SHAPE DISSIMILARITY FOR PRODUCT-SPHERICALLY EMBEDDED TRAJECTORIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Traditional shape analysis in Euclidean space involves constructing pre-shapes by quotienting out Euclidean uniform scaling and translation. However, these operations are fundamentally undefined on spheres or product spheres. To overcome this geometric incompatibility, we introduce a novel framework for the shape analysis of product-spherically embedded trajectories that reside on or can be equivalently mapped to a product sphere. We propose the **P**roduct-**S**pherical **S**hape **D**issimilarity (PSSD) measure to rigorously quantify the shape differences between two such trajectories. Theoretically, we establish that the PSSD measure evaluates to zero if and only if two trajectories are equivalent up to spherical uniform scaling and reference-based spherical translation. Furthermore, we develop two algorithms: one extracts intrinsic spherical pre-shapes that remain embedded in the product sphere, and the other aligns the rotation of these pre-shapes. We also propose a rigorous smoothing estimator that recovers trajectories from discrete observations and guarantees uniform convergence. Extensive empirical evaluations across various synthetic and real-world datasets demonstrate that the PSSD measure achieves superior discriminative power in complex non-Euclidean classification tasks.

1 INTRODUCTION

The representation and analysis of trajectories on non-Euclidean metric spaces have emerged as a fundamental challenge in machine learning. In many real-world domains, dynamic processes naturally involve continuous trajectories that reside on or can be equivalently mapped to spheres or product spheres (Zhu & Müller, 2024; Xu & Shang, 2026). These are collectively referred to as product-spherically embedded trajectories. For instance, energy consumption compositional time series are commonly mapped to the positive orthant of a sphere via a pointwise square-root transformation (Scealy & Welsh, 2011), allowing their underlying smooth trends to be analyzed as spherically embedded trajectories. Similarly, demographic distributional time series, such as life-table death counts, can be mapped to the Hilbert sphere, yielding smooth spherical trends (Dai, 2022). Furthermore, multivariate biomechanical trajectories in human gait analysis (Prakash et al., 2018) naturally embed into product spheres. Across all these domains, isolating and understanding the intrinsic geometric shapes of these product-spherically embedded trajectories are essential for accurately capturing structural evolution and group dynamics in downstream classification tasks.

The established elastic shape framework was primarily developed for Euclidean trajectories, utilizing pre-shape spaces to ensure invariance under uniform scaling and translation (Srivastava et al., 2010; Srivastava & Klassen, 2016). However, such operations are fundamentally ill-defined on product spheres, which lack a global Euclidean vector space structure. Consequently, traditional elastic methodologies cannot be directly applied to spherically embedded data. Notably, while the transported square-root vector field framework (Joshi et al., 2016) and a time-warping framework (Chen & Müller, 2022) provide a Riemannian solution for temporal reparameterization invariance on manifolds, they do not inherently account for spatial scaling or translation invariance.

To overcome these limitations, we introduce a novel geometric framework for spherical shape analysis that evaluates intrinsic shapes directly on the sphere. Specifically, we propose a rigorous PSSD measure designed explicitly for product-spherically embedded trajectories. Crucially, we theoretically establish that the PSSD measure evaluates to zero if and only if two trajectories are equivalent up to spherical uniform scaling and reference-based spherical translation. Building upon this, we further

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

develop algorithms for extracting spherical pre-shapes that are intrinsically invariant to these spherical scaling and reference-based translation operations while remaining embedded in the product sphere, together with a further rotation alignment for spherical pre-shapes to achieve orientation invariance, leading to a novel Spherical Procrustes Dissimilarity (SPD) measure.

Contributions. While our primary focus is on spherical shape analysis, the proposed PSSD measure inherently functions as a rigorous dissimilarity measure. Consequently, it can serve as a foundational tool for broader machine learning applications, including similarity learning, dimensionality reduction, metric learning, and manifold embedding (Kulis, 2013; Saeed et al., 2018; Deng et al., 2024). Our main contributions to shape analysis are summarized as follows:

- We formulate a novel shape analysis framework that explicitly evaluates the shapes of product-spherically embedded trajectories and develop spherical pre-shapes that are invariant to spherical uniform scaling and reference-based spherical translation. To the best of our knowledge, this is the first general framework proposed for this problem.
- We develop theoretical guarantees that a PSSD measure evaluates to zero if and only if the product-spherically embedded trajectories share identical shapes up to reference-based spherical translation and uniform scaling, while the SPD measure returns zero if and only if the two spherical pre-shapes are equivalent up to rotation.
- We provide a rigorous smoothing technique to recover continuous product-spherically embedded trajectories from discrete observations, establishing theoretical uniform convergence results for the proposed estimator.
- Empirically, we demonstrate the superior discriminative power of our proposed PSSD measure on classification tasks across four synthetic and three diverse real-world datasets spanning compositional, distributional, and product-spherical functional data.

2 Preliminary on Euclidean Shape Analysis

Suppose that we have two real-valued trajectories $f_{\mathrm{E}}, g_{\mathrm{E}} : [0,1] \to \mathbb{R}^m$ for a positive integer m, the trajectories f_{E} and g_{E} can be defined to have the same shape if and only if, for every $u \in [0,1]$, $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + b$ for some $a > 0$ and some vector $b \in \mathbb{R}^m$. Thus one needs to construct a Euclidean shape dissimilarity (ESD) measure $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}})$ for f_{E} and g_{E} such that $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$ if and only if f_{E} and g_{E} have the same shape. Figure 1 illustrates two distinct classes of Euclidean-valued trajectories, corresponding to triangular and rectangular shapes. The trajectories within each respective group share identical shapes, so the resulting ESD measure between any two trajectories in the same group should be exactly zero.

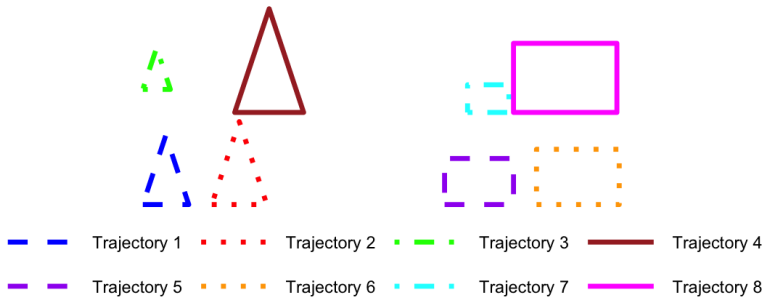


Figure 1: Two distinct groups of Euclidean-valued trajectories that possess identical shapes invariant to translation and uniform scaling. The first group (left) consists of Trajectories 1–4, while the second group (right) consists of Trajectories 5–8.

To achieve the requirement of the ESD measure, first let $\mathcal{H}$ represent a separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and the norm $\|\nu\|_{\mathcal{H}} = \sqrt{\langle \nu, \nu \rangle_{\mathcal{H}}}$. Then, we can construct the ESD measure as follows.

Definition 1. *For real-valued trajectories* $f_{\mathrm{E}}, g_{\mathrm{E}} : [0,1] \to \mathbb{R}^m$, *let* $f_{\mathrm{E}}^c(u) = f_{\mathrm{E}}(u) - \bar{f}_E$, $g_{\mathrm{E}}^c(u) = g_{\mathrm{E}}(u) - \bar{g}_E$, $u \in [0,1]$ *with* $\bar{f}_E = \int_0^1 f_{\mathrm{E}}(u)du$ *and* $\bar{g}_E = \int_0^1 g_{\mathrm{E}}(u)du$, $L_{\mathrm{E}}(f_{\mathrm{E}}^c) = \sqrt{\int_0^1 \|f_{\mathrm{E}}^c(u)\|_{\mathcal{H}}^2 du}$,

and $L_{\mathrm{E}}(g_{\mathrm{E}}^c) = \sqrt{\int_0^1 \|g_{\mathrm{E}}^c(u)\|_{\mathcal{H}}^2 du}$. *The ESD measure is*

$$D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = \begin{cases} 0, & \text{if } L_{\mathrm{E}}(f_{\mathrm{E}}^c) = L_{\mathrm{E}}(g_{\mathrm{E}}^c) = 0 \\ \frac{\pi}{2}, & \text{if one of } L_{\mathrm{E}}(f_{\mathrm{E}}^c), L_{\mathrm{E}}(g_{\mathrm{E}}^c) \text{ is } 0 \\ \arccos\left(\int_0^1 \langle f_{\mathrm{E}}^{cn}(u), g_{\mathrm{E}}^{cn}(u)\rangle_{\mathcal{H}} du\right), & \text{otherwise} \end{cases}$$

where $f_{\mathrm{E}}^{cn}(u) = f_{\mathrm{E}}^c(u)/L_{\mathrm{E}}(f_{\mathrm{E}}^c)$ *and* $g_{\mathrm{E}}^{cn}(u) = g_{\mathrm{E}}^c(u)/L_{\mathrm{E}}(g_{\mathrm{E}}^c)$.

The one-to-one correspondence for the ESD measure is established in the following result with proof in Appendix H.

Proposition 1. *For real-valued trajectories* $f_{\mathrm{E}}, g_{\mathrm{E}} : [0,1] \to \mathbb{R}^m$, *the ESD measure in Definition 1* $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$ *if and only if there exist a constant* $a > 0$ *and a vector* $b \in \mathbb{R}^m$ *such that* $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + b$ *everywhere.*

Proposition 1 ensures that the ESD measure between f and g is zero if and only if f and g are equivalent under translation and uniform scaling.

Notably, in shape analysis, see, e.g., Srivastava & Klassen (2016), f_{E}^{cn} and g_{E}^{cn} are defined as pre-shapes of f_{E} and g_{E}, respectively, as f_{E}^{cn} and g_{E}^{cn} are invariant under translation and uniform scaling.

3 Spherical Shape Dissimilarity Measure

Before constructing the framework of spherical shape analysis, we first provide a formal definition of spherically embedded trajectories. Spherically embedded trajectories take values that either naturally lie on a sphere or can be mapped to the unit sphere $\mathcal{S} = \{\nu \in \mathcal{H} : \|\nu\|_{\mathcal{H}} = 1\}$ of a separable Hilbert space $\mathcal{H}$. Equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, the space $\mathcal{S}$ is endowed with the intrinsic metric $d_{\mathcal{S}}(\nu_1, \nu_2) = \arccos(\langle \nu_1, \nu_2 \rangle_{\mathcal{H}})$. Here we cover both the finite-dimensional sphere $\mathcal{S}^{m-1}$ where $\mathcal{H} = \mathbb{R}^m$, and the infinite-dimensional Hilbert sphere $\mathcal{S}^{\infty}$. Such a framework naturally accommodates complex random objects, including compositional and distributional time series. For instance, following Scealy & Welsh (2011) and Zhu & Müller (2024), compositional data on the simplex $\Delta^{m-1} = \left\{\delta \in \mathbb{R}^m : \delta_l \geq 0, \sum \delta_l = 1\right\}$ can be projected to $\mathcal{S}^{m-1}$ using the pointwise square root $t_p(\delta) = (\sqrt{\delta_1}, \ldots, \sqrt{\delta_m})^{\top}$ where $^{\top}$ denotes the vector transpose. Analogously, a probability density function g supported on $\mathbb{R}^l$ maps to $\mathcal{S}^{\infty}$ through the functional square root $t_f(g)(\omega) = \sqrt{g(\omega)}$ for $\omega \in \mathbb{R}^l$. Because compositional and distributional trajectories admit equivalent spherical representations, our subsequent analysis focuses exclusively on the spherical case.

3.1 Measure for Spherically Embedded Trajectories

Suppose we have two spherically embedded trajectories $f_{\mathrm{S}}, g_{\mathrm{S}} : [0,1] \to \mathcal{S}$, satisfying Assumption A1 in Appendix G, a common assumption on the support of such trajectories. Here, Assumption A1 is used to ensure the existence and uniqueness of the center of each trajectory; see Appendix G for details. Our first objective is to adapt the Euclidean differences $f_{\mathrm{E}}(u) - \bar{f}_E$ and $g_{\mathrm{E}}(u) - \bar{g}_E$ from Definition 1 for spherically embedded trajectories $f_{\mathrm{S}}, g_{\mathrm{S}}$. In Euclidean space, the difference $a - b$ generates the linear translation from b to a. Analogously, for two different points $\nu_1, \nu_2 \in \mathcal{S}$, there exists a rotation generator that traverses the geodesic from ν_1 to ν_2 (Zhu & Müller, 2023; 2024). We express this spherical difference as $\nu_2 \ominus \nu_1 := T_{\nu_1,\nu_2} = \theta(\zeta_2 \otimes \zeta_1 - \zeta_1 \otimes \zeta_2)$, where $\otimes$ is the tensor product, $\theta = d_{\mathcal{S}}(\nu_1, \nu_2)$, $\zeta_1 = \nu_1$, and ζ_2 is obtained via Gram–Schmidt orthogonalization as $\zeta_2 = (\nu_2 - \langle \nu_2, \nu_1 \rangle_{\mathcal{H}} \nu_1)/\|\nu_2 - \langle \nu_2, \nu_1 \rangle_{\mathcal{H}} \nu_1\|_{\mathcal{H}}$. When $\nu_1 = \nu_2$, set $T_{\nu_1,\nu_2} = 0$. Consequently, T_{ν_1,ν_2} operates as a skew-symmetric element within the Hilbert space $\mathcal{C}(\mathcal{H})$, equipped with the Hilbert-Schmidt inner product $\langle h_1 \otimes \widetilde{h}_1, h_2 \otimes \widetilde{h}_2 \rangle_{\mathcal{C}(\mathcal{H})} = \langle h_1, h_2 \rangle_{\mathcal{H}} \langle \widetilde{h}_1, \widetilde{h}_2 \rangle_{\mathcal{H}}$ and the Hilbert-Schmidt norm $\|\cdot\|_{\mathcal{C}(\mathcal{H})}$. By the Rodrigues rotation formula (Zhu & Müller, 2024, Theorem 1), the exponential map of this operator is

$$\mathrm{Exp}(T_{\nu_1,\nu_2}) = \mathrm{id} + \frac{\sin(\theta)}{\theta} T_{\nu_1,\nu_2} + \frac{1 - \cos(\theta)}{\theta^2} T_{\nu_1,\nu_2}^2,$$

where id is the identity map. This optimal transport map, identical to the parallel transport, rotates the sphere by angle θ exclusively within the two-dimensional subspace spanned by ζ_1 and ζ_2. Formally, the connecting geodesic $\gamma : [0,1] \to \mathcal{S}$ is given by $\gamma(\delta) = \mathrm{Exp}(\delta T_{\nu_1,\nu_2})\nu_1$, which correctly yields $\gamma(0) = \nu_1$ and $\gamma(1) = \nu_2$ as shown in Appendix B.

The above construction allows us to define the rotation-based-centered trajectories

$$f_{\mathrm{S}}^{c}(u) = f_{\mathrm{S}}(u) \ominus \bar{f}_{\mathrm{S}} \quad \text{and} \quad g_{\mathrm{S}}^{c}(u) = g_{\mathrm{S}}(u) \ominus \bar{g}_{\mathrm{S}}, u \in [0,1] \tag{1}$$

with $\bar{f}_{\mathrm{S}} = \arg\min_{\nu\in\mathcal{S}} \int_0^1 d_{\mathcal{S}}^2(\nu, f_{\mathrm{S}}(u))du$ and $\bar{g}_{\mathrm{S}} = \arg\min_{\nu\in\mathcal{S}} \int_0^1 d_{\mathcal{S}}^2(\nu, g_{\mathrm{S}}(u))du$. Given that the space of the rotation-based-centered trajectories is Hilbertian, one can obtain that $\int_0^1 f_{\mathrm{S}}^{c}(u)du = \int_0^1 g_{\mathrm{S}}^{c}(u)du = 0$ whose derivation is given in Appendix B. Importantly, however, these two centered operator trajectories are expressed with respect to different centers whenever $\bar{f}_{\mathrm{S}} \neq \bar{g}_{\mathrm{S}}$. To quotient out the effect of spherical location, we therefore need to align the centers before computing the dissimilarity.

