



HAL
open science

Computing regularized splines in the Riemannian manifold of probability measures

Tien Tam Tran, Ines Adouani, Chafik Samir

► **To cite this version:**

Tien Tam Tran, Ines Adouani, Chafik Samir. Computing regularized splines in the Riemannian manifold of probability measures. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2025, 59 (1), pp.73-99. 10.1051/m2an/2024056 . hal-04875435

HAL Id: hal-04875435

<https://hal.science/hal-04875435v1>

Submitted on 8 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMPUTING REGULARIZED SPLINES IN THE RIEMANNIAN MANIFOLD OF PROBABILITY MEASURES

TIEN TAM TRAN¹, INES ADOUANI² AND CHAFIK SAMIR^{3,*} 

Abstract. In this paper, we give new numerical methods to solve C^2 splines for interpolating a finite set of probability measures. Due to the difficulties of generalizing interpolations and solving them with explicit expressions on manifolds, we first introduce a theoretical analysis with different structures on the Riemannian manifold of probability measures $\mathcal{P}_+(I)$. Then, by combining standard numerical methods and a new mathematical modeling we generalize interpolating splines as a minimizer of an energy functional on $\mathcal{P}_+(I)$. Finally, we show the efficiency of the proposed methods and their computational relevance with several applications including real medical data.

Mathematics Subject Classification. 49Q20, 53C05, 49K20, 49M05, 47N40, 65K10.

Received April 20, 2023. Accepted July 4, 2024.

1. INTRODUCTION

Interpolation problems of data lying on non-linear manifolds arise in different topics and scientific applications [3, 5, 7, 17]. While this problem has successful numerical solutions on Euclidean spaces \mathbb{R}^n , much effort has been directed towards rigorous theoretical analysis and efficient numerical methods to encompass the variation within the class of manifolds. In this paper, we introduce a novel mathematical model to fit a given set of probability measures $\mu^{(1)}, \dots, \mu^{(N)} \in \mathcal{P}_+(I)$, where $\mathcal{P}_+(I)$ represents the space of probability measures on a sample finite space I , defined roughly by

$$\mathcal{P}_+(I) = \left\{ \mu = \sum_{i \in I} \mu_i \delta^i \mid \mu_i > 0, \quad \forall i \in I, \text{ and } \sum_{i \in I} \mu_i = 1 \right\}.$$

We formulate splines and give details for rigorous theoretical analysis on $\mathcal{P}_+(I)$ by exploiting the Riemannian structure of the space of probability measures which helps make the solution computationally tractable. It can be seen broadly, as an exploration of curves of measures, which are smooth and in some sense minimize the mean acceleration of the fitting curve while interpolating data, as splines do in Euclidean space.

To make the problem and setting more concrete, consider $\mu^{(0)}, \dots, \mu^{(N)}$ as an indexed set of probability measures associated to a set of observation time t_0, \dots, t_N . Our purpose is to estimate a spline $\sigma : [t_0, t_N] \rightarrow$

Keywords and phrases. Regularized minimization, smooth splines, interpolations, probability measures, Riemannian manifold.

¹ International School, Vietnam National University, Hanoi, Vietnam.

² Institut Supérieur des Sciences Appliquées et de Technologie de Sousse, Sousse, Tunisia.

³ University of Clermont Auvergne, LIMOS, Clermont-Ferrand, France.

*Corresponding author: chafik.samir@uca.fr

$\mathcal{P}_+(I)$ that minimizes the following functional

$$E(\gamma) = \frac{\lambda}{2} \int_{t_0}^{t_N} \left\langle \frac{D^2\sigma(t)}{Dt^2}, \frac{D^2\sigma(t)}{Dt^2} \right\rangle_F + \frac{1}{2} \sum_{i=0}^N d_F^2(\sigma(t_i), \mu^{(i)}), \quad (1)$$

where ($\lambda > 0$) is a smoothing parameter, $\langle \cdot, \cdot \rangle_F$ and d_F denote the Fisher–Rao metric and the geodesic distance on $\mathcal{P}_+(I)$. It is well-known that in the Euclidean setting the solution to this optimization problem under the interpolation constraint $\sum_{i=0}^N d_F^2(\sigma(t_i), \mu^{(i)}) = 0$ are cubic splines. A vast number of methods have appeared in the literature to solve the problem (1) for general Riemannian manifolds, which were motivated by applications in many areas of engineering. In most cases, the Riemannian manifold is simply a Lie group or a sphere.

These approaches includes variational interpretation [25], subdivision schemes [8, 19, 27], Lie-algebraic methods [24], intrinsic polynomial regression [11], extrinsic local regression [18], and global and local Fréchet regression [21]. In this framework, we seek σ , solution of (1) on $\mathcal{P}_+(I)$, to satisfy the following properties: (i) $\sigma(t_i) = \mu^{(i)}$, (ii) σ is of class C^2 . None of these previous methods can be used for constructing a C^2 interpolating splines on $\mathcal{P}_+(I)$. To overcome such limitations, we introduce a generalization of Bézier spline on Riemannian manifolds and construct an interpolation solution of the optimization problem (1) unique to the setting of measures. We focus on Bézier spline interpolations due to their simplicity, flexibility, theoretical guarantees, and mean acceleration minimizing property. Note that various works have been introduced to construct interpolating curves on Riemannian manifolds of class C^1 . Nevertheless, the problem of piecing generalized Bézier curves into a C^2 spline is much more complicated, except for a few cases such as compact Lie groups and spheres [22].

In [23], a C^2 continuity condition was developed on a more general Riemannian manifold. The solution requires the computation of velocities and covariant derivatives at each joint point. Unfortunately, computing the covariant derivative on a general Riemannian manifold is intricate because it brings up the inverse of the derivative of the Riemannian exponential map. In [10], Geir *et al.* simplify the C^2 continuity condition given in [23] in some special cases of Riemannian symmetric spaces. The present work extends this toolbox by developing a theoretical study of measure interpolation problems and seeking a C^2 interpolating Bézier spline on $\mathcal{P}_+(I)$ that reacts to the geometry of the manifold.

Recently much attention has been focused on the space of probability measures $\mathcal{P}_+(I)$ with different metrics including Frobenius, Fisher–Rao, log-Euclidean, Jensen-Shannon and Wasserstein metrics. Works on linear regression [1], transport of statistical model [13, 16], estimation [9], barycenters [14], have been deeply studied and led to computational advances in statistical analysis [15, 26]. Furthermore, the interpolation problem on the space of probability measures $\mathcal{P}_+(I)$ equipped with Wasserstein metrics was well considered concurrently and independently by several authors, see for instance [2, 4, 6]. To our knowledge, this work is the first to address the interpolation problem of the space of probability measures $\mathcal{P}_+(I)$ equipped with Fisher–Rao metric, in the spline context. This metric, remarkably important, has proved immensely fruitful in the statistical analysis of Riemannian manifolds.

To tackle the optimization problem on $\mathcal{P}_+(I)$, we formulate the Riemannian geometric structures of $\mathcal{P}_+(I)$ embedded with the Fisher–Rao metric. We obtain an explicit expression of the Christoffel symbols, and therefore the Levi–Civita connection associated with the Fisher–Rao metric which allows the computation of the geodesic curves joining two points on $\mathcal{P}_+(I)$. In this way, we exhibit an exact equation of the Levi–Civita parallel transport of a tangent vector along a geodesic curve joining two probability measures on $\mathcal{P}_+(I)$. Furthermore, we prove that $\mathcal{P}_+(I)$ is a locally symmetric space. Hence, taking into account these nice results, we present our algorithm to generate a measure interpolating spline of class C^2 on $\mathcal{P}_+(I)$. We prove the effectiveness of our method with potential applications.

This paper is organized as follows. In Section 2, we derive the basic information geometry structure of the space of probability measures $\mathcal{P}_+(I)$ equipped with Fisher–Rao metric. We present several useful theorems on the Levi–Civita connection, minimal geodesics, and parallel transport with detailed proofs. In Section 3, we

show how to take advantage of the results of Section 2 to construct a C^2 Bézier spline on the set of probability measures. In Section 4, the effectiveness of the proposed method is demonstrated by showing several examples of interpolation Bézier spline on $\mathcal{P}_+(I)$. Finally, conclusions and suggested future research are provided in Section 5.

2. A RIEMANNIAN STRUCTURE FOR PROBABILITY MEASURES

In this section, we formulate the Riemannian geometric structures of the space of probability measures on a sample finite space I embedded with the Fisher–Rao metric. In particular, we derive computational tools of interest, namely Levi–Civita connection, geodesics, and parallel transport. For a more detailed exposition of these concepts, see [12].

2.1. Fisher–Rao

Let $I = \{1, \dots, n, n+1\}$, $n \in \mathbb{N}$, be a finite sample space. Let $\mathcal{F}(I) = \{f : I \rightarrow \mathbb{R}\}$ be the algebra of real functions on it. Its unity function \mathcal{K}_I or simply \mathcal{K} is given by $\mathcal{K}(i) = 1$. A canonical basis of $\mathcal{F}(I)$ is defined by

$$e_i(j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

and hence, every $f \in \mathcal{F}(I)$ has the representation

$$f = \sum_{i \in I} f^i e_i, \quad (3)$$

where $f^i = f(i)$. We will denote by $\mathcal{S}(I)$ the dual space of $\mathcal{F}(I)$, the space of \mathbb{R} -valued linear forms on $\mathcal{F}(I)$. With the Riesz representation theorem, this vector space is interpreted as the vector space of signed finite measures on I , namely

$$\mathcal{S}(I) = \left\{ \mu : \mathcal{F}(I) \rightarrow \mathbb{R} \mid \mu = \sum_{i \in I} \mu_i \delta^i \right\}, \quad (4)$$

where $\mu_i = \mu(e_i)$ and δ^i is the Dirac measure supported at $i \in I$. It is also shown that $\mathcal{S}(I)$ is a manifold. In particular, the tangent space at $\mu \in \mathcal{S}(I)$ is given by

$$T_\mu \mathcal{S}(I) = \{\mu\} \times \mathcal{S}(I). \quad (5)$$

Let us consider the following submanifolds of $\mathcal{S}(I)$:

$$\mathcal{S}_\epsilon(I) = \left\{ \mu = \sum_{i \in I} \mu_i \delta^i \mid \sum_{i \in I} \mu_i = \epsilon, \quad \epsilon \in \mathbb{R} \right\}$$

and

$$\mathcal{M}_+(I) = \{\mu \in \mathcal{S}(I) \mid \mu_i > 0, \quad \forall i \in I\}$$

the space of finite strictly positive measures on I .

Definition 2.1. A probability measure on a finite sample space I is a map $\mu : I \rightarrow \mathbb{R}$ defined for any $A \subset I$ by $\mu(A) = \sum_{i \in A} \mu_i$ and which satisfies:

- (1) For all $i \in I$, $\mu_i \geq 0$ and $\mu(\emptyset) = 0$.
- (2) $\sum_{i \in I} \mu_i = 1$.

(3) $\mu(\{i\}) = \mu_i$.

We denote by $\mathcal{P}_+(I)$ the space of strictly positive probability measures on I ,

$$\mathcal{P}_+(I) = \left\{ \mu = \sum_{i \in I} \mu_i \delta^i \mid \mu_i > 0, \quad \forall i \in I, \text{ and } \sum_{i \in I} \mu_i = 1 \right\}.$$

We check at once that $\mathcal{P}_+(I) \subset \mathcal{M}_+(I) \subset \mathcal{S}(I)$. Therefore, as an open submanifold of $\mathcal{S}(I)$, $\mathcal{M}_+(I)$ has the same tangent space at the point $\mu \in \mathcal{M}_+(I)$. $\mathcal{P}_+(I)$ is a submanifold of $\mathcal{S}(I)$, and clearly, for $\mu \in \mathcal{P}_+(I)$, we have:

$$T_\mu \mathcal{P}_+(I) = \{\mu\} \times \mathcal{S}_0(I) = \left\{ (\mu, v) \mid \mu \in \mathcal{P}_+(I) \text{ and } v = \sum_{i \in I} v_i \delta^i \in \mathcal{S}_0(I) \right\}.$$

We want to endow $\mathcal{P}_+(I)$ with a Riemannian metric. To this end, we define a local coordinate map on $\mathcal{P}_+(I)$. Let U be an open set of \mathbb{R}^n given by

$$U = \left\{ x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i > 0, \quad \forall i \in I, \text{ and } \sum_{i=1}^n x_i < 1 \right\}.$$

We define a map φ as

$$\begin{aligned} \varphi : \mathcal{P}_+(I) &\longrightarrow U, \\ \mu = \sum_{i \in I} \mu_i \delta^i &\longmapsto (\varphi^1(\mu), \dots, \varphi^n(\mu)) = (x^1(\mu), \dots, x^n(\mu)), \end{aligned}$$

such that $(\varphi^1(\mu), \dots, \varphi^n(\mu)) = (\mu_1, \dots, \mu_n)$. Clearly, φ is an homomorphism and its inverse is given by

$$\begin{aligned} \varphi^{-1} : U &\longrightarrow \mathcal{P}_+(I), \\ (x_1, \dots, x_n) &\longmapsto \mu = \sum_{i=1}^n x_i \delta^i + \left(1 - \sum_{i=1}^n x_i \right) \delta^{n+1}. \end{aligned}$$

Given a point $\mu \in \mathcal{P}_+(I)$, let $\frac{\partial}{\partial x^i} \Big|_\mu$ be the tangent vector at μ given by

$$\frac{\partial}{\partial x^i} \Big|_\mu = \frac{\partial}{\partial x_i} \Big|_{\varphi(\mu)} \varphi^{-1} = (\delta^i - \delta^{n+1}), \quad \text{for } i = 1, \dots, n.$$

Thus, $\left\{ \frac{\partial}{\partial x^i} \Big|_\mu, i = 1, \dots, n \right\}$ define a local frame field of $T_\mu \mathcal{P}_+(I)$ at a point $\mu \in \mathcal{P}_+(I)$. Similarly we can define the dual basis of $\frac{\partial}{\partial x^i} \Big|_\mu$, the basis of the cotangent bundle $T_\mu^* \mathcal{P}_+(I) = \{\mu\} \times (\mathcal{F}(I)/\mathbb{R})$ by $dx^i = e_i + \mathbb{R}$, $i = 1, \dots, n$.

