The concavity of Rényi entropy power

Giuseppe Savaré and Giuseppe Toscani

***Abstract*—We associate to the p-th Rényi entropy a definition of entropy power, which is the natural extension of Shannon's entropy power and exhibits a nice behaviour along solutions to the p-nonlinear heat equation in $\mathbb{R}^n$. We show that the Rényi entropy power of general probability densities solving such equations is always a concave function of time, whereas it has a linear behaviour in correspondence to the Barenblatt source-type solutions. This result extends Costa's concavity inequality for Shannon's entropy power to Rényi entropies.**

***Index Terms*—Entropy, information measure, information theory, Rényi entropy, nonlinear heat equation.**

I. Introduction

The p-th Rényi entropy of a probability density f in $\mathbb{R}^n$ is defined by (see, e.g. Cover and Thomas [7] and Gardner [12])

$$\mathrm{h}_p(f) := \frac{1}{1-p}\log\int_{\mathbb{R}^n} f^p(x)\,\mathrm{d}x, \tag{1}$$

for $0 < p < +\infty$, $p \neq 1$.

Whenever $p > 1-2/n$, we consider the positive coefficient

$$\mu := 2 + n(p-1), \tag{2}$$

and we associate to the p-th Rényi entropy the entropy power (that we call *p-thRényi entropy power* in the following) given by

$$\mathrm{N}_p(f) := \exp\left(\frac{\mu}{n}\,\mathrm{h}_p(f)\right). \tag{3}$$

The Rényi entropy for $p = 1$ is defined as the limit of $\mathrm{h}_p(f)$ as $p \to 1$. It follows directly from definition (1) that

$$\mathrm{h}_1(f) = \lim_{p\to1}\mathrm{h}_p(f) = \mathrm{h}(f) = -\int_{\mathbb{R}^n} f(x)\log f(x)\,\mathrm{d}x.$$

Therefore, the Shannon's entropy can be identified with the Rényi entropy of index $p = 1$. In this case, the proposed Rényi entropy power of index $p = 1$, given by (3), coincides with Shannon's entropy power

$$\mathrm{N}(f) := \exp\left(\frac{2}{n}\mathrm{h}(f)\right). \tag{4}$$

In 1985 Costa [6] proved that, if u_t, $t > 0$, are probability densities solving the heat equation in the whole space $\mathbb{R}^n$

$$\frac{\partial}{\partial t}u = \Delta u, \tag{5}$$

then

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{N}(u_t) \leq 0. \tag{6}$$

G. Savaré and G. Toscani are with the Department of Mathematics, University of Pavia, 27100 Pavia, Italy (e-mail: giuseppe.savare@unipv.it; giuseppe.toscani@unipv.it) (see http://www-dimat.unipv.it).

The authors acknowledge support by MIUR project "Calculus of Variations".

Manuscript received ; revised .

In (5) $\Delta u = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}u$ and u is the function in $\mathbb{R}^n \times (0,\infty)$ taking the value $u_t(x)$ for $x \in \mathbb{R}^n$ and $t > 0$.

Inequality (6) is referred to as the *concavity of entropy power* theorem. The original proof of Costa has been simplified years later by Dembo [10], [11] with an argument based on the Blachman–Stam inequality [3]. Next, improvements have been obtained by Guo, Shamai and Verdu [13] via the minimum mean-square error (MMSE), and by Rioul [18] using mutual information (see [18] for an exhaustive list of references about more general forms of the concavity property).

A direct proof of (6) in a strengthened form, with an exact error term, has been obtained by Villani [24]. The proof in [24] highlights a strong connection between the concavity of entropy power and some identities of Bakry and Emery [2], established through the so-called Γ_2 calculus as part of their famous work on logarithmic Sobolev inequalities and hyper-contractive diffusions. These connections, together with various consequences of the concavity of entropy power theorem, have been recently discussed in [21].

In this paper we show that the *concavity of entropy power* is a property which is not restricted to Shannon entropy power (4) in connection with the heat equation (5), but it holds for the p-th Rényi entropy power (3), if we put it in connection with the solution to the nonlinear heat equation

$$\frac{\partial}{\partial t}u = \Delta u^p, \tag{7}$$

posed in the whole space $\mathbb{R}^n$. Diffusion processes described by equation (7) are known to arise in different fields of physics such as plasma physics, kinetic theory of gases, solid state and transport in porous medium. In the range $p > 1$, the corresponding nonlinear heat equation, named porous medium equation has been used by J. Boussinesq in the study of groundwater infiltration and in the description of the flow of an isentropic gas through a porous medium. It also applies to the theory of heat radiation in plasmas developed by Ya.B. Zel'dovich and his coworkers. For physical and mathematical details about equation (7), we refer to the recent book by J.L. Vazquez [23], which we will use as reference, by giving, when necessary, precise indications about results we will only quote.

It is remarkable that the range of the parameter p for which we introduced the Rényi entropy power, coincides with the range for which there is mass conservation (i.e. conservation in time of the integral $\int_{\mathbb{R}^n} u(x,t)\,\mathrm{d}x$) for the solution of (7) (cf. [23], Chapter 5, Section 5.10, and [4]).

