DYNAPIN: MECHANISM-PLURALISTIC DYNAMICS PROFILING FOR INTERPRETABLE NODE CLASSIFICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Node classification is a core graph-learning task for citation, hyperlink, purchase, and political-interaction networks, where accurate predictions are valuable only when the learned graph behavior can also be read. Existing graph neural and diffusion models improve message passing, scalability, uncertainty modeling, and non-homophilous learning, but their embeddings rarely state the mechanism by which a node responds to graph-mediated influence. Inspired by recent progress in neural symbolic regression, we introduce DynaPIN (**Dyna**mics-informed **p**rofile learning for **i**nterpretable **n**ode classification), a mechanism-pluralistic framework that uses opinion-dynamics equations as a symbolic hypothesis space for graph classification. DynaPIN constructs graph-induced opinion evolution signals, fits each node with a Candidate Mechanism Bank covering neighbor diffusion, self-retention, selective interaction, and attraction–repulsion, and converts fitted parameters, validation errors, and Soft Mechanism Evidence into a Mechanism Fingerprint. The fingerprint supports downstream classification, while the same fitted quantities instantiate a Soft Equation Lens for mechanism discovery. Experiments on three opinion-oriented voting datasets and five general node-classification benchmarks show that DynaPIN remains competitive on voting graphs and achieves leading or near-leading Macro-F1 on citation, co-purchase, and hyperlink benchmarks. More importantly, the Soft Equation Lens reveals a repeated non-consensus grammar: fitted graph responses concentrate on selective nonlinear coupling and attraction–repulsion-like patterns, with dataset-specific mixtures rather than one universal smoothing law. These results establish opinion dynamics as a predictive and equation-readable mechanism language for interpretable graph representation learning.

Keywords: Node classification; Opinion dynamics; Neural symbolic regression; Mechanism-pluralistic dynamics profiling; Interpretable graph learning.

1 INTRODUCTION

Node classification assigns labels to entities connected by citations, hyperlinks, purchases, votes, or social relations. It supports collective classification in relational data, where labels and relations provide mutually informative context (Sen et al., 2008). Citation and web graphs make this task central to knowledge organization and topical prediction (Namata et al., 2012; Mernyei & Cangea, 2020), while political voting graphs connect relational structure with political positions or affiliations (Voteview, 2026). These applications require more than accurate labels: a graph learner should also expose how attributes and relations generate the evidence behind a prediction.

Graph representation learning has raised the predictive ceiling through message passing, graph transformers, and adaptive propagation. Modern graph neural networks replace hand-crafted relational features with trainable neighborhood aggregation (Khoshraftar & An, 2024). Graph transformers improve long-range representation and scalable attention-style modeling (Rampasek et al., 2022), and tokenized variants adapt this design to large-scale node classification (Chen et al., 2023). Recent heterophily studies show that useful graph learning cannot be reduced to uniform neighbor

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.

averaging (Luan et al., 2022; Platonov et al., 2023). These predictors are powerful, but their learned coordinates rarely state how a node responds to graph-mediated influence.

A dynamic view addresses part of this gap by treating node representations as evolving states. Neural sheaf diffusion connects graph diffusion with geometry-aware evolution under heterophily and oversmoothing (Bodnar et al., 2022). Optimization-induced nonlinear diffusion formalizes graph representations through implicit state evolution (Chen et al., 2022). Reaction–diffusion and stochastic diffusion networks enrich this view with propagation and uncertainty structures (Choi et al., 2023; Lin et al., 2024). These models use dynamics as a representation engine; the response rule behind a node remains hard to read as an equation.

This limitation matters because graph responses are not governed by one universal smoothing law. Diffusion-based graph models expose oversmoothing as a repeated-propagation risk (Bodnar et al., 2022; Choi et al., 2023), and heterophilous graphs violate the assumption that useful signals always become smoother under neighbor aggregation (Luan et al., 2022; Platonov et al., 2023). Homophily itself depends on the task and model design (Ma et al., 2022). Existing graph explainability studies reveal influential features, subgraphs, masks, or decision effects around trained predictors (Yuan et al., 2023; Amara et al., 2022). Evaluation studies and decision-boundary explainers strengthen this literature (Agarwal et al., 2022; Wang & Shen, 2024), but they do not provide a compact equation-level description of the node's graph-mediated response.

Neural symbolic regression provides the bridge from prediction to equation readout. Symbolic models discovered from neural representations show that predictive learning can expose compact governing equations (Cranmer et al., 2020). Transformer-based symbolic regression improves equation search at scale (Kamienny et al., 2022). Recent work extends neural symbolic regression to interacting nodes and graph-coupled evolution (Hu et al., 2025; Yu et al., 2026). DynaPIN transfers this equation-centered view to node classification by using opinion-dynamics equations as a structured symbolic mechanism bank and learning node-wise evidence over that bank.

Equation-driven graph learning already shows the value of dynamical laws, but most designs commit to one governing family. Total-variation graph networks produce sharp vertex clusters through a fixed relaxation (Hansen & Bianchi, 2023). Reaction–diffusion networks fix a reaction–diffusion backbone (Choi et al., 2023), and stochastic graph diffusion models use a diffusion equation for uncertainty-aware classification (Lin et al., 2024). Opinion-dynamics methods face the same issue when a fixed FJ law or bounded-confidence rule is used as the modeling basis (Sun & Zhang, 2023; Bernardo et al., 2024). DynaPIN adopts a pluralistic alternative: multiple opinion-dynamics mechanisms are fitted as diagnostic probes, and node-level evidence decides which response equation explains each local trajectory.

