q-MOMENT MEASURES AND APPLICATIONS: A NEW APPROACH VIA OPTIMAL TRANSPORT

HUYNH KHANH[1]   FILIPPO SANTAMBROGIO[2]

ABSTRACT. In 2017, Bo'az Klartag obtained a new result in differential geometry on the existence of affine hemisphere of elliptic type. In his approach, a surface is associated with every a convex function $\varphi : \mathbb{R}^n \to (0, +\infty)$ and the condition for the surface to be an affine hemisphere involves the 2-moment measure of φ (a particular case of q-moment measures, i.e measures of the form $(\nabla \varphi)_{\#} \varphi^{-(n+q)}$ for $q > 0$). In Klartag's paper, q-moment measures are studied through a variational method requiring to minimize a functional among convex functions, which is studied using the Borell-Brascamp-Lieb inequality. In this paper, we attack the same problem through an optimal transport approach, since the convex function φ is a Kantorovich potential (as already done for moment measures in a previous paper). The variational problem in this new approach becomes the minimization of a local functional and a transport cost among probability measures ϱ and the optimizer ϱ_{opt} turns out to be of the form $\varrho_{\text{opt}} = \varphi^{-(n+q)}$.

Contents

1   Introduction

This paper stems for the analysis performed in [10, 7] and [15]. More precisely, Klartag studied in [10] the connection between the notion of affine hemisphere of elliptic type and that of *q-moment measures*, and proved an existence result thanks to a variational problem in the class of convex functions. To clarify what we mean, let us recall that, given a convex function $\varphi : \mathbb{R}^n \to (0, +\infty)$ with $\lim_{|x| \to \infty} \varphi(x) = +\infty$ and a positive real number $q > 0$, a

---

[1]Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam (`khanh.edu02@gmail.com`).

[2]Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France, and Institut Universitaire de France (`santambrogio@math.univ-lyon1.fr`).

Borel probability measure μ on $\mathbb{R}^n$ is said to be the q-moment measure of φ if the differential of φ pushes forward the measure $\varphi^{-(n+q)}\mathrm{d}x$ towards μ:

$$\mu := (\nabla\varphi)_{\neq}\varrho, \quad \text{where } \mathrm{d}\varrho = \varphi^{-(n+q)}\mathrm{d}x. \tag{1}$$

A similar path was followed in [7] for the more well-known notion of *moment measures*, where (1) is replaced by

$$\mu := (\nabla\varphi)_{\neq}\varrho, \quad \text{where } \mathrm{d}\varrho = e^{-\varphi}\mathrm{d}x. \tag{2}$$

Both [7] and [10] include a characterization of those measures μ which are indeed moment measures (or q-moment measures) of some function φ, by minimizing a convex functional on φ involving μ, and studying its subdifferential so as to prove that the minimizer solves (2) or (1).

The experienced reader will have noticed the connection of these notions with that of optimal transport. Indeed, by Brenier's Theorem, the map $\nabla\varphi$ will be the optimal transport map for the quadratic cost $c(x, y) = \frac{1}{2}|x - y|^2$ from ϱ to μ. Based on this connection, the second author provided in [15] an optimal-transport-based approach to the problem of moment measures. The goal of the present paper is to do the same with q-moment measures.

Before doing this, let us clarify the connection of q-moment measures with affine spheres, which was the motivation of [10]. Indeed, Klartag obtained in [10] a new result on the existence and uniqueness of affine hemisphere of elliptic type (i.e, affine hemispheres whose center is inside a convex set bounded by a hypersurface) by studying the problem of 2-moment measures. hHe showed that for any $q > 1$ any probability measure μ with finite first moment, not supported on a hyperplane, and with 0 as barycenter, can be represented as the q-moment measure for some (unique up to translations) essentially continuous convex function $\varphi : \mathbb{R}^n \to (0, +\infty)$. This can be applied to the problem of finding affine hemispheres of elliptic type if one chooses μ to be the uniform measure on a bounded convex set and takes $q = 2$.

For the reader coming from optimal transport and not from convex geometry, it would be useful to explain the notion of affine spheres and hemispheres. Namely, given a smooth and connected hypersurface $M \subset \mathbb{R}^{n+1}$ which is locally strongly-convex and an arbitrary point $p \in M$, one considers the tangent space P to M at the point p and its parallel translates $(P_t)_{t \geq 0}$ which are defined by $P_0 = P$ and $P_t = P_0 + tv$ where $v \notin P$ is a vector pointing to the convex side of M at the point p. Then, for every t, the plane P_t cuts the hypersurface and selects a compact sector Q_t of finite volume inside M, and we call G_t the barycenter of Q_t. As $t \to 0^+$ the point G_t draws a curve which ends at $G_0 = p$. The tangent to the curve at p selects a special direction, which is the direction of the affine normal. The line $\ell_M(p)$ passing through p and directed as the affine normal at p is called the affine normal line at $p \in M$. The hypersurface M is called affinely-spherical of elliptic type with center at a point $x \in \mathbb{R}^{n+1}$ if all of the affine normal lines of M meet at x and the ray from an arbitrary point of M to the center always passes through the convex side of it. Then, M is called affine hemisphere of elliptic type with anchor K if the following two conditions are satisfied: (i) there exist compact, convex sets $K, \widetilde{K} \subset \mathbb{R}^{n+1}$ with $\dim(K) = n$ and $\dim\left(\widetilde{K}\right) = n+1$ such

that M does not intersect the affine hyperplane spanned by K and $\partial \widetilde{K} = M \cup K$; (ii) the center of M belongs to the relative interior of K. All these notions are invariant for affine transformation (differently from the usual notion of normal to a surface), which explains the choice of the name.
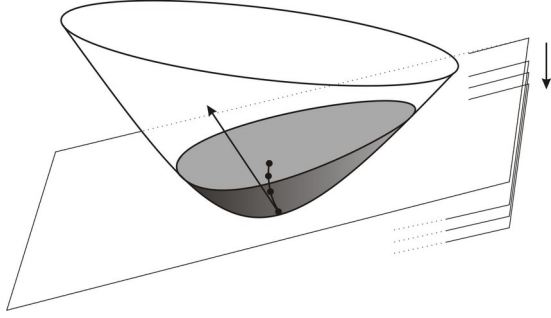

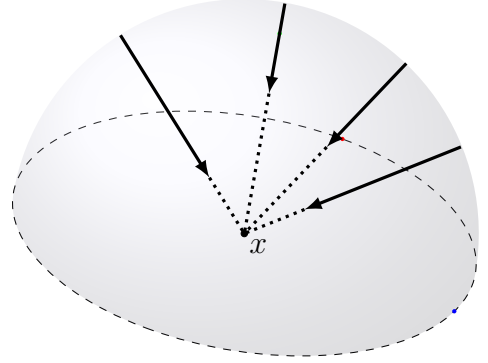FIGURE 1. Geometrical construction of the affine normal.

FIGURE 2. An affine hemisphere of elliptic type

In [10], Klartag proved that *for any n-dimensional, compact and convex set $K \subset \mathbb{R}^{n+1}$, there always exists an $(n+1)$-dimensional, compact and convex set $\widetilde{K} \subset \mathbb{R}^{n+1}$ whose boundary consists of two parts, the convex set K itself is a facet and the rest of the boundary is an affine hemisphere with anchor K, which is centered at the Santaló point[3] of K and the affine hemisphere is uniquely determined up to transformations.* Thanks to affine transformations in $\mathbb{R}^{n+1}$, one may assume that the Santaló point of the set K lies at the origin and $K \subset \{(x,0) : x \in \mathbb{R}^n\}$. The set K in those cases can be rewritten in form $K = L^\circ \times \{0_\mathbb{R}\}$, where L° is the polar body of a certain convex set $L \subset \mathbb{R}^n$. Then the proof of the existence of affine hemispheres of elliptic type relies upon analysis of the following PDE of Monge-Ampère type (see also [8])

$$\begin{cases} \det \nabla^2 \varphi(x) & = \; \dfrac{\mathrm{Vol}_n(L)}{(\varphi(x))^{n+2}}, \quad x \in \mathbb{R}^n \\ \nabla \varphi(\mathbb{R}^n) & = \; L \end{cases} \tag{4}$$

Indeed, it is proven (Theorem 1.2 in [10]) that, if φ is a solution of (4) then we obtain an affine hemisphere with anchor $K = L^\circ \times \{0_\mathbb{R}\}$ setting

$$M = \left\{ \left(\frac{x}{\varphi(x)}, \; \frac{1}{\varphi(x)} \right) \in \mathbb{R}^n \times \mathbb{R} : \; x \in \mathrm{dom}(\varphi) \right\} \tag{5}$$

Equation (4) is quite similar to the moment measures equation of Berman and Berndtsson [3] in their work on Kähler-Einstein metrics in toric manifolds; Cordero-Erausquin and Klartag

---

[3]Given any an n-dimensional, non-empty, bounded and convex set $K \subset \mathbb{R}^n$, then there exists a unique point z in the interior of K such that

$$\mathrm{Vol_n}\left((K-z)^\circ\right) = \inf_{x \in \mathrm{int}(K)} \left[\mathrm{Vol_n}\left((K-x)^\circ\right)\right] \tag{3}$$

here S° is the polar body of a convex set $S \subset \mathbb{R}^n$, defined by $S^\circ = \{x \in \mathbb{R}^n : \sup_{s \in S} \langle x, s \rangle \leq 1\}$. The unique point z in (3) is called the Santaló point of K (see also [2, 12]). Note that the polar body $(K-z)^\circ$ has its barycenter at the origin if and only if z is the Santaló of K. As a corollary, the Santaló point of K lies at the origin if and only if the barycenter of K° lies at the origin.

extended the study of [3] in [7] presenting a functional version of the classical Minkowski problem or the logarithmic Minkowski problem and providing a variational characterization. More recently, the second author provided in [15] a dual counter-part of the results obtained in [7] with ideas coming from the theory of optimal transport, by considering the minimization of an entropy and a transport cost among probability measures.

