Optimal Partial Mass Transportation and Obstacle Monge–Kantorovich Equation

Noureddine Igbida[†] and Van Thanh Nguyen[‡]

Institut de recherche XLIM-DMI, UMR-CNRS 6172
Faculté des Sciences et Techniques
Université de Limoges, France

Abstract. Optimal partial mass transport, which is a variant of the optimal transport problem, consists in transporting effectively a prescribed amount of mass from a source to a target. The problem was first studied by Caffarelli and McCann [7, Ann. of Math., 2010] and Figalli [15, Arch. Ration. Mech. Anal., 2010] with a particular attention to the quadratic cost. Our aim here is to study the optimal partial mass transport problem with Finsler distance costs including the Monge cost given by the Euclidian distance. Our approach is different and our results do not follow from previous works. Among our results, we introduce a PDE of Monge–Kantorovich type with a double obstacle to characterize active submeasures, Kantorovich potential and optimal flow for the optimal partial transport problem. This new PDE enables us to study the uniqueness and monotonicity results for the active submeasures. Another interesting issue of our approach is its convenience for numerical analysis and computations that we develop in a separate paper [17, IMA J. Numer. Anal., 2017].

1. Introduction

The *Monge–Kantorovich (MK) problem* (or optimal transport) aims to find the best way to move all mass from a given source into a prescribed target. The source and target are modelled by two finite Radon measures $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ with $\mu(\mathbb{R}^N) = \nu(\mathbb{R}^N)$ and the problem can be written as

$$\min_{\gamma \in \mathcal{M}_b^+(\mathbb{R}^N \times \mathbb{R}^N)} \left\{ \mathcal{K}(\gamma) := \int_{\mathbb{R}^N \times \mathbb{R}^N} c(x,y) \mathrm{d}\gamma : \pi_x \# \gamma = \mu, \pi_y \# \gamma = \nu \right\},$$

Key words and phrases. Optimal transport, optimal partial transport, Monge–Kantorovich problem, minimum flow problem, obstacle Monge–Kantorovich equation, tangential gradient, nonlinear PDE.

[†]E-mail : `noureddine.igbida@unilim.fr`

[‡]E-mail : `van-thanh.nguyen@unilim.fr`.

where $\pi_x\#\gamma$ and $\pi_y\#\gamma$ are marginals of γ (see Section 2) and $c(x,y)$ is a given ground cost, i.e., $c(x,y)$ is the cost one pays to move one unit of the material from x to y. This problem and related topics are central in the optimal transport theory which has become popular in the last few decades with applications in PDEs, differential geometry, image processing and many other areas. For a further discussion on the MK problem, its history and applications, we refer to pedagogical books [1], [23], [24] and [25].

The *partial Monge–Kantorovich (PMK) problem* (or optimal partial transport) is a very natural extension of the original optimal transport problem. The problem aims to study the case where only a part of the commodity (respectively, consumer demand) of total mass $\mathbf{m}$ needs to be transported (respectively, fulfilled). Given $\mu,\nu\in\mathcal{M}_b^+(\mathbb{R}^N)$ and a prescribed total mass $0\le\mathbf{m}\le\min\left\{\mu(\mathbb{R}^N),\nu(\mathbb{R}^N)\right\}$, the PMK problem reads as follows

$$\min_{\gamma\in\mathcal{M}_b^+(\mathbb{R}^N\times\mathbb{R}^N)}\left\{\mathcal{K}(\gamma):=\int\limits_{\mathbb{R}^N\times\mathbb{R}^N}c(x,y)\mathrm{d}\gamma:\ \pi_x\#\gamma\le\mu,\ \pi_y\#\gamma\le\nu,\ \gamma(\mathbb{R}^N\times\mathbb{R}^N)=\mathbf{m}\right\}.$$

This generalized problem brings out new unknown quantities $\rho_0:=\pi_x\#\gamma$ and $\rho_1:=\pi_y\#\gamma$. Here, ρ_0 and ρ_1 are the sites where the commodity is taken and the consumer demand is fulfilled, respectively. Like the full transfer case (see e.g. [25, Theorem 4.1]), the existence of optimal plan is guaranteed for a lower semicontinuous cost c by the direct method. In Caffarelli and McCann [7], the uniqueness is shown in the case of disjoint supports for cost c satisfying the so-called *left twist condition*, which includes a typical example $c(x,y):=h(x-y)$ with a strictly convex function h. Figalli [15] improves the uniqueness result by removing the disjointness assumption on the supports. For the quadratic cost, the authors in [7] also show local $C^{1,\alpha}$ regularity of the free boundary up to a closed singular set in the case where the supports of μ and ν are separated by a hyperplane; in the case of overlap, Figalli proves local C^1 regularity away from the common region (see [15, Theorem 4.11]) and Indrei [18] obtains local $C^{1,\alpha}$ regularity away from the common region and a set of singular points as well as Hausdorff dimension estimates on the singular set. For the general cost, Chen and Indrei [9] build up the regularity depending gradually on the cost: a locally Lipschitz cost implies a rectifiable free boundary, a C^1 cost implies a locally Lipschitz free boundary, and a C^4 cost satisfying Ma-Trudinger-Wang condition implies local $C^{1,\alpha}$ regularity. Concerning numerical approximations, Barrett and Prigozhin [3] study the problem from the numerical point of view for the case where $c(x,y)=|x-y|$.

In this paper, our aim is to give a rigorous theory study of the PMK problem with the cost given by a Finsler distance $c(x,y)=d_F(x,y)$ (including the case of Euclidean distance cost). We introduce first the Kantorovich-type duality for the PMK problem with general costs. Then, using the triangle inequality satisfied by d_F, we introduce the notion of Kantorovich potential for the PMK problem with Finsler distance costs. Recall that in the case $c(x,y)=|x-y|^2$, the obstacle Monge–Ampère equation (cf. [7,15]) plays an important role to gather many informations on the PMK problem. In our case, we introduce the obstacle Monge–Kantorovich (OMK) equation and show how it is information-rich PDE for the PMK problem and how it can operate effectively. In particular, our variational approach gives an interesting issue for numerical computation (cf. [17] and a forthcoming paper on Lagrangian costs). And the connection with the PDE allows us to

prove the uniqueness of active submeasures and the monotonicity for Finsler costs, which do not satisfy the left twist condition, via L^1 techniques for PDE.

Before giving the plan of the paper, let us take a while to comment our approach and main ideas. It is not difficult to see that the PMK problem is a bilevel optimization problem that aims to find the active submeasures with the constraint on the total mass as well as the optimal plan. The authors in [7] introduce a Lagrange multiplier λ for the mass constraint, add a point at infinity which acts as a tariff-free asreservoir for transporting the extra mass, and study the relations given by classical duality results. In their duality, λ is a parameter to be straightened to study the original PMK problem. In this way, they could deduce existence and uniqueness of minimizers when the supports of μ and ν are disjoint. As to the strategy of [15] is to study directly the minimization problem by studying the convexity of the function that associates to each $\mathbf{m}$ the total Monge–Kantorovich work. In particular, this allows the author to prove the uniqueness without the disjoint support condition on μ and ν. The techniques used in [7,15] also deduce the uniqueness for costs which require existence and uniqueness of optimal plan in the full transfer case (see [7, Proposition 2.9], [15, Renark 2.11] for precise statements). However, these techniques do not work for the uniqueness of active submeasures of the PMK problem with Finsler distance costs. Our point of view is to obtain the uniqueness via the study of the OMK equation. Our approach is different and our results do not follow readily from previous works. We begin by handling directly the problem for general costs by adding two arbitrary sites in $\mathbb{R}^N$ to process the problem into a balanced optimal mass transportation. Taking the cost for free to the new sites, we show that the new total work coincides with the total work of the PMK problem. Moreover, combining this with classical duality results, we introduce a bilevel maximization problem to provide a natural dual partial Monge–Kantorovich (DPMK) problem for the optimal partial transportation. In the case of Finsler distance, the variable of the DPMK problem can be expressed as a couple (λ, u) where u can be interpreted as the Kantorovich potential associated with the PMK problem and λ would be used to give informations on active submeasures. Recall that in the case where the cost is given by the square of the Euclidean distance (cf. [7]), the connection between the obstacle Monge–Ampère PDE and the PMK problem is given by a map that associates to each value parameter λ a solution of the Monge–Ampère PDE. In our case, we introduce a map that associates to each value λ a solution of the OMK equation. Then, we show how a right value $\lambda_{\mathbf{m}}$ enters in connection with the Kantorovich potential to bring out the solution of the PMK problem. Among the main issues of our approach, the uniqueness of the active submeasures holds true in the case where μ and ν are absolutely continuous without disjointness condition of the supports. As a consequence, we also obtain the monotonicity of active submeasures with respect to Lagrange multiplier λ.

The paper is organized as follows: In Section 2, we start in the first part by recalling the Kantorovich duality for a general lower semicontinuous cost function c. In the second part of Section 2, we summarize our main results on the PMK problem for a general cost c as well as for the case where c satisfies the triangle inequality. The third part deserves the results for the case of Finsler distance and the OMK equation. The remaining sections aim to prove the main results. The proof of the duality is given in Section 3. We study the existence and uniqueness issues for the OMK equation in Section 4. In Section 5, we show the connection between the OMK equation and the active submeasures for the PMK problem by using the DPMK problem

and the partial minimum flow problem. Thanks to this connection and the results on the OMK equation, we deduce the uniqueness of active submeasures. To finish the proof of the main results, we study some strong L^1 continuous dependence and monotonicity of the solution of the OMK equation with respect to the obstacle in Section 6. We terminate the paper by an appendix in which we give a chain rule for the tangential gradient to a measure.

2. Preliminaries and main results

In this section, we recall the Kantorovich duality for the optimal transport problem with a general cost $c : \mathbb{R}^N \times \mathbb{R}^N \to [0, +\infty]$. After that, we summarize our main results. The details of the proofs are given in the remaining sections.

Given metric spaces X_1, X_2, a measure $\eta \in \mathcal{M}_b^+(X_1)$ and Borel map $T : X_1 \longrightarrow X_2$, we denote by $T_\#\eta$ the pushforward measure of η by T,

$$T_\#\eta(B) := \eta(T^{-1}(B)) \text{ for every Borel subset } B \subset X_2.$$

2.1. Preliminaries on Monge–Kantorovich problem. To begin with, we assume that $\mu(\mathbb{R}^N) = \nu(\mathbb{R}^N)$. The Monge–Kantorovich (MK) problem reads as

$$\min \left\{ \mathcal{K}(\gamma) := \int\limits_{\mathbb{R}^N \times \mathbb{R}^N} c(x, y) \mathrm{d}\gamma \ : \gamma \in \pi(\mu, \nu) \right\},$$

where

$$\pi(\mu, \nu) := \left\{ \gamma \in \mathcal{M}_b^+(\mathbb{R}^N \times \mathbb{R}^N) \ : \pi_x\#\gamma = \mu, \ \pi_y\#\gamma = \nu \right\}.$$

Here $\pi_x, \pi_y : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ denote the standard projections and are given by $\pi_x(x, y) = x, \pi_y(x, y) = y$ for any $x, y \in \mathbb{R}^N$. The measure $\gamma \in \pi(\mu, \nu)$ is called *transport plan*. One of basic concepts in the optimal transport theory is the Kantorovich duality that can be restated as follows:

Theorem 2.1. *(cf. [25], Chapter 5) Let c be a lower semicontinuous cost function (l.s.c.) and $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be such that $\mu(\mathbb{R}^N) = \nu(\mathbb{R}^N)$. Then*
(i) The MK problem has an optimal plan and the Kantorovich duality holds true, i.e.

$$\min \{\mathcal{K}(\gamma) \ : \gamma \in \pi(\mu, \nu)\} = \sup \left\{ \int_{\mathbb{R}^N} u \mathrm{d}\mu + \int_{\mathbb{R}^N} v \mathrm{d}\nu : \ (u, v) \in \mathcal{S}_c(\mu, \nu) \right\}, \tag{2.1}$$

where

$$\mathcal{S}_c(\mu, \nu) := \left\{ (u, v) \in L^1_\mu(\mathbb{R}^N) \times L^1_\nu(\mathbb{R}^N) \ : \ u(x) + v(y) \le c(x, y) \ \forall x, y \in \mathbb{R}^N \right\}.$$

(ii) It does not change the value of the supremum in the right-hand side of (2.1) *if one restricts the definition of $\mathcal{S}_c(\mu, \nu)$ to those functions (u, v) which are bounded and continuous.*
(iii) If $c(x, y) \le C_\mu(x) + C_\nu(y)$ for some $(C_\mu, C_\nu) \in L^1_\mu \times L^1_\nu$, then the dual problem on the right-hand side (called Kantorovich dual problem) has an optimal solution.

(iv) If the cost function satisfies the triangle inequality and $c(x,x)=0$ for any $x\in\mathbb{R}^N$, then the Kantorovich dual problem can be rewritten as

$$\sup\left\{\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu)\ :\ u\in Lip_c\right\},$$

where $Lip_c := \left\{u:\mathbb{R}^N\mapsto\mathbb{R} : u\in L^1_\mu\cap L^1_\nu,\ u(y)-u(x)\le c(x,y)\ \forall x,y\in\mathbb{R}^N\right\}$. A solution of the Kantorovich dual problem is called Kantorovich potential.

An interesting situation where the triangle inequality is fulfilled corresponds to the case where the cost is proportional to a distance. Monge's original optimal mass transport problem corresponds to the Euclidean distance. The case where the cost c is given by a continuous Finsler distance has been studied recently in [16] (see also [22] for the symmetric case).

Let us begin with a reminder concerning Finsler metric. A continuous function $F:\mathbb{R}^N\times\mathbb{R}^N\longrightarrow[0,+\infty)$ is called *Finsler metric* on $\mathbb{R}^N$ if

- $F(x,.)$ is positively 1-homogeneous $\forall x\in\mathbb{R}^N$, i.e.

$$F(x,tv)=tF(x,v)\ \text{ for every }\ x\in\mathbb{R}^N, v\in\mathbb{R}^N\ \text{ and }\ t>0;$$

- $F(x,.)$ is convex for every $x\in\mathbb{R}^N$.

In addition, in this paper, we assume that F is nondegenerate in the sense that there exist $\beta,\alpha>0$ such that

$$\alpha|v|\le F(x,v)\le\beta|v|\ \text{ for all }\ (x,v)\in\mathbb{R}^N\times\mathbb{R}^N.$$

The *Finsler distance* d_F on $\mathbb{R}^N$ is defined by

$$d_F(x,y)=\inf\left\{\int_0^1 F(\xi(t),\dot\xi(t))\mathrm{d}t:\ \xi\in\Gamma_{x,y}(\mathbb{R}^N)\right\},\tag{2.2}$$

where $\Gamma_{x,y}(\mathbb{R}^N)$ is the set of Lipschitz curves on $[0,1]$ taking values in $\mathbb{R}^N$ joining x to y. Under the above assumptions, it is known that the inf problem (2.2) is actually the minimum and that d_F is not necessary symmetric distance, i.e. d_F satisfies

- $d_F(x,y)\ge 0$; $d_F(x,y)=0$ if and only if $x=y$.
- $d_F(x,y)\le d_F(x,z)+d_F(z,y)$ for any $x,y,z\in\mathbb{R}^N$.

An example of a Finsler metric which is not a norm in $\mathbb{R}$ is given by $F(x,v)=av^-+bv^+$ with $0<a\ne b$. More generally, for each $x\in\mathbb{R}^N$ fixed, given vectors $d_1^x,...,d_k^x\ne 0$ depending on x such that, for any $0\ne v\in\mathbb{R}^N$, $\max\limits_{1\le i\le k}\{\langle v,d_i^x\rangle\}>0$, we define

$$F(x,v):=\max_{1\le i\le k}\{\langle v,d_i^x\rangle\}\quad\text{for any}\ \ v\in\mathbb{R}^N$$

which turns out to be a Finsler metric. Numerical solutions for this kind of Finsler metrics are studied in [17].

