Poincaré GloVe: Hyperbolic Word Embeddings

International Conference on Learning Representations 2019
ICLR 19 - (2019 NLP GROUP STUDY)

Data Mining & Information Systems Lab.
Department of Computer Science and Engineering,
College of Informatics, Korea University

이잉걸

Alexandru Țifrea

FOLLOW

ETH Zurich
Verified email at student.ethz.ch

Machine learning   Natural language understa…

TITLE	CITED BY	YEAR
Poincar\'e GloVe: Hyperbolic Word Embeddings A Tifrea, G Bécigneul, OE Ganea arXiv preprint arXiv:1810.06546	23	2018
Opinion Summarization for Hotel Reviews B Cristian, A Tifrea, MC Volmer, T Rebedea		

Articles 1–2   SHOW MORE

GET MY OWN PROFILE

Cited by

	All	Since 2014
Citations	23	23
h-index	1	1
i10-index	1	1

???

Gary Bécigneul

PHD Researcher at ETH, MIT, MPI
Verified email at mit.edu - Homepage
Geometric Machine Learning   Hyperbolic Deep Learning

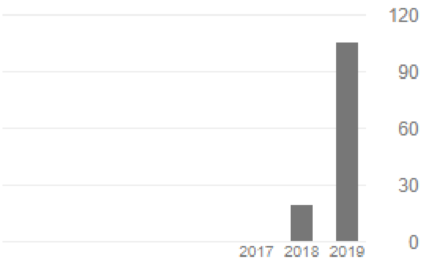
FOLLOW

GET MY OWN PROFILE

TITLE	CITED BY	YEAR
Mixed-curvature Variational Autoencoders O Skopek, OE Ganea, G Bécigneul Under Review at ICLR 2020		2019
Constant Curvature Graph Convolutional Networks G Bachmann*, G Bécigneul*, OE Ganea Under Review at ICLR 2020	1	2019
A Continuous-time Perspective for Modeling Acceleration in Riemannian Optimization F Alimisis, A Orvieto, G Bécigneul, A Lucchi Under Review at AISTATS 2020		2019
Breaking the Softmax Bottleneck via Learnable Monotonic Pointwise Non-linearities OE Ganea, S Gelly, G Bécigneul, A Severyn ICML 2019: International Conference on Machine Learning	1	2019
Poincaré GloVe: Hyperbolic Word Embeddings A Tifrea*, G Bécigneul*, OE Ganea* ICLR 2019: International Conference on Learning Representations	23	2018
Riemannian Adaptive Optimization Methods G Bécigneul, OE Ganea ICLR 2019: International Conference on Learning Representations	21	2018
Escaping Flat Areas via Function-Preserving Structural Network Modifications Y Kilcher*, G Bécigneul*, T Hofmann Submitted at ICLR 2019		2018
Hyperbolic Neural Networks OE Ganea*, G Bécigneul*, T Hofmann NIPS 2018: Neural Information Processing Systems	32	2018
Hyperbolic Entailment Cones for Learning Hierarchical Embeddings OE Ganea, G Bécigneul, T Hofmann ICML 2018: International Conference on Machine Learning	47	2018
Parametrizing filters of a CNN with a GAN Y Kilcher*, G Bécigneul*, T Hofmann Submitted at ICLR 2017	1	2017
On the effect of pooling on the geometry of representations G Bécigneul Submitted at COLT 2017	2	2017

Cited by

	All	Since 2014
Citations	128	128
h-index	4	4
i10-index	4	4



Co-authors

- Octavian Ganea — MIT
- Thomas Hofmann — Professor of Computer Science, ...
- Alexandru Ţifrea — ETH Zurich
- Yannic Kilcher — PhD Student, ETH Zurich
- Sylvain Gelly — Google Brain Zurich
- Aliaksei Severyn — Google
- Aurelien Lucchi — ETH Zurich
- Ondrej Skopek — Software Engineer, Google

Octavian Ganea

MIT
Verified email at csail.mit.edu - Homepage

Machine Learning   Representation Learning   Graph Representation Lear...   Artificial Intelligence

