Heat equation and convolution inequalities

Giuseppe Toscani

Abstract. It is known that many classical inequalities linked to convolutions can be obtained by looking at the monotonicity in time of convolutions of powers of solutions to the heat equation, provided that both the exponents and the coefficients of diffusions are suitably chosen and related. This idea can be applied to give an alternative proof of the sharp form of the classical Young's inequality and its converse, to Brascamp–Lieb type inequalities, Babenko's inequality and Prékopa–Leindler inequality as well as the Shannon's entropy power inequality. This note aims in presenting new proofs of these results, in the spirit of the original arguments introduced by Stam [35] to prove the entropy power inequality.

Mathematics Subject Classification (2010). Primary 39B62; Secondary 94A17.

Keywords. Lyapunov functionals, heat equation, inequalities in sharp form.

1. Introduction

The purpose of this note is to present various results concerned with the monotonicity in time of the convolution of powers of solutions to the heat equation. The main reason behind this investigation is that many functional inequalities can be viewed as the consequence of the tendency of various Lyapunov functionals defined in terms of powers of the solution to the heat equation to reach their extremal values as time tends to infinity. The discovery of a Lyapunov functional which allows to prove Young inequality and its converse [8], is only one of the possible application of this idea (cf. also [36, 37, 38] for a connection of these results with information theory). While the inequalities are not new, and some of the results we present have been obtained before, what is new is the approach to the problem, which takes into

Date: September 16, 2013. This work has been completed within the activities of the GNFM group of INDAM. The author acknowledges support by MIUR project "Optimal mass transportation, geometrical and functional inequalities with applications".

account the information-theoretical meaning of inequalities for convolutions, and consequently allows to obtain clean and relatively simple new proofs.

The prototype of these monotonous in time convolutions is as follows. Let n be an integer, and let α_j, $j = 1, \dots, n$, be positive real numbers such that

$$\sum_{j=1}^{n} \alpha_j = n - 1. \tag{1.1}$$

Let $f_j(x)$, $j = 1, \dots, n$, be non-negative functions on $\mathbb{R}^d$, $d \geq 1$, such that $f_j \in L^1(\mathbb{R}^d)$. For any given j, $j = 1, \dots, n$, we denote by $u_j(x,t)$ the solution to the heat equation (2.1) with the diffusion coefficients κ_j

$$\frac{\partial u_j(x,t)}{\partial t} = \kappa_j \Delta u_j(x,t),$$

such that

$$\lim_{t \to 0^+} u_j(x,t) = f_j(x).$$

We consider the n-th convolution

$$w(x,t) = u_1^{\alpha_1} * u_2^{\alpha_2} * \cdots * u_n^{\alpha_n}(x,t). \tag{1.2}$$

where, as usual the n-th convolution of the functions $g_j(x) \in L^1(\mathbb{R}^d)$, $j = 1, 2, \dots, n$ reads

$$g_1 * g_2 * \cdots * g_n(x) =$$

$$\int_{(\mathbb{R}^d)^{n-1}} g_1(x - x_1) \cdots g_{n-1}(x_{n-2} - x_{n-1}) g_n(x_{n-1})\, dx_1 dx_2 \cdots dx_{n-1}.$$

Then, a natural question arises. Can we fix the diffusion coefficients in the heat equation in such a way that $w(x,t)$ behaves monotonically in time? Note that the choice of condition (1.1) is forced by the fact that we want that the monotonicity of $w(x,t)$, $t > 0$ has to hold at least if $u_j(x,t)$ is the fundamental solution to the heat equation, $j = 1, 2, \dots, n$. In this case, in fact, computations are explicit, and, provided condition (1.1) is satisfied, $w(x,t)$ is increasing in time independently of the choice of the diffusion coefficients (cf. Section 2). In the general case, however, the monotonicity in time of the n-th convolution can be proven under more restrictive assumptions both on the numbers α_j, and only for a unique choice of the diffusion coefficients κ_j (cf. Lemma 3.1).

The interest in the monotonicity of the convolution of powers of solutions to the heat equation is linked to its consequences. Indeed, the discovery of the monotonicity of $w(x,t)$ for a special choice of the diffusion coefficients translates immediately to the proof of an inequality for convolutions in sharp form. Let n be an integer, and let p_j, $j = 1, \dots, n$, be real numbers such that $1 \leq p_j \leq +\infty$ and $\sum_{j=1}^{n} p_j^{-1} = n - 1$. Let $f_j(x)$, $j = 1, \dots, n$, be functions on $\mathbb{R}^d$, $d \geq 1$, such that $f_j \in L^{p_j}(\mathbb{R}^d)$. In Theorem 3.4 we will show that the monotonicity of $w(x,t)$ implies the following inequality for convolutions:

$$\sup_x |f_1 * f_2 * \cdots * f_n| \leq \prod_{j=1}^{n} C_{p_j}^d \|f_j\|_{p_j}. \tag{1.3}$$

In (1.3), the constant C_p which defines the sharp constant is given by

$$C_p^2 = \frac{p^{1/p}}{p'^{1/p'}}, \tag{1.4}$$

where primes always denote dual exponents, $1/p + 1/p' = 1$. Also, the expression of the best constant in (1.3), in the case in which the functions f_j are probability density functions, is obtained by assuming that the functions f_j are suitable Gaussian densities [27]. This expression naturally appears in this monotonicity approach by considering that for large times the solution to the heat equation behaves as the self-similar Gaussian profile.

Note that inequality (1.3) implies

$$\left| \int f_1(x_1) f_2(x_1 - x_2) \cdots f_n(x_{n-1})\, dx_1 dx_2 \cdots dx_{n-1} \right| \leq \prod_{j=1}^{n} C_{p_j}^d \|f_j\|_{p_j}, \tag{1.5}$$

which is a particular case of the general inequalities obtained by Brascamp and Lieb [13], which are nowadays known as the Brascamp–Lieb inequalities.

Inequality (1.3) is closely related to the monotonicity property of the functional given by L^∞-norm of the n-th convolution $w(x,t)$. Naturally one could ask if a similar property holds for the L^r-norm of $w(x,t)$, where $r \geq 0$. Also in this case, the monotonicity in time can be proven under suitable assumptions both on the numbers α_j, and only for a unique choice of the diffusion coefficients κ_j. The study of the monotonicity in time of $\|w(t)\|_r$ is connected with the classical Young's inequality in sharp form ($r > 1$), or with its reverse form ($r < 1$).

Last, the limiting cases $r \to 1$ and $r \to 0$ lead to the monotonicity in time of Shannon's entropy and of the Renyi entropy of order 0 [16]. The monotonicity here leads to the entropy power inequality of Shannon [34], and to the Prékopa–Leindler inequality [25, 31, 32], respectively.

Therefore, all these well-known functional inequalities can be seen into a unified framework, as consequences of the monotonicity in time of the n-convolution of powers of solutions to the heat equation.

As noticed in [38], the heat equation started to be used as a powerful instrument to obtain mathematical inequalities in sharp form in the years between the late fifties to mid sixties. To our knowledge, the first application of this idea can be found in two papers by Linnik [28] and Stam [35] (cf. also Blachman [11]), published in the same year and concerned with two apparently disconnected arguments. Stam [35] was motivated by the finding of a rigorous proof of Shannon's entropy power inequality [34], while Linnik [28] used the information measures of Shannon and Fisher in a proof of the central limit theorem of probability theory. Also, in the same years, the heat equation has been used in the context of kinetic theory of rarefied gases by McKean [30] to investigate that large-time behaviour of Kac caricature of a Maxwell gas. There, various monotonicity properties of the derivatives of Shannon's entropy along the solution to the heat equation have been enlightened.

The huge potentialities of the use of the heat equation to prove inequalities have been rediscovered in more recent times by Carlen, Lieb and Loss [14], that first introduced a Lyapunov functional of solutions to the heat equation which allows to prove Young's inequality and its converse for functions of one variable. Later on, Bennett Carbery Christ and Tao [10] were able to extend the result in [14] to general functions. Other very closely-related works can be found in papers of Bennett and Bez [8], Borell [12], Barthe and Cordero-Erausquin [3] and Barthe-Huet [4]. In particular, Young's inequality and its converse have been proven by Bennett and Bez [8] by showing that a suitable functional of the convolution of powers to the solution to the heat the heat equation exhibits monotonicity properties.

As often happens, however, the seminal ideas of Stam [11, 35] remained confined within the framework of information theory, where, however, functional inequalities gained a lot of interest, in reason of their connections with properties of Shannon's and Renyi's entropies [19]. A notable exception to this confinement is a recent paper by Gardner [20], that clarifies the relationship between the Brunn-Minkowski inequality and other inequalities in geometry and analysis. In [20], clear connections between the entropy power inequality of information theory and Young's inequality and others are described in details, together with an exhaustive list of references.

As far as the classical Young's inequality is concerned, the original proof of the sharp form is due to Beckner [6] and Brascamp and Lieb [13]. In [13] Brascamp and Lieb also proved the sharp form of Young inequality also in the so-called reverse case. A different proof of this sharp reverse Young inequality was subsequently done by Barthe [2]. In their recent paper, Young's inequality has been seen in a different light by Bennett and Bez [8] (cf. also [7, 10, 14]). In this paper, Young's inequality is derived by looking at the monotonicity properties of a suitable functional of the convolution of powers to the solution to the heat the heat equation. In this respect, the arguments of [8] are close to the present ones.

The connections of the sharp form of Young's inequality with the Prékopa–Leindler inequality has been enlightened by Brascamp and Lieb [13]. Then, the connection of Young's inequality with Shannon's entropy power inequality has been noticed by Lieb [26].

