

Concentration of measures supported on the cube

Bo'az Klartag*

Abstract

We prove a log-Sobolev inequality for a certain class of log-concave measures in high dimension. These are the probability measures supported on the unit cube $[0, 1]^n \subset \mathbb{R}^n$ whose density takes the form $\exp(-\psi)$ where the function ψ is assumed to be convex (but not strictly convex) with bounded pure second derivatives. Our argument relies on a transportation-cost inequality á la Talagrand.

1 Introduction

Consider a cube $Q \subset \mathbb{R}^n$ of sidelength ℓ parallel to the axes, that is, Q is a translation of the set $(0, \ell)^n \subset \mathbb{R}^n$ (or of its closure, equivalently). In this paper we prove a concentration inequality for a class of probability measures supported on Q .

Write $|\cdot|$ for the standard Euclidean norm in \mathbb{R}^n and $B^n = \{x \in \mathbb{R}^n; |x| \leq 1\}$ is the Euclidean unit ball centered at the origin. For a subset $A \subset \mathbb{R}^n$ denote $A + \varepsilon B^n = \{x + \varepsilon y; x \in A, y \in B^n\}$, the ε -neighborhood of the set A .

Theorem 1.1 *Let $\ell > 0$, $M \geq 0$ and let $Q \subset \mathbb{R}^n$ be a cube of sidelength ℓ parallel to the axes. Let μ be a probability measure supported on Q with density $\exp(-\psi)$ for a convex function $\psi : Q \rightarrow \mathbb{R}$ such that*

$$\partial^{ii}\psi(x) \leq M \quad \text{for all } x \in Q, i = 1, \dots, n. \quad (1)$$

Suppose that $A \subseteq \mathbb{R}^n$ is a measurable set with $\mu(A) \geq 1/2$. Then, for all $t > 0$,

$$\mu(A + tB^n) \geq 1 - \exp(-t^2/\alpha^2) \quad (2)$$

where $\alpha = \alpha(\ell, M) = 3\ell e^{M\ell^2/8}$.

Theorem 1.1 is equivalent to a logarithmic Sobolev inequality and to a concentration inequality for Lipschitz functions, see Section 4 below. In probabilistic terminology, we consider uniformly bounded random variables X_1, \dots, X_n , possibly dependent, whose joint distribution satisfies the convexity/concavity assumptions of Theorem 1.1. Our results correspond to bounds for the variance and tail distribution of $f(X_1, \dots, X_n)$ where f is a Lipschitz function on \mathbb{R}^n .

We emphasize that we are not assuming any product structure, any symmetries nor strict convexity for the function ψ from Theorem 1.1. There is a vast body of literature pertaining to the case in which the measure μ is an arbitrary product measure in the cube, see Talagrand [24], Marton [17], Dembo and Zeitouni [8], Ledoux [15] and others. When the function ψ from Theorem 1.1 admits a uniform positive lower bound for the Hessian,

*School of Mathematical Sciences, Tel-Aviv University, Tel Aviv 69978, Israel. Email: klartag@tau.ac.il

the conclusion of Theorem 1.1 is well-known and essentially goes back to Bakry and Émery [1].

How can we produce probability measures satisfying the assumptions of Theorem 1.1 with, say, $M = 1$? One may begin with the standard Gaussian density in \mathbb{R}^n , the function

$$\gamma_n(x) = (2\pi)^{-(n/2)} \exp(-|x|^2/2) \quad (x \in \mathbb{R}^n).$$

The restriction of γ_n to any cube $Q \subset \mathbb{R}^n$, normalized to be a probability density, surely satisfies the assumptions of Theorem 1.1 with $M = 1$. Furthermore, begin with any probability density $\rho : \mathbb{R}^n \rightarrow [0, \infty)$ which is log-concave (that is, the function $-\log \rho$ is convex). Consider the convolution

$$f(x) = (\rho * \gamma_n)(x) = \int_{\mathbb{R}^n} \rho(y) \gamma_n(x - y) dy.$$

Then f is a smooth, log-concave probability density according to the Prékopa-Leindler inequality. Furthermore, it is straightforward to verify that for any $x \in \mathbb{R}^n$,

$$(\nabla^2 \log f)(x) \geq -Id \quad (3)$$

in the sense of symmetric matrices, where Id is the identity matrix and $\nabla^2 \log f$ is the Hessian of $\log f$. We conclude that the probability measure on the cube Q whose density is proportional to the restriction of f to Q , satisfies the assumptions of Theorem 1.1 with $M = 1$. It is also possible to view the probability densities that appear in Theorem 1.1 as convex perturbations of probability densities proportional to $x \mapsto \exp(x \cdot v)$ on the cube. Here $x \cdot v$ is the standard scalar product of $x, v \in \mathbb{R}^n$.

One cannot replace $\alpha(\ell, M)$ in Theorem 1.1 by a dimension-free expression that is subexponential in $M\ell^2$, see Remark 4.4 below. We say that a vector $x \in \mathbb{R}^n$ is proportional to one of the standard unit vectors when it has at most one non-zero entry. A unit cube has sidelength one. Theorem 1.1 will be deduced from the following result:

Theorem 1.2 *Let $R \geq 1$ and let $Q \subset \mathbb{R}^n$ be a unit cube parallel to the axes. Let μ be a probability measure supported on Q with a log-concave density f such that*

$$f(\lambda x + (1 - \lambda)y) \leq R[\lambda f(x) + (1 - \lambda)f(y)] \quad (4)$$

for any $0 < \lambda < 1$ and any $x, y \in Q$ for which $x - y$ is proportional to one of the standard unit vectors. Suppose that $A \subseteq \mathbb{R}^n$ is a measurable set with $\mu(A) \geq 1/2$. Then for all $t > 0$,

$$\mu(A + tB^n) \geq 1 - \exp(-t^2/\alpha^2)$$

where $\alpha = \alpha(R) = 3R$.

The inequality (4) holds true with $R = 1$ when f is a convex function. By degenerating Theorem 1.2 to the petty case where $R = 1$ we arrive at the following peculiar corollary:

Corollary 1.3 *Let $Q \subset \mathbb{R}^n$ be a unit cube. Let μ be a probability measure on Q whose density is both log-concave and convex in Q . Then for any measurable $A \subseteq \mathbb{R}^n$ and $t > 0$,*

$$\mu(A) \geq 1/2 \quad \implies \quad \mu(A + tB^n) \geq 1 - \exp(-t^2/9).$$

A moment of reflection reveals that there do exist positive, integrable functions on the cube that are simultaneously log-concave and convex, such as $x \mapsto [b + (x \cdot v)]^p$ for $p \geq 1$. It is also evident that one cannot eliminate neither the log-concavity assumption nor the convexity assumption from Corollary 1.3.

