International Journal of Mathematics and
Computer Science, **15**(2020), no. 4, 1253–1270

Wasserstein Riemannian geometry of Gamma densities

Carlos Ogouyandjou, Nestor Wadagni

Institut de Mathématiques et de Sciences Physiques
Université d'Abomey-Calavi
Porto-Novo, Bénin

email: ogouyandjou@imsp-uac.org, nestor.wadagni@imsp-uac.org

(Received June 2, 2020, Accepted September 9, 2020)

Abstract

A Wasserstein Riemannian Gamma manifold is a space of Gamma probability density functions endowed with the Riemannian Otto metric which is related to the Wasserstein distance. In this paper, we study some geometric properties of such Riemanian manifold. In particular we compute the coefficients of α-connections and the sectional curvature of those manifolds.

1 Introduction

The geometry of Gamma manifold related to the family of Gamma densites and endowed with the Fisher metric has been studied by Amari et al.[1]. Recently this geometry was used in medical imagin by Rebbah et al. [11] where the authors present how the information geometry and the generalized Gamma manifold improved the quality classification of deseases related to persons. Considering the Riemannian structure obtained by the Fisher information on a statistical manifold, Amari [1] defines a one-parameter family of affine connections called α-connections. Hence α-connections have become key tools in information geometry and have been widely investigated by several authors such as Gbaguidi et al. [7] who constructed a family of α-connections on a Hilbert bundle of generalized statistical manifold.

Key words and phrases: Information geometry, Gamma distribution, Wasserstein distance, Otto's metric.
AMS (MOS) Subject Classifications: 15B48, 53C23, 53C25, 60D05.
Corresponding author: Carlos Ogouyandjou
ISSN 1814-0432, 2020, http://ijmcs.future-in-tech.net

Recently Malagó et al. study the geometry of the family of Gaussian densities endowed with Otto metric [8] which is related to wasserstein distance [12, 3, 9]. In this paper, we study the geometry of the family of Gamma densities endowed with the Otto metric. Let $\mathcal{M}$ be a set of probability densities endowed with the Otto Riemannian metric. We construct on $\mathcal{M}$ a family $\nabla^{(\alpha)}$ of torsion-free α-connections that is exactly the Levi-Civita connection on $\mathcal{M}$ when $\alpha = 0$. We also find out that the exponential families and the mixture families are respectively (1)-flat and (-1)-flat. The rest of the paper is organized as follows: we recall some preliminaries on α-connections in section 2, and we present useful results on Otto metric and Wasserstein metric in sections 3 and 4. In section 5 we present the main results : the construction of α-connection using Otto metric. Finally in section 6, we compute the coefficients of α-connections, and the sectional curvature of Gamma manifold.

2 Preliminary on α-connections

For some integer $d \geq 1$, let $\mathcal{X}$ be a non-empty subset of $\mathbb{R}^d$ and $\mathcal{M} = \{p_\theta(\cdot), \theta \in \Theta\}$, with $\Theta \subset \mathbb{R}^n$ be a family of probability distributions on $\mathcal{X}$. Each element of $\mathcal{M}$, can be identified with $\theta = (\theta_1, \cdots, \theta_n) \in \mathit{\Theta}$ a subset of $\mathbb{R}^n$ and the mapping $\theta \mapsto p_\theta$ is injective. $\mathcal{M}$ is a C^∞ differentiable manifold.

Example 2.1. $\mathcal{X} = \mathbb{R}$, $n = 2$, $\theta = (\mu, \sigma)$, $\mathit{\Theta} = \{(\mu, \sigma) : \mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*\}$

$$p(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

Put $\ell(.;\theta) = \log p(.,\theta)$. $\frac{\partial \ell(.;\theta)}{\partial \theta^i}$ for $i = 1, \cdots, n$ are the scores functions.

Definition 2.1. *The Fisher information metric*

The Fisher information matrix of $\mathcal{M}$ at θ is the $n \times n$ matrix $G(\theta) = (\tilde{g}_{ij}(\theta))$ defined by :

$$\tilde{g}_{ij}(\theta) := \mathbb{E}_\theta[\partial_i \ell(X,\theta) \partial_j \ell(X,\theta)] = \int_{\mathcal{X}} \partial_i \ell(x,\theta) \partial_j \ell(x,\theta) p(x;\theta) dx$$

where $\partial_i := \frac{\partial}{\partial \theta_i}$ and $\ell(x,\theta) = \log p(x;\theta)$. In particular, when $n = 1$, we call this the Fisher information.

The inner product of the natural basis of the coordinate system $(\theta_1, \cdots, \theta_n)$

$$\langle \partial_i, \partial_j \rangle = \tilde{g}_{ij}$$

uniquely determines a Riemannian metric $\tilde{g} = \langle \cdot, \cdot \rangle$ such that for all $\theta \in \Theta$, and for all $X, Y \in \top_\theta \mathcal{M}$; $\tilde{g}_\theta(X, Y) = \langle X, Y \rangle_\theta = \mathbb{E}_\theta[(X\ell)(Y\ell)]$. $\tilde{g}$ is called Fisher metric or alternatively, the information metric.

Definition 2.2. *An affine connection ∇ on a differentiable manifold $\mathcal{M}$ is a mapping*

$$\nabla : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \to \mathcal{X}(\mathcal{M})$$

which is denoted by $(X, Y) \to \nabla_X Y$ and which satisfies the following properties:

- $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$
- $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$
- $\nabla_X(fY) = f\nabla_X Y + X(f)Y$ *in which $X, Y, Z \in \mathcal{X}(\mathcal{M})$ and $f, g \in C^\infty(\mathcal{M})$.*

Theorem 2.3. *[5] Given a Riemannian manifold $(\mathcal{M}, g)$, there exists an unique affine connection ∇ on $\mathcal{M}$ satisfing the conditions:*

- ∇ *is symmetric.*
- ∇ *is compatible with the Riemannian metric g.*

This affine connection is the Levi-Civita connection on the manifold $(\mathcal{M}, g)$.

