Asymmetric Covariance Estimates of Brascamp-Lieb Type and Related Inequalities for Log-concave Measures

Eric A. Carlen[1], Dario Cordero-Erausquin[2] and Elliott H. Lieb[3]

1. Department of Mathematics, Hill Center,
Rutgers University, 110 Frelinghuysen Road Piscataway NJ 08854-8019 USA

2. Institut de Mathématiques de Jussieu,
Université Pierre et Marie Curie (Paris 6), 4 place Jussieu, 75252 Paris France

3. Departments of Mathematics and Physics, Jadwin Hall,
Princeton University, P. O. Box 708, Princeton, NJ 08542-0708

October, 2011

Abstract

An inequality of Brascamp and Lieb provides a bound on the covariance of two functions with respect to log-concave measures. The bound estimates the covariance by the product of the L^2 norms of the gradients of the functions, where the magnitude of the gradient is computed using an inner product given by the inverse Hessian matrix of the potential of the log-concave measure. Menz and Otto [13] proved a variant of this with the two L^2 norms replaced by L^1 and L^∞ norms, but only for $\mathbb{R}^1$. We prove a generalization of both by extending these inequalities to L^p and L^q norms and on $\mathbb{R}^n$, for any $n \geq 1$. We also prove an inequality for integrals of divided differences of functions in terms of integrals of their gradients.

Mathematics subject classification number: 26D10

Key Words: convexity, log-concavity, Poincaré inequality

1 Introduction

Let f be a C^2 strictly convex function on $\mathbb{R}^n$ such that e^{-f} is integrable. By strictly convex, we mean that the Hessian matrix, Hess_f, of f is everywhere positive.

[1]Work partially supported by U.S. National Science Foundation grant DMS 0901632.

[2]Work partially supported by U.S. National Science Foundation grant PHY 0965859.

Adding a constant to f, we may suppose that

$$\int_{\mathbb{R}^n} e^{-f(x)}\,\mathrm{d}^n x = 1 \ .$$

Let $\mathrm{d}\mu$ denote the probability measure

$$\mathrm{d}\mu := e^{-f(x)}\,\mathrm{d}^n x \ , \tag{1.1}$$

and let $\|\cdot\|_p$ denote the corresponding $L^p(\mu)$-norm.

For any two real-valued functions $f, g \in L^2(\mu)$, the covariance of f and g is the quantity

$$\mathrm{cov}(g,h) := \int_{\mathbb{R}^n} gh\,\mathrm{d}\mu - \left(\int_{\mathbb{R}^n} g\,\mathrm{d}\mu\right)\left(\int_{\mathbb{R}^n} h\,\mathrm{d}\mu\right) \ , \tag{1.2}$$

and the variance of h is $\mathrm{var}(h) = \mathrm{cov}(h,h)$.

The Brascamp-Lieb (BL) inequality [4] for the variance of h is

$$\mathrm{var}(h) \le \int_{\mathbb{R}^n} (\nabla h, \mathrm{Hess}_f^{-1}\nabla h)\,\mathrm{d}\mu \ , \tag{1.3}$$

where (x,y) denotes the inner product in $\mathbb{R}^n$. (We shall also use $x\cdot y$ to denote this same inner product in simpler expressions where it is more convenient.)

Since $(\mathrm{cov}(g,h))^2 \le \mathrm{var}(g)\mathrm{var}(h)$, an immediate consequence of (1.3) is

$$(\mathrm{cov}(g,h))^2 \le \int_{\mathbb{R}^n} (\nabla g, \mathrm{Hess}_f^{-1}\nabla g)\,\mathrm{d}\mu \int_{\mathbb{R}^n} (\nabla h, \mathrm{Hess}_f^{-1}\nabla h)\,\mathrm{d}\mu \ . \tag{1.4}$$

The one-dimensional variant of (1.4), due to Otto and Menz [13], is

$$|\mathrm{cov}(g,h)| \le \|\nabla g\|_1 \|\mathrm{Hess}_f^{-1}\nabla h\|_\infty \ = \sup_x \left\{\frac{|h'(x)|}{f''(x)}\right\} \int_{\mathbb{R}} |g'(x)|\,\mathrm{d}\mu(x) \tag{1.5}$$

for functions g and h on $\mathbb{R}^1$. They call this an *asymmetric Brascamp-Lieb inequality.* Note that it is asymmetric in *two* respects: One respect is to take an L^1 norm of ∇g and an L^∞ norm of ∇h, instead of L^2 and L^2. The second respect is that the L^∞ norm is weighted with the inverse Hessian – which here is simply a number – while the L^1 norm is not weighted.

Our first result is the following theorem, which generalizes both (1.4) and (1.5).

1.1 THEOREM (Assymetric BL inequality)**.** *Let* $\mathrm{d}\mu(x)$ *be as in* (1.1) *and let* $\lambda_{\min}(x)$ *denote the least eigenvalue of* $\mathrm{Hess}_f(x)$*. For any locally Lipschitz functions* g *and* h *on* $\mathbb{R}^n$ *that are square integrable with respect to* $\mathrm{d}\mu$*, and for* $2 \le p \le \infty$*,* $1/p + 1/q = 1$*,*

$$|\mathrm{cov}(g,h)| \ \le \ \left\|\mathrm{Hess}_f^{-1/p}\,\nabla g\right\|_q \left\|\lambda_{\min}^{(2-p)/p}\mathrm{Hess}_f^{-1/p}\nabla h\right\|_p \ . \tag{1.6}$$

This is sharp in the sense that (1.6) cannot hold, generally, with a constant smaller than 1 on the right side.

For $p = 2$, (1.6) is (1.4). Note that (1.6) implies in particular that for Lipschitz functions g, h on $\mathbb{R}^n$,

$$|\mathrm{cov}(g,h)| \le \left\|\lambda_{\min}^{-1/p}\,\nabla g\right\|_q \left\|\lambda_{\min}^{-1/q}\,\nabla h\right\|_p .$$

For $p = \infty$ and $q = 1$, the latter is

$$|\mathrm{cov}(g,h)| \le \left\|\nabla g\right\|_1 \left\|\lambda_{\min}^{-1}\nabla h\right\|_\infty, \tag{1.7}$$

which for $n = 1$ reproduces exactly (1.5).