TITLE	CITED BY	YEAR
Mixed-curvature Variational Autoencoders O Skopek, OE Ganea, G Bécigneul ICLR'20: International Conference on Learning Representations		2020
Constant Curvature Graph Convolutional Networks G Bachmann, G Bécigneul, OE Ganea https://arxiv.org/pdf/1911.05076	1	2019
Breaking the Softmax Bottleneck via Learnable Monotonic Pointwise Non-linearities OE Ganea, S Gelly, G Bécigneul, A Severyn ICML'19: International Conference on Machine Learning	1	2019
Poincaré GloVe: Hyperbolic Word Embeddings A Tifrea*, G Bécigneul*, OE Ganea* ICLR'19: International Conference on Learning Representations	23	2019
Riemannian Adaptive Optimization Methods G Bécigneul, OE Ganea ICLR'19: International Conference on Learning Representations	21	2019
Hyperbolic Neural Networks OE Ganea, G Bécigneul, T Hofmann NeurIPS'18: Conference on Neural Information Processing Systems	32	2018
Hyperbolic Entailment Cones for Learning Hierarchical Embeddings OE Ganea, G Bécigneul, T Hofmann ICML'18: International Conference on Machine Learning	47	2018
Learning and Evaluating Sparse Interpretable Sentence Embeddings V Trifonov, OE Ganea, A Potapenko, T Hofmann EMNLP'18 Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks ...	2	2018
End-to-End Neural Entity Linking N Kolitsas*, OE Ganea*, T Hofmann CoNLL'18: Conference on Computational Natural Language Learning	21	2018
Web2Text: Deep Structured Boilerplate Removal T Vogels, OE Ganea, C Eickhoff ECIR'18: European Conference on Information Retrieval	7	2018
Deep Joint Entity Disambiguation with Local Neural Attention OE Ganea, T Hofmann EMNLP'17: Conference on Empirical Methods in Natural Language Processing	71	2017
Neural Multi-Step Reasoning for Question Answering on Semi-Structured Tables T Haug, OE Ganea, P Grnarova ECIR'18: European Conference on Information Retrieval	12	2017
Probabilistic Bag-of-Hyperlinks Model for Entity Linking OE Ganea, M Ganea, A Lucchi, C Eickhoff, T Hofmann	85	2016

Cited by

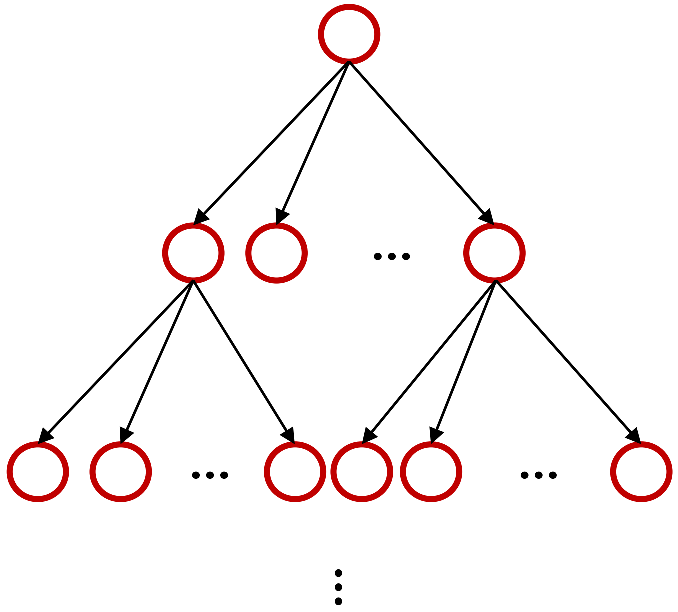
	All	Since 2014
Citations	323	323
h-index	8	8
i10-index	8	8

Co-authors

- Thomas Hofmann — Professor of Computer Science, ...
- Gary Bécigneul — PHD Researcher at ETH, MIT, MPI
- Carsten Eickhoff — Assistant Professor, Brown Univ...
- Aurelien Lucchi — ETH Zurich
- Alexandru Țifrea — ETH Zurich
- Paulina Grnarova — ETH Zurich
- Thijs Vogels — EPFL
- Anna Potapenko — National Research University Hig...
- Sylvain Gelly — Google Brain Zurich
- Aliaksei Severyn — Google
- Ondrej Skopek — Software Engineer, Google

