

A connection of the Brascamp-Lieb inequality with Skorokhod embedding

Yuu Hariya
Mathematical Institute, Tohoku University

1 Introduction and main result

The Brascamp-Lieb moment inequality plays an important role in statistical mechanics, such as in the analysis of gradient interface models; see, e.g., [5, 4, 6]. It asserts that centered moments of a distribution with log-concave density relative to a Gaussian distribution do not exceed those of the original Gaussian's; it is used to derive the tightness of finite-volume Gibbs measures describing the static interface, strict convexity of the associated surface tension, and so on.

The Skorokhod embedding problem is to find a stopping time T for one-dimensional Brownian motion B such that $B(T)$ is distributed as a given probability measure on \mathbb{R} . The problem was proposed by Skorokhod [11] and a number of solutions have been constructed since then ([9]); they are applied to the proof of Donsker's invariance principle, robust pricings of options in mathematical finance (see, e.g., [8]), and so on.

In this article we reveal a connection between the Brascamp-Lieb inequality and the Skorokhod embedding of Bass [1]; as a by-product, we also provide error bounds for the inequality in terms of variance by applying the Itô-Tanaka formula. Let Y be an n -dimensional Gaussian random variable defined on a probability space (Ω, \mathcal{F}, P) with law ν . Let X be an n -dimensional random variable on (Ω, \mathcal{F}, P) , whose law μ is given in the form

$$\mu(dx) = \frac{1}{Z} e^{-V(x)} \nu(dx), \quad Z := \int_{\mathbb{R}^n} e^{-V(x)} \nu(dx) \quad (1.1)$$

with $V : \mathbb{R}^n \rightarrow \mathbb{R}$ a convex function. In what follows, we fix $v \in \mathbb{R}^n$ ($v \neq 0$) arbitrarily. For a one-dimensional random variable ξ , we denote its variance by $\text{var}(\xi)$. We set $a := \text{var}(v \cdot Y)$. Here $a \cdot b$ denotes the inner product of $a, b \in \mathbb{R}^n$. We also set

$$p(t; x) := \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right), \quad t > 0, x \in \mathbb{R}.$$

We state the main result of this article.

Theorem 1.1. *For every convex function ψ on \mathbb{R} , we have the following.*

(i) *It holds that*

$$E[\psi(v \cdot Y - E[v \cdot Y])] \geq E[\psi(v \cdot X - E[v \cdot X])]. \quad (1.2)$$

More precisely, we have

$$E[\psi(v \cdot Y - E[v \cdot Y])] \geq E[\psi(v \cdot X - E[v \cdot X])] + \frac{1}{2} \int_{\mathbb{R}} \psi''(dx) \int_0^{a^{-1}(a - \text{var}(v \cdot X))^2} ds \mathfrak{p}\left(s; \sqrt{x^2 + a}\right), \quad (1.3)$$

where $\psi''(dx)$ denotes the second derivative of ψ in the sense of distribution.

(ii) For every $p > 1$, it holds that

$$E[\psi(v \cdot Y - E[v \cdot Y])] \leq E[\psi(v \cdot X - E[v \cdot X])] + C(a, \psi, q) (a - \text{var}(v \cdot X))^{\frac{1}{2p}}. \quad (1.4)$$

Here $C(a, \psi, q) \in [0, \infty]$ is given by

$$C(a, \psi, q) = (a(1+q))^{\frac{1}{2q}} \int_{\mathbb{R}} \psi''(dx) \mathfrak{p}\left(1; \frac{x}{\sqrt{a(1+q)}}\right)$$

with q the conjugate of p : $p^{-1} + q^{-1} = 1$. Note that $a - \text{var}(v \cdot X) \geq 0$ by (1.2).

The above inequalities (1.2)–(1.4) are understood to hold also in the case that both sides of them are infinity.

