Towards Graph Embedding in Symmetric Spaces

Seminar Presentation

Clemens Fruböse

June 9, 2020

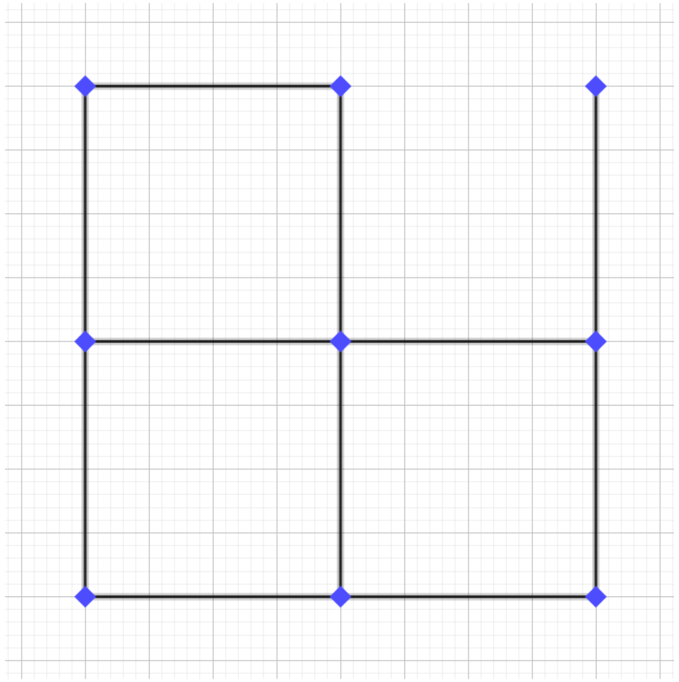
Motivational Example

$$Adj(\mathcal{G}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

- Adjacency matrix is hard to grasp
- link- & node prediction is a combinatorial task

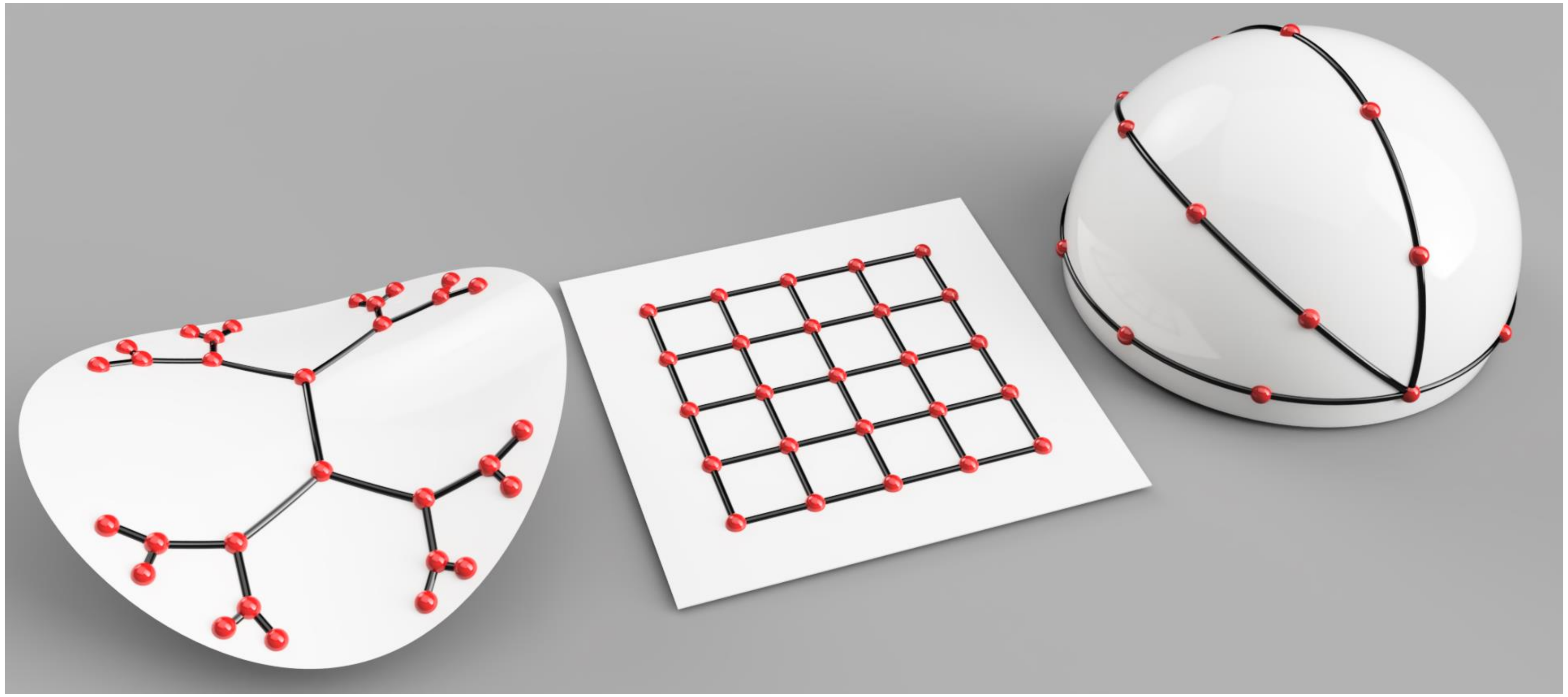
Motivational Example

$$Adj(\mathcal{G}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$



- Assign coordinates to the vertices (embed the graph)
- Predict links according to spatial layout of data