In a coordinate system (U, θ), the function $\overset{\circ}{\Gamma}{}^k_{ij}$ defined on U by $\nabla_{\partial_i}\partial_j = \sum_k \Gamma^k_{ij}\partial_k$ are called the Christoffel symbols of the Levi-Civita connection and we have

$$\overset{\circ}{\Gamma}{}^k_{ij} = \frac{1}{2}\left(\frac{\partial g_{jm}}{\partial \theta^i} + \frac{\partial g_{mi}}{\partial \theta^j} - \frac{\partial g_{ij}}{\partial \theta^m}\right) g^{mk}. \tag{2.1}$$

Amari[2] considers the function $\Gamma^{(\alpha)}_{ij,k}$ which maps each point θ to the following value:

$$\left(\Gamma^{(\alpha)}_{ij,k}\right)_\theta := \mathbb{E}_\theta\left[\left(\partial_i\partial_j\ell(X,\theta) + \frac{1-\alpha}{2}\partial_i\ell(X,\theta)\partial_j\ell(X,\theta)\right)(\partial_k\ell(X,\theta))\right]$$

where α is some arbitrary real number. The α-connection $\nabla^{(\alpha)}$, which is an affine connection, is defined by

$$\langle \nabla^{(\alpha)}_{\partial_i} \partial_j, \partial_k \rangle = \Gamma^{(\alpha)}_{ij,k},$$

where $g = \langle \cdot, \cdot \rangle$ is the Fisher metric and $\nabla^{(\alpha)}_{\partial_i} \partial_j$ is the α covariant derivative of ∂_j in the direction of ∂_i.
Next, we recall some important results on the Otto metric which is a Riemannian metric on the Wasserstein space.

3 Wasserstein space

Let $\mathcal{X}$ be a subset of $\mathbb{R}^n$ and $\mathcal{B}(\mathcal{X})$ be the σ-algebra of Borel sets on $\mathcal{X}$. Let $P(\mathcal{X})$ be the set of probability measures defined on $\mathcal{B}(\mathcal{X})$.

Definition 3.1. *For any two measures $\mu, \nu \in P(\mathcal{X})$, we define the coupling of (μ, ν) by the set*

$$\Pi(\mu, \nu) = \left\{ \eta \in P(\mathcal{X} \times \mathcal{X}) : \eta \circ p_1^{-1} = \mu, \eta \circ p_2^{-1} = \nu \right\}$$

where $p_1, p_2 : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ are the projections yielding the first and second component, respectively.

Definition 3.2. *The Wasserstein space of order p, $p \in [1, \infty[$ is defined as*

$$P_p(\mathcal{X}) = \left\{ \mu \in P(\mathcal{X}); \int_{\mathcal{X}} \|x\|^p d\mu(x) < +\infty \right\}. \tag{3.2}$$

For any two $\mu, \nu \in P_p(\mathcal{X})$, we define

$$W_p(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}} \|x - y\|^p d\pi(x, y) \right)^{1/p}. \tag{3.3}$$

A sequence $(\mu_n)_{n \in \mathbb{N}} \subset P_p(\mathcal{X})$ is said to converge weakly to $\mu \in P_p(\mathcal{X})$ if

$$\lim_{n \to +\infty} \int_{\mathcal{X}} f d\mu_n = \int_{\mathcal{X}} f d\mu$$

for all continuous and bounded map $f : \mathcal{X} \to \mathbb{R}$.

Theorem 3.3. *[4] Let $\mu_0, \mu_1, \mu_2, \cdots$ be a sequence of laws on $\mathcal{X} \times \mathcal{X}$, each of whose marginals is a member of $P_p(\mathcal{X})$, $p \in [1, \infty[$. If $\mu_n \to \mu$ (weakly) as $n \to \infty$, then*

$$\liminf_{n\to\infty} \int \|x-y\|^p d\mu_n(x,y) \geqslant \int \|x-y\|^p d\mu_0(x,y).$$

Theorem 3.4. *[4] Given laws μ and ν in $P_p(\mathcal{X})$, $p \in [1, \infty[$, the infimum in* (3.4) *is attained for some law $\eta \in \Pi(\mu, \nu)$*

The function W_p is called Wasserstein functions.

Theorem 3.5. *[4] The Wasserstein functions W_p are metrics on the sets $P_p(\mathcal{X})$ for $p \in [1, \infty[$.*

Definition 3.6. *The Wasserstein distance of order p between μ and ν in $P_p(\mathcal{X})$ is defined by:*

$$W_p(\mu,\nu) = \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathcal{X}} \|x-y\|^p d\pi(x,y) \right)^{1/p}. \tag{3.4}$$

W_p defines a (finite) distance on $P_p(\mathcal{X})$. For more details on Wasserstein space see [12].

4 Otto metric

We consider a n-dimensional regular statistical manifold

$$\mathcal{M} = \{p(\cdot;\theta); \theta = (\theta_1, \cdots, \theta_n) \in \Theta\}$$

where Θ is an open subset of $\mathbb{R}^n$ and the mapping $\theta \mapsto p_\theta$ is injective.

Motivated by the study of a class of partial differential equation, in [10], Otto considered an inner product defined on smooth functions of the θ-fiber, $\top_\theta\mathcal{M}$ of the tangent bundle, as

$$(u,v) \mapsto \int_{\mathcal{X}} \mathrm{grad}(u)(x) \cdot \mathrm{grad}(v)(x) p(x;\theta) dx \tag{4.5}$$

where grad is the gradient function with respect to x. Then, this inner produit defines on $\mathcal{M}$ a Riemannian metric so called Otto metric g, with coordinate functions:

$$\begin{aligned} g_{ij} &= \int_{\mathcal{X}} \mathrm{grad}\left(\frac{\partial \ell(x;\theta)}{\partial\theta^i}\right) \cdot \mathrm{grad}\left(\frac{\partial \ell(x;\theta)}{\partial\theta^j}\right) p(x;\theta)dx \\ &= \mathbb{E}\left[\mathrm{grad}\left(\partial_{\theta_i}\ell\right) \cdot \mathrm{grad}\left(\partial_{\theta_j}\ell\right)\right]. \end{aligned} \tag{4.6}$$

The following theorem states that the Riemannian Otto metric is related to the Wasserstein metric.

