STUDIA MATHEMATICA 191 (3) (2009)

On Gaussian Brunn–Minkowski inequalities

by

FRANCK BARTHE and NOLWEN HUET (Toulouse)

Abstract. We are interested in Gaussian versions of the classical Brunn–Minkowski inequality. We prove in a streamlined way a semigroup version of the Ehrhard inequality for m Borel or convex sets based on a previous work by Borell. Our method also yields semigroup proofs of the geometric Brascamp–Lieb inequality and of its reverse form, which follow exactly the same lines.

1. Introduction. In this paper, we are interested in Gaussian versions of the classical Brunn–Minkowski inequality on the Lebesgue measure of sum-sets (see e.g. [19, 20]). On $\mathbb{R}^n$ with its canonical Euclidean structure $(\langle\cdot,\cdot\rangle, |\cdot|)$ we consider the standard Gaussian measure $\gamma_n(dx) = (2\pi)^{-n/2}\exp(-|x|^2/2)\,dx$, $x \in \mathbb{R}^n$. Given $\alpha, \beta \in \mathbb{R}$ and sets $A, B \subset \mathbb{R}^n$, we recall that their Minkowski combination is defined by

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

Using symmetrization techniques, Ehrhard [15] proved a sharp lower bound on the Gaussian measure of a convex combination of convex sets: if $\alpha, \beta \geq 0$ satisfy $\alpha + \beta = 1$ and if $A, B \subset \mathbb{R}^n$ are convex, then

$$\Phi^{-1} \circ \gamma_n(\alpha A + \beta B) \geq \alpha \Phi^{-1} \circ \gamma_n(A) + \beta \Phi^{-1} \circ \gamma_n(B),$$

where Φ is the cumulative distribution function of γ_1. This inequality becomes an equality when A and B are parallel half-spaces or the same convex set. Latała [17] showed that the inequality remains valid when A is convex and B is an arbitrary Borel set. In the remarkable paper [9], Borell was able to remove the remaining convexity assumption. He actually derived a functional version of the inequality (in the spirit of the Prékopa–Leindler inequality) by a wonderful interpolation technique based on the heat equation. In a series of papers, Borell extended the inequality to more general combinations:

2000 *Mathematics Subject Classification*: 60E15, 60G15, 52A40, 35K05.
Key words and phrases: Brunn–Minkowski inequality, Gaussian measure, heat equation, Brascamp–Lieb inequalities.

DOI: 10.4064/sm191-3-9

THEOREM (Borell [11]). *Let $\alpha_1, \dots, \alpha_m > 0$. The inequality*

$$\Phi^{-1} \circ \gamma_n\Big(\sum_i \alpha_i A_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ \gamma_n(A_i) \tag{1}$$

holds for all Borel sets $A_1, \dots, A_m$ in $\mathbb{R}^n$ if and only if

$$\sum_i \alpha_i \geq 1 \quad \textit{and} \quad \forall j,\ \alpha_j - \sum_{i \neq j} \alpha_i \leq 1.$$

Moreover, (1) *holds for all convex sets $A_1, \dots, A_m$ in $\mathbb{R}^n$ if and only if*

$$\sum_i \alpha_i \geq 1.$$

Borell established the case $m = 2$ for Borel sets in [10] thanks to his semigroup argument. His proof in [11] of the general case relies on a tricky and somewhat complicated induction. Observe that a linear combination of Borel sets need not be a Borel set; however, it is analytic or Suslin, hence universally measurable (see e.g. [16]).

In this note we give a slight extension of the above statement (the referee pointed out that it can actually be deduced from Borell's theorem, thanks to the Sudakov–Tsirelson inequality $\Phi^{-1} \circ \gamma_n(tA) \geq t\Phi^{-1} \circ \gamma_n(A)$, valid for $t \geq 1$ and A convex. The latter is also a corollary of Borell's general inequality.) More importantly, we propose a streamlined version of the semigroup argument for m functions directly, which allows us to take advantage of convexity type assumptions. This better understanding of the semigroup technique also allows us to study more general situations. The main result is stated next. It involves the heat semigroup, for which we recall the definition: given a Borel non-negative function f on $\mathbb{R}^n$, its evolute at time $t \geq 0$ is the function $P_t f$ given by

$$P_t f(x) = \int f(x + \sqrt{t}\, y)\, \gamma_n(dy) = \mathrm{E}(f(x + B_t))$$

where B is an n-dimensional Brownian motion. By convention $\infty - \infty = -\infty$ so that inequalities like (1), or the one introduced in the next theorem, make sense.

THEOREM 1. *Let $I_{\mathrm{conv}} \subset \{1, \dots, m\}$ and $\alpha_1, \dots, \alpha_m > 0$. The following assertions are equivalent*:

(i) *The parameter $\alpha = (\alpha_1, \dots, \alpha_m)$ satisfies*

$$\sum_i \alpha_i \geq 1 \quad \textit{and} \quad \forall j \notin I_{\mathrm{conv}},\ \alpha_j - \sum_{i \neq j} \alpha_i \leq 1. \tag{2}$$

(ii) *For all Borel sets $A_1, \dots, A_m$ in $\mathbb{R}^n$ such that A_i is convex when $i \in I_{\mathrm{conv}}$,*

$$\Phi^{-1} \circ \gamma\Big(\sum_i \alpha_i A_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ \gamma(A_i).$$

(iii) *For all Borel functions* $h, f_1, \dots, f_m$ *from* $\mathbb{R}^n$ *to* $[0,1]$ *such that* $\Phi^{-1} \circ f_i$ *is concave when* $i \in I_{\mathrm{conv}}$, *if*

$$\forall x_1, \dots, x_m \in \mathbb{R}^n, \quad \Phi^{-1} \circ h\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ f_i(x_i),$$

then

$$\Phi^{-1}\Big(\int h\, d\gamma\Big) \geq \sum_i \alpha_i \Phi^{-1}\Big(\int f_i\, d\gamma\Big).$$

(iv) *For all Borel functions* $h, f_1, \dots, f_m$ *from* $\mathbb{R}^n$ *to* $[0,1]$ *such that* $\Phi^{-1} \circ f_i$ *is concave when* $i \in I_{\mathrm{conv}}$, *if*

$$\forall x_1, \dots, x_m \in \mathbb{R}^n, \quad \Phi^{-1} \circ h\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ f_i(x_i),$$

then for all $t \geq 0$,

$$\forall x_1, \dots, x_m \in \mathbb{R}^n, \quad \Phi^{-1} \circ P_t h\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ P_t f_i(x_i).$$

Remark. Condition (2) can be rephrased as

$$\sum \alpha_i \geq \max(1, \max\{2\alpha_j - 1 : j \notin I_{\mathrm{conv}}\}).$$

Actually, the condition will come up in our argument in the following geometric form: there exist vectors $u_1, \dots, u_m \in \mathbb{R}^m$ such that $|u_i| \leq 1$ for all $i \in I_{\mathrm{conv}}$, $|u_i| = 1$ for all $i \notin I_{\mathrm{conv}}$, and $|\sum \alpha_i u_i| = 1$.

In the next section we show that the condition on α implies the fourth (and formally strongest) assumption in the above theorem, when restricted to smooth enough functions. The third section completes the proof of the theorem. In the final section we discuss related problems.

Before going further, let us introduce some notation.

- We consider functions depending on a time variable t and a space variable x. The time derivative is denoted by ∂_t, while the gradient, Hessian, and Laplacian in x are denoted by ∇_x, Hess_x, and Δ_x; we omit the index x when there is no ambiguity.
- The unit Euclidean (closed) ball and sphere of $\mathbb{R}^d$ are denoted respectively by $\mathbb{B}^d$ and $\mathbb{S}^{d-1}$.
- For $A \subset \mathbb{R}^d$, we set $A^\varepsilon = A + \varepsilon\mathbb{B}^d$. The notation A_i^ε means $(A_i)^\varepsilon$.
- The transposed matrix of the matrix A is denoted by A^*.