Remark. Let $\mu \in \mathcal{P}_+(I)$ and $v = \sum_{i \in I} v_i \delta^i \in T_\mu \mathcal{P}_+(I)$. It can be easily seen that

$$v = \sum_{i=1}^{n+1} v_i \delta^i = \sum_{i=1}^n v_i \delta^i - \sum_{i=1}^n v_i \delta^{n+1} = \sum_{i=1}^n v_i (\delta^i - \delta^{n+1}) = \sum_{i=1}^n v_i \frac{\partial}{\partial x^i}, \quad (6)$$

since $v \in \mathcal{S}_0(I)$.

$S(I)$ is a finite-dimensional linear space, and therefore, it can be naturally equipped with a metric. For $v, w \in T_\mu S(I)$, we define the inner product as

$$\langle v, w \rangle_\mu = \mu \left(\frac{dv}{d\mu} \cdot \frac{dw}{d\mu} \right) = \sum_i \frac{v_i w_i}{\mu_i} \quad (7)$$

where $\frac{dv}{d\mu} = \sum_{i \in I} \frac{v_i}{\mu_i} e_i \in \mathcal{F}(I)$, represents a simple version of the Radon–Nikodym derivative with respect to μ . This metric induces a metric on $\mathcal{M}_+(I)$. The probability manifold $\mathcal{P}_+(I)$ as a submanifold of $\mathcal{M}_+(I)$, is endowed with the Fisher–Rao metric. Hence, following the geometry structures in $\mathcal{M}_+(I)$ equipped with Fisher information metric, we derive the corresponding ones in $\mathcal{P}_+(I)$.

Definition 2.2. Let μ be a probability measure in $\mathcal{P}_+(I)$. Given two tangents vectors v and w in $T_\mu \mathcal{P}_+(I)$, the Fisher–Rao metric $\mathfrak{g}_\mu : T_\mu \mathcal{P}_+(I) \times T_\mu \mathcal{P}_+(I) \rightarrow \mathbb{R}$ is defined by

$$\mathfrak{g}_\mu(v, w) = \sum_{i \in I} \frac{v_i w_i}{\mu_i},$$

and $\|v\|_\mu = \sqrt{\mathfrak{g}_\mu(v, v)}$. With respect to the coordinate map $(\mathcal{P}_+(I), \varphi)$, the Fisher–Rao metric is expressed as

$$g_{ij}(\mu) = \begin{cases} \frac{1}{\mu_i} + \frac{1}{\mu_{i+1}}, & \text{if } i = j, \\ \frac{1}{\mu_{n+1}}, & \text{otherwise,} \end{cases}$$

for $i, j = 1, \dots, n$. And the components of the inverse matrix are given by

$$g^{ij}(\mu) = \begin{cases} \mu_i(1 - \mu_i), & \text{if } i = j, \\ -\mu_i \mu_j, & \text{otherwise.} \end{cases}$$

Our goal to make $\mathcal{P}_+(I)$ as a Riemannian manifold is fully satisfied. Our next goal is to compute explicit expressions of geometric structures on $\mathcal{P}_+(I)$, which will be essential to make our proposed solution of problem (1) on $\mathcal{P}_+(I)$.

2.2. Levi–Civita connection on $\mathcal{P}_+(I)$

Let $\mathcal{X}(\mathcal{P}_+(I))$ denote the set of smooth vector fields on $\mathcal{P}_+(I)$. Essentially, at each point $\mu \in \mathcal{P}_+(I)$, the Levi–Civita connection associated with the Fisher–Rao metric $\nabla : \mathcal{X}(\mathcal{P}_+(I)) \times \mathcal{X}(\mathcal{P}_+(I)) \rightarrow \mathcal{X}(\mathcal{P}_+(I))$ gives a new vector field, notated $\nabla_X Y$, telling us how the vector field Y is changing in the direction X and satisfying for all $X, Y, Z \in \mathcal{X}(\mathcal{P}_+(I))$,

$$\begin{cases} X \mathfrak{g}(Y, Z) = \mathfrak{g}(\nabla_X Y, Z) + \mathfrak{g}(Y, \nabla_X Z), \\ \nabla_X Y - \nabla_Y X = [X, Y]. \end{cases} \quad (8)$$

In the local coordinate map $(\mathcal{P}_+(I), \varphi)$, the Levi–Civita connection is defined by the Christoffel symbols $\Gamma_{ij}^k : \mathcal{P}_+(I) \rightarrow \mathbb{R}$ such that

$$\nabla_{\partial x_i} \partial x_j = \Gamma_{ij}^k \partial x_k. \quad (9)$$

Proposition 2.1. *With respect to the local coordinate map $(\mathcal{P}_+(I), \varphi)$, the Christoffel symbols associated with the Fisher–Rao metric are given by*

$$\Gamma_{ij}^k = \begin{cases} \frac{1}{2} \frac{x_k}{1 - \sum_{h=1}^n x_h}, & i \neq j, \\ \frac{1}{2} \frac{x_k}{1 - \sum_{h=1}^n x_h} + \frac{1}{2} \frac{x_k}{x_i}, & i = j \neq k, \\ \frac{1}{2} \frac{x_k}{1 - \sum_{h=1}^n x_h} - \frac{1}{2} \frac{1 - x_k}{x_k}, & i = j = k, \end{cases} \quad (10)$$

Proof. The smooth functions Γ_{ij}^k are easily computed through the characterization of the Levi–Civita connection by the Koszul formula obtained from (8) computed for all the circular permutations of $X, Y, Z \in \mathcal{X}(\mathcal{P}_+(I))$,

$$\mathfrak{g}(\nabla_X Y, Z) = \frac{1}{2} \{X\mathfrak{g}(Y, Z) + Y\mathfrak{g}(Z, X) - Z\mathfrak{g}(X, Y) + \mathfrak{g}([X, Y], Z) - \mathfrak{g}([Y, Z], X) - \mathfrak{g}([X, Z], Y)\}. \quad (11)$$

Now, in the Koszul formula we set $X = \partial x_i, Y = \partial x_j, Z = \partial x_l$. We get

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l=1}^n g^{kl} (g_{il,j} + g_{jl,i} - g_{ij,l}), \quad \text{for } i, j, k \in \{1, \dots, n\}, \quad (12)$$

where $g_{il,j} = \frac{\partial g_{il}}{\partial x_j}$, $g_{jl,i} = \frac{\partial g_{jl}}{\partial x_i}$, and $g_{ij,l} = \frac{\partial g_{ij}}{\partial x_l}$. In the local coordinate system, the Fisher–Rao metric and its inverse are given by

$$g_{ij} = \begin{cases} \frac{1}{x_i} + \frac{1}{1 - \sum_{h=1}^n x_h}, & \text{if } i = j, \\ \frac{1}{1 - \sum_{h=1}^n x_h}, & \text{if } i \neq j, \end{cases} \quad (13)$$

$$g^{ij} = \begin{cases} x_i(1 - x_i), & \text{if } i = j, \\ -x_i x_j, & \text{if } i \neq j, \end{cases} \quad (14)$$

for $i, j = 1, \dots, n$. Now if we take the derivative of (13) by x_l , we get

$$g_{ij,l} = \begin{cases} -\frac{1}{(x_i)^2} + \frac{1}{(1 - \sum_{h=1}^n x_h)^2}, & \text{if } i = j = l, \\ \frac{1}{(1 - \sum_{h=1}^n x_h)^2}, & \text{otherwise.} \end{cases} \quad (15)$$

Replace (15) in (12), the formula follows. \square

Definition 2.3. Let $X \in \mathcal{X}(\mathcal{P}_+(I))$ be a vector field on $\mathcal{P}_+(I)$. Then in the local coordinate $(\mathcal{P}_+(I), \varphi)$, we have the representation $X = \sum_{i=1}^n X_i \partial x_i$. X is called a constant vector field on $\mathcal{P}_+(I)$ if all X_i are independent of μ .

Theorem 2.1. *Given two constant vector fields X, Y on $\mathcal{P}_+(I)$, the Levi–Civita connection at $\mu \in \mathcal{P}_+(I)$ is given by*

$$\nabla_X Y(\mu) = -\frac{1}{2} \left(\frac{dX}{d\mu} \frac{dY}{d\mu} - \mathfrak{g}_\mu(X, Y) \right) \mu. \quad (16)$$

Proof. Let $X = \sum_{i \in I} X_i \delta^i, Y = \sum_{i \in I} Y_i \delta^i$ and $Z = \sum_{i \in I} Z_i \delta^i$ be constant vector fields on $\mathcal{P}_+(I)$. Thus, we get $[X, Y] = [Y, Z] = [X, Z] = 0$ and consequently (11) gives

$$\mathfrak{g}(\nabla_X Y, Z) = \frac{1}{2} \{X\mathfrak{g}(Y, Z) + Y\mathfrak{g}(X, Z) - Z\mathfrak{g}(X, Y)\}. \quad (17)$$

Set $\mu = \sum_{i \in I} \mu_i \delta^i \in \mathcal{P}_+(I)$, and $\gamma(t) = \mu + vt$, a curve on $\mathcal{P}_+(I)$ such that $\mu(0) = \mu$ and $\dot{\mu}(0) = v = X(\mu)$. We have

$$\begin{aligned} X\mathfrak{g}_\mu(Y, Z) &= \frac{d}{dt} \Big|_{t=0} \mathfrak{g}_{\mu(t)}(Y, Z) \\ &= \frac{d}{dt} \Big|_{t=0} \sum_{i \in I} \frac{Y_i Z_i}{\mu_i + tv_i} \end{aligned}$$

$$= - \sum_{i \in I} \frac{v_i Y_i Z_i}{\mu_i^2} = - \sum_{i \in I} \frac{X_i Y_i Z_i}{\mu_i^2}.$$

Similarly, one obtains formulae for $Y \mathfrak{g}(X, Z)$ and $Z \mathfrak{g}(X, Y)$. Now replacing the above results in (17), we get

$$\begin{aligned} \mathfrak{g}_\mu(\nabla_X Y, Z) &= \frac{1}{2} \left\{ - \sum_{i \in I} \frac{X_i Y_i Z_i}{\mu_i^2} - \sum_{i \in I} \frac{X_i Y_i Z_i}{\mu_i^2} + \sum_{i \in I} \frac{X_i Y_i Z_i}{\mu_i^2} \right\} \\ &= - \frac{1}{2} \sum_{i \in I} \frac{X_i Y_i Z_i}{\mu_i^2}. \end{aligned} \quad (18)$$

On the other hand, we have

$$\sum_{i \in I} \mathfrak{g}_\mu(X, Y) Z_i = \mathfrak{g}_\mu(X, Y) \sum_{i \in I} Z_i = 0, \quad (19)$$

since Z is a vector field on $\mathcal{P}_+(I)$. Then (18) can be written as

$$\begin{aligned} \mathfrak{g}_\mu(\nabla_X Y, Z) &= - \frac{1}{2} \sum_{i \in I} \left(\frac{X_i Y_i}{\mu_i^2} - \mathfrak{g}_\mu(X, Y) \right) \mu_i \frac{Z_i}{\mu_i} \\ &= \mathfrak{g}_\mu \left(- \frac{1}{2} \left(\frac{dX}{d\mu} \frac{dY}{d\mu} - \mathfrak{g}_\mu(X, Y) \right) \mu, Z \right). \end{aligned}$$

This holds for every constant vector field Z , which completes the proof. \square

2.3. Geodesics curves on \mathcal{P}_+

Theorem 2.2. *Let μ be a probability measure in $\mathcal{P}_+(I)$ and $v \in T_\mu \mathcal{P}_+(I)$ a unit tangent vector, i.e., $\|v\|_\mu = 1$. Then the geodesic γ that satisfies $\gamma(0) = \mu$ and $\frac{\partial \gamma}{\partial t}|_{t=0} = v$ has the form represented by $\gamma(t) = \sum_{i \in I} \gamma_i(t) \delta^i$ with*

$$\gamma_i(t) = \left(\cos \frac{t}{2} + \frac{\dot{\gamma}_i(0)}{\gamma_i(0)} \sin \frac{t}{2} \right)^2 \gamma_i(0), \quad (20)$$

where $\gamma_i(0) = \mu_i$ and $\dot{\gamma}_i(0) = v_i$, for $i = 1, \dots, n$.

