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

Dario Cordero-Erausquin and Bernard Maurey
Institut de Mathématiques de Jussieu-PRG, UPMC (Paris 6)

December, 2015

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 A, B Borel subsets of $\mathbb{R}^n$ and $t \in [0,1]$, the volume of the Minkowski combination

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

verifies 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, that 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 (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 $x_0, x_1 \in \mathbb{R}^n$,*

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

Then

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

Accepting the value $+\infty$ enables us to reach directly indicator functions $\mathbf{1}_E = e^{-f_E}$, by letting f_E be 0 on E and $+\infty$ outside. However, we can restrict ourselves to Borel functions that are not Lebesgue-almost everywhere equal to $+\infty$. 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 semi-group 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 2. *Let $f_0, f_1,\ g_0, g_1$ be four Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{+\infty\}$ such that, for every $x_0, x_1 \in \mathbb{R}^n$,*

$$g_0(2x_0/3 + x_1/3) + g_1(x_0/3 + 2x_1/3) \leq f_0(x_0) + f_1(x_1). \tag{1}$$

Then

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

We will see that it is rather natural to arrive to 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 the functions f_i's at some points into the values of two functions g_j's at some related points. This is not interesting in the case the functions f_i's take only the values $+\infty$ and 0, and the functional inequality above does not give anything new when applied to the case where the functions e^{-f_i}'s are indicators of sets, as we will explain in the Section 4 below.

The previous proposition and its proof suggest actually more general inequalities. Writing the conclusion as $\sum_j -\log\left(\int_{\mathbb{R}^n} e^{-g_j}\right) \leq \sum_i -\log\left(\int_{\mathbb{R}^n} e^{-f_i}\right)$, 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), the two values x_0, x_1 at the right 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 $y = \{y(t)\}_{t\in[0,1]}$ at the left will be obtained from a linear transformation A acting on this data x, i.e. $y = Ax$. So, roughly speaking, under appropriate assumptions on the operator A, we may expect that if for "all" $x = \{x(s)\}$ we have

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

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 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 at least one of $F^+ = \max(F,0)$ or $F^- = \max(-F,0)$ is μ-integrable. The integral of F takes then a definite value in $[-\infty,+\infty]$. This assumption is needed for $F(s) = -\log\Big(\int_{\mathbb{R}^n} e^{-f_s}\Big)$ in order to make sense of the preceding integrals.

We introduce the following abstract setting.

Setting 3. 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, $i=1,2$. We assume that Ω_1 is a Polish topological space and Σ_1 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 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 on the constant vector valued functions, *i.e.*, for any $v_0 \in \mathbb{R}^n$, the constant function $\Omega_1 \ni s \to v_0$ is sent by A to the constant function $\Omega_2 \ni t \to v_0$.

- Two families $\{f_s\}_{s\in\Omega_1}$ and $\{g_t\}_{t\in\Omega_2}$ of Borel functions from $\mathbb{R}^n$ to $\mathbb{R}\cup\{+\infty\}$ satisfying

  (iii) the functions $(s,x) \to f_s(x)$ and $(t,x) \to g_t(x)$ are measurable with respect to the σ-algebras $\Sigma_i \otimes \mathcal{B}_{\mathbb{R}^n}$, $i=1,2$, respectively, the notation $\mathcal{B}_{\mathbb{R}^n}$ being for the Borel σ-algebra of $\mathbb{R}^n$,

  (iv) the functions

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

  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 on the functions 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 the Remark 5 and as shown in Corollary 7 below. So, our most abstract version of the Prékopa–Leindler inequality reads as follows.

Theorem 4. *Under the Setting 3 with μ_1 and μ_2 having the same finite mass, $\mu_1(\Omega_1) = \mu_2(\Omega_2) < +\infty$, we make the additional assumptions on the functions:*

- *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 $\varepsilon_0 > 0$, we have that*

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

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

$$\int_{\Omega_2} g_t((A\alpha)(t))\, d\mu_2(t) \le \int_{\Omega_1} f_s(\alpha(s))\, d\mu_1(s), \tag{3}$$

it follows that

$$\int_{\Omega_2} -\log \left(\int_{\mathbb{R}^n} e^{-g_t(x)}\, dx \right) d\mu_2(t) \le \int_{\Omega_1} -\log \left(\int_{\mathbb{R}^n} e^{-f_s(x)}\, dx \right) d\mu_1(s). \tag{4}$$

Remark 5. 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" (3) and the conclusion (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 a reasonable way with respect to the parameters).

Remark 6. 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 (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 with the values of the f_s functions that are equal to $+\infty$. If f_s, g_t are merely non-negative Borel functions, with f_s locally bounded, it is possible to reduce the problem to continuous functions $\{f_s\}$ and $\{g_t\}$ by convolution with non-negative compactly supported continuous kernels, hence reducing this case 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 4. We rather tend to think that the result is true for general Borel functions (as we can prove it in the "discrete case" of Corollary 7; see also Remark 18).

Suppose that the conditions of Setting 3 and Theorem 4 are satisfied for $\{f_s\}$ and $\{g_t\}$. Using the "norm condition" (i), we see that the basic assumption (3) remains true if we add a 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 3 and Theorem 4 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 (2) remains true, with ε_0 replaced by $\varepsilon_0 - \varepsilon$. It follows that the conclusion (4) holds if we replace the inside 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 4 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 (4) then cancels since μ_1 and μ_2 have the same finite mass.

In many applications, X_1 and X_2 are finite probability spaces, and we are then working with finite families of objects parameterized by Ω_1 and Ω_2, or rather by the supports $\mathrm{supp}(\mu_1)$ and $\mathrm{supp}(\mu_2)$; in particular, the mappings $\alpha : \Omega_1 \to \mathbb{R}^n$ are families of $|\Omega_1|$ vectors of $\mathbb{R}^n$ 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 $x \in \mathbb{R}^n$. The proof of the next corollary will be given in Section 6.