The precise result is the following.

Theorem 1: Let $p > 1 - 2/n$ and let u_t, $t > 0$, be probability densities in $\mathbb{R}^n$ solving (7). Then the p-th Rényi entropy power defined in (3) satisfies

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{N}_p(u_t) \le 0, \tag{8}$$

thus $t \mapsto \mathrm{N}_p(u_t)$ is concave.

The relationship between the Rényi entropy power and the solution to the nonlinear heat equation (7) can be fruitfully highlighted owing to its self-similar solution. It was found around 1950 by Zel'dovich and Kompaneets and Barenblatt (cf. Chapter 4, Sections 4.4.2 and 4.4.3 of [23] for an exhaustive derivation). For $p > 0$, the Barenblatt (also called generalized Gaussian) solution departing from $x = 0$ takes the self-similar form (recall the definition of μ in (2))

$$M_{p,t}(x) := \frac{1}{t^{n/\mu}}\tilde{M}_p\left(\frac{x}{t^{1/\mu}}\right), \tag{9}$$

arising from the profile

$$\tilde{M}_p(x) = \left(\tilde{C}_p - \kappa\,|x|^2\right)_+^{\frac{1}{p-1}}; \tag{10}$$

here $(s)_+ = \max\{s, 0\}$, $\kappa := \frac{1}{2\mu}\,\frac{p-1}{p}$.

The mathematical properties of the symmetric function (10) clearly depend on the value of the exponent p which characterizes the nonlinearity in equation (7). While for $p > 1$ $\tilde{M}_p$ has bounded support, and the integral

$$\int_{\mathbb{R}^n} |x|^q \tilde{M}_p(x)\,\mathrm{d}x \tag{11}$$

is bounded for all values of q, in the opposite case $p < 1$ this solution has a polynomial decay at infinity, and the integral (11) can be unbounded. This fact put into evidence two particular intervals of the parameter p, that will be extensively treated in this paper. In the interval $(n-2)/n < p < 1$, the integral (11), with $q = 0$ is bounded. Then, the constant $\tilde{C}_p$ in (10) can be chosen to render source-type Barenblatt solution a probability density.

Second, if $n/(n+2) < p < 1$, the integral (11) is bounded whenever $q = 2$. In this case, by means of the relationship between the second moment $\int_{\mathbb{R}^n} |x|^2\tilde{M}_p(x)\,\mathrm{d}x$ of the Barenblatt solution and its Rényi entropy $\mathrm{h}_p(\tilde{M}_p)$ (see the detailed computations in the Appendix, Part B), the Rényi entropy of $\tilde{M}_p$ is bounded.

In particular, if $p > n/(n+2)$, elementary computations (see also Section III below) show that the bounded p-th Rényi entropy power of the Barenblatt, defined in (3), is a linear function of time

$$\mathrm{N}_p(M_{p,t}) = \mathrm{N}_p(\tilde{M}_p)\,t, \quad \text{so that} \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{N}_p(M_{p,t}) = 0. \tag{12}$$

The concavity property of Rényi entropy power stated in Theorem 1 implies that for all times $t > 0$ the Barenblatt source-type solution maximizes the second derivative of the p-th Rényi entropy power among all possible solutions to the nonlinear heat equation.

Notice that in the case of $p = 1$ the source-type solution of unit mass at time $t > 0$ is represented by the Gaussian density of variance equal to $2t$ [8]

$$M_{1,t}(x) := \frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|x|^2}{4t}\right), \tag{13}$$

representing the heat release from a point source (here the origin $x = 0$). We refer to the book of Vazquez [23], Chapter 4, Section 4.4.2 for an exhaustive discussion on the connections between the Barenblatt profiles (9) of order $p \neq 1$ and the Gaussian (13).

Since the Shannon's entropy of M_t equals

$$\mathrm{h}(M_t) = \frac{n}{2}\log(4\pi\,\mathrm{e}\,t),$$

it follows that also in this case the corresponding entropy power is a linear function of time, i.e.

$$\mathrm{N}(M_{1,t}) = 4\pi\,\mathrm{e}\,t, \qquad \text{hence} \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{N}(M_{1,t}) = 0.$$

If $p > n/(n+2)$, the self-similar Barenblatt solution (9) represents the intermediate asymptotic of solutions to (7) for a large class of initial data $u(x,0)$ such that the second moment $\int_{\mathbb{R}^n} |x|^2 u(x,0)\,\mathrm{d}x$ is finite. (cf. the discussion in [20], and [4] for a recent collections of results in the range $p < 1$). Thus, in the range $p > n/(n+2)$ one can identify in a precise way the large-time behaviour of the solutions to (7), and, consequently, one can obtain from the concavity property a chain of differential inequalities in sharp form.

Indeed, this connection between Rényi entropies and the nonlinear heat equations allows to recover in a simple way a related p-th Fisher information recently considered in [14] [15], and to better understand the distinguished role of self-similar solutions to (7).