Theorem 4.1. *[8] Let* $P^{\infty}(\chi) = \left\{ f : f \in C^{\infty}(\chi), f > 0, \int_{\chi} f(x)d(x) = 1 \right\}$. *If* $c : [0,1] \to P^{\infty}(\chi)$ *is a smooth immersed curve then its length* $L(c)$ *in the Wasserstein space* $P_2(\chi)$ *satisfies*

$$L(c) = \int_0^1 \langle c'(t), c'(t) \rangle^{\frac{1}{2}} dt$$

where

$$L(c) = \sup_{j \in \mathbb{N}} \sup_{0=t_0 \leq t_1 \leq \cdots \leq t_J = 1} \sum_{j=1}^{J} W_2(c(t_{j-1}), c(t_j)).$$

Proposition 4.2. *Let* $\mathcal{M} = \{p(\cdot;\theta); \theta = (\theta_1, \cdots, \theta_n) \in \Theta\}$ *be a statistical manifold endowed with the Otto metric* g *(4.6). For all* $i, j, k \in \{1, \cdots, n\}$

$$\begin{aligned} \frac{\partial g_{jk}}{\partial \theta^i} + \frac{\partial g_{ki}}{\partial \theta^j} - \frac{\partial g_{ij}}{\partial \theta^k} &= \mathbb{E}_{\theta} \left[2grad(\partial_{ij}\ell) grad(\partial_k \ell) \right. \\ &\quad + grad(\partial_j \ell) grad(\partial_k \ell) \partial_i \ell + grad(\partial_k \ell) grad(\partial_i \ell) \partial_j \ell \\ &\quad \left. - grad(\partial_j \ell) grad(\partial_i \ell) \partial_k \ell \right]. \end{aligned} \tag{4.7}$$

Proof. Taking the partial derivative of g_{ij} in Equation (4.6) with respect to θ_i, θ_j, θ_k yields the result. □

5 α-connection related to Otto metric

5.1 Construction of our α-connection

In the remainder, we consider an n-dimensional statistical manifold $\mathcal{M} = \{p(\cdot;\theta); \theta = (\theta_1, \cdots, \theta_n) \in \Theta\}$ where Θ is a open subset of $\mathbb{R}^n$ and the mapping $\theta \mapsto p_\theta$ is injective. We endowed $\mathcal{M}$ with the Riemannian Otto metric g which is related to Wasserstein distance. For any $\alpha \in \mathbb{R}$, $i, j, k \in \{1, \cdots, n\}$, we introduce the function $\Gamma^{(\alpha)}_{ij,k}$ which maps each point θ to the following value:

$$\Gamma^{(\alpha)}_{ij,k} = \mathbb{E}_{\theta} \left[\mathrm{grad}(\partial_{ij}\ell) \mathrm{grad}(\partial_k \ell) \right] + \frac{1-\alpha}{2} \top^{\alpha}_{ij,k} \tag{5.8}$$

where $\top^{\alpha}_{ij,k}$ is a tensor defined by :

$$\top^{\alpha}_{ij,k} = \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_j\ell)\mathrm{grad}(\partial_k\ell)\partial_i\ell\right] + \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_k\ell)\mathrm{grad}(\partial_i\ell)\partial_j\ell\right] + \\ -(1+\alpha)\mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_j\ell)\mathrm{grad}(\partial_i\ell\partial_k\ell)\right]. \quad (5.9)$$

Let ϕ be the parameter of dimension n of some parametrization of $\mathcal{M}$, alternative to that indicated by θ. Coordinates of ϕ will be denoted by $\phi = (\phi_1, \cdots, \phi_n)$, and we write ∂_{ϕ_u} for $\frac{\partial}{\partial \phi_u}$ and $\theta_{i/u} = \frac{\partial \theta_i}{\partial \phi_u}$.

Lemma 5.1. *For any change of coordinate system $\top^{\alpha}_{ij,k}$ satisfies the equation*

$$\top^{\alpha}_{uv,w} = \top^{\alpha}_{ij,k}\theta_{i/u}\theta_{j/v}\phi_{w/k}.$$

Proof. Using (5.9), we have

$$\begin{aligned} \top^{\alpha}_{ij,k}\theta_{i/u}\theta_{j/v}\phi_{w/k} &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_j\ell)\mathrm{grad}(\partial_k\ell)\partial_u\ell\right]\theta_{j/v}\phi_{w/k} \\ &\quad + \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_k\ell)\mathrm{grad}(\partial_i\ell\partial_v\ell)\right]\theta_{i/u}\phi_{w/k} \\ &\quad -(1+\alpha)\mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_j\ell)\mathrm{grad}(\partial_i\ell)\partial_w\ell\right]\theta_{i/u}\theta_{j/v} \\ &=: I_1 + I_2 - (1+\alpha)I_3. \end{aligned} \quad (5.10)$$

We have

$$\begin{aligned} I_1 &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_j\ell)\mathrm{grad}(\partial_k\ell)\partial_u\ell\right]\theta_{j/v}\phi_{w/k} \\ &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_v\ell)\mathrm{grad}(\partial_w\ell)\partial_u\ell\right] \\ &\quad + \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_v\ell)\partial_u\ell\partial_w\ell\right]\left(\mathrm{grad}(\theta_{j/v})\right)\theta_{j/v} \\ &\quad + \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_w\ell)\partial_u\ell\partial_v\ell\right]\left(\mathrm{grad}(\phi_{w/k})\right)\phi_{w/k} \\ &\quad + \mathbb{E}_{\theta}\left[\partial_v\ell\partial_u\ell\partial_w\ell\right]\left(\mathrm{grad}(\theta_{j/v})\right)\left(\mathrm{grad}(\phi_{w/k})\right)\theta_{j/v}\phi_{w/k} \\ &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_v\ell)\mathrm{grad}(\partial_w\ell)\partial_u\ell\right] \end{aligned} \quad (5.11)$$

because $\mathrm{grad}(\theta_{i/u}) = \mathrm{grad}(\theta_{j/v}) = \mathrm{grad}(\phi_{w/k}) = 0$. Similarly, we deduce

$$\begin{aligned} I_2 &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_w\ell)\mathrm{grad}(\partial_u\ell)\partial_v\ell\right] \\ I_3 &= \mathbb{E}_{\theta}\left[\mathrm{grad}(\partial_v\ell)\mathrm{grad}(\partial_u\ell)\partial_w\ell\right]. \end{aligned}$$

Then $\top^{\alpha}_{uv,w} = \top^{\alpha}_{ij,k}\theta_{i/u}\theta_{j/v}\phi_{w/k}$. □

The following result based on the transformation law (see [1]) gives a characterization of affine connections on a Riemannian manifold.