Consider a common target center μ_{S} for the alignment and define the center-alignment rotations

$$R_{f,\mu} = \mathrm{Exp}(\mu_{\mathrm{S}} \ominus \bar{f}_{\mathrm{S}}), \quad \text{and} \quad R_{g,\mu} = \mathrm{Exp}(\mu_{\mathrm{S}} \ominus \bar{g}_{\mathrm{S}}), \tag{2}$$

so that $R_{f,\mu}\bar{f}_{\mathrm{S}} = R_{g,\mu}\bar{g}_{\mathrm{S}} = \mu_{\mathrm{S}}$. The corresponding common-center-aligned trajectories are $f_{\mathrm{S}}^{\mu}(u) = R_{f,\mu}f_{\mathrm{S}}(u)$ and $g_{\mathrm{S}}^{\mu}(u) = R_{g,\mu}g_{\mathrm{S}}(u)$, with common-center-aligned rotation representations are

$$f_{\mathrm{S}}^{c,\mu}(u) = f_{\mathrm{S}}^{\mu}(u) \ominus \mu_{\mathrm{S}} = R_{f,\mu}f_{\mathrm{S}}^{c}(u)R_{f,\mu}^{-1}, \quad \text{and} \quad g_{\mathrm{S}}^{c,\mu}(u) = g_{\mathrm{S}}^{\mu}(u) \ominus \mu_{\mathrm{S}} = R_{g,\mu}g_{\mathrm{S}}^{c}(u)R_{g,\mu}^{-1}. \tag{3}$$

The construction in (3) follows from rotation equivariance of the spherical difference, proved in Appendix B.

We now define the spherical shape dissimilarity (SSD) using these common-center-aligned representations.

Definition 2. *For $f_{\mathrm{S}}, g_{\mathrm{S}} : [0,1] \to \mathcal{S}$ with $f_{\mathrm{S}}^{c,\mu}(u), g_{\mathrm{S}}^{c,\mu}(u)$ in (3), let*

$$L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu}) = \left\{\int_0^1 \|f_{\mathrm{S}}^{c,\mu}(u)\|_{\mathcal{C}(\mathcal{H})}^2 du\right\}^{1/2}, \quad \textit{and} \quad L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu}) = \left\{\int_0^1 \|g_{\mathrm{S}}^{c,\mu}(u)\|_{\mathcal{C}(\mathcal{H})}^2 du\right\}^{1/2}.$$

The SSD is

$$D_{\mathrm{S}}(f_{\mathrm{S}}, g_{\mathrm{S}}) = \begin{cases} 0, & L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu}) = L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu}) = 0, \\ \frac{\pi}{2}, & \textit{if exactly one of } L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu}), L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu}) \textit{ is } 0, \\ \arccos\left\{\int_0^1 \langle f_{\mathrm{S}}^{cn,\mu}(u), g_{\mathrm{S}}^{cn,\mu}(u)\rangle_{\mathcal{C}(\mathcal{H})}\, du\right\}, & \textit{otherwise}, \end{cases}$$

where $f_{\mathrm{S}}^{cn,\mu}(u) = f_{\mathrm{S}}^{c,\mu}(u)/L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu})$ and $g_{\mathrm{S}}^{cn,\mu}(u) = g_{\mathrm{S}}^{c,\mu}(u)/L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu})$.

As shown in Figure 2, the spherically embedded trajectories within the grey spheres share identical spherical shapes, as do those within the blue spheres. The proposed SSD measure between any two spherically embedded trajectories within the same group is then exactly zero.

The following result gives the zero-dissimilarity characterization with proof in Appendix H.

Theorem 1. *For continuous spherically embedded trajectories $f_{\mathrm{S}}, g_{\mathrm{S}} : [0,1] \to \mathcal{S}$, $D_{\mathrm{S}}(f_{\mathrm{S}}, g_{\mathrm{S}}) = 0$ if and only if there exists some constant $a > 0$ such that*

$$f_{\mathrm{S}}^{c,\mu}(u) = a\, g_{\mathrm{S}}^{c,\mu}(u), \qquad u \in [0,1]. \tag{4}$$

Equivalently, after moving both Fréchet centers to the common reference point μ_{S}, the two trajectories differ only by spherical uniform scaling. Under Assumption A1, this can be written pointwise as

$$f_{\mathrm{S}}^{\mu}(u) = \mathrm{Exp}(a\{g_{\mathrm{S}}^{\mu}(u) \ominus \mu_{\mathrm{S}}\})\, \mu_{\mathrm{S}}, \qquad u \in [0,1].$$

Theorem 1 separates the two nuisance operations explicitly. The reference-based spherical translation is removed by the center alignment in (3), while uniform scaling is removed by the global L^2 normalization in Definition 2.

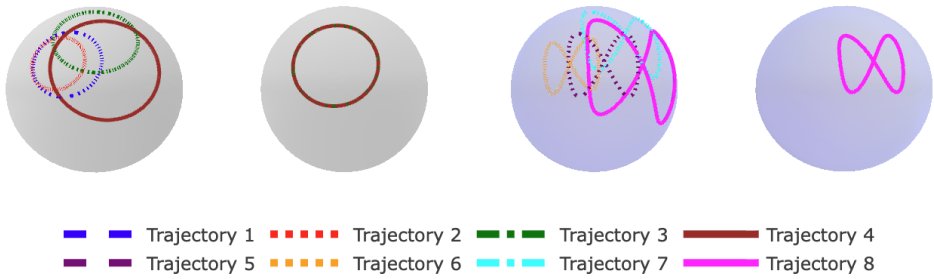


Figure 2: Two distinct groups of spherically embedded trajectories that possess identical geometric shapes up to μ_S-reference spherical translation and spherical uniform scaling. The first group (grey spheres) consists of Trajectories 1–4, while the second group (blue spheres) consists of Trajectories 5–8. The first and third spheres display the unaligned trajectories, whereas the second and fourth spheres display the corresponding common-center pre-shapes using Algorithm 1.

In practice, when we have a collection of trajectories $\{f_{S,i}\}_{i=1}^n$, μ_S is chosen as $\mu_S = \arg\min_{\mu\in\mathcal{S}} n^{-1}\sum_{i=1}^n d_{\mathcal{S}}^2(\mu, \bar{f}_{S,i})$ where $\bar{f}_{S,i}$ is the Fréchet center of the ith trajectory. Moreover, for any center η not antipodal to μ_S, let $R_{\eta,\mu} = \mathrm{Exp}(\mu_S \ominus \eta)$. We define the μ_S-reference spherical translation from η_1 to η_2 by

$$\mathcal{T}^{(\mu)}_{\eta_1\to\eta_2} = (R_{\eta_2,\mu})^{-1} R_{\eta_1,\mu}. \tag{5}$$

Then $\mathcal{T}^{(\mu)}_{\eta_1\to\eta_2}\eta_1 = \eta_2$, and these translations satisfy the composition rule $\mathcal{T}^{(\mu)}_{\eta_2\to\eta_3}\mathcal{T}^{(\mu)}_{\eta_1\to\eta_2} = \mathcal{T}^{(\mu)}_{\eta_1\to\eta_3}$. Thus, trajectories related only by a μ_S-reference spherical translation have exactly the same common-center-aligned representation. See Appendix B for the detailed derivation.

Notably, the second and fourth spheres in Figure 2 display the aligned spherical shapes obtained from the initially unaligned trajectories shown in the first and third spheres. These aligned representations serve as spherical pre-shapes, as they are invariant to both spherical translation and uniform scaling. To transform any spherically embedded trajectory f_S into its corresponding pre-shape with a given scale control parameter $\alpha > 0$ and a given target alignment location $\mu_S \in \mathcal{S}$, we utilize the methodology formalized in Algorithm 1 below.

Algorithm 1 Spherical pre-shape calculation, SPS_Cal.

Input: f_S: spherically embedded trajectory, $\alpha > 0$: scale control parameter, μ_S: target alignment location.
Output: spherical pre-shape p_S.

1: Calculate $\bar{f}_S = \arg\min_{\nu\in\mathcal{S}} \int_0^1 d_{\mathcal{S}}^2(\nu, f_S(u))du$ and $f_S^c(u) = f_S(u) \ominus \bar{f}_S$.
2: Calculate $R_{f,\mu} = \mathrm{Exp}(\mu_S \ominus \bar{f}_S)$ and $f_S^{c,\mu}(u) = R_{f,\mu} f_S^c(u) R_{f,\mu}^{-1}$.
3: Set $L_S = \{\int_0^1 \|f_S^{c,\mu}(u)\|^2_{\mathcal{C}(\mathcal{H})}du\}^{1/2}$.
4: **if** $L_S = 0$ **then**
5: Set $p_S(u) = \mu_S$ for all $u \in [0,1]$.
6: **else**
7: Set $p_S(u) = \mathrm{Exp}(\alpha f_S^{c,\mu}(u)/L_S)\,\mu_S$ for all $u \in [0,1]$.
8: **end if**

While classical shape analysis sometimes needs to discard rotational information to achieve orientation invariance (Rohlf, 1999; Srivastava & Klassen, 2016), our framework in this section mainly focuses on evaluating these spherical pre-shapes without rotational alignment. See Appendix C for a detailed discussion on this design choice driven by the nature of our real-world applications. It is worth noting that when we obtain spherical pre-shapes via Algorithm 1, we can further quotient out rotation with details given in Section 4.1.

3.2 Measure for Product-Spherically Embedded Trajectories

We then extend the SSD measure in Section 3.1 to the product sphere setting. Consider two product-spherically embedded trajectories $f_{\mathrm{PS}}, g_{\mathrm{PS}} : [0,1] \to \mathcal{PS}$ where $\mathcal{PS} = \underbrace{\mathcal{S} \times \cdots \times \mathcal{S}}_{L \text{ times}}$ is the product sphere space for a positive integer L with the product metric $d_{\mathcal{PS}}(\omega_1, \omega_2) = \sqrt{\sum_{l=1}^{L} d^2_{\mathcal{S}}(\omega_{1l}, \omega_{2l})}$ for $\omega_1, \omega_2 \in \mathcal{S}^{(L)}$ and $\omega_m = (\omega_{m1}, \ldots, \omega_{mL})$ for $m = 1, 2$. Extending the idea in Section 3.1, the product-rotation-based-centered trajectories for product-spherically embedded trajectories $f_{\mathrm{PS}} = \left(f^{(1)}_{\mathrm{PS}}, \ldots, f^{(L)}_{\mathrm{PS}}\right)$ and $g_{\mathrm{PS}} = \left(g^{(1)}_{\mathrm{PS}}, \ldots, g^{(L)}_{\mathrm{PS}}\right)$ are obtained via

$$f^{c}_{\mathrm{PS}}(u) = \left(f^{(1)}_{\mathrm{PS}}(u) \ominus \bar{f}^{(1)}_{\mathrm{PS}}, \ldots, f^{(L)}_{\mathrm{PS}}(u) \ominus \bar{f}^{(L)}_{\mathrm{PS}}\right) \text{ and}$$
$$g^{c}_{\mathrm{PS}}(u) = \left(g^{(1)}_{\mathrm{PS}}(u) \ominus \bar{g}^{(1)}_{\mathrm{PS}}, \ldots, g^{(L)}_{\mathrm{PS}}(u) \ominus \bar{g}^{(L)}_{\mathrm{PS}}\right), u \in [0,1] \tag{6}$$

with $\bar{f}_{\mathrm{PS}} = (\bar{f}^{(1)}_{\mathrm{PS}}, \ldots, \bar{f}^{(L)}_{\mathrm{PS}}) = \arg\min_{\omega \in \mathcal{PS}} \int_0^1 d^2_{\mathcal{PS}}(\omega, f_{\mathrm{PS}}(u))du$ and $\bar{g}_{\mathrm{PS}} = (\bar{g}^{(1)}_{\mathrm{PS}}, \ldots, \bar{g}^{(L)}_{\mathrm{PS}}) = \arg\min_{\omega \in \mathcal{PS}} \int_0^1 d^2_{\mathcal{PS}}(\omega, g_{\mathrm{PS}}(u))du$.

Consider a common target center $\mu_{\mathrm{PS}} = (\mu^{(1)}_{\mathrm{PS}}, \ldots, \mu^{(L)}_{\mathrm{PS}})$. For each $l \in \{1, \ldots, L\}$, let $R_{f,l,\mu} = \mathrm{Exp}(\mu^{(l)}_{\mathrm{PS}} \ominus \bar{f}^{(l)}_{\mathrm{PS}})$ and $R_{g,l,\mu} = \mathrm{Exp}(\mu^{(l)}_{\mathrm{PS}} \ominus \bar{g}^{(l)}_{\mathrm{PS}})$, and define

$$f^{c,\mu}_{\mathrm{PS}}(u) = \left(R_{f,1,\mu} f^{c,(1)}_{\mathrm{PS}}(u) R^{-1}_{f,1,\mu}, \ldots, R_{f,L,\mu} f^{c,(L)}_{\mathrm{PS}}(u) R^{-1}_{f,L,\mu}\right),$$
$$g^{c,\mu}_{\mathrm{PS}}(u) = \left(R_{g,1,\mu} g^{c,(1)}_{\mathrm{PS}}(u) R^{-1}_{g,1,\mu}, \ldots, R_{g,L,\mu} g^{c,(L)}_{\mathrm{PS}}(u) R^{-1}_{g,L,\mu}\right). \tag{7}$$

The product Hilbert space $\mathcal{C}^L(\mathcal{H}) = \mathcal{C}(\mathcal{H}) \times \cdots \times \mathcal{C}(\mathcal{H})$ has inner product $\langle q_1, q_2 \rangle_{\mathcal{C}^L(\mathcal{H})} = \sum_{l=1}^{L} \langle q_{1l}, q_{2l} \rangle_{\mathcal{C}(\mathcal{H})}$ and its associated norm.

Definition 3. *For the product-spherically embedded trajectories $f_{\mathrm{PS}}, g_{\mathrm{PS}} : [0,1] \to \mathcal{PS}$ with $f^{c,\mu}_{\mathrm{PS}}(u), g^{c,\mu}_{\mathrm{PS}}(u)$ in (7). Let*

$$L_{\mathrm{PS}}(f^{c,\mu}_{\mathrm{PS}}) = \left\{\int_0^1 \|f^{c,\mu}_{\mathrm{PS}}(u)\|^2_{\mathcal{C}^L(\mathcal{H})} du\right\}^{1/2}, \quad L_{\mathrm{PS}}(g^{c,\mu}_{\mathrm{PS}}) = \left\{\int_0^1 \|g^{c,\mu}_{\mathrm{PS}}(u)\|^2_{\mathcal{C}^L(\mathcal{H})} du\right\}^{1/2}.$$

The product-spherical shape dissimilarity (PSSD) is

$$D_{\mathrm{PS}}(f_{\mathrm{PS}}, g_{\mathrm{PS}}) = \begin{cases} 0, & L_{\mathrm{PS}}(f^{c,\mu}_{\mathrm{PS}}) = L_{\mathrm{PS}}(g^{c,\mu}_{\mathrm{PS}}) = 0, \\ \frac{\pi}{2}, & \text{if exactly one of } L_{\mathrm{PS}}(f^{c,\mu}_{\mathrm{PS}}), L_{\mathrm{PS}}(g^{c,\mu}_{\mathrm{PS}}) \text{ is } 0, \\ \arccos\left\{\int_0^1 \langle f^{cn,\mu}_{\mathrm{PS}}(u), g^{cn,\mu}_{\mathrm{PS}}(u)\rangle_{\mathcal{C}^L(\mathcal{H})}\, du\right\}, & \text{otherwise}, \end{cases}$$

where $f^{cn,\mu}_{\mathrm{PS}} = f^{c,\mu}_{\mathrm{PS}} / L_{\mathrm{PS}}(f^{c,\mu}_{\mathrm{PS}})$ and $g^{cn,\mu}_{\mathrm{PS}} = g^{c,\mu}_{\mathrm{PS}} / L_{\mathrm{PS}}(g^{c,\mu}_{\mathrm{PS}})$.