Proof. Let $\gamma(t) = \sum_{i \in I} \gamma_i(t) \delta^i$ and $\dot{\gamma}(t) = \sum_{i \in I} \dot{\gamma}_i(t) \delta^i$. Then for each t , we have

$$\begin{cases} \sum_{i \in I} \gamma_i(t) = 1, & \text{and } \gamma_i(t) > 0, & \forall i \in I, \\ \sum_{i \in I} \dot{\gamma}_i(t) = 0. \end{cases} \quad (21)$$

Set X a constant vector field in $\mathcal{P}_+(I)$. From the condition (8) of Levi–Civita connection, we have

$$\mathfrak{g}_{\gamma(t)}(\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t), X) = \dot{\gamma}(t) (\mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), X)) - \mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), \nabla_{\dot{\gamma}(t)} X). \quad (22)$$

With the properties of Levi–Civita connection, to compute $\nabla_{\dot{\gamma}(t)} X$, the tangent vector $\dot{\gamma}(t)$ can be considered as a constant vector field on $\mathcal{P}_+(I)$ when t is fixed. Therefore, applying (16) for $\dot{\gamma}(t)$ and X we get,

$$\begin{aligned} \nabla_{\dot{\gamma}(t)} X &= - \frac{1}{2} \left(\frac{d\dot{\gamma}(t)}{d\gamma(t)} \frac{dX}{d\gamma(t)} - \mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), X) \right) \gamma(t) \\ &= - \frac{1}{2} \sum_{i \in I} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \frac{X_i}{\gamma_i(t)} - \sum_{j \in I} \frac{\dot{\gamma}_j(t) X_j}{\gamma_j(t)} \right) \gamma_i(t) \delta^i. \end{aligned} \quad (23)$$

Taking into account of (21), the last term in (22) becomes

$$\begin{aligned} \mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), \nabla_{\dot{\gamma}(t)} X) &= -\frac{1}{2} \sum_{i \in I} \frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \frac{X_i}{\gamma_i(t)} - \sum_{j \in I} \frac{\dot{\gamma}_j(t) X_j}{\gamma_j(t)} \right) \gamma_i(t) \\ &= -\frac{1}{2} \sum_{i \in I} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 X_i. \end{aligned} \quad (24)$$

Now, we compute the second term in (22). We have

$$\begin{aligned} \dot{\gamma}(t)(\mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), X)) &= \frac{d}{dt} \mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), X) \\ &= \sum_{i \in I} \frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) X_i. \end{aligned} \quad (25)$$

Combining (24) and (25) in (22), we get

$$\mathfrak{g}(\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t), X) = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 \right) X_i. \quad (26)$$

Let's define the function $C(t)$ as

$$\begin{aligned} C(t) &= - \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 \right) \gamma_i(t) \\ &= - \sum_{i \in I} \frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) \gamma_i(t) - \frac{1}{2} \mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)). \end{aligned} \quad (27)$$

Hence, the measure

$$v(t) = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 + C(t) \right) \gamma_i(t) \delta^i$$

belongs to $T_{\gamma(t)} \mathcal{P}_+(I)$. In this way, (26) can be written as

$$\mathfrak{g}(\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t), X) = \mathfrak{g}(v(t), X). \quad (28)$$

Since X is an arbitrary constant vector field, we get

$$\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = v(t) = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 + C(t) \right) \gamma_i(t) \delta^i. \quad (29)$$

Therefore, $\gamma(t) = \sum_{i \in I} \gamma_i(t) \delta^i$ is a geodesic if and only if

$$\begin{cases} \frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 + C(t) = 0, & \forall i \in I, \\ \sum_{i \in I} \dot{\gamma}_i(t) = 0, & \forall t. \end{cases} \quad (30)$$

Our next goal is to solve (30). We may remark that if γ is a geodesic then $\mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))$ is constant along $\gamma(t)$. Consequently, taking into account of the assumption that $\|\dot{\gamma}(0)\|_{\mu} = 1$, we can assert that

$$\mathfrak{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) = \sum_{i \in I} \frac{\dot{\gamma}_i^2(t)}{\gamma_i(t)} \equiv 1. \quad (31)$$

Thus

$$\sum_{i \in I} \frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) \gamma_i(t) = \frac{d}{dt} \sum_{i \in I} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \gamma_i(t) \right) - \sum_{i \in I} \left(\frac{\dot{\gamma}_i^2(t)}{\gamma_i(t)} \right) = -1. \quad (32)$$

Which gives that $C(t) = \frac{1}{2}$. Substituting this result in (30), we obtain

$$\frac{d}{dt} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right) + \frac{1}{2} \left(\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} \right)^2 + \frac{1}{2} = 0, \quad \forall i \in I. \quad (33)$$

Set $\omega_i(t) = \frac{\dot{\gamma}_i(t)}{\gamma_i(t)}$. Equation (33) is written as

$$\dot{\omega}_i(t) + \frac{1}{2} \omega_i^2(t) + \frac{1}{2} = 0, \quad \forall i \in I. \quad (34)$$

The solution of this differential equation is given by $\omega_i = \tan(-\frac{t}{2} + A^i)$, where A^i is constant, $i \in I$. Hence, we have

$$\frac{\dot{\gamma}_i(t)}{\gamma_i(t)} = \tan\left(-\frac{t}{2} + A^i\right), \quad \text{for } i \in I. \quad (35)$$

So $\gamma_i(t) = B^i \cos^2(-\frac{t}{2} + A^i)$, where B^i is constant, and $i \in I$. By the initial conditions, we find that

$$A^i = \arctan\left(\frac{\dot{\gamma}_i(0)}{\gamma_i(0)}\right), \quad (36)$$

$$B^i = \frac{\gamma_i^2(0) + \dot{\gamma}_i^2(0)}{\gamma_i(0)}. \quad (37)$$

Which proves the theorem. \square

Corollary 2.1. *The geodesic $\gamma(t)$ with $\gamma(0) = \mu$ and $\dot{\gamma}(0) = v$, where v is a nontrivial tangent vector (not necessary unit), is given by*

$$\gamma(t) = \sum_{i \in I} \left(\cos \frac{t \|v\|_\mu}{2} + \frac{v_i}{\mu_i \|v\|_\mu} \sin \frac{t \|v\|_\mu}{2} \right)^2 \mu_i \delta^i. \quad (38)$$

Proposition 2.2. *The Fisher–Rao distance $d^{FR} : \mathcal{P}_+(I) \times \mathcal{P}_+(I) \rightarrow [0, \pi)$ between two measures $\mu, \nu \in \mathcal{P}_+(I)$ under the Fisher–Rao metric is given by*

$$d^{FR}(\mu, \nu) = 2 \arccos \left(\sum_{i \in I} \sqrt{\mu_i \nu_i} \right). \quad (39)$$

To prove Proposition 2.2, we will show the following lemma given in [1].

Lemma 2.1. *Let*

$$\mathbb{S}_{(0,2),+}(I) = \left\{ f \in \mathcal{F}(I) \mid f^i > 0, \forall i \in I \text{ and } \sum_{i \in I} (f^i)^2 = 4 \right\}$$

be the positive sector of the sphere centered at 0 with radius 2. As a submanifold of $\mathcal{F}(I)$ it carries the induced standard metric of $\mathcal{F}(I)$. That is for a given point $f \in \mathbb{S}_{(0,2),+}(I)$ and two tangents vectors $p, q \in T_f \mathbb{S}_{(0,2),+}(I)$, we have

$$\langle p, q \rangle_f = \sum_{i \in I} p^i q^i. \quad (40)$$

Then the map $\Phi : \mathcal{P}_+(I) \longrightarrow \mathbb{S}_{(0,2),+}(I)$ defined by $\mu = \sum_{i \in I} \mu_i \delta^i \longmapsto 2 \sum_{i \in I} \sqrt{\mu_i} e_i$ is a global isometry.

Proof of the Lemma. It is clear that Φ is bijective. Now, let v, w be in $T_\mu\mathcal{P}_+(I)$. We have

$$\begin{aligned} \left\langle \frac{\partial\Phi}{\partial v}(\mu), \frac{\partial\Phi}{\partial w}(\mu) \right\rangle &= \left\langle \frac{d}{dt}\Phi(\mu + vt) \Big|_{t=0}, \frac{d}{dt}\Phi(\mu + wt) \Big|_{t=0} \right\rangle \\ &= \left\langle \sum_{i \in I} \frac{v_i}{\sqrt{\mu_i}} e_i, \sum_{i \in I} \frac{w_i}{\sqrt{\mu_i}} e_i \right\rangle \\ &= \sum_{i \in I} \frac{v_i w_i}{\mu_i} = \mathfrak{g}_\mu(v, w). \end{aligned}$$

□

Proof of the Proposition. By virtue of Lemma 2.1, we get $d^{FR}(\mu, \nu) = d(\Phi(\mu), \Phi(\nu)) = 2 \arccos(\sum_{i \in I} \sqrt{\mu_i \nu_i})$. □

Lemma 2.2. *Let $t \rightarrow \gamma(t)$ be a curve on $\mathcal{P}_+(I)$ and $v \in T_\mu\mathcal{P}_+(I)$ a vector field along γ . Let $\Phi : \mathcal{P}_+(I) \rightarrow \mathbb{S}_{(0,2),+}(I)$ be the Riemannian isometry, then*

(1) *The covariant derivative of a vector fields along γ is preserved, that is*

$$\Phi^*(D_t v) = D_t(\Phi^* v). \quad (41)$$

(2) *The Riemannian curvature endomorphism is isometry invariant, that is*

$$d\Phi_\mu R(v^{(1)}, v^{(2)}, v^{(3)}) = R(d\Phi_\mu v^{(1)}, d\Phi_\mu v^{(2)}, d\Phi_\mu v^{(3)}), \quad \forall v^{(i)} \in T_\mu\mathcal{P}_+(I), \quad i = 1, 2, 3. \quad (42)$$

Theorem 2.3. *Let μ, ν be two different probability measures in $\mathcal{P}_+(I)$. Then there exists a unique geodesic $\gamma : [0, l] \rightarrow \mathcal{P}_+(I)$, $t \rightarrow \gamma(t)$, joining two points μ and ν , with $\gamma(0) = \mu$, $\gamma(l) = \nu$ and $l = d^{FR}(\mu, \nu)$, given by*

$$\begin{aligned} \gamma(t) &= \left(\cos \frac{t}{2} + \sin \frac{t}{2} \frac{dv}{d\mu} \right)^2 \mu \\ &= \sum_{i \in I} \left(\cos \frac{t}{2} + \sin \frac{t}{2} \frac{dv}{d\mu}(i) \right)^2 \mu_i \delta^i, \end{aligned} \quad (43)$$

where v is the unit tangent vector in $T_\mu\mathcal{P}_+(I)$ defined by

$$v = \frac{1}{\sin \frac{l}{2}} \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) \mu_i \delta^i. \quad (44)$$

Proof. The proof falls naturally into three parts.

Step 1. First, let us check that v is a unit tangent in $T_\mu\mathcal{P}_+(I)$. Indeed,

$$\begin{aligned} &\frac{1}{\sin \frac{l}{2}} \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) \mu_i \\ &= \frac{1}{\sin \frac{l}{2}} \left(\sum_{i \in I} \sqrt{\frac{d\nu}{d\mu}}(i) \mu_i - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) \\ &= 0. \end{aligned} \quad (45)$$

Then, since

$$\left(\sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right)^2 = \left(\sum_{j \in I} \sqrt{\mu_j \nu_j} \right)^2 = \cos^2 \frac{l}{2}, \quad (46)$$

it follows that

$$\begin{aligned} \langle v, v \rangle_\mu &= \frac{1}{\sin^2 \frac{l}{2}} \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right)^2 \mu_i \\ &= \frac{1}{\sin^2 \frac{l}{2}} \left(\sum_{i \in I} \nu(i) - \left(\sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right)^2 \right) \\ &= \frac{1}{\sin^2 \frac{l}{2}} \left(1 - \cos^2 \frac{l}{2} \right) = 1 \end{aligned} \quad (47)$$

hence v is a unit tangent vector.

Step 2. Now let us examine that the curve $\gamma(t)$ defined by equation (43) satisfies $\gamma(0) = \mu$ and $\gamma(l) = \nu$. It is easily seen that for $t = 0$, $\gamma(0) = \mu$. Now for $t = l$, we have

$$\gamma(l) = \sum_{i \in I} \left(\cos \frac{l}{2} + \sin \frac{l}{2} \frac{dv}{d\mu}(i) \right)^2 \mu_i \delta^i. \quad (48)$$

By (44) we get

$$\begin{aligned} \sin \frac{l}{2} \frac{dv}{d\mu} &= \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) e_i \\ &= \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \cos \frac{l}{2} \right) e_i. \end{aligned} \quad (49)$$

Hence,

$$\gamma(l) = \sum_{i \in I} \left(\cos \frac{l}{2} + \sqrt{\frac{d\nu}{d\mu}}(i) - \cos \frac{l}{2} \right)^2 \mu_i \delta^i = \sum_{i \in I} \nu_i \delta^i = \nu. \quad (50)$$

Step 3. Now we go to prove the uniqueness of the curve. Let $\gamma(t)$ and $\tilde{\gamma}(t)$ be unit speed geodesics corresponding with tangent vectors v and \tilde{v} , and satisfying $\gamma(0) = \tilde{\gamma}(0) = \mu$ and $\gamma(l) = \tilde{\gamma}(l) = \nu$. By Theorem 2.2, we have

$$\gamma(t) = \sum_{i \in I} \left(\cos \frac{t}{2} + \frac{dv}{d\mu} \sin \frac{t}{2} \right)^2 \mu_i \delta^i, \quad (51)$$

$$\tilde{\gamma}(t) = \sum_{i \in I} \left(\cos \frac{t}{2} + \frac{d\tilde{v}}{d\mu} \sin \frac{t}{2} \right)^2 \mu_i \delta^i. \quad (52)$$

From later condition, we have

$$\left(\cos \frac{l}{2} + \frac{dv}{d\mu}(i) \sin \frac{l}{2} \right)^2 = \left(\cos \frac{l}{2} + \frac{d\tilde{v}}{d\mu}(i) \sin \frac{l}{2} \right)^2, \quad \forall i \in I \quad (53)$$

$$\Rightarrow \cos \frac{l}{2} + \frac{dv}{d\mu}(i) \sin \frac{l}{2} = \pm \left(\cos \frac{l}{2} + \frac{d\tilde{v}}{d\mu}(i) \sin \frac{l}{2} \right), \quad \forall i \in I. \quad (54)$$

Define

$$I_{\pm} = \left\{ i \in I \mid \cos \frac{l}{2} + \frac{dv}{d\mu}(i) \sin \frac{l}{2} = \pm \left(\cos \frac{l}{2} + \frac{d\tilde{v}}{d\mu}(i) \sin \frac{l}{2} \right) \right\}. \quad (55)$$

Then we have $I_- \cup I_+ = I$. Moreover $I_- \cap I_+ = \emptyset$. Indeed, if there exists $i \in I_- \cap I_+$ then

$$\nu_i = \left(\cos \frac{l}{2} + \frac{dv}{d\mu} \sin \frac{l}{2} \right)^2 \mu_i = 0, \quad (56)$$

contradict the fact that $\nu \in \mathcal{P}_+$. Since $0 < l < \pi$, we have

$$I_+ = \{i \in I \mid v_i = \tilde{v}_i\}, \quad (57)$$

$$I_- = \left\{ i \in I \mid v_i + \tilde{v}_i = -2\mu_i \cot \frac{l}{2} \right\}. \quad (58)$$

Suppose $I_- \neq \emptyset$, since v and \tilde{v} are unit tangent vectors at μ , we have

$$\sum_{i \in I_+} v_i + \sum_{i \in I_-} v_i = \sum_{i \in I_+} \tilde{v}_i + \sum_{i \in I_-} \tilde{v}_i = 0 \quad (59)$$

$$\Rightarrow \sum_{i \in I_-} \left(\tilde{v}_i + 2\mu_i \cot \frac{l}{2} \right) + \sum_{i \in I_-} \tilde{v}_i = 0. \quad (60)$$

Since (60) we see that if $I_- = I$, then $\cot \frac{l}{2} = 0$ contradicts to $0 < l < \pi$. So $I_- \neq I$. We have the claim below.

