# Adaptive Riemannian Graph Neural Networks

**Xudong Wang, Chris Ding, Tongxin Li, Jicong Fan***
School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), China
xudongwang@link.cuhk.edu.cn, {chrisding, litongxin, fanjicong}@cuhk.edu.cn

## Introduction & Motivation

Real-world graphs exhibit **geometric heterogeneity**: tree-like hierarchies and dense communities **coexist** within a network.

**Fundamental Challenge**

How to learn adaptive, continuous geometry that captures diverse local structures in graphs?

### Observations on Existing Geometric GNNs:

- Fixed-curvature GNNs embed into single manifolds, which may not fully capture geometric diversity
- Product manifold methods are limited to block-constant geometries from a discrete set of curvatures
- Scalar curvature approaches remain isotropic, unable to capture directional geometric information

### Evidence of Geometric Heterogeneity:

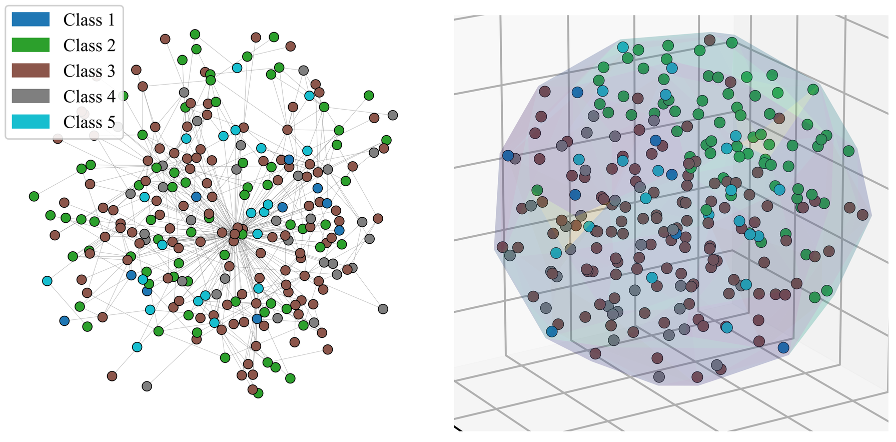



Figure 1: **Wisconsin**: Raw topology (left) vs. 3-D t-SNE with curvature (right). Curvature varies across the Riemannian manifold.

## Core Contributions

**Key Takeaway:** We can move beyond fixed/mixed-curvature paradigms to learn **node-wise metric tensor fields** that precisely characterize local graph geometry.

1. **Novel Framework:** ARGNN learns **continuous, anisotropic Riemannian metric tensor fields** $\{\mathbf{G}_i \in \mathcal{S}^d_{++}\}_{i\in\mathcal{V}}$
2. **Efficient Parameterization:** Diagonal metric tensors with $O(d)$ complexity per node, naturally align with Conformal Transformation
3. **Ricci Flow Regularization:** Stable geometric evolution and smooth metric fields
4. **Theoretical Guarantees:** Convergence, universal approximation. Complexity analysis.
5. **Superior Performance:** On both Node Classification and Link Prediction Tasks across Homophilic and Heterophilic graphs.

## Learning Anisotropic Metric Field

**Key Innovation:** Learn node-wise diagonal metric tensor:

$$\mathbf{G}_i = \mathrm{diag}(\mathbf{g}_i) = \mathrm{diag}(g_{i,1}, g_{i,2}, \ldots, g_{i,d})$$

**Local Metric Estimator:**

$$\mathbf{g}_i = \mathrm{softplus}\left(f_\theta^{(g)}\left([\mathbf{h}_i; \mathbf{a}_i]\right)\right)$$

where $\mathbf{a}_i = \frac{1}{|\mathcal{N}(i)|}\Sigma_{j\in\mathcal{N}(i)}\mathbf{h}_j$ aggregates neighborhood.

