_{\pm 2.1}$	$76.7_{\pm 4.2}$	$81.5_{\pm 4.3}$	$84.6_{\pm 3.6}$	$82.1_{\pm 6.1}$	$82.9_{\pm 3.7}$	$83.2_{\pm 2.9}$	$85.6_{\pm 7.3}$	$82.9_{\pm 4.6}$	$83.0_{\pm 7.1}$
BQN†	$82.3_{\pm 2.1}$	$78.1_{\pm 3.5}$	$76.5_{\pm 5.1}$	$76.8_{\pm 4.2}$	$81.0_{\pm 3.1}$	$76.2_{\pm 4.1}$	$73.8_{\pm 3.8}$	$74.5_{\pm 3.5}$	$85.6_{\pm 2.7}$	$73.6_{\pm 3.2}$	$76.9_{\pm 6.5}$	$76.8_{\pm 5.8}$	$88.0_{\pm 2.7}$	$77.8_{\pm 7.7}$	$77.4_{\pm 7.2}$	$77.3_{\pm 7.4}$
Higher-Order Graph Method																
HGNN	$84.3_{\pm 2.5}$	$88.3_{\pm 1.9}$	$83.3_{\pm 4.9}$	$84.9_{\pm 3.5}$	$80.2_{\pm 1.1}$	$82.1_{\pm 7.2}$	$76.0_{\pm 6.7}$	$78.1_{\pm 6.6}$	$83.7_{\pm 3.4}$	$87.0_{\pm 2.6}$	$80.0_{\pm 5.8}$	$82.1_{\pm 4.1}$	$85.1_{\pm 1.3}$	$88.0_{\pm 1.9}$	$84.5_{\pm 3.4}$	$85.7_{\pm 2.5}$
HNHN	$86.8_{\pm 1.6}$	$89.2_{\pm 1.7}$	$85.9_{\pm 5.0}$	$86.8_{\pm 4.0}$	$84.3_{\pm 2.0}$	$86.9_{\pm 1.8}$	$84.8_{\pm 4.9}$	$85.4_{\pm 3.7}$	$87.6_{\pm 2.0}$	$90.2_{\pm 2.0}$	$85.1_{\pm 5.7}$	$86.5_{\pm 4.3}$	$86.9_{\pm 1.7}$	$87.6_{\pm 2.7}$	$86.5_{\pm 4.2}$	$86.4_{\pm 3.0}$
UniGCNII	$88.6_{\pm 1.9}$	$89.4_{\pm 2.0}$	$86.6_{\pm 4.7}$	$77.5_{\pm 3.7}$	$88.3_{\pm 2.7}$	$87.2_{\pm 4.1}$	$86.3_{\pm 4.9}$	$86.5_{\pm 4.4}$	$89.4_{\pm 2.1}$	$91.3_{\pm 1.6}$	$87.1_{\pm 3.7}$	$88.6_{\pm 2.5}$	$88.3_{\pm 2.0}$	$89.4_{\pm 2.6}$	$86.3_{\pm 4.9}$	$77.3_{\pm 4.1}$
HyperDrop	$71.5_{\pm 1.2}$	$73.2_{\pm 2.8}$	$64.8_{\pm 1.5}$	$67.5_{\pm 2.1}$	$72.8_{\pm 1.5}$	$74.9_{\pm 3.0}$	$66.2_{\pm 1.1}$	$69.1_{\pm 1.9}$	$73.1_{\pm 1.8}$	$75.1_{\pm 2.9}$	$66.5_{\pm 2.0}$	$69.4_{\pm 2.3}$	$72.0_{\pm 1.0}$	$74.0_{\pm 2.5}$	$65.2_{\pm 1.3}$	$68.2_{\pm 1.8}$
dwHGCN†	$88.8_{\pm 2.7}$	$91.4_{\pm 2.7}$	$86.9_{\pm 3.7}$	$88.5_{\pm 3.1}$	$85.5_{\pm 1.5}$	$88.6_{\pm 2.3}$	$85.6_{\pm 4.4}$	$86.5_{\pm 2.9}$	$88.5_{\pm 3.0}$	$90.6_{\pm 2.8}$	$86.7_{\pm 4.3}$	$88.1_{\pm 3.3}$	$88.5_{\pm 2.1}$	$91.3_{\pm 1.9}$	$86.4_{\pm 4.8}$	$88.2_{\pm 3.5}$
HyperGT	$78.3_{\pm 2.9}$	$82.2_{\pm 9.2}$	$69.4_{\pm 7.4}$	$73.4_{\pm 7.3}$	$76.6_{\pm 3.5}$	$74.2_{\pm 9.5}$	$63.0_{\pm 5.2}$	$65.8_{\pm 5.7}$	$82.5_{\pm 2.5}$	$81.8_{\pm 9.8}$	$70.5_{\pm 6.2}$	$74.1_{\pm 7.4}$	$81.2_{\pm 2.1}$	$85.7_{\pm 7.1}$	$69.8_{\pm 2.4}$	$74.6_{\pm 3.6}$
HyBRiD†	$86.5_{\pm 1.5}$	$84.7_{\pm 3.8}$	$86.8_{\pm 4.6}$	$85.2_{\pm 3.8}$	$79.2_{\pm 4.2}$	$80.0_{\pm 6.3}$	$79.8_{\pm 5.9}$	$79.4_{\pm 6.0}$	$85.5_{\pm 5.4}$	$87.3_{\pm 5.4}$	$85.2_{\pm 6.9}$	$85.5_{\pm 5.9}$	$83.8_{\pm 4.4}$	$83.5_{\pm 6.7}$	$81.3_{\pm 6.1}$	$82.1_{\pm 6.4}$
DHHNN	$80.5_{\pm 2.1}$	$88.2_{\pm 3.8}$	$83.1_{\pm 3.5}$	$85.2_{\pm 3.1}$	$79.8_{\pm 2.8}$	$87.5_{\pm 4.2}$	$82.5_{\pm 4.0}$	$84.6_{\pm 3.6}$	$81.8_{\pm 2.3}$	$89.4_{\pm 3.9}$	$85.0_{\pm 3.1}$	$87.0_{\pm 3.2}$	$81.2_{\pm 2.4}$	$88.8_{\pm 4.1}$	$84.2_{\pm 3.3}$	$86.3_{\pm 3.0}$
MuHL†	$91.0_{\pm 2.2}$	$92.9_{\pm 1.9}$	$90.1_{\pm 3.3}$	$91.0_{\pm 3.3}$	$88.9_{\pm 1.9}$	$92.9_{\pm 2.0}$	$87.4_{\pm 5.0}$	$89.4_{\pm 3.9}$	$90.0_{\pm 2.1}$	$93.4_{\pm 1.4}$	$89.5_{\pm 6.4}$	$89.2_{\pm 4.5}$	$91.3_{\pm 1.9}$	$93.2_{\pm 0.9}$	$90.4_{\pm 4.1}$	$90.8_{\pm 3.0}$
HOPE (Ours)	$92.1_{\pm 0.6}$	$93.1_{\pm 2.5}$	$90.8_{\pm 4.3}$	$92.4_{\pm 3.7}$	$91.8_{\pm 1.9}$	$93.5_{\pm 1.3}$	$89.9_{\pm 4.4}$	$91.2_{\pm 3.1}$	$92.3_{\pm 1.6}$	$94.6_{\pm 1.7}$	$90.5_{\pm 4.5}$	$92.0_{\pm 3.3}$	$92.8_{\pm 1.6}$	$94.6_{\pm 1.9}$	$91.1_{\pm 4.9}$	$92.8_{\pm 3.9}$

