AUGMENTED LAGRANGIAN METHOD FOR OPTIMAL PARTIAL TRANSPORTATION

NOUREDDINE IGBIDA† AND VAN THANH NGUYEN‡

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

Abstract. The use of augmented Lagrangian algorithm for optimal transport problems goes back to Benamou and Brenier [5] in the case where the cost corresponds to the square of the Euclidean distance. It was recently extended in [6] to the optimal transport with the Euclidean distance and Mean-Field Games theory, and in [8] to the optimal transportation with Finsler distances. Our aim here is to show how one can use this method to study the optimal partial transport problem with Finsler distance costs. To this aim, we introduce a suitable dual formulation of the optimal partial transport which contains all the information on the active regions and the associated flow. Then, we use a finite element discretization with the FreeFem++ software [21] to provide numerical simulations for the optimal partial transportation. A convergence study for the potential together with the flux and the active regions is given to validate the approach.

1. Introduction

The theory of optimal transportation deals with the problem to find the optimal way to move materials from a given source to a desired target in such a way to minimize the work. The problem was first proposed and studied by G. Monge in 1781 and then L. Kantorovich made fundamental contributions to the problem in the 1940s by relaxing the problem into a linear one. Since the late 80s, this subject has been investigated under various points of view with many applications in image processing, geometry, probability theory, economics, evolution PDEs and other areas. For more informations on the optimal mass transport problem, we refer the reader to the pedagogical books [28], [29], [2] and [27].

The standard optimal transport problem requires that the total mass of the source is equal to the total mass of the target (balance condition of mass) and that all the materials of the source must be transported. Here, we are interested in the optimal partial transportation. That is the case where the balance condition of mass is excluded and the aim is to transport effectively a prescribed amount of mass from the source to the target. In other words, the optimal partial

Key words and phrases. Optimal transport, optimal partial transport, augmented Lagrangian method.

†E-mail : noureddine.igbida@unilim.fr .
‡E-mail : van-thanh.nguyen@unilim.fr.

transport problem aims to study the practical situation where only a part of the commodity (resp. consumer demand) of a prescribed total mass $\mathbf{m}$ needs to be transported (resp. fulfilled).

This generalized problem brings out additional variables called active regions which are not very covered in the literature. The problem was first studied theoretically in [9] (see also [16]) in the case where the work is proportional to the square of the Euclidean distance. Recently in [23], we give a complete theoretical study of the problem in the case where the work is proportional to a Finsler distance d_F (covering by the way the case of the Euclidean distance), where d_F is given by (see Section 2)

$$d_F(x,y) := \inf_{\xi \in Lip([0,1];\overline{\Omega})} \left\{ \int_0^1 F(\xi(t),\dot{\xi}(t))dt : \xi(0)=x, \xi(1)=y \right\}.$$

Concerning numerical approximations for the optimal partial transport, Barrett and Prigozhin [3] studied the case of the Euclidean distance by using an approximation based on nonlinear approximated PDEs and Raviart-Thomas finite elements. Benamou et al. [7] and Chizat et al. [10] introduced general numerical frameworks to approximate solutions to linear programs related to the optimal transport (including the optimal partial transport). Their idea is based on an entropic regularization of the initial linear programs. This is a static approach to optimal transport-type problems and needs to use (approximated) values of $d_F(x,y)$.

In this paper, we use a different approach (based mainly on [5], [6] and [23]) to compute the solution of the optimal partial transport problem. We first show how one can directly reformulate the unknown quantities (variables) of the optimal partial transport into an infinite dimensional minimization problem of the form:

$$\min_{\phi \in V} \mathcal{F}(\phi) + \mathcal{G}(\Lambda\phi),$$

where $\mathcal{F}, \mathcal{G}$ are l.s.c., convex functionals and $\Lambda \in \mathcal{L}(V,Z)$ is a continuous linear operator between two Banach spaces. Thanks to peculiar properties of $\mathcal{F}$ and $\mathcal{G}$ in our situation, an augmented Lagrangian method is applied effectively in the same spirit of [6] and [8]. We show that, for the computation, we just need to solve linear equations (with a symmetric positive definite coefficient matrix) or to update explicit formulations. It is worth to note that this method uses only elementary operations without evaluating d_F.

The paper is organized as follows: In the next section, we introduce the optimal partial transport problem and its equivalent formulations with a particular attention to the Kantorovich dual formulation. In Section 3, we give a finite dimensional approximation of the problem and show that primal-dual solutions of the discretized problems converge to the ones of original continuous problems. The details of the ALG2 algorithm is given in Section 4. Some numerical examples are presented in Section 5. We terminate the paper by an appendix where we give proofs of some facts we need in the paper.

2. Partial transport and its equivalent formulations

Let Ω be a connected bounded Lipschitz domain and F be a continuous Finsler metric on $\overline{\Omega}$, i.e. $F : \overline{\Omega} \times \mathbb{R}^N \longrightarrow [0,+\infty)$ is continuous and $F(x,.)$ is convex, positively homogeneous of degree 1 in the sense

$$F(x,tv)=tF(x,v),\forall t>0,v\in\mathbb{R}^N.$$

We assume moreover that F is non-degenerate in the sense that there exist positive constants M_1, M_2 such that

$$M_1|v|\le F(x,v)\le M_2|v|,\forall x\in\overline{\Omega},v\in\mathbb{R}^N.$$

Let $\mu,\nu\in\mathcal{M}_b^+(\overline{\Omega})$ be two Radon measures on $\overline{\Omega}$ and $\mathbf{m}_{\max}:=\min\{\mu(\overline{\Omega}),\nu(\overline{\Omega})\}$. Given a total mass $\mathbf{m}\in[0,\mathbf{m}_{\max}]$, the optimal partial transport problem (or partial Monge-Kantorovich problem, PMK for short) aims to transport effectively the total mass $\mathbf{m}$ from a supply subregion of the source μ into a subregion of the target ν. The set of subregions of mass $\mathbf{m}$ is given by

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

An element $(\rho_0,\rho_1)\in Sub_{\mathbf{m}}(\mu,\nu)$ is called a couple of active regions.
As for the optimal transport, one can work with different kinds of cost functions for the optimal partial transport, i.e. in the formulation (2.1) below, $d_F(x,y)$ can be replaced by a general measurable cost function $c(x,y)$. However, in this paper, we focus on the case where the cost $c=d_F$. So let us state the problem directly for d_F. The PMK problem ([9], [16], [3], [23]) aims to minimize the following problem

$$\min\left\{\mathcal{K}(\gamma):=\int\limits_{\overline{\Omega}\times\overline{\Omega}}d_F(x,y)d\gamma:\ \gamma\in\pi_{\mathbf{m}}(\mu,\nu)\right\},\tag{2.1}$$

where

- d_F is the Finsler distance on $\overline{\Omega}$ associated with F, i.e.

$$d_F(x,y):=\inf\left\{\int\limits_0^1F(\xi(t),\dot{\xi}(t))dt,\xi(0)=x,\xi(1)=y,\xi\in Lip([0,1];\overline{\Omega})\right\};$$

- $\pi_{\mathbf{m}}(\mu,\nu)$ is the set of transport plans of mass $\mathbf{m}$, i.e.

$$\pi_{\mathbf{m}}(\mu,\nu):=\left\{\gamma\in\mathcal{M}_b^+(\overline{\Omega}\times\overline{\Omega}):\ (\pi_x\#\gamma,\pi_y\#\gamma)\in Sub_{\mathbf{m}}(\mu,\nu)\right\}.$$

Here, $\pi_x\#\gamma$ and $\pi_y\#\gamma$ are the first and second marginals of γ. An optimal γ^* is called an *optimal plan* and $(\pi_x\#\gamma^*,\pi_y\#\gamma^*)$ is called a couple of *optimal active regions.*

Following [23], in order to study the PMK problem we use its dual problem that we call the dual partial Monge-Kantorovich problem (DPMK for short). To this aim, we consider Lip_{d_F} the set of $1-$Lipschitz continuous functions w.r.t. d_F given by

$$Lip_{d_F}:=\{u:\overline{\Omega}\longrightarrow\mathbb{R}\,|\,u(y)-u(x)\le d_F(x,y),\ \forall x,y\in\overline{\Omega}\}.$$

Then, the connection between the PMK problem and DPMK problem is summarized in the following theorem.

Theorem 2.1. *Let $\mu,\nu \in \mathcal{M}_b^+(\overline{\Omega})$ be Radon measures and $\mathbf{m} \in [0, \mathbf{m}_{\max}]$. The partial Monge-Kantorovich problem has a solution $\sigma^* \in \pi_{\mathbf{m}}(\mu,\nu)$ and*

$$\mathcal{K}(\sigma^*) = \max \left\{ \mathcal{D}(\lambda,u) := \int_{\overline{\Omega}} u\, d(\nu-\mu) + \lambda(\mathbf{m} - \nu(\overline{\Omega})) \ :\ \lambda \geq 0 \text{ and } u \in L^\lambda_{d_F} \right\}, \tag{2.2}$$

where

$$L^\lambda_{d_F} := \left\{ u \in Lip_{d_F} \ :\ 0 \leq u(x) \leq \lambda \quad \text{for any } x \in \overline{\Omega} \right\}.$$

Moreover, $\sigma \in \pi_{\mathbf{m}}(\mu,\nu)$ and $(\lambda,u) \in \mathbb{R}^+ \times L^\lambda_{d_F}$ are solutions, respectively if and only if

$$u(x) = 0, \ \text{ for } (\mu - \pi_x\#\sigma)\text{-a.e. } x \in \overline{\Omega}, \quad u(x) = \lambda, \ \text{ for } (\nu - \pi_y\#\sigma)\text{-a.e. } x \in \overline{\Omega},$$
$$\text{and } u(y) - u(x) = d_F(x,y), \ \text{ for } \sigma\text{-a.e. } (x,y) \in \overline{\Omega}\times\overline{\Omega}.$$

Proof. The proof follows in the same way of Theorem 2.4 in [23], where the authors study the case $\Omega = \mathbb{R}^N$. □

The DPMK problem (2.2) contains all the information concerning the optimal partial mass transportation. However, for the numerical approximation of the optimal partial transportation and in order to use the augmented Lagrangian method, we need to rewrite the problem into the form

$$\inf_{\phi \in V} \mathcal{F}(\phi) + \mathcal{G}(\Lambda\phi).$$

To do that, we consider the polar function F^* of F which is defined by

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

Note that $F^*(x,.)$ is not the Legendre-Fenchel transform. It is easy to see that F^* is also a continuous, non-degenerate Finsler metric on $\overline{\Omega}$ and

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

Remark 2.2. *Using the polar function F^*, we can characterize the set Lip_{d_F} as (see the appendix if necessary)*

$$Lip_{d_F} = \left\{ u : \overline{\Omega} \longrightarrow \mathbb{R} \,|\, u \text{ is Lipschitz continuous and } F^*(x,\nabla u(x)) \leq 1,\ \text{a.e. } x \in \Omega \right\}.$$

