



**HAL**  
open science

## A stochastic formula for the entropy and applications

Joseph Lehec

► **To cite this version:**

Joseph Lehec. A stochastic formula for the entropy and applications. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 2013, 49 (3), pp.885-899. 10.1214/11-AIHP464. hal-00534442

**HAL Id: hal-00534442**

**<https://hal.science/hal-00534442>**

Submitted on 9 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A stochastic formula for the entropy and applications

Joseph Lehec\*

June 2010

## Abstract

We prove a stochastic formula for the Gaussian relative entropy in the spirit of Borell's formula for the Laplace transform. As an application, we give unified and short proofs of various well known inequalities.

## 1 Introduction: Borell's formula

Let  $(\Omega, \mathcal{A}, m)$  be a measured space and  $\mu$  be an absolutely continuous probability measure. Letting  $\rho = d\mu/dm$  and assuming that  $\rho \log(\rho)$  is integrable, the relative entropy of  $\mu$  is defined by

$$H_m(\mu) = \int_{\Omega} \rho \log(\rho) dm = \int_{\Omega} \log(\rho) d\mu.$$

Given a function  $f: \Omega \rightarrow \mathbb{R}$  such that  $e^f$  is integrable we let

$$L_m(f) = \log\left(\int e^f dm\right)$$

be the log-Laplace of  $f$ . Then, using Jensen's inequality, it is easily seen that

$$L_m(f) + H_m(\mu) \geq \int f d\mu.$$

Moreover, there is equality as soon as  $e^f$  is a constant multiple of  $\rho$  on the support of  $\mu$ . To sum up, the following Legendre duality relation holds:

$$\begin{aligned} L_m(f) &= \sup_{\mu} \left( \int f d\mu - H_m(\mu) \right) \\ H_m(\mu) &= \sup_f \left( \int f d\mu - L_m(f) \right). \end{aligned} \tag{1}$$

We consider  $\mathbb{R}^n$  equipped with its standard scalar product  $\langle \cdot, \cdot \rangle$ , Euclidean norm  $\|\cdot\|$ , and Gaussian measure  $\gamma_n$ . In [5, 6] Borell proves the following formula. Given a standard  $n$ -dimensional Brownian motion  $B$  and a reasonable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , we have

$$L_{\gamma_n}(f) = \sup_u \left( \mathbb{E} f\left(B(1) + \int_0^1 u(s) ds\right) - \frac{1}{2} \mathbb{E} \int_0^1 \|u(s)\|^2 ds \right),$$

---

\*CEREMADE (UMR CNRS 7534) Université Paris-Dauphine. [lehec@ceremade.dauphine.fr](mailto:lehec@ceremade.dauphine.fr)

where the supremum is taken over all random processes  $u$ , say bounded and progressively measurable with respect to the Brownian filtration. Among other applications, Borell shows that the classical Prékopa-Leindler inequality follows easily from this formula.

Our purpose is to show that a similar formula holds for the Gaussian relative entropy. This should not be too surprising, given the duality relation between the log-Laplace and the entropy. However, this stochastic formula for the entropy provides new, unified and simple proofs of a number of well-known Gaussian inequalities.

## 2 Stochastic formula for the entropy

For simplicity, we only consider measures  $\mu$  on  $\mathbb{R}^n$  having a density with respect to  $\gamma_n$  of the form  $\rho = e^f$  with a function  $f$  twice continuously differentiable, bounded and with bounded first and second derivatives. These hypothesis are certainly not optimal.

Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space equipped with a filtration  $(\mathcal{F}_t)_{t \in [0,1]}$  (we shall only consider processes on the interval  $[0, 1]$  throughout) carrying a standard  $n$ -dimensional Brownian motion  $B$  starting from 0. We call *drift* any progressively measurable random process  $u$  taking values in  $\mathbb{R}^n$  and satisfying

$$\mathbb{E} \int_0^1 \|u(s)\|^2 ds < +\infty.$$

This is the class of processes for which the stochastic integral

$$\int_0^1 \langle u(s), dB(s) \rangle$$

is well defined. Let  $P$  be the semi-group associated to  $B$

$$P_t f(x) = \mathbb{E} f(x + B_t)$$

for every  $t \in \mathbb{R}_+$ ,  $x \in \mathbb{R}^n$  and every bounded (say) function  $f$ . This semi-group evolves according to the heat equation:

$$\partial_t P_t f(x) = \frac{1}{2} \Delta P_t f(x).$$

Our stochastic formula for the entropy states as follows.

