

# On some interrelations of generalized $q$ -entropies and a generalized Fisher information, including a Cramér-Rao inequality

Jean-François Bercher

► **To cite this version:**

Jean-François Bercher. On some interrelations of generalized  $q$ -entropies and a generalized Fisher information, including a Cramér-Rao inequality. Applied Stochastic Models and Data Analysis, Jun 2013, Mataro (Barcelona), Spain. pp.yet\_unknown. hal-00826453

**HAL Id: hal-00826453**

**<https://hal-upec-upem.archives-ouvertes.fr/hal-00826453>**

Submitted on 27 May 2013

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

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

# On some interrelations of generalized $q$ -entropies and a generalized Fisher information, including a Cramér-Rao inequality <sup>★</sup>

Jean-François Bercher<sup>1</sup>

Laboratoire d'informatique Gaspard Monge, UMR 8049 ESIEE-Paris, Université Paris-Est, France  
(E-mail: [jf.bercher@esiee.fr](mailto:jf.bercher@esiee.fr))

**Abstract** In this communication, we describe some interrelations between generalized  $q$ -entropies and a generalized version of Fisher information. In information theory, the de Bruijn identity links the Fisher information and the derivative of the entropy. We show that this identity can be extended to generalized versions of entropy and Fisher information. More precisely, a generalized Fisher information naturally pops up in the expression of the derivative of the Tsallis entropy. This generalized Fisher information also appears as a special case of a generalized Fisher information for estimation problems. Indeed, we derive here a new Cramér-Rao inequality for the estimation of a parameter, which involves a generalized form of Fisher information. This generalized Fisher information reduces to the standard Fisher information as a particular case. In the case of a translation parameter, the general Cramér-Rao inequality leads to an inequality for distributions which is saturated by generalized  $q$ -Gaussian distributions. These generalized  $q$ -Gaussians are important in several areas of physics and mathematics. They are known to maximize the  $q$ -entropies subject to a moment constraint. The Cramér-Rao inequality shows that the generalized  $q$ -Gaussians also minimize the generalized Fisher information among distributions with a fixed moment. Similarly, the generalized  $q$ -Gaussians also minimize the generalized Fisher information among distributions with a given  $q$ -entropy.

**Keywords:** Cramér-Rao inequality, generalized  $q$ -entropy, generalized Gaussians, de Bruijn identity.

Let  $f(x)$  be a probability distribution defined on  $X \subseteq \mathbb{R}^n$ . If  $M_q[f] = \int_X f(x)^q dx$ ,  $q \geq 0$  is the information generating function, then  $S_q[f] = \frac{1}{1-q} (M_q[f] - 1)$  is the so-called Tsallis entropy, or  $q$ -entropy (which can be easily related to the Rényi entropy). Both entropies reduce to the standard Shannon-Boltzmann entropy for  $q = 1$ . It is well-known that the maximum of the Rényi-Tsallis entropy, among all distributions with a fixed moment  $m_\alpha = E_g [\|X\|^\alpha]$ , is obtained for a generalized  $q$ -Gaussian distribution

$$G_\gamma(x) \propto (1 - (q-1)\gamma\|x\|^\alpha)_+^{\frac{1}{q-1}} \quad \text{for } q \neq 1 \quad (1)$$

which reduces to a generalized Gaussian for  $q = 1$ . It is also well-known that the classical Fisher information, which is derived from considerations in estimation

---

<sup>★</sup> To be presented during the conference “Applied Stochastic Models and Data Analysis”, Mataró (Barcelona), Spain, 25-28 June 2013 (and a version published in the proceedings)

theory, is linked to the Shannon entropy via the de Bruijn identity, and that this Fisher information is minimized by a Gaussian distribution, among all distributions with a fixed moment or entropy. It is the main objective of this paper to show that such results can be extended to the  $q$ -entropy and a suitable extension of Fisher information. Incidentally, we derive an interesting new generalized Cramér-Rao inequality in estimation theory, which might prove useful in its own right.

## 1 An extended de Bruijn identity

A fundamental connection between the Boltzmann-Shannon entropy, Fisher information, and the Gaussian distribution is given by the de Bruijn identity [8]. We show here that this important connection can be extended to the  $q$ -entropies, a suitable generalized Fisher information and the generalized  $q$ -Gaussian distributions.

