Information Geometry of κ-Exponential Families: Dually-Flat, Hessian and Legendre Structures

Antonio M. Scarfone [1,*] ![ORCID], **Hiroshi Matsuzoe** [2] ![ORCID] and **Tatsuaki Wada** [3] ![ORCID]

1   Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche (ISC-CNR), c/o Politecnico di Torino, 10129 Torino, Italy
2   Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan; matsuzoe@nitech.ac.jp
3   Region of Electrical and Electronic Systems Engineering, Ibaraki University, Nakanarusawa-cho, Hitachi 316-8511, Japan; tatsuaki.wada.to@vc.ibaraki.ac.jp
*   Correspondence: antoniomaria.scarfone@cnr.it; Tel.: +39-011-090-7339

![check for updates]

Abstract: In this paper, we present a review of recent developments on the κ-deformed statistical mechanics in the framework of the information geometry. Three different geometric structures are introduced in the κ-formalism which are obtained starting from three, not equivalent, divergence functions, corresponding to the κ-deformed version of Kullback–Leibler, "Kerridge" and Brègman divergences. The first statistical manifold derived from the κ-Kullback–Leibler divergence form an invariant geometry with a positive curvature that vanishes in the $\kappa \to 0$ limit. The other two statistical manifolds are related to each other by means of a scaling transform and are both dually-flat. They have a dualistic Hessian structure endowed by a deformed Fisher metric and an affine connection that are consistent with a statistical scalar product based on the κ-escort expectation. These flat geometries admit dual potentials corresponding to the thermodynamic Massieu and entropy functions that induce a Legendre structure of κ-thermodynamics in the picture of the information geometry.

Keywords: κ-generalized statistical mechanics; information geometry; dually-flat geometry; Hessian geometry; Legendre structure; divergence functions

PACS: 02.40.-k; 89.70.Cf; 05.90.+m

1. Introduction

In the study of complex systems, long tailed probability distributions are often discussed. Anomalous statistical behavior is largely observed in physical systems plagued by long-range interactions or long-time memory effects as well as in non-physical systems, such as biological, social and economic ones. In such systems, family of probability distributions characterized by deformed exponentials with asymptotic power-law tails play a relevant role.

To deal with these new phenomenologies, in recent decades, a generalized version of statistical mechanics has been formulated where power-law distributions are expressed in terms of a generalized exponential function which maximizes suitable entropic forms.

In recent years, there is a growing interest in studying statistical physics, including its generalized versions, by means of information geometry [1,2]. In this framework, the methods of the differential geometry are applied to study the properties of statistical manifolds constructed starting from a given parameterized probability distribution function.

The usefulness of the geometric approach in the study of the foundations of statistical mechanics and thermodynamics has been clear since from the early works of Gibbs and Charathéodory [3,4].

On the thermodynamical ground, after the pioneering works of Ruppeiner (see [5] for a review), which introduces a Riemannian metric in the thermodynamical parameter space starting from potentials like entropy and free energy, a rich literature has been developed to study the implications that geometry might have on the thermostatistics and in particular about its contact structure [6,7], the relationships between physical observable [8,9] and the thermodynamics transformations in a physical system [10,11].

Otherwise, the information geometry gives priority to the probabilities distributions. In this framework, one investigates the geometric structures of the statistical manifold generated by parametric families of distributions, by introducing a Riemann metric in the probability distribution space [12], identified with the Fisher metric [13]. After that, many efforts have been done to better understand the role of geometry in statistics, like, to cite a few, Čencov [14], that introduced the affine connection in a statistical manifold, Csizsár [15], that introduced the concept of f-divergence, Efron [16], that studied the role of curvature in a statistical model, Eguchi [17] that relates the affine connection to a divergence function or relative entropy and Amari [18–20], that further developed the differential geometry of statistical models by elucidating its dualistic nature and introducing the Pythagorean theorem and the projective theorem in the framework of information geometry.

More recently, Amari and co-workers showed the existence of a rich geometry in the probability space, the so-called α-geometry, consisting of a dually affine connection endowed by a Hessian structure [21], that is pertinent, in particular, to study of deformed exponential family [22–25].

In recent years, a certain interest has been given to the study of some probability distributions generated by deformed exponential families in which the exponential function is replaced by its generalized version. For instance, in [26] a family of generalized exponentials has been introduced by means of a positive increasing function and the corresponding thermostatistics was studied. Belonging to this family, the κ-exponential [27,28], at the heart of the κ-deformed statistical mechanics introduced by Kaniadakis in [29,30], is useful to study anomalous systems characterized by non-Gibbsian distributions with an asymptotic free-scale behavior.

Up today, many papers have been written on the foundations and the theoretical consistency of κ-statistical mechanics [31–48] (see also [49] and references therein). It has been applied in many research fields, such as statistical physics, thermostatistics, financial physics, social science, statistics and information theory [50–65].

In this paper, inspired by a recent work of Zhang and Naudts [66,67], we study the κ-distribution and its associated statistical manifold in the framework of information geometry, revisiting some results already known [68–70] and deriving new geometrical structures obtained from three different, not equivalent, divergence functions corresponding to the κ-deformed version of Brègman, "Kerridge" and Kullback–Leibler divergences.

It is known that starting from a deformed exponential family, one can introduce several kinds of different statistical manifold [71–73]. These can be derived from a hierarchical structure of escort expectations [74,75]. In particular, a Fisher metric can be obtained by introducing an expectation based on the first order escort distribution while a cubic form can be obtained by introducing an expectation based on the second escort distribution.

Escort averages can be useful to overcome several problems plaguing power-law distributions. For instance, such distributions have not finite momenta of any order as in κ-statistics where $\langle x^n \rangle < \infty$ only for $n < 1/\kappa - 1$ [76]. An expectation defined by means of escort distribution might be more appropriate to solve these and other questions, although other types of nonlinear expectation can also be introduced [77].

In the framework of κ-deformed formalism escort distributions has been previously introduced in [68] to investigate a dually flat geometry in the κ-distribution space and than a double escort distribution has been defined in [70] to study a second dually flat geometry, which is based on the escort expectation instead of the standard expectation. In both these geometric structures a kind of fluctuation-response relation, that could be relevant in the framework of the non-equilibrium

κ-deformed statistical mechanics, has been deduced by using, respectively, escort and double escort expectations. In addition, escort distribution [78] appears recurrently in the research area of multi-fractals and generalized statistical mechanics.

In this paper, we show that within the κ-deformed exponential formalism, one can introduce three kinds of geometrical structures. The first, is an invariant geometry where the Fisher information is the unique Riemannian metric together with a dual pair of invariant affine connections. The other two, are dually-flat geometries that introduce a Legendre structure and a Hessian structure on the corresponding statistical manifold [70] where some potential functions, like the Massieu function ψ_κ and the entropic form S_κ, are related in a way formally similar to what is done in the Kählerian geometry.

The main novelty of this work is to illuminate on the existent of an invariant geometry related to the κ-deformed formalism that has been never explored in the existent literature as well as on the existent of another dually-flat structure different from, but related to the one already studied in [70] that is characterized by scaled potentials, named para-potentials, and by a slimmer expression for the probability distribution.

The structure of the paper is as follows: In the next Section 2 we summarize the main aspect of some κ-deformed functions given by $\exp_\kappa(x)$, $\ln_\kappa(x)$ and $u_\kappa(x)$, that represent the bricks constituting the entire κ-deformed formalism. In Section 3, for the convenience of the reader, we present an overview of the (now) classical information geometry. The three statistical manifolds derived from different version of divergence function are studied in Section 4 and in particular, in Section 4.1 we introduce the statistical manifold derived from the κ-Kullback–Leibler divergence that is an invariant geometry with a positive curvature, while in Sections 4.2 and 4.3 we introduced two dually-flat geometries with their Hessian structures obtained, respectively, from the κ-Kerridge and κ-Brègman divergences. Our conclusions are reported in Section 5.

2. κ-Deformed Functions

To begin with, let us review some preliminary facts concerning the κ-deformed functions remanding to the relevant literature for the details [26,27,37].

Firstly, the κ-logarithm, defined in

$$\ln_\kappa(x) = \frac{x^\kappa - x^{-\kappa}}{2\kappa}\,, \tag{1}$$

where, for $\kappa^2 < 1$, $\ln_\kappa(x) = \ln_{-\kappa}(x)$ is a continuous, monotonic, increasing and concave function for $x \in \mathscr{R}^+$, with $\ln_\kappa(\mathscr{R}^+) \subseteq \mathscr{R}$. It fulfills the relation $\lambda \ln_\kappa(\varepsilon) = 1$ that will be useful in the remainder, with $\varepsilon = \left(\frac{1+\kappa}{1-\kappa}\right)^{1/2\kappa}$ and $\lambda = \sqrt{1+\kappa^2}$. In the $\kappa \to 0$ limit, these constants recover the Neperian natural number and the unity, i.e., $\varepsilon = e$ and $\lambda = 1$, as well as, in the same limit, the κ-logarithm reduces to the standard logarithm $\ln_0(x) = \ln(x)$. Remarkably, $\ln_\kappa(x)$ satisfies the relation $\ln_\kappa(x) = -\ln_\kappa(1/x)$ like does the standard logarithm.

Another κ-deformed function recurrent in this formalism, namely u-function, is given by

$$u_\kappa(x) = \frac{x^\kappa + x^{-\kappa}}{2}\,. \tag{2}$$

For $\kappa^2 < 1$, the u-function $u_\kappa(x) = u_{-\kappa}(x)$ is continuous for $x \in \mathscr{R}^+$, with $u_\kappa(\mathscr{R}^+) \subset \mathscr{R}^+$, $u_\kappa(0) = u_\kappa(+\infty) = +\infty$ and obtains its minimum at $x = 1$.