Corollary 7. *Under the Setting 3, assume in addition that μ_1 and μ_2 are measures with* finite support *and with $\mu_1(\Omega_1) = \mu_2(\Omega_2) < +\infty$. Then the assumption* (3) *implies* (4) *with no further restriction on the functions $\{f_s\}$ and $\{g_t\}$.*

Let us describe a particular instance of Theorem 4. 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

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

verifies the norm condition (i) and the "constant functions condition" (ii) from Setting 3. A simple and 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 4 reads as:

Corollary 8. *Let Ω be a Polish space, Σ its Borel σ-algebra, μ a probability measure on (Ω, Σ). Suppose that we are in Setting 3 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 4, and let g be a bounded below l.s.c. function on $\mathbb{R}^n$. If for every $\alpha \in L^2(\Omega, \Sigma, \mu, \mathbb{R}^n)$, we have*

$$g\left(\int_\Omega \alpha(s)\, d\mu(s)\right) \leq \int_\Omega f_s\big(\alpha(s)\big)\, d\mu(s),$$

then

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

The Prékopa–Leindler inequality follows by taking Ω to be a two point probability space, for instance $\Omega = \{0, 1\}$, and taking μ of the form $\mu = (1-t)\delta_0 + t\delta_1$ for some $t \in [0, 1]$. Actually, it is easily seen and certainly known that the previous corollary can be deduced from the Prékopa–Leindler inequality, by discretizing μ and by a simple induction procedure. Therefore, it only serves as motivation for the abstract setting above, but does not bring new information.

Let us explain how to deduce the more surprising Proposition 2 above. We apply Corollary 7 to $\Omega_1 = \Omega_2 = \{0, 1\}$, to the finite measures $\mu_1 = \mu_2 = \delta_0 + \delta_1$ and to the linear mapping $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ defined by

$$A(x_0, x_1) = (2x_0/3 + x_1/3, x_0/3 + 2x_1/3).$$

We are in Setting 3 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 result therefore follows from Corollary 7.

In the next section, we present Borell's approach and establish in Proposition 12 a Gaussian version of Theorem 4; the proof of this Proposition can be considered as the heart of the present paper. In Section 3, we present the proof of Theorem 4. Then, in the following section, we discuss some consequences of it. In Section 5 we present a generalization of Theorem 4, 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 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 9. *Let $T > 0$ be fixed. For every bounded continuous real function $f : \mathbb{R}^n \to \mathbb{R}$, we have*

$$-\log P_T\left(e^{-f}\right)(0) = \inf_u \mathbb{E}\left[f\Big(B_T + \int_0^T u_r \, dr\Big) + \frac{1}{2}\int_0^T |u_r|^2 \, dr\right], \tag{5}$$

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 with 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 to say

$$f_r(x) = -\log(P_{T-r} e^{-f})(x), \quad x \in \mathbb{R}^n. \tag{6}$$

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

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

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

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

is a submartingale for any drift u in D_2, since

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

This implies the "inequality case" of (5), 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 (5) 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

$$X_0 = 0, \quad dX_r = dB_r - \nabla f_r(X_r) \, dr, \tag{7}$$

then M_r becomes a martingale, thus giving an equality case in (5).

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 (6), 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 $f_{T-\varepsilon} = -\log\bigl(P_\varepsilon e^{-f}\bigr)$ 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 to the "good setting" on $[0,T_\varepsilon]$. Hence the optimal representation by the martingale

$$M_r = f_r\Bigl(B_r + \int_0^r u_\rho\, d\rho\Bigr) + \frac{1}{2}\int_0^r |u_\rho|^2\, d\rho = M_0 + \int_0^r \nabla f_\rho(X_\rho)\cdot dB_\rho,$$

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 that

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

We have $T^{-1}\,\mathbb{E}\bigl(\int_0^T |u_\rho|\, d\rho\bigr)^2 \le \mathbb{E}\int_0^T |u_\rho|^2\, d\rho \le 2f_0(0)$, 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 almost surely to

$$f\Bigl(B_T + \int_0^T u_\rho\, d\rho\Bigr) + \frac{1}{2}\int_0^T |u_\rho|^2\, d\rho \tag{8}$$

when $r \to T$ since f_r converges locally uniformly to the bounded continuous function f. This implies that M_T is equal to the expression in (8), hence the expectation of (8) is equal to $M_0 = f_0(0)$, that is, it implies formula (5) with equality.

In the inequality case of (5), 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 (8), and the result follows. □

Remark 10. 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 to say, $(a) : f_r(y) < f(x) + \varepsilon$. 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 (5), the limit of $(M_r)_{r<T}$ as $r \to T$ is less than or equal to the expression in (8), 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 (5) is true for every bounded Borel function f. Using the reverse inequality (b) for l.s.c. functions, we see that the equality in (5) 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 11. *The formula* (5) *remains valid when f is continuous, bounded below and satisfies an exponential upper bound of the form*

$$f(x) \leq a e^{b|x|}, \quad a, b \geq 0, \quad x \in \mathbb{R}^n. \tag{9}$$

The "inequality case" in (5) *is valid for any bounded below continuous function f.*

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

Proposition 12. *Under the Setting 3 with finite measures μ_1 and μ_2, assume that f_s, g_t are continuous ≥ 0 on $\mathbb{R}^n$ (or bounded from below as in Remark 5), and that f_s satisfies the following exponential bound: there are non-negative measurable functions $a(s), b(s)$ on Ω_1 such that*

$$f_s(x) \leq a(s) e^{b(s)|x|}, \quad x \in \mathbb{R}^n. \tag{10}$$

Then, for every isotropic Gaussian probability measure $\gamma = \gamma_{n,\tau}$ on $\mathbb{R}^n$, the basic assumption (3) *implies that*

$$\int_{\Omega_2} -\log\left(\int_{\mathbb{R}^n} e^{-g_t(x)}\, d\gamma(x)\right) d\mu_2(t) \leq \int_{\Omega_1} -\log\left(\int_{\mathbb{R}^n} e^{-f_s(x)}\, d\gamma(x)\right) d\mu_1(s). \tag{11}$$

As seen from the proof of Corollary 7 (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. The proof will be an application of the formula (5), as extended in Lemma 11, to the functions f_s and g_t. We may assume that the right-hand side of (11) is not equal to $+\infty$. The Gaussian measure $\gamma = \gamma_{n,\tau}$ is the distribution of B_T for $T = \tau > 0$, so that

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

Let $\varepsilon > 0$ be given. By Lemma 11, 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

$$\mathbb{E}\Big[f_s\Big(B_T + \int_0^T U_r(s)\, dr\Big) + \frac{1}{2}\int_0^T |U_r(s)|^2\, dr\Big] < -\log(P_T e^{-f_s})(0) + \varepsilon. \tag{12}$$