2. Functional and semigroup approach. As already mentioned, we follow Borell's semigroup approach to the Gaussian Brunn–Minkowski inequalities (see [9] and [10]): for parameters α satisfying (2), the plan is to show the functional version of the inequality (Theorem 1(iii)) by means of the heat semigroup. Note that (iv) implies (iii) by choosing $t = 1$ and $x_i = 0$ in the last inequality of (iv). So our aim is to establish (iv). More precisely, given Borel functions $h, f_1, \dots, f_m$ from $\mathbb{R}^n$ to $(0,1)$, we define C on $[0,T]\times(\mathbb{R}^n)^m$ by

$$C(t,x) = C(t,x_1,\dots,x_m) = \Phi^{-1}\circ P_t h\Big(\sum \alpha_i x_i\Big) - \sum_i \alpha_i \Phi^{-1}\circ P_t f_i(x_i).$$

Since $P_0 f = f$ the assumption

$$\forall x_i \in \mathbb{R}^n, \quad \Phi^{-1}\circ h\Big(\sum \alpha_i x_i\Big) \ge \sum_i \alpha_i \Phi^{-1}\circ f_i(x_i) \tag{3}$$

translates as $C(0,\cdot)\ge 0$. Our task is to prove

$$C(0,\cdot)\ge 0 \ \Rightarrow\ \forall t\ge 0,\, C(t,\cdot)\ge 0.$$

2.1. *Preliminaries.* When the functions h and f_i are smooth enough, the time evolution of $P_t h$ and $P_t f_i$ is described by the heat equation. This yields a differential equation satisfied by C. Our problem boils down to determining whether this evolution equation preserves non-negative functions. This is clearly related to the maximum principle for parabolic equations (see e.g. [13]). We will use the following lemma.

LEMMA 1. *Assume that C is twice differentiable. If*

$$\begin{cases} \mathrm{Hess}(C)\ge 0 \\ \nabla C = 0 \\ C\le 0 \end{cases} \quad \Rightarrow\ \partial_t C \ge 0 \tag{4}$$

and if for some $T>0$,

$$\liminf_{|x|\to\infty}\Big(\inf_{0\le t\le T} C(x,t)\Big) \ge 0, \tag{5}$$

then

$$C(0,\cdot)\ge 0 \ \Rightarrow\ \forall t\in[0,T],\, C(t,\cdot)\ge 0.$$

Proof. For $\varepsilon>0$, set $C_\varepsilon(t,x) = C(t,x)+\varepsilon t$ on $[0,T]\times(\mathbb{R}^n)^m$. If $C_\varepsilon<0$ at some point, then C_ε reaches its minimum at a point (t_0,x_0) where $\nabla C=0$, $\mathrm{Hess}(C)\ge 0$, $C<0$, and $\partial_t C+\varepsilon\le 0$ ($=0$ if $t_0<T$). By the hypotheses, this implies $\partial_t C\ge 0$, which contradicts $\partial_t C\le -\varepsilon$. So for all $\varepsilon>0$ and $T>0$, C_ε is non-negative on $[0,T]\times(\mathbb{R}^n)^m$, thus C is non-negative everywhere. ■

Property (5) is true under mild assumptions on h and f_i which are related to the initial condition $C(0,\cdot)\ge 0$ in the large:

LEMMA 2. *If there exist $a_1,\dots,a_m\in\mathbb{R}$ such that*

- $\limsup_{|x|\to\infty} f_i(x) \le \Phi(a_i)$,
- $h \ge \Phi(\sum_i \alpha_i a_i)$,

then for all $T > 0$,

$$\liminf_{|x|\to\infty}\Big(\inf_{0\le t\le T} C(x,t)\Big) \ge 0.$$

Proof. Let $\delta > 0$. By continuity of Φ^{-1}, there exists $\varepsilon > 0$ such that

$$\Phi^{-1}(\Phi(a_i)+2\varepsilon) \le a_i + \delta/\sum \alpha_j.$$

Let $r > 0$ be such that $\gamma_n(r\mathbb{B}^n) = 1-\varepsilon$. Then, for $0 \le t \le T$,

$$\begin{aligned} P_t f_i(x_i) &= \int\limits_{r\mathbb{B}^n} f_i(x_i + \sqrt{t}\,y)\,\gamma_n(dy) + \int\limits_{(r\mathbb{B}^n)^{\mathrm{c}}} f_i(x_i+\sqrt{t}\,y)\,\gamma_n(dy) \\ &\le (1-\varepsilon)\sup_{x_i + r\sqrt{t}\,\mathbb{B}^n} f_i + \varepsilon \sup f_i \\ &\le \sup_{x_i + r\sqrt{T}\,\mathbb{B}^n} f_i + \varepsilon \le \Phi(a_i) + 2\varepsilon \quad \text{ for } |x_i| \text{ large enough.} \end{aligned}$$

Moreover, $P_t h \ge \Phi(\sum \alpha_i a_i)$ so for $|x|$ large enough and for $0 \le t \le T$, we have $C(t,x) \ge -\delta$. As $\delta > 0$ was arbitrary, the proof is complete. ■

Checking property (4) of Lemma 1 requires the following lemma:

LEMMA 3. *Let* $d \ge 2$ *and* $\alpha_1,\dots,\alpha_m > 0$. *Let* k *be an integer with* $0 \le k \le m$ *and*

$$\varphi : (\mathbb{S}^{d-1})^k \times (\mathbb{B}^d)^{m-k} \to \mathbb{R}_+, \qquad (v_1,\dots,v_m) \mapsto \Big|\sum_i \alpha_i v_i\Big|.$$

Then the image of φ *is the interval*

$$J := \Big[\max\Big(\{0\} \cup \Big\{\alpha_j - \sum_{i\ne j}\alpha_i : 1 \le j \le k\Big\}\Big), \sum_i \alpha_i\Big].$$

Proof. As φ is continuous on a compact connected set, we have $\mathrm{Im}(\varphi) = [\min\varphi, \max\varphi]$. Plainly $|\sum \alpha_i v_i| \le \sum \alpha_i$, with equality if $v_1 = \dots = v_m$ is a unit vector. So $\max\varphi = \sum_i \alpha_i$. For all $j \le k$, since $|v_j| = 1$, the triangle inequality gives

$$\Big|\sum_i \alpha_i v_i\Big| \ge \alpha_j|v_j| - \sum_{i\ne j}\alpha_i|v_i| \ge \alpha_j - \sum_{i\ne j}\alpha_i.$$

Hence $\mathrm{Im}(\varphi) \subset J$ and these two intervals have the same upper bound. Next we deal with the lower bound. Let us consider a point $(v_1,\dots,v_m)$ where φ achieves its minimum, and differentiate:

For $j \le k$, v_j lies in the unit sphere. Applying the Lagrange multipliers theorem to φ^2 with respect to v_j gives a real number λ_j such that

$$\alpha_j \sum_i \alpha_i v_i = \lambda_j v_j. \tag{6}$$

For $j > k$, the jth variable lives in $\mathbb{B}^d$. If $|v_j| < 1$ the minimum is achieved at an interior point and the full gradient of φ^2 with respect to the jth variable is zero. Hence $\sum_i \alpha_i v_i = 0$. On the other hand, if at the minimum point, $|v_j| = 1$, differentiating in the jth variable only along the unit sphere gives again the existence of $\lambda_j \in \mathbb{R}$ such that (6) is satisfied.

Eventually, we face two cases:

Case 1: $\sum \alpha_i v_i = 0$ and $\min \varphi = 0$. In this case, the triangle inequality gives $0 = |\sum \alpha_i v_i| \geq \alpha_j - \sum_{i \neq j} \alpha_i$ whenever $j \leq k$.

Case 2: the v_i's are collinear unit vectors and there exists a partition $S_+ \cup S_- = \{1, \ldots, m\}$ and a unit vector v such that

$$\min \varphi = \Big|\sum_{S_+} \alpha_i v - \sum_{S_-} \alpha_i v\Big| = \sum_{S_+} \alpha_i - \sum_{S_-} \alpha_i > 0.$$

Assume that S_+ contains two indices j and ℓ. Let e_1 and e_2 be two orthonormal vectors of $\mathbb{R}^d$ and denote by $R(\theta)$ the rotation of angle θ in the plane $\mathrm{Vect}(e_1, e_2)$. The length of the vector $\alpha_j R(\theta) e_1 + \alpha_\ell e_1$ is a decreasing and continuous function of $\theta \in [0, \pi]$. Denote by $U(\theta)$ the rotation in $\mathrm{Vect}(e_1, e_2)$ which maps this vector to $|\alpha_j R(\theta) e_1 + \alpha_\ell e_1| e_1$. Then

$$\alpha_j U(\theta) R(\theta) e_1 + \alpha_\ell U(\theta) e_1 + \sum_{S_+ \setminus \{j,\ell\}} \alpha_i e_1 - \sum_{S_-} \alpha_i e_1 = \lambda(\theta) e_1,$$

where $\lambda(0) = \sum_{S_+} \alpha_i - \sum_{S_-} \alpha_i = \min \varphi > 0$ and λ is continuous and decreasing in $\theta \in [0, \pi]$. This contradicts the minimality of $\min \varphi$. So S_+ contains a single index j and

$$\min \varphi = \Big|\alpha_j v - \sum_{i \neq j} \alpha_i v\Big| = \alpha_j - \sum_{i \neq j} \alpha_i > 0.$$

Note that necessarily $j \leq k$, otherwise one could get a shorter vector by replacing $v_j = v$ by $(1 - \varepsilon) v$. Moreover, the condition $\alpha_j - \sum_{i \neq j} \alpha_i > 0$ ensures that $\alpha_j > \alpha_\ell$ for $\ell \neq j$. This implies that for $\ell \neq j$,

$$\alpha_\ell - \sum_{i \neq \ell} \alpha_i \leq \alpha_\ell - \alpha_j < 0 < \alpha_j - \sum_{i \neq j} \alpha_i.$$

So $\min \varphi = \max(\{0\} \cup \{\alpha_j - \sum_{i \neq j} \alpha_i : 1 \leq j \leq k\})$ as claimed. ∎

2.2. *Semigroup proof for smooth functions.* We deal with smooth functions first, in order to ensure that $P_t f_i$ and $P_t h$ satisfy the heat equation. This restrictive assumption will be removed in Section 3 where the proof of Theorem 1 is completed.

