SPD RICCI FLOW: GRAPH REPRESENTATION LEARNING THROUGH METRIC EVOLUTION

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce **SPD Ricci Flow**, a graph learning framework in which each node's symmetric positive definite (SPD) representation also serves as its local metric and evolves through Ricci-type dynamics. Unlike previous edge-based graph Ricci flows that evolve scalar edge lengths or weights, our framework directly evolves node-wise metric tensors using a matrix-valued Ricci-type curvature operator constructed from AIRM discrepancies between neighboring SPD metrics. This makes curvature-driven metric evolution itself the mechanism for updating node representations. The curvature also enables interpretable analysis of local geometric variation in node representations. We prove continuous-time energy dissipation and show that, under graph-Laplacian consistency, the linearized continuum limit matches the heat-type principal part of Ricci–DeTurck evolution. Experiments on node classification, molecular property prediction, and financial spillover prediction support metric evolution as a mechanism for graph representation learning.

1 INTRODUCTION

Message passing graph neural networks (Kipf & Welling, 2017; Hamilton et al., 2017; Veličković et al., 2018) learn node representations by repeatedly exchanging information along graph edges. In these models, graph connectivity specifies which nodes can interact, while graph geometry can guide how messages are constructed and transformed. One way to describe geometry is through a graph metric, such as the shortest-path distance induced by edge lengths (Ni et al., 2019). Such metrics can guide message passing and evolve alongside node representations (Chen et al., 2025). Beyond using geometry to guide representation learning, we seek to unify representation and local geometry within the same node state, so that metric evolution itself becomes the representation update mechanism. To formulate this local geometry, we draw on the Riemannian notion of a metric: a smooth field of inner products on the tangent spaces, rather than a distance function on the discrete node set (Lee, 2018). As a first step, we ask: *can a node representation itself serve as a local metric tensor on the feature space?*

Symmetric positive definite (SPD) matrices can realize this idea by serving as both node representations and local feature-space metrics. Each $g_i \in \mathrm{SPD}_d$ defines the inner product $\langle u, v\rangle_i = u^\top g_i v$ for vectors u, v in the feature space $\mathbb{R}^d$. Its eigenvalues and eigendirections determine how vectors are compared, allowing the representation to encode anisotropic local geometry. Under a manifold-discretization perspective, graph nodes can be viewed as samples from an underlying manifold, with each node-wise SPD state representing a local Riemannian metric tensor in a chosen frame.

Once the node representation itself serves as a metric, its update becomes a metric evolution problem. Hamilton's Ricci flow (Hamilton, 1982) provides a natural prototype for such an evolution: On a Riemannian manifold $(\mathcal{M}, g)$, it evolves g according to $\partial_t g = -2\,\mathrm{Ric}(g)$, where Ric denotes the Ricci curvature tensor. This tensor-valued update acts directly on the metric and allows direction-dependent changes, motivating a curvature-driven method for updating SPD node representations.

We thus propose **SPD Ricci Flow**, which treats each node's SPD representation as its local metric and evolves this metric directly using matrix-valued Ricci-type curvature, thereby updating the node representation itself. Conventional edge-based discrete Ricci flows evolve edge lengths according to edge curvature, without directly evolving node representations (Ni et al., 2019; Lai et al., 2022), while Chen et al. (2025) connect this evolution to node features through attribute-dependent scalar metrics. In SPD Ricci Flow, metric evolution itself provides the representation update mechanism.

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

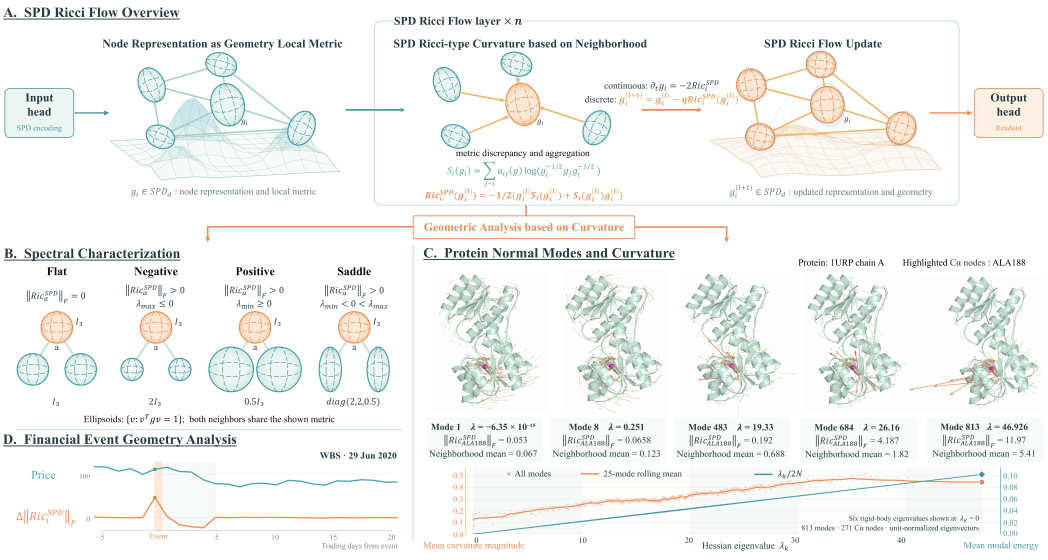


Figure 1: **SPD Ricci Flow.** Node representations serve as local metrics whose curvature-driven evolution enables graph representation learning, while the resulting curvature provides a geometric lens for analyzing local variation in the metric field.

Our contributions are summarized as follows:

(1) Ricci-type dynamics on an SPD metric field. We formulate Bochner–Weitzenböck-inspired Ricci-type evolution dynamics directly on a SPD metric field, where each SPD state serves simultaneously as a node representation and a local metric. We construct the driving Ricci-type curvature operator intrinsically by locally aggregating discrepancies between neighboring metrics under the AIRM. Curvature-driven metric evolution thus provides the representation update mechanism itself.

(2) Theoretical consistency and energy dissipation. We provide theoretical justification for the metric dynamics. Under graph-Laplacian consistency, the linearized flow recovers the heat-type principal mechanism of Ricci–DeTurck evolution, while the continuous-time dynamics dissipate a Dirichlet-type energy over neighboring SPD metrics.

(3) Predictive performance and geometric analysis. Experiments demonstrate strong performance on node classification, molecular property prediction, and financial spillover prediction. Beyond prediction, we use curvature magnitudes and spectra to analyze local geometric variation, illustrating spatial patterns in molecular motion and event-associated geometric changes around financial news.

2 Preliminaries and Related Work

Riemannian Metrics and SPD Geometry. Let M be a smooth manifold. A Riemannian metric g is a smooth symmetric covariant 2-tensor field such that, for every $p \in M$, g_p defines an inner product on the tangent space T_pM. Thus, under any basis of T_pM, g_p admits an SPD matrix representation. A matrix $X \in \mathbb{R}^{d\times d}$ is SPD if $X = X^\top$ and $x^\top Xx > 0$ for all $x \in \mathbb{R}^d \setminus \{0\}$. The space of $d \times d$ SPD matrices forms a smooth manifold, denoted by SPD_d. A widely adopted Riemannian metric on SPD_d is the affine-invariant Riemannian metric (AIRM), defined at $P \in \mathrm{SPD}_d$ as $g_P^{\mathrm{AI}}(U, V) = \mathrm{tr}\big(P^{-1}UP^{-1}V\big)$ where $U, V \in T_P\mathrm{SPD}_d$.

Ricci Curvature and Ricci Flow. On a Riemannian manifold $(\mathcal{M}, g)$, the metric g determines the Ricci curvature tensor, a symmetric covariant 2-tensor defined by $\mathrm{Ric}(g)_p(u, v) = \mathrm{tr}(w \mapsto \mathcal{R}_p(w, u)v)$, where $\mathcal{R}$ is the Riemann curvature tensor of g and $u, v, w \in T_pM$. Ricci flow (Hamilton, 1982) uses this curvature to evolve the metric via $\partial_t g(t) = -2\,\mathrm{Ric}(g(t))$, acting as a geometric smoothing mechanism that tends to reduce local geometric irregularities.

Discrete Ricci Curvature and Ricci Flows. Among discrete graph Ricci curvatures (Forman, 2003), Ollivier–Ricci curvature (ORC) (Ollivier, 2009) assigns each edge a scalar curvature measuring Wasserstein contraction or expansion of neighboring probability measures relative to their

graph distance. The resulting Ollivier Ricci flows evolve scalar edge lengths or weights, making community structure more distinct (Ni et al., 2019; Lai et al., 2022). Extending Ricci flow to graph representation learning, GNRF derives curvature from feature-dependent edge weights and evolves these weights via Ricci flow, inducing node-feature evolution (Chen et al., 2025). Adaptive Riemannian GNN (Wang et al., 2026) assign each node an SPD metric but update representations through standard GNN propagation. Ricci geometry enters only through a Ricci flow-inspired training regularizer, rather than an evolution law for node representations or metric tensors.

SPD Representation Learning. SPDNet (Huang & Van Gool, 2017) develops neural architectures for SPD representations using bilinear mappings, eigenvalue rectification, and matrix-logarithm layers. SPDGNN (Zhao et al., 2023) and SPDSheaf (Peng et al., 2026) extend SPD learning to graphs via neighborhood propagation. Building on SPD representation learning, our method introduces curvature-driven metric evolution as the propagation mechanism.

3 SPD RICCI FLOW

To construct a node-wise SPD Ricci flow, we require a metric field on the graph and a discrete counterpart of Ricci tensor $\mathrm{Ric}(g)$. We first introduce an SPD-valued metric field. The Bochner–Weitzenböck identity then provides the guiding design principle for the curvature construction: we compare neighboring metrics under the AIRM, aggregate their discrepancies into a node-wise Laplacian-like tensor, and construct the corresponding Ricci-type operator and metric flow. Finally, under graph-Laplacian consistency, we relate the linearized discrete flow to the heat-type principal part of Ricci–DeTurck evolution.

3.1 BOCHNER–WEITZENBÖCK MOTIVATION AND SPD RICCI-TYPE CURVATURE.

Let $G = (V, E, w)$ be a finite undirected weighted graph with symmetric nonnegative weights $w_{ij} = w_{ji} \geq 0$. Each node $i \in V$ carries an intrinsic feature in a d-dimensional vector space, together with an SPD matrix $g_i \in \mathrm{SPD}_d$ that defines a positive-definite bilinear form on this space. Moreover, we interpret i as a representative element of its neighbor local coordinate in background geometry space (manifold), and g_i as the *local constant metric* of this coordinate. We take local metrics g_i as the objects of interest and treat them as SPD-valued node representations simultaneously. The collection $g : V \to \mathrm{SPD}_d,\ i \mapsto g_i$ defines the discrete metric field to be evolved over the graph.

Bochner–Weitzenböck principle. In smooth Riemannian geometry, the Bochner–Weitzenböck identity for one-forms states

$$\Delta_H \omega = \nabla^* \nabla \omega + \mathrm{Ric}^{\#}(\omega), \tag{1}$$

where Δ_H is the Hodge Laplacian, $\nabla^*\nabla$ is the rough connection Laplacian, and $\mathrm{Ric}^{\#}$ is the natural Ricci curvature operator acting on one-forms. Equivalently, the identity $\mathrm{Ric}^{\#}(\omega) = \Delta_H \omega - \nabla^*\nabla\omega$ expresses curvature as the tensor-side correction between the Hodge-type diffusion and rough heat diffusion. Guided by this identity, we first construct a Hodge-type Laplacian, then specify the rough connection Laplacian, and use their difference to define the SPD Ricci-type curvature.

Metric discrepancy and Dirichlet aggregation. To construct the Hodge-type Laplacian for the SPD metric field, we follow the graph-Laplacian principle of aggregating neighboring differences. In Euclidean space, this gives $\sum_{j\sim i} w_{ij}(x_j - x_i)$ under the negative Laplacian convention. For SPD-valued node states g_i, we instead measure neighboring metric discrepancies using AIRM relative logarithms. For an oriented edge $i \to j$, define

$$L_{i\to j} = \log(g_i^{-1/2} g_j g_i^{-1/2}), \qquad r_{ij} = \|L_{i\to j}\|_F^2 = d^2_{\mathrm{SPD}}(g_i, g_j). \tag{2}$$

$L_{i\to j}$ records the relative displacement from g_i to g_j in the normalized tangent space $T_I\mathrm{SPD}_d$, while r_{ij} gives its squared AIRM distance.

To aggregate these edge-wise discrepancies in a geometrically principled way, let $\psi \in C^1[0,\infty)$ with $\psi'(r) \geq 0$ for any $r \in [0,\infty)$. Define the Dirichlet-type energy

$$\mathcal{E}_\psi(g) = \sum_{\{i,j\}\in E} w_{ij}\psi(r_{ij}). \tag{3}$$

Its negative AIRM gradient at node i takes the form

$$-\nabla_{g_i}^{\mathrm{AIRM}}\mathcal{E}_\psi(g) = g_i^{1/2} S_i(g) g_i^{1/2}, \qquad S_i(g) = 2\sum_{j\sim i} w_{ij}\psi'(r_{ij}) L_{i\to j}. \tag{4}$$

$S_i(g)$ aggregates the local metric discrepancies around node i into the *Hodge-type Laplacian* used in the subsequent curvature construction. We call $S_i(g)$ the *whitened metric-field tensor*. Meanwhile, under the locally constant metric construction, we take the *rough connection Laplacian* term $\nabla^*\nabla g$ to *vanish*. Thus, we can interpret $S_i(g)$ as a curvature related operator. See Corollary A.1 for details.

In our model, we specifically set $\psi_\beta(r) = \frac{1-e^{-\beta r}}{2\beta}, \quad \beta > 0, \quad \psi_0(r) = \frac{r}{2}$, yielding

$$S_i(g) = \sum_{j\sim i} a_{ij}(g) L_{i\to j}, \qquad a_{ij}(g) = w_{ij} e^{-\beta r_{ij}}, \qquad \beta \geq 0. \tag{5}$$

Hence the aggregation is induced by an increasing Dirichlet-type energy rather than an ad hoc matrix average. The gradient derivation and the corresponding primitive ψ_β are given in Corollary A.1. In particular, in the log-linear case $L_{i\to j} = h_j - h_i$, the aggregation in Eq. (5) satisfies $-S_i(g) = (L_G h)_i$, where $L_G := D - W$ is the usual graph Laplacian.

We then interpret $A_i = -S_i(g)$ as a curvature endomorphism. To turn this endomorphism into a symmetric metric variation at metric g_i for modeling a Ricci-type tensor, we use congruence lift $A_i \longmapsto \frac{1}{2}(g_i A_i + A_i g_i)$. Apply to Eq. (4) gives the following node-wise SPD Ricci-type curvature.

Definition 3.1 (SPD Ricci-type curvature)**.** *The SPD Ricci-type curvature of the SPD metric field at node i is*

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -\frac{1}{2}\left(g_i S_i(g) + S_i(g) g_i\right) \in \mathrm{Sym}_d. \tag{6}$$

The corresponding continuous-time SPD Ricci flow is

$$\partial_t g_i = -2\,\mathrm{Ric}_i^{\mathrm{SPD}}(g) = g_i S_i(g) + S_i(g) g_i. \tag{7}$$

The construction defines an evolution of the metric field itself rather than a scalar edge reweighting. At $g_i = I$, Eq. (6) reduces to $\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -S_i(g) = \nabla_{g_i}^{\mathrm{AIRM}}\mathcal{E}_\psi(g)$. Hence the corresponding flow follows the negative AIRM gradient of the Dirichlet energy, motivating the graph-diffusion analysis in Proposition 3.1.

3.2 Verification of mechanism consistency with Ricci-type Principal Dynamics.

The purpose of this subsection is to verify the mechanism of the proposed SPD Ricci flow. The target is to show that, when written in logarithmic metric coordinates, SPD Ricci flow has the same parabolic principal part as Ricci-DeTurck evolution. The argument proceeds in three steps. First, building on the preceding near-identity observation, we linearize the SPD Ricci flow near a constant metric field on a fixed graph and identify its leading term as graph heat diffusion of the log-metric field. Second, we recall that in harmonic coordinates or after Ricci–DeTurck modification, smooth Ricci flow has a heat-type principal part near a constant background. Third, under a standard graph-Laplacian consistency assumption, the scaled log-metric graph heat operator converges to this principal part. We give a principal-mechanism verification: the SPD Ricci flow reproduces the leading metric-diffusion mechanism of Ricci-type evolution, while noncommutative SPD effects and adaptive weights enter as higher-order corrections.

