TRANSPORT PROOFS OF WEIGHTED POINCARÉ INEQUALITIES FOR LOG-CONCAVE DISTRIBUTIONS

DARIO CORDERO-ERAUSQUIN AND NATHAEL GOZLAN

Abstract. We prove, using optimal transport tools, weighted Poincaré inequalities for log-concave random vectors satisfying some centering conditions. We recover by this way similar results by Klartag and Barthe-Cordero-Erausquin for log-concave random vectors with symmetries. In addition, we prove that the variance conjecture is true for increments of log-concave martingales.

1. Introduction

In all the paper, if $X = (X_1, \ldots, X_n)$ is a random vector defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$ is an Borel (bounded or nonnegative) function, we use the following notation for the conditional expectations:

$$\mathbb{E}_i[h(X)] := \mathbb{E}[h(X)|X_1, \ldots, X_i],$$

with the convention that $\mathbb{E}_0[h(X)] = \mathbb{E}[h(X)]$. To any random vector X, we associate the random vector $\overline{X}$ defined as follows:

$$\overline{X}_i = X_i - \mathbb{E}_{i-1}[X_i], \qquad \forall i \in \{1, \ldots, n\}.$$

This recentering procedure will play an important role in all the paper (see also [9] for another application). We aim at proving Poincaré and transport inequalities for $\overline{X}$, when X is log-concave.

Recall that a random vector X with values in $\mathbb{R}^n$ is log-concave if for all non-empty compact sets $A, B \subset \mathbb{R}^n$, it holds

$$\mathbb{P}(X \in (1-t)A + tB) \geq \mathbb{P}(X \in A)^{1-t}\, \mathbb{P}(X \in B)^t, \qquad \forall t \in [0,1].$$

According to a celebrated result of Borell [11, 12], a random vector X is log-concave if and only if there is an affine map $\ell : \mathbb{R}^k \to \mathbb{R}^n$, $k \leq n$ and a random vector Y taking values in $\mathbb{R}^k$ such that $X = \ell(Y)$ and Y has a density of the form e^{-V} with respect to the Lebesgue measure on $\mathbb{R}^k$, where $V : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ is a convex function. In what follows, by an "n-dimensional log-concave random vector", we will understand a vector X satisfying the conditions above with $k = n$ (and $\ell = \mathrm{Id}$).

The main result of this note is that the class of all random vectors $\overline{X}$ with X n-dimensional and log-concave satisfies a general weighted Poincaré inequality.

Theorem 1.1. *There exists a universal numerical constant $a > 0$ such that for any n and any n-dimensional log-concave random vector X, it holds*

$$\mathrm{Var}(f(\overline{X})) \leq a \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}[\overline{X}_i^2 | \overline{X}_1, \ldots, \overline{X}_{i-1}]\partial_i f(\overline{X})^2\right] \tag{1.2}$$

Date: June 18, 2015.

1991 *Mathematics Subject Classification.* 60E15, 32F32 and 26D10.

Key words and phrases. Transport inequalities, weighted Poincaré inequalities, log-concave measures.

This work was partially supported by the grants ANR 2011 BS01 007 01 (GeMeCOD) and ANR 10 LABX 58.

for all locally-Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ *belonging to* $\mathbb{L}_2(\overline{X})$*, where* $\mathrm{Var}(Y) := \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$ *denotes the variance of a real valued random variable* Y*. In particular, if* X *is such that* $\mathbb{E}_{i-1}[X_i] = 0$ *for all* $i \in \{1, \ldots, n\}$*, then* $\overline{X} = X$ *and it holds*

$$\mathrm{Var}(f(X)) \leq a \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}_{i-1}[X_i^2] \partial_i f(X)^2\right] \tag{1.3}$$

Remark 1.4. *If the operation* $X \mapsto \overline{X}$ *was preserving log-concavity then, of course, (1.2) would follow from (1.3) applied to* $\overline{X}$*. It is not difficult to find examples of log-concave random vectors* X *such that* $\overline{X}$ *is not log-concave anymore. A random vector* X *such that* $\overline{X} = X$ *can be interpreted as a sequence of martingale increments (see Section 3 for more details).*

Theorem 1.1 is reminiscent of recent results of Klartag [25] and of Barthe and Cordero-Erausquin [4] which were based on $\mathbb{L}_2$ methods. The objective of this note is to give alternative proofs of variants of some of the results from [25, 4] using mass transport arguments.

Recall that a random vector X is *unconditional* when $X = (X_1, \ldots, X_n)$ has the same law as $(\varepsilon_1 X_1, \ldots, \varepsilon_n X_n)$ for any choice of $\varepsilon_i = \pm 1$. Since unconditional random vectors satisfy $\mathbb{E}_{i-1}[X_i] = 0$ for all $1 \leq i \leq n$, Theorem 1.1 can be seen as an extension of the following result by Klartag [25]: for any log-concave and unconditional random vector X, it holds

$$\mathrm{Var}(f(X)) \leq c \sum_{i=1}^{n} \mathbb{E}\left[(X_i^2 + \mathbb{E}[X_i^2]) \partial_i f(X)^2\right], \tag{1.5}$$

for all $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough, where $c > 0$ is some absolute constant. Moreover, when f is itself unconditional (i.e $f(\varepsilon_1 x_1, \ldots, \varepsilon_n x_n) = f(x_1, \ldots, x_n)$ for all $\varepsilon_i = \pm 1$), then the terms $\mathbb{E}[X_i^2]$ can be removed from the right hand side of (1.5). Note that in [25], Klartag also obtains weighted Poincaré inequalities for a larger class of unconditional distributions with a density of the form $e^{-\phi}$ with $\phi : \mathbb{R}^n \to \mathbb{R}$ whose restriction to $\mathbb{R}_+^n$ is p convex (i.e $x \mapsto \phi(x_1^{1/p}, \ldots, x_n^{1/p})$ is convex). Inequalities of the form (1.5) were also investigated in details in the recent paper [4]. There, the authors establish general weighted Poincaré inequalities for classes of probability measures invariant by a subgroup of isometries, not only the coordinate reflections.

Note that (1.3) applies to random vectors having less symmetries than unconditional random vectors. For instance, if the X_i are independent mean zero and variance one log-concave random variables then $\mathbb{E}_{i-1}[X_i] = 0$ for all i, whereas X does not have any particular symmetry. In this case, the conclusion (1.3) of Theorem 1.1 is consistent with the Poincaré inequality obtained using the (elementary) tensorisation property of the Poincaré inequality.

Theorem 1.1 also easily implies some variance estimates for log-concave random vectors.

Corollary 1.6. *There exists a universal constant* $b > 0$ *such that if* X *is an* n*-dimensional log-concave random vector, then, denoting by* $|\cdot|$ *the standard Euclidean norm on* $\mathbb{R}^n$*, it holds*

$$\mathrm{Var}\left(|\overline{X}|^2\right) \leq b \sum_{i=1}^{n} \mathbb{E}[\overline{X}_i^4] \leq 16b \sum_{i=1}^{n} \mathbb{E}[X_i^4]. \tag{1.7}$$

In particular, when $\mathbb{E}[X_i^2] = 1$ *for all* $i \in \{1, \ldots, n\}$*, we have*

$$\mathrm{Var}\left(|\overline{X}|^2\right) \leq cn \tag{1.8}$$

and if in addition X *satisfies* $\mathbb{E}_{i-1}[X_i] = 0$ *for all* i*, then*

$$\mathrm{Var}\left(|X|^2\right) \leq cn, \tag{1.9}$$

for some other universal constant c.

The inequality (1.9) on the variance immediately yields to the following concentration in a thin shell estimate

$$\mathbb{P}\left(\left||X|-\sqrt{n}\right| \geq t\sqrt{n}\right) \leq b\, e^{-cn^{1/4}\sqrt{t}}, \qquad \forall t > 0.$$

This type of concentration inequalities plays a central role in the proof of the central limit theorem for log-concave random vectors [2, 21, 23, 7].

Corollary 1.6 is also motivated by the so called *variance conjecture*. Recall that a random vector X is said *isotropic* if $\mathbb{E}[X] = 0$ and $\mathbb{E}[X_iX_j] = \delta_{i,j}$ for all $i, j \in \{1, \ldots, n\}$. The variance conjecture asserts that any log-concave and isotropic random vector X satisfies (1.9) for some universal positive constant b. This conjecture was shown to be true in restriction to the class of unconditional log-concave random vectors by Klartag [22, 25]. We refer to [4] and [1] for other subclasses of log-concave distributions satisfying the variance conjecture. The best (dimensional) estimate in date is due to Guédon and Milman [18] who proved that $\mathrm{Var}(|X|) \leq bn^{2/3}$ for any isotropic log-concave random vector X. The variance conjecture is a weak form of a celebrated conjecture by Kannan, Lovasz and Simonovits [20] stating that any log-concave and isotropic random vector X satisfies a Poincaré inequality

$$\mathrm{Var}(f(X)) \leq a\mathbb{E}\left[|\nabla f|^2(X)\right], \qquad \forall f \text{ smooth enough,}$$

for some universal constant $a > 0$. According to a remarkable recent result of Eldan [14], the variance conjecture implies the KLS conjecture up to some $\log(n)$ factor.

Corollary 1.6 thus shows that the variance conjecture is satisfied on the class of isotropic log-concave random vectors such that $\overline{X} = X$ (see also [4, Theorem 4] and Remark 6.1 below for a related result). It is not difficult to see that this class is strictly larger than the class of unconditional isotropic and log-concave random vectors (some informations on log-concave random vectors such that $\overline{X} = X$ can be found in Proposition 2.3 and Remark 2.4 below). For general log-concave random vectors X, let us mention that it is always at least possible to bound $\mathrm{Var}(|X|^2)$ in terms of $\mathrm{Var}(|\overline{X}|^2)$ and of $\mathrm{Var}(|X'|^2)$, where the "reduced" random vector X' is defined by

$$X'_i = \mathbb{E}_{i-1}[X_i], \qquad \forall i \in \{1, \ldots, n\}.$$

The basic observation behind the following elementary result is that $X = \overline{X} + X'$ is an orthogonal decomposition of X in the space $H := \mathbb{L}_2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$ of square integrable n-dimensional random vectors. More precisely, for any $X \in H$, the vector $\overline{X}$ is the orthogonal projection of X onto the linear subspace $H_0(X) = \{Y \in H; \mathbb{E}[Y_i|X_1, \ldots, X_{i-1}] = 0,\ \forall i \in \{1, \ldots, n\}\}$ (the space of random sequences that are martingale increments with respect to the filtration $\sigma(X_1, \ldots, X_i)$, $0 \leq i \leq n-1$.) We will prove the following useful identity

$$\begin{aligned}(1.10)\quad \mathrm{Var}(|X|^2) = \mathrm{Var}(|\overline{X}|^2) + \mathrm{Var}(|X'|^2) + 2\mathrm{Cov}(|\overline{X}|^2, |X'|^2)\\ + 4\mathbb{E}[(\overline{X} \cdot X')^2] + 4\mathbb{E}[|\overline{X}|^2\overline{X} \cdot X'] + 4\mathbb{E}[|X'|^2\overline{X} \cdot X'],\end{aligned}$$

from which one deduces the following result:

Corollary 1.11. *If X is an isotropic and log-concave random vector in $\mathbb{R}^n$, and X' is defined as above, then*

$$\mathrm{Var}\left(|X|^2\right) \leq a\left(n + \mathrm{Var}\left(|X'|^2\right)\right) \qquad \text{and} \qquad \mathrm{Var}\left(|X'|^2\right) \leq a\left(n + \mathrm{Var}\left(|X|^2\right)\right),$$

for some universal constant a.