Theorem 2. *Let f_i, $i = 1, \ldots, m$, and h be twice continuously differentiable functions from $\mathbb{R}^n$ to $(0,1)$ satisfying the hypotheses of Lemma 2. Assume moreover that for $f = f_i$ or h,*

$$\forall t > 0, \forall x \in \mathbb{R}^n, \qquad |\nabla f(x + \sqrt{t}\, y)| e^{-|y|^2/2} \xrightarrow[|y| \to \infty]{} 0.$$

Let $\alpha_1, \dots, \alpha_m$ be positive real numbers such that

$$\sum_i \alpha_i \geq 1 \quad \text{and} \quad \forall j,\ \alpha_j - \sum_{i \neq j} \alpha_i \leq 1.$$

If

$$\forall x_i \in \mathbb{R}^n, \qquad \Phi^{-1} \circ h\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ f_i(x_i),$$

then

$$\forall t \geq 0,\ \forall x_i \in \mathbb{R}^n, \quad \Phi^{-1} \circ P_t h\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \alpha_i \Phi^{-1} \circ P_t f_i(x_i).$$

Proof. Let us recall that C is defined by

$$C(t,x) = C(t, x_1, \dots, x_m) = H\Big(t, \sum \alpha_i x_i\Big) - \sum \alpha_i F_i(t, x_i)$$

where we have set

$$H(t,y) = \Phi^{-1} \circ P_t h(y) \quad \text{and} \quad F_i(t,y) = \Phi^{-1} \circ P_t f_i(y).$$

In what follows, we omit the variables and write H for $H\big(t, \sum \alpha_i x_i\big)$ and F_i instead of $F_i(t, x_i)$. With this simplified notation,

$$\begin{aligned} C &= H - \sum \alpha_i F_i, \\ \nabla_{x_i} C &= \alpha_i (\nabla H - \nabla F_i), \\ \nabla_{x_i} \nabla^*_{x_j} C &= \alpha_i \alpha_j \operatorname{Hess}(H) - \delta_{ij} \alpha_i \operatorname{Hess}(F_i). \end{aligned}$$

Moreover, one can use the property of the heat kernel to derive a differential equation for F_i and H. Indeed, for any f satisfying the hypotheses of the theorem, we can perform an integration by parts to obtain

$$\partial_t P_t f = \tfrac{1}{2} \Delta P_t f.$$

Then we set $F = \Phi^{-1} \circ P_t f$ and use the identity $(1/\Phi'(x))' = x/\Phi'(x)$ to show

$$\begin{aligned} \partial_t F &= \frac{\partial_t P_t f}{\Phi'(F)} = \frac{\Delta P_t f}{2\,\Phi'(F)}, \\ \nabla F &= \frac{\nabla P_t f}{\Phi'(F)}, \\ \Delta F &= \frac{\Delta P_t f}{\Phi'(F)} + F\,\frac{|\nabla P_t f|^2}{(\Phi'(F))^2}. \end{aligned}$$

We put all this together to get

$$\partial_t F = \tfrac{1}{2}(\Delta F - F|\nabla F|^2)$$

and to deduce the following differential equation for C:

$$\partial_t C = \tfrac{1}{2}(\mathcal{S} + \mathcal{P})$$

where the second order part is

$$\mathcal{S} = \Delta H - \sum \alpha_i \Delta F_i$$

and the terms of lower order are

$$\mathcal{P} = -\Big(H|\nabla H|^2 - \sum \alpha_i F_i |\nabla F_i|^2\Big).$$

We will conclude the proof using Lemma 1. So we need to check condition (4). First we note that $\mathcal{P}$ is non-negative when $\nabla C = 0$ and $C \leq 0$, regardless of α. Indeed, $\nabla C = 0$ implies that $\nabla F_i = \nabla H$ for all i. So $\mathcal{P} = -|\nabla H|^2 C$, which is non-negative if $C \leq 0$.

It remains to deal with the second order part. It is enough to express $\mathcal{S}$ as $\mathcal{E}C$ for some elliptic operator $\mathcal{E}$, since then $\mathrm{Hess}(C) \geq 0$ implies $\mathcal{S} \geq 0$. Such a second order operator can be written as $\mathcal{E} = \nabla^* A \nabla$ where A is a symmetric $nm \times nm$ matrix. Moreover, $\mathcal{E}$ is elliptic if and only if A is positive semidefinite. In view of the structure of the problem, it is natural to look for matrices of the following block form:

$$A = B \otimes I_n = (b_{ij} I_n)_{1 \leq i,j \leq m}\,,$$

where I_n is the identity $n \times n$ matrix and B is a positive semidefinite matrix of size m. Writing $x_i = (x_{i,1}, \ldots, x_{i,n})$, we have

$$\mathcal{E}C = \sum_{i,j=1}^{m} b_{i,j} \left(\sum_{k=1}^{n} \frac{\partial^2}{\partial x_{i,k} \partial x_{j,k}} C \right) = \sum_{i,j=1}^{m} b_{i,j} \big(\alpha_i \alpha_j \Delta H - \delta_{i,j} \alpha_i \Delta F_i \big)$$

$$= \langle \alpha, B\alpha \rangle \Delta H - \sum_{i=1}^{m} b_{i,i} \alpha_i \Delta F_i.$$

Hence there exists an elliptic operator $\mathcal{E}$ of the above form such that $\mathcal{E}C = \mathcal{S} = \Delta H - \sum_{i=1}^{m} \alpha_i \Delta F_i$ if there exists a positive semidefinite matrix B of size m such that

$$\langle \alpha, B\alpha \rangle = \langle e_1, Be_1 \rangle = \cdots = \langle e_m, Be_m \rangle = 1$$

where $(e_i)_i$ is the canonical basis of $\mathbb{R}^m$. Now a positive semidefinite matrix B can be decomposed into $B = V^* V$ where V is a square matrix of size m. Letting $v_1, \ldots, v_m \in \mathbb{R}^m$ be the columns of V, we can translate the latter into conditions on the vectors v_i. Actually, we are looking for vectors $v_1, \ldots, v_m \in \mathbb{R}^m$ with

$$|v_1| = \cdots = |v_m| = \Big| \sum \alpha_i v_i \Big| = 1.$$

By Lemma 3 for $k = m$, this is possible exactly when α satisfies the claimed condition:

$$\sum \alpha_i \geq 1 \quad \text{and} \quad \forall j,\ \alpha_j - \sum_{i \neq j} \alpha_i \leq 1. \ \blacksquare$$

The following corollary will be useful in the next section.

Corollary 1. *Let f be a function on $\mathbb{R}^n$ taking values in $(0,1)$ and vanishing at infinity, i.e.* $\lim_{|x|\to\infty} f(x) = 0$. *Assume also that*

$$\forall t>0,\, \forall x\in\mathbb{R}^n, \quad |\nabla f(x+\sqrt{t}\,y)|e^{-|y|^2/2} \xrightarrow[|y|\to\infty]{} 0.$$

If $\Phi^{-1}\circ f$ is concave, then $\Phi^{-1}\circ P_t f$ is concave for all $t\ge 0$.

Proof. Let $0<\varepsilon<1$ and $\alpha_i>0$ with $\sum\alpha_i=1$. Choosing $h=\varepsilon+(1-\varepsilon)f\ge f$ and $f_i=f$ for $i\ge 1$, one can check that the latter theorem applies. Hence for all $t\ge 0$ and $x_i\in\mathbb{R}^n$,

$$\Phi^{-1}\circ P_t(\varepsilon+(1-\varepsilon)f)\Big(\sum\alpha_i x_i\Big) \ge \sum\alpha_i\Phi^{-1}\circ P_t f(x_i).$$

Letting ε go to 0, we conclude by monotone convergence that $\Phi^{-1}\circ P_t f$ is concave. ■

2.3. *Φ^{-1}-concave functions.* When some of the f_i's are Φ^{-1}-concave, the conditions on the parameters can be relaxed. Such functions allow one to approximate characteristic functions of convex sets, as we will see in Section 3.

