Taxonomy-aware Dynamic Motion Generation on Hyperbolic Manifolds

Luis Augenstein[*1], Noémie Jaquier[*2], Tamim Asfour[1], and Leonel Rozo[3]
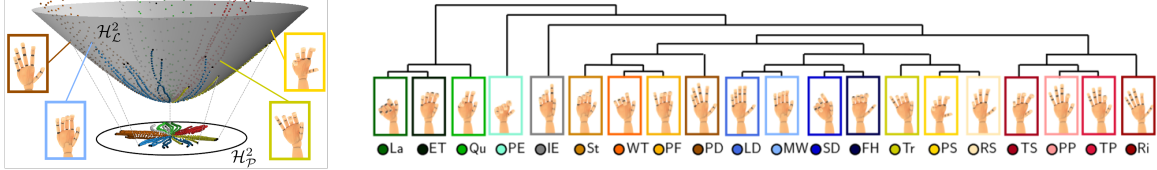
Fig. 1: Illustration of latent trajectories learned by the proposed GPHDM on the Lorentz $\mathbb{H}^2_{\mathcal{L}}$ and Poincaré $\mathbb{H}^2_{\mathcal{P}}$ models of the hyperbolic manifold. The GPHDM preserves the hand grasp taxonomy [1] (*right*) and the temporal dynamics of the demonstrated hand motions.

Abstract— **Human-like motion generation for robots often draws inspiration from biomechanical studies, which often categorize complex human motions into hierarchical taxonomies. While these taxonomies provide rich structural information about how movements relate to one another, this information is frequently overlooked in motion generation models, leading to a disconnect between the generated motions and their underlying hierarchical structure. This paper introduces the Gaussian Process Hyperbolic Dynamical Model (GPHDM), a novel approach that learns latent representations preserving both the hierarchical structure of motions and their temporal dynamics to ensure physical consistency. Our model achieves this by extending the dynamics prior of the Gaussian Process Dynamical Model (GPDM) to the hyperbolic manifold and integrating it with taxonomy-aware inductive biases. Building on this geometry- and taxonomy-aware frameworks, we propose three novel mechanisms for generating motions that are both taxonomically-structured and physically-consistent: two probabilistic recursive approaches and a method based on pullback-metric geodesics. Experiments on generating realistic motion sequences on the hand grasping taxonomy show that the proposed GPHDM faithfully encodes the underlying taxonomy and temporal dynamics, and generates novel physically-consistent trajectories.**

I. INTRODUCTION

Designing robots with human-like capabilities is a long-standing goal in robotics, often drawing inspiration from biomechanics to achieve realistic and functional motions [2]. A critical aspect in this process is analyzing human movements, for which researchers often structure complex actions into hierarchical classifications known as *taxonomies*. These taxonomies, which categorize hand postures [3], full-body poses [4], and manipulation primitives [5], among others, carry a rich hierarchical structure that reflects the relationships among different movements. However, this crucial structural information is frequently ignored in motion generation literature. Recent works [6], [7] have shown that leveraging this structural inductive bias into motion generation models may alleviate the demand of training data, lead to human-like movements, and ultimately generate new motions that comply with the hierarchical structure of a given taxonomy.

Early approaches to generating motions within these taxonomies showed promise but did not directly leverage their hierarchical nature. For example, Romero *et al.* [8] used a Gaussian Process Latent Variable Model (GPLVM) [9], where clusters corresponding to various grasp types of the GRASP taxonomy [3] emerged naturally, despite that the taxonomy's structure was not an explicitly considered during training. They identified the potential for generating new motions via latent space interpolation as potential future work. Separately, Mandery *et al.* [10] abstracted the motion generation problem as a linguistic task, using probabilistic n-gram language models to reproduce whole-body motions based on the types of whole-body poses proposed in [4]. Discrete body poses were represented by words and sequenced using sentences. However, this discrete representation struggled to capture the continuous nature of movement and overlooked the entire hierarchical structure of the associated taxonomy.

To address this gap, recent work by Jaquier *et al.* [6] introduced the Gaussian Process Hyperbolic Latent Variable Model (GPHLVM), which explicitly accounts for the hierarchical structure of taxonomy data. Their key insight was to leverage *hyperbolic geometry* [11], a natural fit for embedding tree-like structures [12], [13], [14], to create a continuous latent representation of the hierarchically-organized taxonomy data. In the GPHLVM, high-dimensional observations (e.g., joint angles of a human body or hand) belonging to the same taxonomy node were embedded closely together, forming distinct clusters in the hyperbolic space. Crucially, they demonstrated that hyperbolic geodesics between these clusters correctly traversed intermediate clusters, mirroring the parent–child relationships of the original taxonomy. These insights confirmed that the GPHLVM could successfully learn a latent space that preserves the hierarchical structure of complex motion data. However, a significant limitation remains: While the GPHLVM can generate novel motions by decoding latent geodesics back to the high-dimensional joint space, some of the resulting motions can be physically impractical. This issue arises because the GPHLVM is trained on data concentrated within the clusters (i.e., static poses) of the taxonomy, leaving data-sparse regions in-between. In these regions, the model lacks information about valid trajectories and thus its predictions revert to non-informative mean, failing to capture the underlying dynamics of target movements.

This paper tackles this challenge. First, we propose the Gaussian Process Hyperbolic Dynamical Model (GPHDM), which learns latent representations that preserve not only the hierarchical structure of motions but also their temporal dynamics to ensure physical consistency (see Fig. 1). To achieve this, we extend the dynamics model of the Gaussian Process Dynamical Model (GPDM) [15] to the hyperbolic manifold and integrate it into the taxonomy-aware framework of GPHLVM. Second, we introduce three novel motion generation mechanisms that are both taxonomically-structured and physically-consistent, namely: two probabilistic recursive approaches, and a pullback-metric geodesics method. We test our approach on the hand grasping taxonomy [1], showing that the proposed GPHDM successfully preserves both the hierarchical structure and temporal dynamics of the data, while allowing us to generate novel physically-consistent motions.

II. BACKGROUND

Riemannian geometry: A Riemannian manifold $\mathcal{M}$ is a smooth manifold equipped with a Riemannian metric, i.e., a smoothly-varying inner product $g_{\boldsymbol{x}}: \mathcal{T}_{\boldsymbol{x}}\mathcal{M} \times \mathcal{T}_{\boldsymbol{x}}\mathcal{M} \to \mathbb{R}$ over tangent spaces $\mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ [16]. When considering coordinates, the Riemannian metric is represented in matrix form as $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\boldsymbol{x}} = \boldsymbol{u}^\top \boldsymbol{G}_{\boldsymbol{x}} \boldsymbol{v}$ with $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{T}_{\boldsymbol{x}}\mathcal{M}$. This metric defines the length of curves in $\mathcal{M}$, leading to the definition of geodesics, defined as locally length-minimizing curves. To operate with Riemannian manifolds, one can leverage the Euclidean tangent spaces and the geodesics. The exponential map $\operatorname{Exp}_{\boldsymbol{x}}(\boldsymbol{u}) = \boldsymbol{y}$ maps $\boldsymbol{u} \in \mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ to a point $\boldsymbol{y} \in \mathcal{M}$, so that $\boldsymbol{y}$ lies on a geodesic starting at $\boldsymbol{x}$ in the direction $\boldsymbol{u}$, and such that the geodesic distance $d_{\mathcal{M}}(\boldsymbol{x}, \boldsymbol{y})$ equals the length of $\boldsymbol{u}$ given by $\|\boldsymbol{u}\|_{\boldsymbol{x}} = \sqrt{\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\boldsymbol{x}}}$. The inverse operation defines the logarithmic map $\operatorname{Log}_{\boldsymbol{x}}(\boldsymbol{y}) = \boldsymbol{u}$. Finally, the parallel transport $\Gamma_{\boldsymbol{x} \to \boldsymbol{y}}(\boldsymbol{u}) = \boldsymbol{v}$ moves a vector $\boldsymbol{u} \in \mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ to $\mathcal{T}_{\boldsymbol{y}}\mathcal{M}$ while preserving the Riemannian inner product.