We will show later (Claim 28) that this process $\{U(s)\}_{s \in \Omega_1}$ can be chosen to be Σ_1-measurable. Note also that (12) and $f_s \geq 0$ ensure that

$$\int_{\Omega_1} \mathbb{E} \int_0^T |U_r(s)|^2\, dr\, d\mu_1(s) \leq 2\Big(\int_{\Omega_1} -\log(P_T e^{-f_s})(0)\, d\mu_1(s) + \varepsilon \mu_1(\Omega_1)\Big) < +\infty. \tag{13}$$

Assume that the Brownian motion (B_r) is defined on a probability space $(E, \mathcal{A}, \mathbb{P})$. We have a $\mathbb{R}^n$-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 (13), 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 $P_T(e^{-g_t})$ using the inequality case of formula (5) given by Lemma 11, with the drift $\{V_r(t)\}_{r \le T} = \{AU_r(t)\}_{r \le T}$, namely

$$V_{r,\omega}(t) := A(U_{r,\omega})(t) = A\big(s \to U_{r,\omega}(s)\big)(t) \in \mathbb{R}^n, \quad t \in \Omega_2, \tag{14}$$

where $U_{r,\omega} \in L^2(X_1, \mathbb{R}^n)$ for almost every r, ω. We shall use the basic assumption (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

$$\beta_\omega(s) := \int_0^T U_r(s,\omega)\, dr, \tag{15}$$

and where we consider $B_T(\omega)$ as a constant function of the s variable. We know by (13) 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

$$(A\alpha_\omega)(t) = A(B_T(\omega) + \beta_\omega)(t) = B_T(\omega) + (A\beta_\omega)(t) = B_T(\omega) + \int_0^T V_r(t,\omega)\, dr. \tag{16}$$

As we said, we apply the inequality case of formula (5) for g_t with the drift $\{V_r\}$ and then we integrate in t, in order to get that

$$\int_{\Omega_2} -\log\big(P_T e^{-g_t}\big)(0)\, d\mu_2(t) \le \int_{\Omega_2} \mathbb{E}\left[g_t\Big(B_T + \int_0^T V_r(t)\, dr\Big) + \frac{1}{2}\int_0^T |V_r(t)|^2\, dr \right] d\mu_2(t).$$

For future reference, we state an obvious consequence of (16),

$$g_t\big((A\alpha_\omega)(t)\big) = g_t\big(A(B_T(\omega)+\beta_\omega)(t)\big) = g_t\big(B_T(\omega)+(A\beta_\omega)(t)\big) = g_t\Big(B_T(\omega)+\int_0^T V_r(t,\omega)\, dr\Big). \tag{17}$$

Using (3) and (17), 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 $r \in [0,T]$,

$$\int_{\Omega_2} |V_r(t,\omega)|^2\, d\mu_2(t) \le \int_{\Omega_1} |U_r(s,\omega)|^2\, d\mu_1(s). \tag{18}$$

Finally

$$\int_{\Omega_2} -\log\big(P_T e^{-g_t}\big)(0)\, d\mu_2(t) \le \int_{\Omega_1} \mathbb{E}\left[f_s\Big(B_T + \int_0^T U_r(s)\, dr\Big) + \frac{1}{2}\int_0^T |U_r(s)|^2\, dr \right] d\mu_1(s)$$

$$< \int_{\Omega_1} -\log\big(P_T e^{-f_s}\big)(0)\, d\mu_1(s) + \varepsilon\mu_1(\Omega_1).$$

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

As the reader has noticed, the proof is rather short, provided one has put forward the abstract properties contained in the four equations (14), (15), (17) and (18) that allow us to run Borell's argument. $\square$

3 Proof of Theorem 4

Suppose that we try to obtain the conclusion (4) of Theorem 4 for $\{f_s\}$ and $\{g_t\}$ as 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 $f_{s,k} \to f_s$, $g_{t,k} \to g_t$ when $k \to +\infty$. 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

$$e^{-H_k(s)} = \int_{\mathbb{R}^n} e^{-h_{s,k}(x)}\, d\nu(x), \quad e^{-H(s)} = \int_{\mathbb{R}^n} e^{-h_s(x)}\, d\nu(x), \tag{19}$$

that are Σ-measurable functions 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 13. *Let $H_k(s), H(s)$ be defined by* (19)*. 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 that $H_0(s) > -\infty$ for μ-almost every $s \in \Omega$. Then:*
  - *(a1) $H_k(s) \nearrow H(s)$ μ-a.e. and $\int_\Omega H(s)\, d\mu(s) \geq \limsup_k \int_\Omega H_k(s)\, d\mu(s)$.*
  - *(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 $\int_\Omega H(s)\, d\mu(s) \leq \liminf_k \int_\Omega H_k(s)\, d\mu(s)$.*
  - *(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 repeatedly one of 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 $\int_\Omega H = \int_\Omega H^+ - \int_\Omega H^-$ (that 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. $\square$

Proof of Theorem 4. 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 $f_s(x)$, $g_t(x) + \varepsilon|x|^2$ decreases to $g_t(x)$ when $\varepsilon \searrow 0$; for $\varepsilon \in [0, \varepsilon_0]$, define $F_\varepsilon(s)$, $G_\varepsilon(t)$ as in (19), with ν being the Lebesgue measure, that is to say, let $e^{-F_\varepsilon(s)} = \int_{\mathbb{R}^n} e^{-f_s(x)-\varepsilon|x|^2}\, dx$ and use the 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 side of (4) is $-\infty$, an obvious case. We also know by the assumption (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 yields that $F_\varepsilon(s)$ and $G_\varepsilon(t)$ are semi-integrable for $\varepsilon \in [0, \varepsilon_0]$. Use Lemma 13, (b) to see that $F_\varepsilon(s) \to F(s)$, $G_\varepsilon(t) \to G(t)$ for every s, t. For the F side, we may use $(b2)$ because $F_{\varepsilon_0}^+$ is integrable. We can therefore pass to the limit as $\varepsilon \to 0$ and conclude that

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

For the analogous expressions with g_t, we use $(b1)$ of Lemma 13.