When $L = 1$, PSSD reduces to SSD. The zero-dissimilarity characterization is the direct product-space analogue of Theorem 1.

Theorem 2. *For continuous product-spherically embedded trajectories $f_{\mathrm{PS}}, g_{\mathrm{PS}} : [0,1] \to \mathcal{PS}$, $D_{\mathrm{PS}}(f_{\mathrm{PS}}, g_{\mathrm{PS}}) = 0$ if and only if there exists some constant $a > 0$ such that*

$$f^{c,\mu}_{\mathrm{PS}}(u) = a\, g^{c,\mu}_{\mathrm{PS}}(u), \qquad u \in [0,1].$$

Thus PSSD is invariant to componentwise μ_{PS}-reference spherical translations of the product center and to one common spherical scale factor across the L components.

Notably, the pre-shape calculation originally developed for spherical trajectories in Algorithm 1 can be generalized to the product sphere setting, with details in Appendix D.

4 Extensions of Spherical Shape Analysis

4.1 Extension to Quotient Out Rotation for Spherical Pre-Shapes

Using Algorithm 1, without loss of generality, suppose we have two spherical pre-shapes $p_{\mathrm{S},1}(u), p_{\mathrm{S},2}(u) \in \mathcal{S}^{m-1}$, $u \in [0,1]$ sharing aligned index points $u_1, u_2, \ldots, u_K$ with the Fréchet

mean at the north pole $\mu_S = (0, \ldots, 0, 1)^\top$. Construct matrices $Z_1, Z_2 \in \mathbb{R}^{m\times K}$ whose k-th columns are $p_{S,1}(u_k)$ and $p_{S,2}(u_k)$, respectively. Let $SO(m)$ be the set of $m \times m$ real orthogonal matrices with determinant 1. The algorithm to quotient out rotation is given as follows.

Algorithm 2 Rotation alignment for the spherically embedded trajectories, Rotation_Align.

Input: Z_1: target configuration in $\mathbb{R}^{m\times K}$, Z_2: configuration to be aligned in $\mathbb{R}^{m\times K}$.
Output: aligned configuration Z_2^{align}.

1: Extract the orthogonal subspace by dropping the m-th dimension. Let $\tilde{Z}_1, \tilde{Z}_2 \in \mathbb{R}^{(m-1)\times K}$ be the submatrices containing the first $m-1$ rows of Z_1 and Z_2.
2: Compute $\tilde{C} = \tilde{Z}_1\tilde{Z}_2^\top$ and perform the singular value decomposition to obtain $\tilde{C} = U\Sigma V^\top$.
3: Compute the optimal constrained sub-rotation $\tilde{R} \in SO(m-1)$ as $\tilde{R} = U\Omega V^\top$ where $\Omega = \text{diag}(1, \ldots, 1, \det(UV^\top))$.
4: Construct the full $m \times m$ rotation matrix $\widehat{R} \in SO(m)$ as $\widehat{R} = \begin{pmatrix} \tilde{R} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}$.
5: Apply the transformation to align the orientation of Z_2 to Z_1 as $Z_2^{\text{align}} = \widehat{R}Z_2$.

Figure 3 displays the spherically embedded trajectories within the grey sphere that share identical spherical shapes invariant to spherical translation, spherical uniform scaling, and spherical rotation. As Algorithm 1 isolates only translation and scaling invariants, the resulting spherical pre-shapes remain unaligned in orientation, as shown in the blue sphere in Figure 3. By subsequently applying Algorithm 2 to these pre-shapes, the purple sphere in Figure 3 indicates that the trajectories are aligned.

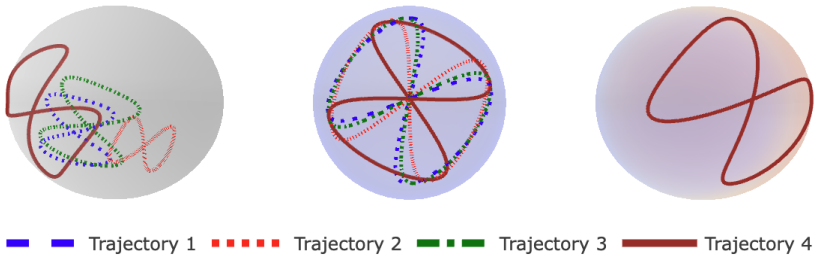


Figure 3: One group of spherically embedded trajectories that possess identical geometric shapes invariant to spherical translation, spherical uniform scaling, and spherical rotation, consisting of Trajectories 1–4. The first sphere displays the unaligned trajectories, the second sphere displays aligned shapes by quotienting out spherical translation and spherical uniform scaling using Algorithm 1, and the third sphere displays aligned shapes by quotienting out spherical translation, spherical uniform scaling, and spherical rotation using Algorithms 1 and 2.

After using Algorithm 2 with outputs Z_1 and Z_2^{align}, the dissimilarity can be measured via $D_{\text{rot}}(Z_1, Z_2) = \arccos\left(K^{-1}\text{tr}\left(Z_1^\top Z_2^{\text{align}}\right)\right)$. The next theorem shows that Algorithm 2 successfully quotients out rotation for spherical pre-shapes, with proof in Appendix H.

Theorem 3. *For spherical pre-shapes $p_{S,1}(u), p_{S,2}(u) \in \mathcal{S}^{m-1}$ obtained via Algorithm 1, let $Z_1, Z_2 \in \mathbb{R}^{m\times K}$ be two matrices whose columns represent corresponding points $p_{S,1}(u_k)$ and $p_{S,2}(u_k)$. The spherical Procrustes dissimilarity measure $D_{\text{rot}}(Z_1, Z_2) = \arccos\left(K^{-1}\text{tr}\left(Z_1^\top Z_2^{\text{align}}\right)\right) = 0$ with Z_2^{align} obtained via Algorithm 2, if and only if there exists a rotation matrix $R^* \in SO(m)$ of the form $R^* = \begin{pmatrix} R_{m-1} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}$ with $R_{m-1} \in SO(m-1)$ such that $Z_1 = R^*Z_2$.*

The above methodology for the sphere can be easily extended to the product sphere setting; see Appendix E for details.

4.2 Smoothing Product-Spherically Embedded Functional Data

In practice, we can only get access to discrete points of a product-spherically embedded trajectory f_{PS} with random measurement errors, denoted as $\tilde{f}_{\mathrm{PS}}(u_1), \tilde{f}_{\mathrm{PS}}(u_2), \ldots, \tilde{f}_{\mathrm{PS}}(u_K)$ for some positive integer K. Suppose that there exist random perturbation maps $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}$ such that for every $k \in \{1, 2, \ldots, K\}$, $\Gamma_k : \mathcal{PS} \to \mathcal{PS}$, $\tilde{f}_{\mathrm{PS}}(u_k) = \Gamma_k\{f_{\mathrm{PS}}(u_k)\}$, and for any fixed $\nu \in \mathcal{PS}$, $\nu = \arg\min_{\omega \in \mathcal{PS}} \mathbb{E}\{d^2_{\mathcal{PS}}(\omega, \Gamma_k(\nu))\}$. This gives $f_{\mathrm{PS}}(u_k) = \arg\min_{\omega \in \mathcal{PS}} \mathbb{E}\{d^2_{\mathcal{PS}}(\omega, \tilde{f}_{\mathrm{PS}}(u_k))\}$ for every time points u_k, $k \in \{1, 2, \ldots, K\}$. To obtain the true but unknown spherically embedded trajectory f_{PS} from the noisy samples $\tilde{f}_{\mathrm{PS}}(u_1), \tilde{f}_{\mathrm{PS}}(u_2), \ldots, \tilde{f}_{\mathrm{PS}}(u_K)$, one could utilize a local Fréchet regression estimator (Petersen & Müller, 2019) that extends the local polynomial regression in Euclidean space to a general metric space, given by

$$\widehat{f}_{\mathrm{PS}}(u) = \underset{\omega \in \mathcal{PS}}{\arg\min}\, \widehat{F}(\omega, u) \tag{8}$$

$$\widehat{F}(\omega, u) = K^{-1} \sum_{k=1}^{K} \widehat{w}(u_k, u) d^2_{\mathcal{PS}}(\omega, \tilde{f}_{\mathrm{PS}}(u_k))$$

for $u \in [0, 1]$, where $\widehat{w}(u_k, u) = \widehat{\sigma}^{-2} \mathcal{K}_h(u_k - u)[\widehat{\tau}_2(u) - \widehat{\tau}_1(u)(u_k - u)]$ with $\widehat{\tau}_l(u) = K^{-1} \sum_{k=1}^{K} [\mathcal{K}_h(u_k - u)(u_k - u)^l]$ for $l \in \{0, 1, 2\}$, and $\widehat{\sigma}^2 = \widehat{\tau}_0(u)\widehat{\tau}_2(u) - \widehat{\tau}_1^2(u)$. Here $\mathcal{K}_h(\cdot) = \mathcal{K}(\cdot/h)/h$ for a smoothing kernel $\mathcal{K}(\cdot)$ and a bandwidth sequence $h > 0$. As noticed in Petersen & Müller (2019), the model in (8) is equivalent to the classical local polynomial regression estimator with polynomial order being 1 if the data are real-valued and the metric is the Euclidean distance. The hyperparameter h can be selected via the cross-validation approach if the noisy samples $\tilde{f}_{\mathrm{PS}}(u_1), \tilde{f}_{\mathrm{PS}}(u_2), \ldots, \tilde{f}_{\mathrm{PS}}(u_K)$ are assumed to be independent (Petersen & Müller, 2019), or via a rolling-window cross-validation approach for temporally dependent samples (Xu & Shang, 2026).

The uniform consistency result of the local Fréchet regression estimator in (8) is given in the following result with proof in Appendix H.

Theorem 4. *Suppose Assumptions A1–A4 in Appendix G hold, if $h \to 0$ and $Kh \to \infty$ as $K \to \infty$, then* $\sup_{u \in [0,1]} d_{\mathcal{PS}}(\widehat{f}_{\mathrm{PS}}(u), f_{\mathrm{PS}}(u)) = o_p(1)$.

5 Experiments

This section evaluates the classification performance of our proposed spherical shape analysis framework. For each sample of product-spherically embedded functional data, we first employ the smoothing methodology described in Section 4.2 to reconstruct the smooth product-spherical trajectories. We then construct a dissimilarity matrix where each entry represents the PSSD measure computed between the corresponding pair of the smooth trajectories. For a given dissimilarity matrix, we finally employ the k-Nearest Neighbor (k-NN) classifier evaluated via two distinct validation protocols covering a randomized (stratified) train-test split with 75% of the samples for training and 25% for testing, and Leave-One-Out Cross-Validation (LOOCV). The details of the usage of the k-NN classifier for a given dissimilarity matrix are given in Appendix I.2. Additional ablation study results when the smoothing methodology is not applied are given in Appendix I.7.

Baselines. We include five baseline dissimilarity measures for comparing product-spherically embedded trajectories f_{PS} and g_{PS}. The first is the Spherical Dissimilarity (SD), defined as $\int_0^1 \sum_{l=1}^{L} \arccos\left(\langle f^{(l)}_{\mathrm{PS}}(u), g^{(l)}_{\mathrm{PS}}(u)\rangle_{\mathcal{H}}\right) du$. The second is the proposed SPD calculated via $D_{\mathrm{rot}}(\cdot, \cdot)$ in Section 4.1. The third is the Euclidean Shape Dissimilarity (ESD) introduced in Definition 1. The fourth is the Euclidean Dissimilarity (ED), calculated as $\int_0^1 \|f_{\mathrm{PS}}(u) - g_{\mathrm{PS}}(u)\|_{\mathcal{H}} du$. The fifth is the Euclidean Procrustes Dissimilarity (EPD) which involves one additional step to quotient out rotation, see Appendix I.1 for details. We also consider two neural-network baselines, a 2D CNN and ManifoldNet Chakraborty et al. (2020), treating each product-spherically embedded trajectory as a matrix. The details of neural network models are given in Appendix I.3.

Table 1 summarizes the key features of PSSD and the baseline dissimilarity measures. Comparing the accuracy of specific pairs of dissimilarity measures allows us to evaluate the contribution of each feature listed in Table 1, serving as ablation studies.

Table 1: Summary of dissimilarity measures.

Feature	PSSD	SD	SPD	ESD	ED	EPD
Account for sphere geometry	✓	✓	✓	✗	✗	✗
Quotient out translation and scaling	✓	✗	✓	✓	✗	✓
Quotient out rotation	✗	✗	✓	✗	✗	✓

In the following, we focus on synthetic data settings, while the real-world data analysis can be found in Appendix I.6.

5.1 Synthetic Data

We consider four distinct synthetic data settings by varying the sample size $n \in \{30, 90\}$ and the dimension $m \in \{6, 101\}$ of the sphere $\mathcal{S}^{m-1}$. The data generation details can be found in Appendix I.5. For each synthetic data setting, we perform 200 independent replications and report the average classification accuracy in Table 2.

Table 2: Evaluation results on classification for different synthetic data settings with sample size $n \in \{30, 90\}$ and the sphere dimension $m \in \{6, 101\}$. All results are in percentages (%).

Split	Setting	PSSD	SD	SPD	ESD	ED	EPD	CNN	ManifoldNet
Random Split	$n = 30, m = 6$	**92.88**	85.50	85.81	85.94	85.25	86.50	37.29	36.57
	$n = 90, m = 6$	**94.50**	89.37	89.57	89.89	90.20	90.33	40.68	78.26
	$n = 30, m = 101$	**97.19**	88.56	88.69	88.62	88.56	89.44	48.36	75.00
	$n = 90, m = 101$	**98.13**	92.39	92.57	92.85	92.93	93.22	40.68	56.52
LOOCV	$n = 30, m = 6$	**93.63**	86.73	86.15	86.98	87.33	86.59	46.51	51.74
	$n = 90, m = 6$	**93.84**	90.63	90.01	90.99	91.24	90.45	44.44	50.52
	$n = 30, m = 101$	**97.82**	90.78	89.12	91.08	90.83	89.67	50.69	46.83
	$n = 90, m = 101$	**97.96**	93.61	93.10	93.99	94.19	93.55	51.02	52.22

As demonstrated in Table 2, our proposed PSSD measure consistently outperforms all the baselines in terms of classification accuracy. Notably, when the product-spherically embedded trajectories f_{PS} and g_{PS} possess fundamentally distinct spherical shapes, the baselines may still yield large dissimilarity values, resulting in deceptively high classification accuracy. However, the primary limitation of these baseline dissimilarity measures, excluding the SPD measure, is their inability to discern whether f_{PS} and g_{PS} share an identical spherical shape. Unlike our PSSD measure, they lack the invariance properties necessary to identify geometric equivalence, leading to a failure in cases where shape preservation is the critical discriminative factor. Because in our synthetic data generation settings, the simulated shapes are not invariant to rotation, leading to a decrease in the classification accuracy produced by the SPD measure. Remarkably, despite employing a straightforward k-NN classifier, our PSSD-based framework outperforms neural network models trained end-to-end on trajectories. This highlights the inherent discriminative power of our proposed measure in capturing the true geometric structure of the data.