Optimizing functions $f: \mathcal{M} \to \mathbb{R}$ requires generalizing the definition of gradients. The Riemannian gradient $\operatorname{grad}_{\boldsymbol{x}} f$ of f at $\boldsymbol{x} \in \mathcal{M}$ is the unique tangent vector in $\mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ satisfying $\mathcal{D}_{\boldsymbol{u}} f(\boldsymbol{x}) = \langle \operatorname{grad}_{\boldsymbol{x}} f, \boldsymbol{u} \rangle_{\boldsymbol{x}}$, with $\mathcal{D}_{\boldsymbol{u}} f(\boldsymbol{x})$ the directional derivative of f along $\boldsymbol{u} \in \mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ [17, Chap. 3]. The Riemannian Jacobian $\boldsymbol{J}^{\mathcal{M}}$ of $f: \mathcal{M} \to \mathbb{R}^D$ is composed by the Riemannian gradients for each output dimension, i.e.,

$$\boldsymbol{J}^{\mathcal{M}} = [\operatorname{grad}_{\boldsymbol{x}} f_1 \ \ldots \ \operatorname{grad}_{\boldsymbol{x}} f_D]^\top. \quad (1)$$

Hyperbolic Manifold: The hyperbolic manifold is the only Riemannian manifold with constant negative curvature [11]. It is often represented by either the Poincaré model $\mathbb{H}_{\mathcal{P}}^D$ [18], using local coordinates within the unit ball, or the Lorentz model $\mathbb{H}_{\mathcal{L}}^D$ [19], [13], using Cartesian coordinates to represent the surface embedded in $\mathbb{R}^{D+1}$ (see Fig. 1-*left*). We mostly rely on the latter as it is numerically more stable. The Lorentz model is defined as $\mathbb{H}_{\mathcal{L}}^D = \{\boldsymbol{x} \in \mathbb{R}^{D+1} \mid \langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\mathcal{L}} = -1, x_0 > 0\}$, where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{L}} = \boldsymbol{x}^\top \boldsymbol{G}^{\mathcal{L}} \boldsymbol{y}$ is the Lorentzian inner product with metric $\boldsymbol{G}^{\mathcal{L}} = \operatorname{diag}(-1, 1, ..., 1)$. Under this model, the Riemannian operations of relevance for this work are,

$$d_{\mathbb{H}_{\mathcal{L}}^D}(\boldsymbol{x}, \boldsymbol{z}) = \operatorname{arccosh}(-\langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathcal{L}}), \quad (2)$$

$$\operatorname{Exp}_{\boldsymbol{x}}(\boldsymbol{u}) = \cosh(\|\boldsymbol{u}\|_{\boldsymbol{x}})\boldsymbol{x} + \sinh(\|\boldsymbol{u}\|_{\boldsymbol{x}})\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|_{\boldsymbol{x}}}, \quad (3)$$

$$\operatorname{Log}_{\boldsymbol{x}}(\boldsymbol{z}) = d_{\mathbb{H}_{\mathcal{L}}^D}(\boldsymbol{x}, \boldsymbol{z})\frac{\boldsymbol{z} + \langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathcal{L}}\boldsymbol{x}}{\sqrt{\langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathcal{L}}^2 - 1}}, \quad (4)$$

$$\Gamma_{\boldsymbol{x} \to \boldsymbol{z}}(\boldsymbol{u}) = \boldsymbol{u} + \frac{\langle \boldsymbol{z}, \boldsymbol{u} \rangle_{\mathcal{L}}}{1 - \langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathcal{L}}}(\boldsymbol{x} + \boldsymbol{z}) \quad (5)$$

$$\operatorname{proj}_{\boldsymbol{x}}(\boldsymbol{w}) = \boldsymbol{P}_{\boldsymbol{x}}\boldsymbol{w} = (\boldsymbol{G}^{\mathcal{L}} + \boldsymbol{x}\boldsymbol{x}^\top)\boldsymbol{w}, \quad (6)$$

where $\operatorname{proj}_{\boldsymbol{x}}(\boldsymbol{w})$ projects $\boldsymbol{w} \in \mathbb{R}^{D+1}$ to $\mathcal{T}_{\boldsymbol{x}}\mathbb{H}_{\mathcal{L}}^D$.

Hyperbolic wrapped Gaussian distribution: Working with probabilistic models on Riemannian manifolds requires probability distributions that account for their geometry. We employ the Wrapped Gaussian distribution (WGD), which maps a normal distribution from a Euclidean tangent space onto a Riemannian manifold $\mathcal{M}$. This *wrapping* is performed by a diffeomorphism ψ, which is often the exponential map $\operatorname{Exp}_{\boldsymbol{\mu}}(\cdot)$, centered at the distribution's mean $\boldsymbol{\mu}$ and restricted to a domain where it is injective. Specifically, a WGD is defined as the pushforward measure $\psi_{\#}(\mathcal{N})$ of a Gaussian distribution with density $\mathcal{N}$. The wrapping operation induces a wrapped density function through the change of variables,

$$\mathcal{N}_W = \mathcal{N} \circ \psi^{-1} \cdot |\det(\partial \psi^{-1})|. \quad (7)$$

Given the Lorentz model as the representation of a hyperbolic manifold, the hyperbolic WGD corresponds to,

$$\mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}} = \mathcal{N}(\operatorname{Log}_{\boldsymbol{\mu}}(\boldsymbol{x}); \boldsymbol{0}, \boldsymbol{\Sigma}) \cdot \left(\frac{\|\operatorname{Log}_{\boldsymbol{\mu}}(\boldsymbol{x})\|_{\boldsymbol{\mu}}}{\sinh(\|\operatorname{Log}_{\boldsymbol{\mu}}(\boldsymbol{x})\|_{\boldsymbol{\mu}})} \right)^{D_x - 1}. \quad (8)$$

Gaussian Process Hyperbolic Latent Variable Model (GPHLVM): A GPLVM defines a stochastic mapping from latent variables $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]^\top$, $\boldsymbol{x}_n \in \mathbb{R}^{D_x}$ to observations $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]^\top$, $\boldsymbol{y}_n \in \mathbb{R}^{D_y}$ with $D_x < D_y$ via a non-linear transformation modeled by a Gaussian Process (GP) [9]. A GPLVM defines a generative model

$$p(\boldsymbol{Y}; \boldsymbol{X}, \Theta) = \mathcal{N}(\boldsymbol{Y}; \boldsymbol{0}, k(\boldsymbol{X}, \boldsymbol{X}) + \sigma_{\boldsymbol{y}}^2 \boldsymbol{I}), \quad (9)$$

with prior $\boldsymbol{x}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, where $k: \mathbb{R}^{D_x} \times \mathbb{R}^{D_x} \to \mathbb{R}$ is the GP kernel function that measures the similarity between pairs $(\boldsymbol{x}_n, \boldsymbol{x}_m)$, and Θ denotes the model hyperparameters, i.e., the kernel parameters and the noise variance $\sigma_{\boldsymbol{y}}^2$.

In the hyperbolic setting, the latent variables $\boldsymbol{x}_n \in \mathbb{H}_{\mathcal{L}}^{D_x}$ lie in a hyperbolic space, so the generative mapping is defined via a GPHLVM [6]. Therefore, the kernel function in (9) is replaced by a hyperbolic kernel $k^{\mathbb{H}_{\mathcal{L}}^{D_x}}: \mathbb{H}_{\mathcal{L}}^{D_x} \times \mathbb{H}_{\mathcal{L}}^{D_x} \to \mathbb{R}$. Here, we employ hyperbolic kernels that build on the definition of the heat kernel on Riemannian manifolds [20], [21], [6], which accurately capture the non-Euclidean geometry of the hyperbolic space. Moreover, the latent variables follow a hyperbolic wrapped Gaussian prior $\boldsymbol{x}_n \sim \mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{\mu}_0, \alpha \boldsymbol{I})$, where $\boldsymbol{\mu}_0 = (1, 0, \ldots, 0)^\top$ is the hyperbolic origin and α controls the spread of the latent variables. The GPHLVM latent variables and hyperparameters are inferred via MAP or variational inference similar as in the Euclidean case [6].

Pullback metrics: An immersion $f: \mathcal{X} \to \mathcal{Y}$ from a latent space $\mathcal{X}$ to a Riemannian manifold $\mathcal{Y}$ equipped with a

Riemannian metric g_y induces a pullback metric g_x^P on $\mathcal{X}$ which, for $x \in \mathcal{X}$ and $u, v \in \mathcal{T}_x\mathcal{X}$, is given by [16, Chap.2],

$$g_x^P(u, v) = g_{f(x)}\big(df_x(u), df_x(v)\big). \qquad (10)$$

In coordinates, the pullback metric is given in matrix form by $G_x^P = J^{\mathcal{X}^\top} G_y J^{\mathcal{X}}$, where $J^{\mathcal{X}}$ is the Riemannian Jacobian (1) of f at x. Intuitively, g_x^P evaluates on tangent vectors of $\mathcal{T}_x\mathcal{X}$ by moving them to $\mathcal{T}_{f(x)}\mathcal{Y}$ to compute their inner product. For an immersion $f : \mathcal{X} \to \mathbb{R}^{D_y}$ with Euclidean co-domain, i.e., $G_y = I$, the pullback metric is defined as $G_x^P = J^{\mathcal{X}^\top} J^{\mathcal{X}}$, where $J^{\mathcal{X}}$ equals the Euclidean Jacobian $J = [\frac{\partial f_1}{\partial x} \cdots \frac{\partial f_{D_y}}{\partial x}]^\top \in \mathbb{R}^{D_y \times D_x}$.