**Theorem 1.** *Let  $\mu$  be a probability measure on  $\mathbb{R}^n$  having a density  $\rho$  with respect to  $\gamma_n$  satisfying the above technical assumptions. Then*

$$H_{\gamma_n}(\mu) = \frac{1}{2} \inf \left( \mathbb{E} \int_0^1 \|u(s)\|^2 ds \right),$$

where the infimum is taken over all bounded drifts  $u$  such that  $B(1) + \int_0^1 u(s) ds$  has law  $\mu$ . Besides, the infimum is attained at some drift  $v$  which solves the differential equation

$$v(t) = \nabla \log P_{1-t} \rho \left( B(t) + \int_0^t v(s) ds \right), \quad t \in [0, 1]. \quad (2)$$

Lastly, the optimal drift  $v$  is a martingale and  $\mathbb{E} v(t) = \int x d\mu(x)$  for all  $t \in [0, 1]$ .

We start the proof with a simple lemma.

**Lemma 2.** *Let  $G: [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  be continuous in the time variable and Lipschitz in the space variable. Then there exists a progressively measurable process  $Y$  satisfying*

$$Y'(t) = G(t, B(t) + Y(t)), \quad t \in [0, 1],$$

almost surely.

For fixed  $\omega \in \Omega$ , the equation above is an ordinary differential equation, the only issue is to show that the solution depends of  $\omega$  in a progressively measurable way. This can be dealt with as follows.

*Proof.* Equivalently, we need to show that the map

$$T: Z \mapsto \left( t \in [0, 1] \mapsto \int_0^t G(s, B(s) + Z(s)) \, ds \right)$$

has a progressively measurable fixed point. A direct (and standard) computation shows that  $T$  is a contraction for the distance induced by the norm

$$\sup_{(t, \omega) \in [0, 1] \times \Omega} \|Z(t, \omega)\| e^{-Kt},$$

where  $K$  is the Lipschitz constant of  $G$ . Recall that a process  $Z$  is said to be progressively measurable when for all  $t \in [0, 1]$  the map

$$(s, \omega) \in [0, t] \times \Omega \mapsto Z(s, \omega)$$

is measurable with respect to the  $\sigma$ -field  $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ . Let  $\mathcal{P}$  be the subspace of the above  $L^\infty$  space consisting of progressively measurable processes. Then clearly  $\mathcal{P}$  is a closed subspace and  $T(\mathcal{P}) \subset \mathcal{P}$ . Therefore  $T$  as a fixed point in  $\mathcal{P}$ .  $\square$

We now prove Theorem 1, most of the ideas are borrowed from Borell's article [5].

*Proof of Theorem 1.* Let  $u$  be a bounded drift such that  $B(1) + \int_0^1 u(s) \, ds$  has law  $\mu$ . Since  $u$  is bounded

$$\mathcal{E}_u(t) := \exp\left(-\int_0^t \langle u(s), dB(s) \rangle - \frac{1}{2} \int_0^t \|u(s)\|^2 \, ds\right)$$

is a martingale and Girsanov's formula applies, see for instance [12]. Under the measure

$$dQ = \mathcal{E}_u(1) \, dP$$

the process  $B(t) + \int_0^t u(s) \, ds$  is a Brownian motion on  $[0, 1]$ . In particular

$$\mathbb{E} \rho\left(B(1) + \int_0^1 u(s) \, ds\right) \mathcal{E}_u(1) = \int_{\mathbb{R}^n} \rho \, d\gamma_n = 1.$$

Taking logarithm and using Jensen's inequality we get

$$\mathbb{E} \log(\rho)\left(B(1) + \int_0^1 u(s) \, ds\right) \leq -\mathbb{E} \log(\mathcal{E}_u(1)).$$

