Low-stress data embedding in the hyperbolic plane using multidimensional scaling

Andrej CVETKOVSKI · Mark CROVELLA

Abstract

Abstract Multidimensional scaling (MDS) is a technique used to produce two- or three-dimensional visualizations of similarities within datasets consisting of multidimensional points or point distances. Metric MDS can also be applied to the problems of graph embedding for approximate encoding of edge or path costs using node coordinates in metric space. Thus, MDS can be used both as a visualization tool and as a data embedding algorithm.
Recent works have noted that for certain datasets, hyperbolic target space may provide a better fit than the traditionally used Euclidean space. In this paper we present several numerical examples of low-stress embedding of synthetic and real-world datasets in the hyperbolic plane. We demonstrate that the hyperbolic plane often but not always accomodates a better fit for the embedded data compared to the Euclidean plane. Therefore, we conclude that the suitability of the hyperbolic space for low-stress data embedding cannot be attributed solely to the properties of the hyperbolic space, but also to the conformity of datasets produced by natural interaction with the structure of the hyperbolic space. The embeddings we present are produced with PD-MDS, a metric MDS algorithm designed specifically for the Poincaré disk (PD) model of the hyperbolic plane.

Keywords Dimensionality reduction, hyperbolic embedding, hyperbolic multidimensional scaling, network graph, steepest descent, visualization, Poincaré disk, approximate line search

1 INTRODUCTION

When points are given by coordinates in a metric space of some dimension, it is trivial to calculate the distance between any two of those points using the distance function of the space. The inverse problem is often times of interest as well: Knowing the distances for all or some pairs of points, choose a metric space of suitable dimensionality and find in it point coordinates such that the distances for pairs in the space are as close as possible to the corresponding known distances. Problems of this type are dealt with by a class of algorithms known as *metric multidimensional scaling* (MDS) [1, 2].

In the setting of metric multidimensional scaling, the given input distances are termed *dissimilarities*, and for a chosen *target metric space* of some dimensionality, the set of output coordinates, termed the output *point configuration,* is calculated by the algorithm. The distances induced by the output configuration ought to match the input distances as close as possible. To arrive at an optimal configuration, MDS usually calls for a numerical iterative optimization procedure. A suitably chosen badness-of-fit (or stress) measure serves as an objective function of the minimization. Ideally, the final output configuration is a point of global minimum of the objective function. The

*Andrej CVETKOVSKI, Mother Teresa University, Skopje, Macedonia; Mark CROVELLA, Boston University, Boston MA, USA. Correspondence to: Andrej CVETKOVSKI, E-mail: andrej at iterbium com

†The authors acknowledge the National Science Foundation for supporting this work under Grant no. CNS-1018266.

badness-of-fit ought to be as small as possible, and is ideally zero – when the output configuration coordinates match the input pairwise distances.

When the input to MDS represents weights of edges of a given graph, MDS can be viewed as a *graph embedding algorithm*, producing coordinates for the graph vertices which encode ideally or approximately the edge weights of the given graph. Additionally, when the target space is 2- or 3-dimensional, e.g. the Euclidean plane or space, the output configuration or graph embedding can also be *visualized* to graphically represent possible clusters of similarity in the input data.

Traditionally, MDS uses Euclidean space as target. A generalization of MDS to curved surfaces can be viewed as MDS restricted to a subspace in Euclidean space [3]. MDS on constant-curvature Riemannian spaces was studied in [4], using the hyperboloid model. The suitability of the *hyperbolic plane* for practical data and graph embedding was investigated for the tasks of network embedding for path cost estimation [5] and routing [6, 7, 8, 9, 10, 11]. These works indicate that the hyperbolic space accommodates embeddings with a better fit than the Euclidean space of the same dimensionality.

In this paper we seek to demonstrate that 2-dimensional hyperbolic space i.e. *the hyperbolic plane* can be used to embed naturally generated graphs using MDS and obtain configurations which encode the input data with a smaller loss (stress) compared to the MDS in the Euclidean plane. The target model for this purpose is chosen to be the Poincaré disk (PD) model of the hyperbolic plane and the PD-MDS algorithm [12] is used to carry out the numerical experiments. This target space, being a model of the hyperbolic plane, is 2-dimensional. However, unlike its Euclidean counterpart, the hyperbolic plane has the known property of 'exponential' expansion of available space as one moves toward infinity, thus allowing for a better fit of the input data. We also discuss the insights that naturally generated graph data are intrinsically suited to low-stress embeddings in hyperbolic space. To help clarify these insights further, we present a well-rounded set of numerical results for synthetically generated graph data, and compare the results.