The de Bruijn identity states that if  $Y_t = X + \sqrt{2t}Z$  where  $Z$  is a standard Gaussian vector and  $X$  a random vector of  $\mathbb{R}^n$ , independent of  $Z$ , then

$$\frac{d}{dt}H[f_{Y_t}] = I_{2,1}[f_{Y_t}] = \phi_{2,1}[f_{Y_t}], \quad (2)$$

where  $f_{Y_t}$  denotes the density of  $Y_t = X + \sqrt{2t}Z$ , and  $I_{2,1}[f_{Y_t}]$ ,  $\phi_{2,1}[f_{Y_t}]$  are two notations for the classical Fisher information (the meaning of which will be made clear in the following). Although the de Bruijn identity holds in a wider context, the classical proof of the de Bruijn identity uses the fact that if  $Z$  is a standard Gaussian vector, then  $Y_t$  satisfies the well-known heat equation  $\frac{\partial f}{\partial t} = \Delta f$ , where  $\Delta$  denotes the Laplace operator.

Nonlinear versions of the heat equation are of interest in a large number of physical situations, including fluid mechanics, nonlinear heat transfer or diffusion. Other applications have been reported in mathematical biology, lubrication, boundary layer theory, etc; see the series of applications presented in [9, chapters 2 and 21] and references therein. The porous medium equation and the fast diffusion equation correspond to the differential equation  $\frac{\partial f}{\partial t} = \Delta f^m$ , with  $m > 1$  for the porous medium equation and  $< 1$  for the fast diffusion. These two equations have been exhaustively studied and characterized by J. L. Vazquez, e.g. in [9,10].

These equations are included as particular cases into the *doubly nonlinear equation*, which involves a  $p$ -Laplacian operator  $\Delta_p f := \operatorname{div}(|\nabla f|^{p-2} \nabla f)$ , and the power  $m$  of the porous medium or fast diffusion equation. This doubly nonlinear equation takes the form

$$\frac{\partial}{\partial t} f = \Delta_\beta f^m = \operatorname{div}(|\nabla f^m|^{\beta-2} \nabla f^m), \quad (3)$$

where we use  $p = \beta$  for convenience and coherence with notation in the paper. The  $\beta$ -Laplacian typically appears in the minimization of a Dirichlet energy like  $\int |\nabla f|^\beta dx$  which leads to the Euler-Lagrange equation. It can be shown, see [10, page 192], that for  $m(\beta-1) + (\beta/n) - 1 > 0$ , (3) has a unique self-similar

solution, called a Barenblatt profile, whose initial value is the Dirac mass at the origin. This fundamental solution is usually given as a function of  $m$ . Here, if we put  $q = m + 1 - \frac{\alpha}{\beta}$ , the solution can be written as a  $q$ -Gaussian distribution:

$$f(x, t) = \frac{1}{t^{\frac{n}{\delta}}} B\left(\frac{x}{t^{\frac{1}{\delta}}}\right), \quad \text{with } B(x) = \begin{cases} (C - k|x|^\alpha)_+^{\frac{1}{q-1}} & \text{for } q \neq 1 \\ \frac{1}{\sigma} \exp\left(-\frac{|\beta-1|}{\beta^\alpha}|x|^\alpha\right) & \text{for } q = 1 \end{cases} \quad (4)$$

with  $\delta = n(\beta - 1)m + \beta - n > 0$ ,  $k = \frac{m(\beta-1)-1}{\beta} \left(\frac{1}{\delta}\right)^{\frac{1}{\beta-1}}$  and  $\alpha = \frac{\beta}{\beta-1}$ .

As mentioned above, the doubly nonlinear diffusion equation allows to derive a nice extension of the de Bruijn identity (2), and leads to a possible definition of a generalized Fisher information. This is stated in the next Proposition. The case  $\beta = 2$  of this result has been given in a paper by Johnson and Vignat [6].

