Interfaces Free Bound. (Online first)
DOI 10.4171/IFB/514

Published by EMS Press

Monge–Kantorovich interpolation with constraints and application to a parking problem

Giuseppe Buttazzo, Guillaume Carlier, and Katharina Eichinger

Abstract. We consider optimal transport problems where the cost for transporting a given probability measure μ_0 to another one (μ_1) consists of two parts: the first one measures the transportation from μ_0 to an intermediate (pivot) measure μ to be determined (and subject to various constraints), and the second one measures the transportation from μ to μ_1. This leads to Monge–Kantorovich interpolation problems under constraints for which we establish various properties of the optimal pivot measures μ. Considering the more general situation where only some part of the mass uses the intermediate stop leads to a mathematical model for the optimal location of a parking region around a city. Numerical simulations, based on entropic regularization, are presented both for the optimal parking regions and for Monge–Kantorovich constrained interpolation problems.

1. Introduction

We consider optimal transport problems where a given probability measure μ_0 in $\mathbb{R}^d$ has to be transported to a given probability measure μ_1 with minimal transportation cost. This cost consists of two parts: the first one measures the transportation from μ_0 to an intermediate measure μ, to be determined, and the second one measures the transportation from μ to μ_1. This situation occurs in some applications where the transport of μ_0 to μ_1 is not directly made but the possibility of an intermediate stop is taken into account. The two parts are described by the Monge–Kantorovich functionals $W_{c_0}(\mu_0, \mu)$ and $W_{c_1}(\mu, \mu_1)$ respectively, where for every pair of probabilities ρ_0, ρ_1 we set

$$W_c(\rho_0, \rho_1) = \inf\Big\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y)\, \mathrm{d}\gamma(x, y) : \gamma \in \Pi(\rho_0, \rho_1)\Big\}. \tag{1.1}$$

Here

$$\Pi(\rho_0, \rho_1) := \big\{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : \pi_{i\#}\gamma = \rho_i,\ i = 0, 1\big\}$$

is the set of transport plans between ρ_0 and ρ_1. Denoting by $\pi_i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ ($i = 0, 1$) the projections on the first and second factor respectively, $\pi_{i\#}\gamma$ are the marginals of γ, so

Mathematics Subject Classification 2020: 49Q22 (primary); 49J45, 49M29, 49K99 (secondary).
Keywords: optimal transport, Monge–Kantorovich distance, measure interpolation, optimal parking regions.

that a probability measure γ on $\mathbb{R}^d \times \mathbb{R}^d$ belongs to $\Pi(\rho_0, \rho_1)$ when

$$\begin{cases} \pi_{0\#}\gamma(A) = \gamma(A \times \mathbb{R}^d) = \rho_0(A), \\ \pi_{1\#}\gamma(A) = \gamma(\mathbb{R}^d \times A) = \rho_1(A) \end{cases} \qquad \text{for all Borel set } A \subset \mathbb{R}^d.$$

Some extra constraints on the *pivot* measure μ can be added, for instance:

- location constraints, where the support of μ, $\operatorname{spt}\mu$, is required to be contained in a given region $K \subset \mathbb{R}^d$;
- density constraints, where the measure μ is required to be absolutely continuous and with a density not exceeding a prescribed function ϕ.

Without additional constraints on the measure μ, the minimization of $W_{c_0}(\mu_0, \cdot) + W_{c_1}(\cdot, \mu_1)$, or its generalizations to more than two prescribed measures, arise in different applied settings such as multi-population matching [5] or Wasserstein barycenters [1]. In particular, in the quadratic case where $c_0(x, y) = c_1(x, y) = |x - y|^2$, minimizers of $W_{c_0}(\mu_0, \cdot) + W_{c_1}(\cdot, \mu_1)$ are the midpoints of McCann's displacement interpolation [14] between μ_0 and μ_1, that is, geodesics for the quadratic Wasserstein metric[1]. Density constraints are important to model congestion effects as in the seminal crowd motion model of Maury, Roudneff-Chupin, and Santambrogio [13]. A first goal of the present paper is to investigate the effect of location and density constraints on such Monge–Kantorovich interpolation problems. Let us also mention that the minimization of $W_{c_0}(\mu_0, \mu)$ with respect to μ in a class of measures which are singular with respect to μ_0 was addressed in [3], whereas the parallel case where the density constraint appears in the definition of congestion penalization for singular measures was studied in [11, 22].

A second goal of the paper is to investigate a more general class of problems as a mathematical model for the optimal location of a parking region around a city. In this context, one is given two probability measures ν_0 and ν_1, which may be interpreted as a distribution of residents and a distribution of services, respectively. A resident living at x_0 reaching a service located at x_1 may either walk directly to x_1 for the cost $c_1(x_0, x_1)$ or drive to an intermediate parking location x and then walk from x to x_1 paying the sum $c_0(x_0, x) + c_1(x, x_1)$. In this model, detailed in Section 6, the pivot/parking measure μ may have total mass less than 1, and one may decompose ν_0 and ν_1 as $\nu_i = \nu_i - \mu_i + \mu_i$ with $0 \le \mu_i \le \nu_i$ denoting the *driving* part of ν_i; the unknowns μ_0, μ and μ_1 (with same total mass) should minimize the overall cost $W_{c_1}(\nu_0 - \mu_0, \nu_1 - \mu_1) + W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1)$ subject to possible additional location and density constraints on μ. Let us remark that if (μ_0, μ_1, μ) solves this parking problem, then μ solves the corresponding Monge–Kantorovich interpolation problem, that is, minimizes $W_{c_0}(\mu_0, \cdot) + W_{c_1}(\cdot, \mu_1)$ so that the qualitative properties established in Sections 4 and 5 will be directly applicable

[1]As kindly pointed out to us by a referee, naming optimal transport distances after Wasserstein is controversial and historically incorrect, even though the use of the name is widely spread in the literature; we preferred to mostly use Monge–Kantorovich instead in the present paper.

to optimal parking measures. We have chosen, as an application of our results, a model for determining the optimal location of parking areas around a city, but other models in different fields use similar frameworks and can be found in the literature: we quote for instance [12, 18], where the transport between singular measures is used to model the behavior of biological membranes.

In Section 2, we consider the general optimization problem in (2.1) and after solving an explicit example, we prove existence and discuss uniqueness of solutions. Dual formulations are introduced in Section 3. In Section 4, the particular case of distance-like costs is studied, while Section 5 deals with the case of strictly convex cost functions; in these sections we study various qualitative properties of the solutions, in particular their integrability. In Section 6, we study a problem related to the optimization of a parking area. Finally, in Section 7, we present some numerical simulations thanks to an entropic approximation scheme and compare the solutions of interpolation and parking problems.

2. Monge–Kantorovich interpolation with constraints

Let $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$ be two probabilities with compact support, and let $c_0, c_1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ be two continuous cost functions. For a class $\mathcal{A} \subset \mathcal{P}(\mathbb{R}^d)$, we are interested in solving the optimization problem

$$\inf\left\{W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1) : \mu \in \mathcal{A}\right\}. \tag{2.1}$$

Here $W_{c_i}(\rho_0, \rho_1)$ denotes the value of the optimal transport problem between two measures $\rho_0, \rho_1 \in \mathcal{P}(\mathbb{R}^d)$, obtained by means of the Monge–Kantorovich functionals defined in (1.1). In order to simplify the presentation, by an abuse of notation, if ρ is a measure and ϕ is a non-negative Lebesgue integrable function, by $\rho \le \phi$ we mean that ρ is absolutely continuous and its density, again denoted ρ, satisfies $\rho \le \phi$ Lebesgue almost everywhere; also, all the integrals with no explicitly defined domain of integration are intended on the whole of $\mathbb{R}^d$.

Typical cases for the class $\mathcal{A}$ of admissible choices are:

(i) no constraint, that is, $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$;

(ii) location constraints, that is, $\mathcal{A} = \mathcal{P}(K)$ for a non-empty compact subset K of $\mathbb{R}^d$;

(iii) density constraints, that is, $\mathcal{A} = \{\rho \in \mathcal{P}_{\mathrm{ac}}(\mathbb{R}^d) : \rho \le \phi\}$ for an L^1-function $\phi : \mathbb{R}^d \to \mathbb{R}_+$ with compact support and $\int \phi\, dx > 1$.

2.1. Explicit one-dimensional examples

Before going to the general case, let us illustrate our problem in a simple one-dimensional case, where optimal solutions can be easily recovered by explicit computations.

Example 2.1. Consider the one-dimensional case and the measures

$$\mu_0(x) = \mathbb{1}_{[0,1]}(x), \quad \mu_1(x) = \mathbb{1}_{[5,6]}(x).$$

We first look at the case where the cost functions are given by distances:

$$c_0(x,y) = (1-t)|x-y|, \quad c_1(x,y) = t|x-y| \quad \text{with } t \in]0,1[.$$

The following results can be easily seen by rephrasing the problem in terms of the distribution functions f, f_0, f_1 of the probabilities μ, μ_0, μ_1 (see, for instance, [21, Chapter 2]):

$$\min\left\{\int_0^6 (1-t)|f_0 - f| + t|f - f_1|\,dx : f \text{ non-decreasing, } f(0) = 0,\ f(6) = 1\right\}$$

with the constraints

(i) no additional constraint;

(ii) $\operatorname{spt} f' \subset [2,4]$;

(iii) $f' \le \theta \mathbb{1}_{[2,4]}$.

Since $f_1 \le f_0$, it is easy to see that in the minimization above, one can always assume that $f_1 \le f \le f_0$ and then remove the absolute values and minimize under the constraint that f is non-decreasing and $f_1 \le f \le f_0$. We then have the following situations:

(i) In the absence of constraints, this becomes the problem of finding the Wasserstein median between μ_0 and μ_1 (see [4] for more on Wasserstein medians). In particular, the optimal solutions μ are characterized as follows:
   - if $t > 1/2$ (respectively $t < 1/2$), the unique solution is given by $\mu = \mu_1$ (respectively $\mu = \mu_0$);
   - if $t = 1/2$, any probability μ whose distribution function f is between the two distribution functions f_0 and f_1 of μ_0 and μ_1, in the sense that

   $$f_1(x) \le f(x) \le f_0(x) \quad \text{for all } x \in \mathbb{R},$$

   is a minimizer.

(ii) In the case of the location constraint $K = [2,4]$, we observe a similar threshold effect:
   - if $t > 1/2$ (respectively $t < 1/2$), the unique solution is given by $\mu = \delta_4$ (respectively $\mu = \delta_2$);
   - if $t = 1/2$, then any probability measure supported on K is a solution.

(iii) In the case of density constraint $\phi(x) := \theta \mathbb{1}_{[2,4]}(x)$ with $\theta > 1/2$ we have:
   - if $t > 1/2$ (respectively $t < 1/2$), the unique solution is given by $\mu = \theta \mathbb{1}_{[4-1/\theta,4]}$ (respectively $\mu = \theta \mathbb{1}_{[2,2+1/\theta]}$);
   - if $t = 1/2$, any probability measure satisfying the constraint is a solution.

The example above relies on the fact that for distance-like costs, optimality somehow forces the triangular inequality to be saturated in dimension 1. We will investigate this phenomenon further in Section 4.

We now consider strictly convex cost functions: as a prototype we take, with the same measures μ_0 and μ_1 above,

$$c_0(x, y) = (1-t)|x-y|^2, \quad c_1(x, y) = t|x-y|^2 \quad \text{with } t \in (0, 1).$$

Also, this case can be rephrased in terms of the so-called *pseudo-inverse* g, g_0, g_1 (see, for instance, [21, Definition 2.1]) of the distribution functions f, f_0, f_1 as:

$$\min\left\{\int_0^1 (1-t)(g-g_0)^2 + t(g_1-g)^2 ds : g \text{ non-decreasing}\right\}$$

with the constraints

(i) no additional constraint;

(ii) $g([0, 1]) \subset [2, 4]$;

(iii) $g' \geq 1/\theta$ and $g([0, 1]) \subset [2, 4]$.

This implies

(i) In the unconstrained case the solution simply corresponds to the Wasserstein-geodesic from μ_0 to μ_1 at time $t \in (0, 1)$ or, equivalently, the weighted barycenter. It is given by

$$\mu_t(x) := \mathbb{1}_{[5t,1+5t]}(x).$$

(ii) Take the constraint $K = [2, 4]$, as above. Here the solution depends on the location of the unconstrained geodesic μ_t. We present a few cases (the other ones are clear by symmetry):

- if $t \leq \frac{1}{5}$, the support of μ_t is contained in $[0, 2]$, hence the optimal solution is δ_2;
- if $\frac{1}{5} < t < \frac{2}{5}$, the optimal solution is $\mu = (2-5t)\delta_2 + \mathbb{1}_{[2,1+5t]}$;
- if $\frac{2}{5} \leq t \leq \frac{3}{5}$, the support of μ_t is contained in $[2, 4]$, hence the solution is simply μ_t.

(iii) Take the function $\phi(x) := \theta \mathbb{1}_{[2,4]}(x)$ with $1 > \theta > \frac{1}{2}$. The solution depends again on the location of the unconstrained geodesic μ_t. We have the following cases (remaining cases are again obtained by symmetry):

- if $t \leq \frac{1}{5}$, the support of μ_t is contained in $[0, 2]$, hence the optimal solution is $\theta \mathbb{1}_{[2,2+1/\theta]}$;
- if $\frac{1}{5} < t < \frac{2}{5}$, the optimal solution is still $\mu = \theta \mathbb{1}_{[2,2+1/\theta]}$;
- if $\frac{2}{5} \leq t \leq \frac{3}{5}$, the support of μ_t is contained in $[2, 4]$, but by the density constraint μ_t is not even feasible this time. So, the solution is of the form $\theta \mathbb{1}_{[a,b]}$ with $2 \leq a < b \leq 4$ and $b - a = 1/\theta$.

2.2. Reformulation, existence, uniqueness

Let us now come back to the constrained Monge–Kantorovich interpolation problem in (2.1). Assuming that the measures μ_0 and μ_1 are compactly supported and the costs c_0 and c_1 are continuous and non-negative, then by the direct method one has

Lemma 2.2. *Assume either case (ii) $\mathcal{A} = \mathcal{P}(K)$ with K non-empty and compact, or case (iii) $\mathcal{A} := \{\rho \in \mathcal{P}(\mathbb{R}^d) : \rho \leq \phi\}$ with $\phi \in L^1$ compactly supported and $\int \phi \, \mathrm{d}x \geq 1$. Then, problem (2.1) admits a solution.*

Proof. In both cases, problem 2.1 consists in minimizing the sum of two Monge–Kantorovich functionals, which is weakly* lower semi-continuous over a fixed weakly* compact set. The conclusion then follows by the direct methods of the calculus of variations. ■

In the unconstrained case where $\mathcal{A} := \mathcal{P}(\mathbb{R}^d)$, one of course needs some coercivity in the problem. We shall therefore assume that there exists a compact subset of $\mathbb{R}^d$, denoted (again) by K, such that for every $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$, one has

$$\operatorname*{argmin}_{x \in \mathbb{R}^d} \big\{c_0(x_0, x) + c_1(x, x_1)\big\} \text{ is non-empty and included in } K. \tag{2.2}$$

We then define, for $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$,

$$c(x_0, x_1) := \inf_{x \in \mathbb{R}^d} \big\{c_0(x_0, x) + c_1(x, x_1)\big\} = \min_{x \in K} \big\{c_0(x_0, x) + c_1(x, x_1)\big\}.$$

In the next proposition, we show that the optimization problem in (2.1), with $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$, is equivalent to the standard transport problem with cost c:

$$\inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x_0, x_1) \, \mathrm{d}\gamma(x_0, x_1), \tag{2.3}$$

which clearly admits a solution, since $c \in C(\mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1))$. We easily deduce the existence of a solution to (2.1) when $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$, as well as the fact that all solutions are supported by K.

We will denote by $\Pi(\mu_0, \mu, \mu_1)$ the set of transport plans in the variables (x_0, x, x_1) with marginals μ_0, μ, μ_1, and we denote by $\pi_{0,\mathrm{piv}}, \pi_{\mathrm{piv},1}, \pi_{0,1}$ the projections on the first and second, second and third, and first and third factors, respectively.