2. Heat equation, Lyapunov functionals and dilation invariance

We begin by recalling some properties of the solution to the heat equation in $\mathbb{R}^d$, $d \geq 1$

$$\frac{\partial u(x,t)}{\partial t} = \kappa \Delta u(x,t), \tag{2.1}$$

where $\kappa > 0$ is the (constant) diffusion coefficient. In the rest of the paper, for the sake of simplicity we will assume that the initial datum is a non-negative

integrable function $f(x)$, that is

$$\int_{\mathbb{R}^d} f(x)\,dx = \mu < +\infty. \tag{2.2}$$

This assumption will not affect the generality of the results that follow. The solution to equation (2.1) is given by the function $u(x,t) = f * M_{2\kappa t}(x)$, convolution of the initial datum with the fundamental solution $M_{2\kappa t}$, where $M_\sigma(x)$, for $\sigma > 0$, denotes the Gaussian density in $\mathbb{R}^d$ of mean 0 and variance $d\sigma$

$$M_\sigma(x) = \frac{1}{(2\pi\sigma)^{d/2}} \exp\left(-\frac{|x|^2}{2\sigma}\right). \tag{2.3}$$

For large times, the solution to the heat equation approaches the fundamental solution. This large-time behaviour can be better specified by saying that the solution to the heat equation (2.1) satisfies a property which can be defined as the *central limit property*. Suppose the initial datum f in equation is such that $(1+|x|^\delta)f(x) \in L^1(\mathbb{R}^d)$ for some constant $\delta > 0$ (typically $\delta = 2$). Then, if

$$U(x,t) = \left(\sqrt{1+2t}\right)^d\, u(x\sqrt{1+2t}, t). \tag{2.4}$$

$U(x,t)$ tends in $L^1(\mathbb{R}^d)$ towards a limit function as time goes to infinity, and this limit function is a Gaussian function

$$\lim_{t\to\infty} U(x,t) = M_\kappa(x) \int_{\mathbb{R}^n} f(x)\,dx = \mu M_\kappa(x). \tag{2.5}$$

This convergence property, as well as convergence in other stronger norms, can be achieved easily by resorting to Fourier transform, or by exploiting the relationship between the heat equation and the Fokker–Planck equation [15, 21] (cf. also [5] for recent results and references). We note that the passage $u(x,t) \to U(x,t)$ defined by (2.4) is mass preserving, that is

$$\int_{\mathbb{R}^d} U(x,t)\,dx = \int_{\mathbb{R}^d} u(x,t)\,dx. \tag{2.6}$$

An important remark concerns the necessity to introduce condition (1.1) in our analysis. Since the fundamental solution is a Gaussian probability density, it is closed under the operation of convolution [24], namely

$$M_{\sigma_1} * M_{\sigma_2}(x) = M_{\sigma_1+\sigma_2}(x).$$

Hence, if we consider at time $t > 0$ the convolution of n powers of the fundamental solutions of heat equations with diffusion coefficients κ_j, $j = 1, 2, \ldots, n$, we obtain

$$M_{2\kappa_1 t}^{\alpha_1} * M_{2\kappa_2 t}^{\alpha_2} * \cdots * M_{2\kappa_n t}^{\alpha_n} =$$

$$\prod_{j=1}^n (4\pi\kappa_j t)^{-\alpha_j d/2} \left(4\pi\frac{\kappa_j}{\alpha_j}t\right)^{d/2} M_{2t\kappa_1/\alpha_1} * M_{2t\kappa_2/\alpha_2} * \cdots * M_{2t\kappa_n/\alpha_n} =$$

$$\prod_{j=1}^n (4\pi\kappa_j t)^{-\alpha_j d/2} \left(4\pi\frac{\kappa_j}{\alpha_j}t\right)^{d/2} M_{2\Sigma t}, \tag{2.7}$$

where

$$\Sigma = \sum_{j=1}^{n} \frac{\kappa_j}{\alpha_j}.$$

In the expression above the time-dependent quantity in front of the exponential is given by

$$\phi(t) = t^{-\frac{d}{2}\sum_{j=1}^{n}\alpha_j + \frac{1}{2}(n-1)}.$$

Therefore, if the exponents α_j satisfy condition (1.1), independently of the values of the diffusion coefficients κ_j, $\phi(t) = 1$, and

$$M^{\alpha_1}_{2\kappa_1 t} * M^{\alpha_2}_{2\kappa_2 t} * \cdots * M^{\alpha_n}_{2\kappa_n t} = \Sigma_1 \exp\left\{-|x|^2/4\Sigma t\right\}, \tag{2.8}$$

where Σ_1 denotes the constant

$$\Sigma_1 = \left(\frac{\kappa_j}{\alpha_j}\right)^{n/2} \Sigma^{-d/2} \prod_{j=1}^{n} (\kappa_j)^{-\alpha_j d/2}.$$

Consequently, independently of the values of the diffusion coefficients κ_j, if the exponents α_j satisfy condition (1.1), for every $x \in \mathbb{R}^d$

$$\frac{d}{dt} M^{\alpha_1}_{2\kappa_1 t} * M^{\alpha_2}_{2\kappa_2 t} * \cdots * M^{\alpha_n}_{2\kappa_n t} \geq 0. \tag{2.9}$$

This property is obviously restricted to a set of positive constants α_j satisfying (1.1).

A second argument is to use the evolution equation for a power of the solution to the heat equation. If $\alpha > 0$ is a positive constant, and $u(x,t)$ solves (2.1), then $u^\alpha(x,t)$ solves

$$\frac{\partial u^\alpha(x,t)}{\partial t} = \kappa \left[\Delta u^\alpha(x,t) + \alpha(1-\alpha) u^\alpha(x,t) |\nabla \log u(x,t)|^2\right]. \tag{2.10}$$

Equation (2.10) is particularly adapted to work with convolutions of powers. Note that equation (2.10) connects in a natural way dual exponents. In fact, if $\alpha = 1/p$, with $p > 1$, equation (2.10) takes the form

$$\frac{\partial u^{1/p}(x,t)}{\partial t} = \kappa \left[\Delta u^{1/p}(x,t) + \frac{1}{pp'} u^{1/p}(x,t) |\nabla \log u(x,t)|^2\right].$$

Our last ingredient is to consider the evolution in time of Lyapunov functionals of solutions to the heat equation which are *dilation invariant*, that is invariant with respect to the scaling

$$f(x) \to f_a(x) = a^d f(ax), \quad a > 0, \tag{2.11}$$

In reason of (2.5), this property allows to reckon immediately the (bounded) limit value of the underlying functional, as time goes to infinity.

One simple example will clarify why dilation invariance is a key ingredient of our strategy. Given a solution to the heat equation (2.1) let us consider its (finite) Shannon's entropy

$$H(u(t)) = -\int_{\mathbb{R}^d} u(x,t) \log u(x,t)\, dx$$

A simple computation shows that the time derivative of $H(u(t))$ is nonnegative [16], and it converges to infinity as time goes to infinity. Indeed (DeBruin's identity [16])

$$\frac{d}{dt}H(u(t)) = \kappa I(u(t)) = \kappa \int_{\{u(t)>0\}} \frac{|\nabla u(x,t)|^2}{u(x,t)}\, dx > 0. \tag{2.12}$$

In (2.12) $I(u(t))$ defines as usual the Fisher infomation of the probability density $u(x,t)$. This time behavior is consequence of the fact that Shannon's entropy is not dilation invariant

$$H(u_a) = H(u) - d\log a. \tag{2.13}$$

It is easily checked that the second moment of a probability density function scales according to

$$E(u_a) = \int_{\mathbb{R}^d} |x|^2 u_a(x)\, dx = \frac{1}{a^2}E(u). \tag{2.14}$$

Hence, if the initial probability density in the heat equation has bounded second moment, a dilation invariant functional of u is obtained by coupling Shannon's entropy of u with the logarithm of the second moment of u

$$\Gamma(t) = H(u(t)) - \frac{d}{2}\log E(u(t)). \tag{2.15}$$

Following [38], let us compute the time derivative of $\Gamma(t)$. We obtain

$$\frac{d}{dt}\Gamma(t) = \kappa\left(I(u(t)) - \frac{d^2}{E(u(t))}\right), \tag{2.16}$$

which is a direct consequence of DeBruin's identity (2.12), and of the time evolution of the second moment of the solution to the heat equation,

$$\frac{d}{dt}E(u(t)) = \frac{d}{dt}\int_{\mathbb{R}^d} |x|^2 u(x,t)\, dx = 2\kappa d \int_{\mathbb{R}^d} u(x,t)\, dx = 2\kappa d.$$

The right-hand side of (2.16) is nonnegative. This can be easily shown by an argument which is often used in this type of proofs. It holds

$$0 \le \int_{\mathbb{R}^d} \left(\frac{\nabla u(x,t)}{u(x,t)} + \frac{dx}{E(u(x,t))}\right)^2 u(x,t)\, dx =$$

$$I(u(t)) + \frac{d^2}{E(u(t))^2}\int_{\mathbb{R}^d} |x|^2 u(x,t)\, dx + 2\frac{d}{E(u(t))}\int_{\mathbb{R}^d} x\cdot\nabla u(x,t)\, dx =$$

$$I(u(t)) + \frac{d^2}{E(u(t))} - 2\frac{d^2}{E(u(t))} = I(u(t)) - \frac{d^2}{E(u(t))}. \tag{2.17}$$

Note that, since $u(x,t)$ is the (smooth) solution to the heat equation (2.1), equality to zero in (2.17) holds if and only if

$$\frac{\nabla u(x)}{u(x)} + \frac{dx}{E(u(x))} = 0$$

for all $x \in \mathbb{R}^d$. This condition can be rewritten as

$$\nabla \left(\log u(x) - \frac{d}{E(u(x))} \frac{x^2}{2} \right) = 0 \tag{2.18}$$

which identifies the probability density $u(x,t)$ as a Gaussian density in $\mathbb{R}^d$. By (2.17), this also shows that, among all densities with variance, Fisher's information takes its minimum value in correspondence to the Gaussian density (2.3), where $\sigma = E(u)/d$.

Thus, unless the initial value is a Gaussian density, the functional $\Gamma(t)$ is monotone increasing, and it will reach its (eventual) maximum value as time $t \to \infty$. The computation of the limit value uses in a substantial way the dilation invariance of Γ. In fact, at each time $t > 0$, the value of $\Gamma(t)$ does not change if we scale $u(x,t)$ according to

$$u(x,t) \to U(x,t) = \left(\sqrt{1+2t}\right)^d \, u(x\sqrt{1+2t}, t). \tag{2.19}$$

By the *central limit property*

$$\lim_{t\to\infty} U(x,t) = M_\kappa(x) \int_{\mathbb{R}^d} f(x)\, dx \tag{2.20}$$

Therefore, passing to the limit one obtains

$$\Gamma(0) = H(f) - \frac{d}{2} \log E(f) \leq H(M_\kappa) - \frac{d}{2} \log E(M_\kappa) = \frac{d}{2} \log \frac{2\pi e}{d}. \tag{2.21}$$

This inequality holds for all probability density functions with bounded second moment, and does not require that the second moment of f equals the second moment of the Gaussian density.

This relatively simple example contains the main ingredients we will use to obtain inequalities in sharp form: monotonicity of a functional of solutions to the heat equation coupled with the dilation invariance property. Following this line of thought, the rest of the paper will be devoted to the proof of various inequalities for convolutions in sharp form. For the sake of simplicity, we will present most of the proofs in dimension $d = 1$. The corresponding higher-dimensional inequalities can be deduced as well by making use of standard properties of the Gaussian function.

3. The monotonicity of convolutions

Let n be an integer, and let p_j, $j = 1, \ldots, n$, be real numbers such that

$$1 \leq p_j \leq +\infty; \quad \sum_{j=1}^{n} \frac{1}{p_j} = n - 1. \tag{3.1}$$

Let $f_j(x)$, $j = 1, \ldots, n$, be non-negative functions on $\mathbb{R}^d$, $d \geq 1$, such that $f_j \in L^1(\mathbb{R}^d)$. For any given j, $j = 1, \ldots, n$, we denote by $u_j(x,t)$ the solution to the heat equation (2.1) with the diffusion coefficients κ_j

$$\frac{\partial u_j(x,t)}{\partial t} = \kappa_j \Delta u_j(x,t), \tag{3.2}$$

such that

$$\lim_{t\to 0^+} u_j(x,t) = f_j(x). \tag{3.3}$$

The following Lemma shows that there is a (unique) choice of the diffusion coefficients in the heat equation such that $w(x,t)$ behaves monotonically in time.