Table 8: AD classification performance (CN/SMC/EMCI/LMCI/AD) between HOPE and baselines for brain network analysis on the ADNI dataset under various scenarios. Each scenario uses two of the four biomarkers (CT, AMY, FDG, and TAU). The best, second-best, and third-best results are highlighted.

Method	CT + AMY				CT + FDG				CT + TAU			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
BrainGNN	$75.4_{\pm 3.9}$	$65.7_{\pm 5.9}$	$63.4_{\pm 5.3}$	$62.3_{\pm 5.3}$	$72.3_{\pm 4.7}$	$63.8_{\pm 5.5}$	$66.3_{\pm 5.7}$	$65.5_{\pm 7.1}$	$70.2_{\pm 1.8}$	$65.3_{\pm 5.8}$	$69.0_{\pm 6.3}$	$69.2_{\pm 6.1}$
IBGNN	$76.3_{\pm 3.7}$	$76.9_{\pm 4.5}$	$72.5_{\pm 4.4}$	$73.7_{\pm 4.1}$	$75.1_{\pm 3.2}$	$78.2_{\pm 5.7}$	$69.4_{\pm 6.7}$	$72.3_{\pm 6.0}$	$76.9_{\pm 2.5}$	$74.0_{\pm 5.4}$	$71.4_{\pm 4.3}$	$71.9_{\pm 4.6}$
BrainGB	$76.5_{\pm 4.5}$	$79.9_{\pm 3.2}$	$75.3_{\pm 5.1}$	$75.8_{\pm 4.4}$	$77.7_{\pm 5.2}$	$80.2_{\pm 4.2}$	$77.0_{\pm 3.9}$	$77.7_{\pm 6.2}$	$78.8_{\pm 3.5}$	$78.4_{\pm 6.5}$	$76.8_{\pm 4.4}$	$77.0_{\pm 5.8}$
BrainNetTF	$83.5_{\pm 3.5}$	$85.5_{\pm 3.0}$	$79.5_{\pm 4.5}$	$81.0_{\pm 4.0}$	$86.5_{\pm 2.5}$	$87.0_{\pm 2.8}$	$84.0_{\pm 3.0}$	$84.5_{\pm 3.0}$	$84.0_{\pm 3.5}$	$83.5_{\pm 3.5}$	$79.5_{\pm 3.8}$	$80.0_{\pm 3.5}$
SGCN	$78.5_{\pm 2.4}$	$82.3_{\pm 3.1}$	$76.3_{\pm 4.8}$	$78.3_{\pm 3.9}$	$76.9_{\pm 3.2}$	$77.9_{\pm 3.7}$	$73.1_{\pm 2.9}$	$74.6_{\pm 3.0}$	$82.2_{\pm 2.9}$	$85.0_{\pm 4.0}$	$80.4_{\pm 3.0}$	$81.7_{\pm 3.4}$
ALTER	$86.2_{\pm 6.1}$	$88.1_{\pm 4.9}$	$81.7_{\pm 5.3}$	$83.5_{\pm 5.0}$	$89.5_{\pm 2.6}$	$89.7_{\pm 3.0}$	$86.8_{\pm 3.2}$	$87.2_{\pm 3.2}$	$86.6_{\pm 4.2}$	$86.0_{\pm 3.9}$	$82.1_{\pm 4.0}$	$82.7_{\pm 3.7}$
BioBGT	$77.4_{\pm 2.0}$	$80.6_{\pm 1.9}$	$76.7_{\pm 5.5}$	$78.0_{\pm 3.9}$	$80.6_{\pm 3.6}$	$84.2_{\pm 1.9}$	$78.8_{\pm 5.9}$	$80.8_{\pm 4.0}$	$80.3_{\pm 1.5}$	$83.6_{\pm 1.4}$	$79.9_{\pm 4.6}$	$81.2_{\pm 3.2}$
BQN	$73.7_{\pm 3.6}$	$74.3_{\pm 8.0}$	$68.7_{\pm 5.3}$	$70.4_{\pm 5.9}$	$84.0_{\pm 2.5}$	$80.3_{\pm 8.5}$	$74.3_{\pm 8.0}$	$76.0_{\pm 7.6}$	$86.3_{\pm 1.6}$	$89.6_{\pm 2.3}$	$82.5_{\pm 6.0}$	$84.6_{\pm 4.2}$
dwHGCN	$83.1_{\pm 3.4}$	$88.3_{\pm 4.1}$	$83.1_{\pm 4.0}$	$84.9_{\pm 3.9}$	$88.3_{\pm 5.2}$	$81.4_{\pm 6.2}$	$86.9_{\pm 5.8}$	$84.5_{\pm 6.0}$	$86.8_{\pm 4.7}$	$85.8_{\pm 5.0}$	$83.9_{\pm 4.9}$	$84.9_{\pm 3.9}$
HyBRiD	$81.4_{\pm 2.7}$	$79.7_{\pm 7.8}$	$74.2_{\pm 3.8}$	$76.2_{\pm 5.2}$	$87.4_{\pm 4.1}$	$88.2_{\pm 6.0}$	$85.1_{\pm 5.1}$	$86.1_{\pm 5.3}$	$85.1_{\pm 3.4}$	$84.6_{\pm 4.9}$	$82.8_{\pm 5.2}$	$83.0_{\pm 4.0}$
MuHL	$87.5_{\pm 2.5}$	$89.1_{\pm 2.1}$	$85.4_{\pm 1.9}$	$86.6_{\pm 1.9}$	$90.9_{\pm 2.6}$	$92.8_{\pm 2.4}$	$89.7_{\pm 2.0}$	$90.7_{\pm 2.1}$	$88.9_{\pm 1.9}$	$87.1_{\pm 2.1}$	$86.5_{\pm 2.4}$	$86.4_{\pm 2.2}$
HOPE (Ours)	$87.4_{\pm 3.9}$	$89.3_{\pm 3.7}$	$87.3_{\pm 5.3}$	$87.6_{\pm 4.0}$	$90.9_{\pm 2.7}$	$94.0_{\pm 1.3}$	$89.1_{\pm 4.3}$	$90.9_{\pm 4.4}$	$89.9_{\pm 1.9}$	$92.6_{\pm 1.5}$	$89.5_{\pm 4.4}$	$90.5_{\pm 3.5}$