Like $\ln_\kappa(x)$, the deformed function $u_\kappa(x)$ satisfies the same relations $\lambda u_\kappa(\varepsilon) = 1$ as well as $u_\kappa(x) = u_\kappa\left(\frac{1}{x}\right)$ while it becomes a constant in the $\kappa \to 0$ limit: $u_0(x) = 1$.

The two functions $\ln_\kappa(x)$ and $u_\kappa(x)$ are closely related each other in several ways. For instance, from the calculus we have

$$\frac{d}{dx}\ln_\kappa(x) = \frac{u_\kappa(x)}{x}, \qquad\qquad \frac{d}{dx}u_\kappa(x) = \kappa^2\frac{\ln_\kappa(x)}{x}. \tag{3}$$

These two relations can be rewritten in an integral form according to

$$\ln_\kappa(x) = \int_1^x \frac{u_\kappa(y)}{y}\,dy, \qquad\qquad u_\kappa(x) = \kappa^2\int_1^x \frac{\ln_\kappa(y)}{y}\,dy + 1, \tag{4}$$

that can be assumed as the defining relations for these κ-deformed functions.

A very general approach to introduce deformed logarithms using an integral relation has been formulated in [26]. Let $\phi(s)$ be a strictly increasing function, we can define a deformed logarithm, namely ϕ-logarithm, in

$$\ln_\phi(x) = \int_1^x \frac{1}{\phi(s)}\,ds. \tag{5}$$

When we choose $\phi(s) = s$, the ϕ-logarithm reduces to the standard logarithm $\ln(x)$. Otherwise, by posing $\phi(s) = s/u_\kappa(s)$, Equation (5) is just the first of Equation (4), and the ϕ-logarithm reproduces the κ-logarithm (1).

It is interesting to note that, according to Equation (4), the function $u_\kappa(x)$ can be introduced as like as the κ-logarithm, by posing $\phi(s) = s/(\kappa^2\ln_\kappa(s))$. However, it should be noted that $u_\kappa(x)$ is not a deformed logarithm since $s/(\kappa^2\ln_\kappa(s))$ is not a monotonic increasing function.

The ϕ-logarithm admits the inverse function, namely the ϕ-exponential, defined in an implicit form as

$$\exp_\phi(x) = 1 + \int_0^x \phi\big(\exp_\phi(s)\big)\,ds. \tag{6}$$

In the κ-formalism, this equation can be solved explicitly so that the κ-deformed exponential takes the expression

$$\exp_\kappa(x) = \left(\kappa x + \sqrt{1 + \kappa^2 x^2}\right)^{1/\kappa}. \tag{7}$$

Its analytic properties follows from those of the κ-logarithm, i.e., for $\kappa^2 < 1$, $\exp_\kappa(x) = \exp_{-\kappa}(x)$ is a continuous, monotonic, increasing and convex function, with $\exp_\kappa(\mathscr{R}) \subseteq \mathscr{R}^+$ and $\exp_\kappa(0) = 1$. Moreover, it satisfies the relation $\exp_\kappa(-x)\exp_\kappa(x) = 1$ like the standard exponential does.

On the physical ground, equilibrium distribution of a system described by an entropic form $S(p)$ constrained by a given observable U, for instance, the internal energy, can be obtained from the optimal problem

$$\text{Max}_p\left(S(p) - \beta\,\mathbb{E}_p[\epsilon]\right) = 0, \tag{8}$$

where β is the Lagrange multipliers associated to the constraint $U = \mathbb{E}_p[\epsilon]$, with $\mathbb{E}_p[x] = \sum_\mu p_\mu x^\mu$ the standard linear average and ϵ^μ are the outcomes of the observable corresponding, for instance, to the available energy levels of the system (hereinafter, Greek indices run on 0 to n; Latin indices run on 1 to n).

In Equation (8) we assumed, at priori, the normalization of pdf so that, we consider $p_0 = 1 - \sum_i p_i$ as a function of the other p_i avoiding to introduce, in this way, the normalization as a further constraint. In the κ-formalism, entropy assumes the following trace-form expression [29,30]

$$S_\kappa(\boldsymbol{p}) = -\sum_\nu p_\nu \ln_\kappa(p_\nu) \equiv -\mathbb{E}_p \left[\ln_\kappa(\boldsymbol{p}) \right] , \tag{9}$$

which reduces to the Boltzmann–Gibbs–Shannon entropy $S^{\mathrm{BGS}}(\boldsymbol{p}) = -\sum_\mu p_\mu \ln(p_\mu)$ in the $\kappa \to 0$ limit. The optimal problem (8) becomes

$$\frac{\delta}{\delta p_i} \left\{ -\left(1 - \sum_i p_i \right) \ln_\kappa \left(1 - \sum_i p_i \right) - \sum_i p_i \ln_\kappa(p_i) - \beta \left[\epsilon^0 \left(1 - \sum_i p_i \right) + \sum_i \epsilon^i p_i \right] \right\} = 0 , \tag{10}$$

that, solved for p_i, gives

$$p_i = \frac{1}{\varepsilon} \exp_\kappa \left(\frac{1}{\lambda} \left(\theta^i - \gamma(\boldsymbol{\theta}) \right) \right) , \tag{11}$$

where we have posed $\boldsymbol{\theta} \equiv \{\theta^i\}$, with

$$\theta^i = -\beta \left(\epsilon^i - \epsilon^0 \right) , \quad \text{and} \quad \gamma(\boldsymbol{\theta}) = -\lambda \ln_\kappa \left(\varepsilon \, p_0 \right) . \tag{12}$$

Here, $\gamma(\boldsymbol{\theta})$ is a function of parameters θ^i and is fixed by the normalization condition.

Strictly related to a given entropic form, the divergence function plays a relevant role in information theory. In a sense, it measures the "distance" between an arbitrary distribution p_μ and the reference distribution q_μ, although it is not rigorously a "distance" function since, in general, is not symmetric in its arguments and does not satisfy the triangle inequality.

Within the Shannon entropy, the Kullback–Leibler divergence $D[\boldsymbol{p}; \boldsymbol{q}]$ [79] is defined as

$$D[\boldsymbol{p}; \boldsymbol{q}] = \sum_\mu p_\mu \ln \left(\frac{p_\mu}{q_\mu} \right) , \tag{13}$$

and can be written in the equivalent forms

$$D[\boldsymbol{p}/\boldsymbol{q}] = \sum_\mu p_\mu \left(\ln(p_\mu) - \ln(q_\mu) \right) , \tag{14}$$

that we will call, with abuse of language, "Kerridge" divergency, because the cross term is known in information theory as Kerridge inaccuracy [80], or also in

$$D[\boldsymbol{p} \| \boldsymbol{q}] = S[\mathbf{q}] - S[\boldsymbol{p}] + \sum_\mu \frac{\partial S[\boldsymbol{q}]}{\partial q_\mu} \left(p_\mu - q_\mu \right) , \tag{15}$$

where, this last expression is known as Brègman divergence [81].

However, the situation is more complicated in the κ-formalism [44] since, in this case, the above expressions, rewritten in

$$D_\kappa[\boldsymbol{p}; \boldsymbol{q}] = \sum_\mu p_\mu \ln_\kappa \left(\frac{p_\mu}{q_\mu} \right) , \tag{16}$$

$$D_\kappa[\boldsymbol{p}/\boldsymbol{q}] = \sum_\mu p_\mu \left(\ln_\kappa(p_\mu) - \ln_\kappa(q_\mu) \right) , \tag{17}$$

$$D_\kappa[\boldsymbol{p} \| \boldsymbol{q}] = S_\kappa[\boldsymbol{q}] - S_\kappa[\boldsymbol{p}] - \lambda \sum_\mu \left(p_\mu - q_\mu \right) \ln_\kappa \left(\varepsilon \, q_\mu \right) , \tag{18}$$

turn out to be no more equivalent, nor related to each other in a simple manner, as in the $\kappa = 0$ case.

Functionals (16)–(18) are non-negative definite and vanish iff $p = q$, a property that follows from the concavity of $\ln_\kappa(x)$. In particular, the κ-Brègman divergence (18) has been introduced and studied in [37].

3. Statistical Manifold and Its Hessian Structure

In this section, we summarize the main aspects of a statistical manifold [1,2].

Let $\boldsymbol{\theta} \equiv \{\theta^1, \dots, \theta^m\} \in \Xi$ a set of m real parameters, where Ξ is an open subset of $\mathcal{R}^m$. We introduce a discrete statistical model $\mathscr{S}$, the $(n+1)$-dimensional simplex of discrete probability functions on the simple space Ω:

$$\mathscr{S} = \left\{ \boldsymbol{p} \equiv \{p_\mu(\boldsymbol{\theta})\} : \Omega \to \mathcal{R} \,\middle|\, \sum_\mu p_\mu(\boldsymbol{\theta}) = 1, \, p_\mu(\boldsymbol{\theta}) > 0 \right\}, \tag{19}$$

where $\mu = 0, \dots, n$, labels the $n+1$ possible outcomes with probability p_μ.

