mathematics

MDPI

*Article*

# Kullback–Leibler Asymmetry and Information Homogeneity in Natural Exponential Families

**Shaul K. Bar-Lev**

Faculty of Industrial Engineering and Technology Management, HIT—Holon Institute of Technology, Holon 5810201, Israel; shaulb@hit.ac.il

**Abstract**

Natural exponential families admit two complementary structural descriptions: through their variance functions and through the information geometry generated by relative entropy and Fisher information. This paper establishes a direct connection between these descriptions. For a regular one-dimensional natural exponential family, we show that a fixed global direction of Kullback–Leibler asymmetry is equivalent to monotonicity of the variance function and, equivalently, to a fixed sign of the standardized skewness. In fact, a local ordering condition around each parameter point already forces the same ordering globally. We then reparametrize the family by Fisher–Rao arc length, which coincides with the classical variance-stabilizing coordinate, and study local spatial homogeneity of information divergence. Local homogeneity of the directed Kullback–Leibler divergence, of its antisymmetric part, of the magnitude of that asymmetry, or local radiality of Jeffreys divergence are shown to be equivalent. Each condition forces constant standardized skewness and consequently characterizes, up to a nonsingular affine transformation, precisely the normal location and fixed-shape gamma natural exponential families. Thus, weak local information conditions determine the global structure of the family. We further show that the global conclusions already follow from a sequential local ordering in the first theorem, and that position-independence of either the leading cubic KL-asymmetry coefficient or the quartic Jeffreys-excess coefficient alone yields the same normal/gamma rigidity.

**Keywords:** natural exponential family; variance function; Kullback–Leibler divergence; Fisher–Rao metric; Jeffreys divergence; skewness; information geometry

**MSC:** 62E10; 94A17

Check for updates

Academic Editor: Mark Kelbert

Received: 21 September 2026
Revised: 26 October 2026
Accepted: 30 October 2026
Published: 2 November 2026

**Copyright:** © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

Natural exponential families (NEFs) form one of the central structural classes of mathematical statistics. Their importance derives not only from the canonical form of their likelihoods but also from the fact that much of their probabilistic and statistical structure is encoded by a single function: the variance expressed as a function of the mean. This variance-function point of view has generated a substantial classification theory. Morris classified the six univariate NEFs with quadratic variance functions and subsequently developed their statistical theory [1,2]. Letac and Mora extended the classification program to cubic variance functions [3], while the broader role of variance functions in exponential dispersion models is developed in Jørgensen [4].

A related line of research has used the variance function as a characterization device for particular structural properties of NEFs. Bar-Lev and Enis linked reproducibility to

power variance functions [5]. The mean-value parametrization and the fact that an NEF is characterized by its variance function on its mean domain were revisited by Bar-Lev and Kokonendji [6]. More recently, Bar-Lev studied NEFs with bounded variance functions despite unbounded support [7]. These works illustrate a recurring principle: a property initially formulated in terms of distributions, support, convolution, or parametrization often becomes more transparent after translation into variance-function language.

A second line of work relevant to the present paper concerns variance-stabilizing transformations. Bar-Lev and Enis studied both the classical choice and the construction of classes of such transformations [8,9]. If an NEF has variance function $V$, the basic variance-stabilizing coordinate is obtained from

$$G'(m) = \frac{1}{\sqrt{V(m)}}.$$

The same coordinate will arise below from a different direction: it is precisely Fisher–Rao arc length when the family is parametrized by its mean. Thus, the classical statistical theory of variance stabilization and the intrinsic geometry of the family meet at the coordinate that is central to our rigidity results.

There are also several characterization results that single out small collections of classical exponential families. Harremoøes showed that exact renormalized saddlepoint approximation in one-dimensional exponential families selects the Gaussian, gamma, and inverse Gaussian families [10]. A recent result of Bar-Lev gives another characterization of the same triple through a converse problem for a mixed parametrization of a two-parameter exponential family [11]. The criterion studied in the present paper is different and more selective: information homogeneity on the intrinsic Fisher scale will retain the normal and gamma families but exclude the inverse Gaussian family.

The second body of literature underlying this paper is information geometry. Kullback and Leibler introduced relative entropy as a fundamental measure of statistical discrepancy [12]. The geometry of exponential families was developed by Efron [13], Eguchi [14], and Amari and Nagaoka [15]. In a regular exponential family, Kullback–Leibler divergence is a Bregman divergence associated with the cumulant function; the close connection between regular exponential families and Bregman divergences is discussed, for example, by Banerjee et al. [16]. Eguchi and Copas emphasized, among other aspects, the statistical interpretation and asymmetry of Kullback–Leibler divergence [17]. Matúš studied variance functions and information divergences simultaneously in exponential families near the boundary of the convex support [18].

Unlike a metric, Kullback–Leibler divergence is directional:

$$D_{\mathrm{KL}}(P\|Q) \neq D_{\mathrm{KL}}(Q\|P)$$

in general. Locally, its quadratic term is governed by Fisher information, whereas its third-order behavior contains the cubic information tensor. In a one-dimensional NEF, this third-order quantity is directly related to standardized skewness. This raises a natural question that is global rather than local: can the sign of a local third-order distributional characteristic determine the direction of relative entropy between every pair of distributions in the family?

Our first result gives an exact affirmative answer. For a regular nondegenerate one-dimensional NEF, the inequality

$$D(P_{\theta_1}\|P_{\theta_2}) \leq D(P_{\theta_2}\|P_{\theta_1}), \qquad \theta_1 < \theta_2,$$

holds for every pair if and only if the variance function is nondecreasing, equivalently if and only if standardized skewness is nonnegative throughout the family. More strongly,

it is enough to impose the corresponding ordering only for sufficiently close symmetric pairs around each parameter point. That weak local requirement forces monotonicity of the variance function and hence the global ordering for all pairs. The first main result is therefore a local-to-global directionality theorem.

Direction alone does not determine the amount of asymmetry. The canonical parameter is not an intrinsic unit of statistical separation, because Fisher information varies along the family. We therefore pass to Fisher–Rao arc length. In the mean parametrization this coordinate is

$$s(m) = \int^{m} \frac{du}{\sqrt{V(u)}},$$

which is exactly the classical variance-stabilizing coordinate. On this intrinsic scale, we ask a second question: if two pairs of distributions have the same sufficiently small Fisher–Rao separation, must they exhibit the same amount of Kullback–Leibler non-reversibility?

