Mixture and Exponential Arcs on Generalized Statistical Manifold

Luiza H. F. de Andrade [1],* , Francisca L. J. Vieira [2], Rui F. Vigelis [3] and Charles C. Cavalcante [4]

[1]   Department of Natural Science, Mathematics and Statistics,Federal Rural University of Semi-Arid Region, Mossoró 59.625-900, Brazil

[2]   Department of Mathematics,Regional University of Cariri, Juazeiro do Norte 63105-000, Brazil; leidmar.vieira@urca.br

[3]   Computer Engineering, Campus Sobral, Federal University of Ceará, Sobral 62010-560, Brazil; rfvigelis@ufc.br

[4]   Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza 60455-900, Brazil; charles@ufc.br

*   Correspondence: luizafelix@ufersa.edu.br

Abstract: In this paper, we investigate the mixture arc on generalized statistical manifolds. We ensure that the generalization of the mixture arc is well defined and we are able to provide a generalization of the open exponential arc and its properties. We consider the model of a φ-family of distributions to describe our general statistical model.

1. Introduction

In the geometry of statistical models, information geometry [1–3] is the part of probability theory dedicated to investigate probability density functions equipped with differential geometry structure. A differential-geometric structure to the multi-parameter families of distributions was provided in [4]. In the mid-1980s, other topics related to the subject, such as fiber bundle theory and duality of connections of statistical models, were investigated by Amari [5] and Amari and Nagaoka [6], respectively. In the parametric case, exponential, mixture and α-connections, as well as their dual structure, are among the most important geometric objects [6], since the dual structure of the α-connections is the key point distinguishing statistical manifolds against arbitrary differential manifolds. Divergence function is an essential topic in information geometry, for both, parametric and non-parametric cases, since a metric and dual connections can be induced from a divergence [7–10]. To find an information-geometrical foundation for multi-parameter families of probability distributions, with a more general description, is one of topics of interest in information geometry [11–14]

Non-parametric statistical models [15] are important in a wide range of areas [16,17]. In the parametric case, the manifold of probability density functions obtains a Euclidian topology from the space of its natural parameters. As for the non-parameter case, a major challenge is to define a convenient topology and a notion of convergence. Pistone and Sempi [18] were the first to formulate a rigorous infinite dimensional extension. In that work, the set $\mathcal{P}_\mu$ of all strictly positive probability densities was endowed with a structure of exponential Banach manifolds, using Orlicz spaces associated to a Young function. In a later work [19], more properties of the statistical manifold were studied, specifically regarding the orthogonality condition.

Similar to in the parametric case, in non-parametric models, the mixture and exponential connections are among the most important geometric objects. To find these connections, it is necessary to guarantee the existence of the open arcs, which are the geodesics of the manifold. Using the notion

of exponential convergence, Gibilisco and Pistone [20] investigated those connections. In that work, the exponential and mixture connections were built in a way that the relation between them is the same as in the parametric case. Another approach was used in [21] where the mixture arc was additionally studied. Moreover, Grasselli [21] proved that two probability densities in the same neighborhood are connected by an open mixture arc if and only if the difference between their random variables is bounded.

The exponential statistical manifold was later studied in [22], with another system of charts, the statistical model $\mathcal{E}(p)$, called the maximal exponential model. Cena and Pistone [22] proved that this model is the set of all positive densities connected to a given positive density p by an open exponential arc and viceversa. In that work, it was used the open mixture arc and the open exponential arc to discuss properties of this model as e-connection and m-connection in the same way that in [6]. This exponential model $\mathcal{E}(p)$ with the open exponential and mixture arcs were also studied recently by Santacroce et al., 2016 [23] and Santacroce et al., 2017 [24], where a proof of duality properties of statistical models was provided. Examples of applications of non-parametric information geometry to statistical physics using the connection by open arcs were studied in [25].

The generalization of the exponential statistical manifold has been an active topic of research in the last years. Pistone [26] used the Kaniadaki's κ-exponential [27] in the construction of a statistical manifold. Vigelis and Cavalcante [28] proposed a φ-family of probability distributions $\mathcal{F}_c^{\varphi}$, which generalizes the exponential family $\mathcal{E}(p)$. This generalization is based on the replacement of the exponential function by a deformed exponential φ which satisfies some properties and provides to the set $\mathcal{P}_\mu$ a Banach manifold structure, so called generalized statistical manifold. In [29], a review of nonparametric information geometry with specific issues of the infinite dimensional setting is provided. In that work, the deformed exponential manifold was studied with a deformed exponential function defined in [30] and a model space was built according to the proposal in [28].

In [31] were given necessary and sufficient conditions for any two probability distributions being connected by a φ-arc. In this work, we ensure the existence of a generalized mixture arc for probability distributions in the same φ-family $\mathcal{F}_c^{\varphi}$, with a deformed exponential function which satisfies some properties. Moreover, we find a generalization of open exponential arcs and we prove, in the same way that in [22], that the φ-family $\mathcal{F}_c^{\varphi}$ is the component connected to a given positive density $p = \varphi(c)$ and viceversa.

The rest of the paper is organized as follows. In Section 2, we revisit results about Musielak–Orlicz space and φ-family of probability distributions. We also briefly recall about the subdifferential of a convex function. In Section 3, where we provide our main results, we ensure that the generalized mixture arc is well-defined. In Section 4, we discuss the generalized, exponential and mixture arcs. Finally, our conclusions and perspectives are stated in Section 5.

2. Preliminary Results

The statistical manifold $\mathcal{P}_\mu$ can be equipped with a structure of C^∞-Banach Manifold, using the Musielak–Orlicz space L^Φ associated to the Musielak–Orlicz function $\Phi_c(t, u) = \varphi(t, c(t) + u(t)) - \varphi(t, c(t))$. Each connected component of the statistical manifold gives rise to a φ-family of probability distributions $\mathcal{F}_c^{\varphi}$. In this section, we provide an introduction of Musielak–Orlicz spaces and the construction of the φ-family of probability distributions.

2.1. φ-Families of Probability Distributions

Let (T, Σ, μ) be a σ-finite, non-atomic measure space. A function $\Phi : T \times [0, \infty) \to [0, \infty]$ is said to be a Musielak–Orlicz function if

(i) $\Phi(t, \cdot)$ is convex and lower semi-continuous for μ-a.e. (almost everywhere) $t \in T$,
(ii) $\Phi(t, 0) = \lim_{u \downarrow 0} \Phi(t, u) = 0$ and $\lim_{u \to \infty} \Phi(t, u) = \infty$ for μ-a.e. $t \in T$,
(iii) $\Phi(\cdot, u)$ is measurable for each $u \geq 0$.

We notice that $\Phi(t,\cdot)$, by (i)-(ii), is not equal to 0 or ∞ on the interval $(0,\infty)$.

Let L^0 be the linear space of all real-value, measurable functions on T. Given a Musielak–Orlicz function Φ, we denote the functional $I_\Phi(u) = \int_T \Phi(t, |u(t)|)d\mu$, for any $u \in L^0$. The Musielak–Orlicz space, Musielak–Orlicz class, Morse–Transue space generated by a Musielak–Orlicz function Φ are defined, respectively, by

$$L^\Phi = \{u \in L^0 : I_\Phi(\lambda u) < \infty \text{ for some } \lambda > 0\}$$

$$\tilde{L}^\Phi = \{u \in L^0 : I_\Phi(u) < \infty\},$$

and

$$E^\Phi = \{u \in L^0 : I_\Phi(\lambda u) < \infty \text{ for all } \lambda > 0\}.$$

The Musielak–Orlicz space L^Φ is a Banach space when it is equipped with the Luxemburg norm given by

$$\|u\|_\Phi = \inf\left\{\lambda > 0 : I_\Phi\left(\frac{u}{\lambda}\right) \le 1\right\},$$

or the Orlicz norm, represented as

$$\|u\|_{\Phi,0} = \sup\left\{\left|\int_T uv\,d\mu\right| : v \in \tilde{L}^{\Phi^*} \text{ and } I_{\Phi^*}(v) \le 1\right\},$$

where $\Phi^*(t,v) = \sup_{u\ge 0}(uv - \Phi(t,u))$ is the Fenchel conjugate of $\Phi(t,\cdot)$, which is also a Musielak–Orlicz function. These norms are equivalent and the inequalities $\|u\|_\Phi \le \|u\|_{\Phi,0} \le 2\|u\|_\Phi$ hold for all $u \in L^\Phi$ [32]. A Musielak–Orlicz function is said to satisfy the Δ_2-condition, or belong to the Δ_2-class (denoted by $\Phi \in \Delta_2$), if we can find a constant $\alpha > 0$ and a non-negative function $f \in \tilde{L}^\Phi$ such that

$$\alpha\Phi(t,u) \le \Phi\left(t,\frac{1}{2}u\right), \qquad \text{for all } u \ge f(t), \quad \text{and } \mu\text{-a.e. } t \in T. \tag{1}$$

If the Musielak–Orlicz function Φ satisfies the Δ_2-condition, then $I_\Phi(u) < \infty$ for every $u \in L^\Phi$ [32]. In this case L^Φ, $\tilde{L}^\Phi$ and E^Φ are equal as sets. Moreover, if the Musielak–Orlicz function Φ does not satisfy the Δ_2-condition, E^Φ is a proper subspace of L^Φ. Every function Φ that satisfies the Δ_2-condition is finite-value. Indeed, we define

$$b_\Phi(t) = \sup\{u \ge 0 : \Phi(t,u) < \infty\}, \tag{2}$$

and assuming that $b_\Phi(t) < \infty$, we get $\Phi(t,\frac{1}{2}u) < \alpha\Phi(t,u) = \infty$ for all $b_\Phi(t) < u < 2b_\Phi(t)$ which implies that Φ cannot satisfy the Δ_2-condition. For more information see for instance [32,33].

