



**HAL**  
open science

## Optimal transportation for the determinant

Guillaume Carlier, Bruno Nazaret

► **To cite this version:**

Guillaume Carlier, Bruno Nazaret. Optimal transportation for the determinant. *ESAIM: Control, Optimisation and Calculus of Variations*, 2008, 14 (4), p678-698. 10.1051/cocv:2008006 . hal-00118459

**HAL Id: hal-00118459**

**<https://hal.science/hal-00118459>**

Submitted on 5 Dec 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimal transportation for the determinant

G. Carlier, B. Nazaret \*

4th december 2006

## Abstract

Among  $\mathbb{R}^3$ -valued triples of random vectors  $(X, Y, Z)$  having fixed marginal probability laws, what is the best way to jointly draw  $(X, Y, Z)$  in such a way that the simplex generated by  $(X, Y, Z)$  has maximal average volume? Motivated by this simple question, we study optimal transportation problems with several marginals when the objective function is the determinant or its absolute value.

**Keywords:** Optimal transportation, multi-marginals problems, determinant, disintegrations.

## 1 Introduction

Given two probability measures  $\mu_1$  and  $\mu_2$  on  $\mathbb{R}^d$  and some objective function  $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ , the classical Monge-Kantorovich optimal transportation problem consists in finding a probability measure  $\gamma$  on  $\mathbb{R}^d \times \mathbb{R}^d$  having  $\mu_1$  and  $\mu_2$  as marginals (i.e. a *transportation plan* between  $\mu_1$  and  $\mu_2$ ) maximizing the total objective  $\int_{\mathbb{R}^d \times \mathbb{R}^d} H(x, y) d\gamma(x, y)$ . In his famous article [1], Brenier solved the case  $H(x, y) = \langle x, y \rangle$  and proved (under mild regularity assumptions) that there is a unique optimal transportation plan which is further characterized by the property of being supported by the graph of the gradient of some convex function. Brenier's seminal results have been extended to the case of more general costs (see McCann and Gangbo [4]) and the subject has received a lot of attention in the last 15 years because of its numerous applications in fluid mechanics, probability and statistics, PDE's, shape optimization, mathematical economics.... The literature on this very active field of research is too vast to give an exhaustive bibliography here, we rather refer to the books of Villani [6] and Rachev and Rüschendorf [5] and the references therein.

In the present article, we are interested in an optimal transportation problem with several marginals. Given  $d$  probability measures  $\mu_1, \dots, \mu_d$

---

\* CEREMADE, UMR CNRS 7534, Université Paris Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE [carlier@ceremade.dauphine.fr](mailto:carlier@ceremade.dauphine.fr), [nazaret@ceremade.dauphine.fr](mailto:nazaret@ceremade.dauphine.fr).

on  $\mathbb{R}^d$  and an objective function  $H : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}$ , the problem is to find a probability measure  $\gamma$  on  $\mathbb{R}^{d \times d}$  having  $\mu_1, \dots, \mu_d$  as marginals maximizing  $\int_{\mathbb{R}^{d \times d}} H(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d)$ . In contrast with the case of two marginals, there are few results on the optimal transportation problem when more than two marginals are involved with the exception of the article of Gangbo and Świąch [3] who fully solved the case  $H(x_1, \dots, x_d) := \sum_{i,j} \langle x_i, x_j \rangle$ . For this particular problem, Gangbo and Świąch proved existence and uniqueness of an optimal transportation plan which is supported by the graph of some transport map. In the present paper, we will pay attention here to different objective functions, namely:  $H(x) = \det(x)$  or  $H(x) = |\det(x)|$ .

The choice of such functions of the determinant is motivated by the following simple question: among random  $\mathbb{R}^3$ -valued vectors  $X, Y$ , and  $Z$ , with fixed marginal probability laws, what is the best way to draw jointly  $(X, Y, Z)$  so that the simplex with vertices  $(0, X, Y, Z)$  has maximal average volume? Denoting by  $\mu_1, \mu_2$  and  $\mu_3$  the (fixed) probability laws of the random vectors  $(X, Y, Z)$ , the previous problem amounts to maximize

$$\int_{\mathbb{R}^{3 \times 3}} |\det(x, y, z)| d\gamma(x, y, z)$$

among joint probability laws  $\gamma$  having  $\mu_1, \mu_2$  and  $\mu_3$  as marginals. In some cases, we will see that solving the problem above amounts to solve the simpler problem where  $|\det|$  is replaced by  $\det$  and we will study this case in dimension  $d \geq 2$ . If  $d = 2$ , we have  $\det(x, y) = \langle x, Ry \rangle$  (with  $R$  the rotation of angle  $-\pi/2$ ), so that, up to the change of variables  $y' = Ry$ , the optimal transportation problem with the determinant is a special case of the problem solved by Brenier [1].

Let us define some notations. In the sequel, given  $X$  a locally compact separable metric space, we denote by  $\mathcal{M}(X)$  (respectively  $\mathcal{M}_+^1(X)$ ) the set of Radon measures (respectively of Radon probability measures) on  $X$ . If  $X$  and  $Y$  are locally compact separable metric spaces,  $\mu \in \mathcal{M}_+^1(X)$ , and  $f : X \rightarrow Y$  is a Borel map we shall denote by  $f\#\mu$  the push forward of  $\mu$  through  $f$  i.e. the element of  $\mathcal{M}_+^1(Y)$  defined by  $f\#\mu(B) = \mu(f^{-1}(B))$  for every Borel subset  $B$  of  $Y$ . If  $\gamma \in \mathcal{M}(\mathbb{R}^{d \times d})$  and  $i \in \{1, \dots, d\}$ ,  $\pi_i\#\gamma \in \mathcal{M}(\mathbb{R}^d)$  is called the  $i$ -th *marginal* of  $\gamma$  (where  $\pi_i$  is the  $i$ -th canonical projection). One can also define  $\pi_i\#\gamma$  by:

$$\int_{\mathbb{R}^d} f(x_i) d(\pi_i\#\gamma)(x_i) = \int_{\mathbb{R}^{d \times d}} f(x_i) d\gamma(x_1, \dots, x_d),$$

for every bounded and continuous function  $f$  on  $\mathbb{R}^d$ . Given  $d$  probability measures on  $\mathbb{R}^d$ ,  $\mu_1, \dots, \mu_d$ , we denote by  $\Pi(\mu_1, \dots, \mu_d)$  the set of probability measures on  $\mathbb{R}^{d \times d}$  having  $\mu_1, \dots, \mu_d$  as marginals. In other words,  $\gamma \in \mathcal{M}_+^1(\mathbb{R}^{d \times d})$  belongs to  $\Pi(\mu_1, \dots, \mu_d)$  if and only if

$$\int_{\mathbb{R}^{d \times d}} f(x_i) d\gamma(x_1, \dots, x_d) = \int_{\mathbb{R}^d} f(x_i) d\mu_i(x_i), \quad \forall i = 1, \dots, d,$$

for every bounded and continuous function  $f$  on  $\mathbb{R}^d$ . Given  $H \in C^0(\mathbb{R}^{d \times d}, \mathbb{R})$ , the Monge-Kantorovich optimal transportation problem with marginals  $\mu_1, \dots, \mu_d$  and objective function  $H$  then reads as:

$$\sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} H(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d).$$

We shall focus here on the two special cases:

$$(\mathcal{MK}) \quad \sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} \det(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d)$$

and

$$(\mathcal{MK})_a \quad \sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} |\det(x_1, \dots, x_d)| d\gamma(x_1, \dots, x_d).$$

The paper is organized as follows. In section 2, we solve a particular example which actually gives insight on the general case. Section 3 is devoted to existence, duality and characterization of minimizers for  $(\mathcal{MK})$ . In section 4, we construct minimizers in the case of radially symmetric marginals. In section 5, we give conditions ensuring that  $(\mathcal{MK})$  and  $(\mathcal{MK})_a$  are in fact equivalent. Section 6 contains various remarks regarding uniqueness issues.

## 2 An elementary example

In this section, we study a simple but illustrative example, which actually contains most of the ideas necessary for the understanding of the general case. Let us consider the problem  $(\mathcal{MK})$  in dimension 3, with  $\mu_1 = \mu_2 = \mu_3 = \mathcal{L}_B^3$ , where  $\mathcal{L}_B^3$  stands for the uniform probability measure on the unit ball  $B$  of  $\mathbb{R}^3$ ,

$$\sup_{\gamma \in \Pi(\mathcal{L}_B^3, \mathcal{L}_B^3, \mathcal{L}_B^3)} \left( \int_{B^3} \det(x, y, z) d\gamma(x, y, z) \right). \quad (1)$$

Then, the following result holds.

**Theorem 1** *The problem (1) admits a solution  $\bar{\gamma}$  given by, for every continuous  $f : B^3 \rightarrow \mathbb{R}$ ,*

$$\int_{B^3} f d\bar{\gamma} = \frac{1}{|B|} \int_B \left( \int_{\mathbb{S}(x)} f(x, |x|y, x \wedge y) \frac{d\mathcal{H}^1(y)}{2\pi} \right) dx, \quad (2)$$

where

$$\mathbb{S}(x) = \{y \in \mathbb{S}^2 \quad \text{s.t.} \quad \langle x, y \rangle = 0\}.$$

**Remark 1.** Let us remark that, the support of the optimal measure  $\bar{\gamma}$  is the set of all the triples  $(x, y, z) \in B^3$ , such that  $|x| = |y| = |z|$  and  $(x, y, z)$  is a direct orthogonal basis of  $\mathbb{R}^3$ . As we shall see later on, this last property comes from the fact that the measures are radially symmetric. In addition, the definition of  $\bar{\gamma}$  through its successive disintegrations has the following probabilistic interpretation in terms of conditional laws. Consider  $\bar{\gamma}$  as the law of a vector  $(X, Y, Z)$  of random vectors in  $\mathbb{R}^3$ , each of them following an uniform law on the unit ball. Then, (2) means that the conditional probability of  $Y$  given  $X$  is uniform on  $S(X)$  and that the conditional probability of  $Z$  given  $(X, Y)$  is the Dirac mass at  $\frac{X \wedge Y}{|X|}$ .