In [10], the existence and the uniqueness of the solution to the equation (4) are proven via a variational method considering a minimization among convex functions. Namely, the author analyzes the subgradient of the functional

$$\phi \mapsto \mathcal{I}_q\left(\phi\right) = \left(\int_{\mathbb{R}^n} \frac{\mathrm{d}x}{\left(\phi^*\left(x\right)\right)^{n+q-1}}\right)^{\frac{-1}{q-1}}, \tag{6}$$

where $\phi^*\left(y\right) := \sup_x \left(x.y - \phi\left(x\right)\right)$ is the Legendre transform of ϕ and is the smallest function compatible with ϕ in the constraint $x.y \leq \phi\left(x\right) + \phi^*\left(y\right)$. The convexity of the functional $\mathcal{I}_q$, which follows from the Borell-Brascamp-Lieb inequality, is the key for the proof. Equation (4) is then obtained imposing that the uniform measure on L belongs to the subdifferential, and taking $q = 2$.

Taking ideas from the purely optimal-transport-based method for moment measures by the second author [15], the main goal of the present work is to reprove the same existence result of [10] with a different method, replacing functional inequalities techniques with ideas from optimal transport. In [15] some heuristics are presented in order to guess which variational problem should be used (inspired by the theory of the JKO scheme for gradient flows, see for instance [1, 9, 16]) and the functional to be minimized involves both a local functional[4] of the form $\varrho \mapsto \int f(\varrho(x))\mathrm{d}x$ and a transport cost $\mathcal{T}\left(\varrho, \mu\right)$. We will not enter here into these considerations and we will directly look at the chosen functional.

We indeed study the following minimization problem

$$(P) \qquad \min\left\{\mathcal{J}\left(\varrho\right) := \mathcal{F}\left(\varrho\right) + \mathcal{T}\left(\varrho, \mu\right) : \quad \varrho \in \mathcal{P}_1\left(\mathbb{R}^n\right)\right\}, \tag{7}$$

where the functional $\mathcal{F}\left(.\right)$ is a very particular local functional (see Proposition 7.7 of [14]), given by

$$\mathcal{F}\left(\varrho\right) := \int_{\mathbb{R}^n} f\left(\varrho^{\mathrm{ac}}\right)\mathrm{d}x, \qquad \text{writing } \varrho = \varrho^{\mathrm{ac}}(x)\mathrm{d}x + \varrho^{\mathrm{sing}} \tag{8}$$

(the decomposition $\varrho = \varrho^{\mathrm{ac}}(x)\mathrm{d}x + \varrho^{\mathrm{sing}}$ being the Radon-Nicodym decomposition of ϱ into an absolutely continuous and a singular part w.r.t. the Lebesgue measure), choosing $f\left(t\right) = -\frac{1}{\alpha}t^\alpha$ (to simplify notations, we set $\alpha = 1 - \frac{1}{n+q} \in (0,1)$). Note that the singular part of ϱ does not appear in the functional since $f'(\infty) = 0$. The functional $\mathcal{T}\left(., \mu\right)$ considered in this case is, instead, the maximal correlation functional (see [15], section 3) and it is defined as follows

$$\mathcal{T}\left(\varrho, \mu\right) := \sup\left\{\int_{\mathbb{R}^n \times \mathbb{R}^n} x.y \mathrm{d}\gamma\left(x,y\right) \mid \gamma \in \Gamma(\mu, \varrho)\right\} \tag{9}$$

---

[4]Local functionals over measures are defined as those functionals $F : \mathcal{P}\left(\Omega\right) \to \mathbb{R}$ such that $F\left(\mu + \nu\right) = F\left(\mu\right) + F\left(\nu\right)$ whenever and μ and ν are mutually singular (i.e., there exists $A, B \subset \Omega$ with $A \cup B = \Omega$, $\mu\left(A\right) = 0$ and $\nu\left(B\right) = 0$).

where $\Gamma(\mu, \varrho)$ is the set of transport plans between two probabily measures μ and ϱ on $\mathbb{R}^n$.

We will see in the rest of the paper that for some parts of the analysis we need $q > 0$ (in particular for geodesic convexity issues) and for some other parts we need a stronger assumption, i.e. $q > 1$ (in particular for the bounds which are needed to prove existence of a minimizer). However, the application to affine hemispheres (that we will not develop here, but is the motivation of the paper) requires $q = 2$, which is fully covered by our results.

2 Technical tools from optimal transport and local functionals on measures

We recall here the main notions and notations that we will use throughout the paper. We refer to [14] (Chapters 1, 5 and 7) and to [1, 17, 18] for more details and complete proofs.

Optimal Transport and Wasserstein distances.

Given two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ we consider the set of transport plans

$$\Gamma(\mu, \nu) = \left\{ \gamma \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) : (\pi_x)_\# \gamma = \mu, \ (\pi_y)_\# \gamma = \nu \right\} \tag{10}$$

i.e. those probability measures on the product space having μ and ν as marginal measures.

For a cost function $c : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty]$ we consider the minimization problem

$$\min \left\{ \int c \, \mathrm{d}\gamma : \ \gamma \in \Gamma(\mu, \nu) \right\} \tag{11}$$

which is called the Kantorovich optimal transport problem for the cost c from μ to ν. In particular, we consider the case $c(x, y) = \frac{1}{2}|x - y|^2$. In this case the above minimal value is finite whenever $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$, where $\mathcal{P}_2(\mathbb{R}^n) := \{ \varrho \in \mathcal{P}(\mathbb{R}^n) : M_2(\varrho) < +\infty \}$ and for $p \geq 1$ we define $M_p(\varrho) := \int |x|^p \mathrm{d}\varrho(x)$.

For the above problem one can prove that the minimal value also equals the maximal value of a dual problem

$$\max \left\{ \int \phi \, \mathrm{d}\mu + \int \psi \, \mathrm{d}\nu : \phi(x) + \psi(y) \leq \frac{1}{2}|x - y|^2 \right\}, \tag{12}$$

and that the optimal function ϕ may be used to construct an optimizer γ. Indeed, the optimal ϕ is locally Lipschitz and semiconcave (more precisely, $x \mapsto \frac{1}{2}|x|^2 - \phi(x)$ is convex), and differentiable μ−a.e. if $\mu \ll \mathcal{L}^n$; one can define a map $T : \mathbb{R}^n \to \mathbb{R}^n$ through $T(x) = x - \nabla\phi(x)$ and this map satisfies $T_\# \mu = \nu \ \gamma_T := (\mathrm{id}, T)_\# \mu$ (i.e. the image measure of μ through the map $x \mapsto (x, T(x))$ belongs to $\Gamma(\mu, \nu)$ and is optimal in the above problem). Moreover, the map T is the gradient of the convex function u given by $u(x) = \frac{1}{2}|x|^2 - \phi(x)$ and is called the optimal transport map (for the quadratic cost $c(x, y) = \frac{1}{2}|x - y|^2$) from μ to ν. The fact that the optimal transport map T exists, is unique, and is the gradient of a convex function is known as Brenier Theorem (see [6]).

The same could be obtained if one withdrew from the cost $\frac{1}{2}|x-y|^2$ the parts $\frac{1}{2}|x|^2$ and $\frac{1}{2}|y|^2$ which only depend on one variable each (hence, their integral w.r.t. γ only depends on its marginals). Doing this we would get to the transport maximization problem

$$\mathcal{T}(\mu,\nu) = \max\left\{\int x.y\,\mathrm{d}\gamma(x,y) \mid \gamma \in \Gamma(\mu,\nu)\right\} \tag{13}$$

and the dual problem would become

$$\inf\left\{\int u\mathrm{d}\mu + \int v\mathrm{d}\nu : u(x) + v(x) \geq x.y\right\}. \tag{14}$$

In this problem it is quite clear that any pair (u,v) can be replaced with (u,u^*) where $u^*(y) = \sup_x\{x.y - u(x)\}$ is the Legendre transform of u, or even with (u^{**}, u^*). We can then assume that both u and v are convex and l.s.c.