6 Conclusion

We presented a novel, mathematically rigorous framework for the shape analysis of dynamic processes that are embedded in, or can be equivalently mapped to, spheres or product spheres. By overcoming the geometric incompatibility of Euclidean scaling and translation on curved manifolds, we developed the PSSD measure, coupled with the ability to extract intrinsic spherical pre-shapes and a theoretically guaranteed smoothing technique. By bridging the gap between spherical shape analysis and classical Euclidean shape analysis, our work not only advances product-spherically embedded data analysis but also opens new avenues for manifold-based metric learning and generalized Procrustes analysis.

AI USE STATEMENT

In this work, we did not use generative AI tools to generate synthetic datasets, develop theoretical models or conceptual frameworks, prove mathematical claims or write proofs, propose or refine hypotheses, provide feedback on research methodology or experiments, implement methods, translate text, clean or reformat datasets, support qualitative analysis, or interpret results. We used generative AI tools only to improve the clarity of the language. The authors reviewed and verified all such output, and take full responsibility for the final content of this work, including all text, claims, and artifacts.

REPRODUCIBILITY STATEMENT

Our code is publicly available on GitHub via https://github.com/d5rhtkp4kc-alt/Shape-dissimilarity-for-product-spherically-embedded-trajectories. All the real datasets are publicly available, with details in Appendix I.6. All proofs of theoretical results can be found in Appendix H.

REFERENCES

Gary Bécigneul and Octavian-Eugen Ganea. Riemannian adaptive optimization methods. In *International Conference on Learning Representations*, 2019. URL: https://openreview.net/forum?id=gde5lYPeOk.

Denis Bosq. *Linear Processes in Function Spaces: Theory and Applications*, volume 149. Springer Science & Business Media, 2000.

Rudrasis Chakraborty, Jose Bouza, Jonathan H Manton, and Baba C Vemuri. Manifoldnet: A deep neural network for manifold-valued data with applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2):799–810, 2020.

Yaqing Chen and Hans-Georg Müller. Uniform convergence of local Fréchet regression with applications to locating extrema and time warping for metric space valued trajectories. *The Annals of Statistics*, 50(3):1573–1592, 2022.

Xiongtao Dai. Statistical inference on the Hilbert sphere with application to random densities. *Electronic Journal of Statistics*, 16(1):700–736, 2022.

Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, Hongbin Sun, and Cheng Xin. Neuc-MDS: Non-Euclidean multidimensional scaling through bilinear forms. *Advances in Neural Information Processing Systems*, 37:121539–121569, 2024.

Nathaniel E Helwig, Sungjin Hong, Elizabeth T Hsiao-Wecksler, and John D Polk. Methods to temporally align gait cycle data. *Journal of Biomechanics*, 44(3):561–566, 2011.

Nathaniel E Helwig, K Alex Shorter, Ping Ma, and Elizabeth T Hsiao-Wecksler. Smoothing spline analysis of variance models: A new tool for the analysis of cyclic biomechanical data. *Journal of Biomechanics*, 49(14):3216–3222, 2016.

Japanese Mortality Database. *National Institute of Population and Social Security Research*, 2025. Available at https://www.ipss.go.jp/p-toukei/JMD/index-en.asp (data downloaded on August 21, 2025).

Shantanu H Joshi, Jingyong Su, Zhengwu Zhang, and Boulbaba Ben Amor. Elastic shape analysis of functions, curves and trajectories. In *Riemannian computing in computer vision*, pp. 211–231. Springer, 2016.

Max Kochurov, Rasul Karimov, and Serge Kozlukov. Geoopt: Riemannian optimization in pytorch. Technical report, arXiv, 2020. URL: https://arxiv.org/abs/2005.02819.

Brian Kulis. Metric learning: A survey. *Foundations and Trends® in Machine Learning*, 5(4): 287–364, 2013.

Russell Lyons. Strong negative type in spheres. *Pacific Journal of Mathematics*, 307(2):383–390, 2020.

Yue-Pok Mack and Bernard W Silverman. Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61(3):405–415, 1982.

A. Petersen and H-G. Müller. Fréchet regression for random objects with Euclidean predictors. *The Annals of Statistics*, 47(2):691–719, 2019.

Chandra Prakash, Rajesh Kumar, and Namita Mittal. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review*, 49(1):1–40, 2018.

F James Rohlf. Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of Classification*, 16(2):197–223, 1999.

Nasir Saeed, Haewoon Nam, Mian Imtiaz Ul Haq, and Dost Bhatti Muhammad Saqib. A survey on multidimensional scaling. *ACM Computing Surveys (CSUR)*, 51(3):1–25, 2018.

J. L. Scealy and A. H. Welsh. Properties of a square root transformation regression model. In *Proceedings of CoDaWork'11: 4th international workshop on Compositional Data Analysis*. CIMNE, Barcelona, 2011. URL: https://congress2.cimne.com/codawork11/Admin/Files/FilePaper/p25.pdf.

Isaac J Schoenberg. On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. *Annals of Mathematics*, 38(4):787–793, 1937.

Isaac J Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44(3):522–536, 1938.

K Alex Shorter, John D Polk, Karl S Rosengren, and Elizabeth T Hsiao-Wecksler. A new approach to detecting asymmetries in gait. *Clinical Biomechanics*, 23(4):459–467, 2008.

Bernard W Silverman. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, 6(1):177–184, 1978.

Anuj Srivastava and Eric P Klassen. *Functional and Shape Data Analysis*. Springer, New York, 2016.

Anuj Srivastava, Eric Klassen, Shantanu H Joshi, and Ian H Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7): 1415–1428, 2010.

Aad W Van Der Vaart and Jon A Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, New York, 1996.

Jiazhen Xu and Han Lin Shang. Spherically embedded time series with unknown trend and periodic components. Technical report, arXiv, 2026. URL: https://arxiv.org/abs/2604.03574v1.

Jiazhen Xu, Andrew TA Wood, and Tao Zou. Robust functional principal component analysis for non-Euclidean random objects. *Biometrika*, 112(4):asaf040, 2025.

C. Zhu and H.-G. Müller. Autoregressive optimal transport models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 85(3):1012–1033, 2023.

C. Zhu and H.-G. Müller. Spherical autoregressive models, with application to distributional and compositional time series. *Journal of Econometrics*, 239(2):105389, 2024.

A Appendix Summary

We include the following appendices as support for the content in the main paper:

B. Derivation for the optimal-transport-related results in Section 3.1 such that $\gamma(0) = \nu_1$ and $\gamma(1) = \nu_2$ for $\gamma(\delta) = \mathrm{Exp}(\delta T_{\nu_1,\nu_2})\nu_1$, and $\int_0^1 f_{\mathrm{S}}^c(u)du = \int_0^1 g_{\mathrm{S}}^c(u)du = 0$.

C. A discussion on when not to quotient out rotation.

D. Product-spherical pre-shape calculation.

E. Quotient out rotation for product-spherical pre-shapes.

F. Reparameterization for product-spherically embedded trajectories.

G. Assumptions needed for the uniform convergence result in Section 4.2.

H. Proofs of Proposition 1, Theorems 1–4.

I. Additional experiment information and results.

B Derivation for the Optimal Transport Map

Part I. We first verify $\gamma(0) = \nu_1$ and $\gamma(1) = \nu_2$ for $\gamma(\delta) = \mathrm{Exp}(\delta T_{\nu_1,\nu_2})\nu_1$. Recall that exponential of T_{ν_1,ν_2} given in Section 3.1 is

$$\mathrm{Exp}(T_{\nu_1,\nu_2}) = \mathrm{id} + \frac{\sin(\theta)}{\theta} T_{\nu_1,\nu_2} + \frac{1-\cos(\theta)}{\theta^2} T_{\nu_1,\nu_2}^2,$$

and one can see that $\gamma(0) = \mathrm{Exp}(0 \times T_{\nu_1,\nu_2})\nu_1 = \nu_1$. For $\gamma(1) = \mathrm{Exp}(T_{\nu_1,\nu_2})\nu_1$, we get

$$\gamma(1) = \nu_1 + \frac{\sin(\theta)}{\theta}(T_{\nu_1,\nu_2}\nu_1) + \frac{1-\cos(\theta)}{\theta^2}(T_{\nu_1,\nu_2}^2\nu_1). \tag{9}$$

Note that $T_{\nu_1,\nu_2}\nu_1 = \theta\{(\zeta_2 \otimes \zeta_1)\nu_1 - (\zeta_1 \otimes \zeta_2)\nu_1\}$. As $\langle \nu_1, \zeta_1\rangle_{\mathcal{H}} = \|\nu_1\|_{\mathcal{H}}^2 = 1$ and $\langle \nu_1, \zeta_2\rangle_{\mathcal{H}} = 0$, one obtains $T_{\nu_1,\nu_2}\nu_1 = \theta\zeta_2$. This also gives $T_{\nu_1,\nu_2}^2\nu_1 = T_{\nu_1,\nu_2}(T_{\nu_1,\nu_2}\nu_1) = T_{\nu_1,\nu_2}(\theta\zeta_2) = \theta(T_{\nu_1,\nu_2}\zeta_2)$ while $T_{\nu_1,\nu_2}\zeta_2 = \theta(\langle\zeta_2,\zeta_1\rangle_{\mathcal{H}}\zeta_2 - \langle\zeta_2,\zeta_2\rangle_{\mathcal{H}}\xi_1) = -\theta\xi_1 = -\theta\nu_1$. Thus, we can conclude that $T_{\nu_1,\nu_2}^2\nu_1 = -\theta^2\nu_1$. The above results, together with (9), show that

$$\begin{aligned}
\gamma(1) &= \nu_1 + \frac{\sin(\theta)}{\theta}(\theta\zeta_2) + \frac{1-\cos(\theta)}{\theta^2}(-\theta^2\nu_1) \\
&= \nu_1 + \sin(\theta)\zeta_2 - (1-\cos(\theta))\nu_1 \\
&= \cos(\theta)\nu_1 + \sin(\theta)\zeta_2 \\
&= \cos(\theta)\nu_1 + \sin(\theta)\frac{\nu_2 - \cos(\theta)\nu_1}{\sin(\theta)} \\
&= \nu_2.
\end{aligned}$$

Part II. We next establish the two identities needed for the common-center construction.

First, we verify that the original rotation-based-centered trajectory integrates to zero. It is sufficient to consider f_{S}. Since $\bar{f}_{\mathrm{S}}$ is the Fréchet mean, thus

$$\int_0^1 \mathrm{Log}_{\bar{f}_{\mathrm{S}}}\{f_{\mathrm{S}}(u)\}\, du = 0.$$

Recall that $f_{\mathrm{S}}^c(u) = \theta(u)\{\zeta_2(u) \otimes \bar{f}_{\mathrm{S}} - \bar{f}_{\mathrm{S}} \otimes \zeta_2(u)\}$, we have $\theta(u)\zeta_2(u) = \mathrm{Log}_{\bar{f}_{\mathrm{S}}}\{f_{\mathrm{S}}(u)\}$. Therefore

$$\int_0^1 f_{\mathrm{S}}^c(u)du = \left\{\int_0^1 \theta(u)\zeta_2(u)du\right\} \otimes \bar{f}_{\mathrm{S}} - \bar{f}_{\mathrm{S}} \otimes \left\{\int_0^1 \theta(u)\zeta_2(u)du\right\} = 0.$$

The same argument gives $\int_0^1 g_{\mathrm{S}}^c(u)du = 0$.

Second, let Q be an orthogonal operator on $\mathcal{H}$. The spherical difference is rotation equivariant:

$$T_{Q\nu_1,Q\nu_2} = QT_{\nu_1,\nu_2}Q^{-1}. \tag{10}$$

Observe that $d_{\mathcal{S}}(Q\nu_1, Q\nu_2) = d_{\mathcal{S}}(\nu_1, \nu_2)$, while the Gram–Schmidt directions transform as $\zeta_1 \to Q\zeta_1$ and $\zeta_2 \to Q\zeta_2$. Using $(Qh_1) \otimes (Qh_2) = Q(h_1 \otimes h_2)Q^{-1}$ yields (10). Note that the coincident case follows from the convention $T_{\nu,\nu} = 0$.

Taking $Q = R_{f,\mu} = \mathrm{Exp}(\mu_{\mathrm{S}} \ominus \bar{f}_{\mathrm{S}})$ gives

$$f_{\mathrm{S}}^{\mu}(u) \ominus \mu_{\mathrm{S}} = T_{R_{f,\mu}\bar{f}_{\mathrm{S}}, R_{f,\mu}f_{\mathrm{S}}(u)} = R_{f,\mu} f_{\mathrm{S}}^{c}(u) R_{f,\mu}^{-1},$$

which proves (3). It also follows that

$$\int_0^1 f_{\mathrm{S}}^{c,\mu}(u)du = R_{f,\mu}\left\{\int_0^1 f_{\mathrm{S}}^{c}(u)du\right\} R_{f,\mu}^{-1} = 0,$$

and similarly for $g_{\mathrm{S}}^{c,\mu}$.

Finally, the reference-based translations in (5) satisfy

$$\mathcal{T}^{(\mu)}_{\eta_2\to\eta_3}\mathcal{T}^{(\mu)}_{\eta_1\to\eta_2} = (R_{\eta_3,\mu})^{-1}R_{\eta_2,\mu}(R_{\eta_2,\mu})^{-1}R_{\eta_1,\mu} = \mathcal{T}^{(\mu)}_{\eta_1\to\eta_3},$$

which motivates the choice of one common reference point for the comparison between multiple trajectories with the sample size $n \geq 3$.

C When not to quotient out rotation

Similar to classical shape analysis, whether we need to discard rotational information to achieve orientation invariance is determined by the data design. In spatial domains like computer vision, rotational invariance is vital because the global coordinate system such as the camera angle is arbitrary. In our applications, however, the coordinates of the spherically embedded trajectories possess absolute, fixed semantic meanings, such as distinct fuel categories or specific demographic age cohorts. Applying a rotation operator would mathematically scramble these physical categories, possibly destroying the underlying geometric structure. Furthermore, the absolute direction of the evolving shape of spherical-embedded functional data is our primary signal of interest. For instance, considering energy compositions, if we allowed rotational alignment, a country transitioning rapidly into coal dependency might possess the exact same geometric evolving shape as a country transitioning into renewables. Thus, comparing pre-shapes directly preserves the semantic integrity and directional context required for accurate downstream classification.

D Product-spherical pre-shape calculation

The procedural details for the product-spherical pre-shape calculation are formalized in the following algorithm.

E Quotient out rotation for product-spherical pre-shapes

By utilizing Algorithm 3, suppose we have two product-spherical pre-shapes $p_{\mathrm{PS},1} = (p^{(1)}_{\mathrm{PS},1}, \ldots, p^{(L)}_{\mathrm{PS},1})$ and $p_{\mathrm{PS},2} = (p^{(1)}_{\mathrm{PS},2}, \ldots, p^{(L)}_{\mathrm{PS},2})$, $u \in [0,1]$ sharing aligned index points $u_1, u_2, \ldots, u_K$ with the Fréchet mean at the north pole $\mu_{\mathrm{S}}^{(l)} = (0, \ldots, 0, 1)^{\top}$ for $l \in \{1, \ldots, L\}$, where $p^{(l)}_{\mathrm{PS,i}} \in \mathcal{S}^{m-1}$ for $i \in \{1, 2\}$ and $l \in \{1, \ldots, L\}$.