It follows that the variance conjecture is (technically) equivalent to the existence of a universal constant $b > 0$ such that for any isotropic log-concave random vector X,

$$\mathrm{Var}\left(|X'|^2\right) \leq bn.$$

It would be of some interest to see if for some specific classes of vectors X, the variance term $\mathrm{Var}(|X'|^2)$ can be estimated by some power of n.

The proof of Theorem 1.1 is based on mass transport. More precisely, we will establish a transport-entropy inequality (Theorem 4.6) which is of independent interest, of the form

$$\mathcal{T}_\mu(\overline{\mu}, \overline{\nu}) \leq D(\nu \,\|\, \mu), \qquad \forall \nu,$$

where $\overline{\mu}$ and $\overline{\nu}$ are the laws of random vectors $\overline{X}$ and $\overline{Y}$, with X, Y distributed according to μ and ν. The optimal transport cost $\mathcal{T}_\mu$ will be of the form

$$\mathcal{T}_\mu(\nu_0, \nu_1) = \inf_{\pi \in C(\nu_0, \nu_1)} \iint c_\mu(x, y)\, \pi(dxdy),$$

for a particular cost function c_μ (precise definitions will be given later). Then, Theorem 1.1 will follow from this transport-entropy inequality by a standard linearization procedure. The argument towards our transport inequality will use an *above tangent lemma* introduced by Cordero-Erausquin [13] which is a handy tool to prove classical functional inequalities (Log-Sobolev, Talagrand) for uniformly log-concave random vectors and to recover the celebrated HWI inequality of Otto and Villani [30].

Let us mention here a byproduct of this approach in terms of transport inequalities involving the classical W_2 distance (definitions are recalled below).

Theorem 1.12. *There exists a universal constant c such that for any n dimensional log-concave random vector X taking values in the hypercube $[-R, R]^n$, $R > 0$, it holds*

$$W_2^2(\overline{\mu}, \overline{\nu}) \leq cR^2 D(\nu \,\|\, \mu),$$

for all probability measures ν on $\mathbb{R}^n$, where $\overline{\mu}$ and $\overline{\nu}$ denote respectively the laws of $\overline{X}$ and $\overline{Y}$, Y being distributed according to ν.

Theorem 1.12 can be considered as a variant of results by Eldan and Klartag [15, Theorem 6.1] and by Klartag [24, Theorem 4.2]. Let us mention that the present paper uses techniques of proof very similar to those involved in [15, 24]. To be more precise, Theorem 6.1 of [15] gives a similar inequality when μ and ν are both unconditional and log-concave. In their statement, the relative entropy is replaced by $\int_{[-R,R]^n} H(f, g) - 1$, where $H(f, g)(y) = \sup_{x \in \mathbb{R}^n} \sqrt{f(x + y)g(x - y)}$, denoting by f, g the densities of μ and ν with respect to Lebesgue. This quantity is relevant in their study of the stability of the Brunn-Minkowski inequality. In Theorem 4.2 of [24], Klartag obtains the inequality

$$W_2^2(\mu, \nu) \leq cL^2 D(\nu \,\|\, \mu), \qquad \forall \nu$$

for all log-concave probability measures μ supported on the hypercube $Q = [-1, 1]^n$ and such that in addition the density f of μ with respect to Lebesgue satisfies for some $L \geq 1$

$$f((1 - t)x + ty) \leq L[(1 - t)f(x) + tf(y)], \qquad \forall t \in [0, 1],$$

for all $x, y \in Q$ with $x - y$ proportional to one of the standard basis vectors e_i. This condition is for instance realized with $L = e^{M/8}$ if $f = e^{-V}$ for some smooth convex function $V : Q \to \mathbb{R}$ such that $\sup_{i \leq n} \sup_{x \in Q} \partial_i^2 V(x) \leq M$ for some $M \geq 0$.

The paper is organized as follows. In Section 2, we gather various observations on the relations between $\overline{X}$ and X for log-concave random vectors. In Section 3, we give some background on the mass transportation tools that are used to establish our general transport-inequality, which is stated and proved in Section 4, together with Theorem 1.12. Then, in Section 5 we linearize this transport-entropy inequality and establish Theorem

1.1. In the final Section 6, we explain how to derive the Corollaries 1.6 and 1.11 on the variance.

Acknowledgment. We warmly thank the referee for his/her careful reading of the manuscript that substantially improved the presentation of the paper.

2. Some observations about log-concave random vectors such that $\overline{X} = X$

First, we begin with a straightforward proposition identifying the class of random vectors such that $\overline{X} = X$ as increments of martingales.

Proposition 2.1. *A random vector $X = (X_1, \ldots, X_n)$ is such that $\overline{X} = X$ if and only if $M = (M_0, M_1, \ldots, M_n)$, with $M_0 = 0$ and $M_k = \sum_{i=1}^k X_i$ is a martingale with respect to the increasing sequence of sub-sigma fields $\mathcal{F}_k = \sigma(M_0, \ldots, M_k)$, $k \in \{0, \ldots, n\}$.*

The proof is left to the reader.

If $M := (M_0, M_1, \ldots, M_n)$ is a martingale, we denote by $\Delta_i = M_i - M_{i-1}$, $i \in \{1, \ldots, n\}$ the increments of M. The quadratic variation process of M is then defined by $[M]_k = \sum_{i=0}^k \Delta_i^2$, for all $k \in \{0, 1, \ldots, n\}$. With these definitions, Corollary 1.7 can be restated as follows.

Proposition 2.2. *There exists a universal constant $c > 0$ such that for all martingale $M = (M_0, M_1, \ldots, M_n)$ such that $M_0 = 0$ and $(M_1, \ldots, M_n)$ has a log-concave density, it holds*

$$\mathrm{Var}([M]_k) \leq c \sum_{i=1}^k \mathbb{E}[\Delta_i^4], \qquad \forall k \leq n.$$

Proof. Since the class of log-concave random vectors is stable under affine transformations, it follows that $(M_1, \ldots, M_n)$ has a log-concave density if and only if $(X_1, \ldots, X_n)$ with $X_i = \Delta_i$ has a log-concave density. The result then follows immediately from Corollary 1.7. □

We now collect in the following proposition some elementary informations on log-concave random vectors X such that $\overline{X} = X$.

Proposition 2.3.

(1) *If X is a log-concave random vector with values in $\mathbb{R}^n$ then $\overline{X} = X$ if and only if $\mathbb{E}[X] = 0$ and for all $k \in \{1, \ldots, n-1\}$, $\mathbb{E}[X|X_1, \ldots, X_k] = (X_1, \ldots, X_k, 0, \ldots, 0)$. In particular, if $C \subset \mathbb{R}^n$ is a convex body and X is uniformly distributed over C, then $\overline{X} = X$ if and only if the barycenter of C is at 0 and for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$*

$$\mathrm{Bar}(C \cap \{(x_1, \ldots, x_k)\} \times \mathbb{R}^{n-i}) = (x_1, \ldots, x_k, 0 \ldots, 0), \qquad \forall k \in \{1, \ldots, n-1\},$$

(whenever this section is not empty). In particular, C is symmetric with respect to the hyperplane $\{x_n = 0\}$.

(2) *Let $C \subset \mathbb{R}^n$ be a convex body satisfying the conditions above. Consider the convex body $D \subset \mathbb{R}^{n+1}$ defined by*

$$D = \{(t, \varphi(t)x) : t \in [a, b], x \in C\},$$

where $\varphi : [a, b] \to [0, \infty)$ is some concave function and let $D' = D - \mathrm{Bar}(D)$. Then a vector X uniformly distributed on D' satisfies $\overline{X} = X$. A typical example is given by a cone of the form

$$D' = D - \mathrm{Bar}(D), \quad \textit{with} \quad D = \mathrm{convex-hull}\Big(0, \{a\} \times C\Big) \subset \mathbb{R}^{n+1}, \ a \in \mathbb{R}.$$

(3) If $C \subset \mathbb{R}^2$ is a bounded convex body with barycenter at 0 and X is uniformly distributed over C, then $\overline{X}$ is uniformly distributed over the convex body $\overline{C}$ obtained from C by applying Steiner symmetrization with respect to the axis $D = \mathbb{R} \times \{0\}$. In particular $\overline{X} = X$ if and only if C is symmetric with respect to D.

(4) If X, Y are two independent log-concave random vectors (defined on the same probability space) such that $\overline{X} = X$ and $\overline{Y} = Y$, then $\overline{X+Y} = X + Y$.