Theorem 3. *Let $I_{\mathrm{conv}}\subset\{1,\dots,m\}$. Let f_i, $i=1,\dots,m$, and h be twice continuously differentiable functions from $\mathbb{R}^n$ to $(0,1)$ satisfying the hypotheses of Lemma 2. Assume also that for $f=f_i$ or h,*

$$\forall t>0,\, \forall x\in\mathbb{R}^n, \quad |\nabla f(x+\sqrt{t}\,y)|e^{-|y|^2/2} \xrightarrow[|y|\to\infty]{} 0.$$

Assume moreover that $\Phi^{-1}\circ f_i$ is concave and decreasing to $-\infty$ at infinity for all $i\in I_{\mathrm{conv}}$.

Let $\alpha_1,\dots,\alpha_m$ be positive numbers satisfying

$$\sum_i\alpha_i\ge 1 \quad\textit{and}\quad \forall j\notin I_{\mathrm{conv}},\ \alpha_j-\sum_{i\ne j}\alpha_i\le 1.$$

If

$$\forall x_i\in\mathbb{R}^n, \quad \Phi^{-1}\circ h\Big(\sum_i\alpha_i x_i\Big)\ge\sum_i\alpha_i\Phi^{-1}\circ f_i(x_i),$$

then

$$\forall t\ge 0,\, \forall x_i\in\mathbb{R}^n, \quad \Phi^{-1}\circ P_t h\Big(\sum_i\alpha_i x_i\Big)\ge\sum_i\alpha_i\Phi^{-1}\circ P_t f_i(x_i).$$

Proof. As in the proof of Theorem 2, we try to apply Lemma 1 to the equation satisfied by C:

$$\partial_t C(t,x)=\tfrac{1}{2}(\mathcal{S}+\mathcal{P}).$$

We have already shown that $\mathcal{P}$ is non-negative when $\nabla C=0$ and $C\le 0$, for any $\alpha_1,\dots,\alpha_m$. We would like to prove that the conditions on α in the theorem imply that $\mathcal{S}$ is non-negative whenever $\mathrm{Hess}(C)\ge 0$.

By Corollary 1, for all $i \in I_{\text{conv}}$ the function F_i is concave, hence $\Delta F_i \le 0$. So we are done if we can write

$$\mathcal{S} = \mathcal{E}C - \sum_{i \in I_{\text{conv}}} \lambda_i \Delta F_i$$

for some elliptic operator $\mathcal{E}$ and some $\lambda_i \ge 0$. As in the proof of the previous theorem, we are looking for operators of the form $\mathcal{E} = \nabla^* A \nabla$ with $A = B \otimes I_n = (b_{ij} I_n)_{1 \le i,j \le m}$ where B is a symmetric positive semidefinite $m \times m$ matrix. Hence our task is to find $B \ge 0$ and $\lambda_i \ge 0$ such that $\lambda_i = 0$ when $i \notin I_{\text{conv}}$ and

$$\Delta H - \sum \alpha_i \Delta F_i = \langle \alpha, B\alpha \rangle \Delta H - \sum (b_{ii}\alpha_i + \lambda_i)\Delta F_i.$$

When $i \in I_{\text{conv}}$, we can find $\lambda_i \ge 0$ such that $b_{ii}\alpha_i + \lambda_i = \alpha_i$ whenever $b_{ii} \le 1$. Consequently, the problem reduces to finding a positive semidefinite $m \times m$ matrix B such that

$$\begin{cases} \langle e_i, Be_i \rangle \le 1, & \forall i \in I_{\text{conv}}, \\ \langle e_i, Be_i \rangle = 1, & \forall i \notin I_{\text{conv}}, \\ \langle \alpha, B\alpha \rangle = 1, \end{cases}$$

where $(e_i)_i$ is the canonical basis of $\mathbb{R}^m$. Equivalently, do there exist $v_1, \ldots, v_m \in \mathbb{R}^m$ such that

$$\begin{cases} |v_i| \le 1, & \forall i \in I_{\text{conv}}, \\ |v_i| = 1, & \forall i \notin I_{\text{conv}}, \\ |\sum \alpha_i v_i| = 1? \end{cases}$$

We conclude with Lemma 3. ∎

3. Back to sets. This section explains how to complete the proof of Theorem 1. The main issue is to get rid of the smoothness assumptions made so far. A key point is that the conditions on α do not depend on the dimension n. The plan of the argument is summed up in the next figure, where we are referring to the assertions of Theorem 1, and "(iv) for smooth functions" means assertion (iv) (of Theorem 1) restricted to functions satisfying all the assumptions of the first paragraph of Theorem 3:

$$\text{(i)} \;\Rightarrow\; \text{(iv)}\ \textit{for smooth functions} \;\Rightarrow\; \text{(ii)} \;\Rightarrow\; \text{(iv)} \;\Rightarrow\; \text{(iii)} \;\Rightarrow\; \text{(i)}.$$

Note that "(iv)⇒(iii)" is obvious, whereas "(i) ⇒ (iv) for smooth functions" was established in Theorem 3. Next we prove the remaining implications.

"(iv) for smooth functions ⇒ (ii)": Fix an arbitrary α. Let $A_1, \ldots, A_m$ be Borel sets in $\mathbb{R}^n$ with A_i convex when $i \in I_{\text{conv}}$. By inner regularity of the measure, we can assume that they are compact. Let $\varepsilon > 0$ and $b > a$ be fixed. Then:

- For $i \notin I_{\mathrm{conv}}$: there exists a smooth function f_i such that $f_i = \Phi(b)$ on A_i, $f_i = \Phi(a)$ off A_i^ε, and $0 < \Phi(a) \le f_i \le \Phi(b) < 1$.
- For $i \in I_{\mathrm{conv}}$: there exists a smooth function f_i such that $F_i = \Phi^{-1} \circ f_i$ is concave, $F_i = b$ on A_i, $F_i \le a$ off A_i^ε, and $F_i \le b$ on $\mathbb{R}^n$.

  For instance, take a point x_i in A_i and define the gauge of $A_i^{\varepsilon/3}$ with respect to x_i by
$$\varrho(x) = \inf\left\{\lambda > 0 : x_i + \frac{1}{\lambda}(x - x_i) \in A_i^{\varepsilon/3}\right\}.$$
  We know that ϱ is convex since A_i is convex (see for instance [20]). Then set
$$\tilde{F}_i(x) = b + c(1 - \max(\varrho(x), 1))$$
  where $c > 0$ is chosen large enough to ensure that $\tilde{F}_i \le a$ off $A_i^{2\varepsilon/3}$. Now, we can take a smooth function g with compact support small enough and of integral 1, such that $f_i = \Phi(\widetilde{F_i} * g)$ is a smooth Φ^{-1}-concave function satisfying the required conditions.
- For h: set
$$a_0 = \max_{\substack{u_i = a \text{ or } b \\ u \neq (b,\dots,b)}} \sum \alpha_i u_i \quad \text{and} \quad b_0 = \sum \alpha_i b.$$
  Again, we can choose a smooth function h such that $h = \Phi(b_0)$ on $\sum \alpha_i A_i^\varepsilon$, $h = \Phi(a_0)$ off $(\sum \alpha_i A_i^\varepsilon)^\varepsilon$, and $0 < \Phi(a_0) \le h \le \Phi(b_0) < 1$.