Under suitable conditions $\mathscr{S}$ can be regarded as a manifold with local coordinates system $\boldsymbol{\theta}$, endowed by a metric tensor

$$g_{ij} = \mathbb{E}_p\left[\partial_i \ell \, \partial_j \ell\right], \tag{20}$$

with $\partial_i = \partial/\partial\theta^i$ and $\boldsymbol{\ell} \equiv \{\ell_\nu(\boldsymbol{\theta})\} = \{\ln p_\nu(\boldsymbol{\theta})\}$, by a cubic form

$$C_{ijl} = \mathbb{E}_p\left[\partial_i \ell \, \partial_j \ell \, \partial_l \ell\right], \tag{21}$$

that is a third-order completely symmetric tensor and by a Riemannian (Levi-Civita) connection

$$\Gamma_{ij,l}^{(0)} = \frac{1}{2}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right). \tag{22}$$

Starting from definitions (21)–(22) we can introduce a family of affine connections, named α-connection $\nabla^{(\alpha)}$, with $\alpha \in \mathcal{R}$, defined in

$$\Gamma_{ij,l}^{(\alpha)} = \Gamma_{ij,l}^{(0)} - \frac{\alpha}{2} C_{ijl}. \tag{23}$$

In particular, 1-connection and −1-connection play a special role in information geometry. They are called, respectively, exponential-connection $\nabla^{(e)} \equiv \nabla^{(1)}$, and mixture-connection $\nabla^{(m)} \equiv \nabla^{(-1)}$ and in a local representation are given by

$$\Gamma_{ij,l}^{(1)} = \mathbb{E}_p\left[\partial_{ij}\ell \, \partial_l \ell\right], \qquad \Gamma_{ij,l}^{(-1)} = \mathbb{E}_p\left[\frac{1}{p}\partial_{ij} p \, \partial_l \ell\right]. \tag{24}$$

We recall that metric (20) is known in statistics as Fisher information and is the only Riemannian metric invariant under sufficient statistics on the manifold of a probability distribution. For this reason, the geometry defined by Equations (20)–(22) is called invariant geometry.

More in general, starting from the affine connection ∇, we can introduce with respect to g a dual connection ∇^*, defined by of the relation

$$\partial_l g_{ij} = \Gamma_{li,j} + \Gamma_{lj,i}^*. \tag{25}$$

We say the triplet (Ξ, g, ∇) is a statistical manifold if ∇g is totally symmetric, that is

$$\partial_j g_{il} - \Gamma_{ji,l} - \Gamma_{jl,i} = \partial_i g_{jl} - \Gamma_{ij,l} - \Gamma_{il,j}, \tag{26}$$

and in addition, if g is torsion-free, also $\nabla^* g$ is totally symmetric. Therefore, (Ξ, g, ∇^*) is again a statistical manifold dual to (Ξ, g, ∇).

We say the statistical manifold (Ξ, g, ∇) is flat if ∇ is curvature-free. In this case, there exist locally on Ξ a coordinate systems θ, named affine coordinate systems, such that the connection $\Gamma_{ij,l}$ vanish on its coordinates neighbourhood.

It can be shown that $\Gamma^{(\alpha)}$ and $\Gamma^{(-\alpha)}$ are dual to each other and this hold also for the exponential and the mixture connection. In particular, as it is well-known, the statistical model defined by the exponential family

$$\mathscr{S}^{\exp} = \left\{ p \equiv \{p_\mu(\theta)\} : \Omega \to \mathscr{R} \,\middle|\, p_i(\theta) = \exp\left[\sum_j \theta^j c_j(x_i) - \gamma(\theta) \right] , \ p_0(\theta) = 1 - \sum_i p_i(\theta) \right\} , \quad (27)$$

where $c_j(\chi)$ are given functions of a random variable χ and $\gamma(\theta)$ is the normalization factor, admits a set of affine coordinates such that the exponential connection vanishes and the corresponding statistical manifold is said exponential-flat. In the same way, the statistical model defined by the mixture family

$$\mathscr{S}^{\mathrm{mix}} = \left\{ p \equiv \{p_\mu(\eta)\} : \Omega \to \mathscr{R} \,\middle|\, p_i(\eta) = \left(1 - \sum_j \eta_i\right) r_0(x_i) + \sum_j \eta_j r_j(x_i), \ p_0(\eta) = 1 - \sum_i p_i(\eta) \right\} , \quad (28)$$

where $r(\chi) = \{r_\mu(\chi)\}$ are given distribution functions of a random variable χ, with $\sum_\mu \eta_\mu = 1$, admits a set of affine coordinates such that the mixture connection vanishes and the corresponding statistical manifold is said mixture flat.

If ∇ is flat on Ξ, we say that the statistical manifold (Ξ, g, ∇) form a Hessian structure on Ξ if there exists a function $\psi \equiv \psi(\theta)$ such that, at least locally, the following formula holds

$$g_{ij}(\theta) = \partial_{ij}\psi . \quad (29)$$

In this case, be θ a ∇-affine coordinate system on Ξ, there exist a dual ∇-affine coordinate system $\eta \equiv \{\eta_j\}$ on the dual space Ξ, such that

$$g(\partial_i, \partial^j) = \delta_i^j , \quad (30)$$

where $\partial^j = \partial/\partial\eta_j$ and in addition, the following relation hold

$$g^{ij}(\eta) = \partial^{ij}\varphi , \quad (31)$$

for some function φ, where g^{ij} is the inverse of the Riemannian metric g_{ij}, with

$$g^{il} g_{lj} = \delta_j^i , \qquad g^{ij} = \partial^i \theta^j , \qquad g_{ij} = \partial_i \eta_j , \quad (32)$$

while the cubic form is given by

$$C_{ijl} = \partial_{ijl}\psi , \qquad \text{and} \qquad C^{ijl} = \partial^{ijl}\varphi . \quad (33)$$

It can be shown that the two dual-affine coordinate systems can be introduced according to

$$\eta_i = \partial_i\psi , \qquad \theta^i = \partial^i\varphi , \quad (34)$$

where the two functions, named respectively ψ-potential and φ-potential, are related by the relation

$$\psi(\theta) + \varphi(\eta) - \sum_i \theta^i \eta_i = 0 , \quad (35)$$

that introduces a Legendre structure on the statistical manifold.

Finally, let us define on $\Xi \times \Xi$ the function

$$D[\boldsymbol{p}, \boldsymbol{q}] = \psi\big(\boldsymbol{\theta}(\boldsymbol{p})\big) + \varphi\big(\boldsymbol{\eta}(\boldsymbol{q})\big) - \sum_i \theta^i(\boldsymbol{p})\, \eta_i(\boldsymbol{q}) \,, \tag{36}$$

called canonical divergence on the Hessian manifold (Ξ, g, ∇), that is a Brégman divergence [81]. It introduces a pseudo-distance that is asymmetric, with $D[\boldsymbol{p}, \boldsymbol{q}] \geq 0$, where equality holds iff $\boldsymbol{p} = \boldsymbol{q}$, in agreement with Equation (35).

If (Ξ, g, ∇) form a Hessian structure, it is easy to verify the following relations

$$\partial_{p^i} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = 0 \,, \qquad\qquad \partial_{p_i} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = 0 \,, \tag{37}$$

$$\partial_{p^i q^j} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -g_{ij} \,, \qquad\qquad \partial_{p_i q_j} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -g^{ij} \,, \tag{38}$$

$$\partial_{p^i p^j q^l} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -C_{ijl} \,, \qquad\qquad \partial_{p_i p_j q_l} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -C^{ijl} \,, \tag{39}$$

$$\partial_{p^i q^j q^l} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = 0 \,, \qquad\qquad \partial_{p_i q_j q_l} D[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = 0 \,, \tag{40}$$

where $\partial_{p^i} \equiv \partial/\partial\theta^i(\boldsymbol{p})$, $\partial_{p_i} \equiv \partial/\partial\eta_i(\boldsymbol{p})$ and similarly for $\boldsymbol{q}$.

More in general, following [82], any definite positive contrast function $\rho[\boldsymbol{p}, \boldsymbol{q}]$, that is a divergence compatible with the structure of the statistical manifold, induces a Riemannian geometry on Ξ, not necessarily flat, whose metric and connections are given by

$$\partial_{p_i q_j} \rho[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -g^{ij} \,, \qquad \partial_{p_i p_j q_l} \rho[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -\Gamma^{ij,l} \,, \qquad \partial_{p_l q_i q_j} \rho[\boldsymbol{p}, \boldsymbol{q}]\big|_{p=q} = -\Gamma^{*\,ij,l} \,, \tag{41}$$

and similarly for the low indices.

Remarkably, different contrast functions that share the same cross term induce the same geometric structure, since, according to the above equations, the crossing derivative vanishes for terms depending only on $\boldsymbol{p}$ or $\boldsymbol{q}$.

4. Statistical Manifolds in κ-Formalism

As known, any discrete distribution is in an exponential family and in a mixture family at the same time [2]. In fact, starting from the statistical model (19) we can define the distribution

$$p_\nu(\boldsymbol{\theta}) = \sum_\mu p_\mu\, \vartheta_\mu(\nu) = \left(1 - \sum_i p_i\right) \vartheta_0(\nu) + \sum_i p_i\, \vartheta_i(\nu) \,, \tag{42}$$

with

$$\vartheta_0(\nu) = 1 - \sum_i \vartheta_i(\nu) \,, \tag{43}$$

where $\vartheta_i(\nu) = 1$ when the discrete random variable $\nu = i$ and zero otherwise, that clearly belongs to the mixture-family (28), with $r_\mu(x_\nu) \equiv \vartheta_\mu(\nu)$ and $\eta_i \equiv p_i$.

Otherwise, by applying the logarithm to Equation (42) we obtain the distribution

$$p_\nu(\boldsymbol{\theta}) = \exp\left(\sum_i \big(\ln(p_i) - \ln(p_0)\big)\, \vartheta_i(\nu) + \ln(p_0)\right) \,, \tag{44}$$

that belongs to the exponential-family (27), with $\theta^\nu \equiv \sum_i(\ln(p_i) - \ln(p_0))$, $c_\nu(x_i) = \vartheta_i(\nu)$ and $\gamma(\boldsymbol{\theta}) \equiv -\ln(p_0)$.