Method	AMY + FDG				AMY + TAU				FDG + TAU			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
BrainGNN	$73.7_{\pm 3.9}$	$69.7_{\pm 5.1}$	$66.1_{\pm 7.5}$	$63.4_{\pm 5.2}$	$75.3_{\pm 5.1}$	$72.2_{\pm 3.2}$	$69.8_{\pm 2.4}$	$70.0_{\pm 2.1}$	$75.2_{\pm 3.9}$	$70.4_{\pm 5.2}$	$66.6_{\pm 7.2}$	$63.7_{\pm 8.4}$
IBGNN	$75.8_{\pm 3.9}$	$79.1_{\pm 4.5}$	$76.4_{\pm 7.4}$	$76.7_{\pm 5.9}$	$75.8_{\pm 2.6}$	$81.9_{\pm 2.3}$	$74.3_{\pm 6.4}$	$76.4_{\pm 5.1}$	$82.2_{\pm 3.9}$	$80.6_{\pm 5.8}$	$75.9_{\pm 5.3}$	$77.4_{\pm 5.1}$
BrainGB	$80.8_{\pm 5.8}$	$83.7_{\pm 6.2}$	$79.7_{\pm 6.4}$	$79.8_{\pm 5.9}$	$82.0_{\pm 5.2}$	$82.3_{\pm 5.9}$	$78.2_{\pm 5.6}$	$79.4_{\pm 5.8}$	$80.6_{\pm 4.9}$	$79.8_{\pm 3.8}$	$72.9_{\pm 4.6}$	$74.6_{\pm 4.0}$
BrainNetTF	$85.0_{\pm 3.5}$	$85.5_{\pm 4.0}$	$83.5_{\pm 4.5}$	$83.0_{\pm 4.5}$	$85.0_{\pm 3.8}$	$87.0_{\pm 3.5}$	$84.5_{\pm 3.5}$	$85.5_{\pm 3.5}$	$87.5_{\pm 2.5}$	$88.5_{\pm 2.5}$	$86.5_{\pm 3.0}$	$87.5_{\pm 3.0}$
SGCN	$80.0_{\pm 3.9}$	$84.1_{\pm 3.5}$	$78.5_{\pm 4.1}$	$80.1_{\pm 3.7}$	$81.4_{\pm 2.8}$	$85.9_{\pm 2.9}$	$80.5_{\pm 3.3}$	$81.8_{\pm 3.0}$	$80.9_{\pm 3.7}$	$84.9_{\pm 3.5}$	$79.4_{\pm 3.3}$	$80.4_{\pm 3.3}$
ALTER	$87.9_{\pm 4.2}$	$88.3_{\pm 5.0}$	$86.3_{\pm 5.2}$	$85.8_{\pm 6.1}$	$87.8_{\pm 4.6}$	$89.7_{\pm 4.9}$	$87.4_{\pm 4.3}$	$88.2_{\pm 4.7}$	$90.2_{\pm 2.8}$	$91.5_{\pm 2.9}$	$89.1_{\pm 3.2}$	$90.2_{\pm 3.3}$
BioBGT	$79.4_{\pm 2.6}$	$81.5_{\pm 2.9}$	$78.3_{\pm 5.7}$	$79.1_{\pm 4.4}$	$78.9_{\pm 1.6}$	$82.2_{\pm 1.9}$	$79.1_{\pm 3.5}$	$80.1_{\pm 2.1}$	$81.2_{\pm 1.6}$	$84.4_{\pm 1.4}$	$79.4_{\pm 5.8}$	$81.2_{\pm 3.8}$
BQN	$78.8_{\pm 3.9}$	$79.8_{\pm 7.6}$	$75.6_{\pm 5.6}$	$76.4_{\pm 5.6}$	$74.2_{\pm 3.7}$	$75.3_{\pm 7.9}$	$71.9_{\pm 8.9}$	$72.7_{\pm 8.1}$	$84.6_{\pm 3.5}$	$86.9_{\pm 4.1}$	$79.7_{\pm 6.7}$	$81.8_{\pm 4.2}$
dwHGCN	$86.5_{\pm 3.5}$	$89.2_{\pm 3.6}$	$84.0_{\pm 3.9}$	$85.9_{\pm 4.0}$	$85.2_{\pm 3.6}$	$89.0_{\pm 3.9}$	$86.0_{\pm 3.5}$	$86.8_{\pm 3.7}$	$88.6_{\pm 2.9}$	$90.5_{\pm 3.5}$	$86.7_{\pm 3.0}$	$88.5_{\pm 3.2}$
HyBRiD	$84.9_{\pm 3.4}$	$84.6_{\pm 4.9}$	$82.8_{\pm 5.2}$	$83.0_{\pm 4.0}$	$80.9_{\pm 3.1}$	$83.0_{\pm 4.6}$	$80.3_{\pm 6.4}$	$81.2_{\pm 5.6}$	$89.8_{\pm 2.5}$	$91.1_{\pm 4.7}$	$88.8_{\pm 5.4}$	$89.6_{\pm 4.8}$
MuHL	$88.9_{\pm 2.5}$	$91.7_{\pm 2.4}$	$87.5_{\pm 2.6}$	$88.9_{\pm 2.5}$	$88.9_{\pm 2.2}$	$92.0_{\pm 2.5}$	$88.1_{\pm 2.0}$	$88.8_{\pm 2.2}$	$91.2_{\pm 1.9}$	$92.7_{\pm 2.2}$	$90.7_{\pm 2.2}$	$91.1_{\pm 2.0}$
HOPE (Ours)	$91.1_{\pm 3.3}$	$93.3_{\pm 2.7}$	$89.0_{\pm 3.1}$	$90.5_{\pm 3.5}$	$89.5_{\pm 2.6}$	$92.2_{\pm 2.1}$	$89.2_{\pm 3.7}$	$90.2_{\pm 2.5}$	$92.0_{\pm 1.9}$	$94.7_{\pm 1.3}$	$89.9_{\pm 4.0}$	$91.6_{\pm 3.2}$

