

The Brunn–Minkowski theorem and related geometric and functional inequalities

Franck Barthe*

Abstract. The Brunn–Minkowski inequality gives a lower bound of the Lebesgue measure of a sum-set in terms of the measures of the individual sets. It has played a crucial role in the theory of convex bodies. This topic has many interactions with isoperimetry or functional analysis. Our aim here is to report some recent aspects of these interactions involving optimal mass transport or the Heat equation. Among other things, we will present Brunn–Minkowski inequalities for flat sets, or in Gauss space, as well as local versions of the theorem which apply to the study of entropy production in the central limit theorem.

Mathematics Subject Classification (2000). Primary 39B62; Secondary 52A40, 26D15, 94A17.

Keywords. Brunn–Minkowski, transport, Heat equation, entropy.

1. Introduction

The Brunn–Minkowski theory studies the relations between addition of vectors and the volume of convex sets. Let us start with some notation. For $\lambda \in \mathbb{R}$ and A a subset of \mathbb{R}^d , one sets $\lambda A = \{\lambda a; a \in A\}$. The Minkowski sum of two sets $A, B \subset \mathbb{R}^d$ is by definition

$$A + B := \{a + b; (a, b) \in A \times B\}.$$

The Brunn–Minkowski inequality gives a lower bound on the volume of a sum-set.

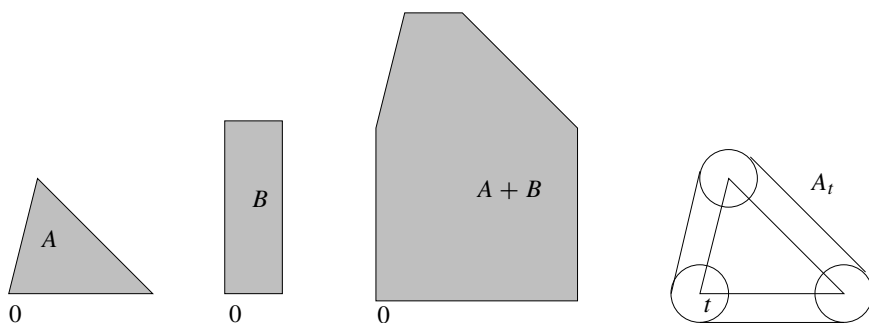
Theorem 1.1. *Let A, B be non-empty compact subsets of \mathbb{R}^d , then*

$$\text{Vol}_d(A + B)^{\frac{1}{d}} \geq \text{Vol}_d(A)^{\frac{1}{d}} + \text{Vol}_d(B)^{\frac{1}{d}}. \quad (1)$$

If A, B are convex homothetic sets, there is equality. Brunn discovered this result in 1887 for A, B convex in dimension at most 3. Minkowski proved the inequality for convex sets in arbitrary dimension and realized the importance of the statement. Indeed, it could be combined with a former result by Steiner, which calculated the volume of the t -enlargement of a convex compact set $A \subset \mathbb{R}^3$ defined for $t > 0$ by

$$A_t = \{x \in \mathbb{R}^3; \text{there exists } y \in A \text{ such that } |x - y| \leq t\}.$$

*Research was supported in part by the European Network PHD, MCRN-511953.



Note that $A_t = A + tB^3$ where B^d denotes the Euclidean unit ball of \mathbb{R}^d . The Steiner formula asserts that for $t > 0$

$$\text{Vol}_3(A + tB^3) = \text{Vol}_3(K) + tS(A) + 2\pi t^2W(A) + \frac{4}{3}\pi t^3,$$

where $S(A)$ is the surface area of A and $W(A)$ is its mean width (the average on unit vectors u of the width of the minimal slab orthogonal to u and containing A). The Brunn–Minkowski theorem provides relations between the above coefficients. Indeed, it implies that

$$\text{Vol}_3(A + tB^3) \geq (\text{Vol}_3(A)^{\frac{1}{3}} + t\text{Vol}_3(B^3)^{\frac{1}{3}})^3$$

with equality at $t = 0$. Comparing derivatives at zero gives

$$S(A) \geq 3\text{Vol}_3(B^3)^{\frac{1}{3}}\text{Vol}_3(A)^{\frac{2}{3}}.$$

This is the classical isoperimetric inequality; it means that among sets of given volume, balls have minimal surface area (the argument actually extends to non-convex sets). Another relation can be obtained by noting that the Brunn–Minkowski inequality shows that $\text{Vol}_3(A + tB^3)^{\frac{1}{3}}$ is a concave function of $t \geq 0$ when A is convex.

Minkowski extended the Steiner formula as follows: for non-empty compact convex sets $K_1, \dots, K_m \subset \mathbb{R}^d$ and $\lambda_1, \dots, \lambda_m \geq 0$, the volume of $\lambda_1 K_1 + \dots + \lambda_m K_m$ is a homogeneous polynomial of the form

$$\text{Vol}_d(\lambda_1 K_1 + \dots + \lambda_m K_m) = \sum_{i_1, \dots, i_d=1}^m \lambda_{i_1} \dots \lambda_{i_d} V(K_{i_1}, \dots, K_{i_d}).$$

Here and by definition $V(K_1, \dots, K_d)$ is the mixed volume of d convex sets in \mathbb{R}^d . The theory of mixed volumes studies the properties of these quantities, their geometric interpretations and the inequalities among them. We refer to the book [53] for more on this topic. See also [52] where such volume estimates are used in the local theory of Banach spaces.

2. Functional extensions, functional tools

There exist several proofs of the Brunn–Minkowski theorem, see the surveys [34], [36] for details and precise references. However the most fruitful approach is probably the one based on the following functional version of the statement:

Theorem 2.1 (Prékopa–Leindler). *Let f, g, h be measurable non-negative functions on \mathbb{R}^d and $\lambda \in [0, 1]$. If for all x, y in \mathbb{R}^d ,*

$$h(\lambda x + (1 - \lambda)y) \geq f^\lambda(x)g^{1-\lambda}(y),$$

then $\int_{\mathbb{R}^d} h \geq (\int_{\mathbb{R}^d} f)^\lambda (\int_{\mathbb{R}^d} g)^{1-\lambda}$.

