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 $\mu_0$ to another one $\mu_1$ consists of two parts: the first one measures the transportation from $\mu_0$ to an intermediate (pivot) measure $\mu$ to be determined (and subject to various constraints), and the second one measures the transportation from $\mu$ to $\mu_1$. This leads to Monge-Kantorovich interpolation problems under constraints for which we establish various properties of the optimal pivot measures $\mu$. 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.

Keywords: optimal transport, Monge-Kantorovich distance, measure interpolation, optimal parking regions.

2010 Mathematics Subject Classification: 49Q22, 49J45, 49M29, 49K99.

1. Introduction

We consider optimal transport problems where a given probability measure $\mu_0$ in $\mathbb{R}^d$ has to be transported to a given probability measure $\mu_1$ with minimal transportation cost. This cost consists of two parts: the first one measures the transportation from $\mu_0$ to an intermediate measure $\mu$, to be determined, and the second one measures the transportation from $\mu$ to $\mu_1$. This situation occurs in some applications, where the transport of $\mu_0$ to $\mu_1$ 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 $\rho_0$, $\rho_1$ we set

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

Here

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

is the set of transport plans between $\rho_0$ and $\rho_1$, 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 $\gamma$, so that a probability measure $\gamma$ on $\mathbb{R}^d \times \mathbb{R}^d$ belongs to $\Pi(\rho_0, \rho_1)$ when

$$\begin{cases} 
\pi_0 \# \gamma(A) = \gamma(A \times \mathbb{R}^d) = \rho_0(A), \\
\pi_1 \# \gamma(A) = \gamma(\mathbb{R}^d \times A) = \rho_1(A)
\end{cases} \quad \text{for all Borel set } A \subset \mathbb{R}^d.$$
Some extra constraints on the *pivot* measure $\mu$ can be added, as for instance:

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

Without additional constraint on the measure $\mu$, the minimization of $W_{c_0}(\mu_0,.) + W_{c_1}(.,\mu_1)$, or its generalizations to more than two prescribed measures, arise in different applied settings such as multi-population matching \[2\] or Wasserstein barycenters \[16\]. In particular, in the quadratic case where $c_0(x, y) = c_1(x, y) = |x - y|^2$, minimizers of $W_{c_0}(\mu_0,.) + W_{c_1}(.,\mu_1)$ are the midpoints of McCann’s displacement interpolation \[14\] between $\mu_0$ and $\mu_1$ i.e. geodesics for the quadratic Wasserstein metric. 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 $\mu$ in a class of measures which are singular with respect to $\mu_0$ 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 \[11\] and \[22\].

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 $\nu_0$ and $\nu_1$, 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 $\mu$ may have total mass less than 1, and one may decompose $\nu_0$ and $\nu_1$ as $\nu_i = \nu_i - \mu_i + \mu_i$ with $0 \leq \mu_i \leq \nu_i$ denoting the driving part of $\nu_i$ and the unknowns $\nu_0$, $\mu$ and $\mu_1$ (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 $\mu$. Let us remark that if $(\mu_0, \mu_1, \mu)$ solves this parking problem, then $\mu$ solves the corresponding Monge-Kantorovich interpolation problem i.e. minimizes $W_{c_0}(\mu_0,.) + W_{c_1}(.,\mu_1)$ so that the qualitative properties established in Sections \[4\] and \[5\] will be directly applicable 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 \[17\] and \[12\], where the transport between singular measures is used to model the behavior of biological membranes.

In Section \[2\] we consider the general optimization problem \[2.1\] 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

\[1\]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.
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 \(A \subset \mathcal{P}(\mathbb{R}^d)\) we are interested in solving the optimization problem

\[
\inf \left\{ W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1) : \mu \in A \right\}.
\]

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 \(\rho \) is a measure and \(\phi\) is a nonnegative Lebesgue integrable function, by \(\rho \leq \phi\) we mean that \(\rho\) is absolutely continuous and its density, again denoted \(\rho\), satisfies \(\rho \leq \phi\) Lebesgue a.e.; also, all the integrals with no domain of integration explicitly defined are intended on the whole \(\mathbb{R}^d\).

Typical cases for the class \(A\) of admissible choices are:

(i) no constraint, that is \(A = \mathcal{P}(\mathbb{R}^d)\);
(ii) location constraints, that is \(A = \mathcal{P}(K)\) for a nonempty compact subset \(K\) of \(\mathbb{R}^d\);
(iii) density constraints, that is \(A = \{ \rho \in \mathcal{P}_{ac}(\mathbb{R}^d) : \rho \leq \phi \}\) 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

\(\mu_0(x) = \mathbb{1}_{(0,1]}(x), \quad \mu_1(x) = \mathbb{1}_{[3,6)}(x)\).

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

\(c_0(x, y) = (1 - t)|x - y|, \quad c_1(x, y) = t|x - y|\) 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 \(\mu, \mu_0, \mu_1\) (see for instance Chapter 2 of [21]):

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

with the constraints

(i) no additional constraint;
(ii) \(\text{spt } f' \subset [2, 4]\);
(iii) \(f' \leq \theta \mathbb{1}_{[2,4]}\).

Since \(f_1 \leq f_0\), it is easy to see that in the minimization above, one can always assume that \(f_1 \leq f \leq f_0\) and then remove the absolute values and minimize under the constraint that \(f\) is nondecreasing and \(f_1 \leq f \leq f_0\). We then have:
(i) In the absence of constraints, this becomes the problem of finding the Wasserstein median between \( \mu_0 \) and \( \mu_1 \) (see [3] for more on Wasserstein medians). In particular, the optimal solutions \( \mu \) are characterized as follows:

- if \( t > \frac{1}{2} \) (respectively \( t < \frac{1}{2} \)), the unique solution is given by \( \mu = \mu_1 \) (respectively \( \mu = \mu_0 \));
- if \( t = \frac{1}{2} \), any probability \( \mu \) whose distribution function \( f \) is between the two distribution functions \( f_0 \) and \( f_1 \) of \( \mu_0 \) and \( \mu_1 \), in the sense that
  \[
  f_1(x) \leq f(x) \leq f_0(x) \quad \text{for all } x \in \mathbb{R},
  \]
  is a minimizer.

(ii) In the case of the location constraint \( K = [2, 4] \), we observe a similar threshold effect:

- if \( t > \frac{1}{2} \) (respectively \( t < \frac{1}{2} \)), the unique solution is given by \( \mu = \delta_4 \) (respectively \( \mu = \delta_2 \));
- if \( t = \frac{1}{2} \), then any probability measure supported on \( K \) is a solution.

(iii) In the case of density constraint \( \phi(x) := \theta \mathbb{I}_{[2, 4]}(x) \) with \( \theta > \frac{1}{2} \) we have:

- if \( t > \frac{1}{2} \) (respectively \( t < \frac{1}{2} \)), the unique solution is given by \( \mu = \theta \mathbb{I}_{[4-\frac{1}{\theta}, 4]} \) (respectively \( \mu = \theta \mathbb{I}_{[2, 2+1/\theta]} \));
- if \( t = \frac{1}{2} \) 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 consider now strictly convex cost functions: as a prototype we take, with the same measures \( \mu_0 \) and \( \mu_1 \) 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 Definition 2.1 of [21]) of the distribution functions \( f, f_0, f_1 \) as:

\[
\min \left\{ \int_0^1 (1 - t)(g - g_0)^2 + t(g_1 - g)^2 ds : g \text{ nondecreasing} \right\}
\]

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 \( \mu_0 \) to \( \mu_1 \) at time \( t \in (0, 1) \), or equivalently the weighted barycenter. It is given by

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

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

- if \( t \leq \frac{1}{5} \) the support of \( \mu_t \) is contained in \([0, 2]\), hence the optimal solution is \( \delta_2 \);
- if \( \frac{1}{5} < t < \frac{2}{5} \) the optimal solution is \( \mu = (2 - 5t)\delta_2 + \mathbb{I}_{[2,1+5t]} \);
2.2. Reformulation, existence, uniqueness. Let us now come back to the constrained Monge-Kantorovich interpolation problem (2.1) assuming that the measures $\mu_0$ and $\mu_1$ are compactly supported and the costs $c_0$ and $c_1$ are continuous and nonnegative, then by the direct method one directly gets:

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

**Proof.** In both cases, one is left to optimize over probabilities over a fixed compact set, the sum of two Monge-Kantorovich terms which are weak* lsc.