Thus we introduce the p-weighted Fisher information

$$\mathrm{I}_p(f) := \frac{1}{\int_{\mathbb{R}^n} f^p\,\mathrm{d}x}\int_{\{f>0\}} \frac{|\nabla f^p|^2}{f}\,\mathrm{d}x, \tag{14}$$

which reduces to the Fisher information of a random variable with density f as $p \to 1$

$$\mathrm{I}(f) := \int_{\{f>0\}} \frac{|\nabla f|^2}{f}\,\mathrm{d}x. \tag{15}$$

A simple computation (see the Appendix and (45)) shows that $\mathrm{I}_p(\tilde{M}_p)$ is finite if and only if the second moment of $\tilde{M}_p$ is finite, i.e. if $p > n/(n+2)$. Like in the Shannon's case, in this range of exponents inequality (8) leads to the following sharp inequality.

Theorem 2: If $p > n/(n+2)$ every smooth, strictly positive and rapidly decaying probability density f satisfies

$$\mathrm{N}_p(f)\,\mathrm{I}_p(f) \ge \mathrm{N}_p(\tilde{M}_p)\,\mathrm{I}_p(\tilde{M}_p) = \gamma_{n,p}, \tag{16}$$

where the value of the strictly positive constant $\gamma_{n,p}$ is given by

$$\gamma_{n,p} = n\pi\frac{2p}{p-1}\cdot \\ \cdot\left(\frac{\Gamma\left(\frac{p}{p-1}\right)}{\Gamma\left(\frac{n}{2}+\frac{p}{p-1}\right)}\right)^{2/n}\left(\frac{(n+2)p-n}{2p}\right)^{\frac{2+n(p-1)}{n(p-1)}} \tag{17}$$

when $p > 1$ and by

$$\gamma_{n,p} = n\pi \frac{2p}{1-p} \cdot \cdot \left(\frac{\Gamma\left(\frac{1}{1-p} - \frac{n}{2}\right)}{\Gamma\left(\frac{1}{1-p}\right)} \right)^{2/n} \left(\frac{(n+2)p - n}{2p} \right)^{\frac{2+n(p-1)}{n(p-1)}}. \quad (18)$$

if $n/(n+2) < p < 1$.

In our opinion the previous results justify the choice to define the p-th Rényi entropy power in the form (3).

Theorem 2 gives an information theoretical validity to the Rényi entropy and to the p-th Rényi entropy power. In analogy with the well-known linear case, in which the concavity of the Shannon entropy power implies the isoperimetric inequality for entropies (cf. Theorem 16 in [11], and the recent discussion in [21])

$$I_1(f)N_1(f) \geq 2\pi n e,$$

in the nonlinear case (when $p > n/(n+2)$) the concavity of the p-th Rényi entropy power implies the analogous sharp isoperimetric inequality (16).

However, at difference with the linear case, in which concavity of the entropy power can be rephrased by asserting that the entropy power of a random variable $X_t = X + \sqrt{t}Z$ is concave with respect to the variance t of a random Gaussian perturbation Z (cf. Corollary 3 in [11]), in the nonlinear case this analogy with respect to perturbations of variance t is completely lost.

In the next two sections we will present the proofs of our main results, Theorems 1 and 2. In Section IV we will discuss an equivalent formulation of inequality (16) and its relationship with Sobolev inequalities in the particular case when $p = 1 - 1/n$. We confine in the Appendix the calculations related to the precise value of the constant $\gamma_{n,p}$ given in (17) and (18).

II. Proof of the concavity of Rényi entropy power

We split the proof in two parts. In the first part we show that the solutions to (7) satisfy two key identities. Arguing as in [22], it is sufficient to consider the case of smooth, strictly positive and rapidly decaying probability densities.

For such a given probability density u we set

$$\mathrm{E}(u) := \int \mathsf{e}(u(x))\,\mathrm{d}x, \quad \text{where } \mathsf{e}(r) := \frac{1}{p-1}\, r^p,$$

and $p > 1 - 2/n$, $p \neq 1$, is a fixed exponent. Consequently, the p-th Rényi entropy of u can be written as

$$\mathrm{h}_p(u) := \frac{1}{1-p} \log\Big((p-1)\mathrm{E}(u)\Big). \quad (19)$$

We now introduce two nonnegative functionals related to the first and second time derivative of E along the nonlinear heat equation (7): the first one is also related to the weighted Fisher information (14)

$$\mathrm{E}'(u) := \int \frac{|\nabla u^p|^2}{u}\,\mathrm{d}x = \Big(\frac{p}{p-1}\Big)^2 \int |\nabla u^{p-1}|^2\, u\,\mathrm{d}x$$
$$= \int \big|\nabla \mathsf{e}'(u)\big|^2 u\,\mathrm{d}x, \qquad \text{since} \quad \mathsf{e}'(u) = \frac{p}{p-1}.$$

The last one is

$$\mathrm{E}''(u) := 2\int u^p \Big(|\nabla^2 \mathsf{e}'(u)|^2 + (p-1)\big(\Delta \mathsf{e}'(u)\big)^2\Big)\,\mathrm{d}x,$$

where $\nabla^2 u$ is the Hessian matrix $\big(\frac{\partial^2 u}{\partial x_i \partial x_j}\big)_{ij}$ and $\big|\nabla^2 u\big|^2 = \sum_{i,j}\big(\frac{\partial^2 u}{\partial x_i \partial x_j}\big)^2$.

