IJCAI 2025 Tutorial
Towards Low-Distortion
Graph Representation Learning

- Session 3 -

Geometry-guided Graph Representation Learning

Presenter: Xingcheng Fu

北京航空航天大學
BEIHANG UNIVERSITY

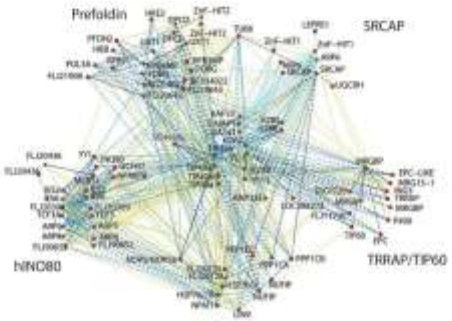
广西師範大學
GUANGXI NORMAL UNIVERSITY

Graph: everything is connected!

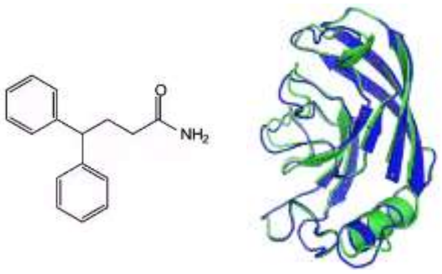
IJCAI 2025 Guangzhou August 29-31

Graphs can model complex environments, and it can help people deeply understand the patterns and mechanisms in the interconnections of all things.

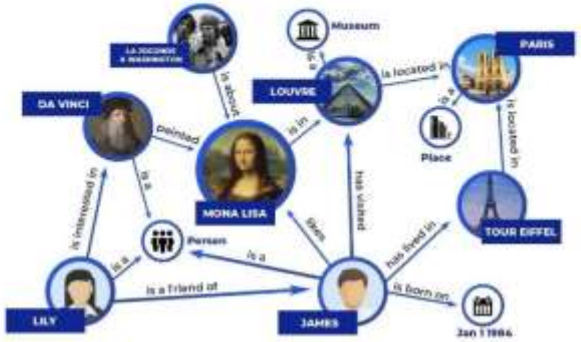
Social Network



Biological Network



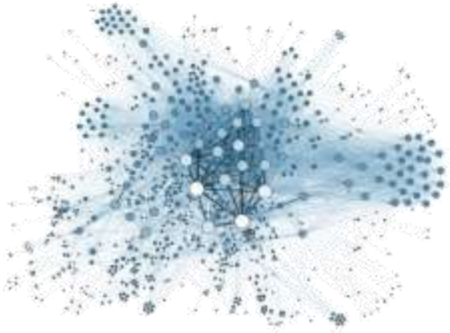
Chemical Molecules



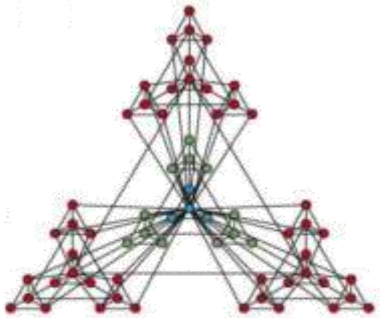
Knowledge Graph

Topological Properties

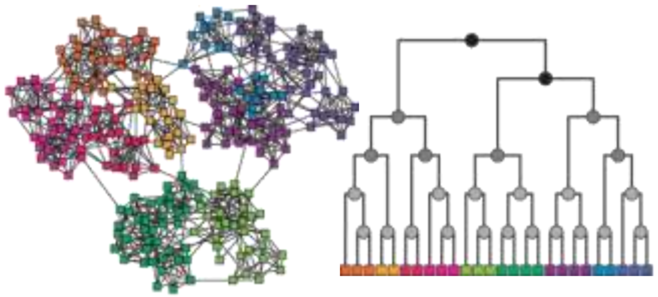
Small Word



Scale-free



Self-similarity

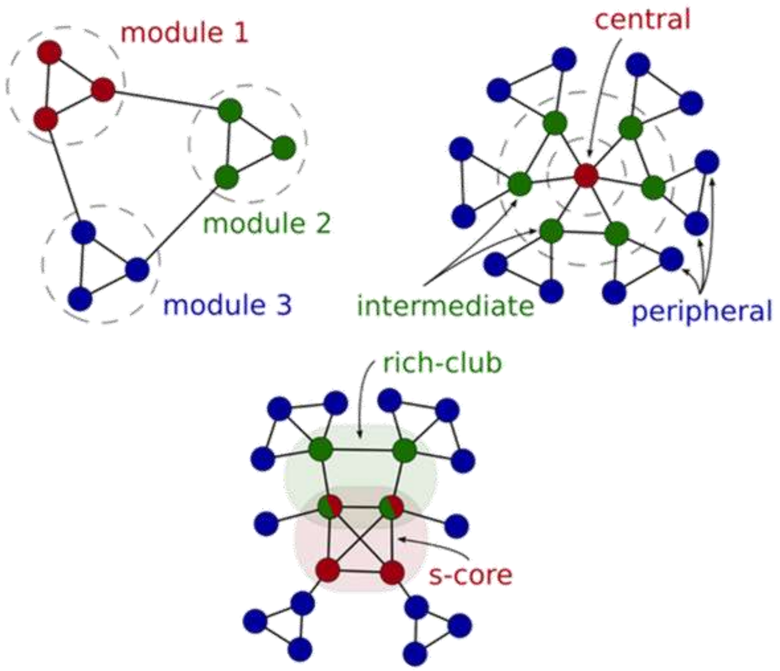


Hierarchy

Perspectives of graph learning

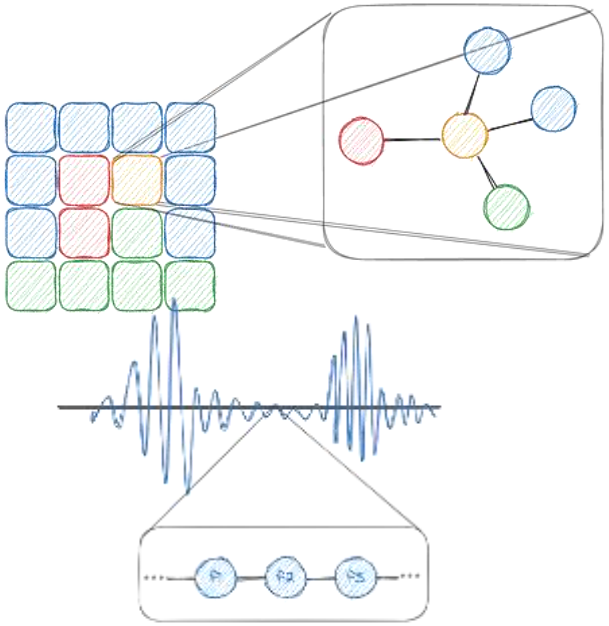
Different theoretical perspectives can provide various inspirations to graph learning.

Graph Theory



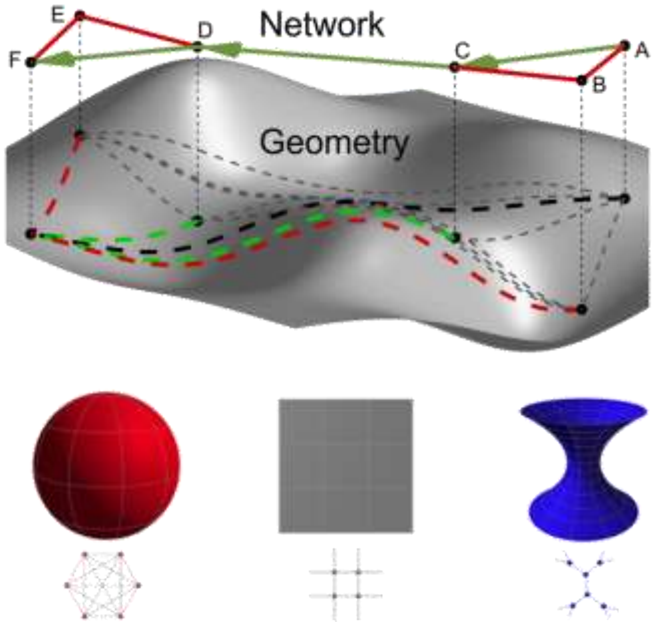
Statistical Characteristics

Information Theory



Signal/Information

Graph Geometry



Topological Properties

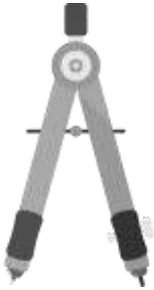
Acknowledgement from https://medium.com/@jlcastrog99/spectral-graph-convolutions-c7241af4d8e2

What is graph geometry?

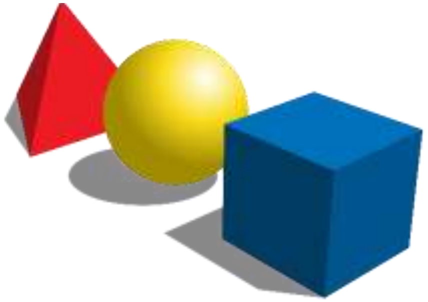
Graph geometry is a mathematical tool that uses geometry to measure various features of graphs.

Classical Geometric Tools

Length

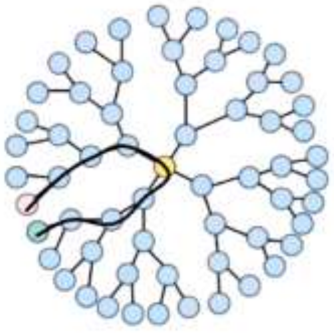


Angle

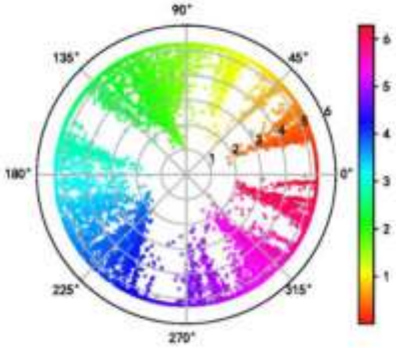


Shape

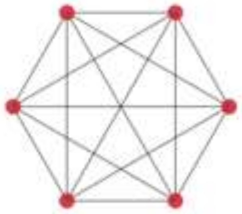
Euclid (fl. 300 BC)



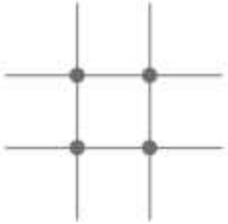
Path



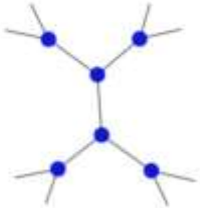
Similarity/Clustering



Clique



Grid



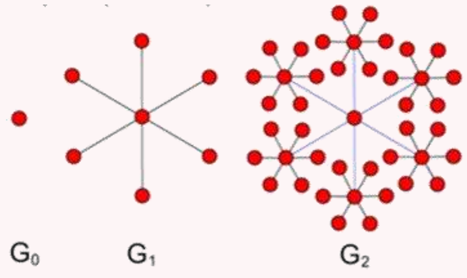
Tree

Topologies

Why do we need graph geometry?

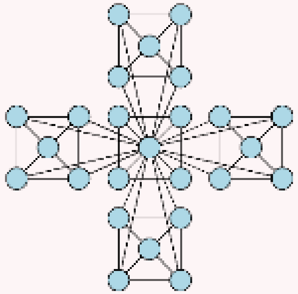
Graph geometry provides geometrically intuitive mathematical tools for measuring topological properties.

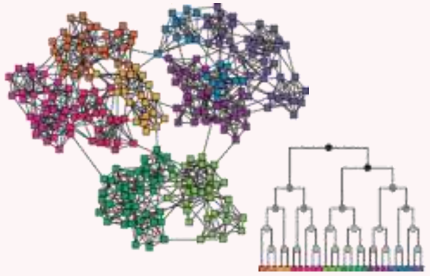
Topologies



Fractal network


Small-world


Self-similarity

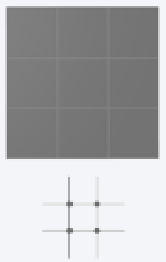

Hierarchy

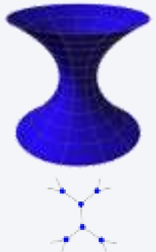
Geometric Intuition

Geometric priors of space

Curved space can be understood as a continuous approximation of the underlying structure.

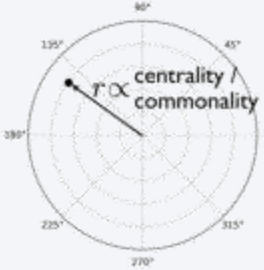

Spherical (Cycle/Clique)


Euclidean (Grid)

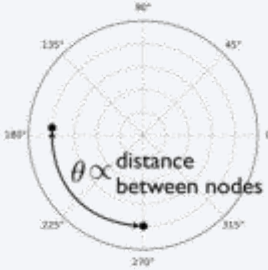

Hyperbolic (Tree-like)

Geometric properties of data

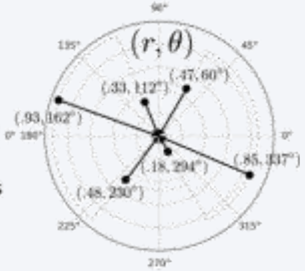
Geometric features carry significant information and semantics



Radius (Centrality)

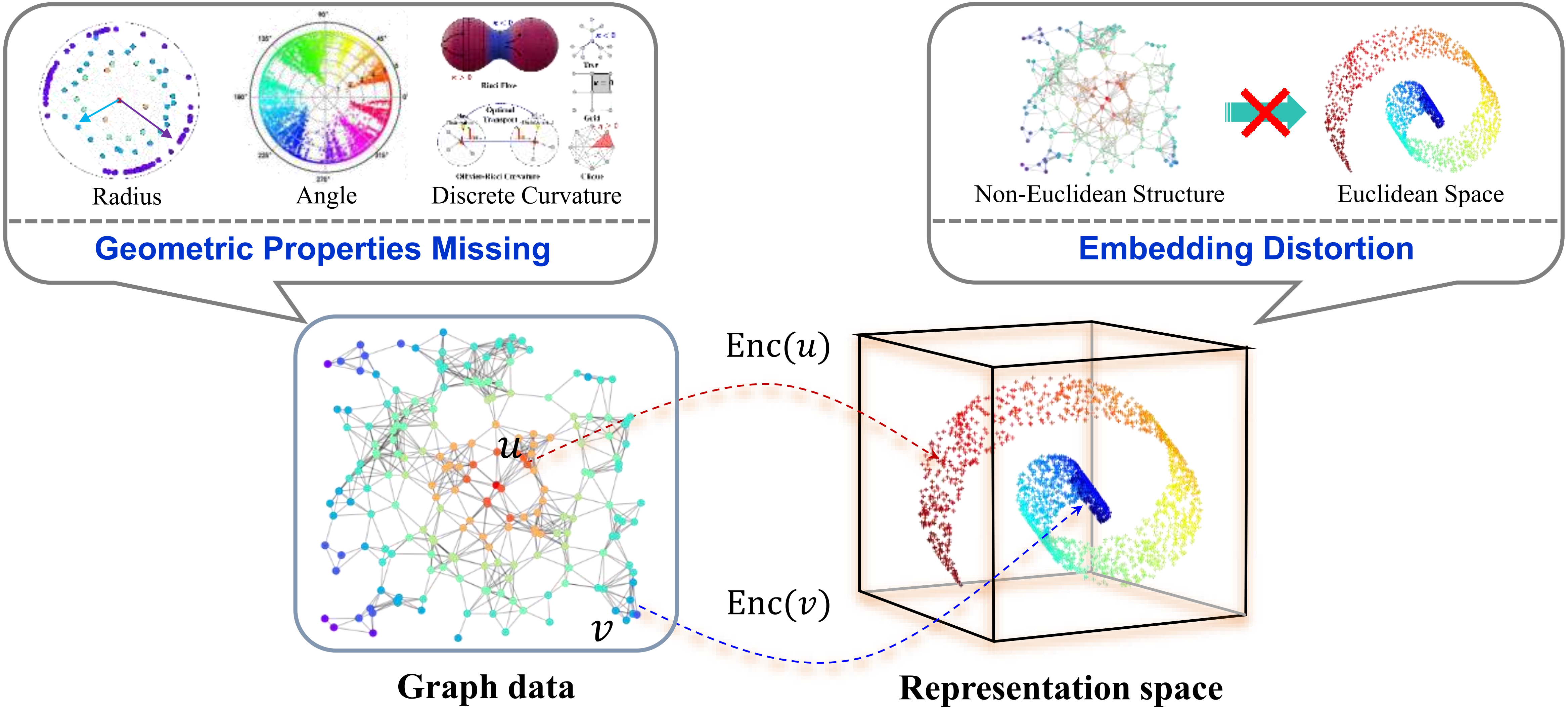


Geodesic (Distance)