In the unconstrained case where $A := \mathcal{P}(\mathbb{R}^d)$, 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 $(x_0, x_1) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1)$ one has

$$\arg\min_{x \in \mathbb{R}^d} \{c_0(x_0, x) + c_1(x, x_1)\}$$

is nonempty and included in $K$.  \hspace{1cm} (2.2)

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

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

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

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

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

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

**Proposition 2.3.** Assume (2.2).

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

$$T(x_0, x_1) \in \arg\min_{x \in K} \{c_0(x_0, x) + c_1(x, x_1)\} \quad \forall (x_0, x_1) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1).$$
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 \( \mu \) 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 \( \tilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1) \) with \( \pi_{0, \text{piv}, \#} \tilde{\gamma} = \gamma_0 \) and \( \pi_{\text{piv}, 1, \#} \tilde{\gamma} = \gamma_1 \) such that \( \pi_{0,1, \#} \tilde{\gamma} \) is optimal for (2.3) and \( c_0(x_0, x) + c_1(x, x_1) = c(x_0, x_1) \) on \( \text{spt}(\tilde{\gamma}) \) so that \( \mu \) 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 Lemma 7.6 in [23]), there is a plan \( \tilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1) \) with \( \pi_{0, \text{piv}, \#} \tilde{\gamma} = \gamma_0 \) and \( \pi_{\text{piv}, 1, \#} \tilde{\gamma} = \gamma_1 \). Hence, since \( \gamma \) solves (2.3) and \( \pi_{0,1, \#} \tilde{\gamma} \in \Pi(\mu_0, \mu_1) \), we have

\[
\int_{\mathbb{R}^d \times \mathbb{R}^d} c_0 \, d\gamma_0 + \int_{\mathbb{R}^d \times \mathbb{R}^d} c_1 \, d\gamma_1 = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \{ c_0(x_0, x) + c_1(x, x_1) \} \, d\gamma(x_0, x, x_1)
\geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} c \, d\pi_{0,1, \#} \tilde{\gamma} \geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} c \, d\gamma
\geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \{ c_0(x_0, T(x_0, x_1)) + c_1(T(x_0, x_1), x_1) \} \, d\gamma(x_0, x_1)
\geq W_{c_0}(\mu_0, T_{\#} \gamma) + W_{c_1}(T_{\#} \gamma, \mu_1)
\]

which, taking the infimum with respect to \( \gamma_0 \in \Pi(\mu_0, \mu) \) and \( \gamma_1 \in \Pi(\mu, \mu_1) \), enable 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 \( \mu \) 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 \( \tilde{\gamma} \in \Pi(\mu_0, \mu, \mu_1) \) with \( \pi_{0, \text{piv}, \#} \tilde{\gamma} = \gamma_0 \) and \( \pi_{\text{piv}, 1, \#} \tilde{\gamma} = \gamma_1 \), and we then have

\[
\inf (2.3) = \inf (2.1) = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \{ c_0(x_0, x) + c_1(x, x_1) \} \, d\gamma(x_0, x, x_1)
\geq \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} c(x_0, x_1) \, d\gamma(x_0, x, x_1) = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} c \, d\pi_{0,1, \#} \tilde{\gamma}.
\]

Therefore \( \pi_{0,1, \#} \tilde{\gamma} \) is optimal for (2.3) and \( c_0(x_0, x) + c_1(x, x_1) = c(x_0, x_1) \) on \( \text{spt}(\tilde{\gamma}) \). \[ \square \]

In other words, the 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 nonempty 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 \( \mu_0 \) and \( \mu_1 \)). If we further assume that \( \mu_0 \) is absolutely continuous and \( c_0 \) is locally Lipschitz and satisfies the twist condition, i.e. it is differentiable in the first coordinate and for every \( x_0 \in \text{spt}(\mu_0) \)

\[
y \mapsto \nabla_{x_0} c_0(x_0, y)
\]

is injective,
then we claim that
\[ \mu \mapsto W_{c_0}(\mu_0, \mu) \quad \text{is strictly convex}. \quad (2.4) \]
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 Proposition 7.19 in [21], we recall the argument for the sake of completeness. Indeed, thanks to the twist condition and the regularity assumptions on \( c_0 \) and \( \mu_0 \), the optimal transport problem between \( \mu_0 \) and any \( \mu \in \mathcal{A} \) has a unique transport plan induced by a map, see Proposition 1.15 of [21] and the discussion after. Assume that \((\mu, \bar{\mu}, t) \in \mathcal{A} \times \mathcal{A} \times (0, 1)\) are such that
\[ W_{c_0}(\mu_0, (1-t)\mu + t\bar{\mu}) = (1-t)W_{c_0}(\mu_0, \mu) + tW_{c_0}(\mu_0, \bar{\mu}) \]
denoting by \( T \) and \( \bar{T} \) the optimal transport maps between \( \mu_0 \) and \( \mu \) and \( \mu_0 \) and \( \bar{\mu} \) respectively, we have
\[ W_{c_0}(\mu_0, (1-t)\mu + t\bar{\mu}) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c_0 \, d\gamma_t \quad \text{with} \quad \gamma_t := (id, (1-t)T + t\bar{T})_#\mu_0 \]
and since \( \gamma_t \in \Pi(\mu_0, (1-t)\mu + t\bar{\mu}) \), we deduce that \( \gamma_t \) is an optimal plan between \( \mu_0 \) and \((1-t)\mu + t\bar{\mu}\), which, by the twist condition and the absolute continuity of \( \mu_0 \) implies that \( \gamma_t \) is induced by a map so that \( T = \bar{T} \) \( \mu_0 \)-a.e. hence \( \mu = T_#\mu_0 = \bar{T}_#\mu_0 = \bar{\mu} \). 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 (ii) and (iii) in the case of convex costs, i.e. when \( c_0(x, y) = l(|x-y|) \) for \( l : \mathbb{R}_+ \to \mathbb{R}_+ \) strictly convex, increasing and differentiable on \((0, +\infty)\), if we assume \( \mu_0 \) absolutely continuous and for (ii) \( K \cap \text{spt} \mu_0 = \emptyset \), or for (iii) \( \text{spt}(\phi) \cap \text{spt} \mu_0 = \emptyset \) (see [10] or [20] for refinements and weaker conditions). All these arguments for uniqueness of course remain true if we replace the assumptions on \( \mu_0 \) and \( c_0 \) by similar assumptions on \( \mu_1 \) and \( c_1 \).

3. Dual formulations

3.1. Location constraints. Thanks to the coercivity condition (2.2) any solution \( \mu \) 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 the convex minimization problem (2.1) by duality as follows. Given \( \varphi \in C(K) \), define the \( c_0 \)-transform of \( \varphi \), \( \varphi^{c_0} \in C(\text{spt}(\mu_0)) \) by
\[ \varphi^{c_0}_0(x_0) := \min_{x \in K} \{ c_0(x, x) - \varphi(x) \} \quad \forall x_0 \in \text{spt}(\mu_0), \quad (3.1) \]
and similarly define the \( c_1 \)-transform of \( \varphi \), \( \varphi^{c_1} \in C(\text{spt}(\mu_1)) \) by
\[ \varphi^{c_1}_1(x_1) := \min_{x \in K} \{ c_1(x, x) - \varphi(x) \} \quad \forall x_1 \in \text{spt}(\mu_1). \quad (3.2) \]
It follows from Theorem 3 in [3] (where the more general multi-marginal case is considered) that the minimum in (2.1) coincides with the value of the dual:
\[ \sup \left\{ \int \varphi^{c_0}_0 \, d\mu_0 + \int \varphi^{c_1}_1 \, d\mu_1 : \varphi_0, \varphi_1 \in C(K), \varphi_0 + \varphi_1 = 0 \right\}, \quad (3.3) \]
and the supremum in (3.3) is attained. Moreover, if \( \varphi_0 \) and \( \varphi_1 \) solve (3.3), then \( \mu \in \mathcal{P}(K) \) solves (2.1) if and only if \( \varphi_0 \) is a Kantorovich potential between \( \mu_0 \) and \( \mu \).
and \( \varphi_1 \) is a Kantorovich potential (see [23] and [21] for more on Kantorovich duality) between \( \mu \) and \( \mu_1 \) i.e. there exist \((\gamma_0, \gamma_1) \in \Pi(\mu_0, \mu) \times \Pi(\mu, \mu_1)\) such that
\[
\varphi_0(x) + \varphi_0^{c_0}(x_0) = c_0(x_0, x), \quad \forall (x_0, x) \in \text{spt}(\gamma_0),
\]
\[
\varphi_1(x) + \varphi_1^{c_1}(x_1) = c_1(x, x_1), \quad \forall (x, x_1) \in \text{spt}(\gamma_1).
\]
Defining the \( c_0 \)-concave envelope of \( \varphi_0 \) and the \( c_1 \)-concave envelope of \( \varphi_1 \) by
\[
\tilde{\varphi}_0(x) := \min_{x_0 \in \text{spt}(\mu_0)} \{c_0(x_0, x) - \varphi_0^{c_0}(x_0)\},
\]
\[
\tilde{\varphi}_1(x) := \min_{x_1 \in \text{spt}(\mu_1)} \{c_1(x, x_1) - \varphi_1^{c_1}(x_1)\},
\]
one has \( \tilde{\varphi}_0 \geq \varphi_0 \) and \( \tilde{\varphi}_1 \geq \varphi_1 \) with an equality on \( \text{spt}(\mu) \) so that \( \tilde{\varphi}_0 + \tilde{\varphi}_1 \geq 0 \) with an equality on \( \text{spt}(\mu) \).

