



**HAL**  
open science

# KANTOROVICH-RUBINSTEIN DUALITY THEORY FOR THE HESSIAN

Karol Bolbotowski, Guy Bouchitté

► **To cite this version:**

Karol Bolbotowski, Guy Bouchitté. KANTOROVICH-RUBINSTEIN DUALITY THEORY FOR THE HESSIAN. 2024. hal-04818344v1

**HAL Id: hal-04818344**

**<https://univ-tln.hal.science/hal-04818344v1>**

Preprint submitted on 4 Dec 2024 (v1), last revised 14 Dec 2024 (v2)

**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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

# KANTOROVICH-RUBINSTEIN DUALITY THEORY FOR THE HESSIAN

KAROL BOLBOTOWSKI AND GUY BOUCHITTÉ

ABSTRACT. The classical Kantorovich-Rubinstein duality theorem establishes a significant connection between Monge optimal transport and the maximization of a linear form on the set of 1-Lipschitz functions. This result has been widely used in various research areas, in particular, to expose the bridge between Monge transport theory and a class of optimal design problems. The aim of this paper is to present a similar theory when the linear form is maximized over all real  $C^{1,1}$  functions whose Hessian is between minus and plus identity matrix. It turns out that this problem can be viewed as the dual of a specific optimal transport problem. The task is to find a minimal three-point plan with the fixed first two marginals, while the third one must be larger than the other two in the sense of convex order. The existence of optimal plans allows to express solutions of the underlying Beckmann problem as a combination of rank-one tensor measures supported by a graph. In the context of two-dimensional mechanics, this graph encodes the optimal configuration of a grillage that transfers a given load system.

**Keywords:** Hessian-constrained problem, Monge optimal transport, tensor valued measures, duality, second-order Beckmann problem, convex order, stochastic dominance, optimal grillage.

**2020 Mathematics Subject Classification:** 49J45, 49K20, 28A50, 74P05

## 1. INTRODUCTION

The classical Kantorovich-Rubinstein duality theorem plays a fundamental role in Monge optimal transport theory. In the Euclidean framework, this theorem states that, for given probability measures  $\mu, \nu$  on  $\mathbb{R}^d$  with finite first-order moments, the Monge-Kantorovich distance

$$W_1(\mu, \nu) = \inf \left\{ \iint |x - y| \gamma(dx dy) : \gamma \in \Gamma(\mu, \nu) \right\}$$

coincides with the maximum in the following linear programming problem

$$\mathcal{I}_1(f) = \sup \left\{ \int u df : u \in C^{0,1}(\mathbb{R}^d), \text{lip}(u) \leq 1 \right\} \quad (1.1)$$

for the signed measure  $f = \nu - \mu$ . Above,  $\Gamma(\mu, \nu)$  stands for the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  whose first and second marginals coincide with  $\mu$  and  $\nu$ , respectively.

The equality  $W_1(\mu, \nu) = \mathcal{I}_1(f)$  is the key point to justify a PDE approach to the optimal transport problem. Moreover, it allows to interpret the Monge distance as the total variation of an optimal vector measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$  which solves the Beckmann's problem:

$$\min \left\{ \int |\sigma| : -\text{div} \sigma = f \right\}.$$

It turns out that problems of Beckmann type as above are directly connected to a class of optimal design problems where the unknown measure  $\sigma$  is associated to the heat flow in a conductor or the stress in a mechanical structure once they are subject to a given source  $f$ . This bridge was exposed in the year 1997 in the case of the compliance minimization problem for the scalar heat equation [6], further extended to the vector case of elasticity in [5]. The geometrical insights of the OT interpretation (through geodesics, transport rays) were very illuminating, as they allowed to derive the solutions to Beckmann's problem from the optimal transport plans  $\gamma$  using the decomposition formula:

$$\sigma = \iint \lambda^{x,y} \gamma(dxdy), \quad \lambda^{x,y} := \frac{y-x}{|y-x|} \mathcal{H}^1 \llcorner [x,y]. \quad (1.2)$$

In particular, it follows that any optimal measure  $\sigma$  is supported on the convex (geodesic) envelope of the support of  $f$ .

In view of recent applications, there is now a strong motivation to look for possible extensions of the Kantorovich-Rubinstein duality principle. Such a question could be formulated as follows. Let  $A$  be a linear differential operator on smooth vector-valued functions  $u : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$  such that  $Au : \Omega \rightarrow \mathbb{R}^n \times \mathbb{R}^d$ , and for  $\varrho$  take a semi-norm on real  $n \times d$  matrices. Is there an optimal transport formulation that we can exploit to address the following maximization problem:

$$\sup \left\{ \langle f, u \rangle : u \in C^\infty(\mathbb{R}^d; \mathbb{R}^n), \quad \varrho(Au) \leq 1 \text{ in } \Omega \right\}, \quad (1.3)$$

where  $\Omega$  is a domain in  $\mathbb{R}^d$ , and  $f$  is a suitable source term supported on  $\overline{\Omega}$ ? By classical duality, the supremum above can be written as the infimum in a Beckmann-type problem:

$$\min \left\{ \int \varrho^0(\sigma) : A^* \sigma = f \text{ in } (\mathcal{D}'(\mathbb{R}^d))^n, \text{ sp } \sigma \subset \overline{\Omega} \right\} \quad (1.4)$$

where  $\varrho^0$  is the polar of  $\varrho$  given by

$$\varrho^0(S) = \sup \{ \langle S, Q \rangle : \varrho(Q) \leq 1 \}. \quad (1.5)$$

In the Rubinstein-Kantorovich framework,  $f$  is a scalar measure ( $n = 1$ ),  $\varrho$  is the Euclidean norm and  $A$  is the gradient operator. In order that the supremum (1.3) is finite,  $f$  must be *balanced*, that is  $\mu = f_+$  and  $\nu = f_-$  must have the same mass. Note that if  $\Omega$  is not convex, the Euclidean distance appearing in the definition of  $W_1(\mu, \nu)$  should be replaced by the geodesic distance induced by  $\Omega$ . For a detailed study see [5] where many explicit examples are given.