We say that a Musielak–Orlicz function Φ satisfies the ∇_2-condition, or belongs to ∇_2-class, if we can find a constant $\gamma > 1$, and a non-negative function $f \in \tilde{L}^\Phi$ such that

$$\gamma\Phi(t,u) \le \Phi\left(t,\frac{1}{2}\gamma u\right), \qquad \text{for all } u > f(t). \tag{3}$$

We notice that, if $\Phi \in \nabla_2$, then

$$d_\Phi = \lim_{u\to\infty}\frac{\Phi(t,u)}{u} = \lim_{u\to\infty}\Phi'_-(t,u) = \lim_{u\to\infty}\Phi'_+(t,u) = \infty.$$

Example 1. *The function* $\Phi : [0,\infty) \to [0,\infty)$ *defined by:*

$$\Phi(u) = \exp(u) - u - 1$$

satisfies the ∇_2-condition and does not satisfy the Δ_2-condition.

The (topological) dual space of L^Φ, is denoted by $(L^\Phi)^*$ and represented in the following way [32,34,35]

$$\left(L^\Phi\right)^* = L^{\Phi^*} \oplus (L^\Phi)^\sim_s,$$

where L^{Φ^*} is the set of the order continuous functionals and $(L^\Phi)^\sim_s$ is formed by singular components. If the Musielak–Orlicz function $\Phi_c \in \Delta_2$ then all functionals in $(L^\Phi)^*$ are order continuous and represented by

$$f_{v^*}(u) := \int_T uv^* d\mu, \quad \text{for all } u \in L^\Phi. \tag{4}$$

Otherwise, if $\Phi \notin \Delta_2$, then the functionals f in $(L^\Phi)^*$ can be uniquely expressed as

$$f = f_c + f_s,$$

where f_c is the order continuous component and f_s is the singular component.

While exponential families are based on the exponential function, φ-families are based on deformed exponential functions. A deformed exponential $\varphi : T \times \mathbb{R} \to (0, \infty)$ is a function that satisfies the following properties, for μ-a.e. $t \in T$ [28]:

(i) $\varphi(\cdot)$ is convex and injective;
(ii) $\lim_{u \to -\infty} \varphi(u) = 0$ and $\lim_{u \to \infty} \varphi(u) = \infty$;
(iii)   There exists a measurable function $u_0 \colon T \to (0, \infty)$ such that

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \qquad \text{for all } \lambda > 0,$$

for every measurable function $c \colon T \to \mathbb{R}$ for which $\int_T \varphi(c) d\mu = 1$.

In de Souza et al. [36], Lemma 1, it was shown that the constraint $\int_T \varphi(c) d\mu = 1$ can be replaced by $\int_T \varphi(c) d\mu < \infty$. Thus, the condition (iii) can be rewritten as:

(iii′) There exists a measurable function $u_0 \colon T \to (0, \infty)$ such that

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \qquad \text{for all } \lambda > 0,$$

for every measurable function $c \colon T \to \mathbb{R}$ for which $\int_T \varphi(c) d\mu < \infty$.

There are many examples of deformed exponential functions. An example of relevance is the exponential function $\varphi(x) = \exp(x)$ that satisfies (i)-(iii) with $u_0 = \mathbf{1}_T$. Another example is Kaniadakis' κ-exponential [26–28]:

Example 2. *The Kaniadakis' κ-exponential* $\exp_\kappa : \mathbb{R} \to (0, \infty)$ *for* $\kappa \in [-1, 1]$ *is defined as*

$$\exp_\kappa(u) = \begin{cases} \left(\kappa u + \sqrt{1 + \kappa^2 u^2}\right)^{\frac{1}{\kappa}}, & \text{if } \kappa \neq 0, \\ \exp(u) & \text{if } \kappa = 0. \end{cases}$$

The inverse of $\exp_\kappa$ *is the Kaniadakis' κ-logarithm*

$$\ln_\kappa(u) = \begin{cases} \frac{u^\kappa - u^{-\kappa}}{2\kappa}, & \text{if } \kappa \neq 0, \\ \ln(u) & \text{if } \kappa = 0. \end{cases}$$

One can easily notice the κ-exponential satisfies (i)–(iii) *[28,36].*

The Musielak–Orlicz function

$$\Phi_c(t, u) = \varphi(t, c(t) + u) - \varphi(t, c(t)) \tag{5}$$

for a measurable function $c : T \to \mathbb{R}$ such that $\varphi(t, c(t))$ is μ-integrable, was defined in [28]. Thus, the sets L^{Φ_c}, $\tilde{L}^{\Phi_c}$ and E^{Φ_c} are denoted by L_c^{φ}, $\tilde{L}_c^{\varphi}$ and E_c^{φ}, respectively, when Φ_c is given by (5). Let

$$\mathcal{P}_\mu = \left\{ p \in L^0 : p > 0 \text{ and } \int_T p d\mu = 1 \right\}$$

be the collection whose φ-family is a subset, where L^0 is the linear space of all real-valued. For each probability density $p \in \mathcal{P}_\mu$, we have a φ-family of probability density associated, $\mathcal{F}_c^{\varphi} = \varphi_c(\mathcal{B}_c^{\varphi}) \subset \mathcal{P}_\mu$ according to

$$\varphi_c(u) = \varphi(c + u - \psi(u)u_0), \quad \text{for each } u \in \mathcal{B}_c^{\varphi}, \tag{6}$$

where the set $\mathcal{B}_c^{\varphi}$ is the intersection of the convex set

$$\mathcal{K}_c^{\varphi} = \left\{ u \in L_c^{\varphi} : \int_T \varphi(c + \lambda u) d\mu < \infty \text{ for some } \lambda > 1 \right\},$$

with the closed subspace

$$B_c^{\varphi} = \left\{ u \in L_c^{\varphi} : \int_T u \varphi_+'(c) d\mu = 0 \right\}, \tag{7}$$

that is $\mathcal{B}_c^{\varphi} = \mathcal{K}_c^{\varphi} \cap B_c^{\varphi}$. The normalizing function $\psi : \mathcal{B}_c^{\varphi} \to [0, \infty)$ is introduced so that expression (6) is a probability distribution in $\mathcal{P}_\mu$. Suppose that the Musielak–Orlicz function Φ_c does not satisfy the Δ_2-condition, we have that the boundary of $\mathcal{B}_c^{\varphi}$, the set $\partial\mathcal{B}_c^{\varphi}$, is not empty. A function $u \in B_c^{\varphi}$ belongs to $\partial\mathcal{B}_c^{\varphi}$ if only if $\int_T \varphi(c + \lambda u) d\mu < \infty$ for all $\lambda \in (0, 1)$, and $\int_T \varphi(c + \lambda u) d\mu = \infty$ for each $\lambda > 1$. The behavior of the normalizing function near the boundary was studied in [33,37].

It is shown that the normalizing function $\psi : \mathcal{K}_c^{\varphi} \to \mathbb{R}$ is a convex function [28]. Assuming that φ is continuously differentiable, the normalizing function is Gâteaux-differentiable and the expression for Gâteaux-derivative is

$$\partial\psi(u)v = \frac{\int_T v\varphi'(c + u - \psi(u)u_0) d\mu}{\int_T u_0 \varphi'(c + u - \psi(u)u_0) d\mu}, \tag{8}$$

with $u \in \mathcal{K}_c^{\varphi}$ and $v \in L_c^{\varphi}$.

In the next section, we recall some differentiability properties of convex functions on infinite dimensional spaces.

2.2. The Subdifferential of a Convex function

In this section, we discuss some properties of extended real-valued convex functions in Banach spaces, i.e., functions with values in $\mathbb{R} \cup \{\pm\infty\}$. Mainly, we recall subdifferentials of lower semicontinuous convex functions and its properties.

Let E be a Banach space. A function f is a convex function on E, with the epigraph [38]

$$\text{epi } f = \{(x, \alpha) : x \in E, \alpha \in \mathbb{R}, \alpha \geq f(x)\}.$$

If $f(x) > -\infty$ for every x and $f(x) < +\infty$ for at least one value of x, we call f a proper function. The set

$$\text{dom } f = \{x \in E : f(x) < \infty\}$$

denotes the effective domain of f. A function $f : E \to (-\infty, \infty]$ is said to be lower semicontinuous (l.s.c.) if for every $\lambda \in \mathbb{R}$ the set

$$[f \leq \lambda] = \{x \in E : f(x) \leq \lambda\}$$

is closed.

Let E^* be the dual space of E. A vector $x^* \in E^*$ is said to be a subgradient of f at $x \in E$ if

$$(x^*, z) \leq f(x + z) - f(x) \quad \text{for all } z \in E.$$

We denote by $\partial f(x)$ the set of subgradients of f at x and the subdifferential of f is the multivalued mapping $x \mapsto \partial f(x)$ from E to E^*. By definition, $\partial f(x)$ is always a closed convex subset of E^* for each x. Suppose f is a convex function finite at x. One has $x^* \in \partial f(x)$ if and only if

$$(z, x^*) \leq f'(x; z), \quad \forall z \in E,$$

where

$$f'(x; z) = \lim_{t \to 0^+} \frac{f(x + tz) - f(x)}{t}$$

is the directional derivative of f at x in direction $z \in E$. The subdifferential may be empty at points of dom f, so we denote by

$$D(\partial f) = \{x \in E : \partial f(x) \neq \varnothing\},$$

the domain of ∂f and we have that $D(\partial f) \subset$ dom f. We say that f is subdifferentiable at x for all $x \in D(\partial f)$.

