Learning Continuous-Time Trajectories of Adolescent Brain Connectomes via Latent Graph Neural ODEs

Executive Summary

Recent advances in **Graph Neural Ordinary Differential Equations (Graph ODEs)** enable continuous-time modeling of graph-structured data, offering a promising solution for irregularly sampled longitudinal connectomes. Traditional dynamic GNNs often assume a fixed graph structure and focus on evolving node features; however, state-of-the-art methods now tackle **evolving adjacency (“edge dynamics”)**. For example, **Coupled Graph ODE (CG-ODE)** learns two coupled ODE functions (one for node states, one for edge weights) to model graphs that gain or lose connections over time [1] [2]. Similarly, **MTGODE** (2023) infers missing or time-varying graph structure by integrating graph learning into a neural ODE, rather than assuming a static adjacency [3] [4]. These continuous-time graph generative frameworks address irregular intervals by naturally interpolating node and edge trajectories between observations [5], which is ideal for the Adolescent Brain Cognitive Development (ABCD) cohort where follow-up scan times vary.

In the neuroimaging domain, most prior works use brain connectivity graphs for **graph regression or classification** (e.g. predicting age or diagnosis) rather than generating future connectivity matrices. Graph neural nets have shown improved accuracy in tasks like sex classification on pediatric connectomes [6], but **full graph generation** (predicting a future $N\times N$ connectivity matrix) remains under-explored. A notable exception is **FLAT-Net (2021)**, a few-shot generative approach that predicts a subject’s future brain network from a single baseline graph [7]. FLAT-Net clusters subjects and trains a geometric GAN per cluster to evolve a “connectional template” over time [8], demonstrating feasibility of connectome generation even with scarce data. More recently, Rekik et al. (2024) proposed a **Federated Multi-trajectory GNN** (FedGmTE-Net++) to predict infant brain connectivity development; it uses an auxiliary regularizer to exploit longitudinal data and a two-step KNN+regression imputation to handle missing time-points [9] [10]. These studies highlight the gap: **no existing method yet models individual brain connectome trajectories in a continuous latent space**. Our project will fill this gap by marrying **variational graph autoencoders with Neural ODEs**, i.e. encoding each subject’s connectome into a latent state and using a graph-ODE to evolve that state forward in time. Such a **latent Graph ODE** approach (inspired by LG-ODE [11]) can capture nonlinear developmental trends (e.g. synaptic pruning manifesting as edge weight decreases) in a subject-specific manner, while naturally handling irregular visits and small time increments.

Another frontier is **multi-modal integration**. Modern studies increasingly fuse brain connectivity with regional morphology, as we intend (using cortical thickness/volume as node features alongside structural connectivity as edges). One approach is to input node-wise MRI features into the GNN and let message-passing occur over the structural connectome [12] [13]. For example, a recent multimodal GNN framework for ADNI data represented each brain region as a node with T1 features (cortical thickness/volume) and edges from DTI (structural) and fMRI (functional) connectivity, then employed attention to fuse modalities [14] [15]. In our context, we will treat morphology as node attributes and structural fiber counts as weighted edges, enabling the model to learn how anatomical properties influence connectivity change. Continuous-time modeling provides further advantage here: rather than imputing intermediate data, the Graph ODE can be conditioned on time (age in months) as a continuous variable, thus directly handling ABCD' s uneven visit schedule. (Notably, prior ABCD studies often resort to discrete time models or simple interpolation. For instance, one MICCAI 2023 method predicted 2-year connectome changes on ABCD data with a graph MLP, but it operated on fixed time intervals [16].) By contrast, our continuous latent trajectory approach can predict connectivity at arbitrary ages, a significant innovation for personalized neurodevelopmental forecasting.

In summary, **state-of-the-art graph generative models** (like diffusion-based graph generators and graph ODEs) are just beginning to be applied to brain networks. Diffusion models have excelled at generating small graphs (e.g. molecules) by learning the distribution of graph structures; for brain-scale graphs, a recent NeurIPS 2025 work introduced a topology-aware graph diffusion (TAGG) that incorporates persistent homology to preserve global network properties [17] [18]. This indicates a trend toward **imposing biological realism constraints** in graph generation. Our proposed research rides on these advances: we will use **Graph Neural ODEs** to capture continuous connectome dynamics and include domain-inspired constraints (sparsity, small-world topology, temporal smoothness) to ensure generated brain networks remain realistic. This fills a timely gap by providing the first continuous-time generative model for individual adolescent structural connectomes, leveraging GNN-based ODEs to respect both the **irregular timing** and the **complex topology** of brain development.

