ON SUDAKOV'S TYPE DECOMPOSITIONS OF TRANSFERENCE PLANS WITH NORM COSTS

STEFANO BIANCHINI AND SARA DANERI

Abstract. We consider the original strategy proposed by Sudakov for solving the Monge transportation problem with norm cost $|\cdot|_{D^*}$

$$\min\bigg\{\int |\mathtt{T}(x)-x|_{D^*}d\mu(x),\ \mathtt{T}:\mathbb{R}^d\to\mathbb{R}^d,\ \nu=\mathtt{T}_\#\mu\bigg\},$$

with μ,ν probability measures in $\mathbb{R}^d$ and μ absolutely continuous w.r.t. $\mathcal{L}^d$. The key idea in this approach is to decompose the optimal transportation problem in regions $Z^k_{\mathfrak{a}}\subset\mathbb{R}^d$ where the construction of an optimal map $\mathtt{T}^k_{\mathfrak{a}}:Z^k_{\mathfrak{a}}\to\mathbb{R}^d$ for the transportation problem in $Z^k_{\mathfrak{a}}\times\mathbb{R}^d$ is simpler than the original problem, and then to obtain $\mathtt{T}$ by piecing together the maps $\mathtt{T}^k_{\mathfrak{a}}$.

The key problems in this kind of approach, when the norm $|\cdot|_{D^*}$ is not strictly convex, are two: on one hand, we need to specify what we means by simpler transportation problems, and on the other hand we have to show that the map $\mathtt{T}^k_{\mathfrak{a}}$ can be constructed.

In this note we hint how these difficulties can be overcome, and that the original idea of Sudakov can be successfully implemented. The analysis requires

(1) the study of directed locally affine partitions $\{Z^k_{\mathfrak{a}},C^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ of $\mathbb{R}^d$, i.e. sets in $Z^k_{\mathfrak{a}}\subset\mathbb{R}^d$ which are relatively open in their affine hull and on which the transport occurs only along directions belonging to a cone $C^k_{\mathfrak{a}}$,
(2) the analysis of transport problem with cost functions which are indicator functions of cones and no potentials can be constructed,
(3) the definition of cyclically connected sets.

The results presented yield a complete characterization of the transport problem, whose straightforward corollary is the solution of the Monge problem in each set $Z^k_{\mathfrak{a}}$ and then in $\mathbb{R}^d$. The strategy is sufficiently powerful to be applied to other optimal transportation problems.

This note contains the results of the paper [6].

1. Introduction

Let $|\cdot|_{D^*}$ be a convex norm in $\mathbb{R}^d$, $\mu,\nu\in\mathcal{P}(\mathbb{R}^d)$ with $\mu\ll\mathcal{L}^d$ and consider the Monge optimal transportation problem

$$\min\bigg\{\int |\mathtt{T}(x)-x|_{D^*}d\mu(x),\ \mathtt{T}:\mathbb{R}^d\to\mathbb{R}^d,\ \nu=\mathtt{T}_\#\mu\bigg\}. \tag{1.1}$$

The main difficulty in this transportation problem is the fact that the function $|\cdot|_{D^*}$ is not strictly convex, so that the optimal transference plans $\pi\in\Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ defined by

$$\begin{gathered}\pi\in\Pi(\mu,\nu):=\Big\{\tilde\pi\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d):(\mathtt{p}_1)_\#\tilde\pi=\mu,(\mathtt{p}_2)_\#\tilde\pi=\nu\Big\}\quad\text{such that}\\ \int|y-x|_{D^*}d\pi(x,y)=\min\bigg\{\int|y-x|_{D^*}d\tilde\pi(y,x),\tilde\pi\in\Pi(\mu,\nu)\bigg\},\end{gathered} \tag{1.2}$$

are not unique, and hence in general they are not concentrated on a graph of a Borel function $\mathtt{T}$. The functions $\mathtt{p}_i:\prod_j X_j\to X_i$ denotes the projection in the i-coordinates in the product space $\prod_j X_j$.

We notice here that by standard theorems in optimal transportation there exists certainly an optimal transference plan, if

$$\mathtt{C}(\mu,\nu):=\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}\mathtt{c}\,d\pi<+\infty. \tag{1.3}$$

Date: May 14, 2013.

In the following we assume that (1.3) holds, otherwise we are in the degenerate situation where every plan $\pi \in \Pi(\mu,\nu)$ and map $\mathtt{T}$, $\nu = \mathtt{T}_{\#}\mu$, are optimal.

The solution to this problem follows mainly two approaches.

The first in chronological order, proposed by Sudakov in [18], is to decompose the transportation problem (1.1) into simpler transportation problem which satisfy additional properties, and for which the existence of the optimal map $\mathtt{T}$ can be proved directly.
An example where this approach has been successfully implemented is the case where $|\cdot|_{D^*}$ is strictly convex: by this we mean that the unit ball $\{x : |x|_{D^*} \leq 1\}$ is strictly convex. In this case one can show that the transport occurs only along one dimensional segments (*optimal rays*) $Z^1_{\mathfrak{a}}$, which intersect at most in a region of negligible Lebesgue measure. Hence the transport problem becomes one dimensional, and some further estimate allows to build optimal transport maps $\mathtt{T}^1_{\mathfrak{a}}$ on each $Z^1_{\mathfrak{a}}$ and then to piece together these maps in order to get a Borel optimal map $\mathtt{T}$, which solves the original problem (1.1).
This idea has been developed in [12], and we will recall it more precisely below in order to explain the difficulties in the general case (i.e. the unit ball of $|\cdot|_{D^*}$ only convex).

A second method is to select among the optimal transference plans $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ a transference plan $\check\pi$ which is also minimizing a secondary cost: more precisely, one selects the (unique) transference plan $\check\pi$ such that

$$\check\pi \text{ is a minimizer of } \inf\bigg\{\int |x-y|^2 d\pi(x,y), \pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)\bigg\}.$$

The key analysis in this strategy is to prove that $\check\pi$ is actually given by a transport plan, i.e. $\check\pi = (\mathbb{I},\check{\mathtt{T}})_\sharp\mu$ for an optimal transport map $\check{\mathtt{T}}$, which clearly satisfies (1.1). This strategy has been successfully adopted in the works of [14, 15], providing a solution to the Monge problem with norm cost.

Other approaches have been used when the norm $|\cdot|_{D^*}$ satisfies additional regularity. In [17], the problem (1.1) has been solved under the assumption that the marginals μ, ν have Lipschitz continuous densities with $\mathcal{L}^d$, and then under milder conditions on the Lebesgue densities of μ, ν in [1, 19]. The case of uniformly convex norms has been solved in [3, 10], and the case of crystalline norms (i.e. where the norm is the supremum of finitely many linear functions) has been solved in [2].

It remains unclear if the original strategy of Sudakov can be successful not only in the case of strictly convex norm. When trying to implement this approach along the same line for the strictly convex norm cost as in [12], we need to define some new concepts, which in the strictly convex case (i.e. when the extremal faces of $|\cdot|_{D^*}$ are 1-dimensional) are trivially satisfied by the decomposition in optimal rays $Z^1_{\mathfrak{a}}$. To better clarify which points have to be studied more carefully, we first give a short overview the approach of [12] to solve the Monge problem for strictly convex norm cost.