Now we have to consider the problem with $\gamma = \gamma_{n,\tau}$ replacing the Lebesgue measure. We know by (2) that $s \to \log^-\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\, d\gamma(x)\Big)$ is μ_1-integrable when $2\tau > 1/\varepsilon_0$. 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 29. Fix $k > 0$ and define $f_{s,k}$, $g_{t,k}$ by inf-convolution of f_s, g_t with $h_k(x) = k|x|^2$,

$$f_{s,k}(x) = \inf_{u\in\mathbb{R}^n} \big(f_s(x+u) + k|u|^2\big), \quad g_{t,k}(x) = \inf_{u\in\mathbb{R}^n} \big(g_t(x+u) + k|u|^2\big). \tag{20}$$

Clearly, $f_{s,k} \le f_s$, $g_{t,k} \le g_t$, and $f_{s,k}$, $g_{t,k}$ are continuous on $\mathbb{R}^n$. By Lemma 29, we may find negligible Borel sets $N_1 \in \Sigma_1$, $N_2 \in \Sigma_2$ such that $(s, x) \to f_{s,k}(x)$, $(t, x) \to g_{t,k}(x)$ 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 that

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

fitting the exponential bound (9) needed to apply Proposition 12. 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 29) such that

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

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

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

which is bounded, using the basic assumption (3) and the norm condition $\|A\| \leq 1$, by

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

Hence, $f_{s,k}$ and $g_{t,k}$ also satisfy the basic assumption. By Proposition 12, we conclude that

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

When k tends to infinity, $f_{s,k}$ increases to a l.s.c. function $\widetilde{f}_s \leq f_s$, and $g_{t,k}(x)$ increases to $g_t(x)$ because $x \to g_t(x)$ is l.s.c. on $\mathbb{R}^n$. We apply again Lemma 13, this time with $\nu = \gamma$, defining $F_k(s)$, $\widetilde{F}(s) \leq F(s)$, $G_k(t)$, $G(t)$ from $f_{s,k}$, $\widetilde{f}_s, f_s$, $g_{t,k}, g_t$ as in (19). Since the functions are ≥ 0, $e^{-f_{s,k}(x)}$, $e^{-g_{t,k}(x)}$ are bounded by 1, we have $F_k(s), G_k(t) \geq 0$ because γ is a probability measure. Thus $F_0(s), G_0(t) > -\infty$ and $F_k(s) \to \widetilde{F}(s)$, $G_k(t) \to G(t)$ by (a). The conclusion follows, by $(a1)$ for F, and by $(a2)$ for G since $G_0^- = 0$. □

Remark 14. Assume that f_s, 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 $(s,x) \to f_{s,k}(x)$, $(t,x) \to g_{t,k}(x)$ 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

Let us first comment on Corollary 8. The classical Prékopa–Leindler inequality corresponds to the case of a two point space. 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 15. *Under the assumptions of Corollary 8 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 5) Borel functions on $\mathbb{R}^n$ and let g be a bounded from below l.s.c. function on $\mathbb{R}^n$. If for every $\alpha \in L^2(S^1, \mathbb{R}^n)$, we have*

$$g\left(\int_0^{2\pi} \alpha(e^{i\theta})\, \frac{d\theta}{2\pi}\right) \leq \int_0^{2\pi} f_{e^{i\theta}}\big(\alpha(e^{i\theta})\big)\, \frac{d\theta}{2\pi},$$

then it follows that

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

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

Corollary 16 (Berndtsson[3]). *Let $\varphi : \mathbb{C} \times \mathbb{C}^n \to \mathbb{R}$ be plurisubharmonic on $\mathbb{C}^{n+1}$ and such that, for every $z \in \mathbb{C}$, $w \in \mathbb{C}^n$, we have that $\varphi(z, w) = \varphi(z, \Re w)$. Then, the function*

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

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 $\psi(0) \leq \int_0^{2\pi} \psi(re^{i\theta}) \frac{d\theta}{2\pi}$ for any fixed $r > 0$ small enough. We take $r = 1$ to simplify notations. Take $g(x) = \varphi(0, x)$ and $f_{e^{i\theta}}(x) := \varphi(e^{i\theta}, x)$ for $x \in \mathbb{R}^n \subset \mathbb{C}^n$. To conclude, it suffices to check that these functions verify the hypothesis in 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 $\widetilde{\alpha}(0) = \int_0^{2\pi} \alpha(e^{i\theta}) \frac{d\theta}{2\pi}$. We can write this $\widetilde{\alpha}$ as the real part $\Re H$ of an holomorphic function $H : \mathbb{D} \to \mathbb{C}^n$ such that $H(0) = \widetilde{\alpha}(0)$. We conclude by noticing that $z \to \varphi(z, H(z))$ is subharmonic on $\mathbb{D}$, and using $\varphi(z, H(z)) = \varphi(z, \widetilde{\alpha}(z))$ we obtain that

$$g\left(\int_0^{2\pi} \alpha(e^{i\theta}) \frac{d\theta}{2\pi}\right) = \varphi(0, H(0)) \leq \int_0^{2\pi} \varphi(e^{i\theta}, H(e^{i\theta})) \frac{d\theta}{2\pi} = \int_0^{2\pi} f_{e^{i\theta}}(\alpha(e^{i\theta})) \frac{d\theta}{2\pi}.$$

□

So far, we have presented consequences of Corollary 8, and therefore more or less straightforward consequences of the Prékopa–Leindler inequality itself. The situation is different with other instances of Theorem 4 such as Proposition 2. It seems difficult to guess the existence of such inequalities without having in mind Borell's proof.

Theorem 17. *Let μ be a Borel probability measure on $[0, 1]$ and set $m := \int t\, d\mu(t) \in [0, 1]$. Under Setting 3 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 bounded from below (as in Remark (5)) Borel functions from $\mathbb{R}^n$ to $\mathbb{R} \cup \{+\infty\}$. Assume g_t l.s.c. for every t in Ω_2. If for every $x_0, x_1 \in \mathbb{R}^n$ we have*

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

then it follows that

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

The deduction of this result from Theorem 4 is as follows. Take

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

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

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

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

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

Remark 18. It is possible to prove Theorem 17 without assuming g_t l.s.c. and f_i bounded from below, see Claim 26 for a sketch of proof.

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

$$\forall x_0, x_1 \in \mathbb{R}^n, \qquad \int_0^1 g\big((1-t)x_0 + tx_1\big) \, d\mu(t) \leq (1-m)f_0(x_0) + mf_1(x_1), \tag{21}$$

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 the functions f_0, f_1, g, one can try to find an optimal μ for which the condition (21) holds. The classical Prékopa–Leindler inequality corresponds to the choice $\mu = \delta_m$, and this is indeed the optimal choice when g is convex, as seen by using Jensen's inequality in (21). 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 convex combination of two Dirac measures) is the following extension of Proposition 2.