Key Literature Table

Title	Year	Methodology	Dataset	Relevance to Project
Learning Continuous System Dynamics from Irregularly-Sampled Partial Observations (LG-ODE) [19]	2020 (NeurIPS)	**Latent Graph ODE**: VAE-based encoder + neural ODE for continuous node dynamics on a fixed graph.	Physics simulations (motion capture, springs)	Introduced latent ODE on graphs for irregular time; inspires using continuous latent trajectories for connectomes [20]. Combines VAE with ODE, hinting at our latent space approach [21].
Coupled Graph ODE for Learning Interacting System Dynamics (CG-ODE) [22] [1]	2021 (KDD)	**Coupled Node-Edge ODEs**: Two intertwined ODEs model node feature evolution and edge weight evolution continuously.	COVID-19 regional data; synthetic social network	State-of-the-art continuous **graph structure** evolution [2]. Validates that edges can be treated as dynamic variables and learned via ODE – a key technique we will adapt for changing connectomes.

Title	Year	Methodology	Dataset	Relevance to Project
Multivariate Time Series Forecasting with Dynamic Graph Neural ODEs (MTGODE) [3]	2022 (TKDE)	**Graph ODE Forecasting**: Learns time-varying graph structure (no fixed adjacency) jointly with node dynamics in an ODE framework.	Traffic & energy time-series benchmarks (5 datasets)	Shows how to infer **unknown or evolving connectivity** in continuous time [4]. We draw on its idea of treating adjacency as a latent function of time, given brain connections may strengthen or weaken during development.
FLAT-Net: Longitudinal Brain Graph Evolution Prediction [7] [23]	2021 (MICCAI Workshop)	**Few-Shot Generative GAN**: Clusters baseline brain networks and trains a geometric GAN per cluster to predict the evolved connectivity at follow-up.	Longitudinal connectomic datasets (e.g. disease vs healthy; few subjects per cluster)	Pioneering **brain graph generation** from a single time-point. Demonstrates data-efficient training via clustering. Highlights need for preserving network traits during generation (each GAN learns a template’s evolution) [7].
FedGmTE-Net++: Federated Multi-Trajectory GNNs for Infant Connectivity [9]	2025 (MedIA)	**Federated GNN + Imputation**: Predicts multi-modal connectivity trajectories (T1, DTI, etc.) using GNNs; employs KNN-based pre-completion and regression refinement to handle missing timepoints.	dHCP infant brain MRI (birth to 1 year), multi-site (federated)	Addresses **longitudinal brain graph prediction under scarce data**. Uses an auxiliary loss to maximize use of all timepoints and explicit imputation [9], underscoring the importance of handling irregular visits (which our continuous model will inherently solve).

Title	Year	Methodology	Dataset	Relevance to Project
Mixing Temporal Graphs with MLP for Longitudinal Connectome Analysis [16]	2023 (MICCAI)	**Spatio-temporal GNN/MLP Hybrid**: Learns separate spatial and temporal projections of brain graphs, fusing them to predict gradual connectome changes.	ABCD Study + ADNI (longitudinal structural connectomes)	One of the first GNN approaches on ABCD longitudinal data. Achieved state-of-the-art accuracy in predicting **connectome changes** over 2-year intervals [16] . We improve upon this by modeling changes continuously (not just at fixed intervals) and by generating whole graphs rather than just predicting certain metrics.
Topology-aware Graph Diffusion (TAGG) [17] [18]	2025 (NeurIPS)	**Diffusion Generative Model + TDA**: Score-based graph generation augmented with persistent homology. Adds a persistence diagram loss and topology-aware attention to ensure generated graphs match global topological features of training graphs.	Graph benchmarks (Ego, Grid) + **Neuroimaging (brain networks)** [24]	Sets new SOTA in **graph generation quality**, emphasizing topological fidelity. Notably applied to brain networks, showing that preserving properties like connectivity patterns and loops improves realism [25] . Inspires us to integrate topological constraints (e.g. degree distribution, small-worldness) into our loss to generate biologically plausible connectomes.

Title	Year	Methodology	Dataset	Relevance to Project
Brain Multigraph Prediction using Topology-aware GAN (TopoGAN) [26]	2021 (MedIA)	**Topology-aware Graph GAN**: Predicts multiple target brain graphs (e.g. multi-modal connectomes) from one source graph, with a discriminator that evaluates local topological features (node centrality, clustering, etc.) to enforce realism.	Multi-modal brain networks (e.g. structural->functional connectivity translation tasks)	Illustrates how **topology constraints** (clustering coefficients, hub connectivity) can be enforced in generation. Reported better preservation of node-level topology than naive graph translators. This guides our use of regularizers for preserving graph metrics (e.g. degree distribution) in longitudinal generation.