– Recap

Characteristics(?) of tree structured graph



Suppose there are b branching factors on each nodes

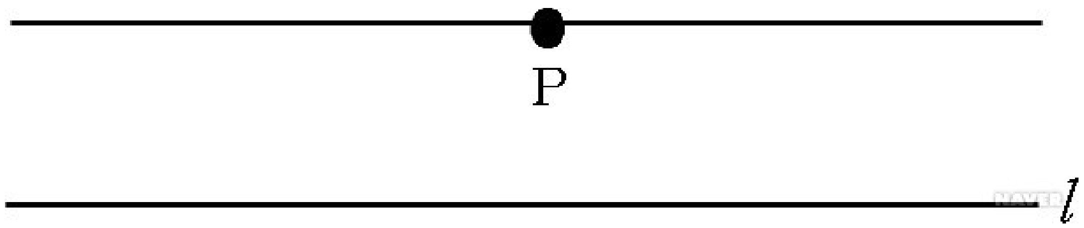
At level l, there are $(b+1)b^l$ nodes

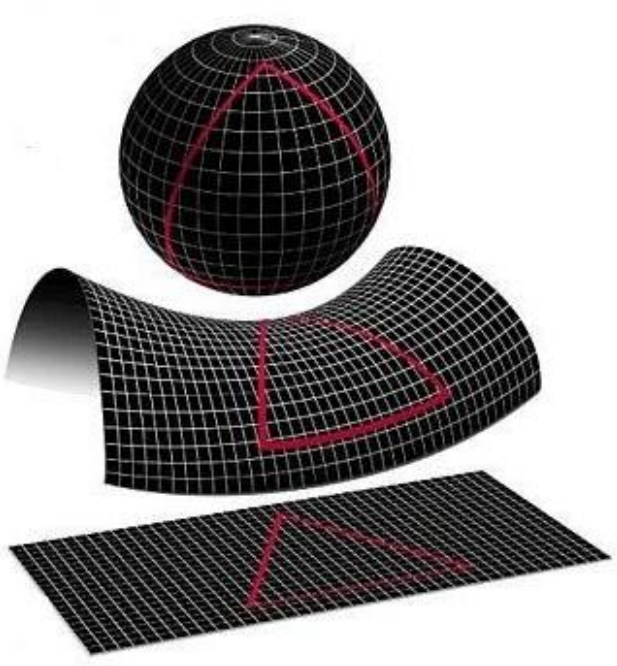
There are $\frac{\left((b+1)b^l-2\right)}{b-1}$ nodes on a level less or equal than l

Note : the # of nodes are exponentially increasing as level l (distance to the root of the tree) increases

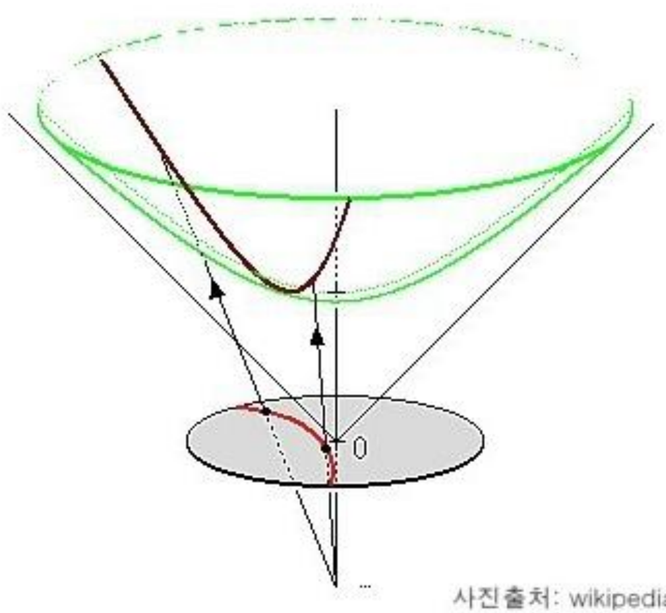
– Recap

> 선 밖의 한 점을 지나 그 직선에 평행한 직선은 단 하나만 존재한다.[1]