More in general, given a generalized logarithm function $\ln_\phi(x)$, from Equation (42) we can introduce the distribution

$$p_v(\boldsymbol{\theta}) = \exp_\phi\left(\sum_i \left(\ln_\phi(p_i) - \ln_\phi(p_0)\right)\vartheta_i(v) + \ln_\phi(p_0)\right),\tag{45}$$

that is in a discrete generalized exponential-family with parameters $\theta^v \equiv \sum_i(\ln_\phi(p_i) - \ln_\phi(p_0))$, $c_v(x_i) = \vartheta_i(v)$ and $\gamma(\boldsymbol{\theta}) \equiv -\ln_\phi(p_0)$.

Therefore, the manifold of a discrete probability distribution is a super-manifold in which any statistical model of a discrete random variable, including ϕ-exponential family *a là Naudts*, is embedded as a submanifold. Different embedding of the same distribution induces different geometric structures which are related to non equivalent divergence functionals. In the following, we explore these alternatives starting from the divergence functionals introduced in Equations (16)–(18).

It is worth noticing that in the above arguments no hypothesis has been done on the origin of distribution $\boldsymbol{p}$. However, if this distribution came from an optimization problem, as in Equation (10), the coordinates θ^i would coincide with those introduced in Equation (12) and are related to the Lagrange multipliers of the corresponding constraints.

4.1. First Statistical Structure

The first geometric structure we are introducing came from the κ-deformed version of Kullback–Leibler divergence (16). It belongs to the family of the f-divergence studied by Csiszár [15]

$$D[\boldsymbol{p}, \boldsymbol{q}] = \sum_\mu p_\mu f\left(\frac{q_\mu}{p_\mu}\right),\tag{46}$$

where, in general, $f(x)$ is any convex differentiable function satisfying the condition $f(1) = 0$. Clearly, Equation (16) follows from (46) for $f(x) = -\ln_\kappa(x)$.

As known, any f-divergence induces a geometry invariant under sufficient statistics and, excepting the Kullback–Leibler divergence, the corresponding statistical manifold has not a Hessian structure nor is dually-flat and thence is characterized by a not vanishing curvature [2].

In this case, the embedding we are implementing is given by the standard logarithmic

$$p_\mu \to \ell_\mu = \ln\left(p_\mu\right),\tag{47}$$

where the natural coordinates are those of the exponential family $\boldsymbol{\eta} \equiv \{p_i\}$ and $\boldsymbol{\theta} \equiv \{\theta^i\}$.

By posing $\rho[\boldsymbol{p}, \boldsymbol{q}] \equiv D_\kappa[\boldsymbol{q}; \boldsymbol{p}]$ from Equations (41) we straightforward derive metric and connection in the Ξ space, according to

$$g^{ij} = \frac{1}{p_0} + \frac{\delta^{ij}}{p_i},\tag{48}$$

$$\Gamma^{ij,l} = (1 + \kappa^2)\left(\frac{1}{p_0^2} - \frac{\delta^{ijl}}{p_i^2}\right),\tag{49}$$

$$\Gamma^{*\,ij,l} = -\kappa^2\left(\frac{1}{p_0^2} - \frac{\delta^{ijl}}{p_i^2}\right),\tag{50}$$

where $\delta^{ijl} = 1$ for $i = j = l$ and zero otherwise, and coincide, as expected, with the definitions of statistical Fisher information matrix and α-connection.

The inverse relations are given by

$$g_{ij} = p_i \left(\delta_{ij} - p_j \right) , \tag{51}$$

$$\Gamma_{ij,l} = \left(1 + \kappa^2 \right) p_i \left[\left(\delta_{ijl} - p_j \, \delta_{jl} \right) - p_j \left(\delta_{il} - p_l \right) - p_l \left(\delta_{lj} - p_j \right) \right] , \tag{52}$$

$$\Gamma^*_{ij,l} = -\kappa^2 \, p_i \left[\left(\delta_{ijl} - p_j \, \delta_{jl} \right) - p_j \left(\delta_{il} - p_l \right) - p_l \left(\delta_{lj} - p_j \right) \right] , \tag{53}$$

while from Equation (23) we can derive the Levi-Civita connection

$$\Gamma^{(0)}_{ij,l} = \frac{1}{2 \left(1 + \kappa^2 \right)} \, \Gamma_{ij,l} , \tag{54}$$

that does not depend on the deformation parameter and the cubic form

$$C^{(\kappa)}_{ij,l} = -\frac{1 + 2\,\kappa^2}{1 + \kappa^2} \, \Gamma_{ij,l} .$$

These relations confirm that κ-Kullback–Leibler divergence actually introduces the α-geometry in the framework of the κ-formalism where

$$\alpha = 1 + 2\,\kappa^2 , \tag{55}$$

in agreement with the general relation $\alpha = 3 + 2\,f'''(1)$ [23].

Remark that this correspondence holds only for $\alpha \geq 1$, considering that the exponential family follows in the $\kappa \to 0$ limit. This should be compared with the results discussed in [25], where an invariant geometry has been derived in the framework of the Tsallis formalism of statistical mechanics starting from the corresponding relative entropy. In that case, the analogue relationships between parameters is given by $\alpha = 1 - 2\,q$, which is consistent for $\alpha < 1$.

The statistical manifold introduced in Equations (48)–(50) has not a dually flat structure, nor a Hessian structure, but it has a constant curvature whose Riemann-Christoffel tensor, given by

$$R^{(\kappa)}_{ijls} = \partial_i \Gamma_{jls} - \partial_j \Gamma_{ils} + \sum_r \left(\Gamma_{irs} \, \Gamma^r_{jl} - \Gamma_{jrs} \, \Gamma^r_{il} \right) , \tag{56}$$

becomes

$$R^{(\kappa)}_{ijls} = c^{(\kappa)} \left(g_{jl} \, g_{is} - g_{il} \, g_{js} \right) , \tag{57}$$

with

$$c^{(\kappa)} = \kappa^2 \left(\kappa^2 - 1 \right) . \tag{58}$$

As expected, $R^{(\kappa)}_{ijls}$ vanishes in the $\kappa \to 0$ limit, where the whole geometric structure collapses to the dually-flat 1-geometry.

The Ricci curvature tensor and the scalar curvature are readily evaluated in

$$R_{ij} = \left(1 - n \right) c^{(\kappa)} \, g_{ij} , \qquad R = \left(n - n^2 \right) c^{(\kappa)} , \tag{59}$$

that are positive definite quantities for $\kappa^2 < 1$.

Finally, in spite of the lack of an entropic form related to this invariant geometry, we can derive a particular family of discrete distributions within the κ-formalism adopting an optimization procedure based on the κ-deformed Kullback–Leibler divergence.

In this way, the equilibrium distribution achieving the minimum divergence of the function (16) is given by the following problem

$$\frac{\delta}{\delta p_i} \left\{ \sum_\mu p_\mu \ln_\kappa \left(\frac{p_\mu}{q_\mu} \right) + \beta \sum_\mu \epsilon^\mu p_\mu \right\} = 0 \,, \tag{60}$$

where we assume, at priori, the normalization for p, that is $p_0 = 1 - \sum_i p_i$, and similar for the target distribution q. We get the κ-distribution in the form

$$p_i = \frac{q_i}{\epsilon} \exp_\kappa \left(\frac{1}{\lambda} \left(\widetilde{\theta^i} - \gamma(\widetilde{\boldsymbol{\theta}}) \right) \right) \,, \tag{61}$$

with $\gamma(\widetilde{\boldsymbol{\theta}}) = -\lambda \ln_\kappa(\epsilon\, p_0/q_0)$ and $\widetilde{\boldsymbol{\theta}}$, defined in a similar manner, as in Equation (12), are related to the natural coordinates $\boldsymbol{\theta}$ in

$$\theta^i = \ln \left(\frac{q_i}{q_0} \right) + \frac{1}{\kappa} \left[\operatorname{arcsinh} \left(\frac{\kappa}{\lambda} (\widetilde{\theta^i} - \gamma(\widetilde{\boldsymbol{\theta}})) \right) + \operatorname{arcsinh} \left(\frac{\kappa}{\lambda} \gamma(\widetilde{\boldsymbol{\theta}}) \right) \right] . \tag{62}$$

Therefore, let q be a reference probability distribution on $\mathscr{S}$, the sub-manifold of the statistical family

$$\mathscr{S}_{(1)}^\kappa = \left\{ \boldsymbol{p} \equiv \{p_\mu(\widetilde{\boldsymbol{\theta}})\} : \Omega \to \mathscr{R} \,\middle|\, p_i = \frac{q_i}{\epsilon} \exp_\kappa \left(\frac{1}{\lambda} (\widetilde{\theta^i} - \gamma(\widetilde{\boldsymbol{\theta}})) \right); \, p_0 = 1 - \sum_i p_i \right\} \,, \tag{63}$$

that optimize the divergence function (16) admits a Riemannian structure described by the Fisher metric (48), dual connection (49) and (50) and constant positive curvature (59). This geometric structure collapses, in the $\kappa \to 0$ limit, in the dually-flat geometry of the exponential distribution families.