Let f be a lower semicontinuous proper convex function, then int dom $f \subset D(\partial f)$ [39] (Corollary 2.38). The conjugate of f is the function $f^* : E^* \to \overline{\mathbb{R}}$ defined by

$$f^*(x^*) = \sup\{(x, x^*) - f(x) : x \in E\}, \quad x^* \in E^*. \tag{9}$$

Observe that, if f is proper, then "sup" in Equation (9) may be restricted to the points $x \in$ dom f. The conjugate f^* is a convex and lower semicontinuous function on E^* and jointly with f satisfy the well known Young's inequality

$$(x, x^*) \leq f(x) + f^*(x^*), \tag{10}$$

with equality holding if and only if $x^* \in \partial f(x)$. If f is a lower semicontinuous function, the subdifferential ∂f^* of the conjugate function f^* coincides with $(\partial f)^{-1}$ ([39], Proposition 2.33).

It is known that, if f is a lower semicontinuous proper convex function, then

$$\text{int dom } f \subset D(\partial f) \subset \text{dom } f,$$

and it was shown in [40] that $D(\partial f)$ is, in fact, dense in dom f.

Fact 1 (([41], Corollary 2.19), ([42], Corollary 7.2.3)). *Suppose $x \in D(\partial f)$. Then $x \in$ int dom f if only if ∂f is locally bounded at x.*

Fact 2 (([41], Lemma 2.20), ([42], Lemma 7.2.4)). *If int dom $f \neq \varnothing$ and $x \in D(\partial f) \setminus$ int dom f, then $\partial f(x)$ is unbounded.*

The subdifferential of a convex function is closely related to Gâteaux-gradient. If the convex function f is Gâteaux-differentiable in $x_0 \in E$, then $\partial f(x_0)$ consists of a single element $x^* = \text{grad } f(x_0)$ ([39], Proposition 2.40), where grad $f(x)$ is the Gâteaux-gradient of f at x.

In the next section, we investigate the subdifferential of the normalizing function ψ. This result will be useful for us to prove that the generalized mixture arc is well defined, which is one of our main goals in this work.

3. Construction of Generalized Mixture Arcs

The normalizing function $\psi : \mathcal{K}_c^{\varphi} \to \mathbb{R}$ is convex and Gâteaux-differentiable and this derivative is given by Equation (8). Hence, with these facts in mind, we can provide the expression for the generalized mixture arc as given by:

$$p(t) = F^{-1}((1-t)F(p) + tF(q)),　　　　　　　(11)$$

where

$$F(p) = \frac{\varphi'(\varphi^{-1}(p))}{\int_T u_0 \varphi'(\varphi^{-1}(p))d\mu},　　　　　　　(12)$$

and p, q belong to a φ-family $\mathcal{F}_c^{\varphi}$. We can rewrite the functional $F(p)$ as

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0)d\mu},　　　u \in \mathcal{B}_c^{\varphi},　　　　(13)$$

with $p = \varphi_c(c + u - \psi(u)u_0)$ and Equation (13) is the Gâteaux-gradient of ψ. Thus, for the generalized mixture arc to be well defined, it is necessary that the set of these functionals in Equation (13) be convex. As mentioned in Section 2.2, the subdifferential and Gâteaux-gradient are closely related. For this reason, we investigate the subdifferential of ψ.

3.1. Subdifferential of the Normalizing Function ψ

Considering that the Musielak–Orlicz function (5) does not satisfy the Δ_2-condition, then we have that $\partial \mathcal{B}_c^{\varphi}$ is not-empty [33]. The effective domain of the normalizing ψ, the set $\text{dom } \psi = \left\{ u \in B_c^{\varphi} : \psi(u) < \infty \right\}$ is

$$\text{dom } \psi = \mathcal{B}_c^{\varphi} \cup \{\partial \mathcal{B}_c^{\varphi}\}_{<\infty},　　　　　(14)$$

where $\{\partial \mathcal{B}_c^{\varphi}\}_{<\infty}$ is the set of points in the boundary of $\mathcal{B}_c^{\varphi}$ such that $\psi(u) < \infty$. The behavior of the normalizing function ψ near the boundary of $\mathcal{B}_c^{\varphi}$ was discussed in [33]. We need to know the subdifferentials of ψ. Hence, we have to prove some properties of ψ, then we have our first result.

Proposition 1. *The normalizing function* $\psi : B_c^{\varphi} \to \mathbb{R} \cup \{\infty\}$ *is lower semicontinuous.*

Proof. Given $\alpha \in \mathbb{R}$, let C_α be the set $C_\alpha = \{u \in B_c^{\varphi} : \psi(u) \leq \alpha\}$. To prove the statement, it suffices to show that C_α is closed. We define a set

$$B = \left\{ u \in B_c^{\varphi} : \int_T \varphi(c + u - \alpha u_0)d\mu \leq 1 \right\},$$

and we are going to prove that B is a closed set and that $B = C_\alpha$. Let $\{u_n\}$ be a sequence which belongs to B, such that $\| u_n - u \|_{\Phi_c} \to 0$. This way, $u_n \to u$, μ-a.e. Since φ is a continuous function, we have that $\varphi(c + u_n - \alpha u_0) \to \varphi(c + u + \alpha u_0)$, μ-a.e. From Fatou's Lemma, it follows that

$$\int_T \varphi(c + u - \alpha u_0)d\mu = \int_T \liminf_{n \to \infty} \varphi(c + u_n - \alpha u_0)d\mu$$
$$\leq \liminf_{n \to \infty} \int_T \varphi(c + u_n - \alpha u_0)d\mu$$
$$\leq 1$$

thus, $u \in B$ and B is a closed set. Now, we prove that $B = C_\alpha$. Let u be a function which belongs to C_α, then $\psi(u) \le \alpha$. The function φ is a strictly increasing function, so that

$$\int_T \varphi(c + u - \alpha u_0)d\mu \le \int_T \varphi(c + u - \psi(u)u_0)d\mu = 1,$$

thus, $u \in B$.

Suppose that there exists $w \in B \setminus C_\alpha$, then $w \in B$, which implies that $\int_T \varphi(c + w - \alpha u_0)d\mu \le 1$ and $w \notin C_\alpha$, which implies that $\psi(w) > \alpha$. Then

$$\int_T \varphi(c + w - \alpha u_0)d\mu > \int_T \varphi(c + w - \psi(w)u_0)d\mu = 1,$$

thus $\int_T \varphi(c + w - \alpha u_0)d\mu > 1$. This contradicts the assumption that $w \in B$. Therefore, $B = C_\alpha$ and C_α is closed. □

The subdifferential of ψ at a function $u \in \text{dom } \psi$ is the set

$$\partial\psi(u) = \left\{ u^* \in (L^{\Phi_c})^* : \int_T u^* v d\mu \le \psi(u + v) - \psi(u), \text{ for all } v \in B_c^\varphi \right\}, \tag{15}$$

where $(L^{\Phi_c})^*$ denotes the dual space of L^{Φ_c}. We know that, for all $u \in \mathcal{B}_c^\varphi$ the normalizing function ψ is Gâteaux-differentiable and the Gâteaux-gradient is given by Equation (13). Hence, $\partial\psi(u)$ consists of a single element and is given by

$$\partial\psi(u) = \frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0 d\mu}, \qquad u \in \mathcal{B}_c^\varphi.$$

In fact, we prove below that Equation (13) belongs to $\partial\psi(u)$, for all $u \in \mathcal{B}_c^\varphi$.

Proposition 2. *Let u be a function in dom ψ. Supposing that the functional*

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0)d\mu} \tag{16}$$

belongs to $L^{\Phi_c^*}$, then (16) belongs to $\partial\psi(u)$.

Proof. We have that the functional (16) belongs to $L^{\Phi_c^*}$. Let v be a function in B_c^φ such that $\int_T \varphi(c + u + v)d\mu < \infty$. In other words, $u + v \in \text{dom } \psi$, so we have that $\int_T \varphi(c + u - \psi(u)u_0)d\mu = 1$ and $\int_T \varphi(c + u + v - \psi(u + v)u_0)d\mu = 1$. Thus, by the convexity of φ, we have

$$\int_T [v + (\psi(u + v) - \psi(u))u_0]\varphi'(c + u - \psi(u)u_0)d\mu \le$$

$$\int_T \varphi(c + u - \psi(u)u_0)d\mu - \int_T \varphi(c + u + v - \psi(u + v)u_0)d\mu = 0.$$

Thus,

$$\int_T v\varphi'(c + u - \psi(u)u_0)d\mu \le \int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu(\psi(u + v) - \psi(u)),$$

and

$$\frac{\int_T v\varphi'(c + u - \psi(u)u_0)d\mu}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \le \psi(u + v) - \psi(u). \tag{17}$$

If $u + v \in B_c^\varphi \setminus \text{dom } \psi$, then $\psi(u + v) = \infty$, and

$$\frac{\int_T v\varphi'(c + u - \psi(u)u_0)d\mu}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} < \psi(u + v) - \psi(u).$$

Consequently, Inequality (17) holds for all $v \in B_c^\varphi$ and the result follows. $\square$

We need to find the subdifferential of ψ for u in the set $\{\partial\mathcal{B}_c^\varphi\}_{<\infty}$. We know that ψ is a proper lower semicontinuous convex function, so

$$\text{int dom } \psi \subset D(\partial\psi) \subset \text{dom } \psi,$$

where int dom $\psi = \mathcal{B}_c^\varphi$ and $D(\partial\psi) \setminus \text{int dom } \psi = \{\partial\mathcal{B}_c^\varphi\}_{<\infty}$. As we have that int dom $\psi \neq \varnothing$, then for $u \in D(\partial\psi) \setminus \text{int dom } \psi$, $\partial\psi(u)$ is unbounded.

Since we are interested to prove that the set of functionals in Equation (13) is convex and these functionals are order continuous, we need to analyze only the order continuous part of the subdifferential, i.e., the part of the subdifferential that belongs to $L^{\Phi_c^*}$. We need to investigate whether the functional in Equation (16) belongs to $L^{\Phi_c^*}$, for $u \in \{\partial\mathcal{B}_c^\varphi\}_{<\infty}$. For this, we will use some results.