Then, it is easy to see by the primal-dual optimality conditions that the optimal γ and the optimal u satisfy

$$\mathrm{spt}(\gamma) \subset \{(x,y) : u(x) + u^*(y) = x.y\} = \{(x,y) : y \in \partial u(x)\}, \tag{15}$$

which shows that γ is concentrated on the graph of a map T given by $T = \nabla u$, which is well-defined $\mu-$a.e provided $\mu \ll \mathcal{L}^n$.

The value of the minimization problem with the quadratic cost may also be used to define a quantity, called Wasserstein distance, over $\mathcal{P}_2(\mathbb{R}^n)$

$$W_2(\mu,\nu) := \left(\min\left\{\int |x-y|^2\mathrm{d}\gamma(x,y) : \gamma \in \Gamma(\mu,\nu)\right\}\right)^{1/2}. \tag{16}$$

This quantity may be proven to be a distance over $\mathcal{P}_2(\mathbb{R}^n)$ and the space $\mathcal{P}_2(\mathbb{R}^n)$ endowed with the distance W_2 is called Wasserstein space of order 2, denoted by $\mathbb{W}_2(\mathbb{R}^n)$. On compact sets, this distance metrizes the usual weak convergence of probability measures (we say that a sequence ϱ_n weakly converges to ϱ if $\int \phi\,\mathrm{d}\varrho_n \to \int \phi\,\mathrm{d}\varrho$ for every bounded and continuous function ϕ, and we write $\varrho_n \rightharpoonup \varrho$), while on $\mathbb{R}^n$ it metrizes a stronger notion of weak convergence, namely $\int \phi\,\mathrm{d}\varrho_n \to \int \phi\,\mathrm{d}\varrho$ for every continuous function ϕ with $|\phi(x)| \leq C(1+|x|^2)$.

The geodesics in this space play an important role in the theory of optimal transport. Given $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^n)$ with $\mu \ll \mathcal{L}^n$, we define $\varrho_t := ((1-t)\,\mathrm{id} + tT)_{\#}\mu$, where T is the optimal transport from μ to ν. This curve ϱ_t happens to be a constant speed geodesic for the distance W_2 connecting μ to ν.

Once we know the geodesics in $\mathbb{W}_2(\mathbb{R}^n)$, one can wonder which functionals $F : \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$ are geodesically convex, i.e. convex along constant speed geodesics. This notion, applied to the case of the Wasserstein spaces, is also called *displacement convexity* and has been introduced by McCann in [11]. It is very useful both to provide uniqueness results for variational problems and to provide sufficient optimality conditions. We will discuss this notion both in what co cerns the functional $\mathcal{F}$ and the functional $\mathcal{T}$.

Maximal correlation functional.

We want to come back to the functional $\mathcal{T}$. It is useful to note that, if we have $\varrho, \mu \in \mathcal{P}_2(\mathbb{R}^n)$, then we also have

$$\mathcal{T}(\varrho, \mu) = \frac{1}{2} \int_{\mathbb{R}^n} |x|^2 \mathrm{d}\varrho(x) + \frac{1}{2} \int_{\mathbb{R}^n} |y|^2 \mathrm{d}\mu(y) - \frac{1}{2} W_2^2(\varrho, \mu). \tag{17}$$

If $\mathcal{T}$ comes from a transport cost, we may also observe that it stands for the maximal correlation between ϱ and μ, in the sense that we have

$$\mathcal{T}(\varrho, \mu) = \sup \left\{ \mathbb{E}[X.Y] \ : \ X \sim \varrho, Y \sim \mu \right\}. \tag{18}$$

For this reason, $\mathcal{T}$ will be called *maximal correlation functional*. In this paper, we will make use of the following properties which were established in [15].

Proposition 2.1. *Suppose that $\mu \in \mathcal{P}(\mathbb{R}^n)$ is such that $M_1(\mu) < +\infty$ and $\int y, \mathrm{d}\mu(y) = 0$. Then we have the following properties.*
(1) For every $\varrho \in \mathcal{P}_1(\mathbb{R}^n)$, we have $0 \leq \mathcal{T}(\varrho, \mu) \leq +\infty$.
(2) If ϱ and $\widetilde{\varrho}$ are one obtained from one another by translation, then $\mathcal{T}(\varrho, \mu) = \mathcal{T}(\widetilde{\varrho}, \mu)$.
(3) (Lower semicontinuity) If $\int x \mathrm{d}\varrho_n(x) = 0$ and $\varrho_n \rightharpoonup \varrho$, then

$$\mathcal{T}(\varrho, \mu) \leq \liminf_{n \to +\infty} \mathcal{T}(\varrho_n, \mu). \tag{19}$$

(4) There exists a sequence μ_n of compactly supported probability measures with $\mu_n \rightharpoonup \mu$ and $\int y \mathrm{d}\mu(y) = 0$ such that for every $\varrho \in \mathcal{P}_1(\mathbb{R}^n)$ we have $\mathcal{T}(\varrho, \mu_n) \to \mathcal{T}(\varrho, \mu)$.
(5) If moreover μ is not supported on a hyperplane then, for every ϱ such that $\int x \mathrm{d}\varrho(x) = 0$, $\mathcal{T}$ satisfies an inequality of the form

$$\mathcal{T}(\varrho, \mu) \geq c \int_{\mathbb{R}^n} |x| \, \mathrm{d}\varrho(x) \tag{20}$$

for $c = c(\mu) > 0$, where

$$c(\mu) := \frac{1}{2n} \inf \left\{ \int_{\mathbb{R}^n} |y.e - \ell| \, \mathrm{d}\mu(y) : e \in \mathbb{S}^{n-1}, \ell \in \mathbb{R} \right\}.$$

(6) (Displacement convexity). Let $\varrho_0, \varrho_1 \in \mathcal{P}_2(\mathbb{R}^n)$ be absolutely continuous measures, and let $\varrho_t = ((1-t)\,\mathrm{id} + tT)_\# \varrho_0$ be the unique constant speed geodesic connecting them for the Wasserstein distance W_2. Then $t \mapsto \mathcal{T}(\varrho_t, \mu)$ is convex on $[0,1]$. Moreover, if ∇v is the optimal transport from ϱ_0 to ϱ_1 and $\nabla \varphi$ is the optimal transport from ϱ_0 to μ then we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \Big(\mathcal{T}(\varrho_t, \mu) \Big)_{|t=0} \geq \int_{\mathbb{R}^n} (\nabla v(x) - x) . \nabla \varphi(x) \, \varrho_0(x) \, \mathrm{d}x. \tag{21}$$

Lower semicontinuity and bounds for the local functional. As we said, in this paper we will make use of a particular local functional $\mathcal{F}$, defined via (8).

The functional $\mathcal{F}$ may be written in the form

$$\varrho \mapsto \int_\Omega f\left(\varrho^{\mathrm{ac}}\left(x\right)\right) \mathrm{d}\lambda\left(x\right) + \left(\lim_{t \to +\infty} \frac{f\left(t\right)}{t}\right) \cdot \varrho^{\mathrm{sing}}\left(\Omega\right) \tag{22}$$

and for local functionals of this form lower semicontinuity results are well-known since at least [5] (see also Chapter 7 in [14] or Chapter 10 in [1]). Yet, these results are true for convex and l.s.c. function $f : \mathbb{R}_+ \to \mathbb{R}$ (which is the case for $f\left(t\right) = -\frac{1}{\alpha}t^\alpha$) but they require the reference measure λ (which is the Lebesgue measure here) to be finite. Unfortunaltely, we need to consider here the whole, unbounded, space. Whenever $f \geq 0$ this difficulty can be easily solved by taking a sup over finite sub-measures of λ, but this is not the case here. We will then prove at the same time semicontinuity (in the spirit of Exercise 45 in [14]) and lower bounds on $\mathcal{F}$ in the following proposition.

Proposition 2.2. *Suppose $q > 1$. Then, for each $\delta \in \left(\frac{n}{n+q-1}, 1\right)$, there exists a constant $C = C(\delta)$ such that the following estimate holds true*

$$\mathcal{F}\left(\varrho\right) \geq -C - M_1(\varrho)^\delta. \tag{23}$$

Moreover, whenever ϱ_n is a sequence weakly-* converging to a measure ϱ and $M_1(\varrho_n)$ is bounded, then we have $\mathcal{F}(\varrho) \leq \liminf_n \mathcal{F}(\varrho_n)$.

Proof. First of all let us notice that when $f\left(t\right) = -\frac{1}{\alpha}t^\alpha$, the Legendre transform f^* of f, that is $f^* : \mathbb{R} \to \mathbb{R}$ defined by $f^*\left(h\right) = \sup_{x \geq 0}\left(x.h - f\left(x\right)\right) = \sup_{x \geq 0}\left(x.h + \frac{1}{\alpha}x^\alpha\right)$ is given by the following formula

$$f^*\left(h\right) = \begin{cases} +\infty & \text{if } h > 0, \\ \left(\frac{1}{\alpha} - 1\right)\left(-h\right)^{\frac{\alpha}{(\alpha - 1)}} & \text{if } h \leq 0. \end{cases} \tag{24}$$

We first exploit the relation between f and f^* to obtain some lower bounds. We note that the functional $\mathcal{F}$ can be represented as follows

$$\mathcal{F}\left(\varrho\right) = \int_{\mathbb{R}^n} \left(f\left(\varrho^{\mathrm{ac}}\left(x\right)\right) + f^*\left(h\left(x\right)\right) - \varrho^{\mathrm{ac}}\left(x\right)h\left(x\right)\right)\mathrm{d}x$$
$$+ \int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}\left(x\right)h\left(x\right)\mathrm{d}x - \int_{\mathbb{R}^n} f^*\left(h\left(x\right)\right)\mathrm{d}x \tag{25}$$

for any function h such that $h \in L^1\left(\varrho^{\mathrm{ac}}\right)$ and $f^*\left(h\right) \in L^1\left(\mathbb{R}^n\right)$.