**Proposition 1.** *[Extended de Bruijn identity [2]] Let  $f(x, t)$  a probability distributions defined on a subset  $X$  of  $\mathbb{R}^n$  and satisfying the doubly nonlinear equation (3). Assume that the domain  $X$  is independent of  $t$ , that  $f(x, t)$  is differentiable with respect to  $t$ , is continuously differentiable over  $X$ , and that  $\frac{\partial}{\partial t} f(x, t)^q$  is absolutely integrable and locally integrable with respect to  $t$ . Then, for  $\beta > 1$ ,  $\alpha$  and  $\beta$  Hölder conjugate of each other, for  $q = m + 1 - \frac{\alpha}{\beta}$ ,  $M_q[f] = \int f^q$  and  $S_q[f] = \frac{1}{1-q} (M_q[f] - 1)$  the Tsallis entropy, we have*

$$\frac{d}{dt} S_q[f] = q m^{\beta-1} \phi_{\beta,q}[f] = \left(\frac{m}{q}\right)^{\beta-1} M_q[f]^\beta I_{\beta,q}[f] \quad (5)$$

$$\text{with } \phi_{\beta,q}[f] = \int_X f(x)^{\beta(q-1)+1} \left(\frac{|\nabla f(x)|}{f(x)}\right)^\beta dx \quad \text{and } I_{\beta,q}[f] = \frac{\phi_{\beta,q}[f]}{M_q[f]^\beta}. \quad (6)$$

In (6),  $\phi_{\beta,q}[f]$  and  $I_{\beta,q}[f]$  are two possible generalization of Fisher information. Of course, the standard Fisher information is recovered in the particular case  $\alpha = \beta = 2$ , and  $q = m = 1$ , and so is the de Bruijn identity (2). The proof of this result relies on integration by part (actually using the Green identity) along the solutions of the nonlinear heat equation (3). This proof can be found in [2] and is not repeated here. A variant of the result for  $\beta = 2$ , which considers a free-energy instead of the entropy above, is well-known in certain circles, see e.g. [5,4]. More than that, by using carefully the calculations in [5], it is possible to check that  $\frac{d}{dt} \phi_{2,q}[f] \leq 0$  for  $q > 1 - \frac{1}{n}$ , which means the Tsallis entropy is a monotone increasing concave function along the solutions of (3). In their recent work [7], Savaré and Toscani have shown that in the case  $\beta = 2$ ,  $m = q$ , the entropy power, up to a certain exponent, is a concave function of  $t$ , thus generalizing the well-known concavity of the (Shannon) entropy power to the case of  $q$ -entropies. This allows to obtain as a by-product a generalized version of the Stam inequality, valid for the solutions of (3). We will come back to this generalized Stam inequality in Proposition 6.

## 2 Extended Cramér-Rao inequalities

Let  $f(x; \theta)$  be a probability distribution, with  $x \in X \subseteq \mathbb{R}^n$  and  $\theta \in \mathbb{R}^k$ . We will deal here with the estimation of a scalar function  $h(\theta)$  of  $\theta$ , with  $T(x)$  the corresponding estimator (the more general case where  $h(\theta)$  and  $T(x)$  are vector valued is a bit more involved; some results are given in [3] with general norms). We extend here the classical Cramér-Rao inequality in two directions: firstly, we give results for a general moment of the estimation error instead of the second order moment, and secondly we introduce the possibility of computing the moment of this error with respect to a distribution  $g(x; \theta)$  instead of  $f(x; \theta)$ : in estimation, the error is  $T(X) - h(\theta)$ , and the bias can be evaluated as  $\int_X (T(x) - h(\theta)) f(x; \theta) dx = E_f [T(X) - h(\theta)] = \eta(\theta) - h(\theta)$ , while a general moment of the error can be computed with respect to another probability distribution  $g(x; \theta)$ , as in  $E_g [|T(X) - h(\theta)|^\beta] = \int_X |T(x) - h(\theta)|^\beta g(x; \theta) dx$ . The two distributions  $f(x; \theta)$  and  $g(x; \theta)$  can be chosen very arbitrary. However, one can also build  $g(x; \theta)$  as a transformation of  $f(x; \theta)$  that highlights, or on the contrary scores out, some characteristics of  $f(x; \theta)$ . An important case is when  $g(x; \theta)$  is defined as the escort distribution of order  $q$  of  $f(x; \theta)$ :

$$f(x; \theta) = \frac{g(x; \theta)^q}{\int g(x; \theta)^q dx} \quad \text{and} \quad g(x; \theta) = \frac{f(x; \theta)^{\bar{q}}}{\int f(x; \theta)^{\bar{q}} dx}, \quad (7)$$

where  $q$  is a positive parameter,  $\bar{q} = 1/q$ , and provided of course that involved integrals are finite. These escort distributions are an essential ingredient in the nonextensive thermostatics context. It is in the special case where  $f(x; \theta)$  and  $g(x; \theta)$  are a pair of escort distributions that we will find again the generalized Fisher information (6) obtained in the extended de Bruijn identity. Our previous results on generalized Fisher information can be found in [3,1] in the case of the direct estimation of the parameter  $\theta$ . We propose here a novel derivation, introducing in particular a notion of generalized Fisher information matrix, in the case of the estimation of a function of the parameters. Let us first state the result.