After this consistency result, we prove an energy-dissipation theorem showing that the same SPD Ricci operator gives a stable relaxation mechanism: the free flow decreases a natural metric-field energy and stops at equilibrium, i.e., at a diffused global section with zero SPD Ricci response. The equilibrium interpretation is discussed in more detail in Corollary A.2 and Corollary A.3.

Linearization and Ricci–DeTurck mechanism. The preceding near–identity observation suggests a graph-heat structure in the SPD Ricci dynamics. We now formalize this observation for perturbations around an arbitrary constant metric background. After normalizing the background to I and using log-metric coordinates, the following proposition shows that the linearized dynamics are governed by graph heat diffusion.

Choosing coordinates in which the background metric is I, we write $g_i = \exp(H_i)$ where $H_i = \log g_i \in \mathrm{Sym}_d$ is the log-metric perturbation from the background.

Proposition 3.1 (Linearized graph-diffusion form of SPD Ricci flow)**.** *Let* $\max_i \|H_i\| \le \varepsilon$*, and let* $S_i(g) = \sum_{j\sim i} w_{ij} \exp\big(-\beta\|L_{i\to j}\|_F^2\big)\, L_{i\to j}$ *where* $\beta \ge 0$*. Then the SPD Ricci flow Eq.* (7)*, expressed in log-metric coordinates, satisfies*

$$\partial_t H_i = 2\sum_{j\sim i} w_{ij}(H_j - H_i) + \mathcal{R}_i(H), \tag{8}$$

where $\mathcal{R}_i(H) = O(\varepsilon^2)$ *for a fixed graph. Hence, to first order,*

$$\partial_t H_i \approx 2\sum_{j\sim i} w_{ij}(H_j - H_i),$$

so the linearized SPD Ricci dynamics are governed by graph heat diffusion of the log-metric field.

The previous proposition identifies the principal operator on a fixed graph. We next compare it with the smooth metric evolution that motivates the term "Ricci-flow-type".

Example 1 (Ricci–DeTurck parabolic principal mechanism of Ricci flow)**.** *Under the Ricci–DeTurck modification, that in harmonic coordinates at the level of principal symbols, smooth Ricci flow can be written as*

$$\partial_t g_{\alpha\beta} = g^{kl}\partial_k\partial_l g_{\alpha\beta} + Q_{\alpha\beta}(g^{-1}, \partial g, \bar{g}), \tag{9}$$

where Q *contains lower-order nonlinear terms but no second derivatives of* g*. After whitening a constant background metric to the identity and writing* $g(x) = \exp(H(x))$*, the matrix field* $H(x) = \log g(x)$ *is simply the logarithmic coordinate of the metric: near the identity,* $g = I + H + O(\|H\|_F^2)$*. Therefore the first-order principal part is the component-wise heat equation*

$$\partial_t H = \Delta H, \qquad \Delta H = \sum_{a=1}^{m} \partial_a^2 H. \tag{10}$$

This heat-type operator is the parabolic principal mechanism of Ricci flow metric evolution.

Let $x_i \in \Omega \subset \mathbb{R}^m$ be graph nodes with sampling scale h. We define the scaled graph Laplacian by $(\mathcal{L}_h F)_i = \frac{1}{h^2}\sum_{j\sim i} w_{ij}^{(h)}\big(F(x_j) - F(x_i)\big)$, where $F : \Omega \to \mathrm{Sym}_m$ is a smooth matrix-valued field, and the theorem below applies this condition to the log-metric field H. We assume Laplacian consistency, i.e.,

$$(\mathcal{L}_h F)_i = c_K \Delta F(x_i) + O(h^\alpha), \quad c_K > 0, \quad \alpha > 0,$$

so that, for the log-metric field H, $\mathcal{L}_h H \to c_K \Delta H$ as $h \to 0$.

Theorem 3.1 (Consistency with Ricci-type principal dynamics)**.** *Let* $g : \Omega \to \mathrm{SPD}_m$ *be a smooth metric field close to a constant background normalized to* I*, with* $g(x) = \exp(H(x))$*. Set* $g_i = g(x_i)$ *and assume that the sampled weighted graphs satisfy the Laplacian-consistency condition above. Consider the scaled SPD Ricci flow*

$$\partial_t g_i = g_i S_i^{(h)} + S_i^{(h)} g_i, \qquad S_i^{(h)} = \frac{1}{h^2}\sum_{j\sim i} w_{ij}^{(h)} e^{-\beta\|L_{i\to j}\|_F^2} L_{i\to j}. \tag{11}$$

Under this consistency assumption, the flow in log-metric coordinates admits

$$\partial_t H_i = 2(\mathcal{L}_h H)_i + \mathcal{N}_h(H)_i, \qquad \mathcal{N}_h(0) = 0, \qquad D\mathcal{N}_h(0) = 0, \tag{12}$$

where $\mathcal{N}_h(H)_i$ *contains only second- and higher-order terms near* $H = 0$*. Applying the same Laplacian-consistency relation to the leading term gives* $(\mathcal{L}_h H)_i = c_K \Delta H(x_i) + O(h^\alpha)$*, hence*

$$\partial_t H_i = 2c_K \Delta H(x_i) + O(h^\alpha) + \mathcal{N}_{h,i}(H).$$

Therefore, after the constant time rescaling $\tau = 2c_K t$*, the linearized continuum limit is*

$$\partial_\tau H = \Delta H,$$

which coincides with the principal dynamics of the Ricci–DeTurck evolution in Example Eq. (1)*.*

This result establishes consistency at the level of the linearized principal dynamics, rather than recovery of the full nonlinear Ricci tensor. Noncommutative SPD effects and state-dependent weights enter only through higher-order terms near the constant background.

Energy dissipation and equilibrium. The same operator also has an intrinsic relaxation principle on the graph. This explains why the free flow smooths metric-field roughness and why curvature residuals decay when the system is not externally forced by a task loss or event perturbation.

Theorem 3.2 (Energy dissipation under SPD Ricci flow)**.** *Let $\mathcal{E}_\psi$ be the energy in Eq.* (3)*, and let $S_i(g)$ be the corresponding whitened tensor in Eq.* (4)*. Along the continuous-time SPD Ricci flow $\partial_t g_i = g_i S_i + S_i g_i$, the energy is nonincreasing:*

$$\frac{d}{dt}\mathcal{E}_\psi(g(t)) \le 0. \tag{13}$$

The dissipation vanishes if and only if $S_i(g) = 0$, equivalently $\mathrm{Ric}_i^{\mathrm{SPD}}(g) = 0$ for every node.

For the ordinary Dirichlet energy $\psi(r) = r/2$, the equilibrium condition $S_i(g) = 0$ says that g_i is the affine-invariant Fréchet barycenter of its neighboring SPD metrics. Under uniform spectral bounds on the node metrics, the aggregate residual $\sum_i \| \mathrm{Ric}_i^{\mathrm{SPD}} \|_F$ is quantitatively comparable to the dissipation rate; see Corollary A.2 for the full identities and proofs. Moreover, we also view the dissipation stopping condition $S_i(g) = 0$ as a *local equilibrium*, which corresponds with the global section of SPD Ricci flow and we provide a preliminary discussion of it in Corollary A.3. Therefore, Frobenius norm and eigenvalue summaries of $\mathrm{Ric}_i^{\mathrm{SPD}}$ are meaningful residual diagnostics. Thus we use this geometric interpretability to analyze real-world data in the next two sections (Fig. 1, bottom).

4 Geometric Analysis and Graph Learning

SPD Ricci-type curvature describes the directional response of a local metric to its neighborhood. Since $\mathrm{Ric}_i^{\mathrm{SPD}}$ is symmetric, it admits an eigendecomposition $\mathrm{Ric}_i^{\mathrm{SPD}} = U_i \,\mathrm{diag}(\lambda_{i,1}, \dots, \lambda_{i,d}) U_i^\top$, with $\lambda_{i,1} \le \cdots \le \lambda_{i,d}$. Under $\partial_t g(t) = -2\,\mathrm{Ric}^{\mathrm{SPD}}(g(t))$, the squared metric length of a fixed vector v satisfies $\partial_t (v^\top g_i v) = -2 v^\top \mathrm{Ric}_i^{\mathrm{SPD}}\, v$. Positive and negative eigenvalues therefore indicate instantaneous contraction and expansion of metric lengths along the corresponding eigendirections. The Frobenius norm $F_i = \| \mathrm{Ric}_i^{\mathrm{SPD}} \|_F = (\sum_{k=1}^d \lambda_{i,k}^2)^{1/2}$ measures the magnitude of the curvature response. As illustrated in Fig. 1B, we distinguish four response types: **Flat** when $F_i = 0$; **Negative** when $F_i > 0$ and $\lambda_{i,d} \le 0$; **Positive** when $F_i > 0$ and $\lambda_{i,1} \ge 0$; and **Saddle** when $\lambda_{i,1} < 0 < \lambda_{i,d}$. These types describe the instantaneous local response.

4.1 Molecular Illustration: Normal-Mode Analysis

To illustrate how SPD Ricci-type curvature responds to spatial motion patterns in a physical system, we examine protein normal modes using the anisotropic network model (ANM) (Eyal et al., 2006; Atilgan et al., 2001; Xia et al., 2015). ANM represents a protein as a network of sites connected by springs and models small displacements around their positions. Each normal mode specifies a collective motion pattern: a displacement vector at each site indicates how that site moves within the pattern, visualized as arrows in Fig. 1C. Mathematically, these patterns are eigenvectors of the Hessian matrix H of the model's elastic energy. The corresponding eigenvalue λ_k measures the mode's stiffness, with larger values indicating a greater energy cost. Higher-eigenvalue modes often exhibit finer-scale or more localized motion. We map these displacement patterns to node-wise SPD metrics and examine how their curvature responses vary across modes. See App. A.2 for details.

For each normalized mode k, let $u_i^{(k)} \in \mathbb{R}^3$ denote the displacement of node i. We encode it as an SPD matrix by $g_i^{(k)} = u_i^{(k)} (u_i^{(k)})^\top + \epsilon I_3$, $\epsilon > 0$. ϵI_3 provides an isotropic background metric, while the rank-one term increases metric weight along the local displacement axis. For nonzero $u_i^{(k)}$, the leading eigenvector of $g_i^{(k)}$ identifies this axis, and its spectral gap $\|u_i^{(k)}\|_2^2$ records the squared displacement magnitude. Together, these states form a discrete metric field encoding the spatial distribution of displacement axes and magnitudes. We then compute SPD Ricci-type curvature on the metric field to characterize neighborhood differences in displacement axes and magnitudes.

At the graph level, the mean curvature magnitude $\overline{F}^{(k)} = N^{-1} \sum_i \| \mathrm{Ric}_i^{\mathrm{SPD},(k)} \|_F$ generally increases with the Hessian eigenvalue λ_k (Fig. 1C bottom). The comparison uses a fixed contact graph and normalized modes. Thus, variations in curvature response reflect the spatial distribution of displacement axes and magnitudes, rather than changes in overall displacement amplitude. The observed trend is consistent with more coherent motion in softer modes and finer-scale or more localized variation in stiffer modes, which can generate stronger neighborhood metric discrepancies. Across the five representative modes shown in Fig. 1C (top), both the curvature magnitude

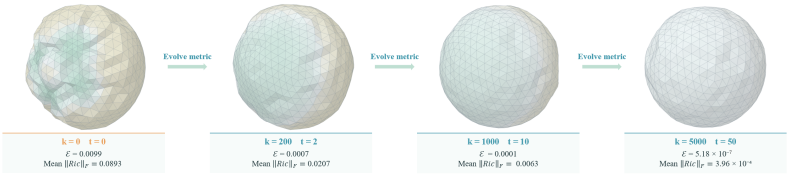


Figure 2: **Metric evolution under SPD Ricci Flow.**

$\| \mathrm{Ric}^{\mathrm{SPD}}_{\mathrm{ALA188}} \|_F$ at the highlighted site ALA188 and the mean curvature magnitude over its neighbors increase with mode stiffness. All these observations illustrate how SPD Ricci-type curvature reflects the spatial organization of displacement axes and magnitudes across molecular motion modes.

4.2 Discrete Updates and Geometric Illustration

Discrete updates. We discretize the continuous-time SPD Ricci Flow Eq. (7) using an explicit Euler step, absorbing the factor of two into the step-size hyperparameter η:

$$g_i^{(k+1)} = g_i^{(k)} - \eta \, \mathrm{Ric}_i^{\mathrm{SPD}}(g^{(k)}), \qquad \eta > 0. \tag{14}$$

At each step, curvature is recomputed from the current metric field, so the evolving node states determine their subsequent updates.

Geometric illustration. Fig. 2 visualizes SPD Ricci Flow on a fixed spherical mesh graph initialized with local metric perturbations. As the metric field evolves, the reconstructed surfaces become smoother as these perturbations relax. At the reported checkpoints, both the metric-field energy $\mathcal{E}_{\psi_\beta}$ (Eq. (57)) and the mean curvature magnitude decrease, illustrating the relaxation behavior consistent with our continuous-time energy-dissipation analysis Thm. 3.2. Details are provided in the App. A.3.

4.3 SPD Ricci Flow Model

We turn the discrete update Eq. (14) into a learnable propagation layer. Given node states $G_i^{(\ell)} \in \mathrm{SPD}_d$, a shared BiMap (Huang & Van Gool, 2017) produces $Z_i^{(\ell)} = W_\ell^\top G_i^{(\ell)} W_\ell$, where $W_\ell \in \mathbb{R}^{d \times d}$ is learnable and invertible. We compute curvature on the transformed metric field $Z^{(\ell)}$, apply the update and a spectral activation, and map back through the tied inverse:

$$G_i^{(\ell+1)} = W_\ell^{-\top} \sigma_{\mathrm{SPD}} \Big(Z_i^{(\ell)} - \eta \, \mathrm{Ric}_i^{\mathrm{SPD}}(Z^{(\ell)}) \Big) \, W_\ell^{-1}, \qquad \eta > 0. \tag{15}$$

The tied inverse exactly reverses the forward BiMap without an independently learned output transformation. For a symmetric matrix $A = Q \, \mathrm{diag}(\mu_1, \ldots, \mu_n) Q^\top$, we define $\sigma_{\mathrm{SPD}}(A) = Q \, \mathrm{diag}(\phi(\mu_1), \ldots, \phi(\mu_n)) Q^\top$, where $\phi(\mu) \geq \delta > 0$. This activation ensures positive definiteness, which the inverse BiMap preserves; each layer thus returns an SPD state still serving as both a node representation and a local metric. For downstream tasks, an encoder maps input attributes to initial SPD states $G^{(0)} = \mathcal{E}_{\mathrm{task}}(X)$; L propagation layers and a task-specific decoder follow. The encoder and decoder vary by tasks, while the curvature-driven propagation retains the same form.

5 Experiments and Results

In this section, we evaluate SPD Ricci Flow on node classification, molecular property prediction, and financial spillover prediction. We also examine curvature responses around financial news events and use controlled ablations to assess the contributions of graph propagation, adaptive weighting, and the curvature update.

5.1 Node Classification

We evaluate SPD Ricci Flow on five node-classification benchmarks, covering heterophilic and citation graphs. We report ROC-AUC on Questions and accuracy on the remaining datasets, with means and standard deviations over ten seeds. As shown in Table 1, SPD Ricci Flow achieves the highest mean score on four benchmarks and ranks second on Cora-ML. Details are provided in App.A.4.

Table 1: **Node classification results** (%). Standard deviations are shown as subscripts. **Bold** and underline indicate the highest and second-highest mean scores, respectively.