**Remark 2.** Let us say a word about the case where the objective function is  $H(x, y, z) = |\det(x, y, z)|$  (volume maximization). For this  $H$ ,  $\bar{\gamma}$  is still a maximizer, but we could have chosen as well

$$Z = -\frac{X \wedge Y}{|X|},$$

for the third variable, and the solution is not unique. Actually, it is not the only source of non uniqueness, and we shall discuss this point in the last section.

**Proof of theorem 1.**

Assuming that  $\bar{\gamma}$  is admissible in (1), we only need to prove that it is optimal. To do so, consider the variational problem

$$\inf_{(\varphi, \psi, \chi) \in \mathcal{E}} \frac{1}{|B|} \left( \int_B \varphi(x) dx + \int_B \psi(y) dy + \int_B \chi(z) dz \right), \quad (3)$$

with  $\mathcal{E}$  the set of triples  $(\varphi, \psi, \chi) \in C^0(B, \mathbb{R})^3$  such that

$$\varphi(x) + \psi(y) + \chi(z) \geq \det(x, y, z), \quad \forall (x, y, z) \in B^3.$$

As we shall see in next section, (3) is dual to (1) in some sense. For all  $(\varphi, \psi, \chi) \in \mathcal{E}$ , and  $\gamma \in \Pi(\mathcal{L}_B^3, \mathcal{L}_B^3, \mathcal{L}_B^3)$ ,

$$\begin{aligned} \int_{B^3} \det(x, y, z) d\gamma(x, y, z) &\leq \int_{B^3} (\varphi(x) + \psi(y) + \chi(z)) d\gamma(x, y, z), \\ &\leq \frac{1}{|B|} \left( \int_B \varphi(x) dx + \int_B \psi(y) dy + \int_B \chi(z) dz \right), \end{aligned}$$

and it immediately follows that

$$\begin{aligned} &\sup_{\Pi(\mathcal{L}_B^3, \mathcal{L}_B^3, \mathcal{L}_B^3)} \int_{B^3} \det(x, y, z) d\gamma(x, y, z) \\ &\leq \inf_{\mathcal{E}} \frac{1}{|B|} \left( \int_B \varphi(x) dx + \int_B \psi(y) dy + \int_B \chi(z) dz \right). \end{aligned} \quad (4)$$

Now, consider  $(\varphi_0, \psi_0, \chi_0)$  given by

$$\forall x \in B, \varphi_0(x) = \psi_0(x) = \chi_0(z) = \frac{|x|^3}{3}.$$

Thanks to the Young inequality, we have

$$\forall (x, y, z) \in B^3, \varphi_0(x) + \psi_0(y) + \chi_0(z) \geq |x||y||z| \geq \det(x, y, z), \quad (5)$$

so that  $(\varphi_0, \psi_0, \chi_0) \in \mathcal{E}$ . In addition, according to remark 1, if  $(x, y, z) \in \text{supp}(\overline{\gamma})$ , then all the inequalities in (5) become equalities and then,

$$\forall (x, y, z) \in \text{supp}(\overline{\gamma}), \varphi_0(x) + \psi_0(y) + \chi_0(z) = \det(x, y, z). \quad (6)$$

It implies that (4) is actually an equality and  $\overline{\gamma}$  is optimal in (1). To complete the proof, it just remains to show that  $\overline{\gamma}$  is admissible, which is done in proposition 1.  $\square$

**Proposition 1** *Let  $\overline{\gamma}$  defined as in theorem 1. Then,  $\overline{\gamma} \in \Pi(\mathcal{L}_B^3, \mathcal{L}_B^3, \mathcal{L}_B^3)$ .*

**Proof.** The fact that  $\pi_1\#\overline{\gamma} = \mathcal{L}_B^3$  is obvious by (2). Now, let us suppose that we proved  $\pi_2\#\overline{\gamma} = \mathcal{L}_B^3$ . Then, we can easily deduce the result for the third marginal. Indeed, since for every fixed  $x \in B$ , the map  $y \mapsto \frac{x \wedge y}{|x|}$  is one-to-one from  $\mathbb{S}(x)$  to itself and is simply a rotation with angle  $\pi/2$ , we have that for all continuous  $f : B \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \int_B f d(\pi_3\#\overline{\gamma}) &= \frac{1}{|B|} \int_B \left( \int_{\mathbb{S}(x)} f(x \wedge y) \frac{d\mathcal{H}^1(y)}{2\pi} \right) dx, \\ &= \frac{1}{|B|} \int_B \left( \int_{\mathbb{S}(x)} f(|x|y) \frac{d\mathcal{H}^1(y)}{2\pi} \right) dx, \\ &= \int_B f d(\pi_2\#\overline{\gamma}). \end{aligned}$$

We then prove that  $\pi_2\#\overline{\gamma} = \mathcal{L}_B^3$ . Let  $f : B \rightarrow \mathbb{R}$  be a continuous function. Then

$$\begin{aligned} \int_B f d(\pi_2\#\overline{\gamma}) &= \frac{1}{|B|} \int_B \left( \int_{\mathbb{S}(x)} f(|x|y) \frac{d\mathcal{H}^1(y)}{2\pi} \right) dx \\ &= \frac{1}{|B|} \int_0^1 r^2 \left( \int_{\mathbb{S}^2} \left( \int_{\mathbb{S}(x)} f(ry) \frac{d\mathcal{H}^1(y)}{2\pi} \right) d\mathcal{H}^2(\sigma) \right) dr. \end{aligned}$$

Applying lemma 2 proved in Appendix A, we get

$$\begin{aligned} \int_B f d(\pi_2\#\overline{\gamma}) &= \frac{1}{|B|} \int_0^1 \left( \int_{\mathbb{S}^2} r^2 f(ry) \left( \int_{S(y)} \frac{d\mathcal{H}^1(\sigma)}{2\pi} \right) d\mathcal{H}^2(y) \right) dr \\ &= \int_B f d\mathcal{L}_B^3, \end{aligned}$$

which ends the proof.  $\square$

Let us notice that the condition (6) completely characterizes the solutions of the dual problem (3). In fact (see subsection 3.3), up to the addition of constants that sum to 0,  $(\varphi_0, \psi_0, \chi_0)$  is the unique solution of (3). Yet, there are infinitely many solutions to (1), indeed any  $\gamma \in \Pi(\mathcal{L}_B^3, \mathcal{L}_B^3, \mathcal{L}_B^3)$  having its support in all the triples  $(x, y, z) \in B^3$ , such that  $|x| = |y| = |z|$  and  $(x, y, z)$  is a direct orthogonal basis of  $\mathbb{R}^3$  is optimal for (3) and we claim that there are infinitely many such probability measures (see section 6).

### 3 Duality, existence and characterization

Given  $d$  probability measures on  $\mathbb{R}^d$ ,  $\mu_1, \dots, \mu_d$ , we consider the problem

$$(\mathcal{MK}) \quad \sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} \det(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d).$$

In this case, a natural assumption on the marginals  $\mu_1, \dots, \mu_d$ , is the existence of  $(p_1, \dots, p_d) \in (1, +\infty)^d$  such that:

$$\sum_{i=1}^d \frac{1}{p_i} = 1, \quad \text{and} \quad a := \sum_{i=1}^d \int_{\mathbb{R}^d} \frac{|x|^{p_i}}{p_i} d\mu_i(x) < +\infty. \quad (7)$$

Defining for all  $x = (x_1, \dots, x_d) \in \mathbb{R}^{d \times d}$ :

$$H_0(x) := \sum_{i=1}^d \frac{|x_i|^{p_i}}{p_i}$$

and using the fact that  $|\det(x)| \leq H_0(x)$ , we immediately get

$$-a \leq \int_{\mathbb{R}^{d \times d}} \det(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d) \leq a, \quad \forall \gamma \in \Pi(\mu_1, \dots, \mu_d).$$

Similarly, defining for all  $x = (x_1, \dots, x_d) \in \mathbb{R}^{d \times d}$ :

$$\overline{H}(x) := \det(x) + H_0(x), \quad \underline{H}(x) := \det(x) - H_0(x), \quad (8)$$

we remark that for all  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$

$$\int_{\mathbb{R}^{d \times d}} \det(x) d\gamma(x) = \int_{\mathbb{R}^{d \times d}} \overline{H} d\gamma - a = \int_{\mathbb{R}^{d \times d}} \underline{H} d\gamma + a.$$

The previous remark implies that when  $H = \det$ , one may without loss of generality replace  $H = \det$  with  $\overline{H} \geq 0$  or  $\underline{H} \leq 0$  in  $(\mathcal{MK})$ .

A key point in Monge-Kantorovich theory, is to remark that (in a sense that will be made precise later),  $(\mathcal{MK})$  is dual to :

$$(\mathcal{D}) \quad \inf_{(\varphi_1, \dots, \varphi_d) \in \mathcal{E}} \sum_{i=1}^d \int_{\mathbb{R}^d} \varphi_i d\mu_i$$

where  $\mathcal{E}$  is the set of  $d$ -uples of lower-semi continuous functions  $(\varphi_1, \dots, \varphi_d)$  from  $\mathbb{R}^d$  to  $\mathbb{R} \cup \{+\infty\}$  such that

$$\sum_{i=1}^d \varphi_i(x_i) \geq \det(x_1, \dots, x_d), \quad \forall (x_1, \dots, x_d) \in \mathbb{R}^{d \times d}. \quad (9)$$

Let us note that one obviously has  $\inf(\mathcal{D}) \geq \sup(\mathcal{MK})$ .