For a Finsler metric F, the polar function F^* of F is defined by

$$F^*(x,p):=\sup\{\langle v,p\rangle:\ F(x,v)\le 1\}\ \text{ for }\ x,\,p\in\mathbb{R}^N.$$

It is easy to see that F^* is also a continuous, non-degenerate Finsler metric and

$$\langle\xi,p\rangle\le F^*(x,p)\,F(x,v)\ \ \forall x,\,v,\,p\in\mathbb{R}^N.$$

Now, we consider the optimal transport problem of moving μ into ν with the cost $c = d_F$. Since the cost d_F satisfies the triangle inequality and $d_F(x,x) = 0$ for any $x \in \mathbb{R}^N$, we have

$$\min\left\{\mathcal{K}(\sigma)\ :\ \sigma \in \pi(\mu,\nu)\right\} = \sup\left\{\int_{\mathbb{R}^N} u\,\mathrm{d}(\nu-\mu) : u \in Lip_{d_F}\right\}.$$

In the case where $F(x,v) = |v|$ for any $(x,v) \in \mathbb{R}^N \times \mathbb{R}^N$, it is clear that d_F is the Euclidean distance. For regular densities μ and ν, Evans and Gangbo proved that the characterization of the Kantorovich potential may be given by a nonlinear PDE of the $p-$Laplacian type with $p = \infty$. Since [6], this PDE is called *Monge–Kantorovich equation* and formally reads

$$\begin{cases} -\nabla\cdot\Phi = \nu - \mu \quad \text{ in } \mathcal{D}'(\mathbb{R}^N) \\ \\ u \in Lip_{|.|},\ \Phi\cdot\nabla u = |\Phi|. \end{cases} \tag{2.3}$$

Roughly speaking, the flux Φ in (2.3) is called the transportation flow and its total variation $|\Phi|$ gives the density of the transportation. Formally, the second line of (2.3) conceives the equation

$$\Phi \in \partial I\!I_{\overline{B}(0,1)}(\nabla u),$$

where $\overline{B}(0,1)$ is the Euclidean closed unit ball.

As we will see, among our results, we give a generalization of the PDE (2.3) for general Finsler metrics. Actually, for a Finsler metric F, (2.3) reads formally

$$\begin{cases} -\nabla\cdot\Phi = \nu - \mu \quad \text{ in } \mathcal{D}'(\mathbb{R}^N) \\ u \in Lip_{d_F},\ \Phi\cdot\nabla u = F(.,\Phi). \end{cases} \tag{2.4}$$

In general, since Φ is a finite Radon vector measure, the second equation of (2.4) needs to be understood in the sense of tangential gradient with respect to a measure. More precisely, a couple (u,Φ) is said to satisfy the PDE (2.4) if $(u,\Phi) \in Lip_{d_F} \times \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)$, $-\nabla\cdot\Phi = \nu - \mu$ in $\mathcal{D}'(\mathbb{R}^N)$ and

$$\frac{\Phi}{|\Phi|}\cdot\nabla_{|\Phi|}u = F\left(.,\frac{\Phi}{|\Phi|}(.)\right), \quad |\Phi| - \text{ a.e. in } \mathbb{R}^N.$$

Here, we denote by $\dfrac{\Phi}{|\Phi|}$ the density of Φ with respect to $|\Phi|$ and by $\nabla_{|\Phi|}u$ the tangential gradient of u with respect to $|\Phi|$ (cf. [5,6,19]) which is well-defined for any Lipschitz continuous function. Following [19, Lemma 4.9], given $\Phi \in \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)$ and $\eta \in \mathcal{M}_b(\mathbb{R}^N)$ such that $-\nabla\cdot\Phi = \eta$ in $\mathcal{D}'(\mathbb{R}^N)$, we have

$$\int_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\cdot\nabla_{|\Phi|}\xi\mathrm{d}|\Phi| = \int_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\cdot\nabla\xi\mathrm{d}|\Phi| = \int_{\mathbb{R}^N} \xi\,\mathrm{d}\eta \quad \text{for all } \ \xi \in C_c^\infty(\mathbb{R}^N). \tag{2.5}$$

Moreover, one can prove that (using approximation and (2.5))

$$\int_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\cdot\nabla_{|\Phi|}\xi\mathrm{d}|\Phi| = \int_{\mathbb{R}^N} \xi\mathrm{d}\eta \quad \text{for all } \ \xi \in Lip(\mathbb{R}^N)\cap C_b(\mathbb{R}^N).$$

2.2. Main result for the PMK with general costs. Assume that $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ are compactly supported and

$$\mathbf{m}_{\max} := \min\{\mu(\mathbb{R}^N), \nu(\mathbb{R}^N)\}.$$

Given a total mass $\mathbf{m} \in [0, \mathbf{m}_{\max}]$, the PMK problem aims to transport effectively a total mass $\mathbf{m}$ from a supply submeasure of μ into a submeasure of ν. The set of submeasures of mass $\mathbf{m}$ is given by

$$Sub_{\mathbf{m}}(\mu,\nu) := \left\{(\rho_0,\rho_1) \in \mathcal{M}_b^+(\mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N) : \ \rho_0 \le \mu, \ \ \rho_1 \le \nu, \ \ \rho_0(\mathbb{R}^N) = \rho_1(\mathbb{R}^N) = \mathbf{m}\right\}.$$

Then the PMK reads as

$$\min\left\{\mathcal{K}(\gamma) := \int\limits_{\mathbb{R}^N\times\mathbb{R}^N} c(x,y)\mathrm{d}\gamma : \ \gamma \in \pi_{\mathbf{m}}(\mu,\nu)\right\}, \tag{2.6}$$

where

$$\pi_{\mathbf{m}}(\mu,\nu) := \Big\{\gamma \in \pi(\rho_0,\rho_1) : \ (\rho_0,\rho_1) \in Sub_{\mathbf{m}}(\mu,\nu)\Big\}.$$

An element $(\rho_0,\rho_1) \in Sub_{\mathbf{m}}(\mu,\nu)$ is called a couple of *active submeasures* if it solves the PMK (2.6). That is, there exists an optimal plan γ of the PMK (2.6) such that $\gamma \in \pi(\rho_0,\rho_1)$.

Our main result concerning duality for the PMK problem with general costs is the following.

Theorem 2.2. *Let $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be measures with compact supports X and Y, $\mathbf{m} \in [0, \mathbf{m}_{\max}]$. Assume that c is lower semi continuous and bounded on $X \times Y$. The PMK problem has a solution $\sigma^* \in \pi_{\mathbf{m}}(\mu,\nu)$ and the Kantorovich duality turns into*

$$\begin{aligned}\mathcal{K}(\sigma^*) &= \min_{\gamma\in\pi_{\mathbf{m}}(\mu,\nu)}\left\{\int\limits_{\mathbb{R}^N\times\mathbb{R}^N} c\,\mathrm{d}\gamma\right\}\\ &= \max\left\{\int\limits_{\mathbb{R}^N}\phi\,\mathrm{d}\mu + \int\limits_{\mathbb{R}^N}\psi\,\mathrm{d}\nu + \lambda\mathbf{m} : \ \lambda\in\mathbb{R}^+, \ (\phi,\psi)\in\mathcal{S}_c^\lambda(\mu,\nu)\right\},\end{aligned} \tag{2.7}$$

where

$$\mathcal{S}_c^\lambda(\mu,\nu) := \left\{(\phi,\psi) \in L^1_\mu \times L^1_\nu : \ \phi \le 0, \ \psi \le 0 \ \textit{and} \ \phi(x)+\psi(y)+\lambda \le c(x,y) \ \ \forall x,y \in \mathbb{R}^N\right\}.$$

Moreover, $\sigma \in \pi_{\mathbf{m}}(\mu,\nu)$ and $(\lambda,\phi,\psi) \in \mathbb{R}^+ \times \mathcal{S}_c^\lambda(\mu,\nu)$ are solutions, respectively, if and only if

$$\begin{aligned}&\phi(x) = 0 \ \textit{for} \ (\mu - \pi_x\#\sigma)\text{-a.e.} \ \ x\in\mathbb{R}^N, \ \ \ \psi(y) = 0 \ \textit{for} \ (\nu - \pi_y\#\sigma)\text{-a.e.} \ \ y\in\mathbb{R}^N\\ &\textit{and} \ \ \phi(x)+\psi(y)+\lambda = c(x,y) \ \ \textit{for} \ \ \sigma\text{-a.e.} \ \ (x,y)\in\mathbb{R}^N\times\mathbb{R}^N.\end{aligned} \tag{2.8}$$

The maximization problem on the right hand side of (2.7) is called dual partial Monge–Kantorovich (DPMK) problem.

Remark 2.3. See that the duality formulation (2.7) is slightly different from Caffarelli–McCann's duality (see [7, Corollary 2.7]) which reads as follows: for any fixed parameter λ,

$$\min_{\gamma\in\Pi_{\leq}(\mu,\nu)}\left\{\int\limits_{\mathbb{R}^N\times\mathbb{R}^N}(c-\lambda)d\gamma\right\} = \max_{\substack{u(x)+v(y)\leq c(x,y)-\lambda\\ u(x)\leq 0, v(y)\leq 0}}\int\limits_{\mathbb{R}^N}u(x)d\mu+\int\limits_{\mathbb{R}^N}v(y)d\nu, \quad (2.9)$$

where

$$\pi_{\leq}(\mu,\nu):=\left\{\gamma\in\mathcal{M}_b^+(\mathbb{R}^N\times\mathbb{R}^N)\ :\pi_x\#\gamma\leq\mu,\ \pi_y\#\gamma\leq\nu\right\}.$$

See that the mass constraint disappears in this formulation, which is not the case for (2.7). However, the connection between both formulations (2.7) and (2.9) is clear in the case where

$$\min_{\gamma\in\Pi_{\leq}(\mu,\nu)}\int\limits_{\mathbb{R}^N\times\mathbb{R}^N}(c-\lambda)d\gamma \quad (2.10)$$

has a unique optimal plan γ for each λ. Indeed, in this case the duality (2.7) can be seen as a consequence of the Caffarelli–McCann one. To see this, we fix a total mass $\mathbf{m}$ and we use [7, Corollary 2.11] to choose λ^* such that the unique solution γ_{λ^*} of (2.10) w.r.t. λ^* satisfies $\gamma_{\lambda^*}(\mathbb{R}^N\times\mathbb{R}^N)=\mathbf{m}$. It follows that the left hand side of (2.7) is less than or equal the right hand side. The inverse inequality can be verified directly. Yet, the uniqueness of the problem (2.10) is not true in general, especially for Finsler distance cost. As to the dual formulation (2.7), it is direct and closely connected to the PMK problem. The parameter λ is a variable and permits to handle the mass constraint. For numerical computation concerning the PMK problem, the formulation (2.7) with λ as a variable is very useful. This issue is discussed in [17].

Following Kantorovich–Rubinstein duality in the case where the cost satisfies the triangle inequality, we prove the following result.

Theorem 2.4. *Under the assumptions and notations of Theorem 2.2, assume moreover that the cost function c satisfies the triangle inequality and $c(x,x)=0$ for any $x\in\mathbb{R}^N$. Then the DPMK problem can be rewritten as*

$$\mathcal{K}(\sigma^*)=\max\left\{\mathcal{D}(\lambda,u):=\int u\,\mathrm{d}(\nu-\mu)+\lambda(\mathbf{m}-\nu(\mathbb{R}^N))\ :\ \lambda\geq 0 \text{ and } u\in L_c^\lambda\right\}, \quad (2.11)$$

where

$$L_c^\lambda:=\left\{u\in Lip_c\ :\ 0\leq u(x)\leq\lambda \quad \text{for any } x\in\mathbb{R}^N\right\}.$$

In addition, $\sigma\in\pi_{\mathbf{m}}(\mu,\nu)$ and $(\lambda,u)\in\mathbb{R}^+\times L_c^\lambda$ are solutions of the PMK and of the DPMK (2.11), respectively, if and only if

$$\begin{aligned}&u(x)=0 \text{ for } (\mu-\pi_x\#\sigma)\text{-a.e. } x\in\mathbb{R}^N,\quad u(x)=\lambda \text{ for } (\nu-\pi_y\#\sigma)\text{-a.e. } x\in\mathbb{R}^N\\ &\text{and } u(y)-u(x)=c(x,y) \text{ for } \sigma\text{-a.e. } (x,y)\in\mathbb{R}^N\times\mathbb{R}^N.\end{aligned} \quad (2.12)$$

The proofs of Theorem 2.2 and Theorem 2.4 are given in Section 3.

2.3. Main results for Finsler distance costs. Coming back to the Finsler distance cost d_F, we introduce a new nonlinear PDE that we call the obstacle Monge–Kantorovich (OMK) equation. Then, we use this PDE to show the uniqueness of active submeasures whenever the data μ and ν are absolutely continuous with respect to Lebesgue measure.

To introduce our PDE, we see that the dual formulation (2.11) may be written as

$$\max_{\lambda\geq 0}\left(\max_{u}\left\{\mathcal{D}(\lambda,u):\ u\in L^{\lambda}_{d_F}\right\}\right).$$

Moreover, formally, for any fixed $\lambda\geq 0$, the Euler–Lagrange equation associated with the problem

$$\max_{u}\left\{\mathcal{D}(\lambda,u)\ :\ u\in L^{\lambda}_{d_F}\right\} \tag{2.13}$$

is given by the following PDE

$$(P_\lambda)\qquad \begin{cases} \theta-\nabla\cdot\Phi=\nu-\mu \quad \text{ in } \ \mathcal{D}'(\mathbb{R}^N)\\[2ex] \Phi\cdot\nabla u=F(.,\Phi)\\[2ex] u\in L^{\lambda}_{d_F},\ \theta\in\partial I\!\!I_{[0,\lambda]}(u). \end{cases}$$

This is a double obstacle problem associated with the PMK problem (2.6) with $c=d_F$. And, formally we conclude that the study of the PMK is closely connected to the study of the dependence of a solution of (P_λ) with respect to λ. Our aim now is to study this connection to get a characterization of active submeasures of the PMK problem. Before going further, let us give the notion of solution to the OMK equation (P_λ).

Definition 2.5. For a fixed $\lambda\geq 0$, a triplet $(\theta,\Phi,u)\in\mathcal{M}_b(\mathbb{R}^N)\times\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times Lip_{d_F}$ is said to be a solution to the OMK equation (P_λ) if $u\in L^{\lambda}_{d_F}$ and we have

$$\begin{cases} \theta-\nabla\cdot\Phi=\nu-\mu \ \text{ in } \ \mathcal{D}'(\mathbb{R}^N)\\ \dfrac{\Phi}{|\Phi|}(x)\cdot\nabla_{|\Phi|}u(x)=F\left(x,\dfrac{\Phi}{|\Phi|}(x)\right),\ |\Phi|\text{-a.e. } \ x\in\mathbb{R}^N\\ u=0,\ \ \theta^-\text{-a.e. in } \ \mathbb{R}^N \quad \text{and} \quad u=\lambda,\ \ \theta^+\text{-a.e. in } \ \mathbb{R}^N, \end{cases}$$

where $\theta^{\pm}$ is the positive and negative part of the measure θ given by the Jordan decomposition.

Without abusing, we also say that a Radon measure $\theta\in\mathcal{M}_b(\mathbb{R}^N)$ is a solution of (P_λ) if there exists $(\Phi,u)\in\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times Lip_{d_F}$ such that (θ,Φ,u) satisfies the OMK equation (P_λ).

To set the connection between the OMK equation and the active submeasures, let us denote by $\mu\wedge\nu$ the measure of common mass of μ and ν. Recall that if $\mu,\nu\in L^1(\mathbb{R}^N)$ then $\mu\wedge\nu\in L^1(\mathbb{R}^N)$ and

$$(\mu\wedge\nu)(x)=\min\{\mu(x),\nu(x)\}\ \text{ for a.e. } \ x\in\mathbb{R}^N.$$

In general, the measure $\mu\wedge\nu$ is defined by (see [2])

$$\mu\wedge\nu(A)=\inf\{\mu(A_1)+\nu(A_2):\ \text{disjoint Borel sets } A_1,A_2,\text{ such that } \ A_1\cup A_2=A\}.$$

Theorem 2.6 (Existence of a solution to the OMK equation)**.** *Given $\mu,\nu\in\mathcal{M}_b^+(\mathbb{R}^N)$ and $\lambda\geq 0$, the OMK equation (P_λ) admits at least one solution (θ,Φ,u). Moreover,*

$$\theta^-\leq\mu-\mu\wedge\nu\leq\mu \ \text{ and } \ \theta^+\leq\nu-\mu\wedge\nu\leq\nu$$

for any solutions (θ,Φ,u).

Because of the degeneracy of the OMK equation, the question of the uniqueness of a solution for (P_λ) is delicate. In fact, one cannot in general expect the uniqueness of the components Φ and u of solutions for the OMK equation (P_λ). However, we can prove the uniqueness of the component θ whenever μ and ν are absolutely continuous with respect to the Lebesgue measure.