The rest of the presentation is organized as follows. Section 2 discusses the existing works on low-stress graph embedding in Riemannian manifolds of constant curvature. Section 3 briefly reviews notions from hyperbolic geometry and Riemannian constant curvature manifolds, and subsequently explains PD-MDS, the embedding method used in this paper to obtain low-stress graph embeddings in the Poincaré disk model of the hyperbolic plane. Section 4 presents numerical examples and results pertaining to multidimensional scaling of synthetic Euclidean, spherical, and hyperbolic graphs in the PD model. Section 5 presents and discusses the results of the embedding of real-world graphs in the PD. Concluding remarks are given in Section 6.

2 RELATED WORK

The advantages of using the hyperbolic instead of the Euclidean plane for data embedding, along with the applicability of various numerical optimization methods, were initially investigated in several works.

The applicability of metric multidimensional scaling to mapping of synthetic and real data in the Poincaré disk model of the hyperbolic plane was studied by [13]. The use of MDS in [13] inherently generalized the applicability of the method from tree structures to continuous-valued multidimensional data or pair distances. It was noted that the "exponential growth of the available space" in the hyperbolic models as one moves towards infinity, makes planar hyperbolic embedding a suitable choice for both hierarchical and high-dimensional data. However, a concrete statement of the used algorithm was not provided.

[5] studied hyperbolic embeddings of the Internet graph for distance estimation and overlay construction. [5] tried to argue that models of the hyperbolic space with circular symmetry may be better suited than Euclidean spaces for embedding network graphs with core-and-tendrils structure. Their work concentrated on applications specific to

communication networks where dissimilarities for each pair of points are derived from the lengths of the shortest paths in the graph. For such applications, the authors' insight about the choice of the Poincaré disk model was that the shortest paths in the studied networks often pass through the core and are therefore longer than the straight-line distance, and this observation empirically matches the behavior of distance function in the chosen model. In order to avoid the constrained nature of the coordinates in the PD, the authors eventually resorted to the hyperboloid model of the hyperbolic plane, omitting the details.

The "big-bang simulation" (BBS) numerical method used in [5], is discussed in [14]. BBS is a variant of a steepest descent method that models the point configuration as an inertial system in a force-generating field. Termination is guaranteed by introducing empirical dampening in the mechanical system. The initial configuration in BBS is always chosen to be a single point in which all particles are collocated, ensuring a fair initial amount of potential energy. Another heuristic feature of BBS is that the objective function changes several times during the minimization in a way that increases the error sensitivity of the penalty terms. The particle inertia in BBS in conjunction with a stepwise changing objective function possibly allows the method to escape a few local minima before termination. However, the advantages of these heuristics in avoiding local minima, compared to a computationally simpler, single phase minimization procedure, were not clearly demonstrated. It is conceivable that the inertial minimum-avoiding mechanism, which comes at an increased computational cost, may as well cause the configuration to leave the global minimum, or a lower local minimum before stopping in a higher one. Finally, since BBS can only be started from one possible initial configuration, it has a deterministic outcome once the heuristic parameters such as friction and time slice are chosen; with this choice, the possibility that the final result is improved by restarting from different initial conditions, is eliminated.

An interesting account of the connections between hyperbolic geometry and the topology of complex networks is given in [15]. It is argued that complex networks are underpinned by naturally arising hierarchies which can be approximated by tree-like structures. The metric properties, of these hierarchies then, allow for successful embedding into hyperbolic spaces. Examples studied in [15] to support these assumptions include social networks, citation networks, biology networks and the evolution of the Internet AS graph. These findings are in accord with the observations we make in Sections 4 and 5 of this paper.

3 PD-MDS: MULTIDIMENSIONAL SCALING IN THE POINCARÉ DISK

To embed graph data in hyperbolic space, in this work we use the Poincaré disk (PD) model of the hyperbolic plane as target and apply the PD-MDS algorithm [12] to the input.