The interest of the above functionals is motivated by the following result, that also justifies the notation we used:

Proposition 3: If u_t, $t > 0$ are smooth probability densities solving the nonlinear heat equation (7), then

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(u_t) = \mathrm{E}'(u_t), \quad (20)$$

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}'(u_t) = \mathrm{E}''(u_t). \quad (21)$$

Proof: Recalling that

$$\mathsf{f}(u) = u^p = u\,\mathsf{e}'(u) - \mathsf{e}(u), \quad \mathsf{f}'(u) = u\,\mathsf{e}''(u) = p u^{p-1},$$

the nonlinear heat equation (7) can be equivalently written as

$$\partial_t u - \nabla \cdot \Big(u\,\nabla \mathsf{e}'(u)\Big) = 0. \quad (22)$$

Using equation (22), and integrating by parts, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(u_t) = \int \mathsf{e}'(u)\,\partial_t u\,\mathrm{d}x$$
$$= \int \mathsf{e}'(u)\,\nabla \cdot \Big(u\,\nabla \mathsf{e}'(u)\Big)\,\mathrm{d}x$$
$$= -\int \nabla \mathsf{e}'(u) \cdot \nabla \mathsf{e}'(u)\,u\,\mathrm{d}x = -\mathrm{E}'(u_t).$$

(21) is based on the Bochner identity or, equivalently, Bakry-Émery Γ-calculus (in their simplest Euclidean form), see [17] for analogous computations of the second derivative of E along geodesics in the Wasserstein space.

One has

$$\partial_t \mathsf{e}'(u) = \mathsf{e}''(u)\partial_t u$$
$$= \mathsf{e}''(u)\,\nabla \cdot \Big(u\,\nabla \mathsf{e}'(u)\Big)$$
$$= u\,\mathsf{e}''(u)\,\Delta \mathsf{e}'(u) + \mathsf{e}''(u)\,\nabla u \cdot \nabla \mathsf{e}'(u)$$
$$= \mathsf{f}'(u)\,\Delta \mathsf{e}'(u) + \big|\nabla \mathsf{e}'(u)\big|^2.$$

Recall the Bochner identity

$$2\Gamma_2(g) = \Delta|\nabla g|^2 - 2\nabla g \cdot \nabla \Delta g = 2|\nabla^2 g|^2, \quad (23)$$

and

$$(r\mathsf{f}'(r) - \mathsf{f}(r))' = r\mathsf{f}''(r),$$

while

$$(r\mathsf{f}'(r) - \mathsf{f}(r))'\mathsf{e}''(r) = \mathsf{f}''(r)\mathsf{f}'(r).$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}'(u_t) = \int \Big(\big|\nabla \mathsf{e}'(u)\big|^2 \partial_t u + 2u\,\nabla \mathsf{e}'(u) \cdot \nabla \partial_t \mathsf{e}'(u)\Big)\,\mathrm{d}x$$
$$= \int \big|\nabla \mathsf{e}'(u)\big|^2 \Delta u^p\,\mathrm{d}x$$
$$+ 2\int u\,\nabla \mathsf{e}'(u) \cdot \nabla\Big(\mathsf{f}'(u)\,\Delta \mathsf{e}'(u) + \big|\nabla \mathsf{e}'(u)\big|^2\Big)\,\mathrm{d}x;$$

repeated integration by parts yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}'(u_t) = \int u^p \Delta\big|\nabla \mathsf{e}'(u)\big|^2 \,\mathrm{d}x$$
$$+ 2\int u\mathsf{f}'(u)\,\nabla\mathsf{e}'(u)\cdot\nabla\Delta\mathsf{e}'(u)\,\mathrm{d}x$$
$$+ 2\int \nabla\mathsf{f}(u)\cdot\nabla\mathsf{f}'(u)\,\Delta\mathsf{e}'(u)\,\mathrm{d}x - 2\int \mathsf{f}(u)\Delta\big|\nabla\mathsf{e}'(u)\big|^2\,\mathrm{d}x$$
$$= -\int \mathsf{f}(u)\Delta\big|\nabla\mathsf{e}'(u)\big|^2\,\mathrm{d}x + 2\int \mathsf{f}(u)\,\nabla\mathsf{e}'(u)\cdot\nabla\Delta\mathsf{e}'(u)\,\mathrm{d}x$$
$$+ 2\int \big(u\mathsf{f}'(u)-\mathsf{f}(u)\big)\,\nabla\mathsf{e}'(u)\cdot\nabla\Delta\mathsf{e}'(u)\,\mathrm{d}x$$
$$+ 2\int \nabla\mathsf{f}(u)\cdot\nabla\mathsf{f}'(u)\,\Delta\mathsf{e}'(u)\,\mathrm{d}x$$

and the Bochner identity (23) eventally gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}'(u_t) = -\int \mathsf{f}(u)\Gamma_2(\mathsf{e}'(u))\,\mathrm{d}x$$
$$- 2\int \big(u\mathsf{f}'(u)-\mathsf{f}(u)\big)\,\big(\Delta\mathsf{e}'(u)\big)^2\,\mathrm{d}x$$
$$= -2\int \mathsf{f}(u)|\nabla^2\mathsf{e}'(u)|^2 + \big(u\mathsf{f}'(u)-\mathsf{f}(u)\big)\,\big(\Delta\mathsf{e}'(u)\big)^2\,\mathrm{d}x$$
$$= -2\int u^p\Big(|\nabla^2\mathsf{e}'(u)|^2 + (p-1)\big(\Delta\mathsf{e}'(u)\big)^2\Big)\,\mathrm{d}x.$$

■

We can now conclude the *proof of Theorem 1*.