Theorem 2.7 (Uniqueness of θ)**.** *Assume that $\mu,\nu\in L^1(\mathbb{R}^N)^+$. Let θ_1 and θ_2 be two solutions to the same OMK equation (P_λ). Then $\theta_1,\theta_2\in L^1(\mathbb{R}^N)$ and $\theta_1=\theta_2$.*

Now, we come to the connection between the OMK equation and the PMK problem.

Theorem 2.8 (Active submeasures and OMK equation)**.** *Let $\mu,\nu\in\mathcal{M}_b^+(\mathbb{R}^N)$ be compactly supported.*
(i) For any $\lambda\geq 0$ and θ_λ a solution of the OMK equation (P_λ), the couple

$$(\rho_0,\rho_1):=(\mu-\theta_\lambda^-,\nu-\theta_\lambda^+)$$

is a couple of active submeasures corresponding to $\mathbf{m}_\lambda=(\mu-\theta_\lambda^-)(\mathbb{R}^N)$.
(ii) Conversely, if $(\rho_0,\rho_1)\in Sub_{\mathbf{m}}(\mu,\nu)$ is a given couple of active submeasures and $\mathbf{m}\in[(\mu\wedge\nu)(\mathbb{R}^N),\mathbf{m}_{\max}]$ then for any $\lambda_{\mathbf{m}}\geq 0$ such that

$$\lambda_{\mathbf{m}}\in\arg\max_{\lambda\geq 0}\left\{\max_u\left\{\mathcal{D}(\lambda,u)\ :u\in L^\lambda_{d_F}\right\}\right\},$$

the measure $\theta_{\lambda_{\mathbf{m}}}$ defined by

$$\theta^-_{\lambda_{\mathbf{m}}}:=\mu-\rho_0 \quad \text{ and } \quad \theta^+_{\lambda_{\mathbf{m}}}:=\nu-\rho_1$$

is a solution of the OMK equation $(P_{\lambda_{\mathbf{m}}})$.

Following Theorems 2.8 and 2.7, we have the following result for the PMK problem.

Corollary 2.9 (Uniqueness of active submeasures)**.** *Let $\mu,\nu\in L^1(\mathbb{R}^N)^+$ be compactly supported and $\mathbf{m}\in[\|\mu\wedge\nu\|_{L^1},\mathbf{m}_{\max}]$. There exists a unique couple of active submeasures.*

To end up this section of main results, we propose to study the maps that associate to each $\lambda\geq 0$ the corresponding active submeasures and their total mass in the case $\mu,\nu\in L^1(\mathbb{R}^N)$. Thanks to Theorems 2.6, 2.8 and 2.7, for any $\lambda\geq 0$ there exist a unique mass $\mathbf{m}_\lambda:=(\mu-\theta_\lambda^-)(\mathbb{R}^N)$ and a unique couple of active submeasures $(\rho_0^\lambda,\rho_1^\lambda):=(\mu-\theta_\lambda^-,\nu-\theta_\lambda^+)$ corresponding to $\mathbf{m}_\lambda$. Let us define the maps

$$\begin{array}{rcl}\mathsf{m}\ :\ [0,\infty) & \to & [(\mu\wedge\nu)(\mathbb{R}^N),\mathbf{m}_{\max}]\\ \lambda & \to & \mathsf{m}(\lambda):=\mathbf{m}_\lambda\end{array}$$

and

$$\begin{array}{rcl}\mathcal{R}\ :\ [0,\infty) & \to & L^1(\mathbb{R}^N)\times L^1(\mathbb{R}^N)\\ \lambda & \to & \mathcal{R}(\lambda):=(\rho_0^\lambda,\rho_1^\lambda).\end{array}$$

To simplify the presentation, let us denote

$$Sub_{opt}(\mu,\nu) := \{(\rho_0,\rho_1) : \rho_0 \text{ and } \rho_1 \text{ are active submeasures for some } \mathbf{m} \in [\|\mu\wedge\nu\|_{L^1}, \mathbf{m}_{\max}]\}.$$

Theorem 2.10. *Let $\mu,\nu \in L^1(\mathbb{R}^N)^+$ be compactly supported. We have that*
(i) the map m is continuous, non-decreasing and surjective;
(ii) the map $\mathcal{R}$ is continuous, non-decreasing and surjective from $[0,\infty)$ to $Sub_{opt}(\mu,\nu)$.

It is known that the monotonicity of active submeasures corresponding to the mass $\mathbf{m}$ is obtained for continuous cost c (see [7, Theorem 3.4]) and that the monotonicity corresponding to Lagrange multiplier λ is guaranteed for costs satisfying the left twist condition (see [7, Sections 2 and 3]). Our result confirms that even if the Finsler distances do not satisfy the left twist condition, the monotonicity w.r.t. Lagrange multiplier λ still holds true. Also, notice that in the quadratic case, Davila and Kim establish a Lipschitz continuous dependence of $\mathbf{m}_\lambda$ on λ (see [10, Theorem 4.5]). We do not know if this kind of results holds true for Finsler distances or not.

Remark 2.11. (i) The uniqueness of active submeasures falls to be true in the case where $\mathbf{m} < (\mu\wedge\nu)(\mathbb{R}^N)$. Indeed, in this case, all submeasures $\rho_0 \equiv \rho_1 \le \mu\wedge\nu$ of mass $\mathbf{m}$ are optimal. This is not a contradiction with our PDE approach. Because there is no such an OMK equation with $\lambda \ge 0$ characterizing the PMK problem in this case.

(ii) In general, the uniqueness of active submeasures does not hold true if both μ and ν are not in L^1. For example, on $\mathbb{R}$, take $\mu = \delta_1 + \delta_3$, $\nu = \delta_2$, where δ_k is the Direct mass at k in $\mathbb{R}$. Then all submeasures of mass $\mathbf{m}$ are optimal for any $\mathbf{m}$. However, we do not know if this holds true if one of μ,ν belongs to $L^1(\mathbb{R}^N)$. Our arguments use PDE techniques and need both μ and ν to be L^1 functions.

(iii) Uniqueness result was obtained in [21, Lemma 5.1, Proposition 5.2] in the quadratic case when $\mathbf{m} = \mathbf{m}_{max}$ by using completely different arguments. These arguments can be adapted to handle more general costs satisfying regularity conditions on the balls $B^1_c(y,R) := \{t \in \mathbb{R}^N : c(t,y) < R\}$ and $B^2_c(x,R) := \{w \in \mathbb{R}^N : c(x,w) < R)\}$. That is, for any $x_0 \in \partial B^i_c(.,R)$ for i=1,2, there exists a sequence of subsets of the ball which shrinks nicely to x_0. One can prove that this is true in the case where $F(x,v) \equiv F(v)$ is independent of x.

3. Kantorovich-type duality

The aim of this section is to prove Theorems 2.2 and 2.4.

Proof of Theorem 2.2. The existence of an optimal plan $\sigma^* \in \pi_{\mathbf{m}}(\mu,\nu)$ can be shown by the direct method. Next, for any $\sigma \in \pi_{\mathbf{m}}(\mu,\nu)$ and $(\lambda,\phi,\psi) \in \mathbb{R}^+ \times \mathcal{S}_c^\lambda(\mu,\nu)$, we have

$$\begin{aligned}\int_{\mathbb{R}^N} \phi(x)\,\mathrm{d}\mu(x) + \int_{\mathbb{R}^N} \psi(y)\,\mathrm{d}\nu(y) + \lambda\mathbf{m} &\le \int_{\mathbb{R}^N} \phi(x)\,\mathrm{d}\pi_x\#\sigma + \int_{\mathbb{R}^N} \psi(y)\,\mathrm{d}\pi_y\#\sigma + \lambda\mathbf{m} \\ &= \int_{\mathbb{R}^N\times\mathbb{R}^N} (\phi(x)+\psi(y)+\lambda)\,\mathrm{d}\sigma \\ &\le \int_{\mathbb{R}^N\times\mathbb{R}^N} c(x,y)\mathrm{d}\sigma.\end{aligned} \tag{3.1}$$

Hence,

$$\sup\left\{\int_{\mathbb{R}^N} \phi\,\mathrm{d}\mu + \int_{\mathbb{R}^N} \psi\,\mathrm{d}\nu + \lambda\mathbf{m} \,:\, \lambda\in\mathbb{R}^+,\ (\phi,\psi)\in\mathcal{S}_c^\lambda(\mu,\nu)\right\} \le \min_{\sigma\in\pi_{\mathbf{m}}(\mu,\nu)} \mathcal{K}(\sigma).$$

In order to prove the converse inequality, we add two points $\hat{x}\in\mathbb{R}^N\setminus X$ and $\hat{y}\in\mathbb{R}^N\setminus Y$ as extra production and consumption positions, respectively. Let us consider $\hat{X} := X\cup\{\hat{x}\}$, $\hat{Y} := Y\cup\{\hat{y}\}$ as metric spaces (induced by the Euclidean distance) and the measures on $\hat{X}$ and $\hat{Y}$ defined, respectively, by

$$\widehat{\mu} = \mu + (\nu(Y)-\mathbf{m}))\delta_{\widehat{x}} \quad \text{and} \quad \widehat{\nu} = \nu + (\mu(X)-\mathbf{m})\delta_{\widehat{y}}.$$

Obviously, $\hat{\mu}(\hat{X}) = \hat{\nu}(\hat{Y})$. Then, let us consider the extra cost on $\hat{X}\times\hat{Y}$

$$\hat{c}(x,y) := \begin{cases} c(x,y) & \text{if} \quad (x,y)\in X\times Y \\ 0 & \text{if} \quad x=\hat{x} \text{ or } y=\hat{y}.\end{cases}$$

From the assumptions on c, we have that $\hat{c}$ is l.s.c. and bounded on the compact metric space $\hat{X}\times\hat{Y}$. Using Theorem 2.1, we have

$$\min_{\hat{\gamma}\in\widehat{\pi}(\hat{\mu},\hat{\nu})} \int_{\widehat{X}\times\widehat{Y}} \hat{c}(x,y)\,\mathrm{d}\hat{\gamma} = \max_{(\hat{u},\hat{v})\in\mathcal{S}_{\hat{c}}(\hat{\mu},\hat{\nu})} \int_{\hat{X}} \hat{u}\,\mathrm{d}\hat{\mu} + \int_{\hat{Y}} \hat{v}\,\mathrm{d}\hat{\nu}.$$

Fixed any $\hat{\gamma}\in\pi(\hat{\mu},\hat{\nu})$, set $\gamma_1 := \hat{\gamma}\llcorner_{X\times Y}$ the restricted measure of $\hat{\gamma}$ on $X\times Y$. It is easy to see that $\pi_x\#\gamma_1 \le \mu,\ \pi_y\#\gamma_1 \le \nu$ and $\gamma_1(X\times Y)\ge\mathbf{m}$. Let us define $\gamma := \dfrac{\mathbf{m}}{\gamma_1(X\times Y)}\gamma_1 \in \pi_{\mathbf{m}}(\mu,\nu)$ and

$$\int_{X\times Y} c(x,y)\mathrm{d}\gamma \le \int_{X\times Y} c(x,y)\mathrm{d}\gamma_1 = \int_{\hat{X}\times\hat{Y}} \hat{c}(x,y)\mathrm{d}\hat{\gamma}.$$

This implies that

$$\min_{\gamma\in\pi_{\mathbf{m}}(\mu,\nu)} \int_{X\times Y} c(x,y)\mathrm{d}\gamma \le \min_{\hat{\gamma}\in\widehat{\pi}(\hat{\mu},\hat{\nu})} \int_{\hat{X}\times\hat{Y}} \hat{c}(x,y)\mathrm{d}\hat{\gamma} = \max_{(\hat{u},\hat{v})\in\mathcal{S}_{\hat{c}}(\hat{\mu},\hat{\nu})} \int_{\hat{X}} \hat{u}\,\mathrm{d}\hat{\mu} + \int_{\hat{Y}} \hat{v}\,\mathrm{d}\hat{\nu}.$$

To finish the proof, for any $(\hat{u},\hat{v})\in\mathcal{S}_{\hat{c}}(\hat{\mu},\hat{\nu})$, we can moreover assume that $\hat{u},\hat{v}$ have finite values. Set $u_1:=\hat{u}+\hat{v}(\hat{y})$, $v_1:=\hat{v}+\hat{u}(\hat{x})$ and $\lambda:=-\hat{u}(\hat{x})-\hat{v}(\hat{y})\geq 0$. Since $\hat{u}(x)+\hat{v}(y)\leq\hat{c}(x,y)$, we see that $u_1\leq 0$ in X, $v_1\leq 0$ in Y and $u_1(x)+v_1(y)\leq c(x,y)-\lambda$ for any $(x,y)\in X\times Y$. So, extending arbitrarily u_1 and v_1 up to $\mathbb{R}^N$ such that $(u_1,v_1)\in\mathcal{S}_c^\lambda(\mu,\nu)$, we get

$$\begin{aligned}\int_{\hat{X}}\hat{u}(x)\,\mathrm{d}\hat{\mu}+\int_{\hat{Y}}\hat{v}(y)\,\mathrm{d}\hat{\nu}&=\int_X\hat{u}(x)\,\mathrm{d}\mu+\int_Y\hat{v}(y)\,\mathrm{d}\nu+(\nu(Y)-\mathbf{m})\hat{u}(\hat{x})+(\mu(X)-\mathbf{m})\hat{v}(\hat{y})\\&=\int_X\left(\hat{u}(x)+\hat{v}(\hat{y})\right)\,\mathrm{d}\mu+\int_Y\left(\hat{v}(y)+\hat{u}(\hat{x})\right)\,\mathrm{d}\nu-(\hat{u}(\hat{x})+\hat{v}(\hat{y}))\mathbf{m}\\&=\int_X u_1(x)\,\mathrm{d}\mu+\int_Y v_1(y)\,\mathrm{d}\nu+\lambda\mathbf{m}.\end{aligned}$$

Thus

$$\begin{aligned}\min_{\gamma\in\Pi_{\mathbf{m}}(\mu,\nu)}\int_{X\times Y}c(x,y)\,\mathrm{d}\gamma&\leq\max_{(\hat{u},\hat{v})\in\mathcal{S}_{\hat{c}}(\hat{\mu},\hat{\nu})}\int_{\hat{X}}\hat{u}\,\mathrm{d}\hat{\mu}+\int_{\hat{Y}}\hat{v}\,\mathrm{d}\hat{\nu}\\&\leq\sup\left\{\int_{\mathbb{R}^N}\phi\,\mathrm{d}\mu+\int_{\mathbb{R}^N}\psi\,\mathrm{d}\nu+\lambda\mathbf{m}\ :\ \lambda\geq 0,\ (\phi,\psi)\in\mathcal{S}_c^\lambda(\mu,\nu)\right\}.\end{aligned}$$

From the above arguments, the supremum is actually the maximum.
At last, by the duality (2.7), $\sigma\in\pi_{\mathbf{m}}(\mu,\nu)$ and $(\lambda,\phi,\psi)\in\mathbb{R}^+\times\mathcal{S}_c^\lambda(\mu,\nu)$ are solutions of the PMK and of the DPMK, respectively, if and only if the two inequalities in (3.1) are equalities. This is equivalent to the optimality condition (2.8). □

Now, we prove the duality for the case where the cost satisfies the triangle inequality.