Proof. (1) The second point follows easily from the fact that for $k \leq i-1$,

$$\mathbb{E}[X_i|X_1,\ldots,X_k] = \mathbb{E}\left[\mathbb{E}[X_i|X_1,\ldots,X_{i-1}]|X_1,\ldots,X_k\right].$$

(2) It is not difficult to check that $\mathrm{Bar}(D) = (m, 0, \ldots, 0)$ for some $m \in \mathbb{R}$ and that

$$\begin{aligned}\mathrm{Bar}(D \cap \{(t, x_1, \ldots x_k)\} \times \mathbb{R}^{n-k}) &= \left(t, \mathrm{Bar}\left(\left(\varphi(t)^{-1}C\right) \cap \{(x_1, \ldots, x_k)\} \times \mathbb{R}^{n-k}\right)\right) \\ &= (t, x_1, \ldots, x_k, 0, \ldots, 0),\end{aligned}$$

where the second line comes from the assumption made on C. This easily implies that D' enjoys the centering conditions of point (2) and so that X satisfies $\overline{X} = X$.
(3) Observe that C can be written as $C = \{(x_1, x_2) \in \mathbb{R}^2; x_1 \in [\alpha, \beta], a(x_1) \leq x_2 \leq b(x_1)\}$, for some $\alpha < \beta$, and where $a : [\alpha, \beta] \to \mathbb{R}$ is a concave function and $b : [\alpha, \beta] \to \mathbb{R}$ is a convex function. Recall that the Steiner symmetrization of C with respect to D is the set $\overline{C}$ defined by

$$\overline{C} = \left\{(x_1, x_2) \in \mathbb{R}^2; x_1 \in [\alpha, \beta], \frac{1}{2}(a(x_1) - b(x_1)) \leq x_2 \leq \frac{1}{2}(b(x_1) - a(x_1))\right\}.$$

Since the function $a-b$ is convex, the set $\overline{C}$ is convex. Moreover $\overline{X} = \left(X_1, X_2 - \frac{1}{2}(a(X_1) + b(X_1))\right)$ and so for all bounded measurable test function $f : \mathbb{R}^2 \to \mathbb{R}$

$$\begin{aligned}\mathbb{E}\left[f(\overline{X})\right] &= \frac{1}{\mathrm{Vol}(C)} \int_\alpha^\beta \int_{a(x_1)}^{b(x_1)} f(x_1, x_2 - \frac{1}{2}(a(x_1) + b(x_1)))\, dx_2 dx_1 \\ &= \frac{1}{\mathrm{Vol}(C)} \int_\alpha^\beta \int_{\frac{1}{2}(a(x_1)-b(x_1))}^{\frac{1}{2}(b(x_1)-a(x_1))} f(x_1, y_2)\, dy_2 dx_1 \\ &= \frac{1}{\mathrm{Vol}(C)} \int_{\overline{C}} f(y_1, y_2)\, dy_1 dy_2.\end{aligned}$$

This shows that $\overline{X}$ is uniformly distributed on $\overline{C}$.
(4) It is well known that $X+Y$ is still log-concave. Let us show that $\overline{X+Y} = X+Y$. Let $i \in \{2, \ldots, n\}$ and take $f : \mathbb{R}^{i-1} \to \mathbb{R}$ a bounded measurable test function, then it holds

$$\mathbb{E}\left[X_i f(X_1 + Y_1, \ldots, X_{i-1} + Y_{i-1})\right] = \mathbb{E}_X[X_i \mathbb{E}_Y[f(X_1 + Y_1, \ldots, X_{i-1} + Y_{i-1})]] = 0.$$

Similarly, $\mathbb{E}\left[Y_i f(X_1 + Y_1, \ldots, X_{i-1} + Y_{i-1})\right] = 0$. Therefore, $\mathbb{E}\left[(X_i + Y_i) f(X_1 + Y_1, \ldots, X_{i-1} + Y_{i-1})\right] = 0$, and since this holds for all test function f, one concludes that $\mathbb{E}_{i-1}[(X+Y)_i] = 0$ for all i and so $\overline{X+Y} = X+Y$.

□

Remark 2.4. *As we already mentioned, the class of log-concave random vectors such that $\overline{X} = X$ already contains unconditional log-concave random vectors and log-concave random vectors with centered independent components. Using the properties above, it is possible to give other examples of log-concave random vectors such that $\overline{X} = X$ in arbitrary large dimension. Namely, observe that if X is a log-concave random vector taking values in $\mathbb{R}^k$ and such that $\overline{X} = X$, then it is easy to check that for all $i \in \{1, \ldots, k+1\}$, the random vector X^i defined by*

$$X^i = (X_1, \ldots, X_{i-1}, 0, X_i, \ldots, X_k) \in \mathbb{R}^{k+1}$$

is still log-concave and verifies $\overline{X^i} = X^i$. *Thanks to point (4) of Proposition 2.3, one thus sees that if* $X_1, \ldots, X_{k+1}$ *are independent log-concave random vectors with values in* $\mathbb{R}^k$ *and such that* $\overline{X_i} = X_i$, *then the random vector* $Y = X_1^1 + X_2^2 + \cdots + X_{k+1}^{k+1}$ *is still log-concave and satisfies* $\overline{Y} = Y$. *Initializing this construction with* $k = 2$ *with the help of point (3) of Proposition 2.3 and iterating the process gives rise to a large class of non-trivial examples of log-concave random vectors such that* $\overline{X} = X$. *Similarly, starting with a two dimensional convex body as in point (3) of Proposition 2.3 one can construct with the help of point (2) convex bodies of arbitrary large dimensions such that a random vector uniformly distributed on them satisfies* $\overline{X} = X$. *This class of convex bodies contains in particular the regular simplex in appropriate coordinates; namely, the regular simplex* $\Delta_n \subset \mathbb{R}^n$ *defined by the sequences of side-one simplexes as follows:* $\Delta_1 = C_1 = [-\frac{1}{2}, \frac{1}{2}] \subset \mathbb{R}$ *and*

$$\Delta_n = C_n - \mathrm{Bar}(C_n), \quad C_n = \mathrm{convex-hull}\left(0, \left\{\sqrt{\frac{n+1}{2n}}\right\} \times \Delta_{n-1}\right) \subset \mathbb{R}^n.$$

3. Some background on mass transport

The key lemma used in [13] is the so called *above tangent lemma* recalled below. In what follows, the relative entropy (also called the Kullback-Leibler distance) of ν with respect to μ is defined by

$$D(\nu \,\|\, \mu) = \int \log \frac{d\nu}{d\mu}\, d\nu, \tag{3.1}$$

if ν is absolutely continuous with respect to μ (otherwise, we set $D(\nu \,\|\, \mu) = \infty$).

Lemma 3.2 ([13])**.** *If* μ *is a probability measure on* $\mathbb{R}^n$ *absolutely continuous with respect to the Lebesgue measure with a density of the form* $\mu(dx) = e^{-V(x)}\, dx$ *where* $V : \mathbb{R}^n \to \mathbb{R}$ *is a function of class* $\mathcal{C}^2$ *such that* $\mathrm{Hess}\, V \geq \rho$, $\rho \in \mathbb{R}$, *then for all compactly supported probability measures* ν_0, ν_1 *absolutely continuous with respect to* μ, *it holds*

$$\begin{aligned} D(\nu_1 \,\|\, \mu) \geq D(\nu_0 \,\|\, \mu) &+ \int \left\langle \nabla \frac{d\nu_0}{d\mu}(x), Tx - x \right\rangle \mu(dx) + \frac{\rho}{2} \int |Tx - x|^2\, \nu_0(dx) \\ &+ \int \left(\mathrm{Tr}\,(DT_x - \mathrm{I}_n) - \log |DT_x|\right) \nu_0(dx), \end{aligned} \tag{3.3}$$

where $T : \mathbb{R}^n \to \mathbb{R}^n$ *pushes forward* ν_0 *onto* ν_1 *and defines a "suitable" change of variables, and where* $|DT_x| \geq 0$ *denotes the determinant of the Jacobian matrix* DT_x

By "suitable" we mean for instance a $\mathcal{C}^1$-diffeomorphism, but actually the inequality (3.3) remains true for less regular transport maps (in particular for Brenier's map which is only almost surely differentiable). We refer to [13] for a precise statement.

First let us recall the classical applications of (3.3). In [13], the inequality (3.3) was applied with the Brenier map T (see [33]), that is to say the ν_0 almost surely unique map T achieving the infimum in the definition of the square Kantorovich distance W_2:

$$\int |Tx - x|^2\, \nu_0(dx) = \inf_{\pi \in C(\nu_0, \nu_1)} \iint |y - x|^2\, \pi(dxdy) := W_2^2(\nu_0, \nu_1),$$

where $C(\nu_0, \nu_1)$ denotes the set of all couplings of ν_0, ν_1, (i.e probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ having ν_0 and ν_1 as marginals). A fundamental property of the Brenier map T is that it is the gradient of a convex function: there exists $\phi : \mathbb{R}^n \to \mathbb{R}$ convex such that $T(x) = \nabla \phi(x)$ for ν_0 almost every $x \in \mathbb{R}^n$. As a consequence of the inequality $\log(\lambda) \leq \lambda - 1$, $\lambda > 0$ and of the fact that $DT_x = \mathrm{Hess}_x\, \phi$ has a non-negative spectrum, the last term in (3.3) is always non-negative (assuming for simplicity that T is smooth).

So (3.3) becomes

$$D(\nu_1 \,\|\, \mu) \geq D(\nu_0 \,\|\, \mu) + \int \left\langle \nabla \frac{d\nu_0}{d\mu}(x), Tx - x \right\rangle \mu(dx) + \frac{\rho}{2} W_2^2(\nu_0, \nu_1). \tag{3.4}$$

Inequality (3.4), which expresses in some sense that the graph of the map $D(\,\cdot \,\|\, \mu)$ lies above its tangent, is also related to the notion of displacement-convexity of the relative entropy along W_2 geodesics (see [29, 33]). When $\rho > 0$, interesting consequences can be drawn from the inequality above. For instance, choosing $\nu_0 = \mu$ yields to the following transport-entropy inequality

$$W_2^2(\nu_1, \mu) \leq \frac{2}{\rho} D(\nu_1 \,\|\, \mu), \qquad \forall \nu_1.$$

This type of inequalities goes back to the works by Marton [28] and Talagrand [32] (see [27, 33, 16] for an introduction to the subject). On the other hand, choosing $\nu_1 = \mu$ it is not difficult to derive from (3.4) the logarithmic-Sobolev inequality (see [13, 3, 16] for details)

$$D(\nu_0 \,||\, \mu) \leq \frac{2}{\rho} \int \frac{|\nabla h_0|^2}{h_0}\, d\mu, \qquad \forall \nu_0 = h_0\, \mu.$$

We refer to [5, 9] for other applications and variants of (3.3) and (3.4).

In this paper, we will use (3.3) with $\rho = 0$ and $\nu_0 = \mu$:

$$D(\nu_1 \,\|\, \mu) \geq \int \left(\mathrm{Tr}\,(DT_x - \mathrm{I}_n) - \log |DT_x|\right) \mu(dx).$$

But as a main difference, we will rather use as T the Knothe map [26] between μ and ν_1.

