STUDIA MATHEMATICA 238 (3) (2017)

Some extensions of the Prékopa–Leindler inequality using Borell's stochastic approach

by

Dario Cordero-Erausquin and Bernard Maurey (Paris)

Abstract. We present an abstract form of the Prékopa–Leindler inequality that includes several known—and a few new—related functional inequalities on Euclidean spaces. The method of proof and also the formulation of the new inequalities are based on Christer Borell's stochastic approach to Brunn–Minkowski type inequalities.

1. Introduction and main statement. The Brunn–Minkowski inequality asserts that for Borel subsets A, B of $\mathbb{R}^n$ and $t \in [0,1]$, the volume of the Minkowski combination

$$(1-t)A + tB = \{(1-t)a + tb : a \in A,\, b \in B\}$$

satisfies the inequality $|(1-t)A + tB| \geq |A|^{1-t}|B|^t$, where $|E|$ denotes the volume of a Lebesgue-measurable subset E of $\mathbb{R}^n$. There is a long story of functional generalizations of this inequality, which we do not recall here; let us just mention that Borell's 1975 paper [5] remains a milestone in the subject. A somewhat definitive form is given by the following theorem.

Theorem 1.1 (Prékopa–Leindler inequality). *Let $t \in [0,1]$ and let $f_0, f_1, g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be Borel functions such that, for every $x_0, x_1 \in \mathbb{R}^n$,*

$$g((1-t)x_0 + tx_1) \leq (1-t)f_0(x_0) + tf_1(x_1).$$

Then

$$\int_{\mathbb{R}^n} e^{-g(x)}\,dx \geq \Big(\int_{\mathbb{R}^n} e^{-f_0(x)}\,dx\Big)^{1-t}\Big(\int_{\mathbb{R}^n} e^{-f_1(x)}\,dx\Big)^{t}.$$

Allowing the value ∞ enables us to reach directly indicator functions $\mathbf{1}_E = e^{-f_E}$, by letting f_E be 0 on E and ∞ outside. However, we can restrict

2010 *Mathematics Subject Classification*: Primary 39B62; Secondary 52A40, 60H30, 52A20.
Key words and phrases: Brunn–Minkowski inequality, Brascamp–Lieb inequalities, geometric functional inequalities, Gaussian stochastic calculus.
Received 19 January 2016; revised 21 July 2016.
Published online 28 April 2017.

ourselves to Borel functions that are not Lebesgue-almost everywhere equal to ∞. Such Borel functions f will be said to be *L-proper*, which is equivalent to saying that $\int_{\mathbb{R}^n} e^{-f(x)}\,dx > 0$.

One can reasonably argue that the interest of the Prékopa–Leindler inequality resides not only in its consequences, which are numerous (some are recalled for instance in [8, 11]), but also in the emphasis it has put on log-concavity, and in the related techniques of proof it has originated, such as mass transportation or semigroup techniques, and more recently L^2-methods (as in [7]).

Here, we will concentrate on Borell's stochastic approach [6] to the inequality above. It somehow reduces the inequalities under study to the convexity of $|\cdot|^2$, the square of the Euclidean norm on $\mathbb{R}^n$. It will allow us to obtain some unexpected inequalities, for instance that of the following proposition.

PROPOSITION 1.2. *Let f_0, f_1, g_0, g_1 be Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{\infty\}$ such that, for every $x_0, x_1 \in \mathbb{R}^n$,*

$$g_0(2x_0/3 + x_1/3) + g_1(x_0/3 + 2x_1/3) \le f_0(x_0) + f_1(x_1). \tag{1.1}$$

Then

$$\Big(\int_{\mathbb{R}^n} e^{-g_0(x)}\,dx\Big)\Big(\int_{\mathbb{R}^n} e^{-g_1(x)}\,dx\Big) \ge \Big(\int_{\mathbb{R}^n} e^{-f_0(x)}\,dx\Big)\Big(\int_{\mathbb{R}^n} e^{-f_1(x)}\,dx\Big).$$

We will see that it is rather natural to arrive at this type of inequality using Borell's stochastic approach, whereas it seems not to be the case with other methods, for instance those based on transportation methods. The point here is to split the values of f_i's at some points into the values of two g_j's at some related points. This is not interesting in the case where the f_i's take only the values ∞ and 0, and the functional inequality above does not give anything new when applied to the case where the functions e^{-f_i} are indicators of sets, as we will explain in Section 4 below.

The previous proposition and its proof actually suggest more general inequalities. Writing the conclusion as

$$\sum_j -\log\Big(\int_{\mathbb{R}^n} e^{-g_j}\Big) \le \sum_i -\log\Big(\int_{\mathbb{R}^n} e^{-f_i}\Big),$$

we may think about an extension of the results, where the finite families of functions f_i, g_j are replaced by families f_s, g_t depending upon continuous parameters s, t (as for instance in [2]). In the "basic assumption" (1.1), the two values x_0, x_1 on the right-hand side of the inequality will be replaced, for example, by a selection $x = \{x(s)\}_{s\in[0,1]}$ of points of $\mathbb{R}^n$, and the values $y = \{y(t)\}_{t\in[0,1]}$ on the left-hand side will be obtained from a linear transformation A acting on this data x, i.e. $y = Ax$. So, roughly speaking, under appropriate assumptions on A, we may expect that if for "all" $x = \{x(s)\}$ we have

$$\int_0^1 g_t\big((Ax)(t)\big)\,dt \le \int_0^1 f_s(x(s))\,ds$$

then it will follow that

$$\int_0^1 -\log\Big(\int_{\mathbb{R}^n} e^{-g_t}\Big)\,dt \le \int_0^1 -\log\Big(\int_{\mathbb{R}^n} e^{-f_s}\Big)\,ds.$$

Actually, there was no reason, when replacing the sums by integrals, to use the "uniform" distributions dt and ds rather than probability measures $\mu(dt)$ and $\nu(ds)$ on $[0,1]$, which include the discrete case when these measures are convex combinations of Dirac measures. Anyway, the main question is to understand what are the appropriate conditions to impose on the linear operator A.

Several points must be set before proceeding. We say that a real function F on a measure space (Ω,Σ,μ) is *μ-semi-integrable* if $F^+ = \max(F,0)$ or $F^- = \max(-F,0)$ is μ-integrable. The integral of F then takes a definite value in $[-\infty,\infty]$. This assumption is needed for $F(s) = -\log(\int_{\mathbb{R}^n} e^{-f_s})$ in order to make sense of the preceding integrals.

We introduce the following abstract setting.

Setting 1. We are given

- Two measure spaces $X_1 = (\Omega_1,\Sigma_1,\mu_1)$ and $X_2 = (\Omega_2,\Sigma_2,\mu_2)$, where Σ_i is a σ-algebra of subsets of Ω_i and μ_i is a *finite* measure, $i = 1,2$. We assume that Ω_1 is a Polish topological space, and Σ_1 is its Borel σ-algebra.
- An integer $n \ge 1$ and a continuous linear operator
$$A : L^2(X_1,\mathbb{R}^n) \to L^2(X_2,\mathbb{R}^n),$$
where the L^2-norms of the $\mathbb{R}^n$-valued functions are computed with respect to the Euclidean norm $|\cdot|$ on $\mathbb{R}^n$ and to the measures μ_1 and μ_2, respectively. We assume that
  - (i) the operator A satisfies the norm condition $\|A\| \le 1$,
  - (ii) the operator A acts as the identity operator on constant vector valued functions, i.e., for any $v_0 \in \mathbb{R}^n$, the constant function $\Omega_1 \ni s \mapsto v_0$ is sent by A to the constant function $\Omega_2 \ni t \mapsto v_0$.
- Two families $\{f_s\}_{s\in\Omega_1}$ and $\{g_t\}_{t\in\Omega_2}$ of Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{\infty\}$ such that
  - (iii) the functions $(s,x) \mapsto f_s(x)$ and $(t,x) \mapsto g_t(x)$ are measurable with respect to the σ-algebras $\Sigma_i \otimes \mathcal{B}_{\mathbb{R}^n}$, $i = 1,2$, respectively, where $\mathcal{B}_{\mathbb{R}^n}$ is the Borel σ-algebra of $\mathbb{R}^n$,

(iv) the functions

$$\Omega_1 \ni s \mapsto \log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\,dx\Big) \quad\text{and}\quad \Omega_2 \ni t \mapsto \log\Big(\int_{\mathbb{R}^n} e^{-g_t(x)}\,dx\Big)$$

are semi-integrable with respect to μ_1, μ_2, respectively.

Unfortunately, there is a tradeoff between the generality of the statement we can reach and the technical assumptions that are required in the proof. We will start with a fairly general situation. In most applications, the technical assumptions on the functions in the statement below are either easy to impose or to discard, as explained for instance in Remark 1.4 and as shown in Theorem 1.6 below. So, our most abstract version of the Prékopa–Leindler inequality reads as follows.

THEOREM 1.3. *Under Setting 1 with μ_1 and μ_2 having the same finite mass, $\mu_1(\Omega_1) = \mu_2(\Omega_2) < \infty$, assume additionally that*

- *for every $s \in \Omega_1$, the function f_s is non-negative, and for every $t \in \Omega_2$, the function g_t is non-negative and lower semicontinuous,*
- *for some $\varepsilon_0 > 0$,*

$$\int_{\Omega_1} \log^-\Big(\int_{\mathbb{R}^n} \exp(-f_s(x) - \varepsilon_0|x|^2)\,dx\Big)\,d\mu_1(s) < \infty. \tag{1.2}$$

If for every $\alpha \in L^2(X_1, \mathbb{R}^n)$ we have

$$\int_{\Omega_2} g_t((A\alpha)(t))\,d\mu_2(t) \le \int_{\Omega_1} f_s(\alpha(s))\,d\mu_1(s), \tag{1.3}$$

then

$$\int_{\Omega_2} -\log\Big(\int_{\mathbb{R}^n} e^{-g_t(x)}\,dx\Big)\,d\mu_2(t) \le \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\,dx\Big)\,d\mu_1(s). \tag{1.4}$$

REMARK 1.4. We can easily relax the restrictions $f_s(x), g_t(x) \ge 0$. Suppose indeed that $f_s(x) \ge -a(s)$, $g_t(x) \ge -b(t)$, where $a(s), b(t)$ are non-negative functions on Ω_1, Ω_2 that are μ_1, μ_2-integrable respectively. Assuming, as we may, that $\int_{\Omega_1} a\,d\mu_1 = \int_{\Omega_2} b\,d\mu_2$, we see that the "basic assumption" (1.3) and the conclusion (1.4) are unchanged when passing from f_s, g_t to the non-negative functions $f_s + a(s)$, $g_t + b(t)$. So, in the theorem above, we are free to assume only that the functions are bounded from below in the way described in the latter discussion.

REMARK 1.5. By classical arguments of measure theory, we may replace f_s Borel by $\widetilde{f}_s$ l.s.c. such that $f_s \le \widetilde{f}_s$ and $\int_{\mathbb{R}^n} e^{-\widetilde{f}_s} \simeq \int_{\mathbb{R}^n} e^{-f_s}$, but we do not know how to do it on the g_t side in a way that (1.3) remains true. However, when the family $\{g_t\}$ consists of a single function g, it is easy to replace g by a l.s.c. function. The real issue is the infinite values of the f_s functions. If f_s, g_t are merely non-negative Borel functions, with f_s locally bounded, it is possible to reduce the problem to the case of continuous functions $\{f_s\}$ and $\{g_t\}$ by using convolution with non-negative compactly supported continuous kernels, hence reducing to the preceding theorem.

On the other hand, it is hard to believe that our l.s.c. assumption is necessary for the validity of Theorem 1.3. We rather think that the result is true for general Borel functions (as we can prove it in the “discrete case” of Theorem 1.6; see also Remark 4.5).

Suppose that the conditions of Setting 1 and Theorem 1.3 are satisfied for $\{f_s\}$ and $\{g_t\}$. Using the norm condition (i), we see that the basic assumption (1.3) remains true if we add the same multiple $\varepsilon|x|^2$ of $|x|^2$, $\varepsilon > 0$, to all the functions f_s and g_t. We can see that after this addition, the other conditions of Setting 1 and Theorem 1.3 remain obviously true, with two exceptions that are either less obvious or not always true: the semi-integrability condition (iv) remains true because f_s and g_t are also assumed to be non-negative, thus $\int_{\mathbb{R}^n} e^{-f_s(x)-\varepsilon|x|^2}\,dx \le C(\varepsilon)$ (and the same for g_t), and if we assume $0 < \varepsilon < \varepsilon_0$, then the condition (1.2) remains true, with ε_0 replaced by $\varepsilon_0 - \varepsilon$. It follows that the conclusion (1.4) holds if we replace the inner integration with respect to the Lebesgue measure on $\mathbb{R}^n$ by integration with respect to the isotropic Gaussian probability measure $\gamma_{n,\tau}$ on $\mathbb{R}^n$ defined by $d\gamma_{n,\tau}(x) = e^{-|x|^2/(2\tau)}\,dx/(2\pi\tau)^{n/2}$, provided $2\tau > \varepsilon_0^{-1}$. Actually, the proof of Theorem 1.3 will start with the Gaussian case and obtain the Lebesgue measure case from it, the Lebesgue case being the “flat” extremal case when $\tau \to \infty$. Note that, indeed, the (log of the) normalization constant $(2\pi\tau)^{n/2}$ on the two sides of (1.4) cancels out since μ_1 and μ_2 have the same finite mass.

In many applications, X_1 and X_2 are finite probability spaces, and then we work with finite families of objects parameterized by Ω_1 and Ω_2, or rather by the supports $\operatorname{supp}(\mu_1)$ and $\operatorname{supp}(\mu_2)$; in particular, the mappings $\alpha : \Omega_1 \to \mathbb{R}^n$ are families of $|\Omega_1|$ vectors in $\mathbb{R}^n$ and the linear operator

$$A : (\mathbb{R}^n)^{|\Omega_1|} \to (\mathbb{R}^n)^{|\Omega_2|}$$

is norm-one for the operator norm associated to the ℓ^2-norms weighted by the μ_i's, with the property that the vector $(x, \dots, x) \in (\mathbb{R}^n)^{|\Omega_1|}$ is sent to $(x, \dots, x) \in (\mathbb{R}^n)^{|\Omega_2|}$, for every $x \in \mathbb{R}^n$. In this case, we can relax the technical assumptions of Theorem 1.3.

THEOREM 1.6. *Under Setting* 1, *assume in addition that* μ_1 *and* μ_2 *are measures with* finite support *and* $\mu_1(\Omega_1) = \mu_2(\Omega_2) < \infty$. *Then the assumption* (1.3) *implies* (1.4) *with no further restriction on* $\{f_s\}$ *or* $\{g_t\}$.