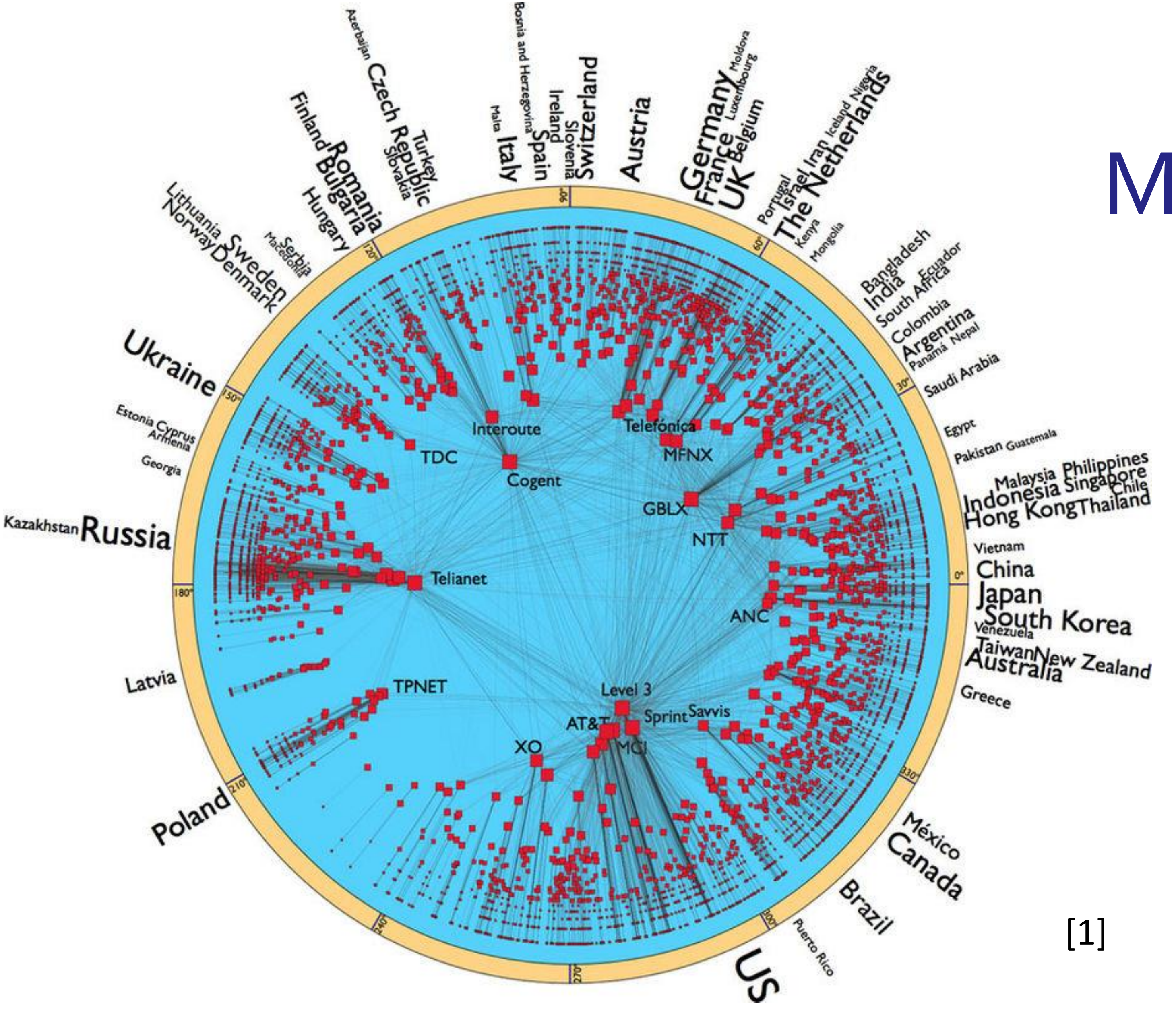
Why choose curved spaces?



- Graphs have *natural* curvature
- Choosing the right embedding space results in more successful embedding

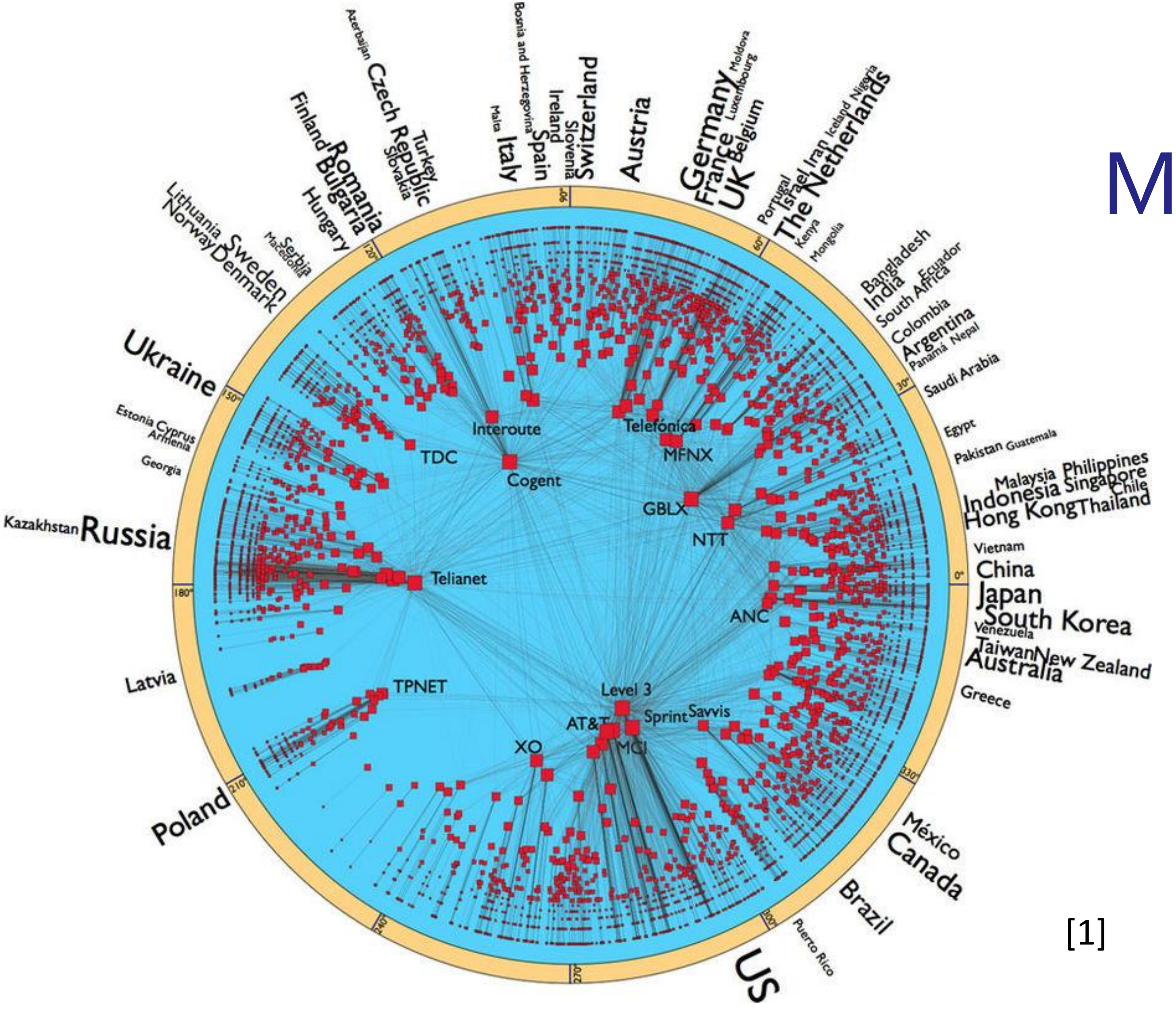
Content

1) Application of meaningful embeddings



[1]