**Advantages of Diagonal Parameterization:**

- **Geometric:** Anisotropic conformal transformation with line element $ds^2 = \Sigma_{k=1}^{d} g_{i,k}(dx^k)^2$
- **Interpretable:** $g_{i,k}$ = geometric importance of $k$-th feature
- **Efficient:** $O(d)$ vs. $O(d^2)$ for full tensors

## Loss Function & Optimization

**Ricci Regularization** (promotes uniform curvature):

$$\mathcal{L}_{\mathrm{Ricci}} = \sum_{i\in\mathcal{V}}\sum_{k=1}^{d}\left(\mathrm{Ric}_{kk}^{(i)}\right)^2, \quad \mathrm{Ric}_{kk}^{(i)} = \frac{1}{2|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)}(g_{i,k} - g_{j,k})$$

**Smoothness Regularization:** $\mathcal{L}_{\mathrm{smooth}} = \Sigma_{(i,j)\in\mathcal{E}}\|\mathbf{g}_i - \mathbf{g}_j\|_2^2$
**Total Loss:** $\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{task}} + \alpha\mathcal{L}_{\mathrm{Ricci}} + \beta\mathcal{L}_{\mathrm{smooth}}$
**Theory-guided Hyperparameters Choice (Theorem 1):**

$$\alpha = \frac{c_1}{L}\min\left(1, \frac{d}{|\mathcal{E}|}\right), \quad \beta = \frac{c_2\sqrt{d}}{|\mathcal{V}|}$$

where $c_1 = (1-\mathcal{H}) + 0.1$, $c_2 = 0.1(1+\mathcal{H})$, and $\mathcal{H}$ is homophily ratio.

## The Proposed ARGNN Framework

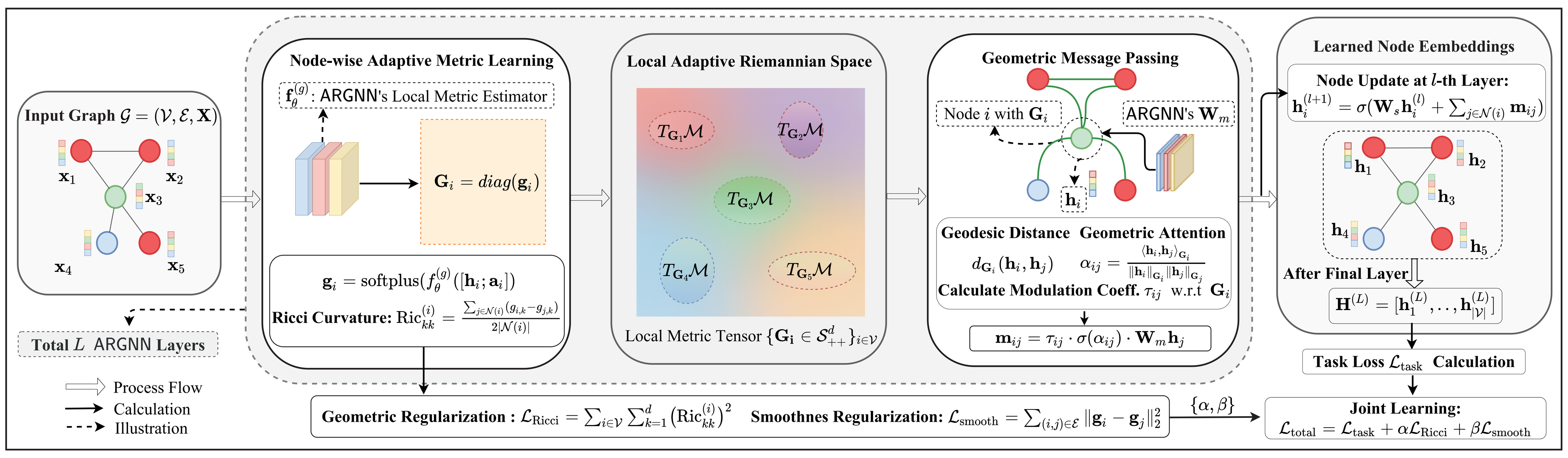



Figure 2: ARGNN jointly learns continuous, anisotropic metric tensor fields $\{\mathbf{G}_i \in \mathcal{S}^d_{++}\}_{i\in\mathcal{V}}$ and node embeddings $\mathbf{H} = \{\mathbf{h}_i\}_{i\in\mathcal{V}}$. The learned $\mathbf{G}_i$ captures nodes' geometric information beyond curvature.