Claim 1. For all $\mu \in \mathcal{P}_+$ and $0 < l < \pi$. If $v, \tilde{v} \in T_{\mu}\mathcal{P}_+$. Let

$$I_+ = \{i \in I \mid v_i = \tilde{v}_i\}, \quad (61)$$

$$I_- = \left\{ i \in I \mid v_i + \tilde{v}_i = -2\mu_i \cot \frac{l}{2} \right\}, \quad (62)$$

then $I_- = \emptyset$.

By means of the claim, we prove the uniqueness of the geodesic (43) defined with the unit tangent vector (44). □

Proof of the Claim. We proof the claim by induction on the degree of I . If $|I|$ is one or two the claim is true since I_+ is not empty. Suppose the claim is true for $|I| = n$. We go to prove the claim for $|I| = n + 1$. Let μ, v, \tilde{v} and l as in the claim. Suppose $I_- \neq \emptyset$ then $|I_-| \geq 2$. Let g, h be two distinct index in I_- , this means $v_g + \tilde{v}_g = -2\mu_g \cot \frac{l}{2}$ and $v_h + \tilde{v}_h = -2\mu_h \cot \frac{l}{2}$. Now let $k \in I_+$ and define three measures v', \tilde{v}', μ' on $I \setminus \{k\}$ as follow

$$v' = \sum_{i \in I, i \neq k, h, g} v_i \delta^i + v_g \delta^g + (v_h + v_k) \delta^h, \quad (63)$$

$$\tilde{v}' = \sum_{i \in I, i \neq k, h, g} \tilde{v}_i \delta^i + (\tilde{v}_g + 2\tilde{v}_k) \delta^g + (\tilde{v}_h - \tilde{v}_k) \delta^h, \quad (64)$$

$$\mu' = \sum_{i \in I, i \neq k, h, g} \mu_i \delta^i + (\mu_g + \mu_k) \delta^g + \mu_h \delta^h. \quad (65)$$

We have $v', \tilde{v}' \in T_{\mu'}\mathcal{P}_+(I \setminus \{k\})$, and $h \in I_- \neq \emptyset$. This contradicts to the hypothesis. This shows the claim for $|I| = n + 1$. □

Corollary 2.2. Let $\varepsilon = \{(\mu, v) \mid \gamma(t, \mu, v) \text{ is defined on an interval containing } [0, l]\}$. The exponential map $\exp_\mu : \varepsilon \rightarrow \mathcal{P}_+(I)$ is defined as

$$\exp_\mu(v) = \sum_{i \in I} \left(\cos \frac{\|v\|_\mu}{2} + \frac{v_i}{\mu_i \|v\|_\mu} \sin \frac{\|v\|_\mu}{2} \right)^2 \mu_i \delta^i. \quad (66)$$

Similarly, given two points μ and ν on $\mathcal{P}_+(I)$, the inverse exponential map (also known as the logarithmic map) at μ is given as follows

$$\begin{aligned} \log_\mu : \mathcal{P}_+(I) &\rightarrow \varepsilon \\ \nu &\mapsto \log_\mu(\nu) = \frac{l}{\sin \frac{l}{2}} \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) \mu_i \delta^i. \end{aligned} \quad (67)$$

Theorem 2.4. $\mathcal{P}_+(I)$ equipped with the Fisher–Rao metric is a locally symmetric space.

To show that $\mathcal{P}_+(I)$ is a locally symmetric space, we need the following proposition.

Proposition 2.3. A locally symmetric space is a Riemannian manifold whose curvature tensor is parallel, i.e., $\nabla R = 0$.

Proof of Theorem 2.4. It follows by Lemma 2.2 as $\mathcal{P}_+(I)$ is isometric to the symmetric space $\mathbb{S}_{(0,2),+}(I)$. Hence, around any point $\mu \in \mathcal{P}_+(I)$ there is a geodesic ball $B_r(\mu)$ which is isometric to a ball $B_\varepsilon(f)$ on $\mathbb{S}_{(0,2),+}(I)$. Therefore, there is an isometry $\kappa_\mu : B_r(\mu) \rightarrow B_r(f)$ such that $\kappa_\mu(\mu) = f$ and the differential $d\kappa_\mu(\mu) = -Id$. Let $v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)} \in T_\mu(\mathcal{P}_+(I))$. Let us show that $\nabla_{v^{(1)}} R(v^{(2)}, v^{(3)}, v^{(4)}) = 0$. Set $v = \nabla_{v^{(1)}} R(v^{(2)}, v^{(3)}, v^{(4)})$ and applying $d\kappa_\mu$ on both sides of this equation, we obtain that v will be changed by $-v$ and $v^{(i)}$ to $-v^{(i)}$, for $i = 1, \dots, 4$. Thus the left-hand side changes sign while the right-hand side stays the same. Moreover, we mention that the curvature tensor is preserved by $d\kappa_\mu$. Consequently, $\nabla R = 0$ and the proof is complete. \square

2.4. Parallel transport on $\mathcal{P}_+(I)$

On the Riemannian manifold $\mathcal{P}_+(I)$, each tangent vector v belongs to a tangent space $T_\mu \mathcal{P}_+(I)$ specific to its root point μ . Hence tangent vectors from different tangent spaces cannot be compared directly. Parallel transport is a unique mathematical tool capable of transporting vectors between tangent spaces while retaining the information they contain. Let us consider two points $\mu, \nu \in \mathcal{P}_+(I)$, a tangent vector $v \in T_\mu \mathcal{P}_+(I)$ and a geodesic curve $\gamma : [0, l] \rightarrow \mathcal{P}_+(I)$ on $\mathcal{P}_+(I)$ such that $\gamma(0) = \mu$ and $\gamma(l) = \nu$. We would like to map v from $T_\mu \mathcal{P}_+(I) = T_{\gamma(0)} \mathcal{P}_+(I)$ to $T_\nu \mathcal{P}_+(I) = T_{\gamma(l)} \mathcal{P}_+(I)$. We introduce X , a vector field defined along the geodesic γ , such that $X(\mu) = v$ and $\nabla_{\dot{\gamma}(t)} X(\gamma(t)) = 0$. We say that the tangent vector v is constant along the geodesic curve γ with respect to ∇ .

Definition 2.4. A metric parallel transport on $\mathcal{P}_+(I)$ is the map

$$\Gamma_{\gamma(0) \rightarrow \gamma(t)} : T_{\gamma(0)} \mathcal{P}_+(I) \rightarrow T_{\gamma(t)} \mathcal{P}_+(I) \quad (68)$$

such that for any $v, w \in T_\mu \mathcal{P}_+(I)$, and for $t \in [0, l]$ we have

$$\mathfrak{g}_{\gamma(0)}(v, w) = \mathfrak{g}_{\gamma(t)}(\Gamma_{\gamma(0) \rightarrow \gamma(t)}(v), \Gamma_{\gamma(0) \rightarrow \gamma(t)}(w)). \quad (69)$$

And let Γ be the Levi–Civita parallel transport of a tangent vector v along a geodesic curve γ on $\mathcal{P}_+(I)$ with the Fisher–Rao metric.

Rewriting equation $\nabla_{\dot{\gamma}(t)}X(\gamma(t)) = 0$, we conclude that computing $X(t) = X(\gamma(t))$ requires solving a linear first order differential equations on $\mathcal{P}_+(I)$ given by

$$\frac{dX_k}{dt} + \sum_{i,j} \Gamma_{ij}^k \frac{d\gamma_i}{dt} X_j = 0, \quad \text{for } k = 1, \dots, n. \quad (70)$$

We check at once that it is difficult to solve equation (70) directly. Hence we will use equation (16).

Theorem 2.5. *Let μ be a probability measure in $\mathcal{P}_+(I)$ and $v \in T_\mu\mathcal{P}_+(I)$ a unit tangent vector, i.e., $\|v\|_\mu = 1$. Let $\gamma : [0, l] \rightarrow \mathcal{P}_+(I)$ be a geodesic curve such that $\gamma(0) = \mu$ and $\dot{\gamma}(0) = v$. The Levi-Civita parallel transport of a vector $w \in T_\mu\mathcal{P}_+(I)$ to $T_{\gamma(t)}\mathcal{P}_+(I)$, is given by*

$$\Gamma_{\gamma(0) \mapsto \gamma(t)}(w) = \sum_{i \in I} \sqrt{\gamma_i(t)} \left(-C(0) \sqrt{\mu_i} \left(2 \sin \frac{t}{2} - 2 \frac{v_i}{\mu_i} \cos \frac{t}{2} \right) + \frac{w_i}{\sqrt{\mu_i}} - 2C(0) \frac{v_i}{\sqrt{\mu_i}} \right) \delta^i, \quad (71)$$

where $C(0) = \frac{1}{2} \mathfrak{g}_\mu(v, w)$.

Proof. We can proceed analogously to the proof of Theorem 2.2. Thus, let $\gamma(t) = \sum_{i \in I} \gamma_i(t) \delta^i$ be a geodesic curve, and define $\dot{\gamma}(t) = \sum_{i \in I} \dot{\gamma}_i(t) \delta^i$. Consider the vector field X on γ defined by $X(\gamma(t)) = \sum_{i \in I} X_i(\gamma(t)) \delta^i$, for $t \in [0, l]$, as the parallel transport of vector w along γ . Then

$$\begin{cases} \nabla_{\dot{\gamma}(t)} X(t) = 0 \\ X(0) = w, \end{cases} \quad (72)$$

where we write $X(\gamma(t))$ simply $X(t)$ when no confusion can arise. Let Y be a constant vector field (in the sense of Def. 2.3) on $\mathcal{P}_+(I)$, we have

$$\mathfrak{g}_{\gamma(t)}(\nabla_{\dot{\gamma}(t)} X(t), Y) = \dot{\gamma}(t)(\mathfrak{g}_{\gamma(t)}(X(t), Y)) - \mathfrak{g}_{\gamma(t)}(X(t), \nabla_{\dot{\gamma}(t)} Y). \quad (73)$$

Applying Theorem 2.1, we get

$$\nabla_{\dot{\gamma}} Y = -\frac{1}{2} \sum_{i \in I} \left(\frac{\dot{\gamma}_i}{\gamma_i} \frac{Y_i}{\gamma_i} - \sum_{j \in I} \frac{\dot{\gamma}_j Y_j}{\gamma_j} \right) \gamma_i \delta^i. \quad (74)$$

Hence the last term in (73) becomes

$$\begin{aligned} \mathfrak{g}_\gamma(X, \nabla_{\dot{\gamma}} Y) &= -\frac{1}{2} \sum_{i \in I} \frac{X_i}{\gamma_i} \left(\frac{\dot{\gamma}_i}{\gamma_i} \frac{Y_i}{\gamma_i} - \sum_{j \in I} \frac{\dot{\gamma}_j Y_j}{\gamma_j} \right) \gamma_i \\ &= -\frac{1}{2} \sum_{i \in I} \frac{X_i Y_i \dot{\gamma}_i}{\gamma_i^2}. \end{aligned} \quad (75)$$

Let us now compute the second term in (73). We obtain

$$\dot{\gamma}(t)(\mathfrak{g}_{\gamma(t)}(X, Y)) = \frac{d}{dt} \mathfrak{g}_{\gamma(t)}(X(t), Y) = \sum_{i \in I} \frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) Y_i. \quad (76)$$

Consequently, equation (73) becomes

$$\mathfrak{g}_\gamma(\nabla_{\dot{\gamma}} X, Y) = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} \right) Y_i. \quad (77)$$

Define the function $C(t)$ by

$$\begin{aligned} C(t) &= - \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} \right) \gamma_i(t) \\ &= - \sum_{i \in I} \frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) \gamma_i(t) - \frac{1}{2} \mathfrak{g}_{\gamma(t)}(X(t), \dot{\gamma}(t)). \end{aligned} \quad (78)$$

Then, $\forall t \in [0, l]$, the probability measure

$$\nu(t) = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} + C(t) \right) \gamma_i \delta^i$$

belongs to $T_{\gamma(t)} \mathcal{P}_+(I)$. Thus, equation (77) can be written as

$$\mathfrak{g}_{\gamma}(\nabla_{\dot{\gamma}} X, Y) = \mathfrak{g}_{\gamma}(\nu, Y). \quad (79)$$

Since Y is an arbitrary constant vector field, we get

$$\nabla_{\dot{\gamma}} X = \nu = \sum_{i \in I} \left(\frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} + C(t) \right) \gamma_i \delta^i. \quad (80)$$

Therefore, $X(t)$ is the parallel transport of the vector w along the geodesic curve $\gamma(t)$ if and only if

$$\begin{cases} \frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} + C(t) = 0, & \forall i \in I, \\ X(0) = w. \end{cases} \quad (81)$$

Our next concern will be to solve equation (81). We remind that $\mathfrak{g}_{\gamma(t)}(X(t), \dot{\gamma}(t)) = \mathfrak{g}_{\gamma(0)}(X(0), \dot{\gamma}(0))$. Moreover

$$\sum_{i \in I} \frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) \gamma_i = \frac{d}{dt} \sum_{i \in I} \left(\frac{X_i}{\gamma_i} \gamma_i \right) - \sum_{i \in I} \left(\frac{X_i \dot{\gamma}_i}{\gamma_i} \right) = -\mathfrak{g}_{\gamma(0)}(X(0), \dot{\gamma}(0)). \quad (82)$$

Which gives that $C(t)$ is a constant function and $C(t) = C(0) = \frac{1}{2} \mathfrak{g}_{\gamma(0)}(X(0), \dot{\gamma}(0))$. Hence, substituting this result in equation (81) we get

$$\frac{d}{dt} \left(\frac{X_i}{\gamma_i} \right) + \frac{1}{2} \frac{X_i \dot{\gamma}_i}{\gamma_i^2} + C(0) = 0, \quad \forall i \in I. \quad (83)$$