achieves the highest F1 score in every setting, surpassing the second-best method by 1.0%p, 0.2%p, 4.1%p, 1.6%p, 1.4%p, and 0.5%p on CT+AMY, CT+FDG, CT+TAU, AMY+FDG, AMY+TAU, and FDG+TAU, respectively, while also attaining the best precision across all six settings. Notably, the largest gains are observed in combinations involving TAU, suggesting that the higher-order energy formulation effectively exploits complementary information between tau pathology and the other biomarkers. Overall, these findings show that HOPE degrades gracefully as modalities are removed and remains effective even when only half of the biomarkers are available.

AD Clinical Symptom Regression. Beyond classification, we investigate whether the representations learned by HOPE carry information about the continuous severity of AD, by regressing four clinical measures, CDR-SOB, ADAS11, ADAS13, and MMSE (Mohs & Cohen, 1988; Kueper et al., 2018), from the last-layer representations via Lasso regression (Tibshirani, 1996), whose sparsity-inducing penalty suits the high-dimensional, low-sample regime of ADNI. For a fair comparison, the

Table 9: Comparison of regression performance between existing brain network-based methods and HOPE in predicting AD-related clinical symptoms. For each method, the predicted scores are evaluated against the ground-truth clinical measures, including CDR-SOB, ADAS11, ADAS13, and MMSE, on the ADNI dataset. The best results are shown in **bold**, and the second-best results are underlined.