Remark 1.1. (1) The inequality (1.2) is called the Brascamp-Lieb inequality. It was originally proved by Brascamp and Lieb [2, Theorem 5.1] in the case $\psi(x) = |x|^p$, $p \geq 1$; it was then extended to general convex ψ 's by Caffarelli [3, Corollary 6] based on analyses of optimal transportation between μ and ν , and the related Monge-Ampère equation. (2) In the case $\psi''(\mathbb{R}) < \infty$, letting $p \rightarrow 1$ in (1.4) yields

$$E[\psi(v \cdot Y - E[v \cdot Y])] - E[\psi(v \cdot X - E[v \cdot X])] \leq \frac{1}{\sqrt{2\pi}} \psi''(\mathbb{R}) (a - \text{var}(v \cdot X))^{\frac{1}{2}}.$$

Taking $\psi(x) = |x|$ and some manipulation show that

$$\frac{E[|v \cdot X - E[v \cdot X]|]}{\text{var}(v \cdot X)} \geq \frac{1}{\sqrt{2\pi \text{var}(v \cdot Y)}}$$

for any convex V .

(3) In the case $\psi(x) = x^2$, inequalities (1.3) and (1.4) hold merely in the trivial manner. We remark that if $V \in C^2(\mathbb{R}^n)$, then $\text{var}(v \cdot X)$ admits the upper bound

$$\int_{\mathbb{R}^n} \mu(dx) v \cdot (\Sigma^{-1} + D^2V(x))^{-1} v,$$

which is less than or equal to $a \equiv v \cdot \Sigma v$; see [2, Theorem 4.1]. Here Σ is the covariance matrix of the Gaussian ν and D^2V is the Hessian of V .

In the sequel, for every real-valued function f on \mathbb{R} and for every $x \in \mathbb{R}$, we denote respectively by $f'_+(x)$ and $f'_-(x)$ the right- and left-derivatives of f at x if they exist. For each $x, y \in \mathbb{R}$, we write $x \wedge y = \min\{x, y\}$ and $x^+ = \max\{x, 0\}$.

2 Proof of Theorem 1.1

In this section we give a proof of Theorem 1.1. Without loss of generality, we may assume that ν is centered: $E[Y] = 0$. Moreover, Theorem 4.3 of [2] reduces the proof to the case $n = 1$; that is, the density of the law $P \circ (v \cdot X)^{-1}$ relative to the one-dimensional Gaussian measure $P \circ (v \cdot Y)^{-1}$ is log-concave. Therefore in what follows, we take the Gaussian measure ν in (1.1) as

$$\nu(dx) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{x^2}{2a}\right) dx, \quad x \in \mathbb{R},$$

and V as a convex function on \mathbb{R} . We accordingly write X and Y for $v \cdot X$ and $v \cdot Y$, respectively; that is, X is distributed as μ and Y as ν .

2.1 Proof of (1.2)

In this subsection we prove the inequality (1.2) in Theorem 1.1. Let F_μ denote the distribution function of μ :

$$F_\mu(x) := \frac{1}{Z} \int_{-\infty}^x e^{-V(y)} \nu(dy), \quad x \in \mathbb{R}.$$

We also set

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{1}{2}y^2\right) dy, \quad x \in \mathbb{R},$$

and

$$g := F_\mu^{-1} \circ \Phi. \tag{2.1}$$

Here $F_\mu^{-1} : (0, 1) \rightarrow \mathbb{R}$ is the inverse function of F_μ . It is obvious that g is strictly increasing. By convexity of V we have moreover

Lemma 2.1. *It holds that $g'(x) \leq \sqrt{a}$ for all $x \in \mathbb{R}$.*

We postpone the proof of this lemma to Subsection 2.3. Once this lemma is shown, the proof of (1.2) is straightforward from the Skorokhod embedding of Bass [1]; for other types of embeddings, refer to the detailed survey [9] by Oblój. Let $\{W_t\}_{t \geq 0}$ be a standard one-dimensional Brownian motion on (Ω, \mathcal{F}, P) .

