BrainKODE: Controlled Koopman Dynamics over Longitudinal Structure–Function Coupling for Early Prediction of Adolescent Substance Use

Badhan Mazumder[1,2], Sir-Lord Wiafe[1,2], Lei Wu[2], Vince D. Calhoun[1,2], and Dong Hye Ye[1,2]

[1] Department of Computer Science, Georgia State University, Atlanta, GA, USA
[2] Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA, USA
dongye@gsu.edu

Abstract. Adolescent substance use initiation (SUI) risk arises from evolving interactions among brain structure, function, and environment, yet most predictive models treat connectivity as static and overlook developmental dynamics. We proposed *BrainKODE* (**K**oopman **O**perator for **D**evelopmental Coupling **E**volution), a geometry-aware deep framework that models longitudinal structure–function coupling as a controlled latent dynamical system. Structural and functional connectomes were aligned within a shared Riemannian space to preserve developmental consistency, while adaptive cross-modal interactions captured network-specific structure–function coupling over time. A temporally constrained Koopman formulation modeled the evolution of latent representations, with contextual factors directly modulating developmental trajectories. On the ABCD cohort, *BrainKODE* improved future SUI prediction over competitive baselines while yielding interpretable maps of time-varying structure–function coupling and contextual influence, highlighting controlled geometric dynamics as a principled approach for forecasting adolescent SUI vulnerability.

Keywords: Structure–Function Coupling · Graph Neural Network · Koopman Operator · Longitudinal Modeling · Substance Use Initiation.

1 Introduction

Adolescence is marked by rapid reorganization of large-scale brain networks that support reward processing, cognitive control, and socio-emotional regulation [1]. This developmental window coincides with heightened vulnerability to substance use initiation (SUI), making early identification of neurodevelopmental risk a critical clinical objective. Functional network connectivity (FNC) from resting-state functional MRI (rs-fMRI) captures coordinated neural activity, while structural connectivity (SC) from diffusion tensor imaging (DTI) reflects the anatomical backbone that constrains information flow [1]. Mounting evidence suggests

that it is not either modality alone, but their evolving interaction, that shapes cognitive and behavioral trajectories [1, 17].

Characterizing this longitudinal structure–function interplay in a principled and predictive manner remains a fundamental challenge. Despite recent advances in graph neural network (GNN) based neuroimaging models, three key limitations persist. First, many approaches [8, 9] treat SC and FNC independently or fuse them statically, failing to model evolving structure-function coupling across development. Second, longitudinal learning is often geometry-agnostic: FNC resides on the symmetric positive-definite (SPD) manifold, yet most methods [11, 18, 20, 14] operate in Euclidean space, distorting intrinsic relationships and weakening cross-visit alignment. Third, contextual factors are typically appended at the classifier level rather than integrated into latent representations, limiting mechanistic insight into how environment shapes developmental trajectories.

To address these gaps, we introduced *BrainKODE* (**K**oopman **O**perator for **D**evelopmental Coupling **E**volution), a geometry-aware deep framework that characterizes longitudinal structure–function coupling as a controlled latent dynamical system. *BrainKODE* aligns structural and functional connectomes within a shared Riemannian tangent space to ensure cross-visit consistency, then constructs a joint multiplex graph with adaptive cross-modal gates to learn network-specific coupling. Developmental trajectories were formulated as a semigroup-constrained Koopman evolution initialized at baseline, enabling stable multi-step forecasting. Contextual variables were incorporated as control inputs that directly modulated latent trajectories, enabling interpretation of their developmental influence. In summary, the main contributions of our work are threefold:

- **Riemannian Longitudinal Structure-Function Coupling:** We proposed a geometry-consistent multiplex graph framework that aligned SC and FNC in a shared tangent space and learns adaptive, network-specific cross-modal coupling over time.
- **Semigroup-Constrained Koopman Dynamics:** We modeled adolescent brain development as a stable continuous-time latent process governed by a Koopman operator, ensuring temporally consistent multi-step evolution from a single baseline initialization.
- **Context-Aligned Controlled Evolution:** We embedded multi-domain contextual variables as structured control inputs within the Koopman evolution, allowing state-dependent modulation of longitudinal brain dynamics instead of post-hoc conditioning.