From these definitions, the functions h and f_i are "smooth" and satisfy
$$\forall x_i \in \mathbb{R}^n, \quad \Phi^{-1} \circ h\Big(\sum \alpha_i x_i\Big) \ge \sum \alpha_i \Phi^{-1} \circ f_i(x_i).$$
By our hypothesis, the inequality remains valid with $P_t h$ and $P_t f_i$ for all $t > 0$. Choosing $t = 1$ and $x_i = 0$ yields
$$\Phi^{-1}\Big(\int h\, d\gamma_n\Big) \ge \sum \alpha_i \Phi^{-1}\Big(\int f_i\, d\gamma_n\Big).$$
By construction, the functions f_i and h actually depend on the parameters a, b, ε (but we decided not to use the heavy notation $h^{(a,b,\varepsilon)}$). Fixing $\varepsilon > 0$ and b, it is easily seen that
$$\liminf_{a \to -\infty} f_i \ge \Phi(b)\mathbb{1}_{A_i} \quad \text{and} \quad \limsup_{a \to -\infty} h \le \Phi(b_0)\mathbb{1}_{(\sum \alpha_i A_i^\varepsilon)^\varepsilon},$$
where $\mathbb{1}_{A_i}$ is the characteristic function of A_i. Hence, using Fatou's lemma twice, we get
$$\Phi^{-1}\Big(\Phi(b_0)\gamma_n\Big(\Big(\sum \alpha_i A_i^\varepsilon\Big)^\varepsilon\Big)\Big) \ge \sum_i \alpha_i \Phi^{-1}(\Phi(b)\gamma_n(A_i)).$$
Next we let b, and consequently b_0, go to $+\infty$. Finally, the compactness of the sets easily yields $\bigcap_{k \ge 1}(\sum \alpha_i A_i^{1/k})^{1/k} = \sum \alpha_i A_i$. Therefore we get, as

expected,

$$\Phi^{-1} \circ \gamma_n\Big(\sum \alpha_i A_i\Big) \geq \sum \alpha_i \Phi^{-1} \circ \gamma_n(A_i).$$

"(ii) (in $\mathbb{R}^{n+1}$) $\Rightarrow$ (iv) (in $\mathbb{R}^n$)". Here we assume that (ii) of Theorem 1 is valid for all Borel sets in $\mathbb{R}^{n+1}$ and we derive (iv) for functions defined on $\mathbb{R}^n$.

For any Borel function f on $\mathbb{R}^n$ taking values in $[0,1]$, $t > 0$, and $x \in \mathbb{R}^n$, we define

$$B_f^{t,x} = \{(u,y) : u \leq \Phi^{-1} \circ f(x + \sqrt{t}\, y)\} \subset \mathbb{R} \times \mathbb{R}^n.$$

Then

$$\gamma_{n+1}\big(B_f^{t,x}\big) = P_t f(x).$$

Let $h, f_1, \ldots, f_n$ be Borel functions on $\mathbb{R}^n$ with values in $[0,1]$ such that $\Phi^{-1} \circ f_i$ is concave when $i \in I_{\mathrm{conv}}$. Assume that

$$\forall x_i \in \mathbb{R}^n, \quad \Phi^{-1} \circ h\Big(\sum \alpha_i x_i\Big) \geq \sum \alpha_i \Phi^{-1} \circ f_i(x_i).$$

Then for (u_i, y_i) in $B_{f_i}^{t,x_i}$, we get

$$\sum \alpha_i u_i \leq \sum \alpha_i \Phi^{-1} \circ f_i(x_i + \sqrt{t}\, y_i) \leq \Phi^{-1} \circ h\Big(\sum \alpha_i (x_i + \sqrt{t}\, y_i)\Big),$$

which means that

$$\sum \alpha_i B_{f_i}^{t,x_i} \subset B_h^{t,\sum \alpha_i x_i}.$$

The same argument shows that $B_f^{t,x}$ is convex if $\Phi^{-1} \circ f$ is concave. Thus, the result for sets in $\mathbb{R}^{n+1}$ implies that

$$\Phi^{-1} \circ P_t h\Big(\sum \alpha_i x_i\Big) \geq \Phi^{-1} \circ \gamma_{n+1}\Big(\sum \alpha_i B_{f_i}^{t,x_i}\Big) \geq \sum \alpha_i \Phi^{-1} \circ P_t f_i(x_i).$$

"(iii)$\Rightarrow$(i)": We will prove the contraposed assertion: if the conditions on α_i are violated, then there exist Borel functions h and f_i such that $\Phi^{-1} \circ f_i$ is concave for $i \in I_{\mathrm{conv}}$ and the relation $\Phi^{-1} \circ h(\sum \alpha_i x_i) \geq \sum \Phi^{-1} \circ f_i(x_i)$ holds for all x_i, but

$$\Phi^{-1}\Big(\int h\, d\gamma\Big) < \sum \alpha_i \Phi^{-1}\Big(\int f_i\, d\gamma\Big).$$

Let $f : \mathbb{R}^n \to (0,1)$ be an even Borel function such that

$$f(0) > \frac{1}{2}, \quad \int f\, d\gamma < \frac{1}{2}, \quad \text{and} \quad F = \Phi^{-1} \circ f \text{ is concave.}$$

For instance, we may take $f(x) = \Phi(1 - |ax|^2)$ for a large enough. Note that for $0 \leq t \leq 1$,

$$F(tx) \geq tF(x) + (1-t)F(0) \geq tF(x). \tag{7}$$

Assume first that $\sum \alpha_i < 1$. Then by concavity and the latter bound, we get, for all x_i,

$$\Phi^{-1} \circ f\Big(\sum_i \alpha_i x_i\Big) = F\Big(\sum_i \alpha_i x_i\Big) \geq \sum_i \frac{\alpha_i}{\sum_j \alpha_j} F\Big(\Big(\sum_j \alpha_j\Big) x_i\Big)$$
$$\geq \sum_i \alpha_i F(x_i) = \sum_i \alpha_i \Phi^{-1} \circ f(x_i).$$

However, since $1 > \sum \alpha_i$ and $\Phi^{-1}(\int f\, d\gamma) < 0$, we have

$$\Phi^{-1}\Big(\int f\, d\gamma\Big) < \sum_i \alpha_i \Phi^{-1}\Big(\int f\, d\gamma\Big).$$

Assume now that there exists $j \notin I_{\mathrm{conv}}$ such that $\alpha_j - \sum_{i \neq j} \alpha_i > 1$. Then using (7) and concavity again, we obtain, for all x_i,

$$\alpha_j F(x_j) \geq \Big(1 + \sum_{i \neq j} \alpha_i\Big) F\bigg(\frac{\alpha_j x_j}{1 + \sum_{i \neq j} \alpha_i}\bigg)$$
$$\geq F\Big(\alpha_j x_j - \sum_{i \neq j} \alpha_i x_i\Big) + \sum_{i \neq j} \alpha_i F(x_i).$$

Let $g = 1 - f$. Since $-F = -\Phi^{-1} \circ f = \Phi^{-1} \circ (1 - f) = \Phi^{-1} \circ g$ and f is even we may rewrite the last inequality as

$$\Phi^{-1} \circ g\Big(\alpha_j x_j + \sum_{i \neq j} \alpha_i(-x_i)\Big) \geq \alpha_j \Phi^{-1} \circ g(x_j) + \sum_{i \neq j} \alpha_i \Phi^{-1} \circ f(-x_i).$$

However, since $\Phi^{-1}(\int g\, d\gamma) = -\Phi^{-1}(\int f\, d\gamma) > 0$ and $\alpha_j - \sum_{i \neq j} \alpha_i > 1$, we also have

$$\Phi^{-1}\Big(\int g\, d\gamma\Big) < \alpha_j \Phi^{-1}\Big(\int g\, d\gamma\Big) + \sum_{i \neq j} \alpha_i \Phi^{-1}\Big(\int f\, d\gamma\Big).$$

Therefore the proof is complete.

4. Further remarks

4.1. *Brascamp–Lieb type inequalities.* In [7, 8], Borell already used his semigroup approach to derive variants of the Prékopa–Leindler inequality. The latter is a functional counterpart to the Brunn–Minkowski inequality for the Lebesgue measure and reads as follows: if $\lambda \in (0, 1)$ and $f, g, h : \mathbb{R}^n \to \mathbb{R}^+$ are Borel functions such that for all $x, y \in \mathbb{R}^n$,

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

then $\int h \geq (\int f)^\lambda (\int g)^{1-\lambda}$ where the integrals are with respect to Lebesgue's measure. Borell actually showed the following stronger fact: for all $t > 0$ and all $x, y \in \mathbb{R}^n$,

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

Setting $H(t,\cdot) = \log P_t h$ and defining F, G similarly, it is proved that $C(t,x,y) := H(t,\lambda x+(1-\lambda)y) - \lambda F(t,x) + (1-\lambda)G(t,y)$ satisfies a positivity-preserving evolution equation. The argument is simpler than for Ehrhard's inequality since the evolution equation for individual functions is simpler: $2\partial_t H = \Delta H + |\nabla H|^2$.

The Brascamp–Lieb inequality [12, 18] is a powerful extension of Hölder's inequality. The so-called reverse Brascamp–Lieb inequality, first proved in [2, 3], appears as an extension of the Prékopa–Leindler inequality. In [4], it was noted that Borell's semigroup method could be used to derive the geometric reverse Brascamp–Lieb inequality for functions of one variable. This observation was also motivated by a proof of the Brascamp–Lieb inequalities based on semigroup techniques (Carlen, Lieb and Loss [14] for functions of one variable, and Bennett, Carbery, Christ and Tao [6] for general functions).