Method	Wisconsin	Questions	Cora-Full	DBLP	Cora-ML
GCN	$61.60_{7.00}$	$50.21_{2.24}$	$68.06_{0.98}$	$83.93_{0.34}$	$87.07_{1.21}$
GAT	$60.00_{11.00}$	$65.47_{0.88}$	$67.55_{1.23}$	$80.61_{1.21}$	$84.12_{0.55}$
DirGNN	$80.50_{5.50}$	$59.95_{0.79}$	$67.80_{0.53}$	$81.22_{0.54}$	$85.66_{0.31}$
FSGNN	$87.60_{5.10}$	$71.41_{0.32}$	$71.90_{0.65}$	$83.31_{0.55}$	$\mathbf{89.44}_{0.43}$
GRAND	$84.00_{7.50}$	$54.90_{2.12}$	$67.66_{1.01}$	$84.60_{0.99}$	$88.49_{0.81}$
GRAND++	$81.50_{6.00}$	$60.14_{0.88}$	$67.53_{0.74}$	$85.21_{0.24}$	$88.44_{0.53}$
ACMP	$86.50_{5.00}$	$71.92_{0.55}$	OOM	$82.31_{0.44}$	$76.11_{2.12}$
HiD-Net	$81.30_{6.60}$	$63.77_{0.85}$	$68.11_{0.64}$	$84.92_{0.31}$	$89.00_{0.51}$
GNRF	$\underline{88.00}_{2.00}$	$\underline{73.86}_{1.18}$	$\underline{72.12}_{0.50}$	$85.73_{0.76}$	$89.18_{0.19}$
GEGCN	$77.40_{13.92}$	$\underline{73.86}_{1.18}$	$69.00_{0.82}$	$85.58_{0.34}$	$87.24_{1.48}$
SPDGNN	$59.20_{8.68}$	$50.34_{11.20}$	$70.84_{0.66}$	$\underline{85.75}_{0.35}$	$88.41_{0.93}$
SPD Ricci Flow	$\mathbf{89.80}_{3.52}$	$\mathbf{75.23}_{1.36}$	$\mathbf{72.67}_{0.75}$	$\mathbf{85.90}_{0.39}$	$\underline{89.20}_{1.08}$

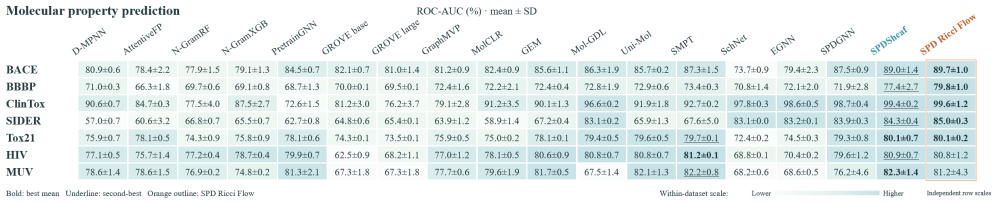


Figure 3: **Molecular property prediction results** (ROC-AUC, %; mean ± standard deviation).

5.2 Molecular Property Prediction

We further evaluate SPD Ricci Flow on graph-level molecular property prediction across seven MoleculeNet benchmarks. We adopt the architecture of SPDSheaf (Peng et al., 2026), replacing its SPD SheafNN layers with SPD Ricci Flow. Fig.3 reports ROC-AUC: our method achieves the highest score on four datasets and ties for best on Tox21. Details are provided in App.A.5.

5.3 Curvature Responses to Financial News

After evaluating SPD Ricci Flow's node- and graph-level predictive performance, we now examine whether SPD Ricci-type curvature captures changes in a stock's local metric relationships around financial news events. Using daily price and news data for 3,273 stocks over 2018–2022 (Dong et al., 2024), we build stock graphs from recent return correlations after removing the market component. Each stock's SPD matrix summarizes the joint variation of price–volume features. News dates provide timestamps for examining metric-field changes.

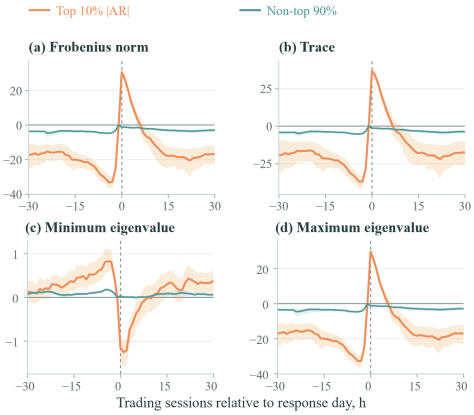


Figure 4: **Adjusted SPD Ricci-type curvature responses around financial news events.**

We identify strong-response events as the 10% of news events with the largest absolute differences between the stock's event-day stock and market returns; the rest form a comparison group. For each curvature statistic, we subtract the daily mean across stocks and measure changes relative to the trading day before the news event. We then subtract the average corresponding change around matched no-news dates for the same stock. Details are provided in App.A.6.

As shown in Fig.4, strong-response events exhibit a positive adjusted response in the SPD curvature Frobenius norm, accompanied by upward and downward responses in the maximum and minimum eigenvalues, respectively. By comparison, the remaining news events show adjusted responses close to zero. These responses are most pronounced around the news date and subsequently weaken. This pattern may reflect a temporary divergence in price–volume behavior between the stock and its

Table 2: **Financial spillover prediction and ablations**. Scores are multiplied by 100; subscripts denote standard deviations over five seeds. **Bold**/underline mark the best/second-best means.

	Method / Variant	Macro-AUPRC	NDCG@5	Precision@5	Hit@5
Baseline	DirGNN	$33.93_{0.70}$	$30.48_{0.76}$	$12.53_{0.33}$	$\underline{63.28}_{0.74}$
	GAT	$34.10_{0.18}$	$31.11_{0.48}$	$\underline{12.85}_{0.14}$	$61.14_{0.94}$
	GraphSAGE	$\underline{34.31}_{0.46}$	$\underline{31.38}_{0.53}$	$12.84_{0.18}$	$61.19_{0.95}$
	GNRF	$34.02_{0.87}$	$30.80_{1.41}$	$12.68_{0.68}$	$62.96_{2.35}$
	HiD-Net	$33.24_{0.14}$	$29.73_{0.18}$	$12.42_{0.08}$	$62.47_{0.58}$
	SPDGNN	$31.94_{1.98}$	$28.20_{2.92}$	$12.04_{1.16}$	$59.21_{6.27}$
	SPDNet	$31.49_{1.59}$	$27.92_{2.23}$	$11.90_{0.81}$	$59.33_{4.33}$
Ablation	Uniform weights	$\underline{34.58}_{0.43}$	$\underline{31.79}_{0.58}$	$\underline{13.32}_{0.21}$	$\underline{65.75}_{1.20}$
	Uniform + $\mathcal{L}_i$	$33.10_{0.21}$	$29.91_{0.38}$	$12.66_{0.21}$	$63.12_{1.08}$
	No Ricci Flow layer	$31.13_{0.45}$	$27.56_{0.39}$	$11.71_{0.18}$	$58.90_{1.31}$
	SPD Ricci Flow	$\mathbf{35.24}_{1.01}$	$\mathbf{32.67}_{1.28}$	$\mathbf{13.57}_{0.40}$	$\mathbf{66.11}_{1.66}$

neighbors, followed by a return toward their usual level of similarity. Similar patterns persist after removing explicit abnormal-return coordinates from the SPD states, including when only range and volume features remain (Appendix A.6). These observations illustrate the curvature's sensitivity to event-associated changes in neighboring metric relationships.

5.4 Predicting News-Driven Spillovers with SPD Ricci Flow

We next test whether curvature-driven propagation helps predict neighboring stocks' responses to news. Given an event associated with stock s, the task is to identify which one- or two-hop neighbors will exhibit unusually large abnormal returns over the following 10 trading days. Inputs consist of the source news and the graph and stock states from the preceding trading day. Labels compare future cumulative abnormal returns with matched no-news periods and historical volatility.

We initialize SPD node representations from pre-event price–volume information, incorporating the source news into the source stock's representation. The SPD Ricci Flow then propagate these states over the graph. After propagation, a task-specific prediction head estimates each candidate's response probability. Details are provided in Appendix A.7.

Predictive performance. In Table 2, SPD Ricci Flow achieves the highest mean scores on all four metrics. Its NDCG@5 and Hit@5 results indicateimproved within-event ranking of responding stocks. It also outperforms SPDNet and SPDGNN, suggesting that curvature-driven metric evolution is more effective in this task than using SPD representations alone or SPD message-passing mechanism. The ablations below further examine the contributions of its propagation components.

Ablation study. Table 2 (bottom) reports the ablation results. Removing SPD Ricci Flow layers causes the largest NDCG@5 drop, from 32.67 to 27.56, demonstrating the substantial contribution of the proposed propagation module beyond the source perturbation and initial stock representations. Removing SPD-dependent edge weighting (*Uniform weights*) yields a smaller drop to 31.79. With the same uniform weighting, replacing the curvature update with $\mathcal{L}_i$ (*Uniform+*$\mathcal{L}_i$) further lowers NDCG@5 to 29.91, supporting the metric coupling through $g_i\mathcal{L}_i + \mathcal{L}_i g_i$. Thus, adaptive weighting alone does not account for the full model's performance. Additional component-wise ablations on node classification and molecular property prediction are further reported in Table 6

6 Conclusion

We presented **SPD Ricci Flow**, which unifies node representations and local metrics in the same SPD state. Matrix-valued Ricci-type curvature evolves these metrics directly, thereby updating the node representations themselves. Our analysis establishes continuous-time energy dissipation and shows that, under graph-Laplacian consistency, the linearized continuum limit matches the heat-type principal part of Ricci–DeTurck evolution. Experiments on node classification, molecular property prediction, and financial spillover prediction support the effectiveness of this propagation mechanism. These results support metric evolution itself as a mechanism for representation learning, with curvature both describing local geometry and driving its evolution.

AI USE STATEMENT

In this work, we used generative AI tools (GPT and Codex) to assist with editing and debugging the model implementation, English-language polishing, sentence rephrasing and shortening, LaTeX formatting, and literature search and summarization. During manuscript revision, GPT also assisted with discussions of reviewer comments, including checking existing mathematical notation and discussing the evaluation protocol and the interpretation of reported results.

The research ideas, mathematical formulations and proofs, and original experimental design were developed by the authors, rather than generated by these tools. Generative AI was not used to create scientific images.

The authors reviewed all AI-assisted work, checked revised text for accuracy and consistency with the intended meaning, verified literature summaries and references against original sources, inspected and tested the modified code, and cross-checked tables and figures against experimental outputs. We take responsibility for the final content of this work, including all text, claims, code, and other artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The code is available at `https://anonymous.4open.science/r/SPDRicci/README.md`.

REFERENCES

A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, and I. Bahar. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophysical Journal*, 80(1):505–515, 2001. ISSN 0006-3495. doi: https://doi.org/10.1016/S0006-3495(01)76033-X. URL https://www.sciencedirect.com/science/article/pii/S000634950176033X.

Aleksandar Bojchevski and Stephan Günnemann. Deep gaussian embedding of graphs: Unsupervised inductive learning via ranking. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=r1ZdKJ-0W.

Ben Chamberlain, James Rowbottom, Maria I Gorinova, Michael Bronstein, Stefan Webb, and Emanuele Rossi. Grand: Graph neural diffusion. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1407–1418. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/chamberlain21a.html.

Jialong Chen, Bowen Deng, Zhen WANG, Chuan Chen, and Zibin Zheng. Graph neural ricci flow: Evolving feature from a curvature perspective. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=7b2JrzdLhA.

Zihan Dong, Xinyu Fan, and Zhiyuan Peng. Fnspid: A comprehensive financial news dataset in time series, 2024.

Pemra Doruker, Ali Rana Atilgan, and Ivet Bahar. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins: Structure, Function, and Bioinformatics*, 40(3):512–524, 2000. doi: https://doi.org/10.1002/1097-0134(20000815)40:3<512::AID-PROT180>3.0.CO;2-M. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/1097-0134%2820000815%2940%3A3%3C512%3A%3AAID-PROT180%3E3.0.CO%3B2-M.

Eran Eyal, Lee-Wei Yang, and Ivet Bahar. Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics*, 22(21):2619–2627, 11 2006. ISSN 1367-4803. doi: 10.1093/bioinformatics/btl448. URL https://doi.org/10.1093/bioinformatics/btl448.

Robin Forman. Bochner's method for cell complexes and combinatorial ricci curvature. *Discrete and Computational Geometry*, 29:323–374, 01 2003. doi: 10.1007/s00454-002-0743-x.

Richard S. Hamilton. Three-manifolds with positive ricci curvature. *Journal of Differential Geometry*, 17:255–306, 1982. URL https://api.semanticscholar.org/CorpusID:117938273.

William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pp. 1025–1035, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964.

Zhiwu Huang and Luc Van Gool. A Riemannian Network for SPD Matrix Learning. In *AAAI Conference on Artificial Intelligence*, pp. 2036–2042, 2017. doi: 10.1609/AAAI.V31I1.10866. URL https://mlanthology.org/aaai/2017/huang2017aaai-riemannian/.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=SJU4ayYgl.

Xin Lai, Shuliang Bai, and Yong Lin. Normalized discrete ricci flow used in community detection. *Physica A: Statistical Mechanics and its Applications*, 597:127251, 2022. ISSN 0378-4371. doi: https://doi.org/10.1016/j.physa.2022.127251. URL https://www.sciencedirect.com/science/article/pii/S0378437122002242.

John M. Lee. *Introduction to Riemannian Manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2018.

Yibo Li, Xiao Wang, Hongrui Liu, and Chuan Shi. A Generalized Neural Diffusion Framework on Graphs. In *AAAI Conference on Artificial Intelligence*, pp. 8707–8715, 2024. doi: 10.1609/AAAI.V38I8.28716. URL https://mlanthology.org/aaai/2024/li2024aaai-generalized/.

Sunil Kumar Maurya, Xin Liu, and Tsuyoshi Murata. Simplifying approach to node classification in graph neural networks. *Journal of Computational Science*, 62:101695, 2022. ISSN 1877-7503. doi: https://doi.org/10.1016/j.jocs.2022.101695. URL https://www.sciencedirect.com/science/article/pii/S1877750322000990.

Chien-Chun Ni, Yu-Yao Lin, Feng Luo, and Jie Gao. Community detection on networks with ricci flow. *Scientific Reports*, 9, 2019. URL https://api.semanticscholar.org/CorpusID:195848169.

Yann Ollivier. Ricci curvature of markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009. ISSN 0022-1236. doi: https://doi.org/10.1016/j.jfa.2008.11.001. URL https://www.sciencedirect.com/science/article/pii/S002212360800493X.

Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-gcn: Geometric graph convolutional networks. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1e2agrFvS.

Yuhan Peng, Junwen Dong, Yuzhi Zeng, Hao Li, Ce Ju, Huitao Feng, Diaaeldin Taha, Anna Wienhard, and Kelin Xia. Sheaf neural networks on SPD manifolds: Second-order geometric representation learning. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=Qbe2cfNfn1.

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 *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=tJbbQfw-5wv.

Emanuele Rossi, Bertrand Charpentier, Francesco Di Giovanni, Fabrizio Frasca, Stephan Günnemann, and Michael M. Bronstein. Edge directionality improves learning on heterophilic graphs. In Soledad Villar and Benjamin Chamberlain (eds.), *Proceedings of the Second Learning on Graphs Conference*, volume 231 of *Proceedings of Machine Learning Research*, pp. 25:1–25:27. PMLR, 27–30 Nov 2024. URL https://proceedings.mlr.press/v231/rossi24a.html.

Matthew Thorpe, Tan Minh Nguyen, Hedi Xia, Thomas Strohmer, Andrea Bertozzi, Stanley Osher, and Bao Wang. GRAND++: Graph neural diffusion with a source term. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=EMxu-dzvJk.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=rJXMpikCZ.

Xudong Wang, Chris Ding, Tongxin Li, and Jicong Fan. Adaptive riemannian graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 26606–26614, 2026.

Yuelin Wang, Kai Yi, Xinliang Liu, Yu Guang Wang, and Shi Jin. ACMP: Allen-cahn message passing with attractive and repulsive forces for graph neural networks. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=4fZc_79Lrqs.

Kelin Xia. Multiscale virtual particle based elastic network model (mvp-enm) for normal mode analysis of large-sized biomolecules. *Physical Chemistry Chemical Physics*, 20(1):658–669, 01 2018. ISSN 1463-9076. doi: 10.1039/c7cp07177a. URL https://doi.org/10.1039/c7cp07177a.