Proposition 2.3. *Assume (2.2). Then, the following statements hold true:*

- *Let $\gamma \in \Pi(\mu_0, \mu_1)$ solve (2.3) and let $T : \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1) \to \mathbb{R}^d$ be measurable and such that*

  $$T(x_0, x_1) \in \operatorname*{argmin}_{x \in K} \{c_0(x_0, x) + c_1(x, x_1)\}, \quad \forall (x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1).$$

  Then, $T_{\#}\gamma$ is a solution of (2.1) with $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$ and the optimal values of (2.1) and (2.3) coincide;

- *Conversely, for any optimal solution μ of* (2.1)*, consider optimal transport plans $\gamma_0 \in \Pi(\mu_0, \mu)$ with respect to the cost c_0 and $\gamma_1 \in \Pi(\mu, \mu_1)$ with respect to the cost c_1. Then, there exists a plan $\widetilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1)$ with $\pi_{0,\mathrm{piv}\#}\widetilde{\gamma} = \gamma_0$ and $\pi_{\mathrm{piv},1\#}\widetilde{\gamma} = \gamma_1$ such that $\pi_{0,1\#}\widetilde{\gamma}$ is optimal for* (2.3) *and $c_0(x_0, x) + c_1(x, x_1) = c(x_0, x_1)$ on $\mathrm{spt}(\widetilde{\gamma})$ so that μ is supported by K.*

The previous equivalence also holds between (2.1) *with $\mathcal{A} = \mathcal{P}(K)$ (with K a given compact subset of $\mathbb{R}^d$) and* (2.3) *with c given by $c(x_0, x_1) = \min_{x \in K}\{c_0(x_0, x) + c_1(x, x_1)\}$.*

Proof. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$, $\gamma_0 \in \Pi(\mu_0, \mu)$, and $\gamma_1 \in \Pi(\mu, \mu_1)$; by the gluing lemma (see [23, Lemma 7.6]), there is a plan $\widetilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1)$ with $\pi_{0,\mathrm{piv}\#}\widetilde{\gamma} = \gamma_0$ and $\pi_{\mathrm{piv},1\#}\widetilde{\gamma} = \gamma_1$. Hence, since γ solves (2.3) and $\pi_{0,1\#}\widetilde{\gamma} \in \Pi(\mu_0, \mu_1)$, we have

$$
\begin{aligned}
\int_{\mathbb{R}^d \times \mathbb{R}^d} c_0 \,\mathrm{d}\gamma_0 + \int_{\mathbb{R}^d \times \mathbb{R}^d} c_1 \,\mathrm{d}\gamma_1 &= \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \big\{c_0(x_0, x) + c_1(x, x_1)\big\} \,\mathrm{d}\widetilde{\gamma}(x_0, x, x_1) \\
&\geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times} c \,\mathrm{d}\pi_{0,1\#}\widetilde{\gamma} \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} c \,\mathrm{d}\gamma \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} \big\{c_0(x_0, T(x_0, x_1)) \\
&\qquad + c_1(T(x_0, x_1), x_1)\big\} \,\mathrm{d}\gamma(x_0, x_1) \\
&\geq W_{c_0}(\mu_0, T_\#\gamma) + W_{c_1}(T_\#\gamma, \mu_1)
\end{aligned}
$$

which, taking the infimum with respect to $\gamma_0 \in \Pi(\mu_0, \mu)$ and $\gamma_1 \in \Pi(\mu, \mu_1)$, enables us to deduce that $T_\#\gamma$ solves (2.1) as well as the equality of the optimal values of (2.1) and (2.3).

Assume now that μ solves (2.1) and consider optimal transport plans $\gamma_0 \in \Pi(\mu_0, \mu)$ with respect to the cost c_0 and $\gamma_1 \in \Pi(\mu, \mu_1)$ with respect to the cost c_1. Using again the gluing lemma, we find $\widetilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1)$ with $\pi_{0,\mathrm{piv}\#}\widetilde{\gamma} = \gamma_0$ and $\pi_{\mathrm{piv},1\#}\widetilde{\gamma} = \gamma_1$, and we then have

$$
\begin{aligned}
\inf(2.3) = \inf(2.1) &= \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \big\{c_0(x_0, x) + c_1(x, x_1)\big\} \,\mathrm{d}\widetilde{\gamma}(x_0, x, x_1) \\
&\geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} c(x_0, x_1) \,\mathrm{d}\widetilde{\gamma}(x_0, x, x_1) = \int_{\mathbb{R}^d \times \mathbb{R}^d \times} c \,\mathrm{d}\pi_{0,1\#}\widetilde{\gamma}.
\end{aligned}
$$

Therefore, $\pi_{0,1\#}\widetilde{\gamma}$ is optimal for (2.3) and $c_0(x_0, x) + c_1(x, x_1) = c(x_0, x_1)$ on $\mathrm{spt}(\widetilde{\gamma})$. ∎

In other words, coercivity condition (2.2) ensures that we can replace $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$ by $\mathcal{A} = \mathcal{P}(K)$ in (2.1) and therefore always optimize over probabilities over a fixed compact subset of $\mathbb{R}^d$.

Remark 2.4. We now discuss uniqueness. Letting K be a non-empty compact subset of $\mathbb{R}^d$ and $\mathcal{A}$ be a convex subset of $\mathcal{P}(K)$, note first that $\mu \in \mathcal{A} \mapsto W_{c_0}(\mu_0, \mu)$ and $\mu \in \mathcal{A} \mapsto W_{c_1}(\mu, \mu_1)$ are convex (regardless of specific assumptions on the costs and the measures μ_0 and μ_1). If we further assume that μ_0 is absolutely continuous and c_0

is locally Lipschitz and satisfies the twist condition, that is, it is differentiable in the first coordinate and for every $x_0 \in \operatorname{spt}(\mu_0)$,

$$y \mapsto \nabla_{x_0} c_0(x_0, y) \quad \text{is injective,}$$

then we claim that

$$\mu \mapsto W_{c_0}(\mu_0, \mu) \quad \text{is strictly convex.} \tag{2.4}$$

This implies, in particular, the strict convexity of functional to be minimized in (2.1), and thus the uniqueness of a minimizer. The proof of strict convexity of (2.4) follows the same lines as [21, Proposition 7.19]; we recall the argument for the sake of completeness. Indeed, thanks to the twist condition and the regularity assumptions on c_0 and μ_0, the optimal transport problem between μ_0 and any $\mu \in \mathcal{A}$ has a unique transport plan induced by a map; see [21, Proposition 1.15] and the discussion following it. Assume that $(\mu, \widetilde{\mu}, t) \in \mathcal{A} \times \mathcal{A} \times (0, 1)$ is such that

$$W_{c_0}(\mu_0, (1-t)\mu + t\widetilde{\mu}) = (1-t)W_{c_0}(\mu_0, \mu) + tW_{c_0}(\mu_0, \widetilde{\mu}).$$

Denoting by T and $\widetilde{T}$ the optimal transport maps between μ_0 and μ and μ_0 and $\widetilde{\mu}$ respectively, we have

$$W_{c_0}(\mu_0, (1-t)\mu + t\widetilde{\mu}) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c_0 \, \mathrm{d}\gamma_t \text{ with } \gamma_t := (\mathrm{id}, (1-t)T + t\widetilde{T})_\# \mu_0$$

and since $\gamma_t \in \Pi(\mu_0, (1-t)\mu + t\widetilde{\mu})$, we deduce that γ_t is an optimal plan between μ_0 and $(1-t)\mu + t\widetilde{\mu}$, which, by the twist condition and the absolute continuity of μ_0, implies that γ_t is induced by a map so that $T = \widetilde{T}$ μ_0-almost everywhere; hence, $\mu = T_\#\mu_0 = \widetilde{T}_\#\mu_0 = \widetilde{\mu}$. This shows the announced strict convexity claim. In particular, this argument gives uniqueness for smooth and strictly convex costs. Note that this also gives uniqueness for cases (ii) and (iii) in the case of concave costs, that is, when $c_0(x, y) = l(|x - y|)$ for $l : \mathbb{R}_+ \to \mathbb{R}_+$ strictly concave, increasing, and differentiable on $(0, +\infty)$, if we assume μ_0 absolutely continuous and for case (ii) $K \cap \operatorname{spt} \mu_0 = \emptyset$, or for case (iii) $\operatorname{spt}(\phi) \cap \operatorname{spt} \mu_0 = \emptyset$ (see [10, 17] for refinements and weaker conditions). All these arguments for uniqueness of course remain true if we replace the assumptions on μ_0 and c_0 by similar assumptions on μ_1 and c_1.

3. Dual formulations

3.1. Location constraints

Thanks to coercivity condition (2.2), any solution μ to (2.1) with $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$ is necessarily concentrated on the compact set K, hence both cases (i) and (ii) can be formulated over $\mathcal{P}(K)$. In this case, it can be convenient to characterize solutions of convex minimization problem (2.1) by duality as follows: given $\varphi \in C(K)$, define the c_0-transform of φ, $\varphi^{c_0} \in C(\mathrm{spt}(\mu_0))$ by

$$\varphi_0^{c_0}(x_0) := \min_{x \in K}\{c_0(x_0, x) - \varphi(x)\}, \quad \forall x_0 \in \mathrm{spt}(\mu_0), \tag{3.1}$$

and similarly define the c_1-transform of φ, $\varphi^{c_1} \in C(\mathrm{spt}(\mu_1))$ by

$$\varphi_1^{c_1}(x_1) := \min_{x \in K}\{c_1(x, x_1) - \varphi(x)\}, \quad \forall x_1 \in \mathrm{spt}(\mu_1). \tag{3.2}$$

It follows from [5, Theorem 3] (where the more general multi-marginal case is considered) that the minimum in (2.1) coincides with the value of the dual:

$$\sup\Big\{\int \varphi_0^{c_0}\, \mathrm{d}\mu_0 + \int \varphi_1^{c_1}\, \mathrm{d}\mu_1 : \varphi_0, \varphi_1 \in C(K),\ \varphi_0 + \varphi_1 = 0\Big\}, \tag{3.3}$$

and the supremum in (3.3) is attained. Moreover, if φ_0 and φ_1 solve (3.3), then $\mu \in \mathcal{P}(K)$ solves (2.1) if and only if φ_0 is a Kantorovich potential between μ_0 and μ and φ_1 is a Kantorovich potential (see [21, 23] for more on Kantorovich duality) between μ and μ_1, that is, there exists $(\gamma_0, \gamma_1) \in \Pi(\mu_0, \mu) \times \Pi(\mu, \mu_1)$ such that

$$\begin{aligned} \varphi_0(x) + \varphi_0^{c_0}(x_0) &= c_0(x_0, x), \quad \forall (x_0, x) \in \mathrm{spt}(\gamma_0), \\ \varphi_1(x) + \varphi_1^{c_1}(x_1) &= c_1(x, x_1), \quad \forall (x, x_1) \in \mathrm{spt}(\gamma_1). \end{aligned}$$

Defining the c_0-concave envelope of φ_0 and the c_1-concave envelope of φ_1 by

$$\begin{aligned} \widetilde{\varphi}_0(x) &:= \min_{x_0 \in \mathrm{spt}(\mu_0)} \{c_0(x_0, x) - \varphi_0^{c_0}(x_0)\}, \\ \widetilde{\varphi}_1(x) &:= \min_{x_1 \in \mathrm{spt}(\mu_1)} \{c_1(x, x_1) - \varphi_1^{c_1}(x_1)\}, \end{aligned}$$

one has $\widetilde{\varphi}_0 \geq \varphi_0$ and $\widetilde{\varphi}_1 \geq \varphi_1$ with an equality on $\mathrm{spt}(\mu)$ so that $\widetilde{\varphi}_0 + \widetilde{\varphi}_1 \geq 0$ with an equality on $\mathrm{spt}(\mu)$.

3.2. Density constraint

We now consider case (iii) where there is a constraint on the density $\mu \leq \phi$. One can characterize minimizers by duality as follows:

Proposition 3.1. *Consider (2.1) in the case (iii) where there is a constraint on the density $\mu \leq \phi$ with $\phi \in L^1(\mathbb{R}^d)$, $\phi \geq 0$, $\int \phi\, dx > 1$, and $\mathrm{spt}(\phi)$ compact (as well as $\mathrm{spt}(\mu_0)$ and $\mathrm{spt}(\mu_1)$). Then, the value of (2.1) coincides with the value of its (pre-)dual formulation*

$$\sup_{\varphi_0, \varphi_1 \in C(\mathrm{spt}(\phi))^2} \int \varphi_0^{c_0}\, \mathrm{d}\mu_0 + \int \varphi_1^{c_1}\, \mathrm{d}\mu_1 + \int \min(\varphi_0 + \varphi_1, 0)\phi\ \mathrm{d}x \tag{3.4}$$

(where $\varphi_i^{c_i}$ are as in formulae (3.1)–(3.2) with K replaced by $\mathrm{spt}(\phi)$). Moreover, the supremum in (3.4) is attained. If (φ_0, φ_1) solves (3.4), then μ solves (2.1) under the constraint $\mu \leq \phi$ if and only if there exist $\gamma_0 \in \Pi(\mu, \mu_0)$ and $\gamma_1 \in \Pi(\mu, \mu_1)$ such that

$$\varphi_0(x) + \varphi_0^{c_0}(x_0) = c_0(x_0, x), \quad \forall (x_0, x) \in \mathrm{spt}(\gamma_0), \tag{3.5}$$

$$\varphi_1(x) + \varphi_1^{c_1}(x_1) = c_1(x, x_1), \quad \forall (x, x_1) \in \mathrm{spt}(\gamma_1) \tag{3.6}$$

(so that γ_0 and γ_1 are optimal plans and φ_0 and φ_1 are Kantorovich potentials) and

$$\varphi_0 + \varphi_1 \geq 0 \quad \textit{on } \operatorname{spt}(\phi - \mu), \qquad \varphi_0 + \varphi_1 \leq 0 \quad \textit{on } \operatorname{spt}(\mu). \tag{3.7}$$

Proof. The fact that the concave maximization problem given by (3.4) is the dual of (2.1) under the constraint $\mu \leq \phi$ follows from the Fenchel–Rockafellar duality theorem and the Kantorovich duality formula. Indeed, we first have

$$\sup (3.4) = - \inf_{\varphi_0, \varphi_1 \in C(\operatorname{spt}(\phi))^2} F(\varphi_0, \varphi_1) + G(-\varphi_0 - \varphi_1)$$

where

$$F(\varphi_0, \varphi_1) := -\int \varphi_0^{c_0} \, \mathrm{d}\mu_0 - \int \varphi_1^{c_1} \, \mathrm{d}\mu_1, \;\; G(\varphi) := \int \max(\varphi, 0)\phi \;\, \mathrm{d}x, \;\; \varphi \in C(\operatorname{spt}(\phi)).$$

Note that F and G are convex and continuous for the uniform convergence topology and it is easy to see that sup (3.4) is finite (see the proof of existence of a solution to (3.4) below) so that by the Fenchel–Rockafellar duality theorem, we have

$$\sup (3.4) = - \sup_{\mu \in C(\operatorname{spt}(\phi))^*} \big\{ -F^*(\mu, \mu) - G^*(\mu) \big\} = \inf_{\mu \in C(\operatorname{spt}(\phi))^*} \big\{ F^*(\mu, \mu) + G^*(\mu) \big\}.$$

By the Kantorovich duality formula ([21, Proposition 1.11]), we have (also see [5] for details)

$$F^*(\mu, \mu) = \begin{cases} W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1) & \text{if } \mu \in \mathcal{P}(\operatorname{spt}(\phi)), \\ +\infty & \text{otherwise} \end{cases}$$

and

$$G^*(\mu) = \sup_{\varphi \in C(\operatorname{spt}(\phi))} \int \varphi \, \mathrm{d}\mu - \int \max(\varphi, 0)\phi \;\, \mathrm{d}x,$$

and when $\mu \in \mathcal{P}(\operatorname{spt}(\phi))$ (which is the case when $F^*(\mu, \mu) < +\infty$), the above maximization can be restricted to non-negative functions φ, yielding

$$G^*(\mu) = \begin{cases} 0 & \text{if } \mu \leq \phi \\ +\infty & \text{otherwise.} \end{cases}$$