We also prove the following theorem. In addition to its intrinsic interest, it gives rise to an alternative proof, which we give later, of Theorem 1.1 in the case $p = \infty$ (though this proof only yields the sharp constant for $\mathbb{R}^1$, which is the original Otto-Menz case (1.5)).

1.2 THEOREM (Divided differences and gradients)**.** *Let μ be a probability measure with log-concave density* (1.1) *. For any locally Lipschitz function h on $\mathbb{R}^n$,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|h(x)-h(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \le 2^n \int_{\mathbb{R}^n} |\nabla h(x)|\,\mathrm{d}\mu \ . \tag{1.8}$$

1.3 Remark. The constant 2^n is not optimal, as indicated by the examples in Section 4 (we will actually briefly mention how to reach the constant $2^{n/2}$). We do not know whether the correct constant grows with n (and then how), or is bounded uniformly in n. We do know that for $n = 1$, the constant is at least $2\ln 2$. We will return to this later.

The rest of the paper is organized as follows: Section 2 contains the proof of Theorem 1.1, and Section 3 contains the proof of Theorem 1.2, as well as an explanation of the connection between the two theorems. Section 4 contains comments and examples concerning the constant and optimizers in Theorem 1.2. Section 5 contains a discussion of an application that motivated Otto and Menz, and finally, Section 6 is an appendix providing some additional details on the original proof of the Brascamp-Lieb inequalities, which proceeds by induction on the dimension, and has an interesting connection with the application discussed in Section 5.

We end this introduction by expressing our gratitude to D. Bakry and M. Ledoux for fruitful exchanges on the preliminary version of our work. We originally proved (1.7) with the constant $n2^n$ using Theorem 1.2, as explained in Section 3. Bakry and Ledoux pointed out to us that using a stochastic representation of the gradient along the semigroup associated to μ (sometimes referred to as the Bismut formula), one could derive inequality (1.7) with the right constant 1. This provided evidence that something more algebraic was at stake. It was confirmed by our general statement Theorem 1.1 and by its proof below.

2 Bounds on Covariance

The starting point of the proof we now give for Theorem 1.1 is a classical dual representation for the covariance which, in the somewhat parallel setting of plurisubharmonic potentials, goes back to the work of Hörmander. We shall then adapt to our L^p setting Hörmander's L^2 approach [8] to spectral estimates.

Let g and h be smooth and compactly supported on $\mathbb{R}^n$. Define the operator L by

$$L = \Delta - \nabla f \cdot \nabla \ , \tag{2.1}$$

and note that

$$\int_{\mathbb{R}^n} g(x) Lh(x)\,\mathrm{d}\mu(x) = -\int_{\mathbb{R}^n} \nabla g(x)\cdot\nabla h(x)\,\mathrm{d}\mu(x)\ , \tag{2.2}$$

so that L is self-adjoint on $L^2(\mu)$. Let us (temporarily) add $\epsilon|x|^2$ to f to make it uniformly convex, so that the Hessian of f is invertible and so that the operator L has a spectral gap. (Actually, L always has a spectral gap since μ is a log-concave probability measure, as noted in [9, 1]. Our simple regularization makes our proof independent of these deep results.)

Then provided

$$\int_{\mathbb{R}^n} h(x)\,\mathrm{d}\mu(x) = 0\ , \tag{2.3}$$

$$u := -\int_0^\infty e^{tL} h(x)\,\mathrm{d}t \tag{2.4}$$

exists and is in the domain of L, and satisfies $Lu = h$.

Thus, assuming (2.3), and by standard approximation arguments,

$$\begin{aligned}\mathrm{cov}(g,h) &= \int_{\mathbb{R}^n} g(x)h(x)\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^n} g(x)Lu(x)\,\mathrm{d}\mu(x)\\ &= -\int_{\mathbb{R}^n} \nabla g(x)\cdot\nabla u(x)\,\mathrm{d}\mu(x)\ .\end{aligned} \tag{2.5}$$

This representation for the covariance is the starting point of the proof we now give for Theorem 1.1.

Proof of Theorem 1.1: Fix $2 \le p < \infty$, and let $q = p/(p-1)$, as in the statement of the theorem. Suppose h satisfies (2.3), and define u by (2.4) so that $Lu = h$. Then from (2.5),

$$\begin{aligned}|\mathrm{cov}(g,h)| &\le \left|\int_{\mathbb{R}^n} \nabla g(x)\cdot\nabla u(x)\,\mathrm{d}\mu(x)\right|\\ &\le \int_{\mathbb{R}^n} \left|\mathrm{Hess}_f^{-1/p}\nabla g(x)\cdot \mathrm{Hess}_f^{1/p}\nabla u(x)\right|\mathrm{d}\mu(x)\\ &\le \|\mathrm{Hess}_f^{-1/p}\nabla g(x)\|_q\, \|\mathrm{Hess}_f^{1/p}\nabla u(x)\|_p\ .\end{aligned} \tag{2.6}$$

Thus, to prove (1.6) for $2 \le p < \infty$, it suffices to prove the following $W^{-1,p}$–$W^{1,p}$ type estimate:

$$\|\mathrm{Hess}_f^{1/p}\nabla u(x)\|_p \le \|\lambda_{\min}^{(2-p)/p}\mathrm{Hess}_f^{-1/p}\nabla h\|_p \ . \tag{2.7}$$

Toward this end, we compute

$$\begin{aligned} L(|\nabla u|^p) &= p|\nabla u|^{p-2}(L\nabla u)\cdot\nabla u \\ &\qquad +p|\nabla u|^{p-2}\mathrm{Tr}(\mathrm{Hess}_u^2) + p(p-2)|\nabla u|^{p-4}|\mathrm{Hess}_u\nabla u|^2 \\ &\ge p|\nabla u|^{p-2}(L\nabla u)\cdot\nabla u \ , \end{aligned} \tag{2.8}$$

where we have used the fact that $p \ge 2$, and where the notation $L(\nabla u)$ refers to the coordinate-wise action $(L\partial_1 u, \ldots, L\partial_n u)$ of L.