The proof of Theorem 1.2 uses transportation of measure in order to analyze the deficit in the Prékopa-Leindler inequality, an idea proposed also in Eldan and Klartag [9]. Rather than working directly with the supremum-convolution, we prefer to analyze another expression that somewhat resembles the relative-entropy functional. Let us shed some light on this expression. Suppose that f and g are non-negative functions defined on \mathbb{R}^n . For a point $x \in \mathbb{R}^n$ in which f is positive and differentiable, and for a point $y \in \mathbb{R}^n$ in which g is positive, we set

$$S_y\{g, f\}(x) = f(x) \log \frac{g(y)}{f(x)} - \nabla f(x) \cdot (y - x). \quad (5)$$

Denote $\text{Supp}(f) = \{x; f(x) \neq 0\}$. For functions $f, g : \mathbb{R}^n \rightarrow [0, \infty)$ and a map $T : \text{Supp}(f) \rightarrow \text{Supp}(g)$ abbreviate

$$S_T\{g, f\}(x) = S_{T(x)}\{g, f\}(x) \quad (x \in \text{Supp}(f)), \quad (6)$$

assuming that the right-hand side is well-defined. Next, suppose that f and g have a finite, positive integral. A measurable map $T : \overline{\text{Supp}(f)} \rightarrow \overline{\text{Supp}(g)}$ is called a *transportation map from f to g* if for any measurable set $A \subseteq \text{Supp}(g)$,

$$\left(\frac{1}{\int g}\right) \int_A g = \left(\frac{1}{\int f}\right) \int_{T^{-1}(A)} f.$$

That is, T pushes forward the probability measure whose density is proportional to f , to the probability measure whose density is proportional to g . Two important examples of transportation maps in \mathbb{R}^n are the Brenier map [4] and the Knothe map [14].

Definition 1.4 *Let f, g be two non-negative functions on \mathbb{R}^n with a finite, positive integral. Assume that f is differentiable almost-everywhere in $\text{Supp}(f)$. Set*

$$\text{Tire}(g || f) = \sup_T \left[\int_{\text{Supp}(f)} S_T\{g, f\}(x) dx - \left(\int f\right) \log \frac{\int g}{\int f} \right], \quad (7)$$

where the supremum runs over all transportation maps T from f to g for which the integral of $S_T\{g, f\}$ is well-defined. Here, *Tire* is an acronym of “Translation-Invariant Relative Entropy”.

The notion is indeed translation-invariant: For functions f, g as in Definition 1.4 and for $x_0 \in \mathbb{R}^n$, denoting $\tau_{x_0}(g)(x) = g(x - x_0)$ we have

$$\text{Tire}(g || f) - \text{Tire}(\tau_{x_0}(g) || f) = \int_{\mathbb{R}^n} (\nabla f(x) \cdot x_0) dx = 0,$$

where we assume that f is locally-Lipschitz and vanishes at infinity in order to justify the integration by parts. In the log-concave case, the quantity $\text{Tire}(g || f)$ is indeed related to relative entropy as is demonstrated in Lemma 4.1 below.

The remainder of this paper is devoted to the proofs of the aforementioned theorems and to related results. We write \overline{A} for the closure of the set A , and \log stands for the natural logarithm. By “measurable” we always mean Borel-measurable. Needless to say, it

is certainly possible to consider $Tire(g || f)$ for non-negative functions defined only on a subset of \mathbb{R}^n by treating such functions as zero outside their original domain of definition.

Acknowledgements. I thank Dario Cordero-Erausquin and Ronen Eldan for interesting, related discussions. I am grateful to Nathael Gozlan and to the anonymous referee for their valuable suggestions and for correcting a mistake in an earlier version of this manuscript. The research was supported in part by the Israel Science Foundation and by a Marie Curie Reintegration Grant from the Commission of the European Communities.

2 Convex functions on an interval

Let $I, J \subset \mathbb{R}$ be two intervals of finite, positive length and let f, g be positive, integrable functions defined on I, J respectively. The *monotone transportation map from f to g* is the map $T : \bar{I} \rightarrow \bar{J}$ defined via

$$\left(\frac{1}{\int_I f} \right) \int_I f(t) 1_{\{t < x\}} dt = \left(\frac{1}{\int_J g} \right) \int_J g(t) 1_{\{t < T(x)\}} dt \quad (x \in \bar{I})$$

where $1_{\{t < x\}}$ equals one when $t < x$ and vanishes otherwise. The map T is uniquely defined, as f, g are positive and integrable. Furthermore, T is an absolutely-continuous, strictly-increasing function. Observe that the monotone transportation in one dimension is indeed a transportation map and that for almost every $x \in I$,

$$T'(x) = \left(\frac{\int_J g}{\int_I f} \right) \frac{f(x)}{g(T(x))}. \quad (8)$$

We will frequently encounter the case where $I = J$. Clearly, in this case $T(x) = x$ for $x \in \partial I$, where ∂I are the two endpoints of the interval I . Our goal in this section is to prove the following transportation-cost inequality in one dimension:

Proposition 2.1 *Let $R \geq 1$ and let $I \subset \mathbb{R}$ be an interval of length one. Let $f : I \rightarrow (0, \infty)$ be an absolutely-continuous function which satisfies*

$$f(\lambda x + (1 - \lambda)y) \leq R[\lambda f(x) + (1 - \lambda)f(y)] \quad \text{for all } x, y \in I, 0 < \lambda < 1.$$

Let g be a positive, integrable function on I , and let T be the monotone transportation map from f to g . Then,

$$\begin{aligned} \int_I |T(x) - x|^2 f(x) dx &\leq CR^2 \left[\int_I S_T\{g, f\} - \left(\int_I f \right) \log \frac{\int_I g}{\int_I f} \right] \\ &\leq CR^2 \cdot Tire(g || f), \end{aligned} \quad (9)$$

where $C \leq 40/9$ is a universal constant.

The proof of Proposition 2.1 requires a few lemmata. Our first lemma is essentially an infinitesimal version of the Prékopa-Leindler inequality, and its proof follows the transportation proofs given by Barthe [2], Cordero-Erausquin [6], Henstock-Macbeath [11] and Talagrand [23]. For $t \in \mathbb{R}$ denote

$$\Lambda(t) = \min\{|t|, t^2\}.$$

Lemma 2.2 *Let $I \subset \mathbb{R}$ be an interval of finite, positive length. Let f, g be positive, integrable functions on I with f being absolutely continuous. Then,*

$$\int_I \Lambda(T'(x) - 1) f(x) dx \leq \frac{10}{3} \left[\int_I S_T\{g, f\} - \left(\int_I f \right) \log \frac{\int_I g}{\int_I f} \right]$$

where T is the monotone transportation map from f to g .

