Gromov-Hausdorff Distances
for Comparing Product Manifolds of Model Spaces

Haitz Sáez de Ocáriz Borde [1]   **Álvaro Arroyo** [2]   **Ismael Morales Lopez** [3]   **Ingmar Posner** [1]   **Xiaowen Dong** [2 4]

Abstract

Recent studies propose enhancing machine learning models by aligning the geometric characteristics of the latent space with the underlying data structure. Instead of relying solely on Euclidean space, researchers have suggested using hyperbolic and spherical spaces with constant curvature, or their combinations (known as *product manifolds*), to improve model performance. However, there exists no principled technique to determine the best latent product manifold signature, which refers to the choice and dimensionality of manifold components. To address this, we introduce a novel notion of distance between candidate latent geometries using the Gromov-Hausdorff distance from metric geometry. We propose using a graph search space that uses the estimated Gromov-Hausdorff distances to search for the optimal latent geometry. In this work we focus on providing a description of an algorithm to compute the Gromov-Hausdorff distance between model spaces and its computational implementation.

1. Introduction

Recent research has shown a growing interest in using concepts from differential geometry and topology to improve learning algorithms (Bortoli et al., 2022; Hensel et al., 2021; Chamberlain et al., 2021; Huang et al., 2022; Barbero et al., 2022b;a). While Euclidean spaces have been commonly used in machine learning, it has been found that using geometries that align better with the underlying data

[1]Oxford Robotics Institute, University of Oxford [2]Oxford-Man Institute, University of Oxford [3]Mathematical Institute, University of Oxford [4]Machine Learning Research Group, University of Oxford. Correspondence to: Haitz Sáez de Ocáriz Borde <haitz@oxfordrobotics.institute>.

structure can lead to significant improvements. One approach is to use constant curvature model spaces like the Poincaré ball (Mathieu et al., 2019), hyperboloid (Chami et al., 2019), or hypersphere (Zhao et al., 2019) to encode latent representations of data. Another approach is to use product spaces (Gu et al., 2018; Skopek et al., 2019; Sáez de Ocáriz Borde et al., 2023b;a; Zhang et al., 2020), which combine multiple model spaces, allowing for more complex representations while maintaining computational tractability. However, determining the optimal configuration of product manifolds for representing data is currently done heuristically, lacking efficiency and a principled framework. In this work, we propose a computational approximation of the Gromov-Hausdorff distances between product manifolds of model spaces. This measure of closeness between manifolds can be useful for machine learning applications where finding an appropriate manifold to represent the embedding space of data or latent representations is crucial.

2. Background

Differential Geometry and Product Manifolds. Constant curvature model spaces refer to Riemannian manifolds that have a consistent sectional curvature, meaning the curvature remains the same in all directions within a 2D plane (and can be extended to higher dimensions) (Gallier & Quaintance, 2020). Examples of such spaces include Euclidean space, hyperbolic space, and the sphere. Euclidean space has zero curvature, hyperbolic space has negative curvature, and the sphere has positive curvature. Constant curvature model spaces are commonly used as reference spaces for comparative analysis and machine learning tasks involving non-Euclidean data. A product manifold can be constructed by taking the Cartesian product $\mathcal{P} = \times_{i=1}^{n_{\mathcal{P}}} \mathcal{M}_{Ki}^{d_i}$ of $n_{\mathcal{P}}$ manifolds with curvature K_i and dimensionality d_i. It is important to note that $n_{\mathcal{P}}$ and d_i are hyperparameters that define the product manifold $\mathcal{P}$ and need to be predetermined. On the other hand, the curvature of each model space K_i can be learned using gradient descent. It is worth mentioning that while the product manifold construction allows for the generation of more complex embedding spaces than the original constant curvature model spaces, it does not enable the generation of arbitrary manifolds or provide control over local curvature. In this study, we are interested in geometries

that can be represented employing model space Riemannian manifolds and Cartesian products of such manifolds.