Opinion dynamics is well-suited as a mechanism language because it separates response modes that generic diffusion merges together. Consensus dynamics names linear neighbor alignment (DeGroot, 1974); the Friedkin–Johnsen model adds self-retention (Friedkin & Johnsen, 1990). Bounded confidence captures selective interaction under state distance (Hegselmann & Krause, 2002), and attraction–repulsion models encode nonlinear agreement and disagreement forces (Cornacchia et al., 2020). These mechanisms correspond to alignment, anchoring, selective coupling, and opposition, which ordinary graph classifiers hide inside embeddings.

DynaPIN builds this probe into a classification framework. Graph-Induced Opinion Evolution turns static attributes and topology into label-free response trajectories. Mechanism-Pluralistic Dynamics Profiling fits every node against a Candidate Mechanism Bank, producing fitted parameters, validation errors, and Soft Mechanism Evidence. These quantities form a Mechanism Fingerprint for downstream prediction and a Soft Equation Lens for mechanism discovery:

$$\dot{\mathbf{x}}_i = \sum_{m \in \mathcal{M}} s_{i,m} F_m(\mathbf{x}, \mathbf{A}; \theta_{i,m}), \tag{1}$$

where m indexes one mechanism in the Candidate Mechanism Bank $\mathcal{M}$, F_m is its dynamics function, $s_{i,m}$ is the Mechanism Evidence of node i, and $\theta_{i,m}$ is the corresponding fitted parameter set. Prediction and interpretation are source-aligned: the classifier uses the Mechanism Fingerprint, and the lens reads the fitted equations that generate it.

The paper makes three contributions. DynaPIN introduces Graph-Induced Opinion Evolution and Mechanism-Pluralistic Dynamics Profiling, connecting graph classification with neural-symbolic

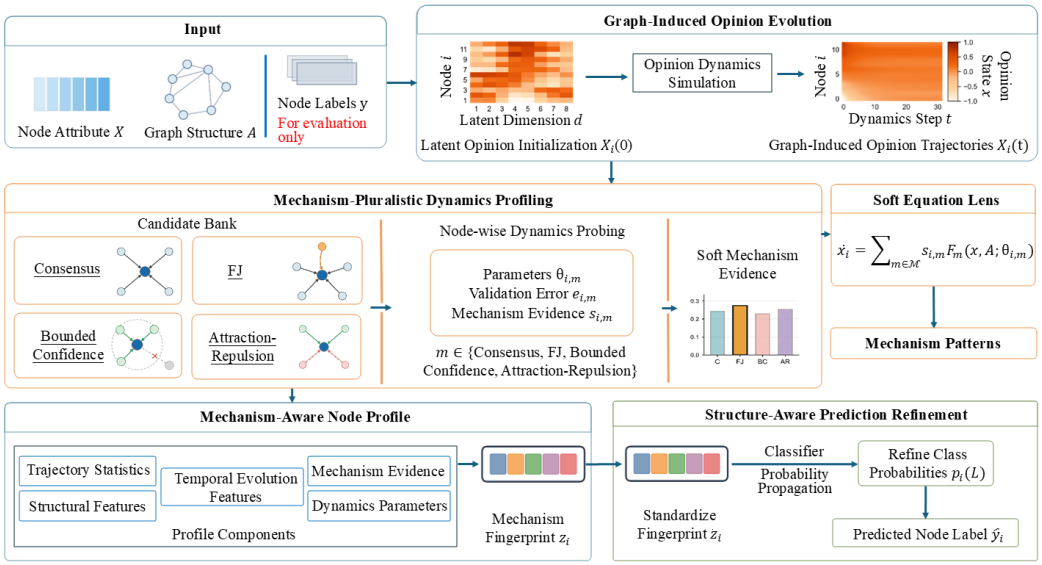


Figure 1: The overview of DynaPIN. Graph-induced opinion evolution creates latent trajectories; mechanism-pluralistic dynamics profiling fits the candidate mechanism bank; the resulting mechanism fingerprint supports structure-aware prediction refinement and the soft equation lens. Node labels are used only for downstream evaluation.

equation discovery through opinion-dynamics mechanisms. DynaPIN discovers a non-convergent mechanism grammar in which fitted responses combine self-retention, linear neighbor coupling, and selective nonlinear coupling rather than collapsing into one consensus-like law. DynaPIN also shows that the same Mechanism Fingerprint supports node classification on opinion-oriented and general graph benchmarks.

2 PRELIMINARIES

This section formalizes the problem studied in this paper and the basic notation of the opinion-dynamics formulation. DynaPIN studies transductive node classification on a static attributed graph, whose intermediate representation is a label-free response profile built from graph-induced opinion evolution. The notation separates original attributes, graph-induced latent opinions, mechanism-profile quantities, and downstream class probabilities.

Task and graph. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an attributed graph with $N = |\mathcal{V}|$ nodes, adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, node attribute matrix $\mathbf{X} \in \mathbb{R}^{N \times F}$ (F is the dimension of attribute vector), and labels $\mathbf{y} \in \{1, \ldots, C\}^N$. Training and testing nodes are denoted as $\mathcal{V}_{\text{tr}}$ and $\mathcal{V}_{\text{te}}$. The full graph and all attributes are visible during representation construction, label supervision is restricted to $\mathbf{y}_{\text{tr}}$, and $\mathbf{y}_{\text{te}}$ is held out for evaluation. The predicted label $\widehat{\mathbf{y}}_{\text{te}}$ of the target node is modelled as:

$$\widehat{\mathbf{y}}_{\text{te}} = \mathcal{C}(\mathcal{G}, \mathbf{X}, \mathbf{y}_{\text{tr}}) . \tag{2}$$

When a later module requires an undirected operator, the stored relation is symmetrized in that preprocessing step.