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 \( \text{spt}(\phi) \) compact (as well as \( \text{spt}(\mu_0) \) and \( \text{spt}(\mu_1) \)). Then the value of (2.1) coincides with the value of its (pre-)dual formulation
\[
\sup_{\varphi_0, \varphi_1 \in C(\text{spt}(\phi))^2} \int \varphi_0^{c_0} \, d\mu_0 + \int \varphi_1^{c_1} \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi \, dx \tag{3.4}
\]
(where \( \varphi_i^{c_i} \) are as in formulae (3.1)-(3.2) with \( K \) replaced by \( \text{spt}(\phi) \)). Moreover, the supremum in (3.4) is attained. If \((\varphi_0, \varphi_1)\) solves (3.4), then \( \mu \) solves (2.1) under the constraint \( \mu \leq \phi \) if and only if there exist \( \gamma_0 \in \Pi(\mu_0, \mu) \) and \( \gamma_1 \in \Pi(\mu, \mu_1) \) such that
\[
\varphi_0(x) + \varphi_0^{c_0}(x_0) = c_0(x_0, x), \quad \forall (x_0, x) \in \text{spt}(\gamma_0),
\]
\[
\varphi_1(x) + \varphi_1^{c_1}(x_1) = c_1(x, x_1), \quad \forall (x, x_1) \in \text{spt}(\gamma_1)
\]
(so that \( \gamma_0 \) and \( \gamma_1 \) are optimal plans and \( \varphi_0 \) and \( \varphi_1 \) are Kantorovich potentials) and
\[
\varphi_0 + \varphi_1 \geq 0 \text{ on } \text{spt}(\phi - \mu), \quad \varphi_0 + \varphi_1 \leq 0 \text{ on } \text{spt}(\mu).
\]

**Proof.** The fact that the concave maximization problem (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
\[
\sup_{(\varphi_0, \varphi_1) \in C(\text{spt}(\phi))^2} \int \varphi_0^{c_0} \, d\mu_0 + \int \varphi_1^{c_1} \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi \, dx = - \inf_{\varphi_0, \varphi_1 \in C(\text{spt}(\phi))^2} F(\varphi_0, \varphi_1) + G(-\varphi_0 - \varphi_1)
\]
where
\[
F(\varphi_0, \varphi_1) := - \int \varphi_0^{c_0} \, d\mu_0 - \int \varphi_1^{c_1} \, d\mu_1, \quad G(\varphi) := \int \max(\varphi, 0) \phi \, dx, \quad \varphi \in C(\text{spt}(\phi)).
\]
Note that \( F \) and \( G \) are convex and continuous for the uniform convergence topology and it is easy to see that \( \sup_{(\varphi_0, \varphi_1) \in C(\text{spt}(\phi))^2} \int \varphi_0^{c_0} \, d\mu_0 + \int \varphi_1^{c_1} \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi \, dx \) is finite (see the proof of existence of a solution to (3.4) below) so that by the Fenchel-Rockafellar duality theorem, we have
\[
\sup_{(\varphi_0, \varphi_1) \in C(\text{spt}(\phi))^2} \int \varphi_0^{c_0} \, d\mu_0 + \int \varphi_1^{c_1} \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi \, dx = \sup_{\mu \in C(\text{spt}(\phi))^*} \left\{ -F^*(\mu, \mu) - G^*(\mu) \right\} = \inf_{\mu \in C(\text{spt}(\phi))^*} \left\{ F^*(\mu, \mu) + G^*(\mu) \right\}
\]
By Kantorovich duality formula (Proposition 1.11 in [21]) we have (also see [5] for details)

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

and

\[
G^*(\mu) = \sup_{\varphi \in C(\text{spt}(\phi))} \int \varphi \, d\mu - \int \max(\varphi, 0) \phi \, dx
\]

and when \( \mu \in \mathcal{P}(\text{spt}(\phi)) \) (which is the case when \( F^*(\mu, \mu) < +\infty \)), the above maximization can be restricted to nonnegative functions \( \varphi \) yielding

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

We thus have

\[
\sup (3.4) = \inf_{\mu \in \mathcal{P}(\text{spt}(\phi)), \mu \leq \phi} W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1)
\]

and (up to extending \( \mu \) by 0 outside \( \text{spt}(\phi) \)) the right-hand side of the previous equality is equivalent to (2.1) in the 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 \((\varphi_0, \varphi_1)\) by \((\varphi_0 + \lambda, \varphi_1 - \lambda)\) where \( \lambda \) is a constant. Moreover, replacing \( \varphi_0 \) and \( \varphi_1 \) by their \( c_0/c_1 \)-concave envelopes defined for every \( x \in \text{spt}(\phi) \) by:

\[
\begin{align*}
\tilde{\varphi}_i(x) &:= \min_{x_0 \in \text{spt}(\mu_0)} \left\{ c_0(x_0, x) - \varphi_0^c(x_0) \right\}, \\
\tilde{\varphi}_i^c(x) &:= \min_{x_1 \in \text{spt}(\mu_1)} \left\{ c_1(x_1, x) - \varphi_1^c(x_1) \right\}
\end{align*}
\]

it is well-known that \( \tilde{\varphi}_i \geq \varphi_i \) and \( \tilde{\varphi}_i^c = \varphi_i^c \) for \( i = 0, 1 \) so that replacing \( \varphi_i \) by \( \tilde{\varphi}_i \) is an improvement in the objective of (3.4), moreover the functions \( \tilde{\varphi}_i \) 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 \( \min_{\text{spt}(\phi)} \varphi_0^n = 0 \) so that \( \varphi_0^n \) 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 \( \varphi_0^n \) we get a uniform lower bound on \( \int (\varphi_1^n)^c \, d\mu_1 \) from which we easily derive a uniform upper bound on \( \varphi_1^n \) thanks to (3.8). To show that \( \varphi_1^n \) is also uniformly bounded from below, we observe that the quantity

\[
\int (\varphi_1^n)^c \, d\mu_1 + \int \min(\varphi_1^n + \varphi_0^n, 0) \phi \, dx
\]

is bounded from below and bounded from above by \( C + (\int \phi \, dx - 1) \min_{\text{spt}(\phi)} \varphi_1^n \) 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 Arzelà-Ascoli theorem.

Let us now look at the optimality conditions which follow from the above duality. If \((\varphi_0, \varphi_1)\) solves (3.4), then \( \mu \) solves (2.1) under the constraint \( \mu \leq \phi \) if and only if

\[
W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1) = \int \varphi_0^c \, d\mu_0 + \int \varphi_1^c \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi.
\]
If \( \gamma_0 \) (respectively \( \gamma_1 \)) is an optimal plan for \( c_0 \) (resp. \( c_1 \)) between \( \mu_0 \) and \( \mu \) (resp. \( \mu \) and \( \mu_1 \)), we thus have

\[
\int \varphi_0^o \, d\mu_0 + \int \varphi_1^c \, d\mu_1 + \int \min(\varphi_0 + \varphi_1, 0) \phi = \int c_0 \, d\gamma_0 + \int c_1 \, d\gamma_1 \\
\geq \int (\varphi_0^o(x_0) + \varphi_0(x)) \, d\gamma_0(x_0, x) + \int (\varphi_1^c(x_1) + \varphi_1(x)) \, d\gamma_1(x, x_1) \\
= \int \varphi_0^o \, d\mu_0 + \int \varphi_1^c \, d\mu_0 + \int (\varphi_0 + \varphi_1) \, d\mu \\
\geq \int \varphi_0^o \, d\mu_0 + \int \varphi_1^c \, d\mu_0 + \int \min(\varphi_0 + \varphi_1, 0) \, d\mu \\
\geq \int \varphi_0^o \, d\mu_0 + \int \varphi_1^c \, d\mu_0 + \int \min(\varphi_0 + \varphi_1, 0) \phi \, dx
\]

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 \( \varphi_0 \) and \( \varphi_1 \) is easily seen to imply (3.5), (3.6), (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 \( \mu \) is optimal for (2.1) under the constraint \( \mu \leq \phi \) and let \( \gamma_0 \) and \( \gamma_1 \) be optimal transport plans. Then, whenever \( x_0, x, x_1 \) are such that \( (x_0, x) \in \text{spt}(\gamma_0) \), \( (x, x_1) \in \text{spt}(\gamma_1) \), \( x \in \text{spt}(\phi - \mu) \), we have

\[
c_0(x_0, x) + c_1(x, x_1) = \min_{y \in \text{spt}(\phi - \mu)} \{c_0(x_0, y) + c_1(y, x_1)\}.
\]