Kelin Xia, Kristopher Opron, and Guowei Wei. Multiscale gaussian network model (mgnm) and multiscale anisotropic network model (manm). *The Journal of chemical physics*, 143 20:204106, 2015. URL https://api.semanticscholar.org/CorpusID:14665300.

Wei Zhao, Federico López, J. Maxwell Riestenberg, Michael Strube, Diaaeldin Taha, and Steve Trettel. Modeling graphs beyond hyperbolic: Graph neural networks in symmetric positive definite matrices. In *Machine Learning and Knowledge Discovery in Databases: Research Track*, volume 14171 of *Lecture Notes in Computer Science*, pp. 122–139. Springer, 2023. doi: 10.1007/978-3-031-43418-1_8. URL https://doi.org/10.1007/978-3-031-43418-1_8.

A Appendix

A.1 Theory of Bochner–Weitzenböck Ricci Metric-Field Diffusion

This appendix develops the theoretical basis of the SPD Ricci-type curvature used in the main text. The theory has two complementary layers. The first is an intrinsic graph-level layer: the proposed response is a tensor-valued relaxation of an SPD metric field, and in log-metric coordinates it reduces exactly to graph heat diffusion of the metric components in the commuting case, and reduces to the first order of this heat diffusion in the non-commuting case. The second is a continuum verification layer: under standard local graph-discretization assumptions, the scaled log-metric graph heat operator converges to the Laplace-type principal part that appears in harmonic-coordinate or Ricci–DeTurck formulations of Ricci flow. These two layers are then connected to an energy-dissipation principle: the same operator decreases a natural Dirichlet-type energy of an SPD-valued metric field, and its vanishing characterizes local harmonic equilibrium.

The terminology "Bochner–Weitzenböck" is used in an operator-structural sense. In the smooth identity

$$\Delta_H = \nabla^*\nabla + \mathrm{Ric},$$

curvature is identified through the part of a Laplace-type operator that acts tensorially on the field. Our graph construction follows this route at the level available to the model: it builds a Laplace-type tensor in whitened SPD coordinates and lifts it back to a symmetric metric variation. The resulting operator is therefore a BW-inspired tensorial metric response, rather than a scalar curvature decoration placed on a fixed graph.

Fixed-frame SPD metric fields on graphs. Let $G = (V, E, w)$ be a finite undirected weighted graph, with symmetric weights $w_{ij} = w_{ji} \geq 0$. Each node $i \in V$ carries a symmetric positive definite matrix $g_i \in \mathrm{SPD}_d$. We interpret g_i simultaneously as a structured node representation and as a local metric tensor on a shared d-dimensional feature frame. In this work the comparison between g_i and g_j is made in this fixed frame. Edge transports between local frames would lead to a gauge-covariant extensor, which is a natural but separate direction.

Throughout this appendix, in all affine-invariant distances, Dirichlet energies, and curvature magnitudes, the norm is the Frobenius norm denoted by $\|\cdot\|_F$. When a matrix norm appears without a subscript in local expansion estimates, it may take to be any fixed matrix norm, but changing this auxiliary norm only changes the constants in the estimates, since all spaces are finite-dimensional.

The manifold SPD_d is equipped with the affine-invariant Riemannian metric (AIRM)

$$\langle A, B\rangle_g = \mathrm{tr}(g^{-1}Ag^{-1}B), \qquad A, B \in T_g\mathrm{SPD}_d \simeq \mathrm{Sym}_d.$$

The corresponding exponential and logarithm maps are

$$\mathrm{Exp}_g(A) = g^{1/2}\exp\Big(g^{-1/2}Ag^{-1/2}\Big)\,g^{1/2},$$

and

$$\mathrm{Log}_g(h) = g^{1/2}\log\Big(g^{-1/2}hg^{-1/2}\Big)\,g^{1/2}.$$

Thus the affine-invariant distance is

$$d_{\mathrm{SPD}}(g, h) = \Big\|\log(g^{-1/2}hg^{-1/2})\Big\|_F.$$

For an oriented edge $i \to j$, define the relative log-metric displacement

$$L_{i\to j} = \log(g_i^{-1/2}g_jg_i^{-1/2}), \qquad r_{ij} = \|L_{i\to j}\|_F^2. \tag{16}$$

The matrix $L_{i\to j} \in \mathrm{Sym}_d$ is the logarithmic discrepancy from g_i to g_j after whitening by the local metric g_i. It is the coordinate form, at the identity, of the tangent vector $\mathrm{Log}_{g_i}(g_j)$:

$$\mathrm{Log}_{g_i}(g_j) = g_i^{1/2}L_{i\to j}g_i^{1/2}.$$

Metric-field energy and whitened tensor. The local roughness of an SPD metric field can be measured by a graph Dirichlet energy. We use a slightly more general form because it also covers the heat-kernel weighting used in the implementation. Let $\psi : [0, \infty) \to \mathbb{R}$ be continuously differentiable with $\psi'(r) \geq 0$ for any $r \in [0, \infty)$. Define

$$\mathcal{E}_\psi(g) = \sum_{\{i,j\}\in E} w_{ij}\,\psi(r_{ij}), \qquad r_{ij} = d^2_{\mathrm{SPD}}(g_i, g_j). \tag{17}$$

The ordinary affine-invariant Dirichlet energy corresponds to $\psi(r) = r/2$.

Corollary A.1 (Graph tensor from the SPD Dirichlet energy)**.** *For the energy in Eq.* (17)*, the negative affine-invariant Riemannian gradient with respect to g_i is*

$$-\nabla^{\mathrm{AIRM}}_{g_i}\mathcal{E}_\psi(g) = g_i^{1/2}S_i(g)g_i^{1/2}, \tag{18}$$

where

$$S_i(g) = 2\sum_{j\sim i} w_{ij}\,\psi'(r_{ij})\,L_{i\to j}. \tag{19}$$

We call $S_i(g)$ the whitened graph tensor of the SPD metric field at node i.

Proof. On any Riemannian manifold, away from the cut locus, the gradient of the squared distance satisfies

$$\nabla_x\frac{1}{2}d^2(x, y) = -\,\mathrm{Log}_x(y).$$

The manifold SPD_d with the affine-invariant metric is geodesically complete and the logarithm is globally defined, so this identity applies for all $g_i, g_j \in \mathrm{SPD}_d$. Therefore,

$$\nabla^{\mathrm{AIRM}}_{g_i}r_{ij} = -2\,\mathrm{Log}_{g_i}(g_j).$$

By the chain rule,

$$\nabla_{g_i}^{\mathrm{AIRM}}\left[w_{ij}\psi(r_{ij})\right] = -2w_{ij}\psi'(r_{ij})\,\mathrm{Log}_{g_i}(g_j).$$

Summing over all neighbors of i gives

$$-\nabla_{g_i}^{\mathrm{AIRM}}\mathcal{E}_\psi(g) = 2\sum_{j\sim i} w_{ij}\psi'(r_{ij})\,\mathrm{Log}_{g_i}(g_j).$$

Using

$$\mathrm{Log}_{g_i}(g_j) = g_i^{1/2} L_{i\to j} g_i^{1/2} \tag{20}$$

then yields Eq. (18) with $S_i(g)$ as in Eq. (19). □

The heat-kernel weighting used by the model is included in this framework. If

$$a_{ij}(g) = w_{ij}\exp(-\beta r_{ij}), \qquad \beta > 0,$$

then choose

$$\psi_\beta(r) = \frac{1-\exp(-\beta r)}{2\beta}. \tag{21}$$

Since $2\psi'_\beta(r) = \exp(-\beta r)$, Eq. (19) becomes

$$S_i(g) = \sum_{j\sim i} w_{ij}\exp(-\beta r_{ij}) L_{i\to j}. \tag{22}$$

Thus the implemented weighted log aggregation is the whitened tensor of a robust, increasing Dirichlet-type energy. The choice ψ_β is concave, so it should not be used to claim global geodesic convexity. Its role here is to identify the variational origin of the state-dependent weighted tensor.

The SPD Ricci-type curvature. A basic principle in Riemannian geometry is that curvature can be detected through the discrepancy between natural second-order operators acting on geometric fields. While scalar functions have a canonical Laplace–Beltrami operator, vector fields, differential forms, and tensor fields carry additional connection data. Consequently, there are several natural Laplace-type operators with the same second-order principal part but different lower-order terms. The difference between these operators is often precisely a curvature action. The Bochner–Weitzenböck formula is the prototypical example of this phenomenon.

We recall the one-form version, which is the form most directly tied to the Ricci tensor. Let (M, g) be a smooth Riemannian manifold. We use the nonnegative Hodge Laplacian

$$\Delta_H = d\delta + \delta d$$

and the rough Laplacian

$$\nabla^*\nabla = -\operatorname{tr}_g \nabla^2 = -\nabla^k\nabla_k.$$

Here ∇ is the Levi–Civita connection and indices are raised using g.

Proposition A.1 (Bochner–Weitzenböck formula for one-forms)**.** *Let (M, g) be a smooth Riemannian manifold and let $\omega \in \Omega^1(M)$ be a smooth one-form. Then*

$$\Delta_H\omega = \nabla^*\nabla\omega + \mathrm{Ric}^{\#}(\omega), \tag{23}$$

where $\mathrm{Ric}^{\#}$ denotes the natural action of the Ricci tensor on one-forms. In local coordinates, this identity reads

$$(\Delta_H\omega)_\alpha = -\nabla^k\nabla_k\omega_\alpha + \mathrm{Ric}_\alpha^{\ \beta}\,\omega_\beta. \tag{24}$$

Equivalently, if X is a vector field and $\omega = X^\flat$ is its metric dual, then

$$\Delta_H(X^\flat) = \left(\nabla^*\nabla X + \mathrm{Ric}(X)\right)^\flat.$$

Thus the difference between the Hodge Laplacian and the rough Laplacian is a zeroth-order curvature action.

Proof sketch. The statement is local, so fix a point $p \in M$ and choose normal coordinates at p. At this point the Christoffel symbols vanish, although their derivatives, and hence the curvature, need not vanish. For a one-form $\omega = \omega_\alpha dx^\alpha$, we have

$$(d\omega)_{\alpha\beta} = \nabla_\alpha \omega_\beta - \nabla_\beta \omega_\alpha, \qquad \delta\omega = -\nabla^\beta \omega_\beta.$$

Therefore,

$$(\delta d\omega)_\alpha = -\nabla^\beta \left(\nabla_\beta \omega_\alpha - \nabla_\alpha \omega_\beta\right),$$

and

$$(d\delta\omega)_\alpha = -\nabla_\alpha \nabla^\beta \omega_\beta.$$

Adding the two terms gives

$$(\Delta_H \omega)_\alpha = -\nabla^\beta \nabla_\beta \omega_\alpha + \nabla^\beta \nabla_\alpha \omega_\beta - \nabla_\alpha \nabla^\beta \omega_\beta.$$

The first term is exactly $(\nabla^*\nabla\omega)_\alpha$. The remaining two terms measure the failure of covariant derivatives to commute:

$$\nabla^\beta \nabla_\alpha \omega_\beta - \nabla_\alpha \nabla^\beta \omega_\beta = \mathrm{Ric}_\alpha^{\ \beta}\, \omega_\beta.$$

This is the Ricci contraction of the curvature tensor. Substituting this commutator identity yields Eq. (24), and hence Eq. (23). □

Proposition A.1 can be read in the operational form $\mathrm{Ric}^{\#} = \Delta_H - \nabla^*\nabla$; thus Ricci curvature is the tensorial action detected by comparing a Hodge-type Laplacian with the purely rough diffusion induced by the connection. This is the sense in which the Bochner–Weitzenböck formula turns a second-order diffusion principle into a curvature operator. In the present graph model, all SPD states are represented in a shared feature frame, and no explicit edge connection/parallel transport is introduced. We therefore interpret this case as the vanishing of the rough Laplacian term, i.e. $\nabla^*\nabla = 0$; for more complicated case where non trivial parallel transport is introduced, the operation of gauge covariance would be induced into the rough Laplacian term, which remains as future work. Thus, we use the intrinsic AIRM relative-log Laplace-type response as the discrete object suggested by the BW principle. This is a mechanism-inspired construction rather than a claim of an exact discrete Hodge-minus-rough Laplacian identity. The corresponding Laplace-type object is the weighted aggregation of affine-invariant relative logarithms $L_{i\to j}$. This aggregation measures how the local metric differs from its neighbors in whitened tangent coordinates. Following the Bochner–Weitzenböck principle, we interpret this graph metric tensor as the metric-evolution response and lift it to a symmetric metric variation.

On a graph carrying an SPD metric field, the object directly available from neighboring metrics is the affine-invariant relative log displacement $L_{i\to j}$. Its weighted aggregation gives the whitened metric-field tensor

$$S_i(g) = \sum_{j\sim i} a_{ij}(g) L_{i\to j}, \qquad a_{ij}(g) = w_{ij} \exp(-\beta \|L_{i\to j}\|_F^2)$$

for the model weighting. With the nonnegative graph-Laplacian sign convention, the corresponding Hodge-type discrete metric Laplacian is

$$\mathcal{B}_i(g) := \left(\Delta_H^{\mathrm{disc}} g\right)_i^{\mathrm{wh}} = -S_i(g) = -\sum_{j\sim i} a_{ij}(g) \log(g_i^{-1/2} g_j g_i^{-1/2}). \tag{25}$$

The superscript "wh" emphasizes that this quantity is written in the whitened tangent coordinates at g_i. In the commuting log-metric case $g_i = \mathrm{diag}(e^{h_i^1}, \ldots, e^{h_i^d})$, Eq. (25) reduces componentwise to

$$\mathcal{B}_i^a(g) = \sum_{j\sim i} a_{ij}(g)(h_i^a - h_j^a),$$

which is exactly the graph Laplacian of the log-metric component. Hence $\mathcal{B}_i(g)$ is the discrete Laplace-type operator that measures how the local metric differs from its neighboring metric field.

Following the Bochner–Weitzenböck principle, we use this Hodge-type metric Laplacian as the curvature-driving endomorphism. However, $\mathcal{B}_i(g)$ is not yet a symmetric $(0,2)$ metric variation; it

is an endomorphism written in whitened coordinates. To convert an endomorphism $A_i \in \mathrm{Sym}_d$ into a metric tensor variation at g_i, we use the infinitesimal congruence action

$$\mathrm{Sym}_{g_i}(A_i) := \frac{1}{2}(g_i A_i + A_i g_i) = \left.\frac{d}{dt}\right|_{t=0} \exp\left(\frac{t}{2}A_i\right) g_i \exp\left(\frac{t}{2}A_i\right). \tag{26}$$

Applying this conversion to the endomorphism $A_i = \mathcal{B}_i(g) = -S_i(g)$ gives

$$\mathrm{Sym}_{g_i}(\mathcal{B}_i(g)) = -\frac{1}{2}\left(g_i S_i(g) + S_i(g) g_i\right).$$

This tensor is the discrete SPD Ricci-type curvature. In particular, when $g_i = I$, the response reduces to

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -S_i(g) = \left(\Delta_H^{\mathrm{disc}} g\right)_i^{\mathrm{wh}},$$

so the definition agrees directly with the Hodge-type Laplacian reading. Away from the identity, the congruence lift is the natural tensorial conversion that turns the whitened BW action into an evolution equation for the metric itself. Above summarized to the following definition.

Definition A.1 (SPD Ricci-type curvature)**.** *Given the whitened tensor $S_i(g)$, define the node-wise SPD Ricci-type curvature by*

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -\frac{1}{2}\left(g_i S_i(g) + S_i(g) g_i\right) \in \mathrm{Sym}_d. \tag{27}$$

The corresponding continuous-time SPD Ricci flow is

$$\partial_t g_i = -2\,\mathrm{Ric}_i^{\mathrm{SPD}}(g) = g_i S_i(g) + S_i(g) g_i. \tag{28}$$

Equivalently, the associated flow $\partial_t g_i = -2\,\mathrm{Ric}_i^{\mathrm{SPD}} = g_i S_i + S_i g_i$ is the infinitesimal congruence deformation generated by the whitened tensor S_i.