Proposition 19. *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 four 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

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

Proof. Although the result is a particular case of Theorem 17 (with $\mu = (1-r)\delta_s + r\delta_t$), it is better to go back to Corollary 7 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 note that for $x_0, x_1 \in \mathbb{R}^n$,

$$\begin{aligned}(1-r)|(1-s)x_0 + sx_1|^2 + r|(1-t)x_0 + tx_1|^2 + ((1-r)s(1-s) + rt(1-t))|x_0 - x_1|^2 \\ = (1-m)|x_0|^2 + m|x_1|^2 \end{aligned} \tag{22}$$

and so in particular

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

which exactly means that $\|A\| \le 1$. $\square$

Let us mention that Proposition 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, one has that

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

and in this case, the result of Proposition 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 that Proposition 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 reduce the problem to sets (ellipsoids, actually) and eventually matrices.

Example 20 (Gaussian self-improvement and generalized τ-property)**.** Let $\alpha \in (0,1)$. We shall comment on Proposition 19 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 (22) 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 $x \in \mathbb{R}^n$ by

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

We assume that $g(x) > -\infty$. Using the following particular case of(22),

$$|(1-\alpha)x_0+\alpha x_1|^2+\alpha(1-\alpha)|x_0-x_1|^2 = (1-\alpha)|x_0|^2+\alpha|x_1|^2, \tag{23}$$

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

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

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 get, with $g_0 = g_{\alpha,\beta,\lambda}$ and $g_1 = g_{1-\alpha,1-\beta,1-\lambda}$, that

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

The identity (22) in the case $s = \alpha$, $t = 1-\alpha$ and $r = 1/2$ (and so $m = 1/2$) rewrites 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

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

Therefore, by Proposition 19 with $s = \alpha$, $t = 1-\alpha$ and $r = 1/2 = m$, we arrive at the following generalized infimal convolution inequality:

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

Example 21 (Exotic situation)**.** All our examples so far of linear maps A are of "convex type", meaning that

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

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, $k \geq 0$. 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 $(A_1 x + B_1 x, C_1 x + D_1 x) = (x, x)$, 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 $\|A\| \leq 1$, each block must have norm ≤ 1, and this implies that the diagonal coefficients of B, C are in $[0,1]$. The condition $\|A\| \leq 1$ means that $A^*A \leq I_{2n}$, which translates to

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

In the simpler case $C = B$, this amounts to $2B^*B \leq B + B^*$, or $\|Bx\|^2 \leq Bx \cdot x$, 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 $b^2 + \varepsilon^2 \leq b$. We see that $B = b\mathrm{Id}_2 - \varepsilon R$, 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

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

for every $x_0, x_1 \in \mathbb{R}^n$, we get the same conclusion as that of Proposition 2.

Remark 22. 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 $|(I_n - B)x + By|^2 \leq (1-\alpha)|x|^2 + \alpha|y|^2$ 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 $B_1, \ldots, B_k$ are $n \times n$ matrices such that $\sum_{j=1}^k B_j = I_n$ and $|\sum_{j=1}^k B_j x_j|^2 \leq \sum_{j=1}^k \alpha_j |x_j|^2$, with $\alpha_j \geq 0$ and $\sum_{j=1}^k \alpha_j = 1$, then $B_j = \alpha_j I_n$, $j = 1, \ldots, k$.

5 Generalized Brascamp–Lieb and reverse Brascamp–Lieb inequalities

By slightly modifying the Setting 3, 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 $m \neq n$, then the linear operator A of Setting 3 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 $A(s \to v)$ of the constant function $\Omega_1 \ni s \to v$ with the constant function $\Omega_2 \ni t \to Qv$. Actually, the new setting will be notably more complicated, introducing a *family* of projections $Q(t)$, $t \in \Omega_2$, and comparing the image $A(s \to v)$ with the function $t \to Q(t)v$. 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 3'. 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 3. 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 the measures μ_1 and μ_2, respectively. The constant functions condition is modified as follows:

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

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

Note that formally, $(ii)_2$ is the adjoint situation to $(ii)_1$. 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 to the following extension of Theorem 4.

Theorem 23. *Under Setting 3', we make the additional assumptions that the functions f_s, g_t are non-negative with g_t lower semicontinuous on $\mathbb{R}^n$, that for some $\varepsilon_0 > 0$ we have*

$$\int_{\Omega_1} \log^- \Big(\int_{\mathbb{R}^m} \exp\big(-f_s(x) - \varepsilon_0 |x|^2\big)\, dx \Big)\, d\mu_1(s) < +\infty, \tag{24}$$

and that 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

$$\int_{\Omega_2} g_t\big((A\alpha)(t)\big)\, d\mu_2(t) \leq \int_{\Omega_1} f_s(\alpha(s))\, d\mu_1(s), \tag{25}$$

then

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

Not only the argument for deriving Theorem 23 follows the proof of Theorem 4, but in its Gaussian version, Theorem 23 is already contained in Proposition 12. Indeed, after reducing to the Gaussian case, we further approximate as before 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 to prove Thorem 4. 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, $T \to +\infty$, in respective dimensions m and n.

For the Gaussian version below, we will have not much to add, except maybe for the case $ii)_2$ of the Setting 3'. So Proposition 12 almost includes all possible geometric situations around the Prékopa–Leindler inequality, including these generalized Brascamp–Lieb inequalities.