### 3.1 The compact case

In this paragraph, for further use, we consider the case of compactly supported marginals and of an arbitrary continuous objective function  $H$ . Denoting by  $B_r$  the closed ball in  $\mathbb{R}^d$ , with center 0 and radius  $r$ , we assume that  $\mu_1, \dots, \mu_d$  are supported in  $B_r$  for some given  $r > 0$  and that  $H$  is an arbitrary continuous function on  $B_r^d$ . We then consider the optimal transportation problem:

$$\sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} H(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d). \quad (10)$$

We define its dual by:

$$\inf_{(\varphi_1, \dots, \varphi_d) \in \mathcal{E}_r} \sum_{i=1}^d \int_{B_r} \varphi_i d\mu_i \quad (11)$$

where  $\mathcal{E}_r$  is the set of  $d$ -uples of continuous functions  $(\varphi_1, \dots, \varphi_d)$  from  $B_r$  to  $\mathbb{R}$  such that

$$\sum_{i=1}^d \varphi_i(x_i) \geq H(x_1, \dots, x_d), \quad \forall (x_1, \dots, x_d) \in B_r^d. \quad (12)$$

**Proposition 2** *Assume that  $\mu_1, \dots, \mu_d$  are supported in the ball  $B_r$  and that  $H$  is continuous on  $B_r^d$  then both (10) and (11) admit solutions and*

$$\max (10) = \min (11).$$

Moreover, (11) admits a solution  $(\varphi_1, \dots, \varphi_d)$  such that for all  $i = 1, \dots, d$  and all  $x_i \in B_r$ , one has:

$$\varphi_i(x_i) = \sup_{(x_j)_{j \neq i} \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j \neq i} \varphi_j(x_j) \right\} \quad (13)$$

and, for all  $x \in B_r$ :

$$\min_{B_r^d} H \leq \varphi_d(x) \leq \max_{B_r^d} H, \quad 0 \leq \varphi_i(x) \leq \max_{B_r^d} H - \min_{B_r^d} H, \quad i = 1, \dots, d-1. \quad (14)$$



**Proof.**

**Step 1 : convex duality**

Equip  $E := C^0(B_r, \mathbb{R})^d$  with the sup norm, and define (the linear continuous operator)  $\Lambda$  by  $\Lambda((\varphi_1, \dots, \varphi_d))(x_1, \dots, x_d) := \sum_{i=1}^d \varphi_i(x_i)$  for all  $(\varphi_1, \dots, \varphi_d) \in C^0(B_r, \mathbb{R})^d$  and all  $(x_1, \dots, x_d) \in B_r^d$ . Remark now that (11) can be rewritten as:

$$\inf_{\varphi \in E} F(\varphi) + G(\Lambda(\varphi)) \quad (15)$$

with  $F(\varphi) = \sum_{i=1}^d \int_{B_r} \varphi_i d\mu_i$  and, for all  $\psi \in C^0(B_r^d, \mathbb{R})$ :

$$G(\psi) = \begin{cases} 0 & \text{if } \psi \geq H \\ +\infty & \text{otherwise .} \end{cases}$$

The dual problem in the usual sense of convex analysis of (15) (see [2]) is then

$$\sup_{\gamma \in \mathcal{M}(B_r^d)} -F^*(\Lambda^* \gamma) - G^*(-\gamma). \quad (16)$$

First, we remark that  $\Lambda^*(\gamma) = (\pi_1 \# \gamma, \dots, \pi_d \# \gamma)$ . Elementary computations then yield:

$$F^*(\Lambda^*(\gamma)) = \begin{cases} 0 & \text{if } \pi_i \# \gamma = \mu_i \text{ for } i = 1, \dots, d, \\ +\infty & \text{otherwise .} \end{cases}$$

and

$$G^*(-\gamma) = \begin{cases} -\int_{\mathbb{R}^{d \times d}} H d\gamma & \text{if } \gamma \geq 0 \\ +\infty & \text{otherwise .} \end{cases}$$

Hence problem (16) is exactly (10). The Fenchel-Rockafellar duality theorem (see [2]) implies then that (10) admits solutions and that

$$\max (10) = \inf (11).$$

**Step 2: convexification trick**

It remains to prove that the infimum is attained in (11). Let  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}_r$  and define for all  $x_1 \in B_r$ :

$$\psi_1(x_1) = \sup_{(x_2, \dots, x_d) \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j=2}^d \varphi_j(x_j) \right\} \quad (17)$$

by construction  $\psi_1 \leq \varphi_1$  and  $(\psi_1, \varphi_2, \dots, \varphi_d) \in \mathcal{E}_r$ . Construct then inductively  $\psi_2, \dots, \psi_{d-1}$  by setting for  $i = 2, \dots, d-1$  and  $x_i \in B_r$

$$\psi_i(x_i) = \sup_{(x_j)_{j \neq i} \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j < i} \psi_j(x_j) - \sum_{j > i} \varphi_j(x_j) \right\}$$

and finally

$$\psi_d(x_d) = \sup_{(x_j)_{j \neq d} \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j=1}^{d-1} \psi_j(x_j) \right\}.$$

By construction,  $\psi_i \leq \varphi_i$  and  $(\psi_1, \dots, \psi_d) \in \mathcal{E}_r$  :  $(\psi_1, \dots, \psi_d)$  is therefore an improvement of  $(\varphi_1, \dots, \varphi_d)$  in problem (11). On the one hand, since  $\psi_j \leq \varphi_j$ , one has for all  $i = 1, \dots, d$  and  $x_i \in B_r$

$$\psi_i(x_i) \leq \sup_{(x_j)_{j \neq i} \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j \neq i} \psi_j(x_j) \right\}.$$

On the other hand, the converse inequality holds because  $(\psi_1, \dots, \psi_d) \in \mathcal{E}_r$ .

**Step 3: existence of minimizers for (11)**

Using the convexification trick of the previous step, we can find a minimizing sequence of (11),  $\varphi^n := (\varphi_1^n, \dots, \varphi_d^n)$  such that for all  $n$ , all  $i$  and all  $x_i \in B_r$ , one has:

$$\varphi_i^n(x_i) = \sup_{(x_j)_{j \neq i} \in B_r^{d-1}} \left\{ H(x_1, \dots, x_d) - \sum_{j \neq i} \varphi_j^n(x_j) \right\} \quad (18)$$

Noting that the objective in (11) is unchanged when changing  $(\varphi_1, \dots, \varphi_d)$  into  $(\varphi_1 + \alpha_1, \dots, \varphi_d + \alpha_d)$  for constants  $\alpha_i$  that sum to 0, we may also assume that

$$\min_{B_r} \varphi_i^n = 0, \quad \forall n \in \mathbb{N}, \quad \forall i = 1, \dots, d-1.$$

Together with (18) we deduce that

$$\min_{B_r^d} H \leq \varphi_d^n \leq \max_{B_r^d} H \text{ on } B_r, \quad \forall n \in \mathbb{N}$$

and

$$\varphi_i^n \leq \max_{B_r^d} H - \min_{B_r^d} H.$$

Denoting by  $\omega$  the modulus of continuity of  $H$  on  $B_r^d$ , we also deduce from (18), that for every  $(x, y) \in B_r^2$ , for every  $n$  and  $i$  one has:

$$|\varphi_i^n(x) - \varphi_i^n(y)| \leq \omega(|x - y|).$$

Thus, the sequence  $\varphi^n$  is bounded and uniformly equicontinuous hence by Ascoli's theorem admits some convergent subsequence. Denoting by  $\varphi = (\varphi_1, \dots, \varphi_d)$  the limit of this subsequence, it is easy to check that  $\varphi$  belongs to  $\mathcal{E}_r$ , solves (11) and satisfies (13) and (14).  $\square$

### 3.2 The general case

We now go back to  $(\mathcal{MK})$  (i.e.  $H = \det$ ) for general marginals that only satisfy (7). By suitable truncation arguments, we have the following result, which is proved in Appendix B:

**Theorem 2** *Assume that (7) is satisfied, then both  $(\mathcal{MK})$  and  $(\mathcal{D})$  admit solutions and*

$$\max (\mathcal{MK}) = \min (\mathcal{D}).$$

Moreover,  $(\mathcal{D})$  admits a solution  $(\varphi_1, \dots, \varphi_d)$  such that for all  $i = 1, \dots, d$  and all  $x_i \in \mathbb{R}^d$ , one has:

$$\varphi_i(x_i) = \sup_{(x_j)_{j \neq i} \in \mathbb{R}^{d \times (d-1)}} \left\{ \det(x_1, \dots, x_d) - \sum_{j \neq i} \varphi_j(x_j) \right\}. \quad (19)$$

### 3.3 Extremality conditions

This paragraph is devoted to optimality conditions for  $(\mathcal{MK})$  that can be derived from Theorem 2 (again we are in the case  $H = \det$  here). For  $(y_1, \dots, y_{d-1}) \in (\mathbb{R}^d)^{d-1}$  we denote by  $\bigwedge_{i=1}^{d-1} y_i$  the vector  $y_1 \wedge \dots \wedge y_{d-1}$  and recall that is characterized by the identity:

$$\det(y_1, \dots, y_{d-1}, x) = \left\langle x, \bigwedge_{i=1}^{d-1} y_i \right\rangle, \quad \forall x \in \mathbb{R}^d.$$

For  $(x_1, \dots, x_d) \in (\mathbb{R}^d)^d$  and  $i \in \{1, \dots, d\}$ ,  $\bigwedge_{j \neq i} x_j$  is defined in a similar way.

At this point, it is useful to remark that the family of l.s.c. functions  $(\varphi_1, \dots, \varphi_d)$  belongs to  $\mathcal{E}$  if and only if for every  $i \in \{1, \dots, d\}$  one has

$$\sum_{j \neq i} \varphi_j(x_j) \geq \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j), \quad \forall (x_1, \dots, x_d) \in (\mathbb{R}^d)^d. \quad (20)$$

It is also obvious that  $(\varphi_1, \dots, \varphi_d)$  belongs to  $\mathcal{E}$  if and only, it satisfies (20) for *some*  $i \in \{1, \dots, d\}$ . Let us finally recall that the convexification trick ensures that one can always improve an element of  $\mathcal{E}$  in the dual problem by replacing it by another element of  $\mathcal{E}$  that satisfies (19). Hence solutions of  $(\mathcal{D})$  have to agree  $\otimes_{i=1}^d \mu_i$  almost everywhere with potentials that satisfy (19). Obviously if  $(\varphi_1, \dots, \varphi_d)$  satisfies (19), each  $\varphi_i$  is a convex potential (as a supremum of a family of affine functions). With no loss of generality, we may therefore restrict ourselves to the subset  $\mathcal{E}_c$  consisting of the elements  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}$  such that each potential  $\varphi_i$  is convex on  $\mathbb{R}^d$ .