Hausdorff and Gromov-Hausdorff distances. The Hausdorff distance (Jungeblut et al., 2021) between two subsets, A and B, in a metric space X with the metric d_X, can be defined as follows: $d_H^X(A, B) = \max\left(\sup_{a \in A} d_X(a, B), \sup_{b \in B} d_X(b, A)\right)$. This quantity is infinite, therefore we will restrict our calculations to compact subsets A and B. In this particular instance, we define $d_H(A, B)$ as the smallest real number $c \geqslant 0$ such that for every $a \in A$ and every $b \in B$ there exist $a' \in A$ and $b' \in B$ such that both $d_X(a, b')$ and $d_X(a', b)$ are at most c. It is worth mentioning that the previous definition does not impose the need for differentiable structures on X, A, and B. They can simply be metric spaces. This enables us to tell apart the metric characteristics of Euclidean, hyperbolic, and spherical geometries beyond analytical concepts like curvature. Nevertheless, the aforementioned definition restricts the comparison of spaces A and B to instances where they are embedded in a particular metric space X. To assess the notion of distance between them, we will employ the Gromov-Hausdorff distance. Given a metric space X and two isometric embeddings $f : A \to X$ and $g : B \to X$, we define $d_H^{X,f,g}(A, B) = d_H^X(f(A), g(B))$. Next, considering two metric spaces A and B, we denote by $ES(A, B)$, the "embedding spaces of A and B", as the triple (X, f, g) where X is a metric space and $f : A \to X$ and $g : B \to X$ are isometric embeddings. The *Gromov-Hausdorff distance between A and B* is defined as: $d_{GH}(A, B) = \inf_{(X,f,g) \in ES(A,B)} d_H^{X,f,g}(f(A), g(B))$.

3. Methodology: Gromov-Hausdorff Distances for Comparing Product Manifolds

The Hausdorff distance is a measure of dissimilarity between two sets of points, in our case representing discretized versions of continuous manifolds. An algorithm proposed by (Taha & Hanbury, 2015) efficiently computes the exact Hausdorff distance between point sets but assumes that the point sets reside in the same space and have the same dimensionality, which limits its applicability. Gromov-Hausdorff distances allow for measuring the distance between metric spaces that are not inherently embedded in a common ambient space. In the application described, distances need to be calculated between Euclidean space, spherical space, and hyperbolic space, as well as products of these. This work focuses on describing an algorithm for obtaining computational upper bounds for the distances between Euclidean space and hyperbolic space, as well as between spherical space and hyperbolic space. The Gromov-Hausdorff distance between Euclidean space and spherical space can be estimated analytically and it is later provided in Table 1, Section 4. These quantities can also be used to infer Gromov-

Hausdorff distances between product manifolds.

Strategy. There exist multiple isometric embeddings of spaces $\mathbb{E}^n$ and $\mathbb{S}^n$ into $\mathbb{E}^{6n-6}$. While these first two spaces can already be embedded in $\mathbb{E}^{n+1}$, considering this higher-dimensional Euclidean space is beneficial because it allows for the isometric embedding of $\mathbb{H}^n$. This common underlying space enables the computation of Hausdorff distances between the geometries $\mathbb{E}^n$, $\mathbb{S}^n$, and $\mathbb{H}^n$, ultimately leading to the estimation of their mutual Gromov-Hausdorff distance. The embedding of $\mathbb{H}^n$ into $\mathbb{E}^{6n-6}$, which we describe, can be found in Henke & Nettekoven (1987) and is made explicit in Blanuša (1955). For a broader context and related advancements in the isometric embedding of homogeneous spaces into higher-dimensional ones, we recommend referring to Chapter 5 of the book by Brander (2003). In our experiments, we will primarily focus on working with product manifolds generated based on constant curvature model spaces of dimension $n = 2$, and we will denote the embedding of $\mathbb{H}^2$ as $F : \mathbb{H}^2 \to \mathbb{E}^6$. Now we can summarise our strategy to estimate $d_{GH}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2})$ as follows (for $d_{GH}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2})$ it will be entirely analogous). The first step consists of approximating our infinite smooth spaces by finite discrete ones. For this, we consider several collections of points $\{P_i\}_{i \in I}$ in $\mathbb{E}^2$ that are sufficiently well distributed. The exponential map can be applied to the collection of points $\exp : T_0\mathbb{H}^2 \cong \mathbb{R}^2 \to B_{\mathbb{H}^2}$ to get several collections of points Q in $B_{\mathbb{H}^2}$ (again, well distributed by construction). In addition, we will consider several isometric embeddings $f_k : B_{\mathbb{E}^2} \to \mathbb{R}^6$. Hence, we take