Let us examine the first situations that one encounters when Ω_1 and Ω_2 are finite, with $\mu_1(\Omega_1) = \mu_2(\Omega_2) = 1$.

- The case $|\Omega_1| = |\Omega_2| = 1$ is trivial. One has $A(x) = x$ by (ii), and the statement amounts to the monotonicity of the integral.
- Assume $|\Omega_1| = 1$ and $|\Omega_2| = 2$. So $\Omega_2 = \{0, 1\}$, say, and $\mu_2 = (1-t)\delta_0 + t\delta_1$, $t \in [0, 1]$. The map $A : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ has to be $A(x) = (x, x)$, which satisfies both the required conditions, and the statement asserts that if the functions g_0, g_1, f satisfy $(1-t)g_0(x) + tg_1(x) \leq f(x)$ then $(\int e^{-g_0})^{1-t}(\int e^{-g_1})^t \geq \int e^{-f}$. This is just Hölder's inequality.
- Assume $|\Omega_1| = 2$ and $|\Omega_2| = 1$. Here $\Omega_1 = \{0, 1\}$, say, and $\mu_1 = (1-t)\delta_0 + t\delta_1$, $t \in [0, 1]$. Then we can take $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ to be $A(x_0, x_1) = (1-t)x_0 + tx_1$ (actually this is the only possible choice, as explained in Remark 4.9). This map will satisfy the required assumptions, by the convexity of the norm squared. The statement asserts that if $g((1-t)x_0 + tx_1) \leq (1-t)f_0(x_0) + tf_1(x_1)$ for all $(x_0, x_1) \in \mathbb{R}^n \times \mathbb{R}^n$, then $\int e^{-g} \geq \big(\int e^{-f_0}\big)^{1-t}\big(\int e^{-f_1}\big)^t$. This is the Prékopa–Leindler inequality.
- Assume $|\Omega_1| = |\Omega_2| = 2$. This is the first case where something new appears. Let us illustrate this by an example. Assume $\Omega_1 = \Omega_2 = \{0, 1\}$ and $\mu_1 = \mu_2 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Consider $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ defined by
$$A(x_0, x_1) = (2x_0/3 + x_1/3, x_0/3 + 2x_1/3).$$
We are in Setting 1 since $A(x, x) = (x, x)$ and $|2x_0/3 + x_1/3|^2 + |x_0/3 + 2x_1/3|^2 \leq |x_0|^2 + |x_1|^2$. The abstract statement then reduces exactly to Proposition 1.2 above.

In the next section, we present Borell's approach and establish in Proposition 2.4 a Gaussian version of Theorem 1.3; the proof of this proposition can be considered the heart of the present paper. In Section 3, we present the proofs of Theorems 1.3 and 1.6. Then, in the following section, we discuss some consequences. In Section 5 we present a generalization of Theorem 1.3 when the functions f_s and g_t live on Euclidean spaces of different dimensions. The result will include as particular cases the Brascamp–Lieb inequality (in the geometric form) and its reverse form devised by Franck Barthe [1]. Several technical proofs that have little geometric interest, and involve mostly measure-theoretic arguments, are gathered in Section 6.

2. Borell stochastic approach and Gaussian inequality. Borell's stochastic proof of the Prékopa–Leindler inequality relies on the representation formula given in the next lemma. Let $(B_r)_{r\geq 0}$ be a standard Brownian motion with values in $\mathbb{R}^n$, starting at 0, with filtration $\mathcal{F} = (\mathcal{F}_r)_{r\geq 0}$. Let $P_r = e^{r\Delta/2}$, $r \geq 0$, be the heat semigroup on $\mathbb{R}^n$ associated with this Brownian motion,

$$(P_r f)(x) = \mathbb{E}\, f(x + B_r) = \int_{\mathbb{R}^n} f(x+y) e^{-|y|^2/(2r)} \frac{dy}{(2\pi r)^{n/2}}, \quad x \in \mathbb{R}^n,$$

for f bounded and continuous on $\mathbb{R}^n$ and $r > 0$. An $\mathbb{R}^n$-valued drift $u = \{u_r\}_{r \le T}$ will be called *of class* D_2 on $[0,T]$ if it is $\mathcal{F}$-progressively measurable on $[0,T]$ and

$$\mathbb{E} \int_0^T |u_r|^2 \, dr < \infty.$$

LEMMA 2.1. *Let $T > 0$. For every bounded continuous function $f : \mathbb{R}^n \to \mathbb{R}$, we have*

$$-\log (P_T e^{-f})(0) = \inf_u \mathbb{E}\Big[f\Big(B_T + \int_0^T u_r \, dr\Big) + \tfrac{1}{2} \int_0^T |u_r|^2 \, dr \Big], \tag{2.1}$$

where the infimum is taken over $\mathbb{R}^n$-valued drifts $\{u_r\}_{r \le T}$ of class D_2. Moreover, the infimum is attained.

Proof (see Borell [6]). We begin by assuming that f is bounded and has bounded derivatives of order ≤ 2. For a drift $u = \{u_r\}_{r \le T}$ of class D_2, we define $X_r^u := B_r + \int_0^r u_\rho \, d\rho$, which satisfies the stochastic differential equation

$$X_0^u = 0, \qquad dX_r^u = dB_r + u_r \, dr$$

on the interval $[0,T]$. For $0 \le r \le T$, define $f_r = f_r^T$ by $e^{-f_r} = P_{T-r} e^{-f}$, that is,

$$f_r(x) = -\log (P_{T-r} e^{-f})(x), \quad x \in \mathbb{R}^n. \tag{2.2}$$

The function $(r,x) \mapsto f_r(x)$ on $(0,T) \times \mathbb{R}^n$ satisfies the partial differential equation

$$\partial_r f_r = -\tfrac{1}{2} \Delta f_r + \tfrac{1}{2} |\nabla f_r|^2.$$

By a direct application of the Itō formula, we see that the process

$$M_r := f_r(X_r^u) + \tfrac{1}{2} \int_0^r |u_\rho|^2 \, d\rho, \qquad r \in [0,T],$$

is a submartingale for any drift u in D_2, since

$$dM_r = \nabla f_r(X_r^u) \cdot dB_r + \tfrac{1}{2} |\nabla f_r(X_r^u) + u_r|^2 dr.$$

This implies the "inequality case" of (2.1), namely, an upper bound of $-\log (P_T e^{-f})(0)$ for every drift u in D_2. Indeed, note that $M_0 = f_0(X_0) = -\log (P_T e^{-f})(0)$ and $f_T = f$, and so the inequality in (2.1) immediately follows by considering $\mathbb{E}\, M_r$ at $r = 0$ and $r = T$. Moreover, if $u_r = -\nabla f_r(X_r^u)$, i.e., if X_r^u is the process solving the stochastic differential equation

$$X_0 = 0, \quad dX_r = dB_r - \nabla f_r(X_r) \, dr, \tag{2.3}$$

then M_r becomes a martingale, thus giving an equality case in (2.1).

Assume now that f is bounded and continuous on $\mathbb{R}^n$. Adding a constant to f, we may suppose that $f \ge 0$. Define f_r as in (2.2), and note that f_r is bounded above and ≥ 0. Since f is bounded, $P_r e^{-f}$ is bounded away from 0, so the function $f_{T-\varepsilon} = -\log(P_\varepsilon e^{-f})$ has bounded derivatives of all orders, for every $\varepsilon \in (0, T]$. Writing $T_\varepsilon = T - \varepsilon$ and $e^{-f_r} = P_{T_\varepsilon - r}(P_\varepsilon e^{-f}) = P_{T_\varepsilon - r} e^{-f_{T_\varepsilon}}$, $0 \le r \le T_\varepsilon$, we are back in the "good setting" on $[0, T_\varepsilon]$. Hence the optimal representation by the martingale

$$M_r = f_r\Big(B_r + \int_0^r u_\rho\, d\rho\Big) + \tfrac12 \int_0^r |u_\rho|^2\, d\rho = M_0 + \int_0^r \nabla f_\rho(X_\rho) \cdot dB_\rho,$$

with $M_0 = f_0(0)$, $u_r = -\nabla f_r(X_r)$ and $X_r = B_r + \int_0^r u_\rho\, d\rho$, is valid for $r \le T_\varepsilon$. Since $f_r \ge 0$, we see that $\int_0^r |u_\rho|^2\, d\rho \le 2M_r$, and we obtain

$$\mathbb{E}\,|M_r - M_0|^2 = \mathbb{E}\Big(\int_0^r |\nabla f_\rho(X_\rho)|^2\, d\rho\Big) = \mathbb{E}\Big(\int_0^r |u_\rho|^2\, d\rho\Big) \le 2\,\mathbb{E}\, M_r = 2 f_0(0).$$

We have $T^{-1}\,\mathbb{E}\big(\int_0^T |u_\rho|\, d\rho\big)^2 \le \mathbb{E}\int_0^T |u_\rho|^2\, d\rho \le 2 f_0(0)$, hence $u \in D_2$, X_r converges a.s. and in L^2 to $X_T = B_T + \int_0^T u_\rho\, d\rho$. We see that $(M_r)_{r<T}$ is an L^2-bounded martingale, thus M_r converges in L^2-norm, as $r \to T$, to a limit M_T such that $\mathbb{E}\, M_T = M_0$. On the other hand, M_r converges a.s. to

$$f\Big(B_T + \int_0^T u_\rho\, d\rho\Big) + \tfrac12 \int_0^T |u_\rho|^2\, d\rho \tag{2.4}$$

as $r \to T$ since f_r converges locally uniformly to the bounded continuous function f. This implies that M_T is equal to the expression in (2.4), hence the expectation of (2.4) is equal to $M_0 = f_0(0)$, that is, this implies formula (2.1) with equality.

In the inequality case of (2.1), the drift u is in D_2 by assumption, hence $(X_r^u)_{r<T}$ defined as above converges a.s. and in L^2 to X_T^u as $r \to T$. The submartingale $(M_r)_{r<T}$ is L^2-bounded, hence converges a.s. and in L^2 to the expression in (2.4), and the result follows. ∎

Remark 2.2. If f is bounded and upper semicontinuous on $\mathbb{R}^n$, then for every $x \in \mathbb{R}^n$ and $\varepsilon > 0$, there is $r_0 < T$ and a neighborhood V of x such that $(P_{T-r} e^{-f})(y) > e^{-f(x)-\varepsilon}$ for $y \in V$, $r_0 < r < T$, that is, (a) $f_r(y) < f(x) + \varepsilon$. When f is lower semicontinuous, the inequality is reversed, (b) $f_r(y) > f(x) - \varepsilon$. When f is u.s.c. it follows from (a) that in the inequality case of (2.1), the limit of $(M_r)_{r<T}$ as $r \to T$ is less than or equal to the expression in (2.4), so the inequality case remains true. For every bounded Borel function f, we can find $\widetilde f$ u.s.c. such that $\widetilde f \le f$ and $\int_{\mathbb{R}^n} e^{-\widetilde f}\, d\gamma \simeq \int_{\mathbb{R}^n} e^{-f}\, d\gamma$. This implies that the inequality case in (2.1) is true for every bounded Borel function f. Using the reverse inequality (b)

for l.s.c. functions, we see that the equality in (2.1) remains true for bounded lower semicontinuous functions (but the infimum need not be achieved by an optimal martingale).

We shall need more than the case of bounded functions. The proof of the next lemma will be given in Section 6; it requires a closer look at the argument given above.

LEMMA 2.3. *Formula* (2.1) *remains valid when f is continuous, bounded below and satisfies an exponential upper bound of the form*

$$f(x) \leq ae^{b|x|}, \quad a, b \geq 0,\ x \in \mathbb{R}^n. \tag{2.5}$$

The "inequality case" in (2.1) *is valid for any lower bounded continuous function f.*

The next proposition provides a rather direct and simple link between Borell's lemma and our theorem.

PROPOSITION 2.4. *Under Setting* 1, *assume that f_s, g_t are continuous ≥ 0 on $\mathbb{R}^n$* (*or bounded from below as in Remark* 1.4), *and that f_s satisfies the following exponential bound: there are non-negative measurable functions $a(s), b(s)$ on Ω_1 such that*

$$f_s(x) \leq a(s)e^{b(s)|x|}, \quad x \in \mathbb{R}^n. \tag{2.6}$$

Then, for every isotropic Gaussian probability measure $\gamma = \gamma_{n,\tau}$ on $\mathbb{R}^n$, the basic assumption (1.3) *implies that*

$$\int_{\Omega_2} -\log\Big(\int_{\mathbb{R}^n} e^{-g_t(x)}\,d\gamma(x)\Big)\,d\mu_2(t) \leq \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\,d\gamma(x)\Big)\,d\mu_1(s). \tag{2.7}$$

As seen from the proof of Theorem 1.6 (which indeed reduces the study to the situation of this proposition), we can remove the additional assumptions on f_s and g_t when the measures μ_1 and μ_2 have finite support.