Rather than compressing $\mathbf{X}$ and $\mathbf{A}$ directly into a task embedding, we model each node as an opinion agent that responds along graph-permitted influence paths. Graph-induced opinion evolution maps attributes to a latent opinion initialization $\mathbf{x}_i(0) \in [-1, 1]^D$ and produces graph-induced opinion trajectories $\mathbf{x}_i(t)$ for $t = 0, \ldots, T$, stored as $\mathcal{X} \in \mathbb{R}^{(T+1) \times N \times D}$. Here $\mathbf{X}$ denotes original node attributes, and $\mathbf{x}_i(t)$ denotes the latent opinion state of node i. For generated general-graph trajectories, $\dot{x}_{i,d}(t)$ is the stored pre-clipping dynamics increment from simulation, and it is the finite difference between adjacent period states for converted voting trajectories. Dynamics probing splits time indices inside $\mathcal{X}$, not labeled nodes.

The response of each node is explained through a bank of named opinion-dynamics families:

$$\mathcal{M} = \{\text{Consensus}, \text{FJ}, \text{Bounded Confidence}, \text{Attraction-Repulsion}\}. \tag{3}$$

DynaPIN fits each node's local rate to these candidates and converts the fit into mechanism-level evidence. Let $\delta_{ij,d}(t) = x_{j,d}(t) - x_{i,d}(t)$, and $\langle u_{ij} \rangle_i = |\mathcal{N}_i|^{-1} \sum_{j \in \mathcal{N}_i} u_{ij}$, with zero returned for isolated nodes. The four probe families are

$$F_{\text{Consensus}}(t) = a \left\langle \delta_{ij,d}(t) \right\rangle_i , \tag{4}$$

$$F_{\text{FJ}}(t) = s\big[x_{i,d}(0) - x_{i,d}(t)\big] + a \left\langle \delta_{ij,d}(t) \right\rangle_i , \tag{5}$$

$$F_{\text{Bounded}}(t) = a \left\langle \frac{\delta_{ij,d}(t)}{1 + b\delta_{ij,d}^2(t)} \right\rangle_i , \tag{6}$$

$$F_{\text{AttractRepel}}(t) = a \left\langle \delta_{ij,d}(t) \right\rangle_i + r \left\langle \tanh(k[x_{i,d}(t) - x_{j,d}(t)]) \right\rangle_i . \tag{7}$$

a measures neighbor alignment, s measures self-retention toward the initial state, b controls distance-based attenuation, and r, k control the strength and saturation of nonlinear opposition. Consensus addresses pure neighbor alignment; FJ adds self-anchoring that resists influence; Bounded Confidence makes influence state-dependent so only similar neighbors matter; Attraction–Repulsion captures the coexistence of agreement pressure and disagreement pressure. The bank is grounded in canonical opinion-dynamics families (DeGroot, 1974; Friedkin & Johnsen, 1990; Hegselmann & Krause, 2002; Cornacchia et al., 2020), but its role is diagnostic: it asks which named response family explains the local rate within the fitted bank.

3 DynaPIN

DynaPIN converts static attributed graphs into mechanism-aware response profiles prior to classification. Figure 1 illustrates its pipeline: graph-driven opinion evolution generates label-free trajectories, multi-mechanism profiling fits a mechanism bank to local rates to build mechanism fingerprints, with two branches sharing fitted quantities for prediction and interpretation. Its innovation lies in feeding the classifier a symbolic-mechanistic fingerprint linked to interpretable dynamics, while the interpretation branch examines the same fitted values via soft equations. This plurality is purposeful: encoding inputs as coordinates over named response equations transfers implicit inductive bias from end-to-end encoders into inspectable, ablatable mechanism quantities, turning multi-equation machinery from opaque complexity into verifiable, interpretable structure.

3.1 Graph-Induced Opinion Evolution

Standard message-passing encoders compress $\mathbf{X}$ and $\mathbf{A}$ into a task embedding optimized end-to-end, so the response mechanism that produced the representation stays implicit and cannot be inspected. DynaPIN therefore does not ask a black-box encoder to do this compression. Instead it treats node attributes as initial opinion states and lets them evolve under graph-mediated opinion dynamics, producing a label-free diagnostic trajectory whose every step is explained by a named equation. Opinion dynamics serves as our analytical lens. Unlike generic diffusion, which leaves response modes implicit, its equations explicitly model neighbor alignment, self-retention, selective attenuation, and attraction–repulsion. Each step of this evolution is therefore attributable to a named equation, so the trajectory is built to be diagnostic rather than merely descriptive.

Firstly, DynaPIN projects the node attribute matrix $\mathbf{X}$ to D latent dimensions and robustly scales the result into $[-1, 1]$, producing $\mathbf{x}_i(0)$. Unlike classifier embeddings optimized via labels, this state acts as the starting condition in a graph-mediated response experiment. For node i, latent dimension index $d \in \{1, \ldots, D\}$, and step t, graph-induced opinion evolution evaluates

$$\dot{x}_{i,d}(t) = f_{\kappa_i}(\mathbf{x}_{\cdot,d}(t), \mathbf{A}, \mathbf{x}_{\cdot,d}(0); \vartheta_i) , \tag{8}$$

where κ_i indexes the generator family assigned to node i and ϑ_i denotes generator parameters. Forward Euler integration gives

$$x_{i,d}(t+1) = \Pi_\Omega[x_{i,d}(t) + \Delta t\, \dot{x}_{i,d}(t)] , \tag{9}$$

with Π_Ω clipping states to the bounded opinion domain. The stored outputs are the trajectory tensor $\mathcal{X}$, and node-wise dynamics probing uses $\dot{x}_{i,d}(t)$ as the derivative target. For general graphs, $\dot{x}_{i,d}(t)$

denotes the pre-clipping generator increment; for converted voting trajectories, it denotes the finite difference between adjacent period states. Labels are absent from both constructions. This step does not claim that a static graph contains observed social time. It creates a controlled diagnostic object: attributes define initial positions, edges define graph-permitted influence paths, and trajectories record how each node responds under graph-mediated evolution.