Method	CDR-SOB				ADAS11				ADAS13				MMSE			
	P value ↓	Corr ↑	RMSE ↓	R^2 ↑	P value ↓	Corr ↑	RMSE ↓	R^2 ↑	P value ↓	Corr ↑	RMSE ↓	R^2 ↑	P value ↓	Corr ↑	RMSE ↓	R^2 ↑
BrainGB	4.6e-09	0.4857	0.7799	0.1332	1.3e-17	0.6597	2.8198	0.4096	4.9e-19	0.6808	4.5031	0.4756	4.0e-04	0.3059	1.5841	0.1587
BrainNetTF	1.2e-09	0.5014	0.7900	0.1635	1.7e-20	0.7004	2.7274	0.4166	7.2e-20	0.6923	4.3246	0.5468	6.7e-07	0.4196	1.5080	0.1678
SGCN	2.3e-08	0.4655	0.8000	0.1537	2.6e-13	0.5853	3.0907	0.4249	3.8e-15	0.6196	4.7065	0.4450	1.4e-03	0.3266	1.5809	0.1768
ALTER	7.8e-10	0.5064	0.7947	0.1799	7.9e-16	0.6315	2.9718	0.4437	7.5e-18	0.6637	4.5317	0.5219	2.8e-05	0.3582	1.5435	0.2272
BioBGT	1.7e-07	0.4390	0.8000	0.2165	2.5e-08	0.4651	3.3318	0.4636	1.6e-08	0.4710	5.2917	0.5268	4.2e-06	0.3790	1.6395	0.2142
BQN	**3.6e-10**	**0.5617**	**0.7359**	0.1958	3.2e-09	0.4902	3.3028	0.4258	7.0e-10	0.5078	5.1982	0.4899	8.2e-03	0.2310	1.6312	0.2292
HyBRiD	1.1e-09	0.5024	0.8167	0.1837	1.0e-11	0.5518	3.2215	0.4198	1.4e-11	0.5485	5.2112	0.5579	8.4e-06	0.3796	1.5493	0.2119
MuHL	2.7e-09	0.4920	0.7625	0.2212	**1.4e-21**	**0.7701**	**2.5205**	0.4699	5.6e-20	0.6340	4.2403	0.5667	8.2e-08	0.4492	1.4953	0.2402
HOPE (Ours)	2.7e-09	0.5123	0.8066	**0.2636**	1.5e-20	0.7013	2.6799	**0.4912**	**5.1e-27**	**0.7726**	**3.7927**	**0.6001**	**5.8e-10**	**0.5099**	**1.4273**	**0.2653**

Table 10: Random-seed robustness of ALTER, MuHL, and HOPE across brain network benchmarks. Each method is evaluated with five seeds under 5-fold cross-validation, reporting mean ± standard deviation and 95% confidence intervals. The best and second-best results are shown in **bold** and underlined, respectively.

Dataset	Method	Acc		Pre		Rec		F1s	
		Mean ± Std	95% CI	Mean ± Std	95% CI	Mean ± Std	95% CI	Mean ± Std	95% CI
ADNI	ALTER	90.31 ± 0.81	[89.30, 91.32]	91.52 ± 1.95	[89.10, 93.94]	88.81 ± 1.43	[87.03, 90.59]	89.79 ± 1.30	[88.18, 91.40]
	MuHL	92.86 ± 0.59	[92.13, 93.59]	95.30 ± 1.34	[93.64, 96.96]	90.86 ± 1.09	[89.51, 92.21]	92.38 ± 1.26	[90.82, 93.94]
	HOPE (Ours)	95.55 ± 0.35	[95.12, 95.98]	96.87 ± 0.63	[96.09, 97.65]	94.05 ± 0.75	[93.12, 94.98]	95.37 ± 0.60	[94.63, 96.11]
PPMI	ALTER	67.73 ± 1.20	[66.24, 69.22]	51.99 ± 5.49	[45.17, 58.81]	51.01 ± 3.74	[46.37, 55.65]	50.28 ± 5.08	[43.97, 56.59]
	MuHL	73.37 ± 1.32	[71.73, 75.01]	64.42 ± 4.60	[58.71, 70.13]	58.48 ± 2.02	[55.97, 60.99]	60.70 ± 2.21	[57.96, 63.44]
	HOPE (Ours)	75.95 ± 1.69	[73.85, 78.05]	65.99 ± 3.69	[61.41, 70.57]	60.65 ± 1.89	[58.30, 63.00]	63.15 ± 2.40	[60.17, 66.13]
ABIDE	ALTER	59.43 ± 1.70	[57.32, 61.54]	59.24 ± 1.69	[57.14, 61.34]	58.69 ± 1.79	[56.47, 60.91]	58.17 ± 1.01	[56.92, 59.42]
	MuHL	62.86 ± 1.05	[61.56, 64.16]	61.27 ± 1.81	[59.02, 63.52]	61.40 ± 1.95	[58.98, 63.82]	61.51 ± 1.38	[59.80, 63.22]
	HOPE (Ours)	64.18 ± 1.75	[62.01, 66.35]	64.19 ± 1.93	[61.79, 66.59]	63.46 ± 1.70	[61.35, 65.57]	63.13 ± 1.61	[61.13, 65.13]

same regression pipeline is applied to representations extracted from brain network-based baselines, and representations are min-max normalized before being fit with 5-fold cross-validation. This analysis complements the classification results by testing whether the discovered higher-order structure remains informative when the target is a continuous clinical score rather than a discrete diagnostic label, thereby probing the clinical relevance of the learned brain-state representations.

Table 9 summarizes the regression results using Pearson correlation (Corr), p-value, RMSE, and R-squared score. For ADAS13 and MMSE, HOPE achieves the best performance across all metrics, with correlations of 0.7726 and 0.5099 and RMSE values of 3.7927 and 1.4273, respectively, while improving R-squared by 3.34%p and 2.51%p over the previous best baselines. For CDR-SOB and ADAS11, BQN and MuHL achieve better results on several metrics, respectively, whereas HOPE attains the highest R-squared scores, improving them by 4.24%p and 2.13%p. Overall, these results show that HOPE captures clinical severity across all four measures, supporting the clinical relevance of its physics-inspired higher-order representations.

Robustness to Random Initialization. To evaluate the robustness of the compared methods to random initialization, we repeated the 5-fold cross-validation using five different random seeds for ALTER, MuHL, and HOPE. As summarized in Table 10, HOPE achieves the highest average performance across all datasets, with relatively small variations across independent runs. For example, HOPE achieves accuracies of 95.55±0.35, 75.95±1.69, and 64.18±1.75 on ADNI, PPMI, and ABIDE, respectively, with 95% confidence intervals of [95.12, 95.98], [73.85, 78.05], and [62.01, 66.35]. Similar trends are observed for precision, recall, and F1-score, indicating that the performance gains of HOPE are consistently maintained across different random initializations.