Proof: We use (8) and compute

$$\begin{aligned}
\int_I S_T\{g, f\} &= \int_I \left[f(x) \log \frac{g(T(x))}{f(x)} - f'(x)(T(x) - x) \right] dx \\
&= \left(\int f \right) \log \frac{\int g}{\int f} + \int_I \left[f(x) \log \frac{1}{T'(x)} - f'(x)(T(x) - x) \right] dx \\
&= \left(\int f \right) \log \frac{\int g}{\int f} + \int_I [-f(x) \log T'(x) + f(x)(T'(x) - 1)] dx
\end{aligned}$$

where the integration by parts is legitimate as $f(x)(T(x) - x)$ is an absolutely-continuous function in I that vanishes on ∂I . In order to complete the proof of the lemma it remains to show that for all $x > 0$,

$$-\log x + (x - 1) \geq \frac{3}{10} \cdot \min\{|x - 1|, (x - 1)^2\}. \quad (10)$$

Indeed, for $0 < x \leq 2$ we use the Cauchy-Schwartz inequality and obtain

$$\begin{aligned}
(x - 1) - \log x &= \int_1^x \frac{t - 1}{t} dt = \int_0^{|x-1|} \frac{t}{1 + \operatorname{sgn}(x - 1)t} dt \geq \int_0^{|x-1|} \frac{t}{1 + t} dt \\
&\geq \left(\int_0^{|x-1|} t dt \right)^2 / \left(\int_0^{|x-1|} (1 + t) t dt \right) = \frac{(x - 1)^2}{2(1 + 2|x - 1|/3)} \geq \frac{3(x - 1)^2}{10},
\end{aligned}$$

where $\operatorname{sgn}(x) = 1$ for $x > 0$ and $\operatorname{sgn}(x) = -1$ for $x < 0$. The inequality (10) is valid in particular for $x = 2$. For $x > 2$ the derivative of the left-hand side in (10) exceeds that of the right-hand side. Hence (10) holds true for all $x > 0$. \square

Remark 2.3 The proof of Lemma 2.2 admits a generalization to n dimensions, in which one utilizes the Brenier map in place of the transportation map T . See Barthe [3] and McCann [18] for related arguments. In this way one obtains the inequality

$$\operatorname{Tire}(g || f) \geq 0, \quad (11)$$

which is valid for any Lipschitz, non-negative, compactly-supported functions f and g on \mathbb{R}^n with a finite, positive integral.

Lemma 2.4 *Let $R \geq 1$ and let $I \subset \mathbb{R}$ be an interval of length one. Assume that ρ is a positive, integrable function on I that satisfies*

$$\rho(\lambda x + (1 - \lambda)y) \leq R[\lambda\rho(x) + (1 - \lambda)\rho(y)] \quad \text{for all } x, y \in I, 0 < \lambda < 1. \quad (12)$$

Then for any $a, b \in I$ with $a < b$,

$$\int_a^b \rho(x) dx \leq \frac{R}{2} [\rho(a) + \rho(b)].$$

Proof: We simply integrate (12) over $\lambda \in [0, 1]$. Since $b - a \leq 1$, then

$$\int_a^b \rho \leq \int_0^1 \rho(\lambda a + (1 - \lambda)b) d\lambda \leq R \int_0^1 [\lambda\rho(a) + (1 - \lambda)\rho(b)] d\lambda = R \frac{\rho(a) + \rho(b)}{2},$$

and the lemma is proven. \square

The following lemma is a one-dimensional Poincaré-type inequality. The proof closely follows the argument by Cheeger [5]. Recall that $\Lambda(t) = \min\{|t|, t^2\}$.

Lemma 2.5 Let $I \subset \mathbb{R}$ be an interval of length one and let $R \geq 1$. Let ρ be a positive, integrable function on I with

$$\rho(\lambda x + (1 - \lambda)y) \leq R[\lambda\rho(x) + (1 - \lambda)\rho(y)] \quad \text{for all } x, y \in I, 0 < \lambda < 1.$$

Then for any absolutely-continuous function $f : \bar{I} \rightarrow \mathbb{R}$ with $f|_{\partial I} = 0$,

$$\int_I \Lambda(f) \rho \leq \frac{4}{3} R^2 \int_I \Lambda(f') \rho. \quad (13)$$

Here, ∂I consists of the two endpoints of the interval I .

Proof: Multiplying ρ by a constant, we may assume that $\int_I \rho = 1$. Let g be an absolutely-continuous, non-negative function with $g|_{\partial I} = 0$. In the first step of the proof we show that

$$\int_I g \rho \leq \frac{R}{2} \int_I |g'| \rho. \quad (14)$$

Denote $J = g(I) = \{g(x); x \in I\}$, an interval whose left boundary point is zero. We apply the change of variables $y = g(x)$ and conclude that

$$\int_I |g'(x)| \rho(x) dx = \int_J \left(\sum_{x \in g^{-1}(y)} \rho(x) \right) dy. \quad (15)$$

See, e.g., Leoni [16, Theorem 3.65] for a proof of this change of variables formula. For any $0 \neq y \in J$ consider the open set $I_y = \{x \in I; g(x) > y\}$. When y is a regular non-zero value of g , the open set I_y is a finite union of intervals with disjoint closures. According to Lemma 2.4, for any such y ,

$$\int_{I_y} \rho \leq \frac{R}{2} \left[\sum_{x \in g^{-1}(y)} \rho(x) \right].$$

The one-dimensional Sard's lemma for absolutely-continuous functions (see, e.g., Leoni [16, Remark 8.9]) implies that almost any $y \in J$ is a regular value of g . Therefore, from (15) we obtain

$$\int_I |g'(x)| \rho(x) dx \geq \frac{2}{R} \int_J \left(\int_{\{x; g(x) > y\}} \rho(x) dx \right) dy = \frac{2}{R} \int_I g \rho$$

where the last equality follows from application of Fubini's theorem. Thus (14) is proven. In order to prove (13), observe that for any $x \geq 0$ and $0 \leq y \leq 1$,

$$xy \leq R\Lambda(x) + \frac{y^2}{4R}. \quad (16)$$

Indeed, (16) holds for $x \geq 1$ since the coefficient in front of $\Lambda(x)$ is at least one, and (16) may be directly proven for $0 \leq x \leq 1$ by completing a square. Let $f : \bar{I} \rightarrow \mathbb{R}$ be an absolutely-continuous function with $f|_{\partial I} = 0$. Applying (14) with $g = \Lambda(|f|)$ and using (16) we see that

$$\begin{aligned} \int_I \Lambda(|f|) \rho &\leq \frac{R}{2} \int_I \Lambda'(|f|) |f'| \rho \leq R \int_I |f'| \min\{|f|, 1\} \rho \\ &\leq R \cdot \left[R \int_I \Lambda(|f'|) \rho + \frac{1}{4R} \int_I \min\{|f|^2, 1\} \rho \right] \\ &\leq R^2 \int_I \Lambda(|f'|) \rho + \frac{1}{4} \int_I \Lambda(|f|) \rho. \end{aligned}$$

By subtracting the right-most summand from the left-hand side, we deduce (13). \square

Proof of Proposition 2.1: Since $T(x) = x$ for $x \in \partial I$ we may invoke Lemma 2.5 and conclude that

$$\begin{aligned} \int_I \Lambda(T(x) - x) f(x) dx &\leq \frac{4}{3} R^2 \int_I \Lambda(T'(x) - 1) f(x) dx \\ &\leq \frac{40}{9} R^2 \left[\int_I S_T\{g, f\} - \left(\int_I f \right) \log \frac{\int_I g}{\int_I f} \right] \end{aligned}$$

where we used Lemma 2.2 in the last passage. Since I is an interval of length one and $T : \bar{I} \rightarrow \bar{I}$, then for any $x \in I$ we have $|T(x) - x| \leq 1$. Consequently, for any $x \in I$,

$$\Lambda(T(x) - x) = \min \{|T(x) - x|^2, |T(x) - x|\} = |T(x) - x|^2.$$

This completes the proof of (9). The proposition now follows from the definition of $Tire(g || f)$. \square

3 Induction on the dimension

In this section we obtain higher-dimensional analogs of Proposition 2.1.