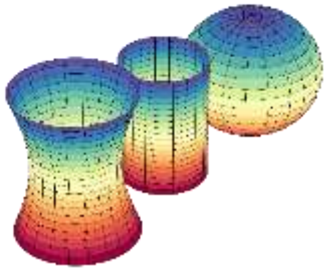
Angle (Clusting)

Geometric perspective of GRL



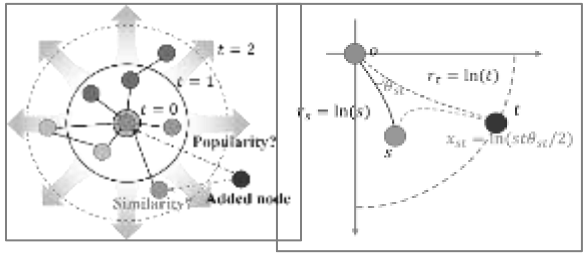
Outline

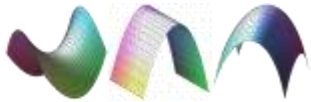
Geometric Space



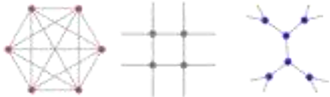
Geometry-guided Low-distortion GRL

Geometric Properties

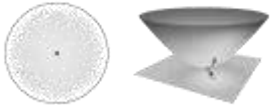




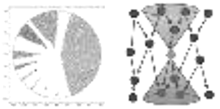
Curvature
- ACE-HGNN(ICDM 2022 Best Paper Candidate)
- CurvGAN(WWW 2022)



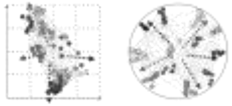
Geo-GFM
- GraphMoRE(AAAI 2025)
- RiemannGFM(WWW 2025)
- GalaxyWalker(CVPR 2025 Highlight Paper)



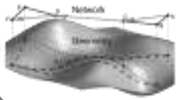
Radius
- HyperIMBA(WWW 2023 Splotlight Paper)
- PoinDP(AAAI 2024)



Angle & Cone
- ConeE(NeurIPS 2021)
- POINE²(ECAI 2024)



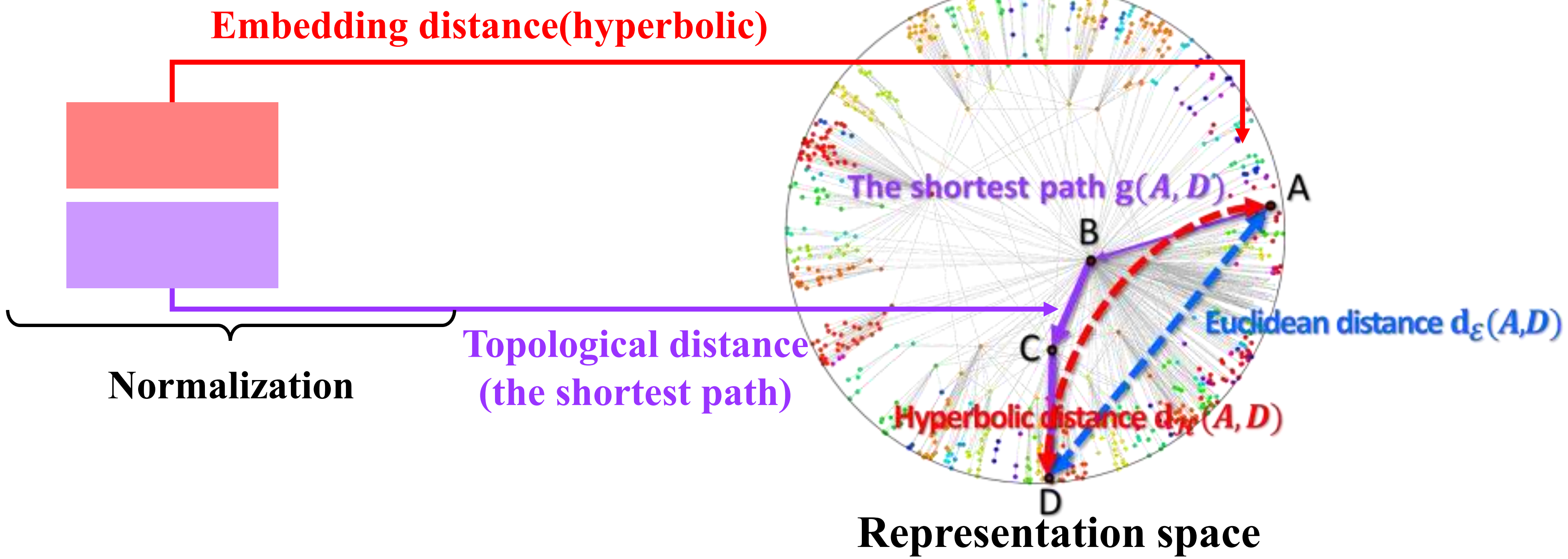
Orientation
- HyerDiff(ICML 2024)



Continualization
- CurvGIB(AAAI 2025)
- LSEnet(ICML 2024 Oral Paper)

Distortion in Geometric Perspective

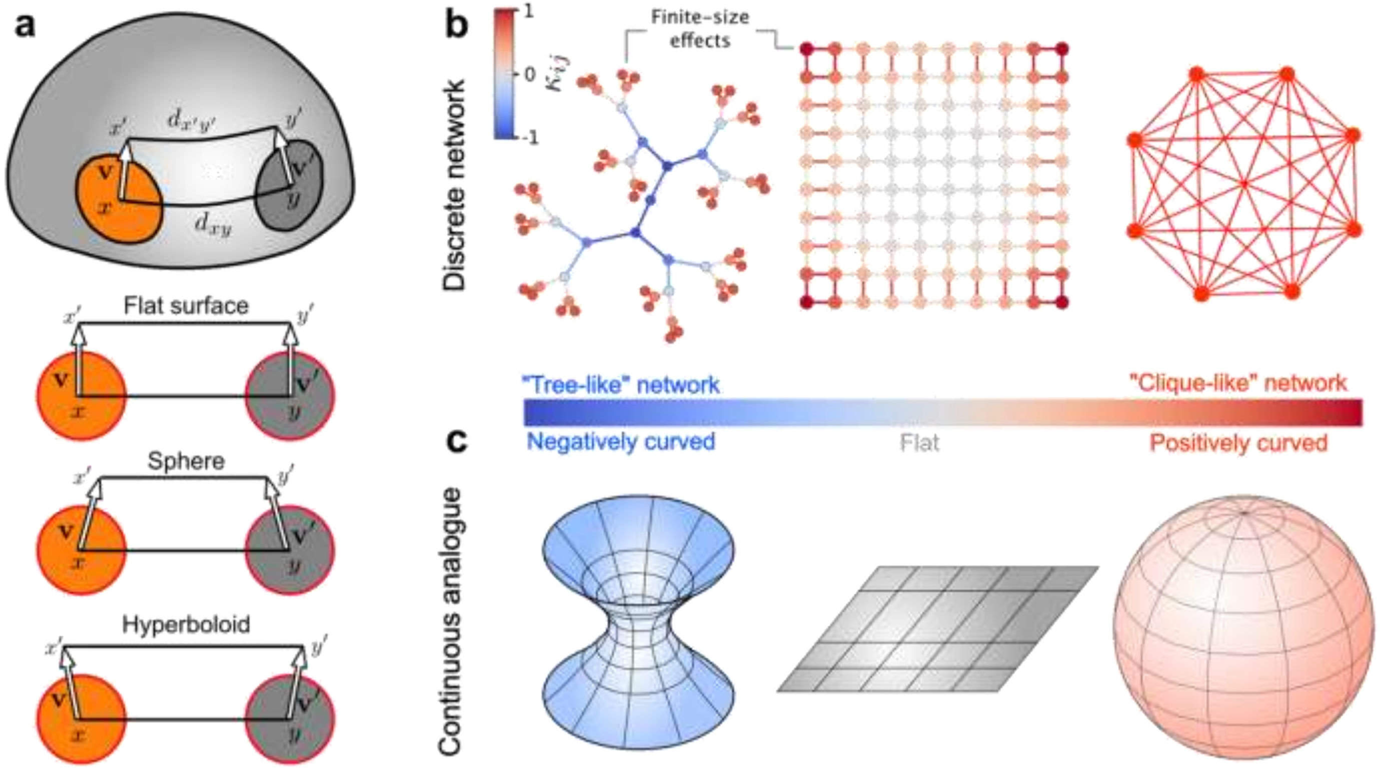
Definition (Embedding Distortion)[1][2]. *Given a graph G with node set V, for each node-pair $(i,\ j) \in V$ the embedding distortion $\mathcal{D}$ in the hyperbolic embedding space $\mathbb{H}$ is defined as:*



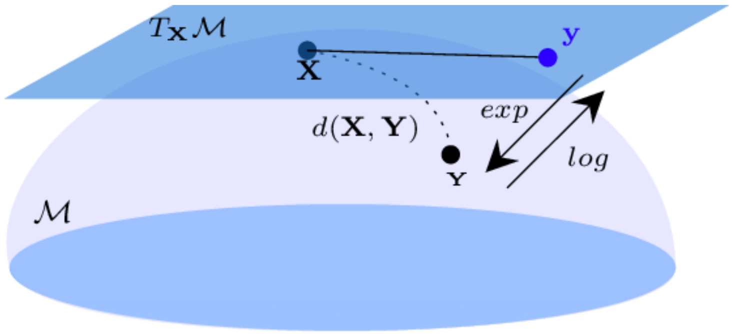
[1] Christopher S, Albert Gu, et al. Representation Tradeoffs for Hyperbolic Embeddings[C]. ICML 2018.
[2] **Xingcheng Fu**, et al. ACE-HGNN: Adaptive curvature exploration hyperbolic graph neural network[C], ICDM 2021, **Best Paper Candidate**.

Riemannian geometry & graphs

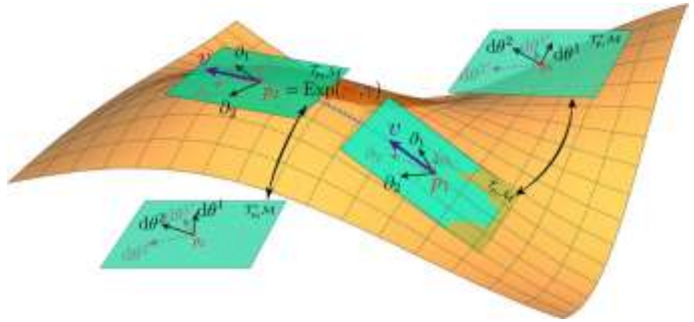
Riemannian geometry provide a powerful and elegant mathematical framework for studying graph deep learning.



Discrete topologies and their continuous analogue manifolds[1]



Exponential & logistical mapping

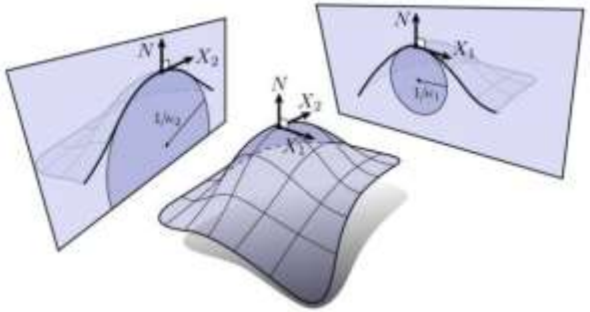


Differential geometry computation

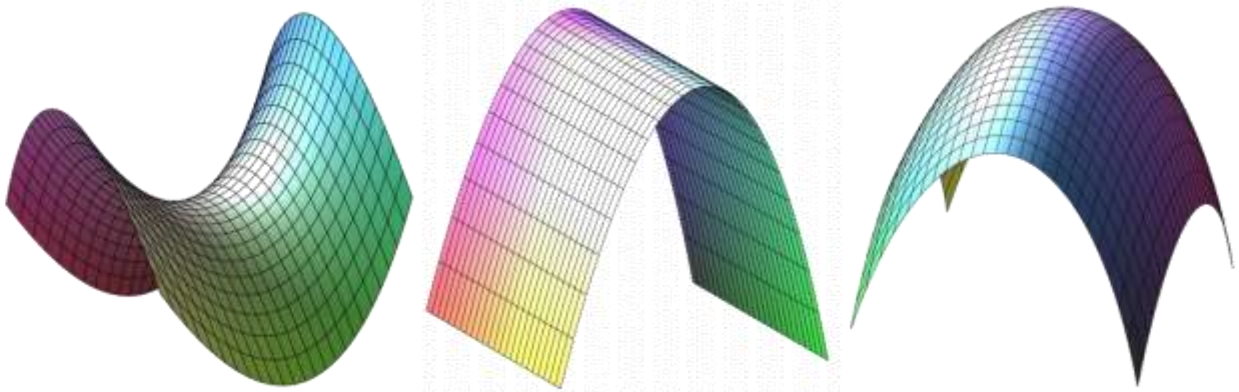
Gosztolai A, Arnaudon A. Unfolding the multiscale structure of networks with dynamical Ollivier-Ricci curvature[J]. Nature Communications, 2021, 12(1): 4561.

Graph curvature

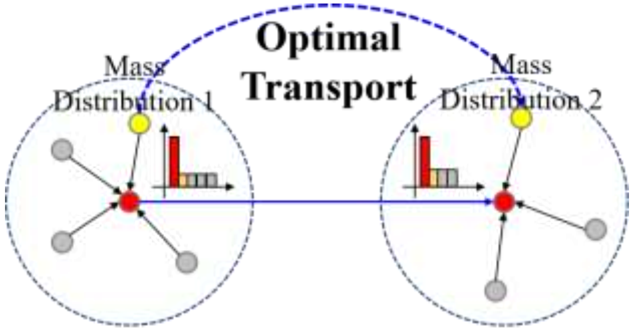
There are two main types of graph curvature: the global ***Gaussian Curvature*** is used to control the curvature of the space; The local ***discrete (Ricci) curvature*** is used to reflect the topological properties of the edges.



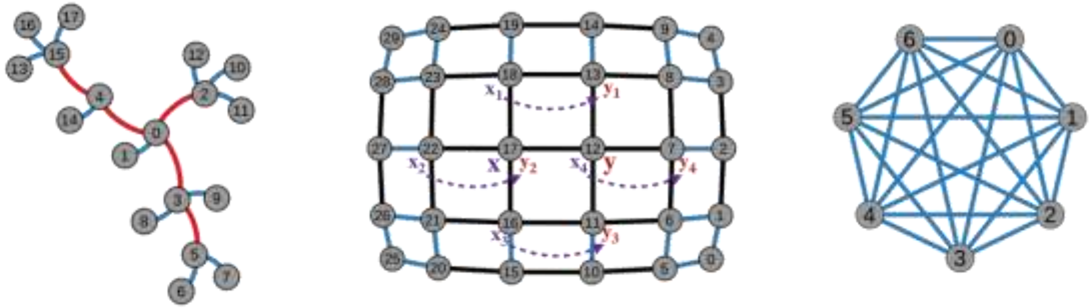
Gaussian curvature:



(a) Surface of Negative Curvature (b) Surface of Zero Curvature (c) Surface of Positive Curvature



Ollivier-Ricci Curvature



(d) Negative Curvature (e) Zero Curvature (f) Positive Curvature

Ni C C, Lin Y Y, Luo F, et al. Community detection on networks with Ricci flow[J]. Scientific reports, 2019, 9(1): 9984.

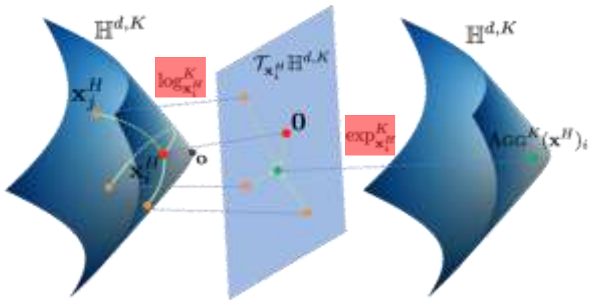
Hyperbolic / Riemannian GNNs

The common non-Euclidean geometric graph neural networks mainly project the non-Euclidean space representation onto the tangent space for GNN aggregation, and then project it back to the non-Euclidean space.