By definition of  $\mathcal{E}$  and the duality result of theorem 2, we deduce that  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$  (respectively  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}$ ) solves  $(\mathcal{MK})$  (respectively solves

( $\mathcal{D}$ ) if and only if there exists  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}$  (respectively  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$ ) such that

$$\sum_{j=1}^d \varphi_j(x_j) = \det(x_1, \dots, x_d) \quad \gamma\text{-a.e.} \quad (21)$$

For all  $\Phi := (\varphi_1, \dots, \varphi_d) \in \mathcal{E}$ , let us define

$$K_\Phi := \{(x_1, \dots, x_d) \in (\mathbb{R}^d)^d : \sum_{j=1}^d \varphi_j(x_j) = \det(x_1, \dots, x_d)\}$$

and remark that the (possibly empty) set  $K_\Phi$  is closed since it is the minimal set of some l.s.c. function. Hence we deduce that (21) is equivalent to  $\gamma$  having its support included in  $K_\Phi$ . If  $\Phi \in \mathcal{E}_c$ , we have the following characterization of  $K_\Phi$ :

**Lemma 1** *Let  $\Phi := (\varphi_1, \dots, \varphi_d) \in \mathcal{E}_c$ , and  $x := (x_1, \dots, x_d) \in (\mathbb{R}^d)^d$ , then  $x \in K_\Phi$  if and only if for all  $i \in \{1, \dots, d\}$ :*

$$\sum_{j \neq i} \varphi_j(x_j) = \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j) \quad (22)$$

and

$$(-1)^{i+1} \bigwedge_{j \neq i} x_j \in \partial \varphi_i(x_i). \quad (23)$$

**Proof.** If  $x \in K_\Phi$  then

$$\sum_{j \neq i} \varphi_j(x_j) = \left\langle x_i, (-1)^{i+1} \bigwedge_{j \neq i} x_j \right\rangle - \varphi_i(x_i) \leq \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j)$$

which proves (22). Now let  $y_i \in \mathbb{R}^d$ , since  $x \in K_\Phi$  and  $\Phi \in \mathcal{E}_c$ , we have:

$$\varphi_i(y_i) - \varphi_i(x_i) \geq \left\langle y_i - x_i, (-1)^{i+1} \bigwedge_{j \neq i} x_j \right\rangle,$$

which proves (23).

Conversely assume that  $x$  satisfies (22)-(23) for some  $i$ . From (22), we get

$$\sum_{k=1}^d \varphi_k(x_k) = \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j) + \varphi_i(x_i)$$

with (23), this yields

$$\sum_{k=1}^d \varphi_k(x_k) = \left\langle x_i, (-1)^{i+1} \bigwedge_{j \neq i} x_j \right\rangle = \det(x),$$

so that  $x \in K_\Phi$ . □

Remark that in fact, for each fixed  $i$ ,  $K_\Phi$  consists of those  $(x \in \mathbb{R}^d)^d$  that satisfy (22)-(23) for that particular index  $i$ . We then immediately deduce the following:

**Proposition 3** *Let  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}_c$ , the following assertions are equivalent:*

1.  $(\varphi_1, \dots, \varphi_d)$  solves  $(\mathcal{D})$ ,
2. there exists  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$  such that for every  $i \in \{1, \dots, d\}$  and for  $\gamma$ -a.e.  $(x_1, \dots, x_d) \in (\mathbb{R}^d)^d$ , (22) and (23) hold,
3. there exists  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$  and  $i \in \{1, \dots, d\}$  such that for  $\gamma$ -a.e.  $(x_1, \dots, x_d)$  (22) and (23) hold,

If the marginals have additional regularity, we immediately deduce the following uniqueness result for the dual problem:

**Proposition 4** *If for every  $i \in \{1, \dots, d\}$ ,  $\mu_i$  is absolutely continuous with respect to  $\mathcal{L}^d$  and has a positive density with respect to  $\mathcal{L}^d$  then  $(\mathcal{D})$  admits a unique solution (up to the addition of constants summing to 0 to each potential).*

**Proof.** Let  $\gamma$  be a solution of  $(\mathcal{MK})$  and  $\Phi := (\varphi_1, \dots, \varphi_d)$  solve  $(\mathcal{D})$ , it follows from the duality relation that the support of  $\gamma$  is included in  $K_\Phi$ . Moreover, it can be assumed that each  $\varphi_i$  satisfies (19) hence  $\Phi \in \mathcal{E}_c$ . Let  $i \in \{1, \dots, d\}$ , and let us desintegrate  $\gamma$  with respect to its  $i$ -th marginal :  $\gamma = \mu_i \otimes \gamma^{x_i}$ . We deduce from lemma 1 that for almost every  $x_i \in \mathbb{R}^d$  and  $\gamma^{x_i}$  almost every  $(x_j)_{j \neq i} \in (\mathbb{R}^d)^{d-1}$  one has:

$$(-1)^{i+1} \bigwedge_{j \neq i} x_j \in \partial \varphi_i(x_i)$$

hence, by convexity

$$(-1)^{i+1} \int_{(\mathbb{R}^d)^{d-1}} \bigwedge_{j \neq i} x_j d\gamma^{x_i}((x_j)_{j \neq i}) \in \partial \varphi_i(x_i).$$

Since  $\varphi_i$  is differentiable  $\mu_i$ -a.e., we get that for  $\mu_i$ -a.e.  $x_i$ , one has

$$\nabla \varphi_i(x_i) = \int_{(\mathbb{R}^d)^{d-1}} \bigwedge_{j \neq i} x_j d\gamma^{x_i}((x_j)_{j \neq i})$$

since the rightmost member of this identity does not depend on  $\Phi$ , we are done. □

To sum up, getting back from  $(\mathcal{D})$  to  $(\mathcal{MK})$ , we obtain the following characterization of optimal transportation plans:

**Theorem 3** *Let  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$ ,  $\gamma$  solves  $(\mathcal{MK})$  if and only if there exists l.s.c. convex functions  $\varphi_i : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  such that for all  $i \in \{1, \dots, d\}$ :*

$$\sum_{j \neq i} \varphi_j(x_j) \geq \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j) \text{ on } (\mathbb{R}^d)^d, \quad (24)$$

$$\sum_{j \neq i} \varphi_j(x_j) = \varphi_i^*((-1)^{i+1} \bigwedge_{j \neq i} x_j) \text{ } \gamma\text{-a.e.}, \quad (25)$$

$$(-1)^{i+1} \bigwedge_{j \neq i} x_j \in \partial \varphi_i(x_i) \text{ } \gamma\text{-a.e.} \quad (26)$$

Once again, one can replace "for all  $i \in \{1, \dots, d\}$ " by "for some  $i \in \{1, \dots, d\}$ " and " $\gamma$ -a.e." by "on the support of  $\gamma$ " in the previous result. Of course, any  $(\varphi_1, \dots, \varphi_d)$  satisfying the previous statements is a solution of  $(\mathcal{D})$ .

To illustrate the previous considerations, let us consider the case  $d = 3$  and assume that  $(\varphi, \psi, \chi)$  is a (known) triple of convex potentials that solve  $(\mathcal{D})$ . For the sake of simplicity, also assume that the marginals  $(\mu_1, \mu_2, \mu_3)$  are absolutely continuous with respect to  $\mathcal{L}^3$ . Then any optimal transport plan  $\gamma$  is characterized by the extremality condition:

$$\varphi(x) + \psi(y) + \chi(z) = \det(x, y, z) \text{ on the support of } \gamma.$$

This implies that,  $\gamma$ -a.e., one has

$$\begin{cases} \psi(y) + \chi(z) = \varphi^*(y \wedge z) \\ \varphi(x) + \chi(z) = \psi^*(-x \wedge z) \\ \varphi(x) + \psi(y) = \chi^*(x \wedge y) \end{cases}$$

and

$$\begin{cases} \nabla \varphi(x) = y \wedge z \\ \nabla \psi(y) = -x \wedge z \\ \nabla \chi(z) = x \wedge y \end{cases} \quad (27)$$

Note that in particular, one has:

$$\begin{aligned} \langle x, \nabla \varphi(x) \rangle &= \langle y, \nabla \psi(y) \rangle = \langle z, \nabla \chi(z) \rangle \\ &= \det(x, y, z) = \det(\nabla \varphi(x), \nabla \psi(y), \nabla \chi(z)) \text{ } \gamma\text{-a.e.} \end{aligned}$$

The previous conditions clearly impose important geometric restrictions on  $\gamma$ . It implies in particular that for  $\mu_1$  almost every  $x$ , the conditional probability of  $y$  or  $z$  given  $x$  is supported by  $\nabla \varphi(x)^\perp$ . The conditional probability of  $z$  given  $(x, y)$  is even more constrained, indeed if  $\langle x, \nabla \varphi(x) \rangle \neq 0$  then the previous conditions impose:

$$z = \frac{\nabla \varphi(x) \wedge \nabla \psi(y)}{\langle x, \nabla \varphi(x) \rangle}.$$

## 4 The radial case

In this section, we focus on  $(\mathcal{MK})$  in the case where the measures  $(\mu_i)$  are radially symmetric. This means that for all  $1 \leq i \leq d$ , and for all  $R$  in the orthogonal group of  $\mathbb{R}^d$ ,

$$R\sharp\mu_i = \mu_i.$$

Then, introducing the measures  $\mu_i^r = |\cdot| \sharp \mu_i$ , we have, for all  $f \in C_c(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} f(x) d\mu_i(x) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} \left( \int_0^{+\infty} f(re) d\mu_i^r(r) \right) d\mathcal{H}^{d-1}(e). \quad (28)$$

In order to solve  $(\mathcal{MK})$ , let us first remark that it is intuitive that the potentials which solve the dual problem  $(\mathcal{D})$  are radially symmetric. Then, noticing that (27) generalizes to every dimension  $d$ , it follows that the support of an extremal measure will be included in the set of orthogonal systems. The only unknown here will be the relations between the norm of each vector. These relations will be obtained solving the following problem.

$$\sup_{\gamma^r \in \Pi(\mu_1^r, \dots, \mu_d^r)} \left( \int_{(\mathbb{R}^d)^d} \left( \prod_{i=1}^d r_i \right) d\gamma^r(r_1, \dots, r_d) \right). \quad (\mathcal{MK}^r)$$

**Proposition 5** *Assume that, for all  $1 \leq i \leq d$ ,  $\mu_i^r$  has no atom. Then, the problem  $(\mathcal{MK}^r)$  admits a unique solution  $\bar{\gamma}^r$  given by*

$$d\bar{\gamma}^r(r_1, \dots, r_d) = d\mu_1^r(r_1) \otimes \delta_{\{r_2=H_2(r_1)\}} \otimes \dots \otimes \delta_{\{r_d=H_d(r_1)\}},$$

where, for each  $2 \leq i \leq d$ ,  $H_i$  is the monotone rearrangement map of  $\mu_1^r$  into  $\mu_i^r$ .