III. GAUSSIAN PROCESS HYPERBOLIC DYNAMICAL MODEL

The standard GPLVM, introduced in Sec. II, does not account for the temporal structure of the data. To address this, the GPDM [15] extends the GPLVM by adding a dynamics prior $p(X)$ over the latent variables. This prior, derived under a first-order Markov assumption, incentivizes the latent points to form smooth trajectories. A GPDM often assumes a linear model $f_A(x_t) = \sum_{i=1}^{N_\phi} a_i \phi_i(x_t) = A^\top \phi_t$ for the dynamics. However, we introduce an alternative formulation that models only the offset required to transition from x_t to x_{t+1} as,

$$f_A(x_t) = x_t + \sum_{i=1}^{N_\phi} a_i \phi_i(x_t) = x_t + A^\top \phi_t. \qquad (11)$$

This can be equivalently written as $\tilde{x}_{t+1} := x_{t+1} - x_t = A^\top \phi_t + \epsilon_t$, with $\epsilon_t \sim \mathcal{N}(0, \Sigma_x)$. While this modification may seem unnecessary in the Euclidean case, it provides a clearer connection for comparing the classical GPDM with our method, which places a hyperbolic dynamics prior on the latent space of the GPHLVM, as explained shortly.

A. Hyperbolic Tangent Vectors in Local Coordinates

The operations introduced for the Lorentz model of the hyperbolic manifold in Sec. II represent tangent vectors as (D_x+1)-dimensional elements embedded in the ambient space. However, the intrinsic dimension of the hyperbolic manifold and its tangent space is D_x, which leads to degenerated covariance matrices $\Sigma \in \mathcal{T}_x\mathbb{H}_{\mathcal{L}}^{D_x}$, e.g., in the hyperbolic WGD (8), since the eigenvalue along the eigenvector orthogonal to $\mathcal{T}_x\mathbb{H}_{\mathcal{L}}^{D_x}$ is always 0. In this paper, we apply a local change of coordinates to represent tangent vectors and covariances intrinsically as D_x-dimensional elements, thus avoiding issues related to degenerated covariances in the WGD akin to [22].

We first define a canonical set of D_x basis vectors $V_{\mu_0} = \begin{pmatrix} e_2 & \cdots & e_{D+1} \end{pmatrix} = \begin{pmatrix} 0 & I_{D_x} \end{pmatrix}^\top \in \mathbb{R}^{(D_x+1) \times D_x}$ in the tangent space $\mathcal{T}_{\mu_0}\mathbb{H}_{\mathcal{L}}^{D_x}$ at the origin $\mu_0 \in \mathbb{H}_{\mathcal{L}}^{D_x}$. A set of basis vectors V_x at any point $x \in \mathbb{H}_{\mathcal{L}}^{D_x}$ can then be obtained by parallel transporting the canonical basis vectors as,

$$V_x = \begin{pmatrix} \Gamma_{\mu_0 \to x}(e_2) & \cdots & \Gamma_{\mu_0 \to x}(e_{D+1}) \end{pmatrix}. \qquad (12)$$

Note that V_x consists of orthonormal basis vectors on $\mathcal{T}_x\mathbb{H}_{\mathcal{L}}^{D_x}$, i.e., $V_x^\top G^{\mathcal{L}} V_x = I_{D_x}$ and $P_x = V_x V_x^\top$. We obtain tangent space vectors u and matrices $\Sigma \in \mathcal{T}_x\mathbb{H}_{\mathcal{L}}^{D_x}$ in local coordinates

using V_x as $\tilde{u} = V_x^\top G^{\mathcal{L}} p$ and $\tilde{\Sigma} = V_x^\top G^{\mathcal{L}} \Sigma G^{\mathcal{L}} V_x$. Tangent space elements in ambient space coordinates are obtained via the inverse operations $u = V_x \tilde{u}$ and $\Sigma = V_x^\top \tilde{\Sigma} V_x$.

B. Hyperbolic Dynamics Prior

To derive the GPHDM, we first generalize the dynamics prior of the GPDM to the hyperbolic manifold. To do so, we define a first-order hyperbolic dynamics model by extending the Markov assumption to the hyperbolic manifold as,

$$x_{t+1} = \text{Exp}_{f_A(x_t)}(V_{f_A(x_t)}\tilde{\epsilon}_t) \text{ with } \tilde{\epsilon}_t \sim \mathcal{N}(0, \tilde{\Sigma}_x), \quad (13)$$

where we define the noise $\tilde{\epsilon}_t \in \mathbb{R}^{D_x}$ in local coordinates and transform it onto the tangent space $\mathcal{T}_{f_A(x_t)}\mathbb{H}_{\mathcal{L}}^{D_x}$ via the basis-vector matrix $V_{f_A(x_t)} \in \mathbb{R}^{(D_x+1) \times D_x}$. Similar to the GPDM in (11), we define f_A via N_ϕ nonlinear basis functions as,

$$f_A(x_t) = \text{Exp}_{x_t}(V_{x_t} A^\top \phi_t), \qquad (14)$$

where $\phi_t = \phi(x_t) = \begin{bmatrix} \phi_1(x_t) & \cdots & \phi_{N_\phi}(x_t) \end{bmatrix}^\top \in \mathbb{R}^{N_\phi}$ is a vector of basis functions and $A \in \mathbb{R}^{N_\phi \times D_x}$ is a weight matrix. Notice that we again perform a change of basis to obtain the Lorentz representation $V_{x_t} A^\top \phi_t \in \mathcal{T}_{x_t}\mathbb{H}_{\mathcal{L}}^{D_x}$ from the vector $A^\top \phi_t \in \mathbb{R}^{D_x}$ represented in local coordinates. In the Euclidean case, we recover the first-order Markov dynamics (11).

The first-order hyperbolic Markov dynamics (13) lead to the following transition probability,

$$p(x_{t+1} \mid x_t, A) = \mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(x_{t+1}; f_A(x_t), \Sigma_{x_t}) \qquad (15)$$
$$= \mathcal{N}(\text{Log}_{x_t}(x_{t+1}); V_{x_t} A^\top \phi_t, \Sigma_{x_t}) \, r_t$$
$$= \mathcal{N}(\tilde{x}_{t+1}; A^\top \phi_t, \tilde{\Sigma}_x) \, r_t, \qquad (16)$$

where the WGD $\mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}$ is defined as in (7) with a pushforward $\psi = \text{Exp}_{x_t}(\cdot)$, resulting into the change of volume $r_t = |\det \partial \psi^{-1}| = \left(\frac{\|\text{Log}_{x_t}(x_{t+1})\|_{x_t}}{\sinh(\|\text{Log}_{x_t}(x_{t+1})\|_{x_t})}\right)^{D_x - 1}$, and noise covariance matrix $\Sigma_{x_t} = V_{x_t} \tilde{\Sigma}_x V_{x_t}^\top \in \mathcal{T}_{x_t}\mathbb{H}_{\mathcal{L}}^{D_x}$ obtained from the common local covariance $\tilde{\Sigma}_x \in \mathbb{R}^{D_x \times D_x}$. Equation (16) represents the Gaussian distribution in local coordinates with $\tilde{x}_{t+1} = V_{x_t}^\top G^{\mathcal{L}} \text{Log}_{x_t}(x_{t+1}) \in \mathbb{R}^{D_x}$.

We derive the hyperbolic dynamics prior $p(X)$ by marginalizing out the parameters A, similar to the Euclidean case. For a trajectory X of N latent variables $x_1, ..., x_N \in \mathbb{H}_{\mathcal{L}}^{D_x}$ the hyperbolic dynamics prior is given as,

$$p(X) = \int p(X \mid A) \, p(A) \, dA. \qquad (17)$$

Incorporating the Markov probability transition (16) leads to,

$$p(X) = p(x_1) \int \prod_{t=1}^{N-1} p(x_{t+1} \mid x_t, A) \, p(A) \, dA$$
$$\overset{(16)}{=} p(x_1) \int \prod_{t=1}^{N-1} \mathcal{N}(\tilde{x}_{t+1}; A^\top \phi_t, \tilde{\Sigma}_x) \, r_t \, p(A) \, dA.$$