Proposition 24. *Under the Setting 3' with finite measures μ_1 and μ_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 5), and that the f_s satisfy the exponential bound* (10) *on $\mathbb{R}^m$. Then, the basic assumption* (25) *implies that*

$$\int_{\Omega_2} -\log\left(\int_{\mathbb{R}^n} e^{-g_t(y)}\, d\gamma_{n,\tau}(y)\right) d\mu_2(t) \leq \int_{\Omega_1} -\log\left(\int_{\mathbb{R}^m} e^{-f_s(x)}\, d\gamma_{m,\tau}(x)\right) d\mu_1(s). \quad (26)$$

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

Proof. We want to apply Proposition 12 but we have to deal with the fact that 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 $(ii)_1$, 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

$$\int_{\mathbb{R}^n} e^{-f_{1,s}(y)}\, d\gamma_{n,\tau}(y) = \int_{\mathbb{R}^m} e^{-f_s(x)}\, d\gamma_{m,\tau}(x). \quad (27)$$

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 $A_1(s \to w_0) = A(s \to P(s)w_0) = w_0$ by $(ii)_1$, so A_1 verifies the constant functions condition (ii). Next, by the basic assumption of Theorem 23 applied to α_1, we get that

$$\int_{\Omega_2} g_t\big((A_1\alpha)(t)\big)\, d\mu_2(t) = \int_{\Omega_2} g_t\big((A\alpha_1)(t)\big)\, d\mu_2(t) \leq \int_{\Omega_1} f_s\big(\alpha_1(s)\big)\, d\mu_1(s)$$

$$= \int_{\Omega_1} f_s\big(P(s)\alpha(s)\big)\, d\mu_1(s) = \int_{\Omega_1} f_{1,s}(\alpha(s))\, d\mu_1(s).$$

We see that A_1, $\{f_{1,s}\}$ and $\{g_t\}$ satisfy the assumptions of Proposition 12, including the basic assumption (3) on $\mathbb{R}^n$. The result follows therefore from (27) and from the conclusion of Proposition 12 for $\{f_{1,s}\}$ and $\{g_t\}$.

In case $(ii)_2$, when $m \geq n$, 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 $g_{1,t}(x) = g_t(Q(t)x)$, $x \in \mathbb{R}^m$. For $\alpha \in L^2(X_1, \mathbb{R}^m)$, we set $(A_1\alpha)(t) = T(t)(A\alpha)(t)$ and get that $Q(t)(A_1\alpha)(t) = (A\alpha)(t)$. We have that $\|A_1\| = \|A\| \leq 1$. If $\alpha_0 = v_0$ is

a constant function from Ω_1 to $\mathbb{R}^m$, we know by $(ii)_2$ 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

$$g_{1,t}\big(A_1(v_0+\alpha)(t)\big) = g_{1,t}\big(v_0 + (A_1\alpha)(t)\big), \quad t \in \Omega_2, \tag{28}$$

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 we have by $(ii)_2$ that

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

The basic assumption (3) is satisfied for the families $\{f_s\}$, $\{g_{1,t}\}$ of functions on $\mathbb{R}^m$ and the mapping A_1: observing that $g_{1,t}\big((A_1\alpha)(t)\big) = g_t\big((A\alpha)(t)\big)$, 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 12 that the equality (28) is enough to run the argument (see equation (17) 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 12 to get the desired conclusion. □

Remark 25. *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

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

The reader will just look at inequality (18) and the inequality before it.

We shall examine now particular cases of Theorem 23. We will see that it contains both Brascamp–Lieb and reverse Brascamp–Lieb inequalities in their geometric form.

Let us give 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 $\mathbb{R}^d$:

$$\sum_{i=1}^{N} c_i\, u_i \otimes u_i = \mathrm{Id}_{\mathbb{R}^d}. \tag{29}$$

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

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

then (29) is equivalent to saying that the mapping

$$U : x \in \mathbb{R}^d \longrightarrow (x \cdot u_1, \dots, x \cdot u_N) \tag{30}$$

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

$$X_1 = (E_2, \nu_2),\ m = 1, \quad X_2 = (E_1, \nu_1),\ n = d, \quad m\mu_1(\Omega_1) = \nu_2(E_2) = d = n\mu_2(\Omega_2).$$

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 (30) between our L^2 spaces, and A verifies the condition $(ii)_1$ of Setting 3', with the family of projections $P_i v = v \cdot u_i$, $i \in \Omega_1$, because we have by (29) that

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

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] (that can be easily derived by approximation, anyway).

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 $A(x) = (x \cdot u_1, \dots, x \cdot u_N)$. Then $A = U$, the into isometry (30) between the corresponding L^2 spaces, and it satisfies the assumption $(ii)_2$ with the family of projections $Q_j(x) = x \cdot u_j$, $j \in \Omega_2 = \{1, \dots, N\}$. The conclusion reads as follows: given $f : \mathbb{R}^d \to \mathbb{R}$ and $g_1, \dots, g_N : \mathbb{R} \to \mathbb{R}$ such that

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

This is the geometric Brascamp–Lieb inequality, usually stated with the best possible f, which is defined by the 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 23 where each function 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 spaces E_s, F_t are subspaces of a given $\mathbb{R}^\ell$. We may consider that the space E_s is $\mathbb{R}^{m(s)}$ and that F_t is $\mathbb{R}^{n(t)}$, with $1 \le m(s), n(t) \le \ell$ and we 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 $s\in\Omega_1$. For every square integrable choice $\alpha\in H$, the image $A\alpha\in L^2(X_2,\mathbb{R}^\ell)$ is such that $(A\alpha)(t)$ belongs to F_t for every $t\in\Omega_2$. We give projections $P(s), Q(t)$ from $\mathbb{R}^\ell$ onto E_s, F_t respectively. The constant functions condition says now that for every $w_0\in\mathbb{R}^\ell$, we have that

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

The proof mixes the two cases $(ii)_1$ and $(ii)_2$ of the proof of Theorem 23. We leave this to the reader.

6 Technicalities

Proof of Corollary 7. Under the Setting 3 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 $x_0, x_1, \dots, x_p$ in $\mathbb{R}^n$ we have

$$\sum_{j=0}^q \nu_j g_j\Big(\sum_{i=0}^p a_{j,i}x_i\Big) \le \sum_{i=0}^p \mu_i f_i(x_i), \tag{31}$$

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). We want to prove that

$$\sum_{j=0}^q -\nu_j \log\Big(\int_{\mathbb{R}^n} e^{-g_j}\Big) \le \sum_{i=0}^p -\mu_i \log\Big(\int_{\mathbb{R}^n} e^{-f_i}\Big). \tag{32}$$

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 (32) is $+\infty$ and this case is obvious. We may therefore assume that all f_i are L-proper. It follows that $\int_{\mathbb{R}^n} e^{-g_j} > 0$, $j = 0, \dots, q$. Indeed, if for

example $\int_{\mathbb{R}^n} e^{-g_0} = 0$, then g_0 is $+\infty$ almost everywhere. We know that $\sum_{i=0}^p a_{0,i} = 1$ since A preserves constant functions by (ii). Assume $a_{0,1} \neq 0$ for example. Since all f_i are L-proper, we may fix x_i with $f_i(x_i) < +\infty$ for $i \neq 1$, and we see from (31) that for almost every x_1,

$$+\infty = \nu_0 g_0\Big(a_{0,1}x_1 + \sum_{i\neq 1} a_{0,i}x_i\Big) + \sum_{j\neq 0} \nu_j g_j\Big(\sum_{i=0}^p a_{j,i}x_i\Big) - \sum_{i\neq 1} \mu_i f_i(x_i) \leq \mu_1 f_1(x_1),$$

hence f_1 is $+\infty$ almost everywhere, a contradiction.