Proof of Proposition 2.4. The proof will be an application of formula (2.1), as extended in Lemma 2.3, to the functions f_s and g_t. We may assume that the right-hand side of (2.7) is not equal to ∞. The Gaussian measure $\gamma = \gamma_{n,\tau}$ is the distribution of B_T for $T = \tau > 0$, so that

$$-\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\,d\gamma(x)\Big) = -\log\,(P_T e^{-f_s})(0) < \infty \qquad \mu_1\text{-a.e.}$$

Let $\varepsilon > 0$. By Lemma 2.3, for almost every $s \in \Omega_1$, we can introduce $U(s) \in D_2$, an almost optimal drift for the function f_s, $s \in \Omega_1$, namely, a drift $U(s) = \{U_r(s)\}_{0\leq r\leq T}$ such that

$$\mathbb{E}\Big[f_s\Big(B_T + \int_0^T U_r(s)\,dr\Big) + \tfrac{1}{2}\int_0^T |U_r(s)|^2\,dr\Big] < -\log\,(P_T e^{-f_s})(0) + \varepsilon. \tag{2.8}$$

We will show later (Claim 6.3) that the process $\{U(s)\}_{s\in\Omega_1}$ can be chosen to be Σ_1-measurable. Note also that (2.8) and $f_s \ge 0$ ensure that

$$\tag{2.9} \int_{\Omega_1} \mathbb{E}\int_0^T |U_r(s)|^2\,dr\,d\mu_1(s) \le 2\Big(\int_{\Omega_1} -\log\,(P_Te^{-f_s})(0)\,d\mu_1(s) + \varepsilon\mu_1(\Omega_1)\Big) < \infty.$$

Assume that the Brownian motion (B_r) is defined on a probability space $(E,\mathcal{A},\mathbb{P})$. We have an $\mathbb{R}^n$-valued random process U, that will be denoted by one of

$$U(s,r,\omega) = U_r(s)(\omega) = U_r(s,\omega) = U_{r,\omega}(s), \quad s\in\Omega_1,\ r\in[0,T],\ \omega\in E.$$

By (2.9), we know that

$$U \in L^2(\Omega_1\times[0,T]\times E,\mathbb{R}^n) = L^2([0,T]\times E, L^2(\Omega_1,\mathbb{R}^n)).$$

We shall estimate $P_Te^{-g_t}$ using the inequality case of formula (2.1) given by Lemma 2.3, with the drift $\{V_r(t)\}_{r\le T} = \{(AU_r)(t)\}_{r\le T}$, namely

$$\tag{2.10} V_{r,\omega}(t) := (AU_{r,\omega})(t) = A(s\mapsto U_{r,\omega}(s))(t) \in \mathbb{R}^n, \quad t\in\Omega_2,$$

where $U_{r,\omega}\in L^2(X_1,\mathbb{R}^n)$ for almost every r,ω. We shall use the basic assumption (1.3) on the families $\{f_s\}$ and $\{g_t\}$ for the random function $\alpha_\omega = B_T(\omega)+\beta_\omega$, where β_ω is defined by

$$\tag{2.11} \beta_\omega(s) := \int_0^T U_r(s,\omega)\,dr,$$

and where we consider $B_T(\omega)$ as a constant function of the s variable. We know by (2.9) that β_ω (and α_ω) are in $L^2(X_1,\mathbb{R}^n)$ for almost every ω. The constant functions condition (ii) on A ensures that

$$\tag{2.12} \begin{aligned}(A\alpha_\omega)(t) &= A(B_T(\omega)+\beta_\omega)(t) = B_T(\omega) + (A\beta_\omega)(t)\\ &= B_T(\omega) + \int_0^T V_r(t,\omega)\,dr.\end{aligned}$$

As already mentioned, we apply the inequality case of formula (2.1) for g_t with the drift $\{V_r\}$ and then we integrate in t to get

$$\int_{\Omega_2} -\log\,(P_Te^{-g_t})(0)\,d\mu_2(t) \le \int_{\Omega_2} \mathbb{E}\Big[g_t\Big(B_T + \int_0^T V_r(t)\,dr\Big) + \tfrac12\int_0^T |V_r(t)|^2\,dr\Big]\,d\mu_2(t).$$

For future reference, we state an obvious consequence of (2.12),

$$g_t\big((A\alpha_\omega)(t)\big) = g_t\big(A(B_T(\omega)+\beta_\omega)(t)\big) = g_t\big(B_T(\omega)+(A\beta_\omega)(t)\big) = g_t\Big(B_T(\omega)+\int_0^T V_r(t,\omega)\,dr\Big). \tag{2.13}$$

Using (1.3) and (2.13), we have on the set E the pointwise inequality

$$\int_{\Omega_2} g_t\Big(B_T(\omega)+\int_0^T V_r(t,\omega)\,dr\Big)\,d\mu_2(t) \le \int_{\Omega_1} f_s\Big(B_T(\omega)+\int_0^T U_r(s,\omega)\,dr\Big)\,d\mu_1(s),$$

and since $\|A\| \le 1$, we have another pointwise inequality: for every $r \in [0,T]$,

$$\int_{\Omega_2} |V_r(t,\omega)|^2\,d\mu_2(t) \le \int_{\Omega_1} |U_r(s,\omega)|^2\,d\mu_1(s). \tag{2.14}$$

Finally,

$$\begin{aligned}\int_{\Omega_2} &-\log\,(P_T e^{-g_t})(0)\,d\mu_2(t)\\ &\le \int_{\Omega_1} \mathbb{E}\Big[f_s\Big(B_T+\int_0^T U_r(s)\,dr\Big)+\tfrac12\int_0^T |U_r(s)|^2\,dr\Big]\,d\mu_1(s)\\ &< \int_{\Omega_1} -\log\,(P_T e^{-f_s})(0)\,d\mu_1(s)+\varepsilon\mu_1(\Omega_1).\end{aligned}$$

We conclude by letting $\varepsilon \to 0$.

As the reader has noticed, the proof is rather short, provided one has at hand the abstract properties contained in the four equations (2.10), (2.11), (2.13) and (2.14) that allow us to run Borell's argument. ∎

3. Proofs of Theorems 1.3 and 1.6. The proofs rely on Proposition 2.4. We shall explain how to pass from a Gaussian measure to the Lebesgue measure, and how to approximate our functions in order to meet the required technical assumptions (continuity and lower/upper bounds).

3.1. Proof of Theorem 1.6. Under Setting 1 in the discrete case, assume that $f_0, f_1, \dots, f_p$ and $g_0, g_1, \dots, g_q$ are Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{\infty\}$, and that for all points $x_0, x_1, \dots, x_p$ in $\mathbb{R}^n$ we have

$$\sum_{j=0}^{q} \nu_j g_j\Big(\sum_{i=0}^{p} a_{j,i}x_i\Big) \le \sum_{i=0}^{p} \mu_i f_i(x_i), \tag{3.1}$$

where $\mu_i, \nu_j > 0$, $\sum_{i=0}^p \mu_i = \sum_{j=0}^q \nu_j = 1$, where the matrix $A = (a_{j,i})$ satisfies the norm condition (i) and the constant functions condition (ii), each $a_{j,i}$ being an $n\times n$ matrix acting from $\mathbb{R}^n$ to $\mathbb{R}^n$. We want to prove that

$$\sum_{j=0}^{q} -\nu_j \log\Big(\int_{\mathbb{R}^n} e^{-g_j(x)}\,dx\Big) \le \sum_{i=0}^{p} -\mu_i \log\Big(\int_{\mathbb{R}^n} e^{-f_i(x)}\,dx\Big). \tag{3.2}$$

If $\int_{\mathbb{R}^n} e^{-f_{i_0}} = 0$ for one i_0, then by the semi-integrability assumption (iv), the other $\int_{\mathbb{R}^n} e^{-f_i}$ are finite, the right-hand side of (3.2) is ∞, and this case is obvious. We may therefore assume that all f_i are L-proper. It follows that all g_j, $j = 0, \dots, q$, are L-proper. Indeed, each $A_i = \{f_i < \infty\}$ has positive Lebesgue measure, so we can find and fix a Lebesgue density point $x_i^{(0)}$ of A_i, for $i = 0, \dots, p$. Then the set U of $u \in \mathbb{R}^n$ such that $x_i^{(0)} + u \in A_i$ for all $i = 0, \dots, p$, i.e., $U = \bigcap_{i=0}^p (A_i - x_i^{(0)})$, has positive measure. We know that $\sum_{i=0}^p a_{j,i} = I_n$ since A preserves constant functions by (ii). We see from (3.1) that

$$\begin{aligned}\sum_{j=0}^{q} \nu_j g_j\Big(u + \sum_{i=0}^{p} a_{j,i} x_i^{(0)}\Big) &= \sum_{j=0}^{q} \nu_j g_j\Big(\sum_{i=0}^{p} a_{j,i}(x_i^{(0)} + u)\Big) \\ &\le \sum_{i=0}^{p} \mu_i f_i(x_i^{(0)} + u) < \infty\end{aligned}$$

for every $u \in U$, hence all g_j are L-proper.

For $\varepsilon > 0$ and $i = 0, \dots, p$, $j = 0, \dots, q$, the functions $f_{i,\varepsilon}(x) = f_i(x) + \varepsilon|x|^2$ and $g_{j,\varepsilon}(x) = g_j(x) + \varepsilon|x|^2$ still satisfy (3.1) by the assumption $\|A\| \le 1$. We may reduce the problem to proving (3.2) for $f_{i,\varepsilon}$, $g_{j,\varepsilon}$, since we can let $\varepsilon \to 0$ in $\int_{\mathbb{R}^n} e^{-g_j(x)-\varepsilon x^2}\,dx$, $\int_{\mathbb{R}^n} e^{-f_i(x)-\varepsilon x^2}\,dx$, obtaining (3.2) by monotone convergence. In other words, we may keep f_i, g_j but replace the Lebesgue measure in (3.2) by a Gaussian measure $d\gamma(x) = e^{-\varepsilon|x|^2}\,dx$. If $\int_{\mathbb{R}^n} e^{-g_{j_0}}\,d\gamma = \infty$ for one j_0, there is nothing to prove: the other integrals $\int_{\mathbb{R}^n} e^{-g_j}\,d\gamma$, $j \ne j_0$, are > 0 and the left-hand side of (3.2) is $-\infty$. Otherwise, for $N \in \mathbb{N}$, define $g_{j,N} = \min(g_j, N)$ and $f_{i,N} = \max(f_i, -N)$, which trivially satisfy (3.1), and observe that it is enough to give the proof for $f_{i,N}$, $g_{j,N}$; indeed, since $e^{-g_{j,0}} = e^{-\min(g_j,0)} \le e^{-g_j} + 1$ is integrable with respect to $d\gamma$, we may again pass to the (decreasing) limit in the integrals $\int_{\mathbb{R}^n} e^{-g_{j,N}}\,d\gamma$ as $N \to \infty$, and use monotone convergence for $\int_{\mathbb{R}^n} e^{-f_{i,N}}\,d\gamma$.

Now we have reduced the question to the case $g_j \le N$ and $f_i \ge -N$. Thanks to the discrete situation, we may further assume that f_{i_0} is bounded above by $2N/\mu_{i_0}$ for each i_0, since

$$\sum_{j=0}^{q} \nu_j g_j\Big(\sum_{i=0}^{p} a_{j,i} x_i\Big) - \sum_{i \ne i_0} \mu_i f_i(x_i) \le N + (1 - \mu_{i_0})N < 2N,$$

and for the same reason, we may also assume g_{j_0} is bounded below by $-2N/\nu_{j_0}$, while still keeping (3.1) true.

Finally, we restricted the problem to bounded Borel functions f_i, g_j and a Gaussian measure γ. If all f_i, g_j are translated by the same vector, then (3.1) remains true by the constant functions condition (ii), therefore (3.1) is stable under convolution with non-negative kernels. We may approximate in $L^1(\gamma)$-norm the functions f_i, g_j by convolution with a compactly supported continuous non-negative kernel, keeping (3.1) and meeting the assumptions of Proposition 2.4. We thus obtain, for a sequence of continuous approximations $f_{i,k}$ and $g_{j,k}$, $k \in \mathbb{N}$, the inequality

$$\sum_{j=0}^{q} -\nu_j \log\Big(\int_{\mathbb{R}^n} e^{-g_{j,k}}\, d\gamma\Big) \leq \sum_{i=0}^{p} -\mu_i \log\Big(\int_{\mathbb{R}^n} e^{-f_{i,k}}\, d\gamma\Big).$$

Some subsequences of the sequence of continuous approximations tend almost everywhere to f_i, g_j respectively, and we finish the proof for the latter restricted problem by using the dominated convergence theorem.

This ends the proof of Theorem 1.6.

3.2. Proof of Theorem 1.3. Suppose that we try to obtain the conclusion (1.4) of Theorem 1.3 for $\{f_s\}$ and $\{g_t\}$ as a limit case of "good cases" for which the conclusion is known, say $\{f_{s,k}\}$ and $\{g_{t,k}\}$, $k \in \mathbb{N}$, such that $f_{s,k} \to f_s$, $g_{t,k} \to g_t$ as $k \to \infty$. The next elementary lemma will help us to do it.

Consider a measure space (Ω, Σ, μ) and a measure ν on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, where μ and ν are σ-finite. Let $h_{s,k}(x)$, $k \in \mathbb{N}$ and $h_s(x)$, $s \in \Omega$, $x \in \mathbb{R}^n$ be $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable from $\Omega \times \mathbb{R}^n$ to $\mathbb{R} \cup \{\infty\}$. Define H_k and H by

$$e^{-H_k(s)} = \int_{\mathbb{R}^n} e^{-h_{s,k}(x)}\, d\nu(x), \quad e^{-H(s)} = \int_{\mathbb{R}^n} e^{-h_s(x)}\, d\nu(x). \tag{3.3}$$

These functions are Σ-measurable by the theory behind Fubini's theorem. Note that

$$H^+(s) = \log^-\Big(\int_{\mathbb{R}^n} e^{-h_s(x)}\, d\nu(x)\Big), \quad H^-(s) = \log^+\Big(\int_{\mathbb{R}^n} e^{-h_s(x)}\, d\nu(x)\Big).$$

LEMMA 3.1. *Let $H_k(s), H(s)$ be defined by* (3.3)*. Assume that H_k, $k \in \mathbb{N}$, and H are μ-semi-integrable.*

(a) *Suppose that $h_{s,k}(x) \nearrow h_s(x)$ pointwise on $\Omega \times \mathbb{R}^n$, and $H_0(s) > -\infty$ for μ-almost every $s \in \Omega$. Then:*

(a1) *$H_k(s) \nearrow H(s)$ μ-a.e., $\int_\Omega H(s)\, d\mu(s) \geq \limsup_k \int_\Omega H_k(s)\, d\mu(s)$.*

(a2) *If in addition $\int_\Omega H_0^-(s)\, d\mu(s) < \infty$, then*

$$\int_\Omega H(s)\, d\mu(s) = \lim_k \int_\Omega H_k(s)\, d\mu(s).$$

(b) *Suppose that $h_{s,k}(x) \searrow h_s(x)$ pointwise on $\Omega \times \mathbb{R}^n$. Then:*

(b1) $H_k(s) \searrow H(s)$ *pointwise on* Ω, *and*

$$\int_\Omega H(s)\, d\mu(s) \le \liminf_k \int_\Omega H_k(s)\, d\mu(s).$$

(b2) *If in addition* $\int_\Omega H_0^+(s)\, d\mu(s) < \infty$, *then*

$$\int_\Omega H(s)\, d\mu(s) = \lim_k \int_\Omega H_k(s)\, d\mu(s).$$

Proof. Apply successively the classical results of integration theory: the Fatou lemma, the monotone convergence theorem and the dominated convergence theorem. In case (a), we have $e^{-h_{s,k}(x)} \searrow e^{-h_s(x)}$, and $H_0(s) > -\infty$ allows us to apply dominated convergence to ν and deduce that $H_k(s) \nearrow H(s)$. Next, $H_k^+(s) \nearrow H^+(s)$ and $H_k^-(s) \searrow H^-(s)$. In $\int_\Omega H = \int_\Omega H^+ - \int_\Omega H^-$ (which makes sense by the semi-integrability assumption), apply monotone convergence for H^+ and Fatou for H^-. If $\int_\Omega H_0^-\, d\mu < \infty$, replace Fatou by dominated convergence for H^-. The proof of (b) is similar and left to the reader. ∎

Proof of Theorem 1.3. Since μ_1 and μ_2 have the same finite mass, it is enough, as mentioned after the statement of the theorem, to have an inequality involving $(P_T e^{-g_t})(0)$ and

$$(P_T e^{-f_s})(0) = \int_{\mathbb{R}^n} e^{-f_s(x) - |x|^2/(2T)} \frac{dx}{(2\pi T)^{n/2}}, \quad 2T > 1/\varepsilon_0.$$

The result will follow by letting $T \to \infty$. Indeed, $f_s(x) + \varepsilon|x|^2$ decreases to $f_s(x)$ and $g_t(x) + \varepsilon|x|^2$ decreases to $g_t(x)$ as $\varepsilon \searrow 0$; for $\varepsilon \in [0, \varepsilon_0]$, define $F_\varepsilon(s)$, $G_\varepsilon(t)$ as in (3.3), with ν being the Lebesgue measure, that is, let $e^{-F_\varepsilon(s)} = \int_{\mathbb{R}^n} e^{-f_s(x) - \varepsilon|x|^2}\, dx$ and use a similar expression for $G_\varepsilon(t)$. By the semi-integrability condition (iv), $G_0(t)$ is semi-integrable, and we may assume that $G_0^-(t)$ is integrable, otherwise the left-hand side of (1.4) is $-\infty$, an obvious case. We also know by the assumption (1.2) that $F_{\varepsilon_0}^+(s) = \log^-\big(\int_{\mathbb{R}^n} e^{-f_s(x) - \varepsilon_0|x|^2}\, dx\big)$ is integrable. This implies that $F_\varepsilon(s)$ and $G_\varepsilon(t)$ are semi-integrable for $\varepsilon \in [0, \varepsilon_0]$. Use Lemma 3.1(b) to see that $F_\varepsilon(s) \to F(s)$ and $G_\varepsilon(t) \to G(t)$ for every s, t. For the F side, we may use (b2) because $F_{\varepsilon_0}^+$ is integrable. We can therefore let $\varepsilon \to 0$ to conclude that

$$\begin{aligned}\int_{\Omega_1} F_\varepsilon(s)\, d\mu_1(s) &= \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x) - \varepsilon|x|^2}\, dx\Big)\, d\mu_1(s) \\ &\to \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s}\Big)\, d\mu_1(s).\end{aligned}$$

For the analogous expressions with g_t, we use (b1) of Lemma 3.1.