3.2 Mechanism-Pluralistic Dynamics Profiling

Real graphs are heterogeneous: different nodes follow different response rules, so assuming a single mechanism biases the whole profile toward that assumption. Mechanism-pluralistic dynamics profiling therefore fits a small bank of canonical opinion-dynamics families to each node's local rate and turns fit quality into soft mechanism evidence, so the model reports which named family actually explains the local dynamics rather than forcing one. Keeping the bank canonical and small also keeps every fitted term readable. The probing step is thus aimed at attributing each node's local rate to a named family rather than fitting a single best curve.

Node-wise dynamics probing fits every candidate mechanism to each node and latent dimension:

$$\widehat{\theta}_{i,d,m} = \arg\min_{\theta\in\Theta_m} \sum_{t\in\mathcal{T}_{\text{fit}}} \left[\dot{x}_{i,d}(t) - F_m(\mathbf{x}_{\cdot,d}(t), \mathbf{A}; \theta)\right]^2 . \tag{10}$$

The following validation interval $\mathcal{T}_{\text{val}}$ gives $e_{i,d,m}$, and the fit is converted into soft mechanism evidence by normalizing inverse-error scores within the bank:

$$q_{i,d,m} = \frac{1}{1+e_{i,d,m}}, \qquad s_{i,d,m} = \frac{q_{i,d,m}}{\sum_{m'\in\mathcal{M}} q_{i,d,m'}}. \tag{11}$$

A mechanism receives high evidence when parameters learned on an earlier trajectory segment remain accurate on a later segment. $\theta_{i,m}$, $e_{i,m}$, and $s_{i,m}$ denote node-level summaries of these dimension-wise quantities.

3.3 Mechanism-Aware Node Profile

The fitted equation-level quantities, trajectories, parameters, and errors, together with mechanism evidence, are distributed across dimensions and nodes and thus cannot be directly supplied to a classifier. The Mechanism-Aware Node Profile assembles them into one mechanism fingerprint whose coordinates are tied to named dynamics, giving the classifier a symbolic-mechanistic input instead of an opaque vector. Each coordinate of the fingerprint is therefore designed to remain tied to a named mechanism, keeping the profile inspectable for the downstream branches.

The profile concatenates five components into a Mechanism Fingerprint:

$$\mathbf{z}_i = \text{concat}\left(\boldsymbol{\tau}_i, \boldsymbol{\phi}_i, \mathbf{s}_i, \mathbf{g}_i, \mathbf{h}_i\right) \in \mathbb{R}^{d_z}. \tag{12}$$

Here $\boldsymbol{\tau}_i$ contains trajectory statistics, $\boldsymbol{\phi}_i$ contains fitted parameters and diagnostics, $\mathbf{s}_i$ contains Soft Mechanism Evidence, $\mathbf{g}_i$ contains structural descriptors, and $\mathbf{h}_i$ contains temporal evolution features. These components state what changed, how named equations explain the change, which mechanisms retain validation support, where the node sits in the graph, and what sequential information remains beyond summary statistics. Standardization is fitted on training nodes only, giving $\bar{\mathbf{z}}_i$.

3.4 Structure-Aware Prediction Refinement

The classifier readout on the fingerprint yields per-node initial probabilities before the final probability propagation step. Structure-Aware Prediction Refinement propagates probabilities on a normalized graph operator, clamping training rows to observed labels while leaving test rows free, so graph smoothness is exploited without leaking test labels. Setting $L = 0$ removes only the propagation and isolates its contribution; the trajectories, fitted mechanisms, and fingerprint remain intact. The propagation step is thus designed to exploit graph smoothness while its clamping rule keeps test labels strictly excluded.

A probabilistic classifier h_ϕ maps $\bar{\mathbf{z}}_i$ to Initial Class Probabilities:

$$\mathbf{p}_i(0) = h_\phi(\bar{\mathbf{z}}_i). \tag{13}$$

Structure-Aware Prediction Refinement then propagates probabilities on a row-normalized graph operator $\mathbf{W}$, formed with self-loops and symmetrization when the benchmark relation requires an undirected refinement operator:

$$\mathbf{P}(\ell+1) = (1-\alpha)\mathbf{P}(\ell) + \alpha\mathbf{W}\mathbf{P}(\ell), \qquad \ell = 0, \ldots, L-1. \tag{14}$$

Rows for $\mathcal{V}_{\text{tr}}$ are clamped to observed one-hot labels after each step, while test rows are never clamped. The Refined Class Probabilities $\mathbf{p}_i(L)$ define

$$\hat{y}_i = \arg\max_{c \in \{1,\ldots,C\}} p_{i,c}(L). \tag{15}$$

3.5 Soft Equation Lens

The prediction branch outputs class probabilities through a trained classifier; interpretation requires a mechanism-level description that is separate from that rule and cannot accidentally become an extra predictor. The soft equation lens aggregates the same fitted quantities into a mechanism-weighted equation, exposing whether responses concentrate on alignment, anchoring, selective influence, or attraction–repulsion. Because it reuses the quantities that enter $\mathbf{z}_i$, the reported patterns describe the information actually used by the prediction branch while staying outside the decision that outputs $\hat{y}_i$. The lens is therefore deliberately built to reuse the fitted quantities without entering the decision that produces $\hat{y}_i$.

The lens aggregates dimension-wise scores and parameters to node-level notation and reads the response of node i as

$$\dot{\mathbf{x}}_i = \sum_{m \in \mathcal{M}} s_{i,m} F_m(\mathbf{x}, \mathbf{A}; \theta_{i,m}). \tag{16}$$

This equation is not a classifier, an auxiliary supervised loss, or a replacement for Eq. (15). It is a mechanism-weighted view of the fitted response profile. At the node, class, and dataset levels, the lens shows whether responses concentrate around alignment, anchoring, selective influence, or attraction–repulsion.