**Proposition 2.** *Let  $f(x; \theta)$  be a multivariate probability density function defined for  $x \in X \subseteq \mathbb{R}^n$ , and with  $\theta \in \Theta \subseteq \mathbb{R}^k$  is a parameter of the density. Let  $g(x; \theta)$  denote another probability density function also defined on  $(X; \Theta)$ . Assume that  $f(x; \theta)$  is a jointly measurable function of  $x$  and  $\theta$ , is integrable with respect to  $x$ , is absolutely continuous with respect to  $\theta$ , and that the derivatives with respect to each component of  $\theta$  are locally integrable. Let  $T(x)$  be an estimator of a function  $h(\theta)$  and set  $\eta(\theta) = E_f [T(X)]$ . Then, for any estimator  $T(x)$  of  $h(\theta)$ , we have*

$$E_g [|T(X) - h(\theta)|^\alpha]^\frac{1}{\alpha} \geq \sup_{A > 0} \frac{\dot{\eta}(\theta)^T A \dot{\eta}(\theta)}{E_g [|\dot{\eta}(\theta)^T A \psi_g(X; \theta)|^\beta]^\frac{1}{\beta}}. \quad (8)$$

*with equality if and only if  $\dot{\eta}(\theta)^T A \psi_g(x; \theta) = c(\theta) \text{sign}(T(x) - h(\theta)) |T(x) - h(\theta)|^{\alpha-1}$ ,  $c(\theta) > 0$  and where  $\alpha^{-1} + \beta^{-1} = 1$ ,  $\alpha > 1$ ,  $A$  is a positive definite matrix and*

$\psi_g(x; \theta)$  a score function given with respect to  $g(x; \theta)$  :

$$\psi_g(x; \theta) := \frac{\nabla_{\theta} f(x; \theta)}{g(x; \theta)}. \quad (9)$$

*Proof.* Let  $\eta(\theta) = E_f[T(X)]$ . Let us first observe that  $E_g[\psi_g(x; \theta)] = \frac{d}{d\theta} \int_X f(x; \theta) dx = 0$ . Differentiating  $\eta(\theta) = E_f[T(X)]$  with respect to each  $\theta_i$  we get

$$\begin{aligned} \dot{\eta}(\theta) &= \nabla_{\theta} \eta(\theta) = \nabla_{\theta} \int_X T(x) f(x; \theta) dx \\ &= \int_X T(x) \frac{\nabla_{\theta} f(x; \theta)}{g(x; \theta)} g(x; \theta) dx \\ &= \int_X (T(x) - h(\theta)) \psi_g(x; \theta) g(x; \theta) dx. \end{aligned}$$

For any positive definite matrix  $A$ , multiplying on the left by  $\dot{\eta}(\theta)^T A$  gives

$$\dot{\eta}(\theta)^T A \dot{\eta}(\theta) = \int_X (T(x) - h(\theta)) \dot{\eta}(\theta)^T A \psi_g(x; \theta) g(x; \theta) dx,$$

and by the Hölder inequality, we obtain

$$E_g [|T(x) - h(\theta)|^{\alpha}]^{\frac{1}{\alpha}} E_g \left[ |\dot{\eta}(\theta)^T A \psi_g(x; \theta)|^{\beta} \right]^{\frac{1}{\beta}} \geq \dot{\eta}(\theta)^T A \dot{\eta}(\theta),$$

with equality if and only if  $(T(x) - h(\theta)) \dot{\eta}(\theta)^T A \psi_g(x; \theta) > 0$  and  $|T(x) - h(\theta)|^{\alpha} = k(\theta) |\dot{\eta}(\theta)^T A \psi_g(x; \theta)|^{\beta}$ ,  $k(\theta) > 0$ . This inequality, in turn, provides us with the lower bound (8) for the moment of order  $\alpha$  and computed wrt to  $g$  of the estimation error.  $\square$

The inverse of the matrix  $A$  which maximizes the right hand side is the Fisher information matrix of order  $\beta$ . Unfortunately, we do not have a closed-form expression for this matrix in the general case. Two particular cases are of interest.