We leverage the properties of the Euclidean Gaussian distributions to reduce the expression as follows,

$$p(X) = p(x_1) \prod_{d=1}^{D_x} \int \prod_{t=1}^{N-1} p(\tilde{x}_{t+1,d}; \phi_t^\top A_d, \sigma_{x,d}^2) \, r_t \, p(A_d) \, dA_d$$

$$= p(\boldsymbol{x}_1) \prod_{t=1}^{N-1} r_t \prod_{d=1}^{D_x} \mathcal{N}(\tilde{X}_{2:N,d}; \mathbf{0}, \boldsymbol{\Phi}\boldsymbol{\Phi}^\top + \sigma_{x,d}^2 \boldsymbol{I}_{N-1})$$

$$= p(\boldsymbol{x}_1) \prod_{t=1}^{N-1} r_t \prod_{d=1}^{D_x} \mathcal{N}(\tilde{X}_{2:N,d}; \mathbf{0}, \boldsymbol{K}_{X,d} + \sigma_{x,d}^2 \boldsymbol{I}_{N-1})$$

$$= p(\boldsymbol{x}_1) \prod_{t=0}^{N-1} r_t \mathcal{N}(\tilde{X}_{2:N}; \mathbf{0}, \boldsymbol{K}_X + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}}),$$

where $\boldsymbol{A}_d$ is the d-th column of $\boldsymbol{A}$, $\tilde{X}_{2:N} = [\tilde{\boldsymbol{x}}_2, \dots, \tilde{\boldsymbol{x}}_N]^\top$, and $\boldsymbol{K}_X = \boldsymbol{\Phi}\boldsymbol{\Phi}^\top$ is the $(N-1)D_x \times (N-1)D_x$ kernel matrix constructed from the $(N-1)$ basis functions associated to $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N-1}$. Also, we used $p(\boldsymbol{A}_d) = \mathcal{N}(\boldsymbol{A}_d; \mathbf{0}, \boldsymbol{I}_{N_\phi})$ and

$$\prod_{t=1}^{N-1} p(\tilde{x}_{t+1,d}; \boldsymbol{\phi}_t^\top \boldsymbol{A}_d, \sigma_{x,d}^2) = \mathcal{N}(\tilde{X}_{2:N,d}; \boldsymbol{\Phi}\boldsymbol{A}_d, \sigma_{x,d}^2 \boldsymbol{I}_{N-1}).$$

Finally, we obtain the hyperbolic dynamics prior as,

$$p(\boldsymbol{X}) = \mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{x}_1; \boldsymbol{\mu}_0, \boldsymbol{V}_{\boldsymbol{\mu}_0} \boldsymbol{V}_{\boldsymbol{\mu}_0}^\top) \tag{18}$$
$$\mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{X}_{2:N}; \boldsymbol{X}_{1:N-1}, \boldsymbol{V}_{\boldsymbol{X}}(\boldsymbol{K}_X + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}})\boldsymbol{V}_{\boldsymbol{X}}^\top),$$

where $p(\boldsymbol{x}_1)$ was set as an isotropic hyperbolic WGD, and $\boldsymbol{V}_{\boldsymbol{X}} = \text{blockdiag}(\boldsymbol{V}_{\boldsymbol{x}_1}, ..., \boldsymbol{V}_{\boldsymbol{x}_{N-1}}) \in \mathbb{R}^{(N-1)(D_x+1) \times (N-1)D_x}$.

The kernel k_X, used as $\boldsymbol{K}_X = k_X(\boldsymbol{X}_{1:N-1}, \boldsymbol{X}_{1:N-1})$ in (18), is key for the hyperbolic dynamics prior as it provides a notion of similarity between consecutive hyperbolic latent variables. Therefore, we must employ hyperbolic kernels that accurately capture the geometry of the hyperbolic space. Here, we leverage the hyperbolic squared exponential (SE) kernels $k^{\mathbb{H}_{\mathcal{L}}^{D_x}}$ [20], [21], [6], and define $k_X = \boldsymbol{I}_{D_x} \otimes k^{\mathbb{H}_{\mathcal{L}}^{D_x}}$ as a multivariate kernel with shared values across output dimensions with $\otimes$ denoting the Kronecker product. The hyperbolic SE kernels in dimensions 2 and 3 are defined as,

$$k^{\mathbb{H}_{\mathcal{L}}^2}(\boldsymbol{x}, \boldsymbol{z}) = \frac{\sigma^2}{C_\infty} \int_\rho^\infty \frac{s e^{-s^2/(2\kappa^2)}}{(\cosh(s) - \cosh(\rho))^{1/2}} ds, \tag{19}$$

$$k^{\mathbb{H}_{\mathcal{L}}^3}(\boldsymbol{x}, \boldsymbol{z}) = \frac{\sigma^2}{C_\infty} \frac{\rho}{\sinh \rho} e^{-\rho^2/(2\kappa^2)}, \tag{20}$$

where $\rho = d_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{x}, \boldsymbol{z})$ is the geodesic distance between $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{H}_{\mathcal{L}}^{D_x}$, κ and σ^2 are the kernel lengthscale and variance, and C_∞ is a normalizing constant. To the best of our knowledge, no closed form expression for $\mathbb{H}_{\mathcal{L}}^2$ is known, therefore we use the Monte-Carlo approximation introduced in [6]. Notice that higher dimensional hyperbolic SE kernels are expressed as derivatives of the kernels (19)-(20) [20].

C. The GPHDM

Our GPHDM extends the GPDM to hyperbolic spaces by combining the GPHLVM with the hyperbolic dynamics prior (18). A GPHDM defines a generative model from low-dimensional latent variables $\boldsymbol{X} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_N]^\top$, $\boldsymbol{x}_t \in \mathbb{H}_{\mathcal{L}}^{D_x}$ to high-dimensional trajectories $\boldsymbol{Y} = [\boldsymbol{y}_1, \dots, \boldsymbol{y}_N]^\top$, $\boldsymbol{y}_t \in \mathbb{R}^{D_y}$, and it is formally described as,

$$p(\boldsymbol{Y}; \boldsymbol{X}, \Theta) = \mathcal{N}(\boldsymbol{Y}; \mathbf{0}, \boldsymbol{K}_Y + \boldsymbol{\Sigma}_{\boldsymbol{y}}), \tag{21}$$

$$p(\boldsymbol{X}; \Theta) = \mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{x}_1; \boldsymbol{\mu}_0, \boldsymbol{V}_{\boldsymbol{\mu}_0} \boldsymbol{V}_{\boldsymbol{\mu}_0}^\top) \tag{22}$$
$$\mathcal{N}_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{X}_{2:N}; \boldsymbol{X}_{1:N-1}, \boldsymbol{V}_{\boldsymbol{X}}(\boldsymbol{K}_X + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}})\boldsymbol{V}_{\boldsymbol{X}}^\top),$$

where $p(\boldsymbol{Y}; \boldsymbol{X}, \Theta)$ is the model likelihood, $p(\boldsymbol{X}; \Theta)$ is the hyperbolic dynamics prior (18) from Sec. III-B, and Θ denotes the model hyperparameters, i.e., the parameters of the hyperbolic kernels k_Y and k_X, and the noise variances $\boldsymbol{\Sigma}_{\boldsymbol{y}}$ and $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}}$.

As in the Euclidean case, training a GPHDM consists of optimizing the latent variables $\boldsymbol{x}_t \in \mathbb{H}_{\mathcal{L}}^{D_x}$ along with the parameters Θ by maximizing the log posterior of the model,

$$\boldsymbol{X}, \Theta = \operatorname*{argmax}_{\boldsymbol{X}, \Theta} \beta_1 \log p(\boldsymbol{Y}; \boldsymbol{X}, \Theta) + \beta_2 \log p(\boldsymbol{X}; \Theta), \tag{23}$$

where β_1, β_2 balance the trade-off between the likelihood and dynamic prior. Analogous to its Euclidean counterpart, the hyperbolic dynamics prior encourages consecutive latent points $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$, to be close in the hyperbolic latent space, therefore promoting smooth latent trajectories. Optimizing the GPHDM requires accounting for the hyperbolic geometry of the latent variables. To this end, we employ Riemannian optimization methods [17]. In this paper, we use Riemannian Adam [23] implemented in Geoopt [24] to optimize (23).

Finally, back constraints, introduced for the GPLVM in [25], define the latent variables as a function of the observations via an encoder function $x_{t,d} = g_d(\boldsymbol{Y}; w_d)$ with parameters $\{w_d\}_{d=1}^{D_x}$. This method facilitates the post-training incorporation of new observations into the latent space while preserving local similarities among the resulting embeddings. To ensure that latent variables lie on the hyperbolic manifold, we use a similar back-constraints mapping as in [6], namely,

$$\boldsymbol{x}_t = \text{Exp}_{\boldsymbol{\mu}_0}(\check{\boldsymbol{x}}_t) \text{ with } \check{x}_{t,d} = \sum_{m=1}^{N} w_{d,m} k^{\mathbb{R}^{D_y}}(\boldsymbol{y}_t, \boldsymbol{y}_m).$$

In this case, the loss (23) is optimized for $\{w_d\}_{d=1}^{D_x}$ and Θ.

D. Incorporating Taxonomy Knowledge

Our GPHDM allows us to learn smooth hyperbolic latent trajectories of high-dimensional motions. When these observations are associated with a robotics taxonomy, the latent variables must preserve the taxonomy's graph structure. This means the hyperbolic distance between pairs of latent variables should match their corresponding graph distances. We enforce this property by introducing the graph-distance information as inductive bias during the GPHDM training.