Having reduced the problem to the Gaussian measure $\gamma = \gamma_{n,\tau}$, our goal is now to show how to relax the continuity assumption from Proposition 2.4. To this end, we shall introduce classical continuous approximations obtained by inf-convolution with large multiples of $x \mapsto |x|^2$. To work with these approximations, we shall need the existence of a selection $\alpha_0 \in L^2(X_1, \mathbb{R}^n)$ such that $\int_{\Omega_1} f_s(\alpha_0(s))\, d\mu_1(s) < \infty$, which is granted by Lemma 6.4 since we know by (1.2) that $s \mapsto \log^-(\int_{\mathbb{R}^n} e^{-f_s(x)}\, d\gamma(x))$ is μ_1-integrable when $2\tau > 1/\varepsilon_0$.

We fix $k > 0$ and define $f_{s,k}$, $g_{t,k}$ by the inf-convolution of f_s, g_t with $h_k(x) = k|x|^2$,

$$f_{s,k}(x) = \inf_{u\in\mathbb{R}^n} \bigl(f_s(x+u)+k|u|^2\bigr), \quad g_{t,k}(x) = \inf_{u\in\mathbb{R}^n} \bigl(g_t(x+u)+k|u|^2\bigr). \tag{3.4}$$

Clearly, $f_{s,k} \le f_s$, $g_{t,k} \le g_t$, and $f_{s,k}$, $g_{t,k}$ are continuous functions on $\mathbb{R}^n$. By Lemma 6.4, we may find negligible Borel sets $N_1 \in \Sigma_1$ and $N_2 \in \Sigma_2$ such that $(s,x) \mapsto f_{s,k}(x)$ and $(t,x) \mapsto g_{t,k}(x)$ are Borel functions on $(\Omega_1 \setminus N_1) \times \mathbb{R}^n$ and $(\Omega_2 \setminus N_2) \times \mathbb{R}^n$. We have

$$0 \le f_{s,k}(x) \le f_s(\alpha_0(s)) + k|x - \alpha_0(s)|^2, \quad x \in \mathbb{R}^n,$$

fitting the exponential bound (2.5) needed to apply Proposition 2.4. Let α be in $L^2(X_1, \mathbb{R}^n)$. For any fixed $\varepsilon > 0$, we may find a measurable selection $u(s)$ (Lemma 6.4) such that

$$f_s(\alpha(s) + u(s)) + k|u(s)|^2 - \varepsilon < f_{s,k}(\alpha(s)) \le f_s(\alpha_0(s)) + k|\alpha(s) - \alpha_0(s)|^2.$$

Since $f_s \ge 0$, this shows that $u \in L^2(X_1, \mathbb{R}^n)$. We can write

$$\int_{\Omega_2} g_{t,k}((A\alpha)(t))\, d\mu_2(t) \le \int_{\Omega_2} \bigl[g_t\bigl((A\alpha)(t) + (Au)(t)\bigr) + k|(Au)(t)|^2\bigr]\, d\mu_2(t),$$

which is bounded, in view of the basic assumption (1.3) and the norm condition $\|A\| \le 1$, by

$$\int_{\Omega_1} \bigl[f_s\bigl(\alpha(s) + u(s)\bigr) + k|u(s)|^2\bigr]\, d\mu_1(s) \le \int_{\Omega_1} f_{s,k}(\alpha(s))\, d\mu_1(s) + \varepsilon\mu_1(\Omega_1).$$

Hence, $f_{s,k}$ and $g_{t,k}$ also satisfy the basic assumption. By Proposition 2.4,

$$\int_{\Omega_2} -\log\Bigl(\int_{\mathbb{R}^n} e^{-g_{t,k}(x)}\, d\gamma(x)\Bigr)\, d\mu_2(t) \le \int_{\Omega_1} -\log\Bigl(\int_{\mathbb{R}^n} e^{-f_{s,k}(x)}\, d\gamma(x)\Bigr)\, d\mu_1(s).$$

When k tends to infinity, $f_{s,k}$ increases to a l.s.c. function $\widetilde{f}_s \le f_s$, and $g_{t,k}(x)$ increases to $g_t(x)$ because $x \mapsto g_t(x)$ is l.s.c. on $\mathbb{R}^n$. We apply again Lemma 3.1, this time with $\nu = \gamma$, defining $F_k(s)$, $\widetilde{F}(s) \le F(s)$, $G_k(t)$, $G(t)$ from $f_{s,k}$, $\widetilde{f}_s, f_s$, $g_{t,k}, g_t$ as in (3.3). Since the functions are ≥ 0 and $e^{-f_{s,k}(x)}$ and $e^{-g_{t,k}(x)}$ are bounded by 1, we have $F_k(s), G_k(t) \ge 0$ because γ is a probability measure. Thus $F_0(s), G_0(t) > -\infty$ and $F_k(s) \to \widetilde{F}(s)$ and

$G_k(t) \to G(t)$ by (a). The conclusion follows, by (a1) for F, and by (a2) for G since $G_0^- = 0$. ∎

REMARK 3.2. Assume that f_s and g_t are continuous ≥ 0 on $\mathbb{R}^n$. Let D be a countable dense subset of $\mathbb{R}^n$. Since f_s and g_t are continuous on $\mathbb{R}^n$, we may define the inf-convolution by

$$f_{s,k}(x) = \inf_{u \in D} \big(f_s(x+u) + k|u|^2\big), \quad g_{t,k}(x) = \inf_{u \in D} \big(g_t(x+u) + k|u|^2\big).$$

It is now clear that $(s,x) \mapsto f_{s,k}(x)$ and $(t,x) \mapsto g_{t,k}(x)$ are $\Sigma_i \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable, as countable infima of measurable functions. Similarly, the possibility of selecting $u(s)$ is now evident.

4. Other formulations and consequences. We start with a simple, natural situation where the assumptions of Theorem 1.3 are satisfied. Assume that Ω is a Polish space, Σ its Borel σ-algebra and μ a probability measure on (Ω, Σ). Take $\Sigma_1 = \Sigma$, let $\Sigma_2 \subset \Sigma$ be a sub-σ-algebra of Σ, and let $\mu_1 = \mu_2 = \mu$. Then the conditional expectation

$$A = \mathbb{E}[\,\cdot \mid \Sigma_2] : L^2(\Sigma, \mathbb{R}^n) \to L^2(\Sigma_2, \mathbb{R}^n) \subset L^2(\Sigma, \mathbb{R}^n)$$

satisfies the norm condition (i) and the constant functions condition (ii) from Setting 1. An already interesting case is when we take Σ_2 to be trivial, $\Sigma_2 = \{\emptyset, \Omega\}$, in which case A is simply the μ-mean,

$$A\alpha = \int_\Omega \alpha(s)\, d\mu(s),$$

and the space Ω_2 can be then considered as being a one-point space, say $\Omega_2 = \{0\}$. In this case, the family $\{g_t\}$ consists of a single function g and Theorem 1.3 reads:

COROLLARY 4.1. *Let Ω be a Polish space, Σ its Borel σ-algebra, and μ a probability measure on (Ω, Σ). Suppose that we are in Setting 1 with $(\Omega_1, \Sigma_1, \mu_1) = (\Omega, \Sigma, \mu)$ and $\Omega_2 = \{0\}$. Let $\{f_s\}_{s\in\Omega}$ satisfy the assumptions of Theorem 1.3, and let g be a lower bounded l.s.c. function on $\mathbb{R}^n$. If for every $\alpha \in L^2(\Omega, \Sigma, \mu, \mathbb{R}^n)$ we have*

$$g\Big(\int_\Omega \alpha(s)\, d\mu(s)\Big) \leq \int_\Omega f_s(\alpha(s))\, d\mu(s),$$

then

$$-\log\Big(\int_{\mathbb{R}^n} e^{-g}\Big) \leq \int_\Omega -\log\Big(\int_{\mathbb{R}^n} e^{-f_s}\Big)\, d\mu(s).$$

The Prékopa–Leindler inequality follows by taking Ω to be a two-point probability space, for instance $\Omega = \{0,1\}$, and taking $\mu = (1-t)\delta_0 + t\delta_1$ for some $t \in [0,1]$. If we replace the two-point space (i.e., Bernoulli variables) by the unit circle $S^1 = \{e^{i\theta} : \theta \in \mathbb{R}\}$ (i.e., Steinhaus variables), then we obtain:

COROLLARY 4.2. *Under the assumptions of Corollary* 4.1 *with* $\Omega = S^1$ *and* $\mu = d\theta/(2\pi)$, *let* $\{f_\xi\}_{\xi\in S^1}$ *be non-negative* (*or properly bounded from below as in Remark* 1.4) *Borel functions on* $\mathbb{R}^n$ *and let* g *be a lower bounded l.s.c. function on* $\mathbb{R}^n$. *If for every* $\alpha \in L^2(S^1, \mathbb{R}^n)$ *we have*

$$g\Big(\int_0^{2\pi} \alpha(e^{i\theta})\,\frac{d\theta}{2\pi}\Big) \le \int_0^{2\pi} f_{e^{i\theta}}(\alpha(e^{i\theta}))\,\frac{d\theta}{2\pi},$$

then

$$-\log\Big(\int_{\mathbb{R}^n} e^{-g}\Big) \le \int_0^{2\pi} -\log\Big(\int_{\mathbb{R}^n} e^{-f_{e^{i\theta}}}\Big)\,\frac{d\theta}{2\pi}.$$

A consequence of the previous corollary is one of Berndtsson's plurisubharmonic extensions of the Prékopa theorem, the relatively easy "tube" case.

COROLLARY 4.3 (Berndtsson [3]). *Let* $\varphi : \mathbb{C} \times \mathbb{C}^n \to \mathbb{R}$ *be plurisubharmonic on* $\mathbb{C}^{n+1}$ *and such that* $\varphi(z, w) = \varphi(z, \Re w)$ *for every* $z \in \mathbb{C}$ *and* $w \in \mathbb{C}^n$. *Then the function*

$$\psi(z) = -\log\Big(\int_{\mathbb{R}^n} e^{-\varphi(z,x)}\,dx\Big)$$

is subharmonic on $\mathbb{C}$.

Proof. We will check that ψ is subharmonic at $z = 0$, say. We can assume that φ is bounded below for all z close to 0. We want to prove that $\psi(0) \le \int_0^{2\pi} \psi(re^{i\theta})\,d\theta/(2\pi)$ for any fixed $r > 0$ small enough. We take $r = 1$ to simplify notation. Set $g(x) = \varphi(0, x)$ and $f_{e^{i\theta}}(x) := \varphi(e^{i\theta}, x)$ for $x \in \mathbb{R}^n \subset \mathbb{C}^n$. It suffices to check that these functions satisfy the hypothesis of the previous corollary.

Let $\alpha \in L^2(S^1, \mathbb{R}^n)$ and let $\widetilde{\alpha}$ be the harmonic extension of α to the unit disc $\mathbb{D}$. In particular $\widetilde{\alpha}(0) = \int_0^{2\pi} \alpha(e^{i\theta})\,d\theta/(2\pi)$. We can write $\widetilde{\alpha}$ as the real part $\Re H$ of a holomorphic function $H : \mathbb{D} \to \mathbb{C}^n$ such that $H(0) = \widetilde{\alpha}(0)$. We conclude by noticing that $z \mapsto \varphi(z, H(z))$ is subharmonic on $\mathbb{D}$, and using $\varphi(z, H(z)) = \varphi(z, \widetilde{\alpha}(z))$ we obtain

$$\begin{aligned} g\Big(\int_0^{2\pi} \alpha(e^{i\theta})\,\frac{d\theta}{2\pi}\Big) &= \varphi(0, H(0)) \\ &\le \int_0^{2\pi} \varphi(e^{i\theta}, H(e^{i\theta}))\,\frac{d\theta}{2\pi} = \int_0^{2\pi} f_{e^{i\theta}}(\alpha(e^{i\theta}))\,\frac{d\theta}{2\pi}. \ \blacksquare \end{aligned}$$

Actually, it is easily seen and certainly known that the previous Corollary 4.1 can be deduced from the Prékopa–Leindler inequality, by discretizing μ and by a simple induction procedure. Therefore, Corollary 4.1 and its consequences only serve as motivation for the abstract setting above, but do not bring new information.

The situation is different with other instances of Theorem 1.3 such as Proposition 1.2 and the following general result. It seems difficult to guess the existence of such inequalities without having Borell's proof in mind.