1.1. **Sudakov approach in the strictly convex case.** First of all, one considers the potentials ϕ, ψ of the transportation problem: in this special case, one can take $\phi(x) = -\psi(x)$ and ψ to be a $|\cdot|_{D^*}$-Lipschitz function on $\mathbb{R}^d$,

$$\big|\psi(x) - \psi(x')\big| \leq |x-x'|_{D^*}. \tag{1.4}$$

Moreover, being a potential, for all $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ it holds

$$\pi\Big(\Big\{(x,y) : \psi(y) - \psi(x) = |y-x|_{D^*}\Big\}\Big) = 1. \tag{1.5}$$

The condition (1.4) of being $|\cdot|_{D^*}$-Lipschitz yields that if the couple (x,y) satisfies

$$\psi(y) - \psi(x) = |y-x|_{D^*}, \tag{1.6}$$

then for all $0 \leq s \leq t \leq 1$

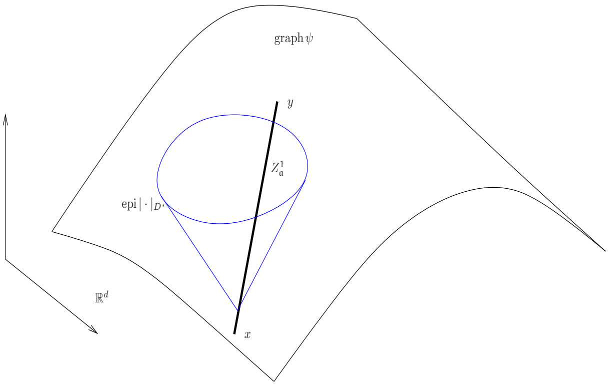
$$\psi(z_t) - \psi(z_s) = |z_t - z_s|_{D^*}, \qquad z_t := (1-t)x + ty.$$

Moreover, if x is a point such that $\nabla\psi(x)$ exists, then if $y \neq x$ satisfies (1.6) one has

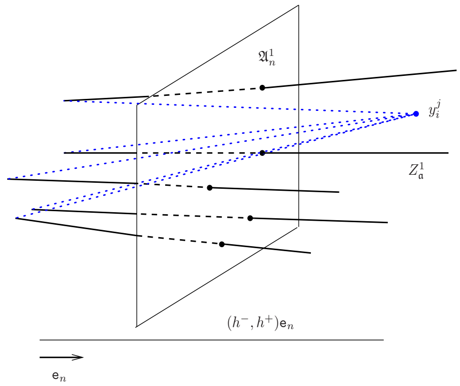
$$\nabla\psi(x) \in \partial|y-x|_{D^*}, \tag{1.7}$$

where $\partial|\cdot|_{D^*}$ is the subdifferential of $|\cdot|_{D^*}$.

By strictly convexity, it follows that the direction of $y-x$ is uniquely defined in the points where $\nabla\psi$ exists. Since one can assume $\mu \perp \nu$ as a consequence of the triangle inequality, it follows that every $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ is concentrated on $\{(x,y), y \neq x\}$, i.e. it moves mass along the optimal rays. In



(A) The construction of optimal rays through the potential ψ and the epigraph of $|\cdot|_{D^*}$.



(B) A set of optimal rays and a simple cone vector field.

particular, up to the μ-negligible set of non-differentiability points of ψ, for any other point passes only one maximal optimal ray $Z^1_{\mathfrak{a}}$: maximal means maximal by set inclusion, and the parameter $\mathfrak{a} \in \mathfrak{A}^1$ is just a parameter of continuum cardinality enumerating the rays $Z^1_{\mathfrak{a}}$.

Typically, one makes a countable partition of the set of optimal rays,

$$\{Z^1_{\mathfrak{a}}\}_{\mathfrak{a}\in\mathfrak{A}^1} = \bigcup_{n\in\mathbb{N}} \{Z^1_{\mathfrak{a}}\}_{\mathfrak{a}\in\mathfrak{A}^1_n},$$

by taking rays $Z^1_{\mathfrak{a}}$ to have direction closed to a given unit vector $\mathrm{e}_n \in \mathbb{S}^{d-1}$ and with projection on $\mathbb{R}\mathrm{e}_n$ containing a given open segment $(h^-, h^+)\mathrm{e}_n$: in that case the set $\mathfrak{A}^1_n$ can be taken to be the range of the map

$$Z^1_{\mathfrak{a}} \mapsto \mathfrak{a} := Z^1_{\mathfrak{a}} \cap \left(\frac{h^+ + h^-}{2}\, \mathrm{e}_n + \big(\mathbb{R}\mathrm{e}_n\big)^{\perp} \right), \tag{1.8}$$

where $V^{\perp}$ is the orthogonal space in $\mathbb{R}^d$ of the subspace $V \subset \mathbb{R}^d$. W.l.o.g. in the following we will restrict the analysis to the set $\{Z^1_{\mathfrak{a}}\}_{\mathfrak{a}\in\mathfrak{A}^1_n}$.

Notice that by (1.6) each ray $Z^1_{\mathfrak{a}}$ is oriented: the direction on which the mass flows on $Z^1_{\mathfrak{a}}$ is given by the extremal face $C^1_{\mathfrak{a}}$ of $|\cdot|_{D^*}$ obtained as the projection on $\mathbb{R}^d$ of the cone

$$\mathrm{span}\big((\mathbb{I}, \psi)(Z^1_{\mathfrak{a}}) - (w, \psi(w))\big) \cap \mathrm{epi}\,|\cdot|_{D^*}, \tag{1.9}$$

for any $w \in \mathrm{int}_{\mathrm{rel}} Z^1_{\mathfrak{a}}$. The above formula is a simple consequence of the definition of $Z^1_{\mathfrak{a}}$ and (1.6). In the strictly convex case these extremal cones are clearly 1-dimensional, and coincides with the exposed faces of epi $|\cdot|_{D^*}$. Given thus $Z^1_{\mathfrak{a}}$, we can speak of initial and final point of $Z^1_{\mathfrak{a}}$ as the points $\mathcal{I}(Z^1_a)$, $\mathcal{E}(Z^1_{\mathfrak{a}})$ in the closure $\bar{Z}^1_{\mathfrak{a}}$ such that

$$\bar{Z}^1_{\mathfrak{a}} = \Big\{ (1-t)\mathcal{I}(Z^1_{\mathfrak{a}}) + t\mathcal{E}(Z^1_{\mathfrak{a}}), t \in [0,1] \Big\}$$

and satisfying $\mathcal{E}(Z^1_{\mathfrak{a}}) - \mathcal{I}(Z^1_{\mathfrak{a}}) \in C^1_{\mathfrak{a}}$.

The next step is to decompose the transport problem in the sets $\{Z^1_{\mathfrak{a}} \times \mathbb{R}^d\}_{\mathfrak{a}\in\mathfrak{A}^1_n}$. This corresponds to the disintegration of the measures μ, $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ on the partitions $\{Z^1_{\mathfrak{a}}\}_{\mathfrak{a}\in\mathfrak{A}^1_n}$, $\{Z^1_{\mathfrak{a}} \times \mathbb{R}^d\}_{\mathfrak{a}\in\mathfrak{A}^1_n}$, respectively:

$$\mu = \int_{\mathfrak{A}^1_n} \mu^1_{\mathfrak{a}}\, dm(\mathfrak{a}), \qquad \pi = \int_{\mathfrak{A}^1_n} \pi^1_{\mathfrak{a}}\, dm(\mathfrak{a}), \tag{1.10}$$

where m is the quotient measure on $\mathfrak{A}^1_n$ obtained through the map (1.8). By the explicit choice (1.8) and the definition of optimal rays by means of (1.4), it is fairly easy to see that the map (1.8) is Borel and thus the disintegrations (1.10) are strongly consistent, which means

$$\mu^1_{\mathfrak{a}}(Z^1_{\mathfrak{a}}) = 1, \qquad \pi^1_{\mathfrak{a}}(Z^1_{\mathfrak{a}} \times \mathbb{R}^d) = 1.$$

One is thus left with the analysis of the reduced transportation problems with marginals $\mu^1_{\mathfrak{a}}$ and $\nu^1_{\mathfrak{a}} := (\mathtt{p}_2)_{\#}\pi^1_{\mathfrak{a}}$.