Let us first notice that for a given function ϕ_t which depends on time, and a positive constant σ we have

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\exp(\sigma\,\phi_t) = \exp(\sigma\phi_t)\Big(\sigma\phi_t'' + (\sigma\phi_t')^2\Big),$$

so that the concavity condition $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\exp(\sigma\phi_t) \le 0$ is equivalent to $-\sigma\phi_t'' \ge (\sigma\phi_t')^2$.

If $\phi_t := \mathrm{h}_p(u_t)$, where $u_t = u_t$, we have

$$\phi_t' = \frac{1}{p-1}\frac{-\mathrm{E}'(u_t)}{\mathrm{E}(u_t)}, \tag{24}$$

and

$$\phi_t'' = \frac{1}{p-1}\frac{\mathrm{E}''(u_t)\,\mathrm{E}(u_t) - \mathrm{E}'(u_t)^2}{\mathrm{E}(u(t))^2}. \tag{25}$$

Hence we end up with the condition

$$\frac{\sigma}{p-1}\Big(\mathrm{E}''(u_t)\,\mathrm{E}(u_t) - \mathrm{E}'(u_t)^2\Big) \ge \Big(\frac{\sigma}{p-1}\Big)^2\mathrm{E}'(u_t)^2,$$

i.e. (suppressing the index t)

$$\sigma\mathrm{E}''(u)\int u^p\,\mathrm{d}x \ge \big(\sigma^2 + \sigma(p-1)\big)\mathrm{E}'(u)^2.$$

Since $\sigma > 0$, the second derivative is non positive if

$$\mathrm{E}''(u)\int u^p\,\mathrm{d}x \ge \big(\sigma + (p-1)\big)\big(\mathrm{E}'(u)\big)^2. \tag{26}$$

Since an integration by parts yields

$$\mathrm{E}'(u) = \int \nabla u^p\cdot\nabla\mathsf{e}'(u)\,\mathrm{d}x = -\int u^p\,\Delta\mathsf{e}'(u)\,\mathrm{d}x,$$

by Cauchy-Schwarz inequality we have

$$\mathrm{E}'(u)^2 \le \int u^p\,\mathrm{d}x\int u^p(\Delta\mathsf{e}'(u))^2\,\mathrm{d}x.$$

It follows that (26) holds if

$$\mathrm{E}''(u) \ge (\sigma + (p-1))\int u^p(\Delta\mathsf{e}'(u))^2\,\mathrm{d}x.$$

On the other hand, the well known trace inequality

$$|\nabla^2 f|^2 \ge \frac{1}{n}\big(\Delta f\big)^2$$

that follows by the elementary inequality $\frac{1}{n}\big(\sum_{i=1}^n a_i\big)^2 \le \sum_{i=1}^n a_i^2,\ a_i\in\mathbb{R}$, yields

$$\mathrm{E}''(u) \ge 2\,(\frac{1}{n} + (p-1))\int u^p(\Delta\mathsf{e}'(u))^2\,\mathrm{d}x,$$

and we end up with the condition

$$\sigma \le \frac{2}{n} + p - 1 = \frac{\mu}{n}. \tag{27}$$

Choosing for σ the upper bound in (27) we conclude.

III. Proof of Theorem 2

We restrict our analysis to the range $p > n/(n+2)$. Owing to the mathematical theory in this range [4], all the convergence results that follow are rigorously justified. Let us first remark that the first derivative of the entropy power along a solution u_t of (7) can be easily obtained by (20) and (19), which yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{h}_p(u_t) = \mathrm{I}_p(u_t), \quad t > 0, \tag{28}$$

where I_p is the p-weighted Fisher information defined by (14). When $p \to 1$, identity (28) reduces to DeBruijn's identity, which connects Shannon's entropy functional with the Fisher information via the heat equation.

Using identity (28) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{N}_p(u_t) = \frac{\mu}{n}\,\mathrm{N}_p(u_t)\,\mathrm{I}_p(u_t),$$

which suggests to introduce the quantity

$$\mathrm{Q}_p(u) := \mathrm{N}_p(u)\,\mathrm{I}_p(u). \tag{29}$$

Notice that (16) is equivalent to

$$\mathrm{Q}_p(u) \ge \mathrm{Q}_p(\tilde{M}_p), \tag{30}$$

and the concavity of entropy power can be rephrased as the decreasing in time property of $t \mapsto \mathrm{Q}_p(u_t)$, since $\mu = 2 + n(p-1) > 0$.

In order to get (30) we just observe that Q_p is invariant with respect to the family of mass-preserving dilations

$$\mathcal{R}_a : f(x) \to \mathcal{R}_a f(x) := a^{-n} f\left(x/a\right), \quad a > 0,\ x\in\mathbb{R}^n,$$

of a given nonnegative density f.