Lemma 1 (Vanishing response and local equilibrium)**.** *For each i, since $g_i \succ 0$,*

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = 0 \quad \Longleftrightarrow \quad S_i(g) = 0. \tag{29}$$

Proof. It is immediate from Eq. (27) that $S_i = 0$ implies $\mathrm{Ric}_i^{\mathrm{SPD}} = 0$. Conversely, suppose $g_i S_i + S_i g_i = 0$. Diagonalize $g_i = U \Lambda U^\top$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ and $\lambda_a > 0$. Let $\widetilde{S} = U^\top S_i U$. In this basis,

$$(\Lambda \widetilde{S} + \widetilde{S}\Lambda)_{ab} = (\lambda_a + \lambda_b)\widetilde{S}_{ab}.$$

Since $\lambda_a + \lambda_b > 0$ for all a, b, the equality $\Lambda\widetilde{S} + \widetilde{S}\Lambda = 0$ implies $\widetilde{S}_{ab} = 0$ for all a, b. Hence $S_i = 0$. □

Exact verification in the commuting case. We first record a simple exact case. This result is not the main theoretical contribution, but it is a useful normalization check: when all local metrics commute, the SPD Ricci flow becomes ordinary graph heat diffusion on the logarithmic metric components.

Lemma 2 (Commuting metrics give log-metric graph heat flow)**.** *Assume all node metrics are diagonal in a common basis,*

$$g_i = \mathrm{diag}\left(e^{h_i^1}, \ldots, e^{h_i^d}\right).$$

If $S_i = \sum_{j\sim i} a_{ij} L_{i\to j}$ for scalar nonnegative coefficients a_{ij}, possibly depending on the current SPD field, then the SPD Ricci flow in Eq. (28) *is equivalent to*

$$\partial_t h_i^a = 2\sum_{j\sim i} a_{ij}(h_j^a - h_i^a), \qquad a = 1, \ldots, d. \tag{30}$$

Proof. In the common diagonal basis,

$$g_i^{-1/2} g_j g_i^{-1/2} = \mathrm{diag}\left(e^{h_j^1 - h_i^1}, \ldots, e^{h_j^d - h_i^d}\right),$$

and therefore

$$L_{i\to j} = \mathrm{diag}\left(h_j^1 - h_i^1, \ldots, h_j^d - h_i^d\right).$$

Thus S_i is diagonal with entries

$$S_i^a = \sum_{j\sim i} a_{ij}(h_j^a - h_i^a).$$

Since g_i and S_i commute,

$$\partial_t g_i = g_i S_i + S_i g_i = 2 g_i S_i.$$

For the a-th diagonal entry, $\dot{g}_i^a = e^{h_i^a}\dot{h}_i^a$, while $2g_i^a S_i^a = 2e^{h_i^a} S_i^a$. Cancelling $e^{h_i^a}$ gives Eq. (30). □

Linearized verification for noncommuting SPD metrics. The previous lemma is exact because all matrices commute. For general SPD matrices, noncommutativity appears in higher-order terms. The principal linearized behavior, however, is still graph heat diffusion in log-metric coordinates. We still need to notice that noncommutativity and state-dependent weights enter higher-order terms; the graph Laplacian is the principal mechanism, not the whole model which additionally contains noncommutative metric coupling and energy structure.

Lemma 3 (Small-log expansion of the relative metric displacement)**.** *Let $g_i = \exp(H_i)$ and $g_j = \exp(H_j)$, with $H_i, H_j \in \mathrm{Sym}_d$ and $\max(\|H_i\|, \|H_j\|) \le \varepsilon$ for ε sufficiently small. Then*

$$L_{i\to j} = \log\left(e^{-H_i/2} e^{H_j} e^{-H_i/2}\right) = H_j - H_i + \mathcal{R}_{ij}, \tag{31}$$

where

$$\|\mathcal{R}_{ij}\| \le C\varepsilon^2 \|H_j - H_i\| \tag{32}$$

for a constant C depending only on the matrix dimension and the chosen operator norm.

Proof. Consider the map

$$F(A, B) = \log(e^{-A/2} e^B e^{-A/2})$$

near $(A, B) = (0, 0)$. It is analytic in a neighborhood of the origin. Its first derivative is

$$DF_{(0,0)}[\dot{A}, \dot{B}] = \dot{B} - \dot{A}, \tag{33}$$

because $e^{-A/2} e^B e^{-A/2} = I + (\dot{B} - \dot{A}) + O(\|(\dot{A}, \dot{B})\|^2)$ at first order.

Moreover $F(A, A) = 0$ for every sufficiently small symmetric A, since

$$e^{-A/2} e^A e^{-A/2} = I.$$

The sharper point is that the symmetric whitening cancels the first commutator correction. Write $B = A + \Delta$ and define

$$G_A(\Delta) = \log(e^{-A/2} e^{A+\Delta} e^{-A/2}).$$

At $\Delta = 0$, the inside of the logarithm is I, so the derivative of the logarithm is the identity map. Hence

$$DG_A(0)[\Delta] = e^{-A/2} D\exp_A[\Delta] e^{-A/2} = \int_{-1/2}^{1/2} e^{-sA} \Delta e^{sA}\, ds.$$

Expanding $e^{-sA}\Delta e^{sA} = e^{-s\,\mathrm{ad}_A}\Delta$ and integrating over the symmetric interval $[-1/2, 1/2]$ removes the odd commutator term. Consequently,

$$DG_A(0)[\Delta] = \Delta + O(\|A\|^2 \|\Delta\|).$$

The remaining Taylor terms are at least quadratic in Δ and vanish when $A = 0$; by analyticity they are bounded by $O(\|A\|\|\Delta\|^2 + \|\Delta\|^3)$ in a sufficiently small neighborhood. Since $\|A\|, \|A + \Delta\| \le \varepsilon$, this is also bounded by $O(\varepsilon^2\|\Delta\|)$. Therefore

$$F(A, B) = B - A + O\big(\|(A, B)\|^2 \|B - A\|\big).$$

which gives Eq. (32). A sufficient version is $L_{i\to j} = H_j - H_i + O(\varepsilon^2)$. □

Lemma 4 (Log-coordinate velocity of the BW lift)**.** *Let $g_i(t) = \exp(H_i(t))$ with $\|H_i(t)\| \le \varepsilon$, and suppose*

$$\partial_t g_i = g_i S_i + S_i g_i$$

for a symmetric matrix $S_i(t)$. Then

$$\dot{H}_i = 2S_i + \mathcal{Q}_i, \qquad \|\mathcal{Q}_i\| \le C\varepsilon^2 \|S_i\|. \tag{34}$$

Proof. The differential of the matrix exponential is

$$D\exp_H[A] = \int_0^1 e^{(1-s)H} A e^{sH}\, ds. \tag{35}$$

Expanding for small H gives

$$D\exp_H[A] = A + \frac{1}{2}(HA + AH) + O(\|H\|^2\|A\|). \tag{36}$$

On the other hand,

$$e^H S + S e^H = 2S + HS + SH + O(\|H\|^2\|S\|). \tag{37}$$

Since $\dot{g}_i = D\exp_{H_i}[\dot{H}_i]$, comparing Eq. (36) with Eq. (37) shows that $\dot{H}_i = 2S_i + O(\|H_i\|^2\|S_i\|)$. □

Proposition A.2 (Linearized graph-diffusion form of SPD Ricci flow)**.**

Let $g_i = \exp(H_i),\ H_i \in \mathrm{Sym}_d,\ \|H_i\| \le \varepsilon$ for all nodes i, with $\varepsilon > 0$ sufficiently small. For each oriented edge $i \to j$, it has been already given that

$$L_{i\to j} = \log(g_i^{-1/2} g_j g_i^{-1/2}), \qquad a_{ij}(g) = w_{ij}\exp(-\beta\|L_{i\to j}\|_F^2),$$

where $\beta \ge 0$ is fixed. The whitened tensor, SPD Ricci response, and SPD Ricci flow are

$$S_i(g) = \sum_{j\sim i} a_{ij}(g) L_{i\to j}, \qquad \mathrm{Ric}_i^{\mathrm{SPD}}(g) = -\frac{1}{2}\left(g_i S_i(g) + S_i(g) g_i\right), \tag{38}$$

$$\partial_t g_i = -2\,\mathrm{Ric}_i^{\mathrm{SPD}}(g) = g_i S_i(g) + S_i(g) g_i. \tag{39}$$

Then the corresponding evolution of the log-metric variables $H_i = \log g_i$ satisfies

$$\partial_t H_i = 2\sum_{j\sim i} w_{ij}(H_j - H_i) + \mathcal{R}_i. \tag{40}$$

The remainder is higher order in the log-metric variation. More precisely, for bounded degree and bounded weights,

$$\|\mathcal{R}_i\| \le C\sum_{j\sim i} w_{ij}\left[\varepsilon^2\|H_j - H_i\| + \beta\|H_j - H_i\|^3\right], \tag{41}$$

where C depends only on the matrix dimension, the chosen norm, and the local bounds used in the expansion.

Proof. By Lemma 3,

$$L_{i\to j} = H_j - H_i + \mathcal{R}_{ij}, \qquad \|\mathcal{R}_{ij}\| \le C\varepsilon^2\|H_j - H_i\|.$$

Also,

$$\|L_{i\to j}\|_F^2 = \|H_j - H_i\|_F^2 + O(\varepsilon^2\|H_j - H_i\|_F^2).$$

For fixed β,

$$\exp(-\beta\|L_{i\to j}\|_F^2) = 1 + O(\beta\|H_j - H_i\|_F^2).$$

Therefore,

$$a_{ij}(g)L_{i\to j} = w_{ij}(H_j - H_i) + O\big(w_{ij}\left[\varepsilon^2\|H_j - H_i\| + \beta\|H_j - H_i\|^3\right]\big).$$

Summing over $j \sim i$ gives the corresponding expansion for S_i. Lemma 4 then converts the metric velocity $\partial_t g_i = g_i S_i + S_i g_i$ into $\partial_t H_i = 2S_i + O(\varepsilon^2\|S_i\|)$, which is absorbed into the stated remainder. □

Proposition A.2 identifies the principal discrete operator generated by the SPD Ricci-type curvature. Although $\mathrm{Ric}_i^{\mathrm{SPD}}$ is defined through the affine-invariant logarithm, state-dependent edge weights, and the tensorial lift $g_i S_i + S_i g_i$, its leading-order action becomes transparent in log-metric coordinates:

$$\partial_t H_i = 2\sum_{j\sim i} w_{ij}(H_j - H_i) + \text{higher-order terms}.$$

Thus, the SPD Ricci flow is a tensor-valued graph diffusion of the local metric field, whose principal discrete operator is the graph Laplacian. This identification is the discrete input for the continuum verification below: standard graph-Laplacian consistency then relates the SPD Ricci flow to the parabolic principal part of smooth Ricci-type metric evolution.

Consistency and verification with the Ricci-type parabolic principal mechanism. The previous proposition identifies the principal discrete operator of the SPD Ricci flow as a graph Laplacian in log-metric coordinates. We now explain why this is the right operator to compare with smooth Ricci-type metric evolution.

We first recall the smooth mechanism that the graph construction is designed to reproduce at the principal level. The Ricci flow $\partial_t g = -2\,\mathrm{Ric}(g)$ is only weakly parabolic in arbitrary coordinates, since the equation is invariant under diffeomorphisms. The Ricci–DeTurck modification fixes this coordinate freedom by adding a Lie-derivative term:

$$\partial_t g = -2\,\mathrm{Ric}(g) + \mathcal{L}_{W(g,\bar{g})} g, \tag{42}$$

where $\bar{g}$ is a fixed background metric and $W^k(g,\bar{g}) = g^{pq}\left(\Gamma^k_{pq}(g) - \Gamma^k_{pq}(\bar{g})\right)$ is the DeTurck vector field. This modification is geometrically equivalent to Ricci flow up to pullback by diffeomorphisms, but it produces a strictly parabolic system. In local coordinates, the equation takes the form

$$\partial_t g_{\alpha\beta} = g^{kl}\partial_k\partial_l g_{\alpha\beta} + Q_{\alpha\beta}(g^{-1}, \partial g, \bar{g}), \tag{43}$$

where Q contains no second derivatives of g. The second-order principal operator $g^{kl}\partial_k\partial_l$ acts component-wise on the metric tensor.

Around a constant background metric $g_0 \in \mathrm{SPD}_m$, affine-invariant congruence whitening sends g_0 to the Euclidean metric:

$$g \longmapsto g_0^{-1/2} g g_0^{-1/2}.$$

Therefore, for the purpose of identifying the linearized principal operator, we may work at the normalized background $g_0 = I$. Writing $H(x) = \log g(x),\ g(x) = \exp(H(x))$, we have

$$g - I = H + O(\|H\|_F^2), \qquad g^{kl} = \delta^{kl} + O(\|H\|_F),$$

where δ^{kl} is the Kronecker delta, i.e. the inverse Euclidean metric. Consequently, the linearized principal part of the Ricci–DeTurck system at the normalized background is the component-wise heat operator

$$\partial_t H = \Delta H, \qquad \Delta H = \sum_{a=1}^{m} \partial_a^2 H.$$

We call this heat-type operator the parabolic principal mechanism of Ricci-type metric evolution. The role of this part is to show that the graph SPD Ricci flow holds this principal mechanism after passing to log-metric coordinates and a consistent graph limit.

Definition A.2 (Graph Laplacian consistency assumption)**.** *Let $\Omega \subset \mathbb{R}^m$ be a coordinate patch and let $x_i \in \Omega$ be graph nodes with mesh scale h. For a smooth matrix-valued field $F : \Omega \to \mathrm{Sym}_q$, define the componentwise Euclidean Laplacian by*

$$\Delta F = \sum_{a=1}^{m} \partial_a^2 F, \qquad (\Delta F)_{\mu\nu} = \sum_{a=1}^{m} \partial_a^2 F_{\mu\nu}.$$

Define the scaled graph Laplacian by

$$(\mathcal{L}_h F)_i = \frac{1}{h^2} \sum_{j\sim i} w_{ij}^{(h)} \left(F(x_j) - F(x_i)\right).$$

We say that the graph weights are Laplacian-consistent on compact subsets of Ω if, for every smooth matrix-valued field $F : \Omega \to \mathrm{Sym}_q$,

$$(\mathcal{L}_h F)_i = c_K \Delta F(x_i) + \mathcal{R}_h[F]_i, \qquad \|\mathcal{R}_h[F]_i\|_F \le C_F h^\alpha, \tag{44}$$

for some constants $c_K > 0$ and $\alpha > 0$ independent of h.

Theorem A.1 (Ricci-type principal-mechanism verification)**.** *Let $\Omega \subset \mathbb{R}^m$ be a coordinate patch. For the direct comparison with smooth metric evolution, take node metrics in SPD_m. After the local background normalization described above, let*

$$g : \Omega \to \mathrm{SPD}_m$$

be a smooth metric field close to the normalized background I, and write the same field in logarithmic SPD coordinates as

$$H(x) = \log g(x) \in \mathrm{Sym}_m, \qquad g(x) = \exp(H(x)).$$

At the graph nodes, set $g_i = g(x_i),\ H_i = H(x_i) = \log g_i$.

Assume the nodes $x_i \in \Omega$ form a family of Laplacian-consistent graphs in the sense of Definition A.2. Define the scaled SPD Ricci flow by

$$\partial_t g_i = g_i S_i^{(h)} + S_i^{(h)} g_i, \qquad S_i^{(h)} = \frac{1}{h^2} \sum_{j \sim i} w_{ij}^{(h)} e^{-\beta \|L_{i\to j}\|_F^2} L_{i\to j}, \tag{45}$$

where $\beta \geq 0$ is fixed. Denote the resulting vector field under log-metric variables by

$$\dot{H}_i = \mathcal{F}_h(H)_i.$$

For any smooth perturbation $K : \Omega \to \mathrm{Sym}_m$, the Fréchet derivative of this log-coordinate vector field at the Euclidean metric $H \equiv 0$ is defined as

$$D\mathcal{F}_h(0)[K]_i := \left.\frac{d}{ds}\right|_{s=0} \mathcal{F}_h(sK)_i = 2(\mathcal{L}_h K)_i. \tag{46}$$

Consequently, on compact subsets where the graph consistency estimate holds,

$$D\mathcal{F}_h(0)[K]_i = 2c_K \Delta K(x_i) + O(h^\alpha), \tag{47}$$

with the remainder measured in Frobenius norm. Equivalently, for sufficiently small smooth log-metric fields H, the log-coordinate SPD Ricci vector field can be decomposed as

$$\mathcal{F}_h(H)_i = 2(\mathcal{L}_h H)_i + \mathcal{N}_h(H)_i, \qquad \mathcal{N}_h(0) = 0, \qquad D\mathcal{N}_h(0) = 0. \tag{48}$$

Here $\mathcal{N}_h$ collects the nonlinear contributions produced by noncommutativity of SPD matrices and by the state-dependent weights $e^{-\beta\|L_{i\to j}\|_F^2}$. The identities $\mathcal{N}_h(0) = 0$ and $D\mathcal{N}_h(0) = 0$ give that nonlinear effects vanish at the constant Euclidean metric and don't change the first-order principal operator.