**Corollary 3.** *[Scalar extended Cramér-Rao inequality] In the scalar case (or the case of a single component of  $\theta$ ), the following inequality holds*

$$E_g [|T(X) - h(\theta)|^{\alpha}]^{\frac{1}{\alpha}} \geq \frac{|\dot{\eta}(\theta)|}{E_g \left[ |\psi_g(X; \theta)|^{\beta} \right]^{\frac{1}{\beta}}}, \quad (10)$$

with equality if and only if  $\psi_g(x; \theta) = c(\theta) \text{sign}(T(x) - h(\theta)) |T(x) - h(\theta)|^{\alpha-1}$ .

In the simple scalar case, we see that  $A > 0$  can be simplified in (8) and thus that (10) follows. Note that for  $\alpha = 2$ , the equality case implies that  $E_g[\psi_g] = 0 = E_g[T(X) - h(\theta)]$ , which means that  $E_g[T(X)] = \eta(\theta) = h(\theta)$ , i.e. the estimator is unbiased (with respect to both  $f$  and  $g$ ). Actually, this inequality recovers at once the generalized Cramér-Rao inequality we presented in the univariate case [1]. The denominator plays the role of the Fisher information

in the classical case, which corresponds to the case  $g(x; \theta) = f(x; \theta)$ ,  $\beta = 2$ . As mentioned above, an extension of this result to the multidimensional case and arbitrary norms has been presented in [3] but it does not seem possible to obtain it here as particular case of (8).

A second interesting case is the multivariate case  $\alpha = \beta = 2$ . Indeed, in that case, we get an explicit form for the generalized Fisher information matrix and an inequality which looks like the classical one.

**Corollary 4.** *[Multivariate Cramér-Rao inequality with  $\alpha = \beta = 2$ ] For  $\alpha = \beta = 2$ , we have*

$$E_g \left[ |T(X) - h(\theta)|^2 \right] \geq \dot{\eta}(\theta)^T J_g(\theta)^{-1} \dot{\eta}(\theta) \quad (11)$$

with  $J_g(\theta) = E_g [\psi_g(X; \theta) \psi_g(X; \theta)^T]$ , and with equality if and only if  $|T(X) - h(\theta)| = k(\theta) |\dot{\eta}(\theta)^T J_g(\theta)^{-1} \psi_g(X; \theta)|$ .

*Proof.* The denominator of (8) is a quadratic form and we have

$$\begin{aligned} E_g \left[ |T(X) - h(\theta)|^2 \right] &\geq \sup_{A>0} \frac{(\dot{\eta}(\theta)^T A \dot{\eta}(\theta))^2}{E_g \left[ |\dot{\eta}(\theta)^T A \psi_g(X; \theta)|^2 \right]} \\ &\geq \sup_{A>0} \frac{(\dot{\eta}(\theta)^T A \dot{\eta}(\theta))^2}{\dot{\eta}(\theta)^T A E_g [\psi_g(X; \theta) \psi_g(X; \theta)^T] A^T \dot{\eta}(\theta)}. \end{aligned} \quad (12)$$

Let  $J_g(\theta) = E_g [\psi_g(X; \theta) \psi_g(X; \theta)^T]$  and set  $z(\theta) = A^{\frac{1}{2}} \dot{\eta}(\theta)$ . With these notations, and using the inequality  $(z^T z)^2 \leq (z^T B z) (z^T B^{-1} z)$  valid for any  $B > 0$ , we obtain that

$$\dot{\eta}(\theta)^T J^{-1} \dot{\eta}(\theta) \geq \sup_{A>0} \frac{(z(\theta)^T z(\theta))^2}{z(\theta)^T A^{\frac{1}{2}} J_g(\theta) \left( A^{\frac{1}{2}} \right)^T z(\theta)}.$$

Since it can be readily checked that the upper bound is attained with  $A = J_g(\theta)^{-1}$ , we finally end with (11). Of course, for  $g = f$ , the inequality (11) reduces to the classical multivariate Cramér-Rao inequality.  $\square$