Lemma 1 ([35], Lemma 3.11). *Let Φ_c be a Musielak–Orlicz function that does not satisfy the Δ_2-condition. In addition, assume that $\Phi_c(t, b_\Phi(t)) = \infty$ for μ-a.e. $t \in T$. Then there exist a strictly increasing sequence $0 < \lambda_n \uparrow 1$, and sequences $\{u_n\}$ and $\{A_n\}$ of finite-valued, non-negative, measurable functions, and pairwise disjoint, measurable sets, respectively, such that*

$$I_{\Phi_c}(u_n\chi_{A_n}) = 1, \quad \text{and} \quad I_{\Phi_c}(\lambda_n u_n\chi_{A_n}) \leq 2^{-n}, \qquad \text{for all } n \geq 1.$$

Proposition 3 ([43], Proposition 2.3). *Let Φ and Ψ be Musielak–Orlicz functions. Suppose that, for constants $\alpha, \lambda > 0$, there exists an integrable function $h : T \to [0, \infty)$ such that*

$$\alpha\Psi(t, u) \leq \Phi(t, \lambda u) + h(t), \quad \text{for all } u \geq 0.$$

Then, for constants $\alpha' \in (0, \alpha)$ and $\lambda' = \lambda$, or $\alpha' = \alpha$ and $\lambda' > \lambda$, a non-negative function $f \in \tilde{L}^\Psi$ can be found such that

$$\alpha'\Psi(t, u) \leq \Phi(t, \lambda'u), \quad \text{for all } u > f(t).$$

Lemma 2. *Let Φ^* and Ψ^* denote the complementary functions to the Musielak–Orlicz functions Φ and Ψ, respectively. Suppose that, for constants $\alpha, \lambda > 0$, there exists a non-negative function $f \in \tilde{L}^\Psi$ such that*

$$\alpha\Psi(t, u) \leq \Phi(t, \lambda u), \qquad \text{for all } u > f(t). \tag{18}$$

Then, for constants $\alpha' = 1/\alpha$ and $\lambda' > \lambda/\alpha$, or $\alpha' \in (0, 1/\alpha)$ and $\lambda' = \lambda/\alpha$, a non-negative function $g \in \tilde{L}^{\Phi^*}$ can be found such that

$$\alpha'\Phi^*(t, v) \leq \Psi^*(t, \lambda'v), \qquad \text{for all } v > g(t). \tag{19}$$

Proof. Defining the function $h(t) = \Psi(t, f(t))$, we can write

$$\alpha\Psi(t, u) \leq \Phi(t, \lambda u) + \alpha h(t), \qquad \text{for all } u \geq 0.$$

Calculating the Fenchel conjugate of the functions in the inequality above, we obtain

$$\frac{1}{\alpha}\Phi^*(t, v) \leq \Psi^*\left(t, \frac{\lambda}{\alpha}v\right) + h(t), \qquad \text{for all } v \geq 0.$$

From Proposition 3, we infer that Equation (19) is satisfied. $\square$

Lemma 3. *The Δ_2-condition is equivalent to the statement that, for every $\lambda \in (0, 1)$, there exist a constant $\alpha_\lambda \in (0, 1)$, and a non-negative function $f_\lambda \in \tilde{L}^\Phi$ such that*

$$\alpha_\lambda \Phi(t, u) \leq \Phi(t, \lambda u), \qquad \text{for all } u > f_\lambda(t). \tag{20}$$

The ∇_2-condition is equivalent to the statement that, for any $\lambda \in (0, 1)$, there exist a constant $\gamma_\lambda > 1$, and a non-negative function $f_\lambda \in \tilde{L}^\Phi$ such that

$$\gamma_\lambda \Phi(t, u) \leq \Phi(t, \lambda \gamma_\lambda u), \qquad \text{for all } u > f_\lambda(t). \tag{21}$$

Proof. Suppose it satisfies the Δ_2-condition. If the natural number $n \geq 1$ is such that $2^{-n} \leq \lambda$, then $\alpha^n \Phi(t, u) \leq \Phi(t, 2^{-n}u) \leq \Phi(t, \lambda u)$, for all $u > 2^{n-1}f(t)$. Conversely, if Φ satisfies Equation (20) and the natural number $n \geq 1$ is chosen so that $\lambda^n \leq 1/2$, then $\alpha_\lambda^n \Phi(t, u) \leq \Phi(t, \lambda^n u) \leq \Phi(t, \frac{1}{2}u)$, for all $u > \lambda^{-n+1}f_\lambda(t)$.

Assume that Equation (3) is satisfied. Let $n \geq 1$ be a natural number such that $2^{-n} \leq \lambda$. Then $\gamma^n \Phi(t, u) \leq \Phi(t, 2^{-n}\gamma^n u) \leq \Phi(t, \lambda \gamma^n u)$, for all $u > f(t)$. Conversely, if Equation (21) holds and the natural number $n \geq 1$ is chosen so that $\lambda^n \leq 1/2$, then $\gamma_\lambda^n \Phi(t, u) \leq \Phi(t, \lambda^n \gamma_\lambda^n u) \leq \Phi(t, \frac{1}{2}\gamma_\lambda^n u)$, for all $u > f(t)$. $\square$

The next result follows from Lemmas 2 and 3.

Theorem 1. *A Musielak–Orlicz function Φ_c satisfies the ∇_2-condition if, and only if, its complementary function Φ^* satisfies the Δ_2-condition.*

Proposition 4. *Let Φ_c be a Musielak–Orlicz function that does not satisfy the Δ_2-condition and that $\Phi(t, b_{\Phi_c}(t)) = \infty$ for μ-a.e. $t \in T$. Then we can find a non-negative function $u \in \tilde{L}^{\Phi_c}$ such that $I_{\Phi_c^*}(\Phi'_{c+}(t, u(t))) = \infty$.*

Proof. Let $\{\lambda_n\}$, $\{u_n\}$ and $\{A_n\}$ be given as in Lemma 1. Select a subsequence $\{\lambda_{n_k}\} \subset \{\lambda_n\}$ for which the series $\sum_{k=1}^\infty (1 - \lambda_{n_k})$ converges, and $(1 - \lambda_{n_k}) + 2^{-n_k} < 1$ for all $k \geq 1$. Because $\lambda \to I_{\Phi_c}(\lambda u_{n_k}\chi_{A_{n_k}})$ is continuous for $\lambda \in [0, 1]$, we can find $\lambda'_k \in (\lambda_{n_k}, 1)$ such that $I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) = (1 - \lambda_{n_k}) + 2^{-n_k}$. Define $u = \sum_{k=1}^\infty \lambda'_k u_{n_k}\chi_{A_{n_k}}$. Then, we can write

$$I_{\Phi_c}(u) = \sum_{k=1}^\infty I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) = \sum_{k=1}^\infty [(1 - \lambda_{n_k}) + 2^{-n_k}] < \infty,$$

and

$$\begin{aligned}
\int_T u(t)\Phi'_{c+}(t, u(t))d\mu &= \sum_{k=1}^\infty \int_{A_{n_k}} \lambda'_k u_{n_k}(t)\Phi'_{c+}(t, \lambda'_k u_{n_k}(t))d\mu \\
&\geq \sum_{k=1}^\infty \lambda'_k \frac{1}{\lambda'_k - \lambda_{n_k}}[I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) - I_{\Phi_c}(\lambda_{n_k} u_{n_k}\chi_{A_{n_k}})] \\
&\geq \sum_{k=1}^\infty \frac{\lambda'_k}{1 - \lambda_{n_k}}[(1 - \lambda_{n_k}) + 2^{-n_k} - 2^{-n_k}] \\
&= \sum_{k=1}^\infty \lambda'_k = \infty.
\end{aligned}$$

Hence, it follows that

$$I_{\Phi_c^*}(\Phi'_{c+}(t, u(t))) = \int_T u(t)\Phi'_{c+}(t, u(t))d\mu - I_{\Phi_c}(u) = \infty,$$

which concludes the proof. ☐

The previous proposition makes it clear that we can find a $u \in \tilde{L}^{\Phi_c}$, but $\Phi'_{c+}(t, u(t)) \notin \tilde{L}^{\Phi^*_c}$. Let u be as in Proposition 4, clearly for $\lambda \in (0,1)$, $I_{\Phi_c}(\lambda u) < \infty$ and for $\lambda > 1$, $I_{\Phi_c}(\lambda u) = \infty$ ([35], Remark 3.12).

Proposition 5. *Let Φ_c be a Musielak–Orlicz function such that, Φ_c satisfies ∇_2-condition, does not satisfy Δ_2-condition and $\Phi_c(t, b_\Phi(t)) = \infty$. Then we can find $w \in \partial\mathcal{B}^\varphi_c$ such that*

$$\frac{\varphi'(c + w - \psi(w)u_0)}{\int_T u_0 \varphi'(c + w - \psi(w)u_0)d\mu} \notin L^{\Phi^*_c} \tag{22}$$

where $\tilde{L}^{\Phi^*_c}$ is the Musielak–Orlicz class of Φ^*_c, the conjugate of Φ_c.