Remark 2.2. When applied to characteristic functions of sets, the above result provides a multiplicative version of the Brunn–Minkowski inequality, which is equivalent to the one we stated. The functional inequality can be written with an outer integral, as

$$\int_{\mathbb{R}^d}^* \sup_{\lambda x + (1-\lambda)y = z} f^\lambda(x)g^{1-\lambda}(y) dz \geq \left(\int_{\mathbb{R}^d} f\right)^\lambda \left(\int_{\mathbb{R}^d} g\right)^{1-\lambda}.$$

It appears as a reverse form of the classical inequality of Hölder which asserts that the right hand side in the latter inequality is at least $\int f^\lambda g^{1-\lambda}$.

An elementary proof of the above inequality appears in [52]. Here we sketch another proof. Its main idea is quite old and appears e.g. in [38]. It contains implicitly the idea of measure transportation which allowed recent developments.

Proof. We work in dimension 1, the general case follows by induction. By approximation arguments one may restrict to positive continuous f and g . One introduces functions $x, y: [0, 1] \rightarrow \mathbb{R}$ satisfying for $t \in [0, 1]$

$$\int_{-\infty}^{x(t)} f = t \int f; \quad \int_{-\infty}^{y(t)} g = t \int g. \tag{2}$$

Consequently for $t \in [0, 1]$ it holds $x'(t)f(x(t)) = \int f$ and $y'(t)g(y(t)) = \int g$. One defines a third function z on $[0, 1]$ by $z(t) = \lambda x(t) + (1 - \lambda)y(t)$. Our three functions are strictly increasing. Comparing geometric mean and arithmetic mean yields $z'(t) \geq (x'(t))^\lambda (y'(t))^{1-\lambda}$. Finally we use z as a change of variables to evaluate the integral of h . Using the above relations

$$\begin{aligned} \int h &\geq \int_0^1 h(z(t))z'(t) dt \\ &\geq \int_0^1 f^\lambda(x(t))g^{1-\lambda}(y(t))(x'(t))^\lambda (y'(t))^{1-\lambda} dt \\ &= \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda}. \quad \square \end{aligned}$$

2.1. Measure transportation. If μ, ν are two probability measures on \mathbb{R}^n , one says that a map $T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ transports μ to ν if $T\mu = \nu$, meaning $\nu(B) = \mu(T^{-1}(B))$ for every Borel set B .

In dimension 1, a canonical choice always exist when the first measure has no atoms: one can choose T non-decreasing. Note that the maps x, y of the previous proof are particular non-decreasing transporting maps. Indeed, if μ is the Lebesgue measure restricted to $[0, 1]$ and $d\nu(t) = f(t)dt / \int f$ the first relation in (2) can be rewritten as follows

$$\mu((-\infty, t]) = \nu((-\infty, x(t)]).$$

Since x is increasing and onto, this is equivalent to $\nu(B) = \mu(T^{-1}(B))$ for $B = (-\infty, s]$, and this relation extends to the Borel σ -field.

In higher dimension a remarkable analogue is available due to the works of Brenier [22] and McCann [48].

Theorem 2.3. *Let μ, ν be probability measures on \mathbb{R}^d . Assume that whenever $B \subset \mathbb{R}^d$ is a Borel set with Hausdorff dimension $d - 1$ one has $\mu(B) = 0$. Then there exists a convex function $\Phi: \mathbb{R}^d \rightarrow \mathbb{R}$ such that the map $T = \nabla\Phi$ (defined a.e.) satisfies $\nu = T\mu$. The map T is uniquely determined almost everywhere too.*

If μ and ν have second moments, then among all maps S with $S\mu = \nu$, T minimizes the quadratic transportation cost

$$\int_{\mathbb{R}^d} |x - S(x)|^2 d\mu(x).$$

As recalled in the second part of the theorem, this monotone transport T is related to the theory of optimal transportation, which looks for the best way to ship some amount of material from a configuration to another one. We refer to the book [59] for more on this fascinating topic.

If we consider measures with densities ρ_μ, ρ_ν with respect to Lebesgue's measure, then Φ is a generalized solution for the Monge–Ampère equation

$$\rho_\mu(x) = \rho_\nu(\nabla\Phi(x)) \det(\text{Hess}\Phi(x)).$$

Weak and strong regularity theory for this equation were developed respectively by McCann [47] and Caffarelli [23]. McCann also introduced the following interpolation between the measures μ and ν : $((1-t)I + tT)\mu = \nabla(x \mapsto (1-t)|x|^2/2 + t\Phi(x))\mu$ for $t \in [0, 1]$. He found applications to equilibrium states (and also to a proof of the Brunn–Minkowski inequality). Optimal transport allows to interpolate between general densities. However it has more structure when Gaussian measures are involved, as Caffarelli proved:

Theorem 2.4 ([24]). *Let Q be a positive definite quadratic form on \mathbb{R}^d let $d\mu(x) = e^{-Q(x)}dx/Z$ be the corresponding Gaussian probability measure. Let $d\nu = \rho d\mu$ be another probability measure with log-concave density ρ with respect to μ (i.e. $\rho(\lambda x + (1-\lambda)y) \geq \rho(x)^\lambda \rho(y)^{1-\lambda}$ for $x, y \in \mathbb{R}^d$ and $\lambda \in [0, 1]$). Then the monotone*

transportation map T such that $T\mu = \nu$ is a contraction for the canonical Euclidean distance.

2.2. Heat equation. Let $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ be the Laplace operator in \mathbb{R}^d . Following the probabilistic normalization we define the heat semigroup as $P_t = e^{t\Delta/2}$. More precisely, for a function f on \mathbb{R}^d the function $u(t, x) = P_t f(x)$ solves the equation $\partial_t u = \frac{1}{2}\Delta u$ on $\mathbb{R}^+ \times \mathbb{R}^d$ with initial condition $u(0, \cdot) = f$. When f has three bounded derivatives, one has

$$P_t f(x) = \int_{\mathbb{R}^d} f(z) e^{-\frac{(z-x)^2}{2t}} \frac{dz}{(2\pi t)^{d/2}} \quad \text{for } t > 0.$$

C. Borell discovered that the heat flow preserves in some sense the hypothesis of the Prékopa–Leindler inequality. More precisely, given $\lambda \in (0, 1)$ and three sufficiently regular non-negative functions $f, g, h: \mathbb{R}^d \rightarrow \mathbb{R}^+$ satisfying

$$h(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda g(y)^{1-\lambda},$$

he proved [18] that for all $t > 0$ and all x, y the following is true

$$P_t h(\lambda x + (1 - \lambda)y) \geq P_t f(x)^\lambda P_t g(y)^{1-\lambda}.$$

The Prékopa–Leindler inequality is obtained in the limit $t \rightarrow +\infty$ since

$$P_t f(x) \sim_{t \rightarrow +\infty} (2\pi t)^{-d/2} \left(\int f(y) dy \right).$$

Borell’s method was recently applied with success to derive other important Brunn–Minkowski type results. We will describe them in the next sections.