Proof of Theorem 2.4. We see that

$$\min_{\gamma\in\Pi_{\mathbf{m}}(\mu,\nu)}\int_{\mathbb{R}^N\times\mathbb{R}^N}c(x,y)\mathrm{d}\gamma\geq\sup\left\{\mathcal{D}(\lambda,u)\ :\ \lambda\geq 0\text{ and }u\in L_c^\lambda\right\}.$$

Indeed, for any $\gamma\in\pi_{\mathbf{m}}(\mu,\nu)$ and $u\in L_c^\lambda$, we have

$$\begin{aligned}\int_{\mathbb{R}^N}u\,\mathrm{d}(\nu-\mu)+\lambda(\mathbf{m}-\nu(\mathbb{R}^N))&=\int_{\mathbb{R}^N}-u(x)\,\mathrm{d}\mu+\int_{\mathbb{R}^N}(u(y)-\lambda)\mathrm{d}\nu+\lambda\mathbf{m}\\&\leq\int_{\mathbb{R}^N}-u\,\mathrm{d}\pi_x\#\gamma+\int_{\mathbb{R}^N}(u(y)-\lambda)\,\mathrm{d}\pi_y\#\gamma+\lambda\mathbf{m}\\&\leq\int_{\mathbb{R}\times\mathbb{R}^N}c(x,y)\,\mathrm{d}\gamma(x,y).\end{aligned}\tag{3.2}$$

Conversely, for a given $\lambda \geq 0$ and $(\phi, \psi) \in \mathcal{S}_c^\lambda(\mu, \nu)$, we consider

$$u_1(x) := \sup_{y \in Y} (\psi(y) + \lambda - c(x, y)) \leq \lambda \quad \text{and} \quad u(x) := \max\{u_1(x), 0\} \quad \forall x \in \mathbb{R}^N.$$

By using the triangle inequality, u is 1-Lipschitz with respect to c. Moreover, $-u \geq \phi$ and $u(y) - \lambda \geq \psi(y) \quad \forall y \in Y$ (where we use the condition $c(y, y) = 0$). So

$$\int_{\mathbb{R}^N} u \, \mathrm{d}(\nu - \mu) + \lambda(\mathbf{m} - \nu(\mathbb{R}^N)) \geq \int_{\mathbb{R}^N} \phi \, \mathrm{d}\mu + \int_{\mathbb{R}^N} \psi \, \mathrm{d}\nu + \lambda \mathbf{m}.$$

By Theorem 2.2, the duality and the existence of a solution (λ^*, u^*) are proven. For the optimality condition (2.12), we use again the duality and (3.2) similarly to the case of general costs c. □

Remark 3.12. If c satisfies the triangle inequality and $c(x, x) = 0$ for any $x \in \mathbb{R}^N$ then the DPMK problem can be also written as

$$\max \left\{ \int_{\mathbb{R}^N} u \, \mathrm{d}(\nu - \mu) + \lambda(\mathbf{m} - \mu(\mathbb{R}^N)) : \ \lambda \in \mathbb{R}^+, \ u \in Lip_c, \ -\lambda \leq u \leq 0 \right\}.$$

Indeed, in the construction of u from (ϕ, ψ), we can take

$$u_1(y) := \inf_{x \in X} (c(x, y) - \phi(x) - \lambda) \quad \text{and} \quad u(y) := \min\{u_1(y), 0\} \ \forall y \in \mathbb{R}^N.$$

4. OMK equation

The aim of this section is to study the existence and uniqueness of solutions for the OMK equation (P_λ). Besides this, we also show some estimates for solutions θ of the OMK equation which are useful for later use. We will make use of variational techniques for the existence while the uniqueness and the estimates of θ are shown by using PDE techniques. In this section, we do not really need the compactness of the supports of μ and ν.

4.1. Existence of solution to the OMK equation. The existence of solution to the OMK equation is based on the dual approach. More precisely, by applying the Fenchel–Rockafellar dual theory to the problem (2.13), we introduce a minimum flow-type problem (for the minimum flow problem and some its variants, one can see [23, Chapter 4]). And then we show that the OMK equation is given by the optimality condition for the two problems.

Proposition 4.13. *Let μ, $\nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ and $\lambda \geq 0$ be fixed. We have*

$$\max_{u \in L^\lambda_{d_F}} \int_{\mathbb{R}^N} u \mathrm{d}(\nu - \mu) = \min \Big\{ \int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x)) \mathrm{d}|\Phi| + \lambda \int_{\mathbb{R}^N} \mathrm{d}\theta^1 : \ (\Phi, \theta^0, \theta^1) \in S \Big\}, \tag{4.1}$$

where

$$S := \Big\{ (\Phi, \theta^0, \theta^1) \in \mathcal{M}_b(\mathbb{R}^N; \mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N) : -\nabla \cdot \Phi = \nu - \theta^1 - (\mu - \theta^0) \Big\}.$$

We need some elementary lemmas concerning the set Lip_{d_F} of 1-Lipschitz functions w.r.t d_F.

Lemma 4.1. *Let F be a continuous, nondegenerate Finsler metric and $u \in Lip_{d_F}$. Set $u_\varepsilon := \rho_\varepsilon \star u$, where ρ_ε is the standard mollifier on $\mathbb{R}^N$. Then*

$$\limsup_{\varepsilon\to 0} F^*(x, Du_\varepsilon(x)) \le 1 \ \text{ for all } \ x \in \mathbb{R}^N. \tag{4.2}$$

Proof. Let $x \in \mathbb{R}^N$ be fixed. There exists some $\|\xi_\varepsilon\| = 1$ such that

$$\begin{aligned} F^*(x, Du_\varepsilon(x)) &= \frac{\langle Du_\varepsilon(x)\cdot \xi_\varepsilon\rangle}{F(x,\xi_\varepsilon)} = \lim_{h\to 0^+} \frac{u_\varepsilon(x+h\xi_\varepsilon) - u_\varepsilon(x)}{F(x, h\xi_\varepsilon)} \\ &= \lim_{h\to 0^+} \frac{\int\limits_{\mathbb{R}^N} \rho_\varepsilon(t)\left(u(x+h\xi_\varepsilon - t) - u(x-t)\right)\,\mathrm{d}t}{F(x,h\xi_\varepsilon)}. \end{aligned}$$

This implies that

$$\begin{aligned} F^*(x, Du_\varepsilon(x)) &\le \limsup_{h\to 0^+} \frac{\int\limits_{\mathbb{R}^N} \rho_\varepsilon(t) d_F(x-t, x-t+h\xi_\varepsilon)\,\mathrm{d}t}{F(x,h\xi_\varepsilon)} \\ &\le \lim_{h\to 0^+} \frac{\int\limits_{\mathbb{R}^N} \rho_\varepsilon(t)\int\limits_0^1 F(x-t+\tau h\xi_\varepsilon, h\xi_\varepsilon)\,\mathrm{d}\tau\,\mathrm{d}t}{F(x,h\xi_\varepsilon)} \\ &= \frac{\int\limits_{\mathbb{R}^N} \rho_\varepsilon(t) F(x-t,\xi_\varepsilon)\,\mathrm{d}t}{F(x,\xi_\varepsilon)}. \end{aligned} \tag{4.3}$$

On the other hand, there is a sequence $\varepsilon_n \to 0$ such that

$$\limsup_{\varepsilon\to 0} F^*(x, Du_\varepsilon(x)) = \lim_{\varepsilon_n\to 0} F^*(x, Du_{\varepsilon_n}(x)). \tag{4.4}$$

Since $\|\xi_{\varepsilon_n}\| = 1$, up to a subsequence of $\{\xi_{\varepsilon_n}\}$, we can assume moreover that

$$\xi_{\varepsilon_n} \to \xi \ \text{ as } \ \varepsilon_n \to 0. \tag{4.5}$$

Thanks to (4.3), we get

$$F^*(x, Du_{\varepsilon_n}(x)) \le \frac{\int\limits_{\mathbb{R}^N} \rho_{\varepsilon_n}(t) F(x-t,\xi_{\varepsilon_n})\,\mathrm{d}t}{F(x,\xi_{\varepsilon_n})}. \tag{4.6}$$

Let $\varepsilon_n \to 0$, using (4.4), (4.6) and (4.5), we obtain

$$\limsup_{\varepsilon\to 0} F^*(x, Du_\varepsilon(x)) = \lim_{\varepsilon_n\to 0} F^*(x, Du_{\varepsilon_n}(x)) \le \lim_{\varepsilon_n\to 0} \frac{\int\limits_{\mathbb{R}^N} \rho_{\varepsilon_n}(t) F(x-t,\xi_{\varepsilon_n})\,\mathrm{d}t}{F(x,\xi_{\varepsilon_n})} = 1. \qquad \square$$

Remark 4.14. The lower semicontinuity of F is not enough to hold (4.2). Indeed, take the lower semicontinuous, non-degenerate Finsler metric F on $\mathbb{R}$ defined by

$$F(x,\xi) = \begin{cases} |\xi| & \text{if } x \le 0 \\ 2|\xi| & \text{if } x > 0 \end{cases} \qquad \text{for } x, \xi \in \mathbb{R}$$

and u is 1-Lipschitz w.r.t. d_F given by

$$u(x) = \begin{cases} x \ \text{ if } \ x \le 0 \\ 2x \ \text{ if } \ x > 0 \end{cases} \qquad \text{for } \ x \in \mathbb{R}.$$

Then,

$$F^*(x,p) = \begin{cases} |p| \ \text{ if } \ x \le 0 \\ \dfrac{1}{2}|p| \ \text{ if } \ x > 0 \end{cases} \qquad \text{and } \ {u'}_\varepsilon(x) = \int\limits_{\mathbb{R}} \rho_\varepsilon(s) u'(x-s)\mathrm{d}s.$$

Therefore, ${u'}_\varepsilon(0) = \int\limits_{[s\ge 0]} \rho_\varepsilon(s)\mathrm{d}s + 2\int\limits_{[s<0]} \rho_\varepsilon(s)\mathrm{d}s = \dfrac{3}{2}$ and $F^*(0, {u'}_\varepsilon(0)) = \dfrac{3}{2} > 1$.

It is known that if $u \in C^1(\mathbb{R}^N)$ then

$$u(y) - u(x) \le d_F(x,y) \ \ \forall x,y \in \mathbb{R}^N (i.e., u \in Lip_{d_F}) \text{ if and only if } \ F^*(x, \nabla u(x)) \le 1 \ \ \forall x \in \mathbb{R}^N.$$

The latter is equivalent to

$$q \cdot \nabla u(x) \le F(x,q) \ \ \forall x \in \mathbb{R}^N, \ \forall q \in \mathbb{R}^N.$$

In the case where u is non-smooth, we have the characterization via the tangential gradient.

Lemma 4.2. *For any $u \in Lip_{d_F}$ and $\Phi \in \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)$ such that $\nabla \cdot \Phi \in \mathcal{M}_b(\mathbb{R}^N)$, we have*

$$\frac{\Phi}{|\Phi|}(x) \cdot \nabla_{|\Phi|} u(x) \le F(x, \frac{\Phi}{|\Phi|}(x)), \ |\Phi|\text{-}a.e. \ \ x \in \mathbb{R}^N.$$

Proof. Taking u_ε as in Lemma 4.1, for any Borel subset B, we have

$$\int\limits_B \frac{\Phi}{|\Phi|} \cdot \nabla_{|\Phi|} u \mathrm{d}|\Phi| = \lim_{\varepsilon\to 0} \int\limits_B \frac{\Phi}{|\Phi|}(x) \cdot \nabla u_\varepsilon(x) \mathrm{d}|\Phi| \le \limsup_{\varepsilon \to 0} \int\limits_B F^*(x, \nabla u_\varepsilon(x)) F(x, \frac{\Phi}{|\Phi|}(x)) \mathrm{d}|\Phi|(x).$$

Letting $\varepsilon \to 0$, using again Fatou's Lemma and Lemma 4.1, we get

$$\int\limits_B \frac{\Phi}{|\Phi|}(x) \cdot \nabla_{|\Phi|} u(x) \mathrm{d}|\Phi| \le \int\limits_B F(x, \frac{\Phi}{|\Phi|}(x)) \mathrm{d}|\Phi|.$$

The proof ends up by the arbitrariness of Borel set B. □

Proof of Proposition 4.13. The case $\lambda = 0$ is obvious. Let us now assume that $\lambda > 0$.
1. We show that

$$\max_{u \in L^\lambda_{d_F}} \int\limits_{\mathbb{R}^N} u \mathrm{d}(\nu - \mu) \le \inf \Big\{ \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x)) \mathrm{d}|\Phi| + \lambda \int\limits_{\mathbb{R}^N} \mathrm{d}\theta^1 : \ \ (\Phi, \theta^0, \theta^1) \in S \Big\}.$$

Fix any $u \in L^{\lambda}_{d_F}$ and $(\Phi, \theta^0, \theta^1) \in S$. Taking u as a test function in the equation $-\nabla \cdot \Phi = \nu - \theta^1 - (\mu - \theta^0)$, using Lemma 4.2, we have

$$\begin{aligned}\int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) &= \int\limits_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\nabla_{|\Phi|}u\mathrm{d}|\Phi| + \int\limits_{\mathbb{R}^N} u\mathrm{d}\theta^1 - \int\limits_{\mathbb{R}^N} u\mathrm{d}\theta^0 \\ &\le \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda \int\limits_{\mathbb{R}^N} \mathrm{d}\theta^1.\end{aligned}$$

So

$$\sup_{u\in L^{\lambda}_{d_F}} \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) \le \inf\Big\{ \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda \int\limits_{\mathbb{R}^N} \mathrm{d}\theta^1 : \ (\Phi, \theta^0, \theta^1) \in S\Big\}.$$

It is not difficult to see that the supremum is actually a maximum by the direct method.

2. Obviously, we have

$$\max\Big\{ \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in L^{\lambda}_{d_F}\Big\} \ge \sup\Big\{ \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in L^{\lambda}_{d_F},\ u \in C^{1,1}(\mathbb{R}^N)\Big\}.$$

It remains to show that

$$\begin{aligned}&\sup\Big\{ \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in C^{1,1}(\mathbb{R}^N)\bigcap L^{\lambda}_{d_F}\Big\} \\ &= \min\Big\{ \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda \int\limits_{\mathbb{R}^N} \mathrm{d}\theta^1 : \ (\Phi, \theta^0, \theta^1) \in S\Big\}.\end{aligned} \tag{4.7}$$

On the other hand,

$$\begin{aligned}&\sup\Big\{ \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in C^{1,1}(\mathbb{R}^N)\bigcap L^{\lambda}_{d_F}\Big\} \\ &= \sup\Big\{ \int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in C^{1,1}(\mathbb{R}^N),\ F^*(x, \nabla u(x)) \le 1,\ 0 \le u(x) \le \lambda\ \forall x \in \mathbb{R}^N\Big\} \\ &= -\inf_{u\in V}\Big\{\mathcal{F}(u) + \mathcal{G}(\Lambda u)\Big\},\end{aligned}$$

where

$$\mathcal{F}(u) := -\int\limits_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) \ \ \forall u \in V := C^{1,1}(\mathbb{R}^N)\bigcap C_b(\mathbb{R}^N),$$

$$\Lambda(u) = (\nabla u, -u, u) \in Z := C_b(\mathbb{R}^N;\mathbb{R}^N) \times C_b(\mathbb{R}^N) \times C_b(\mathbb{R}^N)$$

and, for all $(q, z, w) \in Z$,

$$\mathcal{G}(q, z, w) := \begin{cases} 0 \ \text{ if } \ z(x) \le 0, w(x) \le \lambda \ \text{ and } \ F^*(x, q(x)) \le 1 \ \forall x \in \mathbb{R}^N \\ +\infty \ \ \text{otherwise.}\end{cases}$$

We use the $W^{1,\infty}$-norm and L^∞-norm for the spaces V and Z, respectively, i.e.,

$$\|u\|_V := \|u\|_{L^\infty} + \|\nabla u\|_{L^\infty} \text{ and } \|(q,z,w)\|_Z := \|q\|_{L^\infty} + \|z\|_{L^\infty} + \|w\|_{L^\infty}.$$

Now, applying the Fenchel–Rockafellar dual theory (see e.g. [11], Chapter III, Theorem 4.1 with the choice $u_0 = \dfrac{\lambda}{2}$ there. Note that the assumption $\lambda > 0$ is used in this step), we have

$$\begin{aligned}&\inf_{u\in V} \mathcal{F}(u) + \mathcal{G}(\Lambda(u))\\ &= \max_{(\Phi,\theta^0,\theta^1)\in\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times\mathcal{M}_b(\mathbb{R}^N)\times\mathcal{M}_b(\mathbb{R}^N)} \left(-\mathcal{F}^*(-\Lambda^*(\Phi,\theta^0,\theta^1)) - \mathcal{G}^*(\Phi,\theta^0,\theta^1)\right).\end{aligned}$$

The proof of (4.7) is completed by computing explicitly the quantities in this maximization problem. For completeness, let us give the results of this computation.