For $\varepsilon > 0$ and $i = 0, \dots, p$, $j = 0, \dots, q$, the functions $f_{i,\varepsilon}(x) = f_i(x) + \varepsilon|x|^2$, $g_{j,\varepsilon}(x) = g_j(x) + \varepsilon|x|^2$, still satisfy (31) by the assumption $\|A\| \leq 1$. We may reduce the problem to proving (32) for $f_{i,\varepsilon}$, $g_{j,\varepsilon}$, since we can pass to the limit as $\varepsilon \to 0$ 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 (32) by monotone convergence. In other words, we may keep f_i, g_j but replace the Lebesgue measure in (32) 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 $\int_{\mathbb{R}^n} e^{-g_j}\,d\gamma$, $j \neq j_0$ are > 0, and the left-hand side of (32) is $-\infty$. Otherwise, for $N \in \mathbb{N}$, define $g_{j,N} = \min(g_j, N)$ and $f_{i,N} = \max(f_i, -N)$, that trivially satisfy (31), and observe that it is enough to give the proof for $f_{i,N}$, $g_{j,N}$; indeed, since $e^{-\min(g_j,0)} \leq e^{-g_j} + 1$ 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 reduced the question to the case $g_j \leq N$ and $f_i \geq -N$. 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

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

and for the same reason, we may also assume g_{j_0} bounded below by $-2N/\nu_{j_0}$, while still keeping (31) 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 a same vector, then (31) remains true by the constant functions condition (ii), therefore (31) is stable by 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 (31) and meeting the assumptions of Proposition 12. We obtain thus, 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 the conclusion of the restricted problem follows by the dominated convergence theorem. $\square$

Claim 26. *In Theorem 17, we can remove the assumptions that g_t is l.s.c. and that $f_i \geq 0$ (or bounded from below), $i = 0, 1$.*

Sketch of proof. As in the proof of Theorem 4, 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 reason is the same here for the g_t side, but is easier for f_i, $i = 0, 1$, a simple application of monotone convergence. Next, as in the proof of Corollary 7, 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

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

is the same as at the end of the proof of Theorem 4, 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 $0 \leq g_{t,N} \leq N$ and $f_{0,N}, f_{1,N} \geq -N$. As in the proof of Corollary 7, one can then assume that $f_{0,N}, f_{1,N} \leq 2N/\beta$ with $\beta = \min(\alpha, 1-\alpha)$. Finally, approximation by convolution reduces to the case of bounded continuous functions, and one can conclude by applying Proposition 12. □

Proof of Lemma 11. We may assume that $f \geq 0$. For some integer $N \geq 0$, let $\varphi(x) = \min(f(x), N) \geq 0$. As in (6), define φ_r by $e^{-\varphi_r} = P_{T-r}(e^{-\varphi})$, $0 \leq r \leq T$ and let $\Phi_0 = \varphi_0(0)$, $F_0 = -\log P_T(e^{-f})(0) \geq \Phi_0 > 0$. Note that $\varphi \geq 0$ implies $\varphi_r \geq 0$. Since φ is bounded and continuous, we may by Lemma 9 consider the optimal martingale corresponding to φ,

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

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 $(S_r)_{0\leq r\leq T}$ of the martingale $(M_r - M_0)_{0\leq r\leq T}$ by

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

We know by (8) 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 $0 \leq r < T$, namely

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

By Fatou and the martingale property, we have that

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

Let $Y_T = T^{-1/2}|X_T| \leq T^{-1/2}|B_T| + S_T$ 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 $2\sigma\tau \leq c\sigma^2 + \tau^2/c$ when $c, \sigma, \tau > 0$, we can write

$$\Big(\mathbb{E}\,e^{\lambda Y_T}\Big)^2 \leq \Big(\mathbb{E}\,e^{2\lambda T^{-1/2}|B_T|}\Big)\Big(\mathbb{E}\,e^{2\lambda S_T}\Big) = \Big(\mathbb{E}\,e^{2\lambda|B_1|}\Big)\Big(\mathbb{E}\,e^{2\lambda S_T}\Big)$$

$$\leq e^{4\lambda^2}\,\mathbb{E}\,e^{|B_1|^2/4}e^{8\lambda^2}\,\mathbb{E}\,e^{S_T^2/8} = 2^{n/2}e^{12\lambda^2}\,\mathbb{E}\,e^{S_T^2/8} \leq 2^{n/2}e^{12\lambda^2+\Phi_0/2} \leq e^{12\lambda^2+(F_0+n)/2}. \tag{33}$$

We have by the exponential bound (9) that

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

Using (33) with $\lambda = 2b\sqrt{T}$, we get by Cauchy–Schwarz and Markov

$$\delta \leq a\sqrt{\frac{\mathbb{E}\,M_T}{N}}\sqrt{e^{24b^2T+(F_0+n)/4}} \leq a\sqrt{\frac{F_0}{N}}e^{12b^2T+(F_0+n)/8}. \tag{34}$$

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

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

Conversely, since $1 \geq P_Te^{-\min(f,N)} \searrow P_Te^{-f}$ when $N \to +\infty$, 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

$$F_0-\varepsilon < \Phi_0 \leq \mathbb{E}\Big[\varphi\Big(B_T+\int_0^T u_\rho\,d\rho\Big)+\frac{1}{2}\int_0^T |u_\rho|^2\,d\rho\Big]$$

$$\leq \mathbb{E}\Big[f\Big(B_T+\int_0^T u_\rho\,d\rho\Big)+\frac{1}{2}\int_0^T |u_\rho|^2\,d\rho\Big].$$

This implies the inf formula (5) for the function f. Observe that this final part does not use any upper bound on f, proving thus the last claim of Lemma 11. □