An important consequence of these results is obtained in the case of a translation parameter, where the generalized Cramér-Rao inequality induces a new class of inequalities. Let  $\theta \in \mathbb{R}$  be a scalar location parameter,  $x \in X \subseteq \mathbb{R}^n$ , and define by  $f(x; \theta)$  the family of density  $f(x; \theta) = f(x - \theta \mathbf{1})$ , where  $\mathbf{1}$  is a vector of ones. In this case, we have  $\nabla_{\theta} f(x; \theta) = -\mathbf{1}^T \nabla_x f(x - \theta \mathbf{1})$ , provided that  $f$  is differentiable at  $x - \theta \mathbf{1}$ , and the Fisher information becomes a characteristic of the information in the distribution. If  $X$  is a bounded subset, we will assume that  $f(x)$  vanishes and is differentiable on the boundary  $\partial X$ . Without loss of generality, we will assume that the mean of  $f(x)$  is zero. Set  $h(\theta) = \theta$  and take  $T(X) = \mathbf{1}^T X/n$ , with of course  $\eta(\theta) = E[T(X)] = \theta$  and  $\dot{\eta}(\theta) = 1$ . Finally, let us choose the particular value  $\theta = 0$ . In these conditions, the generalized

Cramér-Rao inequality (10) becomes

$$E_g \left[ |1^T X|^\alpha \right]^{\frac{1}{\alpha}} E_g \left[ \left| 1^T \frac{\nabla_x f(X)}{g(X)} \right|^\beta \right]^{\frac{1}{\beta}} \geq n, \quad (13)$$

with equality if and only if  $1^T \frac{\nabla_x f(x)}{g(x)} = c(\theta) \text{sign}(1^T X) |1^T X|^{\alpha-1}$ . In [3], we have a slightly more general result in the multivariate case:

$$E_g [\|X\|^\alpha]^{\frac{1}{\alpha}} E_g \left[ \left\| \frac{\nabla_x f(X)}{g(X)} \right\|_*^\beta \right]^{\frac{1}{\beta}} \geq n, \quad (14)$$

where  $\|\cdot\|$  is a norm, and the corresponding dual norm is denoted by  $\|\cdot\|_*$ . Finally, let  $f(x)$  and  $g(x)$  be a pair of escort distributions as in (7). In such a case, the following recovers the generalized Fisher information (6) and yields a new characterization of  $q$ -Gaussian distributions.

**Corollary 5.** [*q-Cramér-Rao inequality*] Assume that  $g(x)$  is a measurable differentiable function of  $x$ , which vanishes and is differentiable on the boundary  $\partial X$ , and finally that the involved integrals exist and are finite. Then, for the pair of escort distributions (7), the following  $q$ -Cramér-Rao inequality holds

$$q^\beta E_g [\|X\|^\alpha]^{\frac{1}{\alpha}} I_{\beta,q}[g]^{\frac{1}{\beta}} \geq n, \quad (15)$$

$$\text{with } I_{\beta,q}[g] = \left(1/M_q[g]^\beta\right) E \left[ g(x)^{\beta(q-1)} \left\| \frac{\nabla_x g(x)}{g(x)} \right\|_*^\beta \right],$$

and with equality if and only if  $g(x)$  is a generalized  $q$ -Gaussian, i.e.  $g(x) \propto (1 - \gamma \|x\|^\alpha)_+^{\frac{1}{q-1}}$ , with  $\gamma > 0$ .

*Proof.* The result follows from (14), or (13) with  $n = 1$ , and the fact that

$$\frac{\nabla_x f(X)}{g(X)} = \frac{q}{M_q[g]} g(X)^{q-1} \frac{\nabla_x g(X)}{g(X)}.$$

The case of equality is obtained by solving the general equality conditions in the special case where  $f(x)$  and  $g(x)$  form a pair of escort distributions.  $\square$