$$
\begin{aligned}
d_{GH}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2}) &\approx \min_{i,j,k} d_H^{\mathbb{R}^6, f_k, F}(P_i, Q_j) \\
&= \min_{i,j,k} d_H^{\mathbb{R}^6}(f_k(P_i), F(Q_j)). \quad (1)
\end{aligned}
$$

Note that here we restrict ourselves to $\mathbb{E}^2$, $\mathbb{S}^2$, and $\mathbb{H}^2$ for computational simplicity, but the approach is generalizable to $\mathbb{E}^n$, $\mathbb{S}^n$, and $\mathbb{H}^n$.

Model Space Manifold Discretization. To approximate the mutual Gromov-Hausdorff distance between the smooth spaces $\mathbb{E}^2$ and $\mathbb{H}^2$, we adopt a discretized representation of these spaces. We achieve this by generating well-distributed sets of points in $\mathbb{E}^2$ and $\mathbb{H}^2$ through the application of the exponential map to a collection of points in $\mathbb{R}^2$. Additionally, we consider multiple isometric embeddings of $\mathbb{E}^2$ into $\mathbb{R}^6$. The estimation of the Gromov-Hausdorff distance involves computing the minimum Hausdorff distance between the point sets derived from the collections in $\mathbb{R}^6$ and the isometric embedding of $\mathbb{H}^2$ into $\mathbb{R}^6$. The same approach applies when comparing spherical space. All our computations will be done in dimension two, so we will simply precise how to generate points in $B_{\mathbb{E}^2}$, $B_{\mathbb{S}^2}$ and $B_{\mathbb{H}^2}$. For $B_{\mathbb{E}^2}$ and $B_{\mathbb{S}^2}$, we will use very elementary trigonometry. For $B_{\mathbb{H}^2}$, we will need an explicit de-

scription of the exponential map of $\mathbb{H}^2$. Using the descriptions $B_{\mathbb{E}^2} = \{(r\cos(t), r\sin(t)) : r \in [0,1], t \in [0, 2\pi)\}$, and $B_{\mathbb{S}^2} = \{(\sin(\beta)\cos(\alpha), \sin(\beta)\sin(\alpha), \cos(\beta)) : \alpha \in [0, 2\pi), \beta \in [0,1]\}$. This way, it will be easy to generate collections of points in $B_{\mathbb{E}^2}$ and $B_{\mathbb{S}^2}$. As we anticipated above in the outline of our strategy to estimate $\mathrm{d}_{\mathrm{GH}}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2})$, in order to give explicit well-distributed collection of points in $B_{\mathbb{H}^2}$, it is enough to give a well-distributed collection of points in $B_{\mathbb{E}^2}$ and consider its image under the exponential map $\exp_0 : B_{\mathbb{E}^2} \to B_{\mathbb{H}^2}$, where we have identified $B_{\mathbb{E}^2}$ with the ball of radius one of $\mathbb{R}^2 \cong T_0\mathbb{H}^2$, i.e. the tangent space of $\mathbb{H}^2$ at the point 0.