There are mainly two options for space selection: (1) Heuristically estimating the curvature of the space; (2) Using the product space of mixed positive, zero and negative.

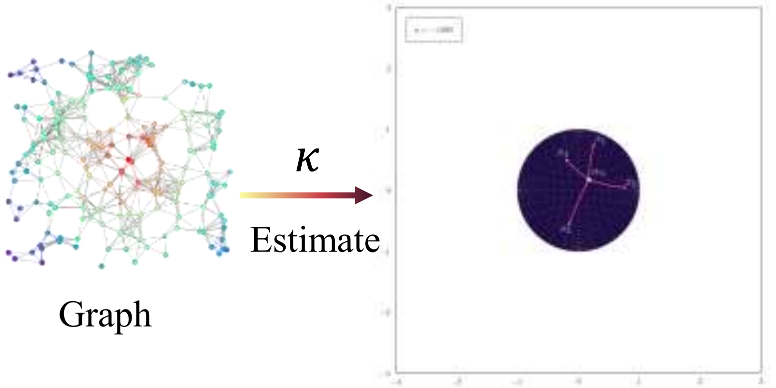
HGNN[1]

Tangent space $\mathbb{H}^n \rightarrow \mathbb{R}^n \rightarrow \mathbb{H}^n$



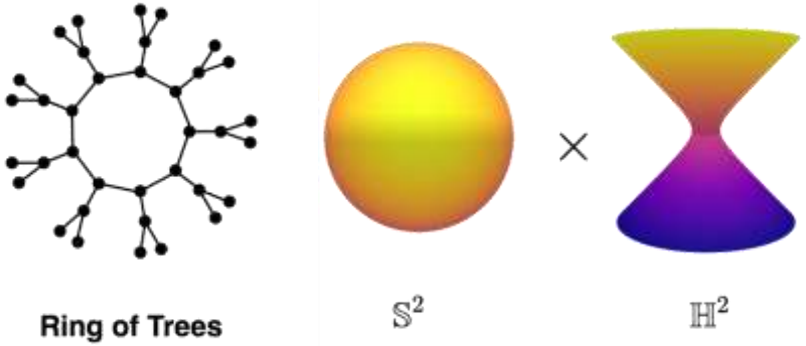
κ-GNN [2]

κ-Stereographic model $\mathfrak{st}^n_\kappa$



Mixed-Curvature [3]

Product space $\mathbb{S}^n \times \mathbb{H}^n$



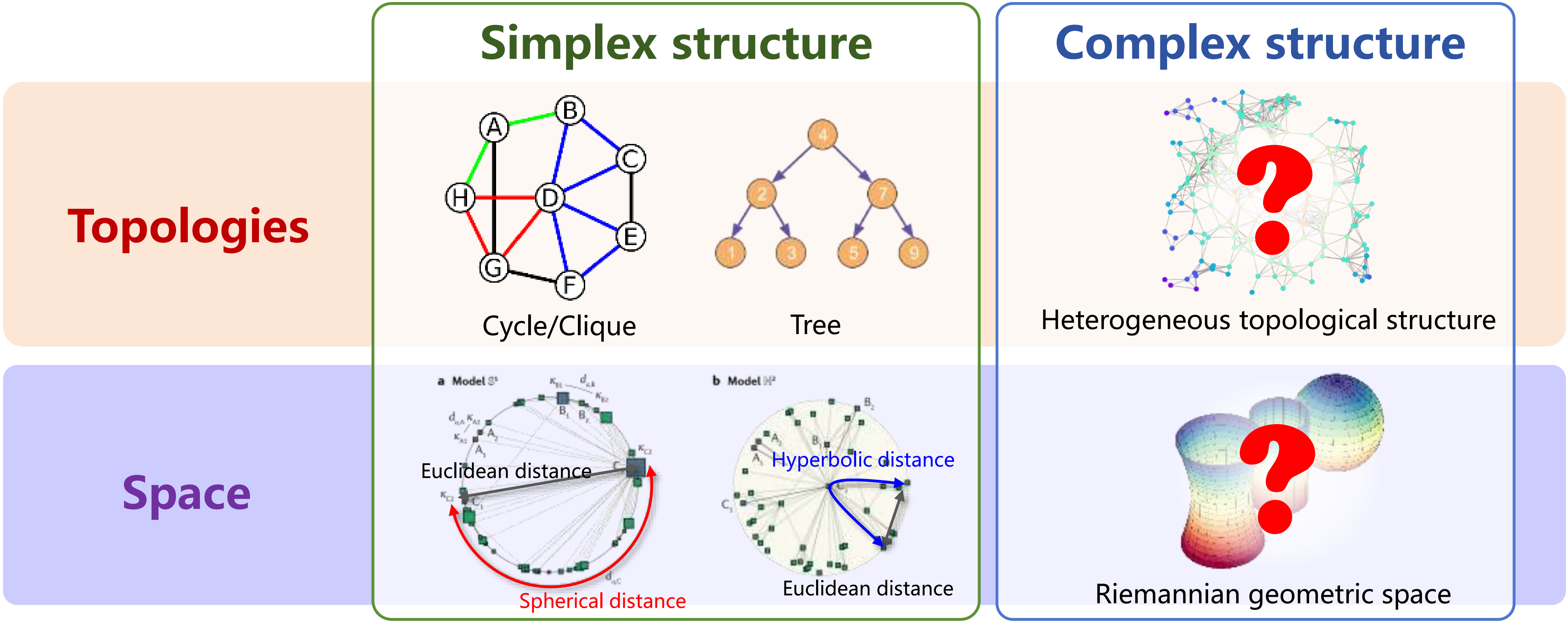
[1] Chami I, Ying Z, Ré C, et al. Hyperbolic graph convolutional neural networks[C]. NeurIPS 2019.

[2] Bachmann G, Bécigneul G, Ganea O. Constant curvature graph convolutional networks[C]. ICML 2020.

[3] Gu A, Sala F, Gunel B, et al. Learning mixed-curvature representations in product spaces[C]. ICLR 2018.

Challenges of embedded space

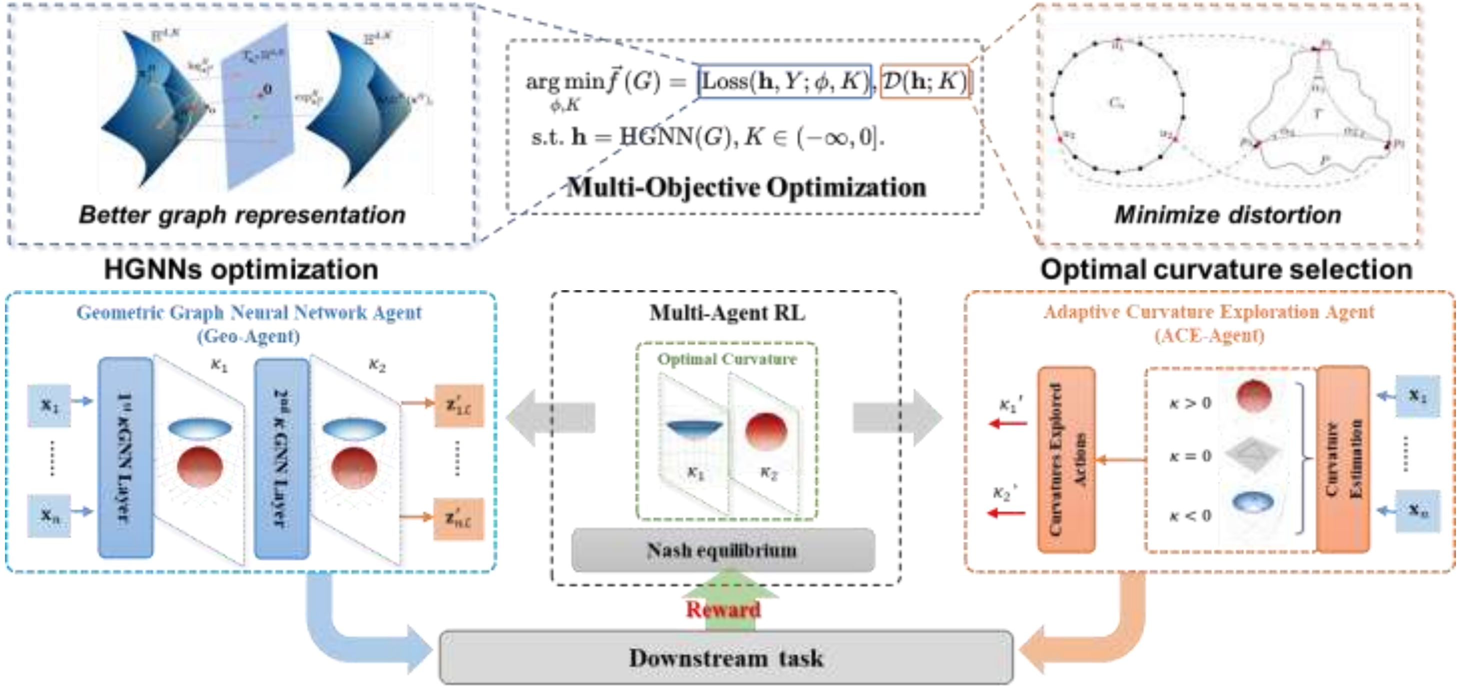
The graph structure in the real world is extremely complex and diverse.
Heuristically selecting the representation space will lead to distortion.



Optimal space exploration

Adaptive Curvature Exploration for any downstream task

ACE-HGNN(hyperbolic) and ACE-GEO(Riemannian) adaptively preserves hierarchical and clique structure for given any graph without any priori knowledge, which allows it to perform supremely across various topology of graphs.



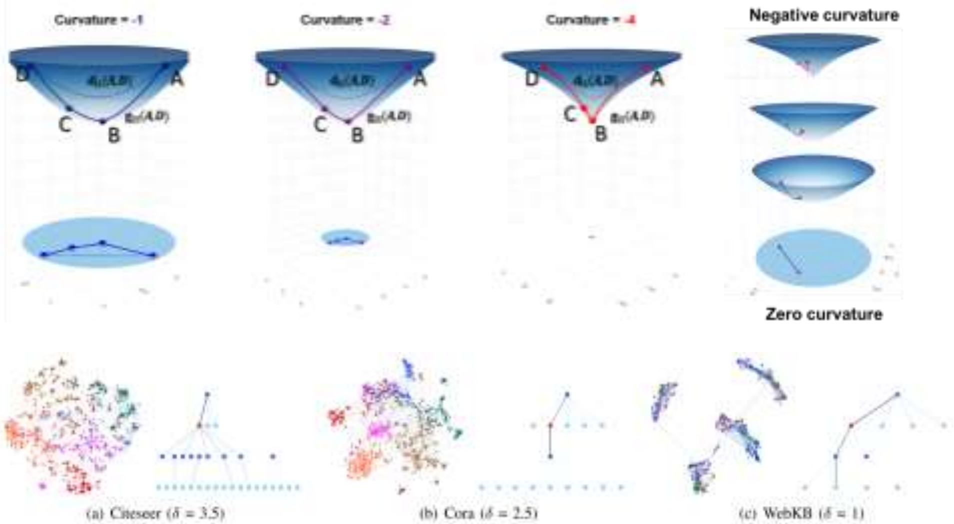


Figure 8. Embeddings (left) and attention weights (right) visualization of ACE-HGNN on datasets with different δ-hyperbolicity.

Paper: https://arxiv.org/pdf/2110.07888.pdf
Code: https://github.com/RingBDStack/ACE-HGNN

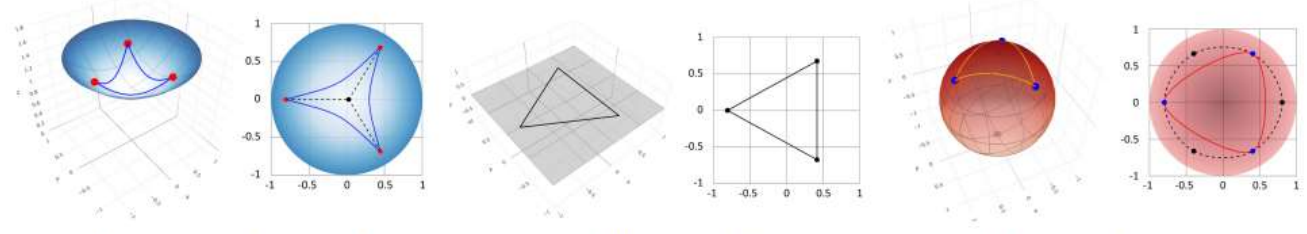
Xingcheng Fu, et al. ACE-HGNN: Adaptive curvature exploration hyperbolic graph neural network[C], ICDM 2021, **Best Paper Candidate**.
Xingcheng Fu, et al. Adaptive Curvature Exploration Geometric Graph Neural Network[J], KAIS 2022, **Best Paper Candidate Invitation Track**

Global-local curvature collaboration

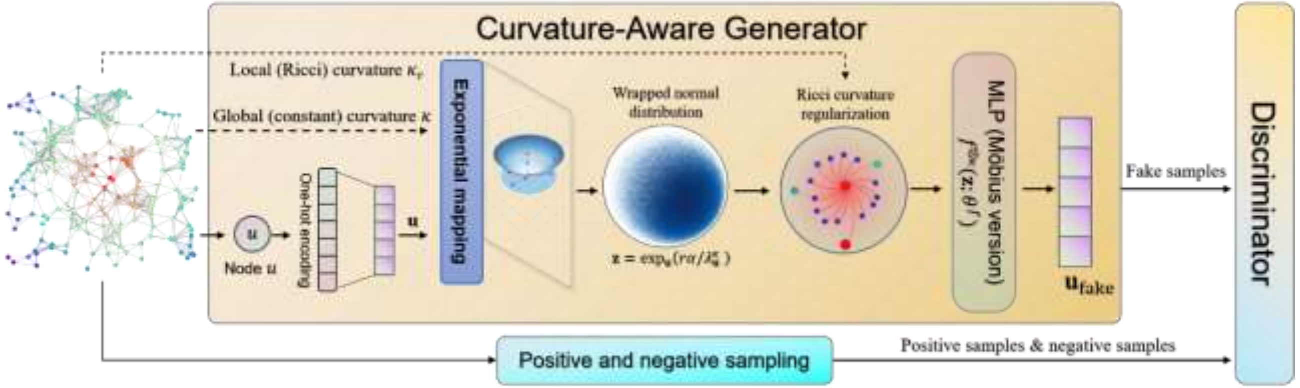
IJCAI 2025 Guangzhou August 29-31

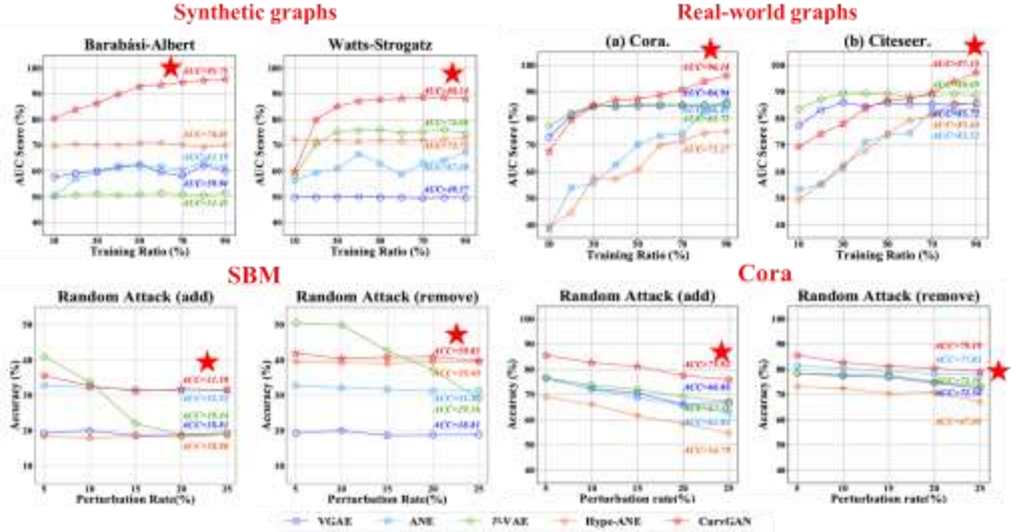
Sampling robust representations by using Gaussian and Ricci curvatures.