Proof. Take $k_0 \geq 1$ and denote $B = T \setminus \bigcup_{k=k_0}^\infty A_{n_k}$, then we define $\tilde{u} = \sum_{k=k_0}^\infty \lambda'_k u_{n_k} \chi_{A_{n_k}}$. We can choose $\lambda' < 0$ such that

$$w = \lambda' u_0 \chi_B + \tilde{u}$$

satisfies $\int_T w\varphi'(c)d\mu = 0$. In other words, $w \in B^\varphi_c$. It is easy to see that $\int_T \varphi(c + \alpha w)d\mu < \infty$ for $\alpha \in (0,1)$ and $\int_T \varphi(c + \alpha w)d\mu = \infty$ for $\alpha > 1$, so $w \in \partial\mathcal{B}^\varphi_c$. The need to show Equation (22) remains. From Proposition 4 we have that

$$\int_T w\Phi'_{c+}(t, w(t))d\mu = \int_B \lambda' u_0(t)\Phi'_{c+}(t, \lambda' u_0)d\mu + \sum_{k=k_0}^\infty \int_{A_{n_k}} I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}})d\mu = \infty,$$

since

$$\int_B \lambda' u_0(t)\Phi'_{c+}(t, \lambda' u_0)d\mu \leq \lambda' \int_B \Phi_c((\lambda' + 1)u_0) - \Phi_c(\lambda' u_0)d\mu$$

$$\leq \lambda' \int_B \varphi(c + (\lambda' + 1)u_0) - \varphi(c + \lambda' u_0)d\mu$$

$$< \infty.$$

Thus,

$$I_{\Phi^*_c}\left(\Phi'_{c+}(t, w(t))\right) = \int_T w(t)\Phi'_{c+}(t, w(t))d\mu - I_{\Phi_c}(w) = \infty,$$

consequently $\varphi'(c + w) \notin \tilde{L}^{\Phi^*_c}$. Since $\Phi_c \in \nabla_2$, we have that $\Phi^*_c \in \Delta_2$ and therefore $\tilde{L}^{\Phi^*_c} = L^{\Phi^*_c}$. We conclude that $\varphi'(c + w) \notin L^{\Phi^*_c}$. Since $L^{\Phi^*_c}$ is a linear set, we have that Equation (22) occurs. ☐

As a consequence of Proposition 5, we have that it is possible to find $u \in \left\{\partial\mathcal{B}^\varphi_c\right\}_{<\infty}$ such that

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0)d\mu} \notin L^{\Phi^*_c},$$

and therefore the functional in Equation (16) does not belong to $\partial\psi(u)$.

We conclude in this section that, if the functional

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0)d\mu} \tag{23}$$

belongs to $L^{\Phi^*_c}$, then the functional belongs to $\partial\psi(u)$ for $u \in$ dom ψ.

In next section we finally prove that the set of functionals formed by Gâteaux gradient of the normalizing function ψ that belongs to $L^{\Phi^*_c}$ is convex, so we can guarantee that the generalized mixture arc is well defined.

3.2. Convexity of the Functionals Set

We already know that, for the generalized mixture arc in Equation (11) to be well defined, it is necessary that the set of functionals

$$\left\{ \frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0) d\mu}, \quad u \in \text{dom } \psi \right\} \bigcap L^{\Phi_c^*} \tag{24}$$

to be convex. From Proposition 2, the set in Equation (24) is contained in the range of $\partial\psi$, the set given by

$$\text{range } \partial\psi = \bigcup \left\{ \partial\psi(u) : u \in \text{dom } \psi \right\}. \tag{25}$$

Let ψ^* be the conjugate function of ψ. By the fact that ψ^* be a l.s.c. proper convex function, int dom ψ^* and dom ψ^* are convex sets and the range of $\partial\psi$ is the effective domain of $\partial\psi^*$, since $(\partial\psi)^{-1} = \partial\psi^*$. Thus

$$\text{int dom } \psi^* \subset D(\partial\psi^*) \subset \text{dom } \psi^* \tag{26}$$

is the same that

$$\text{int dom } \psi^* \subset \text{range } \partial\psi \subset \text{dom } \psi^*. \tag{27}$$

To prove that the set in Equation (24) is convex, we analyze the set in Equation (25) in three cases. Let u^*, v^* be elements in Equation (25) such that

Case 1.　　$u^*, v^* \in \text{int dom } \psi^*$, so by convexity of int dom ψ^*, for $\lambda \in (0,1)$, we have $\lambda u^* + (1 - \lambda)v^* \in \text{int dom } \psi^*$.

Case 2.　　If $u^* \in \text{int dom } \psi^*$ and $v^* \in D(\partial\psi^*) \setminus \text{int dom } \psi^*$, then $\lambda u^* + (1 - \lambda)v^* \in \text{int dom } \psi^*$, for $\lambda \in (0,1)$ ([41], Fact 2.1).

Case 3.　　Let u^*, v^* be elements in Equation (25) belonging to $D(\partial\psi^*) \setminus \text{int dom } \psi^*$.

We want to prove that, for $\lambda \in (0,1)$, $\lambda u^* + (1 - \lambda)v^*$ belongs to Equation (25). To solve this problem, we are going to prove that $D(\partial\psi^*) = \text{int dom } \psi^*$. Supposing φ a strictly convex function, then ψ is a strictly convex function. In next proposition, we show that $\partial\psi^*(u^*)$ is a unitary set.

Proposition 6. *Let ψ be a strictly convex function, then $\partial\psi^*(u^*)$ is a unitary set, where $u^* \in \partial\psi(u)$, with $u \in D(\partial\psi)$.*

Proof. Assuming that ψ is a strictly convex function we have that for $\lambda \in (0,1)$ and $\forall u_1 \neq u_2 \in \text{dom } \psi$

$$\psi(\lambda u_1 + (1 - \lambda)u_2) < \lambda\psi(u_1) + (1 - \lambda)\psi(u_2). \tag{28}$$

Supposing that $\partial\psi^*(u^*)$ is not a unitary set , i.e., $\partial\psi^*(u^*) = \{u_1, u_2, \ldots\}$, where $u_i^* \in D(\partial\psi)$, $i = 1, 2, \ldots$. Taking $u_1, u_2 \in \partial\psi^*(u^*)$. By Young's Inequality (10)

$$(\lambda u_1 + (1 - \lambda)u_2, u^*) \leq \psi(\lambda u_1 + (1 - \lambda)u_2) + \psi^*(u^*), \tag{29}$$

where $\lambda \in (0,1)$ and as a consequence of $u_1, u_2 \in \partial\psi^*(u^*)$ we have

$$\psi(u_1) + \psi^*(u^*) = (u_1, u^*), \tag{30}$$

and

$$\psi(u_2) + \psi^*(u^*) = (u_1, u^*). \tag{31}$$

Taking the product of Equation (30) by λ, the product of Equation (31) by $(1-\lambda)$ and adding the two obtained equations, we have

$$\lambda\psi(u_1) + (1-\lambda)\psi(u_2) + \psi^*(u^*) = (\lambda u_1 + (1-\lambda)u_2, u^*). \tag{32}$$

From Equations (29) and (32), we obtain

$$\lambda\psi(u_1) + (1-\lambda)\psi(u_2) \leq \psi(\lambda u_1 + (1-\lambda)u_2), \tag{33}$$

which is a contradiction by Equation (28). This implies that $\partial\psi^*(u^*)$ is a unitary set and this completes the proof. $\square$

Thus, the set $\partial\psi^*(u^*)$ is unitary, then $\partial\psi^*$ is locally bounded at $u^* \in D(\partial\psi^*)$ and, therefore, by Fact 1, we conclude that $u^* \in \text{int dom } \psi^*$ which implies that $D(\partial\psi^*) \subset \text{int dom } \psi^*$, by Equation (26), we have that

$$\text{range } \partial\psi = D(\partial\psi^*) = \text{int dom } \psi^*. \tag{34}$$

Therefore, by Fact 2, there exists no functional u^* in Equation (25) such that $u^* \in D(\partial\psi^*) \setminus \text{int dom } \psi^*$. Thus Equation (25) is a convex set and, as a consequence, the generalized mixture arc is well defined, since the set in Equation (24) is a convex set. Indeed, let u, v be functions in dom ψ such that

$$u^* = \frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \tag{35}$$

and

$$v^* = \frac{\varphi'(c + v - \psi(v)u_0)}{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu} \tag{36}$$

belong to Equation (24). Clearly,

$$\int_T u_0 u^* d\mu = \frac{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} = 1$$

and

$$\int_T u_0 v^* d\mu = \frac{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu}{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu} = 1.$$

We note that, the functionals in Equation (24) are the only elements in Equation (25) that satisfy $\int_T u_0 u^* d\mu = 1$. For $\lambda \in (0, 1)$ we have

$$\int_T u_0((1-\lambda)u^* + \lambda v^*)d\mu = 1,$$

then there exist functions $w_\lambda \in \text{dom } \psi$ such that

$$\frac{\varphi'(c + w_\lambda - \psi(w_\lambda)u_0)d\mu}{\int_T u_0\varphi'(c + w_\lambda - \psi(w_\lambda)u_0)d\mu} = (1-\lambda)u^* + \lambda v^*, \text{ for each } \lambda \in (0,1).$$

Thus, the set in Equation (24) is a convex set.

In this section, we proved that the generalized mixture arc is well defined for a deformed exponential φ strictly convex. In the next section, we discuss generalized open exponential arcs and generalized open mixture arcs.

4. Generalized Arcs

The concept of arc-connected probability distributions was defined by de Souza et al. [36] defined the concept of arc-connected probability distributions. Fixing any deformed exponential φ

we say that two probability distributions $p, q \in \mathcal{P}_\mu$ are φ-connected if, for each $\alpha \in [0, 1]$, there exists $k(\alpha) := k(\alpha; p, q) \in \mathbb{R}$ such that

$$\int_T \varphi(\alpha \varphi^{-1}(p) + (1 - \alpha) + k(\alpha)u_0)d\mu = 1.$$

In [31], necessary and sufficient conditions for any probability distributions being φ-connected were provided. In this section, we discuss the concept of two probability distributions $p, q \in \mathcal{P}_\mu$ are φ-connected by open arcs. We generalize open exponential arcs and open mixture arcs, defined in [22] and studied later in [23].