We thus have

$$\sup (3.4) = \inf_{\mu \in \mathcal{P}(\operatorname{spt}(\phi)), \; \mu \leq \phi} W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1)$$

and (up to extending μ by 0 outside $\operatorname{spt}(\phi)$) the right-hand side of the previous equality is equivalent to (2.1) in case (iii) where there is a constraint on the density $\mu \leq \phi$. Let us now prove that (3.4) admits a solution. To see this, we remark that the objective is unchanged when one replaces (φ_0, φ_1) by $(\varphi_0 + \lambda, \varphi_1 - \lambda)$, where λ is a constant. Moreover, replacing φ_0 and φ_1 by their c_0/c_1-concave envelopes defined for every $x \in \operatorname{spt}(\phi)$ by

$$\begin{aligned} \widetilde{\varphi}_0(x) &:= \min_{x_0 \in \operatorname{spt}(\mu_0)} \big\{ c_0(x_0, x) - \varphi_0^{c_0}(x_0) \big\}, \\ \widetilde{\varphi}_1(x) &:= \min_{x_1 \in \operatorname{spt}(\mu_1)} \big\{ c_1(x, x_1) - \varphi_1^{c_1}(x_1) \big\}, \end{aligned} \tag{3.8}$$

it is well known that $\widetilde{\varphi}_i \geq \varphi_i$ and $\widetilde{\varphi}_i^{c_i} = \varphi_i^{c_i}$ for $i = 0, 1$, so that replacing φ_i by $\widetilde{\varphi}_i$ is an improvement in the objective of (3.4); moreover, the functions $\widetilde{\varphi}_i$ have a uniform modulus of continuity inherited from the uniform continuity of c_i. From these observations, we can find a uniformly equicontinuous maximizing sequence $(\varphi_0^n, \varphi_1^n)_n$ for which $\min_{\mathrm{spt}(\phi)} \varphi_0^n = 0$ so that φ_0^n is also uniformly bounded. Since $\min(\varphi_1^n + \varphi_0^n, 0) \leq 0$, the fact that $(\varphi_0^n, \varphi_1^n)_n$ is a maximizing sequence together with the uniform bounds on φ_0^n gives a uniform lower bound on $\int (\varphi_1^n)^{c_1} d\mu_1$ from which we easily derive a uniform upper bound on φ_1^n, thanks to (3.8). To show that φ_1^n is also uniformly bounded from below, we observe that the quantity

$$\int (\varphi_1^n)^{c_1} d\mu_1 + \int \min(\varphi_1^n + \varphi_0^n, 0)\phi \, dx$$

is bounded from below and bounded from above by $C + (\int \phi \, dx - 1) \min_{\mathrm{spt}(\phi)} \varphi_1^n$ for some constant C. Since $\int \phi \, dx > 1$, this gives the desired lower bound. Having thus found a uniformly bounded and equicontinuous maximizing sequence, we deduce the existence of a solution to (3.4) from the Arzelà–Ascoli theorem.

Let us now look at the optimality conditions which follow from the above duality. If (φ_0, φ_1) solves (3.4), then μ solves (2.1) under the constraint $\mu \leq \phi$ if and only if

$$W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1) = \int \varphi_0^{c_0} \, \mathrm{d}\mu_0 + \int \varphi_1^{c_1} \, \mathrm{d}\mu_1 + \int \min(\varphi_0 + \varphi_1, 0)\phi.$$

If γ_0 (resp. γ_1) is an optimal plan for c_0 (resp. c_1) between μ_0 and μ (resp. μ and μ_1), we thus have

$$\begin{aligned}
\int \varphi_0^{c_0} \, \mathrm{d}\mu_0 + \int \varphi_1^{c_1} \, \mathrm{d}\mu_1 + \int \min(\varphi_0 + \varphi_1, 0)\phi &= \int c_0 \, \mathrm{d}\gamma_0 + \int c_1 \, \mathrm{d}\gamma_1 \\
&\geq \int (\varphi_0^{c_0}(x_0) + \varphi_0(x)) \, \mathrm{d}\gamma_0(x_0, x) + \int (\varphi_1^{c_1}(x_1) + \varphi_1(x)) \, \mathrm{d}\gamma_1(x, x_1) \\
&= \int \varphi_0^{c_0} \, \mathrm{d}\mu_0 + \int \varphi_1^{c_1} \, \mathrm{d}\mu_0 + \int (\varphi_0 + \varphi_1) \, \mathrm{d}\mu \\
&\geq \int \varphi_0^{c_0} \, \mathrm{d}\mu_0 + \int \varphi_1^{c_1} \, \mathrm{d}\mu_0 + \int \min(\varphi_0 + \varphi_1, 0) \, \mathrm{d}\mu \\
&\geq \int \varphi_0^{c_0} \, \mathrm{d}\mu_0 + \int \varphi_1^{c_1} \, \mathrm{d}\mu_0 + \int \min(\varphi_0 + \varphi_1, 0)\phi \, \mathrm{d}x,
\end{aligned}$$

where we have used that $\mu \leq \phi$ in the last line. All the inequalities above should therefore be equalities which, together with the continuity of φ_0 and φ_1, is easily seen to imply (3.6)–(3.5)–(3.7). This shows the necessity of these conditions; the proof of sufficiency by duality is direct and therefore left to the reader. ■

Corollary 3.2. *Under the same assumptions as in Proposition 3.1, assume that μ is optimal for (2.1) under the constraint $\mu \leq \phi$ and let γ_0 and γ_1 be optimal transport plans. Then, whenever x_0, x, x_1 are such that $(x_0, x) \in \mathrm{spt}(\gamma_0)$, $(x, x_1) \in \mathrm{spt}(\gamma_1)$, and $x \in \mathrm{spt}(\phi - \mu)$, we have*

$$c_0(x_0, x) + c_1(x, x_1) = \min_{y \in \mathrm{spt}(\phi - \mu)} \big\{ c_0(x_0, y) + c_1(y, x_1) \big\}.$$

Proof. Let (φ_0, φ_1) solve (3.4). By construction, for every $(x_0, x_1, y) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1) \times \mathrm{spt}(\phi)$, one has

$$c_0(x_0, y) + c_1(y, x_1) \geq \varphi_0^{c_0}(x_0) + \varphi_1^{c_1}(x_1) + (\varphi_0 + \varphi_1)(y).$$

Together with (3.7), this implies that for every $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$,

$$\min_{y \in \mathrm{spt}(\phi-\mu)} \big\{c_0(x_0, y) + c_1(y, x_1)\big\} \geq \varphi_0^{c_0}(x_0) + \varphi_1^{c_1}(x_1).$$

But now if $x \in \mathrm{spt}(\mu) \cap \mathrm{spt}(\phi - \mu)$, by (3.7) again we have $\varphi_0(x) + \varphi_1(x) = 0$. Hence, by (3.6)–(3.5) whenever $(x_0, x) \in \mathrm{spt}(\gamma_0)$, $(x, x_1) \in \mathrm{spt}(\gamma_1)$, and $x \in \mathrm{spt}(\phi - \mu)$, we have

$$\varphi_0^{c_0}(x_0) + \varphi_1^{c_1}(x_1) = c_0(x_0, x) + c_1(x, x_1) \geq \min_{y \in \mathrm{spt}(\phi-\mu)} \big\{c_0(x_0, y) + c_1(y, x_1)\big\},$$

which yields the desired result. ∎

In the discrete case, we can easily deduce a bang-bang result stating that the constraint $\mu \leq \phi$ is always binding when $\mu > 0$ under mild conditions on the cost. We will give similar bang-bang results for distance-like costs in Section 4.

Corollary 3.3. *Assume that μ_0 and μ_1 are discrete and that for every $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$, $c_0(x_0, \cdot)$, and $c_1(\cdot, x_1)$ are C^1 and M-Lipschitz on* $\mathrm{spt}(\phi)$ *(for some M that does not depend on x_0 and x_1); also assume that the set*

$$\big\{x \in \mathrm{spt}(\phi) : \nabla_x c_0(x_0, x) + \nabla_x c_1(x, x_1) = 0\big\} \tag{3.9}$$

is Lebesgue negligible. Then, if μ is optimal for (2.1) *under the constraint $\mu \leq \phi$, there exists a measurable subset E of* $\mathrm{spt}(\phi)$ *such that $\mu = \phi \mathbb{1}_E$.*

Proof. Let (φ_0, φ_1) solve (3.4). As seen in the proof of Proposition 3.1, we may assume that for every $x \in \mathrm{spt}(\phi)$,

$$\begin{aligned}\varphi_0(x) &:= \min_{x_0 \in \mathrm{spt}(\mu_0)} \big\{c_0(x_0, x) - \varphi_0^{c_0}(x_0)\big\},\\ \varphi_1(x) &:= \min_{x_1 \in \mathrm{spt}(\mu_1)} \big\{c_1(x, x_1) - \varphi_1^{c_1}(x_1)\big\},\end{aligned}$$

so that φ_0 and φ_1 are Lipschitz and, hence, differentiable almost everywhere on $\mathrm{spt}(\phi)$. Since $\varphi_0 + \varphi_1 = 0$ on $\mathrm{spt}(\mu) \cap \mathrm{spt}(\phi - \mu)$, we then have

$$\nabla\varphi_0 + \nabla\varphi_1 = 0 \quad \text{a.e. on } \big\{0 < \mu < \phi\big\},$$

but if φ_0 (resp. φ_1) is differentiable at x and $(x_0, x) \in \mathrm{spt}(\gamma_0)$ (resp. $(x, x_1) \in \mathrm{spt}(\gamma_1)$), where γ_0 and γ_1 are optimal plans, then

$$\nabla\varphi_0(x) = \nabla_x c_0(x_0, x), \quad \nabla\varphi_1(x) = \nabla_x c_1(x, x_1).$$

Hence, denoting by A_i the countable concentration set of μ_i ($i = 0, 1$), almost every x such that $0 < \mu(x) < \phi(x)$ belongs to

$$\bigcup_{(x_0,x_1)\in A_0\times A_1} \big\{x \in \operatorname{spt}(\phi) : \nabla_x c_0(x_0, x) + \nabla_x c_1(x, x_1) = 0\big\},$$

which is negligible by assumption. The desired bang-bang conclusion then readily follows. ■

Remark 3.4. In some cases, for instance, when the costs c_0 and c_1 depend quadratically on or more generally on the p-th power of the distance (with $p > 1$), the set in (3.9) reduces to a single point which depends in a Lipschitz way on x_0 and x_1. The conclusion of Corollary 3.3 then still holds under the weaker assumption that one between μ_0 and μ_1 is discrete and the other one is singular with respect to the Lebesgue measure. More precisely, this still holds if the Hausdorff dimension of the support of μ_0 is h_0, and the Hausdorff dimension of the support of μ_1 is h_1, with $h_0 + h_1 < d$.

4. Distance-like costs

In this section, we pay special attention to the case of distance-like costs

$$c_0(x_0, x) := |x_0 - x|^\alpha, \quad c_1(x, x_1) := \lambda|x - x_1|^\alpha, \tag{4.1}$$

with $0 < \alpha \le 1$ and $\lambda > 0$.

4.1. Location constraint, concentration, and integrability on the boundary

Let us start with the case of a location constraint of type (ii): $\mu \in \mathscr{P}(K)$ for some nonempty compact subset K of $\mathbb{R}^d$.

Lemma 4.1. *Assume K is a compact subset of $\mathbb{R}^d$ and that one of the following assumptions holds:*

- $\alpha = 1$, $\lambda > 1$, *and the interior of K is disjoint from* $\operatorname{spt}(\mu_1)$,
- $\alpha \in (0, 1)$ *and the interior of K is disjoint from* $\operatorname{spt}(\mu_0) \cup \operatorname{spt}(\mu_1)$.

Then, any solution μ of (2.1) under the constraint $\mu \in \mathscr{P}(K)$ is supported by ∂K.

Proof. For $(x_0, x_1) \in \operatorname{spt}(\mu_0) \times \operatorname{spt}(\mu_1)$, set

$$c(x_0, x_1) := \min_{x\in K}\big\{|x_0 - x|^\alpha + \lambda|x - x_1|^\alpha\big\},$$
$$T(x_0, x_1) := \operatorname*{argmin}_{x\in K}\big\{|x_0 - x|^\alpha + \lambda|x - x_1|^\alpha\big\}.$$

We know from Proposition 2.3 that μ is supported by $T(\operatorname{spt}(\mu_0) \times \operatorname{spt}(\mu_1))$. In particular, if $x \in \operatorname{spt}(\mu)$ is an interior point of K, then it is a local minimizer of $c_0(x_0, \cdot) + c_1(\cdot, x_1)$

for some $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$. In the case $\alpha = 1$, $\lambda > 1$, since $x \neq x_1$, this is clearly impossible. In the case $\alpha < 1$, our assumption implies that $x \notin \{x_0, x_1\}$, so that x has to be a critical point of $c_0(x_0, \cdot) + c_1(\cdot, x_1)$. One should have

$$\alpha|x - x_0|^{\alpha-2}(x - x_0) + \lambda\alpha|x - x_1|^{\alpha-2}(x - x_1) = 0,$$

so that $x_0 \neq x_1$ and $x \in [x_0, x_1]$. But $c_0(x_0, \cdot) + c_1(\cdot, x_1)$ is strictly concave on $[x_0, x_1]$, which contradicts x being a local minimizer. ■

Remark 4.2. If $\alpha = \lambda = 1$, the previous result is false: if $d = 1$, $\mu_0 = \delta_0$, $\mu_1 = \delta_1$, and $K = [1/4, 3/4]$, then it follows from the triangle inequality that any probability supported by K is actually optimal.

Now that we know that minimizers are supported by ∂K, one may wonder, if K and μ_1 are regular enough, whether these minimizers are absolutely continuous with respect to the $(d-1)$-Hausdorff measure on ∂K; the answer is positive if μ_0 is discrete, that is, concentrated on a countable set, $\mu_0(K) = 0$, and μ_1 is absolutely continuous with support disjoint from $\mathrm{int}(K)$ (see Proposition 4.4 below). A first step consists in the following result:

Lemma 4.3. *Assume that c_0 and c_1 are as in (4.1) (with $\alpha \in (0, 1]$ and $\lambda > 1$ if $\alpha = 1$), and that K is compact. Then, for every x_0 and (Lebesgue-)almost every $x_1 \in \mathbb{R}^d \setminus K$, the set*

$$T_{x_0}(x_1) := \operatorname*{argmin}_{x\in K} \left\{|x_0 - x|^\alpha + \lambda|x - x_1|^\alpha\right\}$$

is a singleton.

Proof. Fix x_0, set

$$c_{x_0}(x_1) := \min_{x\in K} \left\{|x_0 - x|^\alpha + \lambda|x - x_1|^\alpha\right\},$$

and observe that c_{x_0} is locally Lipschitz on $\mathbb{R}^d \setminus K$. It thus follows from Rademacher's theorem that almost every $x_1 \in \mathbb{R}^d \setminus K$ is a point of differentiability of c_{x_0}, and for such a point, if $x \in T_{x_0}(x_1)$, we have

$$\nabla c_{x_0}(x_1) = \lambda\alpha|x_1 - x|^{\alpha-2}(x_1 - x) \neq 0.$$

If $\alpha \in (0, 1)$, this immediately gives the claim with

$$T_{x_0}(x_1) = \left\{x_1 + (\lambda\alpha)^{\frac{1}{1-\alpha}}|\nabla c_{x_0}(x_1)|^{\frac{2-\alpha}{\alpha-1}}\nabla c_{x_0}(x_1)\right\}.$$

When $\alpha = 1$ and $\lambda > 1$, if both x and x' belong to $T_{x_0}(x_1)$, then x_1, x, and x' are aligned so that the triangle inequality between their differences is saturated. But if $x \in [x_1, x')$, by the definition of $T_{x_0}(x_1)$ and $\lambda > 1$, we should also have

$$c_{x_0}(x_1) = |x - x_0| + \lambda|x - x_1| = |x' - x_0| + \lambda|x' - x_1|$$

$$= |x' - x_0| + \lambda(|x - x_1| + |x' - x|)$$
$$> |x' - x_0| + |x' - x| + \lambda|x - x_1|,$$

which is impossible by the triangle inequality, thus yielding the almost everywhere single-valuedness of T_{x_0} in this case as well. ■

Proposition 4.4. *Assume that either* $\alpha = 1$, $\lambda > 1$ *or* $\alpha \in (0, 1)$ *and*

- K *is the closure of an open, bounded set in* $\mathbb{R}^d$ *with a boundary of class* $C^{1,1}$,
- μ_0 *is discrete and* $\mu_0(K) = 0$,
- μ_1 *is absolutely continuous and* $\operatorname{int}(K) \cap \operatorname{spt} \mu_1 = \emptyset$.