2 Methodology

Let $\mathcal{C} = \{(\{\mathbf{S}_t^{(i)}, \mathbf{F}_t^{(i)}, \mathbf{u}_t^{(i)}\}_{t\in\mathcal{T}}, y^{(i)})\}_{i=1}^{P}$ denote a cohort of P individuals observed at visits $\mathcal{T} = \{0, 2, 4\}$ years, where $t = 0$ denotes the baseline visit. For subject i, $\mathbf{S}_t^{(i)}, \mathbf{F}_t^{(i)} \in \mathbb{R}^{N\times N}$ represent SC and FNC among N brain networks, respectively. Contextual information comprises static covariates $\mathbf{u}_{\text{stat}}^{(i)} \in \mathbb{R}^{M_s}$ and

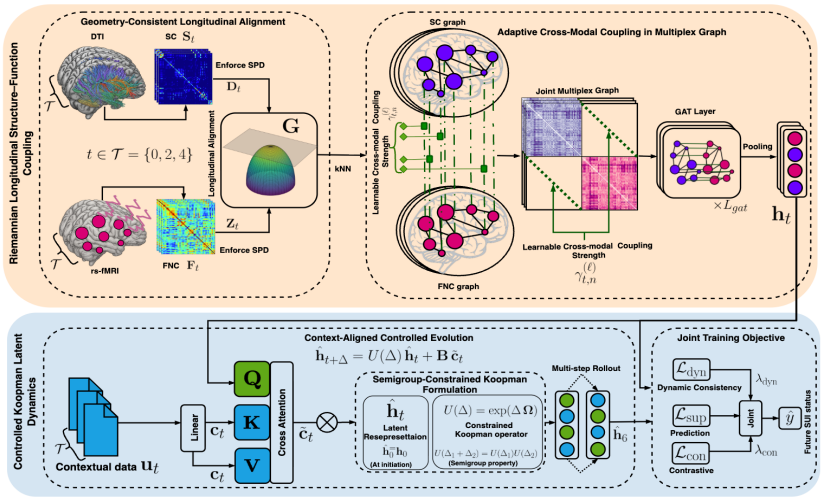

Fig. 1. ***BrainKODE*** **overview:** SC and FNC from baseline (Year-0), Year-2, and Year-4 were aligned in a shared Riemannian space and sparsified into modality-specific graphs. Learnable cross-modal gates modeled structure–function coupling within a joint multiplex graph, where graph attention (GAT) layers learned visit-level embeddings. A semigroup-constrained Koopman operator propagated baseline states forward, with contextual factors modulating the evolution via cross-attention. The controlled rollout yielded the Year-6 representation for SUI prediction under unified end-to-end training.

visit-specific longitudinal variables $\mathbf{u}_{t,\text{long}}^{(i)} \in \mathbb{R}^{M_\ell}$, yielding total dimensionality $M = M_s + M_\ell$ and concatenated input $\mathbf{u}_t^{(i)} = [\mathbf{u}_{\text{stat}}^{(i)} \| \mathbf{u}_{t,\text{long}}^{(i)}] \in \mathbb{R}^M$. The binary label $y^{(i)} \in \{0, 1\}$ indicates SUI status at the future follow-up (Year 6). Our objective is to learn a mapping $(\{\mathbf{S}_t, \mathbf{F}_t, \mathbf{u}_t\}_{t\in\mathcal{T}}) \to \hat{y}$, that generalizes over $\mathcal{C}$ and estimates subject-level future SUI risk by modeling longitudinal structure–function interactions and their contextual modulation as illustrated in Fig. 1.

2.1 Riemannian Longitudinal Structure–Function Coupling

Geometry-Consistent Longitudinal Alignment FNC matrices $\mathbf{F}_t^{(i)}$ are naturally SPD and lie on a Riemannian manifold, where naive Euclidean processing distorts geodesic structure and weakens cross-visit comparability. To enforce longitudinal consistency, we computed a cohort-level reference operator $\mathbf{G} = \text{FM}\Big(\{\mathbf{F}_t^{(i)} \mid i, t\}\Big)$, the Fréchet mean on the SPD manifold across all training subjects and visits within each cross-validation (CV) training fold, which serves as a shared geometric anchor. Each connectome was then mapped to the tangent space at $\mathbf{G}$ as $\mathbf{Z}_t^{(i)} = \log\big(\mathbf{G}^{-1/2}\mathbf{F}_t^{(i)}\mathbf{G}^{-1/2}\big)$, placing all visits in a common geometry-aligned coordinate system centered at the same reference [19]. This anchor ensures that longitudinal differences reflect developmental variation rather

than coordinate mismatch. SC matrices $\mathbf{S}_t^{(i)}$ are not inherently SPD. We therefore constructed diffusion operators $\mathbf{D}_t^{(i)} = \exp(-\chi \mathbf{L}(\mathbf{S}_t^{(i)}))$, which are SPD and allow the same Riemannian treatment [4, 7]. These operators were projected using the identical anchor $\mathbf{G}$, aligning structural–functional representations within the same tangent reference.