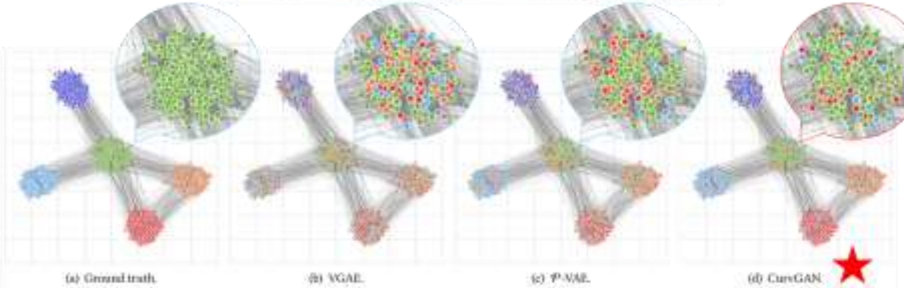
CurvGAN can capture the underlying topology by estimate optimal curvature. It can directly generate fake samples in Riemannian geometric space and refine by Ricci curvature, to obtain more robust representations.



(a) Hyperbolic geometric space ($\kappa < 0$). (b) Euclidean space ($\kappa = 0$). (c) Spherical geometric space ($\kappa > 0$).







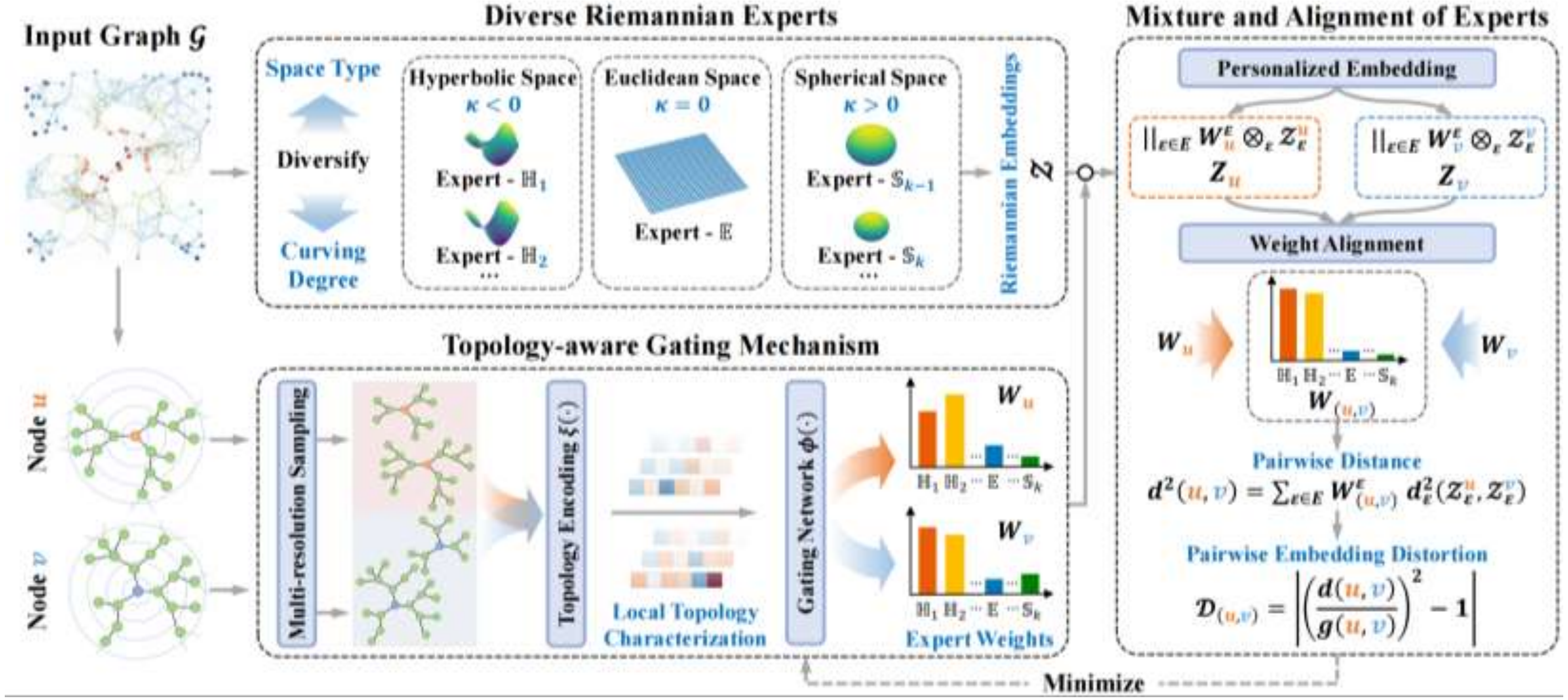
Paper: https://arxiv.org/abs/2203.01604
Code: https://github.com/RingBDStack/CurvGAN

Jianxin Li, **Xingcheng Fu**, et al. Curvature Graph Generative Adversarial Networks[C], WWW 2022.

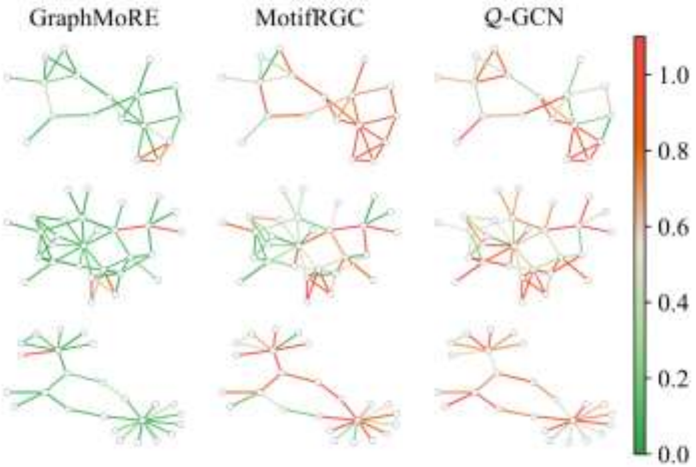
Graph mixture of Riemannian experts

Utilizing Riemannian experts to capture heterogeneous topologies

GraphMoRE construct personalized embedding spaces for nodes, and provide an alignment strategy to calculate pairwise distances, to minimize the distortion of heterogeneous topologies.



Method	Cora	Citeseer	Airport	PubMed	Photo
GraphMoRE	**0.22**	**0.32**	**0.55**	**0.21**	**0.59**
MotifRGC	0.69	0.78	0.64	0.58	OOM
Q-GCN	0.80	0.91	0.70	0.75	0.78



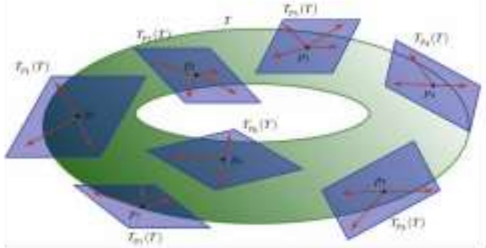
Paper: https://arxiv.org/abs/2412.11085 **Code:** https://github.com/RingBDStack/GraphMoRE

Guo Z, Sun Q, Yuan H, **X Fu**, et al. GraphMoRE: Mitigating Topological Heterogeneity via Mixture of Riemannian Experts[C], AAAI 2025.

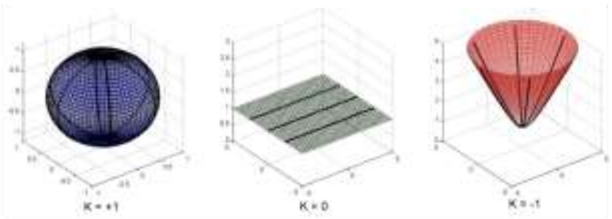
Riemannian graph foundation model

Modeling the structural vocabulary in the Riemannian spaces

Universal pre-trained model (RiemannGFM) on a novel product bundle where the structural vocabulary is learned in Riemannian manifold, offering the shared structural knowledge for cross-domain transferability.



Tangent Bundle

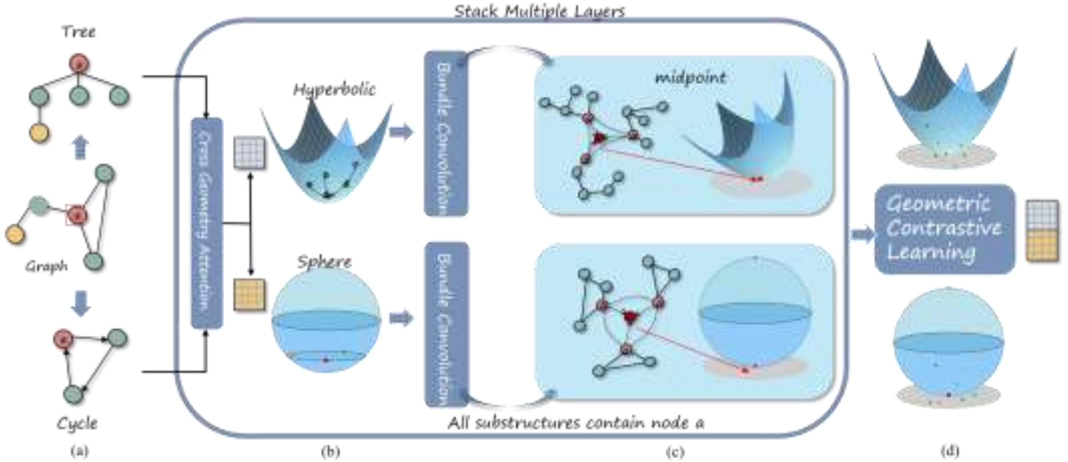


Tangent Bundle

Product Bundle

$$\mathcal{P}^{d_P} = \left(\mathcal{H}^{d_H}_{\kappa_H} \otimes \mathcal{T}\mathcal{H}^{d_H}_{\kappa_H}\right) \otimes \left(\mathcal{S}^{d_S}_{\kappa_S} \otimes \mathcal{T}\mathcal{S}^{d_S}_{\kappa_S}\right), d_P = 2d_H + 2d_S,$$

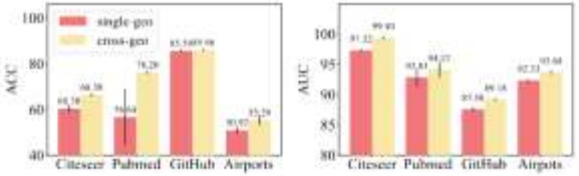
- **The coordinate in the manifold** $p_i \in \mathcal{M}$ contains the relative position in substructures;
- **The encoding in tangent space** $z_i \in \mathcal{T}_{p_i}\mathcal{M}$ carries the information of global structure.



Sample substructures. Vocabulary Learning Module Global Learning Module Geometric Contrast

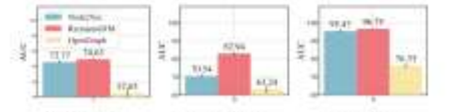
Table 1: Cross-domain transfer learning performance on Citeseer, Pubmed, GitHub and Airport datasets. Node classification and link prediction results are reported. The best results are in boldfaced.

	Method	Node Classification Results								Link Prediction Results							
		Citeseer		Pubmed		GitHub		Airport		Citeseer		Pubmed		GitHub		Airport	
		ACC	F1	ACC	F1	ACC	F1	ACC	F1	AUC	AP	AUC	AP	AUC	AP	AUC	AP
	GCN [3]	70.30	68.56	78.90	77.83	85.68	84.34	50.80	48.09	90.70	92.91	91.16	89.96	87.48	85.34	92.37	94.24
	SAGE [13]	68.24	67.60	77.57	73.61	85.12	77.36	49.16	47.57	87.29	89.03	87.02	86.85	79.13	81.21	92.17	93.56
	DGI [47]	71.30	71.02	76.60	76.52	85.19	84.10	50.10	49.56	96.90	97.05	88.39	87.37	86.39	86.61	92.50	91.63
	GraphMAE2 [14]	**73.40**	71.68	**81.10**	**79.78**	85.23	83.34	52.34	49.02	92.75	89.23	89.46	85.37	87.11	86.23	88.23	90.23
GFM	GCOPE [59]	65.33	62.34	74.15	74.33	82.29	72.89	39.96	36.40	88.60	83.03	90.84	86.45	82.16	83.22	86.17	84.91
GFM	OFA [21]	58.32	65.41	74.40	72.42	-	-	-	-	82.62	83.74	92.26	91.36	-	-	-	-
GFM	GraphAny [60]	66.10	63.01	76.10	70.12	79.45	77.19	47.98	46.88	-	-	-	-	-	-	-	-
GFM	OpenGraph[51]	58.58	**76.78**	58.40	56.49	30.16	30.16	40.45	38.28	76.78	77.35	70.02	72.23	86.72	87.42	85.32	83.25
GFM	LLaGA [5]	59.00	56.91	71.21	63.38	53.33	54.17	38.49	39.89	86.26	83.35	84.04	76.48	71.25	70.63	77.90	74.30
	RiemannGFM	66.38	66.41	76.20	75.83	**85.96**	**85.57**	**55.29**	**53.27**	**99.40**	**98.42**	**94.12**	**91.64**	**89.18**	**93.52**	**93.68**	**96.07**



(a) Node Classification (b) Link Prediction

Figure 3: Ablation on cross-geometric attention.



(a) Airports (b) Pubmed (c) Citeseer

Figure 4: Link prediction results with structural knowledge

(a) Original (b) GCN (c) RiemannGFM

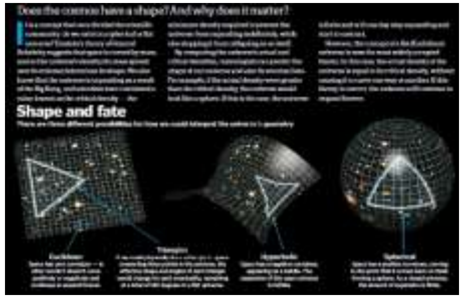
Li Sun, Z Huang, et al. RiemannGFM: Learning a graph foundation model from Riemannian geometry[C]. WWW 2025.

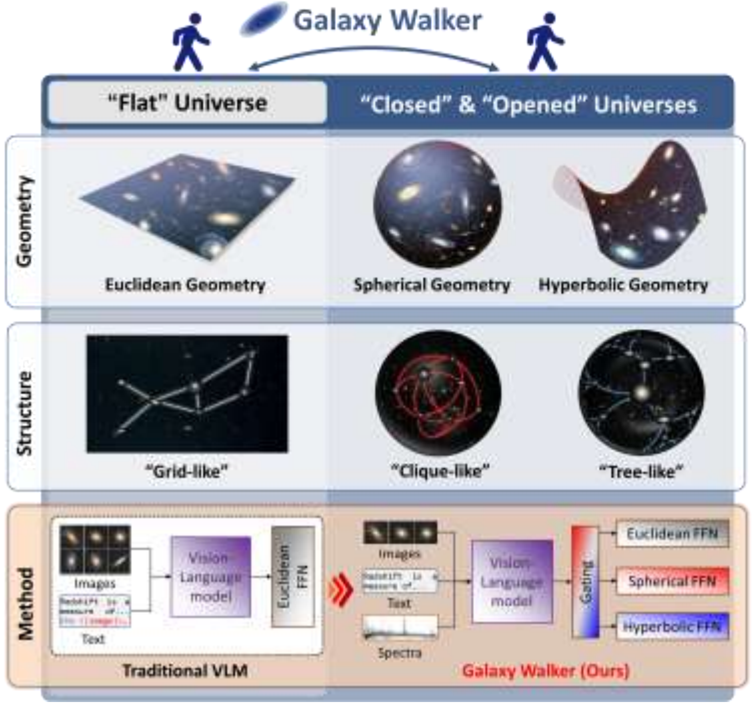
Universe shape understanding

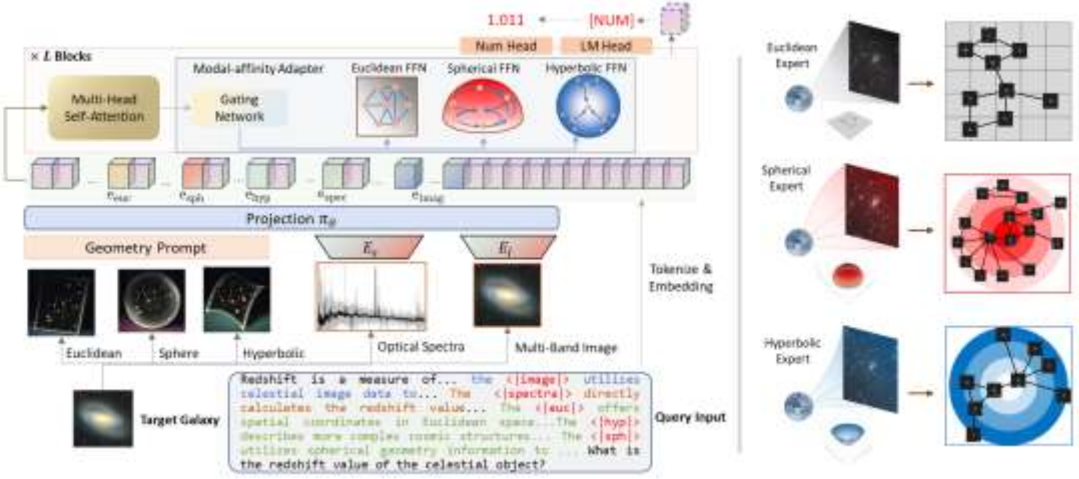
IJCAI 2025 Guangzhou August 29-31

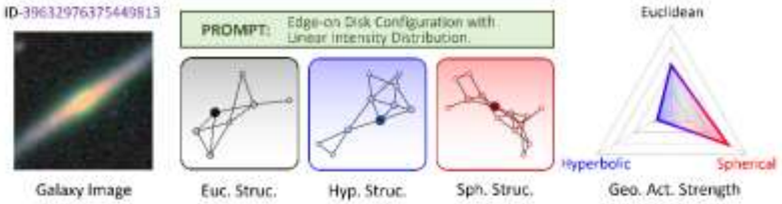
How to understanding the universe shape at galaxy-scale?

