

THE GEOMETRY OF EUCLIDEAN CONVOLUTION INEQUALITIES AND ENTROPY

DARIO CORDERO-ERAUSQUIN AND MICHEL LEDOUX

(Communicated by Marius Junge)

ABSTRACT. The goal of this paper is to show that some convolution type inequalities from Harmonic Analysis and Information Theory, such as Young's convolution inequality (with sharp constant), Nelson's hypercontractivity of the Hermite semi-group or Shannon's inequality, can be reduced to a simple geometric study of frames of \mathbb{R}^2 . We shall derive directly entropic inequalities, which were recently proved to be dual to the Brascamp-Lieb convolution type inequalities.

1. INTRODUCTION

The topic of Brascamp-Lieb and convolution type inequalities was recently renewed by Carlen, Lieb and Loss [11], who proposed a semi-group or heat flow approach to these inequalities. (Soon after, Bennett, Carbery, Christ and Tao [8] independently gave a semi-group approach to multidimensional Brascamp-Lieb inequalities.) Carlen, Lieb and Loss also obtained new inequalities on the sphere, and in particular a subadditivity of the entropy inequality. It was noted in [3] that this inequality can be proved using geometric properties of the Fisher information of the marginal distributions. Pushing forward these investigations, Carlen and Cordero-Erausquin [10] derived a similar geometric treatment for general subadditivity of the entropy inequalities on Euclidean space and proved that these inequalities are dual to the Brascamp-Lieb inequalities. The semi-group approach was recently carried out in the unifying framework of abstract Markov semi-groups in [4].

The abstract geometric argument in \mathbb{R}^n is particularly simple and for self-consistency, we describe it in detail in this introduction.

In the sequel, μ will stand either for the Lebesgue measure on \mathbb{R} or for the standard Gaussian probability measure γ on \mathbb{R} (with density $(2\pi)^{-1/2}e^{-|t|^2/2}$). Consequently, $\mu_n := \mu^{\otimes n}$ will stand for the Lebesgue measure or the standard Gaussian measure γ_n on \mathbb{R}^n . It is convenient to treat these two cases in parallel, although it is possible to derive formally one from another.

We say that f is a probability density with respect to (w.r.t.) μ_n if $f : \mathbb{R}^n \rightarrow \mathbb{R}^+$ is such that $\int f d\mu_n = 1$. Given a random vector $X \in \mathbb{R}^n$ with f as probability density w.r.t. μ_n (a relation written below as $X \sim f d\mu_n$), its *entropy* w.r.t μ_n is

Received by the editors July 16, 2009.

2010 *Mathematics Subject Classification*. Primary 42A85, 52A40, 60E15; Secondary 60G15, 94A17.

©2010 American Mathematical Society
Reverts to public domain 28 years from publication

defined (whenever it makes sense) by

$$S_{\mu_n}(X) := S_{\mu_n}(f) := \int_{\mathbb{R}^n} f \log f \, d\mu_n.$$

In the case that μ_n is the Lebesgue measure we shall use the notation $S(X) = S(f) = \int f \log f$. Throughout the paper, it will be implicitly assumed in all statements that we consider only densities and random vectors with well defined and finite μ_n -entropy.

If f is probability density w.r.t. μ_n on \mathbb{R}^n and $a \in \mathbb{R}^n$ is a fixed non-zero vector, denote by $f_{(a)}$ the marginal probability density w.r.t. $\mu = \mu_1$ on \mathbb{R} ; i.e. $f_{(a)} d\mu$ is the image of $f d\mu_n$ under the map $x \rightarrow x \cdot a$. Thus, $f_{(a)}$ is characterized by the requirement that

$$(1) \quad \int_{\mathbb{R}^n} \phi(x \cdot a) f(x) \, d\mu_n(x) = \int_{\mathbb{R}} \phi(t) f_{(a)}(t) \, d\mu_1(t)$$

for every bounded measurable $\phi : \mathbb{R} \rightarrow \mathbb{R}$. Equivalently, if $X \sim f d\mu_n$, then $f_{(a)} d\mu$ is the density of $X \cdot a$, that is, $X \cdot a \sim f_{(a)} d\mu$. Thus, $S_{\mu}(X \cdot a) = S_{\mu}(f_{(a)}) = \int_{\mathbb{R}} f_{(a)}(t) \log f_{(a)}(t) \, d\mu(t)$. The classical subadditivity of entropy (usually stated with the Lebesgue measure) indicates that for an orthonormal basis (u_1, \dots, u_n) of \mathbb{R}^n and a random vector X ,

$$(2) \quad \sum_{i=1}^n S_{\mu}(X \cdot u_i) \leq S_{\mu_n}(X).$$

The relation between subadditivity inequalities and Brascamp-Lieb inequalities is summarized in the following proposition.

Proposition 1 ([10]). *For non-zero vectors $a_1, \dots, a_m \in \mathbb{R}^n$, $c_1, \dots, c_m \geq 0$ and $D \in \mathbb{R}$, the following assertions are equivalent:*

(1) *For every $f_1, \dots, f_m : \mathbb{R} \rightarrow \mathbb{R}^+$,*

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(x \cdot a_i)^{c_i} \, d\mu_n(x) \leq e^D \prod_{i=1}^m \left(\int_{\mathbb{R}} f_i \, d\mu \right)^{c_i}.$$

(2) *For every random vector $X \in \mathbb{R}^n$, $\sum_{i=1}^m c_i S_{\mu}(X \cdot a_i) \leq S_{\mu_n}(X) + D$.*

We also have a complete equivalence between the equality cases. This result is easy to prove (actually it holds in much more general settings): it formally relies on the fact that the Legendre transform of the entropy functional is the functional $V \rightarrow \log \int e^V d\mu_n$, that is,

$$\log \int e^V d\mu_n = \sup_f \left\{ \int fV \, d\mu_n - S_{\mu_n}(f) \right\}$$

and

$$S_{\mu_n}(f) = \sup_V \left\{ \int fV \, d\mu_n - \log \int e^V d\mu_n \right\},$$

and on how this combines with (1). The dual of the subadditivity inequality (2) is nothing else but Fubini's theorem.