Assuming $M_1(\varrho) < +\infty$, we can fix $\delta < 1$ and take $h\left(x\right) = -(1 + |x|)^\delta$. This guarantees $h \in L^1\left(\varrho^{\mathrm{ac}}\right)$; in order to have $f^*\left(h\right) \in L^1\left(\mathbb{R}^n\right)$ we need $\delta \in \left(\frac{n}{n+q-1}, 1\right)$. Indeed, using polar coordinates, we get

$$0 \leq \int_{\mathbb{R}^n} f^*\left(h\left(x\right)\right)\mathrm{d}x = \left(\frac{1}{\alpha} - 1\right)\int_{\mathbb{R}^n}\left(1 + |x|\right)^{\frac{-\delta\alpha}{1-\alpha}}\mathrm{d}x$$
$$= c\int_0^{+\infty} r^{n-1}(1 + r)^{\frac{-\delta\alpha}{1-\alpha}}\mathrm{d}r$$

and the integral is finite as soon as

$$n - 1 - \frac{\delta\alpha}{1-\alpha} < -1 \iff \delta > \frac{n}{n+q-1}. \tag{26}$$

Let us start now from the first part of the claim, i.e. the lower bound. The Young-Fenchel inequality implies $f(\varrho^{\mathrm{ac}}(x)) + f^*(h(x)) - \varrho^{\mathrm{ac}}(x) h(x) \geq 0$, hence we get

$$\mathcal{F}(\varrho) \geq \int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) h(x)\,\mathrm{d}x - \int_{\mathbb{R}^n} f^*(h(x))\,\mathrm{d}x. \tag{27}$$

The last term in this inequality is a constant independent of ϱ, and for the other term we use

$$\begin{aligned}
\int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) h(x)\,\mathrm{d}x &= - \int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) (1 + |x|)^\delta \mathrm{d}x \\
&\geq - \int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x)\,\mathrm{d}x - \int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) |x|^\delta \mathrm{d}x
\end{aligned} \tag{28}$$

Applying Hölder inequality with exponents $p_1 = \frac{1}{\delta}$ and $p_2 = \frac{1}{1-\delta}$, and using $\varrho^{\mathrm{ac}} \leq \varrho$ and $\int \varrho^{\mathrm{ac}} \leq 1$, we have

$$\int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) |x|^\delta \mathrm{d}x \leq \left(\int_{\mathbb{R}^n} |x|\,\mathrm{d}\varrho(x) \right)^\delta. \tag{29}$$

This provides

$$\int_{\mathbb{R}^n} \varrho^{\mathrm{ac}}(x) h(x)\,\mathrm{d}x \geq -1 - \left(\int_{\mathbb{R}^n} |x|\,\mathrm{d}\varrho(x) \right)^\delta \tag{30}$$

Finally, for each $\delta \in \left(\frac{n}{n+q-1}, 1 \right)$, the following estimate holds true

$$\mathcal{F}(\varrho) \geq -C(\delta) - \left(\int_{\mathbb{R}^n} |x|\,\mathrm{d}\varrho(x) \right)^\delta. \tag{31}$$

We now have to prove the lower semicontinuity. We start from (32) which we re-write as

$$\begin{aligned}
\mathcal{F}(\varrho) = \int_{\mathbb{R}^n} \Big(f(\varrho^{\mathrm{ac}}(x)) + f^*(h(x)) - \varrho^{\mathrm{ac}}(x) h(x) \Big)\mathrm{d}x - \int_{\mathbb{R}^n} h(x)\mathrm{d}\varrho^{\mathrm{sing}}(x) \\
+ \int_{\mathbb{R}^n} h(x)\mathrm{d}\varrho(x) - \int_{\mathbb{R}^n} f^*(h(x))\,\mathrm{d}x.
\end{aligned} \tag{32}$$

Again, the function h is fixed as $h(x) = (1 + |x|)^\delta$, choosing δ as above. We note that $\mathcal{F}$ is then composed of three parts. The last one is a constant, independent of ϱ. The previous one is

$$\varrho \mapsto \int h\mathrm{d}\varrho.$$

We cannot say that this functional is continuous for the weak convergence of ϱ since h is not bounded. This is why we consider a sequence ϱ_n with $M_1(\varrho_n) \leq C$, and we will exploit the sublinear behavior of h. If we fix an arbitrarily large constant M, we have

$$\left| \int h \mathrm{d}\varrho_n - \int h \mathrm{d}\varrho \right| \leq \left| \int \min\{h, M\} \mathrm{d}(\varrho_n - \varrho) \right| + \int_{\{x : h(x) > M\}} h \, \mathrm{d}(\varrho_n + \varrho).$$

Note that we have, for an arbitrary positive measure μ and $\delta < 1$,

$$\int_{\{x : h(x) > M\}} h^{\frac{1}{\delta}} \mathrm{d}\mu \geq M^{\frac{1}{\delta} - 1} \int_{\{x : h(x) > M\}} h \, \mathrm{d}\mu,$$

which, applied to $\mu = \varrho_n + \varrho$ and $h = (1 + |x|)^\delta$, gives

$$2 + 2C \geq \int (1 + |x|) \mathrm{d}(\varrho_n + \varrho) \geq M^{\frac{1}{\delta} - 1} \int_{\{x : h(x) > M\}} h \, \mathrm{d}(\varrho_n + \varrho).$$

Hence we get

$$\left| \int h \mathrm{d}\varrho_n - \int h \mathrm{d}\varrho \right| \leq \left| \int \min\{h, M\} \mathrm{d}(\varrho_n - \varrho) \right| + CM^{1 - \frac{1}{\delta}}.$$

Taking the limsup in n and using that $\min\{h, M\}$ is continuous and bounded we have

$$\limsup_n \left| \int h \mathrm{d}\varrho_n - \int h \mathrm{d}\varrho \right| \leq CM^{1 - \frac{1}{\delta}}.$$

Since M is arbitrary, this provides

$$\lim_n \int h \mathrm{d}\varrho_n = \int h \mathrm{d}\varrho.$$

Hence, this term in the expression of $\mathcal{F}$ is continuous for the convergence we use (weak convergence + bound on M_1). Finally, we just need to prove

$$\varrho \mapsto \int_{\mathbb{R}^n} \Big(f\left(\varrho^{\mathrm{ac}}\left(x\right)\right) + f^*\left(h\left(x\right)\right) - \varrho^{\mathrm{ac}}\left(x\right) h\left(x\right) \Big) \mathrm{d}x - \int_{\mathbb{R}^n} h(x) \mathrm{d}\varrho^{\mathrm{sing}}(x)$$

is l.s.c. for the weak convergence of probability measures. This functional is of the form

$$\int g(s, \varrho^{\mathrm{ac}}(x)) \mathrm{d}x + \int \lim_{t \to \infty} \frac{g(x, t)}{t} \mathrm{d}\varrho^{\mathrm{sing}}(x)$$

where $g(x, t) = f(t) + f^*(h(x)) - h(x)t$. On a set of finite measure, this functional would be l.s.c. because of general results on local functionals (see, for instance, [5]). Here the reference measure has not finite mass, but the integrand g is now positive, and hence the functional can be written as a supremum of local functionals of the same form on sets of finite measure, hence recovering lower semi-continuity. $\qquad\square$

Displacement convexity and strict displacement convexity. Another important property of $\mathcal{F}$ concerns displacement convexity, the notion introduced by McCann in [11] and already presented in this paper concerning $\mathcal{T}$. It is well-known from [11] that whenever $F(\varrho)$ has the the form $\int f(\varrho(x))\mathrm{d}x$, then displacement convexity is guaranteed as soon as $s \mapsto s^n f(s^{-n})$ is convex and decreasing, n being the dimension of the ambient space. For the functional $\mathcal{F}$, this condition is satisfied as soon as $\alpha > 1 - \frac{1}{n}$, i.e. for $q > 0$. We can now state the following proposition.