In this subsection, we take advantage of our streamlined presentation of Borell's method, and quickly reprove the reverse Brascamp–Lieb inequality in geometric form, but for functions of several variables. More surprisingly, we will recover the geometric Brascamp–Lieb inequality from inequalities which are preserved by the heat flow. The result is not new (the inequality for the law of the semigroup appears in the preprint [5]), but it is interesting to have semigroup proofs of the direct and reverse inequalities which follow exactly the same lines. Recall that the transportation argument of [3] provided the direct and reverse inequalities simultaneously.

The setting of the geometric inequalities is as follows: for $i = 1,\dots,m$ let $c_i > 0$ and let $B_i : \mathbb{R}^N \to \mathbb{R}^{n_i}$ be linear maps such that $B_i B_i^* = I_{n_i}$ and

$$\sum_{i=1}^m c_i B_i^* B_i = I_N. \tag{8}$$

These hypotheses were put forward by Ball in connection with volume estimates in convex geometry [1]. Note that B_i^* is an isometric embedding of $\mathbb{R}^{n_i}$ into $\mathbb{R}^N$ and that $B_i^* B_i$ is the orthogonal projection from $\mathbb{R}^N$ to $E_i = \mathrm{Im}(B_i^*)$. The Brascamp–Lieb inequality asserts that for all Borel functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^+$,

$$\int_{\mathbb{R}^N} \prod_{i=1}^m f_i(B_i x)^{c_i}\, dx \le \prod_{i=1}^m \Big(\int_{\mathbb{R}^{n_i}} f_i \Big)^{c_i}.$$

The reverse inequality ensures that

$$\int_{\mathbb{R}^N}^* \sup\Big\{ \prod_{i=1}^m f_i(x_i)^{c_i} : x_i \in \mathbb{R}^{n_i},\ \sum_{i=1}^m c_i B_i^* x_i = x \Big\}\, dx \ge \prod_{i=1}^m \Big(\int_{\mathbb{R}^{n_i}} f_i \Big)^{c_i},$$

where $\int^*$ is the outer integral.

Following [4], we will deduce the latter from the following result.

Theorem 4. *If $h : \mathbb{R}^N \to \mathbb{R}^+$ and $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^+$ satisfy*

$$\forall x_i \in \mathbb{R}^{n_i}, \quad h\Big(\sum_{i=1}^m c_i B_i^* x_i\Big) \geq \prod_{i=1}^m f_i(x_i)^{c_i}$$

then

$$\forall x_i \in \mathbb{R}^{n_i}, \quad P_t h\Big(\sum_{i=1}^m c_i B_i^* x_i\Big) \geq \prod_{i=1}^m P_t f_i(x_i)^{c_i}.$$

The reverse inequality is obtained as $t \to +\infty$ since for f on $\mathbb{R}^d$, $P_t f(x)$ is equivalent to $(2\pi t)^{-d/2} \int_{\mathbb{R}^d} f$. To see this, note that

$$P_t f(x) = (2\pi t)^{-d/2} \int_{\mathbb{R}^d} f(y) \exp\Big(\frac{|x-y|^2}{2t}\Big)\, dy.$$

Note also that taking traces in the decomposition of the identity map yields $\sum_i c_i n_i = N$.

In order to recover the Brascamp–Lieb inequality, we will show the following theorem.

Theorem 5. *If $h : \mathbb{R}^N \to \mathbb{R}^+$ and $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^+$ satisfy*

$$\forall x \in \mathbb{R}^N, \quad h(x) \leq \prod_{i=1}^m f_i(B_i x)^{c_i},$$

then

$$\forall x \in \mathbb{R}^N, \quad P_t h(x) \leq \prod_{i=1}^m P_t f_i(B_i x)^{c_i}.$$

Again, the limit as $t \to +\infty$ yields the Brascamp–Lieb inequality when choosing $h(x) = \prod_{i=1}^m f_i(B_i x)^{c_i}$. We sketch the proofs of the above two statements, omitting the truncation arguments needed to ensure (5).

Proof of Theorem 4. Set $H(t,\cdot) = \log P_t h(\cdot)$ and $F_i(t,\cdot) = \log P_t f_i(\cdot)$. As said above, the functions H and F_i satisfy the equation $2\partial_t U = \Delta U + |\nabla U|^2$. For $(t, x_1, \ldots, x_m) \in \mathbb{R}^+ \times \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$, set

$$C(t, x_1, \ldots, x_m) := H\Big(t, \sum_{i=1}^m c_i B_i^* x_i\Big) - \sum_{i=1}^m c_i F_i(t, x_i).$$

By hypothesis $C(0,\cdot) \geq 0$ and we want to prove that $C(t,\cdot)$ is non-negative as well. As before, we are done if we can show that the three conditions $C \leq 0$, $\nabla C = 0$, and $\mathrm{Hess}(C) \geq 0$ imply that $\partial_t C \geq 0$. Actually, the condition $C \leq 0$ is not used in the following argument. Omitting variables, we have

$$2\partial_t C = \Big(\Delta H - \sum_i c_i \Delta F_i\Big) + \Big(|\nabla H|^2 - \sum_i c_i |\nabla F_i|^2\Big) =: \mathcal{S} + \mathcal{P}.$$

Straightforward calculations give

$$\nabla_{x_i} C = c_i B_i \nabla H - c_i \nabla F_i,$$
$$\mathrm{Hess}_{x_i,x_j}(C) = c_i c_j B_i \,\mathrm{Hess}(H) B_j^* - \delta_{i,j} c_i \,\mathrm{Hess}(F_i).$$

Note that the decomposition (8) implies that for all $v \in \mathbb{R}^N$,

$$|v|^2 = \Big\langle v, \sum c_i B_i^* B_i v \Big\rangle = \sum c_i |B_i v|^2.$$

Hence, if $\nabla C = 0$, the above calculation gives $\nabla F_i = B_i \nabla H$. Consequently $|\nabla H|^2 = \sum c_i |B_i \nabla H|^2 = \sum c_i |\nabla F_i|^2$. So $\nabla C = 0 \Rightarrow \mathcal{P} = 0$.

Next, we deal with the second order term. Using (8) again we obtain

$$\begin{aligned}\Delta H &= \mathrm{Tr}(\mathrm{Hess}(H)) = \mathrm{Tr}\Big(\Big(\sum_i c_i B_i^* B_i\Big) \mathrm{Hess}(H) \Big(\sum_j c_j B_j^* B_j\Big)\Big)\\ &= \sum_{i,j} \mathrm{Tr}(B_i^* (c_i c_j B_i \,\mathrm{Hess}(H) B_j^*) B_j).\end{aligned}$$

Also note that

$$\begin{aligned}\sum_{i,j} \mathrm{Tr}(B_i^* (\delta_{i,j} c_i \,\mathrm{Hess}(F_i)) B_j) &= \sum_i \mathrm{Tr}(B_i^* c_i \,\mathrm{Hess}(F_i) B_i)\\ &= \sum_i c_i \,\mathrm{Tr}(\mathrm{Hess}(F_i) B_i B_i^*) = \sum_i c_i \Delta F_i,\end{aligned}$$

since $B_i B_i^* = I_{n_i}$. Combining the former and the latter and denoting by J_i the canonical embedding of $\mathbb{R}^{n_i}$ into $\mathbb{R}^{n_1+\cdots+n_m}$ we see that

$$\begin{aligned}\mathcal{S} &= \Delta H - \sum_i c_i \Delta F_i = \sum_{i,j} \mathrm{Tr}(B_i^* \,\mathrm{Hess}_{x_i,x_j}(C) B_j)\\ &= \sum_{i,j} \mathrm{Tr}(B_i^* (J_i^* \,\mathrm{Hess}(C) J_j) B_j) = \mathrm{Tr}\Big(\Big(\sum_i J_i B_i\Big)^* \mathrm{Hess}(C) \Big(\sum_j J_j B_j\Big)\Big)\end{aligned}$$

is non-negative when $\mathrm{Hess}(C) \geq 0$. This is enough to conclude that C remains non-negative. ∎

Proof of Theorem 5. As before, we set $H(t,\cdot) = \log P_t h(\cdot)$ and $F_i(t,\cdot) = \log P_t f_i(\cdot)$. For $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^N$,

$$C(t,x) := \sum_{i=1}^m c_i F_i(t, B_i x) - H(t,x).$$

Omitting variables, C evolves according to the equation

$$\partial_t C = \Big(\sum_i c_i \Delta F_i - \Delta H\Big) + \Big(\sum_i c_i |\nabla F_i|^2 - |\nabla H|^2\Big) =: \mathcal{S} + \mathcal{P}.$$

Next,

$$\nabla C = \sum_i c_i B_i^* \nabla F_i - \nabla H, \qquad \mathrm{Hess}(C) = \sum_i c_i B_i^* \,\mathrm{Hess}(F_i) B_i - \mathrm{Hess}(H).$$

Taking traces in the latter equality and since $B_iB_i^* = I_{n_i}$ we obtain

$$\Delta C = \sum_i c_i \operatorname{Tr}(\operatorname{Hess}(F_i)B_iB_i^*) - \Delta H = \sum_i c_i \Delta F_i - \Delta H = \mathcal{S}.$$

Therefore the second order term is clearly elliptic.