Set $\omega_i = \frac{X_i}{\gamma_i}$. Equation (83) can be written as

$$\frac{d}{dt} \omega_i + \frac{1}{2} \frac{\dot{\gamma}_i}{\gamma_i} \omega_i + C(0) = 0, \quad \forall i \in I. \quad (84)$$

Solution of the first order differential equation (84) is given by

$$\omega_i(t) = \frac{1}{\sqrt{\gamma_i(t)}} \left(-C(0) \sqrt{\gamma_i(0)} \left(2 \sin \frac{t}{2} - 2 \frac{\dot{\gamma}_i(0)}{\gamma_i(0)} \cos \frac{t}{2} \right) + A_i \right), \quad \text{for } A_i \text{ constant, } i \in I. \quad (85)$$

Therefore,

$$X_i = \sqrt{\gamma_i(t)} \left(-C(0) \sqrt{\gamma_i(0)} \left(2 \sin \frac{t}{2} - 2 \frac{\dot{\gamma}_i(0)}{\gamma_i(0)} \cos \frac{t}{2} \right) + A_i \right), \quad \text{for } A_i \text{ constant, } i \in I. \quad (86)$$

According to the initial conditions, it follows that

$$A_i = \frac{w_i}{\sqrt{\mu_i}} - 2C(0) \frac{v_i}{\sqrt{\mu_i}}. \quad (87)$$

We conclude that

$$X_i(t) = \sqrt{\gamma_i(t)} \left(-C(0) \sqrt{\mu_i} \left(2 \sin \frac{t}{2} - 2 \frac{v_i}{\mu_i} \cos \frac{t}{2} \right) + \frac{w_i}{\sqrt{\mu_i}} - 2C(0) \frac{v_i}{\sqrt{\mu_i}} \right), \quad i \in I \quad (88)$$

and it is easy to check that, $\forall t \in [0, l]$, $X(t) = \sum_{i \in I} X_i(t) \delta^i \in T_{\gamma(t)} \mathcal{P}_+(I)$ and it is the Levi-Civita parallel transport of the vector w along the geodesic curve $\gamma(t)$. \square

Theorem 2.6. *Given two distinct probability measures μ and ν in $\mathcal{P}_+(I)$, a nontrivial tangent vector $w \in T_\mu \mathcal{P}_+(I)$ and a geodesic curve $\gamma : [0, l] \rightarrow \mathcal{P}_+(I)$ such that $\gamma(0) = \mu$ and $\gamma(l) = \nu$. The Levi-Civita parallel transport, $\Gamma_{\mu \rightarrow \nu} : T_\mu \mathcal{P}_+(I) \rightarrow T_\nu \mathcal{P}_+(I)$, that transports a vector w from $T_\mu \mathcal{P}_+(I) = T_{\gamma(0)} \mathcal{P}_+(I)$ to $T_\nu \mathcal{P}_+(I) = T_{\gamma(l)} \mathcal{P}_+(I)$ given by*

$$\Gamma_{\mu \rightarrow \nu}(w) = \sum_{i \in I} \sqrt{\nu_i} \left(-C(0) \sqrt{\mu_i} \left(2 \sin \frac{l}{2} - 2 \frac{v_i}{\mu_i} \cos \frac{l}{2} \right) + \frac{w_i}{\sqrt{\mu_i}} - 2C(0) \frac{v_i}{\sqrt{\mu_i}} \right) \delta^i, \quad (89)$$

where $l = 2 \arccos \sum_{i \in I} \sqrt{\mu_i \nu_i}$, $C(0) = \frac{1}{2} \mathfrak{g}_\mu(w, v)$, and v is the unit tangent vector

$$v = \frac{1}{\sin \frac{l}{2}} \sum_{i \in I} \left(\sqrt{\frac{d\nu}{d\mu}}(i) - \sum_{j \in I} \sqrt{\frac{d\nu}{d\mu}}(j) \mu(j) \right) \mu_i \delta^i. \quad (90)$$

Proof. It suffices to use the equation of the geodesic curve $\gamma(t)$ joining two points μ and ν given by Theorem 3.2 together with taking $t = l$ in Theorem 2.5, the proof follows. \square

In the next section, we will show how to take advantage of these results to represent our geometric algorithm that generates a C^2 Bézier spline on the space of probability measures $\mathcal{P}_+(I)$.

3. SPLINES ON SPACE OF PROBABILITY MEASURES

3.1. Problem formulation

Now, an interpolation spline on $\mathcal{P}_+(I)$ can be created using geodesic operations introduced in the previous section. A general formulation of the interpolation problem is given by: Given $\mu^{(0)}, \dots, \mu^{(N)}$ an indexed finite set of probability measures associated to a set of observation times t_0, \dots, t_N such that $t_i = i$, for $i = 0, \dots, N$. The aim is to develop a method for smooth interpolation on the space of probability measures $\mathcal{P}_+(I)$. More precisely, we seek a spline $\sigma : [t_0, t_N] \rightarrow \mathcal{P}_+(I)$ that interpolates the given set of probability measures: $\sigma(t_i) = \mu^{(i)}$, $i = 0, \dots, N$ and is of class C^2 . For simplicity we may consider the time interval $[0, N]$ instead of $[t_0, t_N]$. Solutions to this problem will be called measure interpolation spline and attempt to minimize the total cost functional,

$$E(\sigma) = \int_{t_0}^{t_N} \left\langle \frac{D^2 \sigma(t)}{Dt^2}, \frac{D^2 \sigma(t)}{Dt^2} \right\rangle_F dt.$$

To attack the problem of defining and computing splines of measures, we propose a new optimization method based on the generalized De Casteljau algorithm on manifolds. The implementation of the De Casteljau algorithm on $\mathcal{P}_+(I)$ provides the groundwork to obtain the solution of the optimization problem (1) for the probability measures space $\mathcal{P}_+(I)$. The resulting interpolating spline is known in the literature as the Bézier spline. Moreover, we will decouple the interpolation problem into two steps: First, solve a C^1 interpolation spline on $\mathcal{P}_+(I)$. Then, taking into account the local symmetries at interpolation points, we derive equations for control points that generate the requested C^2 Bézier spline. The effectiveness of the method is demonstrated by designing several measure interpolation splines on $\mathcal{P}_+(I)$ and applying them to various applications.

3.2. Measure interpolation spline using De Casteljau algorithm

The De Casteljau algorithm is a well-known algorithm used to generate polynomial spline curves on Euclidean spaces. The method is theoretically sound and excellent, due to its simple geometric construction based on simple successive linear interpolation. The generalizations of De Casteljau's algorithm on Riemannian manifolds are made in an obvious way by simply trading straight lines by minimal geodesic between two points, see [20]. Since the space of probability measures is not geodesically complete, thus an explicit formula for the geodesic that joins two points may be unknown in some particular situations. So, in this case, the implementation of the De Casteljau algorithm is restricted to a convex open subset of the manifold where the expression to compute the geodesic arc joining two points is well-defined.

From now on, we will consider $\sigma_j : [0, 1] \rightarrow \mathcal{P}_+(I)$ the Bézier curve of order j defined by $(j+1)$ control points $V^{(0)}, \dots, V^{(j)} \in \mathcal{P}_+(I)$. Set $(V^{(i)})^0(t) = V^{(i)}$, and for $i = 0, \dots, j-k, k = 1, \dots, j$ we define

$$\left(V^{(i)}\right)^k(t) = \sigma_k\left(t, V^{(i)}, \dots, V^{(i+k)}\right) = \gamma\left(t, \left(V^{(i)}\right)^{k-1}, \left(V^{(i+1)}\right)^{k-1}\right), \quad t \in [0, 1],$$

where γ is the geodesic curve on $\mathcal{P}_+(I)$ given by equation (20). In this framework, the Bézier spline $\sigma : [0, N] \rightarrow \mathcal{P}_+(I)$ will be constructed by a sequence of N Bézier curve σ_j^i of order two and three such that the first and the last ones are quadratic Bézier curves while all the others are cubic. Bézier curves of order j are polynomial functions parametrized by a set of control points such that the first and last control points are the endpoint of the curve, but the intermediate control points are in general not on the curve. Moreover, the number of control points determines the degree j of the polynomial spline. Additionally, let us denote $((\hat{\eta}^{(i)})^-, (\hat{\eta}^{(i)})^+)$ the control points on the left and on the right-hand side of the interpolation point $\mu^{(i)}$ for $i = 1, \dots, (N-1)$. The Bézier spline $\sigma : [0, N] \rightarrow \mathcal{P}_+(I)$ is then given by

$$\sigma(t) = \begin{cases} \sigma_2^0(t; \mu^{(0)}, (\hat{\eta}^{(1)})^-, \mu^{(1)}), & 0 \leq t \leq 1, \\ \sigma_3^i(t-i; \mu^{(i)}, (\hat{\eta}^{(i)})^+, (\hat{\eta}^{(i+1)})^-, \mu^{(i+1)}), & i-1 \leq t \leq i \\ \sigma_2^{N-1}(t-(N-1); \mu^{(N-1)}, (\hat{\eta}^{(N-1)})^+, \mu^{(N)}), & N-1 \leq t \leq N. \end{cases}$$

Since the Bézier spline σ interpolates the first and the last control points of each Bézier curve σ_j^i , $j \in \{2, 3\}, 0 \leq i \leq N-1$, therefore the continuity of σ at joint points is well satisfied. Let us now state the conditions needed to ensure the C^1 continuity along the curve. To address this concern, we propose to shift the problem to the tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$, for $i = 1, \dots, N-1$ and then bring back the solution to the space of probability measures $\mathcal{P}_+(I)$. Our algorithm only requires the Riemannian exponential and logarithm maps. Accurately, given $\mu^{(0)}, \dots, \mu^{(N)}$ an indexed set of probability measures on $\mathcal{P}_+(I)$. By making use of Riemannian logarithmic map given by equation (67), we lift data points $\mu^{(0)}, \dots, \mu^{(N)}$ in each tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$, $i = 1, \dots, N-1$. The mapped data are then defined by $\Phi^i = ((\Phi^{(0)})^i, \dots, (\Phi^{(N)})^i)$ with $(\Phi^{(k)})^i = \text{Log}_{\mu^{(i)}}(\psi^{(k)})$ for $k = 0, \dots, N$. Let $\beta : [0, N] \rightarrow T_{\mu^{(i)}}\mathcal{P}_+(I)$ denote the Bézier spline on $T_{\mu^{(i)}}\mathcal{P}_+(I)$, $i = 1, \dots, N-1$ defined by N Bézier curves β_j^i , $j \in \{2, 3\}, 0 \leq i \leq N-1$. Hence, the optimization problem (1) can be formulated as follows on $T_{\mu^{(i)}}\mathcal{P}_+(I)$:

$$\min_{((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-} E(((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-) \quad (91)$$

$$:= \min_{((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-} \int_0^1 \|\ddot{\beta}_2^0(t; (\Phi^{(0)})^i, ((b^{(1)})^i)^-, (\Phi^{(1)})^i)\|_2^2 \quad (92)$$

$$+ \sum_{i=1}^{N-2} \int_0^1 \|\ddot{\beta}_3^i(t; (\Phi^{(i)})^i, ((b^{(i)})^i)^+, ((b^{(i+1)})^i)^-, (\Phi^{(i+1)})^i)\|_2^2 \quad (93)$$

$$+ \int_0^1 \|\ddot{\beta}_2^{N-1}(t; (\Phi^{(N-1)})^i, ((b^{(N-1)})^i)^+, (\Phi^{(N)})^i)\|_2^2 \quad (94)$$

where $((b^{(1)})^i)^-$ and $((b^{(1)})^i)^+$ denote control points on the left and on the right-hand side of the interpolation point $(\Phi^{(i)})^i$ for $i = 1, \dots, (N-1)$ on $T_{\mu^{(i)}}\mathcal{P}_+(I)$. Since the tangent space is a vector space, the problem (94) is treated similarly as the Euclidean case \mathbb{R}^m . Actually, in this case, we prove that solutions to the problem of minimization of the mean square acceleration of the Bézier curve β are exactly the control points of the curve. Besides, we give conditions under which the Bézier curve β is of class C^2 . The details to obtain equations that govern the control points of the C^2 Bézier spline on \mathbb{R}^m are given in Appendix A. Finally, the Riemannian exponential map $\text{Exp}_{\mu^{(i)}}$ defined on $\mathcal{P}_+(I)$ by equation (66) will move control points of the C^2 Bézier curve β from the tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$ to $\mathcal{P}_+(I)$. The resulting control points in $\mathcal{P}_+(I)$ are optimal, and we assert that the curve is of class C^1 .

Theorem 3.1. *Given $\mu^{(0)}, \dots, \mu^{(N)}$ a set of $(N+1)$ data points on $\mathcal{P}_+(I)$, and $B^i = [((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-]^T$ a matrix of size $(m(N-1) \times m)$ containing the $(N-1)$ control points that generate the C^2 Bézier curve β_i in each tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$, for $i = 1, \dots, N-1$. Then, the Bézier spline $\sigma : [0, N] \rightarrow \mathcal{P}_+(I)$ interpolating the data points $\mu^{(i)}$ on $\mathcal{P}_+(I)$ is of class C^1 and is uniquely defined by the set of control points $\hat{\eta} = [(\hat{\eta}^{(1)})^-, \dots, (\hat{\eta}^{(N-1)})^-]^T \in \mathbb{R}^{m(N-1) \times m}$ given by:*

$$(\hat{\eta}^{(i)})^- = \text{Exp}_{\mu^{(i)}}(\tilde{x}^{(i)}), \quad i = 1, \dots, N-1, \quad (95)$$

where $\tilde{x}^{(i)}$, represent the row i of B^i in $T_{\mu^{(i)}}\mathcal{P}_+(I)$, for $i = 1, \dots, N-1$.