Thanks to this remark, the DPMK problem (2.2) can be written as

$$\max\{\mathcal{D}(\lambda,u) :\ 0 \leq u(x) \leq \lambda,\ u \text{ is Lipschitz continuous, } F^*(x,\nabla u(x)) \leq 1,\ \text{a.e. } x \in \Omega\}.$$

Moreover, we have

Theorem 2.3. *Under the assumptions of Theorem 2.1, setting $V := \mathbb{R} \times C^1(\overline{\Omega})$ and $Z := C(\overline{\Omega})^N \times C(\overline{\Omega}) \times C(\overline{\Omega})$, we have*

$$\mathcal{K}(\sigma^*) = -\inf\left\{ \mathcal{F}(\lambda,u) + \mathcal{G}(\Lambda(\lambda,u)) :\ (\lambda,u) \in V \right\}, \tag{2.3}$$

where $\Lambda \in \mathcal{L}(V,Z)$ is given by

$$\Lambda(\lambda,u) := (\nabla u, -u, u-\lambda),\ \forall(\lambda,u) \in V,$$

and $\mathcal{F} : V \longrightarrow (-\infty, +\infty]$, $\mathcal{G} : Z \longrightarrow (-\infty, +\infty]$ *are the l.s.c. convex functions given by*

$$\mathcal{F}(\lambda, u) := -\int_{\overline{\Omega}} u\, d(\nu - \mu) - \lambda(\mathbf{m} - \nu(\overline{\Omega})),\ \ \forall (\lambda, u) \in V;$$

$$\mathcal{G}(q, z, w) := \begin{cases} 0 & \text{if } z(x) \leq 0,\ w(x) \leq 0,\ F^*(x, q(x)) \leq 1, \forall x \in \overline{\Omega}, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{for } (q, z, w) \in Z.$$

To prove this theorem we need the following lemma.

Lemma 2.1. *Let* $\lambda \geq 0$ *be fixed. For any* $u \in L^{\lambda}_{d_F}$*, there exists a sequence of smooth functions* $u_\varepsilon \in C^\infty_c(\mathbb{R}^N) \bigcap L^{\lambda}_{d_F}$ *such that* $u_\varepsilon \rightrightarrows u$ *uniformly on* $\overline{\Omega}$*.*

The result of the lemma is more or less known in some cases (see [24] for the case where the function u is null on the boundary). The proof in the general case is quite technical and will be given in the appendix.

Proof of Theorem 2.3. Thanks to Remark 2.2 and Lemma 2.1, we have

$$\begin{aligned} -\inf_{(\lambda,u)\in V} \mathcal{F}(\lambda, u) + \mathcal{G}(\Lambda(\lambda, u)) &= \sup \left\{ \int_{\overline{\Omega}} u d(\nu - \mu) + \lambda(\mathbf{m} - \nu(\overline{\Omega})) : \lambda \geq 0,\ u \in C^1(\overline{\Omega}) \cap L^{\lambda}_{d_F} \right\} \\ &= \max \left\{ \mathcal{D}(\lambda, u)\ :\ \lambda \geq 0 \text{ and } u \in L^{\lambda}_{d_F} \right\}. \end{aligned}$$

Using the duality (2.2), the proof is completed. ☐

To end up this section, we prove the following result that will be useful for the proof of the convergence of our discretization.

Theorem 2.4. *Under the assumptions of Theorem 2.1, we have*

$$-\inf_{(\lambda,u)\in V} \mathcal{F}(\lambda, u) + \mathcal{G}(\Lambda(\lambda, u)) = \min \Big\{ \int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x)) d|\Phi|\ :\ (\Phi, \theta^0, \theta^1) \in \ominus_{\mathbf{m}}(\mu, \nu) \Big\}, \tag{2.4}$$

where

$$\ominus_{\mathbf{m}}(\mu, \nu) := \Big\{ (\Phi, \theta^0, \theta^1) \in Z^* = \mathcal{M}_b(\overline{\Omega})^N \times \mathcal{M}_b(\overline{\Omega}) \times \mathcal{M}_b(\overline{\Omega}) : \theta^0 \geq 0,\ \theta^1 \geq 0,\ \theta^1(\overline{\Omega}) = \nu(\overline{\Omega}) - \mathbf{m}$$

$$\text{and } \ -\nabla \cdot \Phi = \nu - \theta^1 - (\mu - \theta^0) \ \text{ with } \ \Phi.n = 0 \ \text{ on } \ \partial\Omega \Big\}.$$

Actually, the minimal flow-type formulation

$$\min \Big\{ \int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x)) d|\Phi|\ :\ (\Phi, \theta^0, \theta^1) \in \ominus_{\mathbf{m}}(\mu, \nu) \Big\} \tag{2.5}$$

introduces the Beckmann problem (see [4]) for the optimal partial transport with Finsler distance costs. See here that in the balanced case, i.e. $\mathbf{m} = \mu(\overline{\Omega}) = \nu(\overline{\Omega})$, the formulation (2.5) becomes

$$\min\Big\{ \int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x))d|\Phi| \, : \, \Phi \in \mathcal{M}_b(\overline{\Omega})^N, -\nabla\cdot\Phi = \nu - \mu \text{ with } \Phi.n = 0 \text{ on } \partial\Omega \Big\}. \tag{2.6}$$

An optimal solution Φ of the problem (2.6) is called an *optimal flow* of transporting μ onto ν. As known from the optimal transport theory, the optimal flow gives a way to visualize the transportation.

To prove Theorem 2.4, we will use the well known duality arguments. For convenience, let us recall here the Fenchel-Rockafellar duality. Let us consider the problem

$$\inf_{\phi\in V} \mathcal{F}(\phi) + \mathcal{G}(\Lambda\phi) \tag{2.7}$$

where $\mathcal{F} : V \longrightarrow (-\infty, +\infty]$ and $\mathcal{G} : Z \longrightarrow (-\infty, +\infty]$ are convex, l.s.c. and $\Lambda \in \mathcal{L}(V, Z)$ the space of linear continuous functions from V to Z. Using $\mathcal{F}^*$ and $\mathcal{G}^*$ the conjugate functions (given by the Legendre-Fenchel transformation) of $\mathcal{F}$ and $\mathcal{G}$ respectively, and Λ^* is the adjoint operator of Λ, it is not difficult to see that

$$\sup_{\sigma\in Z^*} (-\mathcal{F}^*(-\Lambda^*\sigma) - \mathcal{G}^*(\sigma)) \leq \inf_{\phi\in V} \mathcal{F}(\phi) + \mathcal{G}(\Lambda\phi),$$

where Z^* is the topological dual space associated with Z. This is the so called weak duality. For the strong duality, which corresponds to equality we have the following well known result.

Proposition 2.5 (cf. [14]). *In addition, assume that there exists ϕ_0 such that $\mathcal{F}(\phi_0) < +\infty$, $\mathcal{G}(\Lambda\phi_0) < +\infty$, $\mathcal{G}$ being continuous at $\Lambda\phi_0$. Then the Fenchel-Rockafellar dual problem*

$$\sup_{\sigma\in Z^*} (-\mathcal{F}^*(-\Lambda^*\sigma) - \mathcal{G}^*(\sigma)) \tag{2.8}$$

has at least a solution $\sigma \in Z^*$ and $\inf(2.7) = \max(2.8)$. *Moreover, in this case, ϕ is a solution to the primal problem* (2.7) *if and only if*

$$-\Lambda^*\sigma \in \partial\mathcal{F}(\phi) \quad \textit{and} \quad \sigma \in \partial\mathcal{G}(\Lambda\phi). \tag{2.9}$$

Proof of Theorem 2.4. We work with the uniform convergence for the spaces $C(\overline{\Omega})^N$, $C(\overline{\Omega})$ and the norm $\|u\|_{C^1} := \max\{\|u\|_\infty, \|\nabla u\|_\infty\}$ for $C^1(\overline{\Omega})$. It is not difficult to see that the hypotheses of Proposition 2.5 are satisfied. Now, let us compute the Fenchel-Rockafellar dual problem of (2.3). Since $\mathcal{F}$ is linear, $\mathcal{F}^*(-\Lambda^*(\Phi, \theta^0, \theta^1))$ is finite (and always equals to 0) if and only if

$$-\Lambda^*(\Phi, \theta^0, \theta^1) = -(\mathbf{m} - \nu(\overline{\Omega}), \nu - \mu) \text{ in } V^*$$

i.e.

$$\langle\Phi, \nabla u\rangle - \langle\theta^0, u\rangle + \langle\theta^1, u - \lambda\rangle = \lambda(\mathbf{m} - \nu(\overline{\Omega})) + \langle\nu - \mu, u\rangle, \ \forall(\lambda, u) \in V.$$

This implies that

$$\int_{\overline{\Omega}} \nabla u \, d\Phi = \int_{\overline{\Omega}} u \, d(\nu - \theta^1) - \int_{\overline{\Omega}} u \, d(\mu - \theta^0), \text{ for all } u \in C^1(\overline{\Omega})$$

and

$$-\lambda \int_{\overline{\Omega}} d\theta^1 = \lambda(\mathbf{m} - \nu(\overline{\Omega})), \forall \lambda \in \mathbb{R}.$$

These mean that

$$-\nabla \cdot \Phi = \nu - \theta^1 - (\mu - \theta^0) \text{ with } \Phi.n = 0 \text{ on } \partial\Omega,$$

and

$$\theta^1(\overline{\Omega}) = \nu(\overline{\Omega}) - \mathbf{m}.$$

We also have

$$\mathcal{G}^*(\Phi, \theta^0, \theta^1) = \begin{cases} \displaystyle\int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x)) d|\Phi| \text{ if } \theta^0 \geq 0, \theta^1 \geq 0, \\ +\infty \text{ otherwise,} \end{cases} \qquad \text{for any } (\Phi, \theta^0, \theta^1) \in Z^*.$$

Then the proof follows by Proposition 2.5. □

Remark 2.6. *The optimality relations* (2.9) *reads*

$$\begin{cases} -\nabla \cdot \Phi = \nu - \theta^1 - (\mu - \theta^0) \textit{ and } \Phi \cdot n = 0 \textit{ on } \partial\Omega \\ \theta^1(\overline{\Omega}) = \nu(\overline{\Omega}) - \mathbf{m} \\ \langle \Phi, \nabla u \rangle \geq \langle \Phi, q \rangle, \forall q \in C(\overline{\Omega}), F^*(x, q(x)) \leq 1, \forall x \in \overline{\Omega} \\ \lambda \in \mathbb{R}^+, u \in C^1(\overline{\Omega}) \bigcap L^\lambda_{d_F} \\ u = 0, \ \theta^0\textit{-a.e. in } \overline{\Omega} \\ u = \lambda, \ \theta^1\textit{-a.e. in } \overline{\Omega}. \end{cases}$$