It remains to check that $\nabla C = 0$ implies that the first order term $\mathcal{P}$ is non-negative. We will need the following easy consequence of the decomposition (8): if $x_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, m$, then

$$\Big|\sum c_i B_i^* x_i\Big|^2 \le \sum c_i |x_i|^2.$$

The proof is easy: set $v = \sum c_i B_i^* x_i$. Then by Cauchy–Schwarz,

$$\begin{aligned}|v|^2 &= \Big\langle v, \sum c_i B_i^* x_i\Big\rangle = \sum c_i \langle B_i v, x_i\rangle \\ &\le \Big(\sum c_i |B_i v|^2\Big)^{1/2}\Big(\sum c_i |x_i|^2\Big)^{1/2}.\end{aligned}$$

But (8) ensures that $|v|^2 = \sum c_i |B_i v|^2$ so after simplification we get the claim. Finally, note that $\nabla C = 0$ means that $\nabla H = \sum c_i B_i^* \nabla F_i$. Hence $|\nabla H|^2 \le \sum c_i |\nabla F_i|^2$. In other words, $\mathcal{P} \ge 0$. The proof is therefore complete. ∎

Remark. As explained in [6], general (non-geometric) Brascamp–Lieb inequalities can be derived from the geometric form, via change of variables and twisted products. The same methods apply to the reverse inequalities.

4.2. *Looking for Gaussian Brascamp–Lieb inequalities.* It is natural to ask about Gaussian versions of the Brascamp–Lieb or inverse Brascamp–Lieb inequalities. For $0 \le i \le m$, take a non-zero real d_i, a positive integer $n_i \le N$, a linear surjective map $L_i : \mathbb{R}^N \to \mathbb{R}^{n_i}$, and a Borel function f_i on $\mathbb{R}^{n_i}$ taking values in $(0, 1)$. Does the inequality

$$\forall x \in \mathbb{R}^N, \qquad \sum_{i=0}^m d_i \Phi^{-1} \circ f_i(L_i x) \ge 0$$

upgrade for all $t \ge 0$ to

$$\forall x \in \mathbb{R}^N, \qquad \sum_{i=0}^m d_i \Phi^{-1} \circ P_t f_i(L_i x) \ge 0\ ?$$

This general formulation allows negative d_i's and would encompass Gaussian extensions of Theorem 4 or Theorem 5. It also enables a better understanding of the essential properties in the semigroup argument. Note that from now on, the index i goes from 0 to m, the function $f_0 =: h$ playing *a priori* no particular role any more.

As before, for $t \geq 0$ and $x \in \mathbb{R}^N$ we define

$$C(t,x) = \sum_{i=0}^{m} d_i \Phi^{-1} \circ P_t f_i(L_i x) = \sum_{i=0}^{m} d_i F_i(t, L_i x).$$

We are interested in the conditions on the data $(d_i, L_i)_{i=0}^m$ for which $C(0,\cdot) \geq 0$ implies $C(t,\cdot) \geq 0$ for all $t \geq 0$. Assume that our functions are smooth enough for the calculations to follow. We have

$$\begin{aligned} C &= \sum d_i F_i, \\ \nabla C &= \sum d_i L_i^* \nabla F_i, \\ \mathrm{Hess}(C) &= \sum d_i L_i^* \,\mathrm{Hess}(F_i) L_i, \end{aligned}$$

and thanks to the heat equation, C satisfies the differential equation $2\partial_t C = \mathcal{S} + \mathcal{P}$ where

$$\mathcal{S} = \sum d_i \Delta F_i \quad \text{and} \quad \mathcal{P} = -\sum d_i |\nabla F_i|^2 F_i.$$

We require that

$$\begin{cases} \mathrm{Hess}(C) \geq 0 \\ \nabla C = 0 \\ C \leq 0 \end{cases} \quad \Rightarrow \; \mathcal{P} + \mathcal{S} \geq 0$$

in order to apply Lemma 1 (the condition at infinity is satisfied, provided one restricts to good enough functions f_i; we omit the details). In other words, if $(d_i, L_i)_{i=0}^m$ satisfy

$$\inf_{\substack{\sum d_i F_i \leq 0 \\ \sum d_i L_i^* \nabla F_i = 0 \\ \sum d_i L_i^* \,\mathrm{Hess}(F_i) L_i \geq 0}} \Big(-\sum d_i |\nabla F_i|^2 F_i + \sum d_i \,\mathrm{Tr}(\mathrm{Hess}(F_i)) \Big) \geq 0,$$

then $C(0,\cdot) \geq 0$ implies $C(t,\cdot) \geq 0$ for all t. Let us look for workable conditions on $(d_i, L_i)_{i=0}^m$ which ensure that the latter infimum is non-negative.

Given i and (t, x_i), observe that for any $X_i \in \mathbb{R}$, $Y_i \in \mathbb{R}^{n_i}$, and any $n_i \times n_i$ symmetric matrix Z_i, there exists a function f_i such that $F_i(t, x_i) = X_i$, $\nabla F_i(t, x_i) = Y_i$, and $\mathrm{Hess}(F_i)(t, x_i) = Z_i$. Indeed, since

$$\begin{cases} F_i = \Phi^{-1}(P_t f_i), \\ \nabla F_i = \dfrac{\nabla P_t f_i}{\Phi'(F_i)}, \\ \mathrm{Hess}(F_i) = \dfrac{1}{\Phi'(F_i)} \left(\mathrm{Hess}(P_t f_i) - \Phi''(F_i) \nabla F_i \cdot \nabla F_i^* \right), \end{cases}$$

prescribing the values of $(F_i, \nabla F_i, \mathrm{Hess}(F_i))$ boils down to prescribing the values of $(P_t f_i, \nabla P_t f_i, \mathrm{Hess}(P_t f_i))$. Next, one can find a suitable polynomial

f_i of degree 2 satisfying these equations. At this stage, we do not see how the initial condition $C(0,\cdot) = \sum d_i \Phi^{-1}(f_i) \geq 0$ creates relationships between the values of $(F_i, \nabla F_i, \mathrm{Hess}(F_i))$ for different values of i. So it is natural to bound the infimum from below as follows:

$$\inf_{\substack{\sum d_i F_i \leq 0 \\ \sum d_i L_i^* \nabla F_i = 0 \\ \sum d_i L_i^* \mathrm{Hess}(F_i) L_i \geq 0}} \Big(-\sum d_i |\nabla F_i|^2 F_i + \sum d_i \operatorname{Tr}(\mathrm{Hess}(F_i)) \Big)$$
$$\geq \inf_{\substack{\sum d_i X_i \leq 0 \\ \sum d_i L_i^* Y_i = 0}} -\sum d_i |Y_i|^2 X_i + \inf_{\sum d_i L_i^* Z_i L_i \geq 0} \sum d_i \operatorname{Tr}(Z_i),$$

where X_i, Y_i, and Z_i run respectively over $\mathbb{R}$, $\mathbb{R}^{n_i}$, and the set of $n_i \times n_i$ symmetric real matrices. The last quantity is non-negative if and only if there exists a real number K such that

$$\inf_{\substack{\sum d_i X_i \leq 0 \\ \sum d_i L_i^* Y_i = 0}} -\sum d_i |Y_i|^2 X_i \geq -K,$$
$$\inf_{\sum d_i L_i^* Z_i L_i \geq 0} \sum d_i \operatorname{Tr}(Z_i) \geq K.$$

Taking $Y_i = 0$ and $Z_i = 0$ for all i shows that K must be equal to 0. Hence our new problem is to find sufficient conditions for (d_i, L_i) to satisfy

$$\begin{aligned} &\inf_{\substack{\sum d_i X_i \leq 0 \\ \sum d_i L_i^* Y_i = 0}} -\sum d_i |Y_i|^2 X_i \geq 0, \\ &\inf_{\sum d_i L_i^* Z_i L_i \geq 0} \sum d_i \operatorname{Tr}(Z_i) \geq 0. \end{aligned} \tag{9}$$

We remark that this question is related to the formally stronger initial requirement: $(C \leq 0,\ \nabla C = 0) \Rightarrow \mathcal{P} \geq 0$ and $\mathrm{Hess}(C) \geq 0 \Rightarrow \mathcal{S} \geq 0$.