Mapping function to embed $\mathbb{H}^n$ into $\mathbb{E}^{6n-6}$. We want to define an isometric embedding of $B^1_{\mathbb{H}^n}$ into $\mathbb{E}^{6n-6}$, that is, F. This higher dimensional space is our candidate to fit in the three geometries $\mathbb{E}^n$, $\mathbb{H}^n$ and $\mathbb{S}^n$ to compute their Hausdorff distances as subspaces of $\mathbb{E}^{6n-6}$ and hence estimate their Gromov-Hausdorff distances. Before describing such embedding, we introduce several preliminary auxiliary functions. Let $\chi(t) = \sin(\pi t) \cdot e^{-\sin^{-2}(\pi t)}$ for non-integer values of t. A priori, the inverse of $\sin(0) = 0$ does not make sense but since $\lim_{t \to 0^+} \chi(t) = \lim_{t \to 0^-} \chi(t) = 0$, we can set $\chi(0) = 0$ so it is still continuous. In fact, it is smooth and, in particular, integrable. We can say the same at all points when t is an integer, so we set $\chi(t) = 0$ for all integers t and we obtain an smooth function χ defined on $\mathbb{R}$: $A = \int_0^1 \chi(t)dt$. We also define

$$\psi_1(x) = \sqrt{\frac{1}{A} \cdot \int_0^{1+x} \chi(t)\, \mathrm{d}t} \qquad (2)$$

$$\psi_2(x) = \sqrt{\frac{1}{A} \cdot \int_0^{x} \chi(t)\, \mathrm{d}t} \qquad (3)$$

We set c to be the constant $c = 2\max\{G_1, G_2\}$ defined in terms of

$$G_i = \left\| \frac{\mathrm{d}}{\mathrm{d}x}(\sinh(x) \cdot \psi_i(x)) \right\|_{L^\infty[-2,2]}, \; i = \{1,2\} \quad (4)$$

Further, we make use of two auxiliary functions

$$h(x,y) = \frac{\sinh(x)}{c}\Big(\psi_1(x)\cos(c \cdot y), \psi_1(x)\sin(c \cdot y),$$
$$\psi_2(x)\cos(c \cdot y), \psi_2(x)\sin(c \cdot y)\Big) \tag{5}$$

$$\psi(x,y) = \Big(\sinh^{-1}(ye^x), \log(\sqrt{e^{-2x} + y^2})\Big). \quad (6)$$

which allows us to define

$$f_0(x,y) = \Big(\int_0^{\sinh^{-1}(ye^x)} \sqrt{1 - \varepsilon(t)^2}\mathrm{d}t,$$
$$\log(\sqrt{e^{-2x} + y^2}), h(\psi(x,y)) \Big), \quad (7)$$

with ε being $\varepsilon = \frac{G_1^2 + G_2^2}{c^2}$. This way, we can set

$$f(x, y_1, \ldots, y_{n-1}) =$$
$$= \kappa\big(f_0(x, \sqrt{n-1}y_1), \ldots, f_0(x, \sqrt{n-1}y_{n-1}\big) \quad (8)$$

with $\kappa = \frac{1}{\sqrt{n-1}}$. Recall that in our case, we use the function $F(\cdot) = f(n = 2, \cdot)$ to map the set of points Q in $\mathbb{H}^2$ to $\mathbb{R}^6$, and for computing

$$\mathrm{d}_{\mathrm{GH}}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2}) \approx \min_{i,j,k} \mathrm{d}_{\mathrm{H}}^{\mathbb{R}^6, f_k, F}(P_i, Q_j)$$
$$= \min_{i,j,k} \mathrm{d}_{\mathrm{H}}^{\mathbb{R}^6}(f_k(P_i), F(Q_j)). \quad (9)$$

4. Computational Estimation of the Gromov-Hausdorff Distances

Given that the Gromov-Hausdorff distance between Euclidean and Spherical Space can be calculated analytically, we will focus on the numerical algorithm required to compute the distances between $\mathbb{E}^n$ and $\mathbb{H}^n$, and $\mathbb{S}^n$ and $\mathbb{H}^n$.