Proof. Similar to the Euclidean space \mathbb{R}^m , the differentiability condition at the interpolation points allows us to express control points $((b^{(1)})^i)^-$ in terms of $((b^{(1)})^i)^+$ as

$$((b^{(1)})^i)^- = (\Phi^{(i)})^i + \lambda_i(((b^{(1)})^i)^+ - (\Phi^{(i)})^i). \quad (96)$$

Considering that $\log_p(b) = b - p$ in the Euclidean case, hence the generalization of equation (96) on $\mathcal{P}_+(I)$ is given by

$$(\hat{\eta}^{(i)})^+ = \text{Exp}_{\mu^{(i)}}\left(\lambda_i \text{Exp}_{\mu^{(i)}}^{-1}((\hat{\eta}^{(i)})^-)\right) \quad (97)$$

which assert the C^1 differentiability condition on $\mathcal{P}_+(I)$. \square

Remark. The interpolation point $(\Phi^{(N)})^i$ is modified under the C^2 differentiability condition of the curve β_i on $T_{\mu^{(i)}}\mathcal{P}_+(I)$, for $i = 1, \dots, N-1$, therefore the point $\mu^{(N)}$ is changed and the new $(N+1)$ interpolation points on $\mathcal{P}_+(I)$ are given by:

$$\tilde{\mu}^{(k)} = \text{Exp}_{\mu^{(i)}}((\tilde{\Phi}^{(k)})^i), \quad k = 0, \dots, N; \quad i = 1, \dots, N-1, \quad (98)$$

where $\tilde{\Phi}^i = [(\tilde{\Phi}^{(0)})^i, \dots, (\tilde{\Phi}^{(N)})^i]^T$ a matrix of size $m(N+1) \times m$ containing the new $(N+1)$ interpolation points in each tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$. We give more details in Appendix A.

Algorithm 1 synthesizes all steps needed to construct the C^1 solution on $\mathcal{P}_+(I)$.

3.3. C^2 spline on $\mathcal{P}_+(I)$

The task is now to show that Bézier curves σ_j^i , $j \in \{2, 3\}$, $i = 0, \dots, N-1$ can be organized such that the C^2 differentiability at joint points is satisfied and thereby allow σ to be a C^2 spline. The basic geometric ingredient used to achieve this goal is the local symmetries at interpolation points offered on the Riemannian manifold $\mathcal{P}_+(I)$. In fact, since the space of probability measures is a locally symmetric space, hence at every point $\mu \in \mathcal{P}_+(I)$, there exists a local isometry φ of $\mathcal{P}_+(I)$ defined on a neighborhood of U of μ such that $\varphi(\mu) = \mu$ and the differential $d\varphi_\mu = -Id$. Lemma B.1 and Theorem B.1 in Appendix B examine in detail the relation made between local symmetries at interpolation points and the C^2 differentiability of the Bézier spline on symmetric spaces. Once more, we will make use of this relevant material to compute control points of the C^2

Algorithm 1. C^1 solution on $\mathcal{P}_+(I)$.

Input: $n \geq 3$, $\mu = [\mu^{(0)}, \dots, \mu^{(N)}]^T$ a matrix of size $m(N+1) \times m$ containing the $(N+1)$ interpolation points on $\mathcal{P}_+(I)$.

Output: $\hat{\eta}$ and $\tilde{\mu}$.

- 1: **for** $i = 1 : N - 1$ **do**
 - 2: Compute $\Phi^i = [(\Phi^{(0)})^i, \dots, (\Phi^{(N)})^i]^T$ a matrix of size $m(N+1) \times m$ containing the $(N+1)$ interpolation points on $T_{\mu^{(i)}}\mathcal{P}_+(I)$:
 - 3: **for** $k = 0 : N$ **do**
 - 4: $(\Phi^{(k)})^i = \text{Log}_{\mu^{(i)}}(\mu^{(k)})$: $(\Phi^{(k)})^i$ are determined by (67).
 - 5: Compute $B^i = [((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-]^T$ a matrix of size $m(N-1) \times m$ containing the $(N-1)$ control points of the C^2 Bézier curve β_i on $T_{\mu^{(i)}}\mathcal{P}_+(I)$, and $\tilde{\Phi}^i = [(\tilde{\Phi}^{(0)})^i, \dots, (\tilde{\Phi}^{(N)})^i]^T$ a matrix of size $m(N+1) \times m$ containing the new interpolation points on $T_{\mu^{(i)}}\mathcal{P}_+(I)$ using Appendix A.
 - 6: Compute control point $(\hat{\eta}^{(i)})^-$ of C^1 Bézier curve on $\mathcal{P}_+(I)$:
 - 7: $(\hat{\eta}^{(i)})^- = \text{Exp}_{\mu^{(i)}}(((b^{(i)})^i)^-)$.
 - 8: Compute the new interpolation points on $\mathcal{P}_+(I)$:
 - 9: $\tilde{\mu}^{(k)} = \text{Exp}_{\mu^{(i)}}((\tilde{\Phi}^{(k)})^i)$.
 - 10: **end for**
 - 11: **end for**
 - 12: **return** $\hat{\eta}$ and $\tilde{\mu}$.
-

Bézier spline σ on $\mathcal{P}_+(I)$. Let us consider $(\eta^{(i)})^-$ and $(\eta^{(i)})^+$ the new control points on the left and on the right side of the interpolation point $(\tilde{\mu}^{(i)})$. Similar to the Euclidean case \mathbb{R}^m , we might know $(\eta^{(1)})^-$ (and therefore $(\eta^{(1)})^+$ by the C^1 differentiability condition settled on the first step) and wish to define iteratively $(\eta^{(i)})^-$ for $i = 2, \dots, N-1$ (and obviously $(\eta^{(i)})^+$ in much the same way as $(\eta^{(1)})^+$).

Theorem 3.2. Consider $\tilde{\mu}^{(0)}, \dots, \tilde{\mu}^{(N)}$ the new interpolation points on $\mathcal{P}_+(I)$ given by equation (98) at times $t_i = i$ and γ the geodesic joining two points on $\mathcal{P}_+(I)$ given by equation (43). For a given $B^i = [((b^{(1)})^i)^-, \dots, ((b^{(N-1)})^i)^-]^T$, $i = 1, \dots, N-1$, a matrix of size $(m(N-1) \times m)$ containing the $(N-1)$ control points that generate the C^2 Bézier curve β_i in each tangent space $T_{\mu^{(i)}}\mathcal{P}_+(I)$, for $i = 1, \dots, N-1$, the Bézier spline $\sigma : [0, N] \rightarrow \mathcal{P}_+(I)$ is of class C^2 and it is uniquely defined by a set of control points given by the row of the matrix $\eta = [(\eta^{(1)})^-, \dots, (\eta^{(N-1)})^-]^T \in \mathbb{R}^{m(N-1) \times m}$ by

- (1) $(\eta^{(1)})^- = \text{Exp}_{\tilde{\mu}^{(1)}}(((b^{(1)})^1)^-)$,
- (2) $(\eta^{(2)})^- = \text{Exp}_{(\eta^{(1)})^+} \left(\frac{1}{3} \left((d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)})^-} (\dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)})^-)) - 4\dot{\gamma}(0, (\eta^{(1)})^-, \tilde{\mu}^{(1)}) \right) \right)$,
- (3) $(\eta^{(i+1)})^- = \text{Exp}_{(\eta^{(i)})^+} \left(\left((d\varphi_{\tilde{\mu}^{(i)}})_{(\eta^{(i)})^-} (\dot{\gamma}(1, (\eta^{(i-1)})^+, (\eta^{(i)})^-)) - 2\dot{\gamma}(0, (\eta^{(i)})^-, \tilde{\mu}^{(i)}) \right) \right)$, $i = 2, \dots, N-2$.

Proof. The proof strongly depends on the results given in [23]. Hence, all technical details of the proof will be given in Appendix B. For convenience, we remind the main ideas:

- (1) In [23] they compute the covariant derivative of a tangent vector along a curve and provide the explicit C^2 condition on symmetric spaces in terms of the derivative of exponential and symmetry functions. We make use of this result and Theorem 3.1 to simplify the derivative of the inverse of the exponential map.
- (2) We express the derivative of the symmetry as a function of the tangent vector along the Bézier spline σ_i^j at $t = 0$ and $t = 1$ which simplify the C^2 condition in [23] and help us to obtain an explicit expression for control points that generate the Bézier spline σ .

The resulting spline is then reconstructed with the De Casteljau algorithm. □

Algorithm 2 synthesizes all steps needed to construct a C^2 solution on $\mathcal{P}_+(I)$.

Algorithm 2. C^2 solution on $\mathcal{P}_+(I)$.

Input: $n \geq 3$, $\tilde{\mu} = [\tilde{\mu}^{(0)}, \dots, \tilde{\mu}^{(N)}]^T$ a matrix of size $m(N+1) \times m$ containing the $(N+1)$ interpolation points on $\mathcal{P}_+(I)$.

- 1:
 - Output:** Q .
 - 2:
 - 3: Calculate $\hat{\eta} = [(\hat{\eta}^{(1)})^-, \dots, (\hat{\eta}^{(N-1)})^-]^T$ control points of C^1 Bézier spline on $\mathcal{P}_+(I)$ using Algorithm 1.
 - 4: Set $(\eta^{(1)})^- = (\hat{\eta}^{(1)})^-$.
 - 5: Calculate control point $(\eta^{(1)})^+$:
 - 6: $(\eta^{(1)})^+ = \text{Exp}_{\tilde{\mu}^{(1)}}(-\frac{2}{3}\text{Exp}_{\tilde{\mu}^{(1)}}^{-1}((\eta^{(1)})^-))$
 - 7: Calculate control point $(\eta^{(2)})^-$:
 - 8: $(\eta^{(2)})^- = \text{Exp}_{(\eta^{(1)})^+}(\frac{1}{3}((d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)})^-}(\dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)})^-)) - 4\dot{\gamma}(0, (\eta^{(1)})^-, \tilde{\mu}^{(1)})))$
 - 9: **for** $i = 2 : N - 2$ **do**
 - 10: $(\eta^{(i)})^+ = \text{Exp}_{\tilde{\mu}^{(i)}}(-\text{Exp}_{\tilde{\mu}^{(i)}}^{-1}(\mu^{(i)})^-)$
 - 11: $(\eta^{(i+1)})^- = \text{Exp}_{(\mu^{(i)})^+}(((d\varphi_{\tilde{\mu}^{(i)}})_{(\eta^{(i)})^-}(\dot{\gamma}(1, (\eta^{(i-1)})^+, (\eta^{(i)})^-)) - 2\dot{\gamma}(0, (\eta^{(i)})^-, \tilde{\mu}^{(i)})))$
 - 12: **end for**
 - 13: Calculate control point $(\eta^{(N-1)})^+$:
 - 14: $(\eta^{(N-1)})^+ = \text{Exp}_{\tilde{\mu}^{(N-1)}}(-\frac{2}{3}\text{Exp}_{\tilde{\mu}^{(N-1)}}^{-1}((\eta^{(N-1)})^-))$
 - 15: **return** Q .
-

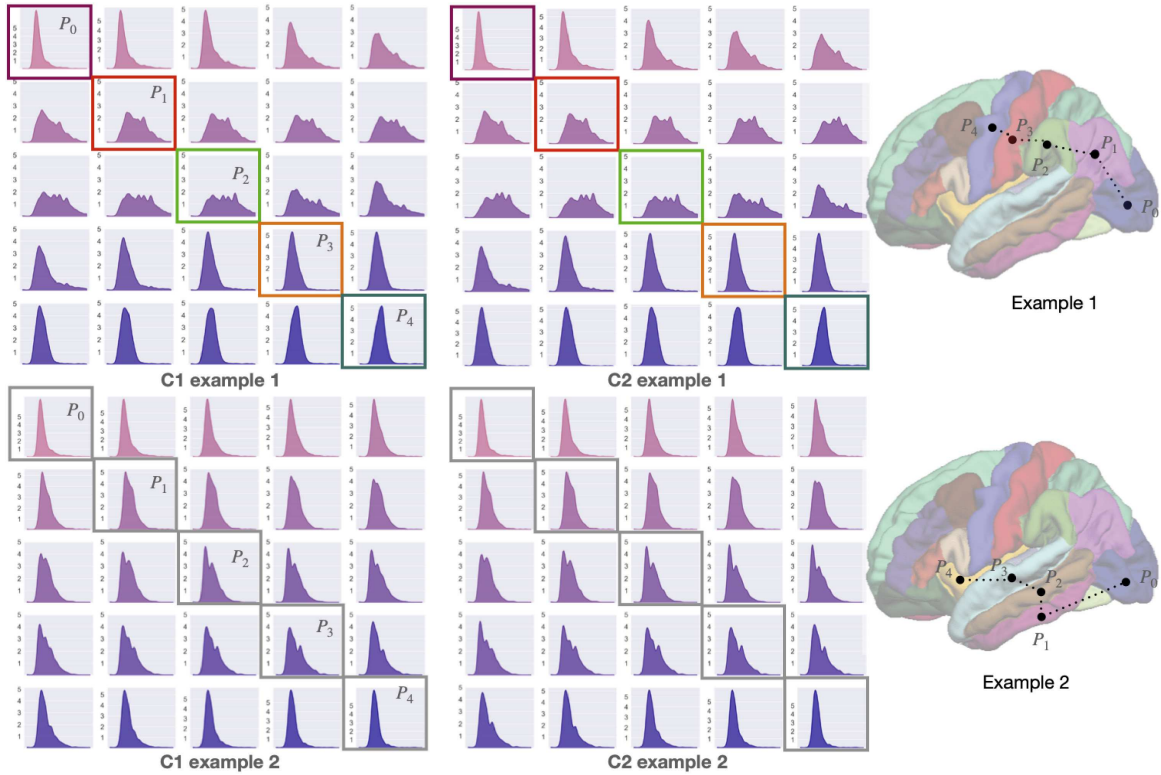


FIGURE 1. Two illustrating examples of splines interpolating different ROIs on the brain.