As shown in [26], the GPLVM latent space can be modified by adding priors of the form $p(\boldsymbol{X}) \propto e^{-\phi(\boldsymbol{X})/\sigma_\phi^2}$, where $\phi(\boldsymbol{X})$ is a function to minimize. This is equivalent to augmenting the GPLVM loss with a regularization term $-\beta_3 \phi(\boldsymbol{X})$. As in [6], we augment the GPHDM loss with such a regularizer aiming at preserving the distances of the taxonomy graph. We define $\phi(\boldsymbol{X})$ as the stress loss,

$$\ell_{\text{stress}}(\boldsymbol{X}) = \sum_{i<j} \left(d_{\mathbb{G}}(c_i, c_j) - d_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \right)^2, \tag{24}$$

where c_i denotes the taxonomy node to which the observation $\boldsymbol{y}_i$ belongs, and $d_{\mathbb{G}}$, $d_{\mathbb{H}_{\mathcal{L}}^{D_x}}$ are the taxonomy graph distance and the geodesic distance on $\mathbb{H}_{\mathcal{L}}^{D_x}$, respectively. The stress loss (24) aims at preserving all the taxonomy graph distances in the latent space $\mathbb{H}_{\mathcal{L}}^{D_x}$. However, for trajectories that transition between taxonomy classes, the intermediate trajectory

points are not assigned to any taxonomy node. In such cases, we adapt the computation by applying the stress loss (24) only to the start and end points of each trajectory, which are anchored to specific nodes. The hyperbolic dynamics prior then ensures that intermediate trajectory points form a smooth path between these points in the latent space $\mathbb{H}_{\mathcal{L}}^{D_x}$, thereby preserving the taxonomy graph globally.

IV. DYNAMIC MOTION GENERATION

The generation of new high-dimensional motions via the GPHDM requires the creation of novel latent trajectories $\boldsymbol{X}^*$ in $\mathbb{H}_{\mathcal{L}}^{D_x}$. Here we introduce three strategies to accomplish this on the hyperbolic manifold. First, we adapt the mean prediction and conditional optimization for recursive trajectory generation of the GPDM [15] to the hyperbolic setting. Then, we propose a new approach that generates trajectories as geodesics on the pullback metric of the learned model.

A. Recursive motion generation

In the Euclidean case, the mean prediction method constructs a latent trajectory sequentially. The prediction for the next step $\boldsymbol{x}_{t+1}^*$ is the mean of the conditional distribution at the current step $\boldsymbol{x}_t^*$. This is straightforward as the conditional of the Gaussian distribution is also a Gaussian, making the mean analytically available. However, this convenient property does not hold for the hyperbolic WGD as the conditional of a WGD is not necessarily a WGD [27]. Consequently, its mean is analytically intractable. To overcome this, we leverage a parallel from the Euclidean case: for a Gaussian distribution, the mean and the Maximum Likelihood Estimate (MLE) are equivalent. We therefore propose to find the MLE of the hyperbolic conditional distribution to determine the next step in the trajectory as follows,

$$\boldsymbol{x}_{t+1}^* = \underset{\boldsymbol{x}_{t+1}^*}{\operatorname{argmax}} \, \mathcal{N}(\tilde{\boldsymbol{x}}_{t+1}^*; \tilde{\boldsymbol{m}}^*, \tilde{\boldsymbol{\Sigma}}^*) \, r_t, \quad \text{with} \quad (25)$$

$$\tilde{\boldsymbol{x}}_{t+1}^* = \boldsymbol{V}_{\boldsymbol{x}_t^*}^\top \boldsymbol{G}^{\mathcal{L}} \mathrm{Log}_{\boldsymbol{x}_t^*}(\boldsymbol{x}_{t+1}^*)$$
$$\tilde{\boldsymbol{m}}^* = \boldsymbol{L}^* \boldsymbol{V}_{\boldsymbol{X}}^\top \boldsymbol{G}^{\mathcal{L}} \mathrm{Log}_{\boldsymbol{X}_{1:N-1}}(\boldsymbol{X}_{2:N}),$$
$$\tilde{\boldsymbol{\Sigma}}^* = k_{\boldsymbol{X}}(\boldsymbol{x}_t^*, \boldsymbol{x}_t^*) - \boldsymbol{L}^* k_{\boldsymbol{X}}(\boldsymbol{X}_{1:N-1}, \boldsymbol{x}_t^*),$$
$$\boldsymbol{L}^* = k_{\boldsymbol{X}}(\boldsymbol{x}_t^*, \boldsymbol{X}_{1:N-1})(\boldsymbol{K}_{\boldsymbol{X}} + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}})^{-1},$$

and r_t is the change of volume defined in Sec. III-B. We solve the optimization problem in (25) using Riemannian Adam [23]. For quick convergence, we initialize the problem at $\boldsymbol{x}_{t+1}^* = \mathrm{Exp}_{\boldsymbol{x}_t^*}(\boldsymbol{V}_{\boldsymbol{x}_t^*} \tilde{\boldsymbol{m}}^*)$, which maximizes the Euclidean Gaussian component of the objective (25). The final solution $\boldsymbol{x}_{t+1}^*$ is usually located in the vicinity of this initial point.

A key limitation of the mean prediction approach, analogous to the Euclidean case, is its inability to specify a desired goal point for the new latent trajectory. To address this, we extend the conditional optimization approach of [15] to the hyperbolic manifold. This method allows us to specify start, goal, and intermediate points, from which it interpolates the remaining trajectory segments according to learned dynamics prior. This approach optimizes the full conditional distribution, mirroring the objective defined for the Euclidean case,

$$\boldsymbol{X}^* = \underset{\boldsymbol{X}^*}{\operatorname{argmax}} \, p(\boldsymbol{Y}^* \mid \boldsymbol{X}^*, \boldsymbol{Y}, \boldsymbol{X}, \Theta) \, p(\boldsymbol{X}^* \mid \boldsymbol{X}, \Theta) \quad (26)$$

$$\text{s.t.} \quad \boldsymbol{Y}^* = k_{\boldsymbol{Y}}(\boldsymbol{X}^*, \boldsymbol{X})(\boldsymbol{K}_{\boldsymbol{Y}} + \boldsymbol{\Sigma}_{\boldsymbol{Y}})^{-1} \boldsymbol{Y} \, .$$

The main difference is that the conditional dynamics prior corresponds to its hyperbolic version, i.e.,

$$p(\boldsymbol{X}^*; \boldsymbol{X}) = \mathcal{N}(\tilde{\boldsymbol{X}}_{2:M}^*; \tilde{\boldsymbol{m}}^*, \tilde{\boldsymbol{\Sigma}}^*) \prod_{t=1}^{M-1} r_t, \quad \text{with} \quad (27)$$

$$\tilde{\boldsymbol{X}}_{2:M}^* = \boldsymbol{V}_{\boldsymbol{X}^*}^\top \boldsymbol{G}^{\mathcal{L}} \mathrm{Log}_{\boldsymbol{X}_{1:M-1}^*}(\boldsymbol{X}_{2:M}^*),$$
$$\tilde{\boldsymbol{m}}^* = \boldsymbol{L}^* \boldsymbol{V}_{\boldsymbol{X}}^\top \boldsymbol{G}^{\mathcal{L}} \mathrm{Log}_{\boldsymbol{X}_{1:N-1}}(\boldsymbol{X}_{2:N}),$$
$$\tilde{\boldsymbol{\Sigma}}^* = \boldsymbol{K}_{\boldsymbol{X}}^* - \boldsymbol{L}^* k_{\boldsymbol{X}}(\boldsymbol{X}_{1:N-1}, \boldsymbol{X}_{1:M-1}^*), \quad \text{and}$$
$$\boldsymbol{L}^* = k_{\boldsymbol{X}}(\boldsymbol{X}_{1:M-1}^*, \boldsymbol{X}_{1:N-1})(\boldsymbol{K}_{\boldsymbol{X}} + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}})^{-1}.$$

We solve the optimization problem (26) using Riemannian Adam [23]. To speed up convergence, we initialize the trajectory segments of $\boldsymbol{X}^*$ as hyperbolic geodesics.

B. Pullback-metric geodesic motion generation

From a Riemannian perspective, motion generation can be cast as geodesics computation [6]. However, geodesics derived from the intrinsic metric of the hyperbolic manifold often risk traversing regions of low data density, where the model's uncertainty is high [7]. To overcome this, we endow the latent space with a pullback metric induced by the stochastic mapping of the GP (see Sec.II). By computing pullback-metric geodesics, we can generate paths confined to the learned data manifold, ensuring that the resulting trajectories comply with the data distribution. Formally, the immersion f represented by the GP is stochastic, i.e., f induces a distribution over its Jacobian $\boldsymbol{J}$, which itself induces a distribution over the pullback metric [28].