## Theoretical Analysis

**Theorem 1 (Convergence & Stability):**
Under mild regularity conditions, learned metrics $\{\mathbf{G}_i\}$ converge to a stationary point with error bound: $\mathbb{E}[\|\nabla\mathcal{L}_{\mathrm{total}}^{(t)}\|^2] = O\left(\frac{1}{\sqrt{t}}e^{-\mu_{\mathrm{eff}}t/L}\right)$ where $\mu_{\mathrm{eff}}$ is the effective curvature of the loss landscape. Optimal regularization hyperparameters satisfy:

$$\alpha^\star = \Theta\left(\frac{\mathcal{H}}{L}\min\left(1, \frac{d}{|\mathcal{E}|}\right)\right), \quad \beta^\star = \Theta\left(\frac{\mathcal{H}\sqrt{d}}{|\mathcal{V}|}\right).$$

where $\mathcal{H} \in (0, 1]$ is the dataset-dependent homophily ratio.

**Proposition 2 (Computational Complexity):**
For $n = |\mathcal{V}|$ nodes, $m = |\mathcal{E}|$ edges, dimension $d$, $L$ layers:

- **Time**: $O(L\cdot(m+n)\cdot d^2)$ per forward pass
- **Space**: $O(n\cdot d + m)$ for features and graph
- **Parameters**: $O(L\cdot d^2 + n\cdot d)$

Matches standard GNNs while providing geometric adaptation.

**Theorem 2 (Universal Geometric Framework):**
ARGNN provides a universal geometric framework that can approximate and generalize existing curvature-based GNNs. Fixed-curvature geometries can be approximated by a constrained parameterization of the learnable diagonal metric $\mathbf{G}_i = \mathrm{diag}(\mathbf{g}_i)$:

- **Euclidean** $\mathbb{E}^d$: $\mathbf{g}_i = \mathbf{1}$
- **Hyperbolic** $\mathbb{H}^d$: $\mathbf{g}_i = c_h\mathbf{1}$, where $0 < c_h < 1$
- **Spherical** $\mathbb{S}^d$: $\mathbf{g}_i = c_s\mathbf{1}$, where $c_s > 1$
- **Product Manifold** e.g. $\mathbb{H}^{d_1}\times\mathbb{S}^{d_2}\times\mathbb{E}^{d_3}$: $\mathbf{g}_i = (\underbrace{c_h, \ldots, c_h}_{d_1}, \underbrace{c_s, \ldots, c_s}_{d_2}, \underbrace{1, \ldots, 1}_{d_3})^\top$
- **ARGNN** (Ours): $\mathbf{g}_i \in \mathbb{R}^d_{++}$ (node-adaptive)

The representation hierarchy of diagonal metric $\mathbf{G}_i$ induced manifold is: $\mathcal{M}_{\mathrm{Euc}} \subset \mathcal{M}_{\mathbb{H}/\mathbb{S}} \subset \mathcal{M}_{\mathrm{Product}} \subset \mathcal{M}_{\mathrm{full}}$, showing ARGNN can approximate the prior geometric GNNs.

## Experimental Results

**Due to space constraints, we present representative results in this poster. Full comparisons with all baselines and metrics are in the paper.**

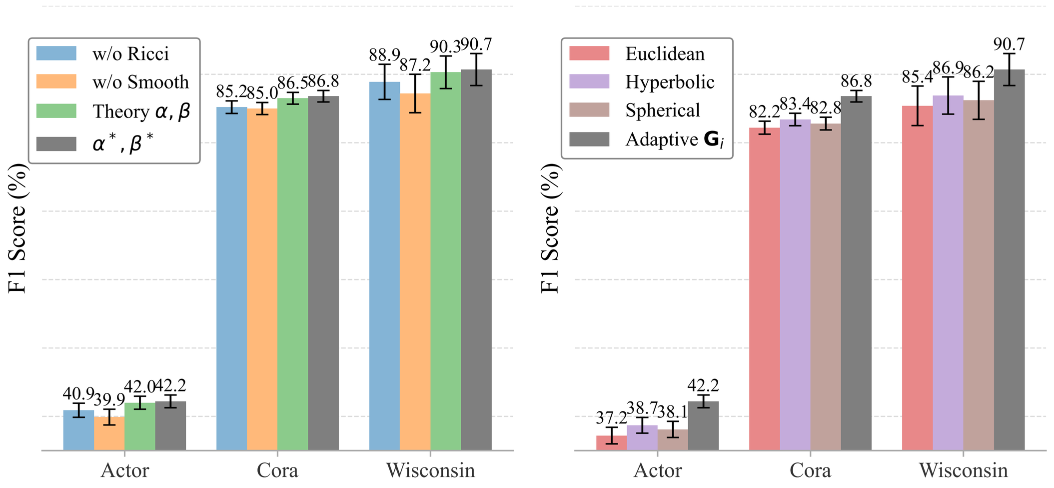
**Node Classification (F1-score %↑ ± 95% Confidence Interval)**, 60%/20%/20% splits from GeomGCN (Pei et al. 2020).