The use of Fisher–Rao arc length is essential to the formulation of this question. For two nearby distributions separated by a Fisher displacement $h$, the quadratic term of Kullback–Leibler divergence is universally $h^2/2$. Thus, equal Fisher displacement places different pairs on the same local information scale. Any remaining dependence of the cubic and higher-order terms on the base point then measures genuine variation in information asymmetry rather than variation in local Fisher information. In this sense, local translation homogeneity asks whether statistically equally separated pairs exhibit the same directed information discrepancy wherever they are located in the family; local radiality of Jeffreys divergence asks the analogous question for the symmetrized discrepancy.

Our principal rigidity theorem shows that even a local version of this homogeneity requirement is extremely restrictive. We prove that each of the following local properties is equivalent to the others: the directed Kullback–Leibler divergence depends only on signed Fisher displacement; its antisymmetric part depends only on signed Fisher displacement; the magnitude of that antisymmetric part depends only on Fisher–Rao distance; and Jeffreys divergence depends only on Fisher–Rao distance. Each condition forces the standardized skewness to be constant. Solving the resulting variance-function equation yields only a positive constant or the square of an affine function. Consequently, up to a nonsingular affine transformation, the family must be either a normal location NEF or a fixed-shape gamma NEF. Once this conclusion is reached, the initially local information homogeneity automatically becomes global. We also prove that Theorem 1 needs only a sequence of local symmetric comparisons approaching each point, and that the normal/gamma conclusion of Theorem 2 already follows if only the leading KL-asymmetry or Jeffreys-excess coefficient is position-independent.

This conclusion is related to, but distinct from, several recent geometric results. Nielsen and Okamura derive invariance of $f$-divergences and Fisher–Rao distance from transformation-group actions [19]. Their direction is from an assumed symmetry of the statistical model to an invariant divergence. Here, the logic is converse-like within the class of NEFs: an intrinsic divergence homogeneity condition forces normal-location or gamma-scale structure. Nielsen and Okamura have also studied dual-connection midpoint means and, in dimension one, the role of a constant cubic form [20]. Tojo and Yoshino recently showed that constant Hessian sectional curvature for an order-one exponential family is essentially equivalent to belonging to the full NEF-QVF class [21]. Our conditions are different and substantially more selective: information homogeneity on the Fisher scale leaves only the normal and affine-gamma cases. General upper bounds of Fisher–Rao distance by Jeffreys–Bregman quantities are also known for Hessian statistical manifolds [22]; we use only their elementary one-dimensional NEF specialization and focus on the associated equality rigidity.

To the best of our knowledge, no previous result establishes that local or global orientation of Kullback–Leibler divergence within a one-dimensional NEF is equivalent to monotonicity of its variance function and to a fixed sign of standardized skewness. More importantly, we are not aware of a characterization showing that local spatial homogeneity of Kullback–Leibler asymmetry, or local radiality of Jeffreys divergence, after Fisher–Rao normalization forces constant standardized skewness and hence, up to affine transformation, precisely the normal and fixed-shape gamma NEFs. The classical local relationship between divergence, Fisher information, and third-order information is not claimed as new, nor are general Bregman or Fisher–Jeffreys inequalities. The novelty lies in the local-to-global NEF consequences of these information-asymmetry conditions.

More precisely, the individual analytic ingredients used below—the Bregman form of Kullback–Leibler divergence in exponential families, its local Fisher expansion, and the relation between the third cumulant and the derivative of the variance function—are classical. The contribution is the exact set of equivalences and rigidity consequences obtained when these ingredients are combined with the one-dimensional NEF structure. In Theorem 1, a condition imposed only on sufficiently close symmetric pairs determines the Kullback–Leibler orientation for every pair in the family. In Theorem 2, the assumptions are exact local functional homogeneity conditions on the divergence functions themselves, not merely constancy of a Taylor coefficient; these local conditions determine the complete full NEF up to affine equivalence.

The paper is organized as follows. Section 2 introduces the NEF framework, variance functions, Kullback–Leibler and Jeffreys divergences, and Fisher–Rao parametrization. It also records the local divergence expansion that exposes standardized skewness. Section 3 establishes the local-to-global directionality theorem. Section 4 proves the information-homogeneity rigidity theorem and the normal/gamma classification, and records the Jeffreys–Fisher equality consequence. Section 5 examines the main classical NEFs and illustrates the difference between directionality and rigidity. Section 6 discusses the relation with the surrounding literature and the interpretation of the results. Section 7 concludes.

## 2. Preliminaries and Local Information Asymmetry

Let $\nu$ be a nondegenerate positive measure on $\mathbb{R}$ and let

$$L_\nu(\theta) = \int_{\mathbb{R}} e^{\theta x}\, \nu(dx).$$

We assume that

$$\Theta = \operatorname{int}\{\theta : L_\nu(\theta) < \infty\}$$

is a nonempty open interval. Throughout the paper, the NEF is the full regular family generated by $\nu$,

$$P_\theta(dx) = \exp\{\theta x - \kappa(\theta)\}\nu(dx), \qquad \theta \in \Theta,$$

where

$$\kappa(\theta) = \log L_\nu(\theta).$$

The cumulant function is analytic and strictly convex on $\Theta$. Hence,

$$m(\theta) = \kappa'(\theta)$$

is strictly increasing and

$$\kappa''(\theta) > 0.$$

Let $M = m(\Theta)$ be the mean domain and let $\psi = m^{-1}$. The variance function is

$$V(m) = \kappa''(\psi(m)), \qquad m \in M.$$

The pair consisting of the mean domain and variance function characterizes the NEF in the usual sense; see [1,4,6].

Differentiating $V(m(\theta)) = \kappa''(\theta)$ gives

$$\kappa'''(\theta) = V'(m)V(m).$$

Therefore, the standardized third cumulant, or skewness, is

$$\gamma(m) = \frac{\kappa'''(\theta)}{\{\kappa''(\theta)\}^{3/2}} = \frac{V'(m)}{\sqrt{V(m)}}. \tag{1}$$

Thus, the sign of skewness is exactly the sign of the derivative of the variance function.