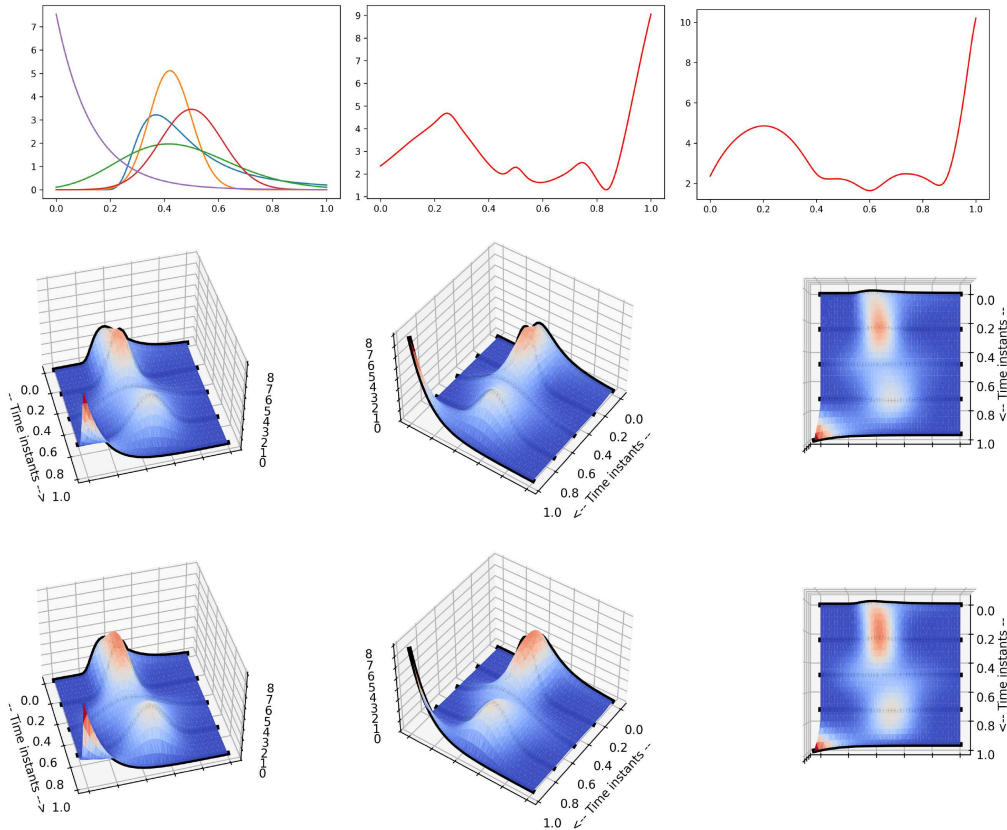


FIGURE 2. Example 1: the first row, from left to right, displays the control PDFs, the velocity along the C^1 path, and the velocity along the C^2 path. The second row displays the C^1 path and last row displays the C^2 path as surfaces.

4. APPLICATIONS

In many applications, it is of great interest to study the changes in densities as a function of time or space positions. In the following examples, we will show how the proposed methods can be used to capture dynamics as a smooth path interpolating the key observations. For example, it is essential to study the changes in densities of Regions Of Interest (ROIs) in the brain. Broad aims in the study of brain dynamics are to investigate how densities as features fold into a 3D functional path and to estimate the full range of such functional path for given functional ROIs. We illustrate this idea in Figure 1 where we display two different examples. For each example, ROI $P_{i \leq 5}$ has a different color and is represented by a PDF. We use our methods to construct C^1 and C^2 splines that interpolate $P_{i \leq 5}$ and we display the results using 25 equally spaced frames.

4.1. Numerical examples

Before we show results on Brain, we illustrate C^1 and C^2 paths on PDFs manifold with several examples where densities vary from simple to complicated shapes. For this application a temporal subsequence of 5 densities equally spaced in time is taken and we consider our method for predicting densities in between the observed ones. We display the C^1 and C^2 paths in Figures 2 and 3.

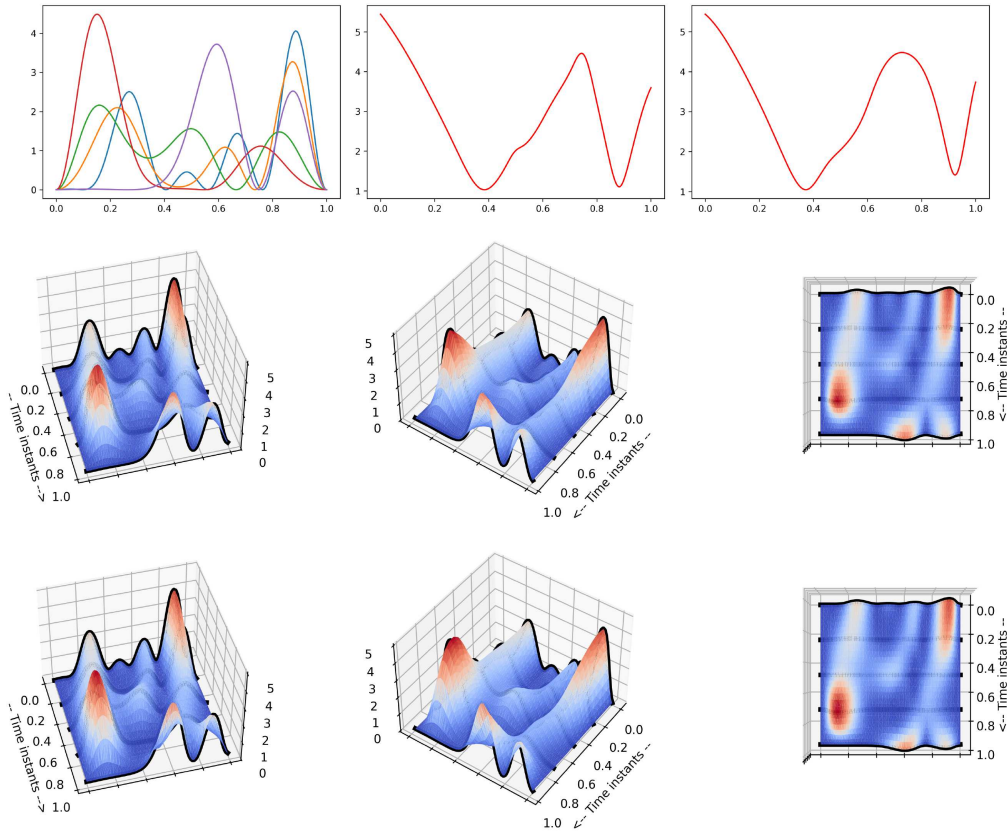


FIGURE 3. Example 2: the first row, from left to right, displays the control PDFs, the velocity along the C^1 path, and the velocity along the C^2 path. The second row displays the C^1 path and last row displays the C^2 path as surfaces.

To better visualize the differences between C^1 and C^2 solutions we first display each path as a surface where the domain of definition and time interval plays the role of x and y axes while z gives the value of the PDF. We show the surfaces from two different points of view for a better illustration. The color bar displays the amplitude of the surface. We then compute and display the norm of the derivative with respect to time.

4.2. Medical examples

An important application of the proposed method is that smooth solutions between different key observations can be explored, visualized and analyzed. This is important information about the evolution (variation) during a trajectory. To show how this could be applied in a medical context we consider a dataset of averages of morphological features sampled from brain cortical surfaces for a population. The morphological measures from surfaces are derived from atlas-based registrations of individual cortices from subjects. Following registration, the cortex is parcellated into distinct regions of interest, and measures such as cortical thickness, gray matter volume, or sulcal depths over the ROIs. In this work we focus on thickness only and show trajectories crossing of the most relevant and neighboring ROIs.

An attractive feature of the spline fit is that smooth paths between different states can be explored, to investigate possible transitions in shape from one step to another. For the smooth prediction we have used the observations at integer times but have predicted at equally spaced time points between observations. In Figure 4

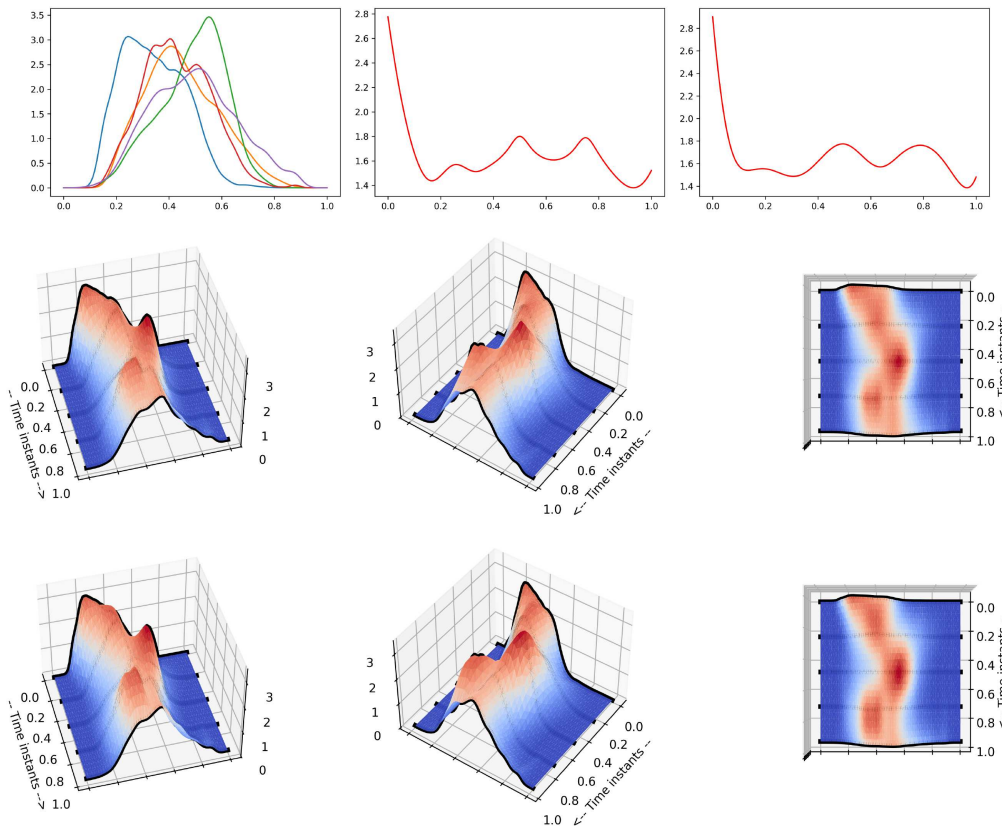


FIGURE 4. Spline paths on Brain’s ROIs: the first row from left to right displays the control PDFs, the velocity along the C^1 path, and the velocity along the C^2 path. The second row displays the C^1 path and last row the C^2 path as surfaces.

we display the predicted shape change in the transition to a state in a later part of the simulation using the cubic spline, at times $t = 1$ to $t = 25$ at equally spaced intervals. We can see that the smooth path predicts the density change between data points well, and that the evolution is seen in the smoothed predicted path in \mathcal{P} .

5. CONCLUSIONS

The paper introduces an effective algorithm for generating a C^2 Bézier spline on the space of probability measures $\mathcal{P}_+(I)$ while interpolating the given ordered set of probability measures. In this work, we consider $\mathcal{P}_+(I)$ as a Riemannian manifold equipped with the Fisher–Rao metric. With the help of the geometric interpretation, we derive explicit theoretical expression of important geometric structures on $\mathcal{P}_+(I)$, namely Levi–Civita connection, minimal geodesic, parallel transport, exponential, and logarithm maps. Using these newly devised geodesic operations, the Bézier splines are generalized on $\mathcal{P}_+(I)$ and applied for medical data. Following the step-by-step construction, The proposed method is simple to implement and easy to compute.

REFERENCES

[1] N. Ay, J. Jost, H. Le and L. Schwachhofer, Information Geometry. Springer, Cham (2017).
 [2] J.D. Benamou, T.O. Gallouët and F.X. Vialard, Second-order models for optimal transport and cubic splines on the wasserstein space. *Found. Comput. Math.* **19** (2019) 1113–1143.

- [3] R. Chakraborty and B.C. Vemuri, Statistics on the compact Stiefel manifold: theory and applications. *Ann. Stat.* **47** (2017) 415–438.
- [4] Y. Chen, G. Conforti and T.T. Georgiou, Measure-valued spline curves: an optimal transport viewpoint. *SIAM J. Math. Anal.* **50** (2018) 5947–5968.
- [5] A. Cherian and J. Wang, Generalized one-class learning using pairs of complementary classifiers. *IEEE TPAMI* **43** (2021) 420–433.
- [6] S. Chewi, J. Clancy, T. Le Gouic, P. Rigollet, G. Stepaniants and A. Stromme, Fast and smooth interpolation on Wasserstein space, in Proceedings of Machine Learning Research. Vol. 130. PMLR (2021).
- [7] I.L. Dryden and K.V. Mardia, Statistical Shape Analysis, with Applications in R. John Wiley and Sons, Chichester (2016).
- [8] N. Dyn, Linear and nonlinear subdivision schemes in geometric modeling, in Foundations of Computational Mathematics. *London Mathematical Society Lecture Note Series*. Vol. 363. Hong Kong (2008) 68–92.
- [9] O. Freifel, S. Hauberg and M.J. Black, Model transport: towards scalable transfer learning on manifolds, in IEEE Conference on Computer Vision and Pattern Recognition. IEEE (2014).
- [10] B. Geir, M. Klas and V. Olivier, Numerical algorithm for c^2 -splines on symmetric spaces. *SIAM J. Numer. Anal.* **56** (2018) 2623–2647.
- [11] J. Hinkle, P.T. Fletcher and S. Joshi, Intrinsic polynomials for regression on Riemannian manifolds. *J. Math. Imaging Vis.* **50** (2014) 32–52.
- [12] M. Itoh and H. Satoh, Geometry of fisher information metric and the barycenter map. *Entropy* **17** (2015) 1814–1849.
- [13] B. Jérémie, G. Raúl, K. Thierry and L. Alfredo, Geodesic PCA in the wasserstein space by convex PCA. *Ann. l’Inst. Henri Poincaré* **53** (2017) 1–26.
- [14] W. Jonathan and B. Francis, Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** (2019) 2620–2648.
- [15] B.V. Julio, F. Joaquin, R. Gonzalo and T. Felipe, Bayesian learning with Wasserstein barycenters. Preprint [arXiv:1805.10833](https://arxiv.org/abs/1805.10833) (2018).
- [16] A. Karimi, L. Ripani and T.T. Georgiou, Statistical learning in wasserstein space. *IEEE Control Syst. Lett.* **5** (2021) 899–904.
- [17] K.R. Kim, I.L. Dryden and H. Le, Smoothing splines on Riemannian manifolds, with applications to 3d shape space. *R. Stat. Soc.* **83** (2020) 108–132.
- [18] L. Lin, B. St.Thomas, H. Zhu and D.B. Dunson, Extrinsic local regression on manifold-valued data. *J. Am. Stat. Assoc.* **112** (2017) 1261–1273.
- [19] G. Mustafa and R. Hameed, Families of non-linear subdivision schemes for scattered data fitting and their non-tensor product extensions. *Appl. Math. Comput.* **359** (2019) 214–240.
- [20] E. Nava-Yazdani and K. Polthier, De Casteljau’s algorithm on manifolds. *Comput. Aided Geom. Des.* **30** (2013) 722–732.
- [21] A. Petersen and H.-G. Müller, Fréchet regression for random objects with Euclidean predictors. *Ann. Stat.* **47** (2019) 691–719.
- [22] T. Popiel and L. Noakes, c^2 spherical Bézier splines. *Comput. Aided Geom. Des.* **23** (2006) 261–275.
- [23] T. Popiel and L. Noakes, Bézier curves and c^2 interpolation in Riemannian manifolds. *J. Approx. Theory* **148** (2007) 111–127.
- [24] T. Shingel, Interpolation in special orthogonal groups. *IMA J. Numer. Anal.* **29** (2009) 731–745.
- [25] F. Silva Leite and L. Machado, Fitting smooth paths on Riemannian manifolds. *Int. J. Appl. Math. Stat.* **6** (2006) 25–53.
- [26] C. Sinho, M. Tyler, R. Philippe and J.S. Austin, Gradient descent algorithms for Bures-Wasserstein barycenters, in Proceedings of Thirty Third Conference on Learning Theory, edited by J. Abernethy and S. Agarwal. Vol. 125. PMLR (2020) 1276–1304.
- [27] J. Wallner, E. Nava Yazdani and P. Grohs, Smoothness properties of lie group subdivision schemes. *Multiscale Model. Simul.* **6** (2007) 493–505.