Remark 2.5. It has been known for many years that the heat equation is a powerful tool to prove functional inequalities of geometric flavor. In particular Bakry and Emery developed a general framework for deriving logarithmic Sobolev inequalities (which ensure Gaussian concentration of measure), or Sobolev type inequalities by semi-group techniques. More recently Bakry and Ledoux were able to prove Bobkov’s functional form of the Gaussian isoperimetric inequality along these lines. It was also observed that the Brunn–Minkowski inequality implies various types of isoperimetric inequality. So morally, the use of the heat equation for Brunn–Minkowski type inequalities is not a complete surprise. The interested reader will find details in [3], [44], [43]. Recently the transportation method also allowed to derive concentration estimates and Sobolev type inequalities, see e.g. [59], [32].

2.3. Riemannian manifolds. McCann [49] has solved the optimal transport problem on a Riemannian manifold when the transportation cost is the square of the geodesic distance. This provides a natural generalization of the monotone map, and allowed remarkable extensions of the Prékopa–Leindler inequality by Cordero-Erausquin, McCann and Schmuckenschläger [31], [30]. The following statement is valid under a curvature assumption in the spirit of Bakry–Emery

Theorem 2.6 ([30]). *Let (M, g) be a Riemannian manifold, and let μ be a measure on M with density e^{-V} with respect to the volume measure. Assume that for $\rho \in \mathbb{R}$, the Ricci curvature and the Hessian of V satisfy*

$$\text{Hess}_x V + \text{Ric}_x \geq \rho g$$

for all $x \in M$. Let $\lambda \in [0, 1]$ and $f, g, h: M \rightarrow \mathbb{R}^+$ such that for all $x, y \in M$ and all z such that $d(x, z) = (1 - \lambda)d(x, y)$ and $d(z, y) = \lambda d(x, y)$ one has

$$h(z) \geq e^{-\rho d^2(x,y)/2} f^\lambda(x) g^{1-\lambda}(y),$$

then one gets: $\int_M h d\mu \geq \left(\int_M f d\mu\right)^\lambda \left(\int_M g d\mu\right)^{1-\lambda}$.

The condition on the intermediate point z (involving geodesic distances) simply means that z is a geodesic barycenter of x, y with weights $\lambda, 1-\lambda$. Unlike in Euclidean spaces, there might be many of them.

3. Multilinear inequalities

The Brascamp–Lieb [21] inequalities are a powerful extension of Hölder’s inequality. Their original motivation was the calculation of the best constant in Young’s convolution inequality. Their most general form was established by Lieb. The setting of the theorem is the following. For $1 \leq i \leq m$, one considers linear surjective maps $B_i: \mathbb{R}^n \rightarrow \mathbb{R}^{n_i}$ and numbers $c_i \in [0, 1]$.

Theorem 3.1 (Lieb [45]). *The best constant $K \in [0, +\infty]$ such that the inequality*

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(B_i x)^{c_i} dx \leq K \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i\right)^{c_i}$$

holds for all integrable functions $f_i: \mathbb{R}^{n_i} \rightarrow \mathbb{R}^+$ can be computed by considering only centered Gaussian functions $f_i(x) = \exp(-\langle A_i x, x \rangle)$ where the A_i ’s are symmetric positive definite matrices of size n_i .

Homogeneity shows that the constant may be finite only when $\sum_{i=1}^m c_i n_i = n$. This condition is assumed in the following. Since Gaussian integral may be computed, one gets $K = D^{-1/2}$ where

$$D = \inf_{A_i > 0} \frac{\det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)}{\prod_{i=1}^m \det(A_i)^{c_i}}. \tag{3}$$

Here B_i^* denotes the adjoint of B_i . The proofs of Brascamp–Lieb and Lieb relied partially on tensorization arguments in higher dimension. We gave another proof using the monotone transport when proving an extension of the Prékopa–Leindler

inequality conjectured by K. Ball. The inequality is a reverse form of the Brascamp–Lieb inequality. The argument of proof is sophistication of the one given in the previous section, and gives both inequalities at a time. It uses the fact that the Jacobian matrices of monotone transport are symmetric positive matrices. This matches exactly the quantity appearing in the calculation of the Gaussian constant (3). The statement is

Theorem 3.2 ([10]). *The best constant $L \geq 0$ such that for all integrable functions $f_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^+$ one has*

$$\int_{\mathbb{R}^n} \sup_{\sum c_i B_i^* x_i = x; x_i \in \mathbb{R}^{n_i}} \prod_{i=1}^m f_i(x_i)^{c_i} dx \geq L \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i \right)^{c_i},$$

can be computed on centered Gaussian functions, and $L = \sqrt{D}$.

Our motivation for studying these inequalities came from convex geometry. Ball first understood the relevance of the Brascamp–Lieb inequality for this topic. In the case $n_i = 1$, $B_i(x) = \langle x, u_i \rangle$ where u_i are unit vectors in \mathbb{R}^n with the additional condition

$$\text{Id}_{\mathbb{R}^n} = \sum_{i=1}^m c_i P_{u_i}$$

(P_u is the orthogonal projection onto the line spanned by u) he showed that $D = 1$. So for non-negative functions on \mathbb{R} one has

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i^{c_i}(\langle x, u_i \rangle) dx \leq \prod_{i=1}^m \left(\int_{\mathbb{R}} f_i \right)^{c_i}.$$

Applied to characteristic functions of intervals this inequality gives an upper bound on the volume of an intersection of slabs. This was one of the crucial ingredients in Ball’s exact estimates on slices of the cubes or on the volume ratios of convex bodies, see [4] for details. The reverse Brascamp–Lieb inequality allows to estimate from below the volumes of convex hull and of sums of flat sets. For example, we obtain the following extension of the Brunn–Minkowski inequality