Then, any solution μ *of* (2.1) *under the constraint* $\mu \in \mathcal{P}(K)$ *is absolutely continuous with respect to the* $(d-1)$*-Hausdorff measure on* ∂K.

Proof. Since μ_0 is discrete, we can write $\mu_0 = \sum_{x_0 \in A_0} p_{x_0}\delta_{x_0}$, with A_0 at most countable, disjoint from K and $p_{x_0} > 0$ for every $x_0 \in A_0$. It follows from Proposition 2.3 and Lemma 4.3 that there exists a transport plan γ between μ_0 and μ_1 which can be written as

$$\gamma = \sum_{x_0 \in A_0} p_{x_0}\delta_{x_0} \otimes \mu_1^{x_0},$$

such that, defining T_{x_0} as in Lemma 4.3 and $T(x_0, x_1) = T_{x_0}(x_1)$, one has

$$\mu = T_\#\gamma = \sum_{x_0 \in A_0} p_{x_0} T_{x_0\#}\mu_1^{x_0}.$$

Since the second marginal of γ is μ_1, we also have

$$\mu_1 = \sum_{x_0 \in A_0} p_{x_0}\mu_1^{x_0},$$

so that all the measures $\mu_1^{x_0}$ are dominated by $1/p_{x_0}\mu_1$ and, hence, absolutely continuous. We are thus left to show that for each fixed x_0 in the countable set A_0, the measure $T_{x_0\#}\mu_1^{x_0}$ (which is supported by ∂K by Lemma 4.1) is absolutely continuous with respect to the $(d-1)$-Hausdorff measure on ∂K which from now on we denote by $\sigma_{(d-1),\partial K}$. We now fix $x_0 \in A_0$ and a Borel subset A of ∂K. Our aim is to bound

$$(T_{x_0\#}\mu_1^{x_0})(A) = \mu_1^{x_0}(T_{x_0}^{-1}(A)).$$

To this end, let us distinguish the two cases $\alpha = 1$, $\lambda > 1$ and $\alpha \in (0, 1)$.

Assume $\alpha = 1$ and $\lambda > 1$. Since $\mu_1(K) = 0$ (because μ_1 is absolutely continuous, ∂K is a smooth hypersurface and thus Lebesgue negligible and $\mu_1(\operatorname{int}(K)) = 0$), we have

$$\mu_1^{x_0}(T_{x_0}^{-1}(A) \setminus K) = \mu_1^{x_0}(T_{x_0}^{-1}(A)).$$

Now take $x = T_{x_0}(x_1) \in \partial K$ with $x_1 \notin K$ which is μ_1-almost everywhere the case (so that $x \notin \{x_0, x_1\}$). By optimality, there exists $\beta \geq 0$ such that

$$\widehat{x - x_0} + \lambda \widehat{x - x_1} + \beta n(x) = 0,$$

where for $\xi \in \mathbb{R}^d \setminus \{0\}$, we have set $\widehat{\xi} = \xi/|\xi|$, and where $n(x)$ is the outward normal to ∂K at x. Using the fact that $\lambda \widehat{x - x_1}$ has norm λ yields

$$\lambda^2 = \beta^2 + 1 + 2\beta n(x) \cdot \widehat{x - x_0},$$

whose only non-negative root is

$$\beta = \beta_{x_0}(x) := -n(x) \cdot \widehat{x - x_0} + \sqrt{\lambda^2 - 1 + (n(x) \cdot \widehat{x - x_0})^2},$$

so that

$$\lambda \widehat{x_1 - x} = \beta_{x_0}(x) n(x) + \widehat{x - x_0}$$

and the right-hand side is a Lipschitz function of x, thanks to our assumptions (∂K being $C^{1,1}$ and x_0 being at a positive distance from K, and hence from x). Using again that $\lambda \widehat{x - x_1}$ has norm λ, this shows that if $x = T_{x_0}(x_1)$, then for some $r \in [0, R]$ with $R = \lambda^{-1} \operatorname{diam}(\operatorname{spt} \mu_1 - K)$, we have

$$x_1 = F_{x_0}(r, x) := x + r[\beta_{x_0}(x) n(x) + \widehat{x - x_0}].$$

Hence,

$$\mu_1^{x_0}(T_{x_0}^{-1}(A)) \leq \mu_1^{x_0}(F_{x_0}([0, R] \times A)).$$

If $\sigma_{(d-1),\partial K}(A) = 0$, the smoothness of K and the fact that F_{x_0} is Lipschitz on $[0, R] \times \partial K$ readily imply that $F_{x_0}([0, R] \times A)$ is Lebesgue negligible. Hence, $\mu_1^{x_0}(T_{x_0}^{-1}(A)) = 0$ and since this holds for any $x_0 \in A_0$, we also have $\mu(A) = 0$, which implies the absolute continuity of μ with respect to $\sigma_{(d-1),\partial K}$.

Let us now assume that $\alpha \in (0, 1)$. To cope with the fact that $c_1(x, x_1)$ is not differentiable if $x = x_1$, it will be convenient to fix $\varepsilon > 0$ and consider $x_1 \in A_1^\varepsilon$, where

$$A_1^\varepsilon := \{x_1 \in \operatorname{spt}(\mu_1) : \ d(K, x_1) \geq \varepsilon\}$$

and

$$d(K, x) := \min_{y \in K} |x - y|$$

is the Euclidean distance to K. If $x_1 \in A_1^\varepsilon \cap T_{x_0}^{-1}(x)$ with $x \in A$, it follows from the first-order optimality condition that there is some $r \geq 0$ such that

$$x_1 = G_{x_0}(r, x) := x + |H_{x_0}(r, x)|^{\frac{2-\alpha}{\alpha-1}} H_{x_0}(r, x),$$

where

$$H_{x_0}(r, x) = r n(x) + \lambda^{-1} |x - x_0|^{\alpha-2}(x - x_0).$$

Now, note that

$$|H_{x_0}(r,x)| = |x_1 - x|^{\alpha-1}.$$

This shows that

$$\begin{aligned}|r| &\le |x_1 - x|^{\alpha-1} + \lambda^{-1}|x - x_0|^{\alpha-1} \\ &\le \varepsilon^{\alpha-1} + \lambda^{-1}\max_{x\in K}|x - x_0|^{\alpha-1} =: R_\varepsilon(x_0).\end{aligned}$$

Hence, $A_1^\varepsilon \cap T_{x_0}^{-1}(x)$ is included in the image by G_{x_0} of the set $\{(r,x),\ x \in A,\ r \in [0, R_\varepsilon(x_0)]\}$. Since G_{x_0} is Lipschitz (with a Lipschitz constant depending on ε), on this set we obtain as soon as $\sigma_{(d-1),\partial K}(A) = 0$ that

$$\begin{aligned}\mu_1^{x_0}(T_{x_0}^{-1}(A)) = \mu_1^{x_0}(T_{x_0}^{-1}(A)\setminus K) &= \lim_{\varepsilon\searrow 0}\mu_1^{x_0}(T_{x_0}^{-1}(A)\cap A_1^\varepsilon) \\ &\le \lim_{\varepsilon\searrow 0}\mu_1^{x_0}(G_{x_0}([0,R_\varepsilon(x_0)]\times A)) = 0.\end{aligned}$$

Thus, we can conclude as before that μ is absolutely continuous. ∎

Proposition 4.5. *Suppose in addition to the assumptions of Proposition 4.4 that μ_0 has finite support and μ_1 has a bounded density with respect to the d-dimensional Lebesgue measure. If $\alpha \in (0,1)$, further assume that $K \cap \operatorname{spt}\mu_1 = \emptyset$. Then μ has a bounded density with respect to the $(d-1)$-Hausdorff measure on ∂K.*

Proof. In the case $\alpha = 1$, $\lambda > 1$ we can continue using the same notation and Lipschitz mapping F_{x_0} and R as in the proof of Proposition 4.4 to conclude, for any Borel subset A of ∂K,

$$\begin{aligned}\mu(A) &= \sum_{x_0\in\operatorname{spt}(\mu_0)} p_{x_0}\mu_1^{x_0}(T_{x_0}^{-1}(A)) \\ &\le \sum_{x_0\in\operatorname{spt}(\mu_0)} p_{x_0}\mu_1^{x_0}(F_{x_0}([0,R]\times A)) \\ &\le \sum_{x_0\in\operatorname{spt}(\mu_0)} \|\mu_1\|_{L^\infty}\mathcal{L}^d(F_{x_0}([0,R]\times A)) \\ &\le C\operatorname{card}(\operatorname{spt}\mu_0)\|\mu_1\|_{L^\infty} R\sigma_{(d-1),\partial K}(A),\end{aligned}$$

where C is a constant that only depends on the $C^{1,1}$ smoothness of ∂K and the maximal (with respect to $x_0 \in \operatorname{spt}(\mu_0)$) Lipschitz constant of F_{x_0} over $[0,R]\times\partial K$. This way we deduce that $\mu \in L^\infty(\partial K, \sigma_{(d-1),\partial K})$.

For the case $\alpha \in (0,1)$, we need in addition $K \cap \operatorname{spt}\mu_1 = \emptyset$ to ensure that, again using the same arguments and notation as in the proof of Proposition 4.4, there is an $\varepsilon_0 > 0$ such that $A_1^{\varepsilon_0} = \operatorname{spt}(\mu_1)$. In this way, all the analysis from the previous proof can be carried through on $A_1^{\varepsilon_0}$ and we obtain

$$\mu(A) = \sum_{x_0\in\operatorname{spt}(\mu_0)} p_{x_0}\mu_1^{x_0}(T_{x_0}^{-1}(A)\cap A_1^{\varepsilon_0})$$

$$
\begin{aligned}
&\le \sum_{x_0 \in \mathrm{spt}(\mu_0)} p_{x_0} \mu_1^{x_0}(G_{x_0}([0, R_{\varepsilon_0}(x_0)] \times A)) \\
&\le \sum_{x_0 \in \mathrm{spt}(\mu_0)} \|\mu_1\|_{L^\infty} \mathcal{L}^d(G_{x_0}([0, R_{\varepsilon_0}(x_0)] \times A)) \\
&\le C \,\mathrm{card}(\mathrm{spt}\,\mu_0) \|\mu_1\|_{L^\infty} R_{\varepsilon_0}(x_0) \sigma_{(d-1),\partial K}(A),
\end{aligned}
$$

where C is a constant that only depends on the $C^{1,1}$ smoothness of ∂K and the maximal (with respect to $x_0 \in \mathrm{spt}(\mu_0)$) Lipschitz constant of G_{x_0} over $[0, R_{\varepsilon_0}(x_0)] \times \partial K$. ■

One might also be interested in the case that the distribution of residents represented by μ_0 and μ_1 is absolutely continuous and discrete, respectively. The case $\alpha \in (0, 1)$ is completely symmetric, as we have not assumed $\lambda > 1$ in the previous proofs. However, for the case $\alpha = 1, \lambda > 1$, the proof slightly differs, as we shall see below. Arguing as in the proof of Lemma 4.3, we have:

Lemma 4.6. *Assume that c_0 and c_1 are as in (4.1) (with $\alpha \in (0, 1]$ and $\lambda > 1$ if $\alpha = 1$), and that K is compact. Then, for (Lebesgue-)almost every $x_0 \in \mathbb{R}^d \setminus K$ and every x_1, the set*

$$
T_{x_1}(x_0) := \operatorname*{argmin}_{x \in K} \left\{ |x_0 - x|^\alpha + \lambda |x - x_1|^\alpha \right\}
$$

is a singleton.

The analogue of Proposition 4.4 then reads

Proposition 4.7. *Assume that either $\alpha = 1$, $\lambda > 1$ or $\alpha \in (0, 1)$ and*

- *K is the closure of an open, bounded set in $\mathbb{R}^d$ with a boundary of class $C^{1,1}$,*
- *μ_0 is absolutely continuous and $\mathrm{int}(K) \cap \mathrm{spt}\,\mu_0 = \emptyset$,*
- *μ_1 is discrete and $\mu_1(K) = 0$.*

Then, any solution μ of (2.1) under the constraint $\mu \in \mathcal{P}(K)$ is absolutely continuous with respect to the $(d-1)$-Hausdorff measure on ∂K.

Proof. As already explained, the case $\alpha > 1$ can be handled exactly as for Proposition 4.4. We shall therefore assume that $\alpha = 1$ and $\lambda > 1$. We write $\mu_1 = \sum_{x_1 \in A_1} p_{x_1} \delta_{x_1}$, with A_1 countable and $p_{x_1} > 0$. It follows from Proposition 2.3 and Lemma 4.6 that there exists a transport plan γ between μ_0 and μ_1 which can be written as

$$
\gamma = \sum_{x_1 \in A_1} \mu_0^{x_1} \otimes p_{x_1} \delta_{x_1}
$$

and is such that, defining T_{x_1} as in Lemma 4.3 and $T(x_0, x_1) = T_{x_1}(x_0)$, one has

$$
\mu = T_\# \gamma = \sum_{x_1 \in A_1} p_{x_1} T_{x_1 \#} \mu_0^{x_1}.
$$

Since the first marginal of γ is μ_0, $\mu_0^{x_1}$ is absolutely continuous for every $x_1 \in A_1$. We are thus left to show that for each fixed x_1 in the countable set A_1, the measure $T_{x_1\#}\mu_0^{x_1}$ is absolutely continuous with respect to the $(d-1)$-Hausdorff measure on ∂K, which from now on we denote by $\sigma_{(d-1),\partial K}$. We now fix $x_1 \in \operatorname{spt}(\mu_1)$ and a Borel subset A of ∂K. Our aim is to bound

$$(T_{x_1\#}\mu_0^{x_1})(A) = \mu_0^{x_1}(T_{x_1}^{-1}(A)).$$

Since $\mu_0(K) = 0$, we have $\mu_0^{x_1}(T_{x_1}^{-1}(A) \setminus K) = \mu_0^{x_1}(T_{x_1}^{-1}(A))$. Now take $x = T_{x_1}(x_0) \in \partial K$ with $x_0 \notin K$. By optimality, there exists $\beta \geq 0$ such that

$$\widehat{x - x_0} + \lambda \widehat{x - x_1} + \beta n(x) = 0 \quad \text{where for } \xi \in \mathbb{R}^d \setminus \{0\}, \text{ we have set } \widehat{\xi} = \xi/|\xi|,$$

where $n(x)$ is the outward normal to ∂K at x. This time our aim is to write, for fixed x_1, x_0 as a Lipschitz function of x and a length factor, so we proceed as follows: Using the fact that $\lambda\widehat{x - x_1}$ has norm λ yields

$$1 = \beta^2 + \lambda^2 + 2\beta\lambda n(x) \cdot \widehat{x - x_1}.$$

This time, it is possible that there are two positive solutions for β. We denote them by

$$\beta_{x_1}^+(x) := -\lambda n(x) \cdot \widehat{x - x_1} + \sqrt{(\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2},$$
$$\beta_{x_1}^-(x) := -\lambda n(x) \cdot \widehat{x - x_1} - \sqrt{(\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2}.$$

Hence, one of the following equalities is satisfied by (x_0, x, x_1):

$$x_0 = x + r(\lambda\widehat{x - x_1} + \beta_{x_1}^+(x)n(x)) =: F_{x_1}^+(r, x),$$
$$x_0 = x + r(\lambda\widehat{x - x_1} + \beta_{x_1}^-(x)n(x)) =: F_{x_1}^-(r, x),$$

where $r \in [0, R]$ and $R = \operatorname{diam}(\operatorname{spt}\mu_0 - K)$.