Theorem 3.1 *Let $R \geq 1$ and let $Q \subset \mathbb{R}^n$ be a unit cube parallel to the axes. Assume that $f : Q \rightarrow (0, \infty)$ is a Lipschitz function with*

$$f(\lambda x + (1 - \lambda)y) \leq R[\lambda f(x) + (1 - \lambda)f(y)] \quad (17)$$

for any $0 < \lambda < 1$ and any $x, y \in Q$ for which $x - y$ is proportional to one of the standard unit vectors in \mathbb{R}^n . Let g be a positive, integrable function on Q . Then there exists a transportation map T from f to g such that

$$\begin{aligned} \int_Q |T(x) - x|^2 f(x) dx &\leq C R^2 \left[\int_Q S_T\{g, f\} - \left(\int_Q f \right) \log \frac{\int_Q g}{\int_Q f} \right] \\ &\leq C R^2 \cdot Tire(g || f), \end{aligned} \quad (18)$$

where $C \leq 40/9$ is a universal constant.

The requirement that f be a Lipschitz function should not be taken too seriously, as it may easily be replaced by other types of regularity assumptions. Theorem 3.1 will be proven by induction on the dimension, where the induction step is going to be Proposition 2.1 in disguise. Throughout this section we use

$$x = (y, r) \in \mathbb{R}^{n-1} \times \mathbb{R}$$

as coordinates in \mathbb{R}^n . For a function f defined on a subset of \mathbb{R}^n and for $y \in \mathbb{R}^{n-1}$, we write

$$f_y(r) = f(y, r)$$

whenever (y, r) is in the domain of definition of f . Abbreviate $\pi(y, r) = y$. For a subset $A \subseteq \mathbb{R}^n$ denote $\pi(A) = \{\pi(x); x \in A\}$. For a non-negative, integrable function f defined on a subset $A \subseteq \mathbb{R}^n$, we set

$$\pi(f)(y) = \int_{-\infty}^{\infty} f_y(r) 1_{\{(y,r) \in A\}} dr \quad (y \in \pi(A)).$$

Let $K \subseteq \mathbb{R}^n$ be a convex set. Let f, g be positive, integrable functions on K . We say that a map $T : K \rightarrow K$ transports the last coordinate monotonically if there exists a map $P : \pi(K) \rightarrow \pi(K)$ such that for almost any $y \in \pi(K)$, the function $g_{P(y)}$ is integrable and furthermore

$$T(y, r) = (P(y), T_y(r)) \quad (19)$$

for any r with $(y, r) \in K$, where T_y is the monotone transportation map from f_y to $g_{P(y)}$. The following lemma is a corollary to Proposition 2.1.

Lemma 3.2 *Let $R \geq 1$. Let $Q = A \times I \subset \mathbb{R}^n$ where $I \subset \mathbb{R}$ is an interval of length one and $A \subset \mathbb{R}^{n-1}$ is a convex set. Assume that f is a positive, Lipschitz function on Q , and that*

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq R[\lambda f(x_1) + (1 - \lambda)f(x_2)] \quad (20)$$

for any $0 < \lambda < 1$ and any $x_1, x_2 \in Q$ for which $x_1 - x_2$ is proportional to one of the standard unit vectors. Let g be a positive, integrable function on Q . Assume that $T : \overline{Q} \rightarrow \overline{Q}$ is a measurable map that transports the last coordinate monotonically. Then,

$$\int_Q |T_y(r) - r|^2 f(y, r) dy dr \leq CR^2 \left[\int_Q S_T\{g, f\} - \int_{\pi(Q)} S_P\{\pi(g), \pi(f)\} \right] \quad (21)$$

where P and T_y are as in (19), and $C \leq 40/9$ is a universal constant.

Proof: According to the definitions (5) and (6), for almost any $(y, r) \in Q$,

$$S_T\{g, f\}(y, r) = S_{T_y}\{g_{P(y)}, f_y\}(r) - \nabla_y f(y, r) \cdot (P(y) - y) \quad (22)$$

where ∇_y is the gradient in the y -variables. Thanks to our assumptions on f we may safely differentiate under the integral sign, thus

$$\nabla \pi(f)(y) \cdot (P(y) - y) = \int_I \nabla_y f(y, r) \cdot (P(y) - y) dr \quad (23)$$

for almost any choice of y . From (22) and (23),

$$\int_I S_{T_y}\{g_{P(y)}, f_y\}(r) dr = \int_I S_T\{g, f\}(y, r) dr + \nabla \pi(f)(y) \cdot (P(y) - y) \quad (24)$$

for almost any choice of y . The equality (24) may be reformulated as

$$\int_I S_{T_y}\{g_{P(y)}, f_y\} - \left(\int_I f_y \right) \log \frac{\int_I g_{P(y)}}{\int_I f_y} = \int_I S_T\{g, f\}(y, r) dr - S_P\{\pi(g), \pi(f)\}(y). \quad (25)$$

We may apply Proposition 2.1 thanks to our assumption (20) and obtain that

$$\int_I |T_y(r) - r|^2 f_y(r) dr \leq CR^2 \left[\int_I S_{T_y}\{g_{P(y)}, f_y\} - \left(\int_I f_y \right) \log \frac{\int_I g_{P(y)}}{\int_I f_y} \right]. \quad (26)$$

By combining (25) and (26) we see that for almost any $y \in \pi(Q)$,

$$\int_I |T_y(r) - r|^2 f_y(r) dr \leq CR^2 \left[\int_I S_T\{g, f\}(y, r) dr - S_P\{\pi(g), \pi(f)\}(y) \right]. \quad (27)$$

We now integrate (27) over $y \in \pi(Q)$ and deduce (21). \square

Remark: The identity (23) is the only place in the proof of Theorem 3.1 where we use the assumption that Q is a cube or a box, rather than, say, a parallelepiped.

Lemma 3.3 *Let $R \geq 1$ and let $Q \subset \mathbb{R}^n$ be a cube parallel to the axes. Assume that $f : Q \rightarrow \mathbb{R}$ is a Lipschitz function on Q , such that*

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq R[\lambda f(x_1) + (1 - \lambda)f(x_2)] \quad (28)$$

for any $0 < \lambda < 1$ and any $x_1, x_2 \in Q$ for which $x_1 - x_2$ is proportional to one of the standard unit vectors in \mathbb{R}^n . Then also

$$\pi(f)(\lambda y_1 + (1 - \lambda)y_2) \leq R[\lambda \pi(f)(y_1) + (1 - \lambda)\pi(f)(y_2)]$$

for any $0 < \lambda < 1$ and any $y_1, y_2 \in \pi(Q)$ for which $y_1 - y_2$ is proportional to one of the standard unit vectors in \mathbb{R}^{n-1} .