In fact, the optimality condition $-\Lambda^*\sigma \in \partial\mathcal{F}(\phi)$ gives the first two equations and $\sigma \in \partial\mathcal{G}(\Lambda\phi)$ gives the last four equations. Moreover, if $\Phi \in L^1(\Omega)^N$, then the condition

$$\langle \Phi, \nabla u \rangle \geq \langle \Phi, q \rangle, \forall q \in C(\overline{\Omega}), F^*(x, q(x)) \leq 1, \forall x \in \overline{\Omega}$$

can be replaced by

$$F(x, \Phi(x)) = \langle \nabla u(x), \Phi(x) \rangle, \quad a.e. \ \ x \in \Omega. \tag{2.10}$$

However, it is not clear in general that Φ belongs to $L^1(\Omega)^N$. In the case where Ω is convex and $F(x,v) := |v|$ the Euclidean norm (or some other uniformly convex and smooth norms), the L^p regularity results are known under suitable assumptions on μ and ν (see e.g. [15], [11], [12], *and* [26]*). To our knowledge, the case of general Finsler metrics is still an open question.*

In the case where Φ is a vector-valued measure, the condition (2.10) *should be adapted to the tangential gradient. Rigorous formulations using the tangential gradient with respect to a measure as well as rigorous proofs in the general case can be found in the paper* [23] *with $\Omega = \mathbb{R}^N$.*

It is expected that $\theta^0 \leq \mu$ and $\theta^1 \leq \nu$ for optimal solutions $(\Phi, \theta^0, \theta^1)$ of the minimal flow formulation (2.5). This is the case whenever $\mathbf{m} \in [(\mu \wedge \nu)(\overline{\Omega}), \mathbf{m}_{\max}]$, where $\mu \wedge \nu$ is the common mass measure of μ and ν, i.e. if $\mu, \nu \in L^1(\Omega)$, then $\mu \wedge \nu \in L^1(\Omega)$ and

$$(\mu \wedge \nu)(x) = \min\{\mu(x), \nu(x)\} \text{ for a.e. } x \in \Omega.$$

In general, the measure $\mu \wedge \nu$ is defined by (see [1])

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

Proposition 2.7. *Let $\mathbf{m} \in [(\mu \wedge \nu)(\overline{\Omega}), \mathbf{m}_{\max}]$ and $(\Phi, \theta^0, \theta^1) \in Z^*$ be an optimal solution of* (2.5)*. Then $\theta^0 \le \mu$ and $\theta^1 \le \nu$. Moreover, $(\mu - \theta^0, \nu - \theta^1)$ is a couple of optimal active regions and Φ is an optimal flow of transporting $\mu - \theta^0$ onto $\nu - \theta^1$.*

Proof. The proof follows in the same way as Theorem 5.21 and Corollary 5.20 in [23]. □

Our next work is to compute an approximation of Φ (in fact, approximations of $\Phi, u, \lambda, \theta^0, \theta^1$). To do that, we will apply an augmented Lagrangian method to the DPMK problem (2.2).

3. Discretization and convergence

Coming back to the DPMK problem (2.2), our aim now is to give, by using a finite element approximation, the discretized problem associated with (2.2). To begin with, let us consider regular triangulations $\mathcal{T}_h$ of $\overline{\Omega}$. For a fixed integer $k \ge 1$, P_k is the set of polynomials of degree less or equal k. Let $E_h \subset H^1(\Omega)$ be the space of continuous functions on $\overline{\Omega}$ and belonging to P_k on each triangle of $\mathcal{T}_h$. We denote by Y_h the space of vectorial functions such that their restrictions belong to $(P_{k-1})^N$ on each triangle of $\mathcal{T}_h$. Let $f = \nu - \mu$ and $f_h \in E_h$ such that $\{f_h\}$ converges weakly* to f in $\mathcal{M}_b(\overline{\Omega})$.

Considering the finite dimensional spaces

$$V_h = \mathbb{R} \times E_h, \qquad Z_h = Y_h \times E_h \times E_h,$$

we set

$$\Lambda_h(\lambda, u) := (\nabla u, -u, u - \lambda) \in Z_h, \text{ for } (\lambda, u) \in V_h,$$

$$\mathcal{F}_h(\lambda, u) := -\langle u, f_h \rangle - \lambda(\mathbf{m} - \nu(\overline{\Omega})), \quad \forall (\lambda, u) \in V_h,$$

and

$$\mathcal{G}_h(q, z, w) := \begin{cases} 0 & \text{if } z \le 0,\ w \le 0,\ F^*(x, q(x)) \le 1, \text{ for a.e. } x \in \Omega, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{for } (q, z, w) \in Z_h.$$

Then the finite dimensional approximation of (2.2) reads

$$\inf_{(\lambda,u) \in V_h} \mathcal{F}_h(\lambda, u) + \mathcal{G}_h(\Lambda_h(\lambda, u)). \tag{3.11}$$

The following result shows that this is a suitable approximation of (2.2).

Theorem 3.8. *Assume that $\mathbf{m} < \nu(\overline{\Omega})$. Let $(\lambda_h, u_h) \in V_h$ be an optimal solution to the approximated problem* (3.11) *and $(\Phi_h, \theta_h^0, \theta_h^1)$ be an optimal dual solution to* (3.11)*. Then, up to a subsequence, (λ_h, u_h) converges in $\mathbb{R} \times C(\overline{\Omega})$ to (λ, u) an optimal solution of the DPMK problem* (2.2) *and $(\Phi_h, \theta_h^0, \theta_h^1)$ converges weakly* in $\mathcal{M}_b(\overline{\Omega})^N \times \mathcal{M}_b(\overline{\Omega}) \times \mathcal{M}_b(\overline{\Omega})$ to $(\Phi, \theta^0, \theta^1)$ an optimal solution of* (2.5)*.*

Proof. Since $\mathbf{m} < \nu(\overline{\Omega})$, $\{\lambda_h\}$ is bounded in $\mathbb{R}$ and $\{u_h\}$ is bounded in $(C(\overline{\Omega}), \|.\|_\infty)$. From the non-degeneracy of F and the definitions of $\mathcal{F}_h, \mathcal{G}_h, \Lambda_h$, we have that $\{u_h\}$ is equi-Lipschitz and

$$u_h(y) - u_h(x) \le d_F(x,y), \forall x, y \in \overline{\Omega}.$$

Using the Ascoli-Arzela Theorem, up to a subsequence, $u_h \rightrightarrows u$ uniformly on $\overline{\Omega}$, and $\lambda_h \to \lambda$. Obviously, $\lambda \ge 0$ and $u \in L^\lambda_{d_F}$. Now, by the optimality of (λ_h, u_h) and of $(\Phi_h, \theta^0_h, \theta^1_h)$, we have

$$-\Lambda^*(\Phi_h, \theta^0_h, \theta^1_h) = -(\mathbf{m} - \nu(\overline{\Omega}), f_h) \text{ in } V^*_h$$

and

$$\mathcal{F}_h(\lambda_h, u_h) + \mathcal{G}_h(\Lambda_h(\lambda_h, u_h)) = -\mathcal{F}^*_h(-\Lambda^*(\Phi_h, \theta^0_h, \theta^1_h)) - \mathcal{G}^*_h(\Phi_h, \theta^0_h, \theta^1_h).$$

More concretely,

$$\langle \Phi_h, \nabla v\rangle - \langle \theta^0_h, v\rangle + \langle \theta^1_h, v - s\rangle = s(\mathbf{m} - \nu(\overline{\Omega})) + \langle f_h, v\rangle, \forall (s,v) \in V_h, \tag{3.12}$$

$$\theta^0_h \ge 0,\ \ \theta^1_h \ge 0,\ \ \theta^1_h(\overline{\Omega}) = \nu(\overline{\Omega}) - \mathbf{m}, \tag{3.13}$$

and

$$\langle u_h, f_h\rangle + \lambda_h(\mathbf{m} - \nu(\overline{\Omega})) = \sup\left\{\langle q, \Phi_h\rangle : q \in Y_h, F^*(x, q(x)) \le 1,\ \text{ a.e. } x \in \Omega\right\}. \tag{3.14}$$

In (3.12), taking $v = 0$ and $s = 1$ (resp. $v = s = 1$), we see that $\{\theta^1_h\}$ (resp. $\{\theta^0_h\}$) is bounded in $\mathcal{M}_b(\overline{\Omega})$. Moreover, using (3.14) and the boundedness of (λ_h, u_h) we deduce that $\{\Phi_h\}$ is bounded in $\mathcal{M}_b(\overline{\Omega})^N$. So, up to a subsequence,

$$(\Phi_h, \theta^0_h, \theta^1_h) \rightharpoonup (\Phi, \theta^0, \theta^1) \ \text{ in } \mathcal{M}_b(\overline{\Omega})^N \times \mathcal{M}_b(\overline{\Omega}) \times \mathcal{M}_b(\overline{\Omega}) - \text{w}^*.$$

Using (3.12) and (3.13), it is clear that $(\Phi, \theta^0, \theta^1)$ satisfies

$$\langle \Phi, \nabla v\rangle - \langle \theta^0, v\rangle + \langle \theta^1, v - s\rangle = s(\mathbf{m} - \nu(\overline{\Omega})) + \langle f, v\rangle, \forall (s,v) \in V,$$

and

$$\theta^0 \ge 0,\ \ \theta^1 \ge 0,\ \ \theta^1(\overline{\Omega}) = \nu(\overline{\Omega}) - \mathbf{m},$$

i.e. $(\Phi, \theta^0, \theta^1)$ is feasible for the minimal flow problem (2.5).
Next, let us show the optimality of (λ, u) and of $(\Phi, \theta^0, \theta^1)$, i.e.

$$\int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x))d|\Phi| = \langle u, \nu - \mu\rangle + \lambda(\mathbf{m} - \nu(\overline{\Omega})). \tag{3.15}$$

We fix $q \in C(\overline{\Omega})^N$ such that $F^*(x, q(x)) \le 1, \forall x \in \overline{\Omega}$, and we consider $q_h \in Y_h$ such that $\|q_h - q\|_{L^\infty(\Omega)} \to 0$ as $h \to 0$. We see that

$$F^*(x, q_h(x)) = F^*(x, q(x)) + F^*(x, q_h(x)) - F^*(x, q(x)) \le 1 + O(h),\ \text{ a.e. } x \in \Omega.$$