For each component $l \in \{1, \ldots, L\}$, construct matrices $Z_1^{(l)}, Z_2^{(l)} \in \mathbb{R}^{m\times K}$ where the k-th columns are $p^{(l)}_{PS,1}(u_k)$ and $p^{(l)}_{PS,2}(u_k)$, respectively. Let $Z_{\mathrm{PS,i}} = \{Z_i^{(l)}\}_{l=1}^{L}$ for $i \in \{1, 2\}$. We then apply Algorithm 2 independently to each pair $(Z_1^{(l)}, Z_2^{(l)})$ to extract L distinct optimal rotation matrices $\widehat{R}^{(l)} \in SO(m)$ and output L aligned configurations $(Z_2^{(l)})^{\mathrm{align}} = \widehat{R}^{(l)}Z_2^{(l)}$. Finally, the dissimilarity can be measured via

$$D_{\mathrm{rotProduct}}(Z_{\mathrm{PS},1}, Z_{\mathrm{PS},2}) = \arccos\left((KL)^{-1}\sum_{l=1}^{L}\mathrm{tr}\left((Z_1^{(l)})^{\top}(Z_2^{(l)})^{\mathrm{align}}\right)\right).$$

Algorithm 3 Product-spherical pre-shape calculation, PSPS_Cal.

Input: $f_{\mathrm{PS}} = (f_{\mathrm{PS}}^{(1)}, \ldots, f_{\mathrm{PS}}^{(L)})$: product-spherically embedded trajectory, $\alpha > 0$: scale control parameter, $\mu_{\mathrm{PS}} = (\mu_{\mathrm{PS}}^{(1)}, \ldots, \mu_{\mathrm{PS}}^{(L)})$: target alignment location.

Output: product-spherical pre-shape $p_{\mathrm{PS}} = (p_{\mathrm{PS}}^{(1)}, \ldots, p_{\mathrm{PS}}^{(L)})$.

Calculate $\bar{f}_{\mathrm{PS}} = \arg\min_{\omega \in \mathcal{PS}} \int_0^1 d_{\mathcal{PS}}^2(\omega, f_{\mathrm{PS}}(u))du$.
for $l = 1$ **to** L **do**
    Set $f_{\mathrm{PS}}^{c,(l)}(u) = f_{\mathrm{PS}}^{(l)}(u) \ominus \bar{f}_{\mathrm{PS}}^{(l)}$.
    Set $R_{f,l,\mu} = \mathrm{Exp}(\mu_{\mathrm{PS}}^{(l)} \ominus \bar{f}_{\mathrm{PS}}^{(l)})$ and $f_{\mathrm{PS}}^{c,\mu,(l)}(u) = R_{f,l,\mu} f_{\mathrm{PS}}^{c,(l)}(u) R_{f,l,\mu}^{-1}$.
end for
Set $L_{\mathrm{PS}} = \left\{ \int_0^1 \sum_{l=1}^{L} \| f_{\mathrm{PS}}^{c,\mu,(l)}(u) \|_{\mathcal{C}(\mathcal{H})}^2 du \right\}^{1/2}$.
if $L_{\mathrm{PS}} = 0$ **then**
    Set $p_{\mathrm{PS}}^{(l)}(u) = \mu_{\mathrm{PS}}^{(l)}$ for all l and u.
else
    for $l = 1$ **to** L **do**
        Set $p_{\mathrm{PS}}^{(l)}(u) = \mathrm{Exp}\left(\alpha f_{\mathrm{PS}}^{c,\mu,(l)}(u) / L_{\mathrm{PS}} \right) \mu_{\mathrm{PS}}^{(l)}$ for all $u \in [0,1]$.
    end for
end if

F Reparameterization for product-spherically embedded trajectories

While reparameterization invariance is out of scope as this is not needed in the real-world datasets considered. If reparameterization is needed, then one can directly follow Section 5 of Chen & Müller (2022). Here, we detail the reparameterization process for product-spherically embedded trajectories as follows.

We establish the piecewise linear formulation used to parameterize the warping functions. Let $u_l = l/(L+1)$ for $l = 1, \ldots, L$ be L equidistant knots in $[0,1]$, with $u_0 = 0$ and $u_{L+1} = 1$. Consider basis functions $A(u) = (A_1(u), \ldots, A_{L+1}(u))^\top$, with $A_l(u) = A_l^{(1)}(u) - A_l^{(2)}(u)$, where $A_l^{(1)}(u) = \frac{u - u_{l-1}}{u_l - u_{l-1}} 1_{[u_{l-1}, u_l)}$, $A_l^{(2)}(u) = \frac{u - u_l}{u_l - u_{l-1}} 1_{[u_{l-1}, u_l)}$ for $l = 1, \ldots, L+1$, and $A_{L+2}^{(2)} = 0$. The coefficient parameter space is $\Theta = \{\theta \in \mathbb{R}^{L+1} : 0 < \theta_1 < \cdots < \theta_{L+1} = 1\}$. The alignment procedure is summarized in Algorithm 4.

Algorithm 4 Reparameterization for the product-spherically embedded trajectories, Reparam_Cal.

Input: f_{PS}: product-spherically embedded trajectory, g_{PS}: product-spherically embedded trajectory for the alignment target, $\lambda > 0$: penalty.
Output: aligned trajectories $f_{\mathrm{PS}}^{\mathrm{align}}$.
Compute the optimal coefficient vector $\widehat{\theta}$ by minimizing the penalized integrated distance as

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} \int_0^1 \left[d_{\mathcal{PS}}^2(f_{\mathrm{PS}}(\theta^\top A(u)), g_{\mathrm{PS}}(u)) + \lambda(\theta^\top A(u) - u)^2 \right] du$$

Construct the temporal transformation function $\widehat{\beta}(u) = \widehat{\theta}^\top A(u)$ for all $u \in [0,1]$
Apply the transformation to align the phase of f_{PS} to g_{PS} as $f_{\mathrm{PS}}^{\mathrm{align}}(u) = f_{\mathrm{PS}}(\widehat{\beta}(u))$.

Theoretical supports and numerical performance for Algorithm 4 can be found in Chen & Müller (2022).

G Assumptions

Before listing the required assumptions, we first introduce the α-mixing condition for the discretely observed product-spherically embedded trajectory. For a given K, without loss of generality, let $u_k =$

k/K, $k = 1, 2, \ldots, K$ and let $Y_{K,k} = \widetilde{f}_{\mathrm{PS}}(u_k)$. For two σ-algebras $\mathcal{A}$ and $\mathcal{B}$, define $\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A},\, B \in \mathcal{B}} |\mathbb{P}(A \cap B) - |\mathbb{P}(A)|\mathbb{P}(B)|$. For $j = 1, \ldots, K$, define $\mathcal{F}^{-}_{K,j} = \sigma(Y_{K,1}, \ldots, Y_{K,j})$ and $\mathcal{F}^{+}_{K,j} = \sigma(Y_{K,j}, \ldots, Y_{K,K})$. For an integer lag $r \geq 1$, the α-mixing coefficient is defined as $\alpha_K(r) = \max_{1 \leq j \leq K-r} \alpha\left(\mathcal{F}^{-}_{K,j}, \mathcal{F}^{+}_{K,j+r}\right)$. Thus, $\alpha_K(r)$ measures the maximal dependence between observations up to index j and observations from index $j + r$ onward.

Assumption A1. *The support* $\mathrm{supp} \subset \mathcal{S}$ *of each individual spherical element* $\{f^{(l)}_{\mathrm{PS}}(u)\}_{u \in [0,1]}, l = 1, \ldots, L$ *in* $\{f_{\mathrm{PS}}(u) = (f^{(1)}_{\mathrm{PS}}(u), \ldots, f^{(L)}_{\mathrm{PS}}(u))\}_{u \in [0,1]}$ *satisfies* $\sup_{\nu_1, \nu_2 \in \mathrm{supp}} d_{\mathcal{S}}(\nu_1, \nu_2) \leq \pi/2$.

Assumption A2. *There exist constants* $C_1 > 0$ *and* $C_2 > 0$*, independent of* K*, such that* $\alpha_K(r) \leq C_1 \exp(-C_2 r)$ *for* $r = 1, 2, \ldots, K-1$.

Assumption A3. *Suppose the kernel function* $K : \mathbb{R} \to \mathbb{R}^+$ *is a uniformly continuous and symmetric probability density. We require the integral constants* $H_{l_1 l_2} = \int_{\mathbb{R}} \mathcal{K}^{l_1}(x) x^{l_2} dx$ *to be finite for combinations of* $l_1, l_2 \in \{1, 2, 4, 6\}$*, with* H_{14} *and* H_{26} *being specifically bounded. Furthermore,* K *must possess a bounded first derivative across its support, satisfying* $\sup_{\{x : \mathcal{K}(x) > 0\}} |\mathcal{K}'(x)| < \infty$*, along with the tail condition* $\int_{\mathbb{R}} x^2 |\mathcal{K}'(x)| \sqrt{|x \log |x||} dx < \infty$.

Assumption A4. *Let* $F(\omega, u) = \mathbb{E}\{d^2_{\mathcal{P}\mathcal{S}}(\omega, \tilde{f}_{\mathrm{PS}}(u))\}$ *is twice continuously differentiable with respect to* u*, uniformly in* ω*, and* $\sup_{u \in [0,1]} \sup_{\omega \in \mathcal{P}\mathcal{S}} \left| \frac{\partial^2}{\partial u^2} F(\omega, u) \right| < \infty$.

Assumption A1 provides the geometric constraints necessary for the population Fréchet mean $f_{\mathrm{PS}}(u)$ to exist and be unique at any given point $u \in [0, 1]$. This is identical to condition (A1) established by Dai (2022), who additionally note that for finite-dimensional spheres, the support restriction can be weakened to $\sup_{\nu_1, \nu_2 \in \mathrm{supp}} d_{\mathcal{S}}(\nu_1, \nu_2) \leq \pi$. Next, Assumption A2 imposes a standard mixing structure typical in time series contexts, and Assumption A2 also covers a special case when the samples of a product-spherically embedded trajectory are independent. Finally, Assumptions A3 and A4 outline the regularity constraints for the kernel necessary for local nonparametric regression. These two assumptions, together with the results in Silverman (1978) and Mack & Silverman (1982), guarantee the asymptotic uniform equicontinuity of the empirical cost function associated with the local Fréchet regression detailed in Section 4.2, aligning with the conditions (K0) and (R0) in Chen & Müller (2022) under the fixed design setting.

H Proofs of main results

Proof of Proposition 1. The proof consists of two parts where we first show the sufficiency and then show the necessity.

Part I: We first show that if there exist a constant $a > 0$ and a vector $b \in \mathbb{R}^m$ such that $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + b$ everywhere, then $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$.

First, observe that $\bar{f}_E = \int_0^1 (a g_{\mathrm{E}}(u) + b) du = a \bar{g}_E + b$, which gives

$$f^c_{\mathrm{E}}(u) = f_{\mathrm{E}}(u) - \bar{f}_E = a(g_{\mathrm{E}}(u) - \bar{g}_E) = a g^c_{\mathrm{E}}(u).$$

Thus one can see that $\|f^c_{\mathrm{E}}(u)\|_{\mathcal{H}} = \|a g^c_{\mathrm{E}}(u)\|_{\mathcal{H}} = a \|g^c_{\mathrm{E}}(u)\|_{\mathcal{H}}$ for any $u \in [0, 1]$.

If $\|g^c_{\mathrm{E}}(u)\|_{\mathcal{H}} = 0$, then we have $\|f^c_{\mathrm{E}}(u)\|_{\mathcal{H}} = 0$ for any $u \in [0, 1]$. Therefore, following the construction of the shape dissimilarity measure, $D(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$. If $\int_0^1 \|g^c_{\mathrm{E}}(u)\|_{\mathcal{H}} du > 0$, then as $a > 0$, we also have $\int_0^1 \|f^c_{\mathrm{E}}(u)\|_{\mathcal{H}} du > 0$. Note that for any $u \in [0, 1]$,

$$f^{cn}_{\mathrm{E}}(u) = \frac{f^c_{\mathrm{E}}(u)}{\|f^c_{\mathrm{E}}\|_{\mathcal{H}}} = g^{cn}_{\mathrm{E}}(u),$$

and thus

$$\langle f^{cn}_{\mathrm{E}}(u), g^{cn}_{\mathrm{E}}(u) \rangle_{\mathcal{H}} = \|g^{cn}_{\mathrm{E}}(u)\|^2_{\mathcal{H}} = 1.$$

This leads to $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$ as $\arccos(1) = 0$.

Part II: We then show that, if $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}}) = 0$, then there exists a constant $a > 0$ and a vector $b \in \mathbb{R}^m$ such that $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + b$ everywhere.

Based on the construction of $D_{\mathrm{E}}(f_{\mathrm{E}}, g_{\mathrm{E}})$, the measure equals 0 happens in two scenarios where (i) $\int_0^1 \|f_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du = 0$ and $\int_0^1 \|g_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du = 0$, and (ii) $\int_0^1 \|f_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du > 0$ and $\int_0^1 \|g_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du > 0$.

For the first case where $\int_0^1 \|f_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du = \int_0^1 \|g_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du = 0$, we obtain $f_{\mathrm{E}}^c(u) = 0$ and $g_{\mathrm{E}}^c(u) = 0$ everywhere. Therefore, $f_{\mathrm{E}}(u) = \bar{f}_E$ and $g_{\mathrm{E}}(u) = \bar{g}_E$ are constant. Choose $a = 1$ and $b = \bar{f}_E - \bar{g}_E$, one has $f_{\mathrm{E}}(u) = \bar{f}_E = a g_{\mathrm{E}}(u) + b$.

For the second case where $\int_0^1 \|f_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du > 0$ and $\int_0^1 \|g_{\mathrm{E}}^c(u)\|_{\mathcal{H}} du > 0$, $D_{\mathrm{E}}(f_{\mathrm{E}}^c, g_{\mathrm{E}}^c) = \arccos\left(\int_0^1 \langle f_{\mathrm{E}}^{cn}(u), g_{\mathrm{E}}^{cn}(u)\rangle_{\mathcal{H}} du\right) = 0$. This implies $\int_0^1 \langle f_{\mathrm{E}}^{cn}(u), g_{\mathrm{E}}^{cn}(u)\rangle_{\mathcal{H}} du = 1$. Utilizing the Cauchy–Schwarz inequality, $|\int_0^1 \langle f_{\mathrm{E}}^{cn}(u), g_{\mathrm{E}}^{cn}(u)\rangle_{\mathcal{H}} du| = 1$ if and only if $f_{\mathrm{E}}^{cn}(u) = c g_{\mathrm{E}}^{cn}(u)$ for some scalar $c > 0$ and any $u \in [0, 1]$. Because their Hilbert norms are 1, it must be that $c = 1$. This gives

$$f_{\mathrm{E}}^c(u) = \frac{\|f_{\mathrm{E}}^c\|_{\mathcal{H}}}{\|g_{\mathrm{E}}^c\|_{\mathcal{H}}} g_{\mathrm{E}}^c(u)$$

Let $a = \|f_{\mathrm{E}}^c\|_{\mathcal{H}} / \|g_{\mathrm{E}}^c\|_{\mathcal{H}}$. Since $\|f_{\mathrm{E}}^c\|_{\mathcal{H}} > 0$, it is guaranteed that $a > 0$. Observe that $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + (\bar{f}_E - a\bar{g}_E)$ and thus choosing $b = \bar{f}_E - a\bar{g}_E$ gives $f_{\mathrm{E}}(u) = a g_{\mathrm{E}}(u) + b$.