Lemma 3.1. *Let $w(x,t)$ be the n-th convolution*

$$w(x,t) = u_1^{1/p_1} * u_2^{1/p_2} * \cdots * u_n^{1/p_n}(x,t) \tag{3.4}$$

where the functions $u_j(x,t)$, $j = 1,2,\ldots,n$, are solutions to the heat equation corresponding to the initial values $0 \le f_j(x) \in L^1(\mathbb{R}^d)$. Then, if for each j the exponents p_j satisfy conditions (3.1) and the diffusion coefficients are given by $\kappa_j = (p_j p_j')^{-1}$, $w(x,t)$ is monotonically increasing in time from

$$w(x,t=0) = f_1^{1/p_1} * f_2^{1/p_2} * \cdots * f_n^{1/p_n}(x).$$

Moreover, $w(x,t)$ remains constant in time if and only if $f_j(x)$, $j = 1,2,\ldots,n$, is a multiple of a Gaussian density of variance $d\kappa_j$.

Proof. For the sake of simplicity, we will prove the Lemma for $d = 1$. As the proof shows, however, analogous computations can be done in higher dimension.

Since $\sum_{j=1}^n p_j^{-1} = n-1$, Hölder inequality implies that

$$\left| \int f_1(x_1)^{1/p_1} \cdots f_n(x_{n-1})^{1/p_n}\, dx_1 dx_2 \cdots dx_{n-1} \right| \le \prod_{j=1}^n \left(\int_{\mathbb{R}} |f_j(x)|\, dx \right)^{1/p_j}.$$

Hence, also

$$f_1^{1/p_1} * f_2^{1/p_2} * \cdots * f_n^{1/p_n}(x) \le \prod_{j=1}^n \left(\int_{\mathbb{R}} |f_j(x)|\, dx \right)^{1/p_j}, \tag{3.5}$$

and, since the right-hand side of (3.5) depends only on the L^1-norms of the functions, which are preserved by the heat equation, the function $w(x,t)$ is bounded for all subsequent times $t > 0$. Also, using basic considerations on the heat equation, it is sufficient to prove the increasing property of $w(t)$ for very smooth initial data f_j, $j = 1,2,\ldots,n$, with fast decay at infinity. This will enable us to have the *central limit property*. In order not to worry about derivatives of logarithms, which will often appear in the proof, we may also impose that $|\frac{d}{dx} \log f_j(x)| \le C(1+|x|^2)$ for some positive constant C. The general case will follow by density [29].

For a given $x \in \mathbb{R}$, let us evaluate the time derivative of the n-th convolution $w(x,t)$. We obtain

$$\frac{\partial w(x,t)}{\partial t} = \left(\sum_{j=1}^n \kappa_j \right) \frac{\partial^2 w(x,t)}{\partial x^2} + \sum_{j=1}^n \frac{\kappa_j}{p_j p_j'} R_j(x,t), \tag{3.6}$$

where, for $j = 1, 2, \dots, n$

$$R_j(x) = \int u_1(x-x_1)^{1/p_1} \cdots u_n(x_{n-1})^{1/p_n} \left| \frac{\partial \log u_j}{\partial x}(x_{j-1} - x_j) \right|^2 dx_1 \cdots dx_{n-1} \tag{3.7}$$

Indeed,

$$\frac{\partial w}{\partial t} = \frac{\partial u_1^{1/p_1}}{\partial t} * u_2^{1/p_2} * \cdots * u_n^{1/p_n} + u_1^{1/p_1} * \frac{\partial u_2^{1/p_2}}{\partial t} * \cdots * u_n^{1/p_n} + \dots$$

$$+ u_1^{1/p_1} * u_2^{1/p_2} * \cdots * \frac{\partial u_n^{1/p_n}}{\partial t},$$

and the time derivative of each term on the right-hand side can be evaluated by considering that the functions $u_j(x,t)$, $j = 1, 2, \dots, n$ satisfy the heat equation (3.2) (with diffusion coefficients κ_j, $j = 1, 2, \dots, n$). Hence

$$\frac{\partial u_1^{1/p_1}}{\partial t} * u_2^{1/p_2} * \cdots * u_n^{1/p_n} = \kappa_1 \frac{\partial^2 u_1^{1/p_1}}{\partial x^2} * u_2^{1/p_2} * \cdots * u_n^{1/p_n} +$$

$$\frac{\kappa_1}{p_1 p_1'} \left(\left| \frac{\partial \log u_1}{\partial x} \right|^2 u_1^{1/p_1} \right) * u_2^{1/p_2} * \cdots * u_n^{1/p_n} =$$

$$\kappa_1 \frac{\partial^2 w}{\partial x^2} + \frac{\kappa_1}{p_1 p_1'} \left(\left| \frac{\partial \log u_1}{\partial x} \right|^2 u_1^{1/p_1} \right) * u_2^{1/p_2} * \cdots * u_n^{1/p_n}. \tag{3.8}$$

An analogous formula holds for the other indexes $j \geq 2$. Note that in (3.8) we used the convolution property

$$\frac{\partial^2}{\partial x^2} f * g(x) = \int f''(x-y) g(y)\, dy = \int f'(x-y) g'(y)\, dy = \int f(x-y) g''(y)\, dy. \tag{3.9}$$

By property (3.9), it holds that, for each pair of indexes (i,j) with $i, j = 1, 2, \dots, n$, with $x_0 = x$ and $x_n = 0$

$$(f_1 * f_2 * \cdots * f_n)'' =$$

$$\int f_1(x-x_1) \dots f_n(x_{n-1}) (\log f_i(x_{i-1} - x_i))' (\log f_j(x_{j-1} - x_j))'\, dx_1 \dots dx_{n-1}.$$

Hence, if we take a set of positive constants $a_{i,j}$'s, $i, j = 1, 2, \dots, n$, such that $\sum_{i \neq j} a_{i,j} = 1$, we can express the second derivative of a convolution as

$$(f_1 * f_2 * \cdots * f_n)'' = \sum_{i \neq j} a_{i,j} \int f_1(x - x_1) \dots f_n(x_{n-1}) \cdot$$

$$\cdot (\log f_i(x_{i-1} - x_i))' (\log f_j(x_{j-1} - x_j))'\, dx_1 \dots dx_{n-1}.$$

This shows that, for any set of positive values $a_{i,j}$ such that $\sum_{i \neq j} a_{i,j} = 1$, with $x_0 = x$ and $x_n = 0$, one has

$$\frac{\partial^2 w}{\partial x^2} = \sum_{i \neq j} \frac{a_{i,j}}{p_i p_j} \int u_1^{1/p_1}(x - x_1) \dots u_n^{1/p_n}(x_{n-1}) \cdot$$

$$\cdot (\log u_i(x_{i-1} - x_i))' (\log u_j(x_{j-1} - x_j))'\, dx_1 \dots dx_{n-1}. \tag{3.10}$$

Finally, by setting, for $j = 1, 2, \ldots, n$

$$L_j = (\log u_j(x_{j-1} - x_j))', \tag{3.11}$$

we can rewrite (3.6) in the following way:

$$\frac{\partial w(x,t)}{\partial t} = \int u_1^{1/p_1}(x - x_1) \ldots u_n^{1/p_n}(x_{n-1}) \cdot$$

$$\left(\sum_{j=1}^{n} \frac{\kappa_j}{p_j p_j'} L_j^2 + \sum_{l=1}^{n} \kappa_l \sum_{i \neq j} \frac{a_{i,j}}{p_i p_j} L_i L_j \right) dx_1 \ldots dx_{n-1}. \tag{3.12}$$

The sign of the time derivative of $w(x,t)$ depends on the quantity

$$\mathfrak{L}(u_1, \cdots u_n) = \sum_{j=1}^{n} \frac{\kappa_j}{p_j p_j'} L_j^2 + \sum_{l=1}^{n} \kappa_l \sum_{i \neq j} \frac{a_{i,j}}{p_i p_j} L_i L_j. \tag{3.13}$$

Let us set the coefficient of diffusion $\kappa_j = (p_j p_j')^{-1}$, and define $Q_j = L_j / p_j$, for $j = 1, 2, \ldots n$. Then

$$\mathfrak{L} = \sum_{j=1}^{n} \left(\frac{1}{p_j'} \right)^2 Q_j^2 + \sum_{l=1}^{n} \frac{1}{p_l p_l'} \sum_{i \neq j} a_{i,j} Q_i Q_j. \tag{3.14}$$

Now, recall that

$$\sum_{j=1}^{n} \frac{1}{p_j} = n - 1$$

implies that, for all $j = 1, 2, \ldots n$

$$\frac{1}{p_j} = \sum_{i \neq j} \frac{1}{p_i'}.$$

Consequently

$$\sum_{l=1}^{n} \frac{1}{p_l p_l'} = \sum_{i \neq j} \frac{1}{p_i' p_j'}.$$

Therefore

$$\mathfrak{L} = \sum_{j=1}^{n} \left(\frac{1}{p_j'} \right)^2 Q_j^2 + \sum_{i \neq j} \frac{1}{p_i' p_j'} \sum_{i \neq j} a_{i,j} Q_i Q_j. \tag{3.15}$$

If we now choose, for $i \neq j$

$$a_{i,j} = \frac{(p_i' p_j')^{-1}}{\sum_{i \neq j} (p_i' p_j')^{-1}}, \tag{3.16}$$

which is such that $\sum_{i \neq j} a_{i,j} = 1$, we end up with

$$\mathfrak{L} = \sum_{j=1}^{n} \left(\frac{1}{p_j'} \right)^2 Q_j^2 + \sum_{i \neq j} \frac{1}{p_i' p_j'} Q_i Q_j = \left(\sum_{j=1}^{n} \frac{Q_j}{p_j'} \right)^2 \geq 0. \tag{3.17}$$

The previous argument shows that the time derivative of $w(x,t)$ can be made non-negative by suitably choosing the diffusion coefficients κ_j, $j = 1, 2, \ldots n$.