As a direct consequence of the  $q$ -Cramér-Rao inequality (15), we obtain that the minimum of the generalized Fisher information among all distributions with a given moment of order  $\alpha$ , say  $m_\alpha = E_g [\|X\|^\alpha]$ , is obtained when  $g$  is a generalized  $q$ -Gaussian distribution, with a parameter  $\gamma$  such that the distribution has the prescribed moment. This parallels, and complements the well known fact that the  $q$ -Gaussians maximize the  $q$ -entropies subject to a moment constraint, and yields new variational characterizations of the generalized  $q$ -Gaussians. As mentioned earlier, the generalized Fisher information also satisfies an extension of Stam's inequality, which links the generalized Fisher



information and the  $q$ -entropy power, defined as an exponential of the Rényi entropy  $H_q[f]$  as

$$N_q[f] = M_q[f]^{\frac{2}{n} \frac{1}{1-q}} = \exp\left(\frac{2}{n} H_q[f]\right) = \left(\int_{\Omega} f(x)^q dx\right)^{\frac{2}{n} \frac{1}{1-q}}, \quad (16)$$

for  $q \neq 1$ . For  $q = 1$ , we set  $N_q[f] = \exp\left(\frac{2}{n} H_1[f]\right)$ , where  $H_1[f]$  is the Boltzmann-Shannon entropy. The generalized Stam inequality is given here without proof (see [2]).

**Proposition 6.** *[Generalized Stam inequality] Let  $n \geq 1$ ,  $\beta$  and  $\alpha$  be Hölder conjugates of each other,  $\alpha > 1$ , and  $q > \max\{(n-1)/n, n/(n+\alpha)\}$ . Then for any probability density on  $\mathbb{R}^n$ , that is continuously differentiable, the following generalized Stam inequality holds*

$$I_{\beta,q}[f]^{\frac{1}{\beta}} N_q[f]^{\frac{1}{2}} \geq I_{\beta,q}[G]^{\frac{1}{\beta}} N_q[G]^{\frac{1}{2}}. \quad (17)$$

with  $\lambda = n(q-1) + 1 > 0$  and with equality if and only if  $f$  is any generalized  $q$ -Gaussian (1).

The generalized Stam inequality implies that the generalized  $q$ -Gaussian minimize the generalized Fisher information within the set of probability distributions with a fixed  $q$ -entropy power.

To sum up and emphasize the main results, let us point out that we have exhibited a generalized Fisher information, both as a by-product of a generalization of de Bruijn identity and as a fundamental measure of information in estimation theory. We have shown that this allows to draw a nice interplay between  $q$ -entropies, generalized  $q$ -Gaussians and the generalized Fisher information. These interrelations yield the generalized  $q$ -Gaussians as minimizers of the  $(\beta, q)$ -Fisher's information under adequate constraints, or as minimizers of functionals involving  $q$ -entropies,  $(\beta, q)$ -Fisher's information and/or moments. This is shown through inequalities and identities involving these quantities and generalizing classical information relations (Cramér-Rao's inequality, Stam's inequality, De Bruijn's identity).

## References

1. J.-F. Bercher, On generalized Cramér-Rao inequalities, generalized Fisher informations and characterizations of generalized  $q$ -Gaussian distributions, *Journal of Physics A: Mathematical and Theoretical*, 45 (2012), p. 255303.
2. J.-F. Bercher, Some properties of generalized Fisher information in the context of nonextensive thermostatics, *Physica A: Statistical Mechanics and its Applications*, (2013), <http://hal.archives-ouvertes.fr/hal-00766699>, in press.
3. J.-F. Bercher, On multidimensional generalized Cramér-Rao inequalities, uncertainty relations and characterizations of generalized  $q$ -Gaussian distributions, *Journal of Physics A: Mathematical and Theoretical*, 46 (2013), p. 095303.
4. J. A. Carrillo and G. Toscani, Asymptotic  $l_1$ -decay of solutions of the porous medium equation to self-similarity, *Indiana University Mathematics Journal*, 49 (2000), pp. 113–142.

5. J. Dolbeault and G. Toscani, Improved interpolation inequalities, relative entropy and fast diffusion equations, *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, (2013), in press.
6. O. Johnson and C. Vignat, Some results concerning maximum Rényi entropy distributions, *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 43 (2007), pp. 339–351.
7. G. Savaré and G. Toscani, The concavity of Rényi entropy power, arXiv:1208.1035, (2012).
8. A. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Information and Control*, 2 (1959), pp. 101–112.
9. J. L. Vázquez, *The Porous Medium Equation: Mathematical Theory*, Oxford University Press, USA, 1 ed., Dec. 2006.
10. J. L. Vázquez, *Smoothing and Decay Estimates for Nonlinear Diffusion Equations: Equations of Porous Medium Type*, Oxford University Press, USA, Oct. 2006.