By taking $\dfrac{q_h}{1 + O(h)}$, we can assume that $q_h \in Y_h, F^*(x, q_h(x)) \le 1,$ a.e. $x \in \Omega$ and $\|q_h - q\|_{L^\infty(\Omega)} \to 0$ as $h \to 0$. Using (3.14), we have

$$\begin{aligned}
\langle q, \Phi\rangle &= \langle q_h, \Phi_h\rangle + \langle q, \Phi - \Phi_h\rangle + \langle q - q_h, \Phi_h\rangle \\
&\le \sup\left\{\langle q_h, \Phi_h\rangle : q_h \in Y_h, F^*(x, q_h(x)) \le 1,\ \text{ a.e. } x \in \Omega\right\} + O(h) \\
&= \langle u_h, f_h\rangle + \lambda_h(\mathbf{m} - \nu(\overline{\Omega})) + O(h).
\end{aligned}$$

Letting $h \to 0$, we get

$$\langle q, \Phi\rangle \le \langle u, \nu-\mu\rangle + \lambda(\mathbf{m} - \nu(\overline{\Omega})) \text{ for any } q \in C(\overline{\Omega})^N, F^*(x, q(x)) \le 1, \forall x \in \overline{\Omega}.$$

Taking supremum in q, we obtain

$$\int_{\overline{\Omega}} F(x, \frac{\Phi}{|\Phi|}(x))d|\Phi| \le \langle u, \nu-\mu\rangle + \lambda(\mathbf{m} - \nu(\overline{\Omega})).$$

At last, thanks to the duality equality (2.4), this implies (3.15), the optimality of (λ, u) and of $(\Phi, \theta^0, \theta^1)$. □

Remark 3.9. *In the case* $\mathbf{m} = \mathbf{m}_{\max}$ *(called the unbalanced transport), the DPMK problem has a simpler formulation. So for the purpose of implementation, we distinguish the two cases: the partial transport and the unbalanced transport. In the unbalanced case, let us assume that* $\mathbf{m} = \mathbf{m}_{\max} = \nu(\overline{\Omega})$ *(i.e.* $\mu(\overline{\Omega}) \ge \nu(\overline{\Omega})$*), the DPMK problem* (2.2) *can be written as*

$$\max_{u \in Lip_{d_F}, u \ge 0} \int_{\overline{\Omega}} ud(\nu - \mu). \tag{3.16}$$

By using $V_h = E_h,\ Z_h = Y_h \times E_h, \Lambda_h u = (\nabla u, -u)$, *and*

$$\mathcal{G}_h(q, z) = \begin{cases} 0 \text{ if } z \le 0, F^*(x, q(x)) \le 1, & a.e. \ x \in \Omega \\ +\infty \text{ otherwise}, \end{cases}$$

a finite dimensional approximation can be given by

$$\inf_{u \in V_h} -\langle u, f_h\rangle + \mathcal{G}_h(\Lambda_h u). \tag{3.17}$$

As in Theorem 3.8, we can prove the convergence of this finite dimensional approximation to the original one (3.16). More precisely, we have

Proposition 3.10. *Assume that* $\mathbf{m} = \nu(\overline{\Omega})$*. Let* $u_h \in V_h$ *be an optimal solution to the approximated problem* (3.17) *and* (Φ_h, θ_h^0) *be an optimal dual solution to* (3.17)*. Then, up to a subsequence and translation by constant,* u_h *converges to* u *an optimal solution of the DPMK problem* (3.16) *and* (Φ_h, θ_h^0) *converges to* (Φ, θ^0) *an optimal solution of* (2.5) *with* $\theta^1 = 0$.

The proof of this proposition is similar to the proof of Theorem 3.8.

4. Solving the discretized problems

Our task now is to solve the finite dimensional problems (3.11) and (3.17). First, let us recall the augmented Lagrangian method we are dealing with.

4.1. ALG2 method.

Assume that V and Z are two Hilbert spaces. Let us consider the problem

$$\inf_{\phi \in V} \mathcal{F}(\phi) + \mathcal{G}(\Lambda\phi) \tag{4.18}$$

where $\mathcal{F} : V \longrightarrow (-\infty, +\infty]$ and $\mathcal{G} : Z \longrightarrow (-\infty, +\infty]$ are convex, l.s.c. and $\Lambda \in \mathcal{L}(V, Z)$.
We introduce a new variable $q \in Z$ to the primal problem (4.18) and we rewrite it in the form

$$\inf_{(\phi, q) \in V \times Z : \Lambda\phi = q} \mathcal{F}(\phi) + \mathcal{G}(q).$$

The augmented Lagrangian is given by

$$L(\phi, q; \sigma) := \mathcal{F}(\phi) + \mathcal{G}(q) + \langle \sigma, \Lambda\phi - q\rangle + \frac{r}{2}|\Lambda\phi - q|^2,\ r > 0.$$

The so called ALG2 algorithm is given as follows: For given $q_0, \sigma_0 \in Z$, we construct the sequences $\{\phi_i\}, \{q_i\}$ and $\{\sigma_i\}, i = 1, 2, ...,$ by

- Step 1: Minimizing $\inf_{\phi} L(\phi, q_i; \sigma_i)$, i.e.

$$\phi_{i+1} \in \arg\min_{\phi\in V} \left\{ \mathcal{F}(\phi) + \langle \sigma_i, \Lambda\phi\rangle + \frac{r}{2}|\Lambda\phi - q_i|^2 \right\}.$$

- Step 2: Minimizing $\inf_{q\in Z} L(\phi_{i+1}, q; \sigma_i)$, i.e.

$$q_{i+1} \in \arg\min_{q\in Z} \left\{ \mathcal{G}(q) - \langle \sigma_i, q\rangle + \frac{r}{2}|\Lambda\phi_{i+1} - q|^2 \right\}.$$

- Step 3: Update the multiplier σ,

$$\sigma_{i+1} = \sigma_i + r(\Lambda\phi_{i+1} - q_{i+1}).$$

For the theory of this method and its interpretation, we refer the reader to [13], [19], [20], [17], [18]. Here, we recall the convergence result of this method which is enough for our discretized problems.

Theorem 4.11 (cf. [13], Theorem 8)**.** *Fixed $r > 0$, assuming that $V = \mathbb{R}^n, Z = \mathbb{R}^m$ and that Λ has full column rank. If there exists a solution to the optimality relations* (2.9) *then $\{\phi_i\}$ converges to a solution of the primal problem* (2.7) *and $\{\sigma_i\}$ converges to a solution of the dual problem* (2.8)*. Moreover, $\{q_i\}$ converges to $\Lambda\phi^*$, where ϕ^* is the limit of $\{\phi_i\}$.*

The proof of this result in the case of finite dimensional spaces V and Z can be found in [13]. The result holds true in infinite dimensional Hilbert spaces under additional assumptions. One can see [20] and [17] for more details in this direction.

Next, we use the ALG2 method for the discretized problems. To simplify the notations, let us drop out the subscript h in (λ_h, u_h) and $(\Phi_h, \theta_h^0, \theta_h^1)$. Thanks to Remark 3.9, we treat separately the case where $\mathbf{m} = \nu(\overline{\Omega})$ and the case where $\mathbf{m} < \nu(\overline{\Omega})$.

4.2. **Partial transport ($\mathbf{m} < \nu(\overline{\Omega})$):** Given $(q_i, z_i, w_i), (\Phi_i, \theta_i^0, \theta_i^1)$ at the iteration i, we compute

- Step 1:

$$\begin{aligned}(\lambda_{i+1}, u_{i+1}) &\in \arg\min_{(\lambda,u)\in V_h} \mathcal{F}_h(\lambda, u) + \langle (\Phi_i, \theta_i^0, \theta_i^1), \Lambda_h(\lambda, u)\rangle + \frac{r}{2}|\Lambda_h(\lambda, u) - (q_i, z_i, w_i)|^2 \\ &= \arg\min_{(\lambda,u)\in V_h} -\langle u, f_h\rangle - \lambda(\mathbf{m} - \nu(\overline{\Omega})) + \langle \Phi_i, \nabla u\rangle + \langle \theta_i^0, -u\rangle + \langle \theta_i^1, u - \lambda\rangle \\ &\quad + \frac{r}{2}|\nabla u - q_i|^2 + \frac{r}{2}|u + z_i|^2 + \frac{r}{2}|u - \lambda - w_i|^2.\end{aligned}$$

- Step 2:

$$(q_{i+1}, z_{i+1}, w_{i+1}) \in \underset{(q,z,w)\in Z_h}{\arg\min}\ \mathcal{G}_h(q,z,w) - \langle(\Phi_i, \theta_i^0, \theta_i^1), (q,z,w)\rangle + \frac{r}{2}|\Lambda_h(\lambda_{i+1}, u_{i+1}) - (q,z,w)|^2$$
$$= \underset{(q,z,w)\in Z_h}{\arg\min}\ \mathbb{I}_{[F^*(.,q(.))\le 1]}(q) + \mathbb{I}_{[z\le 0]}(z) + \mathbb{I}_{[w\le 0]}(w) - \langle \Phi_i, q\rangle - \langle\theta_i^0, z\rangle - \langle\theta_i^1, w\rangle$$
$$+ \frac{r}{2}|\nabla u_{i+1} - q|^2 + \frac{r}{2}|u_{i+1} + z|^2 + \frac{r}{2}|u_{i+1} - \lambda_{i+1} - w|^2.$$

- Step 3: Update the multiplier

$$(\Phi_{i+1}, \theta_{i+1}^0, \theta_{i+1}^1) = (\Phi_i, \theta_i^0, \theta_i^1) + r(\nabla u_{i+1} - q_{i+1}, -u_{i+1} - z_{i+1}, u_{i+1} - \lambda_{i+1} - w_{i+1}).$$

Before giving numerical results, let us take a while to comment the above iteration. Overall, Step 1 is a quadratic programming. Step 2 can be computed easily in many cases and Step 3 updates obviously. We denote by $\text{Proj}_C(.)$ the projection onto a closed convex subset C.

- In Step 1, we split the computation of the couple (λ_{i+1}, u_{i+1}) into two steps: We first minimize w.r.t. u to compute u_{i+1} and then we use u_{i+1} to compute λ_{i+1}. More precisely, we proceed for Step 1 as follows:
  (1) For u_{i+1},

$$u_{i+1} \in \underset{u\in E_h}{\arg\min} -\langle u, f_h\rangle + \langle\Phi_i, \nabla u\rangle + \langle\theta_i^0, -u\rangle + \langle\theta_i^1, u\rangle$$
$$+ \frac{r}{2}|\nabla u - q_i|^2 + \frac{r}{2}|u + z_i|^2 + \frac{r}{2}|u - \lambda_i - w_i|^2.$$

  This is equivalent to

$$r\langle\nabla u_{i+1}, \nabla v\rangle + 2r\langle u_{i+1}, v\rangle = \langle f_h, v\rangle - \langle\Phi_i, \nabla v\rangle + \langle\theta_i^0, v\rangle - \langle\theta_i^1, v\rangle$$
$$+ r\langle q_i, \nabla v\rangle - r\langle z_i, v\rangle + r\langle\lambda_i + w_i, v\rangle, \ \forall v \in E_h.$$

  Remark here that the equation is linear with a symmetric positive definite coefficient matrix.