Then, using the commutation formula (see the remark below)

$$L(\nabla u) = \nabla(Lu) + \mathrm{Hess}_f\nabla u \ , \tag{2.9}$$

we obtain

$$0 = \int_{\mathbb{R}^n} L(|\nabla u|^p)\,\mathrm{d}\mu(x) \ge p\int_{\mathbb{R}^n}|\nabla u|^{p-2}\nabla u\cdot\nabla h\,\mathrm{d}\mu(x) + p\int_{\mathbb{R}^n}|\nabla u|^{p-2}\nabla u\cdot\mathrm{Hess}_f\nabla u\,\mathrm{d}\mu(x) \ ,$$

and hence

$$\int_{\mathbb{R}^n}|\nabla u|^{p-2}|\mathrm{Hess}_f^{1/2}\nabla u|^2\,\mathrm{d}\mu(x) \le \int_{\mathbb{R}^n}|\nabla u|^{p-2}|\mathrm{Hess}_f^{1/p}\nabla u||\mathrm{Hess}_f^{-1/p}\nabla h|\,\mathrm{d}\mu(x) \ . \tag{2.10}$$

We now observe that for any positive $n \times n$ matrix and any vector $v \in \mathbb{R}^n$,

$$|A^{1/p}v|^p \le |v|^{p-2}|A^{1/2}v|^2 \ .$$

To see this, note that we may suppose $|v| = 1$. Then in the spectral representation of A, by Jensen's inequality,

$$|A^{1/p}v| = \left(\sum_{j=1}^n \lambda_j^{1/p}v_j^2\right)^{1/2} \le \left(\sum_{j=1}^n \lambda_j^{1/2}v_j^2\right)^{1/p} \ .$$

Using this on the left side of (2.10), and using the obvious estimate

$$|\nabla u| \le \lambda_{\min}^{-1/p}|\mathrm{Hess}_f^{1/p}\nabla u|$$

on the right, we have

$$\|\mathrm{Hess}_f^{1/p}\nabla u\|_p^p \le \int_{\mathbb{R}^n}|\mathrm{Hess}_f^{1/p}\nabla u|^{p-1}|\lambda_{\min}^{(2-p)/p}\mathrm{Hess}_f^{-1/p}\nabla h|\,\mathrm{d}\mu(x) \ . \tag{2.11}$$

Then by Hölder's inequality we obtain (2.7).

It is now obvious that we can take the limit in which ϵ tends to zero, so that we obtain the inequality without any additional hypotheses on f. Our calculations so far have required $2 \le p < \infty$, however, having obtained the inequality for such p, by taking the limit in which p goes to infinity, we obtain the $p = \infty$, $q = 1$ case of the theorem.

Finally, considering the case in which

$$d\mu(x) = (2\pi)^{-n/2} e^{-|x|^2}\, dx \ ,$$

and $g = h = x_1$, we have that $\mathrm{Hess}_f = \mathrm{I_d}$ and so

$$\lambda_{\min} = |\mathrm{Hess}_f^{-1/p} \nabla g| = |\mathrm{Hess}_f^{-1/p} \nabla h| = 1$$

for all x, and so the constant is sharp, as claimed. □

2.1 Remark. Many special cases and variants of the commutation relation (2.9) are well-known under different names. Perhaps most directly relevant here is the case in which $f(x) = |x|^2/2$. Then ∂_j and its adjoint in $L^2(\mu)$, $\partial_j^* = x_j - \partial_j$, satisfy the canonical commutation relations, and the operator $L = -\sum_{j=1}^n \partial_j^* \partial_j$ is (minus) the Harmonic oscillator Hamiltonian in the ground state representation. This special case of (2.9), in which the Hessian on the right is the identity, is the basis of the standard determination of the spectrum of the quantum harmonic oscillator using "raising and lowering operators".

In the setting of Riemannian manifolds, a commutation relation analogous to (2.9) in which L is the Laplace-Beltrami operator and the Hessian is replaced by Ric, the Ricci curvature tensor, is known as the Bochner-Lichnerowicz formula. Both the Hessian version (2.9) and the Bochner-Lichnerowicz version have been used a number of times to prove inequalities related to those we consider here, for instance in the work of Bakry and Emery on logarithmic Sobolev inequalities.

We note that our proof immediately extends, word for word, to the Riemannian setting if we use, in place of (2.9) the commutation satisfied by the operator L given by (2.1) where f is a (smooth) potential on the manifold; That is, with some abuse of notation, $L(\nabla u) = \nabla(Lu) + \mathrm{Hess}_f \nabla u + \mathrm{Ric} \nabla u$, or rather, more rigorously,

$$L(|\nabla u|^p) \ge p|\nabla u|^{p-2}\big[\nabla(Lu) \cdot \nabla u + \mathrm{Hess}_f \nabla u \cdot \nabla u + \mathrm{Ric} \nabla u \cdot \nabla u\big].$$

Thus, an analog of Theorem 1.1 holds on a Riemannian manifold M equipped with a probability measure

$$d\mu(x) = e^{-f(x)}\, d\mathrm{vol}(x)$$

where dvol is the Riemannian element of volume and f a smooth function on M, provided Hess_f at each point x is replaced in the statement by the symmetric operator

$$H_x = \mathrm{Hess}_f(x) + \mathrm{Ric}_x$$

defined on the tangent space. Of course, the convexity condition on f is accordingly replaced by the assumption that $H_x > 0$ at every point $x \in M$.