Proposition 2.3. *When restricted to $\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^n)$, the functional $\mathcal{F}$ is displacement convex in $\mathbb{W}_2$ and strictly convex on every geodesic $t \mapsto \varrho_t = ((1-t)\,\mathrm{id} + tT)_{\#}\varrho$ unless the optimal map T is such that $DT = I$ a.e. on $\{\varrho > 0\}$. Moreover, if $\varrho_t = ((1-t)\,\mathrm{id} + tT)_{\#}\varrho$, then the derivative at $t = 0$ of $t \mapsto \mathcal{F}(\varrho_t)$ is given by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathcal{F}(\varrho_t)\Big)_{|t=0} = \left(1 - \frac{1}{\alpha}\right)\int_{\mathbb{R}^n}\varrho^\alpha \mathrm{div}\,(T - \mathrm{Id})\,\mathrm{d}x. \tag{33}$$

In Formula (33), the divergence is to be taken in the a.e. sense, as T is countably Lipschitz. Using $T = \nabla v$ with v convex, this divergence is equal to $\Delta^{\mathrm{ac}}v - n$, where Δ^{ac} is the absolutely continuous part of the distributional Laplacian of v, which is a positive measure.

The proof of this statement follows the same line as that of points (3) and (4) in Proposition 2.1 of [15] (also see the computation in the Appendix A2 of [4]).

3 A Variational Principle for q-Moment Measures

As we sketched in the introduction, we consider the following variational problem. We fix $\mu \in \mathcal{P}_1(\mathbb{R}^n)$ with $\int y\mathrm{d}\mu(y) = 0$ and not supported on a hyperplane, and we want to solve

$$(P) \qquad \min\left\{\mathcal{J}(\varrho) = \mathcal{F}(\varrho) + \mathcal{T}(\varrho, \mu): \ \varrho \in \mathcal{P}_1(\mathbb{R}^n)\right\}.$$

3.1 Existence of minimizers

Theorem 3.1. *Problem (P) admits a solution.*

The idea for proving existence relies on the direct method in the calculus of variations, i.e. the fact that the minimized functional is lower semi-continuous and minimizing sequences are compact.

Proof of Theorem 3.1. Let $\varrho_k \in \mathcal{P}_1(\mathbb{R}^n)$ be a minimizing sequence, i.e. $\inf_{\varrho \in \mathcal{P}_1(\mathbb{R}^n)}\mathcal{J}(\varrho) = \lim_{k\to+\infty}\mathcal{J}(\varrho_k)$. We can suppose that all ϱ_k have 0 as their barycenter as translations do not change the value of the two parts of the functional and we may also assume that $\mathcal{J}(\varrho_k) \leq C_0 := \mathcal{J}(\varrho^0)$ with $\varrho^0 \in \mathcal{P}_1(\mathbb{R}^n)$. Using the estimates in Proposition 2.1-part (5) and Proposition 2.3-part (1), we have

$$\mathcal{J}(\varrho) \geq C_1 M_1(\varrho) - M_1(\varrho)^\delta - C_2.$$

This implies that $M_1(\varrho_k)$ must be bounded. According to Remark 5.1.5 in [1], this gives tightness of the sequence ϱ_k and hence we may extract a subsequence such that $\varrho_k \rightharpoonup \overline{\varrho}$. On the other hand, by Proposition 2.3-part (2) we know that the functional $\mathcal{F}$ is l.s.c for the weak convergence when coupled to a bound on M_1, and the semi-continuity of $\mathcal{T}$ along sequences with $\int x \mathrm{d}\varrho_k(x) = 0$ is in Proposition 2.1-part (3). These results imply

$$\mathcal{J}(\overline{\varrho}) \leq \liminf_{k \to +\infty} \mathcal{J}(\varrho_k) = \inf_{\varrho \in \mathcal{P}_1(\mathbb{R}^n)} \mathcal{J}(\varrho)$$

which means that the minimum is indeed attained at $\varrho_{\mathrm{opt}} = \overline{\varrho}$. $\square$

Remark 3.1. *Thanks to Theorem 3.2 in next section, each optimal solution to the problem (P) is absolutely continuous. Hence, we can also obtain uniqueness of the solution from strict displacement convexity, using the geodesic ϱ_t connecting two solutions ϱ_0 and ϱ_1 (see Proposition 2.3-part (3); the case $DT = I$ a.e on $\{\varrho_0 > 0\}$ can be excluded exactly as in Lemma 4.2 in [15]).*

3.2   Properties of the optimal solutions

The main goal of this section is to prove that each solution of the problem (P) is absolutely continuous and show that if $\overline{u}$ is solution to the minimization problem

$$\inf \left\{ \int_{\mathbb{R}^n} u \mathrm{d}\varrho_{\mathrm{opt}} + \int_{\mathbb{R}^n} u^* \mathrm{d}\mu \ \mid \ u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \ \text{ convex and l.s.c} \right\} \tag{34}$$

then there is a constant c such that $\overline{u} > c$ and $\varrho_{\mathrm{opt}} = (\overline{u} - c)^{-(n+q)}$. Finally, we will also prove that $\overline{u}$ is an essentially continuous convex function.

To achieve our goals, we need some lemmas as follows.

Lemma 3.1. *Given two densities $\sigma, \sigma_1 \in L^1(\mathbb{R}^n)$, with $\int_{\mathbb{R}^n} f(\sigma) > -\infty$, we have*

$$\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} \frac{f(\sigma + \varepsilon(\sigma_1 - \sigma)) - f(\sigma)}{\varepsilon} \mathrm{d}x = \int_{\mathbb{R}^n} f'(\sigma)(\sigma_1 - \sigma) \mathrm{d}x.$$

Proof. By convexity of the function f, the inequality

$$\int_{\mathbb{R}^n} \frac{f(\sigma + \varepsilon(\sigma_1 - \sigma)) - f(\sigma)}{\varepsilon} \mathrm{d}x \geq \int_{\mathbb{R}^n} f'(\sigma)(\sigma_1 - \sigma) \mathrm{d}x \tag{35}$$

is straightforward. Note that the right-hand side above could possibly be equal to $-\infty$. For the opposite inequality, we will use Fatou's Lemma. The pointwise convergence of the integrand is trivial and we can get an upper bound by means of the inequality

$$\frac{f(\sigma + \varepsilon(\sigma_1 - \sigma)) - f(\sigma)}{\varepsilon} \leq f(\sigma_1) - f(\sigma) \leq -f(\sigma),$$

valid for $\varepsilon < 1$ (in the last inequality we used $f \leq 0$). Then, since we supposed $f(\sigma) \in L^1$, we can apply Fatou's Lemma and we get

$$\limsup_{\varepsilon \to 0} \int_{\mathbb{R}^n} \frac{f(\sigma + \varepsilon(\sigma_1 - \sigma)) - f(\sigma)}{\varepsilon} \mathrm{d}x \leq \int_{\mathbb{R}^n} f'(\sigma)(\sigma_1 - \sigma) \mathrm{d}x. \tag{36}$$

Combining (35) and (36), the lemma is proven. $\square$

Lemma 3.2. *If $\overline{\varrho}$ is a solution to the problem (P), $\overline{\varrho} = \overline{\varrho}^{\mathrm{ac}}\mathcal{L}^n + \overline{\varrho}^{\mathrm{sing}}$ and $\overline{u} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is solution to (34), then*

$$\int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\overline{\varrho}^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \ \leq \ \int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\varrho^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\varrho^{\mathrm{sing}} \qquad (37)$$

for all $\varrho \in \mathcal{P}_1\left(\mathbb{R}^n\right)$ written as $\varrho = \varrho^{\mathrm{ac}}\mathcal{L}^n + \varrho^{\mathrm{sing}}$.

Proof. Assume that $\overline{\varrho}$ is optimal and $\overline{u}$ is a convex function realizing the minimum in the dual definition of $\mathcal{T}\left(\overline{\varrho}, \mu\right)$, then the functional

$$\varrho \mapsto \int_{\mathbb{R}^n} f\left(\varrho^{\mathrm{ac}}\right) \mathrm{d}x + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\varrho = \int_{\mathbb{R}^n} f\left(\varrho^{\mathrm{ac}}\right) \mathrm{d}x + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\varrho^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\varrho^{\mathrm{sing}}$$

is minimal for $\varrho = \overline{\varrho}$. Now for every $\varepsilon \in (0,1)$ and any $\varrho \in \mathcal{P}_1\left(\mathbb{R}^n\right)$, we define $\varrho_\varepsilon := (1 - \varepsilon)\overline{\varrho} + \varepsilon\varrho$, $\Phi_1\left(\varrho^{\mathrm{ac}}\right) = \int f\left(\varrho^{\mathrm{ac}}\right)\mathrm{d}x$, $\Phi_2\left(\varrho^{\mathrm{ac}}\right) = \int \overline{u}\mathrm{d}\varrho^{\mathrm{ac}}$ and $\Phi_3\left(\varrho^{\mathrm{sing}}\right) = \int \overline{u}\mathrm{d}\varrho^{\mathrm{sing}}$. Using the optimality of $\overline{\varrho}$, we have

$$\lim_{\varepsilon \to 0} \frac{\left\{\left(\Phi_1 + \Phi_2\right)\left(\varrho_\varepsilon^{\mathrm{ac}}\right) - \left(\Phi_1 + \Phi_2\right)\left(\overline{\varrho}^{\mathrm{ac}}\right)\right\} + \left\{\Phi_3\left(\varrho_\varepsilon^{\mathrm{sing}}\right) - \Phi_3\left(\overline{\varrho}^{\mathrm{sing}}\right)\right\}}{\varepsilon} \geq 0. \qquad (38)$$