Combining Parts I and II completes the proof. □

Proof of Theorem 1. Consider the space $\mathbb{H}_{\mathrm{S}} = L^2([0, 1], \mathcal{C}(\mathcal{H}))$ with inner product

$$\langle A, B\rangle_{\mathbb{H}_{\mathrm{S}}} = \int_0^1 \langle A(u), B(u)\rangle_{\mathcal{C}(\mathcal{H})} du.$$

The quantities $L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu})$ and $L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu})$ are then the corresponding $\mathbb{H}_{\mathrm{S}}$ norms.

If both norms are zero, then $f_{\mathrm{S}}^{c,\mu} = g_{\mathrm{S}}^{c,\mu} = 0$ in $\mathbb{H}_{\mathrm{S}}$. Therefore, Definition 2 gives $D_{\mathrm{S}}(f_{\mathrm{S}}, g_{\mathrm{S}}) = 0$. If exactly one norm is zero, one can observe that the dissimilarity equals $\pi/2$.

Now let

$$F = \frac{f_{\mathrm{S}}^{c,\mu}}{L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu})}, \quad G = \frac{g_{\mathrm{S}}^{c,\mu}}{L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu})}.$$

Then $\|F\|_{\mathbb{H}_{\mathrm{S}}} = \|G\|_{\mathbb{H}_{\mathrm{S}}} = 1$. By Definition 2, $D_{\mathrm{S}}(f_{\mathrm{S}}, g_{\mathrm{S}}) = 0$ if and only if $\langle F, G\rangle_{\mathbb{H}_{\mathrm{S}}} = 1$. Since

$$\|F - G\|_{\mathbb{H}_{\mathrm{S}}}^2 = \|F\|_{\mathbb{H}_{\mathrm{S}}}^2 + \|G\|_{\mathbb{H}_{\mathrm{S}}}^2 - 2\langle F, G\rangle_{\mathbb{H}_{\mathrm{S}}},$$

this is equivalent to $F = G$ almost everywhere. Hence $f_{\mathrm{S}}^{c,\mu}(u) = a\, g_{\mathrm{S}}^{c,\mu}(u)$ almost everywhere and $a = \frac{L_{\mathrm{S}}(f_{\mathrm{S}}^{c,\mu})}{L_{\mathrm{S}}(g_{\mathrm{S}}^{c,\mu})} > 0$. As the trajectories are continuous and the spherical-difference and center-alignment maps are continuous under Assumption A1, both sides are continuous in u. Then, equality almost everywhere implies equality everywhere. The converse follows by normalizing both sides. Finally, under Assumption A1, exponentiating the identity $f_{\mathrm{S}}^{c,\mu} = a g_{\mathrm{S}}^{c,\mu}$ at the common base point μ_{S} gives the pointwise spherical-scaling representation stated in the theorem. □

Proof of Theorem 2. Let $\mathbb{H}_{\mathrm{PS}} = L^2([0, 1], \mathcal{C}^L(\mathcal{H}))$. Observe that Definition 3 is the angle between $f_{\mathrm{PS}}^{c,\mu}$ and $g_{\mathrm{PS}}^{c,\mu}$ of $\mathbb{H}_{\mathrm{PS}}$. Following the arguments used in the proof of Theorem 1, we only need to consider the following case, where $D_{\mathrm{PS}} = 0$ if and only if the two unit-norm elements have inner product one, which is equivalent to their difference having squared $\mathbb{H}_{\mathrm{PS}}$ norm zero. Thus

$$\frac{f_{\mathrm{PS}}^{c,\mu}}{L_{\mathrm{PS}}(f_{\mathrm{PS}}^{c,\mu})} = \frac{g_{\mathrm{PS}}^{c,\mu}}{L_{\mathrm{PS}}(g_{\mathrm{PS}}^{c,\mu})} \quad \text{a.e.}$$

and therefore

$$f_{\mathrm{PS}}^{c,\mu}(u) = a\, g_{\mathrm{PS}}^{c,\mu}(u), \qquad a = \frac{L_{\mathrm{PS}}(f_{\mathrm{PS}}^{c,\mu})}{L_{\mathrm{PS}}(g_{\mathrm{PS}}^{c,\mu})} > 0.$$

Borrowing the arguments used in the proof of Theorem 1 completes the proof. □

Proof of Theorem 3. The proof consists of two parts where we first show the sufficiency and then show the necessity.

Part I: We first show that if there exist a rotation matrix $R^* \in SO(m)$ of the specified form such that $Z_1 = R^* Z_2$, then $D_{\text{rot}}(Z_1, Z_2) = 0$.

Observe that Algorithm 2 gives the optimal rotation $\widehat{R}$ that minimizes $\|Z_1 - \widehat{R}Z_2\|_{\text{F}}^2$ where $\|\cdot\|_{\text{F}}$ is the Frobenius norm. Minimizing this squared distance is equivalent to maximizing $\text{tr}(Z_1^\top \widehat{R} Z_2)$. As $Z_1 = R^* Z_2$, the maximum possible trace is achieved when the algorithm outputs $\widehat{R} = R^*$. Thus we get $Z_2^{\text{align}} = \widehat{R}Z_2 = R^* Z_2 = Z_1$. This gives that

$$D(Z_1, Z_2) = \arccos\left(K^{-1}\text{tr}(Z_1^\top Z_1)\right) = \arccos(1) = 0.$$

Part II: We then show that, when $D_{\text{rot}}(Z_1, Z_2) = 0$, then there exist a rotation matrix $R^* \in SO(m)$ of the specified form such that $Z_1 = R^* Z_2$.

Note that $D_{\text{rot}}(Z_1, Z_2) = 0$ gives $\text{tr}(Z_1^\top Z_2^{\text{align}}) = K$. Let v_k and w_k denote the k-th column of Z_1 and Z_2^{align}, respectively. We get

$$\text{tr}(Z_1^\top Z_2^{\text{align}}) = \sum_{k=1}^{K} v_k^\top w_k = K.$$

Note that by construction, all columns v_k in Z_1 represent points on the unit sphere $\mathcal{S}^{m-1}$, meaning $\|v_k\|_{\mathcal{H}} = 1$. Furthermore, because Z_2^{align} is calculated by applying the orthogonal rotation matrix $\widehat{R}$ to Z_2, thus every column w_k is also a unit vector where $\|w_k\|_{\mathcal{H}} = 1$. By the Cauchy–Schwarz inequality, $v_k^\top w_k \leq \|v_k\|_{\mathcal{H}} \|w_k\|_{\mathcal{H}} = 1$, and equality holds if and only if $v_k = w_k$.

For the sum of K such terms to equal exactly K, we must have $v_k^\top w_k = 1$ for all $k \in \{1, \ldots, K\}$. Consequently, this implies that $Z_1 = Z_2^{\text{align}}$. Recall that $Z_2^{\text{align}} = \widehat{R}Z_2$, and $\widehat{R}$ was constructed by Algorithm 2 to be a rotation matrix of the form $\begin{pmatrix} \tilde{R} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}$, we have established that there exists a valid rotation matrix $R^* = \widehat{R}$ such that $Z_1 = R^* Z_2$.

Combining Parts I and II finishes the proof. □

Proof of Theorem 4. Following Theorems 1 and 2 in Xu et al. (2025), it is sufficient to show that, for any compact subset $\overline{\mathcal{PS}}$ of $\mathcal{PS}$,

$$\gamma(\varepsilon) = \inf_{u \in [0,1]} \inf_{\substack{\omega \in \overline{\mathcal{PS}}: \\ d_{\mathcal{PS}}(\omega, f_{\mathcal{PS}}(u)) \geq \varepsilon}} [F(\omega, u) - F(f_{\mathcal{PS}}(u), u)] > 0, \tag{11}$$

for every $\varepsilon > 0$, and

$$\sup_{\omega \in \overline{\mathcal{PS}}} \sup_{u \in [0,1]} \left| \widehat{F}(\omega, u) - F(\omega, u) \right| = o_p(1), \tag{12}$$

as $h \to 0$ where $F(\omega, u) = \mathbb{E}\left[d_{\mathcal{PS}}^2(\omega, \tilde{f}_{\text{PS}}(u))\right]$.

We first verify (11). Observe that $u \to f_{\mathcal{PS}}(u)$ is continuous. To see this, let $u_n \to u$. As $\overline{\mathcal{PS}}$ is compact, every subsequence of $f_{\mathcal{PS}}(u_n)$ contains a subsequence, say $f_{\mathcal{PS}}(u_{n_j})$, converging to some $\omega^\star \in \overline{\mathcal{PS}}$.

For every $\omega \in \overline{\mathcal{PS}}$, $F(f_{\mathcal{PS}}(u_{n_j}), u_{n_j}) \leq F(\omega, u_{n_j})$. Note that the map F is jointly continuous by Assumption A4, while for $\omega_1, \omega_2 \in \overline{\mathcal{PS}}$,

$$|F(\omega_1, u) - F(\omega_2, u)| \leq C d_{\mathcal{PS}}(\omega_1, \omega_2)$$

for some constant C. Thus we get $F(\omega^\star, u) \leq F(\omega, u)$ for any $\omega \in \overline{\mathcal{PS}}$. By Assumption A1, we have $\omega^\star = f_{\mathcal{PS}}(u)$. Thus every convergent subsequence of $f_{\mathcal{PS}}(u_n)$ converges to $f_{\mathcal{PS}}(u)$, which proves $f_{\mathcal{PS}}(u_n) \to f_{\mathcal{PS}}(u)$.

Now define

$$A_\varepsilon = \left\{(\omega, u) \in \overline{\mathcal{PS}} \times [0,1] : d_{\mathcal{PS}}(\omega, f_{\mathcal{PS}}(u)) \geq \varepsilon\right\}.$$

Since $\overline{\mathcal{PS}} \times [0,1]$ is compact and $f_{\mathcal{PS}}(\cdot)$ is continuous, A_ε is compact. Then the function $G(\omega, u) = F(\omega, u) - F(f_{\mathcal{PS}}(u), u)$ is continuous on A_ε. By uniqueness of the minimizer, $G(\omega, u) > 0$ for every $(\omega, u) \in A_\varepsilon$. A continuous strictly positive function on a compact set attains a strictly positive minimum. Hence (11) holds.

We then verify (12). Consider

$$\widetilde{F}(\omega, u) = K^{-1} \sum_{k=1}^{K} \widehat{w}(u_k, u) \mathbb{E}[d^2_{\mathcal{PS}}(\omega, \tilde{f}_{\mathrm{PS}}(u_k))]$$

as the weights are fixed rather than random. Observe that, under Assumptions A3 and A4,

$$\begin{aligned}\left|\widehat{F}(\omega, u) - F(\omega, u)\right| &\leq \left|\widehat{F}(\omega, u) - \widetilde{F}(\omega, u)\right| + \left|\widetilde{F}(\omega, u) - F(\omega, u)\right| \\ &\leq \left|\widehat{F}(\omega, u) - \widetilde{F}(\omega, u)\right| + O(h^2)\end{aligned}$$

as $h \to 0$ by borrowing similar arguments to those used in the proof of (S.13) in Chen & Müller (2022). This means to verify (12), it is sufficient to show that

$$\sup_{\omega \in \overline{\mathcal{PS}}} \sup_{u \in [0,1]} \left|\widehat{F}(\omega, u) - \widetilde{F}(\omega, u)\right| = o_p(1), \tag{13}$$

as $h \to 0$.

We first show that the pointwise convergence such that for any $\omega \in \overline{\mathcal{PS}}$ and $u \in [0,1]$,

$$\left|\widehat{F}(\omega, u) - \widetilde{F}(\omega, u)\right| = o_p(1). \tag{14}$$

The sphere $\mathcal{S}$ is of negative type (Lyons, 2020) and by Schoenberg (1937; 1938), there exist a Hilbert space $\mathcal{H}$ and an embedding map $\phi : \mathcal{S} \to \mathcal{H}$ such that $d_{\mathcal{S}}(\nu, \nu') = \|\phi(\nu) - \phi(\nu')\|^2_{\mathcal{H}}$. For $u \in [0,1]$, let $z(u) = \phi(\tilde{f}(u))$, and one should notice that $\phi(\cdot)$ is a continuous bijection between $(\mathcal{S}, d_{\mathcal{S}})$ and $(\phi(\mathcal{S}), d_{\mathcal{H}})$ with continuous inverse. Then, for any subset $\mathcal{I} \subseteq [0,1]$, the σ-algebra generated by $\{\tilde{f}(u) : u \in \mathcal{I}\}$ is the same as the one generated by $\{z(u) : u \in \mathcal{I}\}$. This means that if $\{\tilde{f}(u)\}_{u \in [0,1]}$ is α-mixing, then one always has $\{z(u)\}_{u \in [0,1]}$ is α-mixing with the same mixing coefficient. Following this discussion, for $\omega = (\omega_1, \ldots, \omega_L) \in \mathcal{PS}$ and $\tilde{f}_{\mathrm{PS}} = \left(\tilde{f}^{(1)}_{\mathrm{PS}}, \ldots, \tilde{f}^{(L)}_{\mathrm{PS}}\right)$,

$$\begin{aligned}\widehat{F}(\omega, u) =& K^{-1} \sum_{k=1}^{K} \widehat{w}(u_k, u) d^2_{\mathcal{PS}}(\omega, \tilde{f}_{\mathrm{PS}}(u)) = K^{-1} \sum_{k=1}^{K} \widehat{w}(u_k, u) \sum_{l=1}^{L} d^2_{\mathcal{S}}(\omega_l, \tilde{f}^{(l)}_{\mathrm{PS}}(u)) \\ =& K^{-1} \sum_{k=1}^{K} \widehat{w}(u_k, u) \sum_{l=1}^{L} \|\phi(\omega_l) - \phi(\tilde{f}^{(l)}_{\mathrm{PS}}(u))\|^4_{\mathcal{H}}\end{aligned}$$

where $\left\{\sum_{l=1}^{L} \|\phi(\omega_l) - \phi(\tilde{f}^{(l)}_{\mathrm{PS}}(u))\|^4_{\mathcal{H}}\right\}_{u \in [0,1]}$ is α-mixing under Assumption A2. Then, by the law of large numbers for the α-mixing sequence, see, e.g., Theorem 2.11 in Bosq (2000), (14) holds for any $\omega \in \overline{\mathcal{PS}}$ and $u \in [0,1]$.