4 Experimental Results and Mechanism Analysis

We first describe the experimental setting and then evaluate the proposed method along the two branches of Fig. 1. Classification verifies that the Mechanism Fingerprint carries a predictive signal across opinion-oriented and general graphs. Mechanism analysis is the main empirical object: the soft equation lens reads the same fitted quantities as equation-level response patterns.

4.1 Experimental Setting

Datasets. The opinion-oriented group contains Voteview, European Parliament, and Riksdag, each built from roll-call voting records with political-group labels. Their trajectories are period-wise voting-position signals, and $\dot{\mathbf{x}}_i$ is the finite difference between consecutive periods. These datasets anchor the mechanism language in its natural domain: nodes are political actors, graph structure records voting affinity, and trajectories describe changing position signals. The general benchmark group contains Cora, Citeseer, PubMed-N5000, Amazon Photo-N5000, and WikiCS-N5000. These graphs supply static attributes, structure, and class labels, and graph-induced opinion evolution supplies controlled latent trajectories from $\mathbf{X}$ and $\mathbf{A}$. Labels are reserved for downstream training and evaluation and do not enter trajectory construction, node-wise dynamics probing, or fingerprint construction.

Metrics. Macro-F1 is the primary metric because several evaluated graphs contain imbalanced classes, and Accuracy is reported as a secondary diagnostic in the appendix and archived summaries.

Protocol. Dynamics fitting uses contiguous 6:2:2 temporal partitions. The fitting interval estimates $\theta_{i,m}$, the validation interval computes $e_{i,m}$ and $s_{i,m}$, and the held-out interval supplies diagnostics. Node classification uses a stratified 7/3 train/test split. Standardization for $\bar{\mathbf{z}}_i$ is fitted on training nodes only. Training labels enter the classifier and training-node anchors in structure-aware prediction refinement; test labels are accessed only for metrics.

Table 1: Macro-F1 comparison on general graph benchmarks. Each baseline entry reports the best archived readout over the shared classifier pool.

Dataset	DynaPIN	DyGLib	TGL	GODNF	DiffGraph	DoG
Cora	**0.8506**	0.7821	0.7640	0.5946	0.6768	0.6917
Citeseer	**0.7170**	0.6925	0.6796	0.6603	0.6596	0.6641
PubMed-N5000	**0.7872**	0.7544	0.5518	0.7753	0.7755	0.7740
Amazon Photo-N5000	0.9399	0.9379	**0.9418**	0.9145	0.9383	0.9402
WikiCS-N5000	**0.7704**	0.7488	0.7118	0.7048	0.7410	0.7465

Baselines. The external comparator pool covers the baseline families used in the paper: DyGLib and TGL temporal graph pipelines (Yu et al., 2023; Zhou et al., 2022), GODNF-adapted opinion diffusion (Hevapathige et al., 2025), DiffGraph-style and DoG-style diffusion representations (Li et al., 2025; Wang et al., 2025), and TT-SBCM-style bounded-confidence temporal modeling (Luo et al., 2026). The archived comparison uses a common downstream interface with logistic regression, random forest, and radial-basis-function support vector machines. DynaPIN uses the same readout pool for the external opinion-dataset comparison and random forest with 200 trees for internal ablations.

Table 2: Macro-F1 comparison on opinion-oriented benchmarks. Each entry reports the best archived readout over the shared classifier pool.

Dataset	DynaPIN	DyGLib	TGL	GODNF	DiffGraph	DoG	TT-SBCM
Voteview	0.9948	0.9948	0.9948	0.9948	0.9948	0.9948	**0.9966**
European Parliament	**0.8631**	0.8494	0.8403	0.8443	0.8614	0.8456	0.8135
Riksdag	0.7156	0.6929	0.7068	0.6892	0.6851	**0.7284**	0.7004

4.2 Classification Results with Equation Readout

Tables 1 and 2 report the full Macro-F1 comparison against the archived baseline families, and Table 3 connects the same classification results to the lens readout. The performance numbers establish that mechanism fingerprints are not detached explanations. DynaPIN reaches near-saturated performance on Voteview, gives the strongest archived Macro-F1 on European Parliament, and remains competitive on Riksdag despite the eight-party class-balance challenge. On general graphs, it gives the highest archived Macro-F1 on Cora, Citeseer, PubMed-N5000, and WikiCS-N5000, and stays within 0.0019 Macro-F1 of the strongest comparator on Amazon Photo-N5000. These results show that the equation-derived profile is predictive on both voting graphs and ordinary node-classification benchmarks.

The equation column is the main reading of Table 3. F_{B} denotes the bounded confidence mechanism in Eq. (6), and F_{AR} denotes the attraction–repulsion mechanism in Eq. (7). Cora and WikiCS-N5000 show mixed attraction–repulsion and bounded confidence behavior, Citeseer and Amazon Photo-N5000 concentrate on attraction–repulsion-like responses, and PubMed-N5000 sits at the boundary between selective bounded coupling and attraction–repulsion. These patterns explain why a mechanism bank is more informative than a single diffusion law. A classifier sees only $\hat{y}_i$; DynaPIN also exposes which fitted response equation supplies the information used by the classifier.

Table 3: Main classification results and soft equation lens readout. The last column reports the equation family that explains the fitted response pattern.