THEOREM 4.4. *Let μ be a Borel probability measure on $[0,1]$ and set $m := \int t\,d\mu(t) \in [0,1]$. Under Setting* 1 *with $\Omega_1 = \{0,1\}$ and $\Omega_2 = [0,1]$, $\mu_2 = \mu$, let f_0, f_1 and $\{g_t\}_{t\in[0,1]}$ be lower bounded* (*as in Remark* 1.4) *Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{\infty\}$. Assume g_t is l.s.c. for every t in Ω_2. If for every $x_0, x_1 \in \mathbb{R}^n$ we have*

$$\int_0^1 g_t\big((1-t)x_0 + tx_1\big)\,d\mu(t) \le (1-m)f_0(x_0) + mf_1(x_1),$$

then

$$-\int_0^1 \log\Big(\int_{\mathbb{R}^n} e^{-g_t}\Big)\,d\mu(t) \le -(1-m)\log\Big(\int_{\mathbb{R}^n} e^{-f_0}\Big) - m\log\Big(\int_{\mathbb{R}^n} e^{-f_1}\Big).$$

The deduction of this result from Theorem 1.3 is as follows. Take

$$X_1 = (\Omega_1, \mu_1) = \big(\{0,1\}, (1-m)\delta_0 + m\delta_1\big), \quad X_2 = (\Omega_2, \mu_2) = ([0,1], \mu)$$

and for $(v_0, v_1) \in \mathbb{R}^n \times \mathbb{R}^n \simeq L^2(X_1, \mathbb{R}^n)$ define $A(v_0, v_1) \in L^2(X_2, \mathbb{R}^n)$ by

$$A(v_0, v_1) : t \mapsto (1-t)v_0 + tv_1.$$

Then the constant functions condition (ii) is satisfied, as $A(v,v) \equiv v$, and by the convexity of the square of the Euclidean norm on $\mathbb{R}^n$, letting $H_i = L^2(X_i, \mathbb{R}^n)$, $i = 1, 2$, we have

$$\begin{aligned}\|A(v_0, v_1)\|_{H_2}^2 &= \int_0^1 |(1-t)v_0 + tv_1|^2\,d\mu(t)\\ &\le (1-m)|v_0|^2 + m|v_1|^2 = \|(v_0, v_1)\|_{H_1}^2.\end{aligned}$$

REMARK 4.5. It is possible to prove Theorem 4.4 without assuming g_t l.s.c. and f_i bounded from below, see Claim 6.1 for a sketch of proof.

As a particular case of the previous proposition, if we take only one function $g_t \equiv g : \mathbb{R}^n \to \mathbb{R}$, and μ is a probability measure on $[0,1]$ with barycenter m and

$$\forall x_0, x_1 \in \mathbb{R}^n,\ \int_0^1 g((1-t)x_0 + tx_1)\,d\mu(t) \le (1-m)f_0(x_0) + mf_1(x_1), \tag{4.1}$$

then

$$\int_{\mathbb{R}^n} e^{-g} \geq \Big(\int_{\mathbb{R}^n} e^{-f_0}\Big)^{1-m}\Big(\int_{\mathbb{R}^n} e^{-f_1}\Big)^{m}.$$

The conclusion is independent of μ, so given f_0, f_1, g, one can try to find an optimal μ for which (4.1) holds. The classical Prékopa–Leindler inequality corresponds to $\mu = \delta_m$, and this is indeed the optimal choice when g is convex, as seen by using Jensen's inequality in (4.1). Does the above result really improve on the Prékopa–Leindler inequality when g is non-convex?

Another particular case of the previous proposition (in the case of a convex combination of two Dirac measures) is the following extension of Proposition 1.2.

PROPOSITION 4.6. *Fix $s, t, r \in [0, 1]$ and set $m := (1-r)s + rt \in [0, 1]$. Let f_0, f_1, g_0, g_1 be Borel functions from $\mathbb{R}^n$ to $\mathbb{R} \cup \{\infty\}$ such that for every $x_0, x_1 \in \mathbb{R}^n$,*

$$(1-r)g_0\big((1-s)x_0+sx_1\big)+rg_1\big((1-t)x_0+tx_1\big) \leq (1-m)f_0(x_0)+mf_1(x_1).$$

Then

$$\Big(\int_{\mathbb{R}^n} e^{-g_0(x)}\,dx\Big)^{1-r}\Big(\int_{\mathbb{R}^n} e^{-g_1(x)}\,dx\Big)^{r} \geq \Big(\int_{\mathbb{R}^n} e^{-f_0(x)}\,dx\Big)^{1-m}\Big(\int_{\mathbb{R}^n} e^{-f_1(x)}\,dx\Big)^{m}.$$

Proof. Although the result is a particular case of Theorem 4.4 (with $\mu = (1-r)\delta_s + r\delta_t$), it is better to go back to Theorem 1.6 since in the case of finitely many functions there are no technical conditions. We take $\Omega_1 = \{\{0, 1\}, (1-m)\delta_0 + m\delta_1\}$ and $\Omega_2 = \{\{0, 1\}, (1-r)\delta_0 + r\delta_1\}$, and the linear mapping $A : H_1 \to H_2$ with $H_i := L^2(\Omega_i, \mathbb{R}^n)$ defined by

$$A(x_0, x_1) := ((1-s)x_0 + sx_1, (1-t)x_0 + tx_1), \quad x_0, x_1 \in \mathbb{R}^n.$$

We have $A(v, v) = (v, v)$ for every $v \in \mathbb{R}^n$ and we note that for $x_0, x_1 \in \mathbb{R}^n$,

$$(4.2)\qquad (1-r)|(1-s)x_0 + sx_1|^2 + r|(1-t)x_0 + tx_1|^2 + ((1-r)s(1-s) + rt(1-t))|x_0 - x_1|^2 = (1-m)|x_0|^2 + m|x_1|^2,$$

and so in particular

$$(1-r)|(1-s)x_0 + sx_1|^2 + r|(1-t)x_0 + tx_1|^2 \leq (1-m)|x_0|^2 + m|x_1|^2,$$

which exactly means that $\|A\| \leq 1$. ∎

Let us mention that Proposition 1.2 (and also the above results) do not give anything more than the Brunn–Minkowski inequality when applied to the case of indicator functions, i.e., to $e^{-f_0} = \mathbf{1}_{A_0}$, $e^{-f_1} = \mathbf{1}_{A_1}$, $e^{-g_0} =$

$\mathbf{1}_{2A_0/3+A_1/3}$, $e^{-g_1} = \mathbf{1}_{A_0/3+2A_1/3}$. Indeed, by Brunn–Minkowski,

$$|2A_0/3 + A_1/3| \geq |A_0|^{2/3}|A_1|^{1/3}, \quad |A_0/3 + 2A_1/3| \geq |A_0|^{1/3}|A_1|^{2/3},$$

and in this case, the result of Proposition 1.2 is obtained by taking the product of these two inequalities. It seems that the extra information comes from applying the results to "true" functions. This is consistent with the fact that we do not know how to reach our functional inequalities using other classical proofs of the Prékopa–Leindler inequalities. For instance, it is not clear whether Proposition 1.2 can be proved using the mass transportation argument of McCann [12]; precisely, this transportation argument uses some form of "localization inside the integral" amounting to reducing the problem to sets (ellipsoids, actually) and eventually matrices.

Example 4.7 (Gaussian self-improvement and generalized τ-property). Let $\alpha \in (0, 1)$. We shall comment on Proposition 4.6 in the case $s = \alpha$, $t = 1 - \alpha = 1 - s$ and $r = 1/2$ (and so $m = 1/2$). The goal is to get improved Gaussian inequalities by exploiting identity (4.2) instead of an inequality. Let f_0, f_1 be real Borel functions on $\mathbb{R}^n$.

We start first with the case of the Prékopa–Leindler inequality and consider the following variant g of inf-convolution, defined for every $x \in \mathbb{R}^n$:

$$g(x) = \inf\{(1-\alpha)f_0(x_0) + \alpha f_1(x_1) + \alpha(1-\alpha)|x_0 - x_1|^2/2 : \\ x = (1-\alpha)x_0 + \alpha x_1\}.$$

We assume that $g(x) > -\infty$. By the following particular case of (4.2),

$$|(1-\alpha)x_0 + \alpha x_1|^2 + \alpha(1-\alpha)|x_0 - x_1|^2 = (1-\alpha)|x_0|^2 + \alpha|x_1|^2, \tag{4.3}$$

it follows that $g(x) + |x|^2/2 \leq (1-\alpha)(f_0(x_0) + |x_0|^2/2) + \alpha(f_1(x_1) + |x_1|^2/2)$ whenever $x = x_\alpha := (1-\alpha)x_0 + \alpha x_1$, and by Prékopa–Leindler we obtain

$$\int_{\mathbb{R}^n} e^{-g}\, d\gamma_n \geq \Big(\int_{\mathbb{R}^n} e^{-f_0}\, d\gamma_n\Big)^{1-\alpha}\Big(\int_{\mathbb{R}^n} e^{-f_1}\, d\gamma_n\Big)^{\alpha},$$

where γ_n is the standard Gaussian measure $\gamma_{n,1}$ on $\mathbb{R}^n$. This infimal convolution inequality ensures the τ-property from [10] for the Gaussian measure.

For $\alpha, \beta, \lambda \in (0, 1)$, we may generalize g as

$$g_{\alpha,\beta,\lambda}(x) = \inf\{(1-\beta)f_0(x_0) + \beta f_1(x_1) + \lambda\alpha(1-\alpha)|x_0 - x_1|^2 : \\ x = (1-\alpha)x_0 + \alpha x_1\},$$

the former g being equal to $g_{\alpha,\alpha,1/2}$. The Prékopa–Leindler inequality does not seem to apply to values of the parameters other than triples of the form $(\alpha, \alpha, 1/2)$. Combining

$$g_{\alpha,\beta,\lambda}(x_\alpha) \leq (1-\beta)f_0(x_0) + \beta f_1(x_1) + \lambda\alpha(1-\alpha)|x_0 - x_1|^2$$

and the corresponding inequality for $g_{1-\alpha,1-\beta,1-\lambda}(x_{1-\alpha})$, we deduce, with $g_0 = g_{\alpha,\beta,\lambda}$ and $g_1 = g_{1-\alpha,1-\beta,1-\lambda}$, that

$$g_0(x_\alpha) + g_1(x_{1-\alpha}) \le f_0(x_0) + f_1(x_1) + \alpha(1-\alpha)|x_0 - x_1|^2.$$

The identity (4.2) in the case $s = \alpha$, $t = 1-\alpha$ and $r = 1/2$ (thus $m = 1/2$) can be rewritten as

$$|(1-\alpha)x_0 + \alpha x_1|^2 + |\alpha x_0 + (1-\alpha)x_1|^2 + 2\alpha(1-\alpha)|x_0 - x_1|^2 = |x_0|^2 + |x_1|^2,$$

and so we find

$$g_0(x_\alpha) + |x_\alpha|^2/2 + g_1(x_{1-\alpha}) + |x_{1-\alpha}|^2/2 \le f_0(x_0) + |x_0|^2/2 + f_1(x_1) + |x_1|^2/2.$$

Therefore, by Proposition 4.6 with $s = \alpha$, $t = 1-\alpha$ and $r = 1/2 = m$, we arrive at the following generalized infimal convolution inequality:

$$\Big(\int_{\mathbb{R}^n} e^{-g_{\alpha,\beta,\lambda}}\, d\gamma_n\Big)\Big(\int_{\mathbb{R}^n} e^{-g_{1-\alpha,1-\beta,1-\lambda}}\, d\gamma_n\Big) \ge \Big(\int_{\mathbb{R}^n} e^{-f_0}\, d\gamma_n\Big)\Big(\int_{\mathbb{R}^n} e^{-f_1}\, d\gamma_n\Big).$$

Example 4.8 (Exotic situation). All our examples so far of linear maps A are of "convex type", meaning that

$$\int_{\Omega_2} \varphi\big(A(\alpha)(t)\big)\, d\mu_2(t) \le \int_{\Omega_1} \varphi(\alpha(s))\, d\mu_1(s)$$

for every convex function φ on $\mathbb{R}^n$, while the norm condition (i) on A ensures this property for $\varphi(x) = k|x|^2$ only, $k \ge 0$. Here is a "non-convex" example, with $\Omega_1 = \Omega_2 = \Omega = \{0,1\}$ and $\mu_1 = \mu_2 = \delta_0 + \delta_1$. The space $L^2(\Omega, \mathbb{R}^n)$ is equal to $\mathbb{R}^n \times \mathbb{R}^n$, the matrix A can be represented by blocks of size $n \times n$,

$$A = \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}.$$

By the constant functions condition (ii), the image of the constant function equal to $x \in \mathbb{R}^n$ is $(A_1x + B_1x, C_1x + D_1x) = (x,x)$, which is equivalent to $A_1 + B_1 = C_1 + D_1 = I_n$, the identity matrix. One can thus write

$$A = \begin{pmatrix} I - B & B \\ C & I - C \end{pmatrix}.$$

Since $\|A\| \le 1$, each block must have norm ≤ 1, and this implies that the diagonal coefficients of B, C are in $[0,1]$. The condition $\|A\| \le 1$ means that $A^*A \le I_{2n}$, which translates to

$$\begin{pmatrix} B^*B + C^*C & -B^*B - C^*C \\ -B^*B - C^*C & B^*B + C^*C \end{pmatrix} \le \begin{pmatrix} B + B^* & -B - C^* \\ -B^* - C & C + C^* \end{pmatrix}.$$

In the simpler case $C = B$, this amounts to $2B^*B \le B + B^*$, or $\|Bx\|^2 \le Bx \cdot x$, for every $x \in \mathbb{R}^n$. For an elementary explicit example, take $n = 2$

and

$$B = \begin{pmatrix} b & \varepsilon \\ -\varepsilon & b \end{pmatrix}$$

with $b \in (0,1)$ and $b^2 + \varepsilon^2 \le b$. We see that $B = bI_2 - \varepsilon R$, with R the rotation by $\pi/2$ in the plane $\mathbb{R}^2$. Then, with $b = 1/3$ and $\varepsilon = \sqrt{2}/3$, if we know that

$$g_0\Big(\frac{2}{3}x_0 + \frac{1}{3}x_1 - \frac{\sqrt{2}}{3}R(x_1 - x_0)\Big) + g_1\Big(\frac{1}{3}x_0 + \frac{2}{3}x_1 + \frac{\sqrt{2}}{3}R(x_1 - x_0)\Big) \le f_0(x_0) + f_1(x_1)$$

for all $x_0, x_1 \in \mathbb{R}^n$, we get the same conclusion as in Proposition 1.2.