The Poincaré disk (PD) model is suitable for our considerations because it is a planar model with circular symmetry having a closed form distance formula [16]. The distance $d_{\mathbb{D}}(z_j, z_k)$ between two points j and k in PD is given by

$$d_{\mathbb{D}}(z_j, z_k) = 2\text{atanh}\frac{\left|z_j - z_k\right|}{\left|1 - z_j\overline{z_k}\right|}, \tag{1}$$

where z_j and z_k are the complex coordinates of the points in the PD

$$\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\},$$

and $\overline{()}$ is the complex conjugate.

The points on the disk $\mathbb{D}$ with $|z| = 1$ make the *boundary at infinity* (or the *horizon*) for this model. These points are termed *ideal points.*

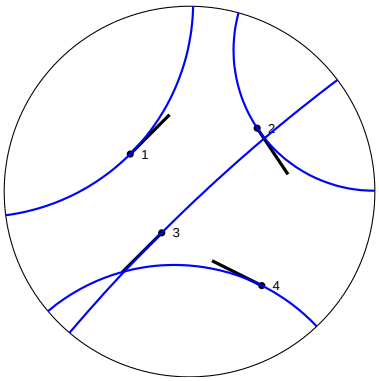


Figure 1: Geodesics (hyperbolic lines) the Poincaré disk and steepest descent direction

In the Poincaré disk model, the paths that realize the hyperbolic distance (1), that is, the *hyperbolic lines* (or *geodesics*) appear as arcs of Euclidean circles in $\mathbb{D}$ that are orthogonal to the horizon at their ideal points (See Fig. 1).

PD-MDS iteratively moves a chosen initial point configuration so as to minimize a badness-of-fit measure. Formally, the algorithm consists of a descent method suited to the Poincaré-disk model and an approximate, binary hyperbolic line search, that together make a simple and computationally efficient minimization method for this model of the hyperbolic plane.

The badness-of-fit (i.e. stress or embedding error) measure to be minimized by PD-MDS is a least-squares function proposed by [17]. The Sammon stress represents a normalized sum of the squared differences between the original dissimilarities δ_{jk} and the embedded distances of the current point configuration d_{jk}. In this work we use the least squares error function of the type

$$E = c \sum_{j=1}^{n} \sum_{k=j+1}^{n} c_{jk} \left(d_{jk} - a\delta_{jk} \right)^2 . \tag{2}$$

Eq. (2) [12] is a general form from which several special embedding error functions can be obtained by substituting appropriate values of the constants c, c_{jk}, and a. Examples include :

- Absolute Differences Squared (ADS)

$$E = \sum_{j=1}^{n} \sum_{k=j+1}^{n} w_{jk} \left(I_{jk} \left(d_{jk} - a\delta_{jk} \right) \right)^2 \tag{3}$$

- Relative Differences Squared (RDS)

$$E = \sum_{j=1}^{n} \sum_{k=j+1}^{n} w_{jk} \left(I_{jk} \frac{d_{jk} - a\delta_{jk}}{a\delta_{jk}} \right)^2 \tag{4}$$

- Sammon stress Criterion (SAM)

$$E = \frac{1}{a \sum_{j=1}^{n} \sum_{k=j+1}^{n} I_{jk}\delta_{jk}} \cdot \sum_{j=1}^{n} \sum_{k=j+1}^{n} w_{jk} \frac{\left(I_{jk} \left(d_{jk} - a\delta_{jk} \right) \right)^2}{a\delta_{jk}} \tag{5}$$

where w_{jk} are individual weights that can be assigned to the input dissimilarities in the most general case, and I_{jk} are indicators to allow for missing dissimilarities in the input.

The objective function can optionally be normalized per pair by dividing with the number of summands $\left(n^2-n\right)/2$.

Ideally, the objective function should be chosen to be *scale-invariant* in the sense that scaling the input dissimilarities and the coordinates of the point configuration with some constant a does not change the embedding error. This is possible for Euclidean space since the Euclidean distance function scales as the point coordinates:

$$\sqrt{\sum_s (a\cdot y_{js} - a\cdot y_{ks})^2} = a\cdot d_{jk}.$$

In the Euclidean case, (4) and (5) are scale-invariant, but (3) is not. In the hyperbolic case (i.e. when d_{jk} is the *hyperbolic* distance function (1)), none of the (3)–(5) are scale-invariant. Therefore, the simplest ADS error function (3) may be a preferable choice for reducing the computational cost in the hyperbolic case. Nevertheless, for our numerical experiments we choose to apply the Sammon criterion (5) so as to facilitate numerical comparison between the final embedding errors for the Sammon map in the Euclidean plane and PD-MDS.