It is possible to consider more general geometric situations than the one of an orthonormal basis (2). Of particular interest is the case of a decomposition of the identity, as put forward by Ball in the context of Brascamp-Lieb inequalities (see

e.g. [1]). Given *unit* vectors u_1, \dots, u_m in the Euclidean space \mathbb{R}^n and real numbers $c_1, \dots, c_m > 0$, we say that they decompose the identity if

$$(3) \quad \sum_{i=1}^m c_i u_i \otimes u_i = \text{Id}_{\mathbb{R}^n},$$

where $u_i \otimes u_i$ stands for the orthogonal projection in the direction u_i . Note that necessarily $c_i \leq 1$ and

$$(4) \quad \sum_{i=1}^m c_i = n.$$

It is easy to derive sharp subadditivity entropy inequalities using (3) because such decompositions combine nicely with the Fisher information, as noted in [3]. The point is that the Fisher information has an L^2 structure which allows for geometric operations such as projections (or equivalently, conditional expectations). A random vector $X \in \mathbb{R}^n$ with f as probability density w.r.t. μ_n is said to have finite μ_n -Fisher information if the following quantity is well defined and finite:

$$I_{\mu_n}(X) := I_{\mu_n}(f) := \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\mu_n.$$

It follows from the Cauchy-Schwarz inequality (see [9, 10]) that for a unit vector $u \in \mathbb{R}^n$,

$$(5) \quad I_{\mu_1}(f(u)) = I_{\mu_1}(X \cdot u) \leq \int_{\mathbb{R}^n} \frac{(\nabla f \cdot u)^2}{f} d\mu_n,$$

with equality if and only if $X \cdot u$ and $X - (X \cdot u)u$ are independent. If we are given a decomposition of the identity, then, rewriting (3) in the form

$$(6) \quad \forall v \in \mathbb{R}^n, \quad \sum_{i=1}^m c_i (v \cdot u_i)^2 = |v|^2,$$

we immediately get from (5) that for any random vector X with finite Fisher information,

$$(7) \quad \sum_{i=1}^m c_i I_{\mu}(X \cdot u_i) \leq I_{\mu_n}(X).$$

In order to get an inequality for entropy, we integrate along the suitable semi-group. Let L stand for the differential operator $Lf = \Delta f$ (Laplacian) when μ_n is the Lebesgue measure, and $Lf = \Delta f - x \cdot \nabla f$ in the case $\mu_n = \gamma_n$. Let $P_t = e^{tL}$ be the corresponding heat semi-group and Ornstein-Uhlenbeck semi-group, which admit as invariant measure the Lebesgue and the Gaussian measure respectively. If X is a random vector with density f with respect to μ_n and finite μ_n -entropy, then $e^{tL}f$ is a smooth probability density with respect to μ_n , with finite μ_n -Fisher information, and

$$\frac{d}{dt} S_{\mu_n}(e^{tL}f) = -I_{\mu_n}(e^{tL}f).$$

Moreover, L (and thus e^{tL}) has the property that it preserves the algebra of functions of the form $f(x) = g(x \cdot u)$, a property also used in the proof of (5). This ensures the following crucial property; namely, for every $t \geq 0$,

$$(8) \quad (e^{tL}f)_{(u)} = e^{tL}(f_{(u)}) \quad \text{on } \mathbb{R},$$

where we used the same notation for the one-dimensional and n -dimensional semi-groups. The heat and Ornstein-Uhlenbeck semi-groups are the two most important diffusion semi-groups sharing property (8), and this explains the particular role played here by the Lebesgue and Gaussian measures.

Now, integration of (7) along the semi-group e^{tL} leads to the inequality $\sum_{i=1}^m c_i S_\mu(X \cdot u_i) \leq S_{\mu_n}(X)$. From the cases of equality in (5) we get that equality holds if and only if for each $i \leq m$, $X \cdot u_i$ and $X - (X \cdot u_i)u_i$ are independent (a property which is preserved along the semi-group). Under mild conditions on the vectors u_i , it is easily seen that this can happen only when X is a Gaussian vector (see [10] for details). The previous discussion is summarized in the next proposition, established in [10].

Proposition 2. *Consider a decomposition of the identity (3) in \mathbb{R}^n . Then, for all random vectors $X \in \mathbb{R}^n$,*

$$(9) \quad \sum_{i=1}^m c_i S_\mu(X \cdot u_i) \leq S_{\mu_n}(X).$$

Furthermore, under the conditions that no two of the unit vectors $\{u_i\}$ are linearly dependent and that if any one vector u_i is removed from $\{u_1, \dots, u_m\}$, the remaining vectors still span \mathbb{R}^n , equality holds in (9) if and only if X is a Gaussian random variable whose covariance is a multiple of the identity.

In view of the duality given by Proposition 1, we recover from the previous inequality Ball's form of the Brascamp-Lieb inequality: for every $f_1, \dots, f_m : \mathbb{R} \rightarrow \mathbb{R}^+$,

$$(10) \quad \int_{\mathbb{R}^n} \prod_{i=1}^m f_i(x \cdot u_i)^{c_i} d\mu_n(x) \leq \prod_{i=1}^m \left(\int_{\mathbb{R}} f_i d\mu \right)^{c_i}.$$

Moreover (under the same hypothesis on the u_i 's as in the previous proposition), equality holds if and only if the measures $f_i d\mu$ are Gaussian with the same covariance: $f_i(t) d\mu(t) = \lambda_i e^{-\alpha^2(t-v \cdot a_i)^2} dt$ with $\lambda_i > 0$, $\alpha \in \mathbb{R}^*$ and $v \in \mathbb{R}^n$ ($v = 0$ if we restrict to centered functions). Note that by using (6) it is also possible to pass from inequalities for the Lebesgue measure to inequalities for the standard Gaussian measure (and *vice versa*) by the correspondence $f_i \longleftrightarrow f_i(t) e^{-|t|^2/2}$.

The goal of this paper is to prove the efficiency of the theoretical aforementioned approach in some meaningful situations. More precisely, we try to understand convolution and information-theoretic inequalities, such as the sharp Young convolution inequality, Nelson's hypercontractivity of the Hermite semi-group, or Shannon's inequality, as functional forms of some particular decompositions of the identity of \mathbb{R}^2 . To do so, we will study in §2 the decompositions of the identity of \mathbb{R}^2 by three vectors. We give a complete description of the relation between the coefficients c_i and the vectors u_i in this case. As a consequence, we obtain the following general inequality which can be viewed as the functional form of Proposition 5 below.

Theorem 3. *Let $p_1, p_2, p_3 > 1$ be such that*

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2,$$

and $\theta_2, \theta_3 \in]0, \pi[$ be defined by

$$(\cos(\theta_2), \sin(\theta_2)) = \left(\sqrt{(p_1 - 1)(p_2 - 1)}, \sqrt{\frac{p_1 p_2 (p_3 - 1)}{p_3}} \right)$$

and

$$(\cos(\theta_3), \sin(\theta_3)) = \left(-\sqrt{(p_1 - 1)(p_3 - 1)}, \sqrt{\frac{p_1 p_3 (p_2 - 1)}{p_2}} \right).$$

Then, for all random variables $X, Y \in \mathbb{R}$,

$$(11) \quad \frac{1}{p_1} S_\mu(X) + \frac{1}{p_2} S_\mu(\cos(\theta_2)X + \sin(\theta_2)Y) + \frac{1}{p_3} S_\mu(\cos(\theta_3)X + \sin(\theta_3)Y) \leq S_{\mu_2}(X, Y)$$

with equality if and only if X and Y are independent identically distributed Gaussian variables.