For $\theta_1, \theta_2 \in \Theta$, direct calculation gives

$$D(P_{\theta_1} \| P_{\theta_2}) = \kappa(\theta_2) - \kappa(\theta_1) - (\theta_2 - \theta_1)\kappa'(\theta_1). \tag{2}$$

The Jeffreys divergence is

$$J(P_{\theta_1}, P_{\theta_2}) = D(P_{\theta_1} \| P_{\theta_2}) + D(P_{\theta_2} \| P_{\theta_1}),$$

and hence

$$J(P_{\theta_1}, P_{\theta_2}) = (\theta_2 - \theta_1)\{m(\theta_2) - m(\theta_1)\}. \tag{3}$$

The Fisher information for the canonical parameter is

$$I(\theta) = \kappa''(\theta) = V(m(\theta)).$$

We introduce increasing Fisher–Rao arc length by

$$s(\theta) = \int_{\theta_0}^{\theta} \sqrt{\kappa''(u)}\, du. \tag{4}$$

Because $dm = V(m)d\theta$,

$$ds = \sqrt{V(m)}\, d\theta = \frac{dm}{\sqrt{V(m)}}. \tag{5}$$

Thus, $s$ is, up to an additive constant, the classical variance-stabilizing coordinate studied in [8,9]. Let $S = s(\Theta)$, define

$$g(s) = \sqrt{V(m(s))},$$

and put

$$q(s) = \log g(s).$$

Then

$$\frac{dm}{ds} = g(s), \qquad \frac{d\theta}{ds} = \frac{1}{g(s)}, \tag{6}$$

and, by (1),

$$q'(s) = \frac{1}{2}\frac{V'(m)}{\sqrt{V(m)}} = \frac{\gamma(s)}{2}. \tag{7}$$

The classical information-geometric theory implies that the first asymmetric term of a smooth divergence is third order [13–15]. For later use, we record its explicit form in Fisher coordinates for a one-dimensional NEF.

In the following proposition, we express the leading local Kullback–Leibler asymmetry directly in terms of standardized skewness.

**Proposition 1.** *For a regular nondegenerate one-dimensional NEF parametrized by Fisher–Rao arc length,*

$$D(P_s\|P_{s+h}) = \frac{h^2}{2} - \frac{\gamma(s)}{12}h^3 + O(h^4), \tag{8}$$

*and*

$$D(P_{s+h}\|P_s) = \frac{h^2}{2} + \frac{\gamma(s)}{12}h^3 + O(h^4). \tag{9}$$

*Consequently,*

$$D(P_s\|P_{s+h}) - D(P_{s+h}\|P_s) = -\frac{\gamma(s)}{6}h^3 + O(h^4), \tag{10}$$

*and*

$$\gamma(s) = 6\lim_{h\to 0}\frac{D(P_{s+h}\|P_s) - D(P_s\|P_{s+h})}{h^3}. \tag{11}$$

**Proof.** From (2) and (6), differentiation with respect to the second Fisher coordinate gives

$$\frac{\partial}{\partial t}D(P_s\|P_t) = \frac{m(t) - m(s)}{g(t)}.$$

Therefore

$$D(P_s\|P_{s+h}) = \int_0^h \frac{\int_0^r g(s+u)\,du}{g(s+r)}\,dr.$$

Taylor expansion of the numerator and denominator at $s$ gives

$$D(P_s\|P_{s+h}) = \frac{h^2}{2} - \frac{g'(s)}{6g(s)}h^3 + O(h^4).$$

Since $g'/g = q' = \gamma/2$ by (7), (8) follows. Reversing the arguments gives (9). Indeed, the reversed expansion may be written with base point $s+h$; since $\gamma(s+h) = \gamma(s) + O(h)$, replacing $\gamma(s+h)$ by $\gamma(s)$ changes the cubic term only by $O(h^4)$. Subtraction and division by $h^3$ give (10) and (11). □

For Jeffreys divergence, (3) and (6) yield

$$J(P_s, P_{s+h}) = \left(\int_0^h g(s+u)\,du\right)\left(\int_0^h \frac{du}{g(s+u)}\right). \tag{12}$$

To make the fourth-order coefficient explicit, put

$$a = \frac{g'(s)}{g(s)}, \qquad b = \frac{g''(s)}{g(s)}.$$

Then,

$$\int_0^h g(s+u)\,du = g(s)\left(h + \frac{a}{2}h^2 + \frac{b}{6}h^3 + O(h^4)\right),$$

whereas

$$\int_0^h \frac{du}{g(s+u)} = \frac{1}{g(s)}\left(h - \frac{a}{2}h^2 + \frac{a^2 - b/2}{3}h^3 + O(h^4)\right).$$

Multiplication shows explicitly that the terms involving $b$ cancel:

$$J(P_s, P_{s+h}) = h^2 + \left\{\frac{b}{6} + \frac{a^2 - b/2}{3} - \frac{a^2}{4}\right\}h^4 + O(h^5) = h^2 + \frac{q'(s)^2}{12}h^4 + O(h^5) = h^2 + \frac{\gamma(s)^2}{48}h^4 + O(h^5). \tag{13}$$

These two local formulas, (10) and (13), are the analytic mechanism behind the rigidity theorem below.

## 3. A Local-to-Global Directionality Theorem

The local expansion (10) shows that skewness determines the infinitesimal direction of Kullback–Leibler asymmetry. We now show that, within an NEF, this local sign information propagates globally.

In the following theorem, we establish that it is enough to impose Kullback–Leibler ordering for sufficiently close symmetric pairs around each canonical parameter value; this local condition is equivalent to monotonicity of the variance function and hence to a global ordering for every pair in the family.

**Theorem 1.** *Let* $\{P_\theta : \theta \in \Theta\}$ *be a regular nondegenerate one-dimensional NEF. The following statements are equivalent.*

*(i) For every* $\theta_1 < \theta_2$,

$$D(P_{\theta_1} \| P_{\theta_2}) \le D(P_{\theta_2} \| P_{\theta_1}). \tag{14}$$

*(ii) For every* $c \in \Theta$, *there exists* $\varepsilon_c > 0$ *such that, whenever* $0 < h < \varepsilon_c$ *and* $c - h, c + h \in \Theta$,

$$D(P_{c-h} \| P_{c+h}) \le D(P_{c+h} \| P_{c-h}).$$

*(iii)* $\kappa''(\theta)$ *is nondecreasing on* $\Theta$.
*(iv)* $V(m)$ *is nondecreasing on* $M$.
*(v)* $\gamma(m) \ge 0$ *for every* $m \in M$.

**Proof.** Statement (i) immediately implies (ii). We prove that (ii) implies (iii). From (2), Taylor's formula with integral remainder gives, for $\theta_1 < \theta_2$,

$$D(P_{\theta_1} \| P_{\theta_2}) = \int_{\theta_1}^{\theta_2} (\theta_2 - t)\kappa''(t)\, dt, \tag{15}$$

and

$$D(P_{\theta_2} \| P_{\theta_1}) = \int_{\theta_1}^{\theta_2} (t - \theta_1)\kappa''(t)\, dt. \tag{16}$$

For $\theta_1 = c - h$ and $\theta_2 = c + h$, subtraction yields

$$D(P_{c+h} \| P_{c-h}) - D(P_{c-h} \| P_{c+h}) = 2 \int_0^h u\{\kappa''(c + u) - \kappa''(c - u)\}\, du. \tag{17}$$

As $h \downarrow 0$,

$$D(P_{c+h} \| P_{c-h}) - D(P_{c-h} \| P_{c+h}) = \frac{4}{3}\kappa'''(c)h^3 + o(h^3). \tag{18}$$

Condition (ii) therefore implies $\kappa'''(c) \ge 0$ for every $c \in \Theta$, and hence $\kappa''$ is nondecreasing. Thus, (ii) implies (iii).