In fact one can easily compute that

$$\mathrm{h}_p(\mathcal{R}_a f) = \mathrm{h}_p(f) + n\log a, \quad \mathrm{N}_p(\mathcal{R}_a f) = a^{\mu}\,\mathrm{N}_p(f), \tag{31}$$

and

$$\mathrm{I}_p(\mathcal{R}_a f) = a^{-\mu}\,\mathrm{I}_p(f), \tag{32}$$

so that

$$\mathrm{Q}_p(\mathcal{R}_a f) = \mathrm{Q}_p(f) \quad \text{for every } a > 0. \tag{33}$$

An application of (31) to (9) with $a := t^{1/\mu}$ yields (12).

Property (33) allows to identify the long-time behavior of the function $t \mapsto \mathrm{Q}_p(u_t)$ along (7). It is nonincreasing, and it will reach its infimum as time $t \to \infty$. The computation of the limit value uses in a substantial way the scaling invariance property. In fact, we can rescale u_t according to

$$\begin{aligned} U_t(x) &= t^{-n/\mu}\, u_t(x\, t^{-1/\mu}) \\ &= \mathcal{R}_{t^{1/\mu}} u_t \end{aligned} \tag{34}$$

where μ is defined in (2) as usual, so that

$$\mathrm{Q}_p(u_t) = \mathrm{Q}_p(U_t) \quad \text{for every } t > 0.$$

On the other hand, if $p > n/(n+2)$

$$\lim_{t\to\infty} U_t(x) = \tilde{M}_p(x), \tag{35}$$

the Barenblatt profile defined in (10). Therefore, denoting by $f(x) := u_0(x)$ the initial probability density, passing to the limit one obtains the isoperimetric inequality of (16).

The calculation of $\gamma_{n,p}$ is postponed to the Appendix.

IV. Sobolev inequality revisited

Inequality (16) can be rewritten in a form more suitable to functional analysis. Let $f(x)$ be a probability density in $\mathbb{R}^n$. Then, if $p > n/(n+2)$

$$\int_{\{f>0\}} \frac{|\nabla f^p(x)|^2}{f(x)}\, \mathrm{d}x \ge \gamma_{n,p} \left(\int_{\mathbb{R}^n} f^p(x)\, \mathrm{d}x \right)^{\frac{2+2n(p-1)}{n(p-1)}}. \tag{36}$$

If $n > 2$, the case $p = (n-1)/n$ is distinguished from the others, since it leads to

$$\frac{2+2n(p-1)}{n(p-1)} = 0, \quad \nu = \frac{1}{n},$$

and

$$\mathrm{N}_{1-1/n}(f) = \int_{\mathbb{R}^n} f^{1-1/n}(x)\, \mathrm{d}x.$$

In this case the concavity of $\mathrm{N}_{1-1/n}$ along (7) has been already known and has a nice geometric interpretation in terms of transport distances, see [16]. Note that the restriction $n > 2$ implies $(n-1)/n > n/(n+2)$. Hence, for $p = (n-1)/n$ we obtain that the probability density f satisfies the inequality

$$\int_{\{f>0\}} \frac{|\nabla f^{(n-1)/n}(x)|^2}{f(x)}\, \mathrm{d}x \ge \gamma_{n,(n-1)/n}. \tag{37}$$

The substitution $f = g^{2n/(n-2)}$ yields

$$\int_{\{f>0\}} \frac{|\nabla f^{(n-1)/n}(x)|^2}{f(x)}\, \mathrm{d}x = \left(\frac{2n-2}{n-2} \right)^2 \int_{\mathbb{R}^n} |\nabla g(x)|^2\, \mathrm{d}x.$$

Therefore, for any given function $g \ge 0$ such that $g(x)^{2n/(n-2)}$ is a probability density in $\mathbb{R}^n$, with $n > 2$, we obtain the inequality

$$\int_{\mathbb{R}^n} |\nabla g(x)|^2\, \mathrm{d}x \ge \left(\frac{n-2}{2n-2} \right)^2 \gamma_{n,(n-1)/n}. \tag{38}$$

A careful computation (see the Appendix) gives

$$\gamma_{n,(n-1)/n} = n\pi\, \frac{2^2(n-1)^2}{n-2} \left(\frac{\Gamma(n/2)}{\Gamma(n)} \right)^{2/n},$$

and a simple scaling argument finally shows that, if $g(x)^{2n/(n-2)}$ has a mass different from 1, g satisfies the *Sobolev* inequality [1], [19]

$$\int_{\mathbb{R}^n} |\nabla g(x)|^2\, \mathrm{d}x \ge \mathcal{S}_n \left(\int_{\mathbb{R}^n} g(x)^{2n/(n-2)}\, \mathrm{d}x \right)^{1-2/n}, \tag{39}$$

where

$$\mathcal{S}_n = n(n-2)\pi \left(\frac{\Gamma(n/2)}{\Gamma(n)} \right)^{2/n}$$

is the sharp Sobolev constant. Hence, Sobolev inequality with the sharp constant is a consequence of the concavity of Rényi entropy power of parameter $p = (n-1)/n$, when $n > 2$.