Equivalently, for every function $g \in L^{p_2}(\mu)$ and $h \in L^{p_3}(\mu)$, setting $p'_1 = \frac{p_1}{p_1 - 1}$,

$$(12) \quad \left\| \int g(\cos(\theta_2)x + \sin(\theta_2)y) h(\cos(\theta_3)x + \sin(\theta_3)y) d\mu(y) \right\|_{L^{p'_1}(d\mu(x))} \leq \|g\|_{L^{p_2}(\mu)} \|h\|_{L^{p_3}(\mu)}$$

with equality if and only if f and g are of constant sign with $|g(t)|^{p_2} d\mu(t) = K_2 e^{-\lambda(t-a_2)^2} dt$ and $|h(t)|^{p_3} d\mu(t) = K_3 e^{-\lambda(t-a_3)^2} dt$ for $a_1, a_2 \in \mathbb{R}$ and $K_2, K_3, \lambda \geq 0$.

We apply this result in §3 to the determination of the sharp constant in the Young convolution inequality. For this, we will work directly with the entropy and exploit the following simple but useful invariance of the entropy (in the case of Lebesgue measure) under linear transformations: for a random vector $X \in \mathbb{R}^n$ and an invertible linear operator A on \mathbb{R}^n ,

$$(13) \quad S(AX) = S(X) - \log(|\det(A)|).$$

In §4 we derive the Shannon inequality from a limit of decompositions of the identity. The hope is to shed in this way new light on the connection between subadditivity of entropy and Shannon’s inequality. We study similarly Gaussian inequalities in §5 and derive in particular hypercontractivity of the Hermite semi-group (and the associated logarithmic Sobolev inequality). For simplicity, we consider here only one-dimensional inequalities. In the last section, §6, we briefly explain how to extend word by word the approach to multidimensional situations.

2. DECOMPOSITION OF THE IDENTITY OF \mathbb{R}^2

As announced, we investigate here decompositions of the identity of \mathbb{R}^2 . A decomposition of the identity with only two vectors u_1 and u_2 holds if and only if these two vectors form an orthonormal basis and $c_1 = c_2 = 1$. In order to get something of interest, consider the case of three distinct unit vectors in \mathbb{R}^2 : u_1, u_2 and u_3 . Note that $u_i \otimes u_i = (-u_i) \otimes (-u_i)$, so here and in the sequel, ‘distinct’ really means that the directions $\mathbb{R}u_i$ are distinct. The first question we address is the following: if the directions are given, can we find *positive* numbers c_1, c_2, c_3 such that the decomposition of the identity

$$(14) \quad c_1 u_1 \otimes u_1 + c_2 u_2 \otimes u_2 + c_3 u_3 \otimes u_3 = \text{Id}_{\mathbb{R}^2}$$

holds? The answer is yes provided the vectors are ‘well enough’ distributed in space.

Proposition 4. *Let $\mathbb{R}u_1, \mathbb{R}u_2$ and $\mathbb{R}u_3$ be three distinct directions of \mathbb{R}^2 . There exist three positive numbers c_1, c_2 and c_3 such that the decomposition of the identity (14) holds if and only if the three geometric angles given by the six angular sectors defined by these directions are all strictly smaller than $\frac{\pi}{2}$. The c_i ’s are then given by*

$$(15) \quad c_i = \frac{\cos(\theta_j - \theta_k)}{\sin(\theta_j - \theta_i) \sin(\theta_k - \theta_i)} = 1 - \cot(\theta_i - \theta_j) \cot(\theta_k - \theta_i)$$

for (i, j, k) a permutation of $(1, 2, 3)$ and $u_i = (\cos(\theta_i), \sin(\theta_i))$ for $i = 1, 2, 3$.

Proof. If two of the vectors are orthogonal, say $u_1 \cdot u_2 = 0$, and if (14) holds, then $0 = u_1 \cdot u_2 = 0 + 0 + c_3 (u_3 \cdot u_1) (u_3 \cdot u_2)$; and therefore, if $c_3 > 0$, then $u_3 \cdot u_1 = 0$ or $u_3 \cdot u_2 = 0$. But this implies $u_3 = \pm u_2$ or $u_3 = \pm u_1$, which is excluded. A genuine three vector situation cannot contain a two vector situation (which is equivalent to an orthonormal basis).

By the assumption, the projections $u_1 \otimes u_1, u_2 \otimes u_2$ and $u_3 \otimes u_3$ span \mathcal{S}_2 , the 3-dimensional space of symmetric operators on \mathbb{R}^2 . Therefore the linear operator

$$(c_1, c_2, c_3) \longrightarrow c_1 u_1 \otimes u_1 + c_2 u_2 \otimes u_2 + c_3 u_3 \otimes u_3$$

is an isomorphism from \mathbb{R}^3 onto \mathcal{S}_2 . Write $u_i = (\cos(\theta_i), \sin(\theta_i))$ for $i = 1, 2, 3$, where all coordinate computations are done in the canonical orthonormal basis of \mathbb{R}^2 ; note that $\theta_i - \theta_j \neq 0[\frac{\pi}{2}]$. Using that

$$u_i \otimes u_i = \begin{pmatrix} \cos(\theta_i)^2 & \cos(\theta_i) \sin(\theta_i) \\ \cos(\theta_i) \sin(\theta_i) & \sin(\theta_i)^2 \end{pmatrix},$$

it is readily checked that the unique solution in \mathbb{R}^3 of (14) is given by

$$c_1 = \frac{\cos(\theta_2 - \theta_3)}{\sin(\theta_2 - \theta_1) \sin(\theta_3 - \theta_1)}, \quad c_2 = \frac{\cos(\theta_3 - \theta_1)}{\sin(\theta_3 - \theta_2) \sin(\theta_1 - \theta_2)}, \quad c_3 = \frac{\cos(\theta_1 - \theta_2)}{\sin(\theta_1 - \theta_3) \sin(\theta_2 - \theta_3)}.$$

It remains to identify when this gives a solution to our problem, i.e. when $c_1, c_2, c_3 > 0$.