Suppose next that (iii) holds. The integrand in (17) is nonnegative for every $u > 0$ for which the points belong to $\Theta$. Therefore, (14) holds for every $\theta_1 < \theta_2$, so (iii) implies (i).

Since $dm/d\theta = V(m) > 0$, the canonical parameter and the mean induce the same ordering. Hence, $\kappa''(\theta) = V(m(\theta))$ is nondecreasing in $\theta$ if and only if $V(m)$ is nondecreasing in $m$. Thus, (iii) and (iv) are equivalent. Finally, (1) and positivity of $V$ show that (iv) and (v) are equivalent. □

The reversed version follows by changing all inequalities: $V$ is nonincreasing if and only if skewness is nonpositive if and only if the reverse Kullback–Leibler ordering holds globally. Thus, Theorem 1 gives a complete sign characterization.

In the following corollary, we record the strict form of the result when the family is nonnormal.

**Corollary 1.** *Assume the equivalent conditions of Theorem 1. If the variance function is not constant, then for every* $\theta_1 < \theta_2$,

$$D(P_{\theta_1} \| P_{\theta_2}) < D(P_{\theta_2} \| P_{\theta_1}).$$

**Proof.** Under monotonicity, the integrand in (17) is nonnegative. Equality for one distinct pair would imply equality of $\kappa''$ at the two endpoints of a nontrivial interval. A monotone function with equal endpoint values is constant on that interval. Since $\kappa''$ is analytic, it would then be constant on the whole connected canonical parameter domain, contrary to the assumption. □

**Corollary 2** (Sequential local criterion)**.** *In Theorem 1 (ii), it is enough to require the Kullback–Leibler ordering along a sequence tending to each parameter point. More precisely, suppose that for every* $c \in \Theta$ *there exists a sequence* $h_n \downarrow 0$, *with* $c - h_n, c + h_n \in \Theta$, *such that*

$$D(P_{c-h_n} \| P_{c+h_n}) \leq D(P_{c+h_n} \| P_{c-h_n}) \qquad \text{for every } n.$$

*Then, the equivalent conclusions of Theorem 1 hold:* $\kappa''$ *and* $V$ *are nondecreasing,* $\gamma \geq 0$, *and the same Kullback–Leibler orientation holds for every ordered pair in the family.*

**Proof.** Using (18) along the sequence $h_n$ gives

$$0 \leq \frac{4}{3}\kappa'''(c)h_n^3 + o(h_n^3).$$

After division by $h_n^3$ and passage to the limit, $\kappa'''(c) \geq 0$. Since $c$ is arbitrary, the remainder follows from Theorem 1. □

The orientation in Theorem 1 is tied to the increasing canonical/mean ordering. If the observation variable is reflected, $Y = -X$, Kullback–Leibler divergence itself is unchanged under this one-to-one transformation, but the induced mean ordering is reversed and standardized skewness changes sign. Accordingly, the displayed orientation is reversed when it is expressed relative to the increasing mean of the reflected family.

The preceding theorem separates two issues. A fixed sign of skewness controls the *orientation* of Kullback–Leibler asymmetry, but it does not make the amount of asymmetry independent of position in the family. That stronger property is considered next.

## 4. Information Homogeneity and Rigidity

We first formalize the local homogeneity conditions used below.

**Definition 1.** *Let* $S$ *be the Fisher-coordinate interval. A function* $B(s,t)$ *is called pointwise locally translation-homogeneous if, for every* $s_0 \in S$, *there exists an open interval* $I_{s_0} \subset S$ *containing* $s_0$ *and a one-variable function* $b_{s_0}$ *such that*

$$B(s,t) = b_{s_0}(t-s), \qquad s,t \in I_{s_0}.$$

*It is called pointwise locally radial if, for every* $s_0 \in S$, *there exists such an interval* $I_{s_0}$ *and a function* $b_{s_0}$ *for which*

$$B(s,t) = b_{s_0}(|t-s|), \qquad s,t \in I_{s_0}.$$

The definition is deliberately weak: both the interval and the representing one-variable function may vary with $s_0$, and no behavior is assumed for widely separated members of the family. The main result shows that these pointwise local requirements already determine the entire NEF.

It is worth stressing that the hypotheses below are functional hypotheses: on each local interval, the full divergence (or its indicated symmetric/antisymmetric component) must be a function only of displacement. They are therefore stronger functional statements than constancy of a single leading Taylor coefficient. The proof uses the local identities to extract constant standardized skewness, after which the NEF structure propagates the pointwise local condition to a global classification.

For the directed Kullback–Leibler divergence, the following exact identity complements the local Taylor expansion and will be useful in the proof of the rigidity theorem.

**Lemma 1.** *For Fisher coordinates $s, t \in S$,*

$$\frac{\partial^2}{\partial s\, \partial t} D(P_s \| P_t) = -\frac{g(s)}{g(t)}. \tag{19}$$

**Proof.** Differentiating (2) with respect to the second Fisher coordinate and using (6) gives

$$\frac{\partial}{\partial t} D(P_s \| P_t) = \frac{m(t) - m(s)}{g(t)}.$$

Differentiation with respect to $s$ and the identity $dm(s)/ds = g(s)$ yields (19). □

In the following theorem, we show that even pointwise local spatial homogeneity of Kullback–Leibler or Jeffreys divergence on the Fisher scale is rigid: a condition imposed separately in arbitrarily small neighborhoods of each Fisher point forces the global family to be normal or gamma.