Computational Implementation of the Gromov-Hausdorff Distance between the Remaining Model Spaces of Constant Curvature. In this section, we explore the process of discretizing the model spaces to generate points within balls of radius one, considering practical considerations. The Euclidean, hyperbolic, and spherical spaces are continuous manifolds, and in order to approximate the Gromov-Hausdorff distance between them, we need to discretize these spaces. For the Euclidean plane, we sample $B_{\mathbb{E}^2} = \{(r\cos(t), r\sin(t)) : r \in [0,1], t \in [0, 2\pi)\}$, with a discretization of $10,000$ points in both r and t. In order to generate points in $B_{\mathbb{H}^2}$ (hyperbolic space), we will use the points in $B_{\mathbb{E}^2}$ (Euclidean space) as a reference. We will employ the exponential map and coordinate transformations to convert the points in $B_{\mathbb{E}^2}$ into points in $B_{\mathbb{H}^2}$. To avoid numerical instabilities, we restrict the points to $B'_{\mathbb{E}^2} = \{(r\cos(t), r\sin(t)) : r \in [0.00000001, 0.97], t \in [0, 2\pi)\}$ and use a discretization of $10,000$. To sample points for the spherical space, $B_{\mathbb{S}^2} = \{(\sin(\beta)\cos(\alpha), \sin(\beta)\sin(\alpha), \cos(\beta)) : \alpha \in [0, 2\pi), \beta \in [0,1]\}$, we use a discretization of 100 for both α and β. The level of discretization was selected as a balance between resolution and computational time. We noticed that the Gromov-Hausdorff distance reached a stable state with our chosen discretization. Nevertheless, due to the inherent characteristics of the Gromov-Hausdorff distance, it is challenging to determine whether increased discretization might reveal any unexpected behavior.

Optimizing the Embedding Functions for the Euclidean and Spherical Spaces. Next to approximate the Gromov-

Hausdorff distances:

$$d_{\mathrm{GH}}(B_{\mathbb{E}^2}, B_{\mathbb{H}^2}) \approx \min_{i,j,k} d_{\mathrm{H}}^{\mathbb{R}^6, f_k, F}(P_i^H, Q_j)$$

$$= \min_{i,j,k} d_{\mathrm{H}}^{\mathbb{R}^6}(f_k(P_i^H), F(Q_j)), \quad (10)$$

$$d_{\mathrm{GH}}(B_{\mathbb{S}^2}, B_{\mathbb{H}^2}) \approx \min_{i,j,k} d_{\mathrm{H}}^{\mathbb{R}^6, g_k, F}(P_i^S, Q_j)$$

$$= \min_{i,j,k} d_{\mathrm{H}}^{\mathbb{R}^6}(g_k(P_i^S), F(Q_j)), \quad (11)$$

we employ the following functions to embed $B_{\mathbb{E}^2}$ (Euclidean space) and $B_{\mathbb{S}^2}$ (spherical space) into $\mathbb{E}^{6n-6}$. To optimize for f_k, we explore all possible permutations of the basis vectors of $\mathbb{E}^6$: $e_1, e_2, e_3, e_4, e_5, e_6$. For each f_k, we consider two elements e_i, e_j and utilize those dimensions to embed $\mathbb{E}^2$ within $\mathbb{E}^6$. In theory, we should also take into account a small offset represented by $F(\mathbf{0}) = \mathbf{0}$, but in this case, the offset is zero. Moreover, during the optimization process, we incorporate a small vector (with all entries, except one dimension, set to zero) into the mapping function. This vector allows us to translate the plane in different directions, with an offset ranging between -0.5 and 0.5. We consider a total of 100 steps within this range. To optimize for g_k, we follow a similar procedure where we explore all permutations of the basis vectors. However, in this case, we consider three basis vectors instead of two. For each permutation family P_i^S, we not only consider the original permutation but also its negative counterpart $-P_i^S$. We also offset the mapping function as described before. It is important to note that, in practice, $B_{\mathbb{E}^2}$ and $B_{\mathbb{S}^2}$ are discretized as P_i^H and P_i^S, respectively.