We proposed the geometry prompt that generates geometry tokens by random walks across diverse spaces on a multi-scale physical graph for galaxy discovery!





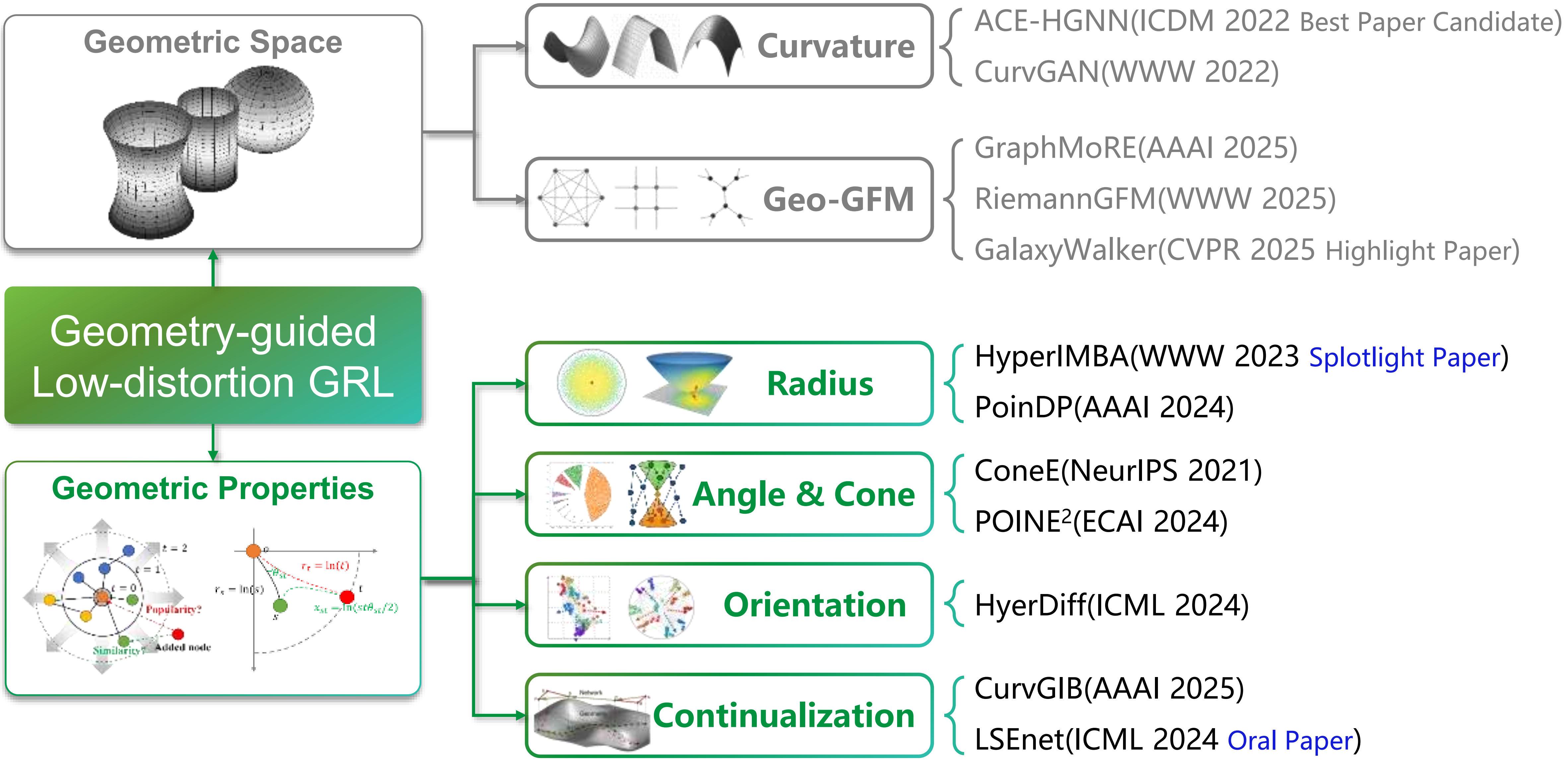




Paper: https://arxiv.org/pdf/2503.18578v1

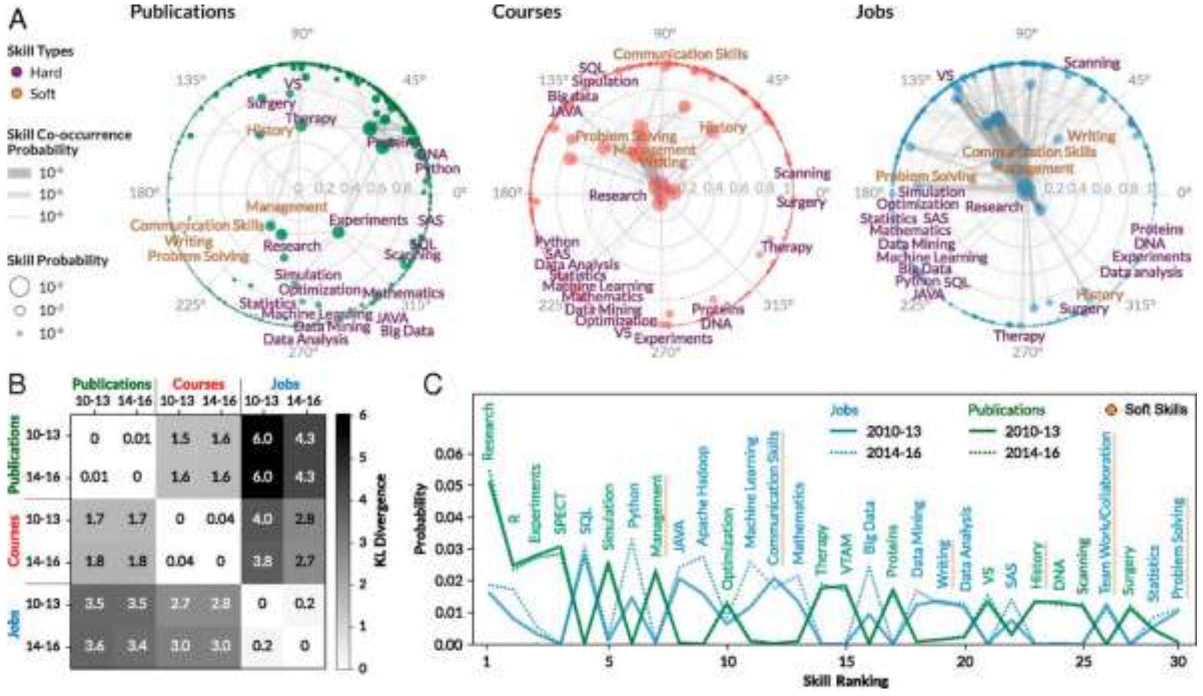
Tianyu Chen, **Xingcheng Fu**, et al. Galaxy Walker: Geometry-aware VLMs For Galaxy-scale Understanding[C], CVPR. 2025, **Highlight Paper**.

Outline

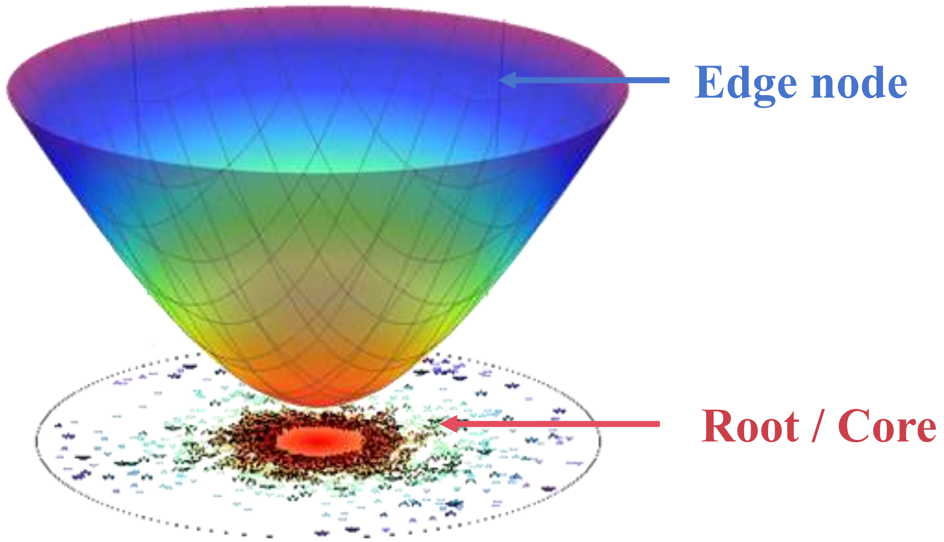

Geometric Space
Curvature
ACE-HGNN(ICDM 2022 Best Paper Candidate)
CurvGAN(WWW 2022)
Geo-GFM
GraphMoRE(AAAI 2025)
RiemannGFM(WWW 2025)
GalaxyWalker(CVPR 2025 Highlight Paper)
Geometry-guided Low-distortion GRL
Geometric Properties
Popularity?
Similarity?
Added node
Radius
HyperIMBA(WWW 2023 Splotlight Paper)
PoinDP(AAAI 2024)
Angle & Cone
ConeE(NeurIPS 2021)
POINE^2(ECAI 2024)
Orientation
HyerDiff(ICML 2024)
Continualization
CurvGIB(AAAI 2025)
LSEnet(ICML 2024 Oral Paper)

Radius and topological roles

Radius of a node—its distance from the center—can powerfully reflect its topological role in the graph. This provides a ***geometric lens*** for understanding hierarchy and influence in complex networks.



Critical need skills mining[1]



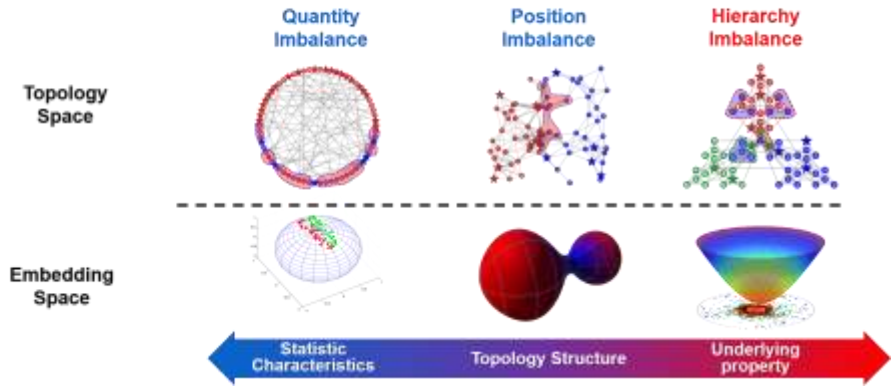
Radius and role of node in hyperbolic space

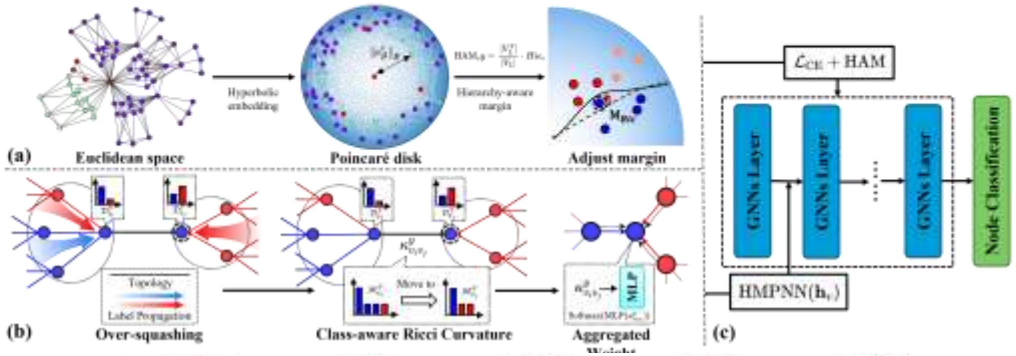
[1] Börner K, et al. Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy[J]. PNAS 2018.

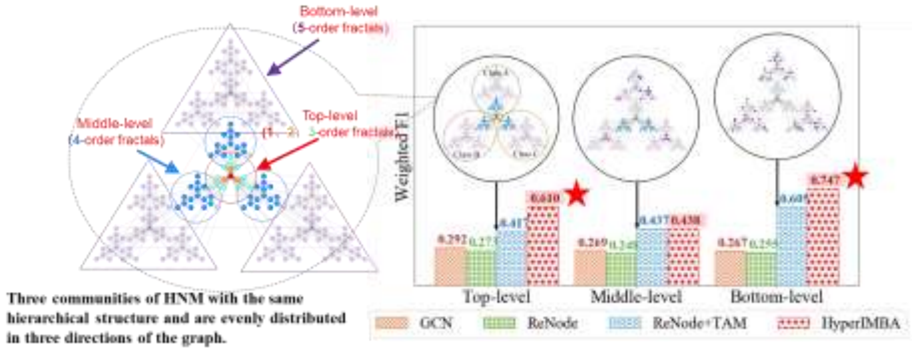
Hierarchy-imbalance learning

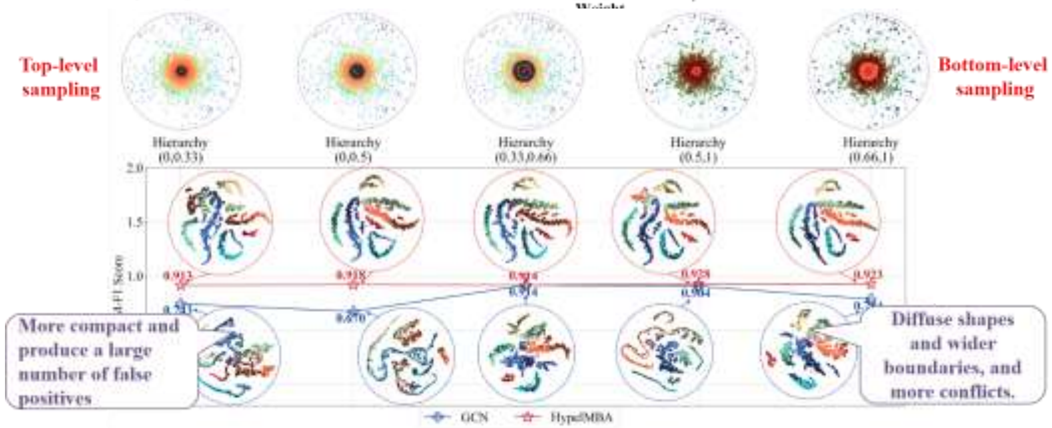
New imbalance issue concept of graph with geometric perspective

This work explore the hierarchy-imbalance as a new topic for semi-supervised node classification, and propose a novel hyperbolic geometric training framework to deal with the imbalance issue.