Now consider a Borel set $A \subset \partial K$ with $\sigma_{(d-1),\partial K}(A) = 0$. We distinguish the cases where the discriminant $(\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2$ is zero or positive:

$$\begin{aligned} A_0 &:= \big\{x \in A : (\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2 = 0\big\}, \\ A_> &:= \big\{x \in A : (\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2 > 0\big\} \\ &= \bigcap_{\delta > 0} \underbrace{\big\{x \in A : (\lambda n(x) \cdot \widehat{x - x_1})^2 + 1 - \lambda^2 \geq \delta\big\}}_{=: A_\delta}. \end{aligned}$$

Since $F_{x_1}^+$ and $F_{x_1}^-$ agree with Lipschitz functions on $[0, R] \times A_0$ and $[0, R] \times A_\delta$, respectively, we obtain for fixed $\delta > 0$

$$\mu_0^{x_1}(T_{x_1}^{-1}(A)) \leq \mu_0^{x_1}(T_{x_1}^{-1}(A_0)) + \lim_{\delta \searrow 0} \mu_0^{x_1}(T_{x_1}^{-1}(A_\delta))$$

$$
\begin{aligned}
&\le \mu_0^{x_1}(F_{x_1}^+([0,R]\times A_0)) \\
&\qquad + \lim_{\delta\searrow 0}\big(\mu_0^{x_1}(F_{x_1}^+([0,R]\times A_\delta)) + \mu_0^{x_1}(F_{x_1}^-([0,R]\times A_\delta))\big) \\
&= 0,
\end{aligned}
$$

as required. ∎

It is unclear whether an L^∞ bound can be obtained with the same proof strategy, since the Lipschitz constant of the maps $F_{x_1}^+$ and $F_{x_1}^-$ may blow up as $\delta \to 0^+$. In addition, in Proposition 4.4 the smoothness of K is crucial, as the example below shows.

Example 4.8. In the two-dimensional case, take as K the square $\{|x|+|y|\le 1\}$ and consider the distance-like cost of Proposition 4.4 with $\alpha = 1$ and $\lambda > 1$. Take as μ_0 the Lebesgue measure on the disk $B(x_0, r)$ and as μ_1 the Lebesgue measure on the disk $B(x_1, r)$, with $x_0 = (-a, 0)$ and $x_1 = (a, 0)$, as in Figure 1. The optimal pivot measure μ has in this case a part proportional to the Dirac mass $\delta_{(1,0)}$ and in some cases, when λ is large, a is large, and r is small, actually reduces to the Dirac mass $\delta_{(1,0)}$.

4.2. Density constrained solutions are bang-bang

We end this section by observing that in the case of a density constraint $\mu \le \phi$, for distance-like costs, minimizers are of bang-bang-type.

Proposition 4.9. *Assume that c_0 and c_1 are as in (4.1) with $\lambda > 1$ if $\alpha = 1$, $\phi \in L^1(\mathbb{R}^d)$ is non-negative with compact support, $\int \phi\, dx > 1$, and both* $\operatorname{spt}(\phi)\cap\operatorname{spt}(\mu_0)$ *and* $\operatorname{spt}(\phi)\cap\operatorname{spt}(\mu_1)$ *are Lebesgue negligible. Then, any solution μ of (2.1) under the constraint $\mu \le \phi$ is of the form $\mu = \phi \mathbb{1}_E$ for some measurable subset E of* $\operatorname{spt}(\phi)$.

Proof. Let us start with the case $\alpha = 1$, $\lambda > 1$ and define $A := \{0 < \mu < \phi\}$. We then consider (Lipschitz) potentials φ_0 and φ_1 as in the proof of Corollary 3.3. Almost every point of A is a differentiability point of φ_0 and φ_1, satisfies $\nabla\varphi_0 + \nabla\varphi_1 = 0$, and lies

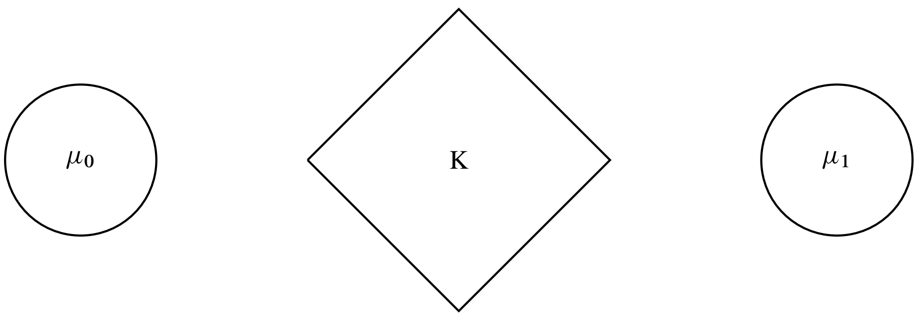

Figure 1. A non-smooth constraint set K may provide a singular optimal pivot measure.

in $\mathbb{R}^d \setminus (\mathrm{spt}(\mu_0) \cup \mathrm{spt}(\mu_1))$. Hence, arguing as in the proof of Corollary 3.3, for almost every x in A, one can find $x_0 \in \mathrm{spt}(\mu_0) \setminus \{x\}$ and $x_1 \in \mathrm{spt}(\mu_1) \setminus \{x\}$ such that

$$0 = \nabla\varphi_0(x) + \nabla\varphi_1(x) = \frac{x - x_0}{|x - x_0|} + \lambda \frac{x - x_1}{|x - x_1|},$$

which is impossible, since $\lambda > 1$. This shows that A is negligible and ends the proof for this case.

Consider now the slightly more complicated case where $\alpha \in (0, 1)$. Since the map $x \mapsto |x - x_0|^\alpha$ is Lipschitz only away from x_0, it is convenient to introduce for $\delta > 0$ the set

$$B_\delta := \big\{x \in \mathrm{spt}(\phi) : d(x, \mathrm{spt}(\mu_0) \cup \mathrm{spt}(\mu_1)) \geq \delta\big\}.$$

On B_δ, the potentials φ_0 and φ_1 are Lipschitz and we can find a subset $\widetilde{B}_\delta$ of B_δ with $B_\delta \setminus \widetilde{B}_\delta$ negligible such that φ_0 and φ_1 are differentiable on $\widetilde{B}_\delta$. Consider now for $\varepsilon > 0$

$$A_\varepsilon := \big\{\varepsilon < \mu < \phi - \varepsilon\big\}$$

and let $\widetilde{A}_\varepsilon$ be the subset (of full Lebesgue measure in A_ε by Lebesgue's density theorem) consisting of its points of density 1, that is,

$$\widetilde{A}_\varepsilon := \left\{y \in A_\varepsilon \ : \ \lim_{r \to 0^+} \frac{\mathcal{L}^d(B(y, r) \cap A_\varepsilon)}{\mathcal{L}^d(B(y, r))} = 1\right\}.$$

Note that $\widetilde{A}_\varepsilon \subset \mathrm{spt}(\phi - \mu)$ and, arguing as before, for almost every $x \in \widetilde{A}_\varepsilon \cap \widetilde{B}_\delta$, we can find $(x_0, x_1) \in \mathrm{spt}(\mu_0) \times \mathrm{spt}(\mu_1)$ such that

$$\nabla\varphi_0(x) + \nabla\varphi_1(x) = \nabla f_{x_0,x_1}(x) = 0,$$

where $f_{x_0,x_1}(x) := |x - x_0|^\alpha + \lambda|x - x_1|^\alpha$. Moreover, we know from Corollary 3.2 that $\mathrm{spt}(\phi - \mu)$ is included in the level set $f_{x_0,x_1} \geq f_{x_0,x_1}(x)$ and so is A_ε, up to a Lebesgue negligible set, by continuity of f_{x_0,x_1}. Since $x \notin \{x_0, x_1\}$ is a critical point of f_{x_0,x_1}, we have $x_1 \neq x_0$, x belongs to $[x_0, x_1]$,

$$e := \widehat{x - x_0} = \widehat{x_1 - x} = \widehat{x_1 - x_0},$$

and the Hessian $D^2 f_{x_0,x_1}$ of f_{x_0,x_1} at x takes the form

$$D^2 f_{x_0,x_1}(x) = (\alpha|x - x_0|^{\alpha-2} + \lambda\alpha|x - x_0|^{\alpha-2})(\mathrm{id} + (\alpha - 2)e \otimes e),$$

which shows that x is a saddle-point of f_{x_0,x_1}, its hessian having a negative eigenvalue with eigenvector e and being positive definite on $e^\perp$. Since for $y \in A_\varepsilon$, we have

$$f_{x_0,x_1}(y) = f_{x_0,x_1}(x) + \frac{1}{2} D^2 f_{x_0,x_1}(x)(y - x, y - x) + o(|y - x|^2) \geq f_{x_0,x_1}(x),$$

we deduce that for $r > 0$ small enough and some positive constant κ, whenever

$y \in A_\varepsilon \cap B(x,r)$, one has $y \in C_{x,e,\kappa}$ where

$$C_{x,e,\kappa} := \big\{ y \in \mathbb{R}^d : |e \cdot (y-x)| \le \kappa |(\mathrm{id} - e \otimes e)(y-x)| \big\}.$$

Hence, for small $r > 0$, $A_\varepsilon \cap B(x,r)$ should lie inside the strict cone $C_{x,e,\kappa}$ so that

$$\limsup_{r\to 0^+} \frac{\mathcal{L}^d(B(x,r)\cap A_\varepsilon)}{\mathcal{L}^d(B(x,r))} \le \limsup_{r\to 0^+} \frac{\mathcal{L}^d(B(x,r)\cap C_{x,e,\kappa})}{\mathcal{L}^d(B(x,r))} < 1,$$

contradicting the fact that x is a point of density 1 of A_ε. This shows that $A_\varepsilon \cap B_\delta$ is negligible, and letting $\delta \to 0^+$, we find that A_ε is negligible; since this is true for every $\varepsilon > 0$, the desired conclusion follows. ■

5. The case of strictly convex costs with a convex location constraint

We now consider (2.1) in the case of the location constraint $\mathcal{A} = \mathcal{P}(K)$, where K is a compact convex subset of $\mathbb{R}^d$ with non-empty interior and c_0 and c_1 satisfy the strong convexity and smoothness assumptions

$$c_i(x,y) := F_i(y-x), \quad F_i \in C^2(\mathbb{R}^d), \quad \lambda\,\mathrm{id} \le D^2 F_i \le \Lambda\,\mathrm{id}, \quad i = 0,1 \tag{5.1}$$

for some constants $0 < \lambda \le \Lambda$. Since these costs are twisted, (2.1) in the case of the location constraint $\mathcal{A} = \mathcal{P}(K)$ admits a unique solution as soon as μ_0 (or μ_1) is absolutely continuous; see Remark 2.4.

Example 5.1. Consider the two-dimensional case with a location constraint given by the square K of Example 4.8; take $\mu_0 = \delta_{(-2,0)}$, μ_1 uniform on the ball of radius 1 centered at $(3,0)$, $c_0(x,y) = |x-y|^2$, and $c_1(x,y) = 2|x-y|^2$. Then, by a direct application of Proposition 2.3, the (unique) solution of (2.1) is explicit: it is the image of the uniform measure on the ball B of radius $2/3$ centered at $(4/3, 0)$ by the projection onto K. It is uniform on $B \cap K$ and has an atom at $(1,0)$, an absolutely continuous part, and a one-dimensional part corresponding to the points of B which project onto the segments $[(0,1),(1,0)]$ and $[(0,-1),(1,0)]$.

This shows that, contrary to the case of distance-like costs, one should expect that μ in general decomposes into a (non-zero) interior part and a boundary part:

$$\mu = \mu^{\mathrm{int}} + \mu^{\mathrm{bd}} \quad \text{where} \quad \mu^{\mathrm{int}}(A) := \mu(A \cap \mathrm{int}(K)),\ \mu^{\mathrm{bd}}(A) := \mu(A \cap \partial K) \tag{5.2}$$

for every Borel subset A of $\mathbb{R}^d$. Regarding μ^{bd}, arguing as in Proposition 4.4, one can show that if μ_0 is absolutely continuous, μ_1 is discrete and K is of class $C^{1,1}$, and μ^{bd} is absolutely continuous with respect to the $(d-1)$-Hausdorff measure on ∂K (and has a bounded density if in addition $\mu_0 \in L^\infty$ and μ_1 is finitely supported; see Proposition 4.5). As for the regularity of μ^{int}, we have

Proposition 5.2. *Assume c_0 and c_1 are of the form in (5.1); μ_0 and μ_1 are compactly supported, with $\mu_0 \in L^\infty$; and K is a compact convex subset of $\mathbb{R}^d$ with non-empty interior. Decomposing the solution μ of (2.1) in the case of the location constraint $\mathcal{A} := \mathcal{P}(K)$ as in (5.2), we have $\mu^{\text{int}} \in L^\infty$ and, more precisely (identifying μ^{int} with its density), we have*

$$\|\mu^{\text{int}}\|_{L^\infty} \leq \|\mu_0\|_{L^\infty} 2^d \lambda^{-d} \Lambda^d, \tag{5.3}$$

where λ and Λ are the positive constants appearing in (5.1).

To establish the L^∞ bound in (5.3), we shall use a penalization strategy, which is detailed in the next paragraph. The proof by a standard Γ-convergence argument is postponed to the end of this section.

5.1. Penalization

Given $g \in C^2(\mathbb{R}^d)$, with g convex and non-negative, let us consider

$$\inf_{\mu \in \mathcal{P}(\mathbb{R}^d)} T(\mu) + \int_{\mathbb{R}^d} g\mu \text{ with } T(\mu) := W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu_1, \mu). \tag{5.4}$$

Then, we have:

Proposition 5.3. *Assuming (5.1) and $\mu_0 \in L^\infty$, (5.4) admits a unique solution μ_g. Moreover, μ_g is absolutely continuous with respect to the Lebesgue measure and its density (still denoted μ_g) satisfies for almost every $x \in \mathbb{R}^d$, the bound*

$$\mu_g(x) \leq \|\mu_0\|_{L^\infty} \lambda^{-d} \det(D^2 g(x) + 2\Lambda \, \text{id}), \tag{5.5}$$

where λ and Λ are the positive constants appearing in (5.1).