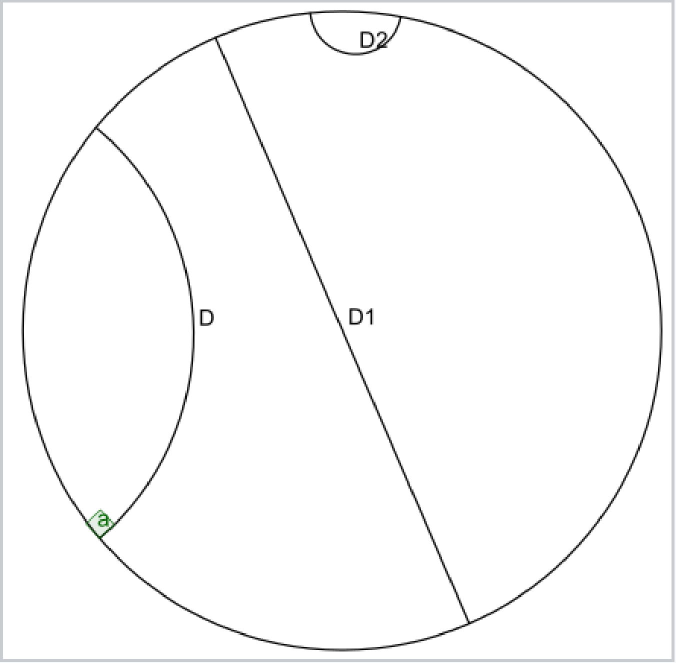




– Recap

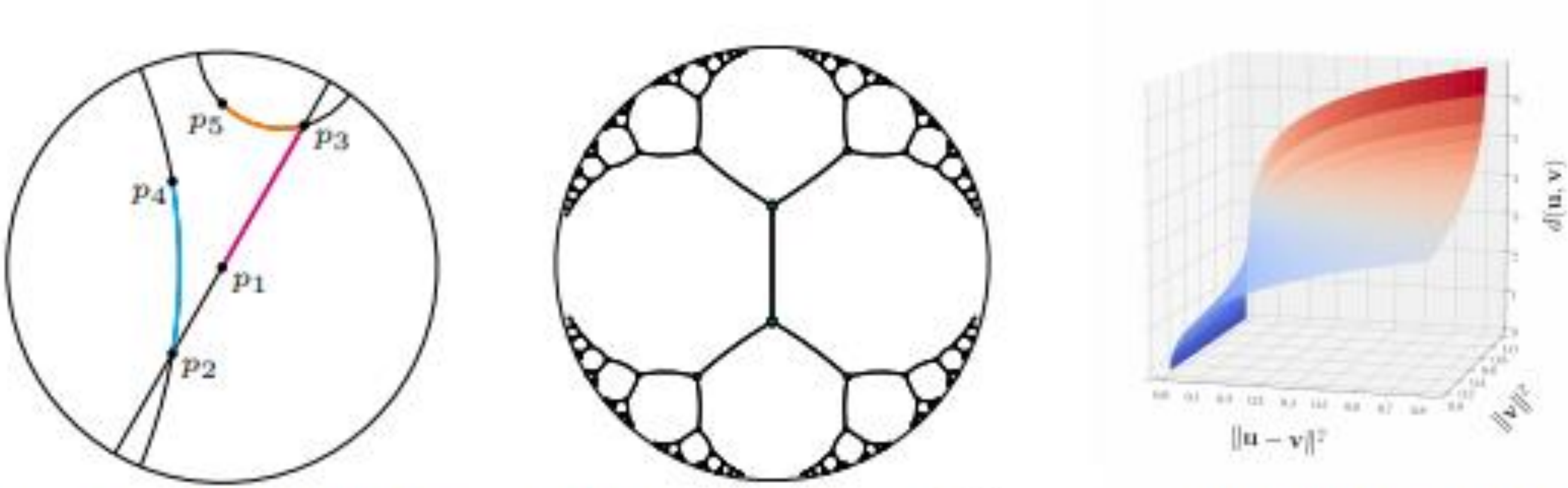


사진출처: wikipedia



Poincaré disk with 3 ultraparallel (hyperbolic) straight lines

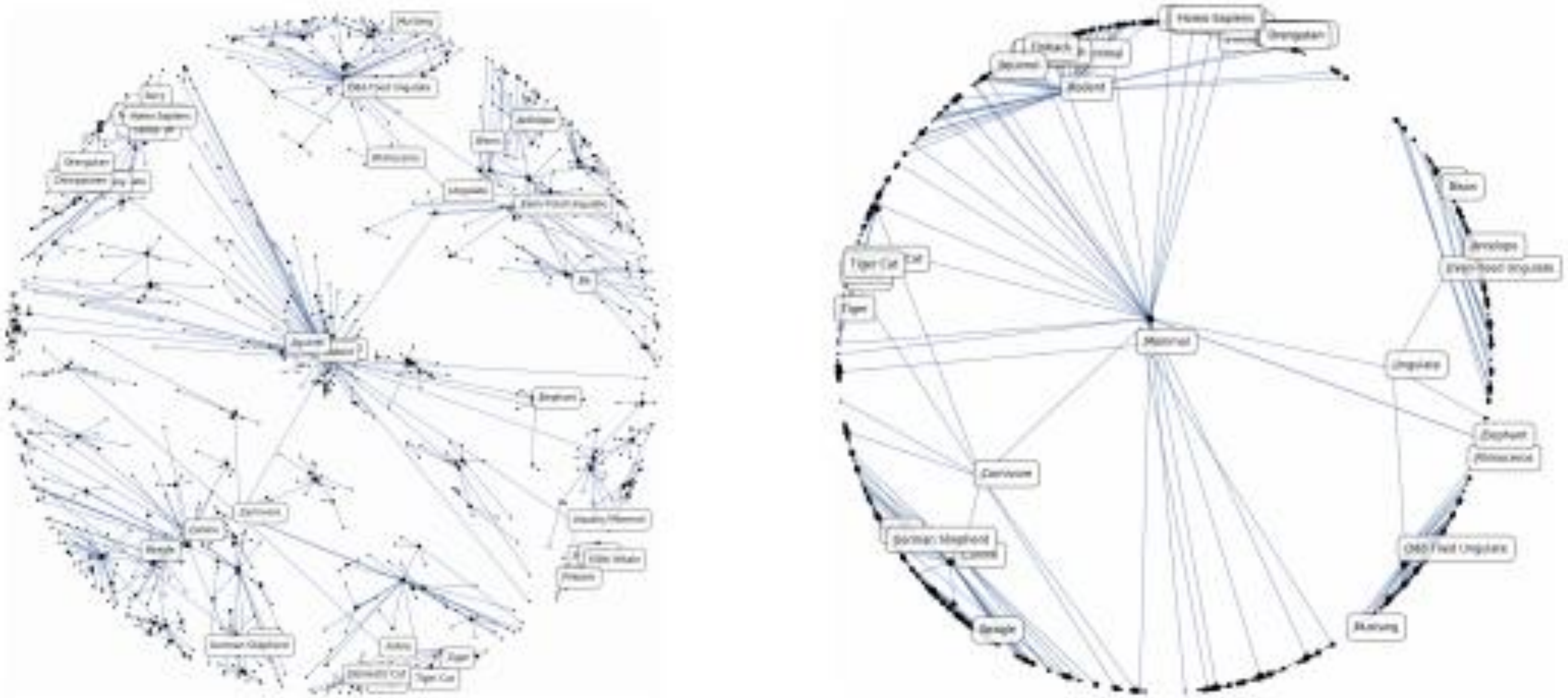
– Recap



(a) Geodesics of the Poincaré disk (b) Embedding of a tree in $\mathcal{B}^2$ (c) Growth of Poincaré distance

Figure 1: (a) Due to the negative curvature of $\mathcal{B}$, the distance of points increases exponentially (relative to their Euclidean distance) the closer they are to the boundary. (c) Growth of the Poincaré distance $d(\boldsymbol{u}, \boldsymbol{v})$ relative to the Euclidean distance and the norm of $\boldsymbol{v}$ (for fixed $\|\boldsymbol{u}\| = 0.9$). (b) Embedding of a regular tree in $\mathcal{B}^2$ such that all connected nodes are spaced equally far apart (i.e., all black line segments have identical hyperbolic length).

– Recap



(a) Intermediate embedding after 20 epochs (b) Embedding after convergence

Figure 2: Two-dimensional Poincaré embeddings of transitive closure of the WordNet mammals subtree. Ground-truth `is-a` relations of the original WordNet tree are indicated via blue edges. A Poincaré embedding with $d = 5$ achieves mean rank 1.26 and MAP 0.927 on this subtree.