Lemma 27. *Assume that μ is a finite measure on (Ω,Σ), ν a probability measure on $\mathbb{R}^n$, $(s,x) \to f_s(x)$ a $\Sigma\otimes\mathcal{B}_{\mathbb{R}^n}$-measurable function, and assume that $s \to \log\Big(\int_{\mathbb{R}^n} e^{-f_s}\,d\nu\Big)$ is μ-integrable on Ω. For every $\varepsilon > 0$, there exists a continuous function $\psi \geq 0$ on $\mathbb{R}^n$, tending to $+\infty$ at infinity, such that $0 \leq \psi(x) \leq |x|^2$ and*

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

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

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

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

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

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 $\sum_{j\ge1} k_j^{-2} < \pi^2/6 - 1 \le 1$ because $k_1 > 1$). □

Claim 28. *The almost optimal drifts $\{U(s)\}_{s\in\Omega_1}$ in (12) can be chosen to be Σ_1-measurable with respect to $s \in \Omega_1$.*

Proof. In the proof of Proposition 12 we may start, using Lemma 27, by replacing $f_s(x) \ge 0$ by $f_s(x) + \psi(x)$, where ψ tends to $+\infty$ at infinity, without changing much the value of $\int_{\Omega_1} -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x)}\, d\gamma(x)\Big)\, d\mu_1(s)$ and without destroying the exponential bound (10) 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 (6). We may find a partition of Ω_1 in countably many subsets $A_k \in \Sigma_1$, $k \in \mathbb{N}$, such that the exponential bound (10) for f_s is uniform on A_k,

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

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

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

By (34), 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 $-\log(P_T e^{-f_s})(0)$. 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 $s \in A_k \to \varphi_s$ 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*), $s \to \varphi_s(x)$ 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 $s \to \varphi_s$ is Σ_1-measurable from Ω_1 to Z. Next, the mappings sending a bounded continuous real function f on $\mathbb{R}^n$ to f_r, defined by (6), and to ∇f_r, $\nabla^2 f_r$ are continuous from Z to $C_b([0, T-\varepsilon], \mathbb{R}^p)$, $p = 1, n, n^2$, for every $\varepsilon > 0$, and the map sending f to the solution $X_{r,f}$ of the stochastic differential equation (7) is continuous, in the precise sense that for every bounded subset B of Z, there is $\kappa = \kappa(B, \varepsilon)$ such that for every $\omega \in E$,

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

Indeed, for every fixed ω and $\theta > 0$, the function $D(r) = \big(|X_{r,f}(\omega) - X_{r,g}(\omega)|^2 + \theta^2\big)^{1/2}$ satisfies a deterministic differential inequality of the form $D' \le \kappa(D + \|f - g\|_\infty)$, with $D(0) = 0$. It follows that the "almost optimal" drift $U : (s, r) \to -\nabla \varphi_{s,r}(X_{s,r})$ is Σ_1-measurable. □

Lemma 29. *Under Setting 3, assume that $f_s \ge 0$ and that for some $\varepsilon_0 > 0$, the function $s \in \Omega_1 \to -\log\Big(\int_{\mathbb{R}^n} e^{-f_s(x) - \varepsilon_0 |x|^2}\, dx\Big)$ is μ_1-integrable. There exists then $\alpha \in L^2(X_1, \mathbb{R}^n)$ such that*

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

The inf-convolution $f_{s,k}$ in (20) is universally measurable, and there exists a μ_1-negligible set $N \in \Sigma_1$ such that $(s, x) \to f_{s,k}(x)$ 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

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

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 that

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

from what the conclusions follow since $f_s \ge 0$.

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 $s \to f_{s,k}(x)$ 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 $s \to f_{s,k}(d)$ is Borel outside N_d. Let $N = \bigcup_d N_d$. Since $x \to f_{s,k}(x)$ is continuous, we get that $(s,x) \to f_{s,k}(x)$ is Borel on $\big(\Omega_1 \setminus N\big) \times \mathbb{R}^n$.

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. $\square$

Remark 30. 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 $h(s,x) = h_s(x)$ such that for every $\varepsilon > 0$ exist two $\Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$-measurable functions $\varphi(s,x) = \varphi_s(x)$, $\psi(s,x) = \psi_s(x)$, such that for every $s \in \Omega$, $\varphi_s \le h_s \le \psi_s$, 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 constants, stable by $\sup(h_1, h_2)$ and stable by 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 $|h_k(s,x)| \le 1$. For every k, let $\psi_{s,k} \ge h_{s,k}$ 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, $\chi_{s,n} = \sup_{k \ge n} \psi_{s,k}$ is l.s.c. and decreases to h_s ν-a.e. Let

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

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 get 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 all indicators of products $B \times C$, $B \in \Sigma$, $C \in \mathcal{B}_{\mathbb{R}^n}$, it follows that $\mathcal{A} = \Sigma \otimes \mathcal{B}_{\mathbb{R}^n}$. The result applies then 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 that $H(s) = \int_{\mathbb{R}^n} h(s,x)\, d\nu(x) > 0$ 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 $0 \le \varphi \le h$ 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 to $h(s,x) = e^{-f_s(x)}$, we get an l.s.c. in x function $\psi_s(x)$ such that $f_s \le \psi_s$ and $-\log\big(\int_{\mathbb{R}^n} e^{-\psi_s}\, d\nu\big) \le -\log\big(\int_{\mathbb{R}^n} e^{-f_s}\, d\nu\big) + \varepsilon$. This shows that in Theorem 4, the function f_s can be assumed to be a l.s.c. function.

References

[1] F. Barthe, *On a reverse form of the Brascamp–Lieb inequality*, Invent. Math. **134** (1998), no. 2, 335–361.

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

[3] B. Berndtsson, *Prekopa's theorem and Kiselman's minimum principle for plurisubharmonic functions*, Math. Ann. **312** (1998), no. 4, 785–792.

[4] V. I. Bogachev, Measure theory, Vol. II, Springer-Verlag, 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), no. 1, 49–71.

[7] D. Cordero-Erausquin, *On Berndtsson's generalization of Prékopa's theorem*, Math. Z. **249** (2005), no. 2, 401–410.

[8] R. J. Gardner, *The Brunn–Minkowski inequality*, Bull. Amer. Math. Soc. (N.S.) 39 (2002), no. 3, 355–405.

[9] J. Lehec, *Short probabilistic proof of the Brascamp-Lieb and Barthe theorems*, Canad. Math. Bull. **57** (2014), no. 3, 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 (2005), no. 299, Exp. No. 928, vii, 95–113.

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