In the Euclidean case, Tosi *et al.* [28] assumed that the probability over $\boldsymbol{J}$ follows a Gaussian distribution, and has independent rows $\boldsymbol{J}_d \in \mathbb{R}^{D_x}$, each with its own mean but shared covariance matrix. Thus, the Jacobian distribution is,

$$p(\boldsymbol{J}) = \prod_{d=1}^{D_y} \mathcal{N}(\boldsymbol{J}_d \mid \boldsymbol{\mu}_{\boldsymbol{J}_d}, \boldsymbol{\Sigma}_{\boldsymbol{J}}) \, . \quad (28)$$

The corresponding metric tensor follows a non-central Wishart distribution [29], $p(\boldsymbol{G}_{\boldsymbol{x}}^{\mathrm{P},\mathbb{R}}) = \mathcal{W}_{D_x}(D_y, \boldsymbol{\Sigma}_{\boldsymbol{J}}, \mathbb{E}[\boldsymbol{J}]^\top \mathbb{E}[\boldsymbol{J}])$, from which the expected metric tensor is computed as $\mathbb{E}[\boldsymbol{G}_{\boldsymbol{x}}^{\mathrm{P},\mathbb{R}}] = \mathbb{E}[\boldsymbol{J}]^\top \mathbb{E}[\boldsymbol{J}] + D_y \boldsymbol{\Sigma}_{\boldsymbol{J}}$. In the hyperbolic case, Augenstein *et al.* [7] follow a similar strategy. In this context, our GPHDM is a stochastic mapping of the form $f : \mathbb{H}_{\mathcal{L}}^{D_x} \to \mathbb{R}^{D_y}$ (as the GPLHVM in [6]), whose Jacobian is formed as in (1). In this case, the Jacobian distribution is,

$$p(\boldsymbol{J}^{\mathcal{L}}) = \prod_{d=1}^{D_y} \mathcal{N}(\boldsymbol{J}_d^{\mathcal{L}}; \boldsymbol{P}_{\boldsymbol{x}} \boldsymbol{\mu}_{\boldsymbol{J}_d}, \boldsymbol{P}_{\boldsymbol{x}} \boldsymbol{\Sigma}_{\boldsymbol{J}} \boldsymbol{P}_{\boldsymbol{x}}^\top) \, . \quad (29)$$

where $\boldsymbol{P}_{\boldsymbol{x}}$ is the projector operator defined in (6). As $\boldsymbol{G}_{\boldsymbol{x}}^{\mathrm{P},\mathbb{R}}$, the metric tensor $\boldsymbol{G}_{\boldsymbol{x}}^{\mathrm{P},\mathcal{L}}$ follows a non-central Wishart distribution $p(\boldsymbol{G}_{\boldsymbol{x}}^{\mathrm{P},\mathcal{L}}) = \mathcal{W}_{D_x}(D_y, \boldsymbol{\Sigma}_{\boldsymbol{J}}^{\mathcal{L}}, \mathbb{E}[\boldsymbol{J}^{\mathcal{L}}]^\top \mathbb{E}[\boldsymbol{J}^{\mathcal{L}}])$ with $\boldsymbol{\Sigma}_{\boldsymbol{J}}^{\mathcal{L}} = \boldsymbol{P}_{\boldsymbol{x}} \boldsymbol{\Sigma}_{\boldsymbol{J}} \boldsymbol{P}_{\boldsymbol{x}}^\top$, leading to the expected metric [7],

$$\mathbb{E}[\boldsymbol{G}_{\boldsymbol{x}^*}^{\mathrm{P},\mathcal{L}}] = \boldsymbol{P}_{\boldsymbol{x}^*}(\boldsymbol{\mu}_{\boldsymbol{J}}^\top \boldsymbol{\mu}_{\boldsymbol{J}} + D_y \boldsymbol{\Sigma}_{\boldsymbol{J}}) \boldsymbol{P}_{\boldsymbol{x}^*}^\top \, . \quad (30)$$

The pullback metric (30) allows us to compute geodesics that reflect the geometry of both the intrinsic hyperbolic

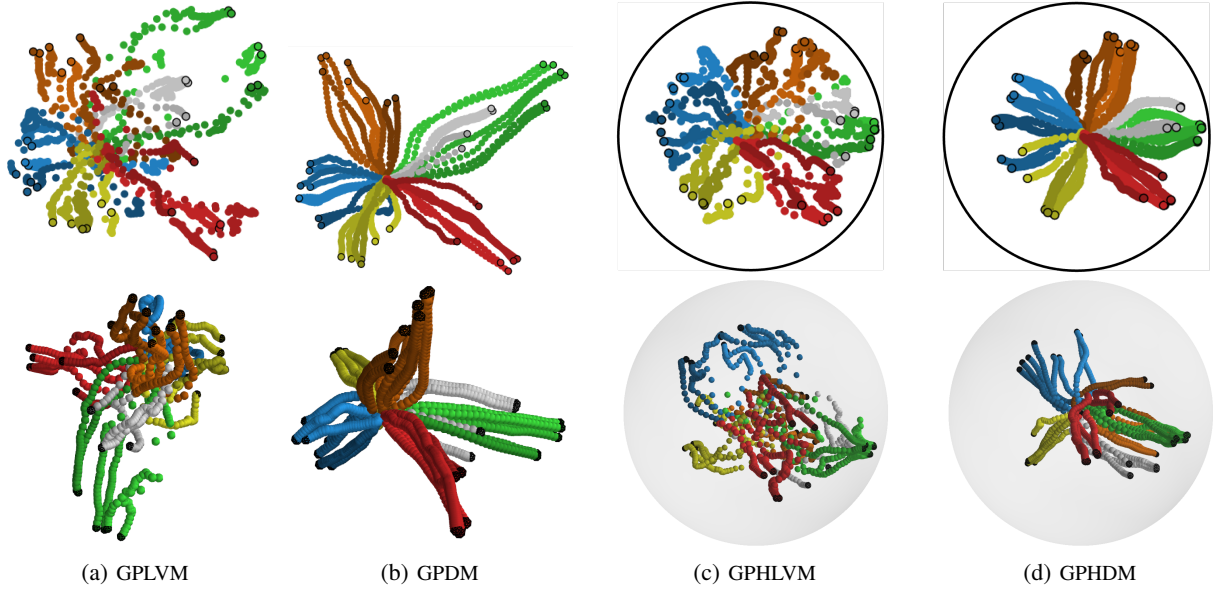(a) GPLVM   (b) GPDM   (c) GPHLVM   (d) GPHDM

Fig. 2: Embeddings of hand grasps colored according to the grasp class of the last trajectory point with colors matching those of Fig. 1. The **top** and **bottom** rows show 2- and 3-dimensional latent spaces, respectively.

space and the learned data distribution. These geodesics are computed by minimizing the curve length or, equivalently, the curve energy E with respect to the pullback metric. For a geodesic discretized into a sequence of M points $\boldsymbol{x}_i \in \mathbb{H}_{\mathcal{L}}^{D_x}$, this problem reduces to minimizing,

$$E = \sum_{i=0}^{M-2} \boldsymbol{v}_i^{\mathsf{T}} \boldsymbol{G}_{\boldsymbol{x}_i}^{\mathrm{P},\mathcal{L}} \boldsymbol{v}_i, \quad \text{with} \quad \boldsymbol{v}_i = \mathrm{Log}_{\boldsymbol{x}_i}(\boldsymbol{x}_{i+1}). \quad (31)$$

Following [7], we regularize the objective (31) with the spline energy $E_{\mathrm{spline}} \approx \sum_{i=1}^{M-2} d_{\mathbb{H}_{\mathcal{L}}^{D_x}}(\boldsymbol{x}_i, \bar{\boldsymbol{x}}_i)^2$, where $\bar{\boldsymbol{x}}_i = \mathrm{Exp}_{\boldsymbol{x}_{i-1}}(\frac{1}{2}\mathrm{Log}_{\boldsymbol{x}_{i-1}}(\boldsymbol{x}_{i+1}))$ represents the geodesic midpoint between $\boldsymbol{x}_{i-1}$ and $\boldsymbol{x}_{i+1}$. This regularization term prevents uneven spacing of the discrete points along the geodesics. Finally, our geodesic optimization problem is,

$$\min_{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{M-1}} E + \lambda E_{\mathrm{spline}}, \quad (32)$$

with λ balancing the regularization. As for the recursive motion generation in Sec.IV-A, we employ Riemannian Adam [23] to solve our geodesic optimization problem.