**Proof.** Let \( (\varphi_0, \varphi_1) \) solve (3.4). By construction, for every \( (x_0, x_1, y) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1) \times \text{spt}(\phi) \) one has

\[
c_0(x_0, y) + c_1(y, x_1) \geq \varphi_0^o(x_0) + \varphi_1^c(x_1) + (\varphi_0 + \varphi_1)(y).
\]

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

\[
\min_{y \in \text{spt}(\phi - \mu)} \{c_0(x_0, y) + c_1(y, x_1)\} \geq \varphi_0^o(x_0) + \varphi_1^c(x_1).
\]

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

\[
\varphi_0^o(x_0) + \varphi_1^c(x_1) = c_0(x_0, x) + c_1(x, x_1) \geq \min_{y \in \text{spt}(\phi - \mu)} \{c_0(x_0, y) + c_1(y, x_1)\},
\]

which yields the desired result. \( \Box \)

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 \( \mu_0 \) and \( \mu_1 \) are discrete, that for every \( (x_0, x_1) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1) \), \( c_0(x_0, \cdot) \) and \( c_1(\cdot, x_1) \) are \( C^1 \) and \( M \)-Lipschitz on \( \text{spt}(\phi) \) (for some \( M \) that does not depend on \( x_0 \) and \( x_1 \)) and that the set

\[
\{ x \in \text{spt}(\phi) : \nabla_x c_0(x_0, x) + \nabla_x c_1(x, x_1) = 0 \}
\]

is Lebesgue negligible. Then if \( \mu \) is optimal for (2.1) under the constraint \( \mu \leq \phi \) there exists a measurable subset \( E \) of \( \text{spt}(\phi) \) such that \( \mu = \phi 1_E \).
Proof. Let \((\varphi_0, \varphi_1)\) solve \((3.4)\). As seen in the proof of Proposition \(3.1\), we may assume that, for every \(x \in \text{spt}(\phi)\)
\[
\varphi_0(x) := \min_{x_0 \in \text{spt}(\mu_0)} \{c_0(x_0, x) - \varphi_0^\alpha(x_0)\},
\]
\[
\varphi_1(x) := \min_{x_1 \in \text{spt}(\mu_1)} \{c_1(x, x_1) - \varphi_1^\alpha(x_1)\},
\]
so that \(\varphi_0\) and \(\varphi_1\) are Lipschitz hence differentiable a.e. on \(\text{spt}(\phi)\). Since \(\varphi_0 + \varphi_1 = 0\) on \(\text{spt}(\mu) \cap \text{spt}(\phi - \mu)\), we then have
\[
\nabla \varphi_0 + \nabla \varphi_1 = 0 \quad \text{a.e. on } \{0 < \mu < \phi\}
\]
but if \(\varphi_0\) (respectively \(\varphi_1\)) is differentiable at \(x\) and \((x_0, x) \in \text{spt}(\gamma_0)\) (resp. \((x, x_1) \in \text{spt}(\gamma_1)\)), where \(\gamma_0\) and \(\gamma_1\) 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 \(\mu_i\) \((i = 0, 1)\), a.e. \(x\) such that \(0 < \mu(x) < \phi(x)\) belongs to
\[
\bigcup_{(x_0, x_1) \in A_0 \times A_1} \{x \in \text{spt}(\phi) : \nabla_x c_0(x_0, x) + \nabla_x c_1(x, x_1) = 0\},
\]
which is negligible by assumption. The desired bang-bang conclusion then readily follows. \(\Box\)

Remark 3.4. In some cases, for instance when the costs \(c_0\) and \(c_1\) depend quadratically or more generally as a \(p\)-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 \(\mu_0\) and \(\mu_1\) 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 \(\mu_0\) is \(h_0\), and the Hausdorff dimension of the support of \(\mu_1\) 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:
\[
c_0(x_0, x) := |x_0 - x|^\alpha, \quad c_1(x, x_1) := \lambda|x - x_1|^\alpha, \quad (4.1)
\]
with \(0 < \alpha \leq 1\) 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): \(\mu \in \mathcal{P}(K)\) 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 assumption holds:
- \(\alpha = 1, \lambda > 1\) and the interior of \(K\) is disjoint from \(\text{spt}(\mu_1)\),
- \(\alpha \in (0, 1)\) and the interior of \(K\) is disjoint from \(\text{spt}(\mu_0) \cup \text{spt}(\mu_1)\).

Then any solution \(\mu\) of \((2.1)\) under the constraint \(\mu \in \mathcal{P}(K)\) is supported by \(\partial K\).
We know from Proposition 2.3 that $\mu$ is supported by $T(\text{spt}(\mu_0) \times \text{spt}(\mu_1))$. In particular, if $x \in \text{spt}(\mu)$ 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 $(x_0, x_1) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1)$. In the case $\alpha = 1$, $\lambda > 1$, since $x \not\in x_1$, this is clearly impossible. In the case $\alpha < 1$, our assumption implies that $x \not\in \{x_0, x_1\}$, so that $x$ has to be a critical point of $c_0(x_0, \cdot) + c_1(\cdot, x_1)$. One should have

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

so that $x_0 \not\in x_1$ 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 $d = 1$, $\mu_0 = \delta_0$ and $\mu_1 = \delta_1$, 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 $\partial K$, one may wonder, if $K$ and $\mu_1$ are regular enough, whether these minimizers are absolutely continuous with respect to the $(d - 1)$-Hausdorff measure on $\partial K$, the answer is positive if $\mu_0$ is discrete, i.e. is concentrated on a countable set, $\mu_0(K) = 0$ and $\mu_1$ is absolutely continuous with support disjoint from $\text{int}(K)$ (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

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

is a singleton.

**Proof.** Fix $x_0$, set

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

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

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

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

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

If $\alpha = 1$ 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

$$c_{x_0}(x_1) = |x - x_0| + \lambda |x - x_1| = |x' - x_0| + \lambda |x' - x_1|
= |x' - x_0| + \lambda (|x - x_1| + |x' - x|)
\geq |x' - x_0| + |x' - x| + \lambda |x - x_1|,$$
which is impossible by the triangle inequality, yielding the a.e. single-valuedness of $T_{x_0}$ in this case as well. \hfill \Box

**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 $C^{1,1}$,
- $\mu_0$ is discrete and $\mu_0(K) = 0$,
- $\mu_1$ is absolutely continuous and $\text{int}(K) \cap \text{spt} \mu_1 = \emptyset$.

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

**Proof.** Since $\mu_0$ 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 $\gamma$ between $\mu_0$ and $\mu_1$, 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 $\gamma$ is $\mu_1$, 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 $1/p_{x_0} \mu_1$ 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 $\partial K$ by Lemma 4.3) is absolutely continuous with respect to the $(d-1)$-Hausdorff measure on $\partial K$ 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 $\partial K$ and 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 $\mu_1$ is absolutely continuous, $\partial K$ is a smooth hypersurface and thus Lebesgue negligible and $\mu_1(\text{int}(K)) = 0$), 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 $\mu_1$-a.e. the case (so that $x \notin \{x_0, x_1\}$). By optimality, there exists $\beta \geq 0$ such that

$$\lambda \xi \cdot \hat{x} - x_1 = \beta n(x) = 0,$$

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

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

whose only nonnegative root is

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

so that

$$\lambda \hat{x}_1 - x = \beta_{x_0}(x)n(x) + \hat{x} - x_0.$$
and the right hand side is a Lipschitz function of $x$ thanks to our assumptions ($\partial K$ being $C^{1,1}$ and $x_0$ being at a positive distance from $K$, hence from $x$). Using again that $\tilde{x} - \bar{x}_1$ has norm $\lambda$, this shows that if $x = T_{x_0}(x_1)$ then for some $r \in [0, R]$ with $R = \lambda^{-1}\text{diam} (\text{spt\mu} - K)$ there holds
\[ x_1 = F_{x_0}(r, x) := x + r[\beta_{x_0}(x) n(x) + \tilde{x} - \bar{x}_0]. \]

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 $\mu$ 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 to consider $x_1 \in A_1^\varepsilon$, where
\[ A_1^\varepsilon := \{ x_1 \in \text{spt}(\mu) ; \ d(K, x_1) \geq \varepsilon \} \]
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, 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
\[ H_{x_0}(r, x) = rn(x) + \lambda^{-1} |x - x_0|^{\alpha - 2}(x - x_0). \]
Now, note that
\[ |H_{x_0}(r, x)| = |x_1 - x|^{\alpha - 1}. \]
This shows that
\[ |r| \leq |x_1 - x|^{\alpha - 1} + \lambda^{-1} |x - x_0|^{\alpha - 1} \leq \varepsilon^{\alpha - 1} + \lambda^{-1} \max_{x \in K} |x - x_0|^{\alpha - 1} =: R_\varepsilon(x_0). \]
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 $\varepsilon$) on this set we obtain as soon as $\sigma_{d-1, \partial K}(A) = 0$
\[ \mu_1^{x_0}(T_{x_0}^{-1}(A)) = \mu_1^{x_0}(T_{x_0}^{-1}(A) \setminus K) = \lim_{\varepsilon \searrow 0} \mu_1^{x_0}(T_{x_0}^{-1}(A) \cap A_1^\varepsilon) \leq \lim_{\varepsilon \searrow 0} \mu_1^{x_0}(G_{x_0}([0, R_\varepsilon(x_0)] \times A)) = 0. \]
Thus we can conclude as before that $\mu$ is absolutely continuous.