The lack of scale-invariance of the hyperbolic distance formula (1) implies an additional degree of freedom in the optimization of the embedding error – the *dissimilarity scaling factor*. In Eqs. (2)–(5) this extra degree of freedom is captured via the parameter a that scales the original entries of the dissimilarity matrix. The dependency of the embedding error of PD-MDS on the dissimilarity scaling factor a varies with the type of input data and is investigated in more detail in Section 4.

Following the specification of PD-MDS described in details in [12], we successfully implemented PD-MDS with the error function (5). In the following sections we show illustrative results of our experimental study using PD-MDS on synthetic as well as real-world data. Some of the methods we used to verify the correctness of our specification are also discussed below.

4 EMBEDDING OF SYNTHETIC GRAPHS

To investigate the dependency of the embedding error on the dissimilarity scaling factor a, we used as input the inter-point distances obtained from random sets of points residing on surfaces of constant positive, zero or negative curvature (i.e. respectively a sphere, a Euclidean plane and the Poincaré disk model of the hyperbolic plane.) The corresponding distances (spherical, Euclidean and hyperbolic) for all pairs were used as dissimilarities in this experiment. The embedding error function was the Sammon criterion (5). We also used noisy inputs obtained by replacing each original dissimilarity δ_{jk} with a value chosen uniformly at random from the interval $\left[(1-e_m)\delta_{jk},\ (1+e_m)\delta_{jk}\right]$ for a chosen noise level $e_m < 1$.

Fig. 2 shows the typical effects of dissimilarity scaling for several Euclidean, spherical, and hyperbolic graphs. Cases (a), (c), and (e) illustrate the variation of the embedding error for noiseless input data, with the number of points as a parameter (20 and 60 points.) Cases (b), (d), and (f) illustrate the variation of the embedding error for noisy input data and are parametrized by the amount of measurement noise ($e_m = 0, 10, 20, 30\%$.)

Each point in the diagrams (a)–(f) was obtained as a minimum Sammon stress in a series of 70 replicates of PD-MDS for different randomly chosen initial configuration in the PD. The smoothness of the obtained curves demonstrates that for the chosen problems, this number of replicates was enough to approach the global minimum achievable embedding error for each simulated value of a. The results are drawn on semilogarithmic axes in order to show more details toward small a values.

Locally, the Poincaré disk model, distance-wise "looks like" the Euclidean plane scaled with some constant factor.

For example, in the vicinity of a point z_0, the hyperbolic distance formula (1) becomes

$$d_{\mathbb{D}}(z_j, z_k) \approx \left|z_j - z_k\right| \cdot 2/(1-|z_0|^2).$$

Therefore, for a sufficiently small scaling factor a and sufficiently many replicates, metric MDS implementations for the PD model and for the Euclidean plane using the same scale-invariant (for Euclidean distances) error function, should return approximately equal embedding errors for the final configurations. (A sufficiently small value of a is one that would make the final configuration land in a sufficiently small neighborhood of a point in the PD.)

In this sense, PD-MDS is a generalization of an Euclidean MDS algorithm. We used these observations to verify that our PD-MDS implementation returned the expected error values for small scaling factors. Indeed, as Fig. 2 shows, for small a values, the Euclidean graphs were embeddable with no error, and the other two graph types had stress that numerically matched the output of other available Euclidean MDS implementations using the Sammon stress criterion.

In the cases (e) and (f), the original configurations are residing on a hyperbolic plane, and therefore are embeddable with zero stress in the PD model for some value of a ($a = 1$ in this synthetic example). For this value, our implementation of PD-MDS was able to find the original configuration up to hyperbolic-distance preserving Möbius transformations.

The diagrams (b), (d), and (f) (Fig. 2) also demonstrate that relatively high noise levels in the measured data do not significantly change the suitability for embedding in the PD in the cases when the original dissimilarity matrix has a natural underlying 2-dimensional space.