G Additional Analysis

G.1 Evolution Dynamics and Model Analysis

Empirical Energy Dissipation. To verify whether the discretized dynamics decrease factor-level energy, we measure the relative per-step decrease $\Delta\mathcal{H}^{(k)}$ as the energy reduction within a state update normalized by its pre-update energy, while keeping the inferred factors fixed. As shown in Fig. 8, $\Delta\mathcal{H}^{(k)}$ remains positive across all diagnostic groups and datasets, confirming that the Euler updates preserve the energy-decreasing behavior of the dissipative flow. The relative decrease becomes progressively smaller over successive steps, indicating larger energy-driven corrections

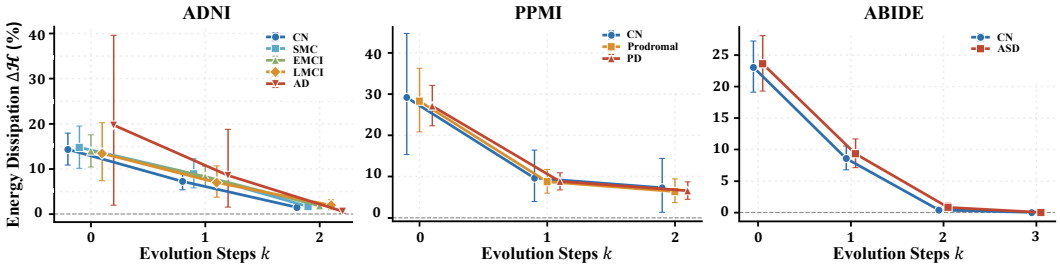


Figure 8: Relative energy decrease $\Delta\mathcal{H}^{(k)}$ across evolution steps for each diagnostic group. Positive values indicate that the discretized fixed-factor updates decrease the corresponding energy.

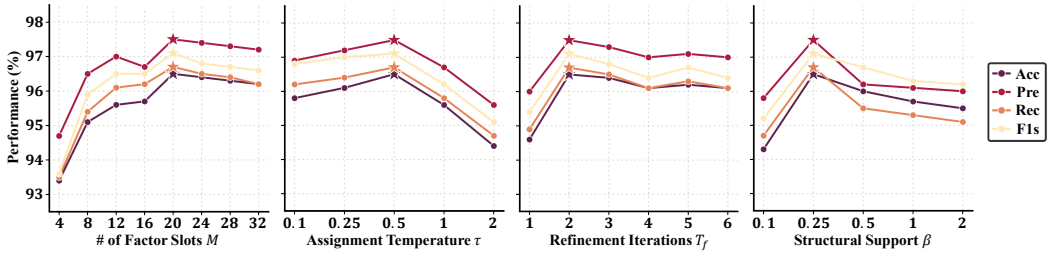


Figure 9: Sensitivity studies of key hyperparameters in HOPE on the ADNI dataset with respect to the number of factor slots M, assignment temperature τ, refinement iterations T_f, and structural support coefficient β. The star marker indicates the configuration achieving the best performance for each metric.

early followed by finer state refinement. Although factors are re-inferred between steps, each fixed-factor update remains dissipative. Thus, this analysis verifies the per-step dissipation guaranteed by Proposition 2, rather than global monotonic descent over the full alternating factor-state trajectory.

Hyper-parameter Sensitivity. We investigate the sensitivity of HOPE to four hyperparameters: the number of factor slots M, assignment temperature τ, refinement iterations T_f, and structural support coefficient β. In Fig. 9, HOPE remains stable across configurations, indicating limited hyperparameter sensitivity. Performance improves with the number of factor slots and peaks around M=20, after which additional slots provide limited gains. The assignment temperature performs best at τ=0.5, suggesting that intermediate granularity balances sparse and shared factor memberships. A small number of refinement updates is sufficient, with T_f=2 achieving the best performance and further refinement providing limited benefits. Finally, β=0.25 performs best, showing that structural information improves factor discovery, while excessive weighting can degrade performance.

Table 11: Ablation study on the structural regularizer of HOPE.

Method	ADNI			ABIDE		
	Acc ↑	Pre ↑	Rec ↑	Acc ↑	Pre ↑	Rec ↑
w/o structural regularization ($\lambda_{\text{str}} = 0$)	$95.2_{\pm 3.1}$	$96.0_{\pm 2.8}$	$94.5_{\pm 2.4}$	$64.2_{\pm 2.4}$	$64.4_{\pm 2.5}$	$64.1_{\pm 2.1}$
HOPE (w/ $\lambda_{\text{str}} = 0.1$)	$96.5_{\pm 2.2}$	$97.5_{\pm 1.6}$	$96.7_{\pm 1.5}$	$65.7_{\pm 2.4}$	$65.7_{\pm 2.3}$	$65.0_{\pm 2.9}$

Effect of Structural Regularization. To further examine the role of the structural regularizer in Eq. 11, we compare the full model with a variant trained using $\lambda_{\text{str}} = 0$. The regularizer encourages ROIs assigned to the same latent factor to remain coherent with the underlying brain connectivity throughout evolution. As shown in Table 11, removing this term consistently degrades performance on both datasets, reducing accuracy from 96.5% to 95.2% on ADNI and from 65.7% to 64.2% on ABIDE, with similar decreases in precision and recall. This suggests that classification supervision alone is insufficient to fully preserve graph-consistent factor organization. By encouraging structural coherence, the regularizer helps the discovered factors better reflect the observed connectome, supporting its complementary role in higher-order factor discovery and predictive performance.