In all the other cases, the concavity of Rényi entropy power leads to Gagliardo-Nirenberg type inequalities with sharp constants, like the ones recently studied by Del Pino and Dolbeault [9], and Cordero-Erausquin, Nazaret, and Villani, [5] with different methods.

Appendix

Computation of the constants $\gamma_{n,p}$

Let us first observe that the rescaling invariance (33) allows for computing the quantity $\gamma_{n,p} = Q_p(\tilde{M}_p)$ along an arbitrary Barenblatt profile with unit mass. Instead of $\tilde{M}_p$ we can thus consider the Barenblatt function

$$\mathcal{B}_p(x) := \begin{cases} (C_p - |x|^2)_+^{1/(p-1)} & \text{if } p > 1, \\ (C_p + |x|^2)^{1/(p-1)} & \text{if } p < 1, \end{cases} \tag{40}$$

and $\gamma_{n,p} = Q(\mathcal{B}_p)$, where we are going to choose C_p so that $\mathcal{B}_p$ is a probability density.

Let us recall some useful formulas. The surface of the $n-1$ dimensional unit sphere $\mathbb{S}^{n-1}$ is given by $|\mathbb{S}^{n-1}| = 2\pi^{n/2}/\Gamma(n/2)$.

A. *The case* $p > 1$

Let us first consider the case $p > 1$. If $a > 0$, using the integral representation of Beta function we have

$$\begin{aligned} \int_{\mathbb{R}^n} (1-|x|^2)_+^a\, \mathrm{d}x &= |\mathbb{S}^{n-1}| \int_0^1 \rho^{n-1}(1-\rho^2)^a\, \mathrm{d}\rho \\ &= \frac{\pi^{n/2}}{\Gamma(n/2)} \int_0^1 t^{n/2-1}(1-t)^a\, \mathrm{d}t \\ &= \frac{\pi^{n/2}}{\Gamma(n/2)} B(\frac{n}{2}, a+1) \\ &= \pi^{n/2} \frac{\Gamma(a+1)}{\Gamma\left(\frac{n}{2}+a+1\right)}, \end{aligned}$$

where we used the identity $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

With this formula, we can evaluate quantities associated to the Barenblatt function (40).

Indeed, if

$$A_p = \int_{\mathbb{R}^n} (1-|x|^2)_+^{1/(p-1)}\,\mathrm{d}x = \pi^{n/2}\frac{\Gamma\left(\frac{p}{p-1}\right)}{\Gamma\left(\frac{n}{2}+\frac{p}{p-1}\right)}, \quad (41)$$

we obtain a Barenblatt of mass equal to one by choosing

$$C_p = A_p^{-\frac{2(p-1)}{n(p-1)+2}}. \quad (42)$$

Also [23]

$$\int_{\mathbb{R}^n} |x|^2 \mathcal{B}_p(x)\,\mathrm{d}x = \frac{n(p-1)}{(n+2)p-n}C_p, \quad (43)$$

and, since

$$\int_{\mathbb{R}^n} \mathcal{B}_p(x)^p\,\mathrm{d}x = \int_{\mathbb{R}^n} (C_p - |x|^2)\mathcal{B}_p(x)\,\mathrm{d}x,$$

one obtains

$$\int_{\mathbb{R}^n} \mathcal{B}_p(x)^p\,\mathrm{d}x = \frac{2p}{(n+2)p-n}C_p. \quad (44)$$

Since

$$\mathrm{I}_p(\mathcal{B}_p) = \frac{4p^2}{(p-1)^2}\frac{1}{\int_{\mathbb{R}^n}\mathcal{B}_p(x)^p\,\mathrm{d}x}\int_{\mathbb{R}^n}|x|^2\mathcal{B}_p(x)\,\mathrm{d}x, \quad (45)$$

thanks to (43) and (44), we reckon the values of the p-Fisher information $\mathrm{I}_p(f)$ defined in (14) and of the Rényi entropy $\mathrm{h}_p(f)$, associated to $\mathcal{B}_p$

$$\mathrm{h}_p(\mathcal{B}_p) = \frac{1}{1-p}\log\frac{2p}{(n+2)p-n}C_p, \quad (46)$$

and

$$\mathrm{I}_p(\mathcal{B}_p) = n\frac{2p}{p-1}. \quad (47)$$

Hence, if $p>1$ the value of the constant $\gamma_{n,p}$ is

$$\gamma_{n,p} = n\pi\frac{2p}{p-1}\cdot \cdot\left(\frac{\Gamma\left(\frac{p}{p-1}\right)}{\Gamma\left(\frac{n}{2}+\frac{p}{p-1}\right)}\right)^{2/n}\left(\frac{(n+2)p-n}{2p}\right)^{\frac{2+n(p-1)}{n(p-1)}}. \quad (48)$$