4.2. Second Statistical Structure

The next geometric structure we are introducing came from κ-deformed version of "Kerridge" divergence (17). It is strictly related to the following statistical model

$$\mathscr{S}_{(2)}^\kappa = \left\{ \boldsymbol{p} \equiv \{p_\mu(\boldsymbol{\theta})\} : \Omega \to \mathscr{R} \,\middle|\, p_i = \exp_\kappa \left(\theta^i - \gamma(\boldsymbol{\theta}) \right); \, p_0 = 1 - \sum_i p_i \right\} \,, \tag{64}$$

consistent with the embedding

$$p_\mu \to \ln_\kappa(p_\mu) \,. \tag{65}$$

To introduce a manifold structure on $\mathscr{S}_{(2)}^\kappa$ we pose $\boldsymbol{\eta} \equiv \{p_i\}$ so that, from relations (41) applied to $\rho[\boldsymbol{p}, \boldsymbol{q}] \equiv D_\kappa[\boldsymbol{q}/\boldsymbol{p}]$, we obtain metric and connection as

$$g^{ij} = \frac{u_\kappa(\eta_0)}{\eta_0} + \frac{u_\kappa(\eta_i)}{\eta_i} \delta^{ij} \,, \tag{66}$$

$$\Gamma^{ij,l} = \lambda \frac{u_\kappa(\eta_0/\epsilon)}{\eta_0^2} - \lambda \frac{u_\kappa(\eta_i/\epsilon)}{\eta_i^2} \delta^{ijl} \,, \tag{67}$$

$$\Gamma^{*\,ij,l} = 0 \,. \tag{68}$$

These quantities define a dually-flat structure on $\mathscr{S}_{(2)}^\kappa$ compatible with a Hessian geometry. In fact, give the φ-potential as

$$\varphi(\boldsymbol{\eta}) = -\frac{1}{\lambda^2} \widetilde{S}_\kappa(\boldsymbol{\eta}) \,, \tag{69}$$

where $\widetilde{S}_\kappa(\boldsymbol{\eta})$ is the parentropy introduced in [34] and related to the κ-entropy in

$$\widetilde{S}_\kappa(\boldsymbol{\eta}) = \lambda\, S_\kappa\left(\frac{\boldsymbol{\eta}}{\varepsilon}\right), \tag{70}$$

metric and connection follows straightforward from Equations (31) and (33), respectively.

Equation (68) states that coordinate $\boldsymbol{\eta}$ is ∇-affine on Ξ, whereas its dual coordinate system is given by

$$\theta^i = \ln_\kappa(\eta_i) + \gamma(\boldsymbol{\theta}), \tag{71}$$

with $\gamma(\boldsymbol{\theta}) = -\ln_\kappa(\eta_0)$.

The ψ-potential can be obtained from the Legendre transform (35) on $\varphi(\boldsymbol{\eta})$ and is given by

$$\psi(\boldsymbol{\theta}) = \frac{1}{\lambda}\widetilde{\mathbb{I}}_\kappa(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta}), \tag{72}$$

where the pull-back of $\widetilde{\mathbb{I}}_\kappa(\boldsymbol{\theta})$ on Ξ, denoted in $\widetilde{\mathbb{I}}_\kappa^*(\boldsymbol{\eta})$, is defined in

$$\widetilde{\mathbb{I}}_\kappa^*(\boldsymbol{\eta}) = \sum_i \eta_i\, u_\kappa\left(\frac{\boldsymbol{\eta}}{\varepsilon}\right) \equiv \mathbb{E}_p\left[u_\kappa\left(\frac{\boldsymbol{\eta}}{\varepsilon}\right)\right]. \tag{73}$$

Function $\widetilde{\mathbb{I}}_\kappa^*(\boldsymbol{\eta})$ is related to $\mathbb{I}_\kappa^*(\boldsymbol{\eta})$, usually introduced in the κ-deformed statistical mechanics [35], according to

$$\widetilde{\mathbb{I}}_\kappa^*(\boldsymbol{\eta}) = \lambda\, \mathbb{I}_\kappa^*\left(\frac{\boldsymbol{\eta}}{\varepsilon}\right), \tag{74}$$

where

$$\mathbb{I}_\kappa^*(\boldsymbol{\eta}) = \sum_i \eta_i\, u_\kappa(\boldsymbol{\eta}) \equiv \mathbb{E}_p\left[u_\kappa(\boldsymbol{\eta})\right], \tag{75}$$

reduces to unity in the $\kappa \to 0$ limit. In [34], functional $\mathbb{I}_\kappa^*(\boldsymbol{\eta})$ has been related to the parentropy $\widetilde{S}_\kappa[\boldsymbol{\eta}]$, through the relation

$$\mathbb{I}_\kappa^*(\boldsymbol{\eta}) = \widetilde{S}_\kappa(\boldsymbol{\eta}) - S_\kappa(\boldsymbol{\eta}) + 1. \tag{76}$$

As will be highlighted in Section 4.3, ψ-potential is related on the physical ground to the κ-logarithm of partition function and to the scaled free energy, the para-free energy, of a physical system described by the parentropy (70), in agreement with the relation

$$\widetilde{F}_\kappa = -\frac{1}{\beta}\left(\widetilde{\mathbb{I}}_\kappa + \gamma\right). \tag{77}$$

By using the expression (72) we can verify the relation

$$\eta_i = \partial_i\left(\frac{1}{\lambda}\widetilde{\mathbb{I}}_\kappa^*\left(\frac{\boldsymbol{\eta}}{\varepsilon}\right) + \gamma(\boldsymbol{\theta})\right), \tag{78}$$

so that metric and connections on the Ξ-space reads

$$g_{ij} = \frac{\eta_i}{u_\kappa(\eta_i)}\left(\delta_{ij} - \partial_j\gamma\right), \tag{79}$$

$$\Gamma_{ij,l} = 0, \tag{80}$$

$$\Gamma^*_{ij,l} = \frac{\eta_i}{u_\kappa(\eta_i)}\left[\lambda\,\frac{u_\kappa\left(\frac{\eta_i}{\varepsilon}\right)}{u_\kappa^2(\eta_i)}\left(\delta_{ij} - \partial_i\gamma\right)\left(\delta_{il} - \partial_j\gamma\right) - \partial_{jl}\gamma\right], \tag{81}$$

confirming the dually-flat structure of $\mathscr{S}^\kappa_{(2)}$, with $\boldsymbol{\theta}$ the ∇-affine coordinate on Ξ.

It is worthy to derive the canonical divergence for $\mathscr{S}^\kappa_{(2)}$, that is a Brégman-like divergence obtainable from Equation (36). In fact, by using κ-parentropy it reads

$$D[\boldsymbol{p}, \boldsymbol{q}] = \frac{1}{\lambda^2}\left(\widetilde{S}_\kappa(\boldsymbol{q}) - \widetilde{S}_\kappa(\boldsymbol{p})\right) - \sum_\mu\left(p_\mu - q_\mu\right)\ln_\kappa\left(q_\kappa\right), \tag{82}$$

and, as expected, although differs from divergence (17), they share the same cross term so that both divergences (17) and (82) give rise to the same geometric structure.

Finally, it is easy to verify that exponential family $\mathscr{S}^\kappa_{(2)}$ can be derived from the optimization problem (8), where the entropic form is given by the parentropy $\widetilde{S}_\kappa(\boldsymbol{\eta})$ and the dual coordinate are related to the Lagrange multiplier according to Equation (12).

To go over in the study of this geometry, we show that the Hessian structure on Ξ is consistent with a scalar product on $\mathscr{S}^\kappa_{(2)}$ induced by the κ-escort probability distribution

$$\langle A, B\rangle = \mathbb{E}_{\widetilde{p}(1)}\left[A\,B\right], \tag{83}$$

where

$$\mathbb{E}_{\widetilde{p}(1)}[f] = \sum_\mu \widetilde{P}^{(1)}_\mu\,f^\mu, \tag{84}$$

with

$$\widetilde{P}^{(1)}_\mu = \frac{1}{\widetilde{\mathscr{U}}_\kappa^{(1)}}\frac{p_\mu}{u_\kappa\left(p_\mu\right)}, \qquad \text{and} \qquad \widetilde{\mathscr{U}}_\kappa^{(1)} = \sum_\mu\frac{p_\mu}{u_\kappa\left(p_\mu\right)}. \tag{85}$$

We employ the embedding

$$p_\mu \to \widetilde{\ell}^{(\kappa)}_\mu = \ln_\kappa\left(p_\mu\right), \tag{86}$$

such that $\mathbb{E}_{\widetilde{p}(1)}\left[\partial^i\widetilde{\ell}^{(\kappa)}\right] = 0$, which is our bias condition [69].

Introducing the tangent vector at point $\boldsymbol{\theta}$ as

$$\boldsymbol{T}^j = \partial^j\widetilde{\ell}^{(\kappa)}, \tag{87}$$

the metric follows from the standard definition

$$\widetilde{g}^{ij} = \left\langle \boldsymbol{T}^i, \boldsymbol{T}^j\right\rangle = \frac{1}{\widetilde{\mathscr{U}}_\kappa^{(1)}}\sum_\mu\frac{u_\kappa\left(p_\mu\right)}{p_\mu}\partial^i p_\mu\partial^j p_\mu, \tag{88}$$

that is conformal equivalent to metric (66) according to [73]

$$\widetilde{g}^{ij} = \frac{g^{ij}}{\widetilde{\mathscr{U}}_\kappa^{(1)}}. \tag{89}$$

In the same manner, recalling that connection is related to the variation of the tangent vectors T^j under infinitesimal changing of the parameters $\theta \to \theta + d\theta$, we consider the quantities $\partial^i T^j \equiv \partial^{ij} \widetilde{\ell}^{(\kappa)}$, that projected on the base vectors T^l, gives

$$\widetilde{\Gamma}^{ij,l} \quad = \quad \left\langle \partial^i T^j, T^l \right\rangle = \mathbb{E}_{\widetilde{P}^{(1)}} \left[\partial^{ij} \widetilde{\ell}^{(\kappa)} \, \partial^l \widetilde{\ell}^{(\kappa)} \right] , \tag{90}$$

which is related to Equation (67) by means of the conformal factor $\mathscr{U}_\kappa^{(1)}$.