Group	Dataset	Macro-F1(DynaPIN)	Best external	Macro-F1(best)	Lens readout
Opinion	Voteview	0.9948	TT-SBCM	0.9966	$\sum_m s_{i,m} F_m$ on roll-call $\Delta x_i(t)$
	European Parliament	**0.8631**	DiffGraph	0.8614	$\sum_m s_{i,m} F_m$ on roll-call $\Delta x_i(t)$
	Riksdag	0.7156	DoG	0.7284	$\sum_m s_{i,m} F_m$ under multi-party imbalance
General	Cora	**0.8506**	DyGLib	0.7821	$0.74F_{\mathrm{AR}} + 0.26F_{\mathrm{B}}$
	Citeseer	**0.7170**	DyGLib	0.6925	$0.85F_{\mathrm{AR}} + 0.15F_{\mathrm{B}}$
	PubMed-N5000	**0.7872**	TT-SBCM	0.7790	$0.50F_{\mathrm{B}} + 0.50F_{\mathrm{AR}}$
	Amazon Photo-N5000	0.9399	TGL	0.9418	$0.97F_{\mathrm{AR}} + 0.03F_{\mathrm{B}}$
	WikiCS-N5000	**0.7704**	DyGLib	0.7488	$0.62F_{\mathrm{AR}} + 0.38F_{\mathrm{B}}$

Table 4: Soft equation lens mechanism-pattern shares across node-dimension equations on general graphs. Shares are computed within the candidate mechanism bank and its simplification rule. Boldface denotes the dominant pattern.

Dataset	Bounded-like	Attr.–Rep.-like
Cora	25.52%	**74.48%**
Citeseer	15.09%	**84.91%**
PubMed-N5000	**50.12%**	49.88%
Amazon Photo-N5000	3.02%	**96.98%**
WikiCS-N5000	37.74%	**62.26%**

The soft equation lens reads the fitted profile as $\dot{\mathbf{x}}_i = \sum_{m \in \mathcal{M}} s_{i,m} F_m(\mathbf{x}, \mathbf{A}; \theta_{i,m})$. Table 4 reports dominant simplified patterns on the five general graphs. Consensus-like and pure FJ-like reductions do not dominate under the archived reduction rule. Instead, Attraction–Repulsion-like equations dominate Cora, Citeseer, Amazon Photo-N5000, and WikiCS-N5000, while PubMed-N5000 splits almost evenly between Bounded-like and Attraction–Repulsion-like patterns.

4.3 Profile and Refinement Contributions

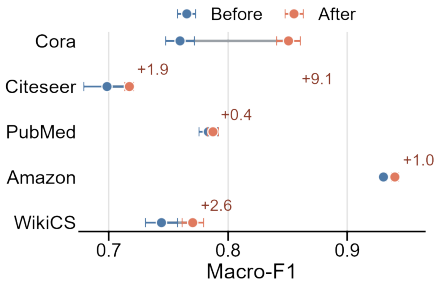


Figure 2: Contribution of structure-aware prediction refinement.

Structure-aware probability refinement improves mechanism-fingerprint readout across five graph benchmarks. Figure 2 separates mechanism profiling from graph-probability refinement by comparing Macro-F1 before and after propagation. Refinement boosts Macro-F1 on all datasets via graph consistency. Importantly, strong profiling performance exists pre-propagation: factorial ablations show DynaPIN profiles surpass raw trajectories, meaning fitted mechanism quantities encode class-related graph structure.

The gain is largest on Cora and remains positive on Citeseer, PubMed-N5000, Amazon Photo-N5000, and WikiCS-N5000. The refinement step therefore acts as a graph-consistent readout on top of the fitted mechanism profile rather than replacing the profile itself. Each component bears a straightforward interpretation: trajectory statistics describe node state evolution, fitted parameters define response equations, Soft Mechanism Evidence verifies equation reliability, structural descriptors capture node graph location, and temporal features preserve sequential behavior. Refinement only propagates class probabilities instead of producing mechanism patterns. These patterns already exist in $\mathbf{z}_i$, so profiling contributes both before and after refinement.

4.4 Soft Equation Lens

Figure 3 connects the predictive and mechanistic readings of the same learned dynamics. In the upper region, sampled nodes are plotted as $y_i(t) = \mathbf{w}^\top \mathbf{x}_i(t)$, where $\mathbf{x}_i(t)$ is the recorded graph-induced opinion trajectory and $\mathbf{w}$ is the final-state class-separating projection direction. This projected state axis makes the classification-side organization visible: trajectories starting from different locations move toward dataset-specific state bands rather than a single consensus value. In the lower region, the same graph-induced trajectories are grouped by the dominant Soft Equation Lens mechanism, with F_{AR} and F_{B} highlighted because they account for the major mechanism patterns reported in Table 4. Thus, Figure 3 complements the table: the table summarizes the discovered equation grammar, while the figure shows how the corresponding dynamics appear as trajectory-level response patterns.

Table 4 and Figure 3 form the central mechanism-discovery result. A pure smoothness prior describes the graph response as movement toward neighbor agreement. The fitted equations show a different grammar: selective neighbor coupling and attraction–repulsion terms dominate the simplified patterns, while their active terms repeatedly decompose into a compact equation-term pattern close to $1/8$ self-retention, $3/8$ linear neighbor coupling, and $1/2$ selective nonlinear coupling. The mechanism is therefore pluralistic and non-convergent: graph learning benefits from equations that allow nodes to retain their own state, follow neighbors selectively, and move against conflicting signals.

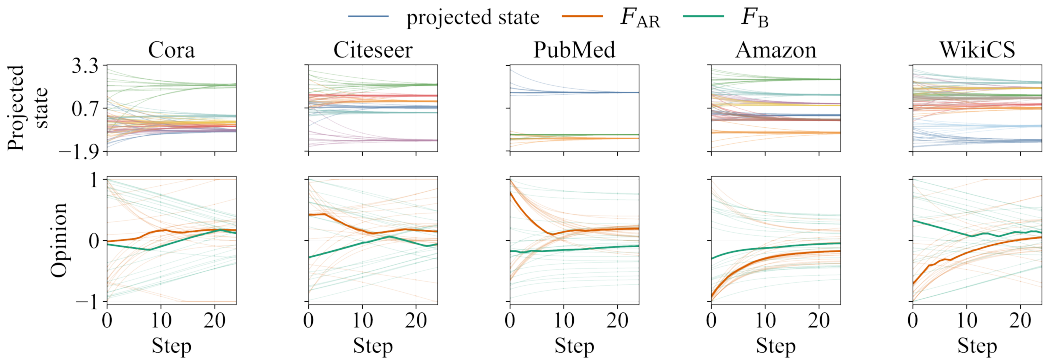


Figure 3: Projected classification trajectories and mechanism-conditioned opinion trajectories on general graphs.