Theorem 3.3. *Let $(E_i)_{i=1}^m$ are vector-subspaces of \mathbb{R}^n and $c_i \in (0, 1]$ be such that $\text{Id}_{\mathbb{R}^n} = \sum_{i=1}^m c_i P_{E_i}$. Set $n_i = \dim(E_i)$. If $K_i \subset E_i$ then*

$$\text{Vol}_n \left(\sum_{i=1}^m c_i K_i \right) \geq \prod_{i=1}^m \text{Vol}_{n_i}(K_i)^{c_i}.$$

It was recently understood that the Brascamp–Lieb inequalities can be derived using the heat equation. This new approach is due to Carlen, Lieb and Loss [26] for functions of one variable and was developed to full generality by Bennett, Carbery,

Christ and Tao [15]. One advantage is that it allows a better description of equality cases. However, contrary to the mass transport approach, the heat equation method requires to know in advance which Gaussian functions are best, and to find a way around when there is no best Gaussian function. It was necessary to understand more precisely the Gaussian optimization problem summed up in Equation (3) and to understand when the constant D is positive (this corresponds to a non-trivial inequality) and when it is achieved (i.e. when a Gaussian maximizer exists). We present answer to the first question, which is of independent interest. The case of functions of one variable has a more explicit solution:

Theorem 3.4 ([10], [26]). *Let $(u_i)_{i \leq m}$ be non-zero vectors in \mathbb{R}^n . There exists a finite constant K such that for all integrable functions $f_i: \mathbb{R} \rightarrow \mathbb{R}^+$,*

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(\langle x, u_i \rangle)^{c_i} dx \leq K \prod_{i=1}^m \left(\int_{\mathbb{R}} f_i \right)^{c_i}$$

if and only if $c = (c_1, \dots, c_m)$ belongs to the set

$$\begin{aligned} \mathcal{C} &= \text{conv}\{\mathbf{1}_I; I \subset \{1, \dots, m\} \text{ and } (u_i)_{i \in I} \text{ is a basis}\} \\ &= \{c \in \mathbb{R}_+^m; \sum_{i=1}^m c_i n_i = n \text{ and for all } S \subset \{1, \dots, m\}, \\ &\quad \sum_{i \in S} c_i \leq \dim(\text{Span}(u_i, i \in S))\}. \end{aligned}$$

Here $\mathbf{1}_I$ is a vector in \mathbb{R}^m whose i th coordinate is 1 if $i \in I$ and 0 otherwise.

In the general case, only a description by facets of the set of exponents leading to a finite constant (*domain of finiteness*) is available

Theorem 3.5 ([16], [15]). *There exist $K < +\infty$ such that for all $f_i: \mathbb{R}^{n_i} \rightarrow \mathbb{R}^+$*

$$\int_{\mathbb{R}^n} \prod_{i=1}^m (f_i \circ B_i)^{c_i} \leq K \prod_{i=1}^m \left(\int_{\mathbb{R}} f_i \right)^{c_i}$$

if and only if for all i , $c_i \geq 0$, $\sum_{i=1}^m c_i n_i = n$ and for all vector subspaces $V \subset \mathbb{R}^n$ it holds

$$\dim V \leq \sum_{i=1}^m c_i \dim(B_i V).$$

Remark 3.6. In the interior of the domain of finiteness, Gaussian maximizers exist and the inequality is equivalent to the multidimensional version of Ball’s form (also called the geometric form) of the Brascamp–Lieb inequality: if vector subspaces E_i and numbers $c_i \in (0, 1]$ are such that $\text{Id}_{\mathbb{R}^n} = \sum_{i=1}^m c_i P_{E_i}$ then for non-negative functions $f_i: E_i \rightarrow \mathbb{R}^+$, one has

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

This is proved using the heat semigroup, showing that $t \mapsto \int \prod_{i=1}^m ((P_t f_i) \circ P_{E_i})^{c_i}$ is non-decreasing and interpolates between the two terms of the above inequality. On the boundary of the finiteness domain, a factorization argument allows to reduce the dimension and conclude by induction.

Remark 3.7. The reverse Brascamp–Lieb inequality can be proved by the heat flow too, along the lines of Borell’s argument for the Prékopa–Leindler inequality. This is written in [12] for functions of one variables and the geometric form. However this easily extends as well as the others steps of the proof.

But the heat equation approach does not only provide us with new proofs. It allows remarkable extensions of the results together with new applications. We refer to [15] for inequalities restricted to special classes of functions. Carlen, Lieb and Loss where able to prove similar inequalities in other spaces as the sphere [26] and the symmetric group [25]. The spherical inequality was motivated by the study of a system of n particles in one dimension, preserving total kinetic energy. Hence their n speeds form a vector in the Euclidean sphere $S^{n-1} \subset \mathbb{R}^n$. In order to know how the information on an individual particle influences the one of the whole system, they established the following: for $f_i : [-1, 1] \rightarrow \mathbb{R}^+$,

$$\int_{S^{n-1}} \prod_{i=1}^m f_i(x_i) d\sigma(x) \leq \prod_{i=1}^m \left(\int_{S^{n-1}} f_i(x_i)^2 d\sigma(x) \right)^{\frac{1}{2}},$$

where σ is the uniform probability measure on the sphere. The surprise here is the 2-norm, which is best possible and in particular does not disappear when $n \rightarrow +\infty$. In [14] this is extended to general decompositions of the identity $\text{Id}_{\mathbb{R}^n} = \sum_{i=1}^m c_i P_{E_i}$, where for functions $f_i : E_i \rightarrow \mathbb{R}^+$ it holds

$$\int_{S^{n-1}} \prod_{i=1}^m f_i(P_{E_i}x)^{c_i/2} d\sigma(x) \leq \prod_{i=1}^m \left(\int_{S^{n-1}} f_i(P_{E_i}x) d\sigma(x) \right)^{c_i/2}.$$

This allows to consider particle systems in d dimension, and also with fixed momentum for example. However the exponents $c_i/2$ may not be best possible in this generality. In the work [13] a general framework of commuting Markov generator is developed to deal with these inequalities in general settings. Also the geometric meaning of the best exponents is better understood in continuous settings, and several new examples are provided.