B. The case $p<1$

Analogous computations can be done in the case $p<1$. In this case

$$A_p = \int_{\mathbb{R}^n}(1+|x|^2)^{1/(p-1)}\,\mathrm{d}x = \pi^{n/2}\frac{\Gamma\left(\frac{1}{1-p}-\frac{n}{2}\right)}{\Gamma\left(\frac{1}{1-p}\right)}, \quad (49)$$

while

$$\int_{\mathbb{R}^n}|x|^2\mathcal{B}_p(x)\,\mathrm{d}x = \frac{n(1-p)}{(n+2)p-n}C_p. \quad (50)$$

Note that, if $p<1$, the second moment and the p-weighted Fisher information of the Barenblatt is bounded if and only if $p>n/(n+2)$. Therefore, the computations that follow are restricted to this domain of p. If this is the case, formula (44) still holds, while

$$\mathrm{I}_p(\mathcal{B}_p) = n\frac{2p}{1-p}. \quad (51)$$

Finally, if $n/(n+2)<p<1$,

$$\gamma_{n,p} = n\pi\frac{2p}{1-p}\cdot \cdot\left(\frac{\Gamma\left(\frac{1}{1-p}-\frac{n}{2}\right)}{\Gamma\left(\frac{1}{1-p}\right)}\right)^{2/n}\left(\frac{(n+2)p-n}{2p}\right)^{\frac{2+n(p-1)}{n(p-1)}}. \quad (52)$$

References

[1] T. Aubin, Problémes isopérimétriques et espaces de Sobolev, *J. Differential Geometry* **11**, pp. 573–598, 1976.

[2] D. Bakry and M. Emery, Diffusions hypercontractives, in *Sém. Proba. XIX* (Lecture Notes in Mathematics). Berlin, Germany: Springer-Verlag, 1985, pp. 177–206.

[3] N.M. Blachman, The convolution inequality for entropy powers, *IEEE Trans. Inform. Theory*, vol. IT-**11**, 267–271, Apr. 1965.

[4] A. Blanchet, M. Bonforte, J. Dolbeault, G. Grillo and J. L Vázquez, Asymptotics of the Fast Diffusion Equation via Entropy Estimates, *Arch. Ration. Mech. Anal.* vol. **191** (2) 347–385, 2009.

[5] D. Cordero-Erausquin, B. Nazaret, and C. Villani, A mass–transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities, *Advances in Mathematics* **182** 307–332, 2004.

[6] M. Costa, A new entropy power inequality, *IEEE Trans. Inform. Theory*, vol. IT-**31** 751–760, Nov. 1985.

[7] T.M. Cover and J.A. Thomas, *Elements of Information Theory* New York, USA: Wiley, 1992.

[8] J. Crank, *The Mathematics of Diffusion*, Oxford, UK: Clarendon Press, 1956.

[9] M. Del Pino, and J. Dolbeault, Best constants for Gagliardo-Nirenberg inequalities and application to nonlinear diffusions, *J. Math. Pures Appl.* **81** 847–875, 2002.

[10] A. Dembo, A simple proof of the concavity of the entropy power with respect to the variance of additive normal noise, *IEEE Trans. Inform. Theory*, vol. **35**, 887–888, July 1989.

[11] A. Dembo, T. Cover, and J. Thomas, Information theoretic inequalities, *IEEE Trans. Inform. Theory*, vol. **37**, 1501–1518, Nov. 1991.

[12] R.J. Gardner, The Brunn-Minkowski inequality, *Bull. Amer. Math. Soc.* **39** 355–405, 2002.

[13] D. Guo, S. Shamai, and S. Verdú, Proof of entropy power inequalities via MMSE, in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, July 2006, 10111015.

[14] O. Johnson, and C. Vignat, Some results concerning maximum Rényi entropy distributions *Ann. I. H. Poincaré*, vol. **43** 339–351, 2007.

[15] E. Lutwak, D. Yang, and G. Zhang, Cramér-Rao and moment-entropy inequalities for Rényi entropy and generalized Fisher information, *IEEE Trans. Inform. Theory*, vol. **51**, 473–478, Feb. 2005.

[16] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Commun. Part. Diff. Eq.*, vol. **26**, 101–174, 2001.

[17] F. Otto, M. Westdickenberg, Eulerian calculus for the contraction in the Wasserstein distance. *SIAM Journal on Mathematical Analysis*, vol. **37**, 1227–1255, 2005.

[18] O. Rioul, Information Theoretic Proofs of Entropy Power Inequalities, *IEEE Trans. Inform. Theory*, vol. **57**, 33–55, Jan. 2011.

[19] G. Talenti, Best constant in Sobolev inequality, *Ann. Mat. Pura Appl.* **110** 353–372, 1976.

[20] G. Toscani, A central limit theorem for solutions of the porous medium equation, *J. Evol. Equ.* **5** 185–203, 2005.

[21] G. Toscani, An information-theoretic proof of Nash's inequality, *Rend. Lincei Mat. Appl.* **24** 83–93, 2013

[22] J.L. Vazquez, Asymptotic behaviour for the porous medium equation posed in the whole space, *J. Evol. Equ.*, **3** 67–118, 2003.

[23] J.L. Vazquez, *The Porous Medium Equation: Mathematical Theory*, Oxford, UK: Oxford University Press, 2007

[24] C. Villani, A Short Proof of the Concavity of Entropy Power, vol. **46**, 1695–1696, July 2000.