Proof: Fix $i = 1, \dots, n - 1$ and let e_i be the i^{th} standard unit vector. Condition (28) implies that for any $y \in \mathbb{R}^{n-1}$, $t_1, t_2, r \in \mathbb{R}$, $0 < \lambda < 1$ such that $(y + t_1 e_i, r) \in Q$ and $(y + t_2 e_i, r) \in Q$,

$$f(y + (\lambda t_1 + (1 - \lambda)t_2) e_i, r) \leq R[\lambda f(y + t_1 e_i, r) + (1 - \lambda)f(y + t_2 e_i, r)].$$

Let I be the interval for which $Q = \pi(Q) \times I$. Integrating with respect to r we have

$$\begin{aligned} \pi(f)(y + (\lambda t_1 + (1 - \lambda)t_2) e_i) &= \int_I f(y + (\lambda t_1 + (1 - \lambda)t_2) e_i, r) dr \\ &\leq R \int_I [\lambda f(y + t_1 e_i, r) + (1 - \lambda)f(y + t_2 e_i, r)] dr \\ &= R[\lambda \pi(f)(y + t_1 e_i) + (1 - \lambda)\pi(f)(y + t_2 e_i)], \end{aligned}$$

and the lemma is proven. \square

Proof of Theorem 3.1: We will prove by induction on the dimension n that there exists a transportation map T from f to g such that

$$\int_Q |T(x) - x|^2 f(x) dx \leq \frac{40}{9} R^2 \left[\int_Q S_T\{g, f\} - \left(\int_Q f \right) \log \frac{\int_Q g}{\int_Q f} \right]. \quad (29)$$

The case $n = 1$ is Proposition 2.1. Assume that the induction hypothesis was proven for dimension $n - 1$, and let us prove it for dimension n . Thus, suppose that we are given a cube $Q \subset \mathbb{R}^n$ and functions f, g which satisfy the assumptions of the theorem. In view of Lemma 3.3, we may apply the induction hypothesis for

$$\pi(Q), \pi(f), \pi(g).$$

Thus, there exists a transportation map $P : \overline{\pi(Q)} \rightarrow \overline{\pi(Q)}$ from $\pi(f)$ to $\pi(g)$ such that

$$\begin{aligned} \int_{\pi(Q)} |P(y) - y|^2 \pi(f)(y) dy & \\ &\leq \frac{40}{9} R^2 \left[\int_{\pi(Q)} S_P\{\pi(g), \pi(f)\} - \left(\int_{\pi(Q)} \pi(f) \right) \log \frac{\int_{\pi(Q)} \pi(g)}{\int_{\pi(Q)} \pi(f)} \right]. \end{aligned} \quad (30)$$

For $y \in \overline{\pi(Q)}$ let T_y be the monotone transportation map from f_y to $g_{P(y)}$, a strictly-increasing function which is well-defined for almost any $y \in \pi(Q)$. We set

$$T(y, r) = (P(y), T_y(r)) \quad \text{for } (y, r) \in \overline{Q}.$$

Then T transports the last coordinate monotonically. Hence, according to Lemma 3.2,

$$\int_Q |T_y(r) - r|^2 f(y, r) dy dr \leq \frac{40}{9} R^2 \left[\int_Q S_T\{g, f\} - \int_{\pi(Q)} S_P\{\pi(g), \pi(f)\} \right]. \quad (31)$$

It is straightforward to verify that the map T is a transportation map from f to g . In fact, the map T is precisely the Knothe transportation map from [14]. By summing (30) and (31), we conclude that

$$\begin{aligned} & \int_Q \left[|P(y) - y|^2 + |T_y(r) - r|^2 \right] f(y, r) dy dr \\ & \leq \frac{40}{9} R^2 \left[\int_Q S_T\{g, f\} - \left(\int_{\pi(Q)} \pi(f) \right) \log \frac{\int_{\pi(Q)} \pi(g)}{\int_{\pi(Q)} \pi(f)} \right] \\ & = \frac{40}{9} R^2 \left[\int_Q S_T\{g, f\} - \left(\int_Q f \right) \log \frac{\int_Q g}{\int_Q f} \right]. \end{aligned} \quad (32)$$

All that remains is to note that when $x = (y, r)$,

$$|T(x) - x|^2 = |P(y) - y|^2 + |T_y(r) - r|^2.$$

The bound (29) follows from (32), and the theorem is proven. \square

4 Log-concavity

We begin this section with a discussion of Definition 1.4. As we shall see, this definition fits very nicely with log-concave functions. Given two probability densities f and g in \mathbb{R}^n we write

$$D(g || f) = \int_{\mathbb{R}^n} \left[\log \frac{g(y)}{f(y)} \right] g(y) dy$$

for the relative entropy or the Kullback-Leibler divergence of g from f .

Lemma 4.1 *Let $f, g : \mathbb{R}^n \rightarrow [0, \infty)$ be probability densities. Assume that f is log-concave. Then,*

$$\text{Tire}(g || f) \leq D(g || f).$$

Proof: The function f is differentiable almost-everywhere in the convex set $\text{Supp}(f)$ as it is a log-concave function. Denote $f = e^{-\psi}$. From the convexity of ψ we see that for any point $x \in \text{Supp}(f)$ in which f is differentiable,

$$\psi(x) + \nabla\psi(x) \cdot (y - x) \leq \psi(y) \quad (y \in \text{Supp}(f)).$$

Let T be any transportation map from f to g . Denoting $\varphi = -\log g$ we find that for almost any $x \in \text{Supp}(f)$,

$$\begin{aligned} S_T\{g, f\}(x) &= \left[f(x) \log \frac{g(T(x))}{f(x)} - \nabla f(x) \cdot (T(x) - x) \right] \\ &= f(x) [\psi(x) - \varphi(T(x)) + \nabla\psi(x) \cdot (T(x) - x)] \\ &\leq f(x) [\psi(T(x)) - \varphi(T(x))] = f(x) \log \frac{g(T(x))}{f(T(x))}. \end{aligned}$$

Since T is a transportation map from f to g , then by applying the change of variables $y = T(x)$ we obtain

$$\begin{aligned} \int_{\text{Supp}(f)} S_T\{g, f\} &\leq \int_{\text{Supp}(f)} \log \frac{g(T(x))}{f(T(x))} f(x) dx \\ &= \int_{\text{Supp}(g)} \left[\log \frac{g(y)}{f(y)} \right] g(y) dy = D(g \| f) \end{aligned}$$

and the lemma is proven. \square

For a log-concave density f , we may think about $\text{Tire}(g \| f)$ as a parameter measuring the proximity of g to a translate of f . Let us mention here additional upper bounds for $\text{Tire}(g \| f)$. Let $f, g : \mathbb{R}^n \rightarrow [0, \infty)$ have finite, positive integrals and denote $\psi = -\log f$ and $\varphi = -\log g$. According to (5),

$$\sup_{y \in \text{Supp}(g)} S_y\{g, f\}(x) = [\nabla f(x) \cdot x - f(x) \log f(x)] + f(x) \varphi^*(\nabla \psi(x)) \quad (33)$$

where $\varphi^*(v) = \sup_{y \in \text{Supp}(g)} [v \cdot y - \varphi(y)]$ is the usual Legendre transform of φ . Consequently, when f is locally-Lipschitz and $x \mapsto |x|f(x)$ vanishes at infinity, we have the bound

$$\text{Tire}(g \| f) \leq \int_{\mathbb{R}^n} \left[\varphi^*(\nabla \psi(x)) - \log \left(\frac{\int g}{\int f} \cdot f(x) \right) - n \right] f(x) dx. \quad (34)$$

Inequality (34) is perhaps less appealing than Lemma 4.1, yet it is applicable also in the non log-concave case.