Therefore, after the constant time rescaling determined by $2c_K$, the linearized continuum limit of the graph SPD Ricci flow is

$$\partial_t H = \Delta H, \tag{49}$$

which matches the Ricci-type parabolic principal mechanism described in Eq. (43)*.*

Proof. We compute the linearization at the constant Euclidean metric. Let $H = sK$, where s is a scalar parameter. Then $g_i(s) = \exp(sK(x_i))$. By Lemma 3,

$$L_{i\to j} = s\,(K(x_j) - K(x_i)) + O(s^2).$$

The state-dependent weight satisfies $e^{-\beta\|L_{i\to j}\|_F^2} = 1 + O(s^2)$, so it has no first-order contribution. Hence

$$S_i^{(h)} = s\frac{1}{h^2} \sum_{j\sim i} w_{ij}^{(h)}\,(K(x_j) - K(x_i)) + O(s^2).$$

By Lemma 4, the SPD Ricci flow $\partial_t g_i = g_i S_i^{(h)} + S_i^{(h)} g_i$ becomes, in log-metric coordinates,

$$\dot{H}_i = 2S_i^{(h)} + O(s^2).$$

Differentiating at $s = 0$ gives

$$D\mathcal{F}_h(0)[K]_i = 2\frac{1}{h^2} \sum_{j\sim i} w_{ij}^{(h)}\,(K(x_j) - K(x_i)) = 2(\mathcal{L}_h K)_i,$$

which proves Eq. (46).

Applying Definition A.2 componentwise to K gives

$$2(\mathcal{L}_h K)_i = 2c_K \Delta K(x_i) + O(h^\alpha),$$

which proves Eq. (47). The decomposition

$$\mathcal{F}_h(H)_i = 2(\mathcal{L}_h H)_i + \mathcal{N}_h(H)_i$$

is obtained by defining $\mathcal{N}_h$ as the difference between the full log-coordinate vector field and its linear graph-Laplacian part. The expansion above shows that the full vector field and $2\mathcal{L}_h$ agree to first order at $H = 0$, hence $\mathcal{N}_h(0) = 0$ and $D\mathcal{N}_h(0) = 0$.

It remains to connect this graph limit with the smooth Ricci-type mechanism described before the theorem. In harmonic coordinates, or under the Ricci–DeTurck modification, Ricci flow has principal second-order operator

$$g^{kl}\partial_k\partial_l g_{\alpha\beta}.$$

At the Euclidean background this reduces to the componentwise Laplacian $\Delta g_{\alpha\beta}$. Since $g = \exp(H)$ and $g - I = H + O(\|H\|_F^2)$ near the background, the same first-order principal operator acts on the log-metric variable H. The graph calculation above shows that the linearization of the scaled SPD Ricci flow converges to $2c_K\Delta H$. After rescaling time by the constant factor $2c_K$, this is precisely the heat-type principal mechanism of the Ricci–DeTurck system. □

Remark (Scope of Theorem A.1)**.** *The theorem is a mechanism-consistency result. It shows that the graph SPD Ricci flow has the same linearized parabolic principal operator as Ricci-type metric evolution after passing to log-metric coordinates and a consistent local graph limit. The result is matched to the implemented SPD Ricci Flow block, whose propagation mechanism is finite-time tensor-valued metric diffusion of the node-wise SPD field.*

Energy dissipation and equilibrium. We now return to the discrete graph and show that the SPD Ricci flow is intrinsically dissipative. This result is not merely a formal analogy: it gives the mechanism used later to interpret metric smoothing, curvature-residual decay, and perturbation relaxation in the experiments.

Theorem A.2 (Energy dissipation under SPD Ricci flow)**.** *Let $\mathcal{E}_\psi$ be defined as in Eq.* (17)*, and let $S_i(g)$ be the corresponding whitened tensor from Eq.* (19)*. Along the SPD Ricci flow*

$$\partial_t g_i = g_i S_i + S_i g_i,$$

the energy is nonincreasing:

$$\frac{d}{dt}\mathcal{E}_\psi(g(t)) \le 0. \tag{50}$$

More precisely, if

$$g_i = U_i \operatorname{diag}(\lambda_{i1}, \dots, \lambda_{id}) U_i^\top, \qquad \widetilde{S}_i = U_i^\top S_i U_i,$$

then

$$-\frac{d}{dt}\mathcal{E}_\psi(g(t)) = \sum_i \sum_{a,b=1}^{d} \frac{\lambda_{ia} + \lambda_{ib}}{\sqrt{\lambda_{ia}\lambda_{ib}}} \left(\widetilde{S}_{i,ab}\right)^2. \tag{51}$$

In particular, equality holds if and only if $S_i(g) = 0$ for every node i, equivalently $\mathrm{Ric}_i^{\mathrm{SPD}}(g) = 0$ for every node i.

Proof. By Corollary A.1,

$$\nabla_{g_i}^{\mathrm{AIRM}}\mathcal{E}_\psi(g) = -g_i^{1/2} S_i g_i^{1/2}.$$

The derivative of the energy along a curve $g(t)$ is

$$\frac{d}{dt}\mathcal{E}_\psi(g(t)) = \sum_i \left\langle \nabla_{g_i}^{\mathrm{AIRM}}\mathcal{E}_\psi, \dot{g}_i \right\rangle_{g_i}.$$

Using the affine-invariant inner product and the SPD Ricci flow velocity gives

$$\frac{d}{dt}\mathcal{E}_\psi(g(t)) = -\sum_i \operatorname{tr}\left[g_i^{-1} \left(g_i^{1/2} S_i g_i^{1/2} \right) g_i^{-1} \left(g_i S_i + S_i g_i \right) \right]. \tag{52}$$

Fix a node i and diagonalize $g_i = U_i \Lambda_i U_i^\top$. By invariance of the trace, Eq. (52) becomes

$$-\operatorname{tr}\left[\Lambda_i^{-1/2} \widetilde{S}_i \Lambda_i^{-1/2} \left(\Lambda_i \widetilde{S}_i + \widetilde{S}_i \Lambda_i \right) \right].$$

The (a,b) entry of $\Lambda_i \widetilde{S}_i + \widetilde{S}_i \Lambda_i$ is $(\lambda_{ia} + \lambda_{ib})\widetilde{S}_{i,ab}$, and the (a,b) entry of $\Lambda_i^{-1/2} \widetilde{S}_i \Lambda_i^{-1/2}$ is $\widetilde{S}_{i,ab}/\sqrt{\lambda_{ia}\lambda_{ib}}$. Summing the trace gives Eq. (51). Since every coefficient

$$\frac{\lambda_{ia} + \lambda_{ib}}{\sqrt{\lambda_{ia}\lambda_{ib}}}$$

is strictly positive, the dissipation vanishes exactly when $\widetilde{S}_i = 0$ for all i, hence $S_i = 0$ for all i. Lemma 1 then gives the equivalent condition $\mathrm{Ric}_i^{\mathrm{SPD}} = 0$ for all i. □

The equality condition $S_i(g) = 0$ corresponds to the local equilibrium performance in experimental realization.

Corollary A.2 (Curvature residual controls the dissipation rate)**.** *Assume the node metrics are uniformly spectrally bounded,*

$$mI \preceq g_i \preceq MI \qquad \textit{for all } i, \tag{53}$$

with $0 < m \le M < \infty$. Let $D(g) = -\frac{d}{dt}\mathcal{E}_\psi(g(t))$ be the dissipation rate along the SPD Ricci flow. Then

$$\frac{2}{M^2} \sum_i \| \mathrm{Ric}_i^{\mathrm{SPD}}(g)\|_F^2 \le D(g) \le \frac{2}{m^2} \sum_i \| \mathrm{Ric}_i^{\mathrm{SPD}}(g)\|_F^2. \tag{54}$$

Proof. In the eigenbasis of g_i, $\mathrm{Ric}_i^{\mathrm{SPD}} = -\frac{1}{2}(g_i S_i + S_i g_i)$ has entries

$$\widetilde{R}_{i,ab}^{\mathrm{BW}} = -\frac{1}{2}(\lambda_{ia} + \lambda_{ib})\widetilde{S}_{i,ab},$$

where

$$\widetilde{S}_i = U_i^\top S_i U_i, \qquad \widetilde{\mathrm{Ric}}_i^{\mathrm{SPD}} = U_i^\top \mathrm{Ric}_i^{\mathrm{SPD}} U_i.$$

Therefore

$$\| \mathrm{Ric}_i^{\mathrm{SPD}} \|_F^2 = \frac{1}{4} \sum_{a,b} (\lambda_{ia} + \lambda_{ib})^2 (\widetilde{S}_{i,ab})^2.$$

Comparing this expression with the dissipation identity in Eq. (51), the ratio of the dissipation coefficient to the squared-residual coefficient is

$$\frac{(\lambda_{ia} + \lambda_{ib})/\sqrt{\lambda_{ia}\lambda_{ib}}}{\frac{1}{4}(\lambda_{ia} + \lambda_{ib})^2} = \frac{4}{(\lambda_{ia} + \lambda_{ib})\sqrt{\lambda_{ia}\lambda_{ib}}}.$$

Under Eq. (53), the denominator lies between $2m^2$ and $2M^2$, which gives Eq. (54). □

Corollary A.2 should be interpreted carefully. It does not say that $\| \mathrm{Ric}_i^{\mathrm{SPD}} \|_F$ or the largest eigenvalue of $\mathrm{Ric}_i^{\mathrm{SPD}}$ must decrease monotonically at every node. It says that, under spectral bounds, the aggregate SPD Ricci residual is quantitatively comparable to the instantaneous energy dissipation rate.

For the ordinary Dirichlet energy $\psi(r) = r/2$, the equilibrium condition $S_i(g) = 0$ has a familiar geometric interpretation.

Corollary A.3 (Harmonic equilibrium and local SPD barycenters)**.** *Let $\psi(r) = r/2$. Then $S_i(g) = 0$ is equivalent to the first-order condition*

$$\sum_{j \sim i} w_{ij} \operatorname{Log}_{g_i}(g_j) = 0. \tag{55}$$

Equivalently, g_i is the affine-invariant Fréchet barycenter of its neighbor with weights w_{ij}:

$$g_i = \arg\min_{h \in \mathrm{SPD}_d} \frac{1}{2} \sum_{j \sim i} w_{ij} d_{\mathrm{SPD}}^2(h, g_j). \tag{56}$$

Proof. For $\psi(r) = r/2$, Eq. (19) becomes $S_i(g) = \sum_{j\sim i} w_{ij} L_{i\to j}$. Multiplying by $g_i^{1/2}$ on both sides gives

$$g_i^{1/2} S_i(g) g_i^{1/2} = \sum_{j\sim i} w_{ij} \operatorname{Log}_{g_i}(g_j).$$

The right-hand side is the negative gradient of the local Fréchet functional in Eq. (56). Since SPD_d with the affine-invariant metric is nonpositively curved, the squared-distance objective is geodesically convex and the first-order condition characterizes the unique weighted barycenter whenever the total weight is positive. □

On a connected graph without boundary constraints or external forcing, any constant field $g_i \equiv g_0$ ($g_0 \in SPD_d$ is constant matrix) has zero energy and hence is a global minimizer of $\mathcal{E}_\psi$ for every increasing ψ with $\psi(0) = 0$. In learning applications, however, the network does not run the free flow to convergence. The learned encoder, decoder, source perturbations, task loss, finite depth, and nonlinear weights all act as forcing or stopping mechanisms. The relevant interpretation is therefore finite-time relaxation toward local equilibrium, not inevitable global collapse to a constant metric.

Relation to Ollivier–Ricci flow. The proposed SPD Ricci response and Ollivier–Ricci curvature address different geometric objects. Standard graph Ollivier–Ricci curvature is a scalar edge quantity determined by the graph metric, edge weights, and node measures. The SPD Ricci response is a matrix-valued node quantity determined by the variation of the SPD metric field over the graph. The distinction between Ollivier–Ricci flow and the proposed SPD Ricci flow is dynamical rather than merely representational.

Ollivier–Ricci curvature is a transport curvature of a metric-measure graph: it measures Wasserstein contraction between probability measures supported around neighboring nodes. Accordingly, Ollivier-type graph Ricci flows naturally close on scalar edge lengths, graph distances, or transport weights. This mechanism is well suited to changing graph-level transport efficiency, separating bottlenecks, and revealing community structure. It is geometrically meaningful, but it is not the same mechanism as Hamilton's Ricci flow, whose evolving variable is the metric tensor itself.

By contrast, SPD Ricci flow closes directly on the node-wise SPD metric field. The local state g_i defines the metric, neighboring metrics determine the affine-invariant tensor $S_i(g)$, and the BW response

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -\frac{1}{2}(g_i S_i(g) + S_i(g) g_i)$$

drives the metric evolution

$$\partial_t g_i = -2\,\mathrm{Ric}_i^{\mathrm{SPD}}(g).$$

Thus the update is not a scalar transport reweighting of the graph, but a tensor-valued evolution of the metric field carried by the nodes. Theorem A.1 verifies this mechanism at the principal level: in log-metric coordinates, the SPD Ricci flow has graph Laplacian principal operator and, under graph-Laplacian consistency, converges to the heat-type parabolic principal part of Ricci–DeTurck metric evolution. In this sense, SPD Ricci flow inherits the metric-diffusion mechanism through which Ricci flow regularizes metric irregularities and induces curvature smoothing, whereas Ollivier-type graph Ricci flow primarily acts through scalar transport geometry.

A.2 Anisotropic Network Model and Normal Modes

The anisotropic network model (ANM) describes small collective motions of a protein around a reference structure using an elastic network (Eyal et al., 2006; Atilgan et al., 2001; Doruker et al., 2000; Xia et al., 2015; Xia, 2018). Each node typically represents a residue through its C_α atom, and nearby nodes are connected by springs. These springs penalize changes in inter-node distances, providing a coarse-grained description of the structure's resistance to deformation.

Let $r_i \in \mathbb{R}^3$ denote the reference position of node i and $x_i = r_i + \Delta x_i$ its displaced position. The contact set is $E = \{(i,j) : i < j,\ \|r_i - r_j\|_2 \le r_c\}$, where r_c is a distance cutoff. With a uniform spring constant $\gamma > 0$, the elastic energy is

$$V(x) = \frac{\gamma}{2} \sum_{(i,j)\in E} \left(\|x_i - x_j\|_2 - \|r_i - r_j\|_2\right)^2.$$

For small displacements, its quadratic approximation is

$$\Delta E(\Delta x) = \frac{1}{2}\Delta x^\top H \Delta x, \qquad H = \nabla_x^2 V(x)\big|_{x=r} \in \mathbb{R}^{3N\times 3N},$$

where $\Delta x = [\Delta x_1^\top, \ldots, \Delta x_N^\top]^\top$. Thus, H describes how displacements at different nodes and along different spatial directions are coupled by the spring network.

Writing $e_{ij} = (r_i - r_j)/\|r_i - r_j\|_2$, the 3×3 blocks of H are

$$H_{ij} = \begin{cases} -\gamma e_{ij} e_{ij}^\top, & \text{if nodes } i \text{ and } j \text{ are connected,} \\ 0, & \text{otherwise,} \end{cases} \qquad i \neq j, \qquad H_{ii} = -\sum_{j\neq i} H_{ij}.$$

The directional dependence of these blocks makes the model anisotropic: the energy cost depends on how neighboring nodes move relative to their connecting axis.