3 Bounds on Differences

Proof of Theorem 1.2: Since $h(x)-h(y)=\int_0^1 \nabla h(x_t)\cdot(x-y)\,\mathrm{d}t$, we have

$$|h(x)-h(y)| \leq |x-y| \int_0^1 |\nabla h(x_t)|\,\mathrm{d}t \quad \text{where} \quad x_t := tx+(1-t)y \ . \tag{3.1}$$

Next, by the convexity of f,

$$e^{-f(x)}e^{-f(y)} = e^{-(1-t)f(x)}e^{-tf(y)}e^{-tf(x)}e^{-(1-t)f(y)} \leq e^{-f(x_t)}e^{-(1-t)f(x)}e^{-tf(y)} \ . \tag{3.2}$$

Introduce the variables

$$\begin{aligned} w &= tx+(1-t)y \\ z &= x-y \ . \end{aligned} \tag{3.3}$$

A simple computation of the Jacobian shows that this change of variables is a measure preserving transformation for all $0 \leq t \leq 1$, and hence

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|h(x)-h(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leq \int_0^1 \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla h(w)| e^{-(1-t)f(w+(1-t)z)}e^{-tf(w-tz)}\,\mathrm{d}z\ \mathrm{d}\mu(w)\right)\mathrm{d}t \ . \tag{3.4}$$

We estimate the right side of (3.4). By Hölder's inequality,

$$\int_{\mathbb{R}^n} e^{-(1-t)f(w+(1-t)z)}e^{-tf(w-tz)}\,\mathrm{d}^n z \leq \left(\int_{\mathbb{R}^n} e^{-f(w+(1-t)z)}\,\mathrm{d}z\right)^{1-t}\left(\int_{\mathbb{R}^n} e^{-f(w-tz)}\,\mathrm{d}z\right)^{t} \ . \tag{3.5}$$

But

$$\int_{\mathbb{R}^n} e^{-f(w+(1-t)z)}\,\mathrm{d}z = (1-t)^{-n} \qquad \text{and} \qquad \int_{\mathbb{R}^n} e^{-f(w-tz)}\,\mathrm{d}z = t^{-n} \ ,$$

and finally, $(1-t)^{-n(1-t)}t^{-nt} = e^{-n(t\log t+(1-t)\log(1-t))} \leq 2^n$. □

A corollary of Theorem 1.2 is a proof of Theorem 1.1 for the special case of $q=1$ and $p=\infty$. This proof is not only restricted to this case, it also has the defect that the constant is not sharp, except in one-dimension. We give it, nevertheless, because it establishes a link between the two theorems.

Alternative Proof of Theorem 1.1 for $q=1$: We shall use the identity

$$\mathrm{cov}(g,h) = \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} [g(x)-g(y)][h(x)-h(y)]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \ , \tag{3.6}$$

and estimate the differences on the right in different ways.

Fix any $x \neq y$ in $\mathbb{R}^n$, and define the vector $v := x - y$, and for $0 \leq t \leq 1$, define $x_t = y + tv = tx + (1-t)y$. Then for any Lipschitz function h,

$$h(x) - h(y) = \int_0^t v \cdot \nabla h(x_t)\, \mathrm{d}t \ . \tag{3.7}$$

Now note that

$$\frac{\mathrm{d}}{\mathrm{d}t} v \cdot \nabla f(x_t) = (v, \mathrm{Hess}_f(x_t) v) \geq |x-y|^2 \lambda_{\min}(x_t) > 0 \ . \tag{3.8}$$

Integrating this in t from 0 to 1, we obtain

$$(x - y, \nabla f(x) - \nabla f(y)) = \int_0^1 (v, \mathrm{Hess}_f(x_t) v)\, \mathrm{d}t > 0 \ , \tag{3.9}$$

which expresses the well-known monotonicity of gradients of convex functions.

Next, multiplying and dividing by $(v, \mathrm{Hess}_f(x_t)v)$ in (3.7), we obtain

$$\begin{aligned}
|h(x) - h(y)| &= \left| \int_0^1 (v, \mathrm{Hess}_f(x_t)v)(v, \mathrm{Hess}_f(x_t)v)^{-1} v \cdot \nabla h(x_t)\, \mathrm{d}t \right| \\
&\leq \int_0^1 (v, \mathrm{Hess}_f(x_t)v) \left|(v, \mathrm{Hess}_f(x_t)v)^{-1} v \cdot \nabla h(x_t)\right| \, \mathrm{d}t \\
&\leq \int_0^1 (v, \mathrm{Hess}_f(x_t)v) \left|(\lambda_{\min}(x_t))^{-1} |x-y|^{-2} v \cdot \nabla h(x_t)\right| \, \mathrm{d}t \\
&\leq \sup_{z \in \mathbb{R}^n} \left\{ \frac{|\nabla h(z)|}{\lambda_{\min}(z)} \right\} |x-y|^{-1} \int_0^1 (v, \mathrm{Hess}_f(x_t)v)\, \mathrm{d}t \\
&= \sup_{z \in \mathbb{R}^n} \left\{ \frac{|\nabla h(z)|}{\lambda_{\min}(z)} \right\} |x-y|^{-1} (x-y, \nabla f(x) - \nabla f(y)) \ .
\end{aligned} \tag{3.10}$$

Define

$$C := \sup_{z \in \mathbb{R}^n} \left\{ \frac{|\nabla h(z)|}{\lambda_{\min}(z)} \right\} \ ,$$

and use (3.10) in (3.6):

$$\begin{aligned}
|\mathrm{cov}(g,h)| &\leq \frac{1}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x) - g(y)||h(x) - h(y)|\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \\
&\leq \frac{C}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x) - g(y)| \frac{1}{|x-y|} (x-y) \cdot [\nabla f(x) - \nabla f(y)]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \\
&= \frac{C}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x) - g(y)| \frac{1}{|x-y|} (x-y) \cdot \left[\nabla_y e^{-f(y)} e^{-f(x)} - \nabla_x e^{-f(x)} e^{-f(y)}\right] \mathrm{d}^n x\, \mathrm{d}^n y \ . \\
&= -C \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x) - g(y)| \frac{1}{|x-y|} (x-y) \cdot \nabla_x e^{-f(x)} e^{-f(y)}\, \mathrm{d}^n x\, \mathrm{d}^n y \ ,
\end{aligned}$$

where, in the last line, we have used symmetry in x and y.