Adaptive Cross-Modal Coupling in Multiplex Graph The tangent-space alignment anchored at $\mathbf{G}$ provides a shared geometric reference across modalities and visits. With both $\mathbf{Z}_t^{(i)}$ and $\mathbf{D}_t^{(i)}$ expressed in this common tangent space, we constructed modality-specific graphs at each visit to model adaptive structure-function interaction. For subject i at visit t, the aligned structural diffusion operators and functional representations were sparsified via k-nearest-neighbor (k-NN) graphs retaining the top-k connections per node. After row normalization, this yielded adjacency matrices $\mathbf{A}_t^{S,(i)}, \mathbf{A}_t^{F,(i)} \in \mathbb{R}^{N\times N}$ that preserved local topology in the aligned tangent space. Each brain network n thus maintained two modality-specific embeddings at layer ℓ: $\mathbf{s}_{t,n}^{(i,\ell)} \in \mathbb{R}^d$ (structure) and $\mathbf{f}_{t,n}^{(i,\ell)} \in \mathbb{R}^d$ (function). To model adaptive structure–function interaction, we introduced network-specific cross-modal coupling gates. At layer ℓ, the coupling strength for network n was computed as follows:

$$\gamma_{t,n}^{(i,\ell)} = \sigma\Big(\mathbf{w}_\ell^\top \Big[\mathbf{s}_{t,n}^{(i,\ell)} \| \mathbf{f}_{t,n}^{(i,\ell)} \| \big|\mathbf{s}_{t,n}^{(i,\ell)} - \mathbf{f}_{t,n}^{(i,\ell)}\big|\Big]\Big) \tag{1}$$

where $\mathbf{w}_\ell$ is a learnable vector and $\sigma(\cdot)$ denotes the sigmoid function. These gates define cross-modal edges between homologous structural and functional nodes, yielding the joint multiplex adjacency: $\mathbf{A}_t^{(i,\ell)} = \begin{bmatrix} \mathbf{A}_t^{S,(i)} & \mathrm{diag}(\boldsymbol{\gamma}_t^{(i,\ell)}) \\ \mathrm{diag}(\boldsymbol{\gamma}_t^{(i,\ell)}) & \mathbf{A}_t^{F,(i)} \end{bmatrix}$.

Message passing followed a Graph Attention Network (GAT)-style [21] formulation. Let $\mathbf{z}_{t,p}^{(i,\ell)}$ denote the embedding of node p (structural or functional) in the multiplex graph at layer ℓ. The update rule is defined as $\mathbf{z}_{t,p}^{(i,\ell+1)} = \sum_{q\in\mathcal{N}(p)} \alpha_{pq}^{(i,\ell)} \mathbf{W}_\ell \mathbf{z}_{t,q}^{(i,\ell)}$, where $\mathcal{N}(p)$ denotes the neighborhood of node p under $\mathbf{A}_t^{(i,\ell)}$, $\mathbf{W}_\ell$ is a learnable weight matrix, and $\alpha_{pq}^{(i,\ell)}$ are attention coefficients normalized over $\mathcal{N}(p)$. After L_{gat} layers, node embeddings were aggregated through global average pooling to obtain a visit-level representation $\mathbf{h}_t^{(i)} \in \mathbb{R}^d$, which summarized structure–function coupling at time t.