All the quantities in the previous equation remain unchanged if we replace some θ_i by $\theta_i + \pi$, which is consistent with the fact that we have been working with directions only, and of course, they also remain unchanged by rotations, i.e. by $\theta_i \rightarrow \theta_i + \alpha$ for $i = 1, 2, 3$ and $\alpha \in \mathbb{R}$. Therefore, up to a relabeling of the directions, we can assume that $0 = \theta_1 < \theta_2 < \theta_3 < \pi$. Then the c_i ’s are positive if and only if

$$\theta_3 - \theta_2 < \pi/2, \quad \theta_3 - 0 > \pi/2, \quad \theta_2 - 0 < \pi/2.$$

Rewriting the second condition as $\pi - \theta_3 < \pi/2$, we get the announced condition on the three angular sectors $\theta_2, \theta_3 - \theta_2, \pi - \theta_3$. □

We now investigate the converse procedure. Given three numbers $c_1, c_2, c_3 \in (0, 1)$ such that

$$(16) \quad c_1 + c_2 + c_3 = 2,$$

we would like to know whether it is possible to find directions $u_i = (\cos(\theta_i), \sin(\theta_i))$ for which the decomposition of the identity (14) holds. The answer is yes, and the construction is unique up to an isometry of \mathbb{R}^2 (which clearly preserves decompositions of the identity).

Proposition 5. *For given $c_1, c_2, c_3 \in (0, 1)$ satisfying (16), there exists a triple of directions $\mathbb{R}u_1, \mathbb{R}u_2$ and $\mathbb{R}u_3$ unique up to isometries such that the decomposition of the identity (14) holds (the solutions $u_i = (\cos(\theta_i), \sin(\theta_i))$ are given by equation (18) below). More explicitly, all solutions are obtained by performing an isometry on*

$$(17) \quad \begin{aligned} u_1 &= (1, 0), \quad u_2 = \left(\sqrt{\frac{(1-c_1)(1-c_2)}{c_1 c_2}}, \sqrt{\frac{1-c_3}{c_1 c_2}} \right), \\ u_3 &= \left(-\sqrt{\frac{(1-c_1)(1-c_3)}{c_1 c_3}}, \sqrt{\frac{1-c_2}{c_1 c_3}} \right). \end{aligned}$$

Proof. Inverting formally (15) we get

$$(18) \quad \begin{aligned} \cot(\theta_2 - \theta_3) &= \varepsilon \sqrt{\frac{(1-c_2)(1-c_3)}{1-c_1}}, \quad \cot(\theta_3 - \theta_1) = \varepsilon \sqrt{\frac{(1-c_3)(1-c_1)}{1-c_2}}, \\ \cot(\theta_1 - \theta_2) &= \varepsilon \sqrt{\frac{(1-c_1)(1-c_2)}{1-c_3}} \end{aligned}$$

and $\varepsilon = \pm 1$. This uniquely determines the directions $\mathbb{R}u_i$ up to isometries. To check this, first perform a rotation ensuring that $\theta_1 = 0$ and $\theta_2, \theta_3 \in (0, \pi)$. We still have an invariance by symmetry with respect to the coordinate axis $x = 0$, which corresponds to the sign of ε . Thus, without loss of generality, we can impose

$$(19) \quad 0 = \theta_1 < \theta_2 < \theta_3 < \pi.$$

The last two equalities in (18) give that $\cot(\theta_2)$ and $\cot(\theta_3)$ are of opposite sign, and thus, by (19), $\varepsilon = -1$, $\theta_2 \in (0, \pi/2)$ and $\theta_3 \in (\pi/2, \pi)$, these angles being uniquely determined by

$$(20) \quad \cot(\theta_2) = \sqrt{\frac{(1-c_1)(1-c_2)}{1-c_3}} \quad \text{and} \quad \cot(\theta_3) = -\sqrt{\frac{(1-c_3)(1-c_1)}{1-c_2}}.$$

Note that the first equality of (18) ensures that $\theta_3 - \theta_2 \in (0, \pi/2)$, which is consistent with the condition in the previous theorem. So long to unicity. It remains now to check that we indeed get a solution. Equivalently, back to the situation (19)-(20), we need to check that once the last two equalities from (18) are used to uniquely determine the angles θ_2 and θ_3 , the first equality of (18) is then automatically verified. This is indeed the case since

$$\sqrt{\frac{1-c_1}{(1-c_2)(1-c_3)}} \times \cot(\theta_3 - \theta_2) = \sqrt{\frac{1-c_1}{(1-c_2)(1-c_3)}} \times \frac{\cot(\theta_2)\cot(\theta_3)+1}{\cot(\theta_2)-\cot(\theta_3)} = \frac{c_1}{(1-c_2)+(1-c_3)}.$$

Therefore the compatibility condition (16) yields the desired equation. Finally, the solution (19)-(20) can be rewritten as (17) in coordinates, and the proof is complete. \square

We can now derive the main theorem stated in the introduction.

Proof of Theorem 3. Introducing, for $i = 1, 2, 3$, $c_i = \frac{1}{p_i}$, the unit vectors $u_i = (\cos(\theta_i), \sin(\theta_i))$ of the previous proposition can be rewritten as $u_1 = (1, 0)$,

$$u_2 = \left(\sqrt{(p_1 - 1)(p_2 - 1)}, \sqrt{\frac{p_1 p_2 (p_3 - 1)}{p_3}} \right),$$

$$u_3 = \left(-\sqrt{(p_1 - 1)(p_3 - 1)}, \sqrt{\frac{p_1 p_3 (p_2 - 1)}{p_2}} \right).$$

Then the result for entropy follows from Proposition 2. Next note that it is enough to prove (12) in the case of non-negative functions, and therefore the integral inequality to be proven is

$$\iint f(x) g(\cos(\theta_2)x + \sin(\theta_2)y) h(\cos(\theta_3)x + \sin(\theta_3)y) d\mu(x)d\mu(y) \leq \|f\|_{L^{p_1}(\mu)} \|g\|_{L^{p_2}(\mu)} \|h\|_{L^{p_3}(\mu)}$$

for $f, g, h : \mathbb{R} \rightarrow \mathbb{R}^+$. But this inequality holds as the dual of inequality (11) by virtue of Proposition 1. The cases of equality follow from the general considerations given in the introduction. \square

3. SHARP YOUNG’S CONVOLUTION INEQUALITY

Here we work with the Lebesgue measure. Let $p, q, r > 1$ be such that

$$(21) \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

which can be rewritten as

$$\frac{1}{r'} + \frac{1}{p} + \frac{1}{q} = 2.$$

Then apply Theorem 3 with $p_1 = r'$, $p_2 = p$, $p_3 = q$. The angles θ_2, θ_3 , or equivalently the unit vectors $u_i = (\cos(\theta_i), \sin(\theta_i))$, $i = 1, 2, 3$, given by Theorem 3 are

$$(22) \quad u_1 = (1, 0), \quad u_2 = \left(\sqrt{(r' - 1)(p - 1)}, \sqrt{\frac{r' p}{q'}} \right), \quad u_3 = \left(-\sqrt{(r' - 1)(q - 1)}, \sqrt{\frac{r' q}{p'}} \right).$$