– Hyperbolic spaces and their Cartesian product

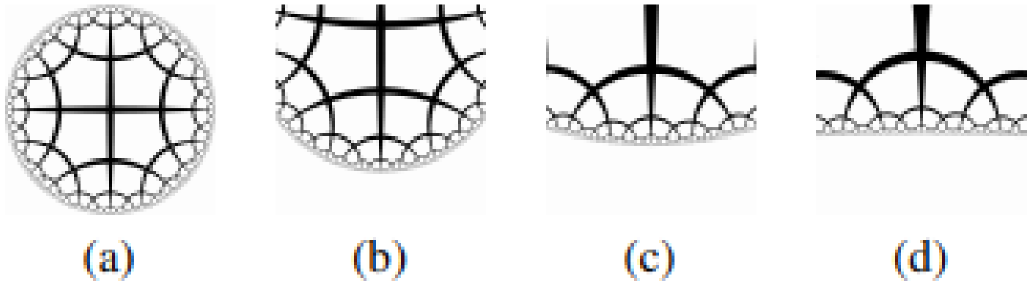


Figure 1: Isometric deformation φ of $\mathbb{D}^2$ into $\mathbb{H}^2$.

$$d_{\mathbb{D}^n}(x, y) = \cosh^{-1}\left(1 + \lambda_x \lambda_y \|x - y\|_2^2 / 2\right)$$

$$\lambda_x := 2/(1 - \|x\|_2^2)$$

$$d_{(\mathbb{D}^n)^p}(x, y)^2 = \sum_{i=1}^{p} d_{\mathbb{D}^n}(x_i, y_i)^2.$$

$$d_{\mathbb{H}^2}(x, y) = \cosh^{-1}\left(1 + \|x - y\|_2^2 / (2 y_1 y_2)\right)$$

– Adapting GloVe

$$X_i = \sum_k X_{ik};\ P_{ij} = X_{ij}/X_i$$

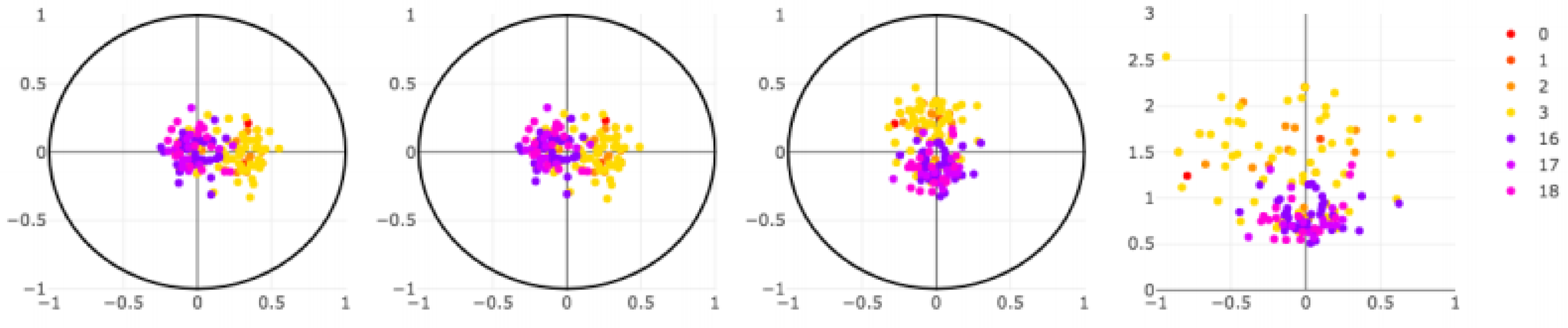
$$J = \sum_{i,j=1}^{V} f(X_{ij}) \left(w_i^T \tilde{w}_j + b_i + \tilde{b}_j - \log X_{ij} \right)^2 ,$$

$$J = \sum_{i,j=1}^{V} f(X_{ij}) \left(-h(d(w_i, \tilde{w}_j)) + b_i + \tilde{b}_j - \log X_{ij} \right)^2 ,$$

– Connecting Gaussian Embeddings & Hyperbolic Embeddings

$$d_F\left(\mathcal{N}(\mu,\sigma^2),\mathcal{N}(\mu',\sigma'^2)\right)=\sqrt{2}d_{\mathbb{H}^2}\left((\mu/\sqrt{2},\sigma),(\mu'/\sqrt{2},\sigma')\right).$$

$$d_F\left(\mathcal{N}(\mu,\Sigma),\mathcal{N}(\mu',\Sigma')\right)=\sqrt{\sum_{i=1}^{n}2d_{\mathbb{H}^2}\left((\mu_i/\sqrt{2},\sigma_i),(\mu_i'/\sqrt{2},\sigma_i')\right)^2}.$$


1
0.5
0
−0.5
−1
−1
−0.5
0
0.5
1
3
2.5
2
1.5
1
0.5
0
0
1
2
3
16
17
18