In general, while from $\nu = (\mathtt{p}_2)_\# \pi$ one deduces that

$$\nu = \int_{\mathfrak{A}^1_n} \nu^1_{\mathfrak{a}} dm(\mathfrak{a}),$$

the above formula is not a disintegration: just take ν to be a Dirac δ-mass and remember that $\mu \ll \mathcal{L}^d$. However in the strictly convex case $\nu^1_{\mathfrak{a}}$ is independent of the particular $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ chosen: indeed, let

$$\nu \llcorner_{\bigcup_{\mathfrak{a}\in\mathfrak{A}^1_n} Z^1_{\mathfrak{a}}} = \int_{\mathfrak{A}^1_n} \tilde{\nu}^1_{\mathfrak{a}} d\tilde{m}(\mathfrak{a}) \tag{1.11}$$

be the disintegration of $\nu \llcorner_{\bigcup_{\mathfrak{a}\in\mathfrak{A}^1_n} Z^1_{\mathfrak{a}}}$. Since $\pi \in \Pi(\mu,\nu)$, it follows that $\tilde{m} \leq m$, and thus we can rewrite it as

$$\nu \llcorner_{\bigcup_{\mathfrak{a}\in\mathfrak{A}^1_n} Z^1_{\mathfrak{a}}} = \int_{\mathfrak{A}^1_n} \hat{\nu}^1_{\mathfrak{a}} dm(\mathfrak{a}), \qquad \tilde{m} = \tilde{\mathtt{h}} m, \ \hat{\nu}^1_{\mathfrak{a}} = \tilde{\mathtt{h}}(\mathfrak{a}) \tilde{\nu}^1_{\mathfrak{a}},$$

with now $\hat{\nu}^1_{\mathfrak{a}}(Z^1_{\mathfrak{a}}) \leq 1$. Hence the mass gap $1 - \hat{\nu}^1_{\mathfrak{a}}(Z^1_{\mathfrak{a}})$ must be concentrated on the end point $\mathcal{E}(Z^1_{\mathfrak{a}})$. Thus we deduce that

$$\nu^1_{\mathfrak{a}} = \hat{\nu}^1_{\mathfrak{a}} + \big(1 - \hat{\nu}^1_{\mathfrak{a}}(Z^1_{\mathfrak{a}})\big)\delta_{\mathcal{E}(Z^1_{\mathfrak{a}})},$$

independently of $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$. As a conclusion, for all $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$, the conditional probabilities $\pi^1_{\mathfrak{a}}$ are concentrated on $Z^1_{\mathfrak{a}} \times Z^1_{\mathfrak{a}} \cup \mathcal{E}(Z^1_{\mathfrak{a}})$, they belong to $\Pi(\mu^1_{\mathfrak{a}}, \nu^1_{\mathfrak{a}})$ and have finite transportation cost w.r.t. the cost

$$\mathtt{c}_{C^1_{\mathfrak{a}}}(y-x) := \mathbb{1}_{C^1_{\mathfrak{a}}}(y-x). \tag{1.12}$$

The last problem one faces is the regularity of the conditional probabilities $\mu^1_{\mathfrak{a}}$. In fact, if $\mu^1_{\mathfrak{a}}$ has no atoms, then the unique transference plan concentrated on a monotone graph in $Z^1_{\mathfrak{a}} \times Z^1_{\mathfrak{a}} \cup \mathcal{E}(Z^1_{\mathfrak{a}})$ is a actually concentrated on a map $\mathtt{T}^1_{\mathfrak{a}}$. In this setting, monotone means monotone w.r.t. the order induced by $C^1_{\mathfrak{a}}$ on $\bar{Z}^1_{\mathfrak{a}}$, and the statement is a well known and simple result for 1-dimensional problems, which can be seen as a particular case of a more general structure result for optimal transportation problems with quadratic cost, see for example [8]. We just observe here that this regularity was a not completely clear point in the original Sudakov paper [18], as pointed out in [2].

In order to prove that $\mu^1_{\mathfrak{a}} \ll \mathcal{H}^1 \llcorner_{Z^1_{\mathfrak{a}}}$, the key idea in [12], adopted also in [7] in order to solve a problem in the calculus of variations, is to approximate the segments $Z^1_{\mathfrak{a}}$, $\mathfrak{a} \in \mathfrak{A}^1_n$, by means of segments $Z^1_{\mathfrak{a},i}$, $\mathfrak{a} \in \mathfrak{A}^1_{n,i}$ and $i \in \mathbb{N}$, with a simpler structure, where the disintegration $\mu = \int \mu^1_{\mathfrak{a},i} dm_i(\mathfrak{a})$ can be done explicitly. Next one shows that the conditional probabilities $\mu^1_{\mathfrak{a},i}$ and image measure m_i satisfy uniform bounds, which are thus inherited by the limits $\mu^1_{\mathfrak{a}}$ and m. A natural choice is to fix the $(d-1)$-dimensional planes

$$P^- := h^- \mathrm{e}_n + (\mathbb{R}\mathrm{e}_n)^\perp \qquad \text{and} \qquad P^+ := h^+ \mathrm{e}_n + (\mathbb{R}\mathrm{e}_n)^\perp,$$

and to consider locally finitely many point $y^j_i \in P^+$, $j = 1, \dots, J_i$.
Next, one solves the transport problem with first marginal $\bar{\mu} \in \mathcal{P}(P^-)$ such that

$$\bar{\mu} \simeq \mathcal{H}^{d-1} \llcorner_{P^-}, \qquad \text{i.e.} \quad \bar{\mu} \ll \mathcal{H}^{d-1} \llcorner_{P^-} \text{ and } \mathcal{H}^{d-1} \llcorner_{P^-} \ll \bar{\mu},$$

and potential ψ_i given by

$$\psi_i(x) := \max_{j=1,\dots,J_i} \big\{ \psi(y^j_i) - |y^j_i - x|_{D^*} \big\}.$$

The optimal rays for ψ_i have a special structure: up to a $\mathcal{H}^{d-1}$-negligible set, they partition P^- into disjoint regions E^j_i such that the direction of the optimal ray for $x \in E^j_i$ is given by

$$\mathtt{d}^j_i(x) = \frac{y^j_i - x}{|y^j_i - x|} \in \mathbb{S}^{d-1}.$$

Such a vector field is called for obvious reasons *finite union of cone vector fields*. The disintegration of the Lebesgue measure on a finite union of cone vector fields is straightforward, since it corresponds to the polar change of coordinates. In particular one has uniform estimates on the conditional probabilities on the slices

$$P^t := t\mathrm{e}_n + (\mathbb{R}\mathrm{e}_n)^\perp, \qquad t \in (h^-, h+),$$

if t is uniformly far from h^+. Since $\psi_i \to \psi$ uniformly in compact sets by Lax formula, it is fairly to see that the optimal rays of P^- converge pointwise to the directions of $Z^1_{\mathfrak{a}}$, up to a $\mathcal{H}^{d-1}$-negligible set, and one can pass these estimates to the disintegration on $Z^1_{\mathfrak{a}}$, yielding in particular the $\mathcal{H}^1$-absolute continuity of $\mu^1_{\mathfrak{a}}$.

Hence one concludes that the transport problem in $Z^1_{\mathfrak{a}} \times Z^1_{\mathfrak{a}} \cup \mathcal{E}(Z^1_{\mathfrak{a}})$ has an optimal map $T^1_{\mathfrak{a}}$. The measurability problem of proving that there exists a map $\mathtt{T}$ such that $\mathtt{T}\llcorner_{Z^1_{\mathfrak{a}}} = \mathtt{T}^1_{\mathfrak{a}}$ is done by means of a measurable selection argument, whose most general statement can be found in [5]. We notice also that from $\mu^1_{\mathfrak{a}} \ll \mathcal{H}^1\llcorner_{Z^1_{\mathfrak{a}}}$ one can takes $Z^1_{\mathfrak{a}}$ to be an open segment: in the following the structure $\{Z^1_{\mathfrak{a}}, C^1_{\mathfrak{a}}\}_{\mathfrak{a}\in\mathfrak{A}^1}$, with $Z^1_{\mathfrak{a}}$ relatively open in his affine hull and $C^1_{\mathfrak{a}}$ extremal face of $|\cdot|_{D^*}$, will be called a *directed locally affine partition*, in this particular case made of *directed relatively open segments*.

1.2. **The general case and the main results.** If one tries now to repeat the above construction by using the set of optimal ray given by (1.6), it follows that (1.7) select only the *exposed faces* of $|\cdot|_{D^*}$. However, the set of exposed faces of a convex body is not a partition of the convex body (and a posteriori it does not generate a partition on $\mathbb{R}^d$): the natural partition is made of the *relative interior* of the extremal faces. Hence one needs stronger requirements on the sets of the directed partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{\substack{k=1,\dots,d-1\\ \mathfrak{a}\in\mathfrak{A}^k}}$, where now k denotes the affine dimension of $Z^k_{\mathfrak{a}}$ and $C^k_{\mathfrak{a}}$ is the extremal cone of $|\cdot|_{D^*}$ corresponding to the directions of the optimal rays of $x \in Z^k_{\mathfrak{a}}$.