When  $A$  is no longer the gradient operator, there are very few results suggesting a possible optimal transport approach. In the recent work [3], the present authors put forward a formulations where the Monge-Kantorovich distance emerges and is maximized with respect to a suitable class of metrics  $d(x, y)$  on  $\Omega$ . Therein, the potential  $u$  is a pair  $(v, w) : \Omega \rightarrow \mathbb{R}^d \times \mathbb{R}$ , and the non-linear operator  $(v, w) \mapsto e(v) + \frac{1}{2} \nabla v \otimes \nabla w$  plays the central role. This operator defines the strain tensor in the Föppl's membrane model [9], rendering the formulation of [3] an optimal membrane problem. Despite the non-linearity, the problem admits the form (1.4) upon the right choice of  $A$  and  $\varrho$ . Prior to the paper [3], an attempt to treat the case where  $A$  is simply the symmetric gradient (i.e.  $Au = e(u)$  for  $u : \Omega \rightarrow \mathbb{R}^d$ ) was undertaken in [8] with the initial motivation that the associated Beckmann-type problem is a measure-theoretic relaxation of the famous Michell problem. However, to our knowledge, the bridge with optimal transport in this case has not yet been established and challenging open problems remain.

The aim of the present paper is to provide an optimal transport approach in the case of the Hessian operator  $Au = \nabla^2 u$  and with  $\varrho$  being the spectral norm. We will limit ourselves to the case when  $\Omega = \mathbb{R}^d$ , and we will assume that  $f$  is a measure, more accurately  $f = \nu - \mu$  for two probabilities  $\mu, \nu$ . The special choice of  $\varrho$  makes it possible to rewrite (1.3) as the second-order counterpart of (1.1):

$$\mathcal{I}(f) := \sup \left\{ \int u d\nu - \int u d\mu : u \in C^{1,1}(\mathbb{R}^d), \text{lip}(\nabla u) \leq 1 \right\}. \quad (1.6)$$

It turns out that the supremum  $\mathcal{I}(f)$  is finite and attained if  $\mu$  and  $\nu$  have finite second-order moments, i.e.  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ , and a common barycentre  $[\mu] = [\nu]$ . This way  $\mathcal{I}(f)$  defines a distance between  $\mu$  and  $\nu$ . We should point out that similar distances (called *ideal metrics*) were introduced years ago by V.M. Zolotarev [25] with the aim of studying continuity and stability of stochastic models in probability theory.

Under the foregoing assumptions on the data  $\mu, \nu$ , we will see that the classical duality theory leads to a well-posed second-order Beckmann-type formulation:

$$\mathcal{I}'(f) := \min \left\{ \int \varrho^0(\sigma) : \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}), \text{div}^2 \sigma = \nu - \mu \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\}. \quad (1.7)$$

and to the equality  $\mathcal{I}(f) = \mathcal{I}'(f)$ . Here  $\varrho^0$  is the Schatten norm on symmetric tensors  $S \in \mathcal{S}^{d \times d}$  given by  $\varrho^0(S) = \sum_{i=1}^d |\lambda_i(S)|$ , and  $\int \varrho^0(\sigma)$  is intended in the sense of convex one-homogeneous functionals on measures [13]. By adapting classical methods, optimality conditions involving pairs  $(u, \sigma)$  can be derived, even in the case of singular measures  $\sigma$  and for general semi-norms  $\varrho$ . For a detailed study we refer for instance to [7] where optimal design problems for plates are considered.

The main novelty of the paper is a connection between the pair (1.6), (1.7) and a new three-marginal optimal transport problem. Let us introduce the cost function defined for the triples  $(x, y, z) \in (\mathbb{R}^d)^3$  by

$$c(x, y, z) := \frac{1}{2} \left( |z - x|^2 + |z - y|^2 \right) \quad (1.8)$$

and look for probabilities  $\pi \in \mathcal{P}_2((\mathbb{R}^d)^3)$  (three-marginal transport plans) solving:

$$\mathcal{J}(\mu, \nu) := \inf \left\{ \iiint c(x, y, z) \pi(dx dy dz) : \pi \in \Sigma(\mu, \nu) \right\}, \quad (1.9)$$

where  $\Sigma(\mu, \nu)$  denotes the subset of  $\mathcal{P}_2((\mathbb{R}^d)^3)$  consisting of 3-plans  $\pi$  whose first and second two marginal agrees with, respectively,  $\mu$  and  $\nu$ , and which satisfy the following equilibrium equations:

$$\iiint \langle z - x, \Phi(x) \rangle \pi(dx dy dz) = \iiint \langle z - y, \Psi(y) \rangle \pi(dx dy dz) = 0 \quad (1.10)$$

for any smooth test functions  $\Phi, \Psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ . As will be seen later, these relations have a natural interpretation in probability theory (convex order) as well as in mechanics (equilibrium condition for a plate in the bending regime). The main result of the paper can be summarized as follows:

**Theorem 1.1.** *Assume  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  sharing the barycentre  $[\mu] = [\nu]$ , and let  $\mathcal{J}(\mu, \nu)$  be defined by (1.9). For  $f = \nu - \mu$  the value  $\mathcal{I}(f)$  is given by (1.6). Then:*

(i) *the equality*

$$\mathcal{I}(f) = \mathcal{J}(\mu, \nu) \quad (1.11)$$

*holds true, while there exist optimal pairs  $(u, \pi)$  solving (1.6) and (1.9), respectively;*

(ii) an admissible pair  $(u, \pi)$  is optimal if and only if the following three-point equality is satisfied  $\pi$ -a.e., with  $c$  defined by (1.8),

$$[u(y) + \langle \nabla u(y), z - y \rangle] - [u(x) + \langle \nabla u(x), z - x \rangle] = c(x, y, z) \quad \text{for } \pi\text{-a.e. } (x, y, z). \quad (1.12)$$

It is worth noticing that the equality (1.12) is in close relation with the admissibility of  $u$  in (1.6). Indeed, following [2], the condition  $\text{lip}(\nabla u) \leq 1$  is equivalent to the the existence of a continuous vector function  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that:

$$[u(y) + \langle \Phi(y), z - y \rangle] - [u(x) + \langle \Phi(x), z - x \rangle] \leq c(x, y, z) \quad \forall (x, y, z) \in (\mathbb{R}^d)^3. \quad (1.13)$$

In addition, the inequality (1.13) can be satisfied only for  $\Phi = \nabla u$ .

As a consequence of Theorem 1.1, we derive the tensor counterpart of the decomposition (1.2) that was valid in the first order gradient case. We show that the optimal measures for the second-order Beckmann problem (1.7) can be decomposed to rank-one tensor measures supported on a network of polygonal lines. More precisely, let us define for any triple  $(x, y, z)$  the following measure valued in the space of symmetric tensors, i.e. an element of  $\mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ :

$$\sigma^{x,y,z}(d\xi) = |\xi - z| \left( \sigma^{z,x}(d\xi) - \sigma^{z,y}(d\xi) \right), \quad (1.14)$$

where we have set

$$\sigma^{a,b} := \frac{b-a}{|b-a|} \otimes \frac{b-a}{|b-a|} \mathcal{H}^1 \llcorner [a, b]. \quad (1.15)$$

Then,  $\sigma^{x,y,z}$  (see Fig. 1) is supported on the set  $[x, z] \cup [z, y]$ , while  $\text{div}^2 \sigma^{x,y,z} = f^{x,y,z}$ , where

$$f^{x,y,z} := \delta_y - \delta_x - \text{div}((z-y)\delta_y - (z-x)\delta_x) \quad (1.16)$$

includes a first-order distribution term. An important observation is that  $\sigma^{x,y,z}$  solves (1.7) for  $f = f^{x,y,z}$  with the equality  $\mathcal{I}(f^{x,y,z}) = c(x, y, z)$  provided that  $z$  belongs to the ball  $B(\frac{x+y}{2}, \frac{|x-y|}{2})$  (see Proposition 4.6). Accordingly, our strategy for solving (1.7) for  $f = \nu - \mu$  is to search for optimal tensor measures  $\sigma$  in the form:

$$\sigma = \iiint \sigma^{x,y,z} \pi(dxdydz)$$

where  $\pi$  is a suitable three-marginal plan, see the convention (1.21) below. If we restrict ourselves to those  $\pi$  which admit  $\mu, \nu$  as first and second marginals, then the differential constraint  $\text{div}^2 \sigma = \iiint f^{x,y,z} d\pi = \nu - \mu$  corresponds exactly to the admissibility condition  $\pi \in \Sigma(\mu, \nu)$  which appears in (1.9).

**Corollary 1.2.** *Let  $\bar{\pi}$  be an optimal plan for (1.9). Then,*

(i) *the tensor measure  $\bar{\sigma} = \iiint \sigma^{x,y,z} \bar{\pi}(dxdydz)$  solves the second-order Beckmann's problem (1.7) for  $f = \nu - \mu$ ;*

(ii) *let  $\bar{\gamma} := \Pi_{1,2}^\#(\bar{\pi})$  be the marginal of  $\bar{\pi}$  with respect to the first two variables, and let  $\bar{u}$  be any solution of (1.6). Then, for any  $\bar{\pi}$ -integrable test function  $\varphi : (\mathbb{R}^d)^3 \rightarrow \mathbb{R}$ , we have the disintegration formula:*

$$\iiint \varphi d\bar{\pi} = \iint \varphi(x, y, z_{\bar{u}}(x, y)) \bar{\gamma}(dxdy)$$

where

$$z_{\bar{u}}(x, y) := \frac{x+y}{2} + \frac{\nabla \bar{u}(y) - \nabla \bar{u}(x)}{2}. \quad (1.17)$$

As a result, the optimal measure  $\bar{\sigma}$  is supported on the closed subset

$$\mathcal{B}(\text{sp } \mu, \text{sp } \nu) := \bigcup \left\{ B\left(\frac{x+y}{2}, \frac{|x-y|}{2}\right) : (x, y) \in \text{sp } \mu \times \text{sp } \nu \right\}. \quad (1.18)$$

We observe that, in the first-order gradient case, a geometric bound on the support of any optimal measure can be recovered from (1.18) if we replace the ball on the right-hand side by the line segment  $[x, y]$ . In contrast, in the Hessian case, a larger set is needed to cover the support of possible optimal measures. Thus, we refute the conjecture in [7] where, assuming mild conditions on the norm  $\rho^0$  entering (1.7), it was suggested that optimal measures are supported on the convex hull of the source  $f$ .