**Proof.** Proceeding exactly as for  $(\mathcal{MK})$ , we get the existence of a minimizer  $\bar{\gamma}^r$  for  $(\mathcal{MK}^r)$ , together with the existence of a solution for its dual problem

$$\inf_{(\psi_1, \dots, \psi_d) \in \mathcal{E}^r} \left( \sum_{i=1}^d \int_0^{+\infty} \psi_i d\mu_i^r \right), \quad (\mathcal{D}^r)$$

where  $\mathcal{E}^r$  is the set of  $(\psi_1, \dots, \psi_d) \in (C_0(\mathbb{R}_+))^d$ , such that, for all  $(r_1, \dots, r_d)$  in  $(\mathbb{R}_+)^d$ ,

$$\sum_{i=1}^d \psi_i(r_i) \geq \prod_{i=1}^d r_i.$$

As for  $(\mathcal{MK})$ , if  $(\psi_1, \dots, \psi_d)$  is such a solution, we can assume that the  $\psi_i$ 's are l.s.c. convex functions on  $\mathbb{R}^+$ , and optimality conditions read as

- $\bar{\gamma}^r \in \Pi(\mu_1^r, \dots, \mu_d^r)$ .

- for  $\bar{\gamma}^r$ -a.e.  $(r_1, \dots, r_d) \in \text{supp}(\bar{\gamma}^r)$ ,

$$\forall 1 \leq i \leq d, \quad \psi'_i(r_i) = \prod_{j \neq i} r_j. \quad (29)$$

Let us introduce the functions  $G_i(r) = r\psi'_i(r)$ . Then, multiplying (29) by  $r_i$ , we get that, for  $\bar{\gamma}^r$ -a.e.  $(r_1, \dots, r_d)$  in  $\text{supp}(\bar{\gamma}^r)$ , and for all  $1 \leq i \leq d$ ,

$$G_i(r_i) = G_1(r_1).$$

Thanks to (29) again, it appears that each  $\psi'_i$  is non negative and then, the function  $G_i$  is non decreasing. It follows that, for  $\mu_1^r$ -a.e.  $r_1$ , and for all  $(r_2, \dots, r_d)$ ,

$$\forall 1 \leq i \leq d, \quad r_i = H_i(r_1) := (G_i^{-1} \circ G_1)(r_1),$$

where  $G_i^{-1}$  stands for the generalized inverse of  $G_i$ , and we get that the support of an optimal measure is necessarily in the closure of the graph of  $(H_i)_{2 \leq i \leq d}$ , where all the  $H_i$ 's are non decreasing. But, the constraint on  $\text{supp}(\bar{\gamma}^r)$  means exactly that, for all  $2 \leq i \leq d$ ,  $H_i$  pushes  $\mu_1^r$  forward to  $\mu_i^r$ . This ends the proof since, the measure  $\mu_1^r$  being non atomic, such a map is unique.  $\square$

The main result of the section is the following.

**Theorem 4** *Let  $(\mu_i)_{1 \leq i \leq d}$  be radially symmetric measures on  $\mathbb{R}^d$ , and assume that the measures  $\mu_i$  on  $\mathbb{R}^+$  have no atom. Define the measure  $\bar{\gamma}$  on  $(\mathbb{R}^d)^d$  by*

$$\bar{\gamma}(x) = \mu_1(x_1) \otimes \bar{\gamma}^{x_1}(x_2) \otimes \bar{\gamma}^{x_1, x_2}(x_3) \otimes \dots \otimes \bar{\gamma}^{x_1, \dots, x_{d-1}}(x_d),$$

with

$$\left\{ \begin{array}{l} \forall 1 \leq i \leq d-2, \quad \bar{\gamma}^{x_1, \dots, x_i} = \mathbf{1}_{H_i(|x_1|)\mathbb{S}(x_1, \dots, x_i)} \frac{d\mathcal{H}^{d-i-1}}{|\mathbb{S}^{d-i-1}|}, \\ \mathbb{S}(x_1, \dots, x_i) = \{x_1, \dots, x_i\}^\perp \cap \mathbb{S}^{d-1}, \\ \bar{\gamma}^{x_1, \dots, x_{d-1}} = \delta_{\{H_d(|x_1|) \wedge_{i=1}^{d-1} \left(\frac{x_i}{|x_i|}\right)\}}, \end{array} \right. \quad (30)$$

with the maps  $(H_i)_{2 \leq i \leq d}$  given by Proposition 5. Then,  $\bar{\gamma}$  is a solution to  $(\mathcal{MK})$ .

**Remark 3.** The previous solution is completely explicit since the  $H_i$ 's are.



**Remark 4.** As in section 2, the previous construction admits a very simple probabilistic interpretation. Indeed, the measure  $\bar{\gamma}$  is the law of a vector  $(X_1, \dots, X_d)$  of random vectors, such that, for all  $1 \leq i \leq d$ ,  $\mu_i$  is the law of  $X_i$ . Then, (30) means that the conditional law of  $X_i$  given  $X_1 = x_1, \dots, X_{i-1} = x_{i-1}$  is uniform on  $H(|x_1|)\mathbb{S}(x_1, \dots, x_{i-1})$  for  $1 \leq i \leq d-1$ , and  $X_d$  is given by

$$X_d = H_d(|X_1|) \bigwedge_{i=1}^{d-1} \left( \frac{X_i}{|X_i|} \right).$$

**Proof.** It is sufficient to prove that  $\bar{\gamma} \in \Pi(\mu_1, \dots, \mu_d)$  and that there exists  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}$  such that (21) holds for all  $(x_1, \dots, x_d)$  in  $\text{supp}(\bar{\gamma})$ . Notice that  $\text{supp}(\bar{\gamma})$  consists of the  $(x_1, \dots, x_d)$  such that

$$\begin{cases} (x_1, \dots, x_d) \text{ is an orthogonal basis of } \mathbb{R}^d, \\ \text{for } \mu_1 \text{ a.e. } x_1, \forall 2 \leq i \leq d, |x_i| = H_i(|x_1|). \end{cases} \quad (31)$$

Let us set

$$\forall 1 \leq i \leq d, \quad \varphi_i(x) = \psi_i(|x|),$$

where  $(\psi_1, \dots, \psi_d)$  is a solution to the radial dual problem  $(\mathcal{D}^r)$ . Then, for all  $(x_1, \dots, x_d)$  in  $(\mathbb{R}^d)^d$ ,

$$\sum_{i=1}^d \varphi_i(x_i) = \sum_{i=1}^d \psi(|x_i|) \geq \prod_{i=1}^d |x_i| \geq \det(x_1, \dots, x_d).$$

In addition, If  $(x_1, \dots, x_d)$  belongs to the support of  $\text{supp}(\bar{\gamma})$ , then by (31), we both have

$$\det(x_1, \dots, x_d) = \prod_{i=1}^d |x_i|$$

and  $(|x_i|)_{1 \leq i \leq d}$  belongs to the support of the solution  $\bar{\gamma}^r$  of  $(\mathcal{MK}^r)$  by proposition 5. Since  $(\psi_i)_{1 \leq i \leq d}$  is optimal in  $(\mathcal{D}^r)$ , it follows that

$$\det(x_1, \dots, x_d) = \sum_{i=1}^d \psi_i(|x_i|) = \sum_{i=1}^d \varphi_i(x_i).$$

To end the proof, we just need to prove that  $\bar{\gamma}$  has its marginals the  $\mu_i$ 's. The first marginal of  $\bar{\gamma}$  is obviously  $\mu_1$ . Let  $f \in C_0(\mathbb{R}^d)$ , then by definition of  $\bar{\gamma}$ ,

$$\int_{\mathbb{R}^d} f d(\pi_2 \# \bar{\gamma}) = \int_{\mathbb{R}^d} \left( \int_{H_2(|x_1|)\mathbb{S}(x_1)} f(x_2) \frac{d\mathcal{H}^{d-2}(x_2)}{H_2(|x_1|)^{d-2} |\mathbb{S}^{d-2}|} \right) d\mu_1(x_1).$$

Using the radially symmetry of  $\mu_1$ , performing the change of variables  $x_2 = H_2(|x_1|)y$  in the inner integral, and then applying Lemma 2, we get

$$\begin{aligned}
\int_{\mathbb{R}^d} f d(\pi_2 \# \bar{\gamma}) &= \int_0^\infty \left( \int_{\mathbb{S}^{d-1}} \left( \int_{\mathbb{S}(\sigma)} f(H_2(r)y) \frac{d\mathcal{H}^{d-2}(y)}{|\mathbb{S}^{d-2}|} \right) \frac{d\mathcal{H}^{d-1}(\sigma)}{|\mathbb{S}^{d-1}|} \right) d\mu_1^r(r) \\
&= \int_0^\infty \left( \int_{\mathbb{S}^{d-1}} f(H_2(r)y) \left( \int_{\mathbb{S}(y)} \frac{d\mathcal{H}^{d-2}(\sigma)}{|\mathbb{S}^{d-2}|} \right) \frac{d\mathcal{H}^{d-1}(y)}{|\mathbb{S}^{d-1}|} \right) d\mu_1^r(r) \\
&= \int_0^\infty \left( \int_{\mathbb{S}^{d-1}} f(H_2(r)y) \frac{d\mathcal{H}^{d-1}(y)}{|\mathbb{S}^{d-1}|} \right) d\mu_1^r(r) \\
&= \int_0^\infty \left( \int_{\mathbb{S}^{d-1}} f(\sigma) \frac{d\mathcal{H}^{d-1}(\sigma)}{|\mathbb{S}^{d-1}|} \right) d(H_2 \# \mu_1^r)(r) \\
&= \int_{\mathbb{R}^d} f d\mu_2,
\end{aligned}$$

by the definition of  $H_2$ . Repeated applications of this argument lead to the same result for the other marginals of  $\bar{\gamma}$  up to the  $(d-1)$ -th. For the last one, we proceed as in the proof of Proposition 1. For a function  $f \in C_0(\mathbb{R}^d)$ , by change of variables,

$$\begin{aligned}
&\int_{H_{d-1}(r)\mathbb{S}(x_1, \dots, x_{d-2})} f \left( H_d(r) \bigwedge_{i=1}^{d-1} \left( \frac{x_i}{|x_i|} \right) \right) \frac{d\mathcal{H}^1(x_{d-1})}{H_d(r)|\mathbb{S}^1|} \\
&= \int_{\mathbb{S}(x_1, \dots, x_{d-2})} f \left( H_d(r) \bigwedge_{i=1}^{d-1} \left( \frac{x_i}{|x_i|} \right) \right) \frac{d\mathcal{H}^1(x_{d-1})}{|\mathbb{S}^1|}, \\
&= \int_{\mathbb{S}(x_1, \dots, x_{d-2})} f(H_d(r)x_d) \frac{d\mathcal{H}^1(x_d)}{|\mathbb{S}^1|},
\end{aligned}$$

noticing that, if  $x_1, \dots, x_{d-2}$  are fixed, the map

$$x_{d-1} \mapsto \bigwedge_{i=1}^{d-1} \left( \frac{x_i}{|x_i|} \right)$$

is a rotation with angle  $\pi/2$  on  $\mathbb{S}(x_1, \dots, x_{d-2})$ . We can then repeat the argument used for the other marginals and then finish the proof.  $\square$