5 EMBEDDING OF REAL-WORLD GRAPHS IN THE HYPERBOLIC PLANE

To further demonstrate the ability of the Poincaré disk to accommodate lower stress 2-dimensional embeddings than classical Euclidean MDS for certain graph types, we applied PD-MDS to dissimilarity matrices obtained from several real-world datasets.

In this section we summarize the results.

5.1 The Iris Dataset

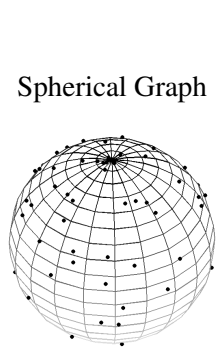
As a first experiment with real-world data, we apply PD-MDS to the Iris dataset [18]. This classical dataset consists of 150 4-dimensional points from which we extract the Euclidean inter-point distances and used them as input dissimilarities. The embedding error as a function of the scaling factor a is shown in Fig. 3. Each value in the diagram is obtained as a minimum embedding error in a series of 100 replicates starting from randomly chosen initial configurations.

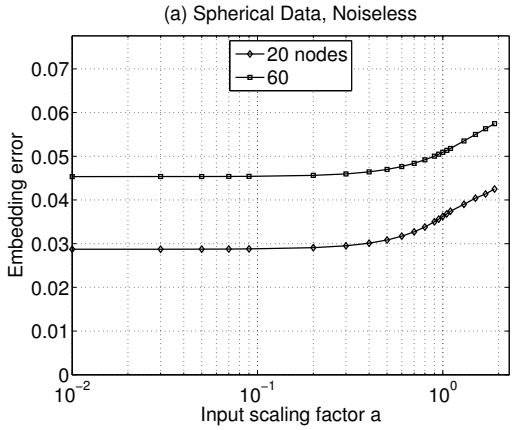
Minimal embedding error overall is achieved for $a \approx 4$. The improvement with respect to the 2-dimensional Euclidean case is 10%. The Iris dataset is an example of dimensionality reduction of an original higher-dimensional dataset that can be done more successfully using the hyperbolic plane than the Euclidean plane.

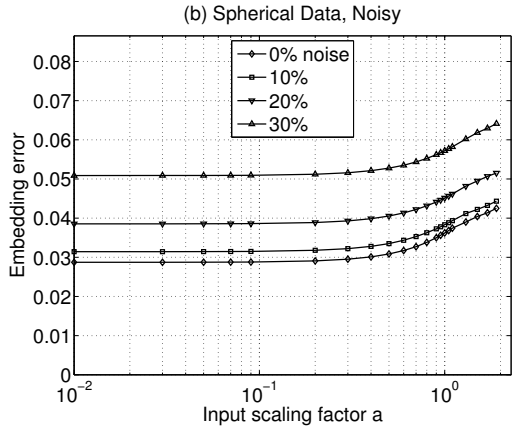
5.2 Political Books

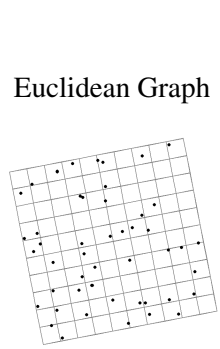
An interesting network was presented by [19], who assembled a connectivity graph of political books frequently bought together during an election campaign.

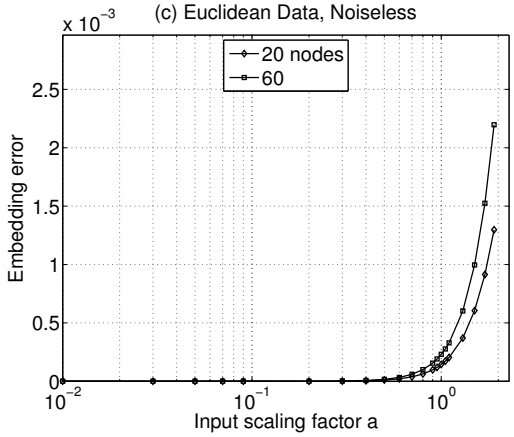
In the graph version we used, there were 105 nodes representing books and a total of 441 undirected, unweighted links between books that were frequently bought together. We obtained dissimilarities by assigning self-dissimilarity