Recalling the definition of Q_j (respectively L_j), equality to zero in (3.17) holds if and only if

$$\frac{1}{p_1 p_1'}(\log u_1(x-x_1))' + \sum_{j=2}^{n-1} \frac{1}{p_j p_j'}(\log u_j(x_{j-1}-x_j))' + \frac{1}{p_n p_n'}(\log u_n(x_{n-1}))' = 0. \tag{3.18}$$

As each variable x_i appears as argument of a pair of functions only, it holds that, for every $j = 1, 2, \dots, n-1$

$$\frac{1}{p_j p_j'}\frac{\partial}{\partial x_j}(\log u_j(x_{j-1}-x_j))' + \frac{1}{p_{j+1} p_{j+1}'}\frac{\partial}{\partial x_j}(\log u_{j+1}(x_j - x_{j+1}))' = 0. \tag{3.19}$$

In (3.19) we set $x_0 = x$ and $x_n = 0$. On the other hand, since

$$(\log u_j(x_{j-1}-x_j))' = \frac{\partial}{\partial x_{j-1}} \log u_j(x_{j-1}-x_j) = -\frac{\partial}{\partial x_j} \log u_j(x_{j-1}-x_j),$$

equation (3.19) coincides with

$$\frac{1}{p_j p_j'}\frac{\partial^2}{\partial x_{j-1}^2} \log u_j(x_{j-1}-x_j) = \frac{1}{p_{j+1} p_{j+1}'}\frac{\partial^2}{\partial x_j^2} \log u_{j+1}(x_j - x_{j+1}). \tag{3.20}$$

Note that (3.20) can be verified if and only if the functions on both sides are constant. Thus, there is equality in (3.20) if and only if

$$\log u_j(x) = c\kappa_j x^2 + c_1 x + d_1, \quad \log u_{j+1}(x) = c\kappa_j x^2 + c_2 x + d_2. \tag{3.21}$$

The integrability of the function u_j then implies that the constant c has to be negative, $c = -C$, where $C > 0$. Hence, there is equality in (3.20) if and only if u_j and u_{j+1} are multiple of Gaussian densities, of variances $C(p_j p_j')^{-1}$ and $C(p_{j+1} p_{j+1}')^{-1}$, respectively, for any given positive constant C. Therefore, equality in (3.17) holds if and only if each function $u_j(x)$, $j = 1, 2, \dots, n$ is a multiple of a Gaussian density of variance $C(p_j p_j')^{-1}$.
Finally, with this choice of the diffusion coefficients, for every $x \in \mathbb{R}$ and $t_1 < t_2$,

$$u_1^{1/p_1} * u_2^{1/p_2} * \cdots * u_n^{1/p_n}(x, t_1) < u_1^{1/p_1} * u_2^{1/p_2} * \cdots * u_n^{1/p_n}(x, t_2), \tag{3.22}$$

unless all initial data are multiple of Gaussian densities with the right variances. Clearly, (3.22) is equivalent to say that the n-th convolution $w(x,t)$ is monotone increasing. An identical proof holds in higher dimension. This concludes the proof of the Lemma. □

Remark 3.2. The result of Lemma 3.1 remains true if each diffusion coefficient k_j is multiplied by a positive constant D. In this case, equality holds if the functions f_j are Gaussian functions with variances Ddk_j.

Remark 3.3. As already specified in the introduction, our quantity $w(x,t)$ is related to a particular geometric Brascamp–Lieb inequality. Results concerning more general Brascamp–Lieb inequalities that are related to Lemma 3.1 have been obtained by Bennett, Carbery, Christ and Tao in [10]. This clearly

indicates that the proof of Lemma 3.1 presented here could be extended to cover more general situations.

Lemma 3.1 has important consequences. Indeed, let us introduce the functional

$$\Psi(t) = \sup_x w(x,t) = \sup_x u_1^{1/p_1} * u_2^{1/p_2} * \cdots * u_n^{1/p_n}(x,t). \tag{3.23}$$

It is a simple exercise to verify that, in view of conditions (3.1) on the constants p_j, the functional $\Psi(t)$ is dilation invariant. In reason of this property we prove:

Theorem 3.4. *Let $\Psi(t)$ be the functional (3.23), where the functions $u_j(x,t)$, $j = 1, 2, \ldots, n$, are solutions to the heat equation corresponding to the initial values $0 \le f_j(x) \in L^1(\mathbb{R}^d)$, $d \ge 1$. Then, if for each j the exponents p_j satisfy conditions (3.1) and the diffusion coefficients are given by $\kappa_j = (p_j p_j')^{-1}$, or by a multiple of them, $\Psi(t)$ is increasing in time from*

$$\Psi(0) = \sup_x f_1^{1/p_1} * f_2^{1/p_2} * \cdots * f_n^{1/p_n}(x)$$

to the limit value

$$\lim_{t\to\infty} \Psi(t) = \prod_{j=1}^n C_{p_j}^d \left(\int_{\mathbb{R}^d} |f_j(x)|\, dx \right)^{1/p_j}. \tag{3.24}$$

The constants C_{p_j} in (3.24) are defined as in (1.4).

Moreover, $\Psi(0) = \lim_{t\to\infty} \Psi(t)$ if and only if $f_j(x)$, $j = 1, 2, \ldots, n$, is a multiple of a Gaussian density of variance $cd\kappa_j$, with $c > 0$.

Proof. Thanks to Lemma 3.1 we know that the functional $\Psi(t)$ is monotonically increasing from $\Psi(t=0)$, unless the initial densities are Gaussian functions with the right variances. To conclude the proof, it remains to show that the functional $\Psi(t)$ converges towards the limit value (3.24) as time converges to infinity. The computation of the limit value uses in a substantial way the scaling invariance of Ψ. In fact, thanks to the dilation invariance, at each time $t > 0$, the value of $\Psi(t)$ does not change if we scale each function $u_j(x)$, $j = 1, 2, \ldots, n$, according to (2.4)

$$u_j(x,t) \to U_j(x,t) = \left(\sqrt{1+2t}\right)^d u(x\sqrt{1+2t}, t).$$

Let us choose, as in Lemma 3.1 smooth initial values rapidly decreasing at infinity. Then, the *central limit property* (2.20) implies that, for $j = 1, 2, \ldots, n$

$$\lim_{t\to\infty} U_j(x,t) = M_{\kappa_j}(x) \int_{\mathbb{R}^d} f_j(x)\, dx \tag{3.25}$$

at least in $L^1(\mathbb{R}^d)$. Therefore, passing to the limit one obtains

$$\lim_{t\to\infty} \Psi(t) = \prod_{j=1}^n \left(\int_{\mathbb{R}^d} |f_j(x)|\, dx \right)^{1/p_j} \sup_x M_{\kappa_1}^{1/p_1} * M_{\kappa_2}^{1/p_2} * \cdots * M_{\kappa_n}^{1/p_n}(x). \tag{3.26}$$

Owing to the identity

$$M_{\kappa_j}^{1/p_j}(x) = C_{p_j}^d (2\pi)^{(2p_j'/d)^{-1}} M_{1/p_j'}, \tag{3.27}$$

and recalling that $\sum_{j=1}^n (p_j')^{-1} = 1$, we obtain

$$M_{\kappa_1}^{1/p_1} * M_{\kappa_2}^{1/p_2} * \cdots * M_{\kappa_n}^{1/p_n}(x) =$$

$$(2\pi)^{-d/2} \prod_{j=1}^n C_{p_j}^d M_1(x) = \prod_{j=1}^n C_{p_j}^d \exp\{-|x|^2/2\}.$$

This implies (3.24), and concludes the proof of the inequality for well chosen initial data. Then, the general case of the theorem follows by standard density arguments. □

Remark 3.5. Theorem 3.4 is related to the monotonicity in time of a dilation invariant functional whose components are solutions to the heat equation. Therefore, the main importance of the theorem is to highlight the existence of a new Lyapunov functional related to the heat equation. This result, however, can be rephrased to give a new proof of known inequalities in sharp form. Let us set, in Theorem 3.4

$$g_j(x) = f_j(x)^{1/p_j},$$

for $j = 1, 2, \dots, n$. Then, it holds

$$\sup_x g_1 * g_2 * \cdots * g_n(x) \le \prod_{j=1}^n C_{p_j}^d \prod_{j=1}^n \|g_j\|_{p_j}. \tag{3.28}$$

Moreover, since

$$\sup_x g_1 * g_2 * \cdots * g_n(x) \ge \int g_1(-x_1) g_2(x_1 - x_2) \dots g_n(x_{n-1})\, dx_1 \dots dx_{n-1},$$

inequality (3.28) implies, under the same conditions on the constants p_j,

$$\int g_1(x_1) g_2(x_1 - x_2) \dots g_n(x_{n-1})\, dx_1 \dots dx_{n-1} \le \prod_{j=1}^n C_{p_j}^d \prod_{j=1}^n \|g_j\|_{p_j}. \tag{3.29}$$

Inequality (3.29) is a particular case of the inequalities obtained by Brascamp and Lieb [13] by a different method.

Remark 3.6. Clearly, the proof of Theorem 3.4 still holds when $n = 2$. In this case, however, the diffusion coefficients κ_j, $j = 1, 2$ coincide. In fact, when $n = 2$, the condition (3.1) reduces to

$$1 \le p_j \le +\infty; \quad \frac{1}{p_1} + \frac{1}{p_2} = 1,$$

so that p_1 and p_2 are dual exponents. Consequently $p_1' = p_2$ and $p_2' = p_1$, which imply $\kappa_1 = \kappa_2 = \kappa = (p_1 p_2)^{-1}$. But in this case the definition (1.4) of the constant C_p implies $C_{p_1} = 1/C_{p_2}$, and the limit (3.24) takes the value

$$\lim_{t\to\infty} \Psi(t) = \left(\int_{\mathbb{R}^d} |f_1(x)|\, dx \right)^{1/p_1} \left(\int_{\mathbb{R}^d} |f_2(x)|\, dx \right)^{1/p_2}. \tag{3.30}$$

Note that in this case inequality (3.29) reduces simply to the classical Hölder inequality.

Remark 3.7. As noticed by Brascamp and Lieb [13], Theorem 3.4 contains as special case the best possible improvement to Young's inequality. If $n = 3$ (3.29) reads

$$\int_{\mathbb{R}^{2d}} f(x)g(x-y)h(y)dx\,dy \le (C_pC_qC_s)^d \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^s}, \tag{3.31}$$

where $1 \le p, q, s \le \infty$, $1/p + 1/q + 1/s = 2$, and equality holds when f, g, h are suitable Gaussian functions. Choosing

$$h(y) = (f * g(y))^{r-1}$$

leads to an equivalent form of (3.31)

$$\|f * g\|_{L^r} \le (C_pC_qC_{r'})^d \|f\|_{L^p}\|g\|_{L^q}, \tag{3.32}$$

namely the standard form of Young's inequality [6, 13].

Also, repeated applications of (3.32) give

$$\|g_1 * g_2 * \cdots * g_n\|_r \le C_{r'}^d \prod_{j=1}^n C_{p_j}^d \|g_j\|_{p_j}, \tag{3.33}$$

where $1 \le p_j \le \infty$ and $\sum_{j=1}^n 1/p_j = n - 1 + 1/r$.

4. Further Lyapunov functionals

Theorem 3.4 shows the monotonicity in time of the L^∞-norm of the n-th convolution of type (3.23), as well as its convergence towards an explicitly computable limit value (in terms of the initial data). The key point in getting this result was the dilation property of the functional $\Psi(t)$.

To get a similar result for the L^r-norm of the n-th convolution $w(x,t)$, $r > 0$, and to obtain the (eventual) limit value, we need that the dilation property still holds for $\|w(t)\|_r$. By applying the same scaling $u_j(x) \to V_j(x) = a^d V(ax)$ to each function $u_j(x)$ in (1.2) we get

$$V_1^{\alpha_1} * V_2^{\alpha_2} * \cdots * V_n^{\alpha_n}(x) = a^{d\gamma} u_1^{\alpha_1} * u_2^{\alpha_2} * \cdots * u_n^{\alpha_n}(ax) = a^{d\gamma} w(ax),$$

where

$$\gamma = \sum_{j=1}^n \alpha_j - n + 1$$

Hence

$$\int_{\mathbb{R}^d} (V_1^{\alpha_1} * V_2^{\alpha_2} * \cdots * V_n^{\alpha_n}(x))^r \, dx = \int_{\mathbb{R}^d} a^{dr\gamma} w^r(ax)\,dx,$$

and dilation invariance occurs if and only if $r\gamma = 1$, that is

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = n - 1 + \frac{1}{r}. \tag{4.1}$$

By analogy with condition (3.1), we will satisfy condition (4.1) in two separate cases. The first refers to fix, for $j = 1, \dots, n$, positive real numbers p_j and r such that

$$p_j < 1, r < 1; \quad \sum_{j=1}^{n} \frac{1}{p_j} = n - 1 + \frac{1}{r}. \tag{4.2}$$

The second refers to fix, for $j = 1, \dots, n$, positive real numbers p_j and r such that

$$1 < p_j \leq \infty, 1 < r \leq \infty; \quad \sum_{j=1}^{n} \frac{1}{p_j} = n - 1 + \frac{1}{r}. \tag{4.3}$$

In the following, we will analyze the time behaviour of $\|w(t)\|_r$ in the case (4.2). Then, the result for the case (4.3) will follow by the same line of proof.