## 5 Volume maximization

So far, we have restricted our attention to the case where the objective function is the determinant although we were initially motivated with a volume maximization problem which corresponds to:

$$(\mathcal{MK})_a \quad \sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} |\det(x_1, \dots, x_d)| d\gamma(x_1, \dots, x_d).$$

instead of  $(\mathcal{MK})$ . In the radial case, treated in the previous section, we have seen that optimal measures for  $(\mathcal{MK})$  give full mass to the set of matrices with nonnegative determinant. In this case, there is no loss of generality in replacing  $(\mathcal{MK})_a$  by  $(\mathcal{MK})$  in the sense that solutions of  $(\mathcal{MK})$  also solve  $(\mathcal{MK})_a$ . This holds true under less restrictive symmetry assumptions on the marginals:

**Proposition 6** *Assume that  $(\gamma)$  is satisfied and that among the marginals  $(\mu_1, \dots, \mu_d)$ , two are symmetric (i.e.  $(-\text{id})\# \mu_i = \mu_i$ ,  $(-\text{id})\# \mu_j = \mu_j$  for two different indices  $i$  and  $j$ ) then any solution of  $\bar{\gamma}$  of  $(\mathcal{MK})$  satisfies  $\det(x_1, \dots, x_d) \geq 0$   $\bar{\gamma}$ -a.e. and actually also solves  $(\mathcal{MK})_a$ .*

**Proof.** Without loss of generality, assume that  $\mu_1$  and  $\mu_2$  are symmetric and that  $d \geq 3$ . Assume that  $(\varphi_1, \dots, \varphi_d) \in \mathcal{E}_c$  solves the dual problem  $(\mathcal{D})$ , then so does  $(\varphi_1 \circ (-\text{id}), \varphi_2 \circ (-\text{id}), \varphi_3, \dots, \varphi_d)$ . Hence  $(\psi_1, \dots, \psi_d)$  also solves  $(\mathcal{D})$  where  $\psi_1$  and  $\psi_2$  are the even part of  $\varphi_1$  and  $\varphi_2$  respectively and  $\psi_i = \varphi_i$  for  $i \geq 3$ . Since  $\psi_1$  is even we actually have for all  $(x_1, \dots, x_d)$ :

$$\psi_1(-x_1) + \sum_{j=2}^d \psi_j(x_j) = \sum_{i=1}^d \psi_i(x_i) \geq |\det(x_1, \dots, x_d)| \quad (32)$$

but, from the duality relations, one deduces that  $\bar{\gamma}$ -a.e.

$$\sum_{i=1}^d \psi_i(x_i) = \det(x_1, \dots, x_d) \geq 0.$$

To show that  $\bar{\gamma}$  solves  $(\mathcal{MK})_a$ , we use a classical duality argument. Indeed, let  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$ , from (32), one deduces:

$$\begin{aligned} \int_{\mathbb{R}^{d \times d}} |\det(x_1, \dots, x_d)| d\gamma(x_1, \dots, x_d) &\leq \sum_{i=1}^d \int \psi_i(x_i) d\mu_i(x_i) \\ &= \int_{\mathbb{R}^{d \times d}} |\det(x_1, \dots, x_d)| d\bar{\gamma}(x_1, \dots, x_d). \end{aligned}$$

□

## 6 Uniqueness issues

We end the paper by some remarks on uniqueness. For this, we will mainly use the very simple example of section 2. First, as noticed at the end of section 2, if the objective function is

$$x \mapsto |\det(x)|,$$

then there are infinitely many possible choices for the third random vector, given the first 2 ones. Indeed, for all  $x$  and  $y$  in the ball  $B$  with  $|x| = |y|$ , and for all  $p \in [0, 1]$ , then we can choose for the conditional law of  $Z$ , given  $X = x$  and  $Y = y$ :

$$p\delta_{\frac{x \wedge y}{|x|}} + (1 - p)\delta_{\frac{y \wedge x}{|x|}}.$$

Note that for the determinant, only the value  $p = 1$  is optimal and it corresponds to the solution given by Theorem 1. In the sequel, we shall work with the determinant as objective function.

As seen in the introduction, in dimension 2, since the problem  $(\mathcal{MK})$  can be reduced to the classical Monge-Kantorovich problem with quadratic cost, the solution is unique, at least for sufficiently regular measures. Actually, we have more, that is, up to a rotation, the optimal measure is supported by the graph of the gradient of a convex function  $\varphi$ . In terms of random pairs  $(X, Y)$ , the optimal coupling is characterized by the relation  $RY = \nabla\varphi(X)$ , with  $\varphi$  convex and  $R$  the rotation with angle  $-\pi/2$ . In higher dimensions, such uniqueness is lost.

First, let us consider the solution given by theorem 1. For  $x \in B$ , and given  $X = x$ , we define the conditional law of  $Y$  by the measure on  $|x|\mathbb{S}(x)$ , absolutely continuous with respect to the 1-dimensional Hausdorff measure, with density

$$\rho(y) = 1 + \frac{\langle y, e \rangle}{|y|},$$

where  $e$  is an arbitrary fixed vector in  $\mathbb{S}^2$ , which is a probability measure since we add to 1 a function which is greater than  $-1$  and has zero mean on any circle  $\mathbb{S}(x)$ . The vector  $Z$  is then chosen exactly as in Theorem 1. The proof of the optimality of this measure is the same as for Theorem 1.

Finally, we study the possibility for the extremal measure to have the same structure as in the 2-dimensional case, that is to be supported by the graph of a map  $T : \mathbb{R}^d \rightarrow (\mathbb{R}^d)^{d-1}$ . Such a  $T$  would then be called a "Monge" solution. Notice that the solution given by Theorems 1 and 4 does not have this structure but the question of the existence of such a solution is quite natural. Let us set the problem on  $\mathbb{R}^d$  with all the measures uniform on the ball, which is actually covered by Theorem 4. Then, according to the extremality conditions, the existence of a Monge solution for  $(\mathcal{MK})$  is equivalent to the existence of  $(d - 1)$  maps  $T_2, \dots, T_{d-1} : B \rightarrow B$  such that:

$$\begin{cases} \text{for a.e. } x \in B, \{x, T_2(x), \dots, T_d(x)\} \text{ is an orthogonal basis of } \mathbb{R}^d, \\ \text{for a.e. } x \in B, \forall 2 \leq i \leq d, |T_i(x)| = |x|, \\ \forall 2 \leq i \leq d, T_i \text{ preserves the Lebesgue measure.} \end{cases} \quad (33)$$

The construction of such maps seems to be very difficult in general and we left it open. However, in dimension 4, we have the following explicit

construction:

$$T_2(x) = \begin{pmatrix} -x_2 \\ x_1 \\ -x_4 \\ x_3 \end{pmatrix}, \quad T_3(x) = \begin{pmatrix} -x_3 \\ x_4 \\ x_1 \\ -x_2 \end{pmatrix}, \quad T_4(x) = \begin{pmatrix} -x_4 \\ -x_3 \\ x_2 \\ x_1 \end{pmatrix}.$$

In contrast, in dimension 3, it is not clear to us whether one can even find a single measure preserving map  $T : B \rightarrow B$  such  $T$  is norm-preserving and a.e. orthogonal to the identity map. If such a map exists, then by the hairy ball Theorem, it cannot be continuous.

## Appendix A

**Lemma 2** *Let  $k \geq 2$  be an integer and  $f \in C_0(\mathbb{S}^k \times \mathbb{S}^k)$ . Then,*

$$\int_{\mathbb{S}^k} \left( \int_{\mathbb{S}(x)} f(x, y) \frac{d\mathcal{H}^{k-1}(y)}{|\mathbb{S}^{k-1}|} \right) \frac{d\mathcal{H}^k(x)}{|\mathbb{S}^k|} = \int_{\mathbb{S}^k} \left( \int_{\mathbb{S}(y)} f(x, y) \frac{d\mathcal{H}^{k-1}(x)}{|\mathbb{S}^{k-1}|} \right) \frac{d\mathcal{H}^k(y)}{|\mathbb{S}^k|}, \quad (34)$$

where

$$\mathbb{S}(\xi) = \left\{ z \in \mathbb{S}^k; \langle z, \xi \rangle = 0 \right\}.$$

**Proof.** Let us introduce the following notations. Let us define the measure  $\mu$  on  $\mathbb{S}^k \times \mathbb{S}^k$  by

$$d\mu(x, y) = d\nu(x) \otimes d\theta^x(y),$$

where

$$d\nu(x) = \frac{d\mathcal{H}^k(x)}{|\mathbb{S}^k|} \quad \text{and} \quad d\theta^x(y) = \mathbf{1}_{\mathbb{S}(x)}(y) \frac{d\mathcal{H}^{k-1}(y)}{|\mathbb{S}^{k-1}|}.$$