It is a duty to observe that $\partial^i T^j$ is actually unbiased because

$$\mathbb{E}_{\widetilde{P}^{(1)}} \left[\partial^i T^j \right] \neq 0 , \tag{91}$$

therefore the right way to obtain the connection elements is by biasing these quantities and then project them on the base vectors T^l according to

$$\widetilde{\Gamma}^{ij,l} \quad = \quad \left\langle \partial^i T^j - \mathbb{E}_{\widetilde{P}^{(1)}} \left[\partial^i T^j \right], T^l \right\rangle , \tag{92}$$

that, nevertheless, reproduce the result (90) due to our bias condition.

Finally, Equation (68) follows from relation (25) and the results (66) and (67).

In a similar way, we can derive the expressions (79)–(81) from the definitions of g_{ij}, $\Gamma_{ij,l}$ and $\Gamma^*_{ij,l}$, that here rewrite in the κ-escort formalism

$$g_{ij} = \widetilde{\mathscr{U}}_\kappa^{(1)} \, \widetilde{P}_i^{(1)} \left(\delta_{ij} - \widetilde{P}_j^{(1)} \right) , \tag{93}$$

$$\Gamma_{ij,l} = 0 , \tag{94}$$

$$\Gamma^*_{ij,l} = \widetilde{\mathscr{U}}_\kappa^{(2)} \, \widetilde{P}_i^{(2)} \left(\delta_{ij} - \widetilde{P}_j^{(1)} \right) \left(\delta_{il} - \widetilde{P}_l^{(1)} \right) - \widetilde{\mathscr{U}}_\kappa^{(1)} \, \widetilde{P}_i^{(1)} \partial_j \widetilde{P}_l^{(1)} , \tag{95}$$

where we introduced the double-escort probability $\widetilde{P}_i^{(2)}$ [70]

$$\widetilde{P}_i^{(2)} = \frac{\lambda}{\widetilde{\mathscr{U}}_\kappa^{(2)}} \frac{p_i \, u_\kappa(p_i/\varepsilon)}{u_\kappa^3(p_i)} , \qquad \text{with} \qquad \widetilde{\mathscr{U}}_\kappa^{(2)} = \lambda \sum_\mu \frac{p_\mu \, u_\kappa(p_\mu/\varepsilon)}{u_\kappa^3(p_\mu)} , \tag{96}$$

and used the relation

$$\partial_i \gamma = \frac{1}{\widetilde{\mathscr{U}}_\kappa^{(1)}} \frac{p_i}{u_\kappa(p_i)} \equiv \widetilde{P}_i^{(1)} , \tag{97}$$

obtained from of the condition $\partial_i(\sum_\mu p_\mu) = 0$.

Remarkably, in the escort formalism, the κ-Brègman divergence can be express in the language of the escort distribution according to

$$D[\boldsymbol{p}, \boldsymbol{q}] = \mathbb{E}_P \left[\ln_\kappa(\boldsymbol{p}) - \ln_\kappa(\boldsymbol{q}) \right] , \tag{98}$$

which has a very straightforward interpretation.

4.3. Third Statistical Structure

The last geometric structure that we are introducing came from the Brègman-like divergence and has been already studied from a different perspective in [68–70].

It is consistent with the following statistical model

$$\mathscr{S}_{(3)}^\kappa = \left\{ \boldsymbol{p} \equiv \{p_\mu(\boldsymbol{\theta})\} : \Omega \to \mathscr{R} \,\middle|\, p_i = \frac{1}{\varepsilon} \exp_\kappa \left(\frac{1}{\lambda} \left(\theta^i - \gamma(\boldsymbol{\theta}) \right) \right); \, p_0 = 1 - \sum_i p_i \right\} , \tag{99}$$

derived from the embedding

$$p_\mu \to \lambda \ln_\kappa (\varepsilon p_\mu) \,. \tag{100}$$

To introduce a manifold structure on $\mathscr{S}^\kappa_{(3)}$, let us identify the η-coordinate in $\boldsymbol{\eta} \equiv \{p_i\}$ and apply relations (41) to $\rho[\boldsymbol{p}, \boldsymbol{q}] \equiv D_\kappa[\boldsymbol{q} \| \boldsymbol{p}]$. We obtain metric and connections as

$$g^{ij} = \lambda \, \frac{u_\kappa (\varepsilon \eta_0)}{\eta_0} + \lambda \, \frac{u_\kappa (\varepsilon \eta_i)}{\eta_i} \, \delta^{ij} \,, \tag{101}$$

$$\Gamma^{ij,l} = \lambda^2 \, \frac{u_\kappa(\eta_0)}{\eta_0^2} - \lambda^2 \, \frac{u_\kappa(\eta_i)}{\eta_i^2} \, \delta^{ijl} \,, \tag{102}$$

$$\Gamma^{* \, ij,l} = 0 \,. \tag{103}$$

Clearly, $\mathscr{S}^\kappa_{(3)}$ has a dually-flat structure compatible with the Hessian geometry related to the φ-potential $\varphi(\boldsymbol{\eta}) \equiv -S_\kappa(\boldsymbol{\eta})$, that is the negentropy of $S_\kappa(\boldsymbol{\eta})$ given in Equation (9), so that the dual ∇-affine coordinate on Ξ are

$$\theta^i = -\frac{\partial S_\kappa(\boldsymbol{\eta})}{\partial \eta_i} \equiv \lambda \ln_\kappa (\varepsilon \eta_i) + \gamma(\boldsymbol{\theta}) \,, \tag{104}$$

while the ψ-potential follows from the Legendre transform on $\varphi(\boldsymbol{\eta})$ and reads

$$\psi(\boldsymbol{\theta}) = \mathbb{I}_\kappa(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta}) \,. \tag{105}$$

Remark that for positive measures, not constrained by $\sum_\mu p_\mu = 1$, $\mathbb{I}_\kappa$ is just the full Legendre transform of S_κ in the probability space [68].

On the physical ground, the potential (105) corresponds to the κ-logarithm of partition function and is related to the free energy of the system according to

$$F_\kappa = -\frac{1}{\beta} \left(\mathbb{I}_\kappa + \gamma \right) \,. \tag{106}$$

Therefore, the third statistical structure, as well as the second one in its scaled formulation, reproduces and confirms the Legendre structure of the κ-deformed thermostatics previously derived by using standard arguments of statistical mechanics [35].

By using the expression of $\psi(\boldsymbol{\theta})$ it is ready to verify the first of Equations (34), that is

$$\eta_i = \partial_i \left(\mathbb{I}_\kappa(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta}) \right) \,, \tag{107}$$

so that metric and connections on the Ξ-space read:

$$g_{ij} = \frac{1}{\lambda} \frac{\eta_i}{u_\kappa (\varepsilon \eta_i)} \left(\delta_{ij} - \partial_j \gamma \right) \,, \tag{108}$$

$$\Gamma_{ij,l} = 0 \,, \tag{109}$$

$$\Gamma^*_{ij,l} = \frac{\eta_i}{u_\kappa(\varepsilon \eta_i)} \left[\frac{1}{\lambda} \frac{u_\kappa(\eta_i)}{u_\kappa^2(\varepsilon \eta_i)} (\delta_{ij} - \partial_i \gamma)(\delta_{il} - \partial_j \gamma) - \partial_{jl} \gamma \right] \,, \tag{110}$$

that confirm the dually-flat structure of $\mathscr{S}^\kappa_{(3)}$.

It is trivial to verify that canonical divergence $D[\boldsymbol{p}, \boldsymbol{q}]$ on the statistical manifold $\mathscr{S}^\kappa_{(3)}$ coincides, as expected, with the κ-Brègman divergence $D_\kappa[\boldsymbol{q} \| \boldsymbol{p}]$ according to [68]

$$D[\boldsymbol{p}, \boldsymbol{q}] = D_\kappa[\boldsymbol{q} \| \boldsymbol{p}] \,. \tag{111}$$

Finally, let us just to observe that, in complete analogy with the discussion done in Section 4.2, the Hessian structure on $\mathscr{S}^\kappa_{(3)}$ is also compatible with the scalar product

$$\langle A, B \rangle = \mathbb{E}_{P^{(1)}} [A\,B] \, , \tag{112}$$

where

$$\mathbb{E}_{P^{(1)}} [f] = \sum_\mu P^{(1)}_\mu f^\mu \, , \tag{113}$$

are the escort average and the κ-escort distribution $P^{(1)}$ is

$$P^{(1)}_\mu = \frac{1}{\mathscr{U}^{(1)}_\kappa} \frac{p_\mu}{\lambda\, u_\kappa\left(\varepsilon\, p_\mu\right)} \, , \qquad \text{with} \qquad \mathscr{U}^{(1)}_\kappa = \sum_\mu \frac{p_\mu}{\lambda\, u_\kappa\left(\varepsilon\, p_\mu\right)} \, . \tag{114}$$

Again, in the escort formalism, the κ-Brègman divergence can be written in the straightforward manner given by

$$D_\kappa[q \| p] = \mathbb{E}_{P^{(1)}} \left[\ln_\kappa(p) - \ln_\kappa(q)\right] \, . \tag{115}$$

5. Concluding Remarks

In the framework of the κ-formalism we have derived several geometric structures starting from non-equivalent κ-divergence functions corresponding, in the $\kappa \to 0$ limit, to the Kullback–Leibler, "Kerridge" and Brégman divergences, respectively.

The main results that characterize the emerging geometries, such as metric, connection, θ-coordinate and φ- and ψ-potentials, when available, are reported in the final Table 1, for an easy comparison.

In particular, the statistical manifold derived from the κ-Kullback–Leibler divergence has an invariant geometry with a dual structure and constant curvature. This geometric structure corresponds to the well-known α-geometry [1,2], with $\alpha = 1 + 2\,\kappa^2$, and is related to the natural embedding of the standard logarithm.