Condition (4.2) implies that $p_j' < 0$ for all $j = 1, 2, \dots, n$, and

$$\frac{1}{p_j} = \sum_{i \neq j} \frac{1}{p_i'} + \frac{1}{r}.$$

Making use of the proof of Lemma 3.1, let us set, for $j = 1, 2, \dots, n$, the (positive) coefficients of diffusion

$$\kappa_j = \frac{1}{p_j |p_j'|}. \tag{4.4}$$

Then, by means of elementary computations we obtain

$$\sum_{l=1}^{n} \kappa_l = \sum_{l=1}^{n} \frac{1}{p_l |p_l'|} = \sum_{i \neq j} \frac{1}{p_i' |p_j'|} + \frac{1}{r|r'|}. \tag{4.5}$$

Since the real numbers p_j now satisfy condition (4.2), the quantity (3.13) considered in Lemma 3.1, with the same choice (3.16) of the coefficients $a_{i,j}$ takes the form

$$\mathfrak{L} = -\left(\sum_{j=1}^{n} \frac{Q_j}{p_j'} \right)^2 + \frac{1}{r|r'|} \sum_{i \neq j} a_{i,j} Q_i Q_j. \tag{4.6}$$

It is evident that in this case we cannot expect that $\mathfrak{L}$ has a definite sign. However, using expression (4.6) into (3.12) we obtain

$$\frac{\partial w(x,t)}{\partial t} = \int u_1^{1/p_1}(x - x_1) \dots u_n^{1/p_n}(x_{n-1}) \mathfrak{L}(u_1, \cdots u_n)\, dx_1 \cdots dx_{n-1} =$$

$$- \int u_1^{1/p_1}(x - x_1) \dots u_n^{1/p_n}(x_{n-1}) \left(\sum_{j=1}^{n} \frac{Q_j}{p_j'} \right)^2 dx_1 \cdots dx_{n-1} + \frac{1}{r|r'|} \frac{\partial^2 w}{\partial x^2}. \tag{4.7}$$

In fact, by formula (3.10)

$$\int u_1^{1/p_1}(x - x_1) \dots u_n^{1/p_n}(x_{n-1}) \sum_{i \neq j} a_{i,j} Q_i Q_j \, dx_1 \cdots dx_{n-1} =$$

$$\int u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})\sum_{i\neq j}\frac{a_{i,j}}{p_ip_j}L_iL_j\,dx_1\cdots dx_{n-1}=\frac{\partial^2 w}{\partial x^2}.$$

Consequently, thanks to (4.7)

$$\frac{d}{dt}\int w^r(x,t)\,dx = r\int w^{r-1}(x,t)\frac{\partial w(x,t)}{\partial t}\,dx = \frac{1}{|r'|}\int w^{r-1}\frac{\partial^2 w}{\partial x^2}\,dx+$$

$$-r\int w^{r-1}\int u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})\left(\sum_{j=1}^n\frac{Q_j}{p_j'}\right)^2 dx_1\cdots dx_{n-1}\,dx =$$

$$-r\int w^{r-1}\,dx\int u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})\left(\sum_{j=1}^n\frac{Q_j}{p_j'}\right)^2 dx_1\cdots dx_{n-1}+$$

$$\frac{(1-r)^2}{r}\int w^{r-2}\left(\frac{\partial w}{\partial x}\right)^2 dx \tag{4.8}$$

Surprisingly, the expression on (4.8) has a sign. This is consequence of the following Lemma, which generalizes a similar result that dates back to Blachman [11]. In case of convolution of two functions, analogous result has been obtained recently in [36]. For the sake of simplicity, we will present the proof in dimension $d=1$.

Lemma 4.1. *Let $w(x)$ be the (smooth) n-th convolution defined by (3.23). Then, for any set of positive constants p_j and r, and positive constants λ_j, $j=1,2,\dots,n$ such that $\sum_{j=1}^n\lambda_j=1$ it holds*

$$\int w^{r-2}\left(\frac{\partial w}{\partial x}\right)^2 \le \int w^{r-1}(x)\int u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})\left(\sum_{j=1}^n\frac{\lambda_j}{p_j}L_j\right)^2. \tag{4.9}$$

Moreover, equality in (4.9) holds if and only if any function u_j, $j=1,2,\dots,n$ is multiple of a Gaussian function of variance λ_j/p_j.

Proof. By property (3.9), if we take a set of positive constants λ_j, $j=1,2,\dots,n$ such that $\sum_{j=1}^n\lambda_j=1$ we can express the first derivative of $w(x)$ as

$$w'(x)=\int u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})\sum_{j=1}^n\frac{\lambda_j}{p_j}L_j dx_1\dots dx_{n-1},$$

where L_j is defined as in (3.11). Therefore, by denoting

$$d\mu_x(x_1,\dots,x_{n-1})=\frac{u_1^{1/p_1}(x-x_1)\dots u_n^{1/p_n}(x_{n-1})}{w(x)}\,dx_1\cdots dx_n, \tag{4.10}$$

we obtain

$$\frac{w'(x)}{w(x)}=\int\sum_{j=1}^n\frac{\lambda_j}{p_j}L_j d\mu_x(x_1,\dots,x_{n-1}).$$

Note that, for any $x \in \mathbb{R}$ the measure $d\mu$ defined in (4.10) is a unit measure on $\mathbb{R}^{n-1}$,

$$\int_{\mathbb{R}^{n-1}} d\mu_x(x_1, x_2, \cdots x_{n-1}) = 1.$$

Jensen's inequality then gives

$$\left(\frac{w'(x)}{w(x)}\right)^2 \leq \int \left(\sum_{j=1}^{n} \frac{\lambda_j}{p_j} L_j\right)^2 d\mu_x(x_1, \dots, x_{n-1}). \tag{4.11}$$

Multiplying both sides of (4.11) by $w^r(x)$, and integrating over x proves the Lemma.

Note that, since equality in Jensen's inequality holds if and only if the argument is constant, equality in (4.11) holds if and only if

$$\sum_{j=1}^{n} \frac{\lambda_j}{p_j} L_j = const.$$

Hence, the reasoning of the last part of Lemma 3.1 can be repeated to show that there is equality in (4.9) if and only if any function u_j, $j = 1, 2, \dots, n$ is multiple of a Gaussian function of variance λ_j/p_j. □

Let us return to formula (4.8). Conditions (4.2) imply that

$$\sum_{j=1}^{n} \frac{1}{|p_j'|} = \frac{1}{|r'|}.$$

Hence

$$\frac{r}{1-r} \sum_{j=1}^{n} \frac{1}{|p_j'|} = 1.$$

Choosing then

$$\lambda_j = \frac{r}{1-r} \frac{1}{|p_j'|},$$

we obtain that (4.9) reads

$$\int w^{r-2} \left(\frac{\partial w}{\partial x}\right)^2 dx \leq \frac{r^2}{(1-r)^2} \int w^{r-1}(x)\, dx \cdot$$

$$\cdot \int u_1^{1/p_1}(x - x_1) \dots u_n^{1/p_n}(x_{n-1}) \left(\sum_{j=1}^{n} \frac{1}{p_j'} Q_j\right)^2 dx_1 \cdots dx_n. \tag{4.12}$$

This shows that the quantity in (4.8) is negative. Hence, we proved that, if the positive constants p_j and r satisfy conditions (4.2), the functional

$$\Lambda(t) = \|w(t)\|_r = \left(\int (u_1^{1/p_1} * u_2^{1/p_2} * \cdots * u_n^{1/p_n})^r(x, t)\, dx\right)^{1/r} \tag{4.13}$$

is monotone decreasing. Since we know that, in view of conditions (4.2) on the constants p_j, the functional $\Lambda(t)$ is dilation invariant, we proved:

Theorem 4.2. *Let $\Lambda(t)$ be the functional (4.13), where the functions $u_j(x,t)$, $j=1,2,\dots,n$, are solutions to the heat equation corresponding to the initial values $0 \le f_j(x) \in L^1(\mathbb{R}^d)$, $d \ge 1$. Then, if for each j the exponents p_j satisfy conditions (4.2) and the diffusion coefficients are given by $\kappa_j = (p_j|p_j'|)^{-1}$, $\Lambda(t)$ is decreasing in time from*

$$\Lambda(0) = \left(\int \left(f_1^{1/p_1} * f_2^{1/p_2} * \cdots * f_n^{1/p_n}(x) \right)^r \, dx \right)^{1/r}$$

to the limit value

$$\lim_{t\to\infty} \Lambda(t) = C_{r'}^d \prod_{j=1}^n C_{p_j}^d \left(\int_{\mathbb{R}^d} |f_j(x)| \, dx \right)^{1/p_j}. \tag{4.14}$$

The constants C_{p_j} in (3.24) are defined by

$$C_p^2 = \frac{p^{1/p}}{|p'|^{1/p'}}, \tag{4.15}$$

Moreover, $\Lambda(0) = \lim_{t\to\infty} \Lambda(t)$ if and only if $f_j(x)$, $j=1,2,\dots,n$, is a multiple of a Gaussian density of variance $d\kappa_j$.