Therefore, for any random vector $(W, Z) \in \mathbb{R}^2$ (with finite entropy),

$$(23) \quad \frac{1}{r'} S(W) + \frac{1}{p} S(\cos(\theta_2)W + \sin(\theta_2)Z) + \frac{1}{q} S(\cos(\theta_3)W + \sin(\theta_3)Z) \leq S(W, Z).$$

We would like to have as random variables in the left-hand side multiples of W , $W - Z$ and Z , respectively. To this task, first perform a linear transformation leaving W invariant so that the last variable is a multiple of Z , and then a diagonal linear operator so that the second one is a multiple of $W - Z$. Readily, perform the linear transformation $(X, Y) = A(W, Z)$ with

$$A := \begin{pmatrix} \cot(\theta_3) & 1 \\ \cot(\theta_3) - \cot(\theta_2) & 0 \end{pmatrix}.$$

Then, by using (13), it follows that (23) is equivalent to the following subadditivity inequality: for all random variables X and Y ,

$$\begin{aligned} \frac{1}{r'} S\left(\frac{1}{\cot(\theta_3) - \cot(\theta_2)} Y\right) + \frac{1}{p} S(\sin(\theta_2)(X - Y)) + \frac{1}{q} S(\sin(\theta_3)X) \\ \leq S(X, Y) + \log |\cot(\theta_3) - \cot(\theta_2)|. \end{aligned}$$

This inequality is also equivalent, using again the scaling of entropy (in dimension 1) to

$$(24) \quad \frac{1}{r'} S(X) + \frac{1}{p} S(X - Y) + \frac{1}{q} S(Y) \leq S(X, Y) + D$$

with

$$D = \left(1 - \frac{1}{r'}\right) \log |\cot(\theta_3) - \cot(\theta_2)| + \frac{1}{p} \log \sin(\theta_2) + \frac{1}{q} \log \sin(\theta_3).$$

Using that $\frac{1}{p'} + \frac{1}{q'} = \frac{1}{r'}$, it follows that $\cot(\theta_2) - \cot(\theta_3) = \frac{1}{r'} \sqrt{\frac{p'q'}{r}}$ and

$$D = -\frac{1}{r} \log \sqrt{r} + \frac{1}{r'} \log \sqrt{r'} + \frac{1}{p} \log \sqrt{p} - \frac{1}{p'} \log \sqrt{p'} + \frac{1}{q} \log \sqrt{q} - \frac{1}{q'} \log \sqrt{q'}.$$

For $t > 1$, set $C_t := \sqrt{\frac{t^{1/t}}{t^{1/t'}}$, where as before t' is the conjugate of t . We have thus derived the following classical result.

Theorem 6 (Sharp Young’s convolution inequality). *Let $p, q, r > 1$ satisfy (21). For all random variables $X, Y \in \mathbb{R}$,*

$$\frac{1}{r'} S(X) + \frac{1}{p} S(X - Y) + \frac{1}{q} S(Y) \leq S(X, Y) + \log \left(\frac{C_p C_q}{C_r}\right).$$

*Furthermore, the inequality is sharp: equality holds if and only if $(X, Y) \in \mathbb{R}^2$ is a Gaussian vector whose covariance matrix is a multiple of A^*A .*

Equivalently, for every $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$,

$$(25) \quad \|f * g\|_{L^r(\mathbb{R})} \leq \frac{C_p C_q}{C_r} \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})}.$$

For the equality cases in the entropic inequality, note that in view of Proposition 2 equality holds in (23) if and only if (W, Z) is a Gaussian vector with covariance a multiple of $\text{Id}_{\mathbb{R}^2}$, and $(X, Y) = A(W, Z)$. Next, note that Young’s inequality (25) reduces to the case of non-negative functions, which is then equivalent to the following dual form of the entropic inequality:

$$(26) \quad \iint f(x)g(x - y)h(y) dx dy \leq \frac{C_p C_q}{C_r} \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})} \|h\|_{L^q(\mathbb{R})}$$

for all non-negative functions $f, g, h : \mathbb{R} \rightarrow \mathbb{R}^+$. It is possible to deduce the equality cases in this inequality (and therefore in Young’s convolution inequality) from the ones in the entropic inequality, as described in [10] (we get some well-chosen Gaussian functions). Actually, it is also possible to use that inequality (26) may be obtained by rescaling the functions (after the change of variables $(x, y) = A(w, z)$) in the Brascamp-Lieb inequality dual to (23), where equality holds if and only the functions are Gaussian with the same covariance.

The sharp Young convolution inequality (25) was obtained independently by Beckner [5] and Brascamp and Lieb [7]. Their proofs rely on rearrangements of

functions and tensorization arguments. Barthe [2] gave a new (simpler) proof using a mass transportation argument. A heat semi-group proof was recently given by Bennett and Bez [6]. The geometric approach based on entropy proposed here might prove useful for the study of Young's inequalities in other contexts.

4. SHANNON'S INEQUALITY

We continue to work with the Lebesgue measure. We now try to reproduce the following classical result in Information Theory (see [12] for details).

Theorem 7 (Shannon's inequality). *Let X and Y be two independent random variables. Then*

$$(27) \quad S\left(\frac{X+Y}{\sqrt{2}}\right) \leq \frac{S(X)+S(Y)}{2}.$$

It is well known in the case that Y (say) is symmetric, $Y \sim -Y$, that Shannon's inequality follows from the classical subadditivity of the entropy (2) since

$$2S\left(\frac{X+Y}{\sqrt{2}}\right) = S\left(\frac{X+Y}{\sqrt{2}}\right) + S\left(\frac{X-Y}{\sqrt{2}}\right) \leq S(X, Y) = S(X) + S(Y),$$

where the last equality expresses the independence of X and Y . However this situation is misleading since in the general case Shannon's inequality seems to be different in nature than an inequality of subadditivity of entropy. One of the obstacles is that we would like to use a decomposition of the identity with the basis vectors e_1 and e_2 together with $\frac{e_1+e_2}{\sqrt{2}}$. But this is not possible. We shall instead approximate such a situation.