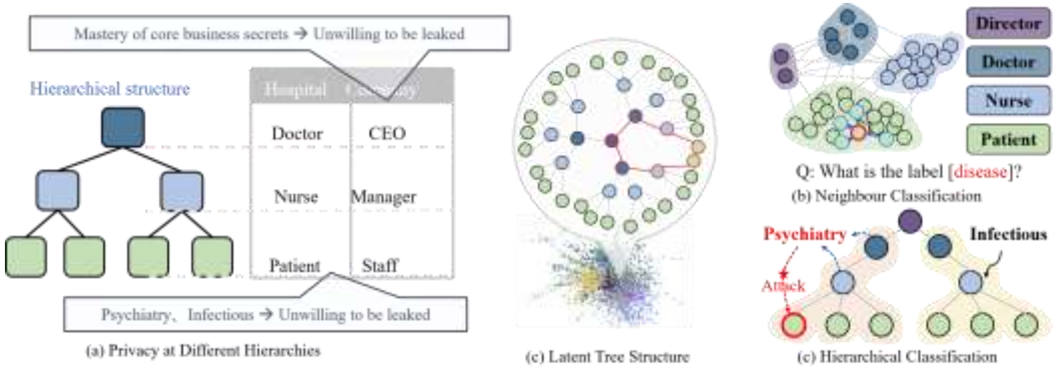
Paper: https://arxiv.org/abs/2304.05059 **Code:** https://github.com/RingBDStack/HyperIMBA

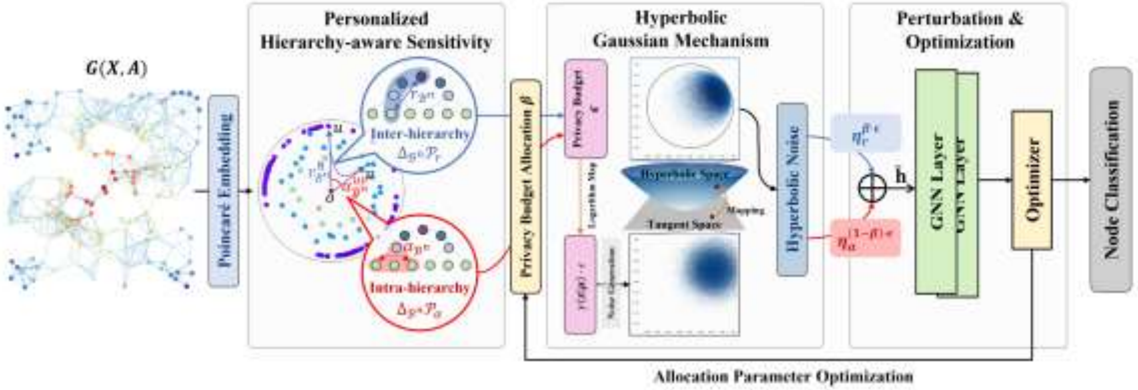
Xingcheng Fu, et al. Hyperbolic Geometric Graph Representation Learning for Hierarchy-imbalance Node Classification. WWW 2023 **Spotlight**.

Hierarchy-aware privacy protection

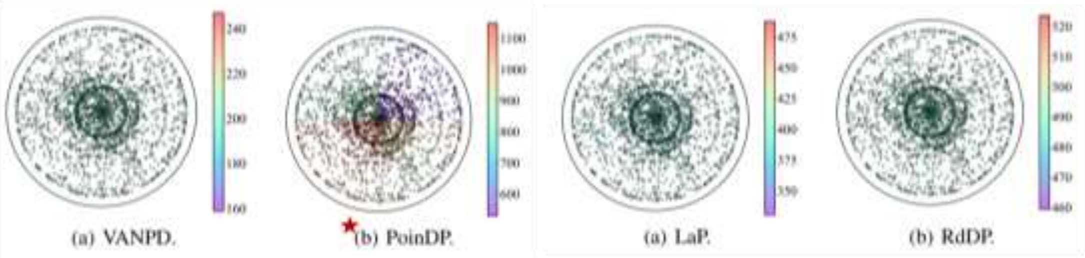
Different topological roles result in more personalized privacy requirements.

We proposed a Hyperbolic Differential Privacy method, it provides much more fine-grained and personalized privacy protection for users.

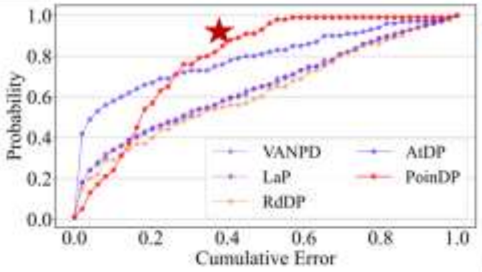




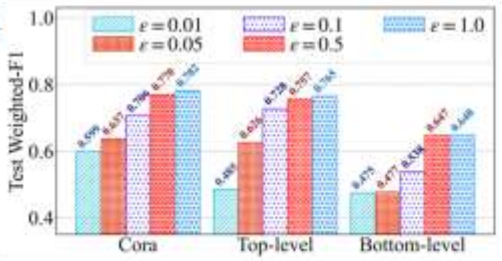
Visualization: PoinDP shows excellent ability to capture inter- and intra-hierarchy information



- Cumulative Error Distribution: PoinDP grows the fastest and ends its accumulation at 0.5.
  - a focused imposition of noise (Personalize)



- Parameter sensitivity ϵ: Smaller values indicating greater privacy protection power, less usability, and more information loss
  - PoinDP provides stronger protection for the *Bottom-level* nodes



Paper: https://arxiv.org/abs/2312.12183 **Code:** https://github.com/RingBDStack/PoinDP

Yuecen Wei, Haonan Yuan, **Xingcheng Fu**, et al. Poincaré Differentially Private for Hierarchy-aware Graph Embedding. AAAI 2024.

Geometry of causality in physics

"Fate lies within the light cone."

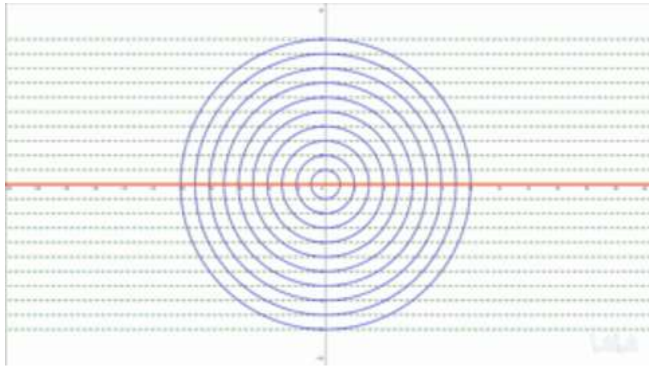
— Liu Cixin, The Dark Forest

"光锥之内皆是命运。"

— 刘慈欣,《三体-黑暗森林》

Minkowski spacetime: taking time to be an imaginary fourth spacetime coordinate ict, where c is the speed of light and i is the imaginary unit.

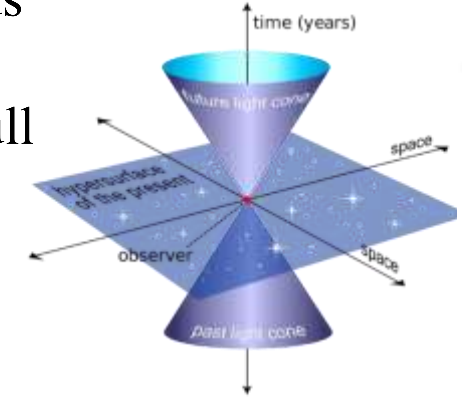
$$x^2 + y^2 + z^2 + ict = \text{const.}$$

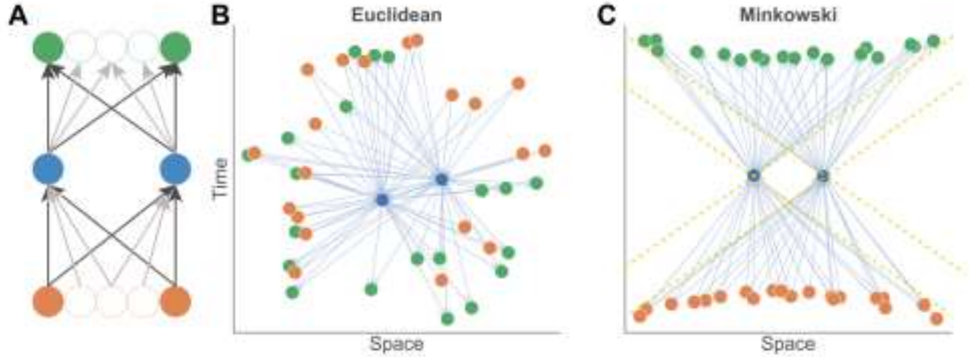



Special relativity (Albert Einstein in 1905)

Light cone and causal structure: Vector fields are called timelike, spacelike or null if the associated vectors are timelike, spacelike or null at each point where the field is defined.

$$\underbrace{(x_0}_{\text{time}}, \underbrace{\mathbf{x})}_{\text{space}} \equiv (\underbrace{x_0}_{\text{causality}}, \underbrace{x_1, \dots, x_n}_{\text{features}})$$



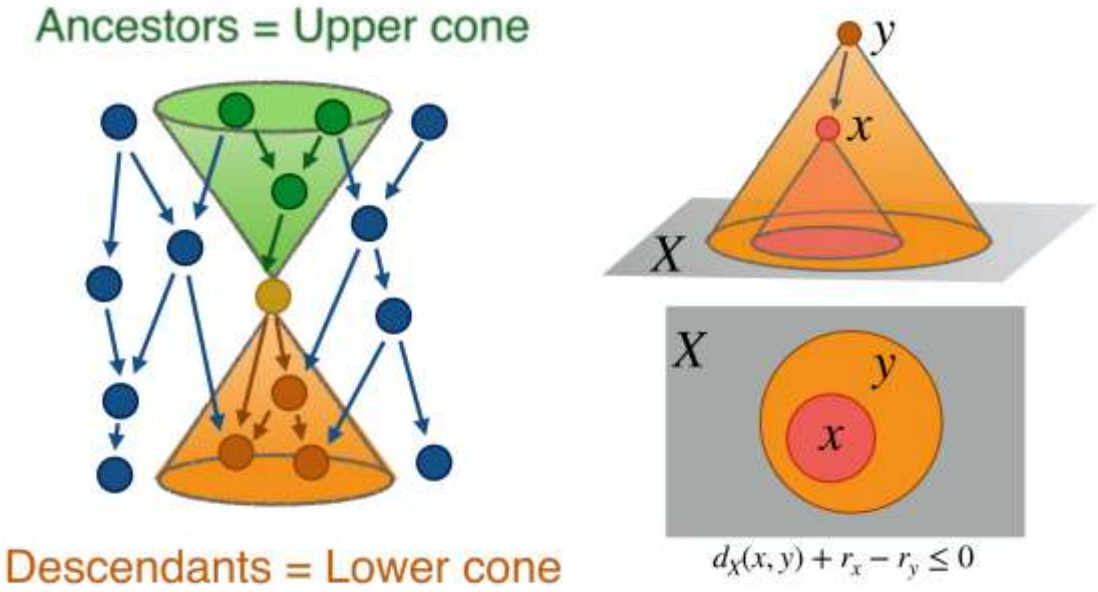


Directed graph in Euclidean space and Minkowski spacetime[1]

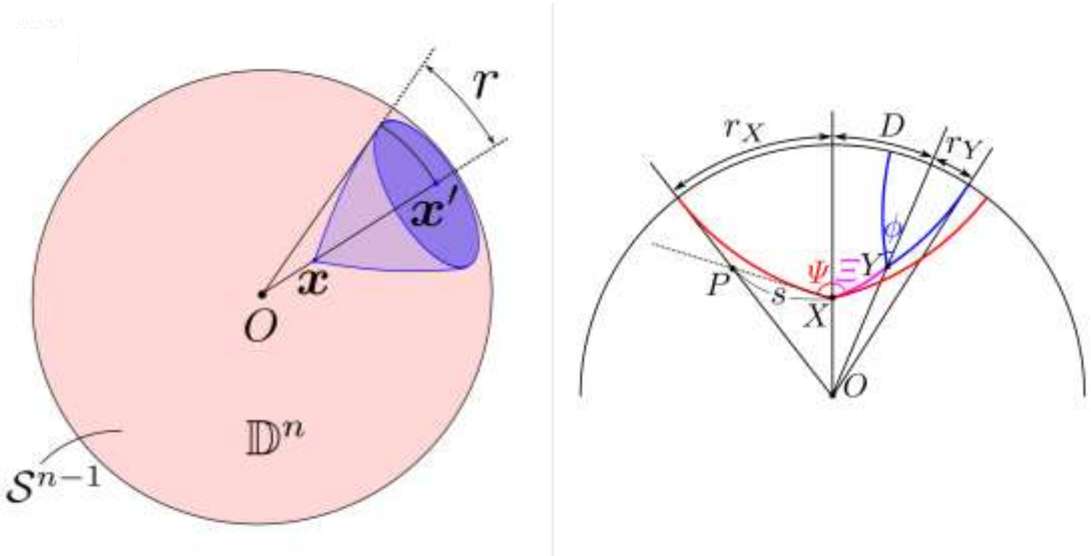
[1] Sim A, Wiatrak M L, Brayne A, et al. Directed graph embeddings in pseudo-riemannian manifolds[C]. ICML 2021.

Angle and cone of DAG modeling

In the geometric perspective, transitive relation of directed acyclic graph (DAG) induces special "light cones" in the embedding space.



Flat cones[1]



Hyperbolic entailment cones[2]

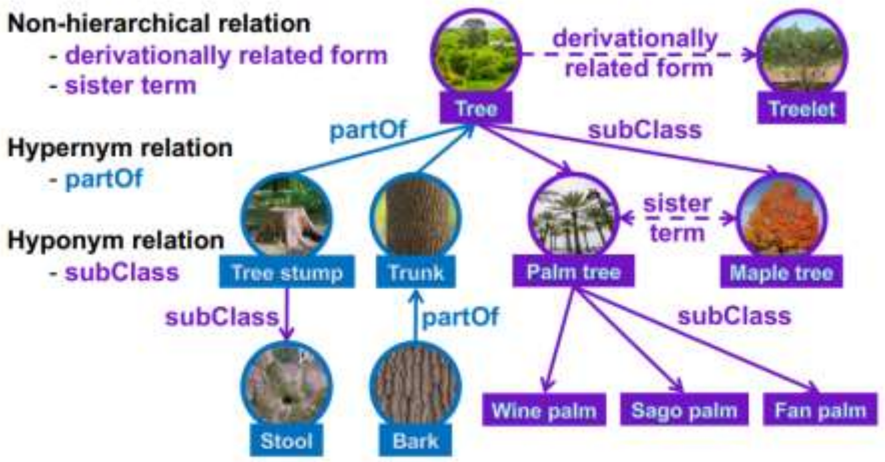
[1] Suzuki R, Takahama R, Onoda S. Hyperbolic disk embeddings for directed acyclic graphs[C]. ICML 2019.

[2] Ganea, O.E., Becigneul, G., and Hofmann, T. Hyperbolic Entailment Cones for Learning Hierarchical Embeddings. ICML 2018.

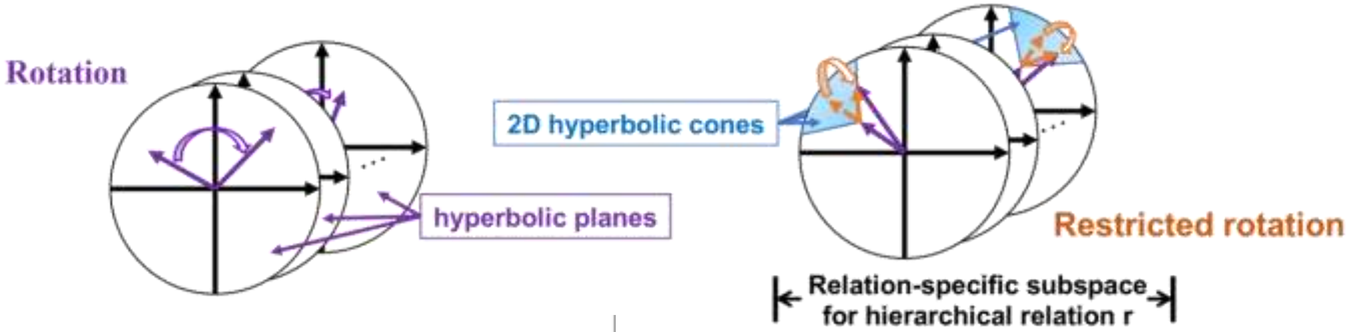
Modeling knowledge hierarchies

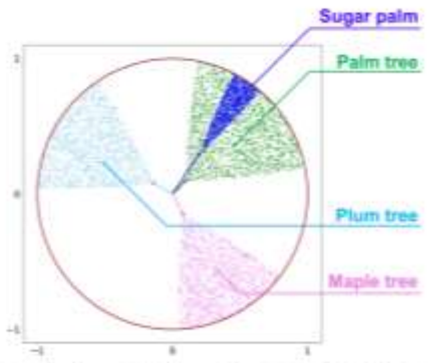
Using restricted rotation transformation to model hierarchical relations

ConE (Cone Embedding for knowledge graphs), the first knowledge graph (KG) embedding method that can capture the transitive closure properties of heterogeneous hierarchical relations as well as other non-hierarchical properties.