Proof. We know that the functional $\Lambda(t)$ is monotonically decreasing from $\Lambda(t=0)$, unless the initial densities are Gaussian functions with the right variances. In addition, $\Lambda(t)$ is dilation invariant. As in Theorem 3.4, let us scale each function $u_j(x)$, $j=1,2,\dots,n$, according to (2.4). Therefore, by the central limit property, passing to the limit one obtains

$$\lim_{t\to\infty} \Lambda(t) = \prod_{j=1}^n \left(\int_{\mathbb{R}^d} |f_j(x)| \, dx \right)^{1/p_j} \left\| M_{\kappa_1}^{1/p_1} * M_{\kappa_2}^{1/p_2} * \cdots * M_{\kappa_n}^{1/p_n} \right\|_r. \tag{4.16}$$

The value of the integral can be evaluated by using formula (3.27) of Theorem 3.4, with the additional difficulty to evaluate the norm of a Gaussian in L^r. Thanks to condition (4.2) we obtain

$$\left\| M_{\kappa_1}^{1/p_1} * M_{\kappa_2}^{1/p_2} * \cdots * M_{\kappa_n}^{1/p_n} \right\|_r = C_{r'}^d \prod_{j=1}^n C_{p_j}^d.$$

This concludes the proof of the theorem. □

The computations leading to Theorem 4.2 can be repeated step-by-step in the case in which the p_j's and r satisfy condition (4.3). In this case, however, the sign of $\mathfrak{L}$ changes, and we obtain

Theorem 4.3. *Let $\Lambda(t)$ be the functional (4.13), where the functions $u_j(x,t)$, $j=1,2,\dots,n$, are solutions to the heat equation corresponding to the initial values $0 \le f_j(x) \in L^1(\mathbb{R}^d)$, $d \ge 1$. Then, if for each j the exponents p_j satisfy conditions (4.3) and the diffusion coefficients are given by $\kappa_j = (p_j p_j')^{-1}$, $\Lambda(t)$ is increasing in time from*

$$\Lambda(0) = \left(\int \left(f_1^{1/p_1} * f_2^{1/p_2} * \cdots * f_n^{1/p_n}(x) \right)^r \, dx \right)^{1/r}$$

to the limit value

$$\lim_{t\to\infty}\Lambda(t) = C_{r'}^d \prod_{j=1}^n C_{p_j}^d \left(\int_{\mathbb{R}^d} |f_j(x)|\,dx\right)^{1/p_j}. \tag{4.17}$$

The constants C_{p_j} in (3.24) are defined by (1.4). Moreover, $\Lambda(0) = \lim_{t\to\infty}\Lambda(t)$ if and only if $f_j(x)$, $j = 1, 2, \dots, n$, is a multiple of a Gaussian density of variance $d\kappa_j$.

Remark 4.4. The monotonicity property of the functional $\Lambda(t)$ defined by (4.13) have been noticed first by Bennett and Bez [8] by means of a different approach. Consequently, the results of both Theorems 4.2 and 4.3 also follow from their arguments. We note, however, that the dilation invariance property of $\Lambda(t)$, which is at the basis of the direct proof of the Theorems, has not explicitly taken into account before.

Remark 4.5. Theorems 4.2 and 4.3 show the monotonicity properties of the L^r-norm of the n-th convolution of powers of solutions to the heat equation. As discussed at the end of Theorem 3.4, apart from its intrinsic physical interest, this monotonicity can be rephrased in the form of inequalities for convolutions in sharp form. In particular, when $n = 2$, Theorem 4.2 contains the sharp form of Young inequality in the so-called reverse case

$$\|f * g\|_r \geq (C_p C_q C_{r'})^d \|f\|_p \|g\|_q, \tag{4.18}$$

where $0 < p, q, r < 1$ while $1/p + 1/q = 1 + 1/r$, and C_p is defined by (4.15).

Remark 4.6. A particular case of Theorem 4.3 implies Babenko's inequality [1] (cf. also Beckner [6]):

$$\|\mathfrak{F}f\|_{p'} \leq C_p^d \|f\|_p, \tag{4.19}$$

where C_p is defined as in (1.4), p' is an even integer $p' = 2, 4, 6, \dots$, and $\mathfrak{F}f$ denotes the Fourier transform of f. Here the Fourier transform is defined for integrable functions by

$$\mathfrak{F}f(\xi) = \int_{\mathbb{R}^d} \exp\{-2\pi i x\cdot\xi\}\, f(x)\,dx$$

Inequality (4.19) follows by choosing in Theorem 4.3 $r = 2$ and $1/p_j = 1/p = (2n-1)/2n$, $j = 1, 2, \dots, n$, which are such that condition (4.3) is satisfied. In this case, in fact, by setting $f_j = f^{2n/(2n-1)}$, for $j = 1, 2, \dots, n$, we obtain that f satisfies the inequality

$$\left(\int (\underbrace{f * f * \cdots * f}_{n})^2\,dx\right)^{1/2} \leq C_p^{dn} \|f\|_p^n.$$

Since

$$\mathfrak{F}\left(\underbrace{f * f * \cdots * f}_{n}\right) = (\mathfrak{F}f)^n,$$

by Plancherel's identity we conclude that

$$\left(\int (\mathfrak{F}f)^{2n}\, d\xi\right)^{1/(2n)} \leq C_p^d \|f\|_p. \tag{4.20}$$

We remark that, as explicitly mentioned in [8], the monotonicity of the quantity in (4.20) also follows from the results in [10] (cf. also [8]). A further inside into Haussdorff–Young inequality, with an interesting discussion about the behavior in time of $\|\mathfrak{F}u^{1/p}(t)\|_{p'}$ can be found in [9].

5. Monotonicity and Prékopa–Leindler inequality

The analysis of the preceding section shows the monotonicity properties of the L^r norm of the n-th convolution of powers of the solutions to the heat equation. In particular Theorem 3.4 covers the L^∞ case, while Theorem 4.2 (respectively Theorem 4.3) cover the case $r < 1$ (respectively $r > 1$). Two limit cases remain to be examined, namely the cases $r \to 0$ and $r \to 1$. Here we will briefly discuss the first case, leaving the second to the next section.

Given a set of positive constants q_j, $j = 1, 2, \ldots, n$, such that $\sum_{j=1}^n 1/q_j = 1$, and a constant $N \gg 1$, we choose in Theorem 4.2

$$p_j = \frac{q_j}{N}, \quad r = \frac{1}{N-(n-1)}. \tag{5.1}$$

Then, if $N \geq \max_j q_j + n$, conditions (4.2) are satisfied and the monotonicity of $\Lambda(t)^r$ is guaranteed. On the other hand, if (5.1) holds, one can write

$$w(x)^r = \left(u_1^{N/q_1} * u_2^{N/q_2} * \cdots * u_n^{N/q_n}(x,t)\right)^{1/(N-n+1)} =$$

$$\left(\int \left(u_1(x-x_1)^{1/q_1} \cdots u_n(x_{n-1})^{1/q_n}\right)^N dx_1 \cdots dx_{n-1}\right)^{1/(N-n+1)}. \tag{5.2}$$

We can render the dependence of $\Lambda(t)^r$ on N explicit by setting

$$\int w(x)^r\, dx = \Upsilon_N(t),$$

where

$$\Upsilon_N(t) = \left(\int \left(u_1(x-x_1)^{1/q_1} \cdots u_n(x_{n-1})^{1/q_n}\right)^N dx_1 \cdots dx_{n-1}\right)^{1/(N-n+1)}. \tag{5.3}$$

By Theorem 4.2, $\Upsilon_N(t)$ is monotonically decreasing in time for all values of $N \geq \max_j q_j + n$, provided the coefficients of diffusion are the correct ones. Moreover, note that, for any given N, the coefficients of diffusions κ_j depend on it, and

$$\kappa_j^N = \frac{N(N-q_j)}{q_j^2}.$$

On the other hand, Theorem 4.2 remains true if we multiply all coefficients of diffusion by the same constant. Therefore, without affecting the monotonicity of $\Upsilon_N(t)$ we can fix the coefficients of diffusion equal to

$$\tilde{\kappa}_j^N = \frac{N(N-q_j)}{N^2}\frac{1}{q_j^2} = \frac{1}{q_j^2} - \frac{1}{Nq_j}. \tag{5.4}$$

By letting $N \to \infty$, the coefficients of diffusion $\tilde{\kappa}_j^N$ tend to the values $\tilde{\kappa}_j^\infty = 1/q_j^2$. Consequently, recalling expression (5.2), for each $t > 0$, as $N \to \infty$ we have pointwise convergence of $w(x,t)^{r(N)}$, and

$$\lim_{N\to\infty} w(x,t)^{r(N)} = ess\ sup_{x_1,\dots,x_{n-1}} u_1(x-x_1,t)^{1/q_1}\cdots u_n(x_{n-1},t)^{1/q_n}\,dx, \tag{5.5}$$

where now $u_j(x,t)$, for $j = 1, 2 \dots, n$ is solution to the heat equation (2.1) with diffusion coefficient $1/q_j^2$.

Suppose now that the initial data $f_j(x)$ in the heat equation are bounded and have bounded support. Then, they are bounded from above by multiples of some Gaussian functions, that is, for $j = 1, 2, \dots, n$

$$f_j(x) \le A_j M_{2\sigma_j}(x),$$

where A_j and σ_j are suitable positive constants. In this case, the solution to the heat equation with diffusion coefficient $\tilde{\kappa}_j^N$ satisfies the bound

$$u_j(x,t) = f_j * M_{2\tilde{\kappa}_j^N t}(x) \le A_j M_{2(\sigma_j+\tilde{\kappa}_j^N t)}. \tag{5.6}$$

Consequently,

$$w(x,t)^r \le \left(\prod_{j=1}^n A_j^{p_j} M_{2\kappa_1 t}^{p_1} * M_{2\kappa_2 t}^{p_2} * \cdots * M_{2\kappa_n t}^{p_n}\right)^r,$$

where the constants p_j and r are given as in (5.1), and

$$\kappa_j t = \sigma_j + \left(\frac{1}{q_j^2} - \frac{1}{Nq_j}\right) t.$$

By applying formula (2.7) we easily obtain that, for sufficiently large N,

$$w^r(x,t) \le C_1(t) M_{2C_2(t)}(x),$$

where the positive functions $C_1(t)$ and $C_2(t)$ do not depend on N. By the dominated convergence theorem it follows that

$$\lim_{N\to\infty} \Upsilon_N(t) = \Upsilon(t),$$

where

$$\Upsilon(t) = \int ess\ sup_{x_1,\dots,x_{n-1}} u_1(x-x_1)^{1/q_1}\cdots u_n(x_{n-1})^{1/q_n}\,dx. \tag{5.7}$$

Moreover, $\Upsilon(t)$ is monotonically decreasing in time if the coefficients of diffusion in the heat equations are given by $\kappa_j = 1/q_j^2$. Since the functional $\Upsilon(t)$ is invariant under dilation, we can pass to the limit to find the lower bound.

By the same argument of the proof of Theorem 3.4, we conclude that the limit value is obtained by setting

$$u_j(x) = \int f_j(x)\,dx M_{1/q_j^2}.$$

Explicit computations then show that

$$\lim_{t\to\infty} \Upsilon(t) = \prod_{j=1}^{n} q_j^{d/q_j} \left(\int f_j(x)\,dx. \right)^{1/q_j} \tag{5.8}$$

The choice $f_j(x) = g_j(q_j x)$ then implies

$$\int_{\mathbb{R}^d} f_j(x)^{1/q_j}\,dx = q_j^{-d/q_j} \left(\int_{\mathbb{R}^d} g_j(x)^{1/q_j}\,dx \right)^{1/q_j}.$$

The case of general data follows by a density argument. We conclude with the following

Theorem 5.1. *Let $\Upsilon(t)$ denote the functional (5.7) where the functions $u_j(x,t)$, $j = 1, 2, \ldots, n$, are solutions to the heat equation corresponding to the initial values $0 \le f_j(x) \in L^1(\mathbb{R}^d)$, $d \ge 1$. Then, if exponents q_j satisfy $\sum_{j=1}^n q_j^{-1} = 1$, and the diffusion coefficients are given by $\kappa_j = q_j^{-2}$, $\Upsilon(t)$ is decreasing in time from*

$$\Upsilon(0) = \int \text{ess sup}_{x_1,\ldots,x_{n-1}} f_1(q_1(x-x_1))^{1/q_1} \cdots f_n(q_n(x_{n-1}))^{1/q_n}\,dx$$

to the limit value

$$\lim_{t\to\infty} \Upsilon(t) = \prod_{j=1}^{n} \left(\int_{\mathbb{R}^d} |f_j(x)|\,dx \right)^{1/q_j}. \tag{5.9}$$

Moreover, $\Upsilon(0) = \lim_{t\to\infty} \Upsilon(t)$ if and only if $f_j(x)$, $j = 1, 2, \ldots, n$, is a multiple of a Gaussian density of variance $d\kappa_j$.