2.2 Controlled Koopman Latent Dynamics

Semigroup-Constrained Koopman Formulation The pooled embeddings $\mathbf{h}_t^{(i)}$ were treated as latent observables of an underlying nonlinear developmental process and modeled via a shared Koopman operator [12]. Rather than learning independent pairwise transitions, we interpreted $\{\mathbf{h}_t^{(i)}\}_{t\in\mathcal{T}}$ as discrete samples of a single latent trajectory $\{\hat{\mathbf{h}}_t^{(i)}\}_{t\in\mathcal{T}}$ governed by the one-parameter operator

family $U(\Delta) = \exp(\Delta\, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega} \in \mathbb{R}^{d\times d}$ is the infinitesimal generator. The exponential parameterization enforces the semigroup property $U(\Delta_1 + \Delta_2) = U(\Delta_1)U(\Delta_2)$, ensuring temporal consistency across horizons [6]. To guarantee stability, we learned $\boldsymbol{\Omega} = \mathbf{P}\,\mathrm{diag}(-\mathrm{softplus}(\boldsymbol{\psi}))\,\mathbf{P}^{-1}$ via spectral parameterization with trainable $\mathbf{P}$ and $\boldsymbol{\psi}$, constraining eigenvalues to non-positive real parts. This bounds latent trajectory energy for stable forecasting rather than implying biological decay. With fixed spacing $\Delta = 2$ years, evolution follows $\hat{\mathbf{h}}^{(i)}_{t+\Delta} = U(\Delta)\,\hat{\mathbf{h}}^{(i)}_{t}$, initialized by $\hat{\mathbf{h}}^{(i)}_{0} = \mathbf{h}^{(i)}_{0}$, yielding $\hat{\mathbf{h}}^{(i)}_{2} = U(\Delta)\mathbf{h}^{(i)}_{0}$ and $\hat{\mathbf{h}}^{(i)}_{4} = U(\Delta)^2\mathbf{h}^{(i)}_{0}$. The generator was identified by jointly enforcing $\mathbf{h}^{(i)}_{2} \approx U(\Delta)\mathbf{h}^{(i)}_{0}$ and $\mathbf{h}^{(i)}_{4} \approx U(\Delta)^2\mathbf{h}^{(i)}_{0}$, such that the rollout from baseline remained consistent with the encoder-derived embeddings at each observed visit. This also implied $\mathbf{h}^{(i)}_{4} \approx U(\Delta)\mathbf{h}^{(i)}_{2}$, coupling first and second-order temporal structure under the same operator. By constraining evolution with the encoder-derived embeddings at observed visits, $\boldsymbol{\Omega}$ was globally anchored and prevented from latent drift. Its eigenvalues encoded developmental contraction rates, while its eigenvectors captured progression modes consistent with baseline, Year 2, and 4, yielding a unified operator that generalizes to Year 6 forecasting.

Context-Aligned Controlled Evolution The operator $U(\Delta)$ captured intrinsic latent progression governed by $\boldsymbol{\Omega}$. Because adolescent development is influenced by contextual factors, we extended the autonomous semigroup formulation to a controlled linear system where contextual variables acted as exogenous inputs without altering the generator. Given contextual variables $\mathbf{u}^{(i)}_{t} \in \mathbb{R}^{M}$ at visit t, we computed a latent control vector $\mathbf{c}^{(i)}_{t} = \phi_{\mathrm{ctx}}(\mathbf{u}^{(i)}_{t}) \in \mathbb{R}^{d_{\mathrm{ctx}}}$, where $\phi_{\mathrm{ctx}}(\cdot)$ is a learnable projection into a lower-dimensional control space. To enable state-dependent modulation, the control was aligned with the current latent state via cross-attention: $\tilde{\mathbf{c}}^{(i)}_{t} = \mathrm{Attn}\Big(Q = \hat{\mathbf{h}}^{(i)}_{t},\, K = \mathbf{c}^{(i)}_{t},\, V = \mathbf{c}^{(i)}_{t}\Big)$, yielding a context vector conditioned on the current developmental state. Assuming piecewise-constant inputs over $[t, t+\Delta]$, the controlled evolution is as follows:

$$\hat{\mathbf{h}}^{(i)}_{t+\Delta} = U(\Delta)\,\hat{\mathbf{h}}^{(i)}_{t} + \mathbf{B}\,\tilde{\mathbf{c}}^{(i)}_{t} \tag{2}$$

where $\mathbf{B} \in \mathbb{R}^{d\times d_{\mathrm{ctx}}}$ injects contextual influence into the latent dynamics. Here $\boldsymbol{\Omega}$ governs autonomous evolution while $\mathbf{B}$ parameterizes exogenous modulation. Because $U(\Delta)$ was unchanged, the semigroup structure and temporal consistency were preserved; contextual inputs perturbed the trajectory without redefining the identified developmental operator.