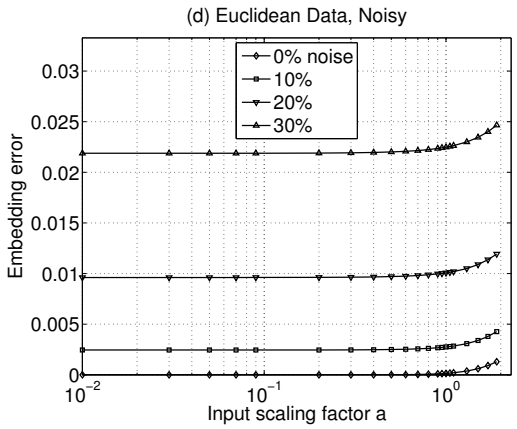


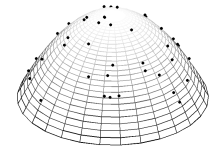


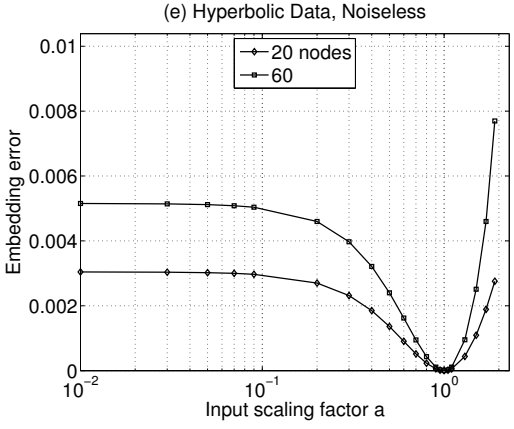












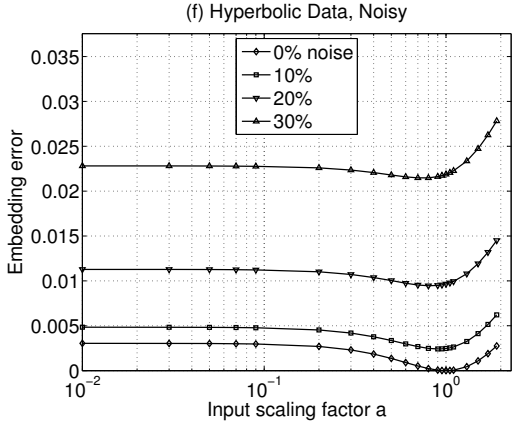


Figure 2: The charts show the *embedding error* when spherical, Euclidean, and hyperbolic graphs are embedded in the Poincaré disk using PD-MDS with a varying scaling factor a. Figures (a), (c), and (e) show the error for noiseless input data with the number of nodes in the graph as a parameter. Figures (b), (d), and (f) show the embedding error for noisy input data for various amounts of measurement noise.

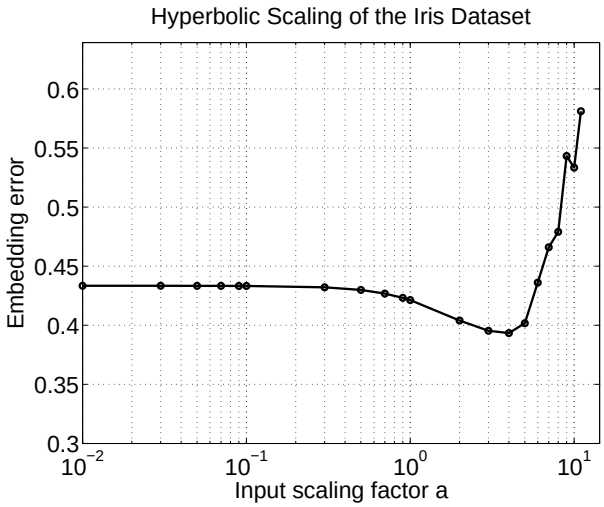


Figure 3: The effect of scaling of the dissimilarities on the embedding error for the Iris Dataset [18]. The input dissimilarities are the Euclidean distances between pairs of original points. This PD-MDS result reveals that the Iris dataset is better suited for embedding to the hyperbolic plane that to the Euclidean plane.

of 0, dissimilarity of 1 for co-purchased books and a missing (unknown) dissimilarity for the remaining pairs and applied PD-MDS to the resulting dissimilarity matrix. We also conducted the experiment using only the liberal and the conservative subgraphs (43 and 49 points respectively).