**Theorem 2.** *Let $\{P_s : s \in S\}$ be a full regular nondegenerate one-dimensional NEF parametrized by increasing Fisher–Rao arc length, and define*

$$A(s,t) = D(P_s \| P_t) - D(P_t \| P_s).$$

*The following statements are equivalent.*

*(i) $D(P_s \| P_t)$ is pointwise locally translation-homogeneous.*
*(ii) $A(s,t)$ is pointwise locally translation-homogeneous.*
*(iii) $|A(s,t)|$ is pointwise locally radial.*
*(iv) $J(P_s, P_t)$ is pointwise locally radial.*
*(v) The standardized skewness is constant on the family.*
*(vi) The variance function has one of the forms*

$$V(m) = \sigma^2, \qquad \sigma^2 > 0, \tag{20}$$

*or*

$$V(m) = a^2 (m - b)^2, \qquad a \neq 0. \tag{21}$$

*(vii) Up to a nonsingular affine transformation of the observation, including reflection when required, the NEF is either a normal location family or a fixed-shape gamma family.*

*Moreover, whenever these equivalent conditions hold, the local homogeneity in (i)–(iv) extends to every admissible pair in $S$.*

**Proof.** Suppose first that (i) holds. Fix $s_0 \in S$ and choose $I_{s_0}$ and $F_{s_0}$ from the definition. For $s, t \in I_{s_0}$,

$$D(P_s \| P_t) = F_{s_0}(t - s).$$

Since the left-hand side is smooth, $F_{s_0}$ is smooth near zero. By Lemma 1,

$$\frac{g(s)}{g(t)} = F''_{s_0}(t-s).$$

Writing $t = s + h$ and taking logarithms gives

$$q(s+h) - q(s) = -\log F''_{s_0}(h),$$

for $s, s+h \in I_{s_0}$ and sufficiently small $h$. Differentiating with respect to $h$ at $h = 0$ shows that $q'(s)$ is constant on $I_{s_0}$. By (7), $\gamma(s) = 2q'(s)$ is therefore locally constant. Because $S$ is connected, a locally constant function on $S$ is constant. Thus, (i) implies (v).

Suppose (ii) holds. Fix $s_0$ and an interval $I_{s_0}$ on which

$$A(s,t) = G_{s_0}(t-s).$$

For every $s \in I_{s_0}$, (11) gives

$$\gamma(s) = -6 \lim_{h\to 0} \frac{G_{s_0}(h)}{h^3}.$$

The right-hand side is independent of $s$, so $\gamma$ is constant on $I_{s_0}$. Connectedness of $S$ again yields global constancy. Hence, (ii) implies (v).

Suppose (iii) holds. Fix $s_0$ and write

$$|A(s,t)| = G_{0,s_0}(|t-s|), \qquad s,t \in I_{s_0}.$$

Then, (10) implies

$$|\gamma(s)| = 6 \lim_{h\to 0} \frac{G_{0,s_0}(|h|)}{|h|^3}, \qquad s \in I_{s_0}.$$

Thus, $|\gamma|$ is constant on $I_{s_0}$. Since $I_{s_0}$ is connected and $\gamma$ is continuous, $\gamma$ itself is constant there: if the common absolute value is positive, its sign cannot change, and if it is zero, then $\gamma$ vanishes. Hence, $\gamma$ is locally constant on $S$ and therefore globally constant. This proves (iii)⇒(v).

Suppose (iv) holds. Fix $s_0$ and write

$$J(P_s, P_t) = H_{s_0}(|t-s|), \qquad s,t \in I_{s_0}.$$

By (13),

$$\gamma(s)^2 = 48 \lim_{h\to 0} \frac{H_{s_0}(|h|) - h^2}{h^4}, \qquad s \in I_{s_0}.$$

Thus, $|\gamma|$ is constant on $I_{s_0}$, and the same continuity argument shows that $\gamma$ is constant there. Hence, $\gamma$ is globally constant on the connected interval $S$. Therefore, (iv) implies (v).

Assume now that (v) holds and write $\gamma(m) = \gamma_0$. By (1),

$$\frac{V'(m)}{\sqrt{V(m)}} = \gamma_0.$$

Therefore,

$$\frac{d}{dm}\sqrt{V(m)} = \frac{\gamma_0}{2},$$

so

$$\sqrt{V(m)} = \frac{\gamma_0}{2} m + c. \tag{22}$$

If $\gamma_0 = 0$, $V$ is constant and (20) holds. If $\gamma_0 \neq 0$, (22) is equivalent to (21). Thus, (v) implies (vi). Conversely, either variance-function form in (vi) gives constant standardized skewness by (1); hence, (v) and (vi) are equivalent.

We next identify the corresponding full NEFs. If $V(m) = \sigma^2$, then $\kappa''(\theta) = \sigma^2$ and

$$\kappa(\theta) = c_0 + c_1\theta + \frac{\sigma^2}{2}\theta^2,$$

which is the cumulant function of a normal location family up to translation. Uniqueness of the Laplace transform identifies the generating measure, and the fullness assumption recovers the corresponding maximal canonical family.

Suppose instead that $V(m) = a^2(m-b)^2$. Translate the observation by $b$. On the connected mean domain, $m - b$ has a fixed sign. After a possible reflection, we may assume $m > 0$ and

$$V(m) = a^2m^2.$$

Put $\alpha = 1/a^2$. Since

$$\frac{d\theta}{dm} = \frac{1}{V(m)} = \frac{\alpha}{m^2},$$

integration gives

$$\theta = c - \frac{\alpha}{m},$$

and hence

$$m = \frac{\alpha}{c-\theta}.$$

A further integration yields

$$\kappa(\theta) = C - \alpha\log(c-\theta),$$

apart from the linear term associated with the removed translation. Hence, the Laplace transform agrees on a nonempty open interval with that of a fixed-shape gamma law. Uniqueness of the Laplace transform identifies the generating measure (up to the already removed affine change) with the corresponding gamma measure. Since the NEF is full, its canonical parameter domain is the interior of the complete domain of that Laplace transform, namely, the maximal half-line. Thus, the full family is the affine image, with possible reflection, of a fixed-shape gamma NEF. This proves (vi) implies (vii). Conversely, normal and fixed-shape gamma NEFs have, respectively, constant zero and constant nonzero standardized skewness, so (vii) implies (v).

It remains to show that constant skewness gives the local homogeneity properties and indeed makes them global. By (7),

$$q'(s) = \frac{\gamma_0}{2}.$$

Put

$$a = \frac{\gamma_0}{2}.$$

If $a = 0$, $g$ is constant and direct integration gives

$$D(P_s\|P_{s+h}) = \frac{h^2}{2}. \tag{23}$$

If $a \neq 0$, then

$$g(s) = Ce^{as}$$

for some $C > 0$. Substitution into the integral representation used in Proposition 1 gives

$$D(P_s\|P_{s+h}) = \frac{e^{-ah} - 1 + ah}{a^2}. \tag{24}$$

The right-hand side depends only on $h$ for every admissible pair. Therefore, (v) implies (i), and consequently (ii) and (iii). Addition of the formula with $h$ replaced by $-h$ shows that Jeffreys divergence depends only on $|h|$, so (iv) also holds. The same calculation proves the final global-extension assertion. □

Theorem 2 is the central rigidity statement. It is important that the assumptions in (i)–(iv) are pointwise local: neither a single common neighborhood of the diagonal nor a single common representing function is assumed in advance. Information homogeneity around arbitrarily small neighborhoods of individual Fisher points nevertheless determines the entire variance function and therefore the entire full NEF.