- Since $\mathcal{F}$ is linear, $\mathcal{F}^*(-\Lambda^*(\Phi,\theta^0,\theta^1))$ is finite (and is always equal to 0) if and only if

$$\langle -\Lambda^*(\Phi,\theta^0,\theta^1), u\rangle = \mathcal{F}(u) = -\int_{\mathbb{R}^N} u \mathrm{d}(\nu-\mu) \text{ for any } u\in V,$$

or equivalently

$$\langle \Phi, \nabla u\rangle - \langle \theta^0, u\rangle + \langle \theta^1, u\rangle = \langle \nu-\mu, u\rangle \text{ for any } u\in V.$$

This means that

$$-\nabla\cdot\Phi = \nu - \theta^1 - (\mu-\theta^0) \text{ in } \mathcal{D}'(\mathbb{R}^N).$$

- For $\mathcal{G}^*(\Phi,\theta^0,\theta^1)$, we have

$$\begin{aligned}\mathcal{G}^*(\Phi,\theta^0,\theta^1) &= \sup_{q\in C_b(\mathbb{R}^N;\mathbb{R}^N):F^*(x,q(x))\le 1,\forall x\in\mathbb{R}^N} \langle\Phi,q\rangle + \sup_{z\in C_b(\mathbb{R}^N):z\le 0}\langle\theta^0,z\rangle + \sup_{w\in C_b(\mathbb{R}^N):w\le\lambda}\langle\theta^1,w\rangle\\ &= \begin{cases}\displaystyle\int_{\mathbb{R}^N} F(x,\frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda\int_{\mathbb{R}^N}\mathrm{d}\theta^1 & \text{if } \theta^0\ge 0 \text{ and } \theta^1\ge 0\\ +\infty & \text{otherwise.}\end{cases}\end{aligned}$$

□

Proposition 4.15. *Given μ, $\nu\in\mathcal{M}_b^+(\mathbb{R}^N)$ and $\lambda\ge 0$. We have that*
(i) if u and (Φ,θ^0,θ^1) are solutions for the duality (4.1) then $(\theta,\Phi,u) := (\theta^1-\theta^0,\Phi,u)$ is a solution to the OMK equation (P_λ). Moreover $\theta^+ = \theta^1$, $\theta^- = \theta^0$ if $\lambda > 0$.
(ii) Conversely, if (θ,Φ,u) is a solution to the OMK equation (P_λ) then u and $(\Phi,\theta^0,\theta^1) := (\Phi,\theta^-,\theta^+)$ are solutions for the duality (4.1).

Proof. (i) Let $u\in L^\lambda_{d_F}$ and $(\Phi,\theta^0,\theta^1)\in S$ be solutions for the duality (4.1). Then $(\theta,\Phi,u) := (\theta^1-\theta^0,\Phi,u)$ is a solution to the OMK equation (P_λ). Indeed, we have

$$\theta - \nabla\cdot\Phi = \nu-\mu \text{ in } \mathcal{D}'(\mathbb{R}^N)$$

and

$$\int u\mathrm{d}(\nu-\mu) = \int \frac{\Phi}{|\Phi|}\nabla_{|\Phi|}u\mathrm{d}|\Phi| + \int u\mathrm{d}\theta^1 - \int u\mathrm{d}\theta^0$$
$$\leq \int F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda \int \mathrm{d}\theta^1 \text{ (by Lemma 4.2)}.$$

From the optimality of u and $(\Phi, \theta^0, \theta^1)$, using Proposition 4.13, we have that

$$\frac{\Phi}{|\Phi|}(x)\nabla_{|\Phi|}u(x) = F(x, \frac{\Phi}{|\Phi|}(x)), |\Phi|\text{-a.e. } x,$$

$$u = 0,\ \theta^0\text{-a.e. and } u = \lambda,\ \theta^1\text{-a.e..}$$

By the Jordan decomposition, we get $\theta^+ \leq \theta^1$, $\theta^- \leq \theta^0$ and thus

$$u = 0,\ \theta^-\text{-a.e. and } u = \lambda,\ \theta^+\text{-a.e..}$$

So (θ, Φ, u) is a solution to the OMK equation (P_λ). It remains to verify that $\theta^- = \theta^0$ and $\theta^+ = \theta^1$ in the case $\lambda > 0$. Since $\lambda > 0$, we deduce that θ^0 and θ^1 are concentrated on two disjoint sets. Thus $\theta^+ = \theta^1$ and $\theta^- = \theta^0$ by the Jordan decomposition.
(ii) Conversely, assume that (θ, Φ, u) is a solution to the OMK equation (P_λ). We see that

$$\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) = \int_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\nabla_{|\Phi|}u\mathrm{d}|\Phi| + \int_{\mathbb{R}^N} u\mathrm{d}\theta$$
$$= \int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \int_{\mathbb{R}^N} \lambda\mathrm{d}\theta^+.$$

The optimality of u and $(\Phi, \theta^-, \theta^+)$ follows immediately from the duality (4.1). □

We have the following estimates for solutions θ of the OMK equation.

Proposition 4.16 (Estimate of solutions θ)**.** *Given $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ and $\lambda \geq 0$. Let (θ, Φ, u) be a solution to the OMK equation (P_λ). Then*

$$\theta^- \leq \mu - \mu\wedge\nu \leq \mu \ \text{ and } \ \theta^+ \leq \nu - \mu\wedge\nu \leq \nu.$$

Proof. Case 1: If $\lambda = 0$, then $u \equiv 0$, $\Phi \equiv 0$ and

$$\theta \equiv \nu - \mu = \nu - \mu\wedge\nu - (\mu - \mu\wedge\nu).$$

By the Jordan decomposition, $\theta^+ \leq \nu - \mu\wedge\nu$ and $\theta^- \leq \mu - \mu\wedge\nu$.
Case 2: Let us now assume that $\lambda > 0$. For $0 < \varepsilon < \lambda$, let us consider the Lipschitz continuous functions of one variable

$$T_\varepsilon^1(r) := \begin{cases} 0 & \text{if } r \leq \lambda - \varepsilon \\ \dfrac{r - (\lambda - \varepsilon)}{\varepsilon} & \text{if } \lambda - \varepsilon \leq r \leq \lambda, \\ 1 & \text{if } r \geq \lambda \end{cases} \quad \forall r \in \mathbb{R}.$$

For $\xi \in C_c^\infty(\mathbb{R}^N)$ such that $\xi \geq 0$, we take $T_\varepsilon^1(u)\xi$ as a test function in the equation $\theta - \nabla \cdot \Phi = \nu - \mu$. We get

$$\int T_\varepsilon^1(u)\xi \mathrm{d}\theta + \int \frac{\Phi}{|\Phi|} \cdot \nabla_{|\Phi|} \left(T_\varepsilon^1(u)\xi\right) \mathrm{d}|\Phi| = \int T_\varepsilon^1(u)\xi \mathrm{d}(\nu - \mu). \tag{4.8}$$

Thanks to Lemma 7.5 (see Appendix), we get

$$\begin{aligned} \int \frac{\Phi}{|\Phi|} \cdot \nabla_{|\Phi|} \left(T_\varepsilon^1(u)\xi\right) \mathrm{d}|\Phi| &= \int (T_\varepsilon^1)'(u) \nabla_{|\Phi|} u \cdot \frac{\Phi}{|\Phi|} \xi \mathrm{d}|\Phi| + \int \frac{\Phi}{|\Phi|} \cdot \nabla \xi T_\varepsilon^1(u) \mathrm{d}|\Phi| \\ &\geq \int \frac{\Phi}{|\Phi|} \cdot \nabla \xi T_\varepsilon^1(u) \mathrm{d}|\Phi|. \end{aligned} \tag{4.9}$$

Using (4.8) and (4.9), we see that

$$\begin{aligned} \int T_\varepsilon^1(u)\xi \mathrm{d}\theta + \int \frac{\Phi}{|\Phi|} \cdot \nabla \xi T_\varepsilon^1(u) \mathrm{d}|\Phi| &\leq \int T_\varepsilon^1(u)\xi \mathrm{d}(\nu - \mu) \\ &= \int T_\varepsilon^1(u)\xi \mathrm{d}(\nu - \mu \wedge \nu - (\mu - \mu \wedge \nu)) \\ &\leq \int T_\varepsilon^1(u)\xi \mathrm{d}(\nu - \mu \wedge \nu). \end{aligned} \tag{4.10}$$

Since $u \leq \lambda$, for any $x \in \mathbb{R}^N$, we have

$$T_\varepsilon^1(u)(x) \to \chi_{[u=\lambda]}(x) \quad \text{as } \varepsilon \to 0.$$

Now, using Lemma 7.5 (ii), the non-degeneracy of F and the definition of solution for (P_λ), we have $|\Phi|([u = \lambda]) = 0$ (i.e. Φ gives no mass on the set $[u = \lambda]$). This implies that

$$\int \frac{\Phi}{|\Phi|} \cdot \nabla \xi T_\varepsilon^1(u) \mathrm{d}|\Phi| \to 0 \quad \text{as } \varepsilon \to 0.$$

Letting $\varepsilon \to 0$ in (4.10), we get

$$\int\limits_{[u=\lambda]} \xi \mathrm{d}\theta \leq \int\limits_{[u=\lambda]} \xi \mathrm{d}(\nu - \mu \wedge \nu) \quad \text{for any } \xi \in C_c^\infty(\mathbb{R}^N),\ \xi \geq 0.$$

Using the definition of solution for (P_λ), we have $u = 0$ for θ^--a.e.. Since $\lambda > 0$, we deduce that

$$\int\limits_{[u=\lambda]} \xi \mathrm{d}\theta^+ = \int\limits_{[u=\lambda]} \xi \mathrm{d}\theta \leq \int\limits_{[u=\lambda]} \xi \mathrm{d}(\nu - \mu \wedge \nu) \quad \text{for any } \xi \in C_c^\infty(\mathbb{R}^N),\ \xi \geq 0.$$

This implies that $\theta^+ \leq \nu - \mu \wedge \nu$ on $[u = \lambda]$ and that $\theta^+ \leq \nu - \mu \wedge \nu$ (since θ^+ is concentrated on $[u = \lambda]$). At last, using $T_\varepsilon^2(u)\xi$ as a test function in the equation $\theta - \nabla \cdot \Phi = \nu - \mu$, where

$$T_\varepsilon^2(r) := \begin{cases} -1 & \text{if } r \leq 0 \\ -1 + \dfrac{r}{\varepsilon} & \text{if } 0 \leq r \leq \varepsilon, \\ 0 & \text{if } r \geq \varepsilon \end{cases} \quad \forall r \in \mathbb{R},$$

we can prove in the same way that $\theta^- \leq \mu - \mu \wedge \nu$. □

Proof of Theorem 2.6. The proof follows directly from Propositions 4.13, 4.15 and 4.16. □

As a consequence of Proposition 4.13, we have the following result that will be useful later.

Corollary 4.17. *Let $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be such that $\nu(\mathbb{R}^N) \leq \mu(\mathbb{R}^N)$. We have*

$$\sup\Big\{\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in Lip_{d_F}, u \geq 0\Big\}$$
$$= \min_{(\Phi,\theta^0)\in\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times\mathcal{M}_b^+(\mathbb{R}^N)} \Big\{\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| : -\nabla\cdot\Phi = \nu - (\mu-\theta^0)\Big\}.$$

Proof. Using the assumption $\nu(\mathbb{R}^N) \leq \mu(\mathbb{R}^N)$, there exists $(\tilde{\Phi}, \tilde{\theta}^0) \in \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N)$ such that $-\nabla\cdot\tilde{\Phi} = \nu - (\mu - \tilde{\theta}^0)$. This implies that

$$\inf_{(\Phi,\theta^0)\in\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times\mathcal{M}_b^+(\mathbb{R}^N)} \Big\{\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| : -\nabla\cdot\Phi = \nu - (\mu-\theta^0)\Big\} := C < +\infty.$$

Now, taking u as a test function in the equation $-\nabla\cdot\Phi = \nu - (\mu-\theta^0)$, we get

$$\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) = \int_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}\nabla_{|\Phi|}u\mathrm{d}|\Phi| - \int_{\mathbb{R}^N} u\mathrm{d}\theta^0 \leq \int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi|.$$

Hence,

$$\begin{aligned}&\sup\Big\{\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in Lip_{d_F}, u \geq 0\Big\}\\ &\leq \inf_{(\Phi,\theta^0)\in\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)\times\mathcal{M}_b^+(\mathbb{R}^N)} \Big\{\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| : -\nabla\cdot\Phi = \nu - (\mu-\theta^0)\Big\} = C < +\infty.\end{aligned} \tag{4.11}$$

Conversely, let us consider a sequence $\lambda_n \to +\infty$ as $n \to +\infty$. Thanks to Proposition 4.13, there exist $u_n \in L_{d_F}^{\lambda_n}$ and $(\Phi_n, \theta_n^0, \theta_n^1) \in S$ such that

$$\begin{aligned}\int_{\mathbb{R}^N} F(x, \frac{\Phi_n}{|\Phi_n|}(x))\mathrm{d}|\Phi_n| + \lambda_n \int_{\mathbb{R}^N} \mathrm{d}\theta_n^1 &= \int_{\mathbb{R}^N} u_n\mathrm{d}(\nu-\mu)\\ &\leq \sup\Big\{\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in Lip_{d_F}, u \geq 0\Big\} \leq C.\end{aligned} \tag{4.12}$$

It is not difficult to see that $\big\{(\Phi_n, \theta_n^0, \theta_n^1)\big\}$ is bounded in $\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N) \times \mathcal{M}_b(\mathbb{R}^N) \times \mathcal{M}_b(\mathbb{R}^N)$. Thus, up to a subsequence, $(\Phi_n, \theta_n^0, \theta_n^1)$ converges to some $(\Phi, \theta^0, \theta^1)$ weakly* in $\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N) \times \mathcal{M}_b(\mathbb{R}^N) \times \mathcal{M}_b(\mathbb{R}^N)$. It is clear that $\theta^1 = 0$, $\theta^0 \geq 0$ and $-\nabla\cdot\Phi = \nu - (\mu-\theta^0)$. Now, using the lower semicontinuity of the functional $\int F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi|$ w.r.t. the weak* convergence in variable Φ (see e.g. [2, Theorem 2.38]) and passing to the limit in (4.12), we obtain

$$\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| \leq \lim_{n\to+\infty} \int_{\mathbb{R}^N} F(x, \frac{\Phi_n}{|\Phi_n|}(x))\mathrm{d}|\Phi_n| \leq \sup\Big\{ \int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) : u \in Lip_{d_F}, u \geq 0\Big\}.$$

The proof is completed by combining this with (4.11). □

4.2. Uniqueness of solution θ to the OMK equation. We usually identify a measure with its density function with respect to Lebesgue measure $\mathcal{L}^N$ when the measure is absolutely continuous w.r.t. $\mathcal{L}^N$. In this subsection, we focus on the proof of the uniqueness for the solution θ of the OMK equation (P_λ) which is then used to show the uniqueness of active submeasures for the PMK problem. The result of uniqueness is somehow optimal in view of Theorem 2.8 and Remark 2.11 (ii). Our proof will be based on the doubling and de-doubling variable technique (the technique was known in PDEs, due to [20], see also [8] and the references therein). It uses mainly the following result.

Lemma 4.3. *Let $\lambda \geq 0$ and $\mu, \nu \in L^1(\mathbb{R}^N)^+$. Suppose that (θ_i, Φ_i, u_i), $i = 1, 2$, are solutions to the same OMK equation (P_λ). Then $\theta_1, \theta_2 \in L^1(\mathbb{R}^N)$ and, for any $\xi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}^N)$ such that $\xi \geq 0$, we have*

$$\begin{aligned}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} (\theta_1(x) - \theta_2(y))^+\xi(x,y)\mathrm{d}x\mathrm{d}y \leq &\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\left(\nabla_x\xi + \nabla_y\xi\right)|\mathrm{d}|\Phi_1|(x)\mathrm{d}y \\ &+ \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\left(\nabla_x\xi + \nabla_y\xi\right)|\mathrm{d}|\Phi_2|(y)\mathrm{d}x \\ &+ \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |(\nu-\mu)(x) - (\nu-\mu)(y)|\xi(x,y)\,\mathrm{d}x\mathrm{d}y.\end{aligned} \tag{4.13}$$

Before giving the proof of this lemma, let us show how it enables us to prove the main result of uniqueness in Subsection 2.3.