**Proposition 4.5.** Suppose in addition to the assumptions of Proposition 4.4 that $\mu_0$ has finite support, $\mu_1$ has a bounded density with respect to the $d$-dimensional Lebesgue measure. If $\alpha \in (0, 1)$, further assume that $K \cap \text{spt} \mu_1 = \emptyset$. Then $\mu$ has a bounded density with respect to the $(d - 1)$-Hausdorff measure on $\partial K$. 

\[ \square \]
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 $\partial K$

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

$$\leq \sum_{x_0 \in \operatorname{spt}(\mu_0)} p_{x_0} \mu_1^{x_0}(F_{x_0}([0, R] \times A))$$

$$\leq \sum_{x_0 \in \operatorname{spt}(\mu_0)} \|\mu_1\|_{L^\infty} \mathcal{L}^d(F_{x_0}([0, R] \times A))$$

$$\leq C \operatorname{card}(\operatorname{spt} \mu_0) \|\mu_1\|_{L^\infty} R \sigma_{(d-1), \partial K}(A)$$

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

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

$$\leq \sum_{x_0 \in \operatorname{spt}(\mu_0)} p_{x_0} \mu_1^{x_0}(G_{x_0}([0, R_{\varepsilon_0}(x_0)] \times A))$$

$$\leq \sum_{x_0 \in \operatorname{spt}(\mu_0)} \|\mu_1\|_{L^\infty} \mathcal{L}^d(G_{x_0}([0, R_{\varepsilon_0}(x_0)] \times A))$$

$$\leq C \operatorname{card}(\operatorname{spt} \mu_0) \|\mu_1\|_{L^\infty} R_{\varepsilon_0}(x_0) \sigma_{(d-1), \partial K}(A),$$

where $C$ is a constant that only depends on the $C^{1,1}$ smoothness of $\partial K$ and the maximal with respect to $x_0 \in \operatorname{spt}(\mu_0)$ 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 $\mu_0$ is absolutely continuous and $\mu_1$ is discrete. 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 for 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

$$T_{x_1}(x_0) := \arg\min_{x \in K} \{ |x_0 - x|^{\alpha} + \lambda |x - x_1|^{\alpha} \}$$

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 $C^{1,1}$,
- $\mu_0$ is absolutely continuous and $\operatorname{int}(K) \cap \operatorname{spt} \mu_0 = \emptyset$, 

\[ \cdot \mu_1 \text{ is discrete and } \mu_1(K) = 0. \]

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

**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 \( \gamma \) between \( \mu_0 \) and \( \mu_1 \), which can be written as

\[ \gamma = \sum_{x_1 \in A_1} \mu_0^{x_1} \otimes p_{x_1} \delta_{x_1}, \]

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 \( \gamma \) is \( \mu_0 \), \( \mu_0^{x_1} \) 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 \( \partial K \), which from now on we denote by \( \sigma_{(d-1), \partial K} \). We now fix \( x_1 \in \text{spt}(\mu_1) \) and a Borel subset \( A \) of \( \partial K \) and 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

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

where \( n(x) \) is the outward normal to \( \partial K \) 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 \( \lambda \) 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 \( \beta \). We denote them

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

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

Hence, we have 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 \( R = \text{diam}(\text{spt } \mu_0 - K) \).
Consider now a Borel set $A \subset \partial K$ with $\sigma_{(d-1),\partial K}(A) = 0$. We distinguish the cases where the discriminant $(\lambda n(x) \cdot x - x_1)^2 + 1 - \lambda^2$ is zero or positive

$$A_0 := \left\{ x \in A : (\lambda n(x) \cdot x - x_1)^2 + 1 - \lambda^2 = 0 \right\},$$

$$A_\delta := \left\{ x \in A : (\lambda n(x) \cdot x - x_1)^2 + 1 - \lambda^2 > 0 \right\} \cap \delta > 0 \left\{ x \in A : (\lambda n(x) \cdot x - x_1)^2 + 1 - \lambda^2 \geq \delta \right\}.$$

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

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

$$\leq \mu_{x_1}^0(F_{x_1}^+([0, R] \times A_0)) + \lim_{\delta \searrow 0} \left( \mu_{x_1}^0(F_{x_1}^+([0, R] \times A_\delta)) + \mu_{x_1}^0(F_{x_1}^-([0, R] \times A_\delta)) \right)$$

$$= 0,$$

as required. \hfill \Box

It is unclear whether an $L^\infty$-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 $\{|x| + |y| \leq 1\}$ and consider the distance-like cost of Proposition 4.4 with $\alpha = 1$ and $\lambda > 1$. Take as $\mu_0$ the Lebesgue measure on the disc $B(x_0, r)$ and as $\mu_1$ the Lebesgue measure on the disc $B(x_1, r)$, with $x_0 = (-a, 0)$ and $x_1 = (a, 0)$ as in Figure 1. The optimal pivot measure $\mu$ has in this case a part proportional to the Dirac mass $\delta_{(1,0)}$ and in some cases, when $\lambda$ is large, $a$ is large, and $r$ is small, actually reduces to the Dirac mass $\delta_{(1,0)}$.

\[ \begin{array}{ccc} \mu_0 & K & \mu_1 \\ \end{array} \]

**Figure 1.** A nonsmooth constraint set $K$ may provide a singular optimal pivot measure.

### 4.2. Density constrained solutions are bang-bang

We end this section by observing that in the case of a density constraint $\mu \leq \phi$, 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 $\alpha = 1$, that $\phi \in L^1(\mathbb{R}^d)$ is nonnegative with compact support, that $\int \phi \, dx > 1$, and that both
Proof. Let us start with the case $\alpha = 1$, $\lambda > 1$ and define $A := \{ 0 < \mu < \phi \}$, we then consider (Lipschitz) potentials $\varphi_0$ and $\varphi_1$ as in the proof of Corollary 3.3. A.e. point of $A$ is a differentiability point of $\varphi_0$ and $\varphi_1$, satisfies $\nabla \varphi_0 + \nabla \varphi_1 = 0$ and lies in $\mathbb{R}^d \setminus (\text{spt}(\mu_0) \cup \text{spt}(\mu_1))$. Hence arguing as in the proof of Corollary 3.3 for a.e. $x$ in $A$ one can find $x_0 \in \text{spt}(\mu_0) \setminus \{ x \}$ and $x_1 \in \text{spt}(\mu_1) \setminus \{ x \}$ 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 $x \mapsto |x - x_0|^\alpha$ is Lipschitz only away from $x_0$, it is convenient for $\delta > 0$ to introduce the set

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

on $B_\delta$ the potentials $\varphi_0$ and $\varphi_1$ are Lipschitz and we can find a subset $\tilde{B}_\delta$ of $B_\delta$ with $B_\delta \setminus \tilde{B}_\delta$ negligible such that $\varphi_0$ and $\varphi_1$ are differentiable on $\tilde{B}_\delta$. Consider now for $\varepsilon > 0$

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

and let $\tilde{A}_\varepsilon$ be the subset (of full Lebesgue measure in $A_\varepsilon$ by Lebesgue’s density Theorem) consisting of its points of density 1, i.e.