Joint Training Objective The framework was trained end-to-end to optimize SUI prediction, dynamical consistency, and latent class separation. Starting from $\mathbf{h}^{(i)}_{0}$, we set $\hat{\mathbf{h}}^{(i)}_{0} = \mathbf{h}^{(i)}_{0}$ and obtained $\hat{\mathbf{h}}^{(i)}_{2}$, $\hat{\mathbf{h}}^{(i)}_{4}$, and the Year 6 forecast $\hat{\mathbf{h}}^{(i)}_{6}$ via iterative application of Eq. (2). A classification head (lightweight MLP) $f_{\mathrm{cls}} : \mathbb{R}^{d} \to \mathbb{R}$ mapped $\hat{\mathbf{h}}^{(i)}_{6}$ to a logit, yielding $\hat{y}^{(i)} = \sigma\big(f_{\mathrm{cls}}(\hat{\mathbf{h}}^{(i)}_{6})\big)$. Binary cross-entropy defines the supervised loss $\mathcal{L}_{\mathrm{sup}} = -\frac{1}{P}\sum_{i=1}^{P}\big[y^{(i)}\log\hat{y}^{(i)} + (1-y^{(i)})\log(1-\hat{y}^{(i)})\big]$. To

enforce temporal consistency, predicted Year-2 and Year-4 states were aligned with observed embeddings through $\mathcal{L}_{\text{dyn}} = \frac{1}{P}\sum_{i=1}^{P}\big(\|\hat{\mathbf{h}}_2^{(i)} - \mathbf{h}_2^{(i)}\|_2^2 + \|\hat{\mathbf{h}}_4^{(i)} - \mathbf{h}_4^{(i)}\|_2^2\big)$. To enhance class separation under imbalance, we introduced a linear projection head $f_{\text{proj}} : \mathbb{R}^d \to \mathbb{R}^{d_c}$ and defined normalized embeddings $\beta^{(i)} = \frac{f_{\text{proj}}(\hat{\mathbf{h}}_6^{(i)})}{\|f_{\text{proj}}(\hat{\mathbf{h}}_6^{(i)})\|_2}$. With positive index set $\gamma(i) = \{j \neq i \mid y^{(j)} = y^{(i)}\}$, supervised contrastive regularization is $\mathcal{L}_{\text{con}} = -\frac{1}{P}\sum_{i=1}^{P}\frac{1}{|\gamma(i)|}\sum_{j\in\gamma(i)}\log\frac{\exp(\beta^{(i)\top}\beta^{(j)}/\tau)}{\sum_{k\neq i}\exp(\beta^{(i)\top}\beta^{(k)}/\tau)}$, where $\tau > 0$ is a temperature parameter [10]. The overall complete training objective balancing all three terms can be formulated as follows:

$$\mathcal{L} = \mathcal{L}_{\text{sup}} + \lambda_{\text{dyn}}\mathcal{L}_{\text{dyn}} + \lambda_{\text{con}}\mathcal{L}_{\text{con}} \tag{3}$$

where λ_{dyn} and λ_{con} weight dynamical alignment and contrastive regularization.

3 Results and Discussion

Dataset and Processing We analyzed the longitudinal Adolescent Brain Cognitive Development (ABCD) [3] dataset including rs-fMRI and DTI data at baseline (Year 0), Year 2, 4 and 6 follow-ups. After quality control and visit matching, the final cohort included 2,487 adolescents (579 SUI by Year-6; 1,908 non-SUI). Contextual data ($M = 10$) were available at each visit and covered demographics (age, household income, caregiver education), family environment (family conflict), neighborhood/environmental risk (deprivation, violent and drug-related crime exposure, lead risk), and behavioral symptoms (CBCL scores).

After rs-fMRI data underwent slice-timing correction, motion realignment, spatial normalization, and smoothing intrinsic connectivity networks ($N = 53$) were extracted via the independent component analysis based *Neuromark* [5] pipeline, and corresponding time courses were band-pass filtered (0.01–0.15 Hz) and z-normalized. FNC matrices (53×53) were computed using normalized Dynamic Time Warping (nDTW) between ICN pairs, with window length determined by the low-frequency cutoff (TR=0.8 s, 110 samples) [15].

For SC construction from DTI, diffusion tensors were estimated using FSL, followed by whole-brain deterministic tractography in CAMINO [16]. Fractional anisotropy map per subject was registered to MNI space via ANTs to warp the *NeuroMark* atlas [5] into native diffusion space; streamlines assigned to atlas regions then yielded 53×53 inter-network fiber-count SC matrices [16].