Map of the Internet in hyperbolic space



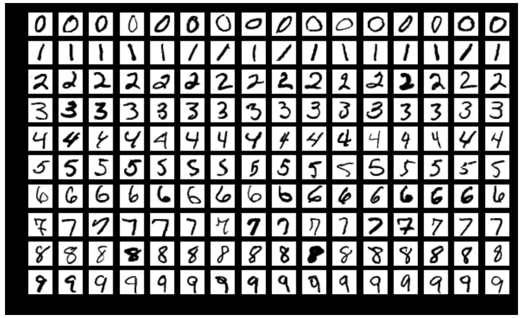
[1]

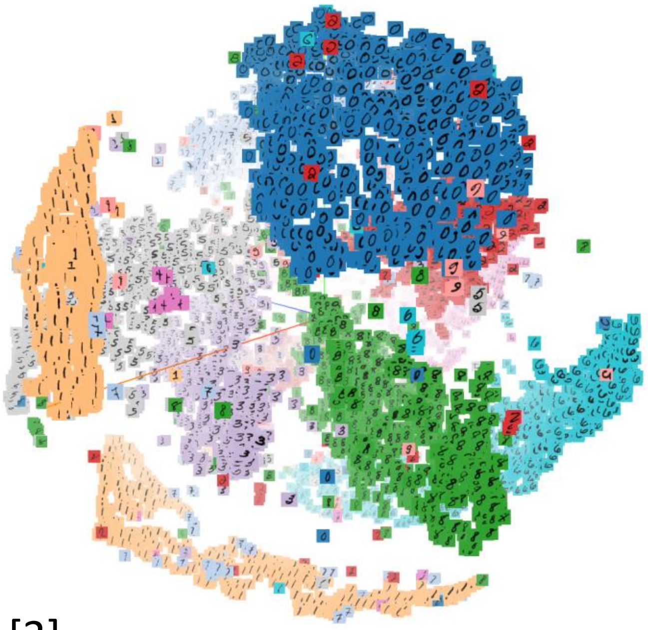
Map of the Internet in hyperbolic space

- Hierarchy becomes evident from embedding
- Greedy rooting is successful

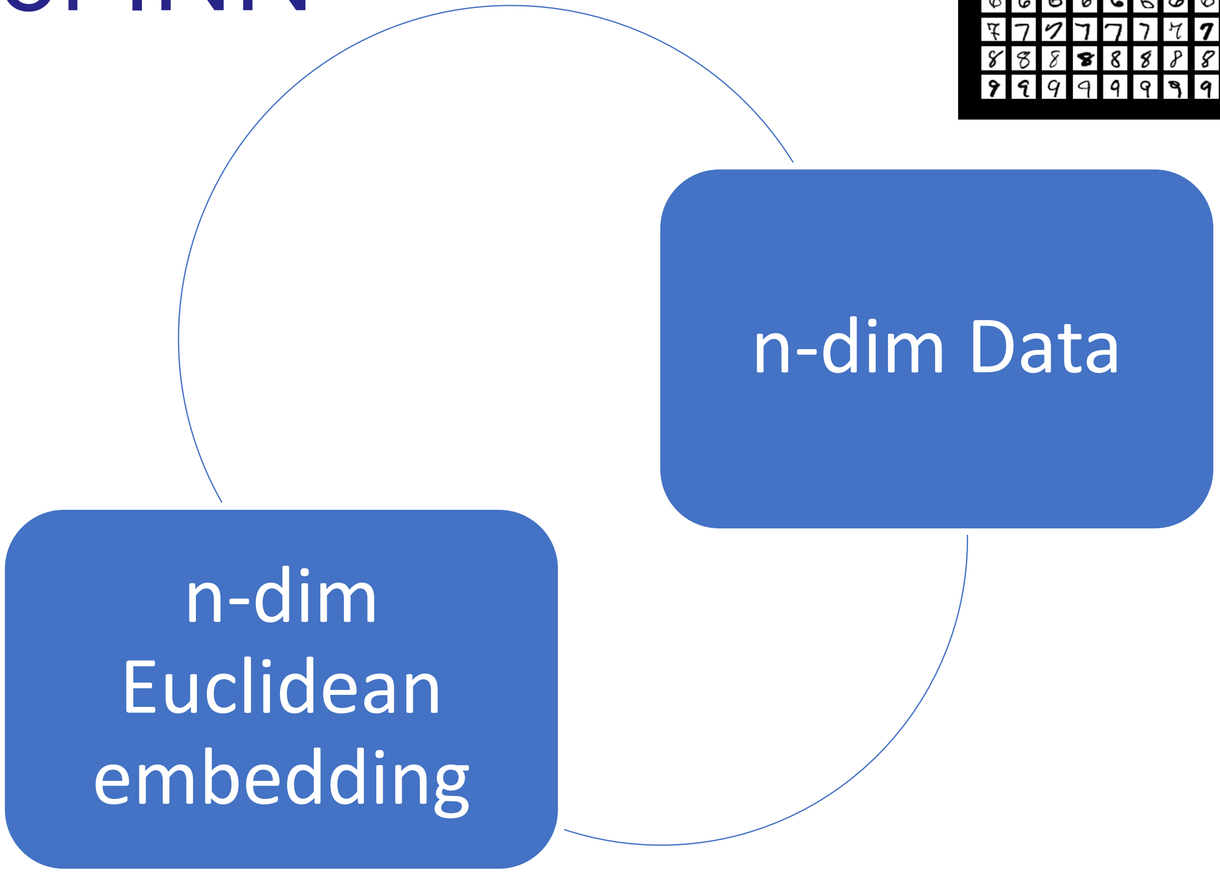
Motivation for INN

[2]



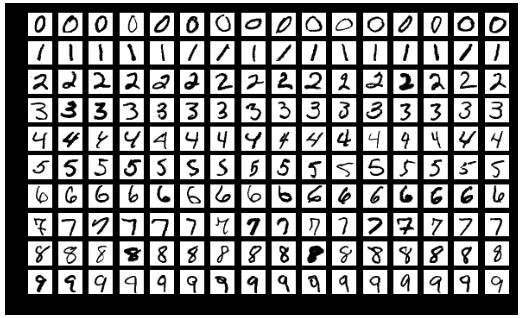


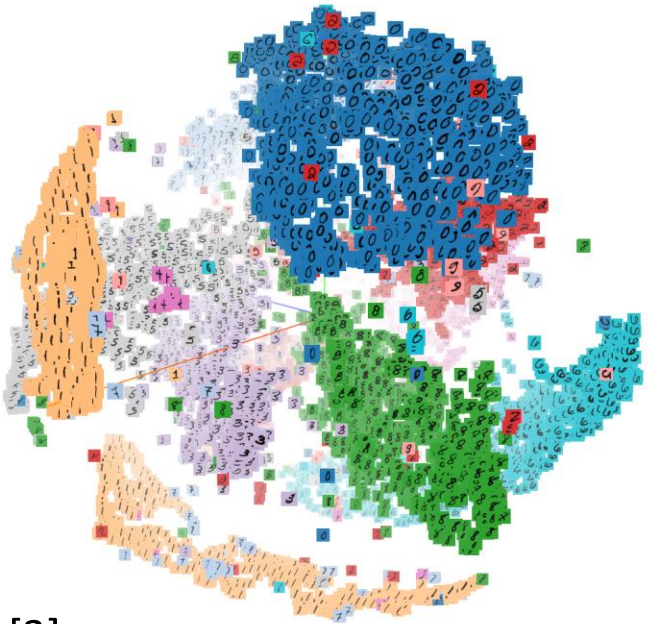
[3]