One is thus led to introduce the definition of regular transport set $\mathcal{R}\theta_\psi$ (the notation will be clear when we study a more general transport problem): these are the points x such that

(1) the set of directions $\mathcal{D}^+\theta_\psi(x)$ of the optimal rays starting in x is convex in $\mathbb{S}^{d-1}$, and the same for the set of directions $\mathcal{D}^-\theta_\psi(x)$ of the optimal rays arriving in x,
(2) the two sets $\mathcal{D}^+\theta_\psi(x)$, $\mathcal{D}^-\theta_\psi(x)$ coincide,
(3) there are points x', x'' such that
$$\frac{x-x'}{|x-x'|} \in \mathrm{int}_{\mathrm{rel}}\mathcal{D}^-\theta_\psi(x), \qquad \frac{x''-x}{|x''-x|} \in \mathrm{int}_{\mathrm{rel}}\mathcal{D}^+\theta_\psi(x)$$
and Points (1-2) hold for x', x'' too.

If one uses the notation
$$\partial^+\psi(x) := \big\{y : \psi(y) = \psi(x) + |y-x|_{D^*}\big\}$$
so that
$$\mathcal{D}^+\theta_\psi(x) = \bigg\{\frac{y-x}{|y-x|}, y \in \partial^+\psi(x) \setminus \{x\}\bigg\},$$
then the sets $Z^k_{\mathfrak{a}}$, $C^k_{\mathfrak{a}}$ are now determined by
$$x \in Z^k_{\mathfrak{a}} \quad \Longrightarrow \quad \begin{cases} Z^k_{\mathfrak{a}} = \mathcal{R}\theta_\psi \cap \mathtt{p}_{\mathbb{R}^d}\Big(\mathrm{graph}\,\psi \cap \Big((x,\psi(x)) + \mathrm{span}\big\{(\mathbb{I},\psi)(\partial^+\psi(x)) - (x,\psi(x))\big\}\Big)\Big), \\ C^k_{\mathfrak{a}} = \mathtt{p}_{\mathbb{R}^d}\Big(\mathrm{epi}\,|\cdot|_{D^*} \cap \Big(\mathrm{span}\big\{(\mathbb{I},\psi)(\partial^+\psi(x)) - (x,\psi(x))\big\}\Big)\Big). \end{cases}$$

It is fairly easy to see that the sets $Z^k_{\mathfrak{a}}$ are relatively open in their affine hull, and that $C^k_{\mathfrak{a}}$ are extremal faces of $|\cdot|_{D^*}$: the latter statement follows from the $|\cdot|_{D^*}$-Lipschitz regularity of ψ and Points (1-3) above. Recall that the index k denotes the affine dimension of $Z^k_{\mathfrak{a}}$, which coincides with the linear dimension of $C^k_{\mathfrak{a}}$, while $\mathfrak{a} \in \mathfrak{A}^k$ is an index of continuum cardinality.

Since the definition of "good points", i.e. of the set $\mathcal{R}\theta_\psi$, is definitely more complicated than in the strictly convex case, it is perfectly understandable that the proof of the μ-negligibility of the set $\mathbb{R}^d \setminus \mathcal{R}$ and the proof the regularity of the conditional probabilities of the disintegration of μ w.r.t. the partition $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ are considerably more intricate, even if they are still based on the cone vector field approximation property. The main reference for the approach used in this part is [13], where it is shown a regularity property of the conditional probabilities of the disintegration of the surface measure on the partition induced by the relative interior of the extremal faces.

The key observation is that the set of points which do not belongs to $\mathcal{R}\theta_\psi$ are either *initial point* for the *forward regular locally affine partition* $\{Z^{k,+}_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ or *final points* for the *backward regular locally affine partition* $\{Z^{k,-}_{\mathfrak{a}}\}_{k,\mathfrak{a}}$. For the forward regular locally affine partition $\{Z^{k,+}_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, the main differences w.r.t. the regular partition are that Point 1 of the definition of regular transport set at Page 5 is required

only for $\mathcal{D}^+\theta_\psi$ and Point 3 only for x', while Point 2 is meaningless in this case; a completely similar requirements for $\mathcal{D}^-\theta_\psi$ and x'' is done for the backward regular locally affine partition.
Hence, one proves the finite cone approximation property for the forward partition and for the backward partition, and concludes that the set of initial/final points is $\mathcal{L}^d$-negligible, and also that the partition $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ is regular, i.e. $\mu^k_{\mathfrak{a}} \ll \mathcal{H}^k \llcorner Z^k_{\mathfrak{a}}$.

The procedure explained above is also presented in the first part of [16].

At this point, we deduce the following Theorem 1.1, which is the analog of the decomposition in the strictly convex norm case: the statement includes also the points which do not belong to any optimal ray, and in that case the dimension k of the elements of the directed locally affine partition they belongs to, is $k = 0$, as well as $C^0_{\mathfrak{a}} = \{0\}$. Moreover, by a careful analysis of the transport problems in the graph of ψ, one can also check that the mass starting in $Z^k_{\mathfrak{a}}$ is moved in $Z^k_{\mathfrak{a}}$ or in the end points $\underset{k,\mathfrak{a}}{\cup} \mathcal{E}(Z^k_{\mathfrak{a}})$.

For notational convenience, if $\mathtt{c} : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ is a Borel cost function, we use the notation

$$\Pi^{\mathrm{f}}_{\mathtt{c}}(\mu,\nu) := \left\{ \pi \in \Pi(\mu,\nu) : \int \mathtt{c}\pi < \infty \right\}.$$

For a locally affine partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, we will write

$$\mathtt{c}_{C^k_{\mathfrak{a}}}(x,y) := \begin{cases} \mathbb{1}_{C^k_{\mathfrak{a}}}(y-x) & x \in Z^k_{\mathfrak{a}}, \\ +\infty & \text{otherwise}, \end{cases} \tag{1.13}$$

and, since we will often write the graph of a directed locally affine partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ as

$$\mathbf{D} := \left\{ (k,\mathfrak{a},z,C^k_{\mathfrak{a}}) : k \in \{0,\dots,d\},\, \mathfrak{a} \in \mathfrak{A}^k,\, z \in Z^k_{\mathfrak{a}} \right\} \subset \bigcup_{k=0}^d \{k\} \times \mathfrak{A}^k \times \mathbb{R}^d \times \mathcal{C}(k,\mathbb{R}^d),$$

we will use also the notation

$$\mathtt{c}_{\mathbf{D}}(x,y) := \begin{cases} \mathbb{1}_{C^k_{\mathfrak{a}}}(y-x) & \exists\, k,\mathfrak{a} \text{ s.t. } x \in Z^k_{\mathfrak{a}}, \\ +\infty & \text{otherwise}. \end{cases} \tag{1.14}$$

Notice that for costs $\mathtt{c}$ of the form (1.13), (1.14) one has clearly $\Pi^{\mathrm{opt}}_{\mathtt{c}}(\mu,\nu) = \Pi^{\mathrm{f}}_{\mathtt{c}}(\mu,\nu)$, since the only values of $\mathtt{c}$ are 0, ∞.

Theorem 1.1. *Let $\mu,\nu \in \mathcal{P}(\mathbb{R}^d)$ with $\mu \ll \mathcal{L}^d$ and let $|\cdot|_{D^*}$ be a convex norm in $\mathbb{R}^d$. Then there exists a locally affine directed partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{\substack{k=0,\dots,d \\ \mathfrak{a}\in\mathfrak{A}^k}}$ in $\mathbb{R}^d$ with the following properties:*

(1) for all $\mathfrak{a} \in \mathfrak{A}^k$ the cone $C^k_{\mathfrak{a}}$ is a k-dimensional extremal face of $|\cdot|_{D^*}$;

(2) $\mathcal{L}^d\left(\mathbb{R}^d \setminus \bigcup_{k,\mathfrak{a}} Z^k_{\mathfrak{a}}\right) = 0$;

(3) the disintegration of $\mathcal{L}^d$ w.r.t. the partition $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, $\mathcal{L}^d \llcorner_{\cup_{k,\mathfrak{a}} Z^k_{\mathfrak{a}}} = \int v^k_{\mathfrak{a}}\, d\eta(k,\mathfrak{a})$, satisfies

$$v^k_{\mathfrak{a}} \simeq \mathcal{H}^k \llcorner_{Z^k_{\mathfrak{a}}};$$

(4) for all $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$, the disintegration $\pi = \int \pi^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ w.r.t. the partition $\{Z^k_{\mathfrak{a}} \times \mathbb{R}^d\}_{k,\mathfrak{a}}$ satisfies

$$\pi^k_{\mathfrak{a}} \in \Pi^{\mathrm{f}}_{\mathtt{c}_{C^k_{\mathfrak{a}}}}\big(\mu^k_{\mathfrak{a}}, (\mathtt{p}_2)_\# \pi^k_{\mathfrak{a}}\big),$$

where $\mu = \int \mu^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ is the disintegration w.r.t. the partition $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, and moreover

$$(\mathtt{p}_2)_\# \pi^k_{\mathfrak{a}} \left(Z^k_{\mathfrak{a}} \cup \left(\mathbb{R}^d \setminus \bigcup_{(k',\mathfrak{a}') \neq (k,\mathfrak{a})} Z^{k'}_{\mathfrak{a}'} \right) \right) = 1.$$