4. Geometry in Gauss space

Let us denote by γ_d the standard Gaussian probability measure on \mathbb{R}^d with density with respect to Lebesgue’s measure given by $\rho(x) = (2\pi)^{-d/2} \exp(-|x|^2/2)$, $x \in \mathbb{R}^d$. There is no need to emphasize its importance, and it is natural and useful to have

Brunn–Minkowski type inequalities for γ_d . Applying the Prékopa–Leindler theorem to $f = \rho \mathbf{1}_A$, $g = \rho \mathbf{1}_B$, where $A, B \subset \mathbb{R}^d$ and using the log-concavity of ρ yields

$$\gamma_d^*(\lambda A + (1 - \lambda)B) \geq \gamma_d(A)^\lambda \gamma_d(B)^{1-\lambda}. \quad (4)$$

This inequality is not sharp. An optimal version was proved by Ehrhard [35] for convex sets, using a symmetrization procedure. Latała [40] showed next that one of the sets may be non-convex and recently Borell [19] completely removed the convexity assumption. His approach is functional and uses the Heat equation. The most general version of his result is given below.

Theorem 4.1 ([20]). *Let $\lambda, \mu \geq 0$ with $\lambda + \mu \geq 1$ and $|\lambda - \mu| \leq 1$. Then for all measurable sets $A, B \subset \mathbb{R}^d$ the following holds:*

$$\Phi^{-1}(\gamma_d^*(\lambda A + \mu B)) \geq \lambda \Phi^{-1}(\gamma_d(A)) + \mu \Phi^{-1}(\gamma_d(B)),$$

where Φ is the distribution function of γ_1 , defined by $\Phi(t) = \int_{-\infty}^t e^{-u^2/2} du / \sqrt{2\pi}$ for $t \in \mathbb{R}$.

Here $\Phi^{-1}: [-1, 1] \rightarrow [-\infty, +\infty]$ is the reciprocal of Φ and by convention $+\infty - \infty = -\infty$. The inequality becomes an equality when A and B are parallel half-spaces. It recovers and unifies classical results on dilates and enlargements of convex sets:

Corollary 4.2 ([56]). *Let A be a convex set in \mathbb{R}^d , and let $H \subset \mathbb{R}^d$ be a half-space with $\gamma_d(A) = \gamma_d(H)$. Then for all $r \geq 1$,*

$$\gamma_d(rA) \geq \gamma_d(rH).$$

This is reversed when $r \in (0, 1]$.

Corollary 4.3 ([56], [17]). *Let $A \subset \mathbb{R}^d$ be measurable, and $H \subset \mathbb{R}^d$ be a half-space such that $\gamma_d(A) = \gamma_d(H)$. Then for all $r \geq 0$,*

$$\gamma_d(A + rB^d) \geq \gamma_d(H + rB^d).$$

The latter statement is the sharp Gaussian isoperimetric inequality. It implies among others the concentration phenomenon (see e.g. [44]). It asserts that every L -Lipschitz function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is close to its median M with high probability:

$$\gamma_d(|f - M| \geq t) \leq 2e^{-t^2/(2L^2)}.$$

This fact is of fundamental importance in particular in the geometry of Banach spaces [50], [52], where it is often applied to norms. This is one of the motivations for studying improvements of the above results for symmetric convex sets. Little is known in this direction. Latała–Oleszkiewicz [41] established a sharp analogue of Corollary 4.2 for symmetric convex sets. Let us mention another result of similar flavor. Caffarelli's Theorem 2.4 on monotone transportation nicely enters its proof.

Theorem 4.4 (Cordero-Erausquin–Fradelizi–Maurey [29]). *Let $A \subset \mathbb{R}^d$ be an origin-symmetric convex set. Then the function $t \mapsto \log(\gamma_d(tA))$ is concave on $(0, +\infty)$.*

In other words for $\lambda \in (0, 1)$ and $s, t > 0$, $\gamma_d(s^\lambda t^{1-\lambda}A) \geq \gamma_d(sA)^\lambda \gamma_d(tA)^{1-\lambda}$. As $s^\lambda t^{1-\lambda}A \subset (\lambda s + (1 - \lambda)t)A$, this is an improvement on what could be obtained before from the Prékopa–Leindler inequality

$$\gamma_d((\lambda s + (1 - \lambda)t)A) \geq \gamma_d(sA)^\lambda \gamma_d(tA)^{1-\lambda}.$$

This fact and the isoperimetric inequality were used to derive the following improvement of the Gaussian concentration of norm, as conjectured by Vershynin.

Theorem 4.5 (Latała–Oleszkiewicz [42]). *Let G be a standard Gaussian vector in $(\mathbb{R}^d, \|\cdot\|)$. Let M be a median of $\|G\|$ and $\sigma^2 = \sup_{\|f\|_* \leq 1} Ef^2(G)$. Then for all $t \in (0, 1]$ one has*

$$P(\|G\| \leq tM) \leq \frac{1}{2}(2t)^{M^2/(4\sigma^2)}.$$

Remark 4.6. Finally let us point out that Caffarelli’s contraction Theorem 2.4 also played a crucial role in the recent progress towards the Gaussian correlation conjecture which predicts that every two origin symmetric convex sets $A, B \subset \mathbb{R}^d$ satisfy $\gamma_d(A \cap B) \geq \gamma_d(A)\gamma_d(B)$. See [28], [37].

5. Shannon entropy

Let X be a random variable with density $f: \mathbb{R} \rightarrow [0, \infty)$ and, to fix ideas, such that $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Its Shannon entropy is by definition

$$\text{Ent}(X) = - \int_{\mathbb{R}} f \log f.$$

This fundamental notion of information theory also plays a crucial role in the study of return to equilibrium of many random systems. In this section we are interested in entropic aspects of the Central Limit Theorem (CLT). A new approach to entropy estimates was developed, which was formally inspired by a local version of the Brunn–Minkowski theorem.