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.

APPENDIX A. C^2 BÉZIER SPLINE ON \mathbb{R}^m

Let us consider the Euclidean case \mathbb{R}^m . Given a list of $(N + 1)$ interpolation points p_0, \dots, p_N and the control points \widehat{b}_i^+ and \widehat{b}_i^- in the right and in the left of p_i , $i = 0, \dots, N$. The C^1 differentiability condition at knots p_i allows us to express control points \widehat{b}_i^+ in terms of \widehat{b}_i^- as:

$$\widehat{b}_1^+ = \frac{5}{3}p_1 - \frac{2}{3}\widehat{b}_1^-, \quad (\text{A.1})$$

$$\widehat{b}_i^+ = 2p_i - \widehat{b}_i^-, \quad i = 2, \dots, N - 2 \quad (\text{A.2})$$

$$\widehat{b}_{N-1}^+ = \frac{5}{2}p_{N-1} - \frac{3}{2}\widehat{b}_{N-1}^-. \quad (\text{A.3})$$

Hence, the task now is reduced to search only control points \widehat{b}_i^- , for $i = 1, \dots, N - 1$, that generate the C^1 Bézier spline β in \mathbb{R}^m . Replacing the new optimization variables in problem (1) gives (94), which is merely the problem of minimization of the mean square acceleration of the Bézier curve β in the Euclidean space \mathbb{R}^m . The optimal solution $Y = [\widehat{b}_1^-, \dots, \widehat{b}_{N-1}^-]^T \in \mathbb{R}^{(N-1) \times m}$ of that problem is the unique solution of a tridiagonal linear system

$$Y = A^{-1}CP = DP \quad \text{with} \quad \sum_{j=0}^{j=N+1} d_{ij} = 1 \quad (\text{A.4})$$

where A is a tridiagonal sparse square matrix of size $(N - 1) \times (N - 1)$ with a dominant diagonal, C a matrix of size $(N - 1) \times (N + 1)$ and P the matrix of p_i 's of size $(N + 1) \times m$ given by:

$$\left| \begin{array}{ll} A_{(1,1:2)} = [16 \ 6] & C_{(1,1:2)} = [16 \ 6] \\ A_{(2,1:3)} = [6 \ 36 \ 9] & C_{(2,2:3)} = [6 \ 36 \ 9] \\ A_{(i,i-1:i+1)} = [9 \ 36 \ 9], & C_{(i,i:i+1)} = [9 \ 36 \ 9], \quad i = 3, \dots, n - 2 \\ A_{(N-1,N-2:N-1)} = [9 \ 36] & C_{(N-1,N-1:N+1)} = [9 \ 36]. \end{array} \right. \quad (\text{A.5})$$

We may now write the C^2 differentiability condition. It is obvious that with this C^2 condition the position of the control points \widehat{b}_i^- and \widehat{b}_i^+ that generate the curve β will be modified. Therefore, it is more convenient to use another notation. Let us denote by b_i^- and b_i^+ the new control points on the left and on the right hand side of the interpolation point p_i , for $i = 1, \dots, N - 1$. Computing the acceleration of β on respective intervals and taking into account that β is C^1 , we shall replace \widehat{b}_1^+ by (A.1), \widehat{b}_i^+ by (A.2), and \widehat{b}_{N-1}^+ by (A.3). We deduce that:

$$b_2^- = \frac{1}{3}p_0 - \frac{1}{2}b_1^- + \frac{8}{3}p_1, \quad (\text{A.9})$$

$$b_{i+1}^- = b_{i-1}^+ + 4p_i - 4b_i^-, \quad i = 2, \dots, N - 2 \quad (\text{A.10})$$

$$p_N = 2p_{N-1} + 2b_{N-1}^+ - 6b_{N-1}^- + 3b_{N-2}^+. \quad (\text{A.11})$$

We see at once that points that will be modified by the additional C^2 condition are \widehat{b}_i^- and hence \widehat{b}_i^+ , for $i = 2, \dots, N - 1$. The point \widehat{b}_1^- remains invariant and consequently it will be the case for \widehat{b}_1^+ . According to the C^1 differentiability condition

ensured at the first step, one can take $b_1^- = \widehat{b}_1^-$, with \widehat{b}_1^- is the first row of the matrix Y obtained as a solution of the optimization problem (94). However, the endpoint p_N is affected as we can deduce from equation (A.11). Nevertheless, it follows that giving the control point b_1^- allows us to find all the other control points including b_2^- with equation (A.9) and hence b_2^+ with (A.2), then b_{i+1}^- for $i = 2, \dots, N-2$ with (A.10) and therefore b_i^+ , for $i = 3, \dots, N-2$ with (A.2) and b_{N-1}^+ with (A.3).

APPENDIX B. PROOF OF THEOREM 3.2

We now prove Theorem 3.2. The proof is based on the following two results given in [23].

Lemma B.1. *Let $\mu_1 \in \mathcal{P}_+(I)$.*

- (1) $(d\varphi_{\mu^{(1)}})_{\mu^{(2)}}^{-1} = (d\varphi_{\mu^{(1)}})_{\varphi_{\mu^{(1)}}(\mu^{(2)})}$, for all $\mu^{(2)} \in \mathcal{P}_+(I)$.
- (2) $(d\varphi_{\mu^{(1)}})_{Exp_{\mu^{(1)}}(H)} \circ (dExp_{\mu^{(1)}})_{-H} = -(dExp_{\mu^{(1)}})_{-H}$ for all $H \in T_{\mu^{(1)}}\mathcal{P}_+(I)$.

Theorem B.1. *Let $t \rightarrow \sigma_j(t, V^{(0)}, \dots, V^{(j)})$ be the Bézier curve of order j on $\mathcal{P}_+(I)$ with a number of control points $V^{(j)}$ for $i = 0, \dots, j$. Then, $\sigma_j(t; V^{(0)}, \dots, V^{(j)})$ satisfies:*

- (1) $\frac{D}{dt} \Big|_{t=0} \dot{\sigma}_j(t; V^{(0)}, \dots, V^{(j)}) = j(j-1)\Omega_0$, where

$$\Omega_0 := \begin{cases} \dot{\gamma}(0, V^{(1)}, V^{(2)}), & \text{if } V^{(0)} = V^{(1)} \\ (dExp_{V^{(0)}})_{\dot{\gamma}(0, V^{(0)}, V^{(1)})}^{-1} \left(\dot{\gamma}(0, V^{(1)}, V^{(2)}) - \dot{\gamma}(1, V^{(0)}, V^{(1)}) \right), & \text{if } V^{(0)} \neq V^{(1)}. \end{cases}$$

- (2) $\frac{D}{dt} \Big|_{t=1} \dot{\sigma}_j(t; V^{(0)}, \dots, V^{(j)}) = j(j-1)\Omega_j$, where

$$\Omega_j := \begin{cases} -\dot{\gamma}(0, V^{(j-2)}, V^{(j-1)}), & \text{if } V^{(j-1)} = V^{(j)} \\ (dExp_{V^{(j)}})_{-\dot{\gamma}(0, V^{(j-1)}, V^{(j)})}^{-1} \left(\dot{\gamma}(0, V^{(j-1)}, V^{(j)}) - \dot{\gamma}(1, V^{(j-2)}, V^{(j-1)}) \right), & \text{if } V^{(j-1)} \neq V^{(j)}. \end{cases}$$

We will exploit a modified form of the Theorem B.1 to obtain the proof of the Theorem 3.2.

Proof of Theorem 3.2. Part (i) follows from Theorem 3.1. We now prove (ii). The Bézier spline σ is C^2 on $\mathcal{P}_+(I)$ if and only if it satisfies the C^2 differentiability condition at joint points $\tilde{\mu}^{(i)}$, for $i = 1, \dots, N-1$. At the point $\tilde{\mu}^{(1)}$, this means:

$$\frac{D}{dt} \Big|_{t=1} \dot{\sigma}_2(t; \tilde{\mu}^{(0)}, (\eta^{(1)})^-, \tilde{\mu}^{(1)}) = \frac{D}{dt} \Big|_{t=0} \dot{\sigma}_3(t; \tilde{\mu}^{(1)}, (\eta^{(1)})^+, (\eta^{(2)})^-, \tilde{\mu}^{(2)}). \quad (\text{B.1})$$

Applying Theorem B.1 yields: σ is C^2 on $\tilde{\mu}^{(1)}$ if and only if $\Omega_2 - 3\Omega_0 = 0$ with:

$$\begin{aligned} \Omega_2 - 3\Omega_0 &= (dExp_{\tilde{\mu}^{(1)}})_{-\dot{\gamma}(1, (\eta^{(1)})^-, \tilde{\mu}^{(1)})}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)})^-, \tilde{\mu}^{(1)}) - \dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)})^-) \right) \\ &\quad - 3(dExp_{\tilde{\mu}^{(1)}})_{\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)})^+, (\eta^{(2)})^-) - \dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)})^+) \right). \end{aligned} \quad (\text{B.2})$$

Since β_1 is a C^1 Bézier curve on $T_{\mu^{(1)}}\mathcal{P}_+(I)$, we get that $\dot{\gamma}(1, (\eta^{(1)})^-, \tilde{\mu}^{(1)}) = \dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+)$. By Lemma B.1, we have

$$\begin{aligned} &(dExp_{\tilde{\mu}^{(1)}})_{\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)})^+, (\eta^{(2)})^-) - \dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)})^+) \right) \\ &= -(dExp_{\tilde{\mu}^{(1)}})_{-\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+}^{-1} \left((d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)})^+} \left(\dot{\gamma}(0, (\eta^{(1)})^+, (\eta^{(2)})^-) - \dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)})^+) \right) \right). \end{aligned}$$

It follows that

$$\begin{aligned} \Omega_2 - 3\Omega_0 &= (dExp_{\tilde{\mu}^{(1)}})_{-\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)})^-, \tilde{\mu}^{(1)}) - \dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)})^-) \right) \\ &\quad + 3(dExp_{\tilde{\mu}^{(1)}})_{-\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)})^+}^{-1} \left((d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)})^+} \left(\dot{\gamma}(0, (\eta^{(1)})^+, (\eta^{(2)})^-) - \dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)})^+) \right) \right) \end{aligned}$$

$$\begin{aligned}
&= (d\text{Exp}_{\tilde{\mu}^{(1)}})_{-\dot{\gamma}(0, \tilde{\mu}^{(1)}, (\eta^{(1)+})}^{-1} \left[3(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) \right. \\
&\quad \left. - 3(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)+}) \right) + \dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}) - \dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)-}) \right) \right].
\end{aligned}$$

Hence, $\Omega_2 - 3\Omega_0 = 0$ if and only if

$$\begin{aligned}
&= 3(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) - 3(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)+}) \right) \\
&\quad + \dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}) - \dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)-}) \right) = 0.
\end{aligned} \tag{B.3}$$

Nevertheless $\varphi_{\tilde{\mu}^{(1)}}(\gamma(t, \tilde{\mu}^{(1)}, (\eta^{(1)+})) = \gamma(1-t, (\eta^{(1)-}), \tilde{\mu}^{(1)})$, $\forall t \in [0, 1]$. Differentiate this identity with respect to t , we obtain

$$(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(1, \tilde{\mu}^{(1)}, (\eta^{(1)+}) \right) = -\dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}).$$

Accordingly, equation (B.3) becomes

$$3(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) = \dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)-}) - 4\dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}). \tag{B.4}$$

Now, Lemma B.1 shows that

$$\begin{aligned}
(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)+})} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) &= (d\varphi_{\tilde{\mu}^{(1)}})_{\varphi_{\tilde{\mu}^{(1)}}((\eta^{(1)-})} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) \\
&= (d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)-})}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right).
\end{aligned}$$

It follows that $(d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)-})}^{-1} \left(\dot{\gamma}(0, (\eta^{(1)+}), (\eta^{(2)-}) \right) = \frac{1}{3} \left(\dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)-}) - 4\dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}) \right)$.

Consequently, with the exponential map at the point η_1^+ , we get

$$(\eta^{(2)-}) = \text{Exp}_{(\eta^{(1)+})} \left(\frac{1}{3} \left((d\varphi_{\tilde{\mu}^{(1)}})_{(\eta^{(1)-})} \left(\dot{\gamma}(1, \tilde{\mu}^{(0)}, (\eta^{(1)-}) \right) - 4\dot{\gamma}(0, (\eta^{(1)-}), \tilde{\mu}^{(1)}) \right) \right). \tag{B.5}$$

The proof of Part (iii) follows in much the same way as Part (ii). \square