Now integrate by parts in x. Suppose first that $n > 1$. Then

$$\operatorname{div}\left(\frac{1}{|z|}z\right) = \frac{n-1}{|z|} ,$$

and $|\nabla_x |g(x) - g(y)|| = |\nabla_x g(x)|$ almost everywhere. Hence we obtain

$$|\operatorname{cov}(g,h)| \le C\left(\int_{\mathbb{R}^n} |\nabla g(x)|\, \mathrm{d}\mu(x) + (n-1)\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|g(x)-g(y)|}{|x-y|}\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)\right) . \quad (3.11)$$

For $n = 1$, $\operatorname{div}\left(\frac{1}{|z|}z\right) = 2\delta_0(z)$ and (3.11) is still valid since $|g(x) - g(y)|\delta_0(x - y) = 0$.

Now, for $n = 1$, (3.11) reduces directly to (1.5). For $n > 1$, it reduces to (1.7) upon application of Theorem 1.2, but with the constant $n2^n$ instead of 1. □

4 Examples and Remarks on Optimizers in Theorem 1.2

Our first examples address the question of the importance of log-concavity.

(1.) Some restriction on μ is necessary: If a measure $d\mu(x) = F(x)dx$ on $\mathbb{R}$ has $F(a) = 0$ for some $a \in \mathbb{R}$, and F has positive mass to the left and right of a, then inequality (1.8) cannot possibly hold with any constant. The choice of h to be the Heaviside step function shows that (1.8) cannot hold with any constant for this μ.

(2.) Unimodality is not enough: Take $d\mu(x) = F(x)dx$, with $F(x) = 1/4\varepsilon$ on $(-\varepsilon, \varepsilon)$ and $F(x) = 1/4(1-\varepsilon)$ otherwise on the interval $(-1, 1)$ and $F(x) = 0$ for $|x| > 1$. Let $g(x) = 1$ for $|x| < \varepsilon + \delta$ and $g(x) = 0$ otherwise. When δ is positive but small,

$$\int_{\mathbb{R}} |\nabla g|\, \mathrm{d}\mu(x) = 1/2(1-\varepsilon)$$

while

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|g(x)-g(y)|}{|x-y|}\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) = O(-\ln(\epsilon)) .$$

(3.) For $n = 1$, the best constant in (1.8) is at least $2\ln 2$: Take $d\mu(x) = F(x)dx$, with $F(x) = 1/2$ on $(-1, 1)$ and $F(x) = 0$ for $|x| > 1$. Let $g(x) = 1$ for $x \ge 0$ and $g(x) = 0$ for $x < 0$. All integrals are easily computed.

(4.) The best constant is achieved for characteristic functions: When seeking the best constant in (1.8), it suffices, by a standard truncation argument, to consider bounded Lipschitz functions h. Then, since neither side of the inequality is affected if we add a constant to h, it suffices to consider non-negative Lipschitz functions. We use the layer-cake representation [11]:

$$h(x) = \int_0^\infty \chi_{\{h>t\}}(x)\,\mathrm{d}t\ .$$

Then

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|h(x)-h(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \le \int_0^\infty \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|\chi_{\{h>t\}}(x)-\chi_{\{h>t\}}(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}t \tag{4.1}$$

Define C_n to be the best constant for characteristic functions of sets A and log-concave measures μ:

$$C_n := \sup_{f,A}\left\{ \frac{\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|\chi_A(x)-\chi_A(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)}{\int_{\partial A} e^{-f(x)}\,\mathrm{d}\mathcal{H}_{n-1}(x)} \right\} \tag{4.2}$$

where $\mathcal{H}_{n-1}$ denotes $n-1$ dimensional Hausdorff measure. Apply this to (4.1) to conclude that

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|h(x)-h(y)|}{|x-y|}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) &\le C_n \int_0^\infty \int_{\partial\chi_{\{h>t\}}} e^{-f(x)}\,\mathrm{d}\mathcal{H}_{n-1}(x)\,\mathrm{d}t \\ &= C_n \int_{\mathbb{R}^n} |\nabla h(x)|\,\mathrm{d}\mu(x)\ ,\end{aligned} \tag{4.3}$$

where the co-area formula was used in the last line. Thus, inequality (1.8) holds with the constant C_n; in short, it suffices to consider characteristic functions as trial functions. Note that the argument is also valid at the level of each measure μ individually, although we are interested here in uniform bounds.

With characteristic functions in mind, let us consider the case that g is the characteristic function of a half-space in $\mathbb{R}^n$. Without loss of generality let us take this to be $\{x\,:\,x_1<0\}$. Clearly, the left side of (1.8) is less than the integral with $|x-y|^{-1}$ replaced by $|x_1-y_1|^{-1}$. Since the marginal (obtained by integrating over $x_2,\ldots,x_n$) of a log concave function is log concave, we see that our inequality reduces to the one-dimensional case. In other words, the constant C_n in (4.2) would equal C_1, independent of n, if the supremum were restricted to half-spaces instead of to arbitrary measurable sets.

(5.) Improved constants and geometry of log-concave measures: With additional assumptions on the measure one can see that the constant is not only bounded in n, but of order $1/\sqrt{n}$. We are grateful to F. Barthe and M. Ledoux for discussions and improvements in particular cases concerning the constant in Theorem 1.2. This relies on the Cheeger constant $\alpha(\mu)^{-1}>0$ associated to the log-concave probability measure $\mathrm{d}\mu$, which is defined to be the best constant in the inequality

$$\forall A\subset\mathbb{R}^n(\text{regular enough}),\qquad \mu(A)(1-\mu(A)) \le \alpha(\mu)\int_{\partial A} e^{-f(x)}\,\mathrm{d}\mathcal{H}_{n-1}(x)$$