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

Note that $\tilde{A}_\varepsilon \subset \text{spt}(\phi - \mu)$ and, arguing as before, for a.e. $x \in \tilde{A}_\varepsilon \cap \tilde{B}_\delta$, we can find $(x_0, x_1) \in \text{spt}(\mu_0) \times \text{spt}(\mu_1)$ 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 $\text{spt}(\phi - \mu)$ is included in the level set $f_{x_0, x_1} \geq f_{x_0, x_1}(x)$ and so is $A_\varepsilon$, 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 $x_1 \neq x_0$ and $x$ belongs to $[x_0, x_1]$.

$$e := \overline{x - x_0} = \overline{x_1 - x} = \overline{x_1 - x_0}$$

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

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

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

$$C_{x, e, \kappa} := \{ y \in \mathbb{R}^d : |e \cdot (y - x)| \leq \kappa |(\text{id} - e \otimes e)(y - x)| \}.$$
Hence, for small \( r > 0 \), \( A_{e} \cap B(x, r) \) should lie inside the strict cone \( C_{x,e,\epsilon} \) so that
\[
\lim_{r \to 0^+} \sup_{r > 0} \frac{\mathcal{L}^d(B(x, r) \cap A_{e})}{\mathcal{L}^d(B(x, r))} \leq \lim_{r \to 0^+} \sup_{r > 0} \frac{\mathcal{L}^d(B(x, r) \cap C_{x,e,\epsilon})}{\mathcal{L}^d(B(x, r))} < 1
\]
contradicting the fact that \( x \) is a point of density 1 of \( A_{e} \). This shows that \( A_{e} \cap B_{\delta} \) is negligible, letting \( \delta \to 0^+ \) we find that \( A_{e} \) is negligible and since this is true for every \( \epsilon > 0 \), the desired conclusion follows. \( \square \)

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

We now consider (2.1) in the case of the location constraint \( A = \mathcal{P}(K) \) where \( K \) is a compact convex subset of \( \mathbb{R}^d \) with nonempty interior and \( c_0 \) and \( c_1 \) satisfy the strong convexity and smoothness assumptions:
\[
c_i(x, y) := F_i(y - x), \quad F_i \in C^2(\mathbb{R}^d), \quad \lambda \text{id} \leq D^2 F_i \leq \Lambda \text{id}, \quad i = 0, 1
\]
for some constants \( 0 < \lambda \leq \Lambda \). Since these costs are twisted, (2.1) in the case of the location constraint \( A = \mathcal{P}(K) \) admits a unique solution as soon as \( \mu_0 \) (or \( \mu_1 \)) 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 \( \mu_0 = \delta_{[-2,0]} \), \( \mu_1 \) 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 has an atom at \((1,0)\), an absolutely continuous part, uniform on \( B \cap K \) 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 \( \mu \) in general decomposes into a (nonzero) interior part and a boundary part:
\[
\mu = \mu^{\text{int}} + \mu^{bd} \quad \text{where } \mu^{\text{int}}(A) := \mu(A \cap \text{int}(K)), \quad \mu^{\text{bd}}(A) := \mu(A \cap \partial K)
\]
for every Borel subset \( A \) of \( \mathbb{R}^d \). Regarding \( \mu^{\text{bd}} \), arguing as in Proposition 4.4 one can show that if \( \mu_0 \) is absolutely continuous, \( \mu_1 \) is discrete and \( K \) is of class \( C^{1,1} \), \( \mu^{\text{bd}} \) is absolutely continuous with respect to the \((d-1)\)-Hausdorff measure on \( \partial K \) (and has a bounded density if in addition \( \mu_0 \in L^\infty \) and \( \mu_1 \) is finitely supported, see Proposition 4.5). As for the regularity of \( \mu^{\text{int}} \), we have:

Proposition 5.2. Assume \( c_0 \) and \( c_1 \) are of the form (5.1), that \( \mu_0 \) and \( \mu_1 \) are compactly supported, with \( \mu_0 \in L^\infty \) and that \( K \) is a compact convex subset of \( \mathbb{R}^d \) with nonempty interior. Decomposing the solution \( \mu \) of (2.1) in the case of the location constraint \( A := \mathcal{P}(K) \) as in (5.2), we have \( \mu^{\text{int}} \in L^\infty \) and more precisely (identifying \( \mu^{\text{int}} \) with its density), we have
\[
\|\mu^{\text{int}}\|_{L^\infty} \leq \|\mu_0\|_{L^\infty} 2^d \lambda^{-d} \Lambda^d
\]
where \( \lambda \) and \( \Lambda \) are the positive constants appearing in (5.1).

To establish the \( L^\infty \) bound in (5.3), we shall use a penalization strategy, detailed in the next paragraph, the proof by a standard \( \Gamma \)-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 nonnegative, let us consider
\[
\inf_{\mu \in P(\mathbb{R}^d)} T(\mu) + \int_{\mathbb{R}^d} g\mu \text{ with } T(\mu) := W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu_1, \mu) \tag{5.4}
\]
then we have:

**Proposition 5.3.** Assuming (5.1) and $\mu_0 \in L^\infty$, (5.4) admits a unique solution $\mu_g$.
Moreover $\mu_g$ is absolutely continuous with respect to the Lebesgue measure and its density (still denoted $\mu_g$) satisfies for a.e. $x \in \mathbb{R}^d$, 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 $\lambda$ and $\Lambda$ 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 $g$ in one of the costs considered there), whereas uniqueness is guaranteed by twistedness of the costs and the absolute continuity of $\mu_0$, see Remark 2.4. Also Proposition 2.3 ensures there is some ball $B$ which contains a neighbourhood of $\text{spt}(\mu_g)$. Then, Theorem 3.3 from Pass [10] guarantees that the minimizer $\mu_g$ is absolutely continuous. The optimality condition derived from the dual formulation of (5.4), (see (3.3)) gives the existence of potentials $\varphi_0$ and $\varphi_1$ such that
\[
\varphi_0 + \varphi_1 + g = 0 \text{ on } B \tag{5.6}
\]
and
\[
W_{c_0}(\mu_0, \mu_g) = \int_{\mathbb{R}^d} \varphi_0^{c_0} \mu_0 + \int_{\mathbb{R}^d} \varphi_0 \mu_g, \quad W_{c_1}(\mu_g, \mu_1) = \int_{\mathbb{R}^d} \varphi_1^{c_1} \mu_1 + \int_{\mathbb{R}^d} \varphi_1 \mu_g
\]
so that defining the $c_i$-concave potentials
\[
\widetilde{\varphi}_0(x) := \inf_{x_0 \in \text{spt}(\mu_0)} \{c_0(x_0, x) - \varphi_0^{c_0}(x_0)\},
\]
\[
\widetilde{\varphi}_1(x) := \inf_{x_1 \in \text{spt}(\mu_1)} \{c_1(x, x_1) - \varphi_1^{c_1}(x)\},
\]
on one should have
\[
\varphi_i \leq \widetilde{\varphi}_i \text{ on } B \quad \text{and} \quad \varphi_i = \widetilde{\varphi}_i \text{ on } \text{spt}(\mu_g). \tag{5.7}
\]
Now observe that thanks to (5.1), $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are semi-concave and more precisely
\[
D^2 \widetilde{\varphi}_i \leq \Lambda \text{ id}, \quad i = 0, 1. \tag{5.8}
\]
In particular $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are everywhere superdifferentiable, but on $\text{spt}(\mu_g)$, thanks to (5.6) and (5.7), $\widetilde{\varphi}_0 + \widetilde{\varphi}_1 + g$ is minimal and since $g$ is differentiable this implies that $\varphi_0 + \varphi_1$ is also subdifferentiable on $\text{spt}(\mu_g)$. This readily implies that $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are differentiable on $\text{spt}(\mu_g)$ and that
\[
\nabla \varphi_0 + \nabla \varphi_1 + \nabla g = 0 \text{ on } \text{spt}(\mu_g).
\]
The functions $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_1$ are semi-concave and, by Alexandrov’s Theorem (see Theorem 6.9 in [8]), they are twice differentiable $\mu_g$-a.e.; the minimality of $\widetilde{\varphi}_0 + \widetilde{\varphi}_1 + g$ on $\text{spt}(\mu_g)$ then gives
\[
D^2 \widetilde{\varphi}_0 + D^2 \widetilde{\varphi}_1 + D^2 g \geq 0 \text{ $\mu_g$-a.e.}. \tag{5.9}
\]
The optimal transport $S_0$ for the cost $c_0$ between $\mu_g$ and $\mu_0$ (see Theorem 3.7 in [10]) is then given by
\[
S_0(x) = x - \nabla F_0^*(\nabla \varphi_0(x)), \quad x \in \text{spt}(\mu_g),
\]
where \( F_0^* \) is the Legendre transform of \( F_0 \). The absolute continuity of \( \mu_g \) enables us to use Theorem 4.8 of Cordero-Erausquin [6] to get the existence of a set of full measure for \( \mu_g \) for which one has the Jacobian equation
\[
\mu_g = \mu_0 \circ S_0 \det(\text{id} - D^2 F_0^* (\nabla \varphi_0) D^2 \varphi_0),
\]
where \( D^2 \varphi_0(x) \) is to be understood in the sense of Alexandrov and the matrix \( \text{id} - D^2 F_0^* (\nabla \varphi_0) D^2 \varphi_0 \) which is diagonalizable with real and nonnegative eigenvalues can be rewritten as
\[
\text{id} - D^2 F_0^* (\nabla \varphi_0) D^2 \varphi_0 = D^2 F_0^* (\nabla \varphi_0) (D^2 F_0 (x - S_0(x)) - D^2 \varphi_0(x)).
\]
Together with (5.10), we then have
\[
- D^2 \varphi_0(x) \leq D^2 g(x) + D^2 \varphi_1(x) \leq D^2 g(x) + \Lambda \text{id}
\]
but since \( D^2 F_0 \) is semi-definite positive, this gives for \( \mu_g \) a.e. \( x \):
\[
\mu_g(x) \leq \|\mu_0\|_{L^\infty} \lambda^{-d} \det(D^2 F_0 (x - S_0(x)) - D^2 \varphi_0(x))
\]
by (5.9) and (5.8), we then have
\[
- D^2 \varphi_0(x) \leq D^2 g(x) + D^2 \varphi_1(x) \leq D^2 g(x) + \Lambda \text{id}
\]
This follows.