Recall that if G stands for a standard Gaussian variable independent of all the variables considered here, then if X has finite entropy, $S(X + \varepsilon G) \rightarrow S(X)$ when $\varepsilon \rightarrow 0$. Therefore we can restrict our study to the case where X and Y have smooth densities (with sub-Gaussian tails). For such regular variables, it is well known and easy to see that when $\varepsilon \rightarrow 0$,

$$(28) \quad S(X + \varepsilon Y) = S(X) + O(\varepsilon^2).$$

Using the notation $u(\theta) = (\cos(\theta), \sin(\theta))$, introduce for fixed $s \in (-\frac{\pi}{2}, 0)$ (s will later tend to 0^-) the unit vectors

$$u_1 = u(s), \quad u_2 = u\left(\frac{\pi}{4}\right), \quad u_3 = u\left(\frac{\pi}{2} - s\right).$$

These vectors define three directions satisfying the assumption of Proposition 4. Let $c_1, c_2, c_3 \in (0, 1)$ be the associated coefficients for which there is a decomposition of the identity so that, by Proposition 2,

$$(29) \quad c_1 S(\cos(s)X + \sin(s)Y) + c_2 S\left(\frac{X+Y}{\sqrt{2}}\right) + c_3 S(\sin(s)X + \cos(s)Y) \leq S(X, Y)$$

for all random variables X and Y with finite entropy. By Proposition 4, as $s \rightarrow 0^-$,

$$\begin{aligned} c_1 &= \frac{\cos(s)}{\sin\left(\frac{\pi}{4} - s\right) \sin\left(\frac{\pi}{2} - 2s\right)} = 1 + 2s + o(s), \\ c_2 &= \frac{\cos\left(\frac{\pi}{2} - 2s\right)}{\sin\left(\frac{\pi}{4} - s\right) \sin\left(s - \frac{\pi}{4}\right)} = -4s + o(s), \\ c_3 &= \frac{\cos\left(\frac{\pi}{2} - s\right)}{\sin\left(2s - \frac{\pi}{2}\right) \sin(s)} = 1 + 2s + o(s). \end{aligned}$$

Note also that when X and Y are (regular enough) random variables and $s \rightarrow 0^-$, we have, in view of (13) and (28), that

$$S(\cos(s)X + \sin(s)Y) = S\left(X + \frac{\sin(s)}{\cos(s)}Y\right) - \log \cos(s) = S(X) + o(s)$$

and similarly $S(\sin(s)X + \cos(s)Y) = S(Y) + o(s)$. Therefore, making a Taylor expansion in (29) when X and Y are independent and $s \rightarrow 0^-$, we obtain that

$$S(X) + S(Y) - s \left[4S\left(\frac{X+Y}{\sqrt{2}}\right) - 2S(X) - 2S(Y) \right] + o(s) \leq S(X, Y) = S(X) + S(Y).$$

The first order in $s < 0$ gives the desired Shannon inequality.

Remark 8. In view of the duality between entropy and Brascamp-Lieb inequalities, one could wonder if the Shannon inequality admits a dual form. However, if we start with the Brascamp-Lieb inequality, in the same situation as above, and perform the Taylor expansion there, then we end up again with the Shannon inequality. We are in a situation where the entropic and Brascamp-Lieb inequalities coincide at the first order.

It should be noted that one can prove along the same lines the Blachmann-Stam inequality (cf. [12]):

$$(30) \quad I\left(\frac{X+Y}{\sqrt{2}}\right) \leq \frac{I(X) + I(Y)}{2}$$

for independent random variables, X and Y with finite Fisher information. Indeed, the decomposition of the identity obtained above and Proposition 2 and (7) give the result once it has been noted that for regular enough random variables, as for entropy, $I(X + \varepsilon Y) = I(X) + O(\varepsilon^2)$ as $\varepsilon \rightarrow 0$. Moreover, for independent random variables, $I(X, Y) = I(X) + I(Y)$. Note that by the scaling of information, inequality (30) is commonly rewritten as $I(X + Y) \leq I(X) + I(Y)$.

Finally, we would like to mention that we could as well have started from Theorem 3. A first order Taylor expansion when $p_2 = p_3 \rightarrow 2$ (and therefore $p_1 \rightarrow 1$) in (11) gives again the Shannon inequality. Having derived the sharp Young inequality from (11), this procedure is reminiscent of Dembo’s proof of Shannon’s inequality, which consisted in a Taylor expansion in (25) around $p = q = 2$ (see [12]).

5. HYPERCONTRACTIVITY AND LOGARITHMIC SOBOLEV INEQUALITY

In this section, the measure $\mu = \gamma$ will be the Gaussian measure and $\mu_n = \gamma_n$. Assume we are given p, q with

$$(31) \quad 1 < p < q$$

and set $\theta \in [0, \frac{\pi}{2})$ such that

$$(32) \quad \cos(\theta) = \sqrt{\frac{p-1}{q-1}}.$$

Write as before $t' = t/(t-1)$ for $t > 1$. Since $q > p$, let $r \in [1, q)$ be such that $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$, or equivalently

$$\frac{1}{q'} + \frac{1}{p} + \frac{1}{r} = 2.$$

Then, introducing $\xi \in (\frac{\pi}{2}, \pi)$ such that

$$(33) \quad \cos(\xi) = -\sqrt{\frac{r-1}{q-1}},$$

the angles θ, ξ are exactly the ones associated to the triple

$$p_1 = q', \quad p_2 = p, \quad p_3 = r$$

in Theorem 3 ($\theta_1 = \theta, \theta_2 = \xi$). Consequently, for all random variables X, Y ,

$$(34) \quad \frac{1}{q'} S_\gamma(X) + \frac{1}{p} S_\gamma(\cos(\theta)X + \sin(\theta)Y) + \frac{1}{r} S_\gamma(\cos(\xi)X + \sin(\xi)Y) \leq S_{\gamma_2}(X, Y).$$

We emphasize in the next proposition the corresponding convolution inequality (12) of independent interest (as a stronger statement than the classical hypercontractivity).

Proposition 9 (Hypercontractivity). *Let $p, r > 1$. If $\frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{q}$ and (32)-(33) hold, then for every function $f \in L^p(\gamma)$ and $g \in L^r(\gamma)$,*

$$\left\| \int f(\cos(\theta)x + \sin(\theta)y) g(\cos(\xi)x + \sin(\xi)y) d\gamma(y) \right\|_{L^q(d\gamma(x))} \leq \|f\|_{L^p(\gamma)} \|g\|_{L^r(\gamma)},$$

with equality if and only if f and g are of constant sign with $|f(x)|^p d\gamma(x) = K_1 e^{-\lambda|x-a_1|^2} dx$ and $|g(x)|^r d\gamma(x) = K_2 e^{-\lambda|x-a_2|^2} dx$, $K_1, K_2, \lambda \geq 0$, $a_1, a_2 \in \mathbb{R}$.

This result contains the hypercontractivity inequality for the Hermite semi-group

$$H_\theta(f)(x) := \int f(\cos(\theta)x + \sin(\theta)y) d\gamma(y).$$

(One may work as well with the Ornstein-Uhlenbeck semi-group $P_t f := e^{tL} f = H_{\arccos(e^{-t})} f$.) Indeed, Proposition 9 applied with $g \equiv 1$ indicates that under (31)-(32),

$$(35) \quad \forall f \in L^p(\gamma), \quad \|H_\theta f\|_{L^q(\gamma)} \leq \|f\|_{L^p(\gamma)}.$$

Equality holds iff f is exponential, $f(x) = Ke^{-ax}$ ($\lambda = 1/2$, since $g = 1$ in the proposition). The hypercontractivity of the Hermite semi-group goes back to the seminal works of Nelson [14] and Gross [13]. The cases of equality were treated by Carlen in [9].