If also $\nu \ll \mathcal{L}^d$, then for all $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$

$$(\mathtt{p}_2)_\# \pi^k_{\mathfrak{a}} = \nu^k_{\mathfrak{a}}$$

where $\nu = \int \nu^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ *is the disintegration w.r.t. the partition* $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$*, and the converse of Point* (4) *holds:*

$$\pi^k_{\mathfrak{a}} \in \Pi^{\mathrm{f}}_{\mathtt{c}_{C^k_{\mathfrak{a}}}}(\mu^k_{\mathfrak{a}}, \nu^k_{\mathfrak{a}}) \quad \Longrightarrow \quad \pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu).$$

A remark is in order here: in Point (4), the conditional second marginals $(\mathtt{p}_2)_\#\pi^k_{\mathfrak{a}}$ are independent on the potential ψ but *depends* on the particular transference plan which we are decomposing. This can be seen with elementary examples. Hence the analysis from now on will be done in the class of transference plans π' which have the same conditional marginals: in fact, by inspection one sees that the next steps of the proof (the key variation is in the construction the subsequent decompositions) are different when changing the conditional marginals.
The notation $\Pi(\mu, \{\nu^k_{\mathfrak{a}}\})$ denotes the set of transference plans π such that the marginals of the conditional probabilities $\pi^k_{\mathfrak{a}}$ w.r.t. the partition $\{Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ satisfy $(\mathtt{p}_2)_\#\pi^k_{\mathfrak{a}} = \nu^k_{\mathfrak{a}}$.

As shown in [11], the decomposition above is not refined enough to give immediately the existence of transport maps in each of the $Z^k_{\mathfrak{a}} \times \mathbb{R}^d$, unless one uses techniques similar to the one adopted in [9] (which means that we are not really simplifying the problem). The key fact is that in general the transport problem in $\Pi(\mu^k_{\mathfrak{a}}, (\mathtt{p}_2)_\#\pi^k_{\mathfrak{a}})$ with cost $\mathtt{c}_{C^k_{\mathfrak{a}}}$ given by (1.13) does not have a potential $\phi^k_{\mathfrak{a}}$ (see the final example of [11]). In order to explain this point, recall the standard formula for constructing a potential ϕ: if Γ is a $\mathtt{c}$-cyclically monotone carriage for π, $\mathtt{c} : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ being the cost function, and $(x_0,y_0) \in \Gamma$, then

$$\phi(x) := \inf \left\{ \sum_{i=0}^{I} \mathtt{c}(x_{i+1},y_i) - \mathtt{c}(x_i,y_i) : I \in \mathbb{N}, (x_i,y_i) \in \Gamma, x_{I+1} = x \right\}, \tag{1.15}$$

yields a potential ϕ provided it is μ-a.e. finite. The sequence of points

$$(x_0,y_0),(x_1,y_0),(x_1,y_1),(x_2,y_1),\dots,(x_i,y_i),(x_{i+1},y_i),(x_{i+1},y_{i+1}),\dots,(x,y_I), \qquad (x_i,y_i) \in \Gamma,$$

is an *axial path* (or Γ-*axial path* if we need to emphasize the dependence w.r.t. Γ), and we say that the axial path has *finite cost* if $\mathtt{c}(x,y) < \infty$ for all couple (x,y) in the axial path: since we can assume that $\Gamma \subset \{\mathtt{c} < \infty\}$ under (1.3), this condition is equivalent to $\mathtt{c}(x_{i+1},y_i), \mathtt{c}(x,y_I) < \infty$. It is a well know fact that if μ-a.a. points belong to an axial path starting from and ending in (x_0,y_0) (which will be called a *(*Γ*-)axial cycle* or *(*Γ*-)cycle* for brevity), then formula (1.15) yields a μ-a.e. finite potential ϕ. Its dual ψ can be obtained by the formula $\psi(y) = \mathtt{c}(x,y) - \phi(x)$, and it turns out that it is finite and independent on x for ν-a.e. $y \in \mathbb{R}^d$.

It becomes then natural to ask for a directed locally affine partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ that the sets $Z^k_{\mathfrak{a}}$ are contained in an axial cycle up to a $\mu^k_{\mathfrak{a}}$-negligible set: the cost in each $Z^k_{\mathfrak{a}}$ is the *cone cost* given by (1.13), namely since now $x \in Z^k_{\mathfrak{a}}$

$$\mathtt{c}(x,y) = \mathbb{1}_{C^k_{\mathfrak{a}}}(y-x). \tag{1.16}$$

If this cyclical connectedness condition is verified, then as in the strictly convex case one can construct a transport map $\mathtt{T}^k_{\mathfrak{a}}$ by standard methods as is [2] and then the optimal $\mathtt{T}$ is obtained by piecing together the maps $\mathtt{T}^k_{\mathfrak{a}}$.

We are now ready to state the key step of the procedure, namely Theorem 1.2: it allows to construct a locally directed affine subpartition $\{\check{Z}^\ell_{\mathfrak{b}}, \check{C}^\ell_{\mathfrak{b}}\}_{\substack{\ell=0,\dots,d\\ \mathfrak{b}\in\mathfrak{B}^\ell}}$ to a directed locally affine partition $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{\substack{k=0,\dots,d\\ \mathfrak{a}\in\mathfrak{A}^k}}$ such that the sets which do not lower their affine dimensions (i.e. for which $\check{Z}^\ell_{\mathfrak{b}} \subset Z^k_{\mathfrak{a}}$ and $\ell = k$) are indecomposable, in the sense that $\mu^k_{\mathfrak{a}}$-a.a. points can be connected by a cycle.
We say that sets are $\Pi^{\mathrm{f}}_{\mathtt{c}_{\mathbf{D}}}(\mu, \{\check{\nu}^\ell_{\mathfrak{b}}\})$-cyclically connected, in the sense that for all $\pi' \in \Pi^{\mathrm{f}}_{\mathtt{c}_{\mathbf{D}}}(\mu, \{\check{\nu}^\ell_{\mathfrak{b}}\})$ and cyclically monotone carriage Γ' of π', each two points of these "indecomposable sets" are Γ'-cyclically connected (or connected by a Γ'-cycle), up to a $\mu^k_{\mathfrak{a}}$-negligible set depending on Γ'. Notice that we need to specify the second marginals of the transference plans, as it should be clear from the discussion after Theorem 1.1.

Theorem 1.2. *Let* $\{Z^k_{\mathfrak{a}}, C^k_{\mathfrak{a}}\}_{\substack{k=0,\dots,d\\ \mathfrak{a}\in\mathfrak{A}^k}}$ *be a regular directed locally affine partition in* $\mathbb{R}^d$ *and let* μ, ν *be probability measures in* $\mathcal{P}(\mathbb{R}^d)$ *such that* $\mu \ll \mathcal{L}^d$ *and* $\Pi^{\mathrm{f}}_{\mathtt{c}_{\mathbf{D}}}(\mu,\nu) \neq \emptyset$.

Then, for all fixed $\check\pi \in \Pi^{\mathrm{f}}_{\mathtt{c_D}}(\mu,\nu)$, there exists a directed locally affine subpartition $\{\check Z^\ell_{\mathfrak b}, \check C^\ell_{\mathfrak b}\}_{\substack{\ell=0,\dots,d\\ \mathfrak b\in\mathfrak B^\ell}}$ of $\{Z^k_{\mathfrak a}, C^k_{\mathfrak a}\}_{k,\mathfrak a}$, up to a μ-negligible set $N_{\check\pi}$, such that

$$\big\{\check Z^\ell_{\mathfrak b}, \check C^\ell_{\mathfrak b}\big\}_{\ell,\mathfrak b} \quad \textit{is regular,}$$

and setting $\check\nu^\ell_{\mathfrak b} := (\mathtt p_2)_\# \check\pi^\ell_{\mathfrak b}$, where $\check\pi^\ell_{\mathfrak b}$ is the conditional probability on the partition $\{\check Z^\ell_{\mathfrak b} \times \mathbb R^d\}_{\ell,\mathfrak b}$, then the sets

$$\Big\{\check Z^\ell_{\mathfrak b} : \check Z^\ell_{\mathfrak b} \subset Z^\ell_{\mathfrak a} \text{ for some } \mathfrak a \in \mathfrak A^\ell, \ell = 1,\dots,d\Big\} \tag{1.17}$$

form a $\Pi^{\mathrm f}_{\mathtt{c_D}}(\mu, \{\check\nu^\ell_{\mathfrak b}\})$-cyclically connected partition.