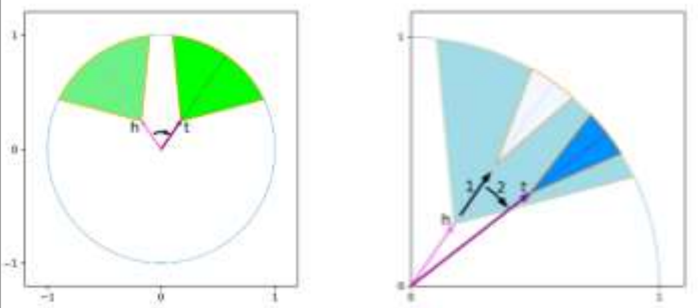


(a) Multiple heterogeneous hierarchies in knowledge graph.





(b) Hyperbolic entailment cones in 2D hyperbolic plane for $K = 0.1$.



(a) Cone rotation (b) Restricted cone rotation

Bai Y, Ying Z, Ren H, et al. Modeling heterogeneous hierarchies with relation-specific hyperbolic cones[C]. NeurIPS, 2021.

Hyperbolic cone for reasoning

IJCAI 2025 Guangzhou August 29-31

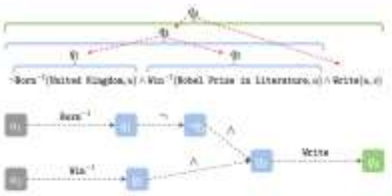
Hyperbolic geometric-based query embedding to handle hierarchies

POINE^2 maps entities and queries as geometric shapes on a Cartesian product space of Poincaré ball spaces, it leaverage the Poincaré radius to represent the different levels of the hierarchy, and the aperture of the angle to indicate semantic differences at the same level of the hierarchy.

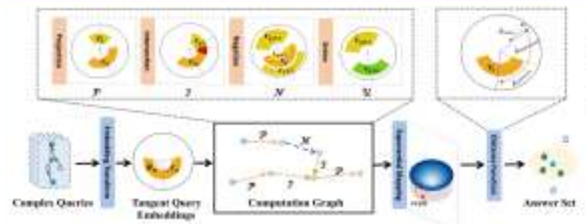
Introduction & Motivation

Reasoning complex logical queries on incomplete and massive knowledge graphs (KGs) remains a significant challenge.

1. How to model entity set and logical operator
   Geometric-based embedding methods can naturally represent answer sets of queries and model logical operations among those sets.
2. How to preserve hierarchical structure of reasoning
   Most existing geometric-based embedding methods have difficulty leveraging hierarchical information of complex query structures



Framework



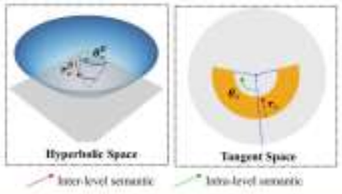
- Initialize query embedding via the embedding of anchor entities
- Compute query embedding progressively using logical operators
- Train or Predict by the distance-based similarity function

We divide the overall distance into inter-distance and intra-distance, which respectively denote inter-level distance and intra-level distance

Query Embedding

- ✓ Poincaré radius to characterize the different levels of the hierarchy
- ✓ Aperture to capture the semantic differences within the same level of the hierarchy,

$$V_q^{\theta} = (r_c^{\theta}, \theta_c^{\theta}, r_o^{\theta}, \theta_o^{\theta})$$



Logical Operators

✓ **Projection Operator**
Perform relational projection from one query embedding to another query embedding

$$V_q' = f(\mathbf{MLP}(r_c + r_c^r; \theta_c + \theta_c^r; r_o + r_o^r; \theta_o + \theta_o^r))$$

✓ **Intersection Operator**
Produce an intersection embedding that represents the intersection of N input conjunctive queries

$$r_{n,2} = \sum_{i=1}^{n} [\alpha_i \cdot r_{i,1}], \alpha_i = \frac{\exp(\mathbf{MLP}(V_{q,i}))}{\sum_{j=1}^{N} \exp(\mathbf{MLP}(V_{q,j}))}, \quad [r_{n,1}/\theta_{n,1}] = \min([r_{n,i}/\theta_{n,i}]_{i=1}^{n}) \cdot \sigma(\mathrm{DeepSetsNet}([r_{n,i}/\theta_{n,i}]_{i=1}^{n}))$$

✓ **Union Operator**
Use the embeddings of the conjunctive queries to represent the disjunction embedding

✓ **Negation Operator**
Perform the difference operation between the entire space and the region indicated by input query embedding

Table 1: Statistics of query structures. For the training set, 'EPFO' query structures contain 1p, 2p, 3p, 2i, 3i, and 'Other' query structures contain 2in, 3in, inp, pni, pin.

Dataset	Train EPFO	Train Other	Valid 1p	Valid Other	Test 1p	Test Other
FB15k	273,710	27,371	59,097	8,000	67,016	8,000
FB15k-237	149,689	14,968	20,101	5,000	22,812	5,000
NELL	107,982	10,798	16,927	4,000	17,034	4,000
WN18RR	103,479	10,347	5,202	1,000	5,356	1,000

Table 2: MRR results (%) on FB15k, FB15k-237, NELL, and WN18RR. **Avg$_p$** means the average MRR performance of EPFO queries, which only contains ∃, ∧, and ∨. **Avg$_n$** means the average MRR performance of logical queries with negation (¬).

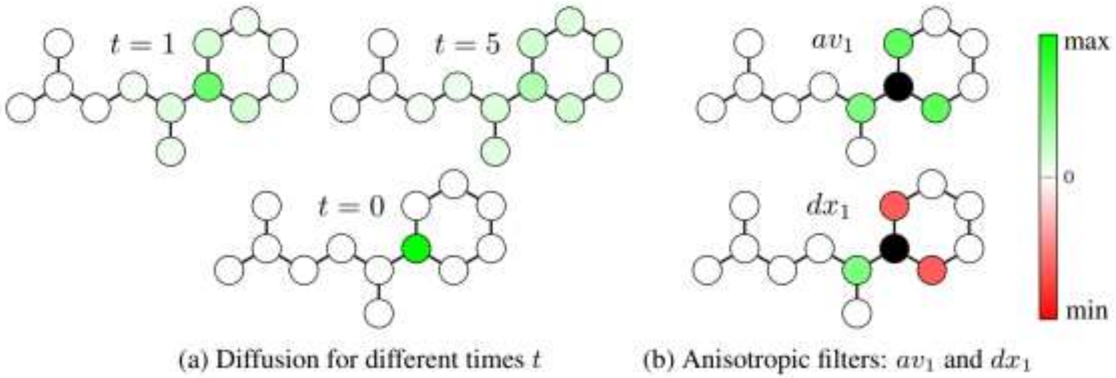
Dataset	Model	Avg$_p$	Avg$_n$	1p	2p	3p	2i	3i	ip	pi	2u	up	2in	3in	inp	pin	pni
FB15k	GQE	28.2	N/A	53.9	15.5	11.1	40.2	52.4	27.5	19.4	22.3	11.7	N/A	N/A	N/A	N/A	N/A
	Q2B	38.4	N/A	70.6	22.5	14.1	55.0	66.7	26.0	39.4	35.0	16.7	N/A	N/A	N/A	N/A	N/A
	ConE	49.3	14.7	73.8	32.0	28.5	64.7	74.0	34.2	49.8	56.4	29.9	18.0	17.9	12.5	9.8	15.1
	POINE^2	**53.9**	**15.4**	**79.1**	**39.2**	**32.8**	**66.4**	**75.6**	**42.4**	**53.7**	**59.8**	**36.2**	**19.1**	**18.5**	**13.4**	**10.2**	**15.8**
FB15k-237	GQE	16.3	N/A	35.0	7.2	5.3	23.3	34.6	10.7	16.5	8.2	5.7	N/A	N/A	N/A	N/A	N/A
	Q2B	20.1	N/A	40.6	9.4	6.8	29.5	42.3	12.6	21.2	11.3	7.6	N/A	N/A	N/A	N/A	N/A
	ConE	23.4	5.9	41.8	12.8	11.0	32.6	47.3	14.0	25.5	14.5	10.8	5.4	8.6	7.8	4.0	3.6
	POINE^2	**24.8**	**6.5**	**42.8**	**13.3**	**11.7**	**35.8**	**50.4**	**15.7**	**27.1**	**14.6**	**11.8**	**6.1**	**9.4**	**8.4**	**4.5**	**4.0**
NELL	GQE	18.6	N/A	32.8	11.9	9.6	27.5	35.2	14.4	18.4	8.5	8.8	N/A	N/A	N/A	N/A	N/A
	Q2B	22.9	N/A	42.2	14.0	11.2	33.3	44.5	16.8	22.4	11.3	10.3	N/A	N/A	N/A	N/A	N/A
	ConE	27.2	6.4	53.1	16.1	13.9	40.0	50.8	17.5	26.3	15.3	11.3	5.7	8.1	10.8	3.5	3.9
	POINE^2	**29.0**	**6.9**	**57.0**	**17.8**	**15.6**	**41.2**	**51.9**	**20.5**	**27.6**	**16.4**	**12.9**	**6.2**	**8.5**	**11.5**	**3.9**	**4.2**
WN18RR	GQE	14.5	N/A	21.1	4.8	3.5	28.8	36.6	14.6	13.6	3.2	4.6	N/A	N/A	N/A	N/A	N/A
	Q2B	21.1	N/A	25.3	5.5	3.9	40.8	70.0	21.1	14.3	4.1	5.1	N/A	N/A	N/A	N/A	N/A
	ConE	34.6	23.5	49.5	19.4	15.9	63.5	81.4	35.8	21.1	10.5	14.5	13.5	67.6	8.9	9.2	18.3
	POINE^2	**36.8**	**25.2**	**51.8**	**21.5**	**18.2**	**65.8**	**83.4**	**37.1**	**25.7**	**11.7**	**16.7**	**15.1**	**70.8**	**10.0**	**9.5**	**20.6**

Paper: https://core.ac.uk/download/pdf/591452439.pdf

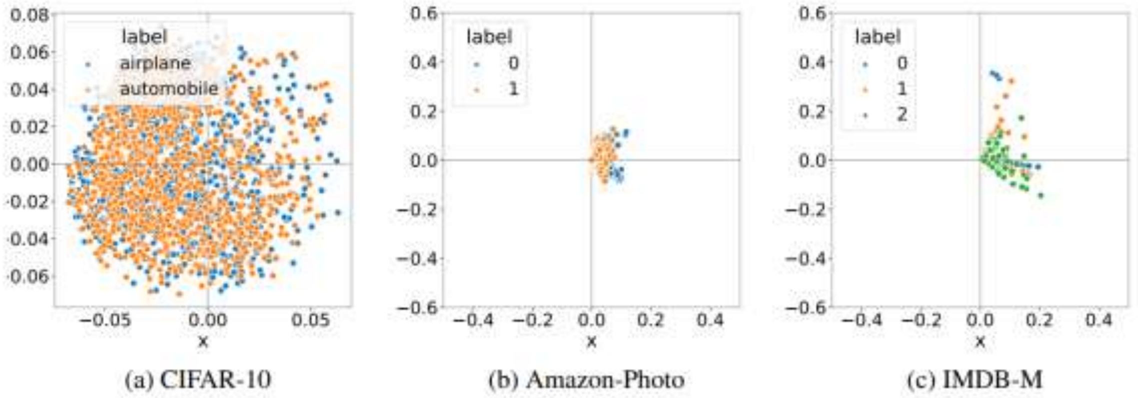
J Liu, Q Mao, J Li, **X Fu**, et al. POINE^2 : Improving Poincaré Embeddings for Hierarchy-Aware Complex Query Reasoning over Knowledge Graphs[C], ECAI 2023.

Orientation and anisotropic of graph

Graph structure is non-Euclidean (irregular), it exhibits significant directional bias and anisotropy both in the discrete topology and the embedding space. Therefore, the geometric property of the graph orientation is important to capture the ground-truth structure.



Anisotorpic message propagation on graph topology[1]



Anisotorpic distribution of graphs on latent space[2]

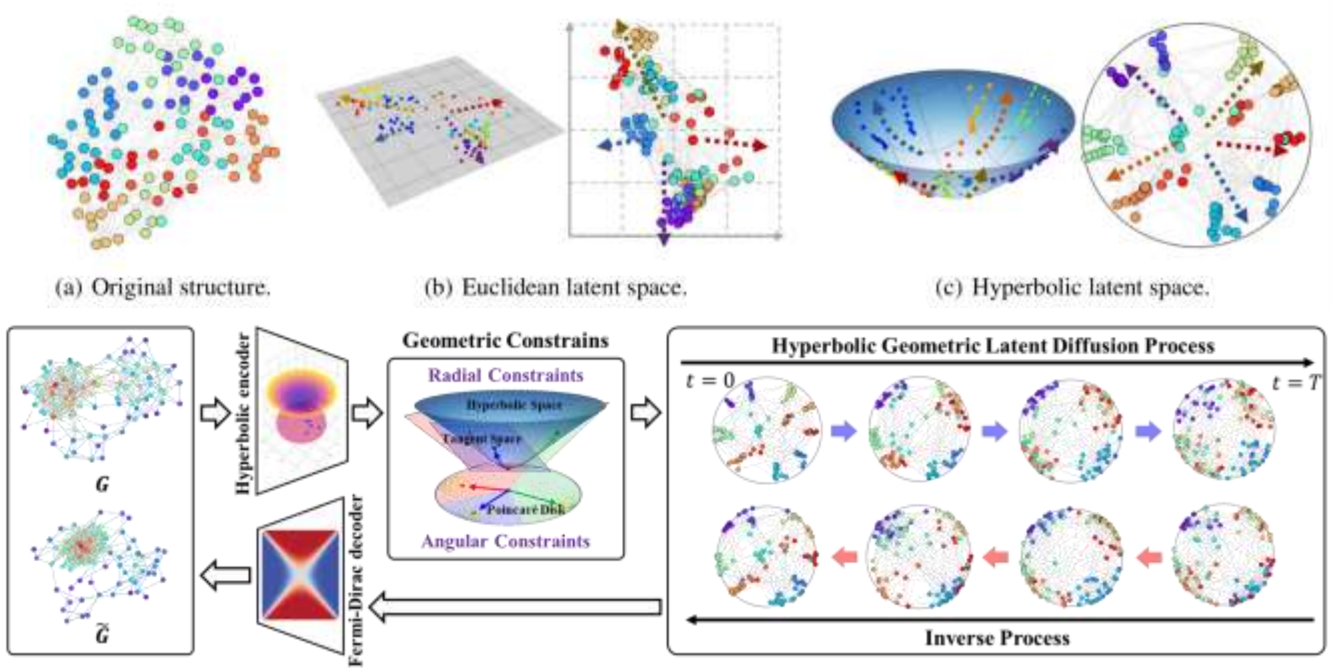
[1] Elhag A A A, Corso G, Stärk H, et al. Graph anisotropic diffusion for molecules[C]. ICLR 2022.

[2] Yang R, Yang Y, Zhou F, et al. Directional diffusion models for graph representation learning[C]. NeurIPS 2023.

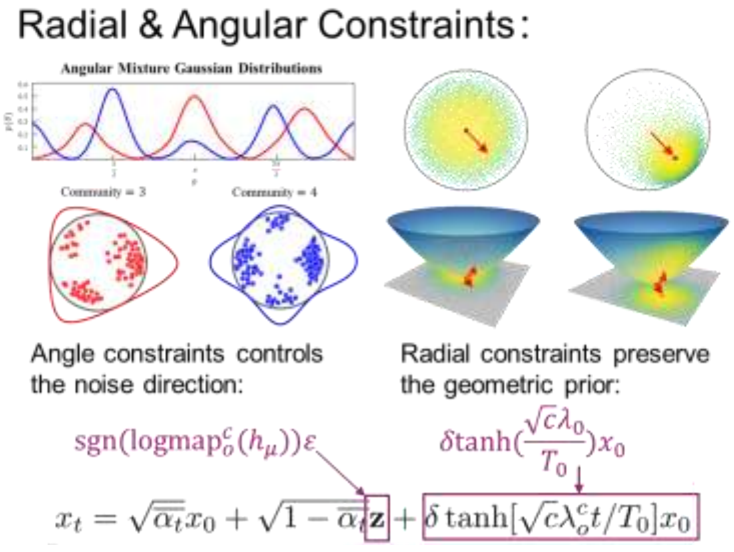
Hyperbolic graph diffusion

Lower-distortion and High-Efficiency Topology Generation

Hyperbolic geometry provides better priors for topological properties of graphs with non-Euclidean structure. **Anisotropic diffusion** provides better and more fine-grained structure details for graph structure generation. Our model has low computational complexity and GPU occupancy.