The proof of Theorem 1.1 relies on an unexpected connection between our three-marginal optimal transport formulation (1.9) and optimization under the convex order dominance conditions. This link rests upon the following observation. The existence of a 3-plan  $\pi \in \Sigma(\mu, \nu)$  which admits  $\rho$  as the third marginal is equivalent to the conditions of convex order  $\rho \succeq_c \mu$ ,  $\rho \succeq_c \nu$ , that is:

$$\int \varphi d\rho \geq \max \left\{ \int \varphi d\mu, \int \varphi d\nu \right\} \quad \text{for all convex } \varphi : \mathbb{R}^d \rightarrow \mathbb{R}. \quad (1.19)$$

**Theorem 1.3.** *Let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  be probability measures satisfying  $[\mu] = [\nu]$ , and set:*

$$\mathcal{V}(\mu, \nu) = \inf \left\{ \text{var}(\rho) : \rho \in \mathcal{P}_2(\mathbb{R}^d), \rho \succeq_c \mu, \rho \succeq_c \nu \right\}, \quad (1.20)$$

where  $\text{var}(\rho)$  is the variance of  $\rho$ .

(i) *The following equality holds true:*

$$\mathcal{J}(\mu, \nu) = \mathcal{V}(\mu, \nu) - \frac{\text{var}(\mu) + \text{var}(\nu)}{2}.$$

Moreover, an admissible 3-plan  $\pi \in \Sigma(\mu, \nu)$  is optimal for (1.9) if and only if its third marginal  $\rho$  is a minimizer in (1.20).

(ii) *The infimum in (1.20) is achieved. Moreover, to any minimal  $\rho$  we can associate at least one 3-plan  $\pi$  that solves (1.9) and whose three marginals are  $\mu, \nu$ , and  $\rho$ , subsequently.*

Here, several comments are in order. First, let us explain how the assertion (i) will eventually facilitate the proof of the central equality (1.11). With  $f = \nu - \mu$ , we shall prove the relation between  $\mathcal{I}(f)$  and  $\mathcal{V}(\mu, \nu)$  passing through a duality result (Proposition 3.3) which allows to identify  $\mathcal{V}(\mu, \nu)$  as the supremum in the following auxiliary problem:

$$\mathcal{V}'(\mu, \nu) := \sup \left\{ \int \varphi d\mu + \int \psi d\nu : \varphi, \psi \text{ are } C^{1,1} \text{ convex functions, } \varphi + \psi \leq |\cdot|^2 \right\}.$$

Subsequently, the relation between  $\mathcal{V}'(\mu, \nu)$  and  $\mathcal{I}(f)$  comes out directly through manipulations on convex functions, see Proposition 3.5.

On another note, the construction of an optimal plan  $\bar{\pi}$  announced in the assertion (ii) will use two martingale optimal transports: one between  $\mu$  and  $\rho$ , and the other between  $\nu$  and  $\rho$ . Their existence follows from the classical theorem of Strassen. An optimal  $\bar{\pi}$  can be then constructed via a gluing argument, see the statement and the proof of Lemma 3.13. In general, it is not unique.

Let us point out that the problem (1.20) falls within a larger class of stochastic optimization problems under dominance constraint, for which there exist many applications in mathematical finance, statistical decision theory, or economics. See, for instance, the recent papers [10], [20], [24].

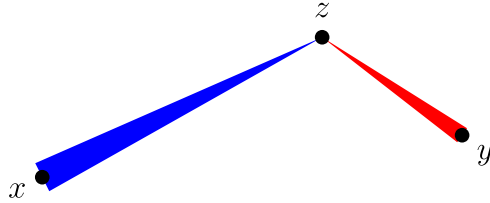


FIGURE 1. The tensor measure  $\sigma^{x,y,z}$ ; density with respect to  $\mathcal{H}^1 \llcorner ([x, z] \cup [z, y])$  is illustrated. Blue and red indicate the positive and the negative part, respectively.

We close this introduction with a comment about the close relation between the results presented in this paper and an optimal design problem in mechanics when  $d = 2$ . Any measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{2 \times 2})$  that satisfies the equation  $\operatorname{div}^2 \sigma = f$  represents a *bending moment tensor* in a plate that is subject to a load  $f$ . If the measure  $\sigma$  is of the form  $\iiint \sigma^{x,y,z} d\pi$ , we speak of a *grillage* – a particular plate that decomposes to straight bars. The bars exhibit linearly varying rank-one bending moments, see Fig. 1 demonstrating the basic two-bar measure  $\sigma^{x,y,z}$ . A natural issue studied in the literature [22, 21, 4] consists in finding an optimal configuration of the grillage, i.e. a coupling  $\bar{\pi}$  that minimizes a certain total energy functional. To date, however, the existence result was not available, and neither were the criteria for the finite support of  $\bar{\pi}$ , which corresponds to practical designs in the form of finite systems of bars.