Proof. The coercivity of c_0, c_1, and $g \geq 0$ easily give the existence of a minimizer as in Proposition 2.3 (incorporating g into one of the costs considered there), whereas uniqueness is guaranteed by twistedness of the costs and the absolute continuity of μ_0; see Remark 2.4. Also, Proposition 2.3 ensures there is some ball B which contains a neighborhood of $\operatorname{spt}(\mu_g)$. Then, the result [16, Theorem 3.3] from Pass guarantees that the minimizer μ_g is absolutely continuous. The optimality condition derived from the dual formulation of (5.4) (see (3.3)) gives the existence of potentials φ_0 and φ_1 such that

$$\varphi_0 + \varphi_1 + g = 0 \quad \text{ on } B \tag{5.6}$$

and

$$W_{c_0}(\mu_0, \mu_g) = \int_{\mathbb{R}^d} \varphi_0^{c_0} \mu_0 + \int_{\mathbb{R}^d} \varphi_0 \mu_g, \; W_{c_1}(\mu_g, \mu_1) = \int_{\mathbb{R}^d} \varphi_1^{c_1} \mu_1 + \int_{\mathbb{R}^d} \varphi_1 \mu_g,$$

so that defining the c_i-concave potentials

$$\widetilde{\varphi}_0(x) := \inf_{x_0 \in \operatorname{spt}(\mu_0)} \{c_0(x_0, x) - \varphi_0^{c_0}(x_0)\},$$
$$\widetilde{\varphi}_1(x) := \inf_{x_1 \in \operatorname{spt}(\mu_1)} \{c_1(x, x_1) - \varphi_1^{c_1}(x)\},$$

one should have

$$\varphi_i \le \widetilde{\varphi}_i \quad \text{on } B \qquad \text{and} \qquad \varphi_i = \widetilde{\varphi}_i \quad \text{on } \mathrm{spt}(\mu_g). \tag{5.7}$$

Now observe that thanks to (5.1), $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are semi-concave and, more precisely,

$$D^2\widetilde{\varphi}_i \le \Lambda\, \mathrm{id}, \quad i = 0, 1. \tag{5.8}$$

In particular, $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are everywhere superdifferentiable, but on $\mathrm{spt}(\mu_g)$, thanks to (5.6) and (5.7), $\widetilde{\varphi}_0 + \widetilde{\varphi}_1 + g$ is minimal and since g is differentiable, this implies that $\widetilde{\varphi}_0 + \widetilde{\varphi}_1$ is also subdifferentiable on $\mathrm{spt}(\mu_g)$. This readily implies that $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are differentiable on $\mathrm{spt}(\mu_g)$ and that

$$\nabla\widetilde{\varphi}_0 + \nabla\widetilde{\varphi}_1 + \nabla g = 0 \text{ on } \mathrm{spt}(\mu_g).$$

The functions $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are semi-concave and, by Alexandrov's theorem (see [8, Theorem 6.9]), they are twice differentiable μ_g-almost everywhere; the minimality of $\widetilde{\varphi}_0 + \widetilde{\varphi}_1 + g$ on $\mathrm{spt}(\mu_g)$ then gives

$$D^2\widetilde{\varphi}_0 + D^2\widetilde{\varphi}_1 + D^2 g \ge 0 \quad \mu_g\text{-a.e.} \tag{5.9}$$

The optimal transport S_0 for the cost c_0 between μ_g and μ_0 (see [10, Theorem 3.7]) is then given by

$$S_0(x) = x - \nabla F_0^*(\nabla\widetilde{\varphi}_0(x)), \quad x \in \mathrm{spt}(\mu_g),$$

where F_0^* is the Legendre transform of F_0. The absolute continuity of μ_g enables us to use Cordero-Erausquin's theorem (see [6, Theorem 4.8]) to get the existence of a set of full measure for μ_g, for which one has the Jacobian equation

$$\mu_g = \mu_0 \circ S_0 \det(\mathrm{id} - D^2F_0^*(\nabla\widetilde{\varphi}_0)D^2\widetilde{\varphi}_0), \tag{5.10}$$

where $D^2\widetilde{\varphi}_0(x)$ is to be understood in the sense of Alexandrov and the matrix $\mathrm{id} - D^2F_0^*(\nabla\widetilde{\varphi}_0)D^2\widetilde{\varphi}_0$, which is diagonalizable with real and non-negative eigenvalues can be rewritten as

$$\mathrm{id} - D^2F_0^*(\nabla\widetilde{\varphi}_0)D^2\widetilde{\varphi}_0 = D^2F_0^*(\nabla\widetilde{\varphi}_0)(D^2F_0(x - S_0(x)) - D^2\widetilde{\varphi}_0(x)).$$

Together with (5.10), since $D^2F_0^* \le \lambda^{-1}\,\mathrm{id}$ and $D^2F_0(x - S_0(x)) - D^2\widetilde{\varphi}_0(x)$ is semi-definite positive, this gives for μ_g-almost every x that

$$\mu_g(x) \le \|\mu_0\|_{L^\infty}\lambda^{-d} \det(D^2F_0(x - S_0(x)) - D^2\widetilde{\varphi}_0(x)).$$

By (5.9) and (5.8), we then have

$$-D^2\widetilde{\varphi}_0(x) \le D^2g(x) + D^2\widetilde{\varphi}_1(x) \le D^2g(x) + \Lambda\,\mathrm{id},$$

but since $D^2F_0 \le \Lambda\,\mathrm{id}$, bound (5.5) follows. ∎

5.2. Proof of the bound by Γ-convergence

Recall that we have assumed that K is a convex compact subset with non-empty interior. For $\varepsilon > 0$, set $K_\varepsilon := K + \varepsilon B$ (where B is the unit Euclidean ball of $\mathbb{R}^d$). Consider the mollifiers $\eta_\varepsilon = \varepsilon^{-d}\eta(\frac{\cdot}{\varepsilon})$ with η a smooth probability density supported on B and consider the smooth and convex function

$$g_\varepsilon := \eta_\varepsilon \star \varepsilon^{-1} d^2_{K_\varepsilon},$$

where d_{K_ε} is the distance to K_ε. Define T as in (5.4). For every $\nu \in \mathcal{P}(\mathbb{R}^d)$,

$$J_\varepsilon(\nu) := T(\nu) + \int_{\mathbb{R}^d} g_\varepsilon \nu, \quad J(\nu) := \begin{cases} T(\nu) & \text{if } \nu \in \mathcal{P}(K), \\ +\infty & \text{otherwise.} \end{cases}$$

It is easy to see that J_ε Γ-converges to J as $\varepsilon \to 0^+$ for the narrow topology. Hence, the tight sequence of minimizers of J_ε, $\mu_\varepsilon := \mu_{g_\varepsilon}$ converges narrowly to μ the minimizer of J, that is, the solution of (2.1) with $\mathcal{A} = \mathcal{P}(K)$. Since $D^2 g_\varepsilon = 0$ on $\operatorname{int}(K)$, we deduce from (5.5) that for every open Ω such that $\Omega \Subset \operatorname{int}(K)$,

$$\|\mu_\varepsilon\|_{L^\infty(\Omega)} \leq \|\mu_0\|_{L^\infty} 2^d \lambda^{-d} \Lambda^d,$$

from which one deduces (5.3) by letting $\varepsilon \to 0^+$.

6. A parking location model

In this section, we introduce a mathematical model for the optimal location of a parking area in a city. We fix

- a compactly supported probability measure on $\mathbb{R}^d$, ν_0, which represents the distribution of residents in a given area;
- a compactly supported probability measure on $\mathbb{R}^d$, ν_1, which represents the distribution of services.

The goal is to determine a measure μ which represents the density of parking places in order to minimize a suitable total transportation cost. All the residents travel to reach the services, but some of them may simply walk (which will cost $c_1(x, y)$ to go from x to y), while some others may use their car to reach a parking place (which will cost $c_0(x, y)$ to go from x to the parking place y) and then walk from the parking place to the services (which will cost $c_1(y, z)$ to go from y to z). We consider two cost functions c_0 and c_1 and the corresponding Monge–Kantorovich functionals W_{c_0} and W_{c_1}, defined as in (1.1), respectively, representing the cost of moving by car and the cost of walking. It may be natural to assume that walking is more costly than driving, that is, $c_1 \geq c_0$; for instance, we may take $p > 0$ and

$$c_0(x, y) = |x - y|^p, \qquad c_1(x, y) = \lambda |x - y|^p \quad \text{with} \quad \lambda \geq 1. \tag{6.1}$$

Assuming that $\mu_0 \leq \nu_0$ denotes the distribution of driving residents and $\mu_1 \leq \nu_1$ the corresponding services they reach for, the total cost we consider is

$$F(\mu_0, \mu_1, \mu) = W_{c_1}(\nu_0 - \mu_0, \nu_1 - \mu_1) + W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1). \tag{6.2}$$

The optimization problems we consider are then the minimization of $F(\mu_0, \mu_1, \mu)$, subject to the constraints

$$0 \leq \mu_0 \leq \nu_0, \quad 0 \leq \mu_1 \leq \nu_1, \quad \int d\mu_0 = \int d\mu_1 = \int d\mu,$$

and additional constraints such as

- no other constraints on the parking density μ;
- location constraints, that is, $\operatorname{spt}\mu \subset K$, with an a priori given compact set $K \subset \mathbb{R}^d$;
- density constraints, that is, $\mu \leq \phi$, for a given non-negative and integrable function ϕ.

This optimization problem in the case of a location constraint can also be reformulated as a linear program in the following way:

$$\inf \int_{\mathbb{R}^d \times \mathbb{R}^d} c_1(x_0, x_1)\, \mathrm{d}\gamma(x_0, x_1) + \int_{\mathbb{R}^d \times \mathbb{R}^d \times K} (c_0(x_0, x) + c_1(x, x_1))\, \mathrm{d}\widetilde{\gamma}(x_0, x, x_1), \tag{6.3}$$

subject to the constraints

$$\gamma, \widetilde{\gamma} \geq 0, \quad \gamma + \pi_{0,1\#}\widetilde{\gamma} \in \Pi(\nu_0, \nu_1).$$

It is indeed easy to see that the optimal solution to minimizing the functional in (6.2) is given by $\pi_{\mathrm{piv}\#}\widetilde{\gamma}$. Hence, to incorporate a density constraint in formulation (6.3) one needs to add the constraint $\pi_{\mathrm{piv}\#}\widetilde{\gamma} \leq \phi$. The problem with location constraint is actually equivalent to a standard optimal transport problem with cost function

$$C(x_0, x_1) := \min\big\{c_1(x_0, x_1), \inf_{x \in K}\{c_0(x_0, x) + c_1(x, x_1)\}\big\}.$$

More precisely, consider

$$\inf_{\beta \in \Pi(\nu_0, \nu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} C(x_0, x_1)\, \mathrm{d}\beta(x_0, x_1). \tag{6.4}$$

Then, both (6.3) and (6.4) admit solutions and they are equivalent in the following sense:

- $\min(6.3) = \min(6.4)$,
- if $\gamma, \widetilde{\gamma}$ are optimal for (6.3), then $\beta := \gamma + \pi_{0,1\#}\widetilde{\gamma}$ is optimal for (6.4),
- if β is optimal for (6.4), then defining

$$V_1 := \big\{(x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d : c_1(x_0, x_1) = C(x_0, x_1)\big\},$$

$\gamma := \beta_{|V_1}$, $P : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ (measurable) by

$$P(x_0, x_1) \in \operatorname*{argmin}_{x \in K}\big\{c_0(x_0, x) + c_1(x, x_1)\big\},$$

and

$$\mathrm{d}\widetilde{\gamma}(x_0, x_1, x) := \delta_{P(x_0,x_1)}(x) \otimes \mathrm{d}\beta_{|\mathbb{R}^d \setminus V_1}(x_0, x_1),$$

then γ and $\widetilde{\gamma}$ are optimal for (6.3).

Remark 6.1. Note that the solutions (μ_0, μ, μ_1) to minimizing (6.2) (resp. the solutions γ and $\widetilde{\gamma}$ to (6.3)) are not necessarily probability measures. The optimal common total mass of $\widetilde{\gamma}$, μ_0, and μ_1 represents the fraction of ν_0 which uses the parking. Thus, the parking problem is a generalization of the interpolation problem from Section 2, which corresponds to imposing that the parking measure is of full mass.

However, under quite general and natural assumptions, it can be shown that the optimal parking measure is non-trivial.

Lemma 6.2. *Assume that $c_0, c_1 \geq 0$ are continuous with $c_0(x, x) = c_1(x, x) = 0$ and $c_1(x, y) > c_0(x, y)$ for $x \neq y \in \mathbb{R}^d$. Consider the density constraint case $\mu \leq \phi$ with $\phi \in L^1(\mathbb{R}^d)$, $0 < \phi < +\infty$ almost everywhere. Then, if $\nu_0 \neq \nu_1$, the optimal μ for the parking problem is non-trivial, that is, $\mu \neq 0$.*

Proof. Assume by contradiction that $\mu = 0$ is an optimal solution. The optimal cost for the parking problem is then given by

$$W_{c_1}(\nu_0, \nu_1).$$

Let γ_1 be an associated optimal transport plan. Since $\nu_0 \neq \nu_1$, there is $(x_0, x_1) \in \operatorname{spt} \gamma_1$ with $x_0 \neq x_1$. Clearly, there exists $x \in \mathbb{R}^d$ (take, for instance, $x = x_1$) such that

$$c_0(x_0, x) + c_1(x, x_1) < c_1(x_0, x_1).$$

Then, by continuity of the cost functions, there exists an open neighborhood of the form $A_0 \times A \times A_1$ of (x_0, x, x_1) such that the inequality remains valid, that is,

$$c_0(y_0, y) + c_1(y, y_1) < c_1(y_0, y_1),$$

for all $(y_0, y, y_1) \in A_0 \times A \times A_1$. By possibly choosing a smaller (open) A, we can assume that

$$\int_A \phi(x)\,\mathrm{d}x \leq \gamma_1(A_0 \times A_1),$$

so that there is $t \in (0, 1]$ such that

$$\int_A \phi(x)\,\mathrm{d}x = t\gamma_1(A_0 \times A_1).$$

But then,

$$W_{c_1}(\nu_0, \nu_1) = \int c_1(y_0, y_1)\,\mathrm{d}\gamma_1(y_0, y_1)$$

$$
\begin{aligned}
&> \int_{(A_0\times A_1)^c} c_1(y_0,y_1)\,\mathrm{d}\gamma_1(y_0,y_1) + (1-t)\int_{A_0\times A_1} c_1(y_0,y_1)\,\mathrm{d}\gamma_1(y_0,y_1)\\
&\qquad + (\gamma_1(A_0\times A_1))^{-1}\int_{A_0\times A_1}\int_A (c_0(y_0,y)+c_1(y,y_1))\phi(y)\,\mathrm{d}y\,\mathrm{d}\gamma_1(y_0,y_1)\\
&\geq W_{c_1}(\nu_0-\mu_0,\nu_1-\mu_1) + W_{c_0}(\mu_0,\mu') + W_{c_1}(\mu',\mu_1),
\end{aligned}
$$

where $\mu' = \phi_{|A}$, and μ_0, μ_1 are the marginals of $t\gamma_{1|A_0\times A_1}$. This gives a contradiction and thus achieves the proof. ∎

6.1. Examples

We first solve a simple particular example in $\mathbb{R}^2$ before giving some numerical simulations. This example shows that in some cases, the optimal choices for μ_0, μ_1, μ are not of unitary mass; these correspond to the cases where it is more efficient for some residents to walk from their residence to the services without using their car.

Example 6.3. Let $\nu_0 = \delta_{x_0}$ and $\nu_1 = \delta_0$ be two Dirac masses in $\mathbb{R}^2$ with $x_0 \neq 0$. We consider the costs c_0 and c_1 as in (6.1) with $p > 0$ and $\lambda > 1$. Then, $\mu_0 = \alpha\delta_{x_0}$ and $\mu_1 = \alpha\delta_0$, for some $\alpha \in [0,1]$ and the optimization problems for the functional F in (6.2) become the minimization of the quantity

$$
\begin{aligned}
\lambda(1-\alpha)|x_0|^p &+ \int (|x-x_0|^p + \lambda|x|^p)\,d\mu\\
&= \lambda|x_0|^p + \int (|x-x_0|^p + \lambda|x|^p - \lambda|x_0|^p)\,d\mu.
\end{aligned}
$$

Since $\lambda|x_0|^p$ is fixed, we are reduced to minimizing the quantity

$$
\int (|x-x_0|^p + \lambda|x|^p - \lambda|x_0|^p)\,d\mu
$$

with the constraint $\int d\mu \leq 1$ and possibly other location and density constraints on μ, as illustrated above. Setting

$$
f(x) = |x-x_0|^p + \lambda|x|^p - \lambda|x_0|^p, \tag{6.5}
$$

it is clear that μ has to be concentrated on the set where $f \leq 0$. The optimization problem with no other constraints on μ has then the trivial solution $\alpha = 1$ and $\mu = \delta_{\operatorname{argmin} f}$ (for instance, $\mu = \delta_0$ if $p = 1$ and $\mu = \delta_{(1+\lambda)^{-1}x_0}$ if $p = 2$). The situation becomes more interesting when other constraints on μ are present. If we impose $\operatorname{spt}\mu \subset K$, let $\overline{x} \in K$ be a minimum point of the function f in (6.5) over K. If $f(\overline{x}) < 0$, then $\alpha = 1$ and $\mu = \delta_{\overline{x}}$ is a solution; if $f(\overline{x}) \geq 0$, then $\alpha = 0$ and $\mu = 0$ is a solution.