REMARK 4.9. Following the previous construction, it is natural to ask if we can construct an "exotic" Prékopa–Leindler situation. The answer is no. Let $\alpha \in [0,1]$. The only $n \times n$ (real) matrix B such that $|(I_n - B)x + By|^2 \le (1-\alpha)|x|^2 + \alpha|y|^2$ for all $x, y \in \mathbb{R}^n$ is $B = \alpha I_n$ (try $x = u + t\alpha v$, $y = u - t(1-\alpha)v$ and $t \to 0$). In other words, there is no "exotic example" in the Prékopa–Leindler case. More generally, if $B_1, \dots, B_k$ are $n \times n$ matrices such that $\sum_{j=1}^k B_j = I_n$ and $|\sum_{j=1}^k B_j x_j|^2 \le \sum_{j=1}^k \alpha_j |x_j|^2$ with $\alpha_j \ge 0$ and $\sum_{j=1}^k \alpha_j = 1$, then $B_j = \alpha_j I_n$ for $j = 1, \dots, k$.

5. Generalized Brascamp–Lieb and reverse Brascamp–Lieb inequalities. By slightly modifying Setting 1, it is possible to recover and extend, almost for free, the Brascamp–Lieb inequalities and their reverse forms. Let us mention that it has been known for some time that the Brascamp–Lieb inequalities can be recovered using Borell's technique (see e.g. [9]).

If the functions $\{f_s\}$ are defined on $\mathbb{R}^m$ and $\{g_t\}$ on $\mathbb{R}^n$, with $m \ne n$, then the linear operator A of Setting 1 should now act from $L^2(X_1, \mathbb{R}^m)$ to $L^2(X_2, \mathbb{R}^n)$ and the constant functions condition (ii) has to be revised. We shall do it by using projections from the larger space, $\mathbb{R}^m$ or $\mathbb{R}^n$, onto the smaller. To be precise, our *projections* are adjoint to isometries from the smaller space into the larger. For example, if $n < m$ and if T is an isometry from $\mathbb{R}^n$ into $\mathbb{R}^m$, its adjoint $Q = T^*$ is a mapping from $\mathbb{R}^m$ onto $\mathbb{R}^n$ such that $QQ^* = \mathrm{Id}_{\mathbb{R}^n}$, and Q^*Q is the orthogonal projection of $\mathbb{R}^m$ onto the range of T. Then, for $v \in \mathbb{R}^m$, we can compare the image $A(s \mapsto v)$ of the constant function $\Omega_1 \ni s \mapsto v$ with the constant function $\Omega_2 \ni t \mapsto Qv$. Actually, the new setting will be notably more complicated, introducing a *family* of projections $Q(t)$, $t \in \Omega_2$, and comparing the image $A(s \mapsto v)$ with the function $t \mapsto Q(t)v$. We can also view the new setting as giving a measurable family of n-dimensional subspaces $X(t)$ of $\mathbb{R}^m$, parameterized by

$T(t) : \mathbb{R}^n \to X(t)$, and $Q(t) = T(t)^*$ being the composition of the orthogonal projection $\pi(t)$ of $\mathbb{R}^m$ onto $X(t)$ with the inverse map of $T(t)$.

Except for what concerns the linear mapping A, the modifications are straightforward, and will be just indicated without rewriting completely the modified assumption.

SETTING 2. The definition of the measure spaces $X_1 = (\Omega_1, \Sigma_1, \mu_1)$, $X_2 = (\Omega_2, \Sigma_2, \mu_2)$ is as in Setting 1, in particular, μ_1 and μ_2 are finite measures. Two integers $m, n \geq 1$ are given, and the linear operator A acts now as

$$A : L^2(X_1, \mathbb{R}^m) \to L^2(X_2, \mathbb{R}^n)$$

with $\|A\| \leq 1$, where the norms are computed with respect to the Euclidean norm $|\cdot|$ on $\mathbb{R}^m$, $\mathbb{R}^n$ and to the measures μ_1 and μ_2, respectively. The constant functions condition is modified as follows:

- $(\mathrm{ii})_1$ if $m \leq n$, then there exists a Σ_1-measurable family $\Omega_1 \ni s \mapsto P(s)$ of projections $P(s)$ from $\mathbb{R}^n$ onto $\mathbb{R}^m$ such that, for every vector $w_0 \in \mathbb{R}^n$, $A(s \mapsto P(s)w_0)(t) = w_0$ for μ_2-almost every $t \in \Omega_2$,
- $(\mathrm{ii})_2$ if $m \geq n$, then there exists a Σ_2-measurable family $\Omega_2 \ni t \mapsto Q(t)$ of projections $Q(t)$ from $\mathbb{R}^m$ onto $\mathbb{R}^n$ such that, for every vector $v_0 \in \mathbb{R}^m$, $A(s \mapsto v_0)(t) = Q(t)v_0$ for μ_2-almost every $t \in \Omega_2$.

Note that formally, $(\mathrm{ii})_2$ is the adjoint situation to $(\mathrm{ii})_1$. Observe that $(\mathrm{ii})_1$ implies that every $w_0 \in \mathbb{R}^n$ can be reconstructed from the family of projections $s \mapsto P(s)w_0 \in \mathbb{R}^m$.

For the conditions on $\{f_s\}$, $\{g_t\}$, we need only replace $\mathbb{R}^n$ by $\mathbb{R}^m$ for what concerns f_s in the measurability condition (iii) and in the semi-integrability condition (iv).

We arrive at the following extension of Theorem 1.3.

THEOREM 5.1. *Under Setting* 2, *assume additionally that f_s, g_t are nonnegative with g_t lower semicontinuous on $\mathbb{R}^n$, for some $\varepsilon_0 > 0$ we have*

$$\int_{\Omega_1} \log^-\Big(\int_{\mathbb{R}^m} \exp(-f_s(x) - \varepsilon_0|x|^2)\,dx\Big)\,d\mu_1(s) < \infty, \tag{5.1}$$

and the measures μ_1, μ_2 are such that $m \cdot \mu_1(\Omega_1) = n \cdot \mu_2(\Omega_2) < \infty$. If for every $\alpha \in L^2(X_1, \mathbb{R}^m)$ we have

$$\int_{\Omega_2} g_t\big((A\alpha)(t)\big)\,d\mu_2(t) \leq \int_{\Omega_1} f_s(\alpha(s))\,d\mu_1(s), \tag{5.2}$$

then

$$\int_{\Omega_2} -\log\Big(\int_{\mathbb{R}^n} e^{-g_t}\Big)\,d\mu_2(t) \leq \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^m} e^{-f_s}\Big)\,d\mu_1(s).$$

Not only does the argument for Theorem 5.1 follow the proof of Theorem 1.3, but in its Gaussian version, Theorem 5.1 is already contained in Proposition 2.4. Indeed, after reducing to the Gaussian case, we further approximate the functions f_s and g_t by inf-convolution, in order to be in a position to apply the proposition below, just as we did when proving Theorem 1.3. The assumption $m\cdot\mu_1(\Omega_1) = n\cdot\mu_2(\Omega_2)$ is needed to pass to the limit from the Gaussian case for P_T, as $T \to \infty$, in respective dimensions m and n.

For the Gaussian version below, we will not have much to add, except maybe for case $(\mathrm{ii})_2$ of Setting 2. So Proposition 2.4 almost includes all possible geometric situations around the Prékopa–Leindler inequality, including these generalized Brascamp–Lieb inequalities.

Proposition 5.2. *Under Setting* 2, *assume that* f_s, g_t *are continuous* ≥ 0 *on* $\mathbb{R}^m$ *and* $\mathbb{R}^n$, *respectively* (*or bounded from below as in Remark* 1.4), *and* f_s *satisfies the exponential bound* (2.6) *on* $\mathbb{R}^m$. *Then the basic assumption* (5.2) *implies that*

$$\int_{\Omega_2} -\log\Big(\int_{\mathbb{R}^n} e^{-g_t(y)}\,d\gamma_{n,\tau}(y)\Big)\,d\mu_2(t) \leq \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^m} e^{-f_s(x)}\,d\gamma_{m,\tau}(x)\Big)\,d\mu_1(s). \tag{5.3}$$

As before, in the case where the measures μ_1 and μ_2 have finite support, we can remove all the additional assumptions on f_s and g_t.

Proof of Proposition 5.2. We want to apply Proposition 2.4 but we have to deal with the fact that the dimensions m and n are now different. We shall do it by reducing to the case when both dimensions are equal to $\max(m,n)$.

In case $(\mathrm{ii})_1$, when $m \leq n$, we "extend" f_s to $\mathbb{R}^n$ by defining $f_{1,s}(y) = f_s(P(s)y)$, $y \in \mathbb{R}^n$, so that

$$\int_{\mathbb{R}^n} e^{-f_{1,s}(y)}\,d\gamma_{n,\tau}(y) = \int_{\mathbb{R}^m} e^{-f_s(x)}\,d\gamma_{m,\tau}(x). \tag{5.4}$$

For every $\alpha \in L^2(X_1,\mathbb{R}^n)$, define $\alpha_1 \in L^2(X_1,\mathbb{R}^m)$ by $\alpha_1(s) = P(s)\alpha(s)$ and set $A_1(\alpha) = A(\alpha_1)$. Then A_1 is linear from $L^2(X_1,\mathbb{R}^n)$ to $L^2(X_2,\mathbb{R}^n)$, and clearly $\|A_1\| \leq \|A\| \leq 1$. If w_0 is a fixed vector in $\mathbb{R}^n$, then $A_1(s\mapsto w_0) = A(s\mapsto P(s)w_0) = w_0$ by $(\mathrm{ii})_1$, so A_1 satisfies the constant functions condition (ii). Next, by the basic assumption of Theorem 5.1 applied to α_1, we get

$$\begin{aligned}\int_{\Omega_2} g_t\big((A_1\alpha)(t)\big)\,d\mu_2(t) &= \int_{\Omega_2} g_t\big((A\alpha_1)(t)\big)\,d\mu_2(t) \leq \int_{\Omega_1} f_s(\alpha_1(s))\,d\mu_1(s)\\ &= \int_{\Omega_1} f_s\big(P(s)\alpha(s)\big)\,d\mu_1(s) = \int_{\Omega_1} f_{1,s}(\alpha(s))\,d\mu_1(s).\end{aligned}$$

We see that A_1, $\{f_{1,s}\}$ and $\{g_t\}$ satisfy the assumptions of Proposition 2.4, including the basic assumption (1.3) on $\mathbb{R}^n$. The result follows therefore from (5.4) and from the conclusion of Proposition 2.4 for $\{f_{1,s}\}$ and $\{g_t\}$.

In case $(\mathrm{ii})_2$, when $m \ge n$, let $T(t)$ be the isometry from $\mathbb{R}^n$ into $\mathbb{R}^m$ adjoint to the projection $Q(t)$, so that $Q(t)T(t) = \mathrm{Id}_{\mathbb{R}^n}$. The function g_t is extended to $\mathbb{R}^m$ by setting $g_{1,t}(x) = g_t(Q(t)x)$ for $x \in \mathbb{R}^m$. For $\alpha \in L^2(X_1, \mathbb{R}^m)$, we set $(A_1\alpha)(t) = T(t)(A\alpha)(t)$ and get $Q(t)(A_1\alpha)(t) = (A\alpha)(t)$. We have $\|A_1\| = \|A\| \le 1$. If $\alpha_0 = v_0$ is a constant function from Ω_1 to $\mathbb{R}^m$, we know by $(\mathrm{ii})_2$ that $(A\alpha_0)(t) = Q(t)v_0$. Then $(A_1\alpha_0)(t) = T(t)(A\alpha_0)(t) = T(t)Q(t)v_0$, not equal to v_0 in general. The constant functions condition (ii) is not satisfied by A_1, but still we have

$$g_{1,t}\big(A_1(v_0+\alpha)(t)\big) = g_{1,t}\big(v_0 + (A_1\alpha)(t)\big), \quad t \in \Omega_2, \tag{5.5}$$

for every $\alpha \in L^2(X_1, \mathbb{R}^m)$. This is because $g_{1,t}(x_1) = g_{1,t}(x_2)$ when $Q(t)x_1 = Q(t)x_2$, and because by $(\mathrm{ii})_2$ we have

$$Q(t)A_1(v_0+\alpha)(t) = A(v_0+\alpha)(t) = Q(t)v_0 + (A\alpha)(t) = Q(t)\big(v_0 + (A_1\alpha)(t)\big).$$

The basic assumption (1.3) is satisfied for the families $\{f_s\}$, $\{g_{1,t}\}$ of functions on $\mathbb{R}^m$ and the mapping A_1: since $g_{1,t}((A_1\alpha)(t)) = g_t((A\alpha)(t))$, we get

$$\int_{\Omega_2} g_{1,t}\big((A_1\alpha)(t)\big)\, d\mu_2(t) = \int_{\Omega_2} g_t\big(A(\alpha)(t)\big)\, d\mu_2(t) \le \int_{\Omega_1} f_s(\alpha(s))\, d\mu_1(s).$$

We stressed during the proof of Proposition 2.4 that the equality (5.5) is enough to run the argument (see equation (2.13) and the lines around it). Finally, as before, observe that

$$\int_{\mathbb{R}^n} e^{-g_t(y)}\, d\gamma_{n,\tau}(y) = \int_{\mathbb{R}^m} e^{-g_{1,t}(x)}\, d\gamma_{m,\tau}(x),$$

and use Proposition 2.4 to get the desired conclusion. ■

Remark 5.3. The normalization $\|A\| \le 1$ can be replaced by the following assumptions. Let $\kappa := \|A\| > 0$ and assume that $\kappa^2 \cdot m \cdot \mu_1(\mathbb{R}^m) = n \cdot \mu_2(\mathbb{R}^n)$. If for every α in $L^2(X_1, \mathbb{R}^m)$, we have

$$\int_{\Omega_2} g_t((A\alpha)(t))\, d\mu_2(t) \le \kappa^2 \int_{\Omega_1} f_s(\alpha(s))\, d\mu_1(s),$$

then

$$\int_{\Omega_2} -\log\Big(\int_{\mathbb{R}^n} e^{-g_t}\Big)\, d\mu_2(t) \le \kappa^2 \int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^m} e^{-f_s}\Big)\, d\mu_1(s).$$

The reader will just look at inequality (2.14) and the one before it.

We shall now examine particular cases of Theorem 5.1. We will see that it contains both the Brascamp–Lieb and reverse Brascamp–Lieb inequalities in their geometric form.

Let us take unit vectors $u_1, \dots, u_N$ in the Euclidean space $\mathbb{R}^d$ and positive reals $c_1, \dots, c_N$ that decompose the identity of $\mathbb{R}^d$, meaning that

$$\sum_{i=1}^{N} c_i\, u_i \otimes u_i = \mathrm{Id}_{\mathbb{R}^d}. \tag{5.6}$$

This is equivalent to saying that $x\cdot x = \sum_{i=1}^{N} c_i |u_i\cdot x|^2$ for every $x \in \mathbb{R}^d$. If we consider the one-point space $E_1 = \{0\}$ equipped with the trivial probability measure ν_1, and the measure space

$$(E_2, \nu_2) = \Big(\{1, \dots, N\}, \sum_{i=1}^{N} c_i \delta_i\Big), \quad \nu_2(E_2) = \sum_{i=1}^{N} c_i = d,$$

then (5.6) is equivalent to saying that the mapping

$$U : \mathbb{R}^d \ni x \mapsto (x\cdot u_1, \dots, x\cdot u_N) \tag{5.7}$$

is an isometry from $L^2(E_1, \nu_1, \mathbb{R}^d) \simeq \mathbb{R}^d$ into $L^2(E_2, \nu_2, \mathbb{R}) \simeq \mathbb{R}^N$.