M. Ledoux suggested the following procedure. Split the function $|x-y|^{-1}$ into two pieces according to whether $|x-y|$ is less than or greater than R, for some $R>0$. With h being the characteristic function of A, the contribution to the left side of (4.3) for $|x-y|>R$ is bounded above by $2R^{-1}\alpha(\mu)\int_{\partial A} e^{-f(x)}\,\mathrm{d}\mathcal{H}_{n-1}(x)$. The contribution for $|x-y|\le R$ is bounded above in the same manner as in the proof of Theorem 1.2, but this time we only have to integrate z over the domain $|z|\le R$ in each of the integrals in (3.5). Thus, our bound 2^n is improved by a factor, which is the $\mathrm{d}\mu$ volume of the ball $B_R=\{|z|\le R\}$, once we used the Brunn-Minkowski inequality for the bound

$$\mu\big((1-t)B_R+w)\big)^{1-t}\,\mu\big(tB_R+w\big)^t\le\mu\big(B_R+w\big)\le\overline{\mu}(B_R):=\sup_x\mu(B_R+x).$$

The final step is to optimize the sum of the contributions of the two terms with respect to R. Thus, if we denote $C_n(\mu)$ the best constant in the inequality (1.8) of Theorem 1.2 for a fixed measure μ, we have

$$C_n(\mu)\le\inf_{R>0}\big\{2^n\overline{\mu}(B_R)+2R^{-1}\alpha(\mu)\big\}\le 2^n. \tag{4.4}$$

Note that if μ is symmetric (i.e. if f is even), then the Brunn-Minkowski inequality ensures that $\overline{\mu}(B_R)=\mu(B_R)$.

Unlike in (1.8), this improved bound depends on μ but there are situation where this gives optimal estimates as pointed out to us by F. Barthe. As an example, consider the case where μ is the standard Gaussian measure on $\mathbb{R}^n$. Using the known value of the Cheeger constant for this μ, and linear trial functions, one finds that the constant is bounded above and below by a constant times $n^{-1/2}$.

Actually, we can use (4.4) to improve the constant from 2^n to $2^{n/2}$ for arbitrary measures using some recent results from the geometry of log-concave measures. Without loss of generality, we can assume, by translation of μ, that $\int|x|\,d\mu(x)=\inf_v\int|x+v|\,d\mu(x)=:M_\mu$. It was proved in [9, 1] that for every log-concave measure on $\mathbb{R}^n$,

$$\alpha(\mu)\le cM_\mu$$

where $c>0$ is some numerical constant (meaning a possibly large, but computable, constant, in particular independent of n and μ, of course). On the other hand, it was proved by Guédon [7] that for every log-concave measure ν on $\mathbb{R}^n$

$$\nu(B_R)\le\frac{C}{\int|x|\,d\nu}R$$

for some numerical constant $C>0$. In the case μ is not symmetric, we pick v such that $\mu(B_r+v)=\overline{\mu}(B_R)$, and then we apply the previous bound to $\nu(\cdot)=\mu(\cdot+v)$ in order to get that $\overline{\mu}(B_R)\le\frac{C}{M_\mu}$. Using these two estimates in (4.4) we see that

$$C_n(\mu)\le\inf_{s>0}\{C2^ns+c/s\}=\kappa\,2^{n/2}$$

for some numerical constant $\kappa > 0$.

The Brascamp-Lieb inequality (1.3), as well as inequality (1.8), have connections with the geometry of convex bodies. It was observed in [2] that (1.3) can be deduced from the Prékopa-Leindler inequality (which is a functional form of the Brunn-Minkowski inequality). But the converse is also true: the Prékopa theorem follows, by a local computation, from the Brascamp-Lieb inequality (see [5] where the procedure is explained in the more general complex setting). To sum up, the Brascamp-Lieb inequality (1.3) can be seen as the *local* form of the Brunn-Minkowski inequality for convex bodies.

5 Application to Conditional Expectations

Otto and Menz were motivated to prove (1.5) for an application that involves a large amount of additional structure that we cannot go into here. We shall however give an application of Theorem 1.1 to a type of estimate that is related to one of the central estimates in [13].

We use the notation in [4], which is adapted to working with a partitioned set of variables. Write a point $x \in \mathbb{R}^{n+m}$ as $x = (y, z)$ with $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$. For a function A on $\mathbb{R}^{n+m}$, let $\langle A \rangle_z(y)$ denote the conditional expectation of A given y, with respect to μ. For a function B of y alone, $\langle B \rangle_y$ is the expected value of B, with respect to μ. As in [4], a subscript y or z on a function denotes differentiation with respect to y or z, while a subscript y or z on a bracket denotes integration. For instance, for a function g on $\mathbb{R}^{n+m}$, g_y denotes the vector $\left(\frac{\partial g}{\partial y_i}\right)_{i \le n}$ in $\mathbb{R}^n$, and for $i \le n$, $g_{y_i z}$ denotes the vector $\left(\frac{\partial^2 g}{\partial y_i \partial z_j}\right)_{j \le m}$ in $\mathbb{R}^m$. Finally, (g_{yz}) denotes the $n \times m$ matrix having the previous vectors as rows.

Let h be non-negative with $\langle h \rangle_x = 1$ so that $h(x)\,\mathrm{d}\mu(x)$ is a probability measure, and so is $\langle h \rangle_z(y)\,\mathrm{d}\nu(y)$, where $\mathrm{d}\nu(y)$ is the marginal distribution of y under $\mathrm{d}\mu(x)$.

A problem that frequently arises [3, 6, 10, 12, 13] is to estimate the Fisher information of $\langle h \rangle_z(y)\,\mathrm{d}\nu(y)$ in terms of the Fisher information of $h(x)\,\mathrm{d}\mu(x)$ by proving an estimate of the form

$$\left\langle \frac{|(\langle h \rangle_z)_y|^2}{\langle h \rangle_z} \right\rangle_y \le C \left\langle \frac{|h_x|^2}{h} \right\rangle_x . \tag{5.1}$$

