Nested Hyperbolic Spaces for Dimensionality Reduction and Hyperbolic NN Design

Xiran Fan
University of Florida
Department of Statistics
fanxiran@ufl.edu

Chun-Hao Yang
National Taiwan University
Institute of Applied Mathematical Sciences
chunhaoy@ntu.edu.tw

Baba C. Vemuri
University of Florida
Department of CISE
vemuri@ufl.edu

Data	Isotropic		Anisotropic		Truncated		Fig. 5	
var.	0.2	1.0	0.2	1.0	0.2	1.0	(i)	(ii)
tPCA	0.34	12.77	2.13	3.53	0.31	11.21	0.12	0.08
Horo	0.36	25.30	3.27	4.85	0.32	23.12	0.18	0.12
EPGA	**0.31**	9.62	1.88	2.62	**0.27**	8.84	0.12	0.08
NH	0.33	**9.53**	**1.87**	**2.60**	0.29	**8.60**	**0.01**	**0.01**

Table 1. Reconstruction errors for data from isotropic, anisotropic and truncated normal distributions respectively with varying variance. Last column depicts results for data in Fig.5 of the paper.

A. Overview

In this supplementary material we include additional experiments on synthetic data (Section B) and discuss the optimization of NHGCN in more detail (Section C).

B. Synthetic Experiments

In the main paper, we claimed that our method performs indistinguishably compared to EPGA in isotropic wrapped normal synthetic data because the data are distributed symmetrically around the Fréchet mean. To verify this claim, we include additional experiments under different data distributions - anisotropic wrapped normal distribution and truncated wrapped normal distribution - which are shown in Table 1. As evident, NH performance is indistinguishable from EPGA for both isotropic and anisotropic normal data, whereas tPCA and HoroPCA fail to capture the data pattern for the *large data variance case*. In addition, we report the reconstruction error for data as shown in Fig.5 (in the paper) in Table 1.The NH model is better than EPGA since data does not lie around a geodesic submanifold in this case.

C. Discussion of Optimization in NHGCN

The optimization of transformation matrix $\boldsymbol{W}$ in Eq. (15) is performed over a semi-Riemannian manifold. To the best of our knowledge, [2] is the only reported work on optimization in a semi-Riemannian manifold setting. However, [2] developed it for the indefinite orthogonal group which does not apply to our case, the Stiefel manifold equipped with an indefinite metric. This is an open research problem and will be the focus of our future work. Our current approach is to decompose the matrix and then update the components sequentially, see Section 3.3 for the decomposition. This approach introduces extra parameters and may be sub-optimal. We compare the computation time for 5 epochs and number of parameters of HGNN [1] and the proposed NHGCN. The results are reported in Table 2. We can see that our method introduces much more parameters than HGCN. However, the computation time is comparable for most cases.

References

[1] Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec. Hyperbolic graph convolutional neural networks. *Advances in Neural Information Processing Systems*, 32:4868–4879, 2019. 1

[2] Tingran Gao, Lek-Heng Lim, and Ke Ye. Semi-riemannian manifold optimization. *arXiv:1812.07643*, 2018. 1

		Disease		Airport		PubMed		Cora	
	Task	LP	NC	LP	NC	LP	NC	LP	NC
HGCN [1]	# parameters	464	∼16.0K	483	260	∼8.3K	∼ 8.0K	∼23.2K	∼23.1K
	time (s/5 epochs)	0.03	0.01	0.05	0.01	0.10	0.01	0.10	0.01
NHGCN(Ours)	# parameters	740	∼1.0M	778	1260	∼257.9K	∼25.9K	∼2.1M	∼2.1M
	time (s/5 epochs)	0.03	0.06	0.05	0.03	0.10	0.02	0.13	0.04

Table 2. Comparison of number of parameters and running times for HGCN and NHGCN in the graph neural networks experiments.