4.1. Generalized Open Exponential Arcs

Let us define the generalized open arcs and prove some of its properties.

Definition 1. *For a fixed deformed exponential φ, we say that p and q in $\mathcal{P}_\mu$ are φ-connected by an open arc if there exists an open interval $I \supset [0, 1]$ and a constant $k(\alpha)$ such that*

$$p(\alpha) = \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q) + k(\alpha)u_0)$$

belongs to $\mathcal{P}_\mu$ for every $t \in I$, where $k(\alpha)$ depends of t, p and q.

In the following proposition, we give an equivalent definition of φ-connection by open arc.

Proposition 7. *$p, q \in \mathcal{P}_\mu$ are φ-connected by an open arc if and only if there exist an open interval $I \supset [0, 1]$ and a random variable $v \in L_c^\varphi$, such that $p(\alpha) \propto \varphi(c + \alpha v)$ belongs to $\mathcal{P}_\mu$, for all $t \in I$ and $p(0) = p$ and $p(1) = q$.*

Proof. Let us assume that p, q are φ-connected, i.e., $\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q))d\mu < \infty$, for all $\alpha \in I$. Since

$$\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q))d\mu = \int_T \varphi(\alpha[\varphi^{-1}(q) - \varphi^{-1}(p)] + \varphi^{-1}(p))d\mu$$

$$= \int_T \varphi(c + \alpha v)d\mu,$$

where $v = \varphi^{-1}(q) - \varphi^{-1}(p)$ and $\varphi(c) = p$, then $v \in L_c^\varphi$. Moreover, $p(\alpha) \propto \varphi(c + \alpha v)$ belongs to $\mathcal{P}_\mu$, for every $\alpha \in I$ and $p(0) = \varphi(c) = p$ and $p(1) = q$. The converse follows immediately. Suppose that $q = p(1)$, we have $\varphi(c + v) = q$, then $v = \varphi^{-1}(q) - \varphi^{-1}(p)$, with $\varphi(c) = p = p(0)$. $\square$

Because of $v \in L_c^\varphi$ the need to define the open arcs arises. As a consequence of Proposition 7, we have that if $p, q \in \mathcal{P}_\mu$ are φ- connected by an open arc, then the random variable $v \in \mathcal{K}_c^\varphi$, since $\int_T \varphi(c + \alpha v)d\mu < \infty$ for all $\alpha \in (-\varepsilon, 1 + \varepsilon)$. With this, we can prove the following results.

Corollary 1. *Let $p, q \in \mathcal{P}_\mu$, where $p = \varphi(c)$. We have that $q \in \mathcal{F}_c^\varphi$ if and only if, p and q are φ-connected by an open arc.*

Proof. Supposing $q \in \mathcal{F}_c^\varphi$, then $q = \varphi(c + v - \psi(v)u_0)$ where $v \in \mathcal{B}_c^\varphi$. Thus, we have $\int_T \varphi(c + \alpha v)d\mu < \infty$ for all $\alpha \in (-\varepsilon, 1 + \varepsilon)$, we deduce that $p(\alpha) \propto \varphi(c + \alpha v)$ is an open arc containing p and q. Conversely, supposing that p and q are φ-connected by an open arc, by Proposition 7, there exist an open interval $I \supset [0, 1]$ and $v \in \mathcal{K}_c^\varphi$ such that $p(\alpha) \propto \varphi(c + \alpha v)$ belongs to $\mathcal{P}_\mu$ with $q = p(1)$. If $v \in B_c^\varphi$, then $q = \varphi(c + v) \in \mathcal{F}_c^\varphi$ and the proof is over. Otherwise, let w be such that

$$w = v - \frac{\int_T v \varphi'(c)d\mu}{\int_T u_0 \varphi'(c)d\mu}u_0,$$

thus $\int_T w\varphi'(c)d\mu = 0$ and $w \in \mathcal{B}_c^\varphi$. Hence, we have $q = \varphi(c + v) = \varphi(c + w)$ and $q \in \mathcal{F}_c^\varphi$. □

With this, we prove that, for $\varphi(c) = p$, the φ-family of probability distributions $\mathcal{F}_c^\varphi$ is the set of all $q \in \mathcal{P}_\mu$ such that q is φ-connected by an open arc to p.

Corollary 2. *Let $p = \varphi(c)$ and $q = \varphi(\tilde{c})$ be such that $p, q \in \mathcal{P}_\mu$ are φ-connected by an open arc. Then the spaces L_c^φ and $L_{\tilde{c}}^\varphi$ are equal as sets.*

Proof. It follows from Corollary 1 that p and q are in the same φ-family, then $\tilde{c} = c + u - \psi(u)u_0$ and by Vigelis and Cavalcante [28], Lemma 5, it follows the result. □

Now, we show that the connection by generalized open exponential arcs is an equivalence relation.

Proposition 8. *The relation in Definition 1 is an equivalence relation.*

Proof. Reflexive and symmetry properties follow from the definition and now, we prove transitivity. Consider $p, q, r \in \mathcal{P}_\mu$

$$p(t) \propto \varphi(c + tu), \qquad r(t) \propto \varphi(c + tv), \qquad t \in (-\varepsilon, 1 + \varepsilon),$$

with $p(0) = \varphi(c) = p$, $p(1) = \varphi(c + u) = q$, $r(0) = \varphi(c) = p$, $r(1) = \varphi(c + v) = r$ with $u, v \in L_c^\varphi$. We have that p is φ-connected to q and r, respectively. We need to prove that q and r are also φ-connected. Consider

$$q(t) \propto \varphi(c + (1 - t)u + tv) \propto \varphi(c + u + t(v - u))$$

is defined with $c + u = \tilde{c}$, $p(t) \propto \varphi(\tilde{c} + t(v - u))$, $v - u \in L_{\tilde{c}}^\varphi$, such that $q(0) = \varphi(\tilde{c}) = \varphi(c + u) = q$, $q(1) = \varphi(\tilde{c} + (v - u)) = \varphi(c + v) = r$. Therefore, q and r are φ-connected. □

We know from Corollary 1 that the φ-family $\mathcal{F}_c^\varphi$ coincides with the set of all $q \in \mathcal{P}_\mu$ which are φ-connected to p by an open arc. We want now to prove that the φ-family $\mathcal{F}_c^\varphi$ is convex for some deformed exponential φ.

Lemma 4. *Let φ be a fixed deformed exponential. Assuming that $(\varphi^{-1})''(x)$ is continuous and*

$$\frac{\alpha\varphi''(\alpha\varphi^{-1}(x) + k)}{\varphi''(\alpha\varphi^{-1}(x) + k)} \geq \frac{\varphi''(\varphi^{-1}(x))}{\varphi''(\varphi^{-1}(x))}, \tag{37}$$

then $F(x) = \varphi(\alpha\varphi^{-1}(x) + k)$, for some fixed $\alpha > 1$ and $k \in \mathbb{R}$ is a convex function.

Proof. We know that, if $F''(x) \geq 0 \; \forall \alpha > 1$ and $\forall x$, then $F(x)$ is a convex function. We have

$$F''(x) = \frac{\alpha^2\varphi''(\alpha\varphi^{-1}(x) + k)\varphi'(\varphi^{-1}(x)) - \alpha\varphi'(\alpha\varphi^{-1}(x) + k)\varphi''(\varphi^{-1}(x))}{[\varphi'(\alpha\varphi^{-1}(x))]^3}$$

by the fact that φ is an increasing function $[\varphi'(\alpha\varphi^{-1}(x))]^3 > 0$. Hence, we have $F''(x) \geq 0$ if and only if

$$\alpha^2\varphi''(\alpha\varphi^{-1}(x) + k)\varphi'(\varphi^{-1}(x)) - \alpha\varphi'(\alpha\varphi^{-1}(x) + k)\varphi''(\varphi^{-1}(x)) \geq 0,$$

which follows from Equation (37). □

Proposition 9. *Let $p \in \mathcal{P}_\mu$ such that $\varphi(c) = p$. Assuming that $(\varphi^{-1})''(x)$ is continuous and*

$$\frac{\alpha\varphi''(\alpha\varphi^{-1}(x) + k)}{\varphi''(\alpha\varphi^{-1}(x) + k)} \geq \frac{\varphi''(\varphi^{-1}(x))}{\varphi''(\varphi^{-1}(x))} \tag{38}$$

for some fixed $\alpha > 1$ and $k \in \mathbb{R}$. Then, the φ-family of probability $\mathcal{F}_c^{\varphi}$ is convex.

Proof. Note that, for any $\varphi(\tilde{c}) = r \in \mathcal{F}_c^{\varphi}$, $\mathcal{F}_c^{\varphi} = \mathcal{F}_{\tilde{c}}^{\varphi}$. Suppose $q \in \mathcal{F}_c^{\varphi}$, and consider $p(\lambda) = \lambda p + (1 - \lambda)q$ for any $\lambda \in [0, 1]$. We show that $p(\lambda) \in \mathcal{F}_c^{\varphi}$, $\forall \lambda \in [0, 1]$ by proving that $\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda))d\mu < \infty$ for $\alpha \in (-\varepsilon, 1 + \varepsilon)$. In the others words, we will show that $p(\lambda)$ and p are φ-connected for all $\lambda \in [0, 1]$.