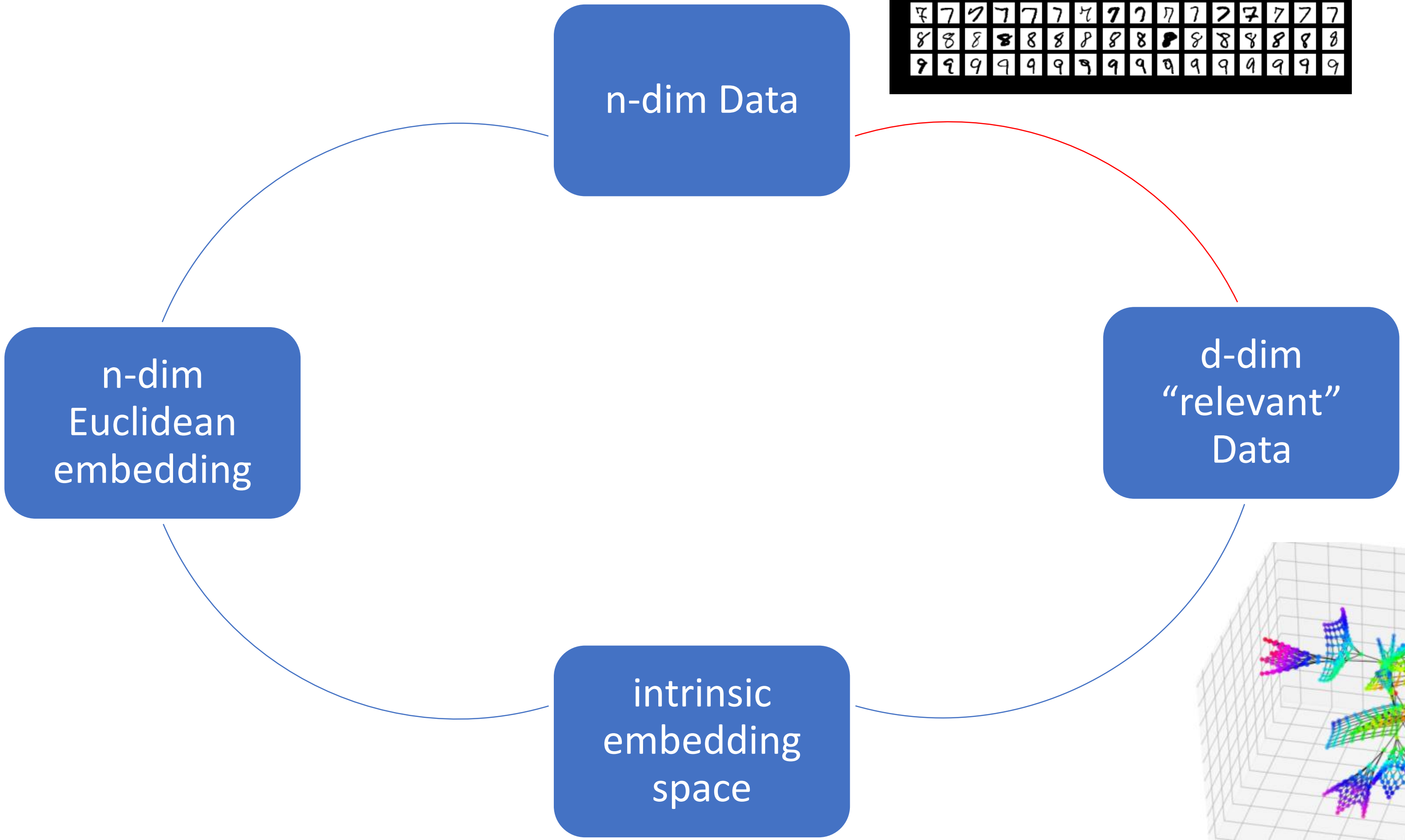
Motivation for INN

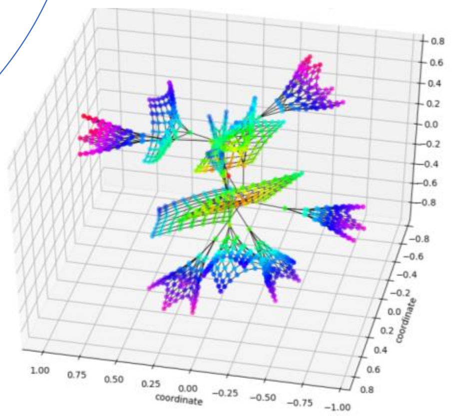
[2]





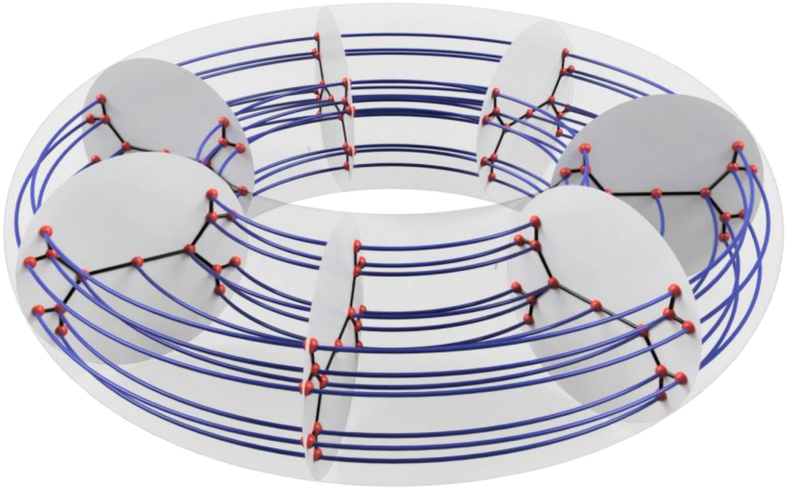
[3]