Since the subpartition

$$\big\{\check Z^\ell_{\mathfrak b}, \check C^\ell_{\mathfrak b}\big\}_{\substack{\ell=0,\dots,d-1\\ \mathfrak b\in\check{\mathfrak B}^\ell}} \quad \text{such that if } \check Z^\ell_{\mathfrak b} \subset Z^k_{\mathfrak a} \text{ then } \ell < k \text{ (equivalently neglecting the sets of (1.17))}$$

is a regular directed locally affine partition, and as a subpartition of $\{Z^k_{\mathfrak a}, C^k_{\mathfrak a}\}_{k,\mathfrak a}$ the index ℓ is decreasing of at least 1 in each $Z^k_{\mathfrak a}$, by a finite iterative argument we obtain the following corollary.

Corollary 1.3. *Let $\{Z^k_{\mathfrak a}, C^k_{\mathfrak a}\}_{\substack{k=0,\dots,d\\ \mathfrak a\in\mathfrak A^k}}$ be a regular directed locally affine partition in $\mathbb R^d$ and let $\mu \ll \mathcal L^d$, $\nu \in \mathcal P(\mathbb R^d)$ such that $\Pi^{\mathrm f}_{\mathtt{c_D}}(\mu,\nu) \neq \emptyset$.*

Then, for all $\check\pi \in \Pi^{\mathrm f}_{\mathtt{c_D}}(\mu,\nu)$ there exists a directed locally affine subpartition $\{\check Z'^{,\ell}_{\mathfrak b}, \check C'^{,\ell}_{\mathfrak b}\}_{\substack{\ell=0,\dots,d\\ \mathfrak b\in\mathfrak B^\ell}}$ of $\{Z^k_{\mathfrak a}, C^k_{\mathfrak a}\}_{k,\mathfrak a}$, up to a μ-negligible set $N'_{\check\pi}$, such that

$$\{\check Z'^{,\ell}_{\mathfrak b}, \check C'^{,\ell}_{\mathfrak b}\}_{\ell,\mathfrak b} \quad \textit{is regular,}$$

and if $\bar\nu'^{,\ell}_{\mathfrak b} := (\mathtt p_2)_\# \check\pi'^{,\ell}_{\mathfrak b}$, where $\check\pi'^{,\ell}_{\mathfrak b}$ is the conditional probability on the partition $\{\check Z'^{,\ell}_{\mathfrak b} \times \mathbb R^d\}_{\ell,\mathfrak b}$, then each set $\check Z'^{,\ell}_{\mathfrak b}$ is $\Pi^{\mathrm f}_{\mathtt{c_D}}(\mu, \{\bar\nu'^{,\ell}_{\mathfrak b}\})$-cyclically connected, for all ℓ, $\mathfrak b$.

An important point in the proof of Theorem 1.2 is that we are now working with cost of the form (1.14), and we cannot rely on the existence of a potential ϕ (or its dual ψ) to prove the regularity of the disintegration, as done in the proof of in Theorem 1.1.

The potential ψ is then replaced by (a Borel extension of) the axial preorder: the construction works as follows. Fixed a cyclically monotone carriage Γ, we say that $x \preccurlyeq_\Gamma y$ is there is a Γ-axial path of finite cost connecting y to x. In this way one obtains only a preorder (i.e. only a transitive relation), but there is a countable procedure to construct a Borel linear preorder (i.e. a preorder with Borel graph and such that every two points can be compared), with the property that the equivalence classes which do not lower their affine dimension are Γ-cyclically connected up to a μ-negligible set. The preorder will be denoted in the following by $\preccurlyeq_{\Gamma,\mathtt W}$, where $\mathtt W$ is related to the countable procedure to construct the Borel preorder.
Unfortunately, by changing Γ the Borel preorder $\preccurlyeq_{\Gamma,\mathtt W}$ varies. Hence we need to use an abstract result on measure theory [4] , assuring that there is a minimal Borel linear preorder: for this one the sets which do not lower the dimension are indecomposable for all carriages Γ of a fixed π. Next, one can apply the uniqueness results proved in [4] for optimal transportation problems on linear preorders with Borel graph , to obtain that the indecomposability property holds also w.r.t. all carriages Γ of optimal plans $\pi \in \Pi^{\mathrm{opt}}_{\mathtt{c_D}}(\mu,\nu)$.

Another advantage of the Borel linear preorder is that by the uniqueness theorem stated in [4] one can prove the finite cone vector field approximation properties for the subpartition. The procedure is similar to the procedure followed in the case a potential ψ is present: the convergence of the optimal rays is now due to the uniqueness of a suitable transference plan.

Applying Corollary 1.3 to the regular locally affine partition given by Theorem 1.1, one obtains immediately the following result. As in the case of Theorem 1.1, the second part of Point (4) of the next theorem is a consequence of the precise analysis of the regions where the mass transport occurs.

Theorem 1.4. *Let μ, $\nu \in \mathcal P(\mathbb R^d)$ with $\mu \ll \mathcal L^d$ and let $|\cdot|_{D^*}$ be a convex norm in $\mathbb R^d$. Then, for all $\check\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ there exists a locally affine directed partition $\{\check Z^k_{\mathfrak a}, \check C^k_{\mathfrak a}\}_{\substack{k=0,\dots,d\\ \mathfrak a\in\mathfrak A^k}}$ in $\mathbb R^d$ with the following properties:*

(1) for all $\mathfrak a \in \mathfrak A^k$ the cone $\check C^k_{\mathfrak a}$ is a k-dimensional extremal face of $|\cdot|_{D^*}$;

(2) $\mu\Big(\mathbb{R}^d \setminus \bigcup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}}\Big) = 0$;

(3) *the disintegration of* $\mathcal{L}^d$ *w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, $\mathcal{L}^d \llcorner_{\cup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}}} = \int \check\upsilon^k_{\mathfrak{a}}\, d\check\eta(k,\mathfrak{a})$, *satisfies*

$$\check\upsilon^k_a \simeq \mathcal{H}^k \llcorner_{\check Z^k_{\mathfrak{a}}};$$

(4) *the disintegration* $\check\pi = \int \check\pi^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ *w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}} \times \mathbb{R}^d\}_{k,\mathfrak{a}}$ *satisfies*

$$\check\pi^k_{\mathfrak{a}} \in \Pi^{\mathrm{f}}_{\mathtt{c}_{C^k_{\mathfrak{a}}}}\big(\check\mu^k_{\mathfrak{a}}, (\mathtt{p}_2)_\#\check\pi^k_{\mathfrak{a}}\big),$$

where $\mu = \int \check\mu^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ *is the disintegration w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, *and moreover*

$$(\mathtt{p}_2)_\#\check\pi^k_{\mathfrak{a}}\bigg(\check Z^k_{\mathfrak{a}} \cup \bigg(\mathbb{R}^d \setminus \bigcup_{(k',\mathfrak{a}')\neq(k,\mathfrak{a})} \check Z^{k'}_{\mathfrak{a}'}\bigg)\bigg) = 1;$$

(5) *the partition* $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ *is* $\Pi^{\mathrm{f}}_{\mathtt{c}_{\mathbf{D}}}(\mu, \{(\mathtt{p}_2)_\#\check\pi^k_{\mathfrak{a}}\})$*-cyclically connected.*

Remark 1.5. We note that the elements of the locally affine partition $\{\check Z^k_{\mathfrak{a}}, \check C^k_{\mathfrak{a}}\}_{\substack{k=1,\dots,d\\ \mathfrak{a}\in\mathfrak{A}^k}}$ given by the above theorem have maximal linear dimension

$$\max\big\{k : \check Z^k_{\mathfrak{a}} \neq \emptyset\big\} \leq \max\big\{\dim C : C \text{ extremal face of epi}|\cdot|_{D^*}\big\}.$$

In particular, if D is strictly convex, the locally affine decomposition is made only of directed rays, and one recovers the results of [12] for strictly convex norms.