Direct differentiation under the integral sign in the variable y_i gives

$$(\langle h \rangle_z)_{y_i} = \langle h_{y_i} \rangle_z - \mathrm{cov}_z(h, f_{y_i}) ,$$

where cov_z denotes the conditional covariance of $h(y, z)$ and $f_{y_i}(y, z)$, integrating in z for each fixed y. Let $u = (u_1, \ldots, u_m)$ be any unit vector in $\mathbb{R}^m$. Then Hence, for each y,

$$\begin{aligned} (\langle h \rangle_z)_y \cdot u = \sum_{i=1}^m (\langle h \rangle_z)_{y_i} u_i &= \sum_{i=1}^m \langle h_{y_i} \rangle_z u_i - \sum_{i=1}^m \mathrm{cov}_z(h, f_{y_i}) u_i \\ &= \langle h_y \rangle_z \cdot u - \mathrm{cov}_z(h, f_y \cdot u) , \end{aligned}$$

and hence, choosing u to maximize the left hand side,

$$|(\langle h \rangle_z)_y|^2 \le 2|\langle h_y \rangle_z|^2 + 2\left(\mathrm{cov}_z(h, f_y \cdot u)\right)^2 . \tag{5.2}$$

By (1.6),

$$|\mathrm{cov}_z(h, f_y \cdot u)| \leq \langle |h_z| \rangle_z \|\lambda_{\min}^{-1} |(f_y \cdot u)_z|\|_\infty \ . \tag{5.3}$$

Note that the least eigenvalue of the $n \times n$ block f_{zz} is at least as large as the least eigenvalue $\lambda_{\min}(y, z)$ of the full Hessian, by the variational principle. Hence, while we are entitled to use the least eigenvalue of the $n \times n$ block f_{zz} of the full $(n + m) \times (n + m)$ Hessian matrix f_{xx}, and this would be important in the application in the one dimensional case made in [13], here, without any special structure to take advantage of, we simply use the least eigenvalue of the full matrix in our bound.

Next note that

$$|(f_y \cdot u)_z|^2 \leq \sum_{i=1}^{m} \left(\sum_{j=1}^{n} (f_{y_i, z_j})^2 \right) u_i^2 \ ,$$

and that $\sum_{j=1}^{n} (f_{y_i, z_j})^2$ is the i, i entry of $f_{yz}^T f_{yz}$ where f_{yz} denotes the upper right corner block of the Hessian matrix. This number is no greater than the i, i entry of the square of the full Hessian matrix. This, in turn, is no greater than $\lambda_{\max}^2$. Then, since u is a unit vector, we have

$$|(f_y \cdot u)_z| \leq \lambda_{\max} \ .$$

Using this in (5.3), we obtain

$$|\mathrm{cov}_z(h, f_y \cdot u)| \leq \langle |h_z| \rangle_z \|\lambda_{\max}/\lambda_{\min}\|_\infty \ , \tag{5.4}$$

and then from (5.2)

$$|(\langle h \rangle_z)_y|^2 \leq 2|\langle h_y \rangle_z|^2 + 2\|\lambda_{\max}/\lambda_{\min}\|_\infty^2 \langle |h_z| \rangle_z^2 \ . \tag{5.5}$$

Then the Cauchy-Schwarz inequality yields

$$(\langle |h_z| \rangle_z)^2 \leq \left\langle \frac{|h_z|^2}{h} \right\rangle_z \langle h \rangle_z \ . \tag{5.6}$$

Use this in (5.5), divide both sides by $\langle h \rangle_z$, and integrate in y. The joint convexity in A and $\alpha > 0$ of A^2/α yields (5.1) with the constant $C = 2\|\lambda_{\max}/\lambda_{\min}\|_\infty^2$.

The bound we have obtained becomes useful when $\lambda_{\max}(x)/\lambda_{\min}(x)$ is bounded uniformly. Suppose that $f(x)$ has the form $f(x) = \varphi(|x|^2)$. Then the eigenvalues of the Hessian of f are $2\varphi'(|x|^2)$, with multiplicity $m + n - 1$, and $4\varphi''(|x|^2)|x|^2 + 2\varphi'(|x|^2)$, with multiplicity 1. Then both eigenvalues are positive, and the ratio is bounded, whenever φ' is positive and, for some $c < 1 < C < \infty$,

$$-c\varphi'(s) \leq s\varphi''(s) \leq C\varphi'(s) \ .$$

5.1 Remark (Other asymmetric variants of the BL inequality)**.** Together, (5.3) and (5.6) yield

$$\frac{(\mathrm{cov}_z(h, f_y \cdot u))^2}{\langle h \rangle_z} \leq \left\langle \frac{|h_z|^2}{h} \right\rangle_z \|\lambda_{\min}^{-1}(f_y \cdot u)_z)|\|_\infty^2 \ .$$

A weaker inequality is

$$\frac{(\mathrm{cov}_z(h, f_y \cdot u))^2}{\langle h \rangle_z} \leq \left\langle \frac{|h_z|^2}{h} \right\rangle_z \|\lambda_{\min}^{-1}\|_\infty^2 \|(f_y \cdot u)_z)\|_\infty^2 \ . \tag{5.7}$$

In the context of the application in [13], finiteness of $\|(f_y \cdot u)_z\|_\infty$ limits f to quadratic growth at infinity. A major contribution of [13] is to remove this limitation in applications of (5.1). The success of this application of (1.5) depended on the full weight of the inverse Hessian being allocated to the L^∞ term.

Nonetheless, once the topic of asymmetric BL inequalities is raised, one might enquire whether an inequality of the type

$$|\mathrm{cov}(g, h)| \leq C\|\nabla g\|_\infty \, \|\mathrm{Hess}_f^{-1} \nabla h\|_1 \tag{5.8}$$

can hold for any constant C. There is no such inequality, even in one dimension. To see this, suppose that for some $a \in \mathbb{R}$ and some $\epsilon > 0$, $f_{xx} > M$ on $(a-\epsilon, a+\epsilon)$. Take $h(x) = 1$ for $x > a$ and $h(x) = 0$ for $x \leq a$. Take $g(x) = x - a$. Suppose that f is even about a. Then $\mathrm{cov}(g, h) = \int_a^\infty (x - a)e^{-f(x)}\,\mathrm{d}x$, while $\|\mathrm{Hess}_f^{-1} \nabla h\|_1 \leq M^{-1}$, and f can be chosen to make M arbitrarily large while keeping $\|\nabla g\|_\infty \leq 1$, and $\mathrm{cov}(g, h)$ bounded away from zero.