Since  $B(1) + \int_0^1 u(s) \, ds$  has law  $\mu$ , the left hand side equals  $H_{\gamma_n}(\mu)$  and we get

$$H_{\gamma_n}(\mu) \leq \frac{1}{2} \mathbb{E} \int_0^1 \|u(s)\|^2 \, ds.$$

This argument shows that if a bounded drift  $v$  satisfies

$$\log(\rho)\left(B(1) + \int_0^1 v(s) \, ds\right) = \int_0^1 \langle v(s), dB(s) \rangle + \frac{1}{2} \int_0^1 \|v(s)\|^2 \, ds \quad (3)$$

almost surely, then  $B(1) + \int_0^1 v(s) \, ds$  has law  $\mu$  and we have equality in the above Jensen inequality, hence

$$H_{\gamma_n}(\mu) = \frac{1}{2} \mathbb{E} \int_0^1 \|v(s)\|^2 \, ds.$$

Let  $F(t, x) = \log(P_{1-t}\rho)(x)$ . The hypothesis made on  $\rho$  easily imply that  $\nabla F$  is bounded, continuous in  $t$  and Lipschitz in  $x$  on  $[0, 1] \times \mathbb{R}^n$ . By Lemma 2 there exists a progressively measurable process  $Y$  satisfying

$$Y'(t) = \nabla F(t, B(t) + Y(t)), \quad t \in [0, 1].$$

Then  $v = Y'$  is progressively measurable, bounded and satisfies (2). We claim that  $v$  satisfies (3). Indeed, let  $M(t) = F(t, B(t) + \int_0^t v(s) \, ds)$  so that  $M(0) = \log(P_1\rho)(0) = 0$  and

$$M(1) = \log(\rho)\left(B(1) + \int_0^1 v(s) \, ds\right).$$

An easy computation shows that  $\partial_t F = -(\Delta F + \|\nabla F\|^2)/2$ . Then Itô's formula and (2) yield (omitting variables)

$$\begin{aligned} dM &= \partial_t F \, dt + \langle \nabla F, dB \rangle + \langle \nabla F, v \rangle \, dt + \frac{1}{2} \Delta F \, dt \\ &= \langle v, dB \rangle + \frac{1}{2} \|v\|^2 \, dt, \end{aligned}$$

hence (3). Another application of Itô's formula and (2) show that  $v$  is a semi-martingale with variation (omitting variables again)

$$dv = \partial_t \nabla F \, dt + \nabla^2 F (dB + v \, dt) + \frac{1}{2} \Delta(\nabla F) \, dt. \quad (4)$$

Observe that

$$\begin{aligned} \partial_t \nabla F &= \nabla \partial_t F = -\frac{1}{2} (\nabla(\Delta F) + \nabla(\|\nabla F\|^2)) \\ &= -\frac{1}{2} \Delta(\nabla F) - \nabla^2 F(\nabla F) \end{aligned}$$

and recall that  $\nabla F = v$ . Therefore the absolutely continuous part in (4) cancels out and  $v$  is a local martingale. It is actually a true martingale since it is bounded. In particular its expectation  $\mathbb{E} v(s)$  is constant on  $[0, 1]$ . Lastly, from the equality

$$\mu = B(1) + \int_0^1 v(s) \, ds$$

in law, we get  $\mathbb{E} v(s) = \int x \, d\mu(x)$  for all  $s \in [0, 1]$ . □

### 3 Applications

Throughout, measures are assumed to satisfy the technical assumptions of the previous section. All the inequalities that we prove can be extended to more general measures by density arguments left to the reader.

#### 3.1 Transportation cost inequality

Recall the definition of the  $L_2$  Wasserstein distance: if  $\mu$  and  $\nu$  are probability measures on  $\mathbb{R}^n$

$$W_2(\mu, \nu) = \inf \left( \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 d\pi(x, y) \right)^{1/2}.$$

where the infimum is taken over all couplings  $\pi$  of  $\mu$  and  $\nu$ , namely all probability measures on the product space  $\mathbb{R}^n \times \mathbb{R}^n$  having marginals  $\mu$  and  $\nu$ . There is a huge literature about this optimisation problem, usually referred to as Monge-Kantorovitch problem. Our modest purpose is to give a simple proof of the following inequality due to Talagrand [13]. The reader willing to know more about optimal transport is referred to the book [14].