We now consider the more realistic case when a density constraint on μ is imposed. We take $\mu \leq 1$. The optimal measure μ for the cost in (6.2) is then the characteristic function 1_{A_c} of a suitable level set $A_c = \{f \leq c\}$ with $c \leq 0$. Thus, the following situations may occur:

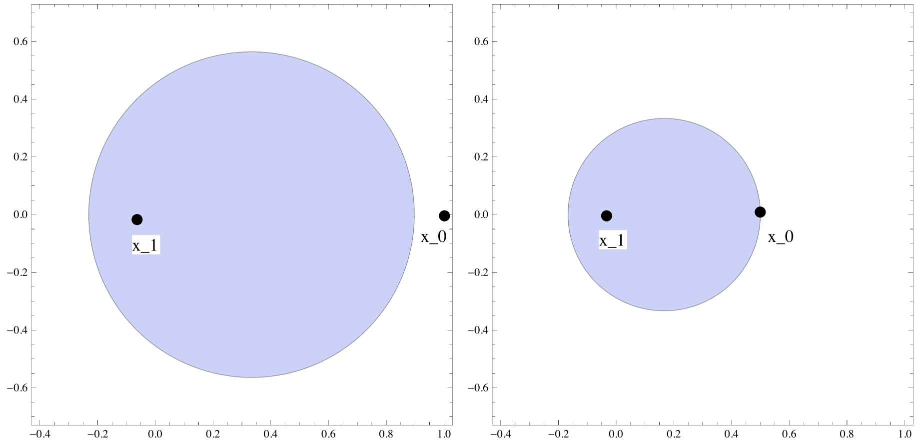

Figure 2. The case $p = \lambda = 2$. On the left, $|x_0| = 1$ gives $|A_{\text{opt}}| = 1$; on the right, $|x_0| = 1/2$ gives $|A_{\text{opt}}| \simeq 0.35$.

- If $|\{f \leq 0\}| \geq 1$, then $\alpha = 1$ and $\mu = 1_{A_c}$, where the level $c \leq 0$ is such that $|A_c| = 1$. Note that, since the function f is convex, the set A_c is convex too. This happens when x_0 is far enough from the origin and all people then drive to the parking area A_c.
- If $|\{f \leq 0\}| < 1$, then $\alpha = |\{f \leq 0\}|$ and $\mu = 1_{A_0}$.

For instance, when $p = 2$, it is easy to see that the set A_0 is the ball centered at $x_0/(1+\lambda)$ with radius $\lambda|x_0|/(1+\lambda)$. Therefore,

- if $|x_0| \geq \pi^{-1/2}(\lambda+1)/\lambda$, we have $\alpha = 1$ and $\mu_{\text{opt}} = 1_A$, where A is the disk centered at $x_0/(1+\lambda)$ of unitary area;
- if $|x_0| < \pi^{-1/2}(\lambda+1)/\lambda$, we have $\alpha = \pi|x_0|^2\lambda^2/(\lambda+1)^2$ and $\mu_{\text{opt}} = 1_{A_0}$. In this case, only the fraction α of people drive to reach the parking area, while the rest of residents walk up to the services.

In Figure 2 the two situations are graphically represented in the cases $p = \lambda = 2$, while in Figure 3 we plot the two optimal parking areas when $p = 1$ and $\lambda = 2$.

7. Numerical simulations

For the numerical simulation of examples in the case of interpolation between measures, given by (2.1), and parking problem (6.2), we replace the optimal transportation costs by their entropically regularized versions. This will enable us to apply some variants of the celebrated Sinkhorn's algorithm, popularized in the context of optimal transport and matching by [7] and [9], respectively. For an introduction to this rapidly developing subject and convergence results, we refer the reader to [15, 20].

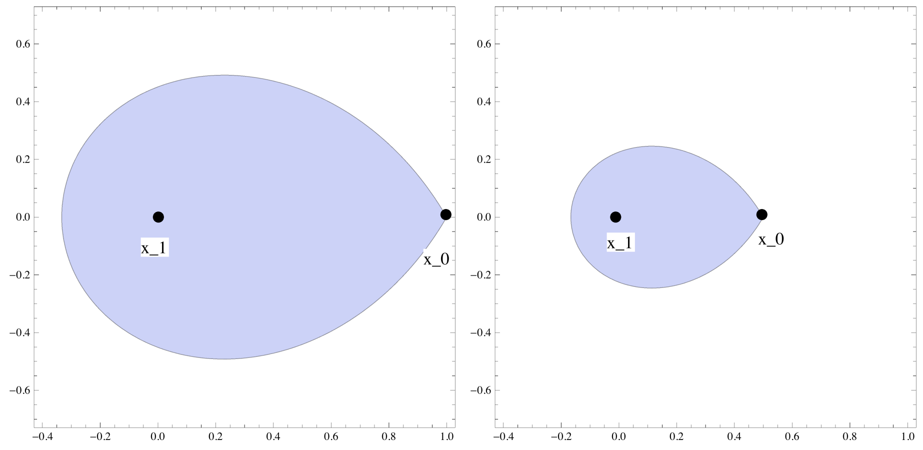

Figure 3. The case $p = 1$ and $\lambda = 2$. On the left, $|x_0| = 1$ gives $|A_{\text{opt}}| \simeq 0.97$; on the right, $|x_0| = 1/2$ gives $|A_{\text{opt}}| \simeq 0.24$.

7.1. Description of the Sinkhorn-like algorithm

The entropically regularized optimal transport cost for a cost function c, a regularizing parameter $\varepsilon > 0$, and a fixed reference measure $Q \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ is given by

$$\inf\Big\{\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\, \mathrm{d}\gamma(x,y) + \varepsilon H(\gamma|Q) : \gamma \in \Pi(\mu_0,\mu_1)\Big\}, \tag{7.1}$$

where the relative entropy $H(P|Q)$ between two non-negative finite measures P, Q on $\mathbb{R}^d$ is defined by

$$H(P|Q) := \begin{cases} \int_{\mathbb{R}^d} \Big(\log\Big(\frac{\mathrm{d}P}{\mathrm{d}Q}\Big) - 1\Big)\, \mathrm{d}P & \text{if } P \ll Q, \\ +\infty & \text{otherwise.} \end{cases}$$

Note that, by setting $R = e^{-c/\varepsilon} Q$, we have

$$\varepsilon H(\gamma|R) = \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\, \mathrm{d}\gamma(x,y) \; + \varepsilon H(\gamma|Q),$$

so that (7.1) amounts to minimizing $H(\cdot|R)$ among transport plans between μ_0 and μ_1. As already observed in [2, Section 3.2], the entropically regularized version of (2.1) becomes, for two suitably chosen reference measures R_0, R_1,

$$\inf\big\{H(\gamma_0|R_0) + H(\gamma_1|R_1) : \mu \in \mathcal{A},\ \gamma_0 \in \Pi(\mu_0,\mu),\ \gamma_1 \in \Pi(\mu,\mu_1)\big\}. \tag{7.2}$$

Cases (i) with no additional constraint and (ii) with location constraint K can be treated by choosing the reference measures to enforce the support of μ being included in K—namely, we choose

$$R_0 = e^{-c_0/\varepsilon} \mu_0 \otimes \mathbb{1}_K, \quad R_1 = e^{-c_1/\varepsilon} \mathbb{1}_K \otimes \mu_1, \tag{7.3}$$

where for case (i) we choose K large enough (yet still compact) as before. The resulting Sinkhorn iterations are standard; see, for instance, [2, Propositions 1 and 2]. case (iii) of a density constraint ϕ requires performing a suitable projection of the estimated interpolation, as specified in [19, Proposition 4.1] in the case of $\phi \equiv \kappa$. We write the corresponding Sinkhorn iterations, including the projection for the density constraint for sake of completeness, in its dual form where the algorithm essentially becomes alternate gradient ascent. For this, note that the dual of (7.2) with R_i as in (7.3) in the case of a density constraint $\mu \leq \phi$ ($\operatorname{spt}\phi \subset K$) is given by

$$\begin{aligned}\sup_{\substack{\varphi_0,\varphi_1,\\ \psi_0,\psi_1}} \Big\{-\sum_{i=0}^{1} \int_{\mathbb{R}^{2d}} \exp(\varphi_i + \psi_i)\,\mathrm{d}R_i + \sum_{i=0}^{1} \int_{\mathbb{R}^d} \varphi_i \,\mathrm{d}\mu_i \\ + \int_{\mathbb{R}^d} (\psi_0 + \psi_1)\phi \,\mathrm{d}x : \psi_0 + \psi_1 \leq 0\Big\}.\end{aligned}$$

Sinkhorn iterations are given by the explicit coordinate ascent updates for this dual formulation:

$$\begin{aligned}\exp(\varphi_i^{l+1}(x_i)) &= \Big(\int_K \exp\Big(-\frac{c_i}{\varepsilon} + \psi_i^l(x)\Big)\,\mathrm{d}x\Big)^{-1},\\ \exp(\psi_i^{l+1}(x)) &= \min\{\mu^l, \phi\}\Big(\int_{\mathbb{R}^d} \exp\Big(-\frac{c_i}{\varepsilon} + \varphi_i^{l+1}(x_i)\Big)\,\mathrm{d}\mu_i(x_i)\Big)^{-1},\end{aligned}$$

where μ^l is the current approximate interpolation which is given by the geometric mean formula (see [2, Proposition 2]):

$$\mu^l = \prod_{j=0}^{1}\Big(\int \exp\Big(-\frac{c_j}{\varepsilon} + \varphi_j^{l+1} + \psi_j^l\Big)\,\mathrm{d}\mu_j(x_j)\Big)^{\frac{1}{2}}.$$

Regularizing parking problem (6.2) in a similar way leads to

$$\inf_{\gamma,\tilde{\gamma}_0,\tilde{\gamma}_1,\tilde{\gamma}\in\mathcal{M}} \big\{H(\gamma|R) + H(\tilde{\gamma}_0|R_0) + H(\tilde{\gamma}_1|R_1)\big\}, \tag{7.4}$$

where

$$\begin{aligned}\mathcal{M} = \big\{\gamma, \tilde{\gamma}_0, \tilde{\gamma}_1, \tilde{\gamma} \in \mathcal{M}_+(\mathbb{R}^{2d})^3 \times \mathcal{M}_+(\mathbb{R}^{3d}) :\\ \gamma + \pi_{0,1\#}\tilde{\gamma} \in \Pi(\nu_0,\nu_1), \pi_{0,\mathrm{piv}\#}\tilde{\gamma} = \tilde{\gamma}_0,\ \pi_{\mathrm{piv},1\#}\tilde{\gamma} = \tilde{\gamma}_1\big\}.\end{aligned}$$

As before, a location constraint on a given set K can be encoded in the choice of the reference measures:

$$R = e^{-c_1/\varepsilon}\nu_0 \otimes \nu_1, \quad R_0 = e^{-c_0/\varepsilon}\nu_0 \otimes \mathbb{1}_K, \quad R_1 = e^{-c_1/\varepsilon}\mathbb{1}_K \otimes \nu_1.$$

For the density constraint, we have to add the condition $\pi_{1\#}\widetilde{\gamma}_0 \le \phi$. The dual of (7.4) in the case of a density constraint is then given by

$$\sup_{\substack{\varphi_0,\varphi_1,\\ \psi_0,\psi_1}} \Big\{ -\int_{\mathbb{R}^{2d}} \exp(\varphi_0+\varphi_1)\,\mathrm{d}R - \sum_{i=0}^{1}\int_{\mathbb{R}^{2d}} \exp(\varphi_i+\psi_i)\,\mathrm{d}R_i \\ + \sum_{i=0}^{1}\int_{\mathbb{R}^d} \varphi_i\,\mathrm{d}\nu_i + \int_{\mathbb{R}^d} (\psi_0+\psi_1)\phi\,\mathrm{d}x : \psi_0+\psi_1 \le 0\Big\}.$$

The Sinkhorn iterations (density constraint included) in the dual variables then become

$$\exp(\varphi_i^{l+1}(x_i)) = \Big(\int \exp\Big(-\frac{c_1}{\varepsilon}+\varphi^l_{i+1 \mod 2}(x)\Big)\,\mathrm{d}x + \int \exp\Big(-\frac{c_i}{\varepsilon}+\psi_i^l(x)\Big)\,\mathrm{d}x\Big)^{-1},$$
$$\exp(\psi_i^{l+1}(x)) = \min\{\mu^l,\phi\}\Big(\int \exp\Big(-\frac{c_i}{\varepsilon}+\varphi_i^{l+1}(x_i)\Big)\,\mathrm{d}\mu_i(x_i)\Big)^{-1},$$

and μ^l, the current approximate parking measure, is again given by an explicit geometric mean expression.

7.2. Numerical results: Comparison of the optimal interpolation and the optimal parking

We now present some numerical results based on the iterative schemes described in the previous paragraph. In all our examples (presented in Figures 4 to 7), we compare the solutions of the interpolation and parking problems with a constant density constraint on the unit square $K=[0,1]^2$. We always take as distribution of services $\mu_1=\nu_1=\delta_{(0.5,0.5)}$, the Dirac at the center of the square; and as distribution of residents, we take a symmetric sum of four Dirac masses:

$$\mu_0=\nu_0=\frac{1}{4}(\delta_{(0.5,0.1)}+\delta_{(0.5,0.9)}+\delta_{(0.1,0.5)}+\delta_{(0.9,0.5)}).$$

We consider power-like costs

$$c_0(x,y)=|x-y|^p, \quad c_1(x,y)=2\,c_0(x,y)$$

for several values of p corresponding to concave, linear, or convex costs and various constant threshold values for the density constraints ϕ. In this setting, we know (Corollary 3.3 for $p>1$ and Proposition 4.9 for $p\le 1$) that the optimal interpolation and the optimal parking are of bang-bang-type. Even with the entropic regularization (which has the effect of blurring the true solution), this is clearly what we observe in these figures with a small regularization $\varepsilon = 5.10^{-4}$. Since the optimal parking may have total mass less than 1, we have indicated its total mass on each figure; of course, if the total mass of the parking is 1 it coincides with the interpolation, a case which is more likely to occur when the threshold level is high. Finally, one can see the influence of the exponent p on the shape

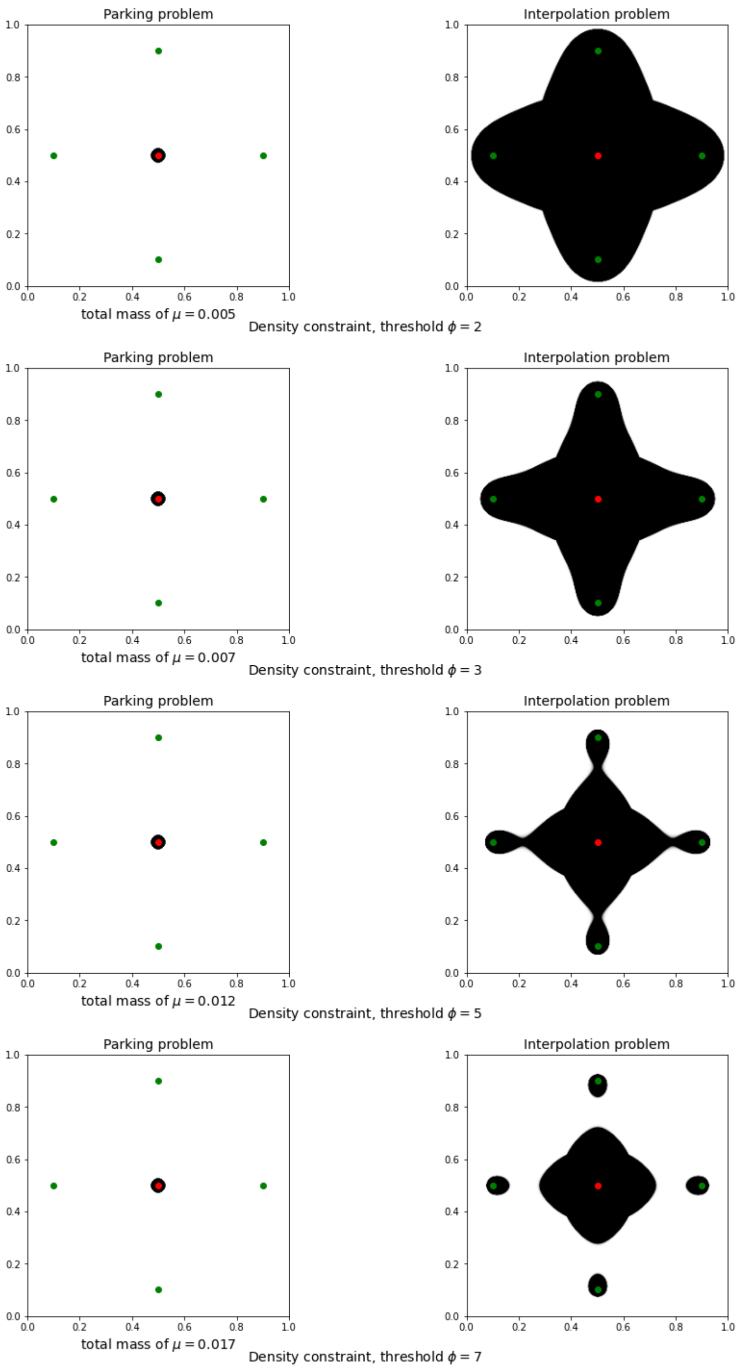

Figure 4. Concave cost $p = 0.25$.