6 Appendix

We recall that the original proof of (1.3), Theorem 4.1 of [4], used dimensional induction, though interesting non-inductive proofs have since been provided [2].

The starting point for the inductive proof is that the proof for $n = 1$ is elementary. The proof of the inductive step is more involved, and we take this opportunity to provide more detail about the passage from eq. (4.9) of [4] to eq. (4.10) of [4]. There is an interesting connection with the application discussed in the previous section, which also concerns $\langle h_y \rangle_z - \mathrm{cov}_z(h, f_y)$. We continue using the notation introduced there, but now $m = 1$ (i.e. $y \in \mathbb{R}$).

Eq. (4.9) reads $\mathrm{var}(h) \leq \langle B \rangle_y$ where

$$B = \mathrm{var}_z(h) + \frac{[\langle h_y \rangle_z - \mathrm{cov}_z(h, f_y)]^2}{\langle f_{yy} \rangle_z - \mathrm{var}_z f_y} \ . \tag{6.1}$$

Our goal is to prove

$$B \leq \langle (h_z, f_{zz}^{-1} h_z) \rangle_z + \frac{\langle h_y - (h_z, f_{zz}^{-1} f_{yz}) \rangle_z^2}{\langle f_{yy} - (f_{yz}, f_{zz}^{-1} f_{yz}) \rangle_z} \ . \tag{6.2}$$

To do this, use the inductive hypothesis; i.e., for any H on $\mathbb{R}^{n-1}$,

$$\mathrm{var}_z(H) \le \langle H_z, f_{zz}^{-1} H_z\rangle_z \ . \tag{6.3}$$

Apply this to arbitrary linear combination $H = \lambda h + \mu f_y$ to conclude the 2×2 matrix inequality

$$\begin{bmatrix} \mathrm{var}_z(h) & \mathrm{cov}_z(h, f_y) \\ \mathrm{cov}_z(h, f_y) & \mathrm{var}_z(f_y) \end{bmatrix} \le \begin{bmatrix} \langle (h_z, f_{zz}^{-1} h_z)\rangle_z & \langle (h_z, f_{zz}^{-1} f_{yz})\rangle_z \\ \langle (f_{yz}, f_{zz}^{-1} h_z)\rangle_z & \langle (f_{yz}, f_{zz}^{-1} f_{yz})\rangle_z \end{bmatrix}$$

Take the determinant of the difference to find that

$$\langle (h_z, f_{zz}^{-1} h_z)\rangle_z - \mathrm{var}_z(h) \ge \frac{[\langle (h_z, f_{zz}^{-1} f_{yz})\rangle_z - \mathrm{cov}_z(h, f_y)]^2}{\langle (f_{yz}, f_{zz}^{-1} f_{yz})\rangle_z - \mathrm{var}_z(f_y)} \ . \tag{6.4}$$

Combine (6.1) and (6.4) to obtain

$$B \le \langle (h_z, f_{zz}^{-1} h_z)\rangle_z + \frac{[\langle h_y\rangle_z - \mathrm{cov}_z(h, f_y)]^2}{\langle f_{yy}\rangle_z - \mathrm{var}_z(f_y)} - \frac{[\langle (h_z, f_{zz}^{-1} f_{yz})\rangle_z - \mathrm{cov}_z(h, f_y)]^2}{\langle (f_{yz}, f_{zz}^{-1} f_{yz})\rangle_z - \mathrm{var}_z(f_y)} \tag{6.5}$$

Since a^2/α is jointly convex in a and $\alpha > 0$, and is homogeneous of degree one, for all $\alpha > \beta > 0$ and all a and b,

$$\frac{a^2}{\alpha} \le \frac{b^2}{\beta} + \frac{(a-b)^2}{\alpha - \beta} \ .$$

That is, $a^2/\alpha - b^2/\beta \le (a-b)^2/(\alpha - \beta)$. Use this on the right side of (6.5) to obtain (6.2), noting that the positivity of $\alpha - \beta = \langle f_{yy}\rangle_z - \langle (f_{yz}, f_{zz}^{-1} f_{yz})\rangle_z$ is a consequence of the positivity of the Hessian of f.

References

[1] S.G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab. **27**, no. 4, 1903–1921 (1999).

[2] S. Bobkov, M. Ledoux, *From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities*, Geom. Funct. Anal. **10**, 102-1052 (2000).

[3] Th. Bodineau and B. Helffer, *On log Sobolev inequalities for unbounded spin systems*, J. Funct. Anal. **166**, 168-178 (1999).

[4] H. J.Brascamp and E. H. Lieb, *On extensions of the Brunn-Minkovski 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).

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

[6] N. Grunewald, F. Otto, C. Villani and M. G. Westdickenberg, *A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit*, Ann. Inst. H. Poincare Prob. Stat. **45**, 302-351 (2009).

[7] O. Guédon, *Kahane-Khinchine type inequalities for negative exponent*, Mathematika **46**, no. 1, 165–173 (1999).

[8] L. Hörmander, *L^2 estimates and existence theorems for the $\bar{\partial}$ operator*, Acta Math. **113**, 89–152 (1965).

[9] R. Kannan, L. Lovász and M. Simonovits, *Isoperimetric problems for convex bodies and a localization lemma*, Discrete Comput. Geom. **13**, 541–559 (1995).

[10] C. Landim, G. Panizo and H. T. Yau, *Spectral gap and logarithmic Sobolev inequality for unbounded conservative spin systems* Ann. Inst. H. Poincaré Prob. Stat. **38**, 739-777 (2002).

[11] E. H. Lieb and M. Loss, *Analysis, Second Edition* Amer. Math. Soc., Providence RI (2001).

[12] F. Otto and M. G. Reznikoff, *A new criterion for the logarithmic Sobolev inequality and two applications*, J. Funct. Anal. **243**, 121-157 (2007).

[13] G. Menz and F. Otto, *Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential*, Leipzig Preprint no. 5 (2011).