This result also explains why opinion dynamics is used rather than an arbitrary dynamical system. Opinion dynamics names exactly the response modes encoded by the fitted equations: agreement, anchoring, selective exposure, and opposition. The voting datasets supply the natural semantic domain for these equations, while the general benchmarks show that the same response language transfers to non-political graphs. DynaPIN therefore produces two linked outputs: a prediction $\hat{y}_i$ and a readable map of mechanism patterns behind the learned profile.

4.5 Soft Mechanism Evidence

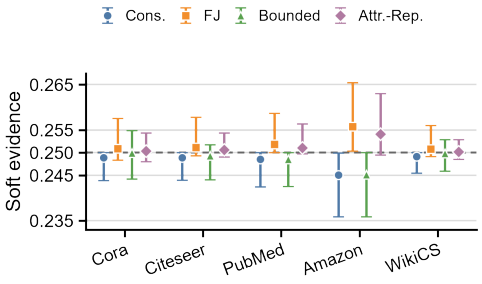


Figure 4: Validation of soft mechanism evidence across node–dimension pairs.

Finally, Figure 4 summarizes the distribution of soft mechanism evidence. Points show means and whiskers show archived 5th–95th percentiles. Values are validation-derived scores that compare candidate mechanisms within the probing bank. Figure 4 shows that DynaPIN does not treat a single dynamics family as universally sufficient, which keeps multiple named mechanisms available and lets node-level fitted responses determine the mechanism profile.

5 Conclusion

We introduce DynaPIN, a mechanism-pluralistic dynamics profiling framework that recasts transductive node classification as graph-mediated response profiling, exposing the response mechanism that end-to-end encoders leave implicit. Graph-induced opinion evolution produces label-free opinion trajectories, a small candidate mechanism bank probes each node's local rate, and the fitted parameters, soft mechanism evidence, and trajectory, structural, and temporal features form a mechanism fingerprint tied to named dynamics. Two branches share this profile: structure-aware prediction refinement for classification, and the soft equation lens for an interpretable equation-level view. DynaPIN attains the highest archived Macro-F1 on four of five benchmarks, within a narrow margin of the leader on Amazon Photo-N5000, and the lens recovers bounded-like and attraction–repulsion-like patterns across datasets. These results establish DynaPIN as a framework using opinion-dynamics equations both to build predictive node representations and to reveal how graph-induced responses vary across domains. Three directions remain: enlarging the mechanism bank toward data-driven or composite mechanisms; extending from static attributed graphs to temporal or heterogeneous graphs; and coupling the equation-level interpretability with human studies to test whether reported patterns yield actionable insight.

REFERENCES

Chirag Agarwal, Marinka Zitnik, and Himabindu Lakkaraju. Probing GNN explainers: A rigorous theoretical and empirical analysis of GNN explanation methods. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 8969–8996, 2022. URL https://proceedings.mlr.press/v151/agarwal22b.html.

Kenza Amara, Zhitao Ying, Zitao Zhang, Zhichao Han, Yang Zhao, Yinan Shan, Ulrik Brandes, Sebastian Schemm, and Ce Zhang. GraphFramEx: Towards systematic evaluation of explainability methods for graph neural networks. In *Proceedings of the First Learning on Graphs Conference*, volume 198 of *Proceedings of Machine Learning Research*, pp. 44:1–44:23, 2022. URL https://proceedings.mlr.press/v198/amara22a.html.

Carmela Bernardo, Claudio Altafini, Anton V. Proskurnikov, and Francesco Vasca. Bounded confidence opinion dynamics: A survey. *Automatica*, 159:111302, 2024. doi: 10.1016/j.automatica.2023.111302. URL https://doi.org/10.1016/j.automatica.2023.111302.

Cristian Bodnar, Francesco Di Giovanni, Benjamin Paul Chamberlain, Pietro Liò, and Michael M. Bronstein. Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in GNNs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 18527–18541, 2022. doi: 10.52202/068431-1346. URL https://papers.neurips.cc/paper_files/paper/2022/hash/75c45fca2aa416ada062b26cc4fb7641-Abstract-Conference.html.

Jinsong Chen, Kaiyuan Gao, Gaichao Li, and Kun He. NAGphormer: A tokenized graph transformer for node classification in large graphs. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=8KYeilT3Ow.

Qi Chen, Yifei Wang, Yisen Wang, Jiansheng Yang, and Zhouchen Lin. Optimization-induced graph implicit nonlinear diffusion. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 3648–3661, 2022. URL https://proceedings.mlr.press/v162/chen22z.html.

Jeongwhan Choi, Seoyoung Hong, Noseong Park, and Sung-Bae Cho. GREAD: Graph neural reaction-diffusion networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 5722–5747, 2023. URL https://proceedings.mlr.press/v202/choi23a.html.

Elisabetta Cornacchia, Noah Singer, and Emmanuel Abbe. Polarization in attraction-repulsion models. In *2020 IEEE International Symposium on Information Theory*, pp. 2765–2770, 2020. doi: 10.1109/ISIT44484.2020.9174010.

Miles Cranmer, Alvaro Sanchez-Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, and Shirley Ho. Discovering symbolic models from deep learning with inductive biases. *Advances in Neural Information Processing Systems*, 33:17429–17442, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/c9f2f917078bd2db12f23c3b413d9cba-Abstract.html.

Morris H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69 (345):118–121, 1974. doi: 10.1080/01621459.1974.10480137.

Noah E. Friedkin and Eugene C. Johnsen. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3–4):193–206, 1990. doi: 10.1080/0022250X.1990.9990069.