Let us recall the definition of the Knothe transport between two probability measures. If μ, ν are two Borel probability on $\mathbb{R}$ and μ has no atom, then there exists a unique non-decreasing and left continuous map $T : \mathbb{R} \to [-\infty, \infty]$ transporting μ on ν in the sense that $\int f(T)\, d\mu = \int f\, d\nu$ for all say bounded continuous function f. This map T is given by

$$T(x) = F_\nu^{-1} \circ F_\mu(x), \qquad \forall x \in \mathbb{R}.$$

where, for $x \in \mathbb{R}$ and $t \in [0, 1]$,

$$F_\mu(x) = \mu(-\infty, x] \qquad \text{and} \qquad F_\mu^{-1}(t) = \inf\{x \in \mathbb{R}; F_\mu(x) \geq t\} \in [-\infty, \infty].$$

The map T takes finite values μ almost surely. Let us mention that the map T achieves the minimum value in a large class of optimal transportation problems (see for instance [31]). This fact will not be explicitly used in the sequel.

The Knothe transport map is a multidimensional extension of this one dimensional transport. To define it properly, we need to introduce the following notation. If μ is a probability measure on $\mathbb{R}^n$ and $X = (X_1, \ldots, X_n)$ is a random vector of law μ, we will denote by μ_i the law of $(X_1, \ldots, X_i)$. For $i \geq 2$ and $x_1, \ldots, x_{i-1} \in \mathbb{R}$, we denote by $\mu_i(\,\cdot\,|x_1, \ldots, x_{i-1})$ the conditional law of X_i knowing $X_1 = x_1, X_2 = x_2, \ldots, X_{i-1} = x_{i-1}$. The conditional probability measure $\mu_i(\,\cdot\,|x_1, \ldots, x_{i-1})$ is well defined for μ_{i-1} almost all $(x_1, \ldots, x_{i-1}) \in \mathbb{R}^{i-1}$. When μ has a positive density h with respect to the Lebesgue measure on $\mathbb{R}^n$, the conditional probability measures $\mu_i(\,\cdot\,|x_1, \ldots, x_{i-1})$ have an explicit density with respect to Lebesgue measure on $\mathbb{R}$ it holds

$$\int f(u_i)\, \mu_i(du_i|x_1, \ldots, x_{i-1}) = \frac{\int f(u_i) h(x_1, \ldots, x_{i-1}, u_i, u_{i+1}, \ldots, u_n)\, du_i \cdots du_n}{\int h(x_1, \ldots, x_{i-1}, u_i, u_{i+1}, \ldots, u_n)\, du_i \cdots du_n},$$

for all bounded continuous $f : \mathbb{R} \to \mathbb{R}$.

The Knothe map $T = (T_1, \ldots, T_n)$ transporting a probability measure μ on $\mathbb{R}^n$ with a positive density on another probability ν, is defined recursively as follows :

- T_1 is the optimal transport map sending μ_1 on ν_1;

- for a given $x \in \mathbb{R}^n$, $T_i(x_1, x_2, \dots, x_{i-1}, \cdot\,)$ is the one dimensional monotone map sending $\mu_i(\,\cdot\,|x_1, \dots, x_{i-1})$ on $\nu_i(\,\cdot\,|T_1(x), \dots, T_{i-1}(x))$.

Note that in particular, T is triangular in the sense that $T_i(x)$ depends only on $x_1, \dots, x_i$.

The following lemma is a formally contained in Lemma 3.2; for completeness, we recall its short proof below.

Lemma 3.5. *Let μ be probability measure on $\mathbb{R}^n$ with $\mu(dx) = e^{-V(x)}\,dx$ with $V : \mathbb{R}^n \to \mathbb{R}$ a convex function of class $\mathcal{C}^1$; for all probability measure ν on $\mathbb{R}^n$ compactly supported with a smooth density, it holds*

$$\begin{aligned} D(\nu \,\|\, \mu) &\geq \int \left[\mathrm{Tr}(DT(x) - I) - \log(|DT(x)|)\right] \mu(dx) \\ &= \int \sum_{i=1}^{n} \left[\partial_i T_i(x) - 1 - \log \partial_i T_i(x)\right] \mu(dx) \end{aligned}$$

where T is the Knothe map transporting μ on ν and $|DT_x| \geq 0$ is the determinant of the Jacobian matrix DT_x.

Proof. Write $g = \frac{d\nu}{dx}$ and $h = \frac{d\nu}{d\mu}$. First assume that T is $\mathcal{C}^1$; according to the change of variable formula, it holds

$$e^{-V(x)} = h(Tx)e^{-V(Tx)}|DT(x)|,$$

so taking the log and integrating with respect to μ, we obtain

$$-\int V(x)\,\mu(dx) = \int \log(h(Tx))\,\mu(dx) - \int V(Tx)\,\mu(dx) + \int \log(|DT(x)|)\,\mu(dx).$$

So

$$D(\nu \,\|\, \mu) = \int V(Tx) - V(x)\,\mu(dx) - \int \log|DT(x)|\,\mu(dx).$$

By assumption,

$$V(y) \geq V(x) + \nabla V(x) \cdot (y - x), \qquad \forall x, y \in \mathbb{R}^n.$$

So,

$$D(\nu \,\|\, \mu) \geq \int \nabla V(x) \cdot (Tx - x)\,\mu(dx) - \int \log|DT(x)|\,\mu(dx).$$

Note that, integrating by parts (and using that ν is compactly supported),

$$\int \nabla V(x) \cdot (Tx - x)\,\mu(dx) = \int \sum_{i=1}^{n} (\partial_i T_i(x) - 1)e^{-V(x)}\,dx = \int \mathrm{Tr}(DT(x) - I)\,\mu(dx)$$

Thus,

$$D(\nu \,\|\, \mu) \geq \int \left[\mathrm{Tr}(DT(x) - I) - \log|DT(x)|\right] \mu(dx). \tag{3.6}$$

Actually the map T is not necessarily of class $\mathcal{C}^1$ so the change of variable formula above needs to be justified. One can consult Section 3 of [8] and invoke for instance [8, Lemma 3.1]. □

4. A general transport inequality for log-concave probability measures

Before introducing our transport cost, we need to briefly discuss on the Cheeger constant (or equivalently, the Poincaré constant) of one-dimensional log-concave densities, a case where optimal bounds are known. If γ is a log-concave probability measure on $\mathbb{R}$, denote by λ_γ its Cheeger's constant, namely the largest constant for which

$$\lambda_\gamma \int |f - m(f)|\,d\gamma \leq \int |f'|\,d\gamma \tag{4.1}$$

holds for all $f : \mathbb{R} \to \mathbb{R}$ locally-Lipschitz, where $m(f)$ denotes a median of f. It was proven by Bobkov [6] that when γ is log-concave probability measure on $\mathbb{R}$, one has

$$\frac{1}{3\mathrm{Var}(X)} \leq \lambda_\gamma^2 \leq \frac{2}{\mathrm{Var}(X)}, \tag{4.2}$$

with $X \sim \gamma$. Note that if X is a constant random variable, $\mathrm{Var}(X) = 0$ and $\lambda = \infty$.

In what follows, μ is a log-concave probability measure on $\mathbb{R}^n$ with full support and $X = (X_1, \ldots, X_n)$ a random vector distributed according to μ.

According to Bobkov's estimate (4.2), for all $x \in \mathbb{R}^n$, the one dimensional (log-concave) probability $\mu_i(\,\cdot\,|x_1, \ldots, x_{i-1})$ verifies Cheeger's inequality (4.1) with a constant (optimal up to universal factor)

$$\lambda_i^2(x) = \lambda_i^2(x_1, \ldots, x_{i-1}) := \frac{1}{3\mathrm{Var}(X_i|X_1 = x_1, \ldots, X_{i-1} = x_{i-1})} \in (0, +\infty] \tag{4.3}$$

where

$$\begin{aligned}\mathrm{Var}(X_i|X_1 = x_1, \ldots, X_{i-1} = x_{i-1})& \\ = \int u^2 \mu_i(du|x_1, \ldots, x_{i-1}) &- \Big(\int u\, \mu_i(du|x_1, \ldots, x_{i-1})\Big)^2 \in [0, +\infty).\end{aligned}$$

In Theorem 4.6 below, we prove that any log-concave probability measure on $\mathbb{R}^n$ verifies some transport-entropy inequality with a cost function c_μ determined by the functions λ_i introduced above. In order to state the result, we need to introduce some additional notation. Recall that if Z is a random vector, we denote by $\overline{Z}$ the random vector defined by

$$\overline{Z}_i = Z_i - \mathbb{E}[Z_i|Z_1, \ldots, Z_{i-1}].$$

Note in particular that $\overline{X} = R(X)$, where the recentering map $R : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $R(x) = (R_1(x), \ldots, R_n(x))$, where

$$R_i(x) = x_i - m_i(x), \quad \text{with} \quad m_i(x) = m_i(x_1, x_2, \ldots, x_{i-1}) = \int u\, \mu_i(du|x_1, \ldots, x_{i-1}) \tag{4.4}$$

It is not difficult to check that the map R is invertible. We will denote by $S = R^{-1}$ its inverse. The cost function $c_\mu : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ is defined as follows,

$$c_\mu(x, y) = \frac{1}{16} \sum_{i=1}^n N\big(\lambda_i(S(x))(x_i - y_i)\big), \qquad \forall x, y \in \mathbb{R}^n,$$

where $N(t) = |t| - \log(1 + |t|)$ (with the conventions $0 \times \infty = 0$ and $a \times \infty =$ sign of a $\times \infty$ for $a \neq 0$). The associated optimal transport cost denoted by $\mathcal{T}_\mu$ is defined by

$$\mathcal{T}_\mu(\nu_1, \nu_2) = \inf_{\pi \in C(\nu_1, \nu_2)} \iint c_\mu(x_1, x_2)\, \pi(dx_1 dx_2),$$

where $C(\nu_1, \nu_2)$ is the set of all probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$\pi(dx_1 \times \mathbb{R}^n) = \nu_1(dx_1) \quad \text{and} \quad \pi(\mathbb{R}^n \times dx_2) = \nu_2(dx_2)$$

Let us mention that the transport inequality below also holds with the cost function $\tilde{c}_\mu : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined as follows

$$\tilde{c}_\mu(x, y) = \frac{1}{16} N\left(\sqrt{\sum_{i=1}^n \lambda_i(S(x))^2 (x_i - y_i)^2}\right), \qquad \forall x, y \in \mathbb{R}^n. \tag{4.5}$$

Indeed the function $x \mapsto N(\sqrt{x})$ is subadditive, since it is concave on $\mathbb{R}^+$ and vanishes at 0, so we have for all $x, y \in \mathbb{R}^n$, $c_\mu(x, y) \geq \tilde{c}_\mu(x, y)$.