Among variables of given variance, Gaussian variables are known to maximize entropy. In other words, if G is a standard Gaussian variable with density given by $g(t) = (2\pi)^{-1} \exp(-|t|^2/2)$, $t \in \mathbb{R}$, it holds that

$$\text{Ent}(X) \leq \text{Ent}(G).$$

Moreover, the difference between these two entropies is a strong distance between the laws of X and G . Indeed, the Pinsker–Csiszar–Kullback inequality [51], [33], [39] asserts that it dominates the square of the total variation distance. If Y, Z are

independent random variables and $\lambda \in (0, 1)$, the Shannon–Stam inequality [54], [55] asserts that

$$\lambda \text{Ent}(Y) + (1 - \lambda) \text{Ent}(Z) \leq \text{Ent}(\sqrt{\lambda} Y + \sqrt{1 - \lambda} Z). \quad (5)$$

In particular if $(X_i)_{i \geq 1}$ are independent copies of X one gets that

$$\text{Ent}(X_1) \leq \text{Ent}\left(\frac{X_1 + X_2}{\sqrt{2}}\right),$$

and by iteration that

$$\text{Ent}\left(\frac{1}{\sqrt{2^k}} \sum_1^{2^k} X_i\right)$$

is non-decreasing k (and bounded from above by the entropy of the standard Gaussian variable). Linnik [46] was the first to prove the CLT using entropy. Next Barron [9] established the CLT with entropic convergence. Obtaining rates for the convergence of the entropy requires to improve on the Shannon–Stam inequality (5). Carlen and Soffer [27] obtained non-explicit results in this direction.

In [5] Ball, Naor and the author developed a new technique to estimate entropy production. It is based on a new representation of the Fisher information of a marginal. Recall that the Fisher information of a variable X with density f is defined as $I(X) = I(f) := \int (f')^2 / f$. It corresponds to the derivative of entropy along the Ornstein–Uhlenbeck semigroup: let G be a standard Gaussian variable independent of X ; set $X_t := \sqrt{e^{-t}} X + \sqrt{1 - e^{-t}} G$ and let f_t denote its density. Then

$$\text{Ent}(G) - \text{Ent}(X) = \int_0^\infty (I(X_t) - I(G)) dt.$$

This classical relation allows to integrate linear inequalities on I in order to derive entropic estimates. The Fisher information representation was inspired by the Brunn–Minkowski theorem, as explained in the following section.

Remark 5.1. There was already a nice connection with Brunn–Minkowski theory. Indeed an equivalent form of the Shannon–Stam inequality (5) known as the entropy power inequality asserts that for independent random variables Y, Z one has

$$e^{2\text{Ent}(Y+Z)} \geq e^{2\text{Ent}(Y)} + e^{2\text{Ent}(Z)}.$$

The similarity with the Brunn–Minkowski theorem was noted early and it was supported by the interpretation of Shannon entropy in terms of volumes of typical sets of values of independent copies of a variable. This analogy as well as the occurrence of the number 2 was explained by Szarek and Voiculescu [57], [58], who derived the entropy power inequality from a *restricted* Brunn–Minkowski inequality. Mass transport allows to establish a functional version of the latter, see [11].

5.1. A local version of the Brunn–Minkowski theorem. Consider a probability density $w(x, y)$ on \mathbb{R}^2 together with the density of its first marginal.

$$h(x) = \int w(x, y) dy.$$

Under appropriate regularity and integrability assumptions, the Fisher information of the marginal is expressed in terms of $(\log h)''$:

$$I(h) = \int \frac{(h')^2}{h} = \int h'(\log h)' = - \int h(\log h)''.$$

A direct consequence of the Prékopa–Leindler theorem is that $\log h$ is concave when $\log w$ is (this fact is actually formally equivalent to Brunn–Minkowski for convex sets). In [5] this is explained in terms of second derivatives. The direct calculation is not conclusive and has to be rearranged as

$$\begin{aligned} & h(x)(-\log h)''(x) \\ &= \int_{\mathbb{R}} w(x, y) \left[(\partial_y \rho)^2(x, y) + D^2(-\log w)_{(x,y)} \cdot \begin{pmatrix} 1 \\ \rho(x, y) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \rho(x, y) \end{pmatrix} \right] dy \end{aligned}$$

where ρ is defined by

$$\rho(x, y) = \frac{1}{w(x, y)} \left(\frac{h'(x)}{h(x)} \int_{-\infty}^y w(x, v) dv - \int_{-\infty}^y \partial_x w(x, v) dv \right).$$

One easily reads on the first formula that $D^2(\log w) \leq 0$ implies $(\log h)'' \leq 0$. Actually the function $y \mapsto \rho(x, y)$ described above minimizes the term on the right. Hence we get the following representation

$$\begin{aligned} & h(x)(-\log h)''(x) \\ &= \inf_{p: \mathbb{R} \rightarrow \mathbb{R}} \int_{\mathbb{R}} w(x, y) \left[(p'(y))^2 + D^2(-\log w)_{(x,y)} \cdot \begin{pmatrix} 1 \\ p(y) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ p(y) \end{pmatrix} \right] dy. \end{aligned}$$

Integrating with respect to x yields an expression of the Fisher information of the marginal.

5.2. Variational expressions for the Fisher information. In subsequent papers with Artstein, Ball and Naor [2], [1], more intrinsic formulations are given: let $w: \mathbb{R}^n \rightarrow \mathbb{R}^+$ be a probability density and let $e \in S^{n-1}$ be a unit vector. One considers the marginal obtained by projection onto $\mathbb{R}e$:

$$h(t) = \int_{te+e^\perp} w.$$

Its Fisher information can be expressed as an infimum in the following ways:

$$\begin{aligned} I(h) &= \inf_k \int_{\mathbb{R}^n} w k^2 \\ &= \inf_q \int_{\mathbb{R}^n} w \left(\frac{\operatorname{div}(wq)}{w} \right)^2 \\ &= \inf_q \int_{\mathbb{R}^n} w [\operatorname{Tr}(Dq)^2 + D^2(-\log w) \cdot q \cdot q] \end{aligned}$$

where the first infimum is over functions $k: \mathbb{R}^n \rightarrow \mathbb{R}$ such that for all $t \in \mathbb{R}$ it holds $\int_{te+e^\perp} \partial_e w = \int_{te+e^\perp} w k$. The last two infima range over applications $q: \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that for all $x \in \mathbb{R}^n$ one has $\langle q(x), e \rangle = 1$ (we have omitted here a few regularity conditions).

These formulas are convenient tools to estimate the Fisher information of $(X_1 + \dots + X_k)/\sqrt{k}$ which is a marginal of (X_1, \dots, X_k) . Applications are presented next.