Our original motivation for Definition 1.4 is that at least in the smooth, log-concave case, the expression in (33) equals $\partial h_\varepsilon(x) / \partial \varepsilon|_{\varepsilon=0}$ where

$$h_\varepsilon(x) = \sup_{y \in \mathbb{R}^n} f(x + \varepsilon y)^{1-\varepsilon} g(x - (1-\varepsilon)y)^\varepsilon \quad (x \in \mathbb{R}^n).$$

In other words, $\text{Tire}(g \| f)$ is related to a kind of ‘‘mixed volume’’ of log-concave functions, see [12, Section 3] for further explanations.

Suppose that μ_1 and μ_2 are Borel probability measures on \mathbb{R}^n . The *transportation cost* between μ_1 and μ_2 is defined to be

$$W_2^2(\mu_1, \mu_2) = \inf_{\gamma} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\gamma(x, y)$$

where the infimum runs over all couplings γ of μ_1 and μ_2 , i.e., all Borel probability measures γ on $\mathbb{R}^n \times \mathbb{R}^n$ whose first marginal is μ_1 and whose second marginal is μ_2 . See, e.g., Villani’s book [25] for more information about the transportation metric W_2 . The following theorem reminds us of Talagrand’s transportation-cost inequalities for product measures from [23].

Theorem 4.2 *Let $R \geq 1$ and let $Q \subset \mathbb{R}^n$ be a unit cube parallel to the axes. Suppose that μ is a probability measure on Q with a log-concave density f . Assume that*

$$f(\lambda x + (1-\lambda)y) \leq R[\lambda f(x) + (1-\lambda)f(y)] \quad (35)$$

for any $0 < \lambda < 1$ and any $x, y \in Q$ for which $x - y$ is proportional to one of the standard unit vectors in \mathbb{R}^n . Let ν be a probability measure on Q that is absolutely continuous with respect to μ . Then,

$$W_2^2(\mu, \nu) \leq CR^2 D(\nu \| \mu)$$

where $D(\nu \| \mu) = \int g(\log g) d\mu$ for $g = d\nu/d\mu$, the usual relative entropy functional, and where $C \leq 40/9$ is a universal constant.

Proof: By a standard approximation argument (e.g., convolve μ with a tiny gaussian and restrict to the cube Q), we may assume that f and g are positive and C^1 -smooth up to the boundary in Q , and in particular both functions are positive and Lipschitz. According to Theorem 3.1 and Lemma 4.1,

$$W_2^2(\mu, \nu) \leq CR^2 \cdot \text{Tire}(g || f) \leq CR^2 \cdot D(\nu || \mu). \quad \square$$

Transportation-cost inequalities such as Theorem 4.2 are the subject of the comprehensive survey by Gozlan and Léonard [10]. The fact that transportation-cost inequalities imply concentration inequalities goes back to Marton [17]. The following proof reproduces her argument, and is included here for completeness.

Proof of Theorem 1.2: Denote $E = Q \setminus (A + tB^n)$. If $\mu(E) = 0$ then there is nothing to prove. Otherwise, we apply Theorem 4.2 for the measure $\nu = \mu|_E$. Thus there exists a coupling γ of μ and $\mu|_E$ with

$$\int_{Q \times E} |y - x|^2 d\gamma(x, y) \leq \frac{40}{9} R^2 D(\nu || \mu) = \frac{40}{9} R^2 \cdot \log \frac{1}{\mu(E)}.$$

According to the Markov-Chebyshev inequality, there exists a subset $F \subseteq Q \times E$ with $\gamma(F) \geq 41/81$ such that for any $(x, y) \in F$,

$$|y - x|^2 \leq 9R^2 \log \frac{1}{\mu(E)}. \quad (36)$$

Since γ is a coupling and $\mu(A) \geq 1/2$ with $\gamma(F) \geq 41/81$, there exists $(x, y) \in F$ with $x \in A$. For such (x, y) ,

$$x \in A, y \in E \quad \text{and} \quad |x - y| \leq 3R \cdot \sqrt{\log \frac{1}{\mu(E)}}$$

where we used (36). However, all points in E are of distance at least t from all points of A . Consequently,

$$t \leq 3R \cdot \sqrt{\log \frac{1}{\mu(E)}}.$$

Therefore $\mu(E) \leq \exp(-t^2/\alpha^2)$ for $\alpha = 3R$ and $\mu(A + tB^n) \geq 1 - \exp(-t^2/\alpha^2)$, as required. \square

Proof of Theorem 1.1: Let $T > 0$. Observe that the validity of both the assumptions and the conclusions of the theorem is not altered under the scaling

$$\ell \mapsto T\ell, \quad M \mapsto T^{-2}M.$$

We may thus normalize so that $\ell = 1$. All that remains is to verify that the assumptions of Theorem 1.2 are satisfied with $R = e^{M/8}$. Fix $i = 1, \dots, n$ and $x \in Q$ and denote $h(t) = \psi(x + te_i)$. Then h is well-defined on a certain interval $I \subset \mathbb{R}$ of length one, and our goal is to show that for any $a, b \in I$ and $0 < \lambda < 1$,

$$e^{-h(\lambda a + (1-\lambda)b)} \leq e^{M/8} \left[\lambda e^{-h(a)} + (1-\lambda)e^{-h(b)} \right]. \quad (37)$$

In view of the arithmetic/geometric means inequality, the desired inequality (37) would follow once we establish that

$$-h(\lambda a + (1-\lambda)b) \leq M/8 - \lambda h(a) - (1-\lambda)h(b). \quad (38)$$

In order to prove (38), we use our assumption that $h''(t) \leq M$ in the interval I . According to the Taylor theorem, for any $x, y \in I$,

$$h(y) - h(x) - h'(x)(y - x) \leq M \frac{(x - y)^2}{2}. \quad (39)$$

We will apply inequality (39) for $y = a, b$ and $x = \lambda a + (1 - \lambda)b$, then add the resulting inequalities with coefficients λ and $1 - \lambda$. This yields

$$\lambda h(a) + (1 - \lambda)h(b) - h(\lambda a + (1 - \lambda)b) \leq M \frac{\lambda(1 - \lambda)(b - a)^2}{2} \leq \frac{M}{8} \quad (40)$$

as $\lambda(1 - \lambda) \leq 1/4$ and $|b - a| \leq 1$. The inequality (38) follows from (40). \square

It is well-known (see, e.g., V. Milman and Schechtman [20, Section 2 and Appendix V]) that Theorem 1.1 implies a concentration inequality for Lipschitz functions as follows:

Corollary 4.3 *Let Q, μ, α be as in Theorem 1.1 (or as in Theorem 1.2). Let $f : Q \rightarrow \mathbb{R}$ be a 1-Lipschitz function, i.e., $|f(x) - f(y)| \leq |x - y|$ for any $x, y \in Q$. Denote $E = \int_Q f d\mu$. Then, for any $t > 0$,*

$$\mu \{x \in Q; |f(x) - E| \geq t\} \leq C e^{-ct^2/\alpha^2}$$

where $c, C > 0$ are universal constants.