In this work we show that, when the load  $f$  is a measure, an optimal grillage can be recast by solving the new three-marginal optimal transport formulation (1.9). Moreover, a finitely supported optimal 3-plan  $\bar{\pi}$  can be selected provided that  $f$  is also finitely supported. As a byproduct, we get the bound  $\mathcal{B}(\operatorname{sp} f_+, \operatorname{sp} f_-)$  on the support of the associated tensor measure  $\bar{\sigma}$ . It should be noted that, despite a similarity to the optimal grillage problem, there is no such OT reformulation for the more popular *optimal truss problem*. In fact, it has been known for 120 years that optimal trusses do not exist even for the simplest load data [17]. In this case, a relaxation in the form of the famous Michell formulation [8, 16] is essential. On top of that, a geometric bound on the support of its solutions is still pending.

Finally, we stress that our results concerning optimal grillages do not immediately extend to the case of a source  $f$  containing a first-order distribution term, or to the case when the support of the induced stress  $\bar{\sigma}$  is confined within a given domain  $\Omega \subset \mathbb{R}^d$ . Such extensions are beyond the scope of this paper and are worthy of future study.

The paper is organized as follows. In the preliminary Section 2 we adapt the classical duality theory to show the no-gap equality (1.7) as well as the existence of an optimal pair  $(u, \sigma)$ . Besides, in view of the forthcoming connection with the stochastic optimization, we give a short background on convex order and its relation with martingale transport. The Section 3 is devoted to the proofs of Theorem 1.1, Corollary 1.2, and Theorem 1.3. In Section 4, we give a series of examples where optimal configurations are determined explicitly. The final Section 5 is devoted to the underlying 2D optimal design formulation. Numerical examples of optimal grillages are given and discussed along with the related open questions.

**Acknowledgements.** The first author would like to thank the Laboratoire IMATH, Université de Toulon for hosting his visit in June 2022. The second author was partially supported by the French project ANR-23-CE40-0017. Both authors are grateful to the Lagrange Mathematics and Computing Research

Center in Paris for facilitating fruitful discussions with Guillaume Carlier, Quentin Mérigot, and Filippo Santambrogio. The authors are thankful for their remarks which have influenced this paper.

**Notations.** Throughout the paper we will use the following notations.

- The Euclidean norm of  $z \in \mathbb{R}^d$  is denoted by  $|z|$ .
- By  $\mathcal{S}^{d \times d}$  we shall denote the space of  $d \times d$  symmetric matrices, while  $\mathcal{S}_+^{d \times d}$  will be its subset whose elements are positive semi-definite. Given  $A, B \in \mathcal{S}^{d \times d}$ , we will write  $A \leq B$  if  $B - A \in \mathcal{S}_+^{d \times d}$ . Moreover,  $\text{Tr}A$  stands for the trace of  $A$ , while  $\text{Id}$  is the identity matrix.
- For natural  $k \in \mathbb{N} \cup \{+\infty\}$ ,  $C^k(\mathbb{R}^d)$  is the spaces of functions on  $\mathbb{R}^d$  that are continuously differentiable up to order  $k$ , while  $\nabla u$  and  $\nabla^2 u$  are the gradient and the Hessian of a function  $u \in C^1(\mathbb{R}^d)$  and  $u \in C^2(\mathbb{R}^d)$ , respectively. Moreover,  $C_0(\mathbb{R}^d) \subset C^0(\mathbb{R}^d)$  denotes the subset of continuous functions that vanish at infinity.
- $\mathcal{D}(\mathbb{R}^d)$  denotes the space of  $C^\infty$  functions that are compactly supported, and  $\mathcal{D}'(\mathbb{R}^d)$  is the space of distributions on  $\mathbb{R}^d$  (the dual of  $\mathcal{D}(\mathbb{R}^d)$ ).
- For a function  $v : \mathbb{R}^d \rightarrow \mathbb{R}^n$ ,  $\text{lip}(v)$  stands for the Lipschitz constant equal to  $\sup_{x \neq y} \frac{|v(x) - v(y)|}{|x - y|}$ .
- By  $C^{0,1}(\mathbb{R}^d)$  (resp.  $C^{1,1}(\mathbb{R}^d)$ ) we understand the Banach space of these functions  $u \in C^0(\mathbb{R}^d)$  (resp.  $u \in C^1(\mathbb{R}^d)$ ) for which  $\text{lip}(u) < +\infty$  (resp.  $\text{lip}(\nabla u) < +\infty$ ).
- For a natural  $k$ ,  $W_{\text{loc}}^{k,\infty}(\mathbb{R}^d)$  is the space of functions  $u$  that belong to the Sobolev space  $W^{k,\infty}(\Omega)$  for any pre-compact domain  $\Omega \subset \mathbb{R}^d$ . For  $u \in W_{\text{loc}}^{2,\infty}(\mathbb{R}^d)$ , the weak Hessian is denoted by  $\nabla^2 u$ .
- $\mathcal{M}_+(\mathbb{R}^d)$  denotes the space of Borel measures on  $\mathbb{R}^d$  with values in  $[0, +\infty]$ . The Banach space of Borel measures valued in a finite dimensional normed vector space  $E$  is denoted by  $\mathcal{M}(\mathbb{R}^d; E)$ . In addition, we agree that  $\mathcal{M}(\mathbb{R}^d) := \mathcal{M}(\mathbb{R}^d; \mathbb{R})$ .
- The topological support of  $\mu \in \mathcal{M}(\mathbb{R}^d; E)$  is denoted by  $\text{sp } \mu$ , while  $\mu \llcorner A$  is the restriction to a Borel subset  $A \subset \mathbb{R}^d$ . By the symbol  $\mu \ll \nu$  one understands the absolute continuity of a measure  $\mu$  with respect to  $\nu \in \mathcal{M}_+(\mathbb{R}^d)$ .
- For a measure  $\mu$  and a  $\mu$ -measurable map  $T$ , by  $T^\#(\mu)$  we understand the push forward, i.e.  $T^\#(\mu)(B) := \mu(T^{-1}(B))$  for every Borel set  $B$ .
- $\mathcal{P}(\mathbb{R}^d) := \{\mu \in \mathcal{M}_+(\mathbb{R}^d) : \mu(\mathbb{R}^d) = 1\}$  is the set of probabilities on  $\mathbb{R}^d$ .
- For  $\gamma \in \mathcal{P}(\mathbb{R}^d \times \dots \times \mathbb{R}^d)$  on the product of  $n$  ambient spaces, by  $\gamma_{k_1, \dots, k_m}$  we understand the marginal  $\Pi_{k_1, \dots, k_m}^\#(\gamma)$  where, for  $m \leq n$ ,  $\Pi_{k_1, \dots, k_m}$  is the projection onto the coordinates  $k_1, \dots, k_m$ .
- Assume  $\mu \in \mathcal{M}_+(\mathbb{R}^n)$  and a map  $x \mapsto \lambda^x \in \mathcal{M}(\mathbb{R}^m; E)$  that is  $\mu$ -measurable in the sense that  $x \mapsto \lambda^x(A)$  is  $\mu$ -measurable for any Borel set  $A \subset \mathbb{R}^m$ . Provided that  $\int |\lambda^x|(\mathbb{R}^d) \mu(dx) < +\infty$ , we will use the notation