We then need to verify the asymptotic equicontinuity in the probability of $\widehat{F}(\omega, u) - \widetilde{F}(\omega, u)$. That is, we aim to show that for any $\theta > 0$,

$$\limsup_K \mathbb{P}\left(\sup_{d_{\mathcal{PS}}(\omega_1, \omega_2) < \zeta_1} \sup_{|t_1 - t_2| < \zeta_2} \left|\widehat{F}(\omega_1, t_1) - \widetilde{F}(\omega_1, t_1) - \widehat{F}(\omega_2, t_2) + \widetilde{F}(\omega_2, t_2)\right| > \theta\right) \to 0, \tag{15}$$

as $\zeta_1, \zeta_2 \to 0$ and $h \to 0$. Observe that

$$\begin{aligned}
&\sup_{d_{\mathcal{PS}}(\omega_1,\omega_2)<\zeta_1} \sup_{|t_1-t_2|<\zeta_2} \left|\widehat{F}(\omega_1,t_1) - \widetilde{F}(\omega_1,t_1) - \widehat{F}(\omega_2,t_2) + \widetilde{F}(\omega_2,t_2)\right| \\
\leq &\sup_{d_{\mathcal{PS}}(\omega_1,\omega_2)<\zeta_1} \sup_{|t_1-t_2|<\zeta_2} \left|\widehat{F}(\omega_1,t_1) - \widetilde{F}(\omega_1,t_1) - \widehat{F}(\omega_1,t_2) + \widetilde{F}(\omega_1,t_2)\right| \\
&+ \sup_{d_{\mathcal{PS}}(\omega_1,\omega_2)<\zeta_1} \sup_{|t_1-t_2|<\zeta_2} \left|\widehat{F}(\omega_1,t_2) - \widetilde{F}(\omega_1,t_2) - \widehat{F}(\omega_2,t_2) + \widetilde{F}(\omega_2,t_2)\right| \\
= &\sup_{\omega\in\overline{\mathcal{PS}}} \sup_{|t_1-t_2|<\zeta_2} \left|\widehat{F}(\omega,t_1) - \widetilde{F}(\omega,t_1) - \widehat{F}(\omega,t_2) + \widetilde{F}(\omega,t_2)\right| \\
&+ \sup_{d_{\mathcal{PS}}(\omega_1,\omega_2)<\zeta_1} \sup_{u\in[0,1]} \left|\widehat{F}(\omega_1,u) - \widetilde{F}(\omega_1,u) - \widehat{F}(\omega_2,u) + \widetilde{F}(\omega_2,u)\right|.
\end{aligned}$$

This indicates that to verify (15), it is sufficient to show for any $\theta > 0$,

$$\limsup_K \mathbb{P}\left(\sup_{\omega\in\overline{\mathcal{PS}}} \sup_{|t_1-t_2|<\zeta} \left|\widehat{F}(\omega,t_1) - \widehat{F}(\omega,t_2)\right| > \theta\right) \to 0, \tag{16}$$

$$\sup_{\omega\in\overline{\mathcal{PS}}} \sup_{|t_1-t_2|<\zeta} \left|\widetilde{F}(\omega,t_1) - \widetilde{F}(\omega,t_2)\right| \to 0, \tag{17}$$

and

$$\limsup_K \mathbb{P}\left(\sup_{d_{\mathcal{PS}}(\omega_1,\omega_2)<\zeta} \sup_{u\in[0,1]} \left|\widehat{F}(\omega_1,u) - \widetilde{F}(\omega_1,u) - \widehat{F}(\omega_2,u) + \widetilde{F}(\omega_2,u)\right| > \theta\right) \to 0, \tag{18}$$

as $\zeta \to 0$ and $h \to 0$. Observe that (17) holds as $\{f(u)\}_{u\in[0,1]}$ is assumed to be continuous. Therefore, we only need to verify (16) and (18).

Regarding (16), note that

$$\begin{aligned}
\left|\widehat{F}(\omega,t_1) - \widehat{F}(\omega,t_2)\right| &= \left|K^{-1}\sum_{k=1}^{K}\{\widehat{w}(u_k,t_1) - \widehat{w}(u_k,t_2)\}\, d^2_{\mathcal{PS}}(\omega,\tilde{f}(u_k))\right| \\
&\leq L\pi\,|\widehat{w}(u,t_1) - \widehat{w}(u,t_2)|
\end{aligned} \tag{19}$$

for any $u \in [0,1]$. First, observe that under Assumption A3, $\mathcal{K}_h'(u) = \frac{d}{du}\mathcal{K}_h(u) = \frac{1}{h^2}\mathcal{K}'(\frac{u}{h})$ and $\frac{\partial}{\partial u}\mathcal{K}_h(x_i - u) = -\mathcal{K}_h'(x_i - u)$. Moreover, we also have

$$\frac{\partial}{\partial u}\widehat{\tau}_l(u) = K^{-1}\sum_{k=1}^{K}\left[-\mathcal{K}_h'(u_k-u)(u_k-u)^l - l\mathcal{K}_h(u_k-u)(u_k-u)^{l-1}\right].$$

This implies that $\widehat{\tau}_l(u) = O_p(h^l)$ and $\widehat{\tau}_l'(u) = \frac{\partial}{\partial u}\widehat{\tau}_l(u) = O_p(h^{l-1})$ when $h \to 0$ and $Kh \to \infty$. Finally, recall that $\widehat{\sigma}^2(u) = \widehat{\tau}_0(u)\widehat{\tau}_2(u) - \widehat{\tau}_1^2(u)$ and thus $\frac{\partial}{\partial u}\widehat{\sigma}^2 = \widehat{\tau}_0'\widehat{\tau}_2 + \widehat{\tau}_0\widehat{\tau}_2' - 2\widehat{\tau}_1\widehat{\tau}_1'$. We can conclude that $\widehat{\sigma}^2(u) = O_p(h^2)$ and $\frac{\partial}{\partial u}\widehat{\sigma}^2 = O_p(h)$. The above analysis on $\mathcal{K}_h(u), \widehat{\tau}_l(u)$, and $\widehat{\sigma}^2$, in conjunction with the mean value theorem, gives

$$|\widehat{w}(u,t_1) - \widehat{w}(u,t_2)| \leq Ch^{-2}|t_1 - t_2| \tag{20}$$

for some constant C by Assumption A3. This, together with (19), implies that (16) holds as $\zeta \to 0$ with $\zeta = o(h^2)$ and $h \to 0$.

For (18), note that, as $[0,1]$ is compact and $\overline{\mathcal{PS}}$ is totally bounded,

$$\begin{aligned}
&\left|\widehat{F}(\omega_1,u) - \widetilde{F}(\omega_1,u) - \widehat{F}(\omega_2,u) + \widetilde{F}(\omega_2,u)\right| \\
\leq &\left|\widehat{F}(\omega_1,u) - \widehat{F}(\omega_2,u)\right| + \left|\widetilde{F}(\omega_1,u) - \widetilde{F}(\omega_2,u)\right| \\
= &O_p(d_{\mathcal{PS}}(\omega_1,\omega_2)) + O(d_{\mathcal{PS}}(\omega_1,\omega_2)),
\end{aligned}$$

where the last equality holds by utilizing techniques similar to those used in the proof of Lemma 2 in Petersen & Müller (2019). This indicates that (18) holds.

Combining (14) and (15), by Theorem 1.5.4 in Van Der Vaart & Wellner (1996), we get (12) as $h \to 0$. This completes the proof. □

I ADDITIONAL EXPERIMENT INFORMATION AND RESULTS

In this appendix we fill the missing details in the experiment section. Our experiments on real-world datasets are implemented with an Intel Core i7 CPU with 32GB of memory for dissimilarity-matrix-based methods and with an NVIDIA H200 GPU of 141 GB memory for neural-network-based methods. The experiments on synthetic datasets are implemented with a high-performance computer with 200 Cascade Lake CPU cores for dissimilarity-matrix-based methods and with two NVIDIA H200 GPUs of 141 GB memory for neural-network-based methods.

I.1 EUCLIDEAN PROCRUSTES DISSIMILARITY

In classical Euclidean shape analysis (Srivastava et al., 2010; Joshi et al., 2016; Srivastava & Klassen, 2016), the standard procedure consists of two main steps: (i) removing the effects of translation and uniform scaling to obtain pre-shapes, and (ii) quotient out rotation. It is worth noting that step (ii) is optional in certain scenarios. Furthermore, when computing dissimilarity scores using only step (i), the resulting measure is equivalent to the ESD used as one of our baselines.

Below we provide details of the full procedure (i) and (ii) used in the baseline EPD. Suppose we have two spherical trajectories $f_{\mathrm{S},1}(u), f_{\mathrm{S},2}(u) \in \mathcal{S}^{m-1}$, $u \in [0,1]$, and suppose these two trajectories share the aligned index points $u_1, u_2, \ldots, u_K$. Construct two matrices $X_1, X_2 \in \mathbb{R}^{K\times m}$ where the k-th rows of X_1, X_2 are $f_{\mathrm{S},1}(u_k)$ and $f_{\mathrm{S},2}(u_k)$, respectively. The translation is first removed using the centering matrix $C = I_K - \frac{1}{K}1_K 1_K^\top$, where I_K is the $K \times K$ identity matrix and 1_K is a vector of K ones, yielding centered configurations $X_{c,1} = CX_1$ and $X_{c,2} = CX_2$. Uniform scaling is then removed by dividing each centered configuration by its Frobenius norm $\|\cdot\|_F$, resulting in the pre-shape matrices Z_1 and Z_2 as $Z_1 = X_{c,1}/\|X_{c,1}\|_F, Z_2 = X_{c,2}/\|X_{c,2}\|_F$. To remove rotational differences, we calculate the optimal rotation matrix $\widehat{R} \in SO(m)$ that minimizes the Frobenius norm between the pre-shapes. Here $SO(m)$ is the set of all $m \times m$ real orthogonal matrices with a determinant of 1. This is equivalent to maximizing the cross-covariance trace

$$\widehat{R} = \arg\max_{R\in SO(m)} \mathrm{tr}(Z_2^\top Z_1 R)$$

In the literature, this can be solved based the singular value decomposition of the cross-covariance matrix $C = Z_2^\top Z_1$. Finally, the Euclidean Procrustes shape distance is calculated as $\|Z_1 - Z_2\widehat{R}\|_F$.

I.2 CLASSIFICATION BASED ON DISSIMILARITY MATRIX

Suppose we have n product-spherically embedded trajectories $f_{\mathrm{PS},1}, f_{\mathrm{PS},2}, , \ldots, f_{\mathrm{PS},n}$, the dissimilarity matrix $D = (D_{ij})_{n\times n}$ is constructed by calculating the $(i,j)^{\text{th}}$ element using a chosen dissimilarity measure such as $D_{ij} = D_{\mathrm{PS}}(f_{\mathrm{PS},i}, f_{\mathrm{PS},j})$. Let $\mathcal{O}_n = \{(f_{\mathrm{PS},i}, \theta_i)\}_{i=1}^n$ denote the training sample, where $\theta_i \in \Theta = \{1, \ldots, C\}$ is the label. For a query observation $f_{\mathrm{PS},*}$, we obtain a vector of dissimilarities to the training set such as $D^* = [D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},1}), D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},2}), \ldots, D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},n})]^\top$. Let (i) denote the index of the i-th smallest distance in D^*, such that $D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},(1)}) \le D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},(2)}) \le \cdots \le D_{\mathrm{PS}}(f_{\mathrm{PS},*}, f_{\mathrm{PS},(n)})$. For a given k, the empirical estimate of the conditional class probability is given by $\widehat{H}(c, f_{\mathrm{PS},*}) = k^{-1}\sum_{j=1}^k \mathbb{I}(\theta_{(j)} = c)$ where $\mathbb{I}(\cdot)$ is the indicator function. The predicted label is obtained via $\widehat{\theta}^* = \arg\max_{c\in\times} \widehat{H}(c, f_{\mathrm{PS},*})$. Moreover, we determine the optimal hyperparameter k through LOOCV within the training set.

I.3 INFORMATION ABOUT CNN AND MANIFOLDNET

For each spherically embedded trajectory f_{S}, we directly embedded it as a matrix whose k^{th} row is $f_{\mathrm{S}}(u_k)$. If the trajectory f_{PS} is product-spherically embedded, the k^{th} row of the matrix is constructed as $(f_{\mathrm{PS}}^{(1)}(u_k), f_{\mathrm{PS}}^{(2)}(u_k), \ldots, f_{\mathrm{PS}}^{(L)}(u_k))$. The input observation then has dimension $\mathbb{R}^{K\times mL}$.

CNN. The feature extraction module consists of two consecutive convolutional blocks. The first block applies 16 filters of size 3×3 with zero-padding, followed by a ReLU activation and a 2×2 max-pooling operation. The second block applies 32 filters of size 3×3, followed by a ReLU activation. To ensure the architecture is invariant to arbitrary input dimensions and to minimize the parameter

count, we utilize a Global Adaptive Average Pooling (GAP) layer. The GAP layer aggressively aggregates the spatial dimensions of the 32 feature maps, yielding a fixed-length representation vector in $\mathbb{R}^{32}$. The final classification is performed by a multi-layer perceptron comprising a hidden layer of dimension 64, regularized via Dropout ($p = 0.3$), which projects the embeddings to the output class logits. The network is trained end-to-end utilizing the standard cross-entropy loss. We optimize the network weights using the Adam optimizer with a learning rate of $\eta = 10^{-3}$ and a mini-batch size of 16. Given the moderate sample size and the high capacity of neural architectures, robust regularization is strictly enforced. Beyond the structural regularization provided by the GAP layer and Dropout, we implement an early stopping criterion during training. Validation loss is monitored continually, and training is halted if no improvement is observed for a patience of 15 epochs. The model weights corresponding to the lowest validation loss are subsequently restored for inference.

ManifoldNet. To take the geometric structure of the data into account, we consider a geometric convolutional neural network, building upon the ManifoldNet framework. For matrix-valued time series where each temporal observation lies on a (product) sphere, the model initializes a distinct, learnable reference point for each of the L spherical components. During the forward pass, the multi-component input trajectories are simultaneously projected onto their respective Euclidean tangent spaces via Riemannian logarithmic map. These independent tangent vectors are subsequently concatenated to form a unified Euclidean representation at each time step. Temporal dependencies are then extracted using a series of 2D convolutional layers equipped with asymmetrical 3×1 kernels, ensuring operations are restricted exclusively to the temporal dimension without conflating the distinct geometric spaces. The temporal sequence is compressed via adaptive average pooling and passed through a fully connected multi-layer perceptron with dropout regularization to yield final class probabilities. To guarantee that the multiple learnable reference points and network parameters strictly adhere to the product manifold surface during backpropagation, the entire architecture is optimized end-to-end using the Riemannian Adam algorithm (Bécigneul & Ganea, 2019), implemented via the geometric optimization library Geoopt (Kochurov et al., 2020).

To ensure determinism and eliminate run-to-run variance for the CNN and ManifoldNet baselines, all random seeds governing weight initialization and stochastic gradient descent were explicitly fixed.

I.4 Computational cost

Table 3 reports the average run time for synthetic data. This table demonstrates that once the smooth trajectories are obtained, the average run time for the proposed PSSD measure and the subsequent classification procedure remains highly scalable. The most time-consuming step is the smoothing procedure, as a rolling-window cross-validation approach is used to select the bandwidth h. For a dataset with n time series where each time series has K time points, suppose the computational complexity of fitting the local Fréchet regression for a fixed bandwidth is $O(\jmath(K))$, the overall computational complexity of the smoothing procedure is then $O(nK\jmath(K))$. It is worth noting that the current smoothing procedure is intrinsic, and its computational cost could be significantly reduced by adopting an extrinsic approach.

Table 3: Average run time for synthetic data with sample size $n \in \{30, 90\}$ and the sphere dimension $m \in \{6, 101\}$ under the random splitting procedure.