Proof of (1.2). Note that $g(W_1)$ is distributed as μ . Applying Clark's formula to $g(W_1)$ yields

$$g(W_1) - E[g(W_1)] = \int_0^1 a(s, W_s) dW_s \quad P\text{-a.s.},$$

where for $0 \leq s \leq 1$ and $y \in \mathbb{R}$,

$$\begin{aligned} a(s, y) &:= \frac{\partial}{\partial y} E[g(y + W_{1-s})] \\ &= E[g'(y + W_{1-s})]. \end{aligned} \tag{2.2}$$

By the Dambis-Dubins-Schwarz theorem (see, e.g., [10, Theorem V.1.6]), there exists a Brownian motion $\{B(t)\}_{t \geq 0}$ on (Ω, \mathcal{F}, P) such that

$$\int_0^t a(s, W_s) dW_s = B \left(\int_0^t a(s, W_s)^2 ds \right) \quad \text{for all } 0 \leq t \leq 1, P\text{-a.s.}$$

We know from [1] that $T := \int_0^1 a(s, W_s)^2 ds$ is a stopping time in the natural filtration of B . Moreover, by (2.2) and Lemma 2.1, we have $T \leq \mathbf{a}$ P -a.s. We denote by $\{L_t^x\}_{t \geq 0, x \in \mathbb{R}}$ the local time process of B . For every $x \in \mathbb{R}$, Tanaka's formula yields

$$E[(B(\mathbf{a}) - x)^+] = E[(B(T) - x)^+] + \frac{1}{2} E[L_{\mathbf{a}}^x - L_T^x], \quad (2.3)$$

$$E[(x - B(\mathbf{a}))^+] = E[(x - B(T))^+] + \frac{1}{2} E[L_{\mathbf{a}}^x - L_T^x]. \quad (2.4)$$

From (2.3) and (2.4), it follows that for every convex ψ ,

$$E[\psi(B(\mathbf{a}))] = E[\psi(B(T))] + \frac{1}{2} \int_{\mathbb{R}} \psi''(dx) E[L_{\mathbf{a}}^x - L_T^x]. \quad (2.5)$$

Indeed, by Fubini's theorem,

$$\begin{aligned} & \int_{[0, \infty)} \psi''(dx) E[(B(\mathbf{a}) - x)^+] + \int_{(-\infty, 0)} \psi''(dx) E[(x - B(\mathbf{a}))^+] \\ &= E[\psi(B(\mathbf{a})) - \psi'_-(0)B(\mathbf{a}) - \psi(0)] \\ &= E[\psi(B(\mathbf{a}))] - \psi(0), \end{aligned}$$

which is equal, by (2.3), (2.4) and $E[B(T)] = 0$, to the right-hand side of (2.5) with $\psi(0)$ subtracted. Hence (2.5) holds. As $T \leq \mathbf{a}$ a.s. and $\psi'' \geq 0$, it is immediate from (2.5) that

$$E[\psi(B(\mathbf{a}))] \geq E[\psi(B(T))], \quad (2.6)$$

which is nothing but (1.2) since

$$B(T) = g(W_1) - E[g(W_1)] \stackrel{(d)}{=} X - E[X] \quad (2.7)$$

and $B(\mathbf{a}) \stackrel{(d)}{=} Y$. The proof is complete. \square

Remark 2.1. (1) For any convex ψ such that $\int_0^\cdot \psi'_-(B(s)) dB(s)$ is a martingale, the identity (2.5) is immediate from the Itô-Tanaka formula.

(2) For any convex ψ such that $E[|\psi(B(\mathbf{a}))|] < \infty$ (i.e., $E[\psi(B(\mathbf{a}))] < \infty$), the inequality (2.6) follows readily from the optional sampling theorem applied to the submartingale $\{\psi(B(t))\}_{0 \leq t \leq \mathbf{a}}$.

2.2 Proof of (1.3) and (1.4)

In this subsection we prove the inequalities (1.3) and (1.4) in Theorem 1.1. We keep the notation in the previous subsection. By (2.5), the proof is reduced to showing the following proposition.

Proposition 2.1. (1) *It holds that for all $x \in \mathbb{R}$,*

$$E[L_a^x - L_T^x] \geq \int_0^{a^{-1}(a - \text{var}(X))^2} ds \mathfrak{p}\left(s; \sqrt{x^2 + a}\right). \quad (2.8)$$

(2) *Let $p > 1$ and let q be such that $p^{-1} + q^{-1} = 1$. It holds that for all $x \in \mathbb{R}$,*

$$E[L_a^x - L_T^x] \leq 2(a(1+q))^{\frac{1}{2q}} \mathfrak{p}\left(1; \frac{x}{\sqrt{a(1+q)}}\right) (a - \text{var}(X))^{\frac{1}{2p}}. \quad (2.9)$$

To prove these estimates, we prepare a lemma; for the proof, see [7].