  (2) For λ_{i+1}, it is computed explicitly

$$\lambda_{i+1} \in \underset{s\in\mathbb{R}}{\arg\min} -s(\mathbf{m} - \nu(\overline{\Omega})) + \langle\theta_i^1, u_{i+1} - s\rangle + \frac{r}{2}\langle u_{i+1} - s - w_i, u_{i+1} - s - w_i\rangle$$
$$= -\frac{\nu(\overline{\Omega}) - \mathbf{m} - \int_{\overline{\Omega}} \theta_i^1 + r\int_\Omega (w_i - u_{i+1})}{r\int_\Omega 1}.$$

- In Step 2, the variables q, z, w are independent. So, we solve them separately:
  (1) For z_{i+1} and w_{i+1}, if we choose P_2 finite element for z_{i+1} and w_{i+1}, at vertex x_k,

$$z_{i+1}(x_k) = \text{Proj}_{[r\in\mathbb{R}:\, r\le 0]}\left(-u_{i+1}(x_k) + \frac{\theta_i^0(x_k)}{r}\right)$$
$$= \min\left(-u_{i+1}(x_k) + \frac{\theta_i^0(x_k)}{r}, 0\right);$$

and

$$w_{i+1}(x_k) = \mathrm{Proj}_{[r\in\mathbb{R}:\, r\le 0]}\left(u_{i+1}(x_k) - \lambda_{i+1} + \frac{\theta_i^1(x_k)}{r}\right)$$
$$= \min\left(u_{i+1}(x_k) - \lambda_{i+1} + \frac{\theta_i^1(x_k)}{r}, 0\right).$$

(2) For q_{i+1}, if we choose P_1 finite element for q_{i+1}, then at each vertex x_l

$$q_{i+1}(x_l) = \mathrm{Proj}_{B_{F^*(x_l,.)}}\left(\nabla u_{i+1}(x_l) + \frac{\Phi_i(x_l)}{r}\right),$$

where $B_{F^*(x,.)} := \left\{q \in \mathbb{R}^N : F^*(x,q) \le 1\right\}$ the unit ball for $F^*(x,.)$.

It remains to explain how we compute the projection onto $B_{F^*(x_l,.)}$. This issue is recently discussed in [8] for Riemann-type Finsler distances and for crystalline norms. For the convenience of the reader, we retake here the case where the unit ball of $F(x,.)$ is (not necessarily symmetric) convex polytope. For short, we ignore the dependence of x in F and F^*. Given $d_1, ..., d_k \neq 0$ such that, for any $0 \neq v \in \mathbb{R}^N$, $\max_{1\le i\le k}\{\langle v, d_i\rangle\} > 0$. We consider the nonsymmetric Finsler metric given by

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

It is not difficult to see that the unit ball B^* corresponding to F^* is exactly the convex hull of $\{d_i\}$,

$$B^* = \mathrm{conv}(d_i, i = 1, ..., k).$$

Thus we need to compute the projection onto the convex hull of finite points. In dimension two, the projection onto B^* can be performed as follows: Compute the successive vertices $S_1, ..., S_n$. If $q \notin B^*$, then compute the projections of q onto the segments $[S_i, S_{i+1}]$ and compare among these projections to chose the right one. Another way is as the one in [8]: Compute outward orthogonal vectors $v_1, ..., v_n$ (see Fig. 1). If q belongs to $[S_i, S_{i+1}] + \mathbb{R}_+ v_i$, then the projection coincides with the one on the line through S_i, S_{i+1}. If q belongs to the sector $S_i + \mathbb{R}_+ v_{i-1} + \mathbb{R}_+ v_i$, the projection is S_i.

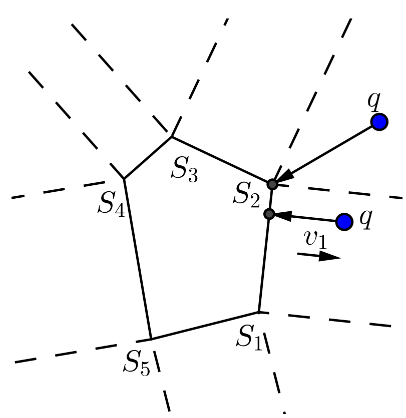


FIGURE 1. Illustration of the projection

4.3. **Unbalanced transport ($\mathbf{m} = \nu(\overline{\Omega})$):** Thanks to Remark 3.9, we can reduce the algorithm in this particular case by ignoring the variable λ. With similar considerations for $\Lambda_h u = (\nabla u, -u)$, we get the following iteration

- Step 1:

$$u_{i+1} \in \operatorname*{arg\,min}_{u \in E_h} -\langle u, f_h\rangle + \langle \Phi_i, \nabla u\rangle + \langle \theta_i^0, -u\rangle + \frac{r}{2}|\nabla u - q_i|^2 + \frac{r}{2}|u + z_i|^2.$$

  Equivalently,

$$r\langle \nabla u_{i+1}, \nabla v\rangle + r\langle u_{i+1}, v\rangle = \langle f_h, v\rangle - \langle \Phi_i, \nabla v\rangle + \langle \theta_i^0, v\rangle + r\langle q_i, \nabla v\rangle - r\langle z_i, v\rangle,\ \forall v \in E_h.$$

- Step 2:
  (1) For z_{i+1}, choosing P_2 finite element for z_{i+1}, then at each vertex x_k,

$$z_{i+1}(x_k) = \operatorname{Proj}_{[r\in\mathbb{R}:\, r\le 0]}\left(-u_{i+1}(x_k) + \frac{\theta_i^0(x_k)}{r}\right) = \min\left(-u_{i+1}(x_k) + \frac{\theta_i^0(x_k)}{r}, 0\right).$$

  (2) For q_{i+1}, choosing P_1 finite element, at vertex x_l,

$$q_{i+1}(x_l) = \operatorname{Proj}_{B_{F^*(x_l,\cdot)}}\left(\nabla u_{i+1}(x_l) + \frac{\Phi_i(x_l)}{r}\right).$$

- Step 3: $(\Phi_{i+1}, \theta_{i+1}^0) = (\Phi_i, \theta_i^0) + r(\nabla u_{i+1} - q_{i+1}, -u_{i+1} - z_{i+1})$.

5. Numerical experiments

For the numerical implementation, we use the FreeFem++ software [21] and base on [5], [6]. We use P_2 finite element for $u_i, z_i, w_i, \theta_i^0, \theta_i^1$ and P_1 finite element for Φ_i, q_i.

5.1. **Stopping criterion.** In computational version, the measures μ and ν are approximated by nonnegative regular functions that we denote again by μ and ν. We use the following stopping criteria:

- For the partial transport:
  (1) MIN-MAX $:= \min\left\{\min_{\overline{\Omega}} u(x), \lambda - \max_{\overline{\Omega}} u(x), \min_{\overline{\Omega}} \theta^0(x), \min_{\overline{\Omega}} \theta^1(x)\right\}$.
  (2) Max-Lip $:= \sup_{\overline{\Omega}}(F^*(x, \nabla u(x)))$.
  (3) DIV $:= \|\nabla\cdot\Phi + \nu - \theta^1 - \mu + \theta^0\|_{L^2}$.
  (4) DUAL $:= \|F(x, \Phi(x)) - \Phi(x)\cdot\nabla u\|_{L^2}$.
  (5) MASS $:= \left|\int(\nu - \theta^1)dx - \mathbf{m}\right|$.
- For the unbalanced transport: We change
  (1) MIN-MAX $:= \min\left\{\min_{\overline{\Omega}} u(x), \min_{\overline{\Omega}} \theta^0(x)\right\}$.
  (2) DIV $:= \|\nabla\cdot\Phi + \nu - \mu + \theta^0\|_{L^2}$.

We expect MIN-MAX ≥ 0, Max-Lip ≤ 1; DIV, DUAL and MASS are small.

5.2. **Some examples.** In all the examples below, we take $\Omega = [0,1] \times [0,1]$. We test for the Riemannian case and the crystalline case. For the later one, we consider the Finsler metric of the form $F(x,v) = \max_{1 \le i \le k} \{\langle v, d_i \rangle\}$ with given directions $d_1, ..., d_k$ such that for any $0 \neq v \in \mathbb{R}^2$,

$$\max_{1 \le i \le k} \{\langle v, d_i \rangle\} > 0.$$

5.2.1. *For the unbalanced transport.*

Example 5.12. Taking $\mu = 3\mathcal{L}^2{}_\Omega$ and $\nu = \delta_{(0.5,0.5)}$ the Dirac mass at $(0.5, 0.5)$. The Finsler metric is the Euclidean one. The optimal flow is given in Figure 2. The stopping criterion at each iteration is given in Figure 3.

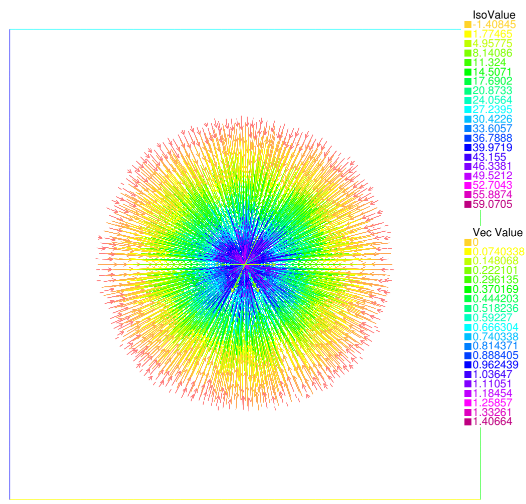
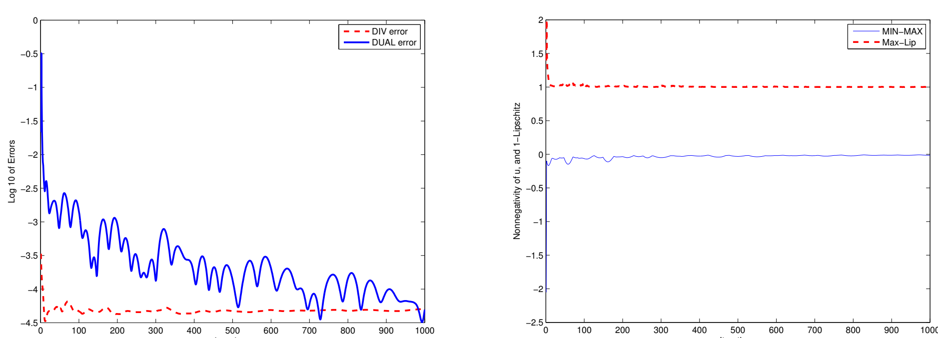


FIGURE 2. Optimal flow for $\mu = 3$, $\nu = \delta_{(0.5,0.5)}$, $F(x,v) = |v|$.