| | Homophilic | | | Heterophilic | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Method** | Cora | CiteSeer | PubMed | Actor | Chameleon | Squirrel | Texas | Cornell | Wisconsin |
| GCN | 75.21±0.28 | 67.30±1.05 | 83.75±0.07 | 31.12±0.96 | 61.16±0.23 | 43.06±0.33 | 75.61±0.07 | 67.72±1.19 | 59.46±3.25 |
| HGCN | 78.50±0.14 | 69.55±0.39 | 83.72±0.21 | 35.89±0.29 | 60.18±0.57 | 39.93±0.35 | 88.11±1.12 | 72.88±1.15 | 86.70±3.70 |
| $\kappa$-GCN | 78.71±1.37 | 68.14±0.34 | 85.18±0.52 | 34.57±0.26 | 62.12±0.49 | 43.04±0.31 | 85.03±0.63 | 86.36±0.64 | 86.90±3.80 |
| CurvDrop | 82.50±0.70 | 72.80±0.60 | 85.20±0.50 | 39.50±1.00 | 67.30±1.40 | 50.10±1.30 | 88.20±2.10 | 89.80±1.80 | 87.50±1.90 |
| CUSP | 83.45±0.15 | 74.21±0.02 | 87.99±0.45 | 41.91±0.11 | 70.23±0.61 | 52.98±0.25 | 89.43±2.72 | 88.31±1.09 | 88.30±0.80 |
| GNRF | 82.10±0.80 | 73.50±0.50 | 86.80±0.40 | 40.80±0.90 | 68.90±1.20 | 49.82±1.50 | 90.80±1.30 | 90.50±1.10 | 88.10±1.40 |
| **ARGNN** | **86.83±0.84** | **74.80±1.26** | **88.59±0.25** | **42.18±0.33** | **70.44±1.27** | **53.12±1.45** | **92.28±1.59** | **90.85±0.33** | **90.65±2.34** |

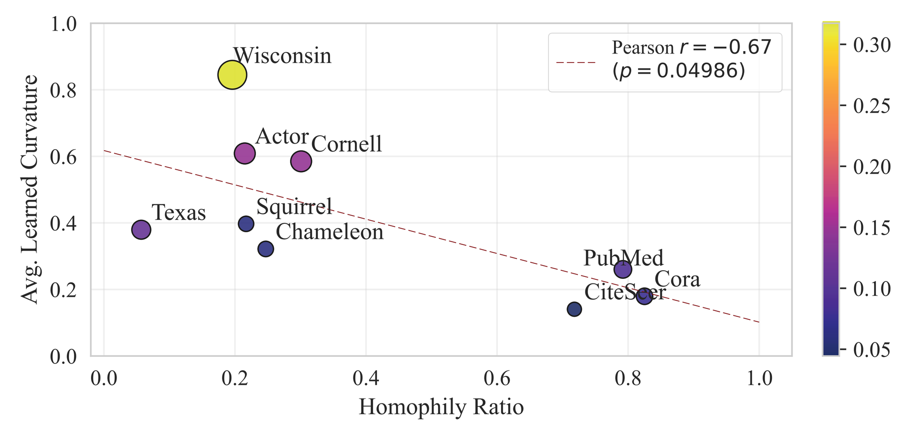
**Link Prediction (AUROC %↑ ± 95% Confidence Interval)**, 80%/5%/15% splits from PyG (He et al. 2024).