Let us deal with the first inequality in (9). It can be rephrased as

$$\begin{cases} \sum d_i X_i \leq 0 \\ \sum d_i L_i^* Y_i = 0 \end{cases} \Rightarrow \sum d_i |Y_i|^2 X_i \leq 0.$$

Reasoning for fixed Y_i's, and viewing the conditions on X_i as equations of half-spaces, we easily see that the latter implication is equivalent to

$$\sum d_i L_i^* Y_i = 0 \ \Rightarrow\ |Y_0|^2_{\mathbb{R}^{n_0}} = \cdots = |Y_m|^2_{\mathbb{R}^{n_m}}. \tag{10}$$

This condition can be worked out a bit more. Let us assume (10) and define $\mathcal{L} : \mathbb{R}^{\sum n_j} \to \mathbb{R}^N$ by

$$\mathcal{L}(Y_0, \ldots, Y_m) = \sum d_i L_i^* Y_i.$$

If $a = (a_0, \ldots, a_m)$ and $b = (b_0, \ldots, b_m)$ belong to $\ker \mathcal{L}$ then $|a_i|^2$, $|b_i|^2$, and by linearity $|a_i+b_i|^2$ are independent of i. Expanding the square of the sum, we deduce that $\langle a_i, b_i\rangle$ is independent of i and therefore equal to the average over i of these quantities. Hence for all i, $(m+1)\langle a_i, b_i\rangle = \langle a, b\rangle$. This means that $u_i : \ker \mathcal{L} \to \mathbb{R}^{n_i}$ defined by $u_i(a) = \sqrt{m+1}\, a_i$ is an isometry. Since $a_i = u_i(u_0^{-1}(a_0))$, we conclude that

$$\ker \mathcal{L} = \{(a_0, u_1(u_0^{-1}(a_0)), \dots, u_m(u_0^{-1}(a_0))) : a_0 \in \mathrm{Im}(u_0)\}.$$

It is then clear that (10) is equivalent to the following: there exists a subspace $X \subset \mathbb{R}^{n_0}$ and linear isometries $R_i : X \to \mathbb{R}^{n_i}$, $i \geq 1$, such that

$$\ker \mathcal{L} = \{(x, R_1 x, \dots, R_m x) : x \in X\}. \tag{11}$$

In order to establish the second inequality in (9), it is sufficient to find a symmetric positive semidefinite $N \times N$ matrix A such that $\sum d_i \operatorname{Tr}(Z_i)$ can be expressed as

$$\operatorname{Tr}\Big(A \sum d_i L_i^* Z_i L_i\Big) = \sum d_i \operatorname{Tr}(L_i A L_i^* Z_i).$$

As we require this identity for arbitrary matrices Z_i, we can conclude that A does the job if and only if for all $0 \leq i \leq m$,

$$L_i A L_i^* = I_{n_i}.$$

We may look for A in the form $A = \sigma^*\sigma$ for some $N \times N$ matrix σ. For $0 \leq i \leq m$ and $1 \leq j \leq n_i$, denote by $u_i^j \in \mathbb{R}^N$ the columns of L_i^*. Rewriting the above conditions in terms of σ we may conclude that the second infimum in (9) is non-negative provided there exists an $N \times N$ matrix σ such that for all $0 \leq i \leq m$ the vectors $(\sigma u_i^j)_{j=1}^{n_i}$ form an orthonormal system in $\mathbb{R}^N$. Note that the first-order condition (11) requires that the linear relations between the vectors u_i^j should have a particular structure.

We have been able to find data (d_i, L_i) satisfying the above conditions, but all of them could be reduced to the Borell theorem, using the rotation invariance of the Gaussian measure and the fact that its marginals remain Gaussian. To conclude this section let us briefly explain why the method does not allow any new Gaussian improvement of Theorems 4 or 5.

For $i = 1, \dots, m$, let $c_i > 0$ and $B_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ be linear surjective maps. If we look for Gaussian versions of the Brascamp–Lieb inequality, we are led to apply the previous reasoning to $N = n$, $B_0 = I_N$, $d_0 = -1$, and for $i \geq 1$, $L_i = B_i$ and $d_i = c_i$. Now, with the above notation, $(Y_0, \dots, Y_m) \in \ker \mathcal{L}$ is equivalent to $Y_0 = \sum_{i=1}^m c_i B_i^* Y_i$. Since this condition can be satisfied even though $|Y_1| \neq |Y_2|$ we conclude that the first-order condition (11) is never satisfied.

Next, we are looking for inequalities of the reverse Brascamp–Lieb type. Hence we choose $N = n_1 + \dots + n_m$, $d_0 = 1$, $L_0(x_1, \dots, x_m) = \sum c_i B_i^* x_i$, and for $i \geq 1$, $d_i = -c_i$ and $L_i(x_1, \dots, x_m) = x_i$. For $x \in \mathbb{R}^n$, $L_0^*(x) = (c_1 B_1 x, \dots, c_m B_m x)$. For $i \geq 1$ and $x_i \in \mathbb{R}^{n_i}$, $L_i^*(x_i) = (0, \dots, 0, x_i, 0, \dots, 0)$ where x_i appears at the ith place. The condition $(Y_0, \dots, Y_m) \in \ker \mathcal{L}$, that is, $L_0^*(Y_0) = \sum_{i \geq 1} c_i L_i^*(Y_i)$, becomes

$$\forall i = 1, \dots, m, \qquad Y_i = B_i Y_0.$$

Hence $\ker \mathcal{L} = \{(Y_0, B_1 Y_0, \dots, B_m Y_0) : Y_0 \in \mathbb{R}^n\}$. So (11) holds only if the B_i's are isometries. This forces $n_i = n$ and up to an isometric change of variables, we are back in the setting of the Gaussian Brunn–Minkowski inequality.

References

[1] K. M. Ball, *Volumes of sections of cubes and related problems*, in: J. Lindenstrauss and V. D. Milman (eds.), Israel Seminar on Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1376, Springer, 1989, 251–260.

[2] F. Barthe, *Inégalités de Brascamp–Lieb et convexité*, C. R. Acad. Sci. Paris Sér. I Math. 324 (1997), 885–888.

[3] —, *On a reverse form of the Brascamp–Lieb inequality*, Invent. Math. 134 (1998), 335–361.

[4] F. Barthe and D. Cordero-Erausquin, *Inverse Brascamp–Lieb inequalities along the heat equation*, in: V. D. Milman and G. Schechtman (eds.), Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1850, Springer, 2004, 65–71.

[5] F. Barthe, D. Cordero-Erausquin, M. Ledoux, and B. Maurey, *Semigroup proofs of Brascamp–Lieb inequalities*, preprint, 2006.

[6] J. Bennett, A. Carbery, M. Christ, and T. Tao, *The Brascamp–Lieb inequalities: finiteness, structure and extremals*, Geom. Funct. Anal. 17 (2008), 1343–1415.

[7] C. Borell, *Geometric properties of some familiar diffusions in* $\mathbb{R}^n$, Ann. Probab. 21 (1993), 482–489.

[8] —, *Diffusion equations and geometric inequalities*, Potential Anal. 12 (2000), 49–71.

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

[10] —, *Minkowski sums and Brownian exit times*, Ann. Fac. Sci. Toulouse Math. (6) 16 (2007), 37–47.

[11] —, *Inequalities of the Brunn–Minkowski type for gaussian measures*, Probab. Theory Related Fields 140 (2008), 195–205.

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

[13] H. Brézis, *Analyse fonctionnelle, théorie et applications*, Masson, 1993.

[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] A. Ehrhard, *Symétrisation dans l'espace de Gauss*, Math. Scand. 53 (1983), 281–301.

[16] H. Federer, *Geometric Measure Theory*, Springer, New York, 1969.

[17] R. Latała, *A note on the Ehrhard inequality*, Studia Math. 118 (1996), 169–174.

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

[19] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge Tracts in Math. 94, Cambridge Univ. Press, Cambridge, 1989.
[20] R. Schneider, *Convex Bodies: the Brunn–Minkowski Theory*, Encyclopedia Math. Appl. 44, Cambridge Univ. Press, Cambridge, 1993.

Institut de Mathématiques de Toulouse
Université Paul Sabatier
31062 Toulouse, France
E-mail: barthe@math.univ-toulouse.fr
huet@math.univ-toulouse.fr

Received April 22, 2008
Revised version September 22, 2008 (6341)