In the case $\mu \ll \mathcal{L}^d$, the decomposition does not depends on the transference plan, as in the strictly convex case. In particular, we can say that it is universal, i.e. it is independent on the particular transference plan $\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$ used.

Theorem 1.6. *Assume that* $\nu \ll \mathcal{L}^d$. *Then the directed locally affine partition of Theorem 1.4 satisfies the following properties:*

(1) *for all* $\mathfrak{a} \in \mathfrak{A}^k$ *the cone* $\check C^k_{\mathfrak{a}}$ *is a k-dimensional extremal face of* $|\cdot|_{D^*}$;

(2') $\mu\Big(\mathbb{R}^d \setminus \bigcup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}}\Big) = \nu\Big(\mathbb{R}^d \setminus \bigcup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}}\Big) = 0$;

(2) *the disintegration of* $\mathcal{L}^d$ *w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$, $\mathcal{L}^d \llcorner_{\cup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}}} = \int \check\upsilon^k_{\mathfrak{a}}\, d\check\eta(k,\mathfrak{a})$, *satisfies*

$$\check\upsilon^k_a \simeq \mathcal{H}^k \llcorner_{\check Z^k_{\mathfrak{a}}};$$

(4') *for all* $\check\pi \in \Pi^{\mathrm{opt}}_{|\cdot|_{D^*}}(\mu,\nu)$, *the disintegration* $\check\pi = \int \check\pi^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ *w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}} \times \mathbb{R}^d\}_{k,\mathfrak{a}}$ *satisfies*

$$\check\pi^k_{\mathfrak{a}} \in \Pi^{\mathrm{f}}_{\mathtt{c}_{C^k_{\mathfrak{a}}}}(\check\mu^k_{\mathfrak{a}}, \check\nu^k_{\mathfrak{a}}),$$

where $\mu = \int \check\mu^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$, $\nu = \int \check\nu^k_{\mathfrak{a}}\, dm(k,\mathfrak{a})$ *are the disintegration of* μ, ν *w.r.t. the partition* $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$;

(5') $\{\check Z^k_{\mathfrak{a}}\}_{k,\mathfrak{a}}$ *is* $\Pi^{\mathrm{f}}_{\mathtt{c}_{\mathbf{D}}}(\mu,\nu)$*-cyclically connected.*

In particular $\check\pi$ lives on $\cup_{k,\mathfrak{a}} \check Z^k_{\mathfrak{a}} \times \check Z^k_{\mathfrak{a}}$.

Finally, from the proof of the above theorems and the precise characterization of the points which are contained in a Γ-cycle , we can deduce the following proposition.

Proposition 1.7. *Let* W *denote any of the sets* $\check Z^\ell_{\mathfrak{b}}$ *in* (1.17), *the sets* $Z'^{,\ell}_{\mathfrak{b}}$ *in Corollary 1.3 and the sets* $\check Z^k_{\mathfrak{a}}$ *of Theorems 1.4 and 1.6. Then it holds*

$$\mathrm{Leb}\big(\mathtt{p}_1(\Gamma) \cap W\big) \quad \textit{is contained in a } (\Gamma \cap Z^\ell_{\mathfrak{a}} \times \mathbb{R}^d)\textit{-cycle},$$

for all carriages Γ of a finite cost/optimal transference plan, and the Lebesgue points are computed w.r.t. the natural Hausdorff measure restricted to W.

To show an application of the above theorems and the ability of our approach to analyze more complicated problem, consider the following situation. Let $|\cdot|_{(D')^*}$ be a convex norm with unit ball D', and consider the secondary minimization problem

$$\min\bigg\{\int |y-x|_{(D')^*}\,d\pi(x,y),\ \pi\in\Pi^{\rm opt}_{|\cdot|_{D^*}}(\mu,\nu)\bigg\}. \tag{1.18}$$

If $\check\pi$ is a minimizer of the above problem, by the fact that $\check\pi$ is also a minimizer of

$$\int \mathtt{c}\pi, \qquad \mathtt{c} := \begin{cases} |y-x|_{(D')^*} & x\in\check Z^k_{\mathfrak a}, y-x\in\check C^k_{\mathfrak a},\\ +\infty & \text{otherwise},\end{cases}$$

and that each $Z^k_{\mathfrak a}$ is $\Pi^{\rm f}_{C^k_{\mathfrak a}}(\mu^k_{\mathfrak a},\nu^k_{\mathfrak a})$-cyclically connected, it is fairly standard to prove that in each $\check Z^k_a$ there exists a potential $\phi_{\mathfrak a}$, ψ_a, and since $\mathtt{c}$ satisfies the triangle inequality we can take $\phi_{\mathfrak a}=-\psi_a$. We will now restrict to a single set Z^k_a, and denote

$$\mathtt{c}^k_{\mathfrak a} := \mathtt{c}\llcorner_{Z^k_a\times\mathbb R^d} = \begin{cases} |y-x|_{(D')^*} & y-x\in\check C^k_{\mathfrak a},\\ +\infty & \text{otherwise},\end{cases} \tag{1.19}$$

We will not address the measurability dependence w.r.t. the parameter $\mathfrak a$ (which in any case follows quite easily from the Borel dependence $(k,\mathfrak a)\mapsto(\check Z^k_{\mathfrak a},\check C^k_{\mathfrak a})$), and we will also not write the index k, $\mathfrak a$ when clear from the context.

The set graph $\psi_{\mathfrak a}$ is a complete epi $\mathtt{c}^k_{\mathfrak a}$-Lipschitz graph, where now Lipschitz means w.r.t. the (degenerate) norm $\mathtt{c}^k_a$. The only variation w.r.t. the proof of Theorem 1.1 is in the cone approximation property, which can be proved by the same techniques using the cost

$$\mathtt{c}''(w,w') = \begin{cases} |w-w'| & w'-w\in\operatorname{epi}\mathtt{c}^k_{\mathfrak a},\\ +\infty & \text{otherwise}.\end{cases} \tag{1.20}$$

Thus we obtain a locally directed affine partition $\{\check Z^\ell_{\mathfrak a,\mathfrak b},\check C^\ell_{\mathfrak a,\mathfrak b}\}_{k,\mathfrak a,\mathfrak b}$ satisfying:

(1) for all $\mathfrak a\in\mathfrak A^k$, $\mathfrak b\in\mathfrak B^\ell_{\mathfrak a}$, the cone $\check C^\ell_{\mathfrak a,\mathfrak b}$ is an ℓ-dimensional extremal face of $\mathtt{c}^k_{\mathfrak a}$;

(2) $\mathcal L^d\Big(\mathbb R^d\setminus\bigcup_{k,\mathfrak a,\mathfrak b}\check Z^k_{\mathfrak a,\mathfrak b}\Big)=0$;

(3) for all $k,\mathfrak a$ the disintegration $\mathcal H^k\llcorner_{\check Z^k_a}=\int v^\ell_{\mathfrak a,\mathfrak b}\,d\eta(\ell,\mathfrak b)$ satisfies
$$v^\ell_{\mathfrak a,\mathfrak b}\simeq\mathcal H^\ell\llcorner\check Z^\ell_{\mathfrak a,\mathfrak b};$$

(4) for all $\pi'\in\Pi^{\rm opt}_{\mathtt{c}^k_{\mathfrak a}}(\mu,\{(\mathtt p_2)_\#\check\pi^k_a\})$, where $\check\pi^k_a$ are the conditional probabilities of the disintegration of $\check\pi$ w.r.t. the partition $\{\check Z^k_a\times\mathbb R^d\}$, the disintegration $\pi'=\displaystyle\int(\pi')^\ell_{\mathfrak a,\mathfrak b}dm(\ell,\mathfrak a,\mathfrak b)$ w.r.t. the partition $\{\check Z^\ell_{\mathfrak a,\mathfrak b}\times\mathbb R^d\}_{\ell,\mathfrak a,\mathfrak b}$ satisfies
$$(\pi')^\ell_{\mathfrak a,\mathfrak b}\in\Pi^{\rm f}_{\mathtt{c}_{\check C^\ell_{\mathfrak a,\mathfrak b}}}\big(\mu^\ell_{\mathfrak a,\mathfrak b},(\mathtt p_2)_\#(\pi')^\ell_{\mathfrak a,\mathfrak b}\big),$$
with $\mu=\displaystyle\int\mu^\ell_{\mathfrak a,\mathfrak b}\,dm(\ell,\mathfrak a,\mathfrak b)$ and moreover
$$(\mathtt p_2)_\#(\pi')^\ell_{\mathfrak a,\mathfrak b}\bigg(\check Z^\ell_{\mathfrak a,\mathfrak b}\cup\bigg(\mathbb R^d\setminus\bigcup_{(\mathfrak a',\ell',\mathfrak b')\neq(\mathfrak a,\ell,\mathfrak b)}\check Z^{\ell'}_{\mathfrak a',\mathfrak b'}\bigg)\bigg)=1.$$

By applying Corollary 1.3 we thus obtain the following theorem.