The other two statistical manifolds, derived from the κ-Kerridge and κ-Brégman divergences, have a dually-flat structure with dual potentials related to each other by the corresponding Legendre transform. Their respective potentials are given by the scaled entropy (or parentropy) and scaled free energy in the first case and by the entropic form and free energy in the latter case. Both these geometries are compatible with an embedding based on the κ-logarithm and a scalar product based on the escort probability average. Moreover, the κ-Brégman functional turns out to be the natural canonical divergence in the own statistical manifold while the κ-Kerridge functional share with the canonical divergence the same cross term and therefore, it is equivalent to the corresponding canonical divergence. These dually-flat structures are together related by a merely scaling transformation of the potentials and the distributions. In this way, the structure based on the κ-Brégman functional has simplest expressions for the potentials with respect to the structure based on the κ-Kerridge functional that has a slimmer expression for the distribution.

Table 1. Geometrical structures emerging from different κ-deformed versions of divergent function.

κ-**Kullback–Leibler**	$D[p, q] = \sum_\mu p_\mu \ln_\kappa \left(\frac{p_\mu}{q_\mu} \right)$
metric	$g^{ij} = \frac{1}{p_0} + \frac{\delta^{ij}}{p_i}$
connection	$\Gamma^{ij,l} = (1 + \kappa^2) \left(\frac{1}{p_0^2} - \frac{\delta^{ijl}}{p_i^2} \right)$
distribution	$p_i = \frac{q_i}{\varepsilon} \exp_\kappa \left(\frac{1}{\lambda} (\widetilde{\theta^i} - \gamma(\widetilde{\boldsymbol{\theta}})) \right)$
$\widetilde{\theta}$-**coordinate**	$\widetilde{\theta^i} = \lambda \left(\ln_\kappa \left(\varepsilon \frac{p_i}{q_i} \right) + \ln_\kappa \left(\varepsilon \frac{p_0}{q_0} \right) \right)$
θ-**coordinate**	$\theta^i = \ln(p_i) - \ln(p_0)$
ϕ-**potential**	\ \
ψ-**potential**	\ \
κ-**Kerridge**	$D[p, q] = \sum_\mu p_\mu \left(\ln_\kappa(p_\mu) - \ln_\kappa(q_\mu) \right)$
metric	$g^{ij} = \frac{u_\kappa(\eta_0)}{\eta_0} + \frac{u_\kappa(\eta_i)}{\eta_i} \delta^{ij}$
connection	$\Gamma^{ij,l} = \lambda \frac{u_\kappa(\eta_0/\varepsilon)}{\eta_0^2} - \lambda \frac{u_\kappa(\eta_i/\varepsilon)}{\eta_i^2} \delta^{ijl}$
distribution	$p_i = \exp_\kappa \left(\theta^i - \gamma(\boldsymbol{\theta}) \right)$
θ-**coordinate**	$\theta^i = \ln_\kappa (p_i) + \ln_\kappa (p_0)$
ϕ-**potential**	$\varphi(\boldsymbol{\eta}) = -\frac{1}{\lambda} S_\kappa \left(\frac{\boldsymbol{\eta}}{\varepsilon} \right)$
ψ-**potential**	$\psi(\boldsymbol{\theta}) = \frac{1}{\lambda} \widetilde{\mathbb{I}}_\kappa(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta})$
κ-**Brégman**	$D[p, q] = S_\kappa(\boldsymbol{q}) - S_\kappa(\boldsymbol{p}) - \lambda \sum_\mu (p_\mu - q_\mu) \ln_\kappa (\varepsilon q_\mu)$
metric	$g^{ij} = \lambda \frac{u_\kappa(\varepsilon \eta_0)}{\eta_0} + \lambda \frac{u_\kappa(\varepsilon \eta_i)}{\eta_i} \delta^{ij}$
connection	$\Gamma^{ij,l} = \lambda^2 \frac{u_\kappa(\eta_0)}{\eta_0^2} - \lambda^2 \frac{u_\kappa(\eta_i)}{\eta_i^2} \delta^{ijl}$
distribution	$p_i = \frac{1}{\varepsilon} \exp_\kappa \left(\frac{1}{\lambda} \left(\theta^i - \gamma(\boldsymbol{\theta}) \right) \right)$
θ-**coordinate**	$\theta^i = \lambda \left(\ln_\kappa (\varepsilon p_i) + \ln_\kappa (\varepsilon p_0) \right)$
ϕ-**potential**	$\varphi(\boldsymbol{\eta}) = -S_\kappa[\boldsymbol{\eta}]$
ψ-**potential**	$\psi(\boldsymbol{\theta}) = \mathbb{I}_\kappa(\boldsymbol{\theta}) + \gamma(\boldsymbol{\theta})$

References

1. Amari, S.-I.; Nagaoka, H. *Methods of Information Geometry*; American Mathematical Society: Providence, RI, USA, 2000.
2. Amari, S.-I. *Information Geometry and Its Applications*; Springer: Tokyo, Japan, 2016.
3. Gibbs, J.W. A Method of Geometrical Representation of the Thermodynamic Properties of Substances by Means of Surfaces. *Trans. Conn. Acad.* **1873**, *II*, 382–404.
4. Charathéodory, C. Untersuchungen über die Grundlagen der Thermodynamik. *Math. Ann.* **1909**, *67*, 355–386. (In German) [CrossRef]
5. Ruppeiner, G. Riemannian geometry in thermodynamic fluctuation theory. *Rev. Mod. Phys.* **1995**, *67*, 605–659. [CrossRef]
6. Janyszek, H. On the Riemannian metrical structure in the classical statistical equilibrium thermodynamics. *Rep. Math. Phys.* **1986**, *24*, 1–10. [CrossRef]
7. Bravetti, A.; Lopez-Monsalvo, C.S.; Nettel, F. Contact symmetries and Hamiltonian thermodynamics. *Ann. Phys.* **2015**, *361*, 377–400. [CrossRef]
8. Quevedo, H. Geometrothermodynamics. *J. Math. Phys.* **2007**, *48*, 013506-14. [CrossRef]
9. Bravetti, A.; Lopez-Monsalvo, C.S.; Nettel, F. Conformal gauge transformations in thermodynamics. *Entropy* **2015**, *17*, 6150–6168. [CrossRef]

10. Mrugala, R. On a special family of thermodynamic processes and their invariants. *Rep. Math. Phys.* **2000**, *46*, 461–468. [CrossRef]

11. Mrugala, R. Continuous contact transformations in thermodynamics. *Rep. Math. Phys.* **1993**, *33*, 149–154. [CrossRef]

12. Rao, C.R. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81–91.

13. Fisher, R.A. *Statistical Methods for Research Workers*; Oliver and Boyd: Edinburgh, UK, 1925.

14. Čencov, N.N. *Statistical Decision Rules and Optimal Conclusion*; Nauka: Moscov, Russian, 1972.

15. Csiszár, I. Information-type measures of difference of probability distributions and indirect observation. *Studia Sci. Math. Hungar.* **1967**, *2*, 229–318.

16. Efron, B. The geometry of exponential families. *Ann. Stat.* **1978**, *6*, 362–376. [CrossRef]

17. Eguchi, S. Second order efficiency of minimum contrast estimators in a curved exponential family. *Ann. Stat.* **1983**, *11*, 793–803. [CrossRef]

18. Amari, S.-I. Differential geometry of curved exponential families-curvature and information loss. *Ann. Stat.* **1982**, *10*, 357–385. [CrossRef]

19. Amari, S.-I. Differential-geometrical methods in statistics. In *Lecture Note in Statistics*; Springer: Berlin/Heidelberg, Germany, 1985; p. 28.

20. Amari, S.-I. Differential geometry of a parametric family of invertible linear systems—Riemannian metric, dual affine connections and divergence. *Math. Syst. Thory* **1987**, *20*, 53–82. [CrossRef]

21. Shima, H. *The Geometry of Hessian Structures*; World Scientific: Singapore, 2007.

22. Ohara, A.; Matsuzoe, H.; Amari, S.-I. A dually flat structure on the space of escort distributions. *J. Phys. Conf. Ser.* **2010**, *201*, 012012. [CrossRef]

23. Amari, S.-I.; Ohara, A.; Matsuzoe, H. Geometry of deformed exponential families: Invariant, dually-flat and conformal geometry. *Phys. A* **2012**, *391*, 4308–4319. [CrossRef]

24. Ohara, A.; Matsuzoe, H.; Amari, S.-I. Conformal geometry of escort probability and its applications. *Mod. Phys. Lett. B* **2012**, *26*, 1250063. [CrossRef]

25. Ohara A. Geometry of distributions associated with Tsallis statistics and properties of relative entropy minimization. *Phys. Lett. A* **2007**, *370*, 184–193. [CrossRef]

26. Naudts, J. *Generalized Thermostatistics*; Springer: Berlin, Germany, 2011.

27. Kaniadakis, G.; Scarfone, A.M. A new one-parameter deformation of the exponential function. *Phys. A* **2002**, *305*, 69–75. [CrossRef]

28. Kaniadakis, G.; Lissia, M.; Scarfone, A.M. Two-parameter deformations of logarithm, exponential, and entropy: A consistent framework for generalized statistical mechanics. *Phys. Rev. E* **2005**, *71*, 046128. [CrossRef] [PubMed]

29. Kaniadakis, G. Statistical mechanics in the context of special relativity. *Phys. Rev. E* **2002**, *66*, 056125. [CrossRef] [PubMed]

30. Kaniadakis, G. Statistical mechanics in the context of special relativity II. *Phys. Rev. E* **2005**, *72*, 036108. [CrossRef] [PubMed]

31. Kaniadakis, G.; Quarati, P.; Scarfone, A.M. Kinetical foundation of non-conventional statistics. *Phys. A* **2002**, *305*, 76–83. [CrossRef]