**Theorem 3.** *Any probability measure  $\mu$  on  $\mathbb{R}^n$  satisfies*

$$W_2(\mu, \gamma_n)^2 \leq 2 H_{\gamma_n}(\mu).$$

*Proof.* Let  $B$  be a standard Brownian motion and  $v$  be the optimal drift for  $\mu$ . Namely, the random vector  $X := B(1) + \int_0^1 v(s) ds$  has law  $\mu$  and

$$H_{\gamma_n}(\mu) = \frac{1}{2} \mathbb{E} \int_0^1 \|v(s)\|^2 ds.$$

Since  $(X, B(1))$  is a coupling of  $(\mu, \gamma_n)$  and by definition of the Wasserstein distance

$$W_2(\mu, \gamma_n)^2 \leq \mathbb{E} \|X - B(1)\|^2.$$

On the other hand, by Jensen's inequality

$$\mathbb{E} \|X - B(1)\|^2 = \mathbb{E} \left\| \int_0^1 v(s) ds \right\|^2 \leq \mathbb{E} \int_0^1 \|v(s)\|^2 ds. \quad \square$$

#### 3.2 Logarithmic Sobolev inequality

Let  $\mu$  be a probability measure on  $\mathbb{R}^n$ . The relative Fisher information of  $\mu$  is defined by

$$I_{\gamma_n}(\mu) = \int_{\mathbb{R}^n} \frac{\|\nabla \rho\|^2}{\rho} d\gamma_n = \int_{\mathbb{R}^n} \|\nabla \log(\rho)\|^2 d\mu,$$

where  $\rho$  is the density of  $\mu$  with respect to  $\gamma_n$ . The logarithmic Sobolev inequality for the Gaussian measure, due to Gross [11], states as follows.

**Theorem 4.** *For all probability measure on  $\mathbb{R}^n$*

$$H_{\gamma_n}(\mu) \leq \frac{1}{2} I_{\gamma_n}(\mu).$$

See the book [1] for an overview on this inequality, we only give a short proof based on Theorem 1.

*Proof.* Let  $B$  be a Brownian motion and let  $v$  be the optimal drift for  $\mu$ . Recall that  $v$  is a martingale, so  $\|v\|^2$  is a sub-martingale, hence the inequality

$$\mathbb{H}_{\gamma_n}(\mu) = \frac{1}{2} \int_0^1 \mathbb{E} \|v(s)\|^2 ds \leq \frac{1}{2} \mathbb{E} \|v(1)\|^2.$$

On the other hand  $v$  solves the equation

$$v(t) = \nabla \log(P_{1-t}\rho) \left( B(t) + \int_0^t v(s) ds \right), \quad t \in [0, 1],$$

and  $B(1) + \int_0^1 v(s) ds$  has law  $\mu$ . Therefore

$$\mathbb{E} \|v(1)\|^2 = \mathbb{E} \left\| \nabla \log(\rho) \left( B(1) + \int_0^1 v(s) ds \right) \right\|^2 = \mathbb{I}_{\gamma_n}(\mu). \quad \square$$

### 3.3 Shannon's inequality

In this section we prove a classical information inequality known as Shannon's inequality, see [10] for an overview on the topic. Let  $\mathcal{L}_n$  be the Lebesgue measure. In the sequel, when  $X$  is a random vector  $\mathbb{H}(X)$  means  $\mathbb{H}(\mu)$  where  $\mu$  is the law of  $X$ .