G.2 Higher-Order Factor Interpretation

Visualization of Factor-State Co-Evolution. To qualitatively examine how higher-order organization adapts during state evolution, we visualize representative factor memberships across successive evolution steps for one control (CN) and one AD subject in Figs. 10 and 11, respectively. For the

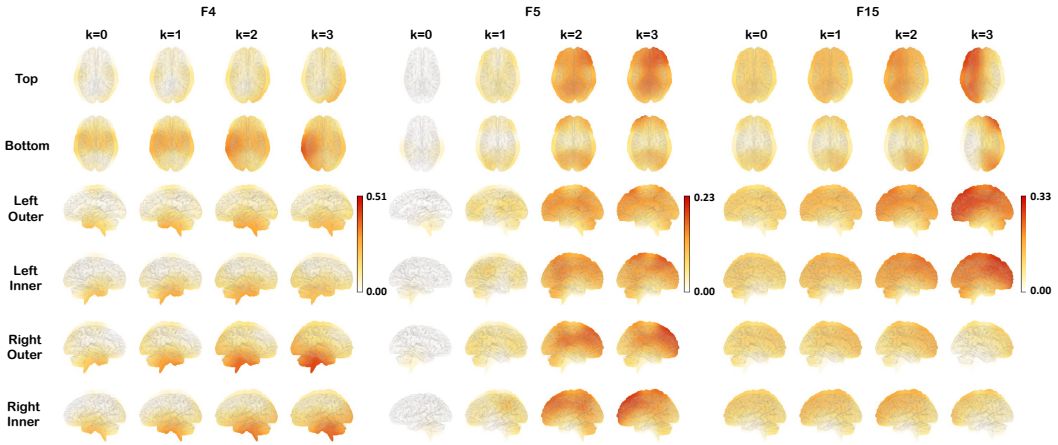


Figure 10: Visualization of factor-membership co-evolution for a representative control (CN) subject *003_S_4441* from ADNI. Factor 4 (F4), Factor 5 (F5), and Factor 15 (F15) are shown across evolution steps $k = 0, \ldots, 3$ from six anatomical views. Color intensity indicates ROI membership strength within each factor.

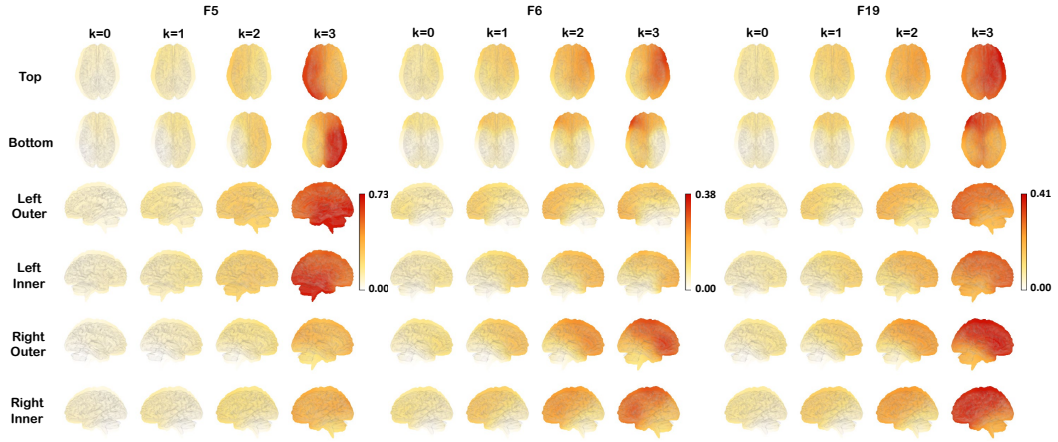


Figure 11: Visualization of factor-membership co-evolution for a representative AD subject *016_S_5057* from ADNI. Factor 5 (F5), Factor 6 (F6), and Factor 19 (F19) are shown across evolution steps $k = 0, \ldots, 3$ from six anatomical views. Color intensity indicates ROI membership strength within each factor.

CN subject, Factor 4 (F4), Factor 5 (F5), and Factor 15 (F15) capture distinct regional organizations and follow different evolution trajectories, with their memberships progressively redistributed and strengthened across cortical regions. Similarly, the AD subject shows evolving organization in Factor 5 (F5), Factor 6 (F6), and Factor 19 (F19), where each factor preserves a distinct spatial role while adapting its regional memberships over successive steps. Importantly, the factors do not remain fixed after initial inference, but co-evolve with the ROI states while maintaining factor-specific organization. These examples support the intended factor-state co-evolution of HOPE, illustrating how distinct higher-order factors adapt their regional composition as the latent brain state evolves.

Higher-Order Structure Analysis. We apply the same analysis used for ADNI in Fig. 3 to PPMI and ABIDE, visualizing the five factors with the highest overall membership strengths across diagnostic groups in Fig. 12. In PPMI, the factor memberships show more pronounced spatial reorganization from CN to PD, with several factors shifting their dominant regional patterns across disease stages. In contrast, ABIDE shows comparatively consistent factor organization between CN and ASD, although several factors exhibit localized changes in their regional memberships. These results show that the learned higher-order factors capture both stable anatomical organization and disease-associated changes in regional participation across heterogeneous brain disorders.