The symmetric positive-semidefinite Hessian admits an orthonormal eigenbasis satisfying

$$Hq^{(k)} = \lambda_k q^{(k)}, \qquad \|q^{(k)}\|_2 = 1.$$

Each eigenvector specifies a normal-mode displacement pattern. Partitioning it as

$$q^{(k)} = \left[(u_1^{(k)})^\top, \ldots, (u_N^{(k)})^\top\right]^\top, \qquad u_i^{(k)} \in \mathbb{R}^3,$$

the vector $u_i^{(k)}$ gives the direction and relative magnitude of node i's displacement within that collective pattern. A general small displacement can be expressed as a superposition $\Delta x = \sum_k a_k q^{(k)}$, with harmonic energy

$$\Delta E = \frac{1}{2}\sum_k \lambda_k a_k^2.$$

In particular, motion along a single mode has energy $\Delta E_k(a) = \frac{1}{2}a^2\lambda_k$. The eigenvalue therefore measures mode stiffness: at the same amplitude, larger λ_k implies a greater energy cost. Softer modes are more readily excited by thermal fluctuations; for a positive-stiffness mode at classical thermal equilibrium, $\mathbb{E}[a_k^2] = k_B T/\lambda_k$.

Our illustration compares unit-normalized mode shapes rather than thermally rescaled displacements. Consequently,

$$\sum_{i=1}^{N} \|u_i^{(k)}\|_2^2 = 1,$$

so all modes have the same total squared displacement, while its distribution across nodes and directions varies. These node-wise vectors are then used to construct the SPD metric fields analyzed in the main text.

For an infinitesimally rigid, non-collinear three-dimensional network, the six zero modes correspond to rigid-body translations and rotations. Insufficiently constrained networks may have additional zero modes. Numerical eigenvectors within the zero eigenspace may mix translations and rotations.

For this illustration, we use a contact cutoff of $d = 1.54$ Å and a uniform spring constant $\gamma = 1$. We construct the node metrics as $g_i = u_i u_i^\top + 10^{-3} I_3$ and compute SPD Ricci-type curvature using unit base edge weights $w_{ij} = 1$ for all contact edges and $\beta = 0.05$.

A.3 Details of the Spherical Metric Evolution

Graph and initialization. We construct an icosphere with subdivision level three, containing 642 vertices, 1920 undirected edges, and 1280 triangular faces. The graph remains fixed throughout the experiment. Each node is initialized with a conformal SPD metric $g_i^{(0)} = \exp(2u_i) I_2$. The scalar field u_i is formed from four localized spherical Gaussian dents and centered to have zero area-weighted mean. Regions with $u_i < 0$ therefore have a smaller metric scale than the identity reference.

Table 3: Metric evolution diagnostics at the displayed checkpoints.

Step k	Time t_k	Energy $\mathcal{E}^{(k)}$	Mean $\|R^{\mathrm{SPD}}\|_F$
0	0	9.88×10^{-3}	8.93×10^{-2}
200	2	7.13×10^{-4}	2.07×10^{-2}
1000	10	1.24×10^{-4}	6.30×10^{-3}
5000	50	5.18×10^{-7}	3.96×10^{-4}

Curvature computation and numerical update. At each iteration, we apply Eq. (14) with $w_i j = 1$, $\beta = 0.5$ and $\eta = 0.01$ for 5000 steps. The projection replaces each eigenvalue μ of the symmetric Euler proposal with $\max(\mu, 10^{-9})$. We record checkpoints at $k = 0, 200, 1000, 5000$, corresponding to rescaled times $t_k = 0, 2, 10, 50$, with rescaled time $t_k = k\eta$..

Surface reconstruction. For visualization, each checkpoint metric field is converted into target lengths for the mesh edges. We then optimize three-dimensional vertex positions to fit these lengths. The reconstruction objective combines an edge-length fitting term with weak radial and tangential regularization and a face-orientation penalty. The radial prior selects an inward-dent visualization branch among potentially nonunique embeddings. The reconstructed coordinates serve only as a visualization of the metric field and are not used to compute subsequent curvature responses. Consequently, the displayed surface changes should be interpreted as fitted realizations of the evolving metric data.

Quantitative diagnostics. Let $\vec{E}$ denote the directed edge set. For each directed edge $i \to j$, let

$$L_{i\to j}^{(k)} = \log\Big((g_i^{(k)})^{-1/2} g_j^{(k)} (g_i^{(k)})^{-1/2}\Big).$$

The displayed energy is the robust edge energy used by the update,

$$\mathcal{E}^{(k)} = \frac{1}{|\vec{E}|} \sum_{(i,j)\in\vec{E}} \frac{1 - \exp\Big(-\beta \|L_{i\to j}^{(k)}\|_F^2\Big)}{2\beta}. \tag{57}$$

We also report the mean Frobenius magnitude of the SPD response. Both of them decrease across these checkpoints, indicating reduced neighborhood variation in this experiment. The reconstruction error remains below 0.4% at all reported checkpoints, supporting the edge-length fidelity of the displayed surfaces. These observations describe the illustrated trajectory rather than establish a general convergence guarantee.

A.4 Details of the SPD Ricci Flow Node Classification Experiments

This appendix provides implementation and evaluation details for the SPD Ricci Flow node classification experiments reported in Table 1.

Baselines, Datasets and Evaluation Protocol. We compare SPD Ricci Flow with GCN (Kipf & Welling, 2017), GAT (Veličković et al., 2018), DirGNN (Rossi et al., 2024), FSGNN (Maurya et al., 2022), GRAND (Chamberlain et al., 2021), GRAND++ (Thorpe et al., 2022), ACMP (Wang et al., 2023), HiD-Net (Li et al., 2024), GNRF (Chen et al., 2025), and a matched SPDGNN baseline (Zhao et al., 2023). The matched baseline uses the same architecture and training setup, with only the propagation operator replaced, as detailed below.

We evaluate on five node classification benchmarks: Wisconsin (Pei et al., 2020), Questions (Platonov et al., 2023), Cora-Full, DBLP, and Cora-ML (Bojchevski & Günnemann, 2018). Wisconsin Cora-Full, DBLP, and Cora-ML are evaluated using classification accuracy. Questions is a binary classification dataset and is evaluated using ROC-AUC.

For all datasets, for each random seed $s \in \{0, \ldots, 9\}$, nodes are randomly partitioned into approximately 60% training, 20% validation, and 20% test nodes following GNRF (Chen et al., 2025). The reported number is the mean test score over the ten seeds, and the subscript is the standard deviation over the same ten seeds.

Before training, each graph is converted to an undirected graph and self-loops are removed. Node features are row-normalized. Let $X \in \mathbb{R}^{n\times f}$ denote the resulting node feature matrix and let E denote the directed edge list obtained from the undirected graph representation.

SPD Node Representation. Each node feature vector is first mapped to a symmetric positive definite matrix. Given a feature vector $x_i \in \mathbb{R}^f$, the FeatureToSPD mapper used in our experiments applies a learned affine map $a_i = Wx_i + b \in \mathbb{R}^{d^2}$. We reshape this vector into a matrix $A_i = \mathrm{unvec}_{d\times d}(a_i) \in \mathbb{R}^{d\times d}$, symmetrize it, $S_i = \frac{1}{2}(A_i + A_i^\top)$, and map it to the SPD manifold by the exponential map at the identity: $G_i^{(0)} = \mathrm{Exp}_I(S_i) = \exp(S_i)$. Thus each node is represented by an SPD matrix $G_i \in \mathrm{SPD}_d$, where d is a dataset-specific SPD dimension.

Classifier Head. After the final SPD Ricci layer, each node matrix is mapped back to the tangent space at the identity:

$$Z_i = \log(G_i).$$

The upper-triangular entries of Z_i are vectorized and passed to a linear classifier. Training uses cross-entropy loss on the training nodes. For Questions, the training loss is still cross-entropy, while evaluation reports ROC-AUC.

A.5 Molecular Property Prediction

Experimental Setup. We further evaluate SPD Ricci Flow on molecular property prediction, where the task is to predict graph-level molecular properties from molecular graphs. We report ROC-AUC scores on seven benchmark datasets: BACE, BBBP, ClinTox, SIDER, Tox21, HIV, and MUV. These datasets cover both single-task and multi-task binary classification settings, with substantial variation in molecular size, label sparsity, and task difficulty.

Each molecule is represented as a graph derived from its SMILES string. Atoms are treated as nodes and chemical bonds are treated as edges. We use RDKit-based molecular preprocessing to construct molecular graphs, generate three-dimensional conformations, and extract geometry-aware information. Atom features are initialized using CGCNN-style atomic descriptors, while bond features encode bond type, bond direction, ring membership, and distance-related quantities. In addition to the ordinary molecular graph connectivity, the preprocessing stores geometric information from the 3D conformer, allowing the model to couple Euclidean molecular features with SPD-valued geometric states.

The datasets are split into training, validation, and test sets using an 80/10/10 scaffold split. Scaffold splitting is used instead of random splitting because it provides a more challenging and chemically meaningful evaluation: molecules in the test set are structurally less similar to those in the training set. For each dataset, we run the model under 5 random seeds and report the mean and standard deviation of ROC-AUC over independent runs. All results are reported in percentage points.

Model Architecture and Training Details. Our molecular model is implemented as a minimal modification of the SPDSheaf architecture of Peng et al. (2026). We keep the same two-stream design: a Euclidean feature stream based on graph message passing and an SPD coordinate stream initialized from the RDKit conformer geometry. We also keep the feature–geometry interaction module, graph-level pooling, and prediction head from the SPDSheaf molecular backbone.

The main difference is the propagation rule used in the SPD coordinate stream. At each layer, after the learned SPD isometry, we replace the original SPDSheaf propagation with an SPD Ricci flow update (Eq. (14)). The coordinate SPD state is then updated by

$$X_i \leftarrow X_i - \eta\,\mathrm{Ric}_i^{\mathrm{SPD}}(X) = X_i + \frac{\eta}{2}\big(X_iS_i(X) + S_i(X)X_i\big).$$

followed by the same SPD stabilization and nonlinearity used in the backbone. Thus, the comparison with SPDSheaf isolates the effect of replacing the SPDSheaf propagation mechanism with a local Ricci-type SPD diffusion rule, while keeping the surrounding molecular architecture unchanged.

The models are trained with binary cross-entropy loss. For multi-task datasets such as Tox21, SIDER, ClinTox, and MUV, the model outputs one sigmoid score per task and the loss is accumulated across tasks. ROC-AUC is computed from the predicted probabilities and ground-truth labels.

Table 4: Molecular property prediction results on MoleculeNet benchmarks. We report ROC-AUC (%); subscripts denote standard deviations. **Bold** and underline indicate the highest and second-highest mean scores, respectively, with ties receiving the same marking.

Model	BACE	BBBP	ClinTox	SIDER	Tox21	HIV	MUV
D-MPNN	$80.9_{0.6}$	$71.0_{0.3}$	$90.6_{0.7}$	$57.0_{0.7}$	$75.9_{0.7}$	$77.1_{0.5}$	$78.6_{1.4}$
AttentiveFP	$78.4_{2.2}$	$66.3_{1.8}$	$84.7_{0.3}$	$60.6_{3.2}$	$78.1_{0.5}$	$75.7_{1.4}$	$78.6_{1.5}$
N-GramRF	$77.9_{1.5}$	$69.7_{0.6}$	$77.5_{4.0}$	$66.8_{0.7}$	$74.3_{0.9}$	$77.2_{0.4}$	$76.9_{0.2}$
N-GramXGB	$79.1_{1.3}$	$69.1_{0.8}$	$87.5_{2.7}$	$65.5_{0.7}$	$75.8_{0.9}$	$78.7_{0.4}$	$74.8_{0.2}$
PretrainGNN	$84.5_{0.7}$	$68.7_{1.3}$	$72.6_{1.5}$	$62.7_{0.8}$	$78.1_{0.6}$	$79.9_{0.7}$	$81.3_{2.1}$
$\text{GROVE}_{\text{base}}$	$82.1_{0.7}$	$70.0_{0.1}$	$81.2_{3.0}$	$64.8_{0.6}$	$74.3_{0.1}$	$62.5_{0.9}$	$67.3_{1.8}$
$\text{GROVE}_{\text{large}}$	$81.0_{1.4}$	$69.5_{0.1}$	$76.2_{3.7}$	$65.4_{0.1}$	$73.5_{0.1}$	$68.2_{1.1}$	$67.3_{1.8}$
GraphMVP	$81.2_{0.9}$	$72.4_{1.6}$	$79.1_{2.8}$	$63.9_{1.2}$	$75.9_{0.5}$	$77.0_{1.2}$	$77.7_{0.6}$
MolCLR	$82.4_{0.9}$	$72.2_{2.1}$	$91.2_{3.5}$	$58.9_{1.4}$	$75.0_{0.2}$	$78.1_{0.5}$	$79.6_{1.9}$
GEM	$85.6_{1.1}$	$72.4_{0.4}$	$90.1_{1.3}$	$67.2_{0.4}$	$78.1_{0.1}$	$80.6_{0.9}$	$81.7_{0.5}$
Mol-GDL	$86.3_{1.9}$	$72.8_{1.9}$	$96.6_{0.2}$	$83.1_{0.2}$	$79.4_{0.5}$	$80.8_{0.7}$	$67.5_{1.4}$
Uni-Mol	$85.7_{0.2}$	$72.9_{0.6}$	$91.9_{1.8}$	$65.9_{1.3}$	$79.6_{0.5}$	$80.8_{0.7}$	$82.1_{1.3}$
SMPT	$87.3_{1.5}$	$73.4_{0.3}$	$92.7_{0.2}$	$67.6_{5.0}$	$\underline{79.7}_{0.1}$	$\mathbf{81.2}_{0.1}$	$\underline{82.2}_{0.8}$
SchNet	$73.7_{0.9}$	$70.8_{1.4}$	$97.8_{0.3}$	$83.1_{0.0}$	$72.4_{0.2}$	$68.8_{0.1}$	$68.2_{0.6}$
EGNN	$79.4_{2.3}$	$72.1_{2.0}$	$98.6_{0.5}$	$83.2_{0.1}$	$74.5_{0.3}$	$70.4_{0.2}$	$68.6_{0.5}$
SPDGNN	$87.5_{0.9}$	$71.9_{2.8}$	$98.7_{0.4}$	$83.9_{0.3}$	$79.3_{0.8}$	$79.6_{1.2}$	$76.2_{4.6}$
SPDSheaf	$\underline{89.0}_{1.4}$	$\underline{77.4}_{2.7}$	$\underline{99.4}_{0.2}$	$\underline{84.3}_{0.4}$	$\mathbf{80.1}_{0.7}$	$\underline{80.9}_{0.7}$	$\mathbf{82.3}_{1.4}$
SPD Ricci Flow	$\mathbf{89.7}_{1.0}$	$\mathbf{79.8}_{1.0}$	$\mathbf{99.6}_{0.12}$	$\mathbf{85.0}_{0.3}$	$\mathbf{80.1}_{0.2}$	$80.8_{1.2}$	$81.2_{4.3}$

The implementation uses double precision for SPD operations to improve numerical stability during eigenvalue decompositions, logarithmic maps, inverse square roots, and SPD projections. Gradient clipping is applied to avoid unstable updates in the SPD stream. The optimization uses a Muon optimizer for matrix-like parameters and Adam for the remaining parameters. Full results are reported in Table 4.

A.6 Financial Curvature Event Study

This appendix gives the construction details for the financial-news curvature experiment in Section 5.3. The goal of the experiment is not to use news text as an input feature, but to use news dates as external timestamps at which real stocks are likely to undergo state changes. We then ask whether graph curvature and SPD Ricci curvature respond systematically around those timestamps.

Stock universe and event dates. We use the daily stock-price and financial-news panel from (Dong et al., 2024). The price panel covers 3,273 stocks over the 2018–2022 trading period. Each stock-day contains adjusted open, high, low, close, and volume information. Raw news timestamps are conservatively mapped to NYSE response sessions. We retain A/B-quality single-ticker event groups and aggregate multiple retained groups assigned to the same ticker and response session into one ticker-day event for the curvature event study. The graph itself is built only from historical price movements; news is used only to select and align event dates.