5.3. The monotonicity of entropy in the CLT. The paper [2] answers an old conjecture of Shannon by showing that if $(X_n)_{n \geq 1}$ are independent copies of square integrable a random variable X with finite entropy then the sequence

$$e_k := \operatorname{Ent} \left(\frac{X_1 + \dots + X_k}{\sqrt{k}} \right)$$

is non-decreasing. The classical Shannon–Stam inequality gives $e_k \leq e_{2k}$, but $e_k \leq e_{k+1}$ is much harder. This is deduced from a similar fact for Fisher information, which is proved using the infimum representation: the best test function for k variables is used to build a suitable test functions for $k + 1$. A corresponding version of the entropy-power inequality is also proved

Theorem 5.2 ([2]). *Let X_1, \dots, X_{n+1} be independent square integrable random variables. Then*

$$\exp \left[2 \operatorname{Ent} \left(\sum_{i=1}^{n+1} X_i \right) \right] \geq \frac{1}{n} \sum_{j=1}^{n+1} \exp \left[2 \operatorname{Ent} \left(\sum_{i \neq j} X_i \right) \right].$$

5.4. Rate of convergence in the entropic CLT. In [5], [1] the rate of entropy production, when adding independent copies, is studied under a spectral gap hypothesis. Recall that a random variable X has a spectral gap (or satisfies a Poincaré inequality) if there exists $c > 0$ such that every smooth function $s: \mathbb{R} \rightarrow \mathbb{R}$ verifies

$$c(E(s(X)^2) - (Es(X))^2) \leq E(s'(X)^2). \quad (6)$$

The strategy of proof was to choose specific functions in the variational formula for the Fisher information. Barron and Johnson were able to recover this result by a different method [8].

Theorem 5.3 ([1]). *Let G be a standard Gaussian random variable and let X be a random variable with variance 1. Assume that X satisfies a spectral gap inequality with constant $c > 0$. If X_1, \dots, X_n are independent copies of X , denote as usual $S_n = (X_1 + \dots + X_n)/\sqrt{n}$. Then*

$$\text{Ent}(G) - \text{Ent}(S_n) \leq \frac{1}{1 + \frac{c}{2}(n-1)} (\text{Ent}(G) - \text{Ent}(X)).$$

The rate $1/n$ is best possible. The spectral gap assumption is easy to decide on the real line. It is however rather strong, as it implies in particular exponential integrability. It is natural to try and replace this assumption with weaker moment conditions. A first quantitative result in this direction was obtained by Ball and Cordero-Erausquin:

Theorem 5.4 ([6]). *Let X be a symmetric random variable with $E(X^2) = 1$. Assume that it has finite Fisher information $I(X)$ and third moment $\tau = (E|X|^3)^{1/3}$. Then for $n \geq 1$*

$$\text{Ent}(G) - \text{Ent}(S_n) \leq \frac{c \sqrt{\tau} I(X)^{3/2}}{n^\alpha},$$

where c, α are universal constants.

Remark 5.5. Entropy production in the case of Markov chains with spectral gap was recently understood [7].

References

- [1] Artstein, S., Ball, K., Barthe, F., and Naor, A., On the rate of convergence in the entropic central limit theorem. *Probab. Theory Related Fields* **129** (2004), 381–390.
- [2] Artstein, S., Ball, K., Barthe, F., and Naor, A., Solution of Shannon’s problem on the monotonicity of entropy. *J. Amer. Math. Soc.* **17** (2004), 975–982.
- [3] Bakry, D., L’hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory* (Saint-Flour, 1992), Lecture Notes in Math. 1581, Springer, Berlin 1994, 1–114.
- [4] Ball, K., Convex geometry and functional analysis. In *Handbook of the geometry of Banach spaces* (ed. by W. B. Johnson et al.), Volume 1, Elsevier, Amsterdam 2001, 161–194.
- [5] Ball, K., Barthe, F., and Naor, A., Entropy jumps in the presence of a spectral gap. *Duke Math. J.* **119** (1) (2003), 41–63.
- [6] Ball, K., and Cordero-Erausquin, D., Convergence of information in the Central Limit Theorem under moment conditions. Preprint, 2005.
- [7] Ball, K., Martin-Marquez, V., and Naor, A., Rapid entropy growth for Markov chains with a spectral gap. In preparation, 2005.
- [8] Barron, A., and Johnson, O., Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields* **129** (3) (2004), 391–409.
- [9] Barron, A. R., Entropy and the central limit theorem. *Ann. Probab.* **14** (1986), 336–342.

- [10] Barthe, F., On a reverse form of the Brascamp-Lieb inequality. *Invent. Math.* **134** (1998), 335–361.
- [11] Barthe, F., Restricted Prékopa-Leindler inequality. *Pacific J. Math.* **189** (1999), 211–222.
- [12] Barthe, F., and Cordero-Erausquin, D., Inverse Brascamp-Lieb inequalities along the Heat equation. In *Geometric Aspects of Functional Analysis* (ed. by V. D. Milman and G. Schechtman), Lecture Notes in Math. 1850, Springer-Verlag, Berlin 2004, 65–71.
- [13] Barthe, F., Cordero-Erausquin, D., Ledoux, M., and Maurey, B., Semigroup proofs of Brascamp-Lieb inequalities. In preparation, 2006.
- [14] Barthe, F., Cordero-Erausquin, D., and Maurey, B., Entropy of spherical marginals. *J. Math. Pures Appl.*, to appear.
- [15] Bennett, J., Carbery, A., Christ, M., and Tao, T., The Brascamp-Lieb inequalities: finiteness, structure and extremals. Preprint, 2005.
- [16] Bennett, J., Carbery, A., Christ, M., and Tao, T., Finite bounds for Hölder-Brascamp-Lieb multilinear inequalities. Preprint, 2005.
- [17] Borell, C., The Brunn-Minkowski inequality in Gauss space. *Invent. Math.* **30** (1975), 207–216.
- [18] Borell, C., Diffusion equations and geometric inequalities. *Potential Anal.* **12** (1) (2000), 49–71.
- [19] Borell, C., The Ehrhard inequality. *C. R. Math. Acad. Sci. Paris* **337** (10) (2003), 663–666.
- [20] Borell, C., Minkowski sums in Gaussian analysis. In *Notes de l'école d'hiver "Probabilistic Methods in High Dimension Phenomena"* (Toulouse, 2005); http://www.lsp.ups-tlse.fr/Proba_Winter_School/.
- [21] Brascamp, H. J., and Lieb, E. H., Best constants in Young's inequality, its converse and its generalization to more than three functions. *Adv. Math.* **20** (1976), 151–173.
- [22] Brenier, Y., Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** (1991), 375–417.
- [23] Caffarelli, L., The regularity of mappings with a convex potential. *J. Amer. Math. Soc.* **4** (1992), 99–104.
- [24] Caffarelli, L. A., Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.* **214** (3) (2000), 547–563; Erratum *ibid.* **225** (2) (2002), 449–450.
- [25] Carlen, E. A., Lieb, E. H., and Loss, M., An inequality of Hadamard type for permanents. Preprint.
- [26] Carlen, E. A., Lieb, E. H., and Loss, M., A sharp analog of Young's inequality on S^N and related entropy inequalities. *J. Geom. Anal.* **14** (3) (2004), 487–520.
- [27] Carlen, E. A., and Soffer, A., Entropy production by block variable summation and central limit theorem. *Comm. Math. Phys.* **140** (2) (1991), 339–371.
- [28] Cordero-Erausquin, D., Applications of mass transport to Gaussian-type inequalities. *Arch. Rational Mech. Anal.* **161** (2002), 257–269.
- [29] Cordero-Erausquin, D., Fradelizi, M., and Maurey, B., The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. *J. Funct. Anal.* **214** (2) (2004), 410–427.