In order to recover the reverse Brascamp–Lieb inequalities, consider

$$X_1 = (E_2, \nu_2),\ m = 1, \qquad X_2 = (E_1, \nu_1),\ n = d,$$

satisfying

$$m\mu_1(\Omega_1) = \nu_2(E_2) = d = n\mu_2(\Omega_2).$$

Define $A : \mathbb{R}^N \simeq L^2(X_1, \mathbb{R}^m) \to \mathbb{R}^d \simeq L^2(X_2, \mathbb{R}^n)$ to be the linear operator

$$A(t_1, \dots, t_N) = \sum_{i=1}^{N} c_i t_i u_i, \quad t_i \in \mathbb{R}.$$

Then A is adjoint to the into isometry U of (5.7) between our L^2 spaces, and A satisfies condition (ii)_1 of Setting 2, with the family of projections $P_i v = v\cdot u_i$, $i \in \Omega_1$, because by (5.6) we have

$$A(i \mapsto P_i v) = \sum_{i=1}^{N} c_i (v\cdot u_i) u_i = v.$$

The conclusion reads as follows. If for every $t_1, \dots, t_N \in \mathbb{R}$ we have

$$g\Big(\sum_{i=1}^{N} t_i c_i u_i\Big) \le \sum_{i=1}^{N} c_i\, f_i(t_i) \quad \text{then} \quad \int_{\mathbb{R}^d} e^{-g} \ge \prod_{i=1}^{N} \Big(\int_{\mathbb{R}} e^{-f_i}\Big)^{c_i}.$$

This is the reverse Brascamp–Lieb inequality of Barthe [1] in its geometric form. Actually, in the same way, we can see that the formulation above contains the "continuous" statement given in [2] (which can be easily derived by approximation, anyway).

In order to recover the classical Brascamp–Lieb inequalities, apply the theorem with $X_1 = (E_1, \nu_1)$, $m = d$, $X_2 = (E_2, \nu_2)$, $n = 1$ and the linear operator $A : L^2(X_1, \mathbb{R}^m) \simeq \mathbb{R}^d \to L^2(X_2, \mathbb{R}^n) \simeq \mathbb{R}^N$ defined by

$A(x) = (x \cdot u_1, \ldots, x \cdot u_N)$. Then $A = U$, the into isometry (5.7) between the corresponding L^2 spaces, and it satisfies the assumption $(\text{ii})_2$ with the family of projections $Q_j(x) = x \cdot u_j$, $j \in \Omega_2 = \{1, \ldots, N\}$. The conclusion reads as follows: given $f : \mathbb{R}^d \to \mathbb{R}$ and $g_1, \ldots, g_N : \mathbb{R} \to \mathbb{R}$,

$$\forall x \in \mathbb{R}^d,\ \sum_{j=1}^{N} c_j g_j(x \cdot u_j) \le f(x) \ \Rightarrow\ \prod_{j=1}^{N} \Big(\int_{\mathbb{R}} e^{-g_j} \Big)^{c_j} \ge \int_{\mathbb{R}^d} e^{-f}.$$

This is the geometric Brascamp–Lieb inequality, usually stated with the best possible f, which is defined by equality in place of inequality above, namely

$$\int_{\mathbb{R}^d} \exp\Big(- \sum_{j=1}^{N} c_j g_j(x \cdot u_j) \Big)\, dx \le \prod_{j=1}^{N} \Big(\int_{\mathbb{R}} e^{-g_j} \Big)^{c_j}.$$

It is possible to state (and prove in the same way) a more general form of Theorem 5.1 where each f_s (resp. g_t) is defined on a different Euclidean space E_s (resp. F_t), and to deduce from it the multidimensional geometric Brascamp–Lieb inequalities. In this situation, we assume that all E_s, F_t are subspaces of a given $\mathbb{R}^\ell$. We may consider E_s to be $\mathbb{R}^{m(s)}$ and F_t to be $\mathbb{R}^{n(t)}$, with $1 \le m(s), n(t) \le \ell$, and assume that

$$\int_{\Omega_1} m(s)\, d\mu_1(s) = \int_{\Omega_2} n(t)\, d\mu_2(t).$$

The mapping A is defined on the closed subspace H of $L^2(X_1, \mathbb{R}^\ell)$ consisting of those α with $\alpha(s) \in E_s$ for every $s \in \Omega_1$. We assume that for every square integrable $\alpha \in H$, the image $A\alpha \in L^2(X_2, \mathbb{R}^\ell)$ is such that $(A\alpha)(t) \in F_t$ for every $t \in \Omega_2$. We consider projections $P(s), Q(t)$ from $\mathbb{R}^\ell$ onto E_s, F_t respectively. The constant functions condition says now that for every $w_0 \in \mathbb{R}^\ell$,

$$A\big(s \mapsto P(s) w_0\big)(t) = Q(t) w_0.$$

The proof mixes the two cases $(\text{ii})_1$ and $(\text{ii})_2$ of the proof of Theorem 5.1. We leave this to the reader.

6. Technicalities

Claim 6.1. *In Theorem 4.4, we can remove the assumptions that g_t is l.s.c. and that $f_i \ge 0$ (or bounded from below), $i = 0, 1$.*

Sketch of proof. As in the proof of Theorem 1.3, we begin by replacing the Lebesgue measure in $\int_{\mathbb{R}^n} e^{-f_i(x)}\, dx$ and $\int_{\mathbb{R}^n} e^{-g_t(x)}\, dx$ by a Gaussian probability measure γ. The justification is the same here for the g_t side, but is easier for f_i, $i = 0, 1$, by a simple application of monotone convergence. Next, as in the proof of Theorem 1.6, one can replace g_t by $g_{t,N} = \min(g_t, N)$ and f_i by $f_{i,N} = \max(f_i, -N)$: since $g_{t,N}$ increases to g_t as $N \to \infty$, the argument for

$$\int_{\Omega_2} \log\Big(\int_{\mathbb{R}^n} e^{-g_{t,N}}\, d\gamma\Big)\, d\mu_2(t) \xrightarrow[N]{} \int_{\Omega_2} \log\Big(\int_{\mathbb{R}^n} e^{-g_t}\, d\gamma\Big)\, d\mu_2(t)$$

is the same as at the end of the proof of Theorem 1.3, and for $f_{i,N}$, decreasing to f_i, the reason is monotone convergence again in $\int_{\mathbb{R}^n} e^{-f_{i,N}}\, d\gamma$, $i = 0, 1$. Then $0 \le g_{t,N} \le N$ and $f_{0,N}, f_{1,N} \ge -N$. As in the proof of Theorem 1.6, one can then assume that $f_{0,N}, f_{1,N} \le 2N/\beta$ with $\beta = \min(m, 1-m)$. Finally, approximation by convolution reduces to the case of bounded continuous functions, and one can conclude by applying Proposition 2.4. ∎

Proof of Lemma 2.3. We may assume that $f \ge 0$. For some integer $N \ge 0$, let $\varphi(x) = \min(f(x), N) \ge 0$. As in (2.2), define φ_r by $e^{-\varphi_r} = P_{T-r}(e^{-\varphi})$, $0 \le r \le T$ and let $\Phi_0 = \varphi_0(0)$ and $F_0 = -\log P_T(e^{-f})(0) \ge \Phi_0 > 0$. Note that $\varphi \ge 0$ implies $\varphi_r \ge 0$. Since φ is bounded and continuous, by Lemma 2.1 we may consider the optimal martingale corresponding to φ,

$$M_r = \varphi_r\Big(B_r + \int_0^r u_\rho\, d\rho\Big) + \frac{1}{2}\int_0^r |u_\rho|^2\, d\rho = M_0 - \int_0^r u_\rho \cdot dB_\rho, \quad 0 \le r \le T,$$

with $M_0 = \Phi_0$, $u_r = -\nabla\varphi_r(X_r)$ and $X_r = B_r + \int_0^r u_\rho\, d\rho$. Define the square function $(S_r)_{0\le r\le T}$ of the martingale $(M_r - M_0)_{0\le r\le T}$ by

$$S_r = \Big(\int_0^r |u_\rho|^2\, d\rho\Big)^{1/2}.$$

We know by (2.4) that S_r, X_r and M_r converge almost surely and in L^2 to S_T, X_T and M_T. Observe that $|u_r|$, S_r and M_r are bounded random variables for each fixed $r < T$. Consider the exponential martingale $e^{\lambda M_r - \lambda^2 S_r^2/2}$ for $\lambda = 1/2$ and $0 \le r < T$, namely

$$\exp(M_r/2 - S_r^2/8) = \exp\big(\varphi_r(X_r)/2 + S_r^2/4 - S_r^2/8\big) = \exp\big(\varphi_r(X_r)/2 + S_r^2/8\big).$$

By Fatou and the martingale property, we have

$$\begin{aligned}\mathbb{E}\exp(S_T^2/8) &\le \mathbb{E}\exp\big(\varphi(X_T)/2 + S_T^2/8\big)\\ &\le \lim_{r\to T} \mathbb{E}\exp\big(\varphi_r(X_r)/2 + S_r^2/8\big) = e^{\Phi_0/2}.\end{aligned}$$

Let $Y_T = T^{-1/2}|X_T| \le T^{-1/2}|B_T| + S_T$ and observe that in $\mathbb{R}^n$, one has $\mathbb{E}\, e^{|B_1|^2/4} = 2^{n/2}$. For every $\lambda > 0$, using Cauchy–Schwarz and the inequality $2\sigma\tau \le c\sigma^2 + \tau^2/c$ for $c, \sigma, \tau > 0$, we can write

$$\begin{aligned}(\mathbb{E}\, e^{\lambda Y_T})^2 &\le (\mathbb{E}\, e^{2\lambda T^{-1/2}|B_T|})(\mathbb{E}\, e^{2\lambda S_T}) = (\mathbb{E}\, e^{2\lambda |B_1|})(\mathbb{E}\, e^{2\lambda S_T})\\ &\le e^{4\lambda^2}\, \mathbb{E}\, e^{|B_1|^2/4} e^{8\lambda^2}\, \mathbb{E}\, e^{S_T^2/8} = 2^{n/2} e^{12\lambda^2}\, \mathbb{E}\, e^{S_T^2/8}\\ &\le 2^{n/2} e^{12\lambda^2 + \Phi_0/2} \le e^{12\lambda^2 + (F_0+n)/2}.\end{aligned} \tag{6.1}$$

By the exponential bound (2.5) we have

$$\delta := \mathbb{E}\big(f(X_T) - \varphi(X_T)\big) \le \mathbb{E}\big(\mathbf{1}_{\{f(X_T)>N\}} f(X_T)\big) \le a\,\mathbb{E}(\mathbf{1}_{\{M_T \ge N\}} e^{b|X_T|}).$$

Using (6.1) with $\lambda = 2b\sqrt{T}$, by Cauchy–Schwarz and Markov we get

$$\delta \le a\sqrt{(\mathbb{E}\, M_T)/N}\sqrt{e^{24b^2T+(F_0+n)/4}} \le a\sqrt{F_0/N}\, e^{12b^2T+(F_0+n)/8}. \tag{6.2}$$

This proves that for any given $\varepsilon > 0$, when N is so large that $\delta < \varepsilon$, the optimal drift u for φ gives

$$\begin{aligned}\mathbb{E}\Big[f\Big(B_T + \int_0^T u_r\, dr\Big) + \tfrac{1}{2}\int_0^T u_r^2\, dr\Big] - \varepsilon \\ \le \mathbb{E}\Big[\varphi\Big(B_T + \int_0^T u_r\, dr\Big) + \tfrac{1}{2}\int_0^T u_r^2\, dr\Big] = \Phi_0 \le F_0.\end{aligned}$$

Conversely, since $1 \ge P_T e^{-\min(f,N)} \searrow P_T e^{-f}$ as $N \to \infty$, we may start by choosing N large enough, so that $F_0 < \Phi_0 + \varepsilon$. For any drift u in D_2, the inequality case for φ gives

$$\begin{aligned}F_0 - \varepsilon < \Phi_0 &\le \mathbb{E}\Big[\varphi\Big(B_T + \int_0^T u_\rho\, d\rho\Big) + \tfrac{1}{2}\int_0^T |u_\rho|^2\, d\rho\Big] \\ &\le \mathbb{E}\Big[f\Big(B_T + \int_0^T u_\rho\, d\rho\Big) + \tfrac{1}{2}\int_0^T |u_\rho|^2\, d\rho\Big].\end{aligned}$$

This implies the inf formula (2.1) for f. Observe that this final part does not use any upper bound on f, thus proving the last claim of Lemma 2.3. ∎

LEMMA 6.2. *Assume that μ is a finite measure on (Ω, Σ), ν a probability measure on $\mathbb{R}^n$, $(s, x) \mapsto f_s(x)$ a $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable function, and assume that $s \mapsto \log(\int_{\mathbb{R}^n} e^{-f_s}\, d\nu)$ is μ-integrable on Ω. For every $\varepsilon > 0$, there exists a continuous function $\psi \ge 0$ on $\mathbb{R}^n$, tending to ∞ at infinity, such that $0 \le \psi(x) \le |x|^2$ and*

$$\int_\Omega -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)-\psi(x)}\, d\nu(x)\Big) d\mu(s) \le \int_\Omega -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\, d\nu(x)\Big) d\mu(s) + \varepsilon.$$

Proof. Consider $\chi(x) = |x|^2/(1 + |x|^2)$ for $x \in \mathbb{R}^n$. Note that $0 \le \chi \le 1$, χ tends to 1 at infinity, and for every x, $\chi(x/k)$ decreases to 0 as $k \to \infty$. Write

$$e^{-F(s)} := \int_{\mathbb{R}^n} e^{-f_s(x)}\, d\nu(x), \ \ s \in \Omega, \qquad I := \int_\Omega F(s)\, d\mu(s) < \infty,$$

and when $F(s) < \infty$, let $u_k(s)$ be defined by

$$e^{-F(s)-u_k(s)} := \int_{\mathbb{R}^n} e^{-f_s(x)-\chi(x/k)}\, d\nu(x) \xrightarrow[k]{} e^{-F(s)}.$$

Then $0 \le u_k(s) \le 1$ and u_k converges μ-almost everywhere to 0; since μ is finite, we can find $k_1 > 1$ such that

$$\int_{\Omega} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)-\chi(x/k_1)}\,d\nu(x)\Big)\,d\mu(s) = \int_{\Omega}\big(F(s)+u_{k_1}(s)\big)\,d\mu(s) < I + \frac{\varepsilon}{2}.$$

In the same way, we can find by induction an increasing sequence (k_j) of integers such that for every integer $p \ge 1$,

$$\int_{\Omega} -\log\Big(\int_{\mathbb{R}^n} \exp\Big(-f_s(x) - \sum_{j=1}^{p}\chi(x/k_j)\Big)\,d\nu(x)\Big)\,d\mu(s) < I + \sum_{j=1}^{p} 2^{-j}\varepsilon,$$

and we can check that

$$\psi(x) = \sum_{j=1}^{\infty}\chi(x/k_j) = \Big(\sum_{j=1}^{\infty}\frac{1}{k_j^2+|x|^2}\Big)|x|^2 \le \Big(\sum_{j=1}^{\infty}k_j^{-2}\Big)|x|^2$$

does the job (note that $\sum_{j\ge 1} k_j^{-2} \le \pi^2/6 - 1 < 1$ because $k_1 > 1$). ∎

Claim 6.3. *The almost optimal drifts $\{U(s)\}_{s\in\Omega_1}$ in (2.8) can be chosen to be Σ_1-measurable with respect to $s \in \Omega_1$.*

Proof. In the proof of Proposition 2.4 we may start, using Lemma 6.2, by replacing $f_s(x) \ge 0$ by $f_s(x) + \psi(x)$, where ψ tends to ∞ at infinity, without changing much the value of $\int_{\Omega_1} -\log(\int_{\mathbb{R}^n} e^{-f_s(x)}\,d\gamma(x))\,d\mu_1(s)$ and without destroying the exponential bound (2.6) for $f_s(x)+\psi(x)$. We shall thus assume that $\psi(x) \le f_s(x)$. Recall that $T = \tau > 0$ is fixed.