The identity (34) expresses the fact that  $\mu$  is symmetric in  $x$  and  $y$ . We set, for  $x \in \mathbb{S}^k$  and  $\varepsilon > 0$ ,

$$\mathbb{S}_\varepsilon(x) = \left\{ z \in \mathbb{S}^k; |\langle z, x \rangle| \leq \varepsilon \right\}.$$

Let us notice first that  $\mathcal{H}^k(\mathbb{S}_\varepsilon(x))$  is independant of  $x$  and denote this quantity by  $\omega_\varepsilon$ . Indeed, for all  $x$  and  $x'$  in  $\mathbb{S}^k$ , let  $R$  be a rotation that maps  $x$  to  $x'$ . Then, we have

$$\mathbb{S}_\varepsilon(x') = \mathbb{S}_\varepsilon(Rx) = R^{-1}(\mathbb{S}_\varepsilon(x)),$$

and then  $\mathcal{H}^k(\mathbb{S}_\varepsilon(x')) = \mathcal{H}^k(\mathbb{S}_\varepsilon(x))$  since the Hausdorff measure is invariant under rotations. Next, we define the probability measure  $\theta_\varepsilon^x$  on  $\mathbb{S}^k$  as

$$d\theta_\varepsilon^x(y) = \mathbf{1}_{\mathbb{S}_\varepsilon(x)}(y) \frac{d\mathcal{H}^k(y)}{\omega_\varepsilon}.$$

For all  $f \in C_0(\mathbb{S}^k, \mathbb{S}^k)$ , and for all  $x \in \mathbb{S}^k$ ,

$$\begin{aligned} \left| \int_{\mathbb{S}^k} f(x, y) d\theta_\varepsilon^x(y) - \int_{\mathbb{S}^k} f(x, y) d\theta^x(y) \right| &\leq \left| \int_{\mathbb{S}^k} (f(x, y) - f(x, T^x(y))) d\theta_\varepsilon^x(y) \right| \\ &+ \left| \int_{\mathbb{S}^k} f(x, T^x(y)) d\theta_\varepsilon^x(y) - \int_{\mathbb{S}^k} f(x, y) d\theta^x(y) \right|, \end{aligned} \quad (35)$$

where

$$T^x(y) = \frac{\pi_x y}{\|\pi_x y\|},$$

and  $\pi_x$  is the orthogonal projection onto  $x^\perp$ . First, we shall estimate the first term in the right-hand side of (35). It is not difficult to see that if  $\varepsilon \leq 1/2$ , then for all  $x \in \mathbb{S}^k$  and  $y \in \mathbb{S}_\varepsilon(x)$ ,

$$\|y - T^x(y)\| \leq 4\varepsilon,$$

thus, denoting by  $\omega_f$  the modulus of continuity of  $f$ ,

$$\left| \int_{\mathbb{S}^k} (f(x, y) - f(x, T^x(y))) d\theta_\varepsilon^x(y) \right| \leq \omega_f(4\varepsilon).$$

For the second term in the right-hand side of (35), we write

$$\int_{\mathbb{S}^k} f(x, T^x(y)) d\theta_\varepsilon^x(y) = \int_{\mathbb{S}^k} f(x, z) d(T^x \# \theta_\varepsilon^x)(z),$$

where

$$T^x(y) = \frac{\pi_x y}{\|\pi_x y\|}.$$

Since  $T^x$  commutes with any rotation that leaves  $x$  invariant and  $\theta_\varepsilon^x$  is invariant under such a rotation, the measure  $T^x \# \theta_\varepsilon^x$  has its support in  $\mathbb{S}(x)$  and is invariant under rotation on  $\mathbb{S}(x)$ . So necessarily, we have

$$T^x \# \theta_\varepsilon^x = \theta^x$$

and

$$\left| \int_{\mathbb{S}^k} f(x, y) d\theta_\varepsilon^x(y) - \int_{\mathbb{S}^k} f(x, y) d\theta^x(y) \right| \leq \omega_f(4\varepsilon).$$

As a consequence,

$$\lim_{\varepsilon \rightarrow 0} \left( \sup_{x \in \mathbb{S}^k} \left| \int_{\mathbb{S}^k} f(x, y) d\theta_\varepsilon^x(y) - \int_{\mathbb{S}^k} f(x, y) d\theta^x(y) \right| \right) = 0. \quad (36)$$

We set  $d\mu_\varepsilon(x, y) = d\nu(x) \otimes d\theta_\varepsilon^x(y)$ . Then, (36) implies that  $(\mu_\varepsilon)_\varepsilon$  converges weakly to  $\mu$ . In addition,  $\mu_\varepsilon$  is symmetric in  $x$  and  $y$ , which ends the proof.  $\square$

## Appendix B

### Proof of theorem 2.

#### Step1 : existence of maximizers for $(\mathcal{MK})$

Let  $k \in \mathbb{N}$  and define:

$$\underline{H}_k := \max(\underline{H}, -k) \quad (37)$$

By construction,  $-k \leq \underline{H}_k \leq 0$  and  $\underline{H}_k$  decreases to  $\underline{H}$ .

Let  $\gamma_n$  be a maximizing sequence of  $(\mathcal{MK})$ . Since  $\Pi(\mu_1, \dots, \mu_d)$  is tight, it follows from Prohorov's Theorem, that (taking a subsequence if necessary) there exists  $\gamma \in \mathcal{M}(\mathbb{R}^{d \times d})$  such that for every continuous bounded function  $f$  on  $\mathbb{R}^{d \times d}$ ,  $\int_{\mathbb{R}^{d \times d}} f \gamma_n$  converges to  $\int_{\mathbb{R}^{d \times d}} f d\gamma$ . Clearly,  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$  and for every  $k$

$$\int_{\mathbb{R}^{d \times d}} \underline{H}_k d\gamma = \lim_n \int_{\mathbb{R}^{d \times d}} \underline{H}_k d\gamma_n \geq \lim_n \int_{\mathbb{R}^{d \times d}} \underline{H} d\gamma_n = \sup(\mathcal{MK}).$$

It thus follows from the monotone convergence theorem that  $\gamma$  is a solution of  $(\mathcal{MK})$ .

#### Step 2 : duality by approximation

Consider the problem:

$$(\mathcal{MK}_k) \quad \sup_{\gamma \in \Pi(\mu_1, \dots, \mu_d)} \int_{\mathbb{R}^{d \times d}} \underline{H}_k(x) d\gamma(x)$$

By the same considerations as in step 1,  $(\mathcal{MK}_k)$  admit solutions, let  $\gamma_k$  be such a solution. Now for  $\varepsilon >$ , let  $r = r_\varepsilon > 0$  be such that

$$\sum_{i=1}^d \mu_i(\mathbb{R}^d \setminus B_r^d) + \int_{\mathbb{R}^{d \times d} \setminus B_r^d} (1 + |H_0(x)|) d\gamma(x) \leq \varepsilon, \quad \forall \gamma \in \Pi(\mu_1, \dots, \mu_d). \quad (38)$$

Define then for  $i = 1, \dots, d$ :

$$\mu_i^{r,k} := \pi_i \# \left( \frac{\gamma_k \mathbb{1}_{B_r^d}}{\gamma_k(B_r^d)} \right).$$

Consider now:

$$(\mathcal{MK}_{r,k}) \quad \sup_{\gamma \in \Pi(\mu_1^{r,k}, \dots, \mu_d^{r,k})} \int_{\mathbb{R}^{d \times d}} \underline{H}_k(x) d\gamma(x)$$

and (defining  $\mathcal{E}_{r,k}$  as in (12) with  $\underline{H}_k$  instead of  $\det$ ) its dual:

$$(\mathcal{D}_{r,k}) \quad \inf_{(\varphi_1, \dots, \varphi_d) \in \mathcal{E}_{r,k}} \sum_{i=1}^d \int_{\mathbb{R}^d} \varphi_i d\mu_i^{r,k}.$$

We know from proposition 2 that the values of  $(\mathcal{MK}_{r,k})$  and  $(\mathcal{D}_{r,k})$  are equal and both attained. Moreover, using again proposition 2, the infimum in  $(\mathcal{D}_{r,k})$  is attained by some  $\varphi^{r,k} = (\varphi_1^{r,k}, \dots, \varphi_d^{r,k})$  such that:

$$\varphi_i^{r,k}(x_i) = \sup_{(x_j)_{j \neq i} \in B_r^{d-1}} \left\{ \underline{H}_k(x_1, \dots, x_d) - \sum_{j \neq i} \varphi_j^{r,k}(x_j) \right\}, \quad \forall x_i \in B_r, \quad (39)$$

and, for all  $x \in B_r$ :

$$-k \leq \varphi_1^{r,k}(x) \leq 0, \quad 0 \leq \varphi_i^{r,k}(x) \leq k, \quad i = 2, \dots, d. \quad (40)$$

Define then for all  $x_1 \in \mathbb{R}^d$ :

$$\psi_1^{r,k}(x_1) = \sup_{(x_2, \dots, x_d) \in B_r^{d-1}} \left\{ \underline{H}_k(x_1, \dots, x_d) - \sum_{j=2}^d \varphi_j^{r,k}(x_j) \right\} \quad (41)$$

Construct then inductively  $\psi_2^{r,k}, \dots, \psi_{d-1}^{r,k}$  by setting for  $i = 2, \dots, d-1$  and  $x_i \in \mathbb{R}^d$ ,

$$\begin{aligned} \psi_i^{r,k}(x_i) &= \sup \left\{ \underline{H}_k(x_1, \dots, x_d) - \sum_{j < i} \psi_j^{r,k}(x_j) - \sum_{j > i} \varphi_j^{r,k}(x_j), \right. \\ &\quad \left. (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \in \mathbb{R}^{d \times (i-1)} \times B_r^{d-i} \right\} \end{aligned}$$

and finally

$$\psi_d^{r,k}(x_d) = \sup_{(x_j)_{j \neq d} \in \mathbb{R}^{d \times (d-1)}} \left\{ \underline{H}_k(x_1, \dots, x_d) - \sum_{j=1}^{d-1} \psi_j^{r,k}(x_j) \right\}.$$