V. EXPERIMENTS

We test the proposed GPHDM and dynamic motion generation strategies on data associated with a hand grasp taxonomy [1] that organizes common grasp types [3] into a tree structure based on their muscular and kinematic properties (see Fig. 1). We use a dataset from the KIT whole-body motion database [30] consisting of 38 motions of 19 common grasp types obtained from recordings of humans grasping different objects. For each motion, a subject (ID 2122 or 2123) reaches out from an initial resting pose to grasp an object placed on a table. The dataset consists of trajectories $\boldsymbol{Y} \in \mathbb{R}^{N \times 24}$ representing the temporal evolution of the 24 degrees of freedom of the wrist and fingers. We preprocessed the recorded data by: (1) Applying a low-pass filter to remove high-frequency noise; (2) Trimming the start and end of each

trajectory to keep only the motion from the initial resting pose to the grasp completion, that we define as the time when the grasped object is first moved by the human; (3) Subsampling the trajectories; and (4) Centering the data to allow for the use of a zero mean function in the GPLVMs. After preprocessing, each trajectory is composed of 30 to 40 data points for a total of $N = 1321$ data points.

A. Hyperbolic Embeddings of Taxonomy-structured Motions

We embed the trajectories associated to the aforementioned grasp taxonomy into 2- and 3-dimensional hyperbolic and Euclidean spaces using back-constrained GPHLVM, GPHDM, GPLVM, and GPDM. We initialize the latent variables of all models by minimizing the stress (24) of the start and end points of each trajectory with respect to their associated taxonomy nodes, using the hyperbolic and Euclidean distance for the respective models. Then, we initialize the intermediate points of each trajectory as equally spaced along the hyperbolic geodesic connecting the obtained start and end points.

Fig. 2 shows the obtained latent spaces. We observe that, for all models, the hand's initial resting pose is located near the latent space origin, from where the embedded trajectories progress outwards until the final grasp. Due to the stress prior, final grasps are organized in the latent space according to the taxonomy with grasps of each node grouped together in the latent space. Importantly, the trajectories appear scattered in the latent spaces of both GPLVM (Fig. 2a) and GPHLVM (Fig. 2c). In contrast, the dynamics priors of the GPDM (Fig. 2b) and GPHDM (Fig. 2d) effectively lead to smooth trajectories between the resting pose and the final grasps.

A quantitative evaluation of the models is presented in Table I. The stress reported in the first columns shows that all hyperbolic models better capture the taxonomy structure that their Euclidean counterparts. The difference is more prominent in the 2-dimensional case, consistent with the ability of hyperbolic geometry to embed tree-like structures in

TABLE I: Average stress, mean squared jerk (MSJ), and reconstruction mean squared error (MSE) per model and geometry.

Model	Stress ↓	MSJ (×100) ↓	MSE (×100) ↓
GPLVM $\mathbb{R}^2$	0.21 ± 0.43	1.87 ± 6.15	$\mathbf{0.50 \pm 1.28}$
GPDM $\mathbb{R}^2$	0.18 ± 0.47	$\underline{0.04 \pm 0.04}$	2.11 ± 5.13
GPHLVM $\mathbb{H}^2_{\mathcal{L}}$	$\mathbf{0.09 \pm 0.19}$	2.53 ± 10.30	1.26 ± 2.70
GPHDM $\mathbb{H}^2_{\mathcal{L}}$	$\underline{0.10 \pm 0.26}$	$\mathbf{0.03 \pm 0.02}$	$\underline{2.03 \pm 4.40}$
GPLVM $\mathbb{R}^3$	0.14 ± 0.21	0.58 ± 2.06	$\mathbf{0.29 \pm 0.76}$
GPDM $\mathbb{R}^3$	0.11 ± 0.21	$\underline{0.02 \pm 0.04}$	0.31 ± 0.99
GPHLVM $\mathbb{H}^3_{\mathcal{L}}$	$\underline{0.09 \pm 0.13}$	1.68 ± 7.47	$\underline{0.32 \pm 0.95}$
GPHDM $\mathbb{H}^3_{\mathcal{L}}$	$\mathbf{0.07 \pm 0.10}$	$\mathbf{0.02 \pm 0.02}$	0.32 ± 1.04

low-dimensional spaces. In general, we observed similar stress values between the GPHLVM and GPHDM. The second column of Table I reports the mean squared jerk (MSJ), which quantifies the smoothness of the latent trajectories. The MSJ of a single trajectory is given by,

$$ \text{MSJ}(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) = \frac{1}{N-4} \sum_{t=4}^{N} |\Delta^2 v_t|^2, \quad \text{with} \quad (33) $$

$$ \Delta^2 v_t = v_t - 2v_{t-1} + v_{t-2} \quad \text{and} \quad v_t = d_{\mathcal{M}}(\boldsymbol{x}_t, \boldsymbol{x}_{t-1}), $$

where $d_{\mathcal{M}}$ is the hyperbolic or Euclidean distance depending on the latent space geometry, v_t is a discrete approximation of the velocity, and Δ^2 is the finite difference operator. We observe a prominent MSJ reduction for the GPHDM and GPDM compared to the GPHLVM and GPLVM, indicating that both hyperbolic and Euclidean dynamics priors effectively promote smooth trajectories. Moreover, the GPHDM achieves the lowest MSJ among all models for both latent space dimensionality. Finally, the last column of Table I reports the mean squared error (MSE) of the decoded trajectories with respect to the training data. All models achieve similar MSE with 3-dimensional latent spaces leading to lower errors. In summary, the GPHDM inherits the superior taxonomy preservation of the GPHLVM, while simultaneously preserving the trajectory dynamical structure akin to the GPDM.

B. Dynamic Motion Generation

We evaluate the three different approaches to generate novel motions by decoding latent space trajectories. Fig. 3 shows the hyperbolic geodesics interpolating from a lateral to a stick grasp in the latent space of 2- and 3-dimensional GPHDMs, akin to the approach proposed in [6], alongside one exemplary dimension of the decoded hand motions. We observe that the 2-dimensional geodesic does not follow the structure of the training data and crosses the training trajectories, thus resulting in jerky hand motions. Moreover, the 3-dimensional geodesic predominantly passes through data-sparse regions, resulting in high-uncertainty motion predictions that revert to the non-informative mean. In both cases, the geodesics result in physically impractical motions that fail to capture the underlying dynamics of the training data.

Fig. 4-*left* displays a latent trajectory obtained via the mean prediction approach introduced in Sec. IV-A. The latent trajectory is initialized in the middle of a training trajectory leading to a lateral grasp. We observe that the dynamics prior induces an outward direction, resulting in motion
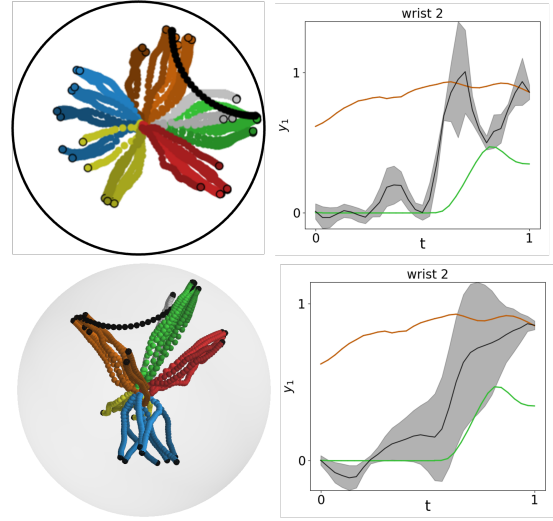

Fig. 3: **Left**: Embeddings of hand grasps and hyperbolic geodesics (——) from a lateral (●) to a stick (●) grasp. **Right**: Representative dimension of the probabilistic hand motion prediction for the geodesics (——) with mean and uncertainty, along with training trajectories for the lateral (reversed) (——) and stick (——) grasps. The **top** and **bottom** rows show 2- and 3-dimensional latent spaces.
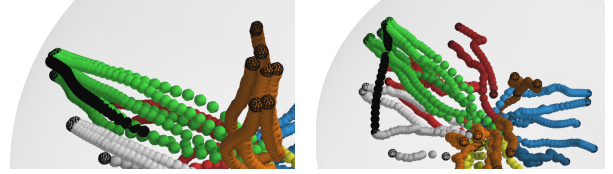

Fig. 4: Latent trajectories (——) obtained via recursive motion generation in a 3-dimensional GPHDM (zoomed-in). **Left**: Mean prediction towards a lateral grasp (●). **Right**: Conditional prediction from an index finger extension (●) to a lateral (●) grasp.

predictions that closely follow the training trajectory. However, as previously discussed, the mean prediction approach does not allow us to specify a desired goal point for the latent trajectory. Fig. 4-*right* shows a latent trajectory obtained via the conditional prediction approach of Sec. IV-A with start and goal points set as an index finger extension and a lateral grasp, respectively. In contrast to geodesics, the obtained trajectory incorporates the hyperbolic dynamics prior to transition between the two grasps. However, the conditional predictions follow the Markov assumption embedded in the hyperbolic dynamics prior, which induces a sense of directionality in the latent trajectories, as illustrated in Fig. 5. To avoid this issue, we trained the GPHDM of Fig. 4-*right* on an augmented dataset by including reverse training motions, i.e., motions from the final grasp onto the initial resting pose. Importantly, the recursive motion generation strategies do not prevent latent trajectories to traverse data-sparse regions, and thus often lead to motion predictions reverting to the mean and featuring high uncertainty.