(Note: All above works are from 2021–2025 in top-tier venues or journals, ensuring a focus on current state-of-the-art techniques.)

Technical Recommendations

- **Latent Graph ODE Model for Edges and Nodes:** We recommend a **variational latent Graph ODE** framework that captures both node attribute dynamics and edge evolution. For example, extend CG-ODE [1] by introducing a **graph variational autoencoder**: an encoder GNN will infer each subject' s latent state $z(t_0)$ from their baseline connectome (and covariates like sex), then a **coupled ODE** $d[z_X, z_E]/dt = F(z_X, z_E; \theta)$ in latent space will govern the joint evolution of node embeddings (z_X) and edge factors (z_E). A decoder GNN can map the latent state at time t back to a predicted structural connectivity matrix. This approach combines the strength of **Neural ODEs for irregular timing** (continuous trajectories [20]) with **VAE-style distribution learning**, to model uncertainty in individual developmental paths. Crucially, by treating adjacency as a latent variable, the model can **add or remove connections smoothly over time** (e.g. simulate strengthening of long-range tracts or weakening of redundant links during pruning) rather than assuming a fixed graph.
- **Incorporate Multi-Modal Node Features:** Leverage the ABCD multimodal data by enriching the node representation. Each brain region (node) at time t can have a feature vector concatenating its cortical thickness, volume, etc., possibly standardized by age. These features enter the ODE as part of the state, so the **node ODE** $dX/dt = f(X, A)$ can learn how morphological changes drive connectivity changes. We may use a two-stream approach: one GNN to update node features and another (or another channel) to update edges (as in CG-ODE) [1]. If certain modalities (e.g. fMRI connectivity or other parcellations) are available at some timepoints and not others, consider a **multi-trajectory loss** (as in FedGmTE-Net++) to make use of all available data [27]. Another idea is to employ an **attention-based fusion**: e.g. use structural connectivity for message passing but modulate messages by the similarity of node morphological features, or vice versa [14] [28], thereby allowing the model to learn cross-talk between structure and morphology in driving network development.

- **Continuous-Time Training and Irregular Sampling:** Exploit the continuous nature of the model to handle ABCD's irregular visits. Instead of imputation, we will train the ODE by backpropagating through a solver (e.g. Dormand-Prince) that can integrate the latent state from one observed time to the next. The loss will sum reconstruction errors at each available timepoint per subject. This way, each subject provides supervision at their actual ages, and the ODE naturally **interpolates or extrapolates** between observations [20]. We should prefer **adaptive-step ODE solvers** to efficiently handle varying time gaps – large gaps will use bigger integration steps, small gaps with dense data will use finer steps. Techniques like **Graph Neural Controlled Differential Equations** (Graph NCDE) or **neural flow matching** could also be explored for stability, but a standard ODE approach is a good starting point given our focus on smooth trajectories.
- **Topology-Preserving Loss Functions:** To ensure generated connectomes remain biologically plausible, we suggest adding **regularization terms that enforce known network properties**. One is a **sparsity loss**: e.g. an L_1 penalty on edge weights or a soft penalty encouraging a target distribution of edge densities. Since brain structural networks are sparse (~10–20% density), we could dynamically threshold small weights to zero during training (as done in diffusion models to preserve sparsity at each generation step [29]). Another is a **topology matching loss**: compare global graph metrics of the generated graph to real ones. For example, use a **Wasserstein distance between degree distributions** or clustering coefficient distributions in the batch. Recent work explicitly matched persistent homology descriptors of graphs [18]; implementing a **Persistence Diagram Matching loss** would encourage correct counts of connected components and loops (which relate to network segregation and cycles) in the generated graphs. Simpler, we can compute the difference in **small-worldness** (ratio of clustering to path length) between generated and real networks and penalize deviations, thereby encouraging the generator to produce small-world like topology. These additions will guide the model to utilize biologically realistic configurations (e.g. preserving hub nodes and community structure) rather than merely fitting edge-by-edge error.
- **Temporal Smoothness Constraint:** Although the ODE inherently produces smooth continuous trajectories, we may still enforce **temporal smoothness** in the learned dynamics to avoid any oscillatory or noisy solutions that fit the data but are biologically implausible (brain connectivity changes are gradual). In practice, we can add a regularizer $R = \lambda \sum_t |A(t+\Delta)-A(t)|^2$ on consecutive generated adjacencies for small Δ (or a penalty on the time-derivative norm $|dA/dt|^2$ directly). This discourages sudden jumps in connectivity from one age to the next. In graph link prediction literature, such temporal smoothness losses have been used to stabilize dynamics [30]. For our continuous model, this could mean penalizing large ODE vector field values that cause rapid change. By tuning this regularizer, we can find a balance where the model is flexible enough to capture individual deviations but still yields **gradual, biophysically plausible evolution** of the connectome.