**Theorem 5.** *Let  $X, Y$  be independant random vectors on  $\mathbb{R}^n$  and  $\lambda \in (0, 1)$*

$$\mathbb{H}_{\mathcal{L}_n}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq (1-\lambda)\mathbb{H}_{\mathcal{L}_n}(X) + \lambda\mathbb{H}_{\mathcal{L}_n}(Y). \quad (5)$$

*Proof.* Observe that

$$\mathbb{H}_{\gamma_n}(X) = \mathbb{H}_{\mathcal{L}_n}(X) + \frac{n}{2} \log(2\pi) + \frac{1}{2} \mathbb{E} \|X\|^2$$

and similarly for  $Y$ . Since  $X$  and  $Y$  are independant, it is easily seen that (5) is equivalent to

$$\begin{aligned} \mathbb{H}_{\gamma_n}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) &\leq (1-\lambda)\mathbb{H}_{\gamma_n}(X) + \lambda\mathbb{H}_{\gamma_n}(Y) \\ &\quad + \sqrt{\lambda(1-\lambda)} \langle \mathbb{E}X, \mathbb{E}Y \rangle \end{aligned} \quad (6)$$

Let  $B$  be a standard Brownian motion on  $\mathbb{R}^n$  and  $u$  be the optimal drift for  $X$ . Recall that it implies in particular that

$$\mathbb{E} u(s) = \mathbb{E} X, \quad s \in [0, 1].$$

Similarly, let  $C$  be a standard Brownian motion on  $\mathbb{R}^n$ , independant of  $B$  and  $v$  be the optimal drift for  $Y$  (built out of  $C$  of course). Now let  $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$  and  $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$ . Then  $W$  is a standard Brownian motion and, since  $(B, u)$  is independant of  $(C, v)$ , the random vector  $W(1) + \int_0^1 w(s) ds$  equals  $\sqrt{1-\lambda}X + \sqrt{\lambda}Y$  in law. Therefore

$$\mathbb{H}_{\gamma_n}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq \frac{1}{2} \mathbb{E} \int_0^1 \|w(s)\|^2 ds.$$

Since  $u$  and  $v$  are independant

$$\mathbb{E} \|w(s)\|^2 = (1-\lambda) \mathbb{E} \|u(s)\|^2 + \lambda \mathbb{E} \|v(s)\|^2 + 2\sqrt{\lambda(1-\lambda)} \langle \mathbb{E}X, \mathbb{E}Y \rangle.$$

Integrating this equality on  $[0, 1]$  yields (6).  $\square$

### 3.4 Brascamp-Lieb inequality

Let us focus on a family of inequalities dating back to Brascamp and Lieb's article [7] on optimal constants in Young's inequality. Since then a number of nice alternate proofs have been discovered, see [4, 8] and the survey article [2], and it would also be possible to use Borell's formula. Still, we believe the following approach based on Theorem 1 is worth noticing.

Let  $E$  be a Euclidean space, let  $E_1, \dots, E_m$  be subspaces and for all  $i$  let  $P_i$  be the orthogonal projection with range  $E_i$ . Let  $c_1, \dots, c_m$  be positive numbers. The crucial hypothesis is the so-called frame condition

$$\sum_{i=1}^m c_i P_i = \text{id}_E. \quad (7)$$

Let  $x \in E$ , we then have  $\|x\|^2 = \langle \sum c_i P_i x, x \rangle$ . Since  $P_i$  is an orthogonal projection we obtain

$$\|x\|^2 = \sum_{i=1}^m c_i \|P_i x\|^2. \quad (8)$$

Now let  $x_1, \dots, x_m$  belong to  $E_1, \dots, E_m$  respectively. For any element  $y \in E$ , using the Cauchy-Schwarz inequality and (8) we get

$$\begin{aligned} \left\langle \sum_{i=1}^m c_i x_i, y \right\rangle &= \sum_{i=1}^m c_i \langle x_i, P_i y \rangle \\ &\leq \left( \sum_{i=1}^m c_i \|x_i\|^2 \right)^{1/2} \left( \sum_{i=1}^m c_i \|P_i y\|^2 \right)^{1/2} \\ &= \left( \sum_{i=1}^m c_i \|x_i\|^2 \right)^{1/2} \|y\|. \end{aligned}$$

Hence

$$\left\| \sum_{i=1}^m c_i x_i \right\|^2 \leq \sum_{i=1}^m c_i \|x_i\|^2. \quad (9)$$

Throughout this subsection we let  $\gamma$  be the standard Gaussian measure on  $E$  and  $\gamma_i$  be the standard Gaussian measure on  $E_i$ , for all  $i \in \{1, \dots, m\}$ . The Brascamp-Lieb inequality states as follows.