If instead we work at the level of entropic inequalities, first note that for every random variable Z , $S_\gamma(Z) \geq 0$, with equality if Z is a standard Gaussian variable. Therefore, inequality (34) implies that, for all random variables X, Y ,

$$(36) \quad \frac{1}{q'} S_\gamma(X) + \frac{1}{p} S_\gamma(\cos(\theta)X + \sin(\theta)Y) \leq S_{\gamma_2}(X, Y),$$

which is dual to the following Brascamp-Lieb inequality, equivalent to (35): for all functions $g, f : \mathbb{R} \rightarrow \mathbb{R}^+$,

$$\iint g(x) f(\cos(\theta)x + \sin(\theta)y) d\gamma(x) d\gamma(y) \leq \|f\|_{L^p(\gamma)} \|g\|_{L^{q'}(\gamma)}.$$

Remark 10 (Particularity of the Gaussian case). It is worth noting that in the Gaussian case, inequalities (9) and (10) hold under the weaker condition

$$(37) \quad \sum_{i=1}^m c_i u_i \otimes u_i \leq \text{Id}_{\mathbb{R}^n}.$$

(So here we can simply use that $\frac{1}{q}e_1 \otimes e_1 + \frac{1}{p}u(\theta) \otimes u(\theta) \leq \text{Id}_{\mathbb{R}^2}$, instead of ‘forgetting’ terms, as we did above.) The argument goes back to [3] and relies on the fact that we are working on a probability space. From the explanations given in the introduction, it is clear that (37) is always sufficient to get (7). When integrating along the heat semi-group, it is necessary to rescale in order to obtain asymptotically a standard Gaussian. So condition (4) is crucial there (see [10]). But in the Gaussian case, there is no need to rescale when integrating along the Ornstein-Uhlenbeck semi-group, and so we can indeed derive inequalities (9) and (10) from (37). Alternatively, starting from (37), we can complete the self-adjoint operator on the RHS in order to get a decomposition of the identity by adding some unit vectors u'_j and coefficients c'_j . We then simply apply (9) and (10) in the case of this decomposition, but with $X_j = G$ (standard Gaussian independent of the rest) and $f_j \equiv 1$, respectively, for the added indices.

It is well known that the Gaussian logarithmic Sobolev inequality is equivalent to the hypercontractivity inequality (35) [13]. To derive the logarithmic Sobolev inequality, one can differentiate (35) at $\theta = 0$. Let us explain how it is even easier to see this implication when working with the dual entropic form. Recall that for every random variable Z , $S_\gamma(Z) \geq 0$ with equality if Z is a standard Gaussian variable. Let X be a random variable and G be a standard Gaussian variable independent of X . For $\theta \in [0, \frac{\pi}{2})$, set

$$H_\theta X := \cos(\theta)X + \sin(\theta)G.$$

Then, inequality (36) gives that, under (31)-(32),

$$\frac{1}{q'}S_\gamma(X) + \frac{1}{p}S_\gamma(H_\theta X) \leq S_{\gamma_2}(X, G) = S_\gamma(X),$$

which can be rewritten as $S_\gamma(H_\theta X) \leq \frac{p}{q}S_\gamma(X)$. Therefore, for any $\theta \in [0, \frac{\pi}{2})$ and $q > 1$, by picking the appropriate p satisfying (36), $S_\gamma(H_\theta X) \leq \frac{1+(q-1)\cos^2(\theta)}{q}S_\gamma(X)$. Letting $q \rightarrow +\infty$,

$$S_\gamma(H_\theta X) \leq \cos^2(\theta)S_\gamma(X),$$

which is the well-known integrated form of the logarithmic Sobolev inequality. Indeed, since there is equality at $\theta = 0$, the θ^2 order term gives the logarithmic Sobolev inequality

$$S_\gamma(X) \leq \frac{1}{2}I_\gamma(X) = -\frac{d^2}{d\theta^2}\Big|_{\theta=0} S_\gamma(H_\theta X).$$

6. HIGHER-DIMENSIONAL INEQUALITIES

We have studied convolution inequalities for functions on \mathbb{R} and for random variables. But the strategy applies word by word to convolution inequalities for functions on \mathbb{R}^n and random vectors. Let us briefly explain why.

All subspaces of \mathbb{R}^n are equipped with the Euclidean structure inherited from the standard Euclidean structure on \mathbb{R}^n . Accordingly, for a subspace $E \subset \mathbb{R}^n$, the measure μ_E will stand for the Lebesgue measure or the standard Gaussian measure on E . Denote by P_E the orthogonal projection in \mathbb{R}^n onto E , and if f is a probability density w.r.t. μ_n , denote by $f_{(E)}$ the probability density w.r.t. μ_E

which is the image of $f d\mu_n$ under the map P_E . For every random vector $X \in \mathbb{R}^n$, we have that $X \sim f d\mu_n \Rightarrow P_E X \sim f_{(E)} d\mu_E$ and

$$\int_E g(y) f_{(E)}(y) d\mu_E(y) = \int_{\mathbb{R}^n} g(P_E x) f(x) d\mu_n(x).$$

The following analogue of (5) immediately holds: for f a (smooth) probability density w.r.t. μ_n (or any random vector with $X \sim f d\mu_n$),

$$(38) \quad I_{\mu_E}(f_{(E)}) = I_{\mu_E}(P_E X) \leq \int_{\mathbb{R}^n} \frac{|P_E \nabla f|^2}{f} d\mu_n.$$

Assume we are given a collection of subspaces $E_1, \dots, E_m \subset \mathbb{R}^n$ and positive numbers $c_1, \dots, c_m > 0$ such that

$$(39) \quad \sum_{i=1}^m c_i P_{E_i} = \text{Id}_{\mathbb{R}^n}.$$

Then, by (38)-(39), $\sum_{i=1}^m c_i I_{\mu_{E_i}}(P_{E_i} X) \leq I_{\mu_n}(X)$. After integration along appropriate semi-groups P_t (noting again that $(P_t f)_{(E_i)} = P_t(f_{(E_i)})$), we get the analogue of Proposition 2:

$$\sum_{i=1}^m c_i S_{\mu_{E_i}}(P_{E_i} X) \leq S_{\mu_n}(X),$$

and by duality, for $f_i : E_i \rightarrow \mathbb{R}^+$, $i = 1, \dots, m$, the classical (multidimensional) geometric Brascamp-Lieb inequality

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i^{c_i}(P_{E_i} x) d\mu_n \leq \prod_{i=1}^m \left(\int_{E_i} f_i d\mu_{E_i} \right)^{c_i}.$$