Potential Pitfalls and Mitigations

- **High-Dimensional State Space (Scalability):** A full structural connectome (e.g. 360 nodes as in Glasser atlas) has ~$360^2 \approx 130k$ possible connections. Modeling each edge as a separate ODE state would be intractable. This is a known challenge – even state-of-the-art graph ODE models usually restrict to smaller graphs or assume known sparsity structure [31]. If we naively include every edge in the latent state, the ODE integration may become prohibitively slow and memory-intensive. **Mitigation:** We will project the graph to a **lower-dimensional latent representation**. For example, use a **low-rank factorization** of the adjacency (learning latent embeddings for nodes such that $A \approx Z Z^T$) and then evolve those embeddings with an ODE, rather than evolving each entry of A individually. This leverages the fact that brain networks have strong community structure and are not fully random – a few latent factors might capture most variability. Alternatively, focus on **edge features for existing edges only**: since the structural connectome is sparse, treat the edge list as a variable-length state (which neural ODE frameworks like torchdiffeq can handle by message-passing form). Using graph sparsity in computations (sparse matrix operations) will also help scale to 300+ nodes.

- **Training Instability and Convergence:** Fitting neural ODEs can be tricky – issues include **gradients through ODE solvers** that are slow or inaccurate, and the risk of the ODE solution diverging if the learned vector field F is too "stiff" or drives explosive growth. Graph ODEs add additional complexity because the GNN inside F could cause **oversmoothing** (node embeddings converging to identical values) or numerical stiffness. **Mitigation:** We will adopt techniques from prior works: e.g. the **GraphCON** method adds skip connections and damping in the ODE to counter oversmoothing. We can incorporate a **reversible Heun or midpoint solver** which is more stable for stiff equations. We will also monitor the **Lipschitz constant** of the ODE function (perhaps by spectral norm regularization on the GNN weights) to ensure the dynamics aren’t too chaotic. Curriculum training – start by modeling shorter time spans or downsampled graphs – might ease the ODE into a stable regime. If gradients through the solver are too noisy, we can explore the **adjoint method** or backpropagation with checkpointing to trade speed for stability. In case Neural ODE training proves unreliable, a backup is to discretize time into small steps (e.g. 6-month increments) and train a sequential graph model (like an RNN) with weight-sharing to approximate the continuous solution; this can later be fine-tuned or analytically linked to an ODE.
- **Subject Variability vs. Data Volume:** ABCD has a large sample (~10k subjects) but each subject has only a few timepoints (so trajectories are short). There is a risk the model might **overfit individual trajectories** or conversely **underfit (learn only population-average trend)**. Learning a distribution of personal trajectories is inherently hard when each individual provides limited longitudinal data. **Mitigation:** We will incorporate **random effects or hierarchical modeling** in the latent ODE. For instance, give the model a population-level ODE F_{θ} but allow each subject i to have a slight deviation in parameters θ_i drawn from a shared distribution. This could be implemented via meta-learning or by encoding subject-specific information (like baseline connectome and covariates) into the initial latent state that the ODE uses. The variational framework will help: by sampling latent initial states for each subject and regularizing them, we encourage the model to explain variation through $z(t_0)$ (the initial brain state) rather than learning completely separate dynamics per subject. We will also perform extensive **cross-validation**: e.g. train on 80% of subjects and evaluate the ability to predict held-out subjects’ follow-up connectomes. If the model is too rigid (can’t fit individuals) or too flexible (overfits seen individuals but fails on new ones), we may adjust the latent dimensionality or regularization strength.
- **Interpretability and Topological Constraints:** There’s a danger that the model learns to predict edges without regard for biological plausibility – e.g. creating unrealistically dense or implausibly structured graphs that still minimize MSE loss. While we plan to enforce some topological losses, choosing the right constraints is non-trivial. Too heavy a constraint (say forcing exact degree distribution match) could harm the fit, while too light allows implausible graphs. Also, metrics like small-worldness or persistent homology are global and may not directly penalize subtle local errors (like a missing hub connection). **Mitigation:** We will validate intermediate and final outputs against known brain network properties. For example, we’ll check if the generated graphs have realistic **degree sequences, clustering coefficients, and modular structure**. If we notice deviations (e.g. the model starts to create overly connected graphs), we can adapt the loss – for instance, increasing the weight on sparsity or adding a penalty for violating known scaling laws (like the degree distribution in structural brain networks typically follows a skewed but not extreme distribution). We might incorporate a **post-processing step**: after generating a probabilistic connectivity matrix, apply a sparsification (like keeping top K edges per node or a density-based threshold) to enforce a reasonable sparsity level, since the real brain’s wiring cost