In particular, we deduce from Corollary 4.3 that in the notation of Theorem 1.1,

$$\text{Cov}(\mu) \leq C\alpha^2 \cdot \text{Id} \quad (41)$$

in the sense of symmetric matrices, where $\text{Cov}(\mu)$ is the covariance matrix of the probability measure μ and $C > 0$ is a universal constant.

Remark 4.4 Regarding the dependence of $\alpha(\ell, M)$ on M in Theorem 1.1: Let X_0, \dots, X_n be independent standard Gaussian random variables. Consider the random vector

$$Y = \frac{(X_1, \dots, X_n)}{100\sqrt{\log n}} + \frac{(X_0, \dots, X_0)}{100\sqrt{\log n}},$$

and let Z be the conditioning of Y to the cube $Q = [-1/2, 1/2]^n$. Denote by μ the distribution of Z , a probability measure on Q . It is not too difficult to verify that μ satisfies the requirements of Theorem 1.1 with $\ell = 1$ and $M = C \log n$. Set $A = \{x \in Q; \sum_i x_i \leq 0\}$. Then $\mu(A) = 1/2$. However, one may compute that for any $t \leq cn^{1/2}/\sqrt{\log n}$,

$$\mu(A + tB^n) \leq 2/3.$$

This shows that $\alpha(1, C \log n) \geq cn^{1/2}/\sqrt{\log n}$. Therefore the exponential dependence of the dimension-free expression $\alpha(\ell, M)$ on $\ell^2 M$ is inevitable. A simple variant of this example shows that it is also impossible to replace the cube Q of sidelength ℓ in Theorem 1.1 by a Euclidean ball of radius $\ell\sqrt{n}$. For another example in which the cube behaves better than the Euclidean ball, see [13, Corollary 3].

It was explained by E. Milman [19] that in the log-concave case, Gaussian concentration inequalities, quadratic transportation-cost inequalities, and log-Sobolev inequalities are all essentially equivalent up to universal constants. In particular, by using the results of Otto and Villani [21, Corollary 3.1], we deduce from Theorem 4.2 the following log-Sobolev and Poincaré inequalities:

Corollary 4.5 *Let ℓ, M, Q, μ be as in Theorem 1.1 (or as in Theorem 1.2, with $\ell = 1$ and $R = e^{M/8}$). Then, for any locally-Lipschitz function $f : Q \rightarrow \mathbb{R}$ with $\int_Q f^2 d\mu = 1$,*

$$\int_Q f^2 \log(f^2) d\mu \leq C_1 \ell^2 e^{M\ell^2/4} \int_Q |\nabla f|^2 d\mu,$$

and for any integrable, locally-Lipschitz function $f : Q \rightarrow \mathbb{R}$ with $\int_Q f d\mu = 0$,

$$\int_Q f^2 d\mu \leq C_2 \ell^2 e^{M\ell^2/4} \int_Q |\nabla f|^2 d\mu.$$

Here, $C_1 \leq 160/9$ and $C_2 \leq 20/9$ are universal constants.

It is conceivable that Theorem 1.1 and Corollary 4.5 will turn out to be relevant to the analysis of lattice models in physics. For instance, one may suggest an Ising model with bounded, real spins as in Royer [22, Section 4.2] in which the assumptions of Theorem 1.1 are satisfied. Essentially, we require that the spins lie in the interval $[-1, 1]$, that the entire Hamiltonian is convex (just convex, not strictly-convex) and that the second derivatives of the pairwise potentials and the self-interactions are bounded. Perhaps the logarithmic Sobolev inequality of Corollary 4.5 may be of some use in this context.

5 Yet another approach for Theorem 4.2

In this section we present a sketch of an alternative proof of Theorem 4.2, in spirit of the transportation arguments of Cordero-Erausquin [6]. The derivation below is applicable for the two types of transportation maps, Brenier and Knothe.

Let f, g, μ and ν satisfy the assumptions of Theorem 4.2. As is explained above, it suffices to consider the case where f and g are positive, Lipschitz functions. In particular, it is well-known that both the Brenier map and the Knothe map from μ to ν are C^1 -smooth up to the boundary (see Cordero-Erausquin [7]).

Denote $\psi = -\log f$, a convex function. Let F be any smooth transportation map from μ to ν . Then, similarly to (8) above, we have

$$\log |\det F'(x)| = -\psi(x) + \psi(F(x)) - \log g(F(x)) \quad (x \in Q)$$

where $F'(x)$ is the $n \times n$ matrix which is the derivative of F . In the case where F is the Brenier map, the matrix $F'(x)$ is symmetric and positive-definite. In the case where F is the Knothe map, the matrix $F'(x)$ is upper-triangular with positive entries on the diagonal. In both cases, denoting $F = (F_1, \dots, F_n)$,

$$\log |\det F'(x)| = \log \det F'(x) \leq \sum_{i=1}^n \log \partial^i F_i(x) \quad (x \in Q). \quad (42)$$

Indeed, in the Knothe case (42) is simply an equality, while in the Brenier case we may use Hadamard's determinant inequality in order to establish (42). Next, denote $\theta(x) = F(x) - x$, so that $\partial^i F_i(x) = 1 + \partial^i \theta_i(x)$. We use the elementary inequality for the logarithm function in (10) and obtain

$$\sum_{i=1}^n \log \partial^i F_i(x) \leq \sum_{i=1}^n \left[\partial^i \theta_i(x) - \frac{3}{10} \Lambda(\partial^i \theta_i(x)) \right].$$

The convexity of ψ implies that

$$\psi(F(x)) - \psi(x) \geq \nabla\psi(x) \cdot (F(x) - x) = \nabla\psi(x) \cdot \theta(x) = \sum_{i=1}^n \theta_i(x) \partial^i \psi(x).$$