For stock i on trading day t, let

$$r_{i,t} = \log P_{i,t}^{\text{close}} - \log P_{i,t-1}^{\text{close}}$$

be the adjusted-close log return. We define the equal-weighted market return over the available stock universe as

$$r_{\text{mkt},t} = |V_t|^{-1} \sum_{j \in V_t} r_{j,t},$$

and use the market-adjusted abnormal return

$$AR_{i,t} = r_{i,t} - r_{\text{mkt},t}.$$

News events are split by the event-day absolute abnormal return. Events in the top decile of $|AR_{i,t_0}|$ are called *strong-response* events, while all remaining events form the non-top-decile comparison group.

Residual-return graph. For every trading day, we construct a stock graph from recent residual return comovement. The market component is removed from individual stock returns using a rolling historical window. Pairwise residual correlations are then estimated with covariance shrinkage, producing a daily correlation $\rho_{ij,t}$ between stocks i and j. The corresponding edge length is

$$\ell_{ij,t} = \sqrt{2(1-\rho_{ij,t})},$$

and the nonnegative similarity weight used by the graph-curvature computation is

$$w_{ij,t} = \max(\rho_{ij,t}, 0)^2.$$

To obtain a connected but sparse daily graph, we first add the minimum spanning tree under the distance $\ell_{ij,t}$ and then add the strongest remaining correlation edges until the graph contains $3|V_t|-6$ edges. This gives the daily graph used throughout the event study.

SPD stock states. Each stock-day is represented by a symmetric positive definite matrix $X_{i,t} \in \mathrm{SPD}_d$ summarizing recent price-volume variation. The full six-dimensional state uses the coordinate vector

$$z_{i,t} = \big(AR_{i,t},\ |AR_{i,t}|,\ r_{i,t}^{\mathrm{overnight}},\ r_{i,t}^{\mathrm{intraday}},\ q_{i,t},\ \Delta v_{i,t}\big),$$

where

$$r_{i,t}^{\mathrm{overnight}} = \log P_{i,t}^{\mathrm{open}} - \log P_{i,t-1}^{\mathrm{close}}, \qquad r_{i,t}^{\mathrm{intraday}} = \log P_{i,t}^{\mathrm{close}} - \log P_{i,t}^{\mathrm{open}},$$

$$q_{i,t} = \log(P_{i,t}^{\mathrm{high}}/P_{i,t}^{\mathrm{low}}), \qquad \Delta v_{i,t} = \log(1+\mathrm{Vol}_{i,t}) - \log(1+\mathrm{Vol}_{i,t-1}).$$

All open, high, and low prices are adjusted using the same adjustment factor as the close price. The raw coordinates are standardized causally using a 252-session rolling median and IQR scale, shifted so that only information before day t is used; standardized values are clipped to $[-10, 10]$. We then accumulate the standardized coordinates into an exponentially weighted second-moment matrix, apply trace-preserving shrinkage toward an isotropic matrix, and enforce a small positive eigenvalue floor. The result is an SPD matrix whose entries encode the local covariance structure of the stock's recent price-volume state.

We then compute the SPD Ricci-type curvature at each node from its SPD metric and those of its neighbors. We summarize each curvature matrix by its Frobenius norm, trace, and minimum and maximum eigenvalues. Before event alignment, each scalar summary is demeaned across the daily graph. For a generic curvature summary $s_{i,t}$, define

$$\widetilde{s}_{i,t} = s_{i,t} - |V_t|^{-1} \sum_{j \in V_t} s_{j,t}.$$

This removes the daily graph-wide mean and expresses each stock's curvature summary relative to that mean.

Mean-placebo event response. For each news event $e = (i, t_0)$, we select $K = 5$ matched placebo dates for the same stock. Placebo dates are stock-days with no assigned news for that stock. Matching uses standardized distances in pre-event curvature, graph, volatility, and global-market state features, with priority given to same-month candidates, then same-year candidates, and finally all same-stock candidates.

Let $\mathcal{P}_e$ denote the placebo set for event e. For horizon h trading sessions, the event change relative to the previous trading day is

$$\Delta_{e,h}^{\mathrm{event}} = \widetilde{s}_{i,t_0+h} - \widetilde{s}_{i,t_0-1}.$$

The corresponding mean-placebo change is

$$\overline{\Delta}_{e,h}^{\mathrm{placebo}} = \frac{1}{|\mathcal{P}_{e,h}|} \sum_{\tau \in \mathcal{P}_{e,h}} \left(\widetilde{s}_{i,\tau+h} - \widetilde{s}_{i,\tau-1}\right),$$

where $\mathcal{P}_{e,h} \subseteq \mathcal{P}_e$ contains placebo dates with the required valid observations at horizon h trading sessions. The event-study outcome is

$$Y_{e,h} = \Delta_{e,h}^{\mathrm{event}} - \overline{\Delta}_{e,h}^{\mathrm{placebo}}.$$

Thus all plotted curves and reported table values are event-minus-average- placebo changes, not raw curvature levels. In Fig 4, the displayed curves average $Y_{e,h}$ over a balanced event set across $h \in [-30, 30]$. Shaded bands are 95% confidence intervals from cluster bootstrap resampling at the stock level.

Table 5: Return-channel robustness of the top-decile Frobenius response. Full uses $(AR, |AR|, r^{\text{overnight}}, r^{\text{intraday}}, q, \Delta v)$; No explicit AR removes $(AR, |AR|)$; AR only keeps $(AR, |AR|)$; Range–volume keeps $(q, \Delta v)$. All rows use the same balanced event set and fixed $K = 5$ placebo matches as Fig. 4. Entries are event-minus-average-placebo means for top-decile news events.

SPD state	$h = 0$	$h = 1$	$h = 3$	$h = 5$	$h = 10$	$h = 20$
Full	30.637	26.591	13.573	3.619	-9.979	-18.605
No explicit AR	17.913	16.072	8.759	2.627	-6.228	-12.113
AR only	13.558	12.745	7.659	2.872	-4.257	-9.150
Range–volume	6.010	5.709	3.169	0.530	-2.426	-4.877

Return-channel robustness. Table 5 tests whether the top-decile Frobenius response is driven only by the explicit abnormal-return coordinates in the SPD state. We keep the event set and the $K = 5$ placebo matches fixed to those used in Fig 4, and recompute the SPD Ricci-type curvature outcome after replacing the full six-coordinate SPD state by reduced coordinate subsets. Thus, the first row reproduces the Fig 4 Frobenius trajectory at the selected horizons, while the remaining rows isolate the effect of changing the SPD state representation.

Removing coordinates attenuates the response magnitude, but does not eliminate the event-time pattern. The state without explicit AR or $|AR|$ still shows a positive event-day response and a similar decay across horizons. Moreover, the range-volume state, which contains neither AR nor $|AR|$, also preserves a positive top-decile response at short horizons. These results support the claim that the curvature response is not simply an algebraic restatement of the abnormal-return coordinate.

A.7 Financial News Spillover Prediction

This appendix provides the construction details for the financial spillover prediction experiment in Section 5.4. The task is designed to test whether an event-conditioned SPD metric field can identify which neighboring stocks become unusually affected after a source-stock news event.

Prediction examples. Each example is a ticker-day news event $e = (s, t_0)$ with source stock s and event date t_0. We use the same 3,273-stock daily price and news panel as in Appendix A.6. The frozen prediction cache keeps events that either have quality-A news or are assigned to a non-other company event category. All graph and stock-state inputs are taken from the trading day immediately preceding the news event, $t_0 - 1$, so the model does not observe future prices or future graph states.

The data are split chronologically. After enforcing a 20-trading-day purge so that neither event labels nor placebo labels cross split boundaries, the retained source-event dates span 2018-01-03–2020-12-02 for training, 2021-01-04–2021-12-02 for validation, and 2022-01-03–2022-12-01 for testing. The training and evaluation loader further removes events without a finite target value or without any valid candidate stock at the primary horizon, leaving 9,249, 1,364, and 1,485 scored source events.

Candidate stocks. For each event, we form a local subgraph from the residual-return graph at $t_0 - 1$. The source stock is excluded from the prediction targets. Candidate stocks are restricted to the one-hop and two-hop shells around the source,

$$\mathcal{C}_e = \{i : \text{dist}_{G_{t_0-1}}(i, s) \in \{1, 2\}\},$$

For the primary horizon $H = 10$, a candidate is also removed if it has its own assigned news during $(t_0, t_0 + H]$. This filter makes the label closer to a spillover response from the source event, rather than a direct response to the candidate's own news.

For the $H = 10$ finance target, the formal cache contains 207,322 training candidate pairs, 33,134 validation pairs, and 29,875 test pairs before the finite-target filtering used by the runner. In the test split, 3,267 candidate pairs are positive under the $q = 0.90$ node-label cutoff, giving a test candidate prevalence of 10.94%. Among the 1,500 cached test events, 891 have at least one positive candidate.

Finance response label. Let $AR_{i,t}$ be the equal-weighted market-adjusted abnormal return of stock i on day t, and let τ_e denote the matched no-news placebo date for the same source event. For a candidate stock i and horizon H, we compare the candidate's future abnormal returns after the real news event with the same candidate's future abnormal returns after the matched placebo date. For $u = 1, \ldots, H$, define

$$D_{e,i}(u) = \sum_{\ell=1}^{u} \left(AR_{i,t_0+\ell} - AR_{i,\tau_e+\ell} \right).$$

The raw finance response score is

$$S_{e,i}^{(H)} = \frac{\max_{u=2,\ldots,H} |D_{e,i}(u)|}{\max(\sigma_{i,t_0}^{(60)}, 10^{-6})},$$

where $\sigma_{i,t_0}^{(60)}$ is the standard deviation of stock i's abnormal return estimated from the previous 60 trading days, using at least 40 valid observations. A candidate-horizon pair is retained only if at least $\max(2, \lceil 0.8H \rceil)$ event-placebo daily return comparisons are finite and $\sigma_{i,t_0}^{(60)}$ is finite and larger than 10^{-6}.

The raw scores are transformed using $\log(1 + S_{e,i}^{(H)})$ and standardized within ticker and horizon using training-set statistics. If a ticker has fewer than 30 valid training observations, or if its ticker-specific scale is too small relative to the global scale, a global horizon-specific standardizer is used instead. The binary label is then

$$y_{e,i}^{(H,q)} = \mathbf{1}\{z_{e,i}^{(H)} > c_{H,q}\},$$

where $c_{H,q}$ is the q-quantile of the standardized training responses among valid candidate pairs. $H = 10$ and $q = 0.90$ are used.

Source-news encoding. The source-news input has a structured component and a title component. The structured news vector contains seven values: log article count, log event count, a quality-A indicator, log quality-A event count, log quality-B event count, and sinusoidal month features. The representative headline is encoded by a 128-dimensional signed hashing vector using up to 48 alphanumeric tokens, followed by ℓ_2 normalization.

Let $g_{i,t_0-1} \in \mathrm{SPD}_6$ be the pre-event SPD stock state. The source encoder maps the news vector, event-type embedding, topic embedding, source log-SPD state, and source scalar graph features to a symmetric tangent matrix T_e. The source perturbation is bounded by applying a hyperbolic tangent and a source scale. The initial post-news metric field is

$$g_s^{(0)} = \exp(\log g_{s,t_0-1} + T_e), \qquad g_i^{(0)} = g_{i,t_0-1} \quad (i \neq s),$$

so the source state remains SPD after injecting the news.

Prediction head For a candidate node i, the readout uses the candidate's propagated log-metric displacement

$$\delta_{e,i} = \log g_i^{(L)} - \log g_{i,t_0-1},$$

the source tangent T_e, the one-hop/two-hop shell embedding, and Frobenius norms of $\delta_{e,i}$ and T_e. These quantities are concatenated and fed to a shared MLP head to produce a probability $\hat{p}_{e,i}$ that candidate i is affected by the source event.

Training objective and evaluation metrics. Models are trained with a node-level weighted binary cross-entropy loss and an event-local pairwise ranking loss. Reported means and standard deviations are computed over seeds 42–46. All metrics are computed over candidate rankings within each source event. Macro-AUPRC averages the average precision over events that contain at least one positive candidate. NDCG@5 is averaged over the same eligible event set. Precision@5 averages the fraction of positives among the top five predicted candidates over all events with candidates; if fewer than five candidates are available, all available candidates are used. Hit@5 is the fraction of eligible events for which at least one positive candidate appears in the top five.

Table 6: Component-wise ablation results on node classification and molecular property prediction. Wisconsin and DBLP report classification accuracy; BACE and SIDER report ROC-AUC. All scores are percentages. Subscripts denote standard deviations over ten seeds for node classification and five seeds for molecular property prediction.

Method / Variant	Wisconsin	DBLP	BACE	SIDER
Ablation study				
Uniform weights	$69.40_{4.10}$	$85.86_{0.44}$	$87.71_{0.47}$	$84.20_{0.34}$
Uniform + $\mathcal{L}_i$	$73.60_{4.27}$	$85.89_{0.44}$	$88.89_{1.10}$	$83.99_{0.39}$
No Ricci Flow layer	$87.40_{2.54}$	$85.88_{0.46}$	$87.96_{1.13}$	$84.03_{0.19}$
SPD Ricci Flow	$\mathbf{89.80}_{3.52}$	$\mathbf{85.90}_{0.39}$	$\mathbf{89.70}_{1.00}$	$\mathbf{85.00}_{0.30}$

Baselines. All baselines use the same event cache, candidate sets, finance-response labels, chronological splits, and validation-selection rule. The comparison includes standard message-passing GNNs, continuous-depth graph models, and SPD-based neural network baselines. Each baseline is adapted to the same source-conditioned candidate-ranking task. For fairness, the source-news encoder, shell information, and candidate-level prediction head are kept aligned with the proposed model whenever the architecture permits; the main difference is the propagation operator used to transmit the source-news signal through the local stock graph.

Ablation across graph domains. To examine the contributions of our propagation components beyond financial spillover prediction, we extend the component-wise ablations to node classification on Wisconsin and DBLP and molecular property prediction on BACE and SIDER. Table 6 compares the full model with three variants: removing SPD-dependent edge weighting (*Uniform weights*), additionally replacing the curvature update with $\mathcal{L}_i$ (*Uniform + $\mathcal{L}_i$*), and removing the Ricci Flow layers (*No Ricci Flow layer*).

A.8 Computational Complexity

Let $n = |V|$ and let $m = |\vec{E}|$ denote the number of directed edges; for an undirected graph represented in both directions, $m = 2|E|$. Each node carries a matrix $g_i \in \mathrm{SPD}_d$. For every directed edge $i \to j$, the SPD Ricci Flow layer computes

$$L_{i\to j} = \log\Big(g_i^{-1/2} g_j g_i^{-1/2}\Big), \qquad a_{ij}(g) = w_{ij} \exp\big(-\beta \|L_{i\to j}\|_F^2\big).$$

The inverse square root and matrix logarithm are evaluated through symmetric eigendecompositions, while the congruence transformation requires dense $d \times d$ matrix multiplications. These operations have $O(d^3)$ time complexity per directed edge. Computing all edge-wise relative logarithms therefore costs $O(md^3)$.

The weighted aggregation

$$S_i(g) = \sum_{j\sim i} a_{ij}(g) L_{i\to j}$$

can be implemented by a scatter operation in $O(md^2)$ time. Constructing

$$\mathrm{Ric}_i^{\mathrm{SPD}}(g) = -\frac{1}{2}\big(g_i S_i(g) + S_i(g) g_i\big)$$

and applying the Euler update and activation require $O(nd^3)$ time. Consequently, the total time complexity of one SPD Ricci Flow layer is

$$O\big(md^3 + nd^3 + md^2\big) = O\big((n+m)d^3\big).$$

For the sparse graphs considered here, this is commonly written as $O(md^3)$.

For comparison, conventional message passing over vectorized $d \times d$ node states requires $O(md^2)$ time for neighborhood aggregation. The additional factor in the proposed layer comes from the spectral matrix operations required by affine-invariant SPD geometry. Nevertheless, since d is small and fixed in our experiments, both the time and memory costs remain linear in the number of nodes and edges.