Theorem 4.6. *Let X be an n-dimensional log-concave random vector and let μ be its law; for all probability measure ν on $\mathbb{R}^n$ with finite first moment, it holds*

$$\mathcal{T}_\mu(\bar{\mu}, \bar{\nu}) \leq D(\nu \,\|\, \mu), \tag{4.7}$$

where $\bar{\mu}$ is the law of $\overline{X}$ and $\bar{\nu}$ is the law of $\overline{Y}$ with Y distributed according to ν.

Note that the transport cost depends on μ, and not $\overline{\mu}$. Indeed, it is given by the values of λ_i, which depend on $X \sim \mu$ through formula (4.3).

Proof. According to a result by Bobkov and Houdré [10], if γ is probability measure on $\mathbb{R}$ verifying Cheeger's inequality (4.1) with constant λ, then for all convex even function $L : \mathbb{R} \to \mathbb{R}^+$ such that $L(0) = 0$, $L(x) > 0$ for all $x > 0$ and $p_L := \sup \frac{xL'(x)}{L(x)} < +\infty$, it holds

$$\int L(f - m(f))\, d\gamma \leq \int L(p_L f'/\lambda)\, d\gamma,$$

where $m(f)$ denotes the median of f. It will be convenient to replace the median of f by its mean denoted by $\gamma(f)$. First observe that Jensen inequality yields $L(\gamma(f) - m(f)) \leq \int L(p_L f'/\lambda)\, d\gamma$. On the other hand, the convexity of L implies that

$$\int L(f - \gamma(f))\, d\gamma \leq \frac{1}{2} \int L(2(f - m(f)))\, d\gamma + \frac{1}{2} L(2(m(f) - \gamma(f))).$$

Finally, it is not difficult to check that the function L^{1/p_L} is subbadditive (see for instance [17, Lemma 4.7]). It follows that $L(2a) \leq 2^{p_L} L(a)$, $a \geq 0$. Therefore,

$$\int L(f - \gamma(f))\, d\gamma \leq 2^{p_L} \int L(p_L f'/\lambda)\, d\gamma.$$

As it is easy to see, for the function N, it holds $p_N \leq 2$. So we have the inequality

$$\frac{1}{16} \int N\left(\lambda(f - \gamma(f))\right)\, d\gamma \leq \int N(f')\, d\gamma. \tag{4.8}$$

First let us assume that $\mu(dx) = e^{-V(x)}\, dx$ where $V : \mathbb{R}^n \to \mathbb{R}$ is a convex function of class $\mathcal{C}^1$ and ν is compactly supported with a smooth density. If X is a random vector of law μ, then using Lemma 3.5, the inequality $t - \log(1 + t) \geq N(t)$, $t > -1$ and the inequality (4.8), it holds

$$\begin{aligned}
D(\nu \,\|\, \mu) &\geq \sum_{i=1}^n \mathbb{E}\left[N(\partial_i T_i(X) - 1)\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\, \mathbb{E}\left[N(\partial_i T_i(X) - 1) \,|\, X_1, \ldots, X_{i-1}\right]\, \right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\, \mathbb{E}\left[N(\partial_i (T_i - x_i)(X)) \,|\, X_1, \ldots, X_{i-1}\right]\, \right] \\
&\geq \frac{1}{16} \sum_{i=1}^n \mathbb{E}\Big\{\, \mathbb{E}\big[\, N\left(\lambda_i(X)\left(T_i(X) - \mathbb{E}[T_i(X)|X_1, \ldots, X_{i-1}] - X_i + \mathbb{E}[X_i|X_1, \ldots, X_{i-1}]\right)\right) \\
&\qquad\qquad |X_1, \ldots, X_{i-1}\big]\Big\} \\
&= \frac{1}{16} \sum_{i=1}^n \mathbb{E}\left[\, N\left(\lambda_i(X)\left(T_i(X) - \mathbb{E}[T_i(X)|X_1, \ldots, X_{i-1}] - X_i + \mathbb{E}[X_i|X_1, \ldots, X_{i-1}]\right)\right)\right].
\end{aligned}$$

Note that, since $T_1(X), \ldots, T_{i-1}(X)$ are functions of $X_1, \ldots, X_{i-1}$, it holds

$$\mathbb{E}[T_i(X)|X_1, \ldots, X_{i-1}] = \mathbb{E}[T_i(X)|T_1(X), \ldots, T_{i-1}(X)]$$

almost surely. It follows, that the vector $\overline{Y}$ defined by $\overline{Y}_i = T_i(X) - \mathbb{E}[T_i(X)|X_1, \ldots, X_{i-1}]$ has law $\bar{\nu}$. Using the definition of our cost, we see that

$$D(\nu \,\|\, \mu) \geq \frac{1}{16}\mathbb{E}\left[\sum_{i=1}^{n} N\left(\lambda_i(S(\overline{X}))\left(\overline{Y}_i - \overline{X}_i\right)\right)\right] = \mathbb{E}\left[c_\mu(\overline{X}, \overline{Y})\right].$$

Therefore, by definition of $\mathcal{T}_\mu$, we have

$$D(\nu \,\|\, \mu) \geq \mathcal{T}_\mu(\bar{\mu}, \bar{\nu}).$$

Using classical approximation arguments, one extends the inequality above to all probability measures ν with finite first moment.

This completes the proof of Theorem 4.6 in the case where $\mu(dx) = e^{-V(x)}\,dx$ with a convex function V of class $\mathcal{C}^1$ on $\mathbb{R}^n$. The conclusion is then extended, using classical approximation arguments, to any $\mu(dx) = e^{-V(x)}\,dx$ where $V : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous convex function whose domain has a non empty interior. A way to do it is to consider the family of convex functions V_s, $s > 0$ defined by

$$V_s(x) = \inf_{y \in \mathbb{R}^n} \left\{ V(y) + \frac{1}{s}|x - y|^2 \right\}, \qquad x \in \mathbb{R}^n,\ s > 0.$$

It is well known that for all $s > 0$, $V_s : \mathbb{R}^n \to \mathbb{R}$ is a $\mathcal{C}^1$ smooth convex function on $\mathbb{R}^n$ converging monotonically to V as $s \to 0$ (see for instance [19, Theorem 4.1.4]). Details are left to the reader. $\square$

Proof of Theorem 1.12. Assume that μ is the law of an n-dimensional log-concave random vector X taking values in the hypercube $Q = [-R, R]^n$. This assumption on the support of μ has for consequence that for all $x \in Q$,

$$\mathrm{Var}(X_i | X_1 = x_1, \ldots, X_{i-1} = x_i) \leq 2R^2.$$

Therefore, $\lambda_i(x) \geq 1/(\sqrt{6}R)$ for all $i \in \{1, \ldots, n\}$ and $x \in Q$. It is not difficult to check that there is an absolute constant $c > 0$ such that $N(u) \geq cu^2$ for all $|u| \leq 2/\sqrt{6}$. So, if ν is a given probability measure on Q, then by Theorem 4.6 it holds

$$D(\nu \,\|\, \mu) \geq \frac{c}{R^2}\mathbb{E}\left[|\overline{X} - \overline{Y}|^2\right],$$

which completes the proof. $\square$

5. Weighted Poincaré inequalities for log-concave probability measures

In this last section, we use a classical linearization technique to prove that the transport cost inequality obtained in Theorem 4.6 implies the weighted Poincaré inequality of Theorem 1.1. Such linearization depends only on the behavior of the cost for small distances. It will be more convenient, notationaly speaking, but equivalent, to use the cost $\tilde{c}_\mu$ defined by (4.5) in the definition of $\mathcal{T}_\mu$ and in Theorem 4.6, rather than c_μ.

Let us introduce the following supremum convolution operator

$$P_t f(x) = \sup_{y \in \mathbb{R}^n} \left\{ f(y) - \frac{1}{t}\tilde{c}_\mu(x, y) \right\}, \qquad \forall x \in \mathbb{R}^n, \qquad \forall t > 0,$$

which is well defined for instance for bounded continuous function f on $\mathbb{R}^n$. It can be shown that the function $u(t, x) = P_t f(x)$ satisfies in some weak sense the following Hamilton-Jacobi equation

$$\partial_t u(t, x) = 8 \sum_{i=1}^{n} \frac{1}{\lambda_i^2(S(x))} \left(\partial_{x_i} u\right)^2 (t, x).$$

Actually, in what follows, we will only need the following elementary inequality:

Lemma 5.1. *For all differentiable bounded Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\limsup_{t\to 0} \frac{1}{t} \int P_t f - f \, d\nu \le 8 \int \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))} \left(\partial_{x_i} f\right)^2(x)\, \nu(dx),$$

for all probability measure ν on $\mathbb{R}^n$ such that $\int \lambda_i^{-2}(S)\, d\nu$ is finite for every $i \in \{1, \dots, n\}$.

Let us admit the lemma for a moment and prove Theorem 1.1.

Proof of Theorem 1.1. Let X be an n-dimensional log-concave random vector whose law is denoted by μ in the sequel and let $g : \mathbb{R}^n \to \mathbb{R}$ be a bounded function such that $\int g\, d\mu = 0$ and define for all $t \ge 0$ the measure $\nu^t(dx) = (1 + tg)\, \mu(dx)$. If t is small enough, then ν^t is a probability measure. Let π be a coupling of $\bar\mu$ and $\overline{\nu^t}$, and $a > 0$ be a parameter whose value will be fixed later on ; for all bounded differentiable Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$, it holds

$$\begin{aligned}\frac{1}{t}\left(\int f d\overline{\nu^t} - \int f\, d\bar\mu\right) &= \frac{1}{t} \int f(y) - f(x)\, \pi(dxdy)\\ &= \frac{1}{t}\int f(y) - P_{at} f(x)\, \pi(dxdy) + \frac{1}{t}\int P_{at} f(x) - f(x)\, \bar\mu(dx)\\ &\le \frac{1}{at^2}\int \tilde c_\mu(x,y)\, \pi(dxdy) + \frac{1}{t}\int P_{at} f(x) - f(x)\, \bar\mu(dx),\end{aligned}$$

where the last line comes from the inequality $f(y) - P_s f(x) \le \frac{1}{s}\tilde c_\mu(x, y)$ for all $s > 0$. So optimizing over $\pi \in C(\bar\mu, \overline{\nu^t})$, we get, for all t small enough,

$$\begin{aligned}\frac{1}{t}\left(\int f d\overline{\nu^t} - \int f\, d\bar\mu\right) &\le \frac{1}{at^2}\mathcal{T}_\mu(\bar\mu, \overline{\nu^t}) + \frac{1}{t}\int P_{at} f(x) - f(x)\, \bar\mu(dx)\\ &\le \frac{1}{at^2} D(\nu^t \,\|\, \mu) + \frac{1}{t}\int P_{at} f(x) - f(x)\, \bar\mu(dx),\end{aligned}$$

where the last inequality comes from Theorem 4.6. A straightforward calculation shows that $t^{-2} D(\nu^t \,\|\, \mu) \to \frac{1}{2}\int g^2\, d\mu$ when t goes to 0. Therefore, using Lemma 5.1, we get

$$\limsup_{t\to 0} \frac{1}{t}\left(\int f d\overline{\nu^t} - \int f\, d\bar\mu\right) \le \frac{1}{2a}\int g^2\, d\mu + 8a \int \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))}\left(\partial_{x_i} f\right)^2(x)\, \bar\mu(dx)$$

and optimizing over $a > 0$

$$\limsup_{t\to 0} \frac{1}{t}\left(\int f d\overline{\nu^t} - \int f\, d\bar\mu\right) \le 4\left(\int g^2\, d\mu\right)^{1/2}\left(\int \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))}\left(\partial_{x_i} f\right)^2(x)\, \bar\mu(dx)\right)^{1/2}$$