For $\alpha \in (0, 1)$, due the convexity of φ, we have

$$\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu \leq \int_T (1 - \alpha)\varphi(\varphi^{-1}(p)) + \alpha\varphi(\varphi^{-1}(p(\lambda)))d\mu$$

$$= \int_T (1 - \alpha)p + \alpha p(\lambda)d\mu$$

$$= (1 - \alpha)\int_T p d\mu + \alpha \int_T p(\lambda)d\mu$$

$$= 1.$$

If $\alpha \in (-\varepsilon, 0)$, according the convexity of $\alpha\varphi^{-1}(x)$ and $\varphi(x)$, we have

$$\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu$$

$$\leq \int_T \varphi(\lambda\alpha\varphi^{-1}(p) + (1 - \lambda)\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p))d\mu$$

$$= \int_T \varphi(\lambda[\alpha\varphi^{-1}(p) + (1 - \alpha)\varphi^{-1}(p)] + (1 - \lambda)[\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p)])d\mu$$

$$\leq \int_T \lambda\varphi(\alpha\varphi^{-1}(p) + (1 - \alpha)\varphi^{-1}(p)) + (1 - \lambda)\varphi(\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p))d\mu$$

$$= \lambda \int_T \varphi(\varphi^{-1}(p))d\mu + (1 - \lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p))d\mu$$

$$= \lambda + (1 - \lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p))d\mu,$$

since $q \in \mathcal{F}_c^{\varphi}$, we have by Corollary 1 that q and p are φ-connected. Hence,

$$\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu < \infty$$

so $p(\lambda)$ and p are φ-connected by an open arc, for all $\alpha \in (-\varepsilon, 0)$.

Now, if $\alpha \in (1, 1 + \varepsilon)$, the Lemma 4, $F(x) = \varphi(\alpha\varphi^{-1}(x) + k)$ is a convex function, so

$$\varphi(\alpha\varphi^{-1}(\lambda x + (1 - \lambda)y) + k) \leq \lambda\varphi(\alpha\varphi^{-1}(x) + k) + (1 - \lambda)\varphi(\alpha\varphi^{-1}(y) + k), \tag{39}$$

where $\lambda \in [0, 1]$ and k a constant. Taking $k = (1 - \alpha)\varphi^{-1}(p)$, we have

$$\int_T \varphi(\alpha\varphi^{-1}(p(\lambda)) + (1 - \alpha)\varphi^{-1}(p))d\mu \leq \lambda \int_T \varphi(\alpha\varphi^{-1}(p) + (1 - \alpha)\varphi^{-1}(p)d\mu$$

$$+ (1 - \lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p)d\mu$$

$$= \lambda + (1 - \lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1 - \alpha)\varphi^{-1}(p))d\mu$$

$$< \infty,$$

since $q \in \mathcal{F}_c^{\varphi}$ and, therefore, p and q are φ-connected by an open arc. ☐

4.2. Generalized Open Mixture Arcs

In Section 3.2, we proved that the generalized mixture arc given by

$$p(\alpha) = F^{-1}\left((1-\alpha)F(p) + \alpha F(q)\right), \tag{40}$$

is well defined for $\alpha \in [0,1]$. In this section, our goal is twofold: firstly, to ensure that the open arc is also well defined; and, secondly, to provide some properties of these arcs. For such objectives, we use Equation (34), which establishes that $D(\partial \psi^*) = \text{range } \partial \psi$ is an open set, so we can extend the convex combination in Equation (40) between $F(p)$ and $F(q)$ beyond these extreme points while maintaining positivity of $(1-\alpha)F(p) + \alpha F(q)$. Indeed, by the fact $D(\partial \psi^*) = \text{range } \partial \psi$ is an open set, so there exists $\varepsilon_1 > 0$ such that $B\left(F(p), \varepsilon_1\right)$ is the open ball of radius ε_1 centered at $F(p)$ with $B\left(F(p), \varepsilon_1\right) \subset D(\partial \psi^*)$. Similarly, there exists $\varepsilon_2 > 0$ such that $B(F(q), \varepsilon_2) \subset D(\partial \psi^*)$. Taking $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ we guarantee that the combination $(1-\alpha)F(p) + \alpha F(q)$ in (40) can be extended to $\alpha \in I = (-\varepsilon, 1+\varepsilon) \supset [0,1]$.

Definition 2. *For a fixed deformed exponential φ, we say that p and q in $\mathcal{P}_\mu$ are φ-connected by an open mixture arc if there exists an open interval $I \supset [0,1]$ such that*

$$p(\alpha) = F^{-1}\left((1-\alpha)F(p) + \alpha F(q)\right) \tag{41}$$

belongs to $\mathcal{P}_\mu$ for every $\alpha \in I$, where $F(p) = \dfrac{\varphi'(\varphi^{-1}(p))}{\int_T u_0 \varphi'(\varphi^{-1}(p))d\mu}$.

In [22], it was shown that densities connected by open mixture arcs have bounded away from zero ratios. Santacroce et al. [23] showed the converse implication, providing a characterization of open mixture models. Here, one can see that the fundamental role for being connected by open mixture arcs is given by ratios $\dfrac{F(p)}{F(q)}$ which have to be bounded. The functional $F(p)$ in the definition of generalized open mixture arc satisfies $F(p) > 0$. Thus the combination $(1-\alpha)F(p) + \alpha F(q)$ in (41) has to satisfy the same property, that is, $(1-\alpha)F(p) + \alpha F(q) > 0$. Assume that p and q are φ-connected by an open mixture arc given according to (41) belong to $\mathcal{P}_\mu$ for all $\alpha \in (-\varepsilon_1, 1+\varepsilon_2) \supset [-\varepsilon, 1+\varepsilon]$ with $\varepsilon > 0$. Since $p(-\varepsilon)$ and $p(1+\varepsilon) \in \mathcal{P}_\mu$, then

$$F(p(-\varepsilon)) = (1+\varepsilon)F(p) + (-\varepsilon)F(q) > 0,$$

which implies that

$$\frac{F(p)}{F(q)} > \frac{\varepsilon}{1+\varepsilon} \tag{42}$$

and

$$F(p(1+\varepsilon)) = (-\varepsilon)F(p) + (1+\varepsilon)F(q) > 0,$$

which give to us

$$\frac{F(p)}{F(q)} < \frac{1+\varepsilon}{\varepsilon}. \tag{43}$$

Combining inequalities (42) and (43), we have

$$\frac{\varepsilon}{1+\varepsilon} < \frac{F(p)}{F(q)} < \frac{1+\varepsilon}{\varepsilon}. \tag{44}$$

Conversely, if we have Equation (44), then $(1-\alpha)F(p) + \alpha F(q) > 0$ and Equation (41) belongs to $\mathcal{P}_\mu$. Thus, we have that p and q in $\mathcal{P}_\mu$ are φ-connected by an open mixture arc if and only if the ratio

$\frac{F(p)}{F(q)}$ is bounded. By the fact that range $\partial \psi$ is an open set, there exists an interval $I \supset [0,1]$ such that $(1 - \alpha)F(p) + \alpha F(q)$ belongs to range $\partial \psi$ and we have

$$\int_T u_0 \left[(1 - \alpha)F(p) + \alpha F(q)\right] d\mu = 1, \text{ for all } \alpha \in I \supset [0,1],$$

for all $\alpha \in I$. Then, there exist functions $w_\alpha \in \text{dom } \psi$ such that

$$(1 - \alpha)F(p) + \alpha F(q) = \frac{\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)}{\int_T u_0 \varphi'(c + w_\alpha - \psi(w_\alpha)u_0)d\mu},$$

with

$$p(\alpha) = F^{-1}\left(\frac{\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)}{\int_T u_0 \varphi'(c + w_\alpha - \psi(w_\alpha)u_0)d\mu}\right), \text{ for all } \alpha \in I \supset [0,1],$$

that is, the convex combination in Equation (41) is also a functional of the type in Equation (12) for all $\alpha \in I$. Then, the open mixture arc is well-defined. Another property of this connection by generalized open mixture arc is that it is an equivalence relation.

Proposition 10. *The relation in Definition 2 is an equivalence relation.*

Proof. Reflexity and symmetry properties follow from definition. As for the transitivity, consider p, q and $r \in \mathcal{P}_\mu$ such that $p(\lambda) = F^{-1}\left((1 - \lambda)F(p) + \lambda F(q)\right) \in \mathcal{P}_\mu$ and $q(\beta) = F^{-1}\left((1 - \beta)F(q) + \beta F(r)\right) \in \mathcal{P}_\mu$ with $\lambda, \beta \in [-\varepsilon, 1 + \varepsilon]$ for some $\varepsilon > 0$. We can take $p(-\varepsilon) = F^{-1}\left((1 + \varepsilon)F(p) + (-\varepsilon)F(q)\right)$ and $q(-\varepsilon) = F^{-1}\left((1 + \varepsilon)F(q) + (-\varepsilon)F(r)\right)$, and define a probability distribution

$$\begin{aligned} p_1 &= F^{-1}\left(\left(1 - \frac{\varepsilon}{1 + 2\varepsilon}\right)F(p(-\varepsilon)) + \frac{\varepsilon}{1 + 2\varepsilon}F(q(-\varepsilon))\right) \\ &= F^{-1}\left(\frac{(1 + \varepsilon)^2}{1 + 2\varepsilon}F(p) - \frac{\varepsilon^2}{1 + 2\varepsilon}F(r)\right). \end{aligned}$$

If we have $p(1 + \varepsilon) = F^{-1}\left((-\varepsilon)F(p) + (1 + \varepsilon)F(q)\right)$ and $q(1 + \varepsilon) = F^{-1}\left((-\varepsilon)F(q) + 1 + \varepsilon)F(r)\right)$, we may define a probability distribution as

$$\begin{aligned} p_2 &= F^{-1}\left(\frac{-\varepsilon}{-1 - 2\varepsilon}F(p(1 + \varepsilon)) + \frac{-1 - \varepsilon}{-1 - 2\varepsilon}F(q(1 + \varepsilon))\right) \\ &= F^{-1}\left(\frac{-\varepsilon^2}{1 + 2\varepsilon}F(p) + \frac{(1 + \varepsilon)^2}{1 + 2\varepsilon}F(r)\right). \end{aligned}$$