The obtained minimum embedding error of 150 replicates as a function of the scaling factor a is shown in Fig. 4. We note that there were remarkable gains of using the PD model instead of the Euclidean plane. For the overall graph, the minimal stress was 7.6 times smaller than the Euclidean stress. The liberal and conservative components alone achieved improvement of 8.8 and 9 times with respect to the Euclidean case.

5.3 A Citation Network

The citation network that we used in this experiment was compiled by [20] from bibliographies of review articles on networking. We extracted the largest connected component from the graph which consisted of 379 nodes representing authors. There were 244 edges in the graph with weights s_{jk} representing the strength of the collaborative ties. We have calculated dissimilarities from these data using $\delta_{jk} = \text{const} - s_{jk}$ and applied the PD-MDS algorithm.

The obtained minimum embedding error of 50 replicates as a function of the scaling factor a is shown in Fig. 5. The overall minimum embedding error was 2.63 times lower than the stress obtained using Euclidean MDS.

6 CONCLUSION

In this paper we presented several numerical examples of low-stress embedding of synthetic and real-world datasets in the hyperbolic plane. The embeddings we presented were produced with PD-MDS, a metric MDS algorithm designed specifically for the Poincaré disk (PD) model of the hyperbolic plane. The quality of the embedding was measured using least squares error functions.

The existence of graphs with a hyperbolic "underlying structure" that are embeddable with notably lower stress in 2-dimensional hyperbolic than in 2-dimensional Euclidean space was demonstrated in Section 4. It is an important illustration of the usefulness of MDS in the Poincaré disk. Candidate graphs having such hyperbolic-like structure are

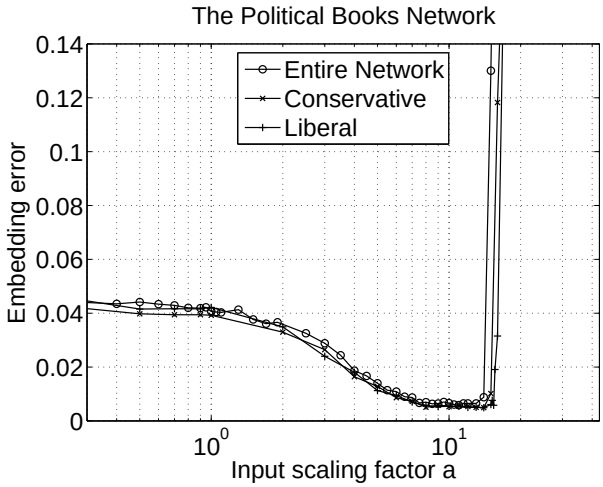


Figure 4: The embedding error as a function of the input scaling factor for the network of political books [19]. The input dissimilarities are simply indicators of the presence or absence of links between the network nodes. The political books network is an example of unweighted, undirected real-world graph data that can be embedded with lower error in the PD model than in the Euclidean plane.

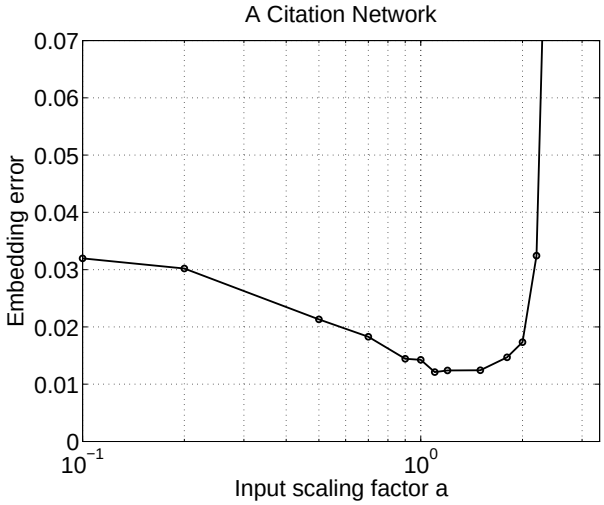


Figure 5: The embedding error as a function of the input scaling factor for the network of citations [20]. The network is an example of weighted graph data that can be encoded with lower embedding error in the PD model than in the Euclidean plane. The input dissimilarities are capturing the strength of the collaboration between pairs of authors.

real-world communication or social networks that tend to have a strongly interconnected core and a sparser periphery of tendrils. PD-MDS can be used in such contexts to investigate the "hyperbolicity" of the input data and arrive at lower stress dissimilarity embedding.