5.2. Proof of the bound by \( \Gamma \)-convergence. Recall that we have assumed that \( K \)
is a convex compact subset with nonempty interior, for \( \varepsilon > 0 \), setting \( K_\varepsilon := K + \varepsilon B \)(where \( B \) is the unit Euclidean ball of \( \mathbb{R}^d \)); consider the mollifiers \( \eta_\varepsilon = \varepsilon^{-d} \eta(\varepsilon \cdot) \) with \( \eta \) a smooth probability density supported on \( B \), consider the smooth and convex function
\[
g_\varepsilon := \eta_\varepsilon \ast \varepsilon^{-1} d_{K_\varepsilon}^2
\]
where \( d_{K_\varepsilon} \) is the distance to \( K_\varepsilon \). Defining \( T \) as in (5.4) and 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 \( J_\varepsilon \) \( \Gamma \)-converges to \( J \) as \( \varepsilon \to 0^+ \) for the narrow topology. Hence the tight sequence of minimizers of \( J_\varepsilon \), \( \mu_\varepsilon := \mu_{g_\varepsilon} \) converges narrowly to \( \mu \) the minimizer of \( J \) i.e. the solution of (2.1) with \( \mathcal{A} = \mathcal{P}(K) \). Since \( D^2 g_\varepsilon = 0 \) on \( \text{int}(K) \), we deduce from (5.5) that for every open \( \Omega \) such that \( \Omega \subseteq \text{int}(K) \)
\[
\|\mu_\varepsilon\|_{L^\infty(\Omega)} \leq \|\mu_0\|_{L^\infty} 2^d \lambda^{-d} \Lambda^d
\]
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 \( \mathbb{R}^d \), \( \nu_0 \) which represents the distribution of residents in a given area;
• a a compactly supported probability measure on \( \mathbb{R}^d \), \( \nu_1 \) which represents the distribution of services.

The goal is to determine a measure \( \mu \) 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 other ones 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 i.e. \(c_1 \geq c_0\), for instance we may take \(p > 0\) and

\[
c_0(x, y) = |x - y|^p, \quad c_1(x, y) = \lambda |x - y|^p \text{ with } \lambda \geq 1.
\]

(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

\[
F(\mu_0, \mu_1, \mu) = W_{c_1}(\nu_0 - \mu_0, \nu_1 - \mu_1) + W_{c_0}(\mu_0, \mu) + W_{c_1}(\mu, \mu_1).
\]

(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 as:

- no other constraints on the parking density \(\mu\);
- location constraints, that is \(\text{spt } \mu \subset K\), with a compact set \(K \subset \mathbb{R}^d\) a priori given;
- density constraints, that is \(\mu \leq \phi\), for a given nonnegative and integrable function \(\phi\).

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

\[
\inf \int_{\mathbb{R}^d \times \mathbb{R}^d} c_1(x, x_1) \, d\gamma(x, x_1) + \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times K} (c_0(x, x) + c_1(x, x_1)) \, d\tilde{\gamma}(x, x, x, x_1)
\]

subject to the constraints:

\[
\gamma, \tilde{\gamma} \geq 0, \quad \gamma + \pi_{0, 1} \# \tilde{\gamma} \in \Pi(\nu_0, \nu_1).
\]

(6.3)

It is indeed easy to see that the optimal solution to minimizing the functional in \((6.2)\) is given by \(\pi_{0, 1} \# \tilde{\gamma}\). Hence to incorporate a density constraint in the formulation \((6.3)\) one needs to add the constraint \(\pi_{0, 1} \# \tilde{\gamma} \leq \phi\). The problem with location constraint is actually equivalent to a standard optimal transport problem with cost function

\[
C(x_0, x_1) := \min \left\{ c_1(x_0, x_1), \inf_{x \in K} \left\{ c_0(x_0, x) + c_1(x, x_1) \right\} \right\}.
\]

More precisely, consider

\[
\inf_{\beta \in \Pi(\nu_0, \nu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} C(x_0, x_1) \, d\beta(x_0, x_1).
\]

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

\[
V_1 := \{(x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d : c_1(x_0, x_1) = C(x_0, x_1)\}
\]
\[ \gamma := \beta_{|V_1} \text{ and } P : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \text{ (measurable)} \]
\[ P(x_0, x_1) \in \arg\min_{x \in K} \{c_0(x_0, x) + c_1(x, x_1)\}, \]
and
\[ d\gamma(x_0, x_1, x) := \delta_{P(x_0, x_1)}(x) \otimes d\beta_{|\mathbb{R}^d \setminus V_1}(x_0, x_1), \]
then \( \gamma \) and \( \gamma \) are optimal for (6.3).

**Remark 6.1.** Note that the solutions \((\mu_0, \mu, \mu_1)\) to minimizing (6.2), (respectively the solutions \( \gamma \) and \( \gamma \) to (6.3)) are not necessarily probability measures. The optimal common total mass of \( \gamma \), \( \mu_0 \) and \( \mu_1 \) represents the fraction of \( \nu_0 \) 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 \( c_1(x, y) > c_0(x, x) \) 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 \( \mu \) for the parking problem is non-trivial, i.e. \( \mu \neq 0 \).

**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 \( \gamma_1 \) be an associated optimal transport plan. Since \( \nu_0 \neq \nu_1 \), there is \((x_0, x_1) \in \text{spt } \gamma_1 \) 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 \( A_0 \times A_1 \) of \((x_0, x, x_1)\) such that the inequality remains valid, i.e.
\[ c_0(y_0, y) + c_1(y, y_1) < c_1(y_0, y_1), \]
for all \((y_0, y, y_1) \in A_0 \times A_1 \). By possibly choosing a smaller (open) \( A \) we can assume that
\[ \int_A \phi(x) \, dx \leq \gamma_1(A_0 \times A_1), \]
so that there is \( t \in (0, 1] \) such that
\[ \int_A \phi(x) \, dx = t\gamma_1(A_0 \times A_1). \]
But then
\[ W_{c_1}(\nu_0, \nu_1) = \int c_1(y_0, y_1) \, d\gamma_1(y_0, y_1) \]
\[ > \int_{(A_0 \times A_1)^c} c_1(y_0, y_1) \, d\gamma_1(y_0, y_1) + (1 - t) \int_{A_0 \times A_1} c_1(y_0, y_1) \, d\gamma_1(y_0, y_1) \]
\[ + (\gamma_1(A_0 \times A_1))^{-1} \int_{A_0 \times A_1} \int_A (c_0(y_0, y) + c_1(y, y_1)) \phi(y) \, dy \, d\gamma_1(y_0, y_1) \]
\[ \geq W_{c_1}(\nu_0 - \mu_0, \nu_1 - \mu_1) + W_{c_0}(\mu_0, \mu') + W_{c_1}(\mu', \mu_1) \]
where \( \mu' = \phi_{|A_c} \) and \( \mu_0, \mu_1 \) are the marginals of \( t_{|A_0 \times A_1} \). This gives a contradiction and achieves the proof. \( \square \)

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 \( \mu_0, \mu_1, \mu \) are not of unitary mass, that corresponds 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 \( p > 1 \) 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

\[
\lambda(1-\alpha)|x_0|^p + \int (|x-x_0|^p + \lambda|x|^p) \, d\mu
\]

\[
= \lambda|x_0|^p + \int (|x-x_0|^p + \lambda|x|^p - \lambda|x_0|^p) \, d\mu.
\]

Since \( \lambda|x_0|^p \) is fixed we are reduced to minimize 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 \( \mu \), as illustrated above. Setting

\[
f(x) = |x-x_0|^p + \lambda|x|^p - \lambda|x_0|^p
\]

(6.5) it is clear that \( \mu \) has to be concentrated on the set where \( f \leq 0 \). The optimization problem with no other constraints on \( \mu \) has then the trivial solution \( \alpha = 1 \) and \( \mu = \delta_{\text{argmin} f} \) (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 \( \mu \) are present. If we impose \( \text{spt} \mu \subset K \), let \( \bar{x} \in K \) be a minimum point of the function \( f \) in (6.5) over \( K \). If \( f(\bar{x}) < 0 \) then \( \alpha = 1 \) and \( \mu = \delta_{\bar{x}} \) is a solution; if \( f(\bar{x}) \geq 0 \) then \( \alpha = 0 \) and \( \mu = 0 \) is a solution.

We consider now the more realistic case when a density constraint on \( \mu \) is imposed, we take \( \mu \leq 1 \). The optimal measure \( \mu \) for the cost (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.

- 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 \( \alpha \) of people drive to reach a parking, 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 $\lambda = 2$.