Lemma 5.2. *On Riemannian manifold $(\mathcal{M}, g)$:*

(a) *all affine connection* ∇ *with connection symbol* Γ_{ij}^k *(i.e.* $\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k$*) are of the form*

$$\Gamma_{ij}^k = \overset{\circ}{\Gamma}{}_{ij}^k + S_{ij}^k \tag{5.12}$$

where $S_{ij,k}$ *satisfies*

$$S_{uv}^w = S_{ij}^k \theta_{i/u}\theta_{j/v}\phi_{w/k}$$

(b) *any set of smooth functions* $\Gamma_{ij,k}$ *on* $\mathcal{M}$ *which satisfies the law (5.12) constitutes the connection symbols of an affine connection on* $\mathcal{M}$*.*

Proof. (a). Let ∇ an affine connection on M. Then

$$\begin{aligned}
\Gamma_{uv}^w \partial_w = \nabla_{\frac{\partial}{\partial \phi^u}} \frac{\partial}{\partial \phi^v} &= \nabla_{\frac{\partial \theta^i}{\partial \phi^u}\frac{\partial}{\partial \theta^i}} \left(\frac{\partial \theta^j}{\partial \phi^v} \frac{\partial}{\partial \theta^j} \right) \\
&= \frac{\partial \theta^i}{\partial \phi^u} \nabla_{\frac{\partial}{\partial \theta^i}} \left(\frac{\partial \theta^j}{\partial \phi^v} \frac{\partial}{\partial \theta^j} \right) \\
&= \frac{\partial \theta^i}{\partial \phi^u} \left[\frac{\partial \theta^j}{\partial \phi^v} \nabla_{\frac{\partial}{\partial \theta^i}} \frac{\partial}{\partial \theta^j} + \frac{\partial}{\partial \theta^i} \left(\frac{\partial \theta^j}{\partial \phi^v} \right) \frac{\partial}{\partial \theta^j} \right] \\
&= \frac{\partial \theta^i}{\partial \phi^u} \left[\frac{\partial \theta^j}{\partial \phi^v} \Gamma_{ij}^k \frac{\partial}{\partial \theta^k} + \frac{\partial}{\partial \theta^i} \left(\frac{\partial \theta^j}{\partial \phi^v} \right) \frac{\partial}{\partial \theta^j} \right] \\
&= \Gamma_{ij}^k \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} \frac{\partial}{\partial \phi^w} + \frac{\partial^2 \theta^j}{\partial \phi^u \partial \phi^v} \frac{\partial}{\partial \theta^j} \\
&= \Gamma_{ij}^k \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} \frac{\partial}{\partial \phi^w} + \frac{\partial^2 \theta^k}{\partial \phi^u \partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} \frac{\partial}{\partial \phi^w} \\
&= \left[\Gamma_{ij}^k \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} + \frac{\partial^2 \theta^w}{\partial \phi^u \partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} \right] \frac{\partial}{\partial \phi^w}.
\end{aligned}$$

Then

$$\Gamma_{uv}^w(\phi) = \Gamma_{ij}^k(\theta) \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} + \frac{\partial^2 \theta^w}{\partial \phi^u \partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k}.$$

It is well known that the Christofell symbol satisfies the transformation law: $\overset{\circ}{\Gamma}{}_{uv}^w(\phi) = \overset{\circ}{\Gamma}{}_{ij}^k(\theta) \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k} + \frac{\partial^2 \theta^w}{\partial \phi^u \partial \phi^v}$ (see [6]). One has

$$\Gamma_{uv}^w(\phi) - \overset{\circ}{\Gamma}{}_{uv}^w(\phi) = \left(\Gamma_{ij}^k(\theta) - \overset{\circ}{\Gamma}{}_{ij}^k(\theta) \right) \frac{\partial \theta^i}{\partial \phi^u} \frac{\partial \theta^j}{\partial \phi^v} \frac{\partial \phi^w}{\partial \theta^k}.$$

The last equation shows that

$$\Gamma_{ij}^{k} = \overset{\circ}{\Gamma}{}_{ij}^{k} + S_{ij}^{k}$$

where S_{ij}^{k} satisfies

$$S_{uv}^{w} = S_{ij}^{k}\theta_{i/u}\theta_{j/v}\phi_{w/k}$$

To proove (b), one shows that the following map:

$$\begin{array}{rlcl} \nabla : & \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) & \to & \mathcal{X}(\mathcal{M}) \\ & (X = x^i\partial_i, Y = y^j\partial_j) & \mapsto & x^i y^j \Gamma_{ij}^{k}\partial_k + x^i\partial_i(y^j)\partial_j \end{array}$$

is an affine connection on $\mathcal{M}$. □

Theorem 5.3. *Let $\mathcal{M} = \{p(\cdot;\theta); \theta = (\theta_1, \cdots, \theta_n) \in \Theta\}$ be a statistical manifold endowed with the Otto metric g (4.6). There exists an affine connection $\nabla^{(\alpha)} : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \to \mathcal{X}(\mathcal{M})$ defined by :*

$$g\left(\nabla^{(\alpha)}_{\partial_i}\partial_j, \partial_j\right) = \Gamma^{(\alpha)}_{ij,k}. \tag{5.13}$$

Proof. Set $\Gamma^{(\alpha),k}_{i,j} = \Gamma^{(\alpha)}_{ij,m} g^{mk}$. By using Lemma 5.2, Lemma 5.1 and Proposition 4.2

$$\begin{array}{rlcl} \nabla^{(\alpha)} : & \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) & \to & \mathcal{X}(\mathcal{M}) \\ & (X = x^i\partial_i, Y = y^j\partial_j) & \mapsto & x^i y^j \Gamma_{ij}^{(\alpha),k}\partial_k + x^i\partial_i(y^j)\partial_j \end{array} \tag{5.14}$$

is an affine connection on $\mathcal{M}$. The proof is completed. □

Now, we prove that this α-connection is torsion-free and for $\alpha = 0$ this connection is the Levi-Civita connection.