Results. The estimated Gromov-Haussdorff distances are presented in Table 1.

Table 1. Estimated Gromov-Hausdorff distances (up to two decimal places) between model spaces and corresponding edge weights in the graph search space.

Comparison Pair	$d_{\mathrm{GH}}(\cdot)$	$w_{(\cdot)}$
$(\mathbb{E}^2, \mathbb{S}^2)$	0.23	4.35
$(\mathbb{E}^2, \mathbb{H}^2)$	0.77	1.30
$(\mathbb{S}^2, \mathbb{H}^2)$	0.84	1.20

5. Applications: Graph Search Space

As outlined in previous sections of this paper, we utilize Cartesian products of model spaces with constant curvature to generate potential latent geometries. Our current emphasis is on constructing a search space specifically tailored for identifying the optimal latent geometry, where *optimal* denotes the latent geometry that exhibits the highest performance for a given downstream task. The search space is represented as a graph, enabling the utilization of established search algorithms. In this setting, we focus on a scenario where the function $\mathfrak{f}(\cdot)$ to be minimized is defined on the graph nodes. The objective of employing a search algorithm is to locate the node corresponding to the minimum value, denoted as $v^* = \arg\min_{v \in \mathcal{V}} \mathfrak{f}(v)$. In our configuration, each node within the graph corresponds to a distinct latent product manifold, and the value assigned to the node represents the validation set performance of a neural network architecture employing that particular latent geometry. To determine edge weights, we utilize the inverse of the Gromov-Hausdorff distance between product manifolds. We denote the edge weight between model spaces $\mathcal{M}_1$ and $\mathcal{M}_2$ as $w_{\mathcal{M}_1, \mathcal{M}_2}$, representing the inverse Gromov-Hausdorff distance. Edge weights are reported in Table 1. Further, the Gromov-Hausdorff distance between manifolds of different dimensions is one, see Appendix B.

To provide additional structure, we suggest imposing connections between only nodes corresponding to product manifolds that differ by one model space. For instance, $\mathbb{E}^2 \times \mathbb{H}^2$ and $\mathbb{E}^2 \times \mathbb{S}^2$ would be connected with edge weighting $w_{\mathbb{H}^2, \mathbb{S}^2}$ while $\mathbb{S}^2 \times \mathbb{S}^2$ and $\mathbb{E}^2 \times \mathbb{E}^2$ would have no connection in the graph. Moreover, the connectivity between product manifolds of different dimensions adheres to a consistent principle. For example, there would exist a connection with a strength of one between $\mathbb{E}^2 \times \mathbb{H}^2$ and $\mathbb{E}^2 \times \mathbb{H}^2 \times \mathbb{S}^2$. However, there would be no connection between $\mathbb{E}^2$ and $\mathbb{E}^2 \times \mathbb{E}^2 \times \mathbb{E}^2$, or between $\mathbb{H}^2$ and $\mathbb{E}^2 \times \mathbb{S}^2$. This construction introduces a sense of directionality to the graph and results in clusters of product manifolds sharing the same dimension. Additionally, it is worth noting that in practice, there are only four distinct edge weights utilized. As an illustration, the connectivity strength between $\mathbb{E}^2 \times \mathbb{H}^2$ and $\mathbb{S}^2 \times \mathbb{H}^2$ is $w_{\mathbb{E}^2, \mathbb{S}^2}$ since $d_{\mathrm{GH}}(\mathbb{E}^2 \times \mathbb{H}^2, \mathbb{S}^2 \times \mathbb{H}^2) = d_{\mathrm{GH}}(\mathbb{E}^2, \mathbb{S}^2) + d_{\mathrm{GH}}(\mathbb{H}^2, \mathbb{H}^2) = d_{\mathrm{GH}}(\mathbb{E}^2, \mathbb{S}^2)$ given that $d_{\mathrm{GH}}(\mathbb{H}^2, \mathbb{H}^2) = 0$. Visual representations of the graph search space can be found in Appendix A.