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

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

7. **Numerical Simulations**

For the numerical simulation of examples in the case of interpolation between measures (2.1) and the 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]. For an introduction to this rapidly developing subject and convergence results, we refer the reader to [15] and [15].
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

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

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

$$H(P|Q) := \int_{\mathbb{R}^d} \log \left( \frac{dP}{dQ} \right) \, dP \text{ if } P \ll Q,$$

$$+ \infty \text{ otherwise.}$$

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

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

so that (7.1) amounts to minimizing $H(\gamma | R)$ among transport plans between $\mu_0$ and $\mu_1$. As already observed in [2] in Section 3.2 the entropically regularized version of (2.1) becomes for two suitably chosen reference measures $R_0, R_1$

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

The cases [i] with no additional constraint and [iii] with location constraint $K$ can be treated by choosing the reference measures to enforce the support of $\mu$ being included in $K$. Namely we choose

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

where for case [i] we choose $K$ large enough (yet still compact) as before. The resulting Sinkhorn iterations are standard, see for instance Proposition 1 and 2 in [2]. The case [iii] of a density constraint $\phi$ requires performing a suitable projection of the estimated interpolation, as specified in Proposition 4.1 in [19] 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 $\mu \leq \phi$ ($\text{spt } \phi \subset K$) is given by

$$\sup_{\varphi_0, \varphi_1, \psi_0, \psi_1} \left\{ -\sum_{i=0}^{1} \int_{\mathbb{R}^d} \exp(\varphi_i + \psi_i) \, dR_i + \sum_{i=0}^{1} \int_{\mathbb{R}^d} \varphi_i \, d\mu_i + \int_{\mathbb{R}^d} (\psi_0 + \psi_1) \, \phi \, dx : \psi_0 + \psi_1 \leq 0 \right\}.$$

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

$$\exp(\varphi_i(x_i)) = \left( \int_K \exp\left( -\frac{c_i}{\varepsilon} + \varphi_i(x) \right) \, dx \right)^{-1},$$

$$\exp(\psi_i(x)) = \min \{ \mu', \phi \} \left( \int_{\mathbb{R}^d} \exp\left( -\frac{c_i}{\varepsilon} + \varphi_i(x_i) \right) \, d\mu_i(x_i) \right)^{-1}.$$
where $\mu^l$ is the current approximate interpolation which is given by the geometric mean formula (see Proposition 2 of [2]):

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

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

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

where

$$\mathcal{M} = \left\{ \gamma, \tilde{\gamma}_0, \tilde{\gamma}_1, \tilde{\gamma} \in \mathcal{M}_+ (\mathbb{R}^{2d})^3 \times \mathcal{M}_+ (\mathbb{R}^{3d}) : \right.$$ 

$$\gamma + \pi_{0,1} \# \tilde{\gamma} \in \Pi(\nu_0, \nu_1), \pi_{0,piv} \# \tilde{\gamma} = \tilde{\gamma}_0, \pi_{piv,1} \# \tilde{\gamma} = \tilde{\gamma}_1 \}.$$ 

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

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

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

$$\sup_{\phi \geq 0, \psi_0, \psi_1} \left\{ -\int_{\mathbb{R}^{2d}} \exp(\phi_0 + \phi_1) \, dR - \sum_{i=0}^{l} \int_{\mathbb{R}^{2d}} \exp(\phi_i + \psi_i) \, dR_i \right.$$ 

$$+ \sum_{i=0}^{l} \int_{\mathbb{R}^{d}} \varphi_i \, d\nu_i + \int_{\mathbb{R}^{d}} (\psi_0 + \psi_1) \phi \, dx : \psi_0 + \psi_1 \leq 0 \right\}.$$ 

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

$$\exp(\varphi_i^{l+1}(x_i)) = \left( \int \exp \left( -\frac{c_i}{\varepsilon} + \varphi_i^{l+1} \mod 2(x) \right) \, dx + \int \exp \left( -\frac{c_i}{\varepsilon} + \psi_i^{l+1}(x) \right) \, dx \right)^{-1},$$

$$\exp(\psi_i^{l+1}(x)) = \min \{ \mu^l, \phi \} \left( \int \exp \left( -\frac{c_i}{\varepsilon} + \varphi_i^{l+1}(x_i) \right) \, d\mu_i(x_i) \right)^{-1},$$

and $\mu^l$, 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 $\mu_1 = \nu_1 = \delta_{0.5,0.0.5}$, the Dirac at the center of the square and as distribution of residents, we take a symmetric sum of four Dirac masses:

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

We consider power-like costs

$$c_0(x, y) = |x - y|^p, \quad c_1(x, y) = 2c_0(x, y)$$
for several values of $p$ corresponding to concave, linear or convex costs and various constant threshold values for the density constraints $\phi$. In this setting, we know (Corollary 3.3 for $p > 1$ and Proposition 4.9 for $p \leq 1$) 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 $\varepsilon = 5 \times 10^{-4}$. 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 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. 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] Martial Agueh and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM J. Math. Anal.*, 43(2):904–924, 2011.
[2] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative Bregman Projections for Regularized Transportation Problems. *SIAM J. Comput.*, 37(2):A1111–A1138, 2015.
[3] Giuseppe Buttazzo, Guillaume Carlier, and Maxime Laborde. On the Wasserstein distance between mutually singular measures. *Adv. Calc. Var.*, 13(2):141–154, 2020.
[4] Guillaume Carlier, Enis Chenchene, and Katharina Eichinger. Wasserstein medians: Robustness, PDE characterization and numerics. arXiv:2307.01765, 2023.
[5] Guillaume Carlier and Ivar Ekeland. Matching for teams. *Econom. Theory*, 42(2):397–418, 2010.
[6] Dario Cordero-Erausquin. Non-smooth differential properties of optimal transport. In *Recent advances in the theory and applications of mass transport*, volume 353 of *Contemp. Math.*, pages 61–71. Amer. Math. Soc., Providence, RI, 2004.
[7] Marco Cuturi. Sinkhorn Distances: Lightspeed Computation of Optimal Transport. In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013.
[8] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
[9] Alfred Galichon and Bernard Salanié. Matching with Trade-Offs: Revealed Preferences over Competing Characteristics. Technical Report info:doi:10.4133/01466212.2010.521551, Sciences Po, 2010. Publication Title: Sciences Po publications.
[10] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.
[11] Hugo Leclerc, Quentin Mérigot, Filippo Santambrogio, and Federico Stra. Lagrangian discretization of crowd motion and linear diffusion. *SIAM J. Numer. Anal.*, 58(4):2093–2118, 2020.
[12] Luca Lussardi, Mark A. Peletier, and Matthias Röger. Variational analysis of a mesoscale model for bilayer membranes. *J. Fixed Point Theory Appl.*, 15(1):217–240, 2014.
[13] Bertrand Maury, Aude Roudneff-Chupin, and Filippo Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.
[14] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
[15] Marcel Nutz. Introduction to Entropic Optimal Transport, 2022. http://www.math.columbia.edu/~mnutz/docs/EOT_lecture_notes.pdf
[16] Brendan Pass. Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions. *Discrete Contin. Dyn. Syst.*, 34(4):1623–1639, 2014.
[17] Mark A. Peletier and Matthias Röger. Partial localization, lipid bilayers, and the elastica functional. *Arch. Ration. Mech. Anal.*, 193(3):475–537, 2009.
[18] Gabriel Peyré and Marco Cuturi. *Computational Optimal Transport: With Applications to Data Science*, volume 11. 2019. Publisher: Now Publishers, Inc.
[19] Gabriel Peyré. Entropic Approximation of Wasserstein Gradient Flows. *SIAM J. Imaging Sci.*, 8(4):2323–2351, 2015. Publisher: Society for Industrial and Applied Mathematics.
[20] Davide Piazzoli, Filippo Santambrogio, and Paul Pegon. Full characterization of optimal transport plans for concave costs. *Discrete Contin. Dyn. Syst.*, 35(12):6113–6132, 2015.
[21] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
[22] Clément Sarrazin. Lagrangian discretization of variational mean field games. *SIAM J. Control Optim.*, 60(3):1365–1392, 2022.
[23] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
Giuseppe Buttazzo: Dipartimento di Matematica, Università di Pisa  
Largo B. Pontecorvo 5, 56127 Pisa - ITALY  
giuseppe.buttazzo@unipi.it  
http://www.dm.unipi.it/pages/buttazzo/  

Guillaume Carlier: CEREMADE, Université Paris Dauphine and Inria-Mokaplan, PSL  
Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16 - FRANCE  
carlier@ceremade.dauphine.fr  
https://www.ceremade.dauphine.fr/carlier/  

Katharina Eichinger: CEREMADE, Université Paris Dauphine and Inria-Mokaplan, PSL  
Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16 - FRANCE  
eichinger@ceremade.dauphine.fr
Figure 4. concave cost $p = 0.25$
Figure 5.concave cost $p = 0.75$
Figure 6. Linear cost $p = 1$
Figure 7. Convex cost $p = 2$