Setting	Smoothing	PSSD	SD	SPD	ESD	ED	EPD	CNN	ManifoldNet
$n = 30, m = 6$	14 min	2 sec	0.04 sec	3 sec	0.03 sec	0.04 sec	0.2 sec	0.6 sec	0.1 sec
$n = 90, m = 6$	42 min	16 sec	0.1 sec	25 sec	0.9 sec	0.2 sec	1 sec	0.6 sec	2 sec
$n = 30, m = 101$	49 min	26 sec	0.05 sec	1 min	0.04 sec	0.1 sec	3 sec	0.7 sec	2 sec
$n = 90, m = 101$	147 min	4 min	0.2 sec	12 min	0.2 sec	0.8 sec	27 sec	3 sec	25 sec

I.5 Synthetic data generation

Data generation consists of three steps where we first generate random noises following the spherical autoregressive (AR) structure in Zhu & Müller (2024), construct three baseline spherical shapes, and finally generate spherical functional data.

Spherical noise process generation. We construct a stationary sequence of skew-symmetric operators via a first-order AR structure. Let $\mu_\epsilon = I_m/\|I_m\|_E$ denote the baseline reference direction for the noise, where I_m is the m-dimensional vector of ones for $m \in \{6, 101\}$. We initialize the sequence by drawing ϵ_1 from a von Mises–Fisher distribution, $\epsilon_1 \sim \text{vMF}(\mu_\epsilon, 5m)$, and obtain the initial skew-symmetric operator $\widetilde{\Xi}_1 = \epsilon_1 \ominus \mu_\epsilon$. For subsequent iterations $k = 2, \ldots, 700$, the process evolves according to $\widetilde{\Xi}_k = 0.5\widetilde{\Xi}_{k-1} + \widetilde{\epsilon}_k$, where $\widetilde{\epsilon}_k = e_k \ominus \mu_\epsilon$, with each e_k drawn independently from $\text{vMF}(\mu_\epsilon, 20m)$. To ensure the simulated operators achieve stationarity, we discard the initial 500 steps as a burn-in period, retaining only $\Xi_k = \widetilde{\Xi}_{k+500}$ for $k = 1, \ldots, 200$. Finally, the corresponding spherically-valued noise sequence is recovered by projecting these operators back onto the sphere via the exponential map: $\{\epsilon_k = \text{Exp}(\Xi_k)\mu_\epsilon\}_{k=1}^{200}$.

Baseline spherical shape construction. We formulate three baseline trajectories on the sphere, denoted as $\{f_{\text{S},\ell}\}_{\ell=1}^3$. For a given vector $w \in \mathbb{R}^m$, we compute a unit-length direction vector within the tangent space at μ_ϵ via the projection $v_{\text{dir}} = (I_{m\times m} - \mu_\epsilon\mu_\epsilon^\top)w/\|(I_{m\times m} - \mu_\epsilon\mu_\epsilon^\top)w\|_{\mathcal{H}}$, where $I_{m\times m}$ represents the identity matrix. To simulate movement along the manifold, we define a time-varying tangent velocity $v(u) = \kappa(u - 0.5)v_{\text{dir}}$. Here, the drift intensity κ governs the total angular distance traversed across the geodesic. We then map this tangent vector back onto the sphere using the standard spherical exponential map to generate the baseline shape

$$f_{\text{S},\ell}(u) = \cos(\|v(u)\|_{\mathcal{H}})\mu_\epsilon + \sin(\|v(u)\|_{\mathcal{H}})\frac{v(u)}{\|v(u)\|_{\mathcal{H}}}.$$

The three baselines are distinguished by their specific configurations for w and κ. For $f_{\text{S},1}$, we set $\kappa = 0.3$ and construct w such that its first three coordinates are 1 and all others are 0. For $f_{\text{S},2}$, we apply $\kappa = -0.3$ and assign a value of 1 to indices $\{1, 2, 6\}$ of w. Finally, for $f_{\text{S},3}$, we set $\kappa = -0.1$ with non-zero elements located at indices $\{1, 2, 5\}$ of w.

Final spherical functional data generation. We simulate datasets of size $n \in \{30, 90\}$, evenly partitioned into three equal classes corresponding to the baseline shapes $f_{\text{S},\ell}$. Specifically, we assign labels $\ell = 1$ for subjects $i = 1, \ldots, n/3$, $\ell = 2$ for $i = n/3+1, \ldots, 2n/3$, and $\ell = 3$ for $i = 2n/3+1, \ldots, n$. For each class, we first use Algorithm 1 to generate a trajectory $\tilde{\tilde{f}}_{\text{S},i} = \text{SPS_Cal}(f_{\text{S},\ell}, \alpha_i, \mu_\epsilon)$ where the scaling parameter α_i is drawn independently from a uniform distribution $U(0, 1)$. These trajectories are generated over an equidistant grid of 200 points, $u_1, u_2, \ldots, u_{200} \in [0, 1]$. Following the framework of Xu & Shang (2026), we introduce random spherical noise to construct the final observed trajectories $\tilde{f}_{\text{S},i}(u_k) = \text{Exp}\big(\tilde{\tilde{f}}_{\text{S},i}(u_k) \ominus \mu_\epsilon\big)\epsilon_{k,i}$, where the noise components $\epsilon_{k,i}$ are generated independently for each subject i using the stationary spherical noise process detailed above.

I.6 Real-World Data

In our real-world data analysis, we consider compositional, distributional, and product-spherical functional data summarized in Table 4.

Table 4: Summary of real-world datasets.

Dataset	Application	Data Type	Sample Size
ECC	energy consumption	compositional time series	65
LTDC	human mortality	distributional time series	39
MultiGait	human gait	product-spherical time series	300

Below we provide detailed information for the three real datasets.

Compositional functional data. To demonstrate our methodology on compositional trajectories, we analyze the global Energy Consumption Composition (ECC) dataset from the Energy Institute via https://www.energyinst.org/statistical-review/resources-and-data-downloads, which provides discrete annual energy records for 65 countries or areas spanning from 1965 to 2024. For each nation, we first construct a discrete compositional time series by calculating the relative fractional shares of six energy sources relative to the total net consumption in each year. Here, the macro-categories of energy sources are coal, natural gas, petroleum, nuclear, conventional hydroelectric, and renewables. At any given discrete time point, this

vector resides in the five-dimensional simplex Δ^5. As discussed in Section 3, we apply the pointwise square root transformation to these compositional time series, mapping them to the positive orthant of the sphere $\mathcal{S}^5$. The underlying smooth trajectory of each compositional time series is obtained via utilizing the methodologies in Section 4.2 based on normalizing the calendar years to a continuous time domain $[0, 1]$.

We utilize membership in the Organisation for Economic Co-operation and Development (OECD) as a natural labeling mechanism to validate our shape-based classification where 33 countries or areas are OECD members in this dataset. This dichotomy serves as a robust proxy for the macroeconomic and regulatory forces driving global energy transitions. Typically, OECD member states possess mature economies characterized by saturated energy demand and stringent decarbonization mandates. Consequently, the evolutionary shapes of their compositional trajectories are expected to reflect a deliberate substitution effect such as a definitive contraction in historical fossil fuel shares offset by a rapid, policy-driven expansion in renewables. Conversely, non-OECD nations predominantly represent industrializing economies experiencing surging aggregate energy demand. We assess the classification accuracy of the OECD versus non-OECD dichotomy using the dissimilarity matrix derived from our proposed PSSD measure. This empirical evaluation tests whether our spherical shape framework successfully encodes the fundamentally divergent topologies of global energy transitions.

Distributional functional data. For our analysis of distributional functional data, we evaluate male Life-Table Death Counts (LTDC) sourced from the Human Mortality Database (https://www.mortality.org). The dataset encompasses 39 countries or areas tracked annually from 1990 to 2018. To achieve higher precision than the standard reported counts (Japanese Mortality Database, 2025), we reconstruct the discrete age-distributional time series by recalculating the estimated death counts (d_x) to six decimal places. This is computed utilizing the probability of dying (q_x) against a fixed life-table radix of 100,000 at age zero, strictly bounding the terminal age group (110+) to one. Normalizing these counts across the 111 age groups $(0, 1, \ldots, 110+)$ yields a discrete probability distribution for each year. We then apply the functional square root transformation to map these discrete distributions onto the Hilbert sphere. We then normalize the calendar years to the continuous time domain $[0, 1]$ and smooth the discrete temporal observations into continuous trajectories using the methodology detailed in Section 4.2. This yields our final continuous spherical functional data.

To validate our shape-based classification, we partition these trajectories into post-communist and non-post-communist cohorts where the number of the post-communist nations or areas is 10 in this dataset. This categorization acts as a proxy for identifying structural shocks in demographic evolution. During the 1990s, post-communist nations or areas endured severe socioeconomic restructuring, triggering a well-documented mortality crisis characterized by unprecedented spikes in working-age male death rates, followed by a sharp recovery in the 2000s. Consequently, we expect the evolutionary shapes of their mortality distributions to exhibit high volatility and non-monotonic shifts. In contrast, non-post-communist nations typically demonstrate smoother, monotonic trajectory shapes driven by steady medical advancements and the gradual aging of the population. We hypothesize that our spherical embedding will successfully distinguish between these fundamentally divergent evolutionary topologies.

Product-spherical functional data. To evaluate our methodology on product-spherical functional data, we analyze the Multivariate Gait (MultiGait) dataset, originally collected by Shorter et al. (2008), preprocessed by Helwig et al. (2011), and publicly disseminated by Helwig et al. (2016) via https://archive.ics.uci.edu/dataset/760/multivariate+gait+data. The study captures the biomechanics of 10 subjects walking on a treadmill under three experimental conditions covering unencumbered normal walking, walking with an immobilized right knee via a brace, and walking with an immobilized right ankle. Each subject performed 10 replicate trials per condition. During each trial, the angular positions of the ankle, knee, and hip joints for both legs were uniformly sampled across 101 discrete time points. Geometrically, each joint angle naturally embeds as a two-dimensional spherical vector. Consequently, the complete lower-body pose at any given time point resides on a product sphere constructed from the Cartesian product of these 6 individual spheres. We then map the discrete temporal sequences to a normalized domain $[0, 1]$ and reconstruct them into continuous curves using the methodology in Section 4.2. This yields a final dataset of 300 continuous product-spherical functional trajectories. The downstream objective is multi-class classification which leverages the intrinsic geometric shape of these trajectories to accurately predict

the subject's walking condition. For the train-test split, we need to ensure at least one subject with all the replicates and conditions is in the test group only.

For real-world data, the random split procedure will be repeated 100 times. The classification performance across our three distinct real-world datasets, summarized in Table 5, demonstrates the robust discriminative power of the proposed PSSD measure. In the ECC dataset, PSSD achieves the highest accuracy, notably outperforming all the baselines. This performance gain validates our hypothesis that the evolutionary shapes of energy structure transitions possess unique topological features that standard spherical or Euclidean metrics fail to fully capture. For the LTDC and MultiGait datasets, our method consistently matches or exceeds the top-performing baselines, reaching perfect classification accuracy in the MultiGait multi-class task. The high accuracy in the LTDC dataset confirms that the PSSD measure effectively encodes the structural changes of post-communist mortality trajectories. Overall, these results indicate that by accounting for spherical translation and scaling invariance, the PSSD measure provides a more refined representation of the intrinsic geometric evolution of functional data, making it a superior metric for identifying complex, non-Euclidean group dynamics in demographic, economic, and biomechanical applications.

Table 5: Evaluation results on classification for three real-world datasets. All results are in percentages (%).

Split	Datasets	PSSD	SD	SPD	ESD	ED	EPD	CNN	ManifoldNet
Random Split	ECC	**73.59**	71.65	53.47	71.29	70.47	55.41	70.59	64.71
	LTDC	**95.64**	95.45	86.09	90.62	90.62	82.36	70.00	90.00
	MultiGait	**100.00**	100.00	91.67	100.00	100.00	78.33	98.67	98.67
LOOCV	ECC	**78.46**	70.77	41.54	64.62	67.69	52.31	64.62	70.77
	LTDC	**97.44**	94.87	84.62	97.44	97.44	84.62	74.36	74.36
	MultiGait	**100.00**	100.00	100.00	100.00	100.00	99.00	45.67	58.33

A comparison of the classification performance produced by the PSSD and SPD measures shows a marked decrease for the SPD measure on the ECC and LTDC datasets, consistent with the discussion of when not to quotient out rotation in Appendix C. Importantly, MultiGait provides reliable ground-truth labels for distinct walking conditions. It serves as a necessary real-world sanity check, confirming that when underlying pre-shapes are genuinely and substantially different, PSSD successfully and accurately detects this difference. Moreover, in shape analysis, if the underlying pre-shapes are radically different, the raw configurations often exhibit massive differences as well. In the MultiGait dataset, the structural differences are so pronounced that they completely overpower nuisance variations like translation and scale. Consequently, naive baselines (SD, ESD, and ED) also output large distances and achieve perfect accuracy, but they do so accidentally, as observed in Table 2 for synthetic data.

I.7 Additional results

To evaluate the necessity of our smoothing procedure in Section 4.2, we conduct an ablation study by directly applying the classification framework to the raw observations.

Synthetic data. The results of the analysis on synthetic datasets are presented in Table 6. A direct comparison with the smoothed results in Table 2 reveals a consistent degradation in classification accuracy when the smoothing step is omitted. This performance drop underscores a fundamental characteristic of spherical shape analysis in which raw empirical observations are inherently corrupted by stochastic noise. These findings empirically validate the necessity of the smoothing methodology for reliably evaluating product-spherically embedded trajectories.

Real-world data. Table 7 presents the classification results when evaluating the raw, unsmoothed real-world datasets. Compared with Table 5, we observe that applying the smoothing technique yields a distinct performance improvement for the ECC and MultiGait datasets. This confirms that for noisier real-world applications, extracting the smooth trajectories is a critical step for maximizing the discriminative power of the PSSD measure.

Table 6: Evaluation results on classification for different synthetic data settings with sample size $n \in \{30, 90\}$ and the sphere dimension $m \in \{6, 101\}$ without smoothing. All results are in percentages (%).

Split	Setting	PSSD	SD	SPD	ESD	ED	EPD	CNN	ManifoldNet
Random Split	$n = 30, m = 6$	**91.50**	81.44	81.62	81.19	81.69	82.25	37.50	44.79
	$n = 90, m = 6$	**95.00**	86.37	86.48	86.09	87.39	87.04	73.91	69.57
	$n = 30, m = 101$	**93.38**	85.00	85.19	85.62	86.38	85.81	62.50	37.50
	$n = 90, m = 101$	**98.30**	90.80	90.91	90.63	91.26	91.48	39.13	82.61
LOOCV	$n = 30, m = 6$	**92.92**	82.93	82.05	82.85	83.90	82.36	46.74	45.02
	$n = 90, m = 6$	**94.26**	88.01	86.87	87.96	89.01	87.30	51.11	47.78
	$n = 30, m = 101$	**95.22**	87.40	85.70	87.67	88.08	86.12	46.99	43.93
	$n = 90, m = 101$	**97.43**	91.83	91.37	91.76	92.27	91.79	31.50	38.67

Table 7: Evaluation results on classification for three real-world datasets without smoothing. All results are in percentages (%).

Split	Datasets	PSSD	SD	SPD	ESD	ED	EPD	CNN	ManifoldNet
Random Split	ECC	71.71	**72.24**	56.12	71.59	72.24	54.00	64.71	76.47
	LTDC	95.27	95.64	79.09	95.45	**97.18**	81.18	70.00	90.00
	MultiGait	96.67	98.33	88.33	95.00	**100.00**	73.33	96.00	98.67
LOOCV	ECC	75.38	**80.00**	61.54	73.85	69.23	47.69	63.08	70.77
	LTDC	**97.44**	94.87	79.49	94.87	97.44	84.62	74.36	74.36
	MultiGait	**100.00**	100.00	100.00	100.00	100.00	99.00	46.33	65.00