(A) log_{10} of DIV and DUAL errors

(B) Feasibility of u and θ^0

FIGURE 3. Stopping criterion at each iteration

Example 5.13. We take μ and ν as in the previous example, and the Finsler metric given by $F(x, v) := |v_1| + |v_2|$, for $v = (v_1, v_2) \in \mathbb{R}^2$. This corresponds to the crystalline norm with $d_1 = (1, 1), d_2 = (-1, 1), d_3 = (-1, -1), d_4 = (1, -1)$. The optimal flow is given in Figure 4 and the stopping criterion at each iteration is given in Figure 5.

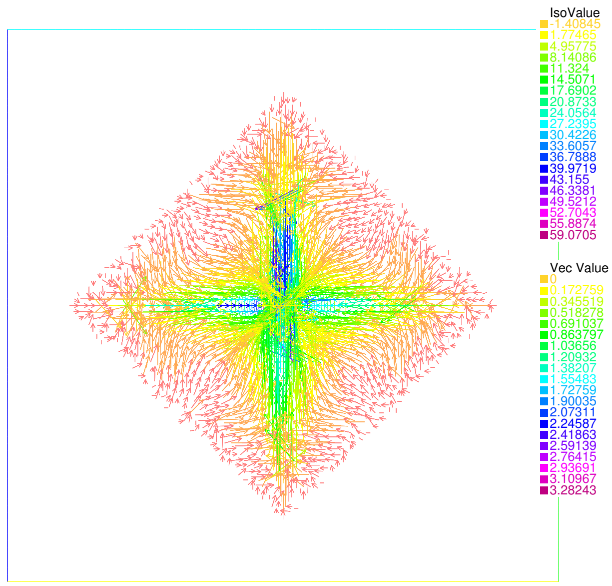
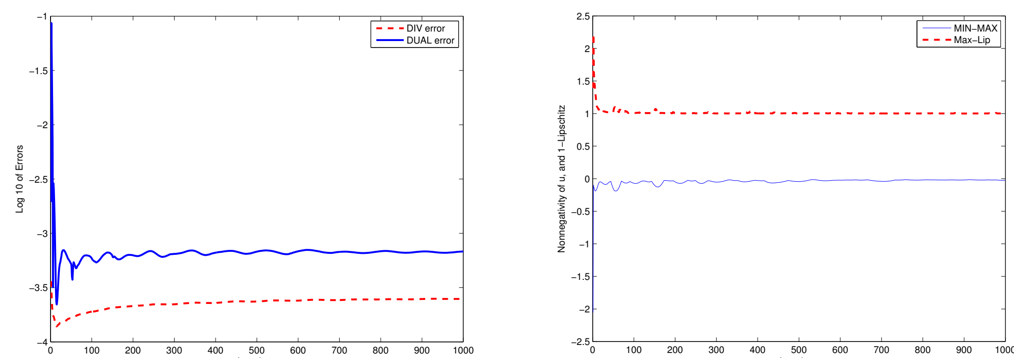


FIGURE 4. Optimal flow for $\mu = 3,\ \nu = \delta_{(0.5,0.5)},\ F(x, (v_1, v_2)) = |v_1| + |v_2|$.



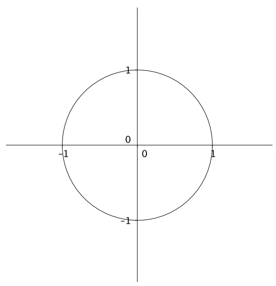
(A) log_{10} of DIV and DUAL errors

(B) Feasibility of u and θ^0

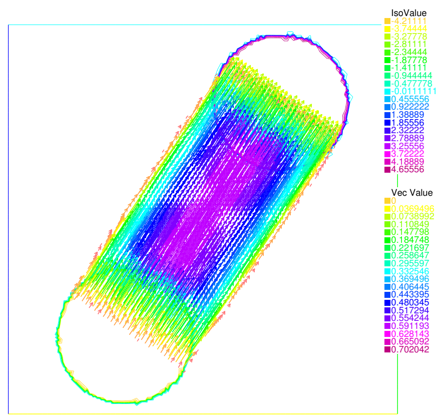
FIGURE 5. Stopping criterion at each iteration

5.2.2. *For the partial transport.*

Example 5.14. Taking $\mu = 4\chi_{[(x-0.3)^2+(y-0.2)^2<0.03]}$, and $\nu = 4\chi_{[(x-0.7)^2+(y-0.8)^2<0.03]}$. The mass of the transport is $\mathbf{m} := \dfrac{\nu(\overline{\Omega})}{2}$. We test for different Finsler metrics. On each figure below, the subfigure at left illustrates the unit ball of F and the subfigure at right gives the numerical result (see Figures 6, 7, 8 and 9). The stopping criteria are summarized in Table 1.

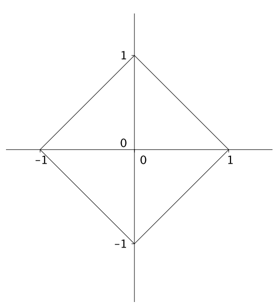


(A) The unit ball of F

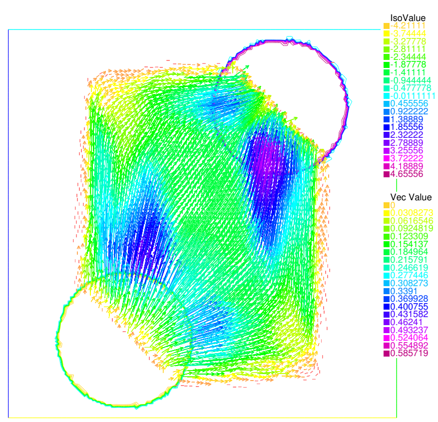


(B) Optimal flow

FIGURE 6. Case 1: $F(x, v) = |v|$.

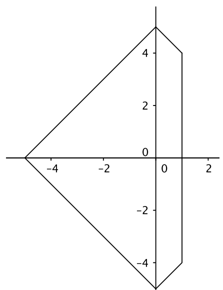


(A) The unit ball of F

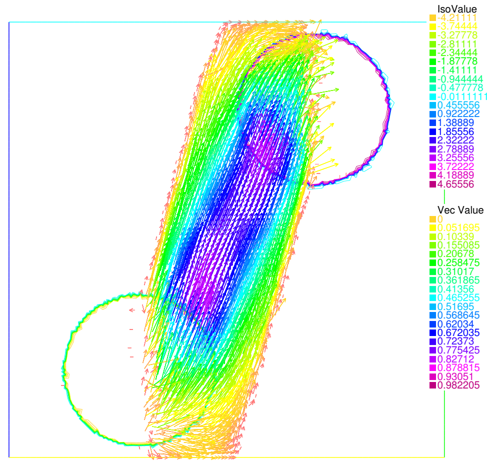


(B) Optimal flow

FIGURE 7. Case 2: The crystalline case with $d_1 = (1, 1), d_2 = (-1, 1), d_3 = (-1, -1), d_4 = (1, -1)$.

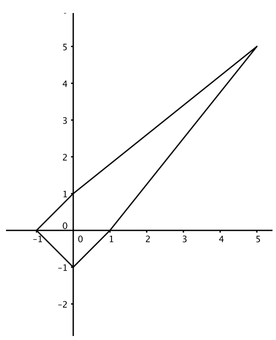


(A) The unit ball of F

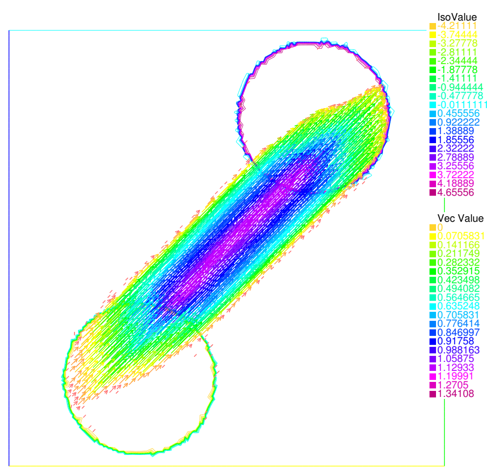


(B) Optimal flow

FIGURE 8. Case 3: The crystalline case with $d_1 = (1,0), d_2 = (\frac{1}{5}, \frac{1}{5}), d_3 = (-\frac{1}{5}, \frac{1}{5}), d_4 = (-\frac{1}{5}, -\frac{1}{5}), d_5 = (\frac{1}{5}, -\frac{1}{5})$ makes the transport more expensive in the direction of the vector $(1,0)$.



(A) The unit ball of F



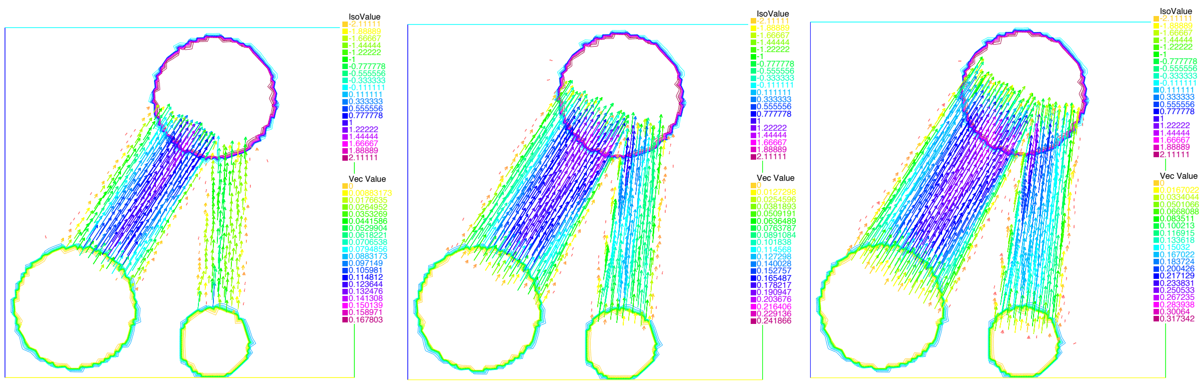
(B) Optimal flow

FIGURE 9. Case 4: The crystalline case with $d_1 = (1,-1), d_2 = (1, -\frac{4}{5}), d_3 = (-\frac{4}{5}, 1), d_4 = (-1,1), d_5 = (-1,-1)$ makes the transport cheaper in the direction of the vector $(1,1)$.