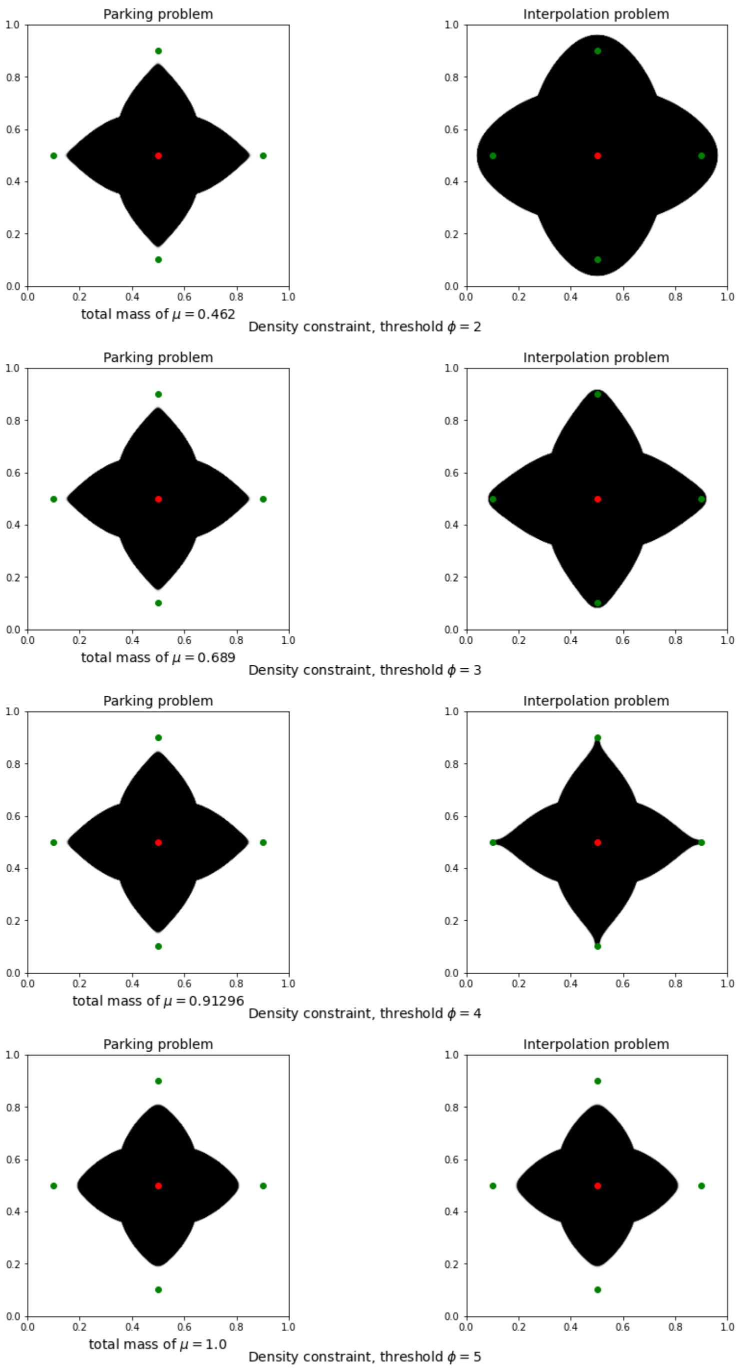

Figure 5. Concave cost $p = 0.75$.

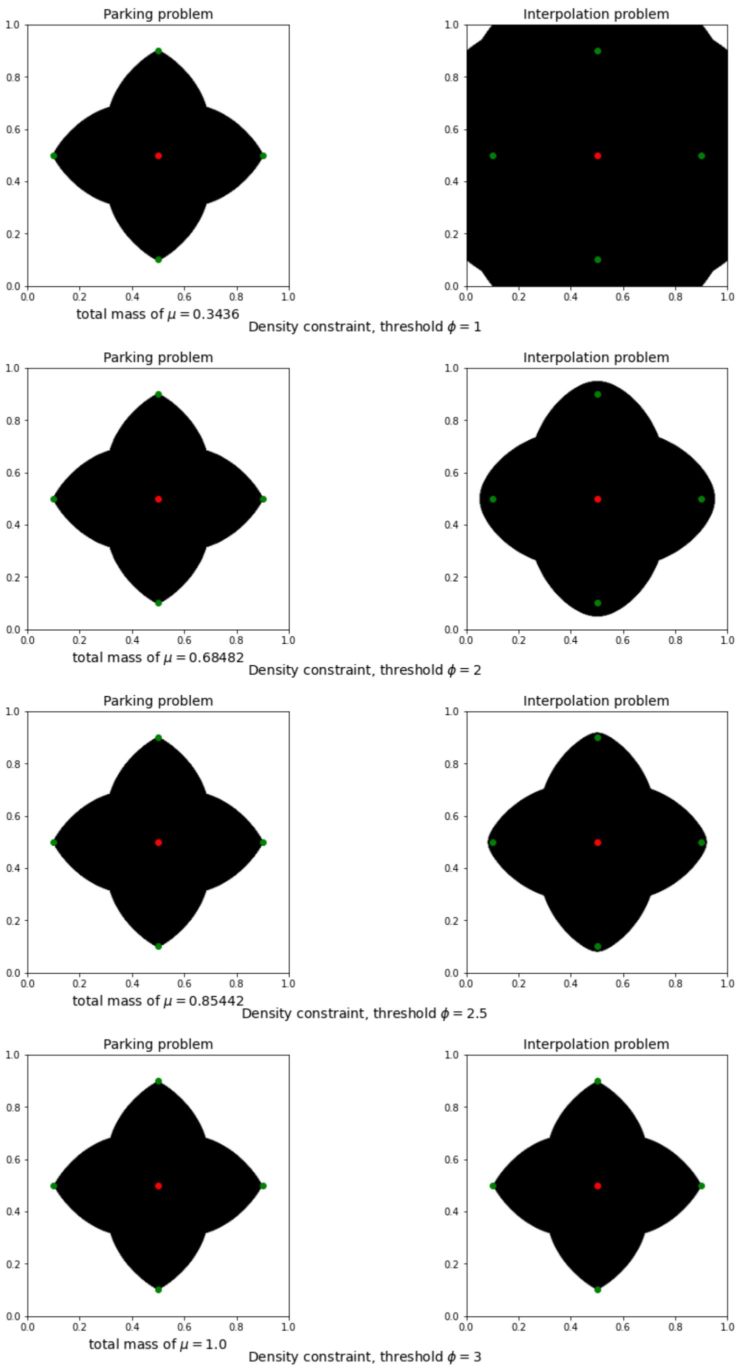

Figure 6. Linear cost $p = 1$.

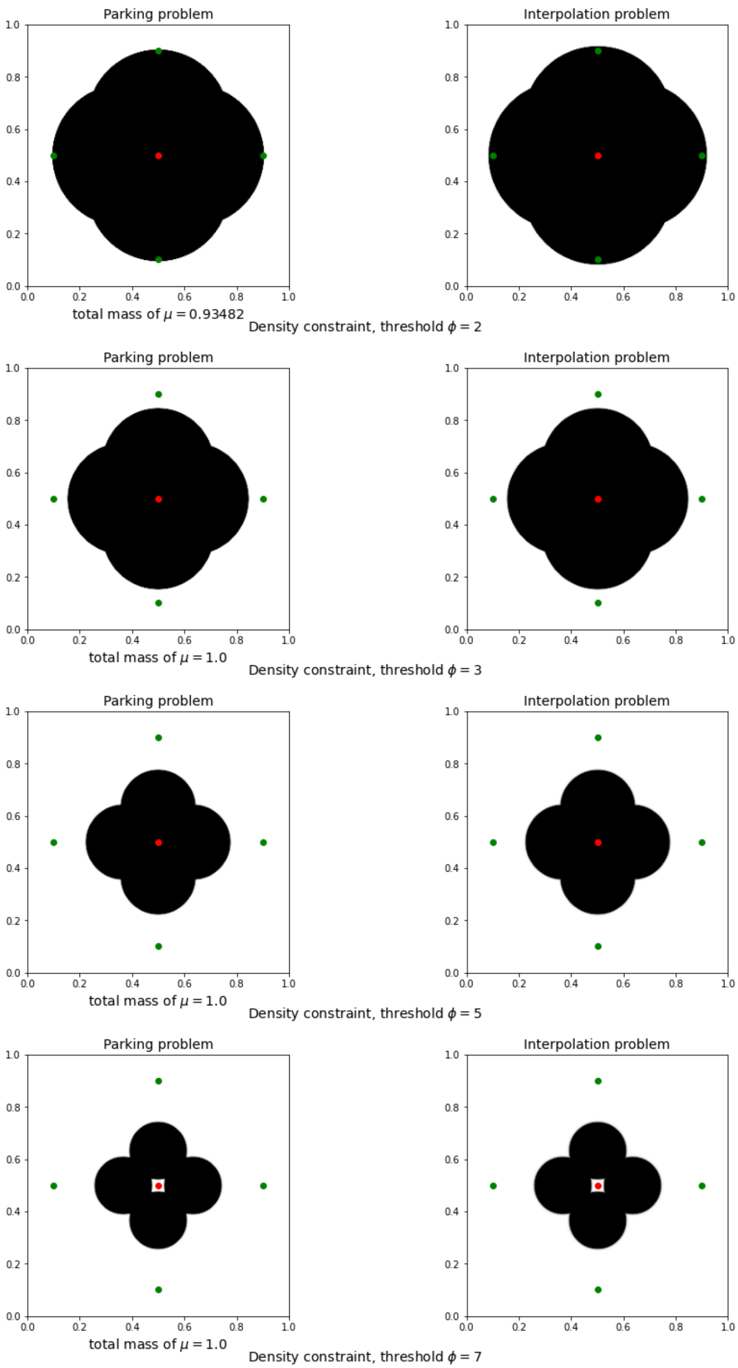

Figure 7. Convex cost $p = 2$.

of the support of the optimal measure and, in particular, recognize for $p = 1$ (Figure 6) the drop-like shape which was explicitly computed and plotted in Figure 3 and balls for $p = 2$ (Figure 7).

Acknowledgments. The authors are grateful to an anonymous referee for insightful and kind suggestions that directly inspired Lemma 6.2.

Funding. The first author is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM); his work is part of the project 2017TEXA3H *"Gradient flows, Optimal Transport and Metric Measure Structures"* funded by the Italian Ministry of Research and University. This work was initiated during a visit of GC and KE at the Dipartimento di Matematica of the University of Pisa, the hospitality of this institution is gratefully acknowledged as well as the support from the Agence Nationale de la Recherche through the project MAGA (ANR-16-CE40-0014). G.C. acknowledges the support of the Lagrange Mathematics and Computing Research Center. KE acknowledges that this project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754362.

References

[1] M. Agueh and G. Carlier, Barycenters in the Wasserstein space. *SIAM J. Math. Anal.* **43** (2011), no. 2, 904–924 Zbl 1223.49045 MR 2801182

[2] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.* **37** (2015), no. 2, A1111–A1138 Zbl 1319.49073 MR 3340204

[3] G. Buttazzo, G. Carlier, and M. Laborde, On the Wasserstein distance between mutually singular measures. *Adv. Calc. Var.* **13** (2020), no. 2, 141–154 Zbl 1446.49010 MR 4085708

[4] G. Carlier, E. Chenchene, and K. Eichinger, Wasserstein medians: Robustness, PDE characterization and numerics. 2023, arXiv:2307.01765

[5] G. Carlier and I. Ekeland, Matching for teams. *Econom. Theory* **42** (2010), no. 2, 397–418 Zbl 1183.91112 MR 2564442

[6] D. Cordero-Erausquin, Non-smooth differential properties of optimal transport. In *Recent advances in the theory and applications of mass transport*, pp. 61–71, Contemp. Math. 353, Amer. Math. Soc., Providence, RI, 2004 Zbl 1141.49040 MR 2079070

[7] M. Cuturi, Sinkhorn Distances: Lightspeed Computation of Optimal Transport. Proceedings of the 26th International Conference on Neural Information Processing Systems, **2** (2013), 2292–2300

[8] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*. Stud. Adv. Math., CRC Press, Boca Raton, FL, 1992 Zbl 0804.28001 MR 1158660

[9] A. Galichon and B. Salanié. Matching with Trade-Offs: Revealed Preferences over Competing Characteristics. Technical Report info:hdl:2441/1293p84sf58s482v2dpn0gsd67, Sciences Po, 2010.

[10] W. Gangbo and R. J. McCann, The geometry of optimal transportation. *Acta Math.* **177** (1996), no. 2, 113–161 Zbl 0887.49017 MR 1440931

[11] H. Leclerc, Q. Mérigot, F. Santambrogio, and F. Stra, Lagrangian discretization of crowd motion and linear diffusion. *SIAM J. Numer. Anal.* **58** (2020), no. 4, 2093–2118 Zbl 1456.65137 MR 4123686

[12] L. Lussardi, M. A. Peletier, and M. Röger, Variational analysis of a mesoscale model for bilayer membranes. *J. Fixed Point Theory Appl.* **15** (2014), no. 1, 217–240 Zbl 1306.49020 MR 3282788

[13] B. Maury, A. Roudneff-Chupin, and F. Santambrogio, A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.* **20** (2010), no. 10, 1787–1821 Zbl 1223.35116 MR 2735914

[14] R. J. McCann, A convexity principle for interacting gases. *Adv. Math.* **128** (1997), no. 1, 153–179 Zbl 0901.49012 MR 1451422

[15] M. Nutz. Introduction to Entropic Optimal Transport, URL, accessed on 20 July 2022

[16] B. Pass, Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions. *Discrete Contin. Dyn. Syst.* **34** (2014), no. 4, 1623–1639 Zbl 1278.49054 MR 3121634

[17] P. Pegon, F. Santambrogio, and D. Piazzoli, Full characterization of optimal transport plans for concave costs. *Discrete Contin. Dyn. Syst.* **35** (2015), no. 12, 6113–6132 Zbl 1334.49146 MR 3393269

[18] M. A. Peletier and M. Röger, Partial localization, lipid bilayers, and the elastica functional. *Arch. Ration. Mech. Anal.* **193** (2009), no. 3, 475–537 Zbl 1170.74034 MR 2525110

[19] G. Peyré, Entropic approximation of Wasserstein gradient flows. *SIAM J. Imaging Sci.* **8** (2015), no. 4, 2323–2351 Zbl 1335.90068 MR 3413589

[20] G. Peyré and M. Cuturi. Computational Optimal Transport: With Applications to Data Science. Found. Trends Mach. Learn. 11, Now Publishers, 2011 Zbl 1475.68011

[21] F. Santambrogio, *Optimal transport for applied mathematicians*. Progr. Nonlinear Differential Equations Appl. 87, Birkhäuser, Cham, 2015 Zbl 1401.49002 MR 3409718

[22] C. Sarrazin, Lagrangian discretization of variational mean field games. *SIAM J. Control Optim.* **60** (2022), no. 3, 1365–1392 Zbl 1498.91041 MR 4426883

[23] C. Villani, *Topics in optimal transportation*. Grad. Stud. Math. 58, American Mathematical Society, Providence, RI, 2003 Zbl 1106.90001 MR 1964483

Received 27 July 2022; revised 16 October 2023.

Giuseppe Buttazzo
Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy;
giuseppe.buttazzo@unipi.it

Guillaume Carlier
CEREMADE, Université Paris Dauphine-PSL, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16; Mokaplan, Inria, 2 Rue Simone Iff, 75012 Paris, France;
carlier@ceremade.dauphine.fr

Katharina Eichinger
CMAP, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France;
katharina.eichinger@polytechnique.edu