**Theorem 6.** *Under the frame condition, for all  $f_i: E_i \rightarrow \mathbb{R}$ , and  $g: E \rightarrow \mathbb{R}$  satisfying*

$$g(x) \leq \sum_{i=1}^m c_i f_i(P_i x)$$

for all  $x \in E$ , we have

$$L_\gamma(g) \leq \sum_{i=1}^m c_i L_{\gamma_i}(f_i).$$

For the many consequences of this inequality, we refer to [9]. As observed by Carlen and Cordero in [8], this theorem admits an equivalent formulation in terms of entropy.



**Theorem 7.** *Under the frame condition, we have*

$$\mathbb{H}_\gamma(X) \geq \sum_{i=1}^m c_i \mathbb{H}_{\gamma_i}(P_i X)$$

for all random vector  $X$  on  $E$ .

*Proof.* Let  $X$  be a random vector. Let  $B$  be a standard Brownian motion on  $E$ , and  $u$  be the optimal drift for  $X$ . Then  $P_i B$  is a standard Brownian motion on  $E_i$  and since  $B(1) + \int_0^1 u(s) ds = X$  in law, we have  $P_i B(1) + \int_0^1 P_i u(s) ds = P_i X$  in law. Therefore

$$\mathbb{H}_{\gamma_i}(P_i X) \leq \frac{1}{2} \mathbb{E} \int_0^1 \|P_i u(s)\|^2 ds.$$

Using (8), we get

$$\begin{aligned} \sum_{i=1}^m c_i \mathbb{H}_{\gamma_i}(P_i X) &\leq \frac{1}{2} \mathbb{E} \int_0^1 \sum_{i=1}^m c_i \|P_i u(s)\|^2 ds \\ &= \frac{1}{2} \mathbb{E} \int_0^1 \|u(s)\|^2 ds = \mathbb{H}_\gamma(X). \end{aligned} \quad \square$$

The Brascamp-Lieb inequality admits a reversed version due to Barthe [3].

**Theorem 8.** *Assuming (7), for all  $f_i: E_i \rightarrow \mathbb{R}$  and  $h: E \rightarrow \mathbb{R}$  satisfying*

$$\sum_{i=1}^m c_i f(x_i) \leq h\left(\sum_{i=1}^m c_i x_i\right) \quad (10)$$

for all  $(x_1, \dots, x_m) \in E_1 \times \dots \times E_m$ , we have

$$\sum_{i=1}^m c_i \mathbb{L}_{\gamma_i}(f_i) \leq \mathbb{L}_\gamma(h).$$

This reversed version has also many implications, it is in particular related to the Brunn-Minkowski inequality. As far as the author know, there is no entropic version of this reversed inequality in the litterature, here is an attempt.

**Theorem 9.** *Under condition (7), for all random vectors  $X_1, \dots, X_m$  on  $E_1, \dots, E_m$  respectively, there exist  $Y_1, \dots, Y_m$  satisfying  $X_i = Y_i$  in law for all  $i \in \{1, \dots, m\}$  and*

$$\mathbb{H}_\gamma\left(\sum_{i=1}^m c_i Y_i\right) \leq \sum_{i=1}^m c_i \mathbb{H}_{\gamma_i}(X_i).$$

*Proof.* We first prove Theorem 9, then we prove that it implies Theorem 8.

Let  $X_1, \dots, X_m$  be random vectors on  $E_1, \dots, E_m$  respectively. Again let  $B$  be a standard Brownian motion on  $E$ . Let  $i \in \{1, \dots, m\}$ , since  $P_i B$  is a standard Brownian motion on  $E_i$ , there exists a drift  $v_i$  satisfying  $P_i B(1) + \int_0^1 v_i(s) ds = X_i$  in law and

$$\mathbb{H}_{\gamma_i}(X_i) = \frac{1}{2} \mathbb{E} \int_0^1 \|P_i v_i(s)\|^2 ds.$$