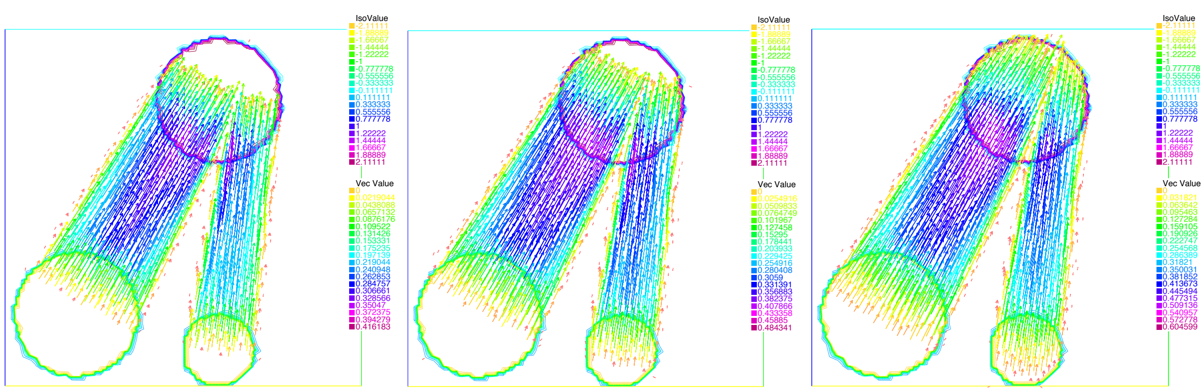
TABLE 1. Stopping criteria for 800 iterations

Case	DIV	DUAL	MASS	MIN-MAX	Max-Lip	Time execution (s)
1	2.48182e-05	9.5294e-06	0.000161361	-0.0149942	1.00068	357
2	3.38395e-05	5.58717e-05	0.000195881	-0.00120123	1.00248	867
3	7.44768e-05	5.5997e-05	6.66404e-06	-0.00272389	1.00351	1269
4	6.33726e-05	3.20691e-05	0.000120909	-0.0104915	1.02572	1123

Example 5.15. Let $\mu = 2\chi_{[(x-0.2)^2+(y-0.2)^2<0.03]} + 2\chi_{[(x-0.6)^2+(y-0.1)^2<0.01]}$, and $\nu = 2\chi_{[(x-0.6)^2+(y-0.8)^2<0.03]}$. In this example, we take the Euclidean norm and we let $\mathbf{m}$ vary by taking the values $\mathbf{m}_i = \dfrac{i}{6}\min\{\mu(\Omega), \nu(\Omega)\}$, $i = 1, ..., 6$. The results are given in Figure 10.



(A) $\mathbf{m}_1$ (B) $\mathbf{m}_2$ (C) $\mathbf{m}_3$



(D) $\mathbf{m}_4$ (E) $\mathbf{m}_5$ (F) $\mathbf{m}_6$

FIGURE 10. Optimal flows

6. Appendix

Our aim here is to show Lemma 6.2 which gives a smooth approximation of 1-d_F Lipschitz continuous function for continuous non-degenerate Finsler metrics F. This result is more or less known in some particular cases. However, we couldn't find any rigorous proofs for the general case in the litterature.

Lemma 6.2. *Let Ω be a connected bounded Lipschitz domain and F be a continuous non-degenerate Finsler metric on $\overline{\Omega}$. For any Lipschitz continuous function u on $\overline{\Omega}$ satisfying*

$$F^*(x, \nabla u(x)) \leq 1, \quad a.e. \;\; x \in \Omega, \tag{6.19}$$

there exists a sequence of functions $u_\varepsilon \in C_c^\infty(\mathbb{R}^N)$ such that

$$F^*(x, \nabla u_\varepsilon(x)) \leq 1, \forall x \in \overline{\Omega},$$

and

$$u_\varepsilon \rightrightarrows u \;\; \textit{uniformly on} \;\; \overline{\Omega}.$$

Note that F and F^* are defined only in $\overline{\Omega}$ and that the gradient of u is controlled only inside of Ω by (6.19). If we use the standard convolution to define u_ε, the value of $u_\varepsilon(x)$ is affected by the value of $u(y)$ outside of $\overline{\Omega}$ which remains uncontrolled. To overcome this difficulty, if x is near the boundary, we move it a little into inside of Ω before taking the convolution. To this aim, we use the smooth partition of unity tool to deal with approximation of u near the boundary.

Proof. Set

$$\forall x \in \mathbb{R}^N, \, \tilde{u}(x) := \begin{cases} u(x) \;\text{ if }\; x \in \overline{\Omega} \\ 0 \;\text{ otherwise.} \end{cases}$$

Step 1: Fix $z \in \partial\Omega$. Since Ω is a Lipschitz domain, there exist $r_z > 0$ and a Lipschitz continuous function $\gamma_z : \mathbb{R}^{N-1} \longrightarrow \mathbb{R}$ such that (up to rotating and relabeling if necessary)

$$\Omega \cap B(z, r_z) = \{x \,|\, x_N > \gamma_z(x_1, ..., x_{N-1})\} \cap B(z, r_z).$$

Set $U_z := \Omega \cap B(z, \frac{r_z}{2})$. For any $x \in \mathbb{R}^N$, taking

$$x_z^\varepsilon := x + \varepsilon\lambda_z e_n \tag{6.20}$$

where we choose a sufficiently large fixed λ_z and all small ε, say fixed $\lambda_z \geq \mathrm{Lip}(\gamma_z) + 1$, $0 < \varepsilon < \dfrac{r_z}{2(\lambda_z + 1)}$. By this choice and the Lipschitz property of γ_z, we see that

$$B(x_z^\varepsilon, \varepsilon) \subset \Omega \cap B(z, r_z) \;\text{ for all }\; x \in U_z. \tag{6.21}$$

Defining

$$\tilde{u}_\varepsilon(x) := \int\limits_{\mathbb{R}^N} \rho_\varepsilon(y)\tilde{u}(x_z^\varepsilon - y)dy = \int\limits_{B(x_z^\varepsilon, \varepsilon)} \rho_\varepsilon(x_z^\varepsilon - y)\tilde{u}(y)dy, \;\text{ for all }\; x \in \mathbb{R}^N, \tag{6.22}$$

where ρ_ε is the standard mollifier on $\mathbb{R}^N$. Obviously, $\tilde{u}_\varepsilon \in C_c^\infty(\mathbb{R}^N)$. Using (6.21), (6.22) and the continuity of u on $\overline{\Omega}$, we get

$$\tilde{u}_\varepsilon \rightrightarrows u \text{ on } \overline{U}_z.$$

Step 2: Now, using the compactness of $\partial\Omega$ and $\partial\Omega \subset \bigcup_{z\in\partial\Omega} B(z, \frac{r_z}{2})$, there exist $z_1, ..., z_n \in \partial\Omega$ such that

$$\partial\Omega \subset \bigcup_{i=1}^n B(z_i, \frac{r_{z_i}}{2}).$$

For short, we write r_i, U_i, x_i instead of $r_{z_i}, U_{z_i}, x_{z_i}$. Taking an open set $U_0 \Subset \Omega$ such that

$$\overline{\Omega} \subset \bigcup_{i=1}^n B(z_i, \frac{r_i}{2}) \bigcup U_0.$$

Let $\{\phi\}_{i=0}^n$ be a smooth partition of unity on $\overline{\Omega}$, subordinate to $\Big\{U_0, B(z_1, \frac{r_1}{2}), ..., B(z_n, \frac{r_n}{2})\Big\}$, that is,

$$\begin{cases} \phi_i \in C_c^\infty(\mathbb{R}^N), 0 \le \phi_i \le 1, \forall i = 0, ..., n \\ \text{supp}(\phi_i) \Subset B(z_i, \frac{r_i}{2}), \ \forall i = 1, ..., n, \ \text{supp}(\phi_0) \Subset U_0 \\ \sum_{i=0}^n \phi_i(x) = 1 \ \text{ for all } \ x \in \overline{\Omega}. \end{cases}$$

Due to Step 1, there exist $\tilde{u}_\varepsilon^1, ..., \tilde{u}_\varepsilon^n \in C_c^\infty(\mathbb{R}^N)$ such that

$$\tilde{u}_\varepsilon^i \rightrightarrows u \text{ on } \overline{U}_i, i = 1, ..., n.$$

For $i = 0$, since $U_0 \Subset \Omega$, we can take $\tilde{u}_\varepsilon^0 := \rho_\varepsilon \star \tilde{u} \in C_c^\infty(\mathbb{R}^N)$ and $\tilde{u}_\varepsilon^0 \rightrightarrows u$ on $\overline{U}_0$. Set

$$u_\varepsilon := \frac{1}{1 + C\varepsilon + w(\varepsilon)} \sum_{i=0}^n \phi_i \tilde{u}_\varepsilon^i,$$

where C is chosen later and

$$w(\varepsilon) := \sup\{|F^*(x,p) - F^*(y,p)| : x, y \in \overline{\Omega}, |x - y| \le M\varepsilon, |p| \le \|\nabla u\|_{L^\infty}\},$$

with constant $M := \max_{1\le i\le n}\{\lambda_{z_i} + 1\}$, λ_{z_i} is given in Step 1. We show that u_ε satisfies all the desired properties. By the construction, $u_\varepsilon \in C_c^\infty(\mathbb{R}^N)$ and

$$u_\varepsilon \rightrightarrows \sum_{i=0}^n \phi_i u = u \text{ on } \overline{\Omega}.$$

At last, we show that $F^*(x, \nabla u_\varepsilon(x)) \le 1, \forall x \in \overline{\Omega}$. Indeed, for any $x \in \Omega$, if $x \in U_i, i = 1, ..., n$ (near the boundary of Ω), we move x a bit into inside of Ω to $x_i^\varepsilon := x_{z_i}^\varepsilon$ (see (6.20) and (6.21)), if $x \in U_0$, set $x_0^\varepsilon = x$. We have

$$\begin{aligned}\nabla u_\varepsilon(x) &= \frac{1}{1+C\varepsilon+w(\varepsilon)}\left(\sum_{i=0}^n \nabla\phi_i(x)\tilde{u}_\varepsilon^i(x) + \sum_{i=0}^n \phi_i(x)\nabla\tilde{u}_\varepsilon^i(x)\right)\\ &= \frac{1}{1+C\varepsilon+w(\varepsilon)}\left(\sum_{i=0}^n \nabla\phi_i(x)\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(y)dy\right.\\ &\qquad\left. + \sum_{i=0}^n \phi_i(x)\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)\nabla u(y)dy\right).\end{aligned}$$

The first sum on the right hand side has a small norm. Indeed, using the fact that

$$\sum_{i=0}^n \nabla\phi_i(x)u(x) = 0, \text{ for all } x \in \Omega,$$

we have

$$\sum_{i=0}^n \nabla\phi_i(x)\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(y)dy = \sum_{i=0}^n \nabla\phi_i(x)\left(\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(y)dy - u(x)\right). \quad (6.23)$$

Moreover,

$$\begin{aligned}\left|\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(u)\,dy - u(x)\right| &\le \left|\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)\left(u(y) - u(x_i^\varepsilon)\right)dy\right| + |u(x_i^\varepsilon) - u(x)|\\ &\le C_1\varepsilon, \forall i = 0, ..., n,\end{aligned}$$

where the constant C_1 depends only on $\mathrm{Lip}(\gamma_{z_i})$ and the Lipschitz constant of u on $\overline{\Omega}$. Thus, by combing this with (6.23),

$$\left|\sum_{i=0}^n \nabla\phi_i(x)\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(y)dy\right| \le C_2\varepsilon, \forall x \in \Omega,$$

where C_2 depends only on C_1 and $\|\nabla\phi_i\|_{L^\infty}$.