Experimental Setting All experiments used 5-fold stratified CV; after grid-based hyperparameter search, a unified configuration was fixed across folds. AdamW optimization was used (LR: 3×10^{-4}, weight decay: 5×10^{-4}) with batch size 24 for up to 100 epochs. Graph sparsification used top-$k = 7$, diffusion scale $\chi = 1.0$, and SPD regularization 10^{-4}. A lightweight encoder comprised a single GAT layer (hidden 64, latent $d = 128$, context $d_{\text{ctx}} = 64$, 4 heads) on multiplex graphs. Loss weights were $\lambda_{\text{dyn}} = 0.2$, $\lambda_{\text{con}} = 0.15$, with temperature $\tau = 0.1$.

Table 1. Comparison against representative baselines (Mean ± Standard Deviation).

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
SPDNet [8]	76.54 ± 0.21	68.12 ± 0.84	79.10 ± 0.50
STAGIN [11]	78.73 ± 0.59	74.09 ± 1.26	80.14 ± 0.38
EvolveGCN [18]	75.05 ± 0.17	69.57 ± 1.45	76.71 ± 0.66
RBGM [20]	76.80 ± 0.27	72.19 ± 2.98	78.20 ± 1.21
BNT [9]	79.89 ± 0.31	74.65 ± 1.59	81.48 ± 0.60
NeuroKoop [14]	82.55 ± 2.15	74.78 ± 4.39	84.91 ± 1.48
BrainKODE	**85.64 ± 0.23**†	**85.18 ± 0.37**†	**85.78 ± 0.29**

$^{\dagger}p < 0.05$ vs. NeuroKoop (paired Wilcoxon)

Table 2. Ablation study of *BrainKODE* (Mean ± Standard Deviation).

RQ	Method Variant	Accuracy (%)	Sensitivity (%)	Specificity (%)
RQ1 - Geometry	w/o Riemannian Alignment	76.56 ± 0.19	73.68 ± 1.91	77.43 ± 0.58
	w/o Longitudinal Anchor ($\mathbf{G}$)	81.88 ± 0.16	80.69 ± 0.71	82.24 ± 0.39
	w/o Diffusion Kernel	82.31 ± 0.15	81.07 ± 0.91	82.68 ± 0.43
RQ2 - Coupling	w/o Adaptive Coupling (γ)	81.37 ± 0.05	80.14 ± 0.76	81.74 ± 0.24
RQ3 - Dynamics	w/o Koopman Dynamics ($U(\Delta)$)	77.42 ± 0.03	76.96 ± 0.54	77.56 ± 0.13
	w/o Semigroup Constraint ($\boldsymbol{\Omega}$)	82.50 ± 0.26	81.04 ± 0.65	82.95 ± 0.39
	w/o Dynamic Alignment Loss ($\mathcal{L}_{\text{dyn}}$)	84.13 ± 0.57	83.18 ± 1.06	84.41 ± 0.43
	w/o Supervised Contrastive ($\mathcal{L}_{\text{con}}$)	79.82 ± 0.39	72.12 ± 0.80	82.15 ± 0.27
RQ4 - Context	w/o Context Control ($\mathbf{B}$)	82.65 ± 0.47	80.73 ± 0.61	83.23 ± 0.46
	w/o Context Cross-Attention	83.40 ± 0.18	83.32 ± 0.50	83.43 ± 0.37
Proposed		**85.64 ± 0.23**	**85.18 ± 0.37**	**85.78 ± 0.29**

3.1 Quantitative Evaluation

Baseline comparison: We benchmarked *BrainKODE* against six representative baselines: a manifold network (SPDNet [8]), temporal GNNs (STAGIN [11], EvolveGCN [18], RBGM [20]), a transformer model (BNT [9]), and a Koopman framework (NeuroKoop [14]). All baselines were trained on identical longitudinal multimodal connectomes and contextual variables across visits (Year: 0, 2, 4), with fusion and temporal modeling following their native designs. As shown in Table 1, *BrainKODE* achieved the best overall performance with closely balanced sensitivity and specificity (85.18%, 85.78%). Relative to SPDNet, accuracy and sensitivity improved by +9.1 percentage points (pp) and +17.06 pp, highlighting the benefit of modeling longitudinal structure–function evolution beyond static manifold learning. Temporal GNNs (STAGIN, EvolveGCN, RBGM) captured sequential change but lack geometry-consistent alignment and adaptive cross-modal coupling, resulting in lower sensitivities. Compared with BNT, sensitivity increased by +10.53 pp, and compared with STAGIN by +11.09