6. Conclusion

We have described an algorithm for estimating the Gromov-Hausdorff distance between model spaces and product manifolds of these. Our computational approximations of these distances can be leveraged to compare and provide a mathematically grounded sense of *closeness* between possible latent geometry candidates for machine learning applications. In particular, we suggest that a Gromov-Hausdorff-informed graph search space can be constructed and combined with search algorithms.

References

Barbero, F., Bodnar, C., de Ocáriz Borde, H. S., and Liò, P. Sheaf attention networks. In *NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations*, 2022a.

Barbero, F., Bodnar, C., de Ocáriz Borde, H. S., Bronstein, M., Veličković, P., and Liò, P. Sheaf neural networks with connection laplacians, 2022b.

Blanuša, D. Über die Einbettung hyperbolischer Räume in euklidische Räume. *Monatsh. Math.*, 59:217–229, 1955.

Bortoli, V. D., Mathieu, E., Hutchinson, M., Thornton, J., Teh, Y. W., and Doucet, A. Riemannian score-based generative modeling. *ArXiv*, abs/2202.02763, 2022.

Brander, D. Isometric embeddings between space forms. 2003. Master thesis.

Chamberlain, B., Rowbottom, J., Eynard, D., Di Giovanni, F., Dong, X., and Bronstein, M. Beltrami flow and neural diffusion on graphs. *Advances in Neural Information Processing Systems*, 34:1594–1609, 2021.

Chami, I., Ying, R., Ré, C., and Leskovec, J. Hyperbolic graph convolutional neural networks, 2019.

Gallier, J. and Quaintance, J. Differential geometry and lie groups: A computational perspective, 2020.

Gu, A., Sala, F., Gunel, B., and Ré, C. Learning mixed-curvature representations in product spaces. In *International Conference on Learning Representations*, 2018.

Henke, W. and Nettekoven, W. The hyperbolic n-space as a graph in Euclidean $(6n - 6)$-space. *Manuscripta Math.*, 59(1):13–20, 1987.

Hensel, F., Moor, M., and Rieck, B. A. A survey of topological machine learning methods. *Frontiers in Artificial Intelligence*, 4, 2021.

Huang, C.-W., Aghajohari, M., Bose, A., Panangaden, P., and Courville, A. C. Riemannian diffusion models. *ArXiv*, abs/2208.07949, 2022.

Jungeblut, P., Kleist, L., and Miltzow, T. The complexity of the hausdorff distance. In *International Symposium on Computational Geometry*, 2021.

Mathieu, E., Le Lan, C., Maddison, C. J., Tomioka, R., and Teh, Y. W. Continuous hierarchical representations with poincaré variational auto-encoders. *Advances in neural information processing systems*, 32, 2019.

Sáez de Ocáriz Borde, H., Arroyo, A., and Posner, I. Projections of model spaces for latent graph inference. In *ICLR 2023 Workshop on Physics for Machine Learning*, 2023a.

Sáez de Ocáriz Borde, H., Kazi, A., Barbero, F., and Lio, P. Latent graph inference using product manifolds. *The Eleventh International Conference on Learning Representations*, 2023b.

Skopek, O., Ganea, O.-E., and Bécigneul, G. Mixed-curvature variational autoencoders. *arXiv preprint arXiv:1911.08411*, 2019.

Taha, A. A. and Hanbury, A. An efficient algorithm for calculating the exact hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37: 2153–2163, 2015.

Zhang, S., Moscovich, A., and Singer, A. Product manifold learning. In *International Conference on Artificial Intelligence and Statistics*, 2020.

Zhao, D., Zhu, J., and Zhang, B. Latent variables on spheres for autoencoders in high dimensions. *arXiv: Learning*, 2019.