Statistical Validation of ROI Ranking Stability. To assess the reproducibility of ROI patterns identified by HOPE, we performed 20 independent runs with different random seeds, as summarized in Table 12. Factors were first aligned across runs based on their ROI membership profiles, after which ROI rankings were compared for each aligned factor and diagnostic group. Kendall's τ (Sen,

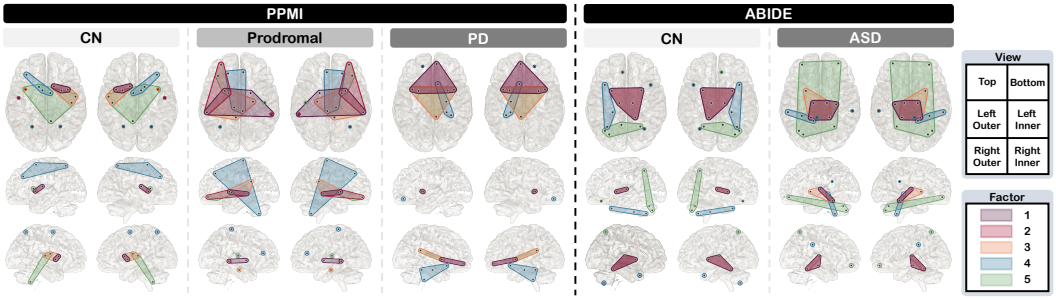


Figure 12: Visualization of the five selected higher-order factors across diagnostic labels on PPMI (Left) and ABIDE (Right). Each factor represents a distinct group of representative ROIs identified by HOPE. The visualizations highlight how factor-specific spatial patterns vary across diagnostic groups and datasets, illustrating the higher-order organization captured by the learned factors.

1968) showed consistent pairwise agreement, with values of 0.650, 0.637, and 0.631 on ADNI, PPMI, and ABIDE, respectively. Kendall's W (Legendre, 2005) further showed overall ranking agreement of 0.794, 0.765, and 0.770, while Top-5 Jaccard similarity (Niwattanakul et al., 2013) reached 0.573, 0.495, and 0.499. These results demonstrate substantial consistency in the overall ROI rankings and strong agreement across runs, while approximately three to four of the Top-5 representative ROIs are repeatedly preserved across independent initializations.

Table 12: Statistical consistency of ROI rankings across 20 independent runs. Kendall's τ measures pairwise rank consistency between runs, Kendall's W quantifies overall agreement across all runs, and Top-5 Jaccard similarity evaluates the reproducibility of the most highly ranked ROIs.

Dataset	Kendall's $\tau \uparrow$	Kendall's $W \uparrow$	Top-5 Jaccard $\uparrow$
ADNI	0.650	0.794	0.573
PPMI	0.637	0.765	0.495
ABIDE	0.631	0.770	0.499

G.3 Complexity and Efficiency

Model Complexity. HOPE consists of three main computational components: 1) structure-aware latent factor discovery requires $O(T_f(NMd + |E|M))$ per evolution step, where $O(NMd)$ arises from the content-affinity and slot updates and $O(|E|M)$ from propagating factor assignments over the graph for structural support; 2) higher-order energy construction requires $O(|E|Md)$, since the factor-wise adjacency in Eq. (6) reweights the observed edges for each of the M factors and Eq. (7) evaluates their state discrepancies over d dimensions; 3) applying the structure-aware mobility to the energy gradient requires $O(|E|d)$ per state update. Since these components are repeated over K evolution steps, the overall complexity is $O\left(K\left[T_f(NMd + |E|M) + |E|Md + |E|d\right]\right)$. Here, N, M, d, $|E|$, K, and T_f denote the number of nodes, factor slots, latent dimensions, observed edges, evolution steps, and factor-refinement iterations, respectively. Since M, d, K, and T_f are small and fixed, the resulting complexity remains linear in the number of nodes and observed edges.

Table 13: Efficiency comparison between HOPE and representative higher-order graph-based methods, measured in terms of parameter count, FLOPs, and average inference time over 5-fold cross-validation.

Efficiency	**Methods**						
	HGNN	UniGCNII	dwHGCN	HyperGT	HyBRiD	MuHL	HOPE
Params (M)	3.28	3.29	3.28	3.29	4.31	5.54	3.32
FLOPs (G)	0.43	0.44	0.43	0.52	0.69	0.88	0.55
Runtime (ms)	1.84	1.52	1.72	97.06	22.13	9.98	9.87

Empirical Efficiency. Table 13 compares efficiency across representative higher-order methods. The factor slots increase parameter counts, while FLOPs mainly arise from alternating factor-state updates and energy-gradient computation at each step. Despite these costs, HOPE runs at 9.87 ms, substantially faster than HyperGT (97.06 ms) and HyBRiD (22.13 ms), and comparable to MuHL (9.98 ms), while constructing higher-order energies to drive state evolution. Although higher-order factor modeling increases memory usage relative to pairwise methods, all methods, including HOPE, run smoothly on a single GPU. Overall, HOPE achieves a favorable trade-off between modeling capacity and runtime efficiency while delivering state-of-the-art performance.

H Discussions on Our Framework

Discussion on Higher-Order Energy Dynamics. Unlike conventional graph and hypergraph models that primarily use pairwise or group-wise structures as aggregation operators, HOPE converts instance-specific latent higher-order factors into factor-wise energies that jointly govern node-state evolution. Because the factors are repeatedly inferred from the evolving states, the interaction structure can adapt throughout inference rather than remaining fixed after its initial construction. Together with the structure-aware mobility operator, this formulation separates the learned energy that drives state evolution from the graph structure that modulates how the states evolve.

Limitations and Future Directions. While HOPE provides a general formulation for energy-driven higher-order graph dynamics, several limitations remain. First, the number of latent factors is predefined for each dataset, although the appropriate granularity may vary across graph instances. Adaptive factor discovery could therefore be explored in future work. Second, the current formulation assumes an undirected, non-negative, fixed input graph to construct the Laplacian and structure-aware mobility. Extending the framework to signed, directed, or dynamic graphs would broaden its applicability. Finally, although the formulation is graph-generic, our empirical evaluation focuses on structural and functional brain networks. Evaluation on other domains with meaningful higher-order interactions would further establish its generality.