32. Biro, T.S.; Kaniadakis, G. Two generalizations of the Boltzmann equation. *Eur. Phys. J. B* **2006**, *50*, 3–6. [CrossRef]

33. Kaniadakis, G. Maximum entropy principle and power-law tailed distributions. *Eur. J. Phys. B* **2009**, *70*, 3–13. [CrossRef]

34. Kaniadakis, G.; Scarfone, A.M.; Sparavigna, A.; Wada, T. Composition law of kappa-entropy for statistically independent systems. *Phys. Rev. E* **2017**, *95*, 052112. [CrossRef] [PubMed]

35. Scarfone, A.M.; Wada, T. Canonical partition function for anomalous systems described by the κ-entropy. *Prof. Theor. Phys. Suppl.* **2006**, *162*, 45–52. [CrossRef]

36. Scarfone, A.M. Entropic forms and related algebras. *Entropy* **2013**, *15*, 624–649. [CrossRef]

37. Scarfone, A.M.; Wada, T. Thermodynamic equilibrium and its stability for microcanonical systems described by the Sharma-Taneja-Mittal entropy. *Phys. Rev. E* **2005**, *72*, 026123. [CrossRef] [PubMed]

38. Silva, R. The H-theorem in κ-statistics: Influence on the molecular chaos hypothesis. *Phys. Lett. A* **2006**, *352*, 17–20. [CrossRef]

39. Wada, T. Thermodynamic stability conditions for nonadditive composable entropies. *Contin. Mech. Thermodyn.* **2004**, *16*, 263–267. [CrossRef]

40. Wada, T.; Scarfone, A.M. Asymptotic solutions of a nonlinear diffusive equation in the framework of κ-generalized statistical mechanics. *Eur. Phys. J. B* **2009**, *70*, 65–71. [CrossRef]

41. Wada, T. A nonlinear drift which leads to kappa-generalized distributions. *Eur. Phys. J. B* **2010**, *73*, 287–291. [CrossRef]

42. Matsuzoe, H.; Wada, T. Deformed algebras and generalizations of independence on deformed exponential families. *Entropy* **2015**, *17*, 5729–5751. [CrossRef]

43. Wada, T.; Suyari, H. The κ-generalizations of Stirling approximation and multinominal coefficients. *Entropy* **2013**, *15*, 5144–5153. [CrossRef]

44. Scarfone, A.M. A maximal entropy distribution derivation of the Sharma-Taneja-Mittal entropic form. *Open Syst. Inf. Dyn.* **2018**, *25*, 1850002. [CrossRef]

45. Wada, T. Generalized log-likelihood functions and Bregman divergences. *J. Math. Phys.* **2009**, *50*, 113301. [CrossRef]

46. Pistone, G. κ-exponential models from the geometrical viewpoint. *Eur. Phys. J. B* **2009**, *70*, 29–37. [CrossRef]

47. Guo, L.; Du, J. The κ-parameter and κ-distribution in κ-deformed statistics for the systems in an external field. *Phys. Lett. A* **2007**, *362*, 368–370. [CrossRef]

48. Santos, A.P.; Silva, R.; Alcaniz, J.S.; Anselmo, D.H.A.L. Non-Gaussian effects on quantum entropies. *Phys. A* **2012**, *391*, 2182–2192. [CrossRef]

49. Kaniadakis, G. Theoretical foundations and mathematical formalism of the power-law tailed statistical distributions. *Entropy* **2013**, *15*, 3983–4010. [CrossRef]

50. Wada, T.; Suyari, H. κ-Generalization of Gauss's law of error. *Phys. Lett. A* **2006**, *348*, 89–93. [CrossRef]

51. Scarfone, A.M. κ-deformed Fourier transform. *Phys. A* **2017**, *480*, 63–78. [CrossRef]

52. Scarfone, A.M. On the κ-deformed cyclic functions and the generalized fourier series in the framework of the κ-algebra. *Entropy* **2015**, *17*, 2812–2833. [CrossRef]

53. Cravero, M.; Iabichino, G.; Kaniadakis G.; Miraldi, E. Scarfone, A.M. A κ-entropic approach to the analysis of the fracture problem. *Phys. A* **2004**, *340*, 410–417. [CrossRef]

54. Rossani, A.; Scarfone, A.M. Generalized kinetic equations for a system of interacting atoms and photons: Theory and simulations. *J. Phys. A Math. Theor.* **2004**, *37*, 4955–4975. [CrossRef]

55. Lapenta, G.; Markidis, S.; Kaniadakis, G. Computer experiments on the relaxation of collisionless plasmas. *J. Phys. A Math. Gen.* **2009**. [CrossRef]

56. Pereira, F.I.M.; Silva, R.; Alcaniz, J.S. Non-Gaussian statistics and the relativistic nuclear equation of state. *Nucl. Phys. A* **2009**, *828*, 136–148. [CrossRef]

57. Olemskoi, A.I.; Borysov, S.S.; Shuda, I.A. Statistical field theories deformed within different calculi. *Eur. Phys. J. B* **2010**, *77*, 219–231. [CrossRef]

58. Carvalho, J.C.; Silva, R.; do Nascimento, J.D., Jr.; Soares, B.B.; de Medeiros, J.R. Observational measurement of open stellar clusters: A test of Kaniadakis and Tsallis statistics. *Europhys. Lett.* **2010**, *91*, 69002. [CrossRef]

59. Clementi, F.; Gallegati, M.; Kaniadakis, G. A generalized statistical model for the size distribution of wealth. *J. Stat. Mech. Theory Exp.* **2012**. [CrossRef]

60. Clementi, F.; Gallegati, M.; Kaniadakis, G. κ-generalized statistical mechanics approach to income analysis. *J. Stat. Mech. Theory Exp.* **2009**. [CrossRef]

61. Moretto, E.; Pasquali, S.; Trivellato, B. A non-Gaussian option pricing model based on Kaniadakis exponential deformation. *Eur. Phys. J. B* **2017**, *90*, 179. [CrossRef]

62. Trivellato, B. Deformed exponentials and applications to finance. *Entropy* **2013**, *15*, 3471–3489. [CrossRef]

63. Tapiero, O.J. A maximum (non-extensive) entropy approach to equity options bid-ask spread. *Phys. A* **2013**, *392*, 3051–3060. [CrossRef]

64. Bertotti, M.L.; Modenese, G. Statistics of binary exchange of energy or money. *Entropy* **2017**, *19*, 465. [CrossRef]

65. Bertotti, M.L.; Modenese, G. Economic inequality and mobility in kinetic models for social sciences. *Eur. Phys. J. ST* **2016**, *225*, 1945–1958. [CrossRef]

66. Zhang, J.; Naudts, J. Information geometry under monotone embedding, Part I: Divergence functions. In Proceedings of the Geometric Science of Information, GSI2017, Paris, France, 7–9 November 2017; Nielson, F., Barbaresco, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 205–214.

67. Zhang, J.; Naudts, J. Information geometry under monotone embedding, Part II: Geometry. In Proceedings of the Geometric Science of Information, GSI2017, Paris, France, 7–9 November 2017; Nielson, F., Barbaresco, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 215–222.

68. Scarfone, A.M.; Wada, T. Legendre structure of κ-thermostatistics revisited in the framework of information geometry. *J. Phys. A Math. Theor.* **2014**, *47*, 275002. [CrossRef]

69. Wada, T.; Scarfone, A.M. Information geometry on the κ-thermostatistics. *Entropy* **2015**, *17*, 1204–1217. [CrossRef]

70. Wada, T.; Matsuzoe, H.; Scarfone, A.M. Dualistic hessian structures among the thermodynamic potentials in the κ-thermostatistics. *Entropy* **2015**, *17*, 7213–7229. [CrossRef]

71. Matsuzoe, H.; Henmi, M. Hessian structures and divergence functions on deformed exponential families. In *Geometric Theory of Information, Signals and Communication Technology*; Nielsen, F., Ed.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 57–80.

72. Matsuzoe, H.; Henmi, M. Hessian structures on deformed exponential families. In *Geometric Science of Information*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8085, pp. 275–282.

73. Matsuzoe, H. Hessian structures on deformed exponential families and their conformal structures. *Diff. Geom. Appl.* **2014**, *35*, 323–333. [CrossRef]

74. Matsuzoe, H. A sequence of escort distributions and generalizations of expectations on q-exponential family. *Entropy* **2017**, *19*, 7. [CrossRef]

75. Matsuzoe, H.; Scarfone, A.M.; Wada, T. A sequential structure of statistical manifolds on deformed exponential family. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2017; Volume 10589, pp. 223–230.

76. Scarfone, A.M.; Wada, T. Lie symmetries and related group-invariant solutions of a nonlinear Fokker-Planck equation based on the Sharma-Taneja-Mittal entropy. *Braz. J. Phys.* **2009**, *39*, 475–482. [CrossRef]

77. Scarfone, A.M.; Matsuzoe, H.; Wada, T. Consistency of the structure of Legendre transform in thermodynamics with the Kolmogorov–Nagumo average. *Phys. Lett. A* **2016**, *380*, 3022–3028. [CrossRef]

78. Beck, C.; Schögl, F. *Thermodynamics of Chaotic Systems: An Introduction*; Cambridge University Press: Cambridge, UK, 1993.

79. Kullback, S. *Information Theory and Statistics*; Wiley: New York, NY, USA, 1959.

80. Kerridge, D.F. Inaccuracy and inference. *J. R. Stat. Soc. Ser. B* **1961**, *23*, 184–194.

81. Brégman, L.M. The relaxation method to find the common point of convex sets and its applications to the solution of problems in convex programming. *USSR Comp. Math. Phys.* **1967**, *7*, 200–217. [CrossRef]

82. Eguchi, S. Geometry of minimum contrast. *Hiroshima Math. J.* **1992**, *22*, 631–647.