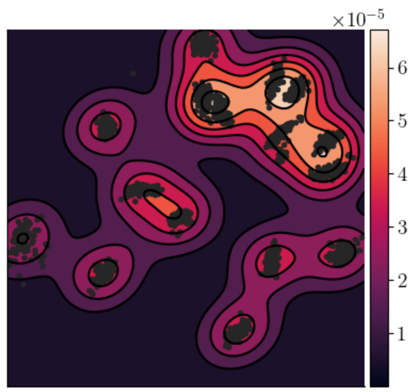
Inherent uncertainty of hyperbolic embeddings of complex networks

Simon Lizotte,[1,2] Jean-Gabriel Young,[2,3,4] and Antoine Allard[1,2,4]
[1] *Département de physique, de génie physique et d'optique, Université Laval, Québec, QC, G1V 0A6 Canada*
[2] *Centre interdisciplinaire en modélisation mathématique, Université Laval, Québec, QC, G1V 0A6 Canada*
[3] *Department of Mathematics and Statistics, University of Vermont, Burlington VT, 05405, USA*
[4] *Vermont Complex Systems Center, University of Vermont, Burlington VT, 05405, USA*

Network geometry is a versatile yet simple framework that captures several of the observed properties of empirical networks, such as their non-vanishing clustering, sparsity, and power-law degree distribution [1]. This accurate description is achieved by placing vertices in a metric space (usually hyperbolic) and by connecting them according to their proximity in that space. To use this framework to describe observed networks, one must find vertex coordinates that best reproduce the observed topology, a difficult task generally solved with a mixture of greedy algorithms, simplifying heuristics and machine learning techniques [2–5].

By approaching the embedding task from a Bayesian perspective, we quantify the uncertainty of the inferred positions and parameters [6], thereby going beyond the point-wise estimates obtained with current state-of-the-art embedding techniques. We make use of recent advances in probabilistic programming and implement a Hamiltonian Monte Carlo algorithm to estimate a complete posterior distribution for the embedding [7]. In so doing, we overcome several technical challenges: 1) we introduce a continuous approximation of the likelihood gradients to manage discontinuities caused by vanishing angular distances; 2) we make use of standardization techniques to handle graph and space symmetries and align samples; and 3) we explore the influence of initial conditions.

With these techniques, we uncover genuine multimodality in the posterior distributions that cannot be explained by algorithmic issues or structural or spatial symmetries (see Figure). As such, our work highlights the irreducible uncertainty inherent to the hyperbolic embedding task, thereby paving the way for more comprehensive and accurate embedding algorithms of empirical networks in hyperbolic space.



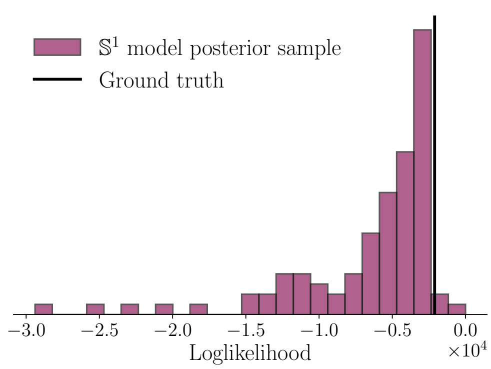


Figure: (left) Multidimensional scaling (MDS) of the angular coordinates sampled from the posterior distribution. Each point is an element of the posterior sample, and similar elements are close by in this plot. The heatmap displays the probability density of a Gaussian kernel density estimation. (right) Likelihood distribution of the posterior embeddings. The sample of the posterior distribution is obtained from 16 chains initialized randomly for a graph generated with the $\mathbb{S}^1$ model [1]. Rotation and reflection symmetries are removed by fixing the angular position of one vertex and by constraining that of a second vertex to $[0, \pi]$. Graph symmetries are removed by enumerating all automorphisms and picking the permutation that minimizes the MDS distance. The left panel highlights the rich multimodality of the posterior distribution. Even though it shows that the Markov chains are unable to exit their local maxima, the right panel shows that the embeddings obtained by those local maxima have a similar likelihood to the ground truth parameters, thus suggesting that there exist many different competing ways to embed the same network.

[1] Nat. Rev. Phys. **3**, 114–135 (2021).
[2] New J. Phys. **21**, 123033 (2019).
[3] Phys. Rev. E **104**, 044315 (2021).
[4] Nat. Commun. **8**, 1615 (2017).
[5] IEEE/ACM Trans. Netw. **26**, 920–933 (2018).
[6] Nat. Commun. **13**, 6794 (2022).
[7] J. Stat. Softw. **76**, 1–32 (2017).