It can be checked easily that each  $\psi_i^{r,k}$  extends  $\varphi_i^{r,k}$  to the whole of  $\mathbb{R}^d$ . By (40) it can also be checked that  $\|\psi^{r,k}\|_\infty = O(k)$ . By construction  $\psi^{r,k} = (\psi_1^{r,k}, \dots, \psi_d^{r,k}) \in \mathcal{E}^k$ , where  $\mathcal{E}^k$  is the set of  $d$ -uples of lower-semi continuous functions  $(\varphi_1, \dots, \varphi_d)$  from  $\mathbb{R}^d$  to  $\mathbb{R} \cup \{+\infty\}$  such that

$$\sum_{i=1}^d \varphi_i(x_i) \geq \underline{H}_k(x_1, \dots, x_d), \quad \forall (x_1, \dots, x_d) \in \mathbb{R}^{d \times d}. \quad (42)$$

In other words,  $\mathcal{E}^k$  is the admissible set of the dual problem of  $(\mathcal{MK}_k)$ :

$$(\mathcal{D}_k) \quad \inf_{(\varphi_1, \dots, \varphi_d) \in \mathcal{E}^k} \sum_{i=1}^d \int_{\mathbb{R}^d} \varphi_i d\mu_i.$$

Let  $\gamma_{r,k}$  be a solution of  $(\mathcal{MK}_{r,k})$  and define:

$$\bar{\gamma}_{r,k} := \gamma_k(B_r^d) \gamma_{r,k} + \mathbf{1}_{\mathbb{R}^{d \times d} \setminus B_r^d}.$$



It is easy to verify that  $\bar{\gamma}_{r,k} \in \Pi(\mu_1, \dots, \mu_d)$  hence we get:

$$\begin{aligned} \sup(\mathcal{MK}_k) &= \int_{\mathbb{R}^{d \times d}} \underline{H}_k d\gamma_k \geq \int_{\mathbb{R}^{d \times d}} \underline{H}_k d\bar{\gamma}_{r,k} \\ &= \gamma_k(B_r^d) \int_{B_r^d} \underline{H}_k d\gamma_{r,k} + \int_{\mathbb{R}^{d \times d} \setminus B_r^d} \underline{H}_k d\gamma_k \\ &= \gamma_k(B_r^d) \left( \sum_{i=1}^d \int_{B_r} \varphi_i^{r,k} d\mu_i^{r,k} \right) + \int_{\mathbb{R}^{d \times d} \setminus B_r^d} \underline{H}_k d\gamma_k \end{aligned}$$

Using (38), we thus get:

$$\begin{aligned} \sup(\mathcal{MK}_k) &\geq \gamma_k(B_r^d) \left( \sum_{i=1}^d \int_{B_r} \varphi_i^{r,k} d\mu_i^{r,k} \right) - \varepsilon = \sum_{i=1}^d \int_{B_r^d} \varphi_i^{r,k} d\gamma_k - \varepsilon \\ &= \sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \psi_i^{r,k} d\gamma_k - \sum_{i=1}^d \int_{\mathbb{R}^{d \times d} \setminus B_r^d} \psi_i^{r,k} d\gamma_k - \varepsilon \geq \inf(\mathcal{D}_k) + O(k)\varepsilon. \end{aligned}$$

Letting  $\varepsilon$  go to 0 we then get  $\sup(\mathcal{MK}_k) \geq \inf(\mathcal{D}_k)$ . Using Prohorov's Theorem, taking a subsequence if necessary, we may assume that there exists  $\gamma \in \Pi(\mu_1, \dots, \mu_d)$  such that for every continuous bounded function  $f$  on  $\mathbb{R}^{d \times d}$ ,  $\int_{\mathbb{R}^{d \times d}} f \gamma_k$  converges to  $\int_{\mathbb{R}^{d \times d}} f d\gamma$ . Since  $\underline{H}_k \geq \underline{H} = \det -H_0$ , we then have for all  $k$ :

$$\sup(\mathcal{MK}_k) = \int_{\mathbb{R}^{d \times d}} \underline{H}_k d\gamma_k \geq \inf(\mathcal{D}) - a$$

Since  $\underline{H}_k$  is nonincreasing in  $k$ , for all  $k_0$ , we then have:

$$\int_{\mathbb{R}^{d \times d}} \underline{H}_{k_0} d\gamma = \lim_k \int_{\mathbb{R}^{d \times d}} \underline{H}_{k_0} d\gamma_k \geq \int_{\mathbb{R}^{d \times d}} \underline{H}_k d\gamma_k \geq \inf(\mathcal{D}) - a$$

taking the infimum in  $k_0$  and using the monotone convergence theorem we then get:

$$\sup(\mathcal{MK}) - a \geq \int_{\mathbb{R}^{d \times d}} \underline{H} d\gamma \geq \inf(\mathcal{D}) - a.$$

Since we already know that  $\sup(\mathcal{MK}) \leq \inf(\mathcal{D})$ , we have  $\sup(\mathcal{MK}) = \inf(\mathcal{D})$ .

### Step 3: existence of minimizers for $(\mathcal{D})$

It remains to prove that the infimum is attained in  $(\mathcal{D})$ . By the convexification trick, we can find a minimizing sequence of  $(\mathcal{D}_r)$ ,  $\varphi^n := (\varphi_1^n, \dots, \varphi_d^n)$  such that for all  $n$ , all  $i$  and all  $x_i \in \mathbb{R}^d$ , one has:

$$\varphi_i^n(x_i) = \sup_{(x_j)_{j \neq i} \in \mathbb{R}^{d \times d-1}} \left\{ \det(x_1, \dots, x_d) - \sum_{j \neq i} \varphi_j^n(x_j) \right\} \quad (43)$$

Note the functions  $\varphi_i^n$  are convex l.s.c. and not identically equal to  $+\infty$ . By (43),  $\varphi_i^n$  admits therefore some affine minorant hence  $\varphi_i^n(x) + |x|^{p_i}/p_i$  achieves its minimum. Using again that the objective in  $(\mathcal{D})$  is unchanged when changing  $(\varphi_1, \dots, \varphi_d)$  into  $(\varphi_1 + \alpha_1, \dots, \varphi_d + \alpha_d)$  for constants  $\alpha_i$  that sum to 0, we may then also assume that for all  $i = 1, \dots, d - 1$ :

$$\min_{x \in \mathbb{R}^d} \left\{ \varphi_i^n(x) + \frac{|x|^{p_i}}{p_i} \right\} = 0. \quad (44)$$

We easily deduce from the previous and (43) that

$$\varphi_d^n(x) + \frac{|x|^{p_d}}{p_d} \geq 0, \quad \forall x \in \mathbb{R}^d.$$

Let  $m > 0$  and define truncated potentials  $\varphi_i^{n,m}$  and  $\psi_i^{n,m}$  by:

$$\psi_i^{n,m}(x) = \varphi_i^{n,m}(x) + \frac{|x|^{p_i}}{p_i} = \min(\varphi_i^n(x) + \frac{|x|^{p_i}}{p_i}, m) \quad \forall x \in \mathbb{R}^d. \quad (45)$$

By construction,  $0 \leq \psi_i^{n,m} \leq m$  and there exists  $C$  such that for all  $(n, m)$ , one has:

$$\sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \psi_i^{n,m} d\mu_i \leq C. \quad (46)$$

and since  $\psi_i^{n,m} \geq 0$ , for all  $(x_1, \dots, x_d) \in \mathbb{R}^{d \times d}$ , one has:

$$\sum_{i=1}^d \psi_i^{n,m}(x_i) \geq \min(\overline{H}(x_1, \dots, x_d), m). \quad (47)$$

For fixed  $m$ , taking subsequence if necessary, we may assume that  $\psi_i^{n,m}$  weakly converges in  $L^1(\mu_i)$  to some limit  $\psi_i^m$  as  $n \rightarrow +\infty$ . By Mazur's Lemma, there is some sequence of convex combinations of the  $(\psi_i^{l,m})_{l \geq n}$  that converges in  $L^1(\mu_i)$  and (possibly after an extraction)  $\mu_i$ -almost everywhere to  $\psi_i^m$  as  $n \rightarrow +\infty$ , there is no loss of generality of setting  $\psi_i^m = +\infty$  outside the set where there is convergence. Hence, passing to the limit in (47), we get, for all  $(x_1, \dots, x_d) \in \mathbb{R}^{d \times d}$

$$\sum_{i=1}^d \psi_i^m(x_i) \geq \min(\overline{H}(x_1, \dots, x_d), m). \quad (48)$$

Now,  $\psi_i^m$  is a nondecreasing sequence of nonnegative functions that is bounded in  $L^1(\mu_i)$ . By the monotone convergence theorem,  $\psi_i^m$  converges pointwise and in  $L^1(\mu_i)$  to some function  $\psi_i$ . Defining  $\varphi_i(x) := \psi_i(x) - |x|^{p_i}/p_i$ , and passing to the limit in (48) first yields:

$$\sum_{i=1}^d \varphi_i^m(x_i) \geq \det(x_1, \dots, x_d), \quad \forall (x_1, \dots, x_d) \in \mathbb{R}^{d \times d}. \quad (49)$$

Since  $\psi_i^n \geq \psi_i^{n,m}$ , we also get:

$$\begin{aligned} \inf(\mathcal{D}) &= \lim_n \sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \varphi_i^n d\mu_i \geq \sup_m \left( \lim_n \sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \psi_i^{n,m} d\mu_i \right) - a \\ &= \sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \psi_i d\mu_i - a = \sum_{i=1}^d \int_{\mathbb{R}^{d \times d}} \varphi_i d\mu_i \end{aligned}$$

Thanks to (49), we can apply the convexification trick to  $(\varphi_1, \dots, \varphi_d)$ . We then obtain an element of  $\mathcal{E}$  that satisfies (43) and solves  $(\mathcal{D})$ .

□

## References

- [1] Y. Brenier. *Polar Factorization and monotone rearrangements of vector valued functions*, Communications in Pure and Applied Mathematics, **44** (1991).
- [2] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, Classics in Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, (1999).
- [3] W. Gangbo, A. Święch, *Optimal maps for the multidimensional Monge-Kantorovich problem*, Comm. Pure Appl. Math., **51**, no. 1, pp. 23–45 (1998).
- [4] W. Gangbo, R.J. Mc Cann. *The Geometry of Optimal Transportation*, Acta Math., vol. 177 (1996).
- [5] S.T. Rachev, L. Rüschendorf. *Mass Transportation Problems. Vol. I: Theory; Vol. II : Applications*, Springer-Verlag (1998).
- [6] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, (2003).