Now let us evaluate the left hand side. Consider the map R^t defined by

$$R^t(x) = \left[x_1 - \int u_1\, \nu_1^t(du), x_2 - \int u_2\, \nu_2^t(du_2|x_1), \dots, x_n - \int u_n\, \nu_n^t(du_n|x_1, \dots, x_{n-1})\right].$$

For $t = 0$, $R^0 = R$ is the map introduced in (4.4). Then $\overline{\nu^t}$ is the image of ν^t by the map R^t and $\bar\mu$ the image of μ by the map R. Therefore,

$$\begin{aligned}\frac{1}{t}\left(\int f d\overline{\nu^t} - \int f\, d\bar\mu\right) &= \frac{1}{t}\left(\int f(R^t)(1 + tg)\, d\mu - \int f(R)\, d\mu\right)\\ &\to -\int \nabla f(R)\cdot U\, d\mu + \int f(R) g\, d\mu,\end{aligned} \tag{5.2}$$

when t goes to 0, where $U := \lim_{t\to 0}\frac{1}{t}(R^t - R)$. Let us calculate U. Writing the definition, it is not difficult to see that,

$$\int u_i\, \nu_i^t(du_i|x_1, \dots, x_{i-1}) = \frac{a_i + tb_i}{c_i + td_i},$$

with

$$\begin{aligned}
a_i &= \int u_i e^{-V(x_1,\ldots,x_{i-1},u_i,\ldots,u_n)}\, du_i \cdots du_n,\\
b_i &= \int u_i g(x_1,\ldots,x_{i-1},u_i,\ldots,u_n) e^{-V(x_1,\ldots,x_{i-1},u_i,\ldots,u_n)}\, du_i \cdots du_n,\\
c_i &= \int e^{-V(x_1,\ldots,x_{i-1},u_i,\ldots,u_n)}\, du_i \cdots du_n,\\
d_i &= \int g(x_1,\ldots,x_{i-1},u_i,\ldots,u_n) e^{-V(x_1,\ldots,x_{i-1},u_i,\ldots,u_n)}\, du_i \cdots du_n.
\end{aligned}$$

Therefore,

$$\begin{aligned}
U_i(x) &= \lim_{t\to 0} \frac{1}{t}\left(\int u_i\, d\nu_i^t(du_i|x_1,\ldots,x_{i-1}) - \int u_i\, d\mu_i(du_i|x_1,\ldots,x_{i-1})\right)\\
&= \lim_{t\to 0} \frac{1}{t}\left(\frac{a_i+tb_i}{c_i+td_i} - \frac{a_i}{c_i}\right) = \frac{b_i}{c_i} - \frac{a_i}{c_i}\frac{d_i}{c_i}\\
&= \mathbb{E}[X_i g(X)|X_1 = x_1,\ldots,X_{i-1}=x_{i-1}]\\
&\quad - \mathbb{E}[X_i|X_1 = x_1,\ldots,X_{i-1}=x_{i-1}]\cdot \mathbb{E}[g(X)|X_1 = x_1,\ldots,X_{i-1}=x_{i-1}].
\end{aligned}$$

It is easy to check that $\frac{1}{t}\left|\frac{a_i+tb_i}{c_i+td_i} - \frac{a_i}{c_i}\right| \leq \frac{2M}{1-tM}\frac{|a_i|}{c_i}$ for t sufficiently small, where $M = \sup|g|$. This can be used to justify the limit in (5.2). Details are left to the reader.

According to what precedes,

$$U(X) = \mathbb{E}_{i-1}[X_i g(X)] - \mathbb{E}_{i-1}[X_i]\mathbb{E}_{i-1}[g(X)] = \mathbb{E}_{i-1}[\overline{X}_i g(X)].$$

So putting everything together, we get

$$\begin{aligned}
\mathbb{E}[f(\overline{X})g(X)] &\leq 4\mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \frac{1}{\lambda_i^2(S(\overline{X}))}\left(\partial_i f(\overline{X})\right)^2\right]^{1/2} + \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}_{i-1}[\overline{X}_i g(X)]\partial_i f(\overline{X})\right]\\
&= 4\sqrt{3}\mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^2|\overline{X}_1,\ldots,\overline{X}_{i-1}\right]\left(\partial_i f(\overline{X})\right)^2\right]^{1/2} + \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}_{i-1}\left[\overline{X}_i g(X)\right]\partial_i f(\overline{X})\right]
\end{aligned}$$

where the second line comes from the definition of the λ_i's and the identity

$$\begin{aligned}
\mathrm{Var}(X_i|X_1 = S_1(\bar{x}),\ldots,X_{i-1} = S_{i-1}(\bar{x})) &= \mathbb{E}[\overline{X}_i^2|X_1 = S_1(\bar{x}),\ldots,X_{i-1} = S_{i-1}(\bar{x})]\\
&= \mathbb{E}\left[\overline{X}_i^2|\overline{X}_1 = \bar{x}_1,\ldots,\overline{X}_{i-1} = \overline{x}_{i-1}\right],
\end{aligned}$$

for all $\bar{x} = (\bar{x}_1,\ldots,\overline{x}_n) \in \mathbb{R}^n$. Finally let us bound the last term. Using Cauchy-Schwarz, it holds

$$\begin{aligned}
\sum_{i=1}^n \mathbb{E}\left[\mathbb{E}_{i-1}[\overline{X}_i g(X)]\partial_i f(\overline{X})\right] &= \sum_{i=1}^n \mathbb{E}\left[\overline{X}_i g(X)\mathbb{E}_{i-1}[\partial_i f(\overline{X})]\right]\\
&\leq \mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\left(\sum_{i=1}^n \overline{X}_i\mathbb{E}_{i-1}[\partial_i f(\overline{X})]\right)^2\right]^{1/2}.
\end{aligned}$$

Now observe that if $i \leq j-1$, then, since $\overline{X}_i\mathbb{E}_{i-1}[\partial_i f(\overline{X})]\mathbb{E}_{j-1}[\partial_i f(\overline{X})]$ is a function of $X_1,\ldots,X_{j-1}$, it holds

$$\mathbb{E}\left[\overline{X}_i\mathbb{E}_{i-1}[\partial_i f(\overline{X})]\cdot \overline{X}_j\mathbb{E}_{j-1}[\partial_i f(\overline{X})\right] = \mathbb{E}\left[\overline{X}_i\mathbb{E}_{i-1}[\partial_i f(\overline{X})]\mathbb{E}_{j-1}[\partial_i f(\overline{X})]\cdot \mathbb{E}_{j-1}[\overline{X}_j]\right] = 0,$$

since $\mathbb{E}_{j-1}[\overline{X}_j] = 0$. Therefore,

$$\begin{aligned}\sum_{i=1}^n \mathbb{E}\left[\mathbb{E}_{i-1}[\overline{X}_i g(X)]\partial_i f(\overline{X})\right] &\leq \mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \overline{X}_i^2 \mathbb{E}_{i-1}[\partial_i f(\overline{X})]^2\right]^{1/2} \\ &\leq \mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \overline{X}_i^2 \mathbb{E}_{i-1}[\partial_i f(\overline{X})^2]\right]^{1/2} \\ &= \mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \mathbb{E}_{i-1}\left[\overline{X}_i^2\right]\partial_i f(\overline{X})^2\right]^{1/2}.\end{aligned}$$

Using again that $\mathbb{E}_{i-1}\left[\overline{X}_i^2\right] = \mathbb{E}\left[\overline{X}_i^2 | \overline{X}_1, \dots, \overline{X}_{i-1}\right]$, we get

$$\mathbb{E}[f(\overline{X})g(X)] \leq \left(4\sqrt{3}+1\right)\mathbb{E}[g^2(X)]^{1/2}\mathbb{E}\left[\sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^2|\overline{X}_1,\dots,\overline{X}_{i-1}\right]\left(\partial_i f(\overline{X})\right)^2\right]^{1/2}.$$

Taking $g = f \circ R$ with f such that $\int f\, d\bar{\mu} = 0$, we obtain

$$\mathbb{E}[f(\overline{X})^2] \leq \left(4\sqrt{3}+1\right)^2 \mathbb{E}\left[\sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^2|\overline{X}_1,\dots,\overline{X}_{i-1}\right]\left(\partial_i f(\overline{X})\right)^2\right].$$

The inequality is then extended by density to all locally Lipschitz functions $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int f^2\, d\mu < \infty$. □

It remains to prove Lemma 5.1.

Proof of Lemma 5.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable bounded Lipschitz function and denote by $M = 1 + \sup|f|$. For all $x \in \mathbb{R}^n$, we denote by $\|\cdot\|_x$ the quantity defined by

$$\|u\|_x = \sqrt{\sum_{i=1}^n \lambda_i^2(S(x))u_i^2}, \qquad \forall u \in \mathbb{R}^n.$$

When x is such that $\lambda_i(S(x)) < \infty$ for all i, then $\|\cdot\|_x$ is a norm on $\mathbb{R}^n$. With this notation $\tilde{c}_\mu(x,y) = \frac{1}{16}N(\|x-y\|_x)$. and

$$P_t f(x) = \sup_{y\in\mathbb{R}^n}\left\{f(y) - \frac{1}{16t}N\left(\|y-x\|_x\right)\right\}.$$

First, note that, for all $x \in \mathbb{R}^n$, the supremum in the definition of $P_t f(x)$ is attained on the ball $\{y \in \mathbb{R}^n; \|y-x\|_x \leq N^{-1}(48Mt)\}$. Namely, if y is outside the ball, it holds

$$f(y) - f(x) - \frac{1}{16t}N\left(\|y-x\|_x\right) \leq -M < 0.$$

Since $P_t f(x) \geq f(x)$, we conclude that the supremum is reached inside the ball.