Remark 5.2. If $n = 2$ the monotonicity of the functional Υ proven in Theorem 5.1 implies the classical Prékopa–Leindler inequality. In this case, in fact one obtains the Prékopa–Leindler theorem [25, 31, 32] that reads

$$\|h\|_1 \ge \|f\|_1^{\lambda} \|g\|_1^{1-\lambda},$$

where

$$h(x|f,g) = ess\ sup_y f\left(\frac{x-y}{\lambda}\right)^{\lambda} g\left(\frac{y}{1-\lambda}\right)^{1-\lambda}.$$

The derivation of Prékopa–Leindler inequality from the Young's inequality has been obtained by Brascamp and Lieb [13]. Our result, however, enlightens a new meaning of this inequality, that is viewed as a consequence of the monotonicity of a Lyapunov functional of the convolution of two powers of the solution to the heat equation.

Remark 5.3. Theorem 5.1 is a corollary of the general result of Theorem 4.2. However, a direct proof of monotonicity could be possible by looking at the functional (5.7) directly.

6. A short proof of entropy power inequality

In its original version, Shannon's entropy power inequality (*EPI*) [34] gives a lower bound on Shannon's entropy functional of the sum of independent random variables X, Y with densities

$$\exp\left(\frac{2}{d}H(X+Y)\right) \geq \exp\left(\frac{2}{d}H(X)\right) + \exp\left(\frac{2}{d}H(Y)\right), \tag{6.1}$$

with equality if X and Y are Gaussian random variables. Shannon's entropy functional of the probability density function $f(x)$ of X is

$$H(X) = H(f) = -\int_{\mathbb{R}^d} f(v)\log f(v)\,dv. \tag{6.2}$$

Note that Shannon's entropy functional coincides to Boltzmann's entropy up to a change of sign. The entropy-power

$$N(X) = N(f) = \exp\left(\frac{2}{d}H(X)\right)$$

(variance of a Gaussian random variable with the same Shannon's entropy functional) is maximum and equal to the variance when the random variable is Gaussian, and thus, the essence of (6.1) is that the sum of independent random variables tends to be *more Gaussian* than one or both of the individual components.

The first rigorous proof of inequality (6.1) was given by Stam [35] for the case $d = 1$ (see also Blachman [11] for the generalization of *EPI* to d-dimensional random vectors), and was based on an identity which couples Fisher's information with Shannon's entropy functional [16].

Making use of the relationship between mutual information and minimum mean-square error for additive Gaussian channels [22], a different and simpler proof of *EPI* based on an elementary estimation–theoretic reasoning which sidesteps invoking Fisher's information, and makes use of a result of Lieb [26], was recently given in [23] (see also Rioul [33] for a unified view of proofs of *EPI* via Fisher's information and minimum mean-square errors).

Other variations of the entropy–power inequality are present in the literature. Costa's strengthened entropy–power inequality [17], in which one of the variables is Gaussian, and a generalized inequality for linear transforms of a random vector due to Zamir and Feder [40].

Also, other properties of Shannon's entropy-power $N(f)$ have been investigated so far. In particular, the *concavity of entropy power* theorem, which asserts that

$$\frac{d^2}{dt^2}\left(N(u(t))\right) \leq 0 \tag{6.3}$$

provided that $u(t)$ is the solution to the heat equation (2.1). Inequality (6.3) is due to Costa [17]. Later, the proof has been simplified in [18, 19], by an argument based on the Blachman-Stam inequality [11]. More recently, a short

and simple proof has been obtained by Villani [39], using an old idea by McKean [30]. Various consequences of inequality (6.3), including the logarithmic Sobolev inequality and Nash's inequality have been recently discussed in [37].

As noticed by Lieb [26], the *EPI* can also be proven as a limit case of the Young inequality in the sharp form (3.32), by letting the parameters p, q and r tend to one. This result can be obtained as follows. Let $0 < a < 1$ denote a fixed constant. For a given (small) positive χ, let us consider Young's inequality (3.32) with

$$r = 1 + \chi, \qquad p = \frac{1+\chi}{1+a\chi}, \qquad q = \frac{1+\chi}{1+(1-a)\chi}, \tag{6.4}$$

which are such that

$$\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}.$$

Note that, as $\chi \to 0$, $p, q, r \to 1$. Let f, g, h be smooth probability densities, and let us define $z(\chi) = \|h\|_{1+\chi}$. Then $z(0) = 1$, and thanks to the identity

$$z(\chi) = \exp\left\{ \frac{1}{1+\chi} \log \int_{\mathbb{R}^d} h^{1+\chi}\, dv \right\},$$

one evaluates straightforwardly

$$z'(\chi) = z(\chi) \left[-\frac{1}{(1+\chi)^2} \log \int h^{1+\chi} + \frac{1}{1+\chi} \frac{\int h^{1+\chi} \log h}{\int h^{1+\chi}} \right]. \tag{6.5}$$

Hence,

$$z'(0) = \int_{\mathbb{R}^d} h \log h\, dv = -H(h).$$

Owing to the smoothness of $f * g$, we can expand $\|f * g\|_{1+\chi}$ in Taylor's series of χ up to order one, to obtain

$$\|f * g\|_{1+\chi} = 1 - H(f * g)\chi + o_1(\chi), \tag{6.6}$$

where $o_1(\chi)$ is such that $o_1(\chi)/\chi \to 0$ as $\chi \to 0$. Analogous computations for the function

$$\omega(\chi) = \exp\left\{ d \log(C_p C_q C_{r'}) + \frac{1}{p} \log \int_{\mathbb{R}^d} f^p\, dv + \frac{1}{q} \log \int_{\mathbb{R}^d} g^q\, dv \right\}$$

where p and q are defined in (6.4), allow to conclude that

$$\omega'(0) = \frac{d}{2} \left(a \log a + (1-a) \log(1-a) \right) - (1-a) H(f) - a H(g). \tag{6.7}$$

Therefore, expanding again in Taylor's series of χ, we obtain

$$\omega(\chi) = 1 + \left(\frac{d}{2} \left(a \log a + (1-a) \log(1-a) \right) - (1-a) H(f) - a H(g) \right) \chi + o_2(\chi), \tag{6.8}$$

where again $o_2(\chi)/\chi \to 0$ as $\chi \to 0$. It is interesting to remark that the sharp constant $(C_p C_q C_{r'})^d$ furnishes an important contribution to formula (6.6). This contribution can be derived straightforwardly using the identity

$$\frac{d}{d\chi} \left(\frac{1}{p} \right) = -\frac{d}{d\chi} \left(\frac{1}{p'} \right).$$

This gives

$$\frac{d}{d\chi}\log C_p^2 = \frac{d}{d\chi}\left(\frac{1}{p}\log p - \frac{1}{p'}\log p'\right) = \left(-2+\log\frac{p}{p'}\right)\frac{d}{d\chi}\left(\frac{1}{p}\right) =$$

$$(-2+\log(p-1))\,\frac{d}{d\chi}\left(\frac{1}{p}\right) = \frac{1-a}{(1+\chi)^2}\left(2-\log\frac{(1-a)\chi}{1+a\chi}\right),$$

and

$$\frac{d}{d\chi}\log(C_pC_qC_{r'})^2 = \frac{1}{(1+\chi)^2}\left((1-a)\log\frac{1-a}{1+a\chi} + a\log\frac{a}{1+(1-a)\chi}\right).$$

In conclusion we have the following [26]:

Lemma 6.1. *Let the probability densities $f(x)$ and $g(x)$ $x \in \mathbb{R}^d$ possess bounded Shannon's entropy functional. Then, for any positive constant $0 < a < 1$ the following inequality holds*

$$H(f*g) \geq (1-a)H(f) + aH(g) - \frac{d}{2}\left(a\log a + (1-a)\log(1-a)\right). \quad (6.9)$$

Proof. The proof is a direct consequence of the sharp Young inequality (3.32). With our notations, Young inequality can be rephrased as $z(\chi) - \omega(\chi) \leq 0$. Using expansions (6.6) and (6.8), and letting $\chi \to 0$, inequality (6.9) follows for smooth densities. A standard density argument then concludes the proof. □

Shannon's entropy power inequality then follows by maximizing the right-hand side of inequality (6.9). A simple computation shows that the right-hand side, say $A(a, H(f), H(g))$ attains the maximum when

$$a = \bar{a} = \frac{\exp\left\{2\left(H(g) - H(f)\right)/d\right\}}{1+\exp\left\{2\left(H(g) - H(f)\right)/d\right\}}, \quad (6.10)$$

and, for $a = \bar{a}$

$$A(\bar{a}, H(f), H(g)) = \frac{d}{2}\log\left\{\exp\left(2H(f)/d\right) + \exp\left(2H(g)/d\right)\right\}. \quad (6.11)$$

With analogous computations, Shannon's entropy-power inequality can be easily extended to a convolution of n probability densities by means of Theorem 4.3.

While the result of Lieb [26] outlines an interesting connection between Young's inequality and the entropy power inequality, the proof of *EPI* via Young's inequality does not contain any connection with our idea about monotonicity properties of Lyapunov functionals for the solution to the heat equation. Indeed, a much simpler direct proof is available by making use of this idea. For the moment, let us fix the dimension equal to 1.

Let as usual $w(x,t)$ denote the n-th convolution

$$w(x,t) = u_1 * u_2 * \cdots * u_n(x,t), \quad (6.12)$$

where the functions $u_j(x,t)$, $j = 1, 2, \ldots, n$, are solutions to the heat equations, with coefficients of diffusion κ_j, corresponding to the initial probability densities $0 \leq f_j(x)$ with bounded Shannon's entropy. It is important to note

that, in view of the closure property of the Gaussian density (2.3) with respect to convolutions, $w(x,t)$ itself satisfies the heat equation (2.1) with coefficient of diffusion $\kappa = \sum_{j=1}^{n} \kappa_j$. For any set of positive values γ_j, $j = 1, 2, \dots, n$, such that $\sum_{j=1}^{n} \gamma_j = 1$, we introduce the functional

$$\Phi(t) = H(w(t)) - \sum_{j=1}^{n} \gamma_j H(u_j(t)). \tag{6.13}$$

Let f_α be the scaled function defined as in (2.11). Since, for $\alpha > 0$

$$H(f_\alpha) = H(f) - \log \alpha,$$

the functional $\Phi(t)$ is dilation invariant. Given $t > 0$, let us evaluate the time derivative of $\Phi(t)$. We obtain

$$\frac{d}{dt} H(w(t)) = \kappa I(w(t)) - \sum_{j=1}^{n} \gamma_j \kappa_j I(u_j(t)), \tag{6.14}$$

where we defined by $I(f)$ the Fisher information of the density f, given in any dimension $d \geq 1$ by

$$I(f) = \int_{\mathbb{R}^d} \frac{|\nabla f(x)|^2}{f(x)}\, dx. \tag{6.15}$$

By setting in Lemma 4.1 $r = 1$ and $p_j = 1$, $j = 1, 2, \dots, n$, which satisfy conditions (4.3), inequality (4.9) assumes the form

$$I(w) \leq \int dx \int u_1(x - x_1) \dots u_n(x_{n-1}) \left(\sum_{j=1}^{n} \lambda_j L_j \right)^2 = \sum_{j=1}^{n} \lambda_j^2 I(u_j). \tag{6.16}$$

Formula (6.16) follows simply owing to the definition of L_j, and applying Fubini's theorem. The proof of (6.16) in the case of the convolution of two functions goes back to Blachman [11].