2) Mixed-curvature of graphs

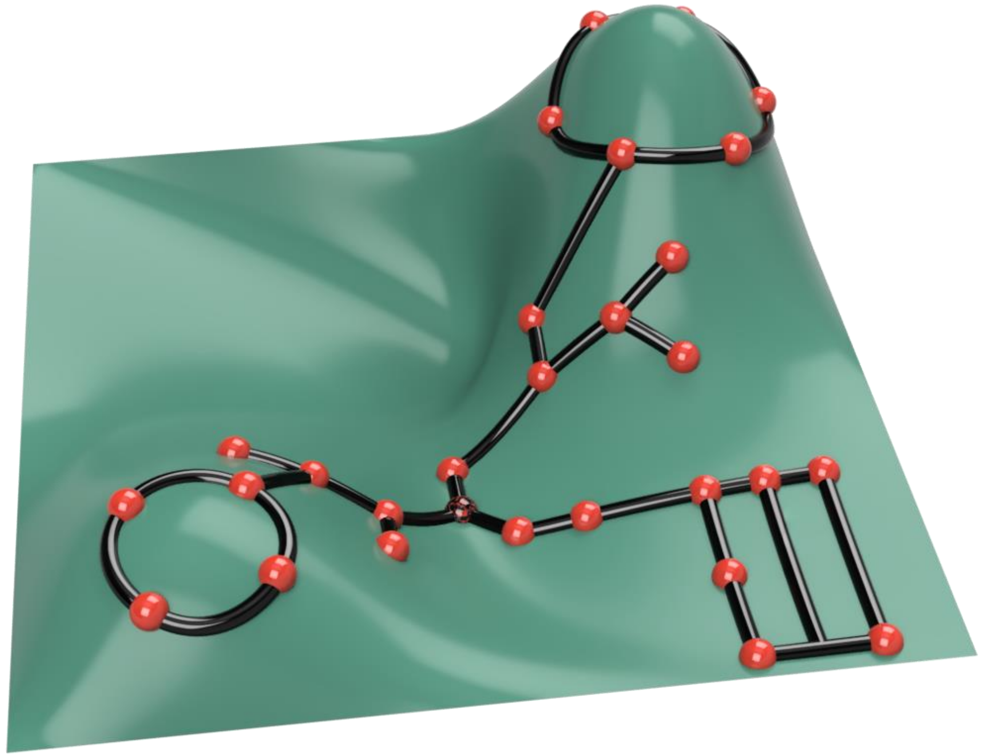
Interpreting Data



Graph with different structures (tree-like, cyclical)

- Real-world graphs do not fit perfectly to uniformly curved space
- Graphs express structures of different fashion

Embed on curved spaces



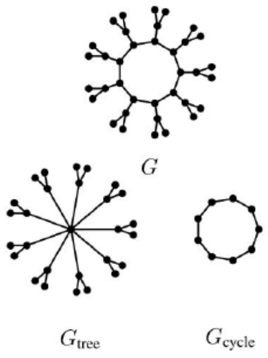
Graph embedded on surface of non-constant curvature

- Adapt the curvature of the embedding space to the graph

→ Interpretation and computations within embedding space become difficult

Computationally “easy” approach to mixed-curvature

Graph Embeddings (IV)



- [Gu et al, 2019]: Products of spaces of mixed-curvatures
  - Data may have varying structure, in some regions tree-like, in others cyclical
  - Combine hyperbolic, spherical and Euclidean spaces

$$\mathcal{P} = \mathbb{S}^{s_1} \times \mathbb{S}^{s_2} \times \cdots \times \mathbb{S}^{s_m} \times \mathbb{H}^{h_1} \times \mathbb{H}^{h_2} \times \cdots \times \mathbb{H}^{h_n} \times \mathbb{E}^{e}$$

$$X = [\underbrace{s_{11}, s_{12}, \ldots, s_{1n}}_{\text{Spheric space}}, s_{21}, \ldots, s_{2k}, \ldots, \underbrace{s_{m1}, \ldots, s_{mm}}_{\text{Spheric space}}, \underbrace{h_{11}, h_{12}, \ldots, h_{1n}}_{\text{Hyperbolic space}}, h_{21}, \ldots, h_{2k}, \ldots, \underbrace{h_{n1}, \ldots, h_{nn}}_{\text{Hyperbolic space}}, \underbrace{e_1, e_2, \ldots, e_e}_{\text{Euclidean space}}]$$

- Exponential maps, distances, gradient calculations and gradient projections are decomposed per space:

$$\mathrm{Exp}_p(v) = (\mathrm{Exp}_{p_1}(v_1), \ldots, \mathrm{Exp}_{p_k}(v_k)), \qquad d^2_{\mathcal{P}}(x, y) = \sum_{i=1}^{k} d_i^2(x_i, y_i).$$

- Loss function:

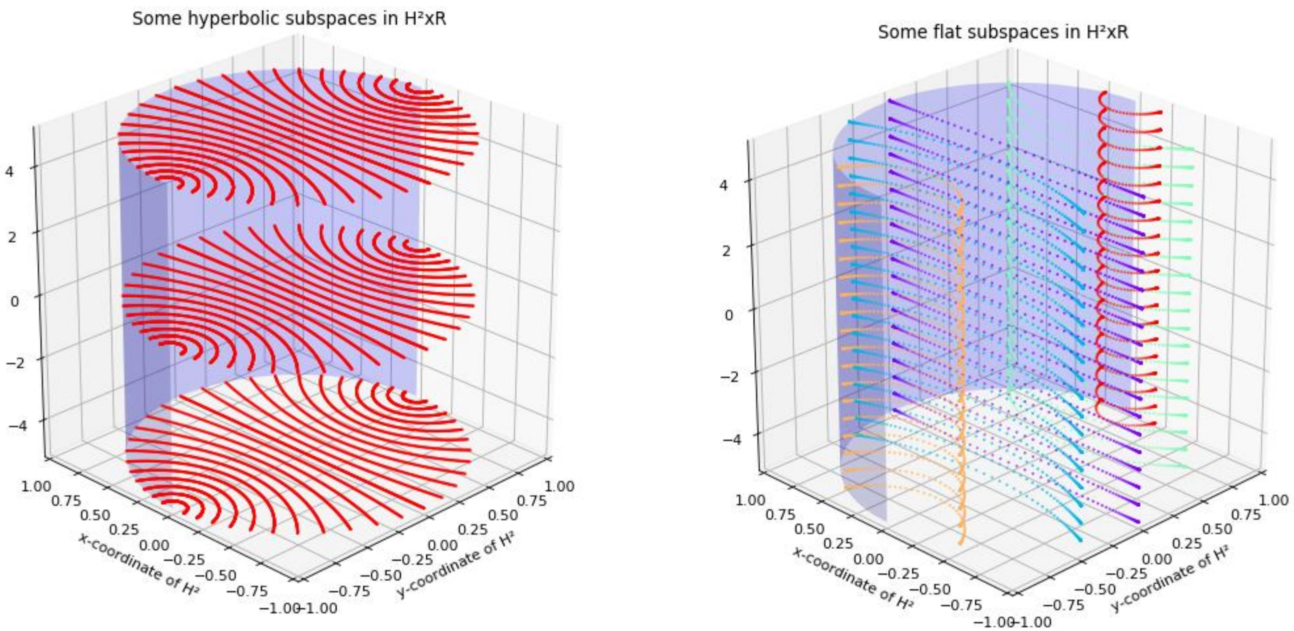
$$\mathcal{L}(x) = \sum_{1 \le i \le j \le n} \left| \left(\frac{d_{\mathcal{P}}(x_i, x_j)}{d_G(X_i, X_j)} \right)^2 - 1 \right|$$