HyperDiff Architecture

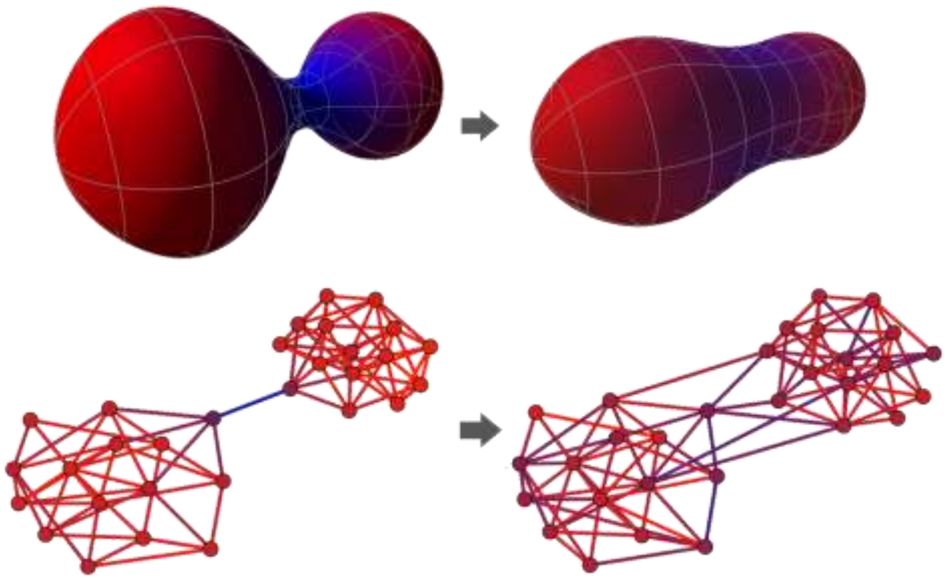


Paper: https://arxiv.org/pdf/2405.03188
Code: https://github.com/RingBDStack/HyperDiff

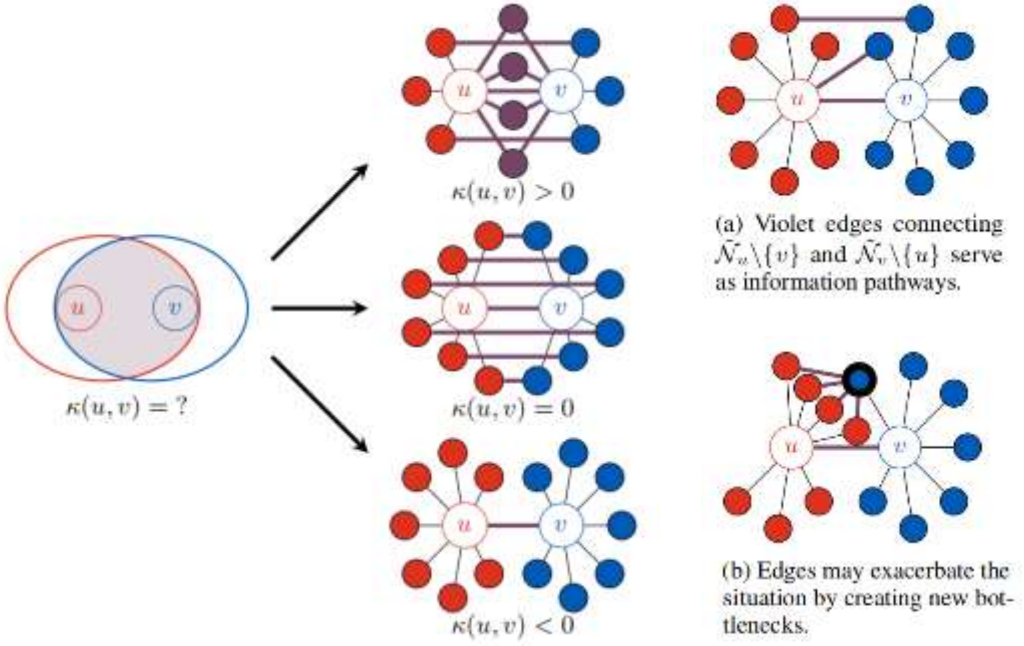
Xingcheng Fu, Yisen Gao, et al. Hyperbolic Geometric Latent Diffusion Model for Graph Generation [C], ICML 2024.

Geometric intuition of GNNs

Graph geometry can provide a powerful geometrically intuitive mathematical tools for underlying mechanism understanding of GNNs.



Over-squashing problem of MPNNs[1]



Understanding of Over-smoothing and Over-squashing[2]

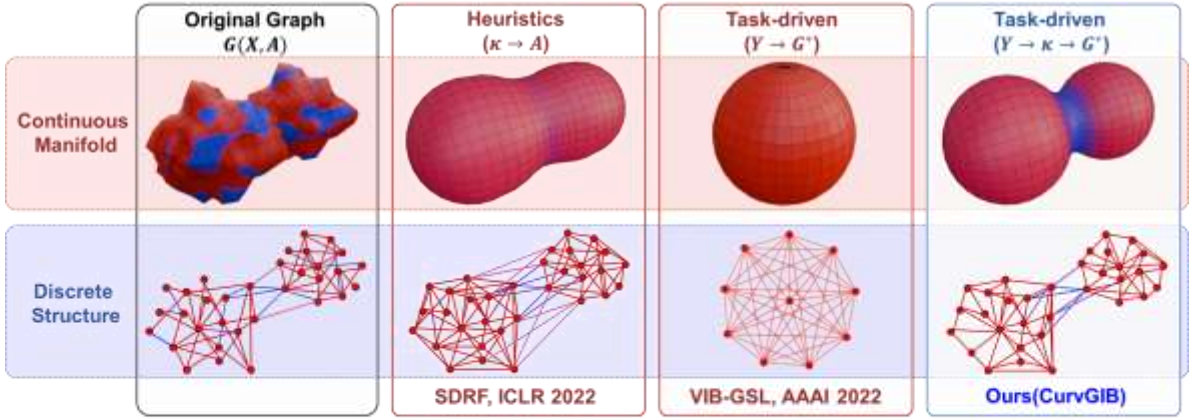
[1] Topping J, Di Giovanni F, Chamberlain B P, et al. Understanding over-squashing and bottlenecks on graphs via curvature[C]. ICLR 2022.

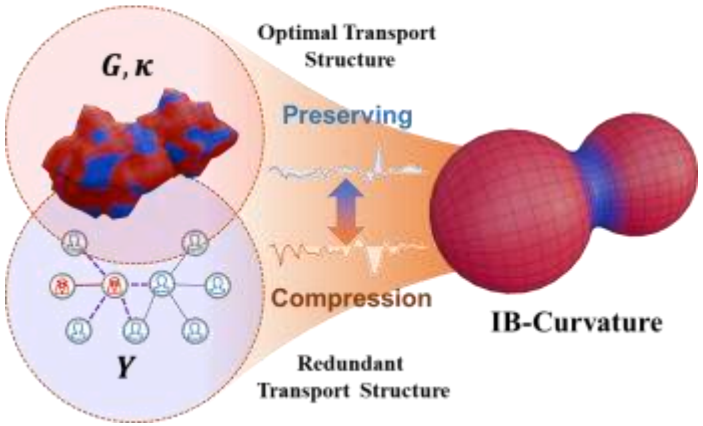
[2] Nguyen K, Hieu N M, Nguyen V D, et al. Revisiting over-smoothing and over-squashing using ollivier-ricci curvature[C]. ICML, 2023.

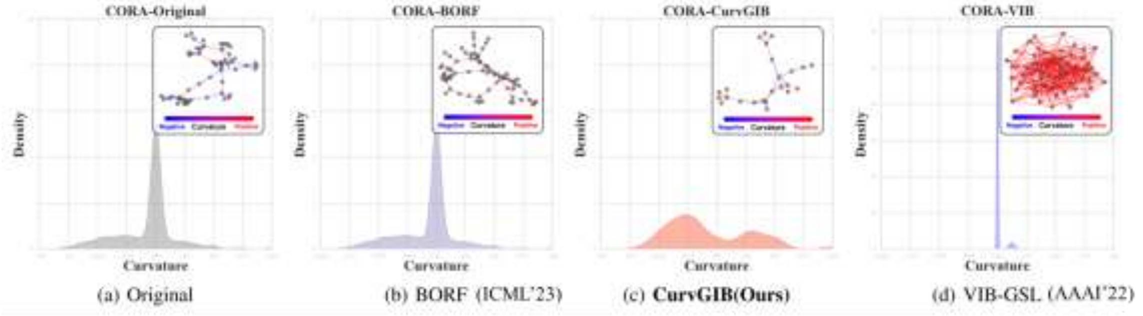
Curvature information bottleneck

Comparison between Information Theory and Graph Geometry

IB-based methods only focuses on finding task-relevant information, and lacks a unified perspective to comprehensively understand the *underlying task-relevant optimal transport structures* in GNNs.







Definition (Curvature Graph Information Bottleneck)

$$\mathbf{Z}_{\mathrm{IB}}, \kappa_{\mathrm{IB}} = \arg\min_{\mathbf{Z},\kappa} \mathrm{CurvGIB}(\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \kappa)$$

$$\triangleq \arg\min_{\mathbf{Z},\kappa} \Big[\underbrace{-I(\mathbf{Z} \mid \kappa; \mathbf{Y})}_{\text{Preserving}} + \underbrace{\beta I(\mathbf{Z} \mid \kappa; \mathbf{X})}_{\text{Compression (Smoothing)}}\Big],$$

Paper: https://arxiv.org/pdf/2412.19993
Code: https://github.com/RingBDStack/CurvGIB

Xingcheng Fu, Jian Wang, et al. Discrete Curvature Graph Information Bottleneck[C], AAAI 2025.

Continuous structural entropy

Generalizing the classic theory to the continuous realm

It present the *Differentiable Structural Information* (DSI), generalizing the classic theory to the continuous realm. DSI emerges as a new graph clustering objective, not requiring the cluster number.

- **Q1:** We propose a new formulation of structural information (Li & Pan, 2016), providing a **continuous relaxation** to the differentiable realm.

$$\mathcal{H}^{T}(G;h) = -\frac{1}{V}\sum_{k=1}^{N_h}\Big(V_k^h - \sum_{(i,j)\in\mathcal{E}} S_{ik}^h S_{jk}^h w_{ij}\Big)\log_2\frac{V_k^h}{V_k^{h-1}}$$

Defined on S which is given by *the level-wise assignment*, a real number in [0,1]

Level-wise Assignment

The probability of the i-th node of T at the h-th level being the parent node of j-th node at (h-1)-th level.

Theorem: Equivalence

The proposed formulation with the level-wise assignment is equivalent to the formulation of Li & Pan, 2016.

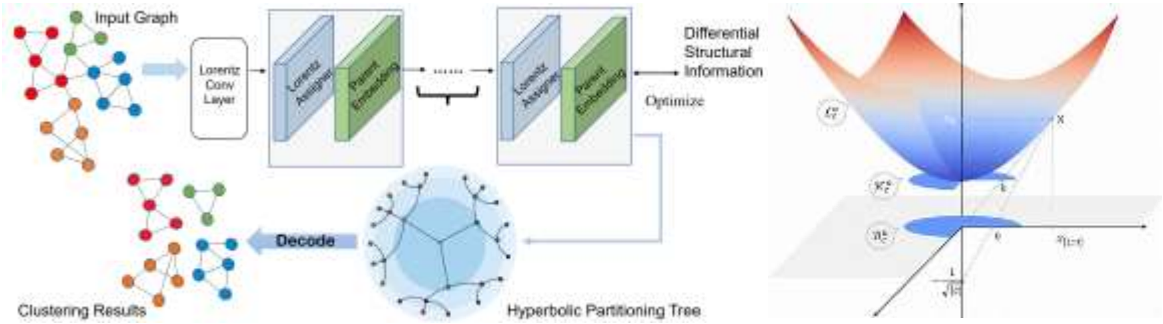
- **Q2:** We establish the connection between structural information and graph clustering theoretically.

Theorem: Connection

Given a graph G, the normalized H-structural entropy of G is proved to be **the upper bound** *of the graph conductance, with the property of Additivity.*

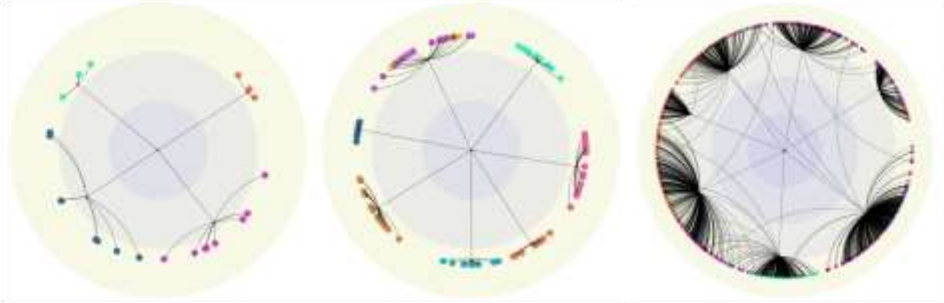
Additivity

The 1-dimensional case $\mathcal{H}^1(G) = \sum_{h=1}^{H}\sum_{j=1}^{N_{h-1}}\frac{V_j^{h-1}}{V}E\Big(\Big[\frac{C_{kj}^h V_k^h}{V_j^{h-1}}\Big]_{k=1,\dots,N_h}\Big)$



	Cora		Citeseer		Football		Karate	
	NMI	ARI	NMI	ARI	NMI	ARI	NMI	ARI
CSE	60.82±0.30	59.36±0.17	47.12±1.01	45.67±0.67	79.23±0.12	54.06±0.05	81.92±0.29	69.77±0.02
LSEnet	62.57±0.59	61.80±0.67	49.35±0.20	46.91±1.12	80.37±1.02	54.72±2.05	82.19±3.16	70.31±0.80
Performance Gain	**1.75**	**2.44**	**2.23**	**1.24**	**1.14**	**0.66**	**0.27**	**0.54**

Comparison between the classic structural entropy (CSE) and the proposed LSEnet.



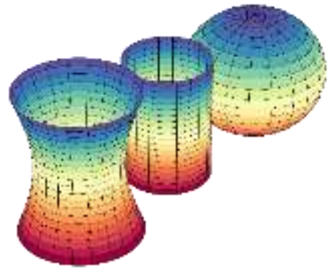
Visualization of Karate, Football and Cora in the corresponding Poincare discs.

Li Sun, Z Huang, H Peng, et al. LSEnet: Lorentz structural entropy neural network for deep graph clustering[C]. ICML, 2024, Oral Paper.

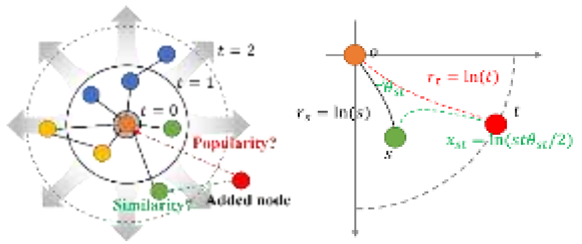
Recap

Geometry-guided Low-distortion GRL

Geometric Space



Geometric Properties



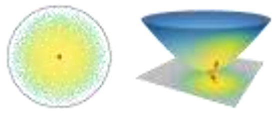
Curvature

- ACE-HGNN(ICDM 2022 Best Paper Candidate)
- CurvGAN(WWW 2022)

Geo-GFM

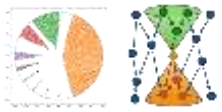
- GraphMoRE(AAAI 2025)
- RiemannGFM(WWW 2025)
- GalaxyWalker(CVPR 2025 Highlight Paper)

Radius



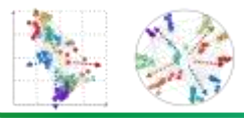
- HyperIMBA(WWW 2023 Splotlight Paper)
- PoinDP(AAAI 2024)

Angle & Cone



- ConeE(NeurIPS 2021)
- POINE^2(ECAI 2024)

Orientation



- HyerDiff(ICML 2024)

Continualization

- CurvGIB(AAAI 2025)
- LSEnet(ICML 2024 Oral Paper)