Let $\delta > 0$ be given and let $e^{-f_{s,r}} = P_{T-r}e^{-f_s}$ as in (2.2). We can find a partition of Ω_1 into countably many subsets $A_k \in \Sigma_1$, $k \in \mathbb{N}$, such that on A_k, the exponential bound (2.6) for f_s is uniform,

$$f_s(x) \le a_k e^{b_k|x|}, \quad \text{and} \quad f_{s,0}(0) \le F_k, \qquad s \in A_k.$$

It is enough to prove the measurability on each A_k separately. We can choose N_k so large that

$$a_k\sqrt{F_k/N_k}\, e^{12b_k^2T+(F_k+n)/8} < \delta.$$

By (6.2), we may replace f_s, $s \in A_k$, by $\varphi_s = \min(f_s, N_k)$ with an error $< \delta$ in the estimate of $-\log\big(P_Te^{-f_s}\big)(0)$. The "almost optimal drift" $U(s) \in D_2$ for f_s is chosen equal to the optimal drift for φ_s.

After this reduction, φ_s is "constant at infinity" since $f_s \ge \psi$, hence $A_k \ni s \mapsto \varphi_s$ is a map to the separable Banach space Z of continuous functions on $\mathbb{R}^n$ tending to a limit at infinity, equipped with the sup norm, and by the measurability assumption (iii), $s \mapsto \varphi_s(x)$ is Σ_1-measurable for every $x \in \mathbb{R}^n$. By a classical theorem of Lusin—that Hausdorff topologies weaker than a Polish topology have the same Borel σ-algebra—this implies that $s \mapsto \varphi_s$ is Σ_1-measurable from Ω_1 to Z. Next, consider the four mappings sending $f \in Z$ to e^{-f}, to $(f_r)_{0\le r\le T}$ defined by (2.2), and, for any given

$\varepsilon \in (0,T)$, to $(\nabla f_r)_{0\le r\le T-\varepsilon}$ and $(\nabla^2 f_r)_{0\le r\le T-\varepsilon}$. On every bounded subset B of Z, these mappings are Lipschitz from B, equipped with the norm of Z, respectively to Z, $C_b([0,T]\times\mathbb{R}^n)$, or $C_b([0,T-\varepsilon]\times\mathbb{R}^n,\mathbb{R}^p)$, $p=n,n^2$. The Lipschitz constants depend on B, and on ε for the last two. The map sending $f\in Z$ to the solution $X_{r,f}$ of the stochastic differential equation (2.3) is continuous, in the precise sense that for every bounded subset B of Z, there is $\kappa=\kappa(B,\varepsilon)$ such that for every $\omega\in E$,

$$\sup_{0\le r\le T-\varepsilon}\bigl|X_{r,f}(\omega)-X_{r,g}(\omega)\bigr|\le e^{\kappa T}\|f-g\|_\infty,\quad f,g\in B.$$

Indeed, for every fixed ω, we have the deterministic differential equation in the r variable

$$X'_{r,f}(\omega)-X'_{r,g}(\omega)=-\nabla f_r(X_{r,f}(\omega))+\nabla g_r(X_{r,g}(\omega)).$$

Writing $X'_{r,f}-X'_{r,g}$ as $-(\nabla f_r(X_{r,f})-\nabla f_r(X_{r,g}))-(\nabla f_r(X_{r,g})-\nabla g_r(X_{r,g}))$, using the Lipschitz properties mentioned above (implying that the second derivatives of f_r are uniformly bounded for $f\in B$), we see that for every fixed $\theta>0$, the function $D(r)=(|X_{r,f}(\omega)-X_{r,g}(\omega)|^2+\theta^2)^{1/2}$ satisfies on $[0,T-\varepsilon]$ a differential inequality of the form $D'\le\kappa(D+\|f-g\|_\infty)$ with $D(0)=\theta$. It follows that the "almost optimal" drift $U:(s,r)\mapsto -\nabla\varphi_{s,r}(X_{s,r})$ is Σ_1-measurable. ∎

LEMMA 6.4. *Under Setting* 1, *assume that* $f_s\ge 0$ *and for some* $\varepsilon_0>0$, *the function* $\Omega_1\ni s\mapsto -\log(\int_{\mathbb{R}^n}e^{-f_s(x)-\varepsilon_0|x|^2}\,dx)$ *is* μ_1*-integrable. Then there exists* $\alpha\in L^2(X_1,\mathbb{R}^n)$ *such that*

$$\int_{\Omega_1} f_s(\alpha(s))\,d\mu_1(s)<\infty.$$

The inf-convolution $f_{s,k}$ *in* (3.4) *is universally measurable, and there exists a* μ_1*-negligible set* $N\in\Sigma_1$ *such that* $(s,x)\mapsto f_{s,k}(x)$ *is* $\Sigma_1\otimes\mathcal{B}_{\mathbb{R}^n}$*-measurable on* $(\Omega_1\setminus N)\times\mathbb{R}^n$. *It is possible to find for every* $\alpha\in L^2(X_1,\mathbb{R}^n)$ *and every* $\varepsilon>0$ *a measurable selection* $u(s)$ *such that*

$$f_s\bigl(\alpha(s)+u(s)\bigr)+k|u(s)|^2<f_{s,k}(\alpha(s))+\varepsilon,\quad s\in\Omega_1\setminus N.$$

Proof. Let

$$e^{-F(s)}=\kappa\int_{\mathbb{R}^n}e^{-f_s(x)-\varepsilon_0|x|^2}\,dx=\int_{\mathbb{R}^n}e^{-f_s(x)-\varepsilon_0|x|^2/2}\,d\nu(x),$$

where $\kappa=(2\pi/\varepsilon_0)^{-n/2}$ is chosen so that $\kappa\int_{\mathbb{R}^n}e^{-\varepsilon_0|x|^2/2}\,dx=1$, making ν a probability measure. Consider the Borel set

$$C=\{(s,x):f_s(x)+\varepsilon_0|x|^2/2<F(s)+1\}$$

in the Polish space $\Omega_1\times\mathbb{R}^n$. The projection of this set on Ω_1 is an analytic set, and contains the Borel set $B:=\{F<\infty\}$. We have $\mu_1(\Omega_1\setminus B)=0$. By the Jankoff–von Neumann selection theorem (see [4, Theorem 6.9.1]), we can find a universally measurable section $\sigma(s) = (s, \alpha(s)) \in C$ defined on B. We get

$$f_s(\alpha(s)) + \varepsilon_0 |\alpha(s)|^2/2 < F(s) + 1,$$

from which the conclusions follow since $f_s \geq 0$.

Similarly, for every real c, the set $\{(s,x) : f_{s,k}(x) < c\}$ is the projection of the Borel set

$$\{(s,x,u) : f_s(x+u) + k|u|^2 < c\}.$$

It follows that for every fixed x, the function $s \mapsto f_{s,k}(x)$ is universally measurable, hence μ_1-equivalent to a Borel function on Ω_1. Let D be a countable dense set in $\mathbb{R}^n$. For every $d \in D$, there is a negligible set $N_d \in \Sigma_1$ such that $s \mapsto f_{s,k}(d)$ is Borel outside N_d. Let $N = \bigcup_d N_d$. Since $x \mapsto f_{s,k}(x)$ is continuous, we deduce that $(s,x) \mapsto f_{s,k}(x)$ is Borel on $(\Omega_1 \setminus N) \times \mathbb{R}^n$.

Let α be Borel from Ω_1 to $\mathbb{R}^n$. Then $\Omega_1 \setminus N$ is the projection of the Borel set

$$\{(s,u) : s \notin N,\ f_s(\alpha(s)+u) + k|u|^2 < f_{s,k}(\alpha(s)) + \varepsilon\}.$$

Consider as before a universally measurable section $\sigma(s) = (s, u(s))$ defined on $\Omega_1 \setminus N$. Then $u(s)$ provides the promised selection. ∎

Remark 6.5. Let (Ω, Σ) be a measurable space and let ν be a probability measure on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$. Let V be the set of bounded real $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable functions $h(s,x) = h_s(x)$ with the property that for every $\varepsilon > 0$, there exist $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable functions $\varphi(s,x) = \varphi_s(x)$ and $\psi(s,x) = \psi_s(x)$ such that for every $s \in \Omega$, $\varphi_s \leq h_s \leq \psi_s$, φ_s is u.s.c. on $\mathbb{R}^n$, ψ_s is l.s.c. and $\int_{\mathbb{R}^n} (\psi_s - \varphi_s)\, d\nu < \varepsilon$. Then V is a vector space of functions on $\Omega \times \mathbb{R}^n$, containing the constants, stable by $\sup(h_1, h_2)$ and stable under pointwise convergence of uniformly bounded sequences.

Indeed, suppose that $h_k \in V$, $k \in \mathbb{N}$, and that $h_k(s,x) \to h(s,x)$ pointwise with $|h_k(s,x)| \leq 1$. For every k, let $\psi_{s,k} \geq h_{s,k}$ be l.s.c. and $\int_{\mathbb{R}^n} (\psi_{s,k} - h_{s,k})\, d\nu < 2^{-k}$. Then, for every s, we have $\psi_{s,k} - h_{s,k} \to 0$ ν-a.e., hence $\psi_{s,k} - h_s \to 0$ ν-a.e. Next, for every s, $\chi_{s,n} = \sup_{k \geq n} \psi_{s,k}$ is l.s.c. and decreases to h_s ν-a.e. Let

$$B_n = \Big\{ s \in \Omega : \int_{\mathbb{R}^n} (\chi_n(s,x) - h(s,x))\, d\nu(x) < \varepsilon \Big\}.$$

Then $B_n \in \Sigma$ increases to Ω. Define $\psi(s,x) = \chi_n(s,x)$ when $s \in B_n \setminus B_{n-1}$, and similarly on the φ side. We see that the pointwise limit h belongs to V.

It follows that V is the space of all bounded $\mathcal{A}$-measurable functions for some sub-σ-algebra $\mathcal{A}$ of $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$. Since V contains the indicators of all products $B \times C$ with $B \in \Sigma$ and $C \in \mathcal{B}_{\mathbb{R}^n}$, it follows that $\mathcal{A} = \Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$. The result then applies to all $\mathcal{A}$-measurable functions h such that h_s is bounded for every s, by cutting Ω into pieces $B_k \in \Sigma$ where the bound of h_s is in $[k, k+1)$.

Suppose that h is bounded and ≥ 0, and $H(s) = \int_{\mathbb{R}^n} h(s,x)\,d\nu(x) > 0$ for every $s \in \Omega$. Applying the result to the function $H(s)^{-1}h(s,x)$, we find a function $\varphi(s,x)$ u.s.c. in x such that $0 \leq \varphi \leq h$ and

$$\log\Big(\int_{\mathbb{R}^n} \varphi(s,x)\,d\nu(x)\Big) > \log\Big(\int_{\mathbb{R}^n} h(s,x)\,d\nu(x)\Big) - \varepsilon$$

for every s. Applying this to $h(s,x) = e^{-f_s(x)}$, we get a l.s.c. (in x) function $\psi_s(x)$ such that $f_s \leq \psi_s$ and $-\log\big(\int_{\mathbb{R}^n} e^{-\psi_s}\,d\nu\big) \leq -\log\big(\int_{\mathbb{R}^n} e^{-f_s}\,d\nu\big) + \varepsilon$. This shows that in Theorem 1.3, f_s can be assumed to be l.s.c.

References

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

[2] F. Barthe, *A continuous version of the Brascamp–Lieb inequalities*, in: Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1850, Springer, Berlin, 2004, 53–63.

[3] B. Berndtsson, *Prekopa's theorem and Kiselman's minimum principle for plurisubharmonic functions*, Math. Ann. 312 (1998), 785–792.

[4] V. I. Bogachev, *Measure Theory*, Vol. II, Springer, Berlin, 2007.

[5] C. Borell, *Convex set functions in d-space*, Period. Math. Hungar. 6 (1975), 111–136.

[6] C. Borell, *Diffusion equations and geometric inequalities*, Potential Anal. 12 (2000), 49–71.

[7] D. Cordero-Erausquin, *On Berndtsson's generalization of Prékopa's theorem*, Math. Z. 249 (2005), 401–410.

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

[9] J. Lehec, *Short probabilistic proof of the Brascamp–Lieb and Barthe theorems*, Canad. Math. Bull. 57 (2014), 585–597.

[10] B. Maurey, *Some deviation inequalities*, Geom. Funct. Anal. 1 (1991), 188–197.

[11] B. Maurey, *Inégalité de Brunn–Minkowski–Lusternik, et autres inégalités géométriques et fonctionnelles*, Astérisque 299 (2005), exp. 928, vii, 95–113.

[12] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. 128 (1997), 153–179.

Dario Cordero-Erausquin, Bernard Maurey
Institut de Mathématiques de Jussieu-PRG
UPMC (Paris 6)
F-75252 Paris Cedex 05, France
E-mail: dario.cordero@imj-prg.fr