13

[4]

3) Choosing symmetric spaces for graph embeddings

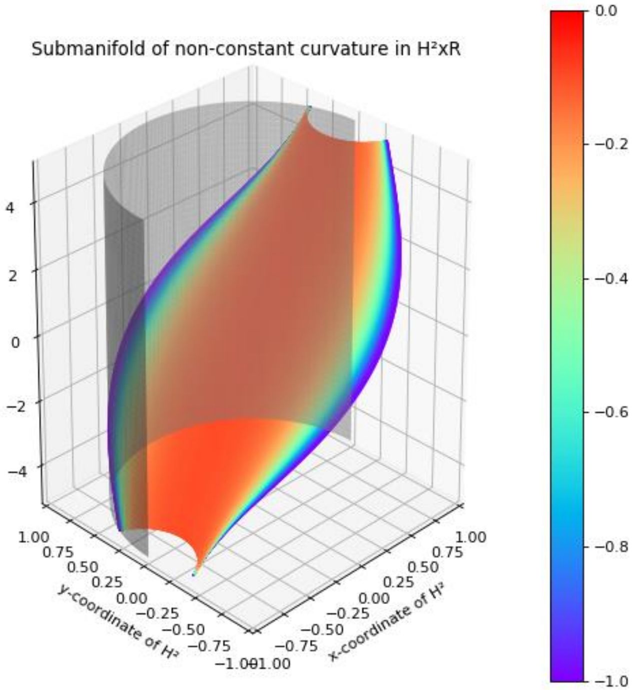
Subspaces of product spaces – Example: $H^2 \times R$

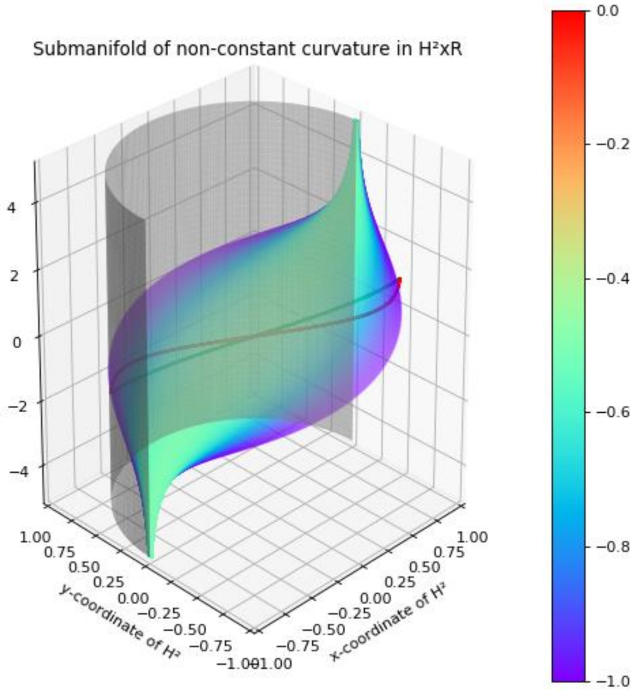


Constant curvature subspaces of $H^2 \times R$

- $H^2 \times R$ has 2-dim. flat and hyperbolic subspaces

Subspaces of mixed-curvature





- There are subspaces of non-constant curvature
- The shortest path does not lie on that subspace

Non-constant curvature subspaces of $H^2 \times R$

Totally geodesic submanifolds

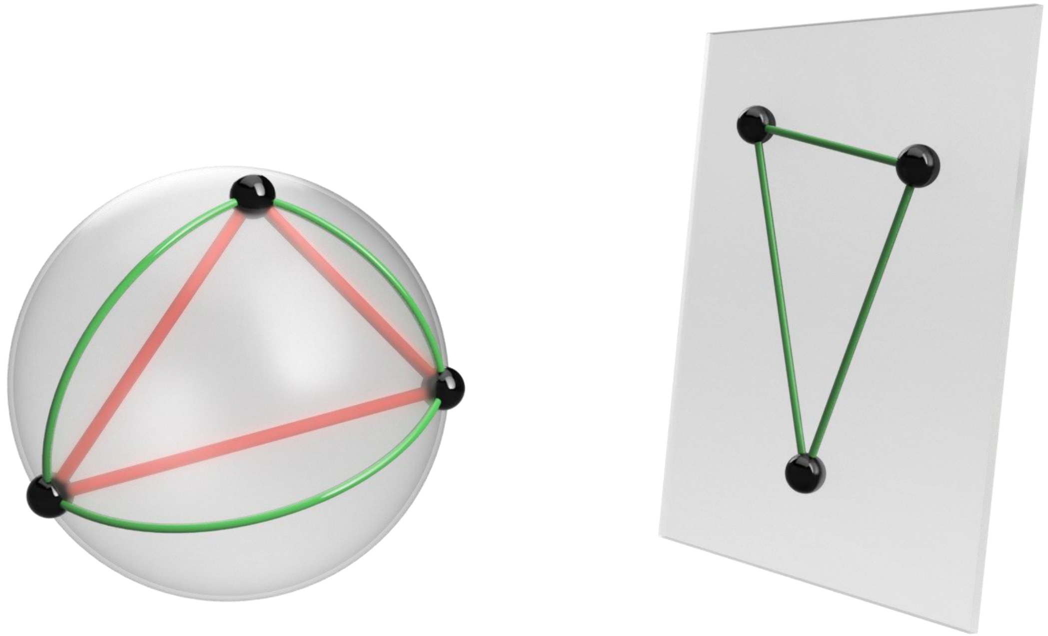


Illustration of the concept of totally geodesic submanifolds

- A structure corresponding to a particular curvature has to be contained as totally geodesic submanifold