Theorem 5.4. *1. $\nabla^{(\alpha)}$ is a torsion-free affine connection.*

2. The 0-connection is the Levi-Civita connection with respect to the Otto metric.

Proof. 1. We have

$$\begin{aligned} \nabla^{(\alpha)}_{\partial_i}\partial_j - \nabla^{(\alpha)}_{\partial_j}\partial_i &= \Gamma_{ij}^{(\alpha),k}\partial_k - \Gamma_{ji}^{(\alpha),k}\partial_k \\ &= \Gamma_{ij}^{(\alpha),k}\partial_k - \Gamma_{ij}^{(\alpha),k}\partial_k \\ &= 0 \end{aligned}$$

where $\Gamma_{ij}^{(\alpha),k} = \Gamma^{(\alpha)}_{ij,m} g^{mk}$.

2. Taking the partial derivative of g_{ij} in Equation (4.6) with respect to θ_k, we obtain $\partial_k g_{ij} = \Gamma^{(0)}_{ij,k} + \Gamma^{(0)}_{kj,i}$.

□

5.2 Flatness of exponential and mixture families

Let's introduce now the notion of exponential family. In general, if an n-dimensional model $\mathcal{M} = \{p(\cdot;\theta), \theta \in \Theta\}$ can be expressed in terms of functions $\{C, F_1, \cdots, F_n\}$ on $\mathcal{X}$ and a function ψ on Θ such that

$$p(x;\theta) = \exp\left[C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta)\right], \tag{5.15}$$

then we say that $\mathcal{M}$ is an exponential family, and that the $[\theta^i]$ are its natural or its canonical parameters. Next, let's consider the case where an n-dimensional model $\mathcal{M}$ can be expressed in terms of functions $\{C, F_1, \cdots, F_n\}$ on $\mathcal{X}$ as

$$p(x;\theta) = C(x) + \sum_{i=1}^{n} \theta^i F_i(x), \tag{5.16}$$

then we say that $\mathcal{M}$ is a mixture family, and that the $[\theta^i]$ are its mixture parameters.
The following theorem gives the flatness result of exponential family.

Theorem 5.5. *An exponential family*

$$\mathcal{M} = \left\{p(\cdot;\theta) = \exp\left(C(\cdot) + \sum_{i=1}^{n} \theta^i F_i(\cdot) - \psi(\theta)\right), \theta \in \Theta\right\}$$

equipped with Otto metric is (1)*-flat.*

Proof. Let $p(\cdot;\theta) \in \mathcal{M}$ with $\mathcal{M}$ the exponential family. We have $p(x;\theta) = \exp\left\{C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta)\right\}$. One has

$$\ell(x) = C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta)$$

Then

$$\partial_i \ell(x) = F_i(x) - \partial_i \psi(\theta); \partial_{ij} \ell(x) = -\partial_{ij} \psi(\theta); \mathrm{grad}(\partial_{ij}\ell)(x) = 0.$$

Thus

$$\Gamma_{ij,k}^{(1)} = -\mathrm{grad}(\partial_{ij}\psi)(\theta) \cdot \mathbb{E}_\theta\left[\mathrm{grad}(\partial_k l)\right] = 0.$$

This completes the proof. ∎

Similarly, we state the flatness for a mixture family.

Theorem 5.6. *A mixture familly*

$$\mathcal{M} = \left\{ p(\cdot;\theta) = C(\cdot) + \sum_{i=1}^{n} \theta^i F_i(\cdot) - \psi(\theta), \theta \in \Theta \right\}$$

equiped with Otto metric is (-1)*-flat.*

Proof. Let $p(\cdot;\theta) \in \mathcal{M}$ with $\mathcal{M}$ the mixture family. We have $p(x;\theta) = C(x) + \sum_{i=1}^{n} \theta^i F_i(x)$. One has

$$\ell(x) = \log p(x;\theta).$$

Then

$$\partial_i \ell(x) = \frac{F_j(x)}{p(x;\theta)}; \partial_{ij}\ell(x) = -\frac{F_i(x)F_j(x)}{p^2(x;\theta)},$$

$$\begin{aligned} \mathrm{grad}(\partial_{ij}\ell) &= -\mathrm{grad}([\partial_i \ell \partial_j \ell]) \\ &= -\mathrm{grad}(\partial_i \ell)\partial_j \ell - \partial_i \ell \mathrm{grad}(\partial_j \ell). \end{aligned}$$

Thus

$$\mathrm{grad}(\partial_{ij}\ell)(x) \cdot \mathrm{grad}(\partial_k \ell)(x) = -\mathrm{grad}(\partial_i \ell) \cdot \mathrm{grad}(\partial_k \ell)(x)\partial_j \ell - \mathrm{grad}(\partial_j \ell)\mathrm{grad}(\partial_k \ell)(x)\partial_i \ell \quad (5.17)$$

Using the previous equations and the definition of $\Gamma_{ij,k}^{(\alpha)}$ (5.8) we have

$$\Gamma_{ij,k}^{(-1)} = 0.$$

We conclude that the mixture family is (-1)-flat. □

6 Wasserstein Gamma manifold

Let M be the set of Gamma distributions, that is,

$$M = \left\{ p(\cdot;\mu,\beta) | p(x;\mu,\beta) = \left(\frac{\beta}{\mu}\right)^{\beta} \frac{x^{\beta-1}}{\Gamma(\beta)} e^{-\frac{x\beta}{\mu}}; \mu > 0, \beta > 0 \right\}. \quad (6.18)$$

Identifying (μ, β) as a local coordinate system, $\mathcal{M}$ can be regarded as a manifold, the Gamma manifold.