Proof of Theorem 2.7. Fixed any $\alpha \in C_c^\infty(\mathbb{R}^N)$, $\alpha \geq 0$, let us choose

$$\xi_\varepsilon(x,y) := \rho_\varepsilon(x-y)\alpha(x+y)$$

as test functions in (4.13). Note that $\nabla_x\xi_\varepsilon + \nabla_y\xi_\varepsilon = 2\rho_\varepsilon(x-y)\nabla\alpha(x+y)$. We have

$$\begin{aligned}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\nabla_x\xi_\varepsilon + \nabla_y\xi_\varepsilon|\mathrm{d}|\Phi_1|(x)\mathrm{d}y &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\nabla_x\xi_\varepsilon + \nabla_y\xi_\varepsilon|\mathrm{d}y\mathrm{d}|\Phi_1|(x) \\ &= 2\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \rho_\varepsilon(x-y)|\nabla\alpha(x+y)|\,\mathrm{d}y\mathrm{d}|\Phi_1|(x) \\ &= 2\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \rho_\varepsilon(t)|\nabla\alpha(2x-t)|\,\mathrm{d}t\mathrm{d}|\Phi_1|(x)\end{aligned}$$

$$\to 2\int\limits_{\mathbb{R}^N} |\nabla\alpha(2x)|\mathrm{d}|\Phi_1|(x).$$

Similarly, $\int\limits_{\mathbb{R}^N}\int\limits_{\mathbb{R}^N} |\nabla_x\xi_\varepsilon + \nabla_y\xi_\varepsilon|\mathrm{d}|\Phi_2|(y)\mathrm{d}x \to 2\int\limits_{\mathbb{R}^N} |\nabla\alpha(2y)|\mathrm{d}|\Phi_2|(y)$. Next, since $f := \nu - \mu \in L^1$, we have

$$\begin{aligned}\int\limits_{\mathbb{R}^N}\int\limits_{\mathbb{R}^N} |f(x)-f(y)|\xi_\varepsilon(x,y)\,\mathrm{d}x\mathrm{d}y &= \int\limits_{\mathbb{R}^N}\int\limits_{\mathbb{R}^N} |f(x)-f(y)|\rho_\varepsilon(x-y)\alpha(x+y)\,\mathrm{d}x\mathrm{d}y\\ &\le \|\alpha\|_\infty \int\limits_{\mathbb{R}^N}\int\limits_{\mathbb{R}^N} |f(x)-f(y)|\rho_\varepsilon(x-y)\,\mathrm{d}y\mathrm{d}x\\ &= \|\alpha\|_\infty \int\limits_{\mathbb{R}^N}\int\limits_{\mathbb{R}^N} |f(x)-f(x-t)|\rho_\varepsilon(t)\,\mathrm{d}t\mathrm{d}x\\ &= \|\alpha\|_\infty \int\limits_{\mathbb{R}^N} F_\varepsilon(x)\mathrm{d}x \to 0,\end{aligned}$$

by the fact that $F_\varepsilon(x) := \int\limits_{\mathbb{R}^N} |f(x)-f(x-t)|\rho_\varepsilon(t)\,\mathrm{d}t$, and $F_\varepsilon \to 0$ in L^1. Thus (4.13) leads to

$$\int\limits_{\mathbb{R}^N} (\theta_1(x)-\theta_2(x))^+\alpha(2x)\mathrm{d}x \le 2\int\limits_{\mathbb{R}^N} |\nabla\alpha(2x)|\mathrm{d}|\Phi_1|(x) + 2\int\limits_{\mathbb{R}^N} |\nabla\alpha(2y)|\mathrm{d}|\Phi_2|(y). \tag{4.14}$$

Take a sequence $\alpha_n \in C_c^\infty(\mathbb{R}^N)$ such that $\chi_{B(0,n)} \le \alpha_n \le \chi_{B(0,n+1)}$ and $|\nabla\alpha_n| \le C$. Substituting α_n into (4.14) and letting $n \to +\infty$, using the finiteness of Φ_i, we get $\int\limits_{\mathbb{R}^N} (\theta_1(x)-\theta_2(x))^+\,\mathrm{d}x \le 0$. Hence $\theta_1 \le \theta_2$. Since θ_1 and θ_2 have the same role, we obtain $\theta_1 = \theta_2$. □

Now, we give the proof of Lemma 4.3. Let us consider the Lipschitz continuous function

$$H_\varepsilon(r) := \min(r^+/\varepsilon, 1) \quad \text{for any } r \in \mathbb{R}.$$

Proof of Lemma 4.3. Thanks to Proposition 4.16, we have $\theta_1, \theta_2 \in L^1(\mathbb{R}^N)$. Let us consider the test functions $\xi_\varepsilon(x,y) := H_\varepsilon(u_1(x) - u_2(y) + \varepsilon\rho(x,y))\xi(x,y)$ where $\xi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}^N)$, $\xi \ge 0$, $\rho \in C^\infty(\mathbb{R}^N \times \mathbb{R}^N)$ and $0 \le \rho \le 1$. For each y, considering $\xi_\varepsilon(.,y)$ as a test function, we have

$$\begin{aligned}&\int\limits_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\theta_1(x)\,\mathrm{d}x\\ &+\int\limits_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x).\nabla_{|\Phi_1|,x}\left(H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\right)\,\mathrm{d}|\Phi_1|(x)\\ &= \int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\,(\nu-\mu)(x)\mathrm{d}x.\end{aligned}$$

Integrating with respect to y, we get

$$\begin{aligned}
&\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\theta_1(x)\,\mathrm{d}x\mathrm{d}y\\
&+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x).\nabla_{|\Phi_1|,x}\left(H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\right)\,\mathrm{d}|\Phi_1|(x)\mathrm{d}y\\
&=\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\,(\nu-\mu)(x)\mathrm{d}x\mathrm{d}y.
\end{aligned}\tag{4.15}$$

Similarly, applying for (θ_2,Φ_2,u_2), we get

$$\begin{aligned}
&\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\theta_2(y)\,\mathrm{d}y\mathrm{d}x\\
&+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_2}{|\Phi_2|}(y).\nabla_{|\Phi_2|,y}\left(H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\right)\,\mathrm{d}|\Phi_2|(y)\mathrm{d}x\\
&=\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\,(\nu-\mu)(y)\mathrm{d}y\mathrm{d}x.
\end{aligned}\tag{4.16}$$

From (4.15) and (4.16), we have

$$I_1(\varepsilon)+I_2(\varepsilon)+I_3(\varepsilon)=0,\tag{4.17}$$

where

$$\begin{aligned}
I_1(\varepsilon):=&\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)(\theta_1(x)-\theta_2(y))\,\mathrm{d}x\mathrm{d}y\\
&-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\,(\nu-\mu)(x)\mathrm{d}x\mathrm{d}y\\
&+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\,(\nu-\mu)(y)\mathrm{d}x\mathrm{d}y;
\end{aligned}$$

$$I_2(\varepsilon):=\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x).\nabla_{|\Phi_1|,x}\left(H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\right)\,\mathrm{d}|\Phi_1|(x)\mathrm{d}y;$$

$$\text{and }\ I_3(\varepsilon):=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_2}{|\Phi_2|}(y).\nabla_{|\Phi_2|,y}\left(H_\varepsilon(u_1(x)-u_2(y)+\varepsilon\rho(x,y))\xi(x,y)\right)\,\mathrm{d}|\Phi_2|(y)\mathrm{d}x.$$

Recall that

$$\int_{\mathbb{R}^N}\nabla g(x)dx=0\ \text{ for any }\ g\in Lip(\mathbb{R}^N)\cap C_c(\mathbb{R}^N).\tag{4.18}$$

For short, in the following computation, we denote by $H_\varepsilon := H_\varepsilon(u_1(x) - u_2(y) + \varepsilon\rho(x,y))$ and $H_\varepsilon^{'} := H_\varepsilon^{'}(u_1(x) - u_2(y) + \varepsilon\rho(x,y))$. Using the chain rule in Lemma 7.5, we have

$$\begin{aligned} I_2(\varepsilon) &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)\left(\nabla_x\xi H_\varepsilon + \nabla_{|\Phi_1|}u_1 H_\varepsilon^{'}\xi + \varepsilon\nabla_x\rho H_\varepsilon^{'}\xi\right)\mathrm{d}|\Phi_1|(x)\mathrm{d}y \\ &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)\left((\nabla_x\xi + \nabla_y\xi)H_\varepsilon + (\nabla_{|\Phi_1|}u_1 - \nabla u_2(y))H_\varepsilon^{'}\xi + \varepsilon(\nabla_x\rho + \nabla_y\rho)H_\varepsilon^{'}\xi\right)\mathrm{d}|\Phi_1|(x)\mathrm{d}y \\ &\geq \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)\left(\nabla_x\xi + \nabla_y\xi\right)H_\varepsilon\mathrm{d}|\Phi_1|(x)\mathrm{d}y + \varepsilon\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)(\nabla_x\rho + \nabla_y\rho)H_\varepsilon^{'}\xi\mathrm{d}|\Phi_1|(x)\mathrm{d}y, \end{aligned} \tag{4.19}$$

where, in the second equality, we used (4.18) and the fact that $\xi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}^N)$ as follows:

$$\begin{aligned} &\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)\left(\nabla_y\xi H_\varepsilon - \nabla u_2(y)H_\varepsilon^{'}\xi + \varepsilon\nabla_y\rho H_\varepsilon^{'}\xi\right)\mathrm{d}|\Phi_1|(x)\mathrm{d}y \\ &= \int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)\int_{\mathbb{R}^N}\nabla_y\left(H_\varepsilon(u_1(x) - u_2(y) + \varepsilon\rho(x,y))\xi(x,y)\right)\mathrm{d}y\mathrm{d}|\Phi_1|(x) = 0. \end{aligned}$$

On the other hand,

$$\begin{aligned} &\varepsilon\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)(\nabla_x\rho + \nabla_y\rho)H_\varepsilon^{'}(u_1(x) - u_2(y) + \varepsilon\rho(x,y))\xi\mathrm{d}|\Phi_1|(x)\mathrm{d}y \\ &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)(\nabla_x\rho + \nabla_y\rho)\chi_{[-\varepsilon\rho\leq u_1(x)-u_2(y)\leq\varepsilon(1-\rho)]}\xi\mathrm{d}|\Phi_1|(x)\mathrm{d}y \to 0. \end{aligned} \tag{4.20}$$

Indeed, since Φ_1 gives no mass on the set $[u_1 = u_2(y)]$, for each y (using again Lemma 7.5 (ii), the non-degeneracy of F and the definition of solution for (P_λ)),

$$\begin{aligned} F_\varepsilon(y) &:= \int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)(\nabla_x\rho + \nabla_y\rho)\chi_{[-\varepsilon\rho\leq u_1(x)-u_2(y)\leq\varepsilon(1-\rho)]}\xi\mathrm{d}|\Phi_1|(x) \\ &\to \int_{\mathbb{R}^N} \frac{\Phi_1}{|\Phi_1|}(x)(\nabla_x\rho + \nabla_y\rho)\chi_{[u_1(x)=u_2(y)]}\xi\mathrm{d}|\Phi_1|(x) = 0; \end{aligned}$$

and moreover,

$$|F_\varepsilon(y)| \leq \int_{\mathbb{R}^N} |(\nabla_x\rho + \nabla_y\rho)|\xi\mathrm{d}|\Phi_1|(x) \in L^1(\mathbb{R}^N).$$

Using the Lebesgue Dominated Convergence Theorem gives (4.20). Next, by (4.19) and (4.20),

$$\liminf_\varepsilon I_2 \geq -\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\left(\nabla_x\xi + \nabla_y\xi\right)|\mathrm{d}|\Phi_1|(x)\mathrm{d}y. \tag{4.21}$$

In the same way, we have

$$\liminf_{\varepsilon} I_3 \geq -\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |\left(\nabla_x\xi + \nabla_y\xi\right)| \mathrm{d}|\Phi_2|(y)\mathrm{d}x. \tag{4.22}$$

Concerning $I_1(\varepsilon)$, we have the convergence in pointwise (x, y),

$$H_\varepsilon(u_1(x) - u_2(y) + \varepsilon\rho(x, y)) \to Sign_0^+(u_1(x) - u_2(y)) + \rho(x, y)\chi_{[u_1(x)=u_2(y)]},$$

where

$$Sign_0^+(r) = \begin{cases} 1 \text{ if } r > 0 \\ 0 \text{ if } r \leq 0. \end{cases}$$

Since $\nu - \mu \in L^1$, then

$$\begin{aligned} I_1(\varepsilon) \to & \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} (\theta_1(x) - \theta_2(y)) \left(Sign_0^+(u_1(x) - u_2(y)) + \rho(x, y)\chi_{[u_1(x)=u_2(y)]}\right) \xi \,\mathrm{d}x\mathrm{d}y \\ & - \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} ((\nu - \mu)(x) - (\nu - \mu)(y)) \left(Sign_0^+(u_1(x) - u_2(y)) + \rho(x, y)\chi_{[u_1(x)=u_2(y)]}\right) \xi \,\mathrm{d}x\mathrm{d}y \\ & \geq \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} (\theta_1(x) - \theta_2(y)) \left(Sign_0^+(u_1(x) - u_2(y)) + \rho(x, y)\chi_{[u_1(x)=u_2(y)]}\right) \xi \,\mathrm{d}x\mathrm{d}y \\ & - \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |(\nu - \mu)(x) - (\nu - \mu)(y)|\xi(x, y) \,\mathrm{d}x\mathrm{d}y, \end{aligned}$$

where we used the assumption $0 \leq \rho(x, y) \leq 1$ and therefore

$$Sign_0^+(u_1(x) - u_2(y)) + \rho(x, y)\chi_{[u_1(x)=u_2(y)]} \leq 1.$$

Now, by density, we can choose $\rho(x, y) := \mathrm{Sign}_0^+(\theta_1(x) - \theta_2(y))$, so that

$$\liminf_{\varepsilon} I_1(\varepsilon) \geq \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} (\theta_1(x) - \theta_2(y))^+\xi - \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |(\nu - \mu)(x) - (\nu - \mu)(y)|\xi(x, y) \,\mathrm{d}x\mathrm{d}y.$$

Combining this with (4.17), (4.21) and (4.22), we obtain Lemma 4.3. □

5. OMK equation vs active submeasures

5.1. Partial minimum flow problem. Recall that in the connection between balanced MK problem and the Monge–Kantorovich equation the so called minimum flow problem is a key ingredient. For the PMK problem, the definition of minimum flow problem, that we call here the partial minimum flow problem, as well as its connection with the PMK problem are given in the following proposition.

Proposition 5.18 (Partial minimum flow problem)**.** *Let $\mu,\ \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be compactly supported. For any $\mathbf{m} \in [0, \mathbf{m}_{\max}]$, we have*

$$\begin{aligned}\min\{\mathcal{K}(\sigma):\ \sigma \in \pi_{\mathbf{m}}(\mu,\nu)\} &= \max\left\{\mathcal{D}(\lambda,u) : (\lambda,u) \in \mathbb{R}^+ \times L^\lambda_{d_F}\right\} \\ &= \min\left\{\int_{\mathbb{R}^N} F\left(x, \frac{\Phi}{|\Phi|}(x)\right)\, \mathrm{d}|\Phi|(x) :\ (\Phi,\theta^0,\theta^1) \in \Psi_{\mathbf{m}}(\mu,\nu)\right\},\end{aligned} \tag{5.1}$$

where

$$\begin{aligned}\Psi_{\mathbf{m}}(\mu,\nu) := \Big\{&(\Phi,\theta^0,\theta^1) \in \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N) \times \mathcal{M}_b^+(\mathbb{R}^N) :\ \theta^0(\mathbb{R}^N) = \mu(\mathbb{R}^N) - \mathbf{m}, \\ &\theta^1(\mathbb{R}^N) = \nu(\mathbb{R}^N) - \mathbf{m} \ \ \textit{and} \ \ -\nabla\cdot\Phi = \nu - \theta^1 - (\mu - \theta^0) \ \ \textit{in} \ \ \mathcal{D}'(\mathbb{R}^N)\Big\}.\end{aligned}$$

The last minimization problem in (5.1) is called the *partial minimum flow* (PMF). We have an immediate consequence.

Corollary 5.19. *If $(\Phi,\theta^0,\theta^1) \in \Psi_{\mathbf{m}}(\mu,\nu)$ is an optimal solution to the PMF problem and $\theta^0 \le \mu$, $\theta^1 \le \nu$ then $\rho_0 := \mu - \theta^0$ and $\rho_1 := \nu - \theta^1$ are active submeasures of the PMK problem. Conversely, if ρ_0 and ρ_1 are active submeasures to the PMK problem then there is a vector measure Φ such that $(\Phi,\theta^0,\theta^1) := (\Phi, \mu - \rho_0, \nu - \rho_1)$ is a solution to the PMF problem.*

Note that we do not have any constraints of type $\theta^0 \le \mu$ or $\theta^1 \le \nu$ in the definition of the PMF problem. However, following Theorem 5.20 and Proposition 4.16, these constraints are automatically satisfied for any optimal solutions (Φ,θ^0,θ^1) whenever $\mathbf{m} \in [(\mu\wedge\nu)(\mathbb{R}^N), \mathbf{m}_{\max}]$. The case $\mathbf{m} < (\mu\wedge\nu)(\mathbb{R}^N)$ is not interesting for the optimal partial transport problem because of the obviousness of solutions.