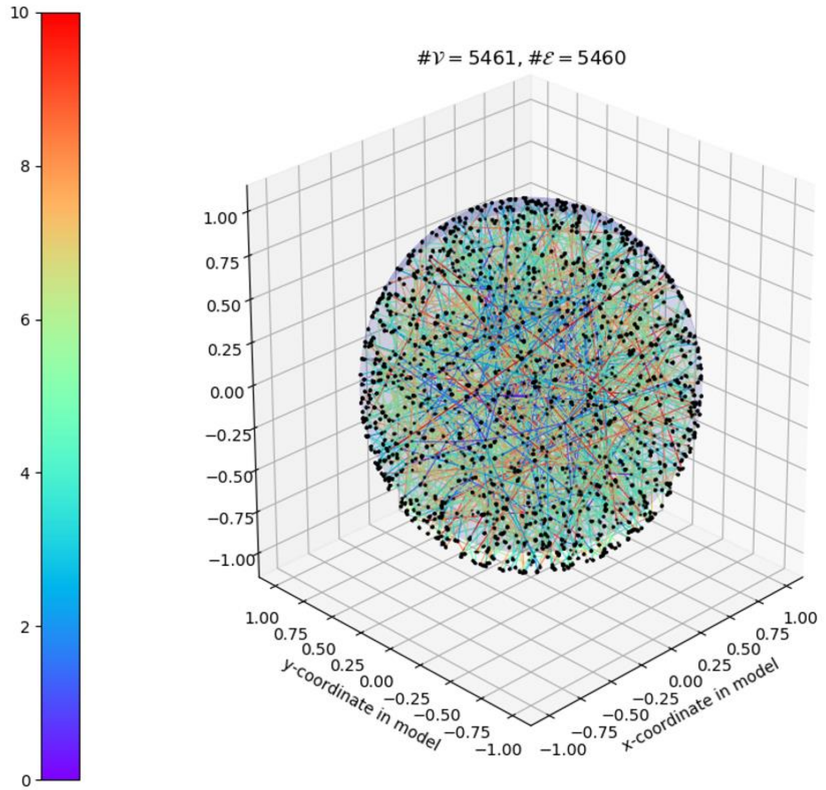
→ Embedding space has to have correct totally geodesic subspaces

What are these symmetric spaces?

- Posses a high degree of symmetry, i.e. fit to computational requirements
- Have subspaces of constant curvature as tot. geodesic subspaces
- Example:
  - Matrix version of hyperbolic space: Siegel upper half space Sp(2n,R) / SpO(2n,R)

4) Experiments in spaces of constant curvature

Experiments in simple setting

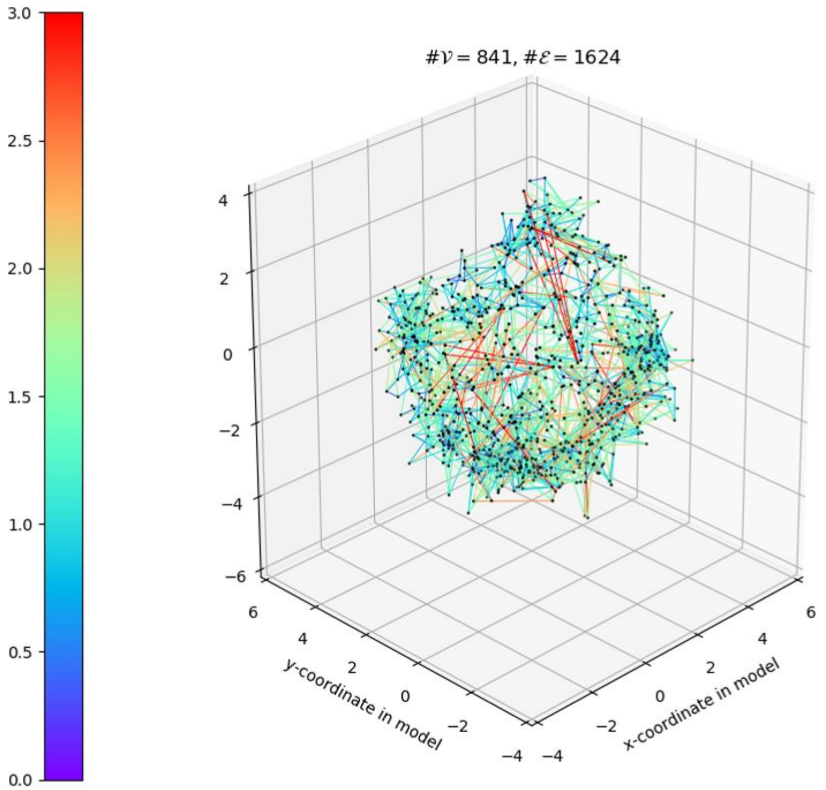


Perfect tree embedded in 3-dim hyperbolic space
with distortion indicated by color

- mAP is acceptable
- Intrinsic geometry not detectable

Code from [5] modified

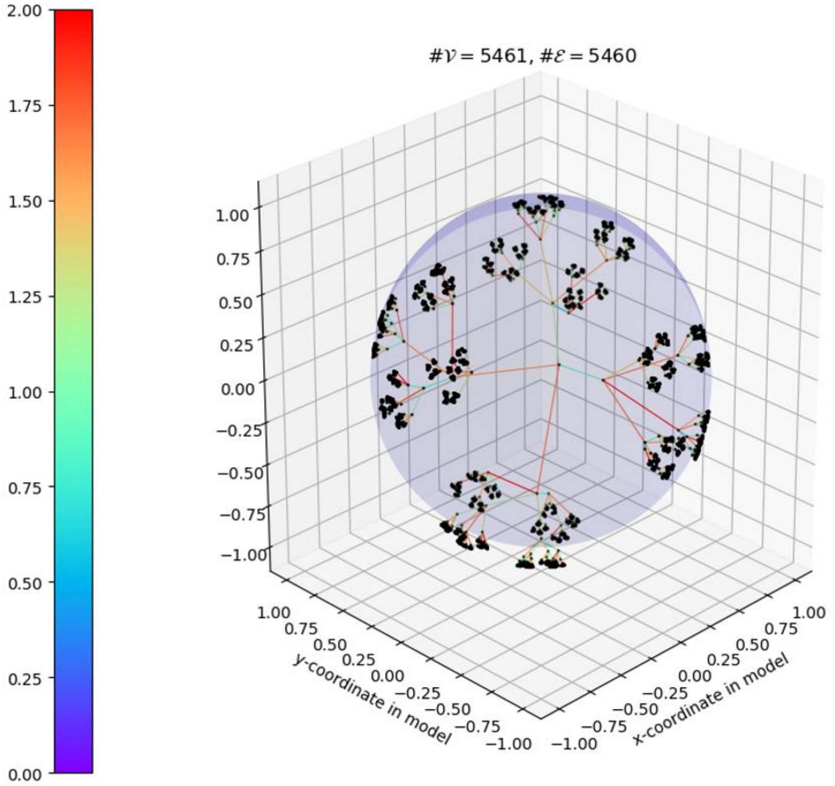
Experiments in simple setting



Perfect lattice embedded in 3-dim Euclidean space with distortion indicated by color