| | Homophilic | | | Heterophilic | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Method** | Cora | CiteSeer | PubMed | Actor | Chameleon | Squirrel | Texas | Cornell | Wisconsin |
| GCN | 82.15±0.80 | 79.84±0.92 | 83.42±0.85 | 70.78±0.95 | 81.83±0.74 | 84.61±0.68 | 64.70±1.15 | 65.90±1.05 | 74.20±0.96 |
| HGCN | 86.48±0.70 | 84.92±0.78 | 86.98±0.72 | 73.82±0.91 | 85.35±0.70 | 86.25±0.64 | 65.82±1.12 | 67.12±1.07 | 74.82±0.92 |
| $\kappa$-GCN | 87.15±0.64 | 85.52±0.69 | 87.48±0.70 | 73.94±0.89 | 85.90±0.68 | 86.52±0.62 | 66.92±1.05 | 67.35±1.02 | 75.22±0.93 |
| CurvDrop | 87.10±0.66 | 85.60±0.67 | 87.00±0.60 | 72.60±0.78 | 86.80±0.66 | 86.30±0.63 | 65.40±0.99 | 66.80±0.97 | 74.10±0.89 |
| CUSP | 89.85±0.60 | 88.50±0.62 | 87.90±0.58 | 74.20±0.74 | 87.20±0.66 | 86.60±0.61 | 67.50±0.95 | 67.80±0.92 | 74.50±0.85 |
| GNRF | 87.70±0.65 | 86.90±0.64 | 87.10±0.60 | 73.50±0.75 | 86.90±0.67 | 86.10±0.62 | 66.70±0.98 | 66.90±0.95 | 73.30±0.88 |
| **ARGNN** | **91.03±0.72** | **90.13±0.82** | **88.62±0.55** | **76.40±0.70** | **91.60±0.65** | **88.10±0.60** | **69.30±0.53** | **69.25±0.90** | **77.48±3.06** |

**Ablation Studies: Complete ablations including regularization components, hyperparameter sensitivity, network depth, embedding dims, and computational efficiency are detailed in our paper.**

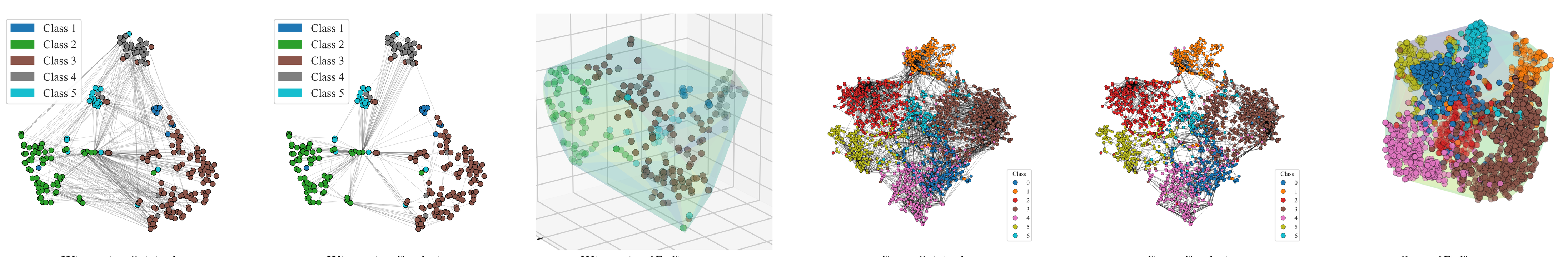



(a) Regularization impact (b) Fixed/Adaptive geometries




**(c) Homophily $\mathcal{H}$ vs. learned geometry.** Avg. learned curvature across datasets with marker size/colour encodes the mean *Neighbour-Relative Metric Dispersion* ($\mathrm{NRMD} = \frac{1}{|\mathcal{E}|}\Sigma_{(i,j)\in\mathcal{E}}\frac{\|\mathbf{g}_i - \mathbf{g}_j\|_2}{\frac{1}{2}(\|\mathbf{g}_i\|_2 + \|\mathbf{g}_j\|_2)}$). Heterophilic graphs exhibit 2–4× higher curvature and greater metric dispersion, confirming our design motivation and ARGNN's capability.

## Learned Geometry Analysis & Visualization




**Visualization of learned geometry on Wisconsin (heterophilic, $\mathcal{H} = 0.196$) and Cora (homophilic, $\mathcal{H} = 0.825$).** The geodesic rewiring based on learned metric distances reveals clearer class separation than the original topology, while the 3-D t-SNE embedding with curvature visualization (violet→flat, yellow→strongly curved) demonstrates that ARGNN learns dataset-adaptive geometry—bending more strongly on heterophilic graphs where diverse local structures require richer geometric representations.