Now let us bound from above the derivative of $P_t f$ with respect to the t variable. Let $x \in \mathbb{R}^n$ be such that $\lambda_i(x) < \infty$ for all i. Using the preceding remark and the inequality

$uv \le \frac{u^2}{2} + \frac{v^2}{2}$, we see that

$$
\begin{aligned}
\frac{1}{t}(P_t f(x) - f(x)) &= \sup_{\|y-x\|_x \le N^{-1}(48Mt)} \left\{ \frac{f(y)-f(x)}{t} - \frac{1}{16t^2} N(\|y-x\|_x) \right\} \\
&\le 4 \sup_{\|y-x\|_x \le N^{-1}(48Mt)} \left\{ \frac{(f(y)-f(x))^2}{N(\|y-x\|_x)} \right\} \qquad (5.3) \\
&= 4 \sup_{\|u\|_x \le N^{-1}(48Mt)} \left\{ \frac{(\nabla f(x)\cdot u)^2 + o(\|u\|_x^2)}{\frac{\|u\|_x^2}{2} + o(\|u\|_x^2)} \right\} \\
&\le 8 \sup_{\|u\|_x \le N^{-1}(48Mt)} \left\{ \frac{\sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))}(\partial_i f)(x)^2 + o(1)}{1+o(1)} \right\} \to 8 \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))} (\partial_i f)(x)^2,
\end{aligned}
$$

when t goes to 0, where the last inequality follows from Cauchy-Schwarz. So we conclude that if x is such that $\lambda_i(S(x)) < \infty$ for all i, then

$$
\limsup_{t\to 0} \frac{1}{t}(P_t f(x) - f(x)) \le 8 \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))}(\partial_i f)(x)^2.
$$

If x is such that $\lambda_i(x) = \infty$ for some i, then $P_t f(x) = f(x)$ and so the inequality above is still true.

Moreover, denoting by $\lambda^*(x) = \min\{\lambda_i(S(x))\} > 0$ a.s. (as regards the positivity, we recall that $1/\lambda_i(y) = \mathrm{Var}(X_i|X_1 = y_1, \ldots, X_{i-1} = y_{i-1})$ and is therefore finite for almost every $y \in \mathbb{R}^n$), it follows from (5.3) and from the inequality $\|u\|_x \ge \lambda^*(x)|u|$, $u \in \mathbb{R}^n$, that for all $t \le 1/(48M)$

$$
\frac{1}{t}(P_t f(x) - f(x)) \le 4 \sup_{\lambda^*(x)|y-x| \le N^{-1}(1)} \left\{ \frac{(f(y)-f(x))^2}{N(\lambda^*(x)|y-x|)} \right\} \le a \frac{L^2}{\lambda^*(x)^2}, \qquad (5.4)
$$

where L the Lipschitz constant of f and $a = 4\sup_{0<v\le N^{-1}(1)} \frac{v^2}{N(v)}$ is some universal constant.

Now, let ν be a probability measure on $\mathbb{R}^n$ such that $\int \frac{1}{\lambda_i^2(S(x))}\,\nu(dx) < +\infty$ for all i. Then $1/\lambda^*$ is also square integrable with respect to ν. Therefore, thanks to (5.4), one can apply Fatou's lemma in its lim sup form:

$$
\limsup_{t\to 0} \int \frac{1}{t}(P_t f - f)\, d\nu \le \int \limsup_{t\to 0} \frac{1}{t}(P_t f - f)\, d\nu \le 8 \int \sum_{i=1}^n \frac{1}{\lambda_i^2(S(x))}(\partial_i f)(x)^2\, d\nu.
$$

□

6. Variance estimates

Here we prove Corollary 1.6, identity (1.10) and Corollary 1.11.

Proof of Corollary 1.6. According to Theorem 1.1 and standard properties of conditional expectations, it holds

$$
\begin{aligned}
\mathrm{Var}(|\overline{X}|^2) &\le 4a \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}[\overline{X}_i^2|\overline{X}_1, \ldots, \overline{X}_{i-1}]\overline{X}_i^2\right] = 4a \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}[\overline{X}_i^2|\overline{X}_1, \ldots, \overline{X}_{i-1}]^2\right] \\
&\le 4a \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}[\overline{X}_i^4|\overline{X}_1, \ldots, \overline{X}_{i-1}]\right] = 4a \sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^4\right].
\end{aligned}
$$

Observe that $\mathbb{E}[\overline{X}_i^4] \le 8\mathbb{E}[X_i^4] + 8\mathbb{E}\left[\mathbb{E}_{i-1}[X_i]^4\right] \le 16\mathbb{E}[X_i^4]$. We conclude using Borell's reverse Hölder inequality [11]: there exists some universal constant a' such that for any log-concave random variable Y, it holds $\mathbb{E}[Y^4] \le a'\mathbb{E}[Y^2]^2$. So,

$$\mathrm{Var}(|\overline{X}|^2) \le 64aa'n.$$

□

Remark 6.1. *Our main result Theorem 1.1 is closely related to a result by Barthe and Cordero-Erausquin [4]. Namely, it follows from [4, Theorem 4] that if X is a random vector following a law $\mu(dx) = e^{-V(x)}\,dx$ on $\mathbb{R}^n$ with $\mathrm{Hess}\,V \ge \rho\mathrm{Id}$ for some $\rho \ge 0$, then for all smooth function $f : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\mathbb{E}[\partial_i f(X)|X_1 \dots, X_{i-1}, X_{i+1}, \dots, X_n] = 0, \qquad \forall i \in \{1, \dots, n\} \tag{6.2}$$

it holds

$$\mathrm{Var}_\mu(f(X)) \le \sum_{i=1}^n \mathbb{E}\left[\frac{1}{\rho + 1/C_i(X)}\partial_i f^2(X)\right],$$

where, for all $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $C_i(x)$ denotes the Poincaré constant of the conditional distribution of X_i knowing $X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \dots, X_n = x_n$. Note that the conclusion of [4, Theorem 4] is more general than what we state above. In the general formulation, to any decomposition of the identity $\mathrm{Id} = \sum_{i=1}^m c_i P_{E_i}$ where $c_i > 0$ and P_{E_i} is the orthogonal projection on a subspace E_i corresponds a weighted Poincaré inequality involving the Poincaré constants of the conditional distributions of X knowing $P_{F_i}(X)$, with $F_i = E_i^\perp$.

It is well known (see for instance Theorem 4.2 below) that Poincaré constants of one dimensional log-concave probability measures can be estimated by the variance. In particular, it holds

$$C_i(x) \le 3\mathrm{Var}(X_i|X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \dots, X_n = x_n), \qquad \forall i \in \{1, \dots, n\}.$$

Therefore, taking $\rho = 0$, it holds

$$\mathrm{Var}_\mu(f(X)) \le 3\sum_{i=1}^n \mathbb{E}\left[\mathrm{Var}(X_i|X_1 \dots, X_{i-1}, X_{i+1}, \dots, X_n)\partial_i f^2(X)\right], \tag{6.3}$$

for all smooth f enjoying 6.2. The difference between this result and Theorem 1.1 (besides the fact that we condition only with respect to the first variables) is that our result is true for all f but for $\overline{X}$ instead of X.

Let us see how to recover the conclusion of Corollary 1.6 from (6.3). Let us assume that X is such that

$$\mathbb{E}[X_i|X_1 \dots, X_{i-1}, X_{i+1}, \dots, X_n] = 0, \qquad \forall i \in \{1, \dots, n\}.$$

This condition (which is actually a bit stronger than the condition $\overline{X} = X$) exactly amounts to require that the function $f(x) = |x|^2$ satisfies (6.2). So applying (6.3) to this function and reasoning as in the proof of Corollary 1.6 we thus get from (6.3) that $\mathrm{Var}(|X|^2) \le an$ for some universal constant a.

Proof of Corollary 1.11. Let us start with identity (1.10). For all $i \in \{1, \dots, n\}$, it holds

$$\mathbb{E}\left[\overline{X}_i\, \mathbb{E}_{i-1}[X_i]\right] = \mathbb{E}\Big[(X_i - \mathbb{E}_{i-1}[X_i])\, \mathbb{E}_{i-1}[X_i]\Big] = \mathbb{E}\Big[\mathbb{E}_{i-1}[X_i - \mathbb{E}_{i-1}[X_i]]\; \mathbb{E}_{i-1}[X_i]\Big] = 0.$$

As a result $\overline{X}$ and X' are orthogonal in $\mathbb{L}_2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^n)$. Therefore, it holds

$$\mathbb{E}[|X|^2] = \mathbb{E}[|\overline{X}|^2] + \mathbb{E}[|X'|^2] \quad \text{and} \quad \mathbb{E}[|X|^2]^2 = \mathbb{E}[|\overline{X}|^2]^2 + 2\mathbb{E}[|\overline{X}|^2]\mathbb{E}[|X'|^2] + \mathbb{E}[|X'|^2]^2.$$

Since

$$\mathbb{E}[|X|^4] = \mathbb{E}[|\overline{X}|^4] + 2\mathbb{E}[|\overline{X}|^2]\mathbb{E}[|X'|^2] + \mathbb{E}[|X'|^4] + 4\mathbb{E}[(\overline{X}\cdot X')^2] + 4\mathbb{E}[|\overline{X}|^2\overline{X}\cdot X'] + 4\mathbb{E}[|X'|^2\overline{X}\cdot X'],$$

we get that

$$\mathrm{Var}(|X|^2) = \mathrm{Var}(|\overline{X}|^2) + \mathrm{Var}(|X'|^2) + 2\mathrm{Cov}(|\overline{X}|^2, |X'|^2) + 4\mathbb{E}[(\overline{X}\cdot X')^2] + 4\mathbb{E}[|\overline{X}|^2\overline{X}\cdot X'] + 4\mathbb{E}[|X'|^2\overline{X}\cdot X']$$