- mAP is bad
- Intrinsic geometry not detectable

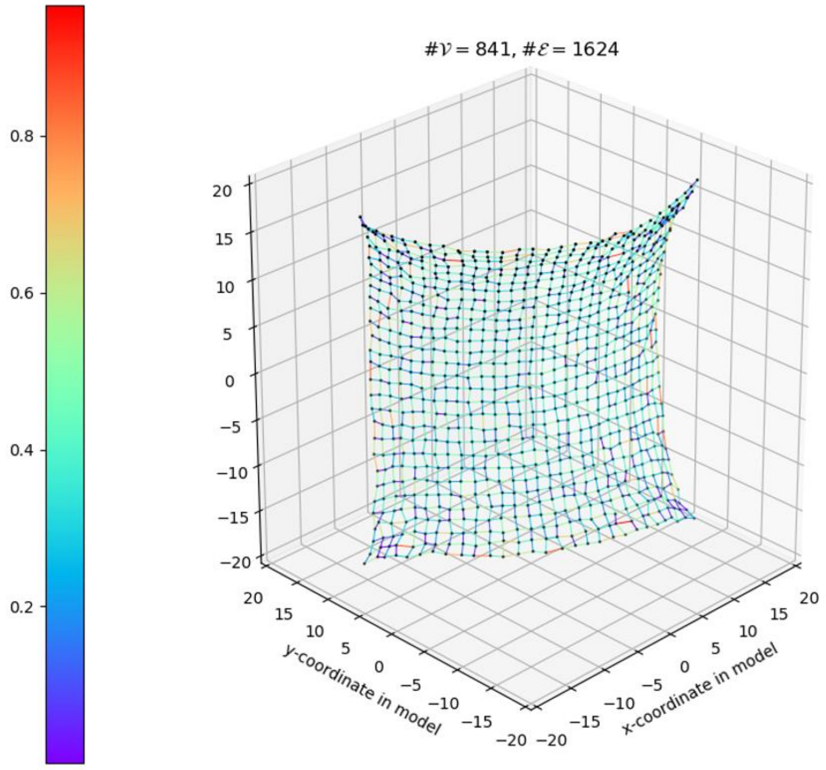
Experiments in simple setting



Perfect tree embedded in 3-dim hyperbolic space
with distortion indicated by color

- mAP is good
- Intrinsic geometry clearly detectable

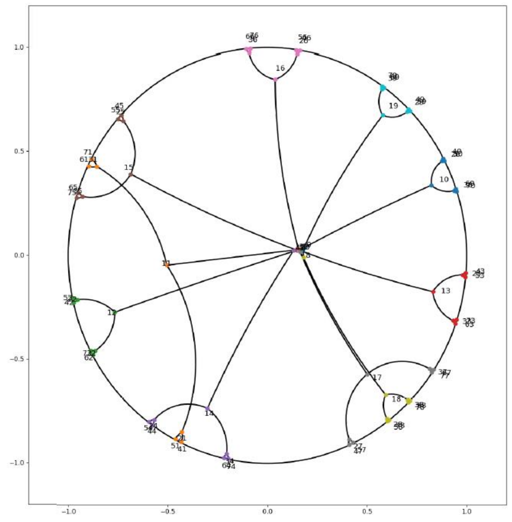
Experiments in simple setting

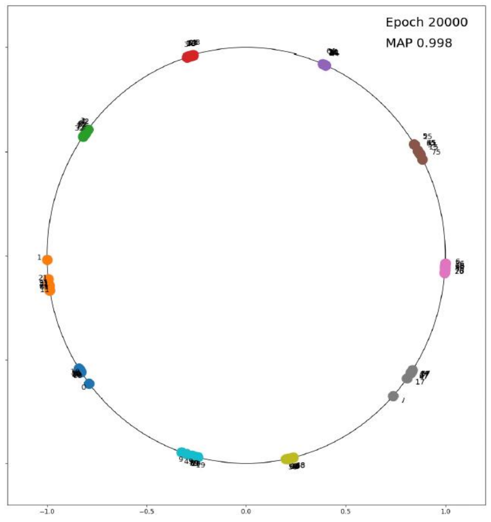


Perfect lattice embedded in 3-dim Euclidean space
with distortion indicated by color

- mAP is very good
- Intrinsic geometry clearly detectable

Best results of current code





[6]

- Very small graph #V=80

→ Do not expect to much from current code

Overview of results

Dataset	fidelity measure	expanding loss function	expanding loss function (FR)	distance-preserving loss function	distance-preserving loss function (FR)
Tree graph	D_{avg}	0.455	0.092	0.143	0.039
	RD_{avg}	1.450	0.960	0.992	0.995
	mAP	0.652	0.851	0.620	0.798
Lattice graph	D_{avg}	0.329	0.293	0.770	0.090
	RD_{avg}	0.754	0.757	0.234	1.034
	mAP	0.191	0.473	0.151	0.982

- Meaningful embedding highly dependent on starting point

→ Current code reaches local minima

5) Discussion / Outlook / Summary

Summary

- Graph embedding has numerous applications
- "meaningful" embeddings require suiting geometry (curvature)
- mixed-curvature properties can be represented in symmetric spaces
  (totally geodesic submanifolds)
- Problem: Embedding requires decent starting embedding
- Solution (?) :
  1. Adapting the optimization
  2. Suitable preprocessing

Images

[1] taken from Boguná, M., Papadopoulos, F. and Krioukov, D., 2010. Sustaining the internet with hyperbolic mapping. Nature communications, 1, p.62., available at https://www.nature.com/articles/ncomms1063

[2] taken from https://upload.wikimedia.org/wikipedia/commons/2/27/MnistExamples.png

[3] taken from https://projector.tensorflow.org/preview.png

[4] taken from presentation by Federico Lopez: "The shortest history of geometric deep learning", 26.05.2020

[5] code modified from https://papers.nips.cc/paper/7213-poincare-embeddings-for-learning-hierarchical-representations, 23.05.19

[6] taken from https://openreview.net/pdf?id=HJxeWnCcF7