Lemma 2.2. *For every $t > 0$ and $x \in \mathbb{R}$, we have*

$$E[L_t^x] = \int_0^t ds \mathfrak{p}(s; x) \quad (2.10)$$

$$= 2 \int_0^\infty dy (y - |x|)^+ \mathfrak{p}(t; y) \quad (2.11)$$

$$= 2 \int_0^\infty dy (\sqrt{t}y - |x|)^+ \mathfrak{p}(1; y). \quad (2.12)$$

The proof of the proposition proceeds as follows: Recall $T \leq a$ a.s.

Proof of Proposition 2.1. (1) By the strong Markov property of Brownian motion,

$$E[L_a^x - L_T^x] = E\left[E[L_{a-t}^{x-z}] \Big|_{(t,z)=(T,B(T))}\right]. \quad (2.13)$$

By (2.12), this is rewritten as

$$2E\left[\int_0^\infty dy (\sqrt{a-T}y - |x - B(T)|)^+ \mathfrak{p}(1; y)\right].$$

Using Fubini's theorem and Jensen's inequality, we estimate this from below by

$$2 \int_0^\infty dy \left(E[\sqrt{a-T}]y - E[|x - B(T)|]\right)^+ \mathfrak{p}(1; y).$$

By the optional sampling theorem and Schwarz's inequality,

$$\begin{aligned} E[|x - B(T)|] &\leq E[|x - B(a)|] \\ &\leq \sqrt{x^2 + a}. \end{aligned}$$

Plugging this and using the identity between (2.12) and (2.10) lead to

$$E[L_a^x - L_T^x] \geq \int_0^{E[\sqrt{a-T}]^2} ds \mathfrak{p}(s; \sqrt{x^2 + a}).$$

Since $\sqrt{a-t} \geq a^{-1/2}(a-t)$ for $0 \leq t \leq a$, we see that

$$\begin{aligned} E[\sqrt{a-T}]^2 &\geq a^{-1}(a - E[T])^2 \\ &= a^{-1}(a - \text{var}(X))^2, \end{aligned}$$

where the equality follows from Wald's identity

$$E[T] = E[B(T)^2] \tag{2.14}$$

and from (2.7). This proves (2.8).

(2) First we show that for every $t > 0$ and $x \in \mathbb{R}$,

$$E\left[\int_0^t ds \mathfrak{p}(s; |x - B(T)|)\right] \leq \int_0^{a+t} ds \mathfrak{p}(s; x). \tag{2.15}$$

By the identity between (2.10) and (2.11), and by Fubini's theorem, the left-hand side is equal to

$$2 \int_0^\infty dy E[(y - |x - B(T)|)^+] \mathfrak{p}(t; y). \tag{2.16}$$

We note the identity $(y - |x - z|)^+ = (z - x + y)^+ \wedge (x + y - z)^+$ for $z \in \mathbb{R}$, to bound the expectation in the integrand from above by

$$\begin{aligned} &E[(B(T) - x + y)^+] \wedge E[(x + y - B(T))^+] \\ &\leq E[(B(a) - x + y)^+] \wedge E[(x + y - B(a))^+] \\ &= E[(B(a) + y - |x|)^+]. \end{aligned}$$

Here for the inequality, we used the optional sampling theorem; the equality follows from the monotonicity of $E[(B(a) - x + y)^+]$ in x and the symmetry in the sense that $E[(B(a) - (-x) + y)^+] = E[(x + y - B(a))^+]$. Therefore (2.16) is dominated by

$$\begin{aligned} &2 \int_0^\infty dy \int_{\mathbb{R}} dz (z + y - |x|)^+ \mathfrak{p}(a; z) \mathfrak{p}(t; y) \\ &= 2 \int_{\mathbb{R}} du (\sqrt{a+tu} - |x|)^+ \mathfrak{p}(1; u) \int_{-\infty}^{\sqrt{a^{-1}t}u} dv \mathfrak{p}(1; v) \\ &\leq 2 \int_0^\infty du (\sqrt{a+tu} - |x|)^+ \mathfrak{p}(1; u), \end{aligned}$$

where we changed variables with $u = \frac{z+y}{\sqrt{a+t}}$ and $v = \frac{tz-ay}{\sqrt{at(a+t)}}$ for the equality. Now (2.15) follows from the identity between (2.12) and (2.10).