We demonstrated that the hyperbolic plane often but not always accomodates a better fit for the embedded data than the Euclidean plane. We noted that for naturally arising networks, in our experiments, the hyperbolic plane was always able to provide for a better fit of the input data compared to its Euclidean counterpart. We conclude that the suitability of the hyperbolic space for low-stress data embedding cannot be attributed solely to the properties (i.e. the less-restricted axiomatic foundation) of the hyperbolic space, but also to the conformity of datasets to the structure of the hyperbolic space. As noted in this and other works, this conformity is, as a rule, observed in networks arising from natural interaction of a number of independent entities.

Acknowledgement

The authors acknowledge the National Science Foundation for supporting this work under Grant no. CNS-1018266.

References

[1] T. F. Cox and M. A. A. Cox. *Multidimensional Scaling (Monographs on Statistics and Applied Probability)*. Chapman & Hall/CRC, 2nd edition, 2000.

[2] I. Borg and P. J. F. Groenen. *Modern Multidimensional Scaling: Theory and Applications (Springer Series in Statistics)*. Springer, Berlin, 2nd edition, 2005.

[3] T. F. Cox and M. A. A. Cox. Multidimensional scaling on a sphere. *Communications in Statistics*, 20:2943–2953, 1991.

[4] H. Lindman and T. Caelli. Constant curvature Riemannian scaling. *J. Math. Psychol.*, 2:89–109, 1978.

[5] Y. Shavitt and T. Tankel. Hyperbolic embedding of Internet graph for distance estimation and overlay construction. *IEEE/ACM Trans. Netw.*, 16(1):25–36, Feb. 2008.

[6] R. Kleinberg. Geographic routing using hyperbolic space. In *Proceedings of IEEE Infocom 2007*, pages 1902–1909, May 2007.

[7] D. Krioukov, F. Papadopoulos, M. Boguñá, and A. Vahdat. Greedy forwarding in scale-free networks embedded in hyperbolic metric spaces. *SIGMETRICS Performance Evaluation Review*, 37:15–17, October 2009.

[8] A. Cvetkovski and M. Crovella. Hyperbolic embedding and routing for dynamic graphs. In *Proceedings of IEEE Infocom 2009*, pages 1647–1655, Apr 2009.

[9] F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *INFOCOM, 2010 Proceedings IEEE*, March 2010.

[10] A. Cvetkovski and M. Crovella. Low-stretch greedy embedding heuristics. In *Fourth International Workshop on Network Science for Communication Networks (NetSciCom '12)*, pages 232 –237, In conjunction with IEEE Infocom '12, Orlando, FL, USA, march 2012.

[11] A. Cvetkovski and M. Crovella. On the choice of a spanning tree for greedy embedding of network graphs. *Networking Science*, pages 1–11, 2013.

[12] A. Cvetkovski and M. Crovella. Multidimensional scaling in the Poincaré disk. *Applied Mathematics & Information Sciences*, 10(1):125–133, 2016.

[13] J. A. Walter. H-MDS: a new approach for interactive visualization with multidimensional scaling in the hyperbolic space. *Information Systems*, 29(4):273 – 292, 2004.

[14] Y. Shavitt and T. Tankel. Big-bang simulation for embedding network distances in Euclidean space. *IEEE/ACM Trans. Netw.*, 12(6):993–1006, 2004.

[15] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010.

[16] J. W. Anderson. *Hyperbolic Geometry*. Springer, 2nd edition, 2007.

[17] J. W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.*, C-18(5):401–409, May 1969.

[18] E. Anderson. The irises of the Gaspé peninsula. *Bulletin of the American Iris Society*, 59:2–5, 1935.

[19] V. Krebs. Case studies - Political Book Networks, 2008. Available online at www.orgnet.com.

[20] M. E. J. Newman. Scientific collaboration networks II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64(1):016132, 1–7, Jun 2001.