- [30] Cordero-Erausquin, D., McCann, R. J., and Schmuckenschläger, M., Prékopa-Leindler type inequalities on riemannian manifolds, mass transport and Jacobi fields. Preprint.
- [31] Cordero-Erausquin, D., McCann, R. J., and Schmuckenschläger, M., A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.* **146** (2) (2001), 219–257.
- [32] Cordero-Erausquin, D., Nazaret, B., and Villani, C., A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. *Adv. Math.* **182** (2) (2004), 307–332.
- [33] Csiszar, I., Informationstheoretische Konvergenzbegriffe im Raum der Wahrscheinlichkeitsverteilungen. *Magyar Tud. Akad. Mat. Kutató Int. Közl. Ser. A* **7** (1962), 137–157.
- [34] Das-Gupta, S., Brunn-Minkowski inequality and its aftermath. *J. Multivariate Anal.* **10** (1980), 296–318.
- [35] Ehrhard, A., Symétrisation dans l’espace de Gauss. *Math. Scand.* **53** (1983), 281–301.
- [36] Gardner, R. J., The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)* **3** (2002), 355–405.
- [37] Hargé, G., A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields* **130** (3) (2004), 415–440.
- [38] Henstock, R., and Macbeath, A. M., On the measure of sum sets. (I) the theorems of Brunn, Minkowski and Lusternik. *Proc. London Math. Soc.* **3** (1953), 182–194.
- [39] Kullback, S., A lower bound for discrimination information in terms of variation. *IEEE Trans. Inform. Theory* **4** (1967), 126–127.
- [40] Latała, R., A note on the Ehrhard inequality. *Studia Math.* **118** (2) (1996), 169–174.
- [41] Latała, R., and Oleszkiewicz, K., Gaussian measures of dilatations of convex symmetric sets. *Ann. Probab.* **27** (4) (1999), 1922–1938.
- [42] Latała, R., and Oleszkiewicz, K., Small ball probability estimates in terms of width. *Studia Math.* **169** (3)(2005), 305–314.
- [43] Ledoux, M., The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math.* (6) **9** (2) (2000), 305–366.
- [44] Ledoux, M., *The concentration of measure phenomenon*. Math. Surveys Monogr. 89, Amer. Math. Soc., Providence, RI, 2001.
- [45] Lieb, E. H., Gaussian kernels have only gaussian maximizers. *Invent. Math.* **102** (1990), 179–208.
- [46] Linnik, Ju. V., An information theoretic proof of the central limit theorem with lindeberg conditions. *Theory Probab. Appl.* **4** (1959), 288–299.
- [47] McCann, R. J., A Convexity Theory for Interacting Gases and Equilibrium Crystals. PhD thesis, Princeton University, 1994.
- [48] McCann, R. J., Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80** (2) (1995), 309–323.
- [49] McCann, R. J., Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11** (3) (2001), 589–608.
- [50] Milman, V., and Schechtman, G., *Asymptotic Theory of Finite Dimensional Normed Spaces*. Lecture Notes in Math. 1200, Springer-Verlag, Berlin 1986.
- [51] Pinsker, M. Š., *Information and information stability of random variables and processes*. Holden-Day, San Francisco, 1964.

- [52] Pisier, G., *The volume of convex bodies and Banach space geometry*. Cambridge Tracts in Math. 94, Cambridge University Press, Cambridge 1989.
- [53] Schneider, R., *Convex bodies: the Brunn-Minkowski theory*. Encyclopedia Math. Appl. 44, Cambridge University Press, Cambridge 1993.
- [54] Shannon, C. E., and Weaver, W., *The mathematical theory of communication*. University of Illinois Press, Urbana, IL, 1949.
- [55] Stam, A. J., Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inform. and Control* **2** (1959), 101–112.
- [56] Sudakov, V. N., and Tsirel'son, B. S., Extremal properties of half-spaces for spherically invariant measures. *J. Soviet Math.* **9** (1978), 9–18; Russian original *Zap. Nauchn. Sem. Leningrad. Otdel. Math. Inst. Steklova.* **41** (1974) 14–24.
- [57] Szarek, S., and Voiculescu, D., Volumes of restricted Minkowski sums and the free analogue of the entropy power inequality. *Comm. Math. Phys.* **178** (3) (1996), 563–570.
- [58] Szarek, S., and Voiculescu, D., Shannon's entropy power inequality via restricted Minkowski sums. In *Geometric aspects of functional analysis*, Lecture Notes in Math. 1745, Springer-Verlag, Berlin 2000, 257–262.
- [59] Villani, C., *Topics in optimal transportation*. Grad. Stud. Math. 58, Amer. Math. Soc., Providence, RI, 2003.

Institut de Mathématiques, Université Paul Sabatier, 31062 Toulouse Cedex 09, France

E-mail: barthe@math.ups-tlse.fr