By (2.13), (2.10) and Hölder's inequality,

$$\begin{aligned} E[L_a^x - L_T^x] &\leq \left(\frac{1}{2\pi}\right)^{\frac{1}{2p}} E\left[\int_0^{a-T} \frac{ds}{\sqrt{s}}\right]^{\frac{1}{p}} E\left[\int_0^a ds \rho(s; \sqrt{q}|x - B(T)|)\right]^{\frac{1}{q}} \\ &= \left(\frac{2}{\pi}\right)^{\frac{1}{2p}} q^{\frac{1}{2q}} E\left[\sqrt{a-T}\right]^{\frac{1}{p}} E\left[\int_0^{aq^{-1}} ds \rho(s; |x - B(T)|)\right]^{\frac{1}{q}}. \end{aligned}$$

By Jensen's inequality, and by (2.14) and (2.7),

$$E\left[\sqrt{a-T}\right] \leq (a - E[T])^{\frac{1}{2}} = (a - \text{var}(X))^{\frac{1}{2}}.$$

Moreover, by (2.15),

$$\begin{aligned} E\left[\int_0^{aq^{-1}} ds \rho(s; |x - B(T)|)\right] &\leq \int_0^{a(1+q^{-1})} ds \rho(s; x) \\ &\leq \left(\frac{2a(1+q)}{\pi q}\right)^{\frac{1}{2}} \exp\left\{-\frac{qx^2}{2a(1+q)}\right\}. \end{aligned}$$

Combining these leads to (2.9) and ends the proof of Proposition 2.1. \square

Proof of (1.3) and (1.4). They are immediate from (2.5) and Proposition 2.1. \square

2.3 Proof of Lemma 2.1

We conclude this article with the proof of Lemma 2.1; the assertion itself is nothing but that of [3, Theorem 11] in one-dimension, where a more direct proof is possible. To begin with, note that we only need to consider the case $a = 1$; indeed, setting

$$\tilde{V}(x) := V(\sqrt{a}x), \quad \tilde{F}_\mu(x) := \frac{\sqrt{a}}{Z} \int_{-\infty}^x \exp\left(-\frac{1}{2}y^2 - \tilde{V}(y)\right) dy,$$

we have $F_\mu(x) = \tilde{F}_\mu(x/\sqrt{a})$, from which it follows that

$$F_\mu^{-1} \circ \Phi(x) = \sqrt{a} \tilde{F}_\mu^{-1} \circ \Phi(x).$$

Therefore the assertion of Lemma 2.1 is equivalent to

$$\left(\tilde{F}_\mu^{-1} \circ \Phi\right)' \leq 1.$$

Note that \tilde{V} remains convex.

From now on we let $a = 1$. We utilize the following:

Lemma 2.3. *It holds that for all $x \in \mathbb{R}$,*

$$F_\mu'(x) \geq \Phi'(x + V_-'(x)).$$

Proof. Since $V(y) - V(x) \geq V'_-(x)(y - x)$ for all $x, y \in \mathbb{R}$, we have

$$\begin{aligned} \frac{1}{F'_\mu(x)} &= \int_{\mathbb{R}} \exp\left(-\frac{1}{2}y^2 - V(y)\right) dy \times \exp\left(\frac{1}{2}x^2 + V(x)\right) \\ &\leq \exp\left(\frac{1}{2}x^2\right) \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}y^2 - V'_-(x)(y - x)\right\} dy \\ &= \exp\left\{\frac{1}{2}(x + V'_-(x))^2\right\} \times \sqrt{2\pi}, \end{aligned}$$

which is the desired inequality. \square

Lemma 2.1 follows readily from the above lemma.