Combining all of the above, we arrive at the inequality

$$\log g(F(x)) \geq \frac{3}{10} \sum_{i=1}^n \Lambda(\partial^i \theta_i(x)) - \sum_{i=1}^n [\partial^i \theta_i(x) - \partial^i \psi(x) \cdot \theta_i(x)], \quad (43)$$

valid pointwise in Q . Here comes a fundamental property of both the Brenier map and the Knothe map: In both cases, the map F preserves each of the $(n-1)$ -dimensional facets of the cube Q . In other words, let A_1^\pm, \dots, A_n^\pm be an enumeration of all of the $2n$ facets of dimension $n-1$ of the cube Q . Assume that $\pm e_i$ is the outer unit normal to the cube Q at the facet A_i^\pm . We claim that for any i ,

$$x \in A_i^\pm \implies \theta_i(x) = 0. \quad (44)$$

It is quite clear that (44) holds in the case of the Knothe map. In order to argue for (44) in the Brenier case, recall that here

$$(F(x) - F(y)) \cdot (x - y) > 0 \quad (x, y \in Q, x \neq y)$$

as F is the gradient of a strictly-convex function. In particular, when $F(x) \in A_i^\pm$ then necessarily $x \pm t e_i \notin Q$ for $t \in (0, \varepsilon)$ for some $\varepsilon > 0$. Hence $F(x) \in A_i^\pm$ implies that $x \in A_i^\pm$. Arguing similarly for the inverse map F^{-1} , which is the Brenier map from ν to μ , we conclude that (44) holds true in the Brenier case as well.

We may now multiply (43) by $e^{-\psi}$ and integrate over the cube Q . Observe that for any $i = 1, \dots, n$,

$$\int_Q [\partial^i \theta_i(x) - \partial^i \psi(x) \cdot \theta_i(x)] e^{-\psi} = \int_Q \partial^i (\theta_i e^{-\psi}) = 0$$

thanks to the boundary condition (44). Furthermore, this boundary condition allows us to use the one-dimensional Lemma 2.5, and conclude that

$$\int_Q \Lambda(\theta_i(x)) d\mu(x) \leq \frac{4R^2}{3} \int_Q \Lambda(\partial^i \theta_i(x)) d\mu(x)$$

for $i = 1, \dots, n$. We therefore obtain

$$\int_Q \sum_{i=1}^n \Lambda(\theta_i(x)) d\mu(x) \leq \frac{40}{9} R^2 \int_Q [\log g(F(x))] d\mu(x) = \frac{40}{9} R^2 \int_Q [\log g(y)] d\nu(y).$$

It remains to note that always $|\theta_i(x)| = |F_i(x) - x_i| \leq 1$ since Q is a unit cube. Consequently $\Lambda(\theta_i(x)) = |\theta_i(x)|^2$ and hence,

$$\int_Q |F(x) - x|^2 d\mu(x) = \int_Q |\theta(x)|^2 d\mu(x) \leq \frac{40}{9} R^2 \int_Q [\log g(y)] d\nu(y) = \frac{40}{9} R^2 \cdot D(\nu || \mu).$$

This finishes the sketch of the alternative proof of Theorem 4.2. \square

References

- [1] Bakry, D., Émery, M., *Diffusions hypercontractives*. Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., Vol. 1123, Springer, Berlin, (1985), 177–206.
- [2] Barthe, F., *Inégalités de Brascamp-Lieb et convexité*. C. R. Acad. Sci. Paris Sér. I Math., Vol. 324., No. 8, (1997), 885–888.
- [3] Barthe, F., *On a reverse form of the Brascamp-Lieb inequality*. Invent. Math., Vol. 134, No. 2, (1998), 335–361.
- [4] Brenier, Y., *Polar factorization and monotone rearrangement of vector-valued functions*. Comm. Pure Appl. Math., Vol. 44, No. 4, (1991), 375–417.
- [5] Cheeger, J., *A lower bound for the smallest eigenvalue of the Laplacian*. Problems in analysis (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, NJ, (1970), 195–199.
- [6] Cordero-Erausquin, D., *Some applications of mass transport to Gaussian-type inequalities*. Arch. Ration. Mech. Anal., Vol. 161, No. 3, (2002), 257–269.
- [7] Cordero-Erausquin, D., *Sur le transport de mesures périodiques*. C. R. Acad. Sci. Paris Sér. I Math., Vol. 329, No. 3, (1999), 199–202.
- [8] Dembo, A., Zeitouni, O., *Transportation approach to some concentration inequalities in product spaces*. Electron. Comm. Probab., Vol. 1, No. 9, (1996), 83–90.
- [9] Eldan, R., Klartag, B., *Dimensionality and the stability of the Brunn-Minkowski inequality*. To appear in Ann. Sc. Norm. Super. Pisa. Available under <http://arxiv.org/abs/1110.6584>
- [10] Gozlan, N., Léonard, C., *Transport inequalities. A survey*. Markov Processes and Related Fields, Vol. 16, (2010), 635–736.
- [11] Henstock, R., Macbeath, A. M., *On the measure of sum-sets. I. The theorems of Brunn, Minkowski, and Lusternik*. Proc. London Math. Soc. (3), Vol. 3, (1953), 182–194.
- [12] Klartag, B., *Marginals of geometric inequalities*. Geometric Aspects of Functional Analysis, Lecture Notes in Math., Vol. 1910, Springer (2007), 133–166.
- [13] Klartag, B., *A Berry-Esseen type inequality for convex bodies with an unconditional basis*. Probab. Theory Related Fields, Vol. 145, No. 1-2, (2009), 1–33.
- [14] Knothe, H., *Contributions to the theory of convex bodies*. Michigan Math. J., Vol. 4, (1957), 39–52.
- [15] Ledoux, M., *On Talagrand’s deviation inequalities for product measures*. ESAIM: Probab. Statist., Vol. 1, (1997), 63–87.
- [16] Leoni, G., *A first course in Sobolev spaces*. Graduate Studies in Mathematics, 105. Amer. Math. Soc., Providence, RI, 2009.
- [17] Marton, K., *A measure concentration inequality for contracting Markov chains*. Geom. Funct. Anal., Vol. 6, No. 3, (1996), 556–571.
- [18] McCann, R. J., *A convexity principle for interacting gases*. Adv. Math., Vol. 128, No. 1, (1997), 153–179.
- [19] Milman, E., *Isoperimetric and concentration inequalities - equivalence under curvature lower bound*. Duke Math. J., Vol. 154, No. 2, (2010), 207–239.
- [20] Milman, V. D., Schechtman, G., *Asymptotic theory of finite-dimensional normed spaces*. Lecture Notes in Math., Vol. 1200, Springer, Berlin, 1986.

- [21] Otto, F., Villani, C., *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*. J. Funct. Anal., Vol. 173, No. 2, (2000), 361–400.
- [22] Royer, G., *An initiation to logarithmic Sobolev inequalities*. SMF/AMS texts and monographs, 14. American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2007.
- [23] Talagrand, M., *Transportation cost for Gaussian and other product measures*. Geom. Funct. Anal., Vol. 6, No. 3, (1996), 587–600.
- [24] Talagrand, M., *Concentration of measure and isoperimetric inequalities in product spaces*. Inst. Hautes Études Sci. Publ. Math., No. 81, (1995), 73–205.
- [25] Villani, C., *Topics in optimal transportation*. Graduate Studies in Mathematics, 58. Amer. Math. Soc., Providence, RI, 2003.