$$\nu = \int \lambda^x \mu(dx), \quad \gamma = \mu \otimes \lambda^x \quad (1.21)$$

to define measures  $\nu \in \mathcal{M}(\mathbb{R}^m; E)$  and  $\gamma \in \mathcal{M}(\mathbb{R}^n \times \mathbb{R}^m; E)$  that satisfy

$$\nu(A) := \int \lambda^x(A) \mu(dx), \quad \gamma(B) := \int \left( \int \chi_B(x, y) \lambda^x(dy) \right) \mu(dx)$$

for every Borel sets  $A \subset \mathbb{R}^m$  and  $B \subset \mathbb{R}^n \times \mathbb{R}^m$ , where  $\chi_B$  is the characteristic function of the latter.

- $[\mu]$  stands for the barycentre of a probability  $\mu \in \mathcal{P}(\mathbb{R}^d)$  of finite first-order moments, whilst  $\text{var}(\mu) := \int |x - [\mu]|^2 \mu(dx)$  is its variance.



- $\nu \succeq_c \mu$  denotes the convex order between two probability measures  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$  of finite first-order moments.
- $\mu \star \nu$  stands for the convolution of two probabilities  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ .
- $\langle \cdot, \cdot \rangle$  shall be used to denote a canonical scalar product in a finite dimensional space of vectors or matrices, whilst in the case of infinite dimensional spaces it will stand for the duality bracket.
- The double distributional divergence  $\text{div}^2$  of a matrix measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$  is an element of  $\mathcal{D}'(\mathbb{R}^d)$  that is defined as follows:

$$\text{div}^2 \sigma = f \quad \text{in } \mathcal{D}'(\mathbb{R}^d) \quad \Leftrightarrow \quad \int \langle \nabla^2 \varphi, \sigma \rangle = \langle \varphi, f \rangle \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^d). \quad (1.22)$$

- Given a tensor-valued measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ ,  $\int \rho^0(\sigma)$  will denote the integral in the sense of the Goffman-Serrin convention [13], namely,

$$\int \rho^0(\sigma) := \int \rho \left( \frac{d\sigma}{d\theta} \right) d\theta,$$

where  $\theta$  is any non-negative Radon measure  $\theta \in \mathcal{M}_+(\mathbb{R}^d)$  such that  $\sigma \ll \theta$ . Due to the one-homogeneity of  $\rho^0$ , the above expression does not depend on  $\theta$ .

## 2. PRELIMINARIES

**2.1. The classical duality framework.** The duality theory involving the linear constraint problem  $\mathcal{I}(f)$  in (1.3) and the general Beckmann's formulation (1.4) is well understood in the case of the Hessian operator as far as  $\Omega$  is a bounded domain of  $\mathbb{R}^d$  and when  $\varrho$  is any norm on  $\mathcal{S}^{d \times d}$  (see for instance [7]). Since we are concerned with the case  $\Omega = \mathbb{R}^d$ , some specific functional spaces are needed for proving the existence of solutions and for using suitable duality arguments. Therefore, in addition to the general notations already given in the introduction, for  $p \geq 1$  we introduce the following Banach spaces:

- $\mathcal{M}_p(\mathbb{R}^d)$ , the space of Borel signed measures  $\mu$  on  $\mathbb{R}^d$  such that  $\|\mu\|_p := \int (1 + |x|^p) |\mu|(dx) < +\infty$ . Then,  $\mathcal{P}_p(\mathbb{R}^d)$  denotes the subset of  $\mathcal{M}_p(\mathbb{R}^d)$  consisting of probability measures with finite  $p$ -moment. The definition extends naturally to  $\mathcal{M}_p(\mathbb{R}^d; E)$ , where  $E$  is a finite dimensional normed vector space.
- $X_p(\mathbb{R}^d)$ , the set of continuous functions  $u \in C^0(\mathbb{R}^d)$  such that  $\|u\|_{X_p} := \sup \frac{|u(x)|}{1+|x|^p} < +\infty$ . The closed subspace  $X_{p,0}$ , consisting of those  $u$  such that  $\lim_{|x| \rightarrow +\infty} \frac{|u(x)|}{1+|x|^p} = 0$ , is a separable Banach space.