constrains it. As long as this is done consistently, it can be integrated into the training loop (using a differentiable approximation of thresholding, e.g. sigmoid). Additionally, engaging neuroscientific expertise to define acceptable ranges for properties (e.g. global efficiency, modularity) will guide us in tuning these constraints. In summary, an iterative approach of "generate – evaluate topology – adjust loss" will be used to ensure the learned trajectories are not only statistically sound but also neuroscientifically plausible.

- **Generality to Different Parcellations:** The choice of brain parcellation (ROI definition) can impact model performance. Our plan using the Glasser 360 atlas yields a moderately large graph; if we move to finer parcellations (e.g. 1000+ nodes) the graph becomes huge, whereas coarser ones (e.g. 68-node Desikan atlas) might lose important detail. There is a risk that a model tuned for one scale struggles on another. **Mitigation:** We will **include multiple parcellation resolutions** in our analysis. This means training and testing the model on, say, both a 200-node and a 360-node atlas to see if the method is robust. The latent ODE approach should in principle handle both, but computational adjustments (like using sparser methods for 1000-node graphs) may be needed. If the model does not scale to a high-resolution parcellation, one strategy is to **learn at a moderate resolution and then upsample** the generated graph to a finer resolution using a secondary model (perhaps leveraging known correspondences between atlases). Conversely, to ensure we're not missing effects due to over-compression, we might first validate on a simpler 68-node atlas (where conventional methods could also be applied for sanity check) before moving to 360. By comparing outcomes, we can identify if any failure modes are simply due to parcellation issues.

In conclusion, while training a Graph Neural ODE on large, longitudinal brain networks presents challenges, our careful design – combining the latest continuous-time graph generation techniques with domain-specific constraints – and our mitigation strategies – dimensionality reduction, regularization, and multi-scale testing – will together maximize the likelihood of a successful model that predicts individual adolescent connectome trajectories both **accurately and realistically**. With this foundation, our proposal stands to advance the state-of-the-art in developmental connectomics modeling, filling a crucial gap between data-driven graph learning and neuroscientific validity.

---

[1] [22] Coupled Graph ODE for Learning Interacting System Dynamics
https://web.cs.ucla.edu/~yzsun/papers/2021_KDD_CG_ODE.pdf

[2] [5] [19] [20] Graph ODEs and Beyond: A Comprehensive Survey on Integrating Differential Equations with Graph Neural Networks
https://www.cs.emory.edu/~jyang71/files/graphode.pdf

[3] [4] [2202.08408] Multivariate Time Series Forecasting with Dynamic Graph Neural ODEs
https://arxiv.org/abs/2202.08408

[6] Graph Neural Network Learning on the Pediatric Structural Connectome - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC11861995/

[7] [8] [23] (PDF) FLAT-Net: Longitudinal Brain Graph Evolution Prediction from a Few Training Representative Templates
https://www.researchgate.net/publication/354875262_FLAT-Net_Longitudinal_Brain_Graph_Evolution_Prediction_from_a_Few_Training_Representative_Templates

[9] [10] [27] [2401.01383] Predicting Infant Brain Connectivity with Federated Multi-Trajectory GNNs using Scarce Data
https://arxiv.org/abs/2401.01383

11 21 [2011.03880] Learning Continuous System Dynamics from Irregularly-Sampled Partial Observations
https://arxiv.org/abs/2011.03880

12 13 14 15 28 Multimodal Graph Neural Networks for Prognostic Modeling of Brain Network Reorganization
https://arxiv.org/html/2512.06303v1

16 Mixing Temporal Graphs with MLP for Longitudinal Brain Connectome Analysis | MICCAI 2023 - Accepted Papers, Reviews, Author Feedback
https://conferences.miccai.org/2023/papers/425-Paper3659.html

17 18 24 25 29 31 openreview.net
https://openreview.net/pdf/72bdeb175c74739bd71db96679700189412c2573.pdf

26 Brain multigraph prediction using topology-aware adversarial graph neural network - PubMed
https://pubmed.ncbi.nlm.nih.gov/34004494/

30 [PDF] Temporal Smoothness Regularisers for Neural Link Predictors - arXiv
https://arxiv.org/pdf/2309.09045