Fig. 6 shows a latent trajectory obtained as a geodesic with respect to the GPHDM's pullback metric introduced in Sec. IV-B. The corresponding hyperbolic geodesic (akin to Fig. 3) is depicted as a reference. We observe that, in contrast to the hyperbolic geodesic and conditionally-optimized latent trajectories, the pullback geodesic closely adheres to the
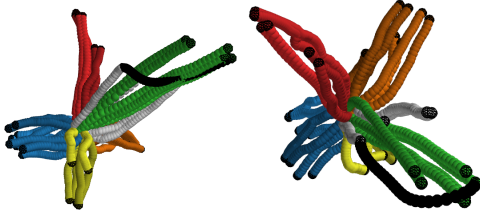
Fig. 5: Illustration of the directionality induced by the GPDM's Markov prior on latent trajectories obtained via conditional predictions. The outward transition (**left**) from an index finger extension (●) to a lateral (●) grasp follows a training trajectory, while the reverse transition (**right**) avoids the training data.
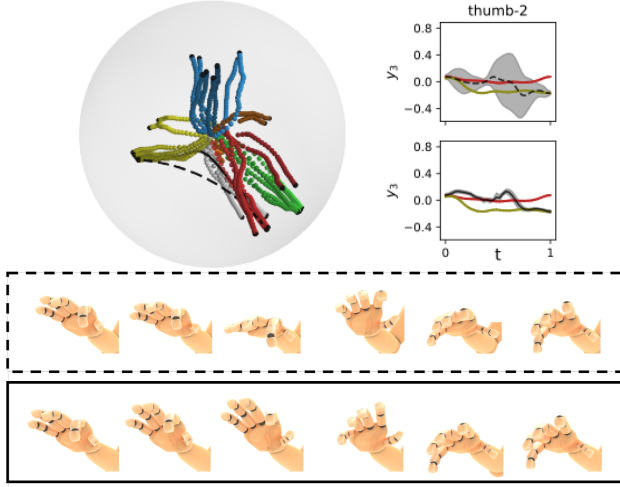

Fig. 6: **Top left**: 3-dimensional embeddings of hand grasps via a GPHDM with a hyperbolic (- -) and a pullback (——) geodesic from a ring (●) to a spherical (●) grasp. **Top right**: Representative dimension of the probabilistic hand motion prediction with mean and uncertainty along with training trajectories for the spherical (——) and ring (reversed) (——) grasps. **Bottom**: Generated hand motions from the decoded geodesics.

data support. Thus, the pullback-metric geodesic results in low-uncertainty motion predictions. As shown at the bottom of Fig. 6, the decoded hyperbolic geodesic leads to hand motions displaying large deviations from the start ring and goal spherical grasps. In contrast, by capturing the underlying motion dynamics and adhering to the training data, the decoded pullback geodesic produces physically-plausible motions with little deviation from the start and goal grasps.

VI. CONCLUSIONS

This paper proposed the GPHDM, a model that leverages hyperbolic geometry, human-designed taxonomy structures, and dynamics priors as inductive biases to learn latent spaces that preserve the hierarchical structure and temporal dynamics of human motions. We showed that these three forms of inductive biases are essential to learn taxonomy-aware dynamically-consistent latent spaces. Moreover, we introduced three novel mechanisms for generating taxonomy-aware and physically-consistent motions. Our results showed that trajectories obtained as geodesics on the pullback metric of the learned model produced low-uncertainty, physically-consistent motions that capture hierarchical structure and temporal dynamics of the motion data.

REFERENCES

[1] F. Stival, S. Michieletto, M. Cognolato, E. Pagello, H. Müller, and M. Atzori, "A quantitative taxonomy of human hand grasps," *J. Neuroeng Rehabil.*, vol. 16, no. 28, 2019.

[2] F. Iida and A. J. Ijspeert, "Biologically inspired robotics," in *Springer Handbook of Robotics*, 2016, pp. 2015–2034.

[3] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The GRASP taxonomy of human grasp types," *IEEE Trans. on Human-Machine Systems*, vol. 46, no. 1, pp. 66–77, 2016.

[4] J. Borràs, C. Mandery, and T. Asfour, "A whole-body support pose taxonomy for multi-contact humanoid robot motions," *Sci. Robot.*, vol. 2, no. 13, 2017.

[5] D. Paulius, N. Eales, and Y. Sun, "A motion taxonomy for manipulation embedding," in *R:SS*, 2020.

[6] N. Jaquier, L. Rozo, M. González-Duque, V. Borovitskiy, and T. Asfour, "Bringing motion taxonomies to continuous domains via GPLVM on hyperbolic manifolds," in *ICML*, 2024.

[7] L. Augenstein, N. Jaquier, T. Asfour, and L. Rozo, "On probabilistic pullback metrics on latent hyperbolic manifolds," *arXiv preprint arXiv:2410.20850*, 2024.

[8] J. Romero, T. Feix, H. Kjellström, and D. Kragic, "Spatio-temporal modeling of grasping actions," in *IEEE/RSJ IROS*, 2010, pp. 2103–2108.

[9] N. D. Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data," in *NeurIPS*, 2003.

[10] C. Mandery, J. Borràs, M. Jöchner, and T. Asfour, "Using language models to generate whole-body multi-contact motions," in *IEEE/RSJ IROS*, 2016, pp. 5411–5418.

[11] J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*. Springer, 2019.

[12] M. Nickel and D. Kiela, "Poincaré embeddings for learning hierarchical representations," in *NeurIPS*, 2017.

[13] W. F. Reynolds, "Hyperbolic geometry on a hyperboloid," *Am. Math. Mon.*, vol. 100, no. 5, pp. 442–455, 1993.

[14] I. Chami, A. Gu, V. Chatziafratis, and C. Ré, "From trees to continuous embeddings and back: Hyperbolic hierarchical clustering," in *NeurIPS*, 2020.

[15] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Gaussian process dynamical models for human motion," *IEEE PAMI*, vol. 30, no. 2, pp. 283–298, 2008.

[16] J. Lee, *Introduction to Riemannian Manifolds*. Springer, 2018.

[17] N. Boumal, *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023.

[18] H. Poincaré, "Théorie des groupes fuchsiens," *Acta Mathematica*, vol. 1, pp. 1 – 62, 1900.

[19] H. Jansen, "Abbildung der hyperbolischen geometrie auf ein zweischaliges hyperboloid," *Mitt. Math. Ges. Hamburg*, vol. 4, pp. 409–440, 1909.

[20] A. Grigoryan and M. Noguchi, "The heat kernel on hyperbolic space," *Bull. Lond. Math. Soc.*, vol. 30, no. 6, pp. 643–650, 1998.

[21] N. Jaquier, V. Borovitskiy, A. Smolensky, A. Terenin, T. Asfour, and L. Rozo, "Geometry-aware Bayesian optimization in robotics using Riemannian Matérn kernels," in *CoRL*, 2021.

[22] L. Rozo, M. González-Duque, N. Jaquier, and S. Hauberg, "Riemann^2: Learning Riemannian submanifolds from Riemannian data," in *AISTATS*, 2025.

[23] G. Bécigneul and O.-E. Ganea, "Riemannian adaptive optimization methods," in *ICLR*, 2019.

[24] M. Kochurov, R. Karimov, and S. Kozlukov, "Geoopt: Riemannian optimization in PyTorch," *arXiv preprint arXiv:2005.02819*, 2020.

[25] N. D. Lawrence and J. Quiñonero Candela, "Local distance preservation in the GP-LVM through back constraints," in *ICML*, 2006, p. 513–520.

[26] R. Urtasun, D. J. Fleet, A. Geiger, J. Popović, T. J. Darrell, and N. D. Lawrence, "Topologically-constrained latent variable models," in *ICML*, 2008, p. 1080–1087.

[27] A. Mallasto and A. Feragen, "Wrapped Gaussian process regression on Riemannian manifolds," in *CVPR*, 2018, pp. 5580–5588.

[28] A. Tosi, S. Hauberg, A. Vellido, and N. D. Lawrence, "Metrics for probabilistic geometries," in *UAI*, 2014.

[29] T. Anderson, "The non-central Wishart distribution and certain problems of multivariate statistics," *Ann. Math. Stat.*, vol. 17, no. 4, pp. 409–431, 1946.

[30] C. Mandery, O. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "Unifying representations and large-scale whole-body motion databases for studying human motion," *IEEE T-RO*, vol. 32, no. 4, pp. 796–809, 2016.