Proof of Lemma 2.1. Since

$$g'(x) = \frac{\Phi'(x)}{F'_\mu \circ F_\mu^{-1}(\Phi(x))},$$

the assertion of the lemma with $a = 1$ is equivalent to

$$G(\xi) := F'_\mu \circ F_\mu^{-1}(\xi) - \Phi' \circ \Phi^{-1}(\xi) \geq 0 \quad \text{for all } \xi \in (0, 1). \quad (2.17)$$

First note that

$$G(0+) = G(1-) = 0 \quad (2.18)$$

because both $F'_\mu \circ F_\mu^{-1}$ and $\Phi' \circ \Phi^{-1}$ are zero at $\xi = 0+$ and $\xi = 1-$. Next, G is both right- and left-differentiable since F'_μ is and since F_μ^{-1} is monotone. Suppose now that G has a local minimum at some $\xi_0 \in (0, 1)$. Then $G'_-(\xi_0) \leq 0 \leq G'_+(\xi_0)$. Since

$$\begin{aligned} G'_\pm(\xi) &= \frac{(F'_\mu)_\pm'}{F'_\mu} \circ F_\mu^{-1}(\xi) + \Phi^{-1}(\xi) \\ &= -(x + V'_\pm(x)) \Big|_{x=F_\mu^{-1}(\xi)} + \Phi^{-1}(\xi), \end{aligned}$$

we have

$$(x + V'_+(x)) \Big|_{x=F_\mu^{-1}(\xi_0)} \leq \Phi^{-1}(\xi_0) \leq (x + V'_-(x)) \Big|_{x=F_\mu^{-1}(\xi_0)},$$

which indicates that

$$\Phi^{-1}(\xi_0) = (x + V'_-(x)) \Big|_{x=F_\mu^{-1}(\xi_0)}.$$

Hence by Lemma 2.3

$$G(\xi_0) = \{F'_\mu(x) - \Phi'(x + V'_-(x))\} \Big|_{x=F_\mu^{-1}(\xi_0)} \geq 0.$$

Combining this observation with (2.18), we conclude (2.17). The proof is complete. \square

References

- [1] R.F. Bass, Skorokhod imbedding via stochastic integrals, in Séminaire de Probabilités, XVII, 221–224, Lect. Notes in Math. **986**, Springer, Berlin (1983).
- [2] H.J. Brascamp, E.H. Lieb, On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation, J. Funct. Anal. **22**, 366–389 (1976).
- [3] L.A. Caffarelli, Monotonicity properties of optimal transportation and the FKG and related inequalities, Commun. Math. Phys. **214**, 547–563 (2000); Erratum, Commun. Math. Phys. **225**, 449–450 (2002).
- [4] J.-D. Deuschel, G. Giacomin, D. Ioffe, Large deviations and concentration properties for $\nabla\phi$ interface models, Probab. Theory Relat. Fields **117**, 49–111 (2000).
- [5] T. Funaki, H. Spohn, Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model, Commun. Math. Phys. **185**, 1–36 (1997).
- [6] G. Giacomin, S. Olla, H. Spohn, Equilibrium fluctuations for $\nabla\phi$ interface model, Ann. Probab. **29**, 1138–1172 (2001).
- [7] Y. Hariya, A connection of the Brascamp-Lieb inequality with Skorokhod embedding, arXiv:1309.2727 [math.PR] (2013).
- [8] D. Hobson, The Skorokhod embedding problem and model-independent bounds for option prices, in Paris-Princeton Lectures on Mathematical Finance 2010, 267–318, Lecture Notes in Math. **2003**, Springer, Berlin (2011).
- [9] J. Oblój, The Skorokhod embedding problem and its offspring, Probab. Surv. **1**, 321–392 (2004).
- [10] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, 3rd ed., Springer, Berlin (1999).
- [11] A.V. Skorokhod, Issledovaniya po teorii sluchaïnykh protsessov (Stokhasticheskie differentsial'nye uravneniya i predel'nye teoremy dlya protsessov Markova), Izdat. Kiev. Univ., Kiev (1961).

Mathematical Institute
 Tohoku University
 Sendai 980-8578
 JAPAN

E-mail address: hariya@math.tohoku.ac.jp

東北大学大学院理学研究科数学専攻 針谷 祐