The generalized open mixture arc, $r(\alpha) = F^{-1}\left((1 - \alpha)F(p_1) + \alpha F(p_2)\right)$, $\alpha \in (0, 1)$, connects $r\left(\frac{(1 + \varepsilon)^2}{2\varepsilon^2 + 2\varepsilon + 1}\right) = p$ and $r\left(\frac{\varepsilon^2}{2\varepsilon^2 + 2\varepsilon + 1}\right) = r$. $\square$

5. Conclusions

In this work, we have generalized open exponential arc and open mixture arc for probability distributions. Moreover, we ensure that the generalization of open mixture arc is well-defined for deformed exponential strictly convex. From two φ-connected probability distributions p_1 and p_2, we can define the generalized parallel transport $\tau^{(1)}_{p_1, p_2}$ between the tangent spaces $T_{p_1}\mathcal{P}_\mu$ and $T_{p_2}\mathcal{P}_\mu$ given by

$$u \mapsto u - \frac{\int_T u \varphi'(\varphi^{-1}(p_2))d\mu}{\int_T u_0 \varphi'(\varphi^{-1}(p_2))d\mu} u_0,$$

where $T_p \mathcal{P}_\mu \simeq B_c^\varphi$ with $p = \varphi(c)$. A next step is to find a generalized parallel transport $\tau_{p_1,p_2}^{(-1)}$ that is dual to $\tau_{p_1,p_2}^{(1)}$. Another goal is to investigate if the generalized Rényi divergence $\mathcal{D}_\varphi^\alpha(\cdot \| \cdot)$ defined in [36] from two probability distributions φ-connected, can be related to the statistical divergence associated with $(\tau_{p_1,p_2}^{(-1)}, \tau_{p_1,p_2}^{(1)}, \langle \cdot, \cdot \rangle)$.

Author Contributions: All authors contributed equally to the design of the research. The research was carried out by all authors. Rui F. Vigelis and Charles C. Cavalcante gave the central idea of the paper and managed the organization of it. Luiza H.F. de Andrade wrote the paper. All the authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Amari, S.-I. Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* **2001**, *47*, 1701–1711.

2.　Calin, O.; Udrişte, C. *Geometric Modeling in Probability and Statistics*; Springer: Cham, Switzerland, 2014.

3.　Amari, S.-I. Information geometry and its applications. In *Applied Mathematical Sciences*; Springer: Tokyo, Japan, 2016; Volume 194.

4.　Amari, S.-I. Differential Geometry of Curved Exponential Families-Curvatures and Information Loss. *Ann. Stat.* **1982**, *10*, 357–385.

5.　Amari, S.-I. *Differential-Geometrical Methods in Statistics*; Lecture Notes in Statistics; Springer: New York, NY, USA, 1985; Volume 28.

6.　Amari, S.-I.; Nagaoka, H. Methods of information geometry. In *Translations of Mathematical Monographs*; American Mathematical Society, Providence, RI; Translated from the 1993 Japanese original by Daishi Harada; Oxford University Press: Oxford, UK, 2000; Volume 191.

7.　Amari, S-I.; Cichocki, A. Information geometry of divergence functions. *Bull. Pol. Acad. Sci. Tech. Sci.* **2010**, *58*, 183–195.

8.　Amari, S.-I. α-Divergence Is Unique, Belonging to Both f-Divergence and Bregman Divergence Classes. *IEEE Trans. Inf. Theory* **2009**, *55*, 4925–4931.

9.　Zhang, J. Divergence Function, Duality, and Convex Analysis. *Neural Comput.* **2004**, *16*, 159–195.

10.　Nielsen, F.; Nock, R. On w-mixtures: Finite convex combinations of prescribed component distributions. *arXiv* **2017**, arXiv:1708.00568v1.

11.　Amari, S.-I.; Ohara, A.; Matsuzoe, H. Geometry of deformed exponential families: Invariant, dually-flat and conformal geometries. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 4308–4319.

12.　Harsha, K.V.; Moosath, K.S.S. Dually flat geometries of the deformed exponential family. *Phys. A Stat. Mech. Appl.* **2015**, *433*, 136–147.

13.　Matsuzoe, H. Hessian structures on deformed exponential families and their conformal structures. *Differ. Geom. Appl.* **2014**, *35*, 323–333.

14.　Matsuzoe, H.; Wada, T. Deformed Algebras and Generalizations of Independence on Deformed Exponential Families. *Entropy* **2015**, *17*, 5729–5751.

15.　Giné, E.; Nickl, R. *Mathematical Foundations of Infinite-Dimensional Statistical Models*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press: Cambridge, UK, 2015.

16.　Townsend, J.; Solomon, B.; Smith, J. The perfect gestalt: Infinite dimensional Riemannian face spaces and other aspects of face perception. In *Computacional, Geometric and Process Perspectives on Facial Cognition: Contexs and Challenges*; Wenger, M.J., Townsend, J.T., Eds.; Society for Mathematical Psychology: Washington, DC, USA, 2001; pp. 39–82.

17.　Trivellato, B. Deformed exponentials and applications to finance. *Entropy* **2013**, *15*, 3471–3489.

18. Pistone, G.; Sempi, C. An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. *Ann. Stat.* **1995**, *23*, 1543–1561.

19. Pistone, G.; Rogantin, M.P. The exponential statistical manifold: Mean parameters, orthogonality and space transformations. *Bernoulli* **1999**, *5*, 721–760.

20. Gibilisco, P.; Pistone, G. Connections on Non-Parametric Statistical Manifolds by Orlicz Space Geometry. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1998**, *1*, 325–347.

21. Grasselli, M.R. Dual connections in nonparametric classical information geometry. *Ann. Inst. Statist. Math.* **2010**, *62*, 873–896.

22. Cena, A.; Pistone, G. Exponential statistical manifold. *Ann. Inst. Statist. Math.* **2007**, *59*, 27–56.

23. Santacroce, M.; Siri, P.; Trivellato, B. New results on mixture and exponential models by Orlicz spaces. *Bernoulli* **2016**, *22*, 1431–1447.

24. Santacroce, M.; Siri, P.; Trivellato, B. On Mixture and Exponential Connection by Open Arcs. In *Geometric Science of Information*; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 577–584.

25. Pistone, G. Examples of the application of nonparametric information geometry to statistical physics. *Entropy* **2013**, *15*, 4042–4065.

26. Pistone, G. kappa-exponential models from the geometrical viewpoint. *Eur. Phys. J. B* **2009**, *70*, 29–37.

27. Kaniadakis, G. Non-linear kinetics underlying generalized statistics. *Phys. A Stat. Mech. Appl.* **2001**, *296*, 405–425.

28. Vigelis, R.F.; Cavalcante, C.C. On ϕ-families of probability distributions. *J. Theoret. Probab.* **2013**, *26*, 870–884.

29. Pistone, G. Nonparametric information geometry. In *Geometric Science of Information*; Lecture Notes in Comput. Sci.; Springer: Heidelberg, Germany, 2013; Volume 8085, pp. 5–36.

30. Naudts, J. *Generalised Thermostatistics*; Springer-London, Ltd.: London, UK, 2011.

31. Vigelis, R.F.; de Andrade, L.H.F.; Cavalcante, C.C. On the Existence of Paths Connecting Probability Distributions. In Proceedings of the Geometric Science of Information: Third International Conference, GSI 2017, Paris, France, 7–9 November 2017; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Cham, Switzerland , 2017; pp. 801–808.

32. Musielak, J. *Orlicz Spaces and Modular Spaces*; Lecture Notes in Mathematics, Springer: Berlin, Germany, 1983; Volume 1034.

33. Vigelis, R.F.; Cavalcante, C.C. The Δ_2-condition and ϕ-families of probability distributions. In *Geometric Science of Information*; Lecture Notes in Comput. Sci.; Springer: Heidelberg, Germany, 2013; Volume 8085, pp. 729–736.

34. Hudzik, H.; Zbaszyniak, Z. Smoothness in Musielak-Orlicz spaces equipped with the Orlicz norm. *Collect. Math.* **1997**, *48*, 543–561.

35. Vigelis, R.F.; Cavalcante, C.C. Smoothness of the Orlicz norm in Musielak-Orlicz function spaces. *Math. Nachr.* **2014**, *287*, 1025–1041.

36. de Souza, D.C.; Vigelis, R.F.; Cavalcante, C.C. Geometry Induced by a Generalization of Rényi Divergence. *Entropy* **2016**, *18*, 407.

37. De Andrade, L.H.F.; Vigelis, R.F.; Vieira, F.L.J.; Cavalcante, C.C. Normalization and ϕ -function: Definition and Consequences. In Proceedings of the Geometric Science of Information: Third International Conference, GSI 2017, Paris, France, 7–9 November 2017; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 231–238.

38. Asplund, E.; Rockafellar, R.T. Gradients of convex functions. *Trans. Am. Math. Soc.* **1969**, *139*, 443–467.

39. Barbu, V.; Precupanu, T. *Convexity and Optimization in Banach Spaces*, 4th ed.; Springer Monographs in Mathematics; Springer: Dordrecht, The Netherlands, 2012.

40. Brøndsted, A.; Rockafellar, R.T. On the subdifferentiability of convex functions. *Proc. Am. Math. Soc.* **1965**, *16*, 605–611.

41. Bauschke, H.H.; Borwein, J.M.; Combettes, P.L. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. *Commun. Contemp. Math.* **2001**, *3*, 615–647.

42. Borwein, J.M.; Vanderwerff, J.D. Convex functions: Constructions, characterizations and counterexamples. In *Encyclopedia of Mathematics and Its Applications*; Cambridge University Press: Cambridge, UK, 2010; Volume 109.

43. Vigelis, R.F. On Musielak-Orlicz Spaces and Applications to Information Geometry. Ph.D. Thesis, Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza, Brazil, 2011.