**Corollary 3** (Coefficient-level rigidity)**.** *Let* $\{P_s : s \in S\}$ *be as in Theorem 2. Define the local Kullback–Leibler asymmetry coefficient and the local Jeffreys excess coefficient by*

$$C_A(s) = \lim_{h\to 0} \frac{D(P_{s+h}\|P_s) - D(P_s\|P_{s+h})}{h^3}$$

*and*

$$C_J(s) = \lim_{h\to 0} \frac{J(P_s, P_{s+h}) - h^2}{h^4}.$$

*If either* $C_A(s)$ *is independent of* $s$ *or* $C_J(s)$ *is independent of* $s$*, then the standardized skewness is constant and hence, up to a nonsingular affine transformation, the full NEF is normal or fixed-shape gamma. Conversely, the normal and fixed-shape gamma NEFs satisfy both coefficient-homogeneity properties.*

**Proof.** By (11), $C_A(s) = \gamma(s)/6$. Hence, constancy of $C_A$ is exactly constancy of standardized skewness. By (13), $C_J(s) = \gamma(s)^2/48$. If $C_J$ is constant, then $|\gamma|$ is constant; continuity on the connected interval $S$ implies that $\gamma$ itself is constant. The classification and converse now follow from Theorem 2. □

In the following corollary, we give the divergence functions explicitly and display the normal family as the zero-skewness limit of the gamma geometry.

**Corollary 4.** *Under the equivalent conditions of Theorem 2, let* $h = t - s$*. If the family is normal, then*

$$D(P_s\|P_{s+h}) = \frac{h^2}{2},$$

$$A(s, s+h) = 0,$$

*and*

$$J(P_s, P_{s+h}) = h^2.$$

*If the constant standardized skewness is nonzero and* $a = \gamma_0/2$*, then*

$$D(P_s\|P_{s+h}) = \frac{e^{-ah} - 1 + ah}{a^2}, \tag{25}$$

$$A(s, s+h) = \frac{2\{ah - \sinh(ah)\}}{a^2}, \tag{26}$$

*and*

$$J(P_s, P_{s+h}) = \frac{2\{\cosh(ah) - 1\}}{a^2}. \tag{27}$$

*As* $a \to 0$*, these expressions converge to the normal formulas.*

**Proof.** Equation (25) is (24). Replacing $h$ by $-h$ gives the reversed divergence. Subtraction and addition yield (26) and (27). The limits as $a \to 0$ follow from the Taylor series of the exponential and hyperbolic cosine. □

General Jeffreys–Bregman upper bounds on Fisher–Rao distance are known in Hessian statistical geometry [22]. In one dimension, the NEF specialization follows immediately from (12). The next proposition is included because its equality case provides a useful normal rigidity statement.

In the following proposition, we record the one-dimensional Jeffreys–Fisher inequality and show that equality at even one nontrivial pair forces normality.

The inequality itself is a direct Cauchy–Schwarz consequence and is not claimed as new. The NEF-specific point used here is the rigidity of equality: equality for a single distinct pair forces constancy on an interval, and analyticity then propagates that constancy to the full connected family.

**Proposition 2.** *For any two members $P_s$ and $P_t$ of a regular nondegenerate one-dimensional NEF,*

$$J(P_s, P_t) \geq d_{\mathrm{FR}}(P_s, P_t)^2. \tag{28}$$

*If equality holds for one pair with $s \neq t$, then the NEF is a normal location family up to a nonsingular affine transformation.*

**Proof.** Assume $s < t$. From (3) and (6),

$$J(P_s, P_t) = \left(\int_s^t g(u)\, du\right)\left(\int_s^t \frac{du}{g(u)}\right).$$

By the Cauchy–Schwarz inequality,

$$(t-s)^2 \leq \left(\int_s^t g(u)\, du\right)\left(\int_s^t \frac{du}{g(u)}\right).$$

Since $s$ is Fisher–Rao arc length,

$$d_{\mathrm{FR}}(P_s, P_t) = t - s,$$

and (28) follows. Equality in Cauchy–Schwarz holds if and only if $g$ is constant on $[s, t]$. Analyticity then forces $g$, and hence $V$, to be constant on the connected domain. The NEF is therefore normal. □

## 5. Classical Examples

The examples below show why the directionality theorem and the rigidity theorem are genuinely different. Several important NEFs have a fixed global Kullback–Leibler direction but fail information homogeneity because their standardized skewness is not constant.

### *5.1. Normal Family*

For a normal location NEF with fixed variance $\sigma^2$,

$$V(m) = \sigma^2, \qquad \gamma(m) = 0.$$

The Kullback–Leibler divergence is symmetric and the family satisfies all conditions of Theorem 2. It is the homogeneous reversible case.

### *5.2. Gamma Family*

For a fixed-shape gamma NEF with shape parameter $\alpha$,

$$V(m) = \frac{m^2}{\alpha}$$

and

$$\gamma(m) = \frac{2}{\sqrt{\alpha}}.$$

The skewness is positive and constant. The gamma family therefore satisfies both the global directionality theorem and the full rigidity theorem. It is the homogeneous nonreversible case: after Fisher normalization, Kullback–Leibler asymmetry is nonzero but independent of position.

### *5.3. Poisson Family*

For the Poisson NEF,

$$V(m) = m, \qquad m > 0,$$

so

$$\gamma(m) = m^{-1/2} > 0.$$

Theorem 1 therefore yields a strict global Kullback–Leibler direction, but the skewness is not constant. Hence, Poisson fails the information-homogeneity conditions of Theorem 2.

The contrast with the gamma family is particularly transparent at a fixed Fisher displacement $h$. For gamma, Corollary 4 gives

$$A(s, s+h) = \frac{2\{ah - \sinh(ah)\}}{a^2},$$

which is independent of the base point $s$. For Poisson, the Fisher coordinate may be chosen as $s = 2\sqrt{m}$, so that $m = s^2/4$ and $\gamma(s) = 2/s$. Hence, from (10),

$$A(s, s+h) = -\frac{h^3}{3s} + O(h^4).$$

Thus, Poisson has the same KL orientation throughout the family, but the magnitude of the local asymmetry varies with position; gamma has both fixed orientation and Fisher-scale homogeneity.