Using Cauchy-Schwarz, and the orthogonality of $\overline{X}$ and X' we get

$$\left|\mathrm{Cov}\left(|\overline{X}|^2, |X'|^2\right)\right| \leq \sqrt{\mathrm{Var}\left(|\overline{X}|^2\right)}\sqrt{\mathrm{Var}\left(|X'|^2\right)}$$

$$\left|\mathbb{E}\left[|\overline{X}|^2\, \overline{X}\cdot X'\right]\right| = \left|\mathbb{E}\left[\left(|\overline{X}|^2 - \mathbb{E}[|\overline{X}|^2\right)\, \overline{X}\cdot X'\right]\right| \leq \sqrt{\mathrm{Var}(|\overline{X}|^2)}\sqrt{\mathbb{E}[(\overline{X}\cdot X')^2]}$$

$$\left|\mathbb{E}\left[|X'|^2\, \overline{X}\cdot X'\right]\right| = \left|\mathbb{E}\left[\left(|X'|^2 - \mathbb{E}[|X'|^2\right)\, \overline{X}\cdot X'\right]\right| \leq \sqrt{\mathrm{Var}(|X'|^2)}\sqrt{\mathbb{E}[(\overline{X}\cdot X')^2]}.$$

Moreover, note that if $i<j$ the random variable $\overline{X}_i\mathbb{E}_{i-1}[X_i]\mathbb{E}_{j-1}[X_j]$ is measurable with respect to the σ field generated by $X_1,\ldots,X_{j-1}$. Therefore

$$\mathbb{E}\left[\overline{X}_i\mathbb{E}_{i-1}[X_i]\overline{X}_j\mathbb{E}_{j-1}[X_j]\right] = \mathbb{E}\left[\overline{X}_i\mathbb{E}_{i-1}[X_i]\mathbb{E}_{j-1}[\overline{X}_j]\mathbb{E}_{j-1}[X_j]\right] = 0$$

So, it holds

$$\begin{aligned}
\mathbb{E}[(\overline{X}\cdot X')^2] &= \sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^2\mathbb{E}_{i-1}[X_i]^2\right] + 2\sum_{i<j}\mathbb{E}\left[\overline{X}_i\mathbb{E}_{i-1}[X_i]\overline{X}_j\mathbb{E}_{j-1}[X_j]\right]\\
&= \sum_{i=1}^n \mathbb{E}\left[\overline{X}_i^2\mathbb{E}_{i-1}[X_i]^2\right] = \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}_{i-1}[X_i^2]\mathbb{E}_{i-1}[X_i]^2 - \mathbb{E}_{i-1}[X_i]^4\right]\\
&\leq \sum_{i=1}^n \mathbb{E}[X_i^4] \leq a'n,
\end{aligned}$$

where a' is some universal constant such that $\mathbb{E}[Y^4]\leq a'\mathbb{E}[Y^2]^2$ for all log-concave random variable Y. We conclude from the inequalities above that

$$\begin{aligned}
\mathrm{Var}(|X|^2) &\leq \left(\sqrt{\mathrm{Var}(|\overline{X}|^2)} + \sqrt{\mathrm{Var}(|X'|^2)}\right)^2 + 4a'n + 4\sqrt{a'n}\left(\sqrt{\mathrm{Var}(|\overline{X}|^2)} + \sqrt{\mathrm{Var}(|X'|^2)}\right)\\
&= \left(\sqrt{\mathrm{Var}(|\overline{X}|^2)} + \sqrt{\mathrm{Var}(|X'|^2)} + 2\sqrt{a'n}\right)^2 \leq \left(\sqrt{\mathrm{Var}(|X'|^2)} + (2\sqrt{a'}+b)\sqrt{n}\right)^2\\
&\leq 2\mathrm{Var}(|X'|^2) + 2(2\sqrt{a'}+b)^2 n
\end{aligned}$$

where in the last inequalities b is the universal constant given by Corollary 1.6.

Similarly,

$$\mathrm{Var}(|X|^2) \geq \left(\sqrt{\mathrm{Var}(|X'|^2)} - \sqrt{\mathrm{Var}(|\overline{X}|^2)}\right)^2 - 4\sqrt{a'n}\left(\sqrt{\mathrm{Var}(|X'|^2)} + \sqrt{\mathrm{Var}(|\overline{X}|^2)}\right).$$

Therefore, expanding the square, we see that the number $\sqrt{V'} := \sqrt{\mathrm{Var}(|X'|^2)}$ is less than or equal the positive root of the equation

$$x^2 - 2x\left(\sqrt{\bar{V}} + 2\sqrt{a'n}\right) + \bar{V} - 4\sqrt{a'n}\sqrt{\bar{V}} - V = 0,$$

with $V = \mathrm{Var}(|X|^2)$ and $\bar{V} = \mathrm{Var}(|\overline{X}|^2)$. An easy calculation thus gives

$$\sqrt{V'} \leq \sqrt{\bar{V}} + 2\sqrt{a'n} + \sqrt{4a'n + V},$$

which together with Corollary 1.6 easily gives the desired inequality. □

References

1. David Alonso-Gutiérrez and Jesús Bastero, *The variance conjecture on some polytopes*, Asymptotic geometric analysis, Fields Inst. Commun., vol. 68, Springer, New York, 2013, pp. 1–20. MR 3076144
2. Milla Anttila, Keith Ball, and Irini Perissinaki, *The central limit problem for convex bodies*, Trans. Amer. Math. Soc. **355** (2003), no. 12, 4723–4735. MR 1997580 (2005b:52010)
3. Franck Barthe, *Transportation techniques and Gaussian inequalities*, Optimal transportation, geometry and functional inequalities, CRM Series, vol. 11, Ed. Norm., Pisa, 2010, pp. 1–44. MR 2649000 (2011k:49096)
4. Franck Barthe and Dario Cordero-Erausquin, *Invariances in variance estimates*, Proc. Lond. Math. Soc. (3) **106** (2013), no. 1, 33–64. MR 3020738

5. Franck Barthe and Alexander V. Kolesnikov, *Mass transport and variants of the logarithmic Sobolev inequality*, J. Geom. Anal. **18** (2008), no. 4, 921–979. MR 2438906 (2009i:58048)
6. Sergey G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab. **27** (1999), no. 4, 1903–1921. MR 1742893 (2001h:60026)
7. ______, *On concentration of distributions of random weighted sums*, Ann. Probab. **31** (2003), no. 1, 195–215.
8. ______, *Large deviations and isoperimetry over convex probability measures with heavy tails*, Electron. J. Probab. **12** (2007), 1072–1100. MR 2336600 (2008g:60066)
9. Sergey G. Bobkov, Nathael Gozlan, Cyril Roberto, and Paul-Marie Samson, *Bounds on the deficit in the logarithmic sobolev inequality*, Preprint, 2013.
10. Sergey G. Bobkov and Christian Houdré, *Isoperimetric constants for product probability measures*, Ann. Probab. **25** (1997), no. 1, 184–205. MR 1428505 (98g:60032)
11. Christer Borell, *Convex measures on locally convex spaces*, Ark. Mat. **12** (1974), 239–252. MR 0388475 (52 #9311)
12. ______, *Convex set functions in d-space*, Period. Math. Hungar. **6** (1975), no. 2, 111–136. MR 0404559 (53 #8359)
13. Dario Cordero-Erausquin, *Some applications of mass transport to Gaussian-type inequalities*, Arch. Ration. Mech. Anal. **161** (2002), no. 3, 257–269. MR 1894593 (2003h:49076)
14. Ronen Eldan, *Thin shell implies spectral gap up to polylog via a stochastic localization scheme*, Geom. Funct. Anal. **23** (2013), no. 2, 532–569.
15. Ronen Eldan and Bo'az Klartag, *Dimensionality and the stability of the brunn-minkowski inequality*, To appear in Ann. Sc. Norm. Super. Pisa., 2013.
16. Nathael Gozlan and Christian Léonard, *Transport inequalities. A survey*, Markov Process. Related Fields **16** (2010), no. 4, 635–736. MR 2895086
17. Nathael Gozlan, Cyril Roberto, and Paul-Marie Samson, *Characterization of Talagrand's transport-entropy inequalities in metric spaces*, Ann. Probab. **41** (2013), no. 5, 3112–3139.
18. Olivier Guédon and Emanuel Milman, *Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures*, Geom. Funct. Anal. **21** (2011), no. 5, 1043–1068. MR 2846382
19. Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal, *Convex analysis and minimization algorithms. II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 306, Springer-Verlag, Berlin, 1993, Advanced theory and bundle methods. MR 1295240 (95m:90002)
20. Ravi Kannan, László Lovász, and Miklós Simonovits, *Isoperimetric problems for convex bodies and a localization lemma*, Discrete Comput. Geom. **13** (1995), no. 3-4, 541–559. MR 1318794 (96e:52018)
21. Bo'az Klartag, *A central limit theorem for convex sets*, Invent. Math. **168** (2007), no. 1, 91–131. MR 2285748 (2008d:60034)
22. ______, *A Berry-Esseen type inequality for convex bodies with an unconditional basis*, Probab. Theory Related Fields **145** (2009), no. 1-2, 1–33. MR 2520120 (2010h:60072)
23. ______, *High-dimensional distributions with convexity properties*, European Congress of Mathematics, Eur. Math. Soc., Zürich, 2010, pp. 401–417. MR 2648334 (2011d:46028)
24. ______, *Concentration of measures supported on the cube*, To appear in Israel J. of Math., 2012.
25. ______, *Poincaré inequalities and moment maps*, Ann. Fac. Sci. Toulouse Math. **22** (2013), no. 1, 1–41.
26. Herbert Knothe, *Contributions to the theory of convex bodies*, Michigan Math. J. **4** (1957), 39–52. MR 0083759 (18,757b)
27. Michel Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. MR MR1849347 (2003k:28019)
28. Katalin Marton, *A simple proof of the blowing-up lemma*, IEEE Trans. Inform. Theory **32** (1986), no. 3, 445–446. MR MR838213 (87e:94018)
29. Robert J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179. MR 1451422 (98e:82003)
30. Felix Otto and Cédric Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR 1760620
31. Svetlozar T. Rachev and Ludger Rüschendorf, *Mass transportation problems. Vol. I*, Probability and its Applications (New York), Springer-Verlag, New York, 1998, Theory. MR 1619170 (99k:28006)
32. Michel Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600. MR MR1392331 (97d:60029)
33. Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454 (2010f:49001)

DCE: Institut de Mathématiques de Jussieu and Institut Universitaire de France, Université Pierre et Marie Curie – Paris 6, 4, place Jussieu, 75252 Paris Cedex 05,France
E-mail address: `dario.cordero@imj-prg.fr`

NG: Université Paris Est Marne la Vallée - Laboratoire d'Analyse et de Mathématiques Appliquées (UMR CNRS 8050), 5 bd Descartes, 77454 Marne la Vallée Cedex 2, France
E-mail address: `natael.gozlan@u-pem.fr`