Thanks to (6.16), by setting the constants $\gamma_j = k_j/k$, we have at once that these constants satisfy the condition $\sum_{j=1}^{n} \gamma_j = 1$, and that the sign of the derivative (6.14), consequent to this choice, is negative, unless the functions u_j are Gaussian. Since the functional $\Phi(t)$ is dilation invariant, we can pass to the limit $t \to \infty$ obtaining

$$\lim_{t \to \infty} \Phi(t) = H(M_\kappa) - \sum_{j=1}^{n} \frac{\kappa_j}{\kappa} H(M_{\kappa_j}). \tag{6.17}$$

Since

$$H(M_\sigma) = \frac{1}{2} \log 2\pi\sigma,$$

we obtain from (6.17)

$$\lim_{t \to \infty} \Phi(t) = -\frac{1}{2} \sum_{j=1}^{n} \frac{\kappa_j}{\kappa} \log \frac{\kappa_j}{\kappa}. \tag{6.18}$$

Clearly, the same result holds in dimension $d \geq 1$. Hence we proved the following:

Theorem 6.2. *Let $\gamma_j \geq 0$, $j = 1, 2, \dots, n$ be such that $\sum_{j=1}^n \gamma_j = 1$, and let $\Phi(t)$ be the functional (6.13), where the functions $u_j(x,t)$, $j = 1, 2, \dots, n$, are solutions to the heat equation corresponding to the initial probability densities $f_j(x) \in L^1(\mathbb{R}^d)$, $d \geq 1$. Then, if the diffusion coefficients $\kappa_j = C\gamma_j$, $j = 1, 2, \dots, n$ and $C > 0$, $\Phi(t)$ is decreasing in time from*

$$\Phi(0) = H(f_1 * f_2 * \cdots * f_n) - \sum_{j=1}^n \gamma_j H(f_j)$$

to the limit value

$$\lim_{t\to\infty} \Phi(t) = -\frac{d}{2} \sum_{j=1}^n \gamma_j \log \gamma_j. \tag{6.19}$$

Moreover, $\Phi(0) = \lim_{t\to\infty} \Phi(t)$ if and only if $f_j(x)$, $j = 1, 2, \dots, n$, is a Gaussian density of variance $d\kappa_j$.

Theorem 6.2 shows the monotonicity of a dilation invariant functional linked to the Shannon's entropy of a n-th convolution of probability density functions. A direct consequence of this monotonicity is the entropy power inequality. Indeed, the monotonicity of $\Phi(t)$ implies that, for any choice of the constants γ_j, with $\sum_{j=1}^n \gamma_j = 1$

$$H(f_1 * f_2 * \cdots * f_n) \geq \sum_{j=1}^n \gamma_j H(u_j(t)) - \frac{d}{2} \sum_{j=1}^n \gamma_j \log \gamma_j. \tag{6.20}$$

Inequality (6.20) generalizes to n functions the result of Lemma 6.1. Shannon's entropy power inequality then follows by maximizing the right-hand side of (6.20) over the sequence γ_j.

7. Conclusions

In this paper we studied the monotonicity properties of various functionals related to convolutions of powers of solutions to the heat equation. This monotonicity is at the basis of a new proof of many well-known inequalities in sharp form, which are viewed in our picture as consequence of a unique well understandable physical principle, in the form of time monotonicity of a Lyapunov functional. Partial results of this strategy were presented in [36, 37, 38].

This idea has been applied here to prove classical Young's inequality and its converse, Brascamp–Lieb type inequalities, Babenko's inequality and Prékopa–Leindler inequality. In addition, a new direct proof of Shannon's entropy power inequality is shown to follow by the same argument.

Unlike similar results obtained in recent years (cf. [2, 3, 4, 7, 8, 9, 10, 12, 14]), we were inspired by some relatively old papers by people working on

information theory [11, 35] and kinetic theory of rarefied gases [30], mainly connected with classical Shannon's entropy and its monotonicity properties.

References

[1] K.I. Babenko, *An inequality in the theory of Fourier analysis.* Izv. Akad. Nauk SSSR, Ser. Mat **25** (1961), 531–542; English transl. Amer. Math. Soc. Transl. (2) **44** (1965), 115–128.

[2] F. Barthe, *Optimal Young's inequality and its converse: a simple proof.* Geom. Funct. Anal. **8** (1998), 234–242.

[3] F. Barthe and D. Cordero-Erausquin, *Inverse Brascamp-Lieb inequalities along the heat equation.* In: Geometric Aspects of Functional Analysis. Lecture Notes in Math., vol. 1850, 65–71. Springer, Berlin (2004)

[4] F. Barthe and N. Huet, *On Gaussian Brunnï£¡Minkowski inequalities.* Stud. Math. **191** (2009), 283–304

[5] J-P. Bartier, A. Blanchet, J. Dolbeault and M. Escobedo, *Improved intermediate asymptotics for the heat equation.* Appl. Math. Letters **24** (2011), 76–81.

[6] W. Bechner, *Inequalities in Fourier Analysis on* $\mathbb{R}^d$. Proc. Nat. Acad. Sci. USA **72** (1975), 638–641.

[7] J. Bennett, *Heat-flow monotonicity related to some inequalities in euclidean analysis.* Contemporary Mathematics **505** (2010), 85–96.

[8] J. Bennett and N. Bez, *Closure Properties of Solutions to Heat Inequalities.* J. Geom. Anal. **19** (2009), 584–600.

[9] J. Bennett, N. Bez and A. Carbery, *Heat-flow monotonicity related to the Hausdorff–Young inequality*, Bull. London Math. Soc. **41** (6) (2009), 971–979.

[10] J. Bennett, A. Carbery, M. Christ and T. Tao, *The Brascamp–Lieb Inequalities: Finiteness, Structure and Extremals.* Geometric And Functional Analysis **17** (2007), 1343–1415.

[11] N.M. Blachman, *The convolution inequality for entropy powers.* IEEE Trans. Inform. Theory **2** (1965), 267–271.

[12] C. Borell, *The Ehrhard inequality.* C.R. Math. Acad. Sci. Paris **337** (2003), 663–666

[13] H.J. Brascamp and E.H. Lieb, *Best constants in Young's inequality, its converse and its generalization to more than three functions.* Adx. Math. **20** (1976), 151–173.

[14] E.A. Carlen, E.H. Lieb, and M. Loss, *A sharp analog of Young's inequality on* S^N *and related entropy inequalities.* J. Geom. Anal. **14** (2004), 487–520.

[15] J.A. Carrillo, G. Toscani, *Exponential convergence toward equilibrium for homogeneous Fokker-Planck-type equations.* Math. Methods Appl. Sci., **21** (1998), 1269-1286.

[16] T.M. Cover and J.A. Thomas, *Elements of Information Theory.* 2nd ed. John Wiley & Sons, New York, 2006.

[17] M.H.M. Costa, *A new entropy power inequality.* IEEE Trans. Inf. Theory, **IT-31**, (6) (1985), 751–760.

[18] 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 **35**, (1989), 887–888.

[19] A. Dembo, T.M. Cover, and J.A. Thomas, *Information theoretic inequalities*. IEEE Trans. Inf. Theory **37** (6) (1991), 1501–1518.

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

[21] T. Goudon, S. Junca, and G. Toscani, *Fourier-based distances and Berry–Esseen like inequalities for smooth densities*. Monatsh. Math. **135** (2002) 115–136.

[22] D. Guo, S. Shamai, and S. Verdú, *Mutual information and minimum mean-square error in Gaussian channels*. IEEE Trans. Inf. Theory, **51**, (4) (2005), 1261–1282.

[23] D. Guo, S. Shamai, and S. Verdú, *A Simple Proof of the Entropy-Power Inequality*. IEEE Trans. Inf. Theory, **52**, (5) (2006), 2165–2166.

[24] R. G. Laha and V. K. Rohatgi, *Probability theory*. John Wiley & Sons, New York, 1979.

[25] L. Leindler, *On a certain converse of Hölder's inequality. II*. Acta. Sci. Math. Szeged **33** (1972), 217–223.

[26] E.H. Lieb, *Proof of an entropy conjecture of Wehrl*. Commun. Math. Phys., **62** (1978), 35–41.

[27] E.H. Lieb, *Gaussian kernels have only gaussian maximizers*. Invent. Math., **102** (1990), 179–208.

[28] Yu.V. Linnik, *An information-theoretic proof of the central limit theorem with the Lindeberg condition*. Theory Probab. Appl. **4** (1959), 288-299.

[29] P.L. Lions, G. Toscani, *A strenghened central limit theorem for smooth densities*. J. Funct. Anal. **129** (1995), 148–167.

[30] H.P. McKean Jr., *Speed of approach to equilibrium for Kacï£¡s caricature of a Maxwellian gas*. Arch. Rat. Mech. Anal. **21** (1966), 343ï£¡367.

[31] A. Prékopa, *Logarithmic concave measures with application to stochastic programming*. Acta. Sci. Math. Szeged **32** (1971), 301–315.

[32] A. Prékopa, *On logarithmic measures and functions*. Acta. Sci. Math. Szeged **34** (1973), 335–343.

[33] O. Rioul, *A simple proof of the entropy-power inequality via properties of mutual information*, in Proceedings of the 2007 IEEE International Symposium on Information Theory (ISIT 2007), Nice, France, June 24-29-th, 2007.

[34] C.E. Shannon, *A mathematical theory of communication*. Bell Syst. Tech. J. **27** Jul.(1948), 379–423; Oct. (1948), 623–656.

[35] A.J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*. Inf. Contr. **2** (1959), 101–112.

[36] G. Toscani, *Heat equation and the sharp Young's inequality*. arXiv:1204.2086 (2012)

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

[38] G. Toscani, *Lyapunov functionals for the heat equation and sharp inequalities*. Atti Acc. Peloritana Pericolanti, Classe Sc. Fis. Mat. e Nat., **91** (2013), 1–10.

[39] C. Villani, *A short proof of the concavity of entropy power*. IEEE Trans. Info. Theory **46**, (4) (2000), 1695-1696.

[40] R. Zamir and M. Feder, *A generalization of the entropy power inequality with applications*. IEEE Trans. Inf. Theory, **39** (5) (1993), 1723–1728.

Giuseppe Toscani
Department of Mathematics
University of Pavia
Via Ferrata 1
27100 Pavia
Italy
e-mail: giuseppe.toscani@unipv.it