A pairing between  $X_p(\mathbb{R}^d)$  and  $\mathcal{M}_p(\mathbb{R}^d)$  is defined by  $\langle u, \mu \rangle = \int_{\mathbb{R}^d} u d\mu$ . Noticing that  $X_{p,0} = \frac{1}{1+|\cdot|^p} C_0$ , it is easy to see that the topological dual of  $X_{p,0}$  can be identified with  $\mathcal{M}_p(\mathbb{R}^d)$  through this duality bracket. As a consequence of dominated convergence, we have a useful convergence criterium for a sequence  $(v_n)$  in  $X_p(\mathbb{R}^d)$ , namely:

$$\sup_n \|v_n\|_{X_p} < +\infty \quad \text{and} \quad \lim_{n \rightarrow \infty} v_n(x) = 0 \quad \forall x \in \mathbb{R}^d \quad \Rightarrow \quad \langle v_n, \mu \rangle \rightarrow 0 \quad \forall \mu \in \mathcal{M}_p(\mathbb{R}^d). \quad (2.1)$$

The next result applies to general first order distributional source terms of the kind  $f = f_0 - \text{div} F$ , where  $(f_0, F)$  is any pair in  $\mathcal{M}_2(\mathbb{R}^d) \times \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d)$  such that the following balance condition is met:

$$\int f_0 = 0, \quad \int x f_0 + \int F = 0. \quad (2.2)$$

The two conditions mean that  $f$  is orthogonal to affine functions, which is clearly necessary for the finiteness of (1.6), namely,

$$\mathcal{I}(f) := \sup \left\{ \langle u, f \rangle : u \in C^{1,1}(\mathbb{R}^d), \text{lip}(\nabla u) \leq 1 \right\}.$$

We recall the dual problem (1.7) (the second-order Beckmann's formulation):

$$\mathcal{I}'(f) := \inf \left\{ \int \varrho^0(\sigma) : \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}), \text{div}^2 \sigma = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\},$$

where  $\int \varrho^0(\sigma)$  is intended in the sense of convex one-homogeneous functionals on measures [13].

**Proposition 2.1.** *Assume that  $f$  given as above satisfy (2.2). Let  $\varrho$  be any norm on  $\mathcal{S}^{d \times d}$  and  $\varrho^0$  its polar defined by (1.5). Then, the supremum in (1.6) is reached. Furthermore, the infimum in (1.7) is a minimum, and we have the equality:*

$$\mathcal{I}(f) = \mathcal{I}'(f). \quad (2.3)$$

*Proof.* We begin by proving the existence of a maximizer for  $\mathcal{I}(f)$ . By the orthogonality conditions (2.2), we may restrict the supremum to functions  $u$  belonging to the subset

$$K_0 := \left\{ u \in C^{1,1}(\mathbb{R}^d) : \text{lip}(\nabla u) \leq 1, u(0) = 0, \nabla u(0) = 0 \right\}.$$

Let  $(u_n)$  be a maximizing sequence in  $K_0$ . Then,  $|u_n| \leq \frac{1}{2}|x|^2$ ,  $|\nabla u_n| \leq |x|$ . By applying Arzela-Ascoli compactness theorem, we can assume that  $(u_n, \nabla u_n) \rightarrow (u, \nabla u)$  uniformly on compact subsets, where  $u$  is a suitable element of  $K$ . To prove that  $u$  is optimal, we only need to check that  $\langle u_n, f \rangle \rightarrow \langle u, f \rangle$ , which, due to the particular form of  $f$ , reduces to showing that

$$\langle u_n - u, f_0 \rangle \rightarrow 0, \quad \langle \nabla(u_n - u), F \rangle \rightarrow 0.$$

Let  $v_n = u_n - u$ . Then  $(v_n, \nabla v_n) \rightarrow (0, 0)$  pointwisely, while  $|v_n| \leq |x|^2$  and  $|\nabla v_n| \leq 2|x|$ . Therefore,  $(v_n, \nabla v_n)$  is bounded in  $X_2(\mathbb{R}^d) \times X_1(\mathbb{R}^d)$ , and the convergence criterium (2.1) applies.

The existence of a minimal  $\sigma$  on the right hand side of (2.3) follows from the direct method. Indeed, the convex functional  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}) \mapsto \int \varrho^0(\sigma)$  is coercive (hence inf-compact for the weak-\* topology of  $\mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ ), while the distributional constraint  $\text{div}^2 \sigma = f$  is weakly-\* closed.

We prove now the equality (2.3) within two steps.

**Step 1:**  $\mathcal{I}(f) \leq \mathcal{I}'(f)$ . It is enough to prove the following inequality:

$$\langle f, u \rangle \leq \int \varrho^0(\sigma) \quad \text{for every } (u, \sigma) \in K_0 \times \mathcal{S}_f, \quad (2.4)$$

where  $\mathcal{S}_f := \left\{ \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}) : \text{div}^2 \sigma = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\}$ .