### *5.4. Negative Binomial Family*

For a negative binomial NEF with fixed shape parameter $r > 0$,

$$V(m) = m + \frac{m^2}{r}, \qquad m > 0.$$

Thus,

$$\gamma(m) = \frac{1 + 2m/r}{\sqrt{m + m^2/r}},$$

which is positive but nonconstant. The family has a fixed global Kullback–Leibler direction but does not satisfy the rigidity theorem.

### *5.5. Inverse Gaussian Family*

For the inverse Gaussian NEF in the usual fixed-dispersion normalization,

$$V(m) = \frac{m^3}{\lambda}, \qquad m > 0,$$

and therefore

$$\gamma(m) = 3\sqrt{\frac{m}{\lambda}}.$$

Again, the variance function is increasing, so the global Kullback–Leibler direction is fixed, but standardized skewness is not constant. This comparison is especially informative because the inverse Gaussian joins normal and gamma in the exact-saddlepoint characterization of Harremoøes [10] and in the mixed-parametrization characterization of Bar-Lev [11]; the information-homogeneity criterion developed here separates it from the other two.

*5.6. Binomial Family*

For a binomial NEF with size parameter $n$,

$$V(m) = m\left(1 - \frac{m}{n}\right), \qquad 0 < m < n.$$

Its derivative changes sign at $m = n/2$. Consequently standardized skewness changes sign and no single Kullback–Leibler orientation holds on the full family. The direction reverses as the mean passes through the midpoint.

*5.7. Generalized Hyperbolic Secant Family*

In a standard normalization of the generalized hyperbolic secant NEF,

$$V(m) = 1 + m^2, \qquad m \in \mathbb{R},$$

so

$$\gamma(m) = \frac{2m}{\sqrt{1 + m^2}}.$$

The skewness changes sign at zero. Thus, the family belongs to the Morris NEF-QVF class [1,2] but has neither a single global Kullback–Leibler direction nor Fisher-scale information homogeneity. This provides a simple illustration of the distinction between quadratic variance and the much stronger perfect-square condition in Theorem 2.

For convenience, the examples may be summarized as follows:

| **Family** | **Variance Function** | **KL Orientation** | **Fisher Homogeneity** |
|---|---|---|---|
| Normal | $\sigma^2$ | symmetric | yes |
| Gamma | $m^2/\alpha$ | fixed, positive-skew direction | yes |
| Poisson | $m$ | fixed, positive-skew direction | no |
| Negative binomial | $m + m^2/r$ | fixed, positive-skew direction | no |
| Inverse Gaussian | $m^3/\lambda$ | fixed, positive-skew direction | no |
| Binomial | $m(1 - m/n)$ | changes at $m = n/2$ | no |
| Hyperbolic secant | $1 + m^2$ | changes at $m = 0$ | no |

These examples yield a useful hierarchy. Poisson, negative binomial, and inverse Gaussian satisfy the directionality theorem but not the rigidity theorem. Binomial and generalized hyperbolic secant fail even global directionality on their full mean domains. Only normal and gamma satisfy the homogeneity theorem.

## 6. Discussion

The main results connect three descriptions of a one-dimensional NEF that are usually treated separately:

variance-function structure, standardized skewness, relative-entropy geometry.

Equation (1) is the local bridge between the first two. Theorem 1 shows that this local identity has a global consequence: a fixed sign of standardized skewness is equivalent to a coherent Kullback–Leibler orientation across every pair of distributions in the family. The local assumption in Theorem 1 is particularly noteworthy. One need not compare widely separated distributions. The sign of the asymmetry for sufficiently close symmetric pairs already forces monotonicity of the variance function, which then propagates the direction globally.

Theorem 2 is stronger. Fisher–Rao parametrization removes the local scale variation caused by changing information along the family. In this coordinate, the normalized third-order information is simply $2q'(s) = \gamma(s)$. Requiring the amount of Kullback–Leibler

non-reversibility to be locally independent of position forces this quantity to be constant. The resulting differential equation for the variance function is elementary once it has been reached, but the substantive point is that it is forced by a purely information-geometric homogeneity condition.

The theorem may be interpreted as a local-to-global rigidity principle. Local translation homogeneity of the directed divergence, local translation homogeneity of its antisymmetric component, local radiality of the magnitude of that asymmetry, and local radiality of Jeffreys divergence all produce the same conclusion. Once constant skewness is obtained, the entire NEF is fixed up to affine equivalence, and the local invariance automatically extends globally.

The normal and gamma families admit a particularly simple interpretation in this framework. Normality corresponds to zero standardized skewness and hence to homogeneous *reversible* information geometry:

$$D(P_s \| P_{s+h}) = D(P_{s+h} \| P_s) = \frac{h^2}{2}.$$

Gamma corresponds to constant nonzero standardized skewness and hence to homogeneous *nonreversible* information geometry. In both cases, the divergence is translation-homogeneous in Fisher coordinate, but only the normal case is symmetric.

The variance-stabilizing literature gives the Fisher coordinate an additional statistical interpretation. The transformation (5), studied in particular by Bar-Lev and Enis [8,9], was introduced to remove first-order heteroscedasticity. Here, the same transformation removes local Fisher scaling and exposes whether the remaining information asymmetry varies with position. The rigidity theorem therefore identifies precisely those NEFs whose relative-entropy geometry becomes homogeneous after variance stabilization.

The result should also be distinguished from the classical NEF-QVF classification. Morris's six families [1,2] are characterized by quadratic variance functions. Theorem 2 retains only a constant variance function or the square of an affine function. Thus, quadraticity itself is far too weak. This distinction is reinforced by the recent result of Tojo and Yoshino [21]: constant Hessian sectional curvature leads essentially to the full NEF-QVF class, whereas local KL/Jeffreys homogeneity on the Fisher scale selects only normal and gamma.

There is also a useful contrast with transformation-model invariance. Nielsen and Okamura [19] derive divergence invariance from an assumed group action. In the present setting, no location or scale group is postulated. Instead, intrinsic divergence homogeneity is assumed locally, and the normal-location or gamma-scale structure emerges as a conclusion. Their more recent midpoint-rigidity work [20] is closer in spirit because it links a local geometric invariance condition to constancy of a cubic form, but the invariant, functional equation, and resulting statistical statement are different.

At the time of this revision, the Nielsen–Okamura works cited here remain arXiv preprints, as does the Tojo–Yoshino constant-Hessian-curvature work. The comparisons above are therefore stated only at the level of the specific results currently available in those manuscripts: group-induced divergence invariance, midpoint/constant-cubic-form rigidity, and the equivalence between constant Hessian sectional curvature and the NEF-QVF class, respectively.