The convolution inequalities on \mathbb{R}^n are obtained by using appropriate projections onto three n -dimensional subspaces of \mathbb{R}^{2n} . Given an angle $\theta \in [0, \pi]$, denote by P_θ the projection in \mathbb{R}^{2n} obtained by tensoring the projection on the direction $u(\theta)$ in \mathbb{R}^2 :

$$P_\theta = U_\theta^* U_\theta, \quad \text{with } U_\theta = (\cos(\theta)\text{Id}_{\mathbb{R}^n} \quad \sin(\theta)\text{Id}_{\mathbb{R}^n}) : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n.$$

Identifying \mathbb{R}^{2n} with $\mathbb{R}^n \times \mathbb{R}^n$, $P_\theta(x, y) = U_\theta^*(\cos(\theta)x + \sin(\theta)y)$ for $x, y \in \mathbb{R}^n$. The projection onto the first \mathbb{R}^n is P_0 : $P_0(x, y) = x$. Note that the image subspace $E_\theta := \text{Im}(P_\theta)$ is normally parametrized by \mathbb{R}^n as $E_\theta = \{U_\theta^* z \mid z \in \mathbb{R}^n\}$. Therefore, for a random vector $(X, Y) \in \mathbb{R}^{2n}$,

$$S_{\mu_{E_\theta}}(P_\theta(X, Y)) = S_{\mu_n}(\cos(\theta)X + \sin(\theta)Y).$$

Assume then we are given $p_1, p_2, p_3 > 1$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2$, and let θ_2, θ_3 be the angles given by Theorem 3. These angles came from the decomposition of the identity of Proposition 5, which extends to a decomposition of the identity of \mathbb{R}^{2n} :

$$\frac{1}{p_1} P_0 + \frac{1}{p_2} P_{\theta_2} + \frac{1}{p_3} P_{\theta_3} = \text{Id}_{\mathbb{R}^{2n}}.$$

By the previous considerations, Theorem 3 immediately extends to random vectors on \mathbb{R}^n and functions on \mathbb{R}^n : for all random vectors $X, Y \in \mathbb{R}^n$,

$$\begin{aligned} \frac{1}{p_1} S_{\mu_n}(X) + \frac{1}{p_2} S_{\mu_n}(\cos(\theta_2)X + \sin(\theta_2)Y) + \frac{1}{p_3} S_{\mu_n}(\cos(\theta_3)X + \sin(\theta_3)Y) \\ \leq S_{\mu_{2n}}(X, Y). \end{aligned}$$

Similarly, for all functions $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ with $g \in L^{p_2}(\mu_n)$ and $h \in L^{p_3}(\mu_n)$,

$$\left\| \int g(\cos(\theta_2)x + \sin(\theta_2)y) h(\cos(\theta_3)x + \sin(\theta_3)y) d\mu_n(y) \right\|_{L^{p'_1}(d\mu_n(x))} \leq \|g\|_{L^{p_2}(\mu_n)} \|h\|_{L^{p_3}(\mu_n)}.$$

The cases of equality are also the same. Then, the multidimensional forms of Young's convolution inequality, Shannon's inequality, hypercontractivity and the logarithmic Sobolev inequality are obtained by the same computations we have performed previously.

Of course, it is also possible to derive these inequalities from the one-dimensional ones by standard tensorization techniques. But as pointed out earlier, the geometric approach used here might prove useful in some other contexts.

REFERENCES

- [1] K. Ball, *Convex geometry and functional analysis*, in *Handbook of the Geometry of Banach Spaces*, Vol. I, 161–194, eds. W. Johnson and J. Lindenstrauss, North-Holland, Amsterdam, 2001. MR1863692 (2003c:52001)
- [2] F. Barthe, *Optimal Young's inequality and its converse, a simple proof*, *Geom. Funct. Analysis* **80** (1998), 234–242. MR1616143 (99f:42021)
- [3] F. Barthe, D. Cordero-Erausquin and B. Maurey, *Entropy of spherical marginals and related inequalities*, *J. Math. Pures Appl.* **86** (2006), 89–99. MR2247452 (2008a:62004)
- [4] F. Barthe, D. Cordero-Erausquin, M. Ledoux and B. Maurey, *Correlation and Brascamp-Lieb inequalities for Markov semigroups*, preprint (2009).
- [5] W. Beckner, *Inequalities in Fourier analysis*, *Ann. of Math. (2)* **102** (1975), 159–182. MR0385456 (52:6317)
- [6] J. Bennett and N. Bez, *Closure properties of solutions to heat inequalities*, *J. Geom. Anal.* **19** (2009), 584–600. MR2496567
- [7] H. Brascamp and E. Lieb, *The best constant in Young's inequality and its generalization to more than three functions*, *Advances in Math.* **20** (1976), 151–173. MR0412366 (54:492)
- [8] J. Bennett, A. Carbery, M. Christ and T. Tao, *The Brascamp-Lieb inequalities: Finiteness, structure, and extremals*, *Geom. Funct. Analysis* **17** (2008), 1343–1415. MR2377493 (2009c:42052)
- [9] E. Carlen, *Superadditivity of Fisher information and logarithmic Sobolev inequalities*, *J. Funct. Analysis* **101** (1991), 194–211. MR1132315 (92k:94006)
- [10] E. Carlen and D. Cordero-Erausquin, *Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities*, *Geom. Funct. Analysis* **19** (2009), 373–405. MR2545242
- [11] E. Carlen, E. Lieb and M. Loss, *A sharp form of Young's inequality on S^N and related entropy inequalities*, *Jour. Geom. Analysis* **14** (2004), 487–520. MR2077162 (2005k:82046)
- [12] A. Dembo, T. Cover and J. Thomas, *Information-theoretic inequalities*, *IEEE Trans. Inform. Theory* **37** (1991), 1501–1518. MR1134291 (92h:94005)
- [13] L. Gross, *Logarithmic Sobolev inequalities*, *Amer. J. Math.* **97** (1975), 1061–1083. MR0420249 (54:8263)
- [14] E. Nelson, *A quartic interaction in two dimensions*, in *Mathematical Theory of Elementary Particles* (Proc. Conf., Dedham, Mass., 1965), M.I.T. Press, 1966, pp. 69–73. MR0210416 (35:1309)

INSTITUT DE MATHÉMATIQUES DE JUSSIEU, UNIVERSITÉ PIERRE ET MARIE CURIE (PARIS 6),
4 PLACE JUSSIEU, 75252 PARIS CEDEX 05, FRANCE

E-mail address: `cordero@math.jussieu.fr`

INSTITUT DE MATHÉMATIQUES DE TOULOUSE, UNIVERSITÉ DE TOULOUSE, 31062 TOULOUSE,
FRANCE

E-mail address: `ledoux@math.univ-toulouse.fr`