Theorem 1.8. *Let μ, $\nu\in\mathcal P(\mathbb R^d)$ with $\mu\ll\mathcal L^d$ and let $\check\pi$ be an optimal transport plan for the problem (1.18). Then there exists a locally affine directed partition $\{\check Z^{k,\ell}_{\mathfrak a,\mathfrak b},\check C^{k,\ell}_{\mathfrak a,\mathfrak b}\}_{\substack{k=0,\dots,d,\,\mathfrak a\in\mathfrak A^k\\ \ell=0,\dots,k,\,\mathfrak b\in\mathfrak B^{k,\ell}_{\mathfrak a}}}$ in $\mathbb R^d$ with the following properties:*

(1) *for all* k, $\mathfrak{a} \in \mathfrak{A}^k$, *the cone* $\check{C}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}$ *is an* ℓ*-dimensional cone of the cost* $\mathtt{c}^k_{\mathfrak{a}}$ *given by* (1.19), *i.e. the intersection of a* k*-dimensional face of* $|\cdot|_{D^*}$ *with an extremal face of* $|\cdot|_{(D')^*}$*;*

(2) $\mu\Big(\mathbb{R}^d \setminus \bigcup_{k,\mathfrak{a},\ell,\mathfrak{b}} \check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\Big) = 0$*;*

(3) *the disintegration* $\mathcal{L}^d \llcorner_{\cup_{k,\mathfrak{a},\ell,\mathfrak{b}} \check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}} = \int \check{\upsilon}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\, d\eta(k,\mathfrak{a},\ell,\mathfrak{b})$ *w.r.t. the partition* $\{\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\}_{k,\mathfrak{a},\ell,\mathfrak{b}}$ *satisfies*

$$\upsilon^{k,\ell}_{\mathfrak{a},\mathfrak{b}} \simeq \mathcal{H}^\ell \llcorner_{\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}};$$

(4) *the disintegration* $\check{\pi} = \int \check{\pi}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\, dm(k,\mathfrak{a},\ell,\mathfrak{b})$ *w.r.t. the partition* $\{\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}} \times \mathbb{R}^d\}_{k,\mathfrak{a},\ell,\mathfrak{b}}$ *satisfies*

$$\pi^{k,\ell}_{\mathfrak{a},\mathfrak{b}} \in \Pi^{\mathrm{f}}_{\mathtt{c}_{C^{k,\ell}_{\mathfrak{a},\mathfrak{b}}}}\big(\check{\mu}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}, (\mathtt{p}_2)_\# \check{\pi}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\big),$$

where $\mu = \int \check{\mu}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\, dm(k,\mathfrak{a},\ell,\mathfrak{b})$ *is the disintegration w.r.t. the partition* $\{\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\}_{k,\mathfrak{a},\ell,\mathfrak{b}}$*, and moreover*

$$(\mathtt{p}_2)_\# \check{\pi}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\bigg(\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}} \cup \bigg(\mathbb{R}^d \setminus \bigcup_{(k',\mathfrak{a}',\ell',\mathfrak{b}') \neq (k,\mathfrak{a},\ell,\mathfrak{b})} \check{Z}^{k',\ell'}_{\mathfrak{a}',\mathfrak{b}'}\bigg)\bigg) = 1;$$

(5) *the partition* $\{\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\}_{k,\mathfrak{a},\ell,\mathfrak{b}}$ *is* $\Pi^{\mathrm{f}}_{\mathtt{c}_{\bar{\mathbf{D}}}}(\mu, \{(\mathtt{p}_2)_\# \check{\pi}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}\})$*-cyclically connected.*

A completely similar extension can be given to Theorem 1.6.

A particular case is when each extremal face of $|\cdot|_{(D')^*}$ is contained in an extremal face of $|\cdot|_{D^*}$: in this case condition (1) becomes

(1') for all k, $\mathfrak{a} \in \mathfrak{A}^k$, the cone $\check{C}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}$ is an ℓ-dimensional extremal face of $|\cdot|_{(D')^*}$.

The only difference w.r.t. Theorem 1.4 is that now $\check{\pi}$ is a minimum for the secondary minimization problem, not a transference plan in $\Pi^{\mathrm{opt}}_{|\cdot|_{(D')^*}}(\mu,\nu)$.

The case (1') above happens if for example $|\cdot|_{(D')^*}$ is strictly convex, so that the $\check{Z}^{k,\ell}_{\mathfrak{a},\mathfrak{b}}$ are now directed segments, i.e. $\ell = 1$. By the standard analysis on transportation problems in 1-d, and the measurable dependence on k, $\mathfrak{a}$, $\mathfrak{b}$, the next corollary follows.

Corollary 1.9. *There exists an optimal transport map* $\mathtt{T}$*, and the map can be selected as follows: the restriction* $\mathtt{T}\llcorner_{\check{Z}^{k,1}_{\mathfrak{a},\mathfrak{b}}}$ *is a monotone increasing map on* $\operatorname{aff} \check{Z}^{k,1}_{\mathfrak{a},\mathfrak{b}}$*, for all* k*,* $\mathfrak{a}$*,* $\mathfrak{b}$*.*

References

[1] L. Ambrosio. Lecture notes on optimal transportation problems. In *Mathematical aspects of evolving interfaces*, volume 1812 of *Lecture Notes in Math.*, pages 1–52. Springer, 2003.

[2] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.

[3] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In *Optimal Transportation and Applications*, volume 1813 of *Lecture notes in mathematics*, pages 123–160. Springer, 2001.

[4] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 4(4):353–454, 2009.

[5] S. Bianchini and F. Cavalletti. The Monge problem in metric spaces. *Comm. Math. Phys.*, 318:615–673, 2013.

[6] S. Bianchini and S. Daneri. On sudakov's type decompositions of transference plans with norm costs. in preparation.

[7] S. Bianchini and M. Gloyer. On the Euler-Lagrange equation for a variational problem: the general case II. *Math. Z.*, 265(4):889–923, 2010.

[8] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44:375–417, 1991.

[9] F. Santambrogio C. Jimenez. Optimal transportation for a quadratic cost with convex constraints and applications. *J. Math. Pures et Appl.*, 2011.

[10] L. Caffarelli, M. Feldman, and R.J. McCann. Constructing optimal maps for monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2002.

[11] L. Caravenna. A proof of Monge problem in $\mathbb{R}^n$ by stability. *Rend. Istit. Mat. Univ. Trieste*, 43:31–51, 2011.

[12] L. Caravenna. A proof of Sudakov theorem with strictly convex norms. *Math. Z.*, 268:371–407, 2011.

[13] L. Caravenna and S. Daneri. The disintegration of the Lebesgue measure on the faces of a convex function. *J. Funct. Anal.*, 258(11):3604–3661, 2010.

[14] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. *Journ. Eur. Math. Soc.*, 12(6):1355–1369, 2010.

[15] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. *Duke Math. J.*, 157(3):551–572, 2011.
[16] S. Daneri. *Dimensional Reduction and Approximation of Measures and Weakly Differentiable Homeomorphisms*. PhD thesis, SISSA, 2011.
[17] L.C. Evans and W. Gangbo. Differential equations methods for the monge-kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.
[18] V.N. Sudakov. Geometric problems in the theory of dimensional distributions. *Proc. Steklov Inst. Math.*, 141:1–178, 1979.
[19] N. Trudinger and X.J. Wang. On the monge mass transfer problem. *Calc. Var. PDE*, 13:19–31, 2001.

SISSA, via Beirut 2, IT-34014 Trieste (ITALY)
E-mail address: bianchin@sissa.it

University of Zurich
E-mail address: sara.daneri@math.uzh.ch