Theorem 6.1. *The Otto metric and the coefficients of the α-connection on the Gamma manifold with respect to (μ, β) coordinate are given by :*

$$g_{11} = \frac{\beta^2}{\mu^4};\ g_{12} = g_{21} = 0;\ g_{22} = \frac{1}{\mu^2(\beta-1)}.$$

$$\Gamma_{11}^{(\alpha),1} = -(1+\alpha)\frac{3-\alpha}{2\mu}.$$

$$\Gamma_{11}^{(\alpha),2} = (\alpha-1)\frac{\beta(\beta-1)(2\mu^2-(1+\alpha))}{\mu^2}.$$

$$\Gamma_{22}^{(\alpha),1} = (\alpha-1)\frac{1+2\alpha(1-\beta^2)}{2\beta^2(1-\beta)}.$$

$$\Gamma_{22}^{(\alpha),2} = (1-\alpha)^2\frac{\left(\log\left(\mu\beta^{2\beta-1}\right)+2\beta^2-\beta+1\right)}{2\beta}.$$

$$\Gamma_{12}^{(\alpha),1} = \Gamma_{21}^{(\alpha),1} = -(\alpha-1)^2\frac{1}{2\beta}.$$

$$\Gamma_{12}^{(\alpha),2} = \Gamma_{21}^{(\alpha),2} = (\alpha-1)\frac{1+2\alpha(1-\beta^2)}{2\mu}.$$

Proof. One has

$$p(x;\mu,\beta) = \left(\frac{\beta}{\mu}\right)^{\beta}\frac{x^{\beta-1}}{\Gamma(\beta)}e^{-\frac{x\beta}{\mu}},$$

$$\ell(x) := \log p(x;\mu,\beta) = \beta\log\left(\frac{\beta}{\mu}\right) + (\beta-1)\log(x) - \log(\Gamma(\beta)) - \frac{x\beta}{\mu},$$

$$\partial_1\ell(x) = -\frac{\beta}{\mu} + \frac{x\beta}{\mu^2};\ \partial_2\ell(x) = \log(\beta) + 1 - \log(\mu) + \log(x) - \frac{\Gamma'(\beta)}{\Gamma(\beta)} - \frac{x}{\mu},$$

$$grad(\partial_1\ell)(x) = \frac{\beta}{\mu^2},\ grad(\partial_2\ell)(x) = \frac{1}{x} - \frac{1}{\mu}$$

for $x \in \mathbb{R}_+^*$.

$$\begin{aligned} g_{11} &= \int_{\mathbb{R}_+} grad(\partial_1\ell)(x)grad(\partial_1\ell)(x)p(x;\mu,\beta)dx \\ &= \int_{\mathbb{R}_+}\frac{\beta}{\mu^2}\frac{\beta}{\mu^2}p(x;\mu,\beta)dx \\ &= \frac{\beta^2}{\mu^4}. \end{aligned}$$

$$\begin{aligned} g_{21} = g_{12} &= \int_{\mathbb{R}_+} grad(\partial_1 \ell)(x) grad(\partial_2 \ell)(x) p(x;\mu,\beta)dx \\ &= \int_{\mathbb{R}_+} \frac{\beta}{\mu^2}\left(\frac{1}{x} - \frac{1}{\mu}\right) p(x;\mu,\beta)dx \\ &= \frac{\beta}{\mu^2}\int_{\mathbb{R}_+} \left(\frac{1}{x} - \frac{1}{\mu}\right) p(x;\mu,\beta)dx \\ &= \frac{\beta}{\mu^2}\left(\frac{1}{\mu} - \frac{1}{\mu}\right) = 0. \end{aligned}$$

$$\begin{aligned} g_{22} &= \int_{\mathbb{R}_+} grad(\partial_2 \ell)(x) grad(\partial_2 \ell)(x) p(x;\mu,\beta)dx \\ &= \int_{\mathbb{R}_+} \left(\frac{1}{x} - \frac{1}{\mu}\right)\left(\frac{1}{x} - \frac{1}{\mu}\right) p(x;\mu,\beta)dx \\ &= \int_{\mathbb{R}_+} \left(\frac{1}{x^2} - \frac{2}{x\mu} + \frac{1}{\mu^2}\right) p(x;\mu,\beta)dx \\ &= \frac{1}{\mu^2} + \frac{\beta}{\mu^2(\beta-1)} - \frac{2}{\mu^2} \\ &= \frac{1}{\mu^2(\beta-1)}. \end{aligned}$$

- Expression of $\Gamma_{11}^{(\alpha),1}$.

$$\Gamma_{11,1}^{(\alpha)} = \mathbb{E}\left(grad(\partial_{11}\ell) grad(\partial_1 \ell)\right) + \frac{1-\alpha}{2}\top_{11,1}^{\alpha}$$

where $\top_{11,1}^{\alpha}$ is defined by the relation 5.9 with $i = 1; j = 1; k = 1$.

$$\mathbb{E}\left(grad(\partial_{11}\ell) grad(\partial_1 \ell)\right) = -\frac{2\beta}{\mu^5}.$$

$$\begin{aligned} \mathbb{E}\left(grad(\partial_1 \ell) grad(\partial_1 \ell)\partial_1 \ell\right) &= \frac{\beta^2}{\mu^4}\int_{\mathbb{R}_+}\left(-\frac{\beta}{\mu} + \frac{x\beta}{\mu^2}\right) p(x;\mu,\beta)dx \\ &= \frac{\beta^2}{\mu^4}\left(-\frac{\beta}{\mu} + \frac{\beta}{\mu^2}\int_{\mathbb{R}_+} xp(x;\mu,\beta)dx\right) \\ &= \frac{\beta^2}{\mu^4}\left(-\frac{\beta}{\mu} + \frac{\beta}{\mu^2}\frac{(\beta+1)\mu}{\beta}\right) \\ &= \frac{\beta^2}{\mu^5}. \end{aligned}$$

Then

$$\Gamma_{11,1}^{(\alpha)} = -(1+\alpha)\frac{\beta^2(3-\alpha)}{2\mu^5}.$$

However

$$\begin{aligned}\Gamma_{11}^{(\alpha),1} &= \Gamma_{11,1}^{(\alpha)} g^{11} \\ &= -(1+\alpha)\frac{\beta^2(3-\alpha)}{2\mu^5} \times \frac{\mu^4}{\beta^2} \\ &= -(1+\alpha)\frac{3-\alpha}{2\mu}.\end{aligned}$$