Proof of Proposition 5.18. The first equality has been shown in Theorem 2.4. Let us prove the second equality. First, for any $(\lambda,u) \in \mathbb{R}^+ \times L^\lambda_{d_F}$ and a triplet $(\Phi,\theta^0,\theta^1) \in \Psi_{\mathbf{m}}(\mu,\nu)$, using Lemma 4.2, we have

$$\begin{aligned}\int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) + \lambda(\mathbf{m} - \nu(\mathbb{R}^N)) &= \int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) - \lambda\int_{\mathbb{R}^N}\mathrm{d}\theta^1 \\ &\le \int_{\mathbb{R}^N} u\mathrm{d}(\nu-\mu) + \int_{\mathbb{R}^N} u\mathrm{d}\theta^0 - \int_{\mathbb{R}^N} u\mathrm{d}\theta^1 \\ &= \int_{\mathbb{R}^N} \nabla_{|\Phi|}u(x)\frac{\Phi}{|\Phi|}(x)\mathrm{d}|\Phi| \le \int_{\mathbb{R}^N} F(x,\frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi|.\end{aligned}$$

This shows that

$$\max\left\{\mathcal{D}(\lambda,u) : (\lambda,u) \in \mathbb{R}^+ \times L^\lambda_{d_F}\right\} \le \inf\left\{\int_{\mathbb{R}^N} F\left(x, \frac{\Phi}{|\Phi|}(x)\right)\, \mathrm{d}|\Phi|(x) :\ (\Phi,\theta^0,\theta^1) \in \Psi_{\mathbf{m}}(\mu,\nu)\right\}.$$

Now, let (ρ_0, ρ_1) be a couple of active submeasures for the PMK problem w.r.t. $\mathbf{m}$. Thanks to Corollary 4.17, there exists $\Phi \in \mathcal{M}_b(\mathbb{R}^N; \mathbb{R}^N)$ such that $-\nabla \cdot \Phi = \rho_1 - \rho_0$ and

$$\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| = \sup\Big\{ \int_{\mathbb{R}^N} u\mathrm{d}(\rho_1 - \rho_0) : u \in Lip_{d_F}, u \geq 0\Big\}$$
$$= \min\left\{\mathcal{K}(\sigma) :\ \sigma \in \pi_{\mathbf{m}}(\rho_0, \rho_1)\right\}.$$

Let us set

$$\theta^0 := \mu - \rho_0 \ \text{ and } \ \theta^1 := \nu - \rho_1.$$

Then $(\Phi, \theta^0, \theta^1) \in \Psi_{\mathbf{m}}(\mu, \nu)$ and

$$\int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| = \min\left\{\mathcal{K}(\sigma) :\ \sigma \in \pi_{\mathbf{m}}(\rho_0, \rho_1)\right\} = \max\Big\{\mathcal{D}(\lambda, u) : (\lambda, u) \in \mathbb{R}^+ \times L^{\lambda}_{d_F}\Big\}. \ \square$$

5.2. Link between the OMK equation and the PMK problem. The connection between the OMK equation and the PMK problem appears when we deal with the extremal condition between the PMF problem and DPMK problem. Roughly speaking, the optimality condition in the duality of the DPMK and PMF problems corresponds to (P_λ) for some λ.

Theorem 5.20. *Let $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be compactly supported.*
(i) Given $\mathbf{m} \in [0, \mathbf{m}_{\max}]$ and a solution $(\Phi, \theta^0, \theta^1)$ to the PMF problem and (λ, u) is a solution to the DPMK problem. Setting $\theta := \theta^1 - \theta^0$, the triplet (θ, Φ, u) is a solution to the OMK equation (P_λ). Moreover, $\theta^+ = \theta^1$ and $\theta^- = \theta^0$ if $\mathbf{m} \geq (\mu \wedge \nu)(\mathbb{R}^N)$.
(ii) Given $\lambda \geq 0$ and (θ, Φ, u) a solution to the OMK equation (P_λ). Then (λ, u) is a solution to the DPMK problem corresponding to $\mathbf{m} = (\mu - \theta^-)(\mathbb{R}^N)$ and $(\Phi, \theta^0, \theta^1) := (\Phi, \theta^-, \theta^+)$ is a solution to the associated PMF problem.

Proof. (i) From the optimality of $(\Phi, \theta^0, \theta^1)$ and of (λ, u), using Proposition 5.18, we have

$$\int_{\mathbb{R}^N} u\mathrm{d}(\nu - \mu) + \lambda(\mathbf{m} - \nu(\mathbb{R}^N)) = \int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi|,$$

or

$$\int_{\mathbb{R}^N} u\mathrm{d}(\nu - \mu) = \int_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| + \lambda \int_{\mathbb{R}^N} \mathrm{d}\theta^1.$$

Thanks to Proposition 4.13, we have that u and $(\Phi, \theta^0, \theta^1)$ are solutions for the duality (4.1). Using Proposition 4.15, we have that (θ, Φ, u) is a solution to the OMK equation (P_λ). Now, let us show that $\theta^+ = \theta^1$ and $\theta^- = \theta^0$ for the case $\mathbf{m} \geq (\mu \wedge \nu)(\mathbb{R}^N)$. We divide into two cases: If $\mathbf{m} = (\mu \wedge \nu)(\mathbb{R}^N)$, then the total cost of the associated optimal partial transport problem is zero. This implies that $\Phi \equiv 0$ and $\theta := \theta^1 - \theta^0 = \nu - \mu = \nu - \mu \wedge \nu - (\mu - \mu \wedge \nu)$. By the Jordan decomposition, we have

$$\theta^+ = \nu - \mu \wedge \nu \leq \theta^1 \quad \text{and} \quad \theta^- = \mu - \mu \wedge \nu \leq \theta^0.$$

Using the constraints on the total mass of θ^0 and of θ^1, we obtain

$$\theta^+ = \nu - \mu \wedge \nu = \theta^1 \quad \text{and} \quad \theta^- = \mu - \mu \wedge \nu = \theta^0.$$

If $\mathbf{m} > (\mu \wedge \nu)(\mathbb{R}^N)$ then $\lambda > 0$ and the conclusion follows from Proposition 4.15.
(ii) The proof is similar to the one of Proposition 4.15 (ii) with the use of the duality (5.1). □

We are now ready to give the proof of the connection between active submeasures and solutions θ of the OMK equation.

Proof of Theorem 2.8. First, let θ_λ be a solution of the OMK equation (P_λ). Thanks to Proposition 4.16, $0 \leq \mu - \theta_\lambda^- \leq \mu$ and $0 \leq \nu - \theta_\lambda^+ \leq \nu$. Then, using Theorem 5.20 (ii) and Corollary 5.19, we deduce that $\rho_0 := \mu - \theta_\lambda^-$ and $\rho_1 := \nu - \theta_\lambda^+$ are active submeasures.
Conversely, let $\mathbf{m} \in [(\mu \wedge \nu)(\mathbb{R}^N), \mathbf{m}_{\max}]$ and (ρ_0, ρ_1) be a couple of active submeasures. Let $(\lambda_{\mathbf{m}}, u_{\mathbf{m}})$ be a solution of the DPMK problem. Thanks to Corollary 5.19, there exists a flow Φ such that $(\Phi, \mu - \rho_0, \nu - \rho_1)$ is a solution of the corresponding PMF problem. And, thanks to Theorem 5.20 (i), $\theta_{\lambda_{\mathbf{m}}} := \nu - \rho_1 - \mu + \rho_0$ is a solution of the OMK equation $(P_{\lambda_{\mathbf{m}}})$ and

$$\theta_{\lambda_{\mathbf{m}}}^+ = \nu - \rho_1, \ \ \theta_{\lambda_{\mathbf{m}}}^- = \mu - \rho_0. \qquad \square$$

Thanks to the above connection, let us give the proof of the uniqueness of active submeasures by using the result of the OMK equation.

Proof of Corollary 2.9. Assume that (ρ_0, ρ_1) and $(\eta_0, \eta_1) \in Sub_{\mathbf{m}}(\mu, \nu)$ are two pairs of active submeasures. We will show that $\rho_0 = \eta_0$ and $\rho_1 = \eta_1$. Let $\lambda_{\mathbf{m}} \geq 0$ be fixed such that

$$\lambda_{\mathbf{m}} \in \arg\max_{\lambda \geq 0} \left\{ \max_u \left\{ \mathcal{D}(\lambda, u) \ : u \in L^\lambda_{d_F} \right\} \right\}.$$

Let θ_1, θ_2 be Lebesgue functions with negative and positive parts defined by

$$\begin{aligned} &\theta_1^+ = \nu - \rho_1, \ \ \theta_1^- = \mu - \rho_0, \\ \text{and} \quad &\theta_2^+ = \nu - \eta_1, \ \ \theta_2^- = \mu - \eta_0. \end{aligned}$$

Thanks to Theorem 2.8, θ_1 and θ_2 are solutions to the same OMK equation $(P_{\lambda_{\mathbf{m}}})$. So, using the uniqueness in Theorem 2.7, we deduce that $\theta_1 = \theta_2$ and that $\theta_1^- = \theta_2^-$, $\theta_1^+ = \theta_2^+$. This implies that $\rho_0 = \eta_0$ and $\rho_1 = \eta_1$. □

6. Monotonicity

In order to study the maps m and $\mathcal{R}$ defined in Section 2, we study the monotone and continuous dependence of the solution θ_λ of the OMK equation (P_λ) on the parameter λ.

Proposition 6.21 (Monotonicity and continuity of θ_λ)**.** *Let $\mu, \nu \in \mathcal{M}_b^+(\mathbb{R}^N)$ be compactly supported and absolutely continuous. Let $(\theta_\lambda, \Phi_\lambda, u_\lambda)$ be a solution to the OMK equation (P_λ).*
(i) Let $0 \leq \lambda_1 \leq \lambda_2$ and $\theta_{\lambda_1}, \theta_{\lambda_2}$ be solutions to the OMK equations (P_{λ_1}) and (P_{λ_2}), respectively. Then

$$\theta_{\lambda_1}^+ \geq \theta_{\lambda_2}^+ \ \ \textit{and} \ \ \theta_{\lambda_1}^- \geq \theta_{\lambda_2}^-.$$

(ii) If a nonnegative sequence $\lambda_n \to \lambda$ then $\theta_{\lambda_n} \to \theta_\lambda$ strongly in $L^1(\mathbb{R}^N)$.

Lemma 6.4 (Monotonicity of total mass)**.** *For any $\lambda \geq 0$, let θ_λ be the solution of the OMK equation (P_λ) and $\mathbf{m}_\lambda := (\mu - \theta_\lambda^-)(\mathbb{R}^N) = (\nu - \theta_\lambda^+)(\mathbb{R}^N)$. If $0 \leq \lambda_1 \leq \lambda_2$ then*

$$(\mu \wedge \nu)(\mathbb{R}^N) \leq \mathbf{m}_{\lambda_1} \leq \mathbf{m}_{\lambda_2} \leq \mathbf{m}_{\max}.$$

Proof. Thanks to Proposition 4.16, we see that $\mu \wedge \nu \leq \mu - \theta_\lambda^-$ and therefore $(\mu \wedge \nu)(\mathbb{R}^N) \leq \mathbf{m}_\lambda$. Since $\mu - \theta_\lambda^- \leq \mu$ and $\nu - \theta_\lambda^+ \leq \nu$, we have $\mathbf{m}_\lambda \leq \mathbf{m}_{\max}$. For the monotonicity, due to Theorem 5.20, $(\lambda_1, u_{\lambda_1})$ and $(\lambda_2, u_{\lambda_2})$ are solutions to the DPMK problem w.r.t. $\mathbf{m}_{\lambda_1}$ and $\mathbf{m}_{\lambda_2}$. By optimality, we have

$$\int u_{\lambda_1} \mathrm{d}(\nu - \mu) + \lambda_1(\mathbf{m}_{\lambda_1} - \nu(\mathbb{R}^N)) \geq \int u_{\lambda_2} \mathrm{d}(\nu - \mu) + \lambda_2(\mathbf{m}_{\lambda_1} - \nu(\mathbb{R}^N)),$$

and

$$\int u_{\lambda_2} \mathrm{d}(\nu - \mu) + \lambda_2(\mathbf{m}_{\lambda_2} - \nu(\mathbb{R}^N)) \geq \int u_{\lambda_1} \mathrm{d}(\nu - \mu) + \lambda_1(\mathbf{m}_{\lambda_2} - \nu(\mathbb{R}^N)).$$

Adding both sides, we obtain

$$\lambda_1 \mathbf{m}_{\lambda_1} + \lambda_2 \mathbf{m}_{\lambda_2} \geq \lambda_2 \mathbf{m}_{\lambda_1} + \lambda_1 \mathbf{m}_{\lambda_2},$$

or

$$(\lambda_2 - \lambda_1)(\mathbf{m}_{\lambda_2} - \mathbf{m}_{\lambda_1}) \geq 0. \qquad \square$$

To prove Proposition 6.21, we use the following result whose proof is given in [7].

Theorem 6.22. *([7, Theorem 3.4]) Let $\Gamma_{\text{opt}}^{\mathbf{m}}$ be the set of optimal transport plans of the mass $\mathbf{m} \geq 0$. There is a curve $\mathbf{m} \in [0, \min\{\|\mu\|_{L^1}, \|\nu\|_{L^1}\}] \longrightarrow \gamma^{\mathbf{m}} \in \Gamma_{\text{opt}}^{\mathbf{m}}$ along which the left and right marginals $\gamma^{\mathbf{m}+\varepsilon}$ dominate those of $\gamma^{\mathbf{m}}$ whenever $\varepsilon > 0$.*

Proof of Proposition 6.21. (i) Set $\mathbf{m}_i := \mathbf{m}_{\lambda_i} \geq (\mu \wedge \nu)(\mathbb{R}^N), i = 1, 2$. Since $\lambda_1 \leq \lambda_2$ and Lemma 6.4, we have $\mathbf{m}_1 \leq \mathbf{m}_2$. Thanks to Theorem 6.22, there exist pairs of active submeasures $(\rho_0^{\lambda_i}, \rho_1^{\lambda_i})$ corresponding to the mass $\mathbf{m}_i, i = 1, 2$, such that

$$\rho_0^{\lambda_1} \leq \rho_0^{\lambda_2} \text{ and } \rho_1^{\lambda_1} \leq \rho_1^{\lambda_2}. \tag{6.1}$$

By Theorem 5.20 (ii), $(\lambda_1, u_{\lambda_1})$ is a solution to the DPMK with mass $\mathbf{m}_1$. Set $\theta := \nu - \rho_1^{\lambda_1} - \mu + \rho_0^{\lambda_1}$. By Theorem 5.20 (i), there is Φ such that $(\theta, \Phi, u_{\lambda_1})$ is a solution to the OMK equation (P_{λ_1}). Due to the uniqueness in Theorem 2.7, we get

$$\theta_{\lambda_1} \equiv \theta = \nu - \rho_1^{\lambda_1} - \mu + \rho_0^{\lambda_1}.$$

Following the proof of Theorem 2.8, we obtain

$$\theta_{\lambda_1}^- = \mu - \rho_0^{\lambda_1} \text{ and } \theta_{\lambda_1}^+ = \nu - \rho_1^{\lambda_1}.$$

In the same way, we have

$$\theta_{\lambda_2}^- = \mu - \rho_0^{\lambda_2} \text{ and } \theta_{\lambda_2}^+ = \nu - \rho_1^{\lambda_2}.$$

Combining these with (6.1), we get $\theta_{\lambda_1}^- \geq \theta_{\lambda_2}^-$ and $\theta_{\lambda_1}^+ \geq \theta_{\lambda_2}^+$.
(ii) Since $\theta_{\lambda_n}^- \leq \mu$, $\theta_{\lambda_n}^+ \leq \nu$ as in Proposition 4.16, we have that $|\theta_{\lambda_n}| \leq \mu + \nu \in L^1$ and therefore $\{\theta_{\lambda_n}\}$ is equi-integrable. By Dunford–Pettis theorem, up to a subsequence, θ_{λ_n} converges weakly to some $\theta \in L^1(\mathbb{R}^N)$. Now, let us show that θ is a solution of the OMK equation (P_λ). Once this

is done, by the uniqueness in Theorem 2.7, we deduce that $\theta \equiv \theta_\lambda$ and thus the whole sequence $\theta_{\lambda_n} \to \theta_\lambda$ weakly in $L^1(\mathbb{R}^N)$. By the non-degeneracy of F and the definition of solution for the OMK equation (P_{λ_n}), it is clear that $\{u_{\lambda_n}\}$ is bounded and equi-Lipschitz; and that $\{\Phi_{\lambda_n}\}$ is bounded in $\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)$. So, up to subsequence,

$$u_{\lambda_n} \to u \ \text{ uniformly on each compact subset of } \ \mathbb{R}^N,$$

and

$$\Phi_{\lambda_n} \to \Phi \quad \text{ weakly* in } \mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N).$$