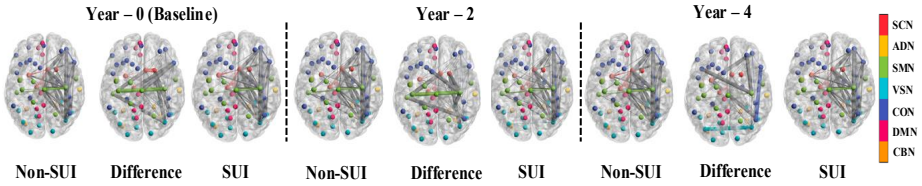

Fig. 2. Axial demonstration of group-averaged structure-function coupling (Non-SUI, SUI, difference) across Year 0, 2 and 4, displaying the top 3% connections by *Neuromark* networks: cerebellar (CBN), default mode (DMN), cognitive control (CON), visual (VSN), sensorimotor (SMN), auditory (ADN), and subcortical (SCN) networks. Node colors denote network membership; intra-network edges follow network colors, cross-network edges are shown in gray, and edge thickness scales with coupling strength.

pp, indicating improved minority detection under class imbalance. Against NeuroKoop, we improved accuracy by +3.09 pp and sensitivity by +10.4 pp, yielding more stable performance. Additionally, leave-one-site-out evaluation across major ABCD sites achieved 83.22% ± 0.11 accuracy, 82.73% ± 0.24 sensitivity, and 83.36% ± 0.10 specificity, supporting robust cross-site generalization.

Ablation study: As shown in Table 2, our ablation addressed four research questions (RQs). *RQ1–Geometry consistency:* Replacing Riemannian projection with Euclidean processing reduced accuracy to 76.56%, and removing the anchor $\mathbf{G}$ or diffusion mapping further degraded performance, underscoring the necessity of geometry-consistent alignment. *RQ2–Adaptive structure–function coupling:* Disabling the multiplex gate (γ) lowered accuracy to 81.37%, demonstrating the benefit of adaptive cross-modal interaction over static fusion. *RQ3–Semigroup Koopman dynamics:* Replacing $U(\Delta)$ with a static mapping reduced accuracy to 77.42%; removing the semigroup constraint ($\boldsymbol{\Omega}$) or dynamic alignment loss weakened temporal coherence, and omitting the contrastive term notably reduced sensitivity to 72.12%. *RQ4–Context-controlled evolution:* Removing contextual injection ($\mathbf{B}$) decreased sensitivity to 80.73%, while replacing cross-attention further weakened performance, highlighting the role of contextual modulation.

3.2 Qualitative Evaluation

Longitudinal structure–function coupling patterns: For each visit, we computed group-level structure-function coupling by averaging cross-modal gate strengths $\gamma_{t,n}^{(i,\ell)}$ within groups, deriving between-group differences, and visualized the top 3% strongest connections in Fig. 2. At baseline, both groups exhibited a CON-centered backbone linking SMN and VSN [13]; however, SUI showed stronger CON–SMN and CON–VSN integration, whereas Non-SUI demonstrated more balanced coupling involving DMN and ADN. By Year-2, SUI displayed enhanced cross-network interactions among CON, VSN, and SCN, while Non-SUI maintained comparatively distributed DMN–SMN coupling [1]. At Year-4, group differences were most pronounced: SUI was marked by strengthened

CON-centered and SCN-related links alongside reduced DMN integration relative to controls [2]. Collectively, these patterns indicate progressively amplified control–salience–sensorimotor coupling in SUI, consistent with altered maturation of executive and affective regulatory systems during adolescence [13, 2, 1].

Contextual significance: We assessed contextual impact using fold-averaged absolute gradient sensitivity of the predicted probability with respect to visit-specific inputs ($|\partial\hat{y}/\partial\mathbf{u}_t|$), which reflects model sensitivity to each covariate under the controlled dynamics. Across visits, the dominant contributors were family conflict and age, followed by socioeconomic indicators (income, caregiver education) and behavioral symptoms, with neighborhood risk factors showing smaller but consistent effects. These patterns suggest that proximal family and psychosocial domains most strongly modulate developmental risk trajectories.