- Expression of $\Gamma_{11}^{(\alpha),2}$.

$$\Gamma_{11,2}^{(\alpha)} = \mathbb{E}\left(grad(\partial_{11}\ell)grad(\partial_2\ell)\right) + \frac{1-\alpha}{2}\top_{11,2}^{\alpha}$$

where $\top_{11,2}^{\alpha}$ is defined by the relation 5.9 with $i=1; j=1; k=2$.

$$\begin{aligned}\mathbb{E}\left(grad(\partial_{11}\ell)grad(\partial_2\ell)\right) &= \frac{2\beta}{\mu^3}\int_{\mathbb{R}_+}\left(\frac{1}{x}-\frac{1}{\mu}\right)p(x;\mu,\beta)dx \\ &= \frac{2\beta}{\mu^3}\left(\int_{\mathbb{R}_+}\frac{1}{x}p(x;\mu,\beta)dx - \frac{1}{\mu}\right) \\ &= \frac{2\beta}{\mu^3}\left(\frac{1}{\mu}-\frac{1}{\mu}\right) \\ &= 0.\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(grad(\partial_1\ell)grad(\partial_2\ell)\partial_1\ell\right) &= \frac{\beta}{\mu^2}\int_{\mathbb{R}_+}\left(\frac{1}{x}-\frac{1}{\mu}\right)\left(-\frac{\beta}{\mu}+\frac{x\beta}{\mu^2}\right)p(x;\mu,\beta)dx \\ &= -\frac{\beta}{\mu^4},\end{aligned}$$

$$\mathbb{E}\left(grad(\partial_1\ell)grad(\partial_1\ell)\partial_2\ell\right) = -\frac{\beta}{\mu^4},$$

then

$$\Gamma_{11,2}^{(\alpha)} = (\alpha-1)\frac{\beta(2\mu^2-(1+\alpha))}{\mu^4}.$$

Consequently

$$\Gamma_{11}^{(\alpha),2} = \Gamma_{11,2}^{(\alpha)} g^{22} = (\alpha-1)\frac{\beta(\beta-1)(2\mu^2-(1+\alpha))}{\mu^2}.$$

- Expression of $\Gamma_{12}^{(\alpha),1}$.

$$\Gamma_{12,1}^{(\alpha)} = \mathbb{E}\left(grad(\partial_{12}\ell)grad(\partial_1\ell)\right) + \frac{1-\alpha}{2}\top_{12,1}^{\alpha}$$

where $\top_{12,1}^{\alpha}$ is defined by the relation 5.9 with $i=1; j=2; k=1$.

$$\mathbb{E}\left(grad(\partial_{12}\ell)grad(\partial_1\ell)\right) = \frac{\beta}{\mu^2}\times\frac{\beta}{\mu^2} = \frac{\beta^2}{\mu^4},$$

$$\mathbb{E}\left(grad(\partial_1\ell)grad(\partial_2\ell)\partial_1\ell\right) = -\frac{\beta}{\mu^4},$$

$$\begin{aligned}\mathbb{E}\left(grad(\partial_1\ell)grad(\partial_1\ell)\partial_2\ell\right) &= \frac{\beta^2}{\mu^4}\int_{\mathbb{R}_+}\left(\log(\beta e\mu^{-1}) + \log(x) - \frac{\Gamma'(\beta)}{\Gamma(\beta)} - \frac{x}{\mu}\right)p(x;\mu,\beta)dx \\ &= -\frac{\beta}{\mu^4}.\end{aligned}$$

Then

$$\Gamma_{12,1}^{(\alpha)} = -(\alpha-1)^2\frac{\beta}{2\mu^4}.$$

Consequently

$$\Gamma_{12}^{(\alpha),1} = \Gamma_{12,1}^{(\alpha)}g^{11} = -(\alpha-1)^2\frac{1}{2\beta}.$$

- Expression of $\Gamma_{12}^{(\alpha),2}$.

$$\Gamma_{12,2}^{(\alpha)} = \mathbb{E}\left(grad(\partial_{12}\ell)grad(\partial_2\ell)\right) + \frac{1-\alpha}{2}\top_{12,2}^{\alpha}$$

where $\top_{12,2}^{\alpha}$ is defined by the relation 5.9 with $i=1; j=2; k=2$.

$$\begin{aligned}\mathbb{E}\left(grad(\partial_{12}\ell)grad(\partial_2\ell)\right) &= \frac{1}{\mu^2}\int_{\mathbb{R}_+}\left(\frac{1}{x} - \frac{1}{\mu}\right)p(x;\mu,\beta)dx \\ &= \frac{1}{\mu^2}\left(\frac{1}{\mu} - \frac{1}{\mu}\right) = 0.\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(grad(\partial_2\ell)grad(\partial_2\ell)\partial_1\right) &= \int_{\mathbb{R}_+}\left(\frac{1}{x} - \frac{1}{\mu}\right)^2\left(-\frac{\beta}{\mu} + \frac{x\beta}{\mu^2}\right)p(x;\mu,\beta)dx \\ &= \frac{1}{\mu^3(1-\beta)}.\end{aligned}$$

$$\mathbb{E}\left(grad(\partial_2\ell)grad(\partial_1\ell)\partial_2\right) = -2\frac{1+\beta}{\mu^3}.$$

Then

$$\Gamma^{(\alpha)}_{12,2} = (1-\alpha)\frac{1+2\alpha(1-\beta^2)}{2(1-\beta)\mu^3}.$$

Consequently

$$\Gamma^{(\alpha),2}_{12} = \Gamma^{(\alpha)}_{12,2}g^{22} = (\alpha-1)\frac{1+2\alpha(1-\beta^2)}{2\mu}.$$

- Expression of $\Gamma^{(\alpha),1}_{22}$.

$$\Gamma^{(\alpha)}_{22,1} = \mathbb{E}\left(grad(\partial_{22}\ell)grad(\partial_1\ell)\right) + \frac{1-\alpha}{2}\top^{\alpha}_{22,1}$$

where $\top^{\alpha}_{22,1}$ is defined by the relation 5.9 with $i=2; j=2; k=1$.

