Hyperbolic Associative Memory Networks

ICML 2026

Boliang Hao, Bailing Zhang, Fangyu Wu

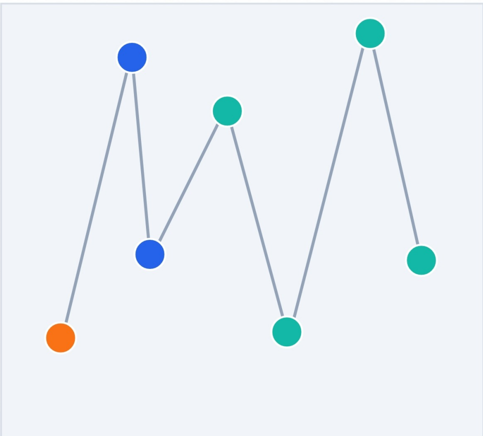
A geometry-aware Hopfield memory layer for hierarchical retrieval

Poincaré ball

Motivation: flat memories distort hierarchy

Modern Hopfield retrieval is attention-like, but its geometry is usually Euclidean.

Euclidean MHN



Move memory
retrieval to curved space

Hyperbolic space

Poincaré ball

Hierarchy-sensitivity hypothesis

Deeper hierarchy → larger benefit from hyperbolic retrieval;
weak/shallow hierarchy → near parity.

HAMNs: Hopfield energy on a hyperbolic manifold

We move both the similarity and the retrieval dynamics from flat space to negative curvature.

Euclidean features → Exp map to ball → Hyperbolic similarity → **Hopfield energy** → CCCP/MM retrieval → Task head

Energy

$$E(\xi) = -\frac{1}{\theta}\log\sum_i \exp(\theta\langle x_i, \xi\rangle_M) + \frac{1}{2}d_{\mathcal{M}}(\xi, p)^2$$

$$\langle x, y\rangle_M = -\cosh(d_{\mathcal{M}}(x, y))$$

Retrieval update

$$\xi^{t+1} = \mathrm{Exp}_p\left(-\eta \cdot \mathrm{PT}_{\xi^t \to p}(\mathrm{grad}\, E_{\mathrm{cave}}(\xi^t))\right)$$

Softmax weights choose memories;
Riemannian steps keep states on the manifold.

Not just a new similarity

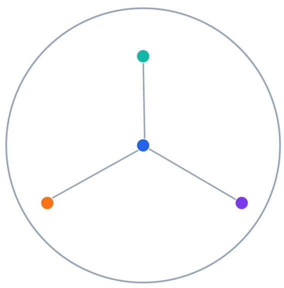
HAMNs define a global energy landscape and iterative attractor-style retrieval in hyperbolic space.

Theory: why hyperbolic memory helps hierarchy

Negative curvature supports more well-separated recall wells

Euclidean space

- Polynomial volume growth
- Packing bound grows only polynomially
- Limited separation for deep hierarchies

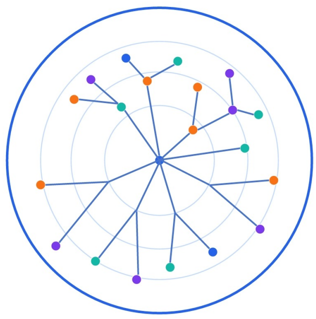


flat geometry

More room for separated recall wells

Hyperbolic space

- Exponential volume growth
- Recall-well count can grow exponentially
- Better separation for hierarchical memories



negative curvature

Takeaway

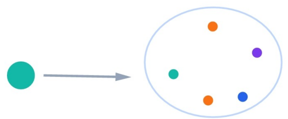
Hyperbolic geometry gives HAMNs a structural advantage when hierarchy becomes deep.

Three plug-and-play HAMN modules

Same retrieval core, different roles

HypHopfield

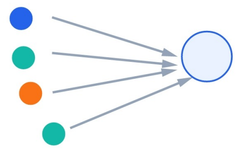
Query–memory association



- Input query + memory set
- Outputs retrieved hyperbolic state
- General associative retrieval

HypPooling

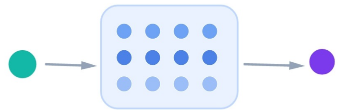
Set / bag aggregation



- Learnable queries retrieve from instance memories
- Produces fixed-size summary
- Useful for multi-instance learning

HypLayer

Learnable memory layer



- Retrieval through a learnable memory matrix
- End-to-end trainable
- Useful for classification, matching, aggregation

Shared retrieval core: exp map → hyperbolic similarity → Hopfield energy → CCCP/MM update

Experiments: hierarchy-sensitivity is validated

Best gains on deep hierarchies; near parity on weak/shallow tasks

CIFAR-100 hierarchical classification (4-layer)

Model	Top	Super	Coarse	Coph corr
MHNs	89.39	76.97	65.56	0.5902
U-Hop	88.62	75.46	62.46	0.7154
HypAttn	90.13	78.23	67.74	0.6795
HAMNs	**90.98**	**79.48**	**68.51**	**0.7184**

Best top / super / coarse accuracy and best cross-level consistency on the deepest label tree.

WordNet hypernym prediction (+OntEuc)

HAMNs: 96.97% accuracy

Avg. hierarchical distance: 0.037

Best among compared decoders

Weak / shallow hierarchy tasks

- MIL and MoleculeNet: competitive / near parity
- Consistent with the hierarchy-sensitivity hypothesis

Deep hierarchy → clear benefit; weak hierarchy → comparable performance

Conclusion: when hierarchy matters, geometry matters

HAMNs make associative memory hierarchy-aware without replacing the whole model.

Take-home message

Hyperbolic retrieval turns modern Hopfield networks into a practical memory module for hierarchical representation learning.

1. **Problem**
   Euclidean MHNs are powerful but flat.
2. **Method**
   Define Hopfield energy + CCCP retrieval on hyperbolic manifolds.
3. **Evidence**
   Gains increase on deeper CIFAR hierarchy and WordNet taxonomies.
4. **Boundary**
   Weak/shallow hierarchy: near parity; current hyperbolic ops add runtime overhead.

Thank you

Code: github.com/hbl66/HAMNs