On the other hand,

$$\begin{cases} \dfrac{\Phi_2\left(\varrho_\varepsilon^{\mathrm{ac}}\right) - \Phi_2\left(\overline{\varrho}^{\mathrm{ac}}\right)}{\varepsilon} & = \displaystyle\int_{\mathbb{R}^n} \overline{u}\mathrm{d}\left(\varrho^{\mathrm{ac}} - \overline{\varrho}^{\mathrm{ac}}\right) \\[3mm] \dfrac{\Phi_3\left(\varrho_\varepsilon^{\mathrm{sing}}\right) - \Phi_3\left(\overline{\varrho}^{\mathrm{sing}}\right)}{\varepsilon} & = \displaystyle\int_{\mathbb{R}^n} \overline{u}\mathrm{d}\left(\varrho^{\mathrm{sing}} - \overline{\varrho}^{\mathrm{sing}}\right) \end{cases} \qquad (39)$$

and by Lemma 3.1,

$$\lim_{\varepsilon \to 0} \frac{\Phi_1\left(\varrho_\varepsilon^{\mathrm{ac}}\right) - \Phi_1\left(\overline{\varrho}^{\mathrm{ac}}\right)}{\varepsilon} = \int_{\mathbb{R}^n} f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\mathrm{d}\left(\varrho^{\mathrm{ac}} - \overline{\varrho}^{\mathrm{ac}}\right). \qquad (40)$$

Combining (38), (39) and (40), we obtain (37). $\qquad\qquad\square$

Lemma 3.3. *The solution $\overline{\varrho}$ to the problem (P) cannot be such that $\overline{\varrho}^{\mathrm{ac}} = 0$ a.e.*

Proof. Indeed, assume that $\overline{\varrho}^{\mathrm{ac}} = 0$, then $\overline{\varrho} = \overline{\varrho}^{\mathrm{sing}}$, $\mathcal{F}\left(\overline{\varrho}\right) = \mathcal{F}\left(\overline{\varrho}^{\mathrm{sing}}\right) = 0$ and hence

$$\mathcal{J}\left(\overline{\varrho}\right) = \mathcal{T}\left(\overline{\varrho}, \mu\right) \geq 0.$$

On the other hand, for $\eta > 0$, we take $\varrho^\eta = \frac{\mathbf{1}_{B(0,\eta)}}{|B(0,\eta)|}\mathcal{L}^n$. Then $\varrho^\eta \in \mathcal{P}_1\left(\mathbb{R}^n\right)$ and

$$\mathcal{T}\left(\varrho^\eta, \mu\right) = \max\left\{\int_{\mathrm{spt}(\varrho^\eta) \times \mathbb{R}^n} x.y\mathrm{d}\gamma : \gamma \in \Gamma\left(\mu, \varrho^\eta\right)\right\} \leq \eta \int_{\mathbb{R}^n} |y|\,\mathrm{d}\mu =: C\eta$$

and value of the local functional $\mathcal{F}\left(\varrho^\eta\right) = -\frac{|B(0,\eta)|}{\alpha|B(0,\eta)|^\alpha} = -\alpha^{-1}|B(0,\eta)|^{1-\alpha} = -C_1\eta^{n(1-\alpha)}$. It follows that

$$\mathcal{J}\left(\varrho^\eta\right) = C\eta - C_1\eta^{n(1-\alpha)}.$$

Thanks to the assumption $\alpha > 1 - \frac{1}{n}$ (corresponding to $q > 0$), the above quantity can be made strictly negative as soon as η is small enough, which gives a contradiction to the optimality of $\overline{\varrho}$. $\qquad\qquad\square$

Theorem 3.2. *If $\overline{\varrho} = \varrho_{\mathrm{opt}}$ is a solution to the problem (P) and $\overline{u}$ is a solution to (34), then there exists a constant c such that $\overline{u} \geq c$ and*

$$\begin{cases} \overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) = c \ \text{ a.e on } \ \{\overline{\varrho}^{\mathrm{ac}} > 0\} \\ \\ \overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) \geq c \ \text{ a.e on } \ \{\overline{\varrho}^{\mathrm{ac}} = 0\} \end{cases} \tag{41}$$

and $\overline{\varrho}^{\mathrm{sing}}$ is concentrated on $\{\overline{u} = c\}$. Moreover, the set $\{\overline{u} = c\}$ is empty and thus $\overline{\varrho}^{\mathrm{sing}} = 0$, the optimal $\overline{\varrho}$ is absolutely continuous, with bounded density given by

$$\overline{\varrho} = \overline{\varrho}^{\mathrm{ac}} = \frac{1}{\left(\overline{u} - c\right)^{n+q}}. \tag{42}$$

Proof. Define the essential infimum

$$c = \operatorname{ess\,inf}\left\{\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right\} := \sup\left\{\ell \ : \ \mathcal{L}^n\left(\{x : \overline{u}\left(x\right) + f'\left(\overline{\varrho}^{\mathrm{ac}}\left(x\right)\right) < \ell\}\right) = 0\right\}.$$

Then, we have $\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) \geq c$ a.e, and so $\overline{u} \geq \overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) \geq c$ a.e (since $f' \leq 0$). Hence we have

$$\int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\overline{\varrho}^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \geq \int_{\mathbb{R}^n} c\mathrm{d}\overline{\varrho}^{\mathrm{ac}} + \int_{\mathbb{R}^n} c\mathrm{d}\overline{\varrho}^{\mathrm{sing}} = c. \tag{43}$$

Take any $c' > c$. By definition of essential infimum, the set $\{\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) < c'\}$ has positive Lebesgue measure, and so we can choose $\varrho \in \mathcal{P}_1^{\mathrm{ac}}\left(\mathbb{R}^n\right)$ such that ϱ concentrated on the set $\{\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) < c'\}$. Then we get

$$\int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\varrho^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\varrho^{\mathrm{sing}} = \int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\varrho < c'. \tag{44}$$

Combining (37) in Lemma 3.2 and (44), we obtain

$$c' > \int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\overline{\varrho}^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \tag{45}$$

Letting $c' \to c$ in (45), we get

$$c \geq \int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\overline{\varrho}^{\mathrm{ac}} + \int_{\mathbb{R}^n} \overline{u}\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \tag{46}$$

From (43) and (46), we deduce that

$$\begin{cases} \displaystyle\int_{\mathbb{R}^n} \left(\overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right)\right) \mathrm{d}\overline{\varrho}^{\mathrm{ac}} = \int_{\mathbb{R}^n} c\mathrm{d}\overline{\varrho}^{\mathrm{ac}} \\ \displaystyle\int_{\mathbb{R}^n} \overline{u}\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \qquad\quad = \int_{\mathbb{R}^n} c\mathrm{d}\overline{\varrho}^{\mathrm{sing}} \end{cases}$$

It follows that

$$\begin{cases} \overline{u} + f'\left(\overline{\varrho}^{\mathrm{ac}}\right) = c \ \ \overline{\varrho}^{\mathrm{ac}}-\text{a.e} \\ \overline{u} \qquad\qquad\quad = c \ \ \overline{\varrho}^{\mathrm{sing}}-\text{a.e} \end{cases}$$

This implies that (41) is satisfied and $\overline{\varrho}^{\text{sing}}$ is concentrated on $\{\overline{u} = c\}$.

Setting $\varphi := \overline{u} - c$ we would like to prove that the set $\{\varphi = 0\}$ is empty.

First, let us note that the interior of $\{\varphi < +\infty\}$ is not empty. Indeed, should it be empty, since the set $\{\varphi < +\infty\}$ is a convex set, then it would be negligible, which means that in this case $\varphi = +\infty$ a.e. and hence $\overline{\varrho} = 0$, which is impossible. Thus the interior of $\{\varphi < +\infty\}$ has to be non-empty. Next we suppose that there exists a point $x_0 \in \mathbb{R}^n$ such that $\varphi(x_0) = 0$. First, we exclude the case where x_0 belongs to the interior of $\{\varphi < +\infty\}$. Indeed, in this case there exists a neighborhood of x_0, denoted by $\mathcal{N}(x_0) \left(\subset \{\varphi < +\infty\} \right)$ where φ is locally Lipschitz, so that we have $\varphi \le L |x - x_0|$ for $x \in \mathcal{N}(x_0)$, which implies

$$1 > \int_{\mathcal{N}(x_0)} \overline{\varrho} \mathrm{d}x = \int_{\mathcal{N}(x_0)} \varphi^{-(n+q)} \mathrm{d}x \ge \int_{\mathcal{N}(x_0)} \left(L |x - x_0| \right)^{-(n+q)} \mathrm{d}x = +\infty.$$

This is impossible.

The case where x_0 lies on the boundary of the set $\{\varphi < +\infty\}$ is more subtle. In this case we choose n points $x_1, x_2, ..., x_n$ in the interior of $\{\varphi < +\infty\}$ so as to built a symplex Δ whose $(n+1)$ vertices are $x_0, x_1, x_2, ...$ and x_n, such that its interior is non-empty. On this symplex, the convex function φ is finite and satisfies an inequality of the form $\varphi \le L |x - x_0|$ for all $x \in \Delta$, where the constant L depends on $(x_i)_i$ and $(\varphi(x_i))_i$. Then, we find a similar contradiction as in the previous case.