$$\mathbb{E}\left(grad(\partial_{22}\ell)grad(\partial_1\ell)\right) = 0.$$

$$\mathbb{E}\left(grad(\partial_2\ell)grad(\partial_1\ell)\partial_2\right) = -2\frac{1+\beta}{\mu^3}.$$

$$\mathbb{E}\left(grad(\partial_2\ell)grad(\partial_2\ell)\partial_1\right) = -\frac{1}{\mu^3(\beta-1)}.$$

Then

$$\Gamma^{(\alpha)}_{22,1} = (1-\alpha)\frac{4\beta^2-\alpha-5}{2\mu^3(1-\beta)}.$$

Consequently

$$\Gamma^{(\alpha),1}_{22} = \Gamma^{(\alpha)}_{22,1}g^{11} = (\alpha-1)\frac{1+2\alpha(1-\beta^2)}{2\beta^2(1-\beta)}.$$

- Expression of $\Gamma^{(\alpha),2}_{22}$.

$$\Gamma^{(\alpha)}_{22,2} = \mathbb{E}\left(grad(\partial_{22}\ell)grad(\partial_2\ell)\right) + \frac{1-\alpha}{2}\top^{\alpha}_{22,2}.$$

where $\top^{\alpha}_{22,2}$ is defined by the relation 5.9 with $i = 2; j = 2; k = 2$.

$$\mathbb{E}\left(grad(\partial_{22}\ell)grad(\partial_2\ell)\right) = \frac{1}{\mu^3\beta(\beta-1)}\left[\log\left(\frac{\beta}{\mu}\right)^{\beta} + 2\beta^2 - \beta + 1\right]$$

Then

$$\Gamma^{(\alpha)}_{22,2} = (1-\alpha)^2 \frac{\log\left(\frac{\beta}{\mu}\right)^{\beta} + 2\beta^2 - \beta + 1}{2\mu^2\beta(\beta-1)}.$$

Consequently

$$\Gamma^{(\alpha),2}_{22} = \Gamma^{(\alpha)}_{22,2}g^{22} = (1-\alpha)^2 \frac{\log\left(\frac{\beta}{\mu}\right)^{\beta} + 2\beta^2 - \beta + 1}{2\beta}.$$

□

Theorem 6.2. *The sectional curvature of the Gamma manifold M endowed with the Otto metric with respect to (μ, β) coordinate is given by:*

$$\begin{aligned} K &= \frac{(1-\alpha)\mu^4}{4\beta^4(1-\beta)}\left[(1-\beta)(1-\beta)^3\left(\log\left(\frac{\beta}{\mu}\right)^{\beta} + 2\beta^2 - \beta + 1\right)\right. \\ &+ \left.(4\beta^2 - \alpha - 5)(3\alpha^2 + 4\alpha\beta^2(1-\beta) + 1) - 8\beta^2 + 2\beta(\alpha - 1) + 12\right] \end{aligned}$$

Proof. One has

$$K = \frac{R^1_{122}g_{11}}{g_{11}g_{22} - g^2_{22}}$$

with

$$R^{(\alpha),1}_{122} = \sum_{l=1}^{2}\Gamma^{(\alpha),l}_{12}\Gamma^{(\alpha),1}_{l2} - \sum_{l=1}^{2}\Gamma^{(\alpha),l}_{22}\Gamma^{(\alpha),1}_{1l} + \partial_2\Gamma^{(\alpha),1}_{12} - \partial_1\Gamma^{(\alpha),1}_{22}.$$

using the Theorem 6.1, the result follows. □

Theorem 6.3. *Let M the Gamma manifold endowed with the Otto metric defined by relation 6.18. Set $\nu = \beta/\mu$. Then (ν, β) is a natural coordinate system of 1-connection.*

Proof. Set $\nu = \beta/\mu$.

$$\begin{aligned} p(x;\nu,\beta) &= (\nu)^{\beta}\frac{x^{\beta-1}}{\Gamma(\beta)}e^{-x\nu} \\ &= \exp\left[-\log x + (\beta\log x - \nu x) - (\log\Gamma(\beta) - \beta\log\nu)\right]. \end{aligned}$$

Hence the set of all Gamma distribution is an exponential family. Then, using the Theorem 5.5 the result follows. □

Acknowledgments

This work was supported by CEA-SMA. The authors would like to thank the referees for their insightful and constructive comments and suggestions.

References

[1] S. I. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen, C. R. Rao, *Differential geometry in statistical inference*, Lecture notes-monograph series (S. Shanti, Series Editor) **10**.

[2] S. Amari, H. Nagaoko, *Methods of information geometry*, AMS, (2007).

[3] Y. Brenier, *Polar factorization and monotone rearrangement of vector valued functions*, Comm. Pure Appl. Math., **44**, no. 4, (1991), 375–417.

[4] R. Clark, Givens and Rae Michael Shortt: *A class of Wasserstein metrics for probability distribution*, Michigan Math. J., **31,** (1984).

[5] M. Do Carmo, *Riemannian Geometry*, Birkhauser Inc., Boston, (1992).

[6] E. de Souza, *Tensor Calculus for Engineers and Physicists*, (2016), Springer.

[7] A. Gbaguidi Amoussou, F. Djibril Moussa, C. Ogouyandjou, M. A. Diop, *New connections on the fiber-bundle of generalized statistical manifolds*, Balkan Society of Geometers Proc., **26**, (2019), 23–32.

[8] J. Lott: Some geometric calculation on Wasserstein space, Commun. Math. Phys., **277**, (2008), 423–437.

[9] R. J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11**, no. 3, (2001), 589–608.

[10] I. Olkin, F., Pukelsheim, *The distance between two random vectors with given dispersion matrices*, Linear Algebra Appl., **48**, (1982), 257–263.

[11] Sana Rebbah, Florence Nicol and Stéphane Puechmorel: *The Geometry of the Generalized Gamma Manifold and an Application to Medical Imaging*, Mathematics, (2019).

[12] C. Villani, *Topics in Optimal Transportation. Graduate studies in Mathematics,* AMS, **58**, (2003).