Now let  $Y_i = P_i B(1) + \int_0^1 v_i(s) ds$ . By the frame condition

$$\sum_{i=1}^m c_i Y_i = B(1) + \int_0^1 \sum_{i=1}^m c_i v_i(s) ds.$$

Using Theorem 1 and (9) we get

$$\begin{aligned} \mathbb{H}_\gamma\left(\sum_{i=1}^m c_i Y_i\right) &\leq \frac{1}{2} \mathbb{E} \int_0^1 \left\| \sum_{i=1}^m c_i v_i(s) \right\|^2 ds \\ &\leq \frac{1}{2} \mathbb{E} \int_0^1 \sum_{i=1}^m c_i \|v_i(s)\|^2 ds \\ &= \sum_{i=1}^m c_i \mathbb{H}(X_i). \end{aligned}$$

We now prove Theorem 8, the argument relies on the duality (1). Let  $f_1, \dots, f_m, h$  satisfy (10). Let  $X_1, \dots, X_m$  be random vectors on  $E_1, \dots, E_m$  and  $Y_1, \dots, Y_m$  be the vectors given by Theorem 9. Writing

$$\begin{aligned} \sum_{i=1}^m c_i (\mathbb{E} f_i(X_i) - \mathbb{H}_{\gamma_i}(X_i)) &= \mathbb{E} \left( \sum_{i=1}^m c_i f_i(Y_i) \right) - \sum_{i=1}^m c_i \mathbb{H}_{\gamma_i}(X_i) \\ &\leq \mathbb{E} h \left( \sum_{i=1}^m c_i Y_i \right) - \mathbb{H}_\gamma \left( \sum_{i=1}^m c_i Y_i \right) \\ &\leq L_\gamma(h). \end{aligned}$$

and taking suprema in the vectors  $X_i$  yield the desired inequality.  $\square$

*Remark.* The equivalence between Theorem 6 and Theorem 7 relies only on the duality relation (1), it has nothing to do with the frame condition nor the Gaussian measure. We refer to [8] for a general statement. Similarly, it is possible to formulate a general entropic reversed Brascamp-Lieb inequality along the lines of Theorem 9 and prove that it implies a general reversed Brascamp-Lieb inequality. It is not clear though (at least to the author) whether the converse implication holds true.

## References

- [1] C. Ané et al. Sur les inégalités de Sobolev logarithmiques. Panoramas et Synthèses 10. SMF, Paris, 2000.
- [2] K. Ball. Convex geometry and functional analysis, in *Handbook of the Geometry of Banach Spaces*, vol. I, edited by W.B. Johnson and J. Lindenstrauss, North Holland (2001) 161–194.
- [3] F. Barthe. On a reverse form of the Brascamp-Lieb inequality, *Invent. Math.* 134 (2) (1998) 335–361.
- [4] F. Barthe and N. Huet. On Gaussian Brunn-Minkowski inequalities, *Studia Math.* 191 (3) (2009) 283–304.

- [5] C. Borell. Diffusion equations and geometric inequalities, *Potential Anal.* 12 (1) (2000) 49–71.
- [6] C. Borell. Isoperimetry, log-concavity, and elasticity of option prices, in *New Directions in Mathematical Finance*, edited by P. Wilmott and H. Rasmussen, Wiley (2002) 73–91.
- [7] H. J. Brascamp and E. H. Lieb. Best constants in Youngs inequality, its converse and its generalization to more than three functions. *Adv. Math.* 20 (1976) 151–173.
- [8] E. Carlen and D. Cordero-Erausquin. Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities, *Geom. Funct. Anal.* 19 (2009) 373–405.
- [9] D. Cordero-Erausquin and M. Ledoux. The geometry of Euclidean convolution inequalities and entropy, *Proc. A.M.S.* (to appear).
- [10] A. Dembo, T.M. Cover and J.A. Thomas. Information Theoretic Inequalities, *IEEE Transactions on Information Theory* 37 (6) (1991) 1501–1518.
- [11] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* 97 (4) (1975) 1061–1083.
- [12] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, 3rd edition, Springer, 1999.
- [13] M. Talagrand. Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.* 6 (3) (1996) 587–600.
- [14] C. Villani. *Topics in optimal transportation*. Graduate studies in mathematics 58. AMS, 2003.