We conclude that we cannot have $x_0 \in \mathbb{R}^n$ satisfying $\varphi(x_0) = 0$. Since φ is l.s.c. and we have $\lim_{|x| \to \infty} \varphi(x) = +\infty$ (this is a consequence of the integrability of $\overline{\varrho} = \varphi^{-(n+q)}$, then it admits a minimum on the space $\mathbb{R}^n$, and this minimum should be strictly positive. This proves the boundedness of $\overline{\varrho}$. $\qquad\square$

We have now proven that the optimal $\overline{\varrho}$ can be expressed as $\overline{\varrho} = \varphi^{-(n+q)}$, and in order to fit the theory of Klartag we just need to prove that φ is essentially continuous. This means that we want to prove $\lim_{x \to x_0} \varphi(x) = +\infty$ for $\mathcal{H}^{d-1}$-a.e. $x_0 \in \partial\{\varphi < +\infty\}$.

In order to do this, given an optimal solution $\overline{\varrho}$, we choose a precise representative of it, and more precisely we take $\overline{\varrho} = \varphi^{-(n+q)}$, with φ convex and l.s.c. We need to prove that $\overline{\varrho}$ vanishes on almost every point of the boundary.

Theorem 3.3. *Let $\overline{\varrho}$ be the precise representative above of a solution. Set $\Omega = \{\varphi < +\infty\}$. Then $\overline{\varrho} = 0$ holds $\mathcal{H}^{d-1}$-a.e. on $\partial\Omega$.*

Proof. The proof will strongly follow that of Theorem 4.3 of [15].

Suppose $\overline{\varrho} > 0$ on a set of positive $\mathcal{H}^{d-1}$ measure on $\partial\Omega$. Writing locally Ω as $\{(x_1, x') : x_1 > h(x')\}$ we assume that this set is given by $A = \{(x_1, x') \in \mathbb{R} \times \mathbb{R}^{d-1} \, x' \in B, x_1 = h(x')\}$, where $B \subset \mathbb{R}^{d-1}$. Up to reducing the sets A and B, we can suppose $\inf_{x \in A} \overline{\varrho}(x) > 0$ for $x \in A$ and that A, B are compact. We define $A_\varepsilon := \{(x_1, x') \in \mathbb{R} \times \mathbb{R}^{d-1} \, x' \in B, x_1 \in [h(x'), h(x') + \varepsilon]\}$. Then, by the continuity of $\overline{\varrho}$ inside Ω, we can also assume, for small $\varepsilon > 0$, that we have $\overline{\varrho}(x_1, x') \ge c > 0$ for $(x_1, x') \in A_\varepsilon$. From the fact that $\overline{\varrho}$ is bounded, we also have the opposite inequality $\overline{\varrho}(A_\varepsilon) \le C_1 \varepsilon$.

Now, we define a new density ϱ_ε as a competitor by taking $T : A_\varepsilon \to \mathbb{R}^d$ by $T(x_1, x') = (x_1 - \varepsilon, x')$ and setting

$$\varrho_\varepsilon = \bar\varrho \llcorner (A_\varepsilon^c) + \frac{1}{2}\bar\varrho \llcorner (A_\varepsilon) + \frac{1}{2}T_\#(\bar\varrho \llcorner (A_\varepsilon)).$$

By computing the density of ϱ_ε we can check

$$\mathcal{F}(\varrho_\varepsilon) = \mathcal{F}(\bar\varrho) - (2^{\frac{1}{\alpha}} - 1)\int_{A_\varepsilon} \bar\varrho\,(x)^\alpha\,\mathrm{d}x \le \mathcal{F}(\bar\varrho) - C_2\varepsilon.$$

In order to estimate $\mathcal{T}(\varrho_\varepsilon, \mu)$, take the optimal function φ (realizing $\mathcal{T}(\bar\varrho, \mu) = \int \varphi\,\mathrm{d}\bar\varrho + \int \varphi^*\,\mathrm{d}\mu$) and modify it into a function φ_δ as follows: take a convex, positive and superlinear function $\chi : \mathbb{R}^d \to \mathbb{R}$ with $\int \chi(x)\,\mathrm{d}\mu(x) < +\infty$ (which exists because $\mu \in \mathcal{P}_1(\mathbb{R}^d)$), choose $\delta > 0$ and set $\varphi_\delta = (\varphi^* + \delta\chi)^*$. We have

$$\varphi_\delta(x) \le \varphi(x) \text{ for every } x \in \Omega \quad \varphi_\delta(T(x)) \le \varphi(x) + \delta\chi^*\left(\frac{\varepsilon e_1}{\delta}\right) \text{ for every } x \in A_\varepsilon$$

(where e_1 denotes te first vector of the canonical basis $e_1 = (1, 0, \dots, 0)$). Hence,

$$\mathcal{T}(\varrho_\varepsilon, \mu) \le \int \varphi_\delta\,\mathrm{d}\varrho_\varepsilon + \int (\varphi^* + \delta\chi)\,\mathrm{d}\mu \le \mathcal{T}(\bar\varrho, \mu) + \frac{1}{2}\bar\varrho(A_\varepsilon)\delta\chi^*\left(\frac{\varepsilon e_1}{\delta}\right) + \delta\int \chi\,\mathrm{d}\mu.$$

The optimality of $\bar\varrho$ compared to ϱ_ε provides

$$C_2\varepsilon \le \frac{1}{2}\bar\varrho(A_\varepsilon)\delta\chi^*\left(\frac{\varepsilon e_1}{\delta}\right) + \delta\int \chi\,\mathrm{d}\mu.$$

Now, use $\bar\varrho(A_\varepsilon) \le C_1\varepsilon$ and choose $\delta = c\varepsilon$: we obtain, after dividing by ε,

$$C_2 \le \frac{C_1 c}{2}\varepsilon\delta\chi^*\left(\frac{e_1}{c}\right) + c\int \chi\,\mathrm{d}\mu.$$

If we choose c small enough, such that $c\int \chi\,\mathrm{d}\mu < \frac{1}{2}C_2$ we obtain a contradiction as $\varepsilon \to 0$. $\quad\square$

3.3 Sufficient optimality conditions

To complete the current study, it remains to prove that every density $\varrho = \varphi^{-(n+q)}$ such that $(\nabla\varphi)_\#\varrho = \mu$ and $\varphi : \mathbb{R}^n \to (0, +\infty)$ is an essentially continuous convex function is necessarily a minimizer of $\mathcal{J}(\varrho)$. This would explain that the variational principle of the previous section finds exactly all the desired functions φ.

First of all, let us remind that for ϱ to be integrable on $\mathbb{R}^n$ it is necessary that the convex function φ satisfies $\lim_{|x|\to\infty} \varphi(x) = +\infty$ and that, because of convexity, the growth at infinity should be at least linear. Moreover, to be locally integrable around each point, the same computations as those that we showed at the end of the proof of Theorem 3.2 prove that φ should be bounded from below by a strictly positive constant. In particular, ϱ mus be a bounded density.

The result that we will prove is the following.

Theorem 3.4. *Let $\varphi : \mathbb{R}^n \to (0, +\infty)$ be an essentially continuous convex function, consider $\overline{\varrho} = \varphi^{-(n+q)}$ and suppose that $\mu = (\nabla\varphi)_{\#}\overline{\varrho}$. Then $\overline{\varrho} \in \mathcal{P}_1(\mathbb{R}^n)$ and $\mathcal{J}(\overline{\varrho}) = \min_{\varrho \in \mathcal{P}_1(\mathbb{R}^n)} \mathcal{J}(\varrho)$.*

Exactly as in the case of moment measures ([15]), the ideas to deal with the sufficient conditions come from displacement convexity (as it was the case in [4]). Take an arbitrary ϱ with compact support, and the geodesic curve $\varrho_t = ((1 - t)\,\mathrm{Id} + tT)_{\#}\overline{\varrho}$, where $T = \nabla v$ is the optimal transport from $\overline{\varrho}$ to ϱ. Then we know, from displacement convexity, that the following inequality holds true

$$\mathcal{J}(\varrho) \ - \ \mathcal{J}(\overline{\varrho}) \ \geq \ \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathcal{J}(\varrho_t)\Big)_{|t=0}$$

It is sufficient to show that the derivative in the right hand side above is non-negative. Using Proposition 2.1-part (6) and Proposition 2.3-part (4) we have

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\mathcal{F}(\varrho_t)\Big)_{|t=0} & = \left(1 - \dfrac{1}{\alpha}\right) \displaystyle\int_{\mathbb{R}^n} \overline{\varrho}^{\,\alpha}\mathrm{div}\,(T - \mathrm{Id})\,\mathrm{d}x \\[2mm] \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\mathcal{T}(\varrho_t, \mu)\Big)_{|t=0} & \geq \displaystyle\int_{\mathbb{R}^n} (T(x) - x).\nabla\varphi(x)\,\overline{\varrho}(x)\,\mathrm{d}x. \end{cases} \tag{47}$$

After considering ϱ with compact support, we can use part (5) of Proposition 2.3 to show that the optimality of $\overline{\varrho}$ is also valid when compared to non-compactly supported measures, by approximation.