Let us show that (θ, Φ, u) is a solution to the OMK equation (P_λ). First, it is clear that $u \in L^\lambda_{d_F}$,

$$\int u \,\mathrm{d}\theta^- = \lim_{\lambda_n \to \lambda} \int u_{\lambda_n} \,\mathrm{d}\theta^-_{\lambda_n} = 0,$$

and

$$\int (u-\lambda)\,\mathrm{d}\theta^+ = \lim_{\lambda_n \to \lambda} \int (u_{\lambda_n} - \lambda_n)\,\mathrm{d}\theta^+_{\lambda_n} = 0.$$

Moreover,

$$\int \xi \,\mathrm{d}\theta + \int \frac{\Phi}{|\Phi|}\nabla\xi \,\mathrm{d}|\Phi| = \lim_{\lambda_n \to \lambda} \int \xi \,\mathrm{d}\theta_{\lambda_n} + \int \frac{\Phi_{\lambda_n}}{|\Phi_{\lambda_n}|}\nabla\xi \,\mathrm{d}|\Phi_{\lambda_n}| = \int \xi \mathrm{d}(\nu-\mu) \ \ \forall \xi \in C_c^\infty(\mathbb{R}^N),$$

which implies that

$$\theta - \nabla\cdot\Phi = \nu - \mu \ \text{ in } \ \mathcal{D}'(\mathbb{R}^N).$$

It remains to check that $\frac{\Phi}{|\Phi|}(x)\nabla_{|\Phi|}u(x) = F(x, \frac{\Phi}{|\Phi|}(x))$, $|\Phi|$-a.e. x in $\mathbb{R}^N$. Thanks to Lemma 4.2, this is equivalent to

$$\int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| \le \int\limits_{\mathbb{R}^N} \frac{\Phi}{|\Phi|}(x)\nabla_{|\Phi|}u(x)\mathrm{d}|\Phi|. \tag{6.2}$$

Since $\Phi_{\lambda_n} \to \Phi$ weakly* in $\mathcal{M}_b(\mathbb{R}^N;\mathbb{R}^N)$, we have (see e.g. [2, Theorem 2.38])

$$\int\limits_{\mathbb{R}^N} F(x, \frac{\Phi}{|\Phi|}(x))\mathrm{d}|\Phi| \le \liminf_{\lambda_n \to \lambda} \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi_{\lambda_n}}{|\Phi_{\lambda_n}|}(x))\mathrm{d}|\Phi_{\lambda_n}|. \tag{6.3}$$

On the other hand,

$$\begin{aligned}
\lim_{\lambda_n \to \lambda} \int\limits_{\mathbb{R}^N} F(x, \frac{\Phi_{\lambda_n}}{|\Phi_{\lambda_n}|}(x))\mathrm{d}|\Phi_{\lambda_n}| &= \lim_{\lambda_n \to \lambda} \int \frac{\Phi_{\lambda_n}}{|\Phi_{\lambda_n}|}\nabla_{|\Phi_{\lambda_n}|}u_{\lambda_n}\mathrm{d}|\Phi_{\lambda_n}| \\
&= \lim_{\lambda_n \to \lambda} \int u_{\lambda_n}\,\mathrm{d}(\nu-\mu) + \int u_{\lambda_n}\mathrm{d}\theta_{\lambda_n} \\
&= \int u\,\mathrm{d}(\nu-\mu) + \int u\,\mathrm{d}\theta = \int \frac{\Phi}{|\Phi|}(x)\nabla_{|\Phi|}u(x)\mathrm{d}|\Phi|.
\end{aligned} \tag{6.4}$$

From (6.3) and (6.4), we deduce (6.2). We have just proved that $\theta_{\lambda_n} \to \theta_\lambda$ weakly in $L^1(\mathbb{R}^N)$. At last, by the monotonicity of the first part, we deduce the strong convergence in $L^1(\mathbb{R}^N)$. $\square$

Proof of Theorem 2.10. The fact that $\mathbf{m}_\lambda \in [(\mu \wedge \nu)(\mathbb{R}^N), \mathbf{m}_{\max}]$ and the monotonicity of $\mathbf{m}_\lambda$ are given in Lemma 6.4 while the continuity of $\mathbf{m}_\lambda$ follows from the continuity of θ_λ. Let us show the surjectivity of $\mathbf{m}_\lambda$. Fix any $\mathbf{m} \in [(\mu \wedge \nu)(\mathbb{R}^N), \mathbf{m}_{\max}]$. Let (ρ_0, ρ_1) be a couple of active submeasures w.r.t. $\mathbf{m}$. Take $\lambda := \lambda_{\mathbf{m}}$ as in Theorem 2.8 (ii), then $\mathbf{m}_\lambda = \mathbf{m}$.
Now, for the properties of $\mathcal{R}$, the proof follows again from Theorem 2.8 and Proposition 6.21. □

7. Appendix

In this section, we prove a chain rule for the tangential gradient which was used in the paper.

Lemma 7.5 (Chain rule for the tangential gradient)**.** *Let $\eta \in \mathcal{M}_b^+(\mathbb{R}^N)$ and u be a Lipschitz continuous function defined on $\mathbb{R}^N$. Let G be a Lipschitz continuous function on $\mathbb{R}$ such that the set of non-differentiable points of G is finite. Then*

$$\nabla_\eta G(u)(x) = G^{'}(u(x))\nabla_\eta u(x), \ \ \textit{for } \eta\textit{-a.e. } x, \tag{7.1}$$

where $G^{'}(u(x))$ is the usual derivative with convention $G^{'}(u(x))\nabla_\eta u(x) = 0$ if $\nabla_\eta u(x) = 0$ even G is not differentiable at $u(x)$. In particular, we have
(i) $\nabla_\eta u^+ = \chi_{[u>0]}\nabla_\eta u$ and $\nabla_\eta u^- = -\chi_{[u<0]}\nabla_\eta u$, η-a.e. in $\mathbb{R}^N$;
(ii) $\nabla_\eta u = 0$, η-a.e. on the set $[u = c] := \{x \in \mathbb{R}^N : u(x) = c\}$ for any constant $c \in \mathbb{R}$.

Proof. 1. Let us first assume that G is continuously differentiable. To prove (7.1), it is enough to show that

$$\int_{\mathbb{R}^N} \nabla_\eta G(u) \cdot \Phi \, \mathrm{d}\eta = \int_{\mathbb{R}^N} G^{'}(u)\nabla_\eta u \cdot \Phi \, \mathrm{d}\eta,$$

for every $\Phi \in L^1_\eta(\mathbb{R}^N; \mathbb{R}^N)$ such that $\Phi(x) \in T_\eta(x)$, η-a.e. $x \in \mathbb{R}^N$, where $T_\eta(x)$ is the tangential space w.r.t. η (see [5, 6, 19]). Let $u_\varepsilon \in C^\infty(\mathbb{R}^N)$ be the regularization of u by convolution. Since u and G are Lipschitz, we have that u_ε and $G \circ u_\varepsilon$ converge uniformly to u and $G \circ u$ on $\mathbb{R}$, respectively. Thus (see e.g. [19, Proposition 4.5]) $\nabla_\eta u_\varepsilon$ and $\nabla_\eta G(u_\varepsilon)$ converge to $\nabla_\eta u$ and $\nabla_\eta G(u)$ in $L^\infty_\eta(\mathbb{R}^N; \mathbb{R}^N)$-w*, respectively. Since $\Phi(x) \in T_\eta(x)$, η-a.e. $x \in \mathbb{R}^N$, we have

$$\begin{aligned}
\int_{\mathbb{R}^N} \nabla_\eta G(u) \cdot \Phi \, \mathrm{d}\eta &= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^N} \nabla_\eta G(u_\varepsilon) \cdot \Phi \, \mathrm{d}\eta \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^N} \nabla G(u_\varepsilon) \cdot \Phi \, \mathrm{d}\eta \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^N} G^{'}(u_\varepsilon)\nabla u_\varepsilon \cdot \Phi \, \mathrm{d}\eta \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^N} G^{'}(u_\varepsilon)\nabla_\eta u_\varepsilon \cdot \Phi \, \mathrm{d}\eta = \int_{\mathbb{R}^N} G^{'}(u)\nabla_\eta u \cdot \Phi \, \mathrm{d}\eta.
\end{aligned}$$

This gives the result (7.1) whenever G is continuously differentiable by taking

$$\Phi = \nabla_\eta G(u) - G^{'}(u)\nabla_\eta u.$$

For (i), consider the function $G_\varepsilon(r) := \begin{cases} \sqrt{r^2+\varepsilon^2}-\varepsilon & \text{if } r>0 \\ 0 & \text{if } r\le 0 \end{cases}$. Then G_ε is continuously differentiable and Lipschitz on $\mathbb{R}$. Thus we have

$$\int_{\mathbb{R}^N} \nabla_\eta G_\varepsilon(u)\cdot\Phi\,\mathrm{d}\eta = \int_{\mathbb{R}^N} G_\varepsilon^{'}(u)\nabla_\eta u\cdot\Phi\,\mathrm{d}\eta = \int_{\{[u>0]\}} \frac{u}{\sqrt{u^2+\varepsilon^2}}\nabla_\eta u\cdot\Phi\,\mathrm{d}\eta$$

for every $\Phi\in L^1_\eta(\mathbb{R}^N;\mathbb{R}^N)$ such that $\Phi(x)\in T_\eta(x)$ for η-a.e. x. Letting $\varepsilon\to 0$, we obtain

$$\int_{\mathbb{R}^N} \nabla_\eta u^+.\Phi\,\mathrm{d}\eta = \int_{\{[u>0]\}} \nabla_\eta u.\Phi\,\mathrm{d}\eta = \int_{\mathbb{R}^N} \chi_{[u>0]}\nabla_\eta u.\Phi\,\mathrm{d}\eta.$$

The proof of the positive part ends up by choosing $\Phi := \nabla_\eta u^+ - \chi_{[u>0]}\nabla_\eta u$. A similar proof is done for the negative part. For (ii), we can assume that $c=0$. The proof follows from

$$\nabla_\eta u = \nabla_\eta u^+ - \nabla_\eta u^-.$$

2. Now, let us deal with a general Lipschitz function G satisfying our assumptions. Let us call $\{r_1, r_2, ..., r_n\}$ the set of non-differentiable points of G and set open subsets $\Omega_i := u^{-1}(\mathbb{R}\setminus\{r_i\})$ and $\Omega := \bigcap_{i=1}^n \Omega_i$. In this case, since u is a constant on the set $\mathbb{R}^N\setminus\Omega_i, i=1,...,n$, we have

$$\nabla_\eta G(u)(x) = G^{'}(u(x))\nabla_\eta u(x) = 0 \quad \text{for } \eta\text{-a.e. } x\in\mathbb{R}^N\setminus\Omega_i, i=1,...,n.$$

It remains to verify that

$$\nabla_\eta G(u)(x) = G^{'}(u(x))\nabla_\eta u(x) \quad \text{for } \eta\text{-a.e. } x\in\Omega. \tag{7.2}$$

Let us assume that $\Omega\neq\emptyset$ (if not, there is nothing to prove). Let G_ε be a smooth approximation of G by convolution. Let $\Phi\in L^1_\eta(\mathbb{R}^N;\mathbb{R}^N)$ be such that $\Phi(x)=0$ for η-a.e. x in $\mathbb{R}^N\setminus\Omega$. Then

$$\begin{aligned}
\int_{\mathbb{R}^N} \nabla_\eta G(u)\Phi\mathrm{d}\eta &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^N} \nabla_\eta G_\varepsilon(u)\Phi\mathrm{d}\eta \\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^N} G_\varepsilon^{'}(u)\nabla_\eta u\Phi\mathrm{d}\eta \\
&= \lim_{\varepsilon\to 0}\int_{\Omega} G_\varepsilon^{'}(u)\nabla_\eta u\Phi\mathrm{d}\eta \,(\text{since } \Phi(x)=0, \eta\text{-a.e. } x \text{ in } \mathbb{R}^N\setminus\Omega) \\
&= \int_{\Omega} G^{'}(u)\nabla_\eta u\Phi\mathrm{d}\eta,
\end{aligned}$$

where we used the Lebesgue Dominated Convergence Theorem. Next, choosing

$$\Phi = \nabla_\eta G(u) - G^{'}(u)\nabla_\eta u$$

as a test function, we obtain (7.2). □

References

[1] L. Ambrosio. Lecture notes on optimal transport problems. *in Mathematical aspects of evolving interfaces (Funchal, 2000), 1–52, Lecture Notes in Math., 1812*, Springer, Berlin, 2003.

[2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.

[3] J. W. Barrett and L. Prigozhin. Partial L^1 Monge-Kantorovich problem: Variational formulation and numerical approximation. *Interfaces Free Bound.*, **11** (2009), 201-238.

[4] J.D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84** (2000), 375–393.

[5] G. Bouchitté, G. Buttazzo and P. Seppercher. Energy with respect to a measure and applications to low dimensional structures. *Calc. Var. Partial Differential Equations*, **5** (1997), 37–54.

[6] G. Bouchitté, G. Buttazzo, and P. Seppecher. Shape optimization solutions via Monge–Kantorovich equation. *C. R. Acad. Sci. Paris Sér. I Math*, **324** (1997), 1185–1191.

[7] L. A. Caffarelli and R. J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math.*, **171** (2010), 673–730.

[8] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.* **147** (1999), 269-361.

[9] S. Chen and E. Indrei. On the regularity of the free boundary in the optimal partial transport problem for general cost functions. *J. Differential Equations*, **258** (2015), 2618-2632.

[10] G. Davila and Y. H. Kim. Dynamics of Optimal Partial Transport. *Calc. Var. Partial Differential Equations* (2016).

[11] I. Ekeland and R. Teman. *Convex analysis and variational problems.* North-Holland American Elsevier (1976).

[12] L. C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Mem. Amer. Math. Soc.*, 137 (1999).

[13] M. Feldman and R. J. McCann. Monge's transport problem on a Riemannian manifold. *Trans. Amer. Math. Soc.* **354** (2002), 1667–1697.

[14] A. Figalli. A note on the regularity of the free boundaries in the optimal partial transport problem. *Rendiconti del Circolo Matematico di Palermo*, **58** (2009), 283-286.

[15] A. Figalli. The Optimal Partial Transport Problem. *Arch. Ration. Mech. Anal.* **195** (2010), 533-560.

[16] N. Igbida, J. M. Mazón, J. D. Rossi and J. Toledo. Optimal mass transportation for costs given by Finsler distances via p-Laplacian approximations. *Adv. Calc. Var.* (2016).

[17] N. Igbida and V. T. Nguyen. Augmented Lagrangian method for Optimal Partial Transportation. *IMA J. Numer. Anal.* (2017), https://doi.org/10.1093/imanum/drw077, https://hal.archives-ouvertes.fr/hal-01518536.

[18] E. Indrei. Free boundary regularity in the optimal partial transport problem. *J. Funct. Anal.*, **264** (2013), 2497-2528.

[19] C. Jimenez. Dynamic formulation of optimal transport problems. *J. Convex Anal.* **15** (2008), 593-622.

[20] S. N. Kruzkov. First order quasilinear equations in several independent variables. *Mathematics of the USSR-Sbornik*, **10** (1970), 217-243.

[21] G. D. Philippis, A. R. Mészáros, F. Santambrogio and B. Velichkov. BV estimates in optimal transportation and applications. *Arch. Ration. Mech. Anal.* **219** (2016), 829-860.

[22] A. Pratelli. Equivalence between some definitions for the optimal mass transport problem and for the transport density on manifolds. *Ann. mat. Pura Appl.* **184** (2005), 215–238.

[23] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, Birkäuser (2015).

[24] C. Villani. *Topics in Optimal Transportation.* Graduate Studies in Mathematics. Vol. 58, 2003.

[25] C. Villani. *Optimal Transport, Old and New.* Grundlehren des Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), Vol. 338, Springer-Verlag, Berlin-New York, 2009.