Using the non-degeneracy of F, we have

$$F^*\left(x, \sum_{i=0}^n \nabla\phi_i(x)\int_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y)u(y)dy\right) \le C_3\varepsilon \text{ for all } x \in \Omega.$$

Fixed any $x \in \Omega$, if $y \in B(x_i^\varepsilon, \varepsilon)$ then $|x-y| \le |x-x_i^\varepsilon| + |x_i^\varepsilon - y| \le M\varepsilon$. So we obtain

$$\begin{aligned}
F^*(x, \nabla u_\varepsilon(x)) &\le \frac{1}{1+C\varepsilon+w(\varepsilon)} \Big[F^*\Big(x, \sum_{i=0}^n \nabla\phi_i(x) \int\limits_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y) u(y) dy\Big) \\
&\qquad + F^*\Big(x, \sum_{i=0}^n \phi_i(x) \int\limits_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y) \nabla u(y) dy\Big)\Big] \\
&\le \frac{1}{1+C\varepsilon+w(\varepsilon)} \left(C_3\varepsilon + \sum_{i=0}^n \phi_i(x) \int\limits_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y) F^*(x, \nabla u(y)) dy \right) \\
&\le \frac{1}{1+C\varepsilon+w(\varepsilon)} \Big[C_3\varepsilon + \sum_{i=0}^n \phi_i(x) \int\limits_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y) F^*(y, \nabla u(y))\, dy \\
&\qquad + \sum_{i=0}^n \phi_i(x) \int\limits_{B(x_i^\varepsilon,\varepsilon)} \rho_\varepsilon(x_i^\varepsilon - y) \left(F^*(x, \nabla u(y)) - F^*(y, \nabla u(y))\right) dy \Big] \\
&\le \frac{C_3\varepsilon + 1 + w(\varepsilon)}{1+C\varepsilon+w(\varepsilon)} \\
&\le 1 \text{ (choose a constant } C \ge C_3).
\end{aligned}$$

By the continuity of ∇u_ε and of F^*, we also have $F^*(x, \nabla u_\varepsilon(x)) \le 1$, $\forall x \in \overline{\Omega}$. □

Proposition 6.16. *Let F be a continuous non-degenerate Finsler metric on a connected bounded Lipshitz domain Ω. We have*

$$Lip_{d_F} = \left\{u : \overline{\Omega} \longrightarrow \mathbb{R} \,|\, u \text{ is Lipschitz continuous and } F^*(x, \nabla u(x)) \le 1, \text{ a.e. } x \in \Omega\right\} := \mathcal{B}_{F^*}.$$

As a consequence, for any 1-d_F Lipschitz continuous function u, there exists a sequence of 1-d_F Lipschitz continuous functions $u_\varepsilon \in C_c^\infty(\mathbb{R}^N)$ and $u_\varepsilon \rightrightarrows u$ uniformly on $\overline{\Omega}$.

Lemma 6.3. *We have $Lip_{d_F} \subset \mathcal{B}_{F^*}$.*

Proof. Let $u \in Lip_{d_F}$. Then u is Lipschitz and u is differentiable a.e. in Ω. Let $x \in \Omega$ be any point where u is differentiable. We have, for any $v \in \mathbb{R}^N$,

$$\begin{aligned}
\frac{\langle \nabla u(x), v\rangle}{F(x,v)} &= \lim_{h\to 0} \frac{u(x+hv) - u(x)}{F(x,hv)} \\
&\le \limsup_{h\to 0} \frac{d_F(x, x+hv)}{F(x,hv)} \\
&\le \limsup_{h\to 0} \frac{\int_0^1 F(x+thv, hv)dt}{F(x,hv)} = 1.
\end{aligned}$$

Hence, $F^*(x, \nabla u(x)) \leq 1$. So $u \in \mathcal{B}_{F^*}$. □

Lemma 6.4. *We have* $\mathcal{B}_{F^*} \subset Lip_{d_F}$.

Proof. Fix any $u \in \mathcal{B}_{F^*}$.
Case 1: If u is smooth, then $F^*(x, \nabla u(x)) \leq 1, \forall x \in \overline{\Omega}$. For any $x, y \in \overline{\Omega}$ and any Lipschitz curve ξ in $\overline{\Omega}$ joining x and y, we have

$$\begin{aligned} u(y) - u(x) &= \int_0^1 \nabla u(\xi(t))\dot{\xi}(t)dt \\ &\leq \int_0^1 F^*(\xi(t), \nabla u(\xi(t)))F(\xi(t), \dot{\xi}(t))dt \\ &\leq \int_0^1 F(\xi(t), \dot{\xi}(t))dt. \end{aligned}$$

Hence $u \in Lip_{d_F}$.
Case 2: For general Lipschitz continuous function u satisfying $F^*(x, \nabla u(x)) \leq 1$, a.e. $x \in \Omega$, thanks to Lemma 6.2, there exist $u_\varepsilon \in \mathcal{B}_{F^*} \bigcap C_c^\infty(\mathbb{R}^N)$ such that $u_\varepsilon \rightrightarrows u$ on $\overline{\Omega}$. According to Case 1 above, $u_\varepsilon \in Lip_{d_F}$. Since $u_\varepsilon \rightrightarrows u$ on $\overline{\Omega}$, we obtain $u \in Lip_{d_F}$. □

Proof of Proposition 6.16. The proof follows by Lemma 6.3 and Lemma 6.4. □

Proof of Lemma 2.1. Since $0 \leq u \leq \lambda$, the sequence u_ε in the proof of Lemma 6.2 satisfies $0 \leq u_\varepsilon \leq \lambda$. So $u_\varepsilon \in C_c^\infty(\mathbb{R}^N) \cap L^\lambda_{d_F}$ and $u_\varepsilon \rightrightarrows u$ on $\overline{\Omega}$. □

Remark 6.17. *The results still hold true if* Ω *is connected, bounded and has the segment property.*

Acknowledgements

The authors are grateful to J. D. Benamou and G. Carlier who provide some codes of ALG2 on the link `https://team.inria.fr/mokaplan/software/`. Some parts of our codes are inspired from their work.

References

[1] L. Ambrosio, N. Fusco and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford Mathematical Monographs, (2000).
[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics, Birkhäuser (2005).
[3] J. W. Barrett and L. Prigozhin. Partial L^1 Monge-Kantorovich problem: Variational formulation and numerical approximation. *Interfaces and Free Boundaries*, 11(2), (2009), 201-238.
[4] M. Beckmann. A continuous model of transportation. *Econometrica: Journal of the Econometric Society*, 20(4), (1952), 643-660.
[5] J. D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3), (2000), 375–393.

[6] J. D. Benamou and G. Carlier. Augmented Lagrangian Methods for Transport Optimization, Mean Field Games and Degenerate Elliptic Equations. *J. Optim. Theory Appl.*, 167(1), (2015), 1-26.
[7] J. D. Benamou, G. Carlier, M. Cuturi, L. Nenna and G. Peyré. Iterative Bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2), (2015), A1111-A1138.
[8] J. D. Benamou, G. Carlier and R. Hatchi. A numerical solution to Monge's problem with a Finsler distance cost. Submitted.
[9] L. Caffarelli and R.J. McCann. Free boundaries in optimal transport and Monge-Ampere obstacle problems. *Annals of Mathematics*, 171(2), (2010), 673-730.
[10] L. Chizat, G. Peyré, B. Schmitzer and F. X. Vialard. Scaling algorithms for unbalanced transport problems. arXiv preprint arXiv:1607.05816 (2016).
[11] L. De Pascale, L. C. Evans and A. Pratelli. Integral estimates for transport densities. *Bulletin of the London Mathematical Society*, 36(3), (2004), 383-395.
[12] L. De Pascale and A. Pratelli. Sharp summability for Monge transport density via interpolation. *ESAIM: Control, Optimisation and Calculus of Variations*, 10(4), (2004), 549-552.
[13] J. Eckstein and D. P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1), (1992), 293-318.
[14] I. Ekeland and R. Teman. *Convex analysis and variational problems.* North-Holland American Elsevier (1976).
[15] M. Feldman and R. J. McCann. Uniqueness and transport density in Monge's mass transportation problem. *Calc. Var. Partial Differential Equations*, 15(1), (2002), 81-113.
[16] A. Figalli. The Optimal Partial Transport Problem. *Arch. Rational Mech. Anal.*, 195(2), (2010), 533-560.
[17] M. Fortin and R. Glowinski. *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems.* North-Holland, Vol 15 (1983).
[18] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers and Mathematics with Applications*, 2(1), (1976), 17-40.
[19] R. Glowinski and J. L. Lions and R. Trémolières. *Numerical Analysis of Variational Inequalities.* North-Holland Publishing Company (1981).
[20] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics.* SIAM, Vol 9 (1989).
[21] F. Hecht. New development in freefem++. *Journal of Numerical Mathematics*, 20(3-4), (2012), 251-266.
[22] N. Igbida, J. M. Mazón, J. D. Rossi and J. Toledo. Optimal mass transportation for costs given by Finsler distances via p-Laplacian approximations. To appear in *Advances in Calculus of Variations.*
[23] N. Igbida and V. T. Nguyen. Optimal partial mass transportation and obstacle Monge-Kantorovich equation. Submitted.
[24] N. Igbida and N. N. Ta Thi. Sub-gradient Diffusion Operator. To appear in *J. Differential Equations.*
[25] G. D. Philippis, A. R. Mészáros, F. Santambrogio and B. Velichkov. BV estimates in optimal transportation and applications. *Arch. Rational Mech. Anal.*, 219(2), (2016), 829-860.
[26] F. Santambrogio. Absolute continuity and summability of transport densities: simpler proofs and new estimates. *Calc. Var. Partial Differential Equations*, 36(3), (2009), 343–354.
[27] F. Santambrogio. *Optimal Transport for Applied Mathematicians.* Birkäuser (2015).
[28] C. Villani. *Topics in Optimal Transportation.* Graduate Studies in Mathematics. Vol. 58, (2003).
[29] C. Villani. *Optimal Transport, Old and New.* Grundlehren des Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), Vol. 338, Springer-Verlag, Berlin-New York, (2009).

(**Noureddine IGBIDA**) Institut de recherche XLIM-DMI, UMR-CNRS 6172, Faculté des Sciences et Techniques, Université de Limoges, France.
E-mail address: noureddine.igbida@unilim.fr

(**Van Thanh NGUYEN**) Institut de recherche XLIM-DMI, UMR-CNRS 6172, Faculté des Sciences et Techniques, Université de Limoges, France.
E-mail address: van-thanh.nguyen@unilim.fr