A. Additional figures

In this appendix we provide visualizations of the suggested Gromov-Hausdorff-informed graph search space.
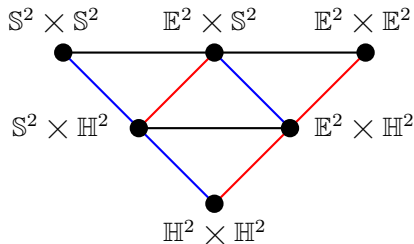

Figure 1. Slice of the graph search space for latent geometries of dimension 4: product manifolds obtained using 2 models spaces of dimension 2. The graph edges are shown in different colours to depict a different degree of connectivity (black: $w_{\mathbb{E}^2,\mathbb{S}^2}$, red: $w_{\mathbb{E}^2,\mathbb{H}^2}$, blue: $w_{\mathbb{S}^2,\mathbb{H}^2}$), this is determined by the inverse of the Gromov-Hausdorff distance between the different product manifolds.
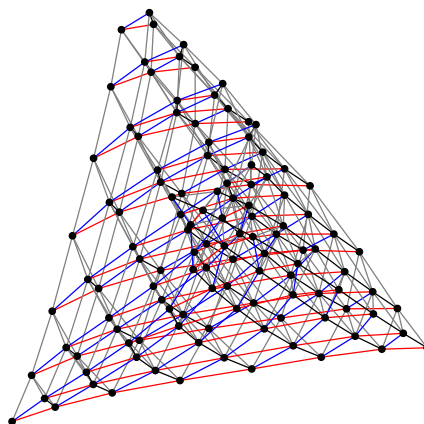

Figure 2. Example graph search space for product manifolds composed of up to seven model spaces. Manifolds of different dimensionality are connected with edges coloured grey. Node labels have been omitted for visual clarity.

B. Upper Bound for the Gromov-Hausdorff Distance

It is important to acknowledge that when considering both A and B as compact sets, there exists a straightforward upper bound for their Gromov-Hausdorff distance, which can be expressed in terms of their respective diameters. The *diameter* of a metric space Y is defined to be $\mathrm{diam}(Y) = \sup_{y,y' \in Y} d_Y(y, y')$. Given $a_0 \in A$ and $b_0 \in b$, we can define the isometric embeddings $f : A \to A \times B$ and $g : B \to A \times B$ given by $f(a) = (a, b_0)$ and $g(b) = (a_0, b)$. It is trivial that $\mathrm{d}_{\mathrm{H}}^{A \times B, f, g}(A, B) \leqslant \max\left(\mathrm{diam}(A), \mathrm{diam}(B)\right)$. Since the triple $(A \times B, f, g)$ belongs to $\mathrm{ES}(A, B)$, we can estimate $\mathrm{d}_{\mathrm{GH}}(A, B) \leqslant \max\left(\mathrm{diam}(A), \mathrm{diam}(B)\right)$. In order to compare $\mathbb{E}^n$, $\mathbb{H}^n$, and $\mathbb{S}^n$, we suggest considering closed balls with a radius of one in each space. Since these balls, regardless of the space, are homogeneous Riemannian manifolds, they are isometric to one another. By estimating or establishing an upper bound for their Gromov-Hausdorff distance, we can effectively compare these spaces. With exactly an analogous argument as before, we can notice that given two compact balls of radius one B and B', of centres x_0 and x_0', we can embed B into $B \times B'$ by the mapping $f : b \mapsto (b, x_0')$. From here, it is obvious to see that $\mathrm{d}_{\mathrm{H}}^{B \times B', f, \mathrm{id}}(B, B \times B') = 1$. Effectively, this observation provides us with the bound $\mathrm{d}_{\mathrm{GH}}(B, B') \leqslant 1$.

We will consider this as the estimation of Gromov-Hausdorff distances for product manifolds that differ solely in one model space, such as $\mathbb{H}^2$ and $\mathbb{S}^2 \times \mathbb{H}^2$.