The arguments to be used to prove non-negativity of the derivative are very similar to the case of moment measures ([15], Proposition 5.1), up to some modifications. The reader is invited to compare to [15], and also to [7]. In particular, we first treat the terms involving the identity map in the above derivatives. This requires a computation similar to that in [7], Lemma 5, i.e. the inequality

$$\frac{n}{n+q-1}\int_{\mathbb{R}^n} \frac{\mathrm{d}x}{(\varphi(x))^{n+q-1}} \geq \int_{\mathbb{R}^n} x.\nabla\varphi(x)\,\mathrm{d}\overline{\varrho}(x) \tag{48}$$

This is were we need the convex function φ to be essentially continuous, as this guarantees $\int_{\mathbb{R}^n} \nabla\left(\dfrac{1}{\varphi^{n+q-1}}\right)(x)\,\mathrm{d}x = 0$ along the lines of ([7], Lemma 4).

Lemma 3.4. *The inequality (48) holds for $\overline{\varrho}$ and φ as in Theorem 3.4.*

Proof. The computation is very similar to that developed for moment measures in [7], Lemma 5. Thanks to the essentially-continuous property of the function φ, we have

$$\int_{\mathbb{R}^n} \nabla\varphi(x)\,\mathrm{d}\overline{\varrho}(x) \ = \ \frac{-1}{n+q-1}\int_{\mathbb{R}^n} \nabla\left(\frac{1}{\varphi^{n+q-1}}\right)(x)\,\mathrm{d}x \ = \ 0.$$

Pick now a point x_0 in the interior of $\{\varphi < +\infty\}$ and define $\mathcal{K}$ to be the class of all convex, smooth, compact sets K, contained in the interior of $\{\varphi < +\infty\}$ and containing x_0 in their

interior. Using the fact that the function $x \mapsto \nabla\varphi(x) \cdot (x - x_0)$ is bounded from below (by $\inf \varphi - \varphi(x_0)$), we get

$$\int_{\mathbb{R}^n} x.\nabla\varphi(x)\,\mathrm{d}\overline{\varrho}(x) \;=\; \int_{\mathbb{R}^n} \nabla\varphi(x).(x - x_0)\,\mathrm{d}\overline{\varrho}(x) \leq \sup_{K \in \mathcal{K}} \left\{ \int_K \nabla\varphi(x).(x - x_0)\,\mathrm{d}\overline{\varrho}(x) \right\}$$

We then observe that, integrating by parts, we have

$$\int_K \nabla\varphi(x).(x - x_0)\,\mathrm{d}\overline{\varrho}(x) = \frac{n}{n + q - 1} \int_K \frac{\mathrm{d}x}{\varphi^{n+q-1}} - \frac{1}{n + q - 1} \int_{\partial K} \frac{(x - x_0).\mathbf{n}_x}{\varphi^{n+q-1}}\mathrm{d}\mathcal{H}^{n-1}(x).$$

The second term of the right hand side is negative, using the inequality $(x - x_0) \cdot \mathbf{n}_x \geq 0$, valid for K convex, $x_0 \in K$ and $x \in \partial K$, and the first can be estimated by the integral on the whole space, which gives the claim. $\qquad\square$

Proof of Theorem 3.4. Similar to Proposition 5.1 in [15], using (47) and (48) above, we obtain

$$\mathcal{J}(\varrho) - \mathcal{J}(\overline{\varrho}) \geq \frac{-1}{n + q - 1} \int_{\mathbb{R}^n} (\Delta^{\mathrm{ac}}v)\,\overline{\varrho}^{\alpha}\mathrm{d}x + \int_{\mathbb{R}^n} \nabla v(x).\nabla\varphi(x)\,\overline{\varrho}(x)\,\mathrm{d}x. \qquad (49)$$

Note that we suppose that ϱ is compactly supported, which implies $|\nabla v| \leq C$. We will prove that the right hand side is positive, exactly as in Proposition 5.1 in [15], by first using $\Delta^{\mathrm{ac}}v \leq \Delta v$ (where Δv is the distributional derivative of v). By integrating by parts we have

$$\int_{B(0,R)} \Delta^{\mathrm{ac}}v \leq \int_{B(0,R)} \Delta v \leq \int_{\partial B(0,R)} |\nabla v| \leq CR^{n-1}.$$

Using $\varrho \leq cR^{-(n+q)}$ for large R, this implies $\int \Delta^{\mathrm{ac}}v\overline{\varrho} < \infty$, since the integral on the annulus $B(0, 2^k) \setminus B(0, 2^{k-1})$ can be estimated by $C(2^k)^{n-1-(n+q)}$, which is summable. This shows that the first integral in (49) is well-defined and can be approximated with integrals on finite balls. To do the same for the second integral, we observe that we have $\nabla v \in L^\infty$ and $\nabla\varphi \in L^1(\overline{\varrho})$ since this is equivalent to $M_1(\mu) < +\infty$, which gives $\nabla v \cdot \nabla\varphi\overline{\varrho} \in L^1(\mathbb{R}^n)$.

Hence, we have

$$\mathcal{J}(\varrho) - \mathcal{J}(\overline{\varrho})$$

$$\geq \lim_{R \to +\infty} \left(\frac{-1}{n + q - 1} \int_{B(0,R)} (\Delta v)\,\overline{\varrho}^{\alpha}\mathrm{d}x + \int_{B(0,R)} \nabla v(x).\nabla\varphi(x)\,\overline{\varrho}(x)\,\mathrm{d}x \right)$$

$$= \lim_{R \to +\infty} \left(\frac{-1}{n + q - 1} \int_{B(0,R)} (\Delta v)\,\overline{\varrho}^{\alpha}\mathrm{d}x + \int_{B(0,R)} \nabla v(x).\nabla_x \left(\frac{-1}{n + q - 1}.\varphi^{-(n+q-1)} \right)(x)\,\mathrm{d}x \right)$$

$$= \frac{-1}{n + q - 1}.\lim_{R \to +\infty} \left(\int_{\partial B(0,R)} \nabla v(x).\mathbf{n}\big(\varphi(x)\big)^{-(n+q-1)}\mathrm{d}\mathcal{H}^{n-1}(x) \right) = 0,$$

where the fact that the last integral tends to 0 is justified using $|\nabla v| \leq C$ and bounding it from above by $CR^{n-1}R^{-(n+q-1)}$, which tends to 0 thanks to $q > 0$.

$\qquad\square$

References

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures.* Second edition. Lecture notes in Mathematics ETH Zürich. Birkhäuser Verlag, Bassel, (2008).

[2] S. Artstein-Avidan, B. Klartag and M. Vitali, The Santaló point of a function, and a functional form of the Santaló inequality, *Mathematika*, 51:33−48, 2

[3] Robert J. Berman, B. Berndtsson, Real Monge-Ampère equations and Kähler-Ricci solitons on toric log Fano varieties. *Ann. Fac. Sci. Toulouse Math.* (6), Vol. 22, No. 4, (2013), 649−711.

[4] A. Blanchet, P. Mossay, F. Santambrogio Existence and uniqueness of equilibrium for a spatial model of social interactions, *Int. Econ. Rev.*, Vol 57, Issue 1, 31–60, 2016.

[5] G. Bouchitté, G. Buttazzo, New lower semicontinuity results for nonconvex functionals defined on measures. *Nonlinear Analysis*, 15:679−692, 1990.

[6] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.* 305 (19), 805−808, 1987.

[7] D. Cordero-Erausquin, B. Klartag, Moment measures, *J. Functional Analysis*, Vol. 268, No. 12, (2015), 3834−3866.

[8] A. Figalli. *The Monge-Ampère equation and its applications.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2017. x+200

[9] R. Jordan, D. Kinderlehrer, F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. An.*, 1998.

[10] B. Klartag, Affine hemispheres of elliptic type, *Algebra i Analiz*, 29:145−188, 2017.

[11] R. J. McCann, A convexity principle for interacting gases. *Adv. Math.* 128 (1) 153−159, 1997.

[12] M. Meyer, A. Pajor, On the Blaschke-Santaló inequality. *Arch. Math.*, Vol 55, 82−93 (1990).

[13] R. T. Rockafellar, *Convex Analysis,* Princeton University Press, 1970.

[14] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, book, dans Progress in Nonlinear Differential Equations and Their Applications 87, Birkhäuser Basel (2015).

[15] F. Santambrogio, Dealing with moment measures via entropy and optimal transport, *Journal of Functional Analysis,* 271 (2016), no. 2, 418−436.

[16] F. Santambrogio, {Euclidean, Metric, and Wasserstein} Gradient Flows: an overview, *Bulletin of Mathematical Sciences 7*, 87−154, 2017.

[17] C. Villani, *Topics in Optimal Transportation*, no. 58, American Mathematical Soc., 2003.

[18] C. Villani, *Optimal Transport. Old and New.* Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.