Proposition 1 also provides a direct information-theoretic recovery formula for standardized skewness, (11). This is not presented as a replacement for moment-based skewness, but it clarifies the role of skewness geometrically: on the Fisher scale, standardized skewness is exactly the leading infinitesimal coefficient of Kullback–Leibler nonreversibility. Theorem 1 then says that a fixed sign of this coefficient controls global orientation, while Theorem 2 says that constancy of its magnitude and sign is equivalent to full information homogeneity.

Finally, Proposition 2 places the normal family at an equality boundary. The general comparison of Fisher–Rao and Jeffreys–Bregman quantities is known in broader Hessian geometry [22]; within a regular one-dimensional NEF, equality at one distinct pair already implies that the variance function is constant and hence that the family is normal. This equality rigidity is consistent with the broader message of the paper: apparently weak information conditions become highly restrictive once combined with NEF structure.

The scope of the present rigidity theorem is deliberately one-dimensional. In one dimension, Fisher–Rao arc length supplies a canonical scalar coordinate and a signed displacement between nearby models. In higher dimensions, Fisher information is a Riemannian metric, the third-order information object is tensorial, and equal Fisher distance does not by itself specify a direction or a unique comparison path. A multidimensional analogue would therefore require a directional or geodesic formulation of homogeneity rather than a direct replacement of $t - s$ by distance. We do not claim such an extension here. The assumptions of regularity and nondegeneracy are used to ensure a smooth Fisher coordinate, while fullness enters only when the variance-function conclusion is converted into the maximal normal or gamma NEF.

## 7. Conclusions

This paper has studied two related questions concerning Kullback–Leibler asymmetry in one-dimensional natural exponential families.

The first question concerns orientation. A local Kullback–Leibler ordering for sufficiently close symmetric pairs is equivalent to a fixed sign of standardized skewness and to monotonicity of the variance function. These conditions in turn imply the same direction of Kullback–Leibler divergence for every ordered pair in the family. Thus, a local third-order property determines a global information ordering.

The second question concerns homogeneity. After the family is parametrized by Fisher–Rao arc length, local spatial homogeneity of the directed Kullback–Leibler divergence, of its antisymmetric component, of the magnitude of this asymmetry, or local radiality of Jeffreys divergence are all equivalent. Each condition forces constant standardized skewness. The resulting variance function is either constant or the square of an affine function, and the full NEF is therefore, up to a nonsingular affine transformation, normal or fixed-shape gamma. The local information property then extends automatically to all admissible pairs.

The central structural chain is therefore

local information asymmetry $\longrightarrow$ skewness structure $\longrightarrow$ variance-function structure $\longrightarrow$ global NEF geometry.

In this sense, local information asymmetry determines global NEF structure. The normal family is the homogeneous reversible case and the gamma family is the homogeneous nonreversible case.

**Funding:** This research received no external funding

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** The author is grateful to the two anonymous reviewers for their careful and critical reading of the manuscript and for their constructive comments and suggestions. Their remarks helped improve both the presentation and the clarity of the paper.

**Conflicts of Interest:** The author declares no conflict of interest.

## References


1. Morris, C.N. Natural exponential families with quadratic variance functions. *Ann. Stat.* **1982**, *10*, 65–80. [CrossRef]
2. Morris, C.N. Natural exponential families with quadratic variance functions: Statistical theory. *Ann. Stat.* **1983**, *11*, 515–529. [CrossRef]
3. Letac, G.; Mora, M. Natural real exponential families with cubic variance functions. *Ann. Stat.* **1990**, *18*, 1–37. [CrossRef]
4. Jørgensen, B. *The Theory of Dispersion Models*; Chapman & Hall: London, UK, 1997.
5. Bar-Lev, S.K.; Enis, P. Reproducibility and natural exponential families with power variance functions. *Ann. Stat.* **1986**, *14*, 1507–1522. [CrossRef]
6. Bar-Lev, S.K.; Kokonendji, C.C. On the mean value parametrization of natural exponential families—A revisited review. *Math. Methods Stat.* **2017**, *26*, 159–175. [CrossRef]
7. Bar-Lev, S.K. Boundedness of variance functions of natural exponential families with unbounded support. *Mathematics* **2025**, *13*, 2045. [CrossRef]
8. Bar-Lev, S.K.; Enis, P. On the classical choice of variance stabilizing transformations and an application for a Poisson variate. *Biometrika* **1988**, *75*, 803–804. [CrossRef]
9. Bar-Lev, S.K.; Enis, P. On the construction of classes of variance stabilizing transformations. *Stat. Probab. Lett.* **1990**, *10*, 95–100. [CrossRef]
10. Harremoøes, P. Statistical inference and exact saddle point approximations. In Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018; pp. 1395–1399. [CrossRef]
11. Bar-Lev, S.K. An additional characterization of the triple consisting of normal, gamma, and inverse Gaussian distributions. *Mathematics* **2026**, *14*, 2758. [CrossRef]
12. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
13. Efron, B. The geometry of exponential families. *Ann. Stat.* **1978**, *6*, 362–376. [CrossRef]
14. Eguchi, S. Second order efficiency of minimum contrast estimators in a curved exponential family. *Ann. Stat.* **1983**, *11*, 793–803. [CrossRef]
15. Amari, S.; Nagaoka, H. *Methods of Information Geometry*; American Mathematical Society and Oxford University Press: Providence, RI, USA, 2000. [CrossRef]
16. Banerjee, A.; Merugu, S.; Dhillon, I.S.; Ghosh, J. Clustering with Bregman divergences. *J. Mach. Learn. Res.* **2005**, *6*, 1705–1749.
17. Eguchi, S.; Copas, J. Interpreting Kullback–Leibler divergence with the Neyman–Pearson lemma. *J. Multivar. Anal.* **2006**, *97*, 2034–2040. [CrossRef]
18. Matúš, F. On limiting towards the boundaries of exponential families. *Kybernetika* **2015**, *51*, 725–738. [CrossRef]
19. Nielsen, F.; Okamura, K. Group invariance of $f$-divergences and the Fisher–Rao distance. *arXiv* **2026**, arXiv:2606.25790.
20. Nielsen, F.; Okamura, K. Dual-connection midpoint matrix means and their Gauss composition. *arXiv* **2026**, arXiv:2609.07551.
21. Tojo, K.; Yoshino, T. Equivalence on exponential families as Hessian manifolds and classification of exponential families of order 1 with constant Hessian sectional curvature. *arXiv* **2026**, arXiv:2609.14562.
22. Nielsen, F. Approximation and bounding techniques for the Fisher–Rao distances between parametric statistical models. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 2024; Volume 51, pp. 67–116. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.