Jonas Berg Hansen and Filippo Maria Bianchi. Total variation graph neural networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 12445–12468, 2023. URL https://proceedings.mlr.press/v202/hansen23a.html.

Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3):2, 2002. URL https://www.jasss.org/5/3/2.html.

Akila Hevapathige, Asiri Wijesinghe, and Ahad N. Zehmakan. Graph neural diffusion via generalized opinion dynamics. *arXiv preprint arXiv:2508.11249*, 2025. doi: 10.48550/arXiv.2508.11249.

Jiao Hu, Jiaxu Cui, and Bo Yang. Learning interpretable network dynamics via universal neural symbolic regression. *Nature Communications*, 16:6226, 2025. doi: 10.1038/s41467-025-61575-7. URL https://doi.org/10.1038/s41467-025-61575-7.

Pierre-Alexandre Kamienny, Stéphane d'Ascoli, Guillaume Lample, and François Charton. End-to-end symbolic regression with transformers. In *Advances in Neural Information Processing Systems*, volume 35, pp. 10269–10281, 2022. doi: 10.52202/068431-0746. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/42eb37cdbefd7abae0835f4b67548c39-Abstract-Conference.html.

Shima Khoshraftar and Aijun An. A survey on graph representation learning methods. *ACM Transactions on Intelligent Systems and Technology*, 15(1):1–55, 2024. doi: 10.1145/3633518.

Zhiwei Li, Lianghao Xia, Hao Hua, Shijie Zhang, Shuangyang Wang, and Chao Huang. DiffGraph: Heterogeneous graph diffusion model. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining*, pp. 40–49, 2025. doi: 10.1145/3701551.3703590.

Xixun Lin, Wenxiao Zhang, Fengzhao Shi, Chuan Zhou, Lixin Zou, Xiangyu Zhao, Dawei Yin, Shirui Pan, and Yanan Cao. Graph neural stochastic diffusion for estimating uncertainty in node classification. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 30457–30478, 2024. URL https://proceedings.mlr.press/v235/lin24x.html.

Sitao Luan, Chenqing Hua, Qincheng Lu, Jiaqi Zhu, Mingde Zhao, Shuyuan Zhang, Xiao-Wen Chang, and Doina Precup. Revisiting heterophily for graph neural networks. In *Advances in Neural Information Processing Systems*, volume 35, pp. 1362–1375, 2022. doi: 10.52202/068431-0100. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/092359ce5cf60a80e882378944bf1be4-Abstract.html.

Jing Luo, Yue Yao, Zhi Liu, Wei Tang, and Xiang Zhang. Sociologically-informed opinion prediction: Fusing bounded confidence theory with TabTransformer. *Neurocomputing*, 674:132875, 2026. doi: 10.1016/j.neucom.2026.132875.

Yao Ma, Xiaorui Liu, Neil Shah, and Jiliang Tang. Is homophily a necessity for graph neural networks? In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=ucASPPD9GKN.

Péter Mernyei and Cătălina Cangea. Wiki-CS: A wikipedia-based benchmark for graph neural networks. In *Graph Representation Learning and Beyond Workshop at the International Conference on Machine Learning*, 2020. URL https://grlplus.github.io/papers/32.pdf.

Galileo Namata, Ben London, Lise Getoor, and Bert Huang. Query-driven active surveying for collective classification. In *Proceedings of the Workshop on Mining and Learning with Graphs*, 2012. URL https://berthuang.com/papers/namata-mlg12.pdf.

Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, and Liudmila Prokhorenkova. A critical look at the evaluation of GNNs under heterophily: Are we really making progress? In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=tJbbQfw-5wv.

Ladislav Rampasek, Michael Galkin, Vijay Prakash Dwivedi, Anh Tuan Luu, Guy Wolf, and Dominique Beaini. Recipe for a general, powerful, scalable graph transformer. In *Advances in Neural Information Processing Systems*, volume 35, 2022. doi: 10.52202/068431-1054.

Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Gallagher, and Tina Eliassi-Rad. Collective classification in network data. *AI Magazine*, 29(3):93–106, 2008. doi: 10.1609/aimag.v29i3.2157.

Haoxin Sun and Zhongzhi Zhang. Opinion optimization in directed social networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 4623–4632, 2023. doi: 10.1609/aaai.v37i4.25585. URL https://doi.org/10.1609/aaai.v37i4.25585.

Voteview. Voteview: Congressional roll-call votes and DW-NOMINATE data. https://voteview.com/data, 2026. Accessed: 2026-09-07.

Xiaoqi Wang and Han-Wei Shen. GNNBoundary: Towards explaining graph neural networks through the lens of decision boundaries. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=WIzzXCVYiH.

Yiming Wang, Chen Liu, and Yixin Yang. Diffusion on graph: Augmentation of graph structure for node classification. *Transactions on Machine Learning Research*, 2025. URL https://openreview.net/forum?id=tzW948kU6x.

Le Yu, Leilei Sun, Bowen Du, and Weifeng Lv. Towards better dynamic graph learning: New architecture and unified library. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-2960.

Zihan Yu, Jingtao Ding, and Yong Li. Discovering network dynamics with neural symbolic regression. *Nature Computational Science*, 6:156–168, 2026. doi: 10.1038/s43588-025-00893-8. URL https://doi.org/10.1038/s43588-025-00893-8.

Hao Yuan, Haiyang Yu, Shurui Gui, and Shuiwang Ji. Explainability in graph neural networks: A taxonomic survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5): 5782–5799, 2023. doi: 10.1109/TPAMI.2022.3204236.

Haoteng Zhou, Da Zheng, Israt Nisa, Vassilis Ioannidis, Xiang Song, and George Karypis. TGL: A general framework for temporal GNN training on billion-scale graphs. *Proceedings of the VLDB Endowment*, 15(8):1572–1580, 2022. doi: 10.14778/3529337.3529342.