4 Conclusions

We introduced *BrainKODE*, a geometry-aware GNN framework that models longitudinal structure–function coupling as a semigroup-constrained, context-controlled latent dynamical system. By combining Riemannian alignment, adaptive cross-modal coupling, and multi-timepoint-consistent Koopman evolution, it improved future SUI prediction and yielded interpretable maps of network coupling and contextual influence, highlighting the value of modeling adolescent neurodevelopment as a structured dynamical process rather than independent snapshots. Future work will address fixed-interval assumptions, validate on independent longitudinal cohorts, and extend to broader behavioral outcomes.

Disclosure of Interests

The authors confirm that they have no competing interests related to this work.

References

1. Baum, G.L., Cui, Z., Roalf, D.R., Ciric, R., Betzel, R.F., Larsen, B., Cieslak, M., Cook, P.A., Xia, C.H., Moore, T.M., et al.: Development of structure–function coupling in human brain networks during youth. Proceedings of the National Academy of Sciences **117**(1), 771–778 (2020)
2. Casey, B.J., Getz, S., Galvan, A.: The adolescent brain. Developmental review **28**(1), 62–77 (2008)
3. Chaarani, B., Hahn, S., Allgaier, N., Adise, S., Owens, M., Juliano, A., Yuan, D., Loso, H., Ivanciu, A., Albaugh, M., et al.: Baseline brain function in the preadolescents of the abcd study. Nature neuroscience **24**(8), 1176–1186 (2021)
4. Chung, F.R.: Spectral graph theory, vol. 92. American Mathematical Soc. (1997)
5. Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., Salman, M., Abrol, A., Rahaman, M.A., Chen, J., et al.: Neuromark: An automated and adaptive ica based pipeline to identify reproducible fmri markers of brain disorders. NeuroImage: Clinical **28**, 102375 (2020)

6. Engel, K.J., Nagel, R.: One-parameter semigroups for linear evolution equations. Springer (2000)
7. Higham, N.J.: Functions of matrices: theory and computation. SIAM (2008)
8. Huang, Z., Van Gool, L.: A riemannian network for spd matrix learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
9. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. Advances in Neural Information Processing Systems **35**, 25586–25599 (2022)
10. Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., Krishnan, D.: Supervised contrastive learning. Advances in neural information processing systems **33**, 18661–18673 (2020)
11. Kim, B.H., Ye, J.C., Kim, J.J.: Learning dynamic graph representation of brain connectome with spatio-temporal attention. Advances in Neural Information Processing Systems **34**, 4314–4327 (2021)
12. Koopman, B.O.: Hamiltonian systems and transformation in hilbert space. Proceedings of the National Academy of Sciences **17**(5), 315–318 (1931)
13. Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., Chahal, R.: An integrative model of the maturation of cognitive control. Annual review of neuroscience **38**(1), 151–170 (2015)
14. Mazumder, B., Kotoski, A., Calhoun, V.D., Ye, D.H.: Neurokoop: Neural koopman fusion of structural–functional connectomes for identifying prenatal drug exposure in adolescents. In: 2025 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). pp. 1–7. IEEE (2025)
15. Mazumder, B., Wiafe, L., Calhoun, V.D., Ye, D.H.: Neurobridge: Behavior-conditioned koopman dynamics with riemannian alignment for early substance use initiation prediction from longitudinal functional connectome. In: ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 8097–8101. IEEE (2026)
16. Mazumder, B., Wu, L., Wiafe, L., Calhoun, V.D., Ye, D.H.: Kocobrain: Kuramoto-guided graph network for uncovering structure-function coupling in adolescent prenatal drug exposure. In: 2026 IEEE 23rd International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2026)
17. Miller, A.P., Baranger, D.A., Paul, S.E., Hatoum, A.S., Rogers, C., Bogdan, R., Agrawal, A.: Characteristics associated with cannabis use initiation by late childhood and early adolescence in the adolescent brain cognitive development (abcd) study. JAMA pediatrics **177**(8), 861–863 (2023)
18. Pareja, A., Domeniconi, G., Chen, J., Ma, T., Suzumura, T., Kanezashi, H., Kaler, T., Schardl, T., Leiserson, C.: Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 5363–5370 (2020)
19. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. International Journal of computer vision **66**(1), 41–66 (2006)
20. Tekin, A., Nebli, A., Rekik, I.: Recurrent brain graph mapper for predicting time-dependent brain graph evaluation trajectory. In: MICCAI Workshop on Domain Adaptation and Representation Transfer. pp. 180–190. Springer (2021)
21. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)