First we observe that we need only to show (2.4) for  $u \in K_0 \cap C^\infty$ . Indeed, we may approximate any  $u \in K_0$  by  $u_n = u * \rho_n$ , where  $\rho_n = n^d \rho(nx)$  is a sequence of mollifiers ( $\rho \in \mathcal{D}(\mathbb{R}^d; \mathbb{R}_+)$ , and  $\int \rho = 1$ ). Then,  $\text{lip}(\nabla u_n) \leq \text{lip}(\nabla u) \leq 1$ , and, therefore, the sequence  $(u_n, \nabla u_n)$  is bounded in  $X_2(\mathbb{R}^d) \times X_1(\mathbb{R}^d)$ . The convergence  $\langle u_n, f \rangle \rightarrow \langle u, f \rangle$  can be obtained by applying once more the criterium (2.1) to  $v_n = u_n - u$  and to  $\nabla v_n$ .

Let us now consider an element  $u \in K_0 \cap C^\infty$  and a generic  $\sigma \in \mathcal{S}_f$ . Then, recalling (1.5), we have

$$\int \langle \nabla^2 u, \sigma \rangle \leq \int \varrho(\nabla^2 u) \varrho^0(\sigma) \leq \int \varrho^0(\sigma) < +\infty.$$

Then, our claim (2.4) follows from Lemma B.1 (see Appendix B) which states that

$$\int \langle \nabla^2 u, \sigma \rangle = \langle u, f \rangle = \langle u, f_0 \rangle + \langle \nabla u, F \rangle \quad \forall \sigma \in \mathcal{S}_f. \quad (2.5)$$

This concludes Step 1.

**Step 2:**  $\mathcal{I}(f) \geq \mathcal{I}'(f)$ . We are going to show the equality:

$$\mathcal{I}_{\text{reg}}(f) := \sup \left\{ \langle u, f \rangle : u \in \mathcal{D}(\mathbb{R}^d), \text{lip}(\nabla u) \leq 1 \right\} = \mathcal{I}'(f).$$

Clearly  $\mathcal{I}_{\text{reg}}(f)$  is not larger than  $\mathcal{I}(f)$  and, thanks to Step 1, the equality above will imply that the three quantities coincide<sup>1</sup>. We introduce the value function  $h : C_0(\mathbb{R}^d; \mathcal{S}^{d \times d}) \rightarrow \mathbb{R} \cup \{+\infty\}$  defined by:

$$h(\zeta) := \inf \left\{ -\langle u, f \rangle : u \in \mathcal{D}(\mathbb{R}^d), \varrho(\nabla^2 u + \zeta) \leq 1 \right\} \quad (\text{with } h(\zeta) = +\infty \text{ if no admissible } u \text{ exists}).$$

Then,  $h$  is a convex proper functional such that  $h(\zeta) \leq 0$  whenever  $\sup(\varrho(\zeta)) \leq 1$  ( $u = 0$  is then an admissible competitor). It follows that  $h$  is continuous at 0 (with respect to the norm topology of  $C_0(\mathbb{R}^d; \mathcal{S}^{d \times d})$ ), where it takes the value  $h(0) = -\mathcal{I}_{\text{reg}}(f)$ . By a classical result of convex analysis (see Appendix A), it holds that

$$h(0) = h^{**}(0) = -\min h^*.$$

Then, the wished equality  $\mathcal{I}_{\text{reg}}(f) = \mathcal{I}'(f)$  follows if we can identify the polar of  $h$  as:

$$h^*(\sigma) = \int \varrho^0(\sigma) \quad \text{if } \text{div}^2 \sigma = f \text{ in } \mathcal{D}'(\mathbb{R}^d), \quad h^*(\sigma) = +\infty \text{ otherwise.} \quad (2.6)$$

Let  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ , and assume that  $h^*(\sigma) < +\infty$ . Then,

$$\begin{aligned} h^*(\sigma) &= \sup \left\{ \langle \zeta, \sigma \rangle - h(\zeta) : \zeta \in C_0(\mathbb{R}^d; \mathcal{S}^{d \times d}) \right\} \\ &= \sup \left\{ \langle \zeta, \sigma \rangle + \langle u, f \rangle : (u, \zeta) \in \mathcal{D}(\mathbb{R}^d) \times C_0(\mathbb{R}^d; \mathcal{S}^{d \times d}), \varrho(\nabla^2 u + \zeta) \leq 1 \right\} \\ &= \sup \left\{ \langle \chi, \sigma \rangle - \langle \nabla^2 u, \sigma \rangle + \langle u, f \rangle : (u, \chi) \in \mathcal{D}(\mathbb{R}^d) \times C_0(\mathbb{R}^d; \mathcal{S}^{d \times d}), \varrho(\chi) \leq 1 \right\} \\ &= \int \varrho^0(\sigma) + \sup \left\{ -\langle \nabla^2 u, \sigma \rangle + \langle u, f \rangle : u \in \mathcal{D}(\mathbb{R}^d) \right\}, \end{aligned}$$

where:

- in the third line, we put  $\chi = \nabla^2 u + \zeta$  which runs over the whole  $C_0(\mathbb{R}^d; \mathcal{S}^{d \times d})$ ;
- in the last line, we have taken for fixed  $u$  the supremum with respect to  $\chi$  recovering  $\int \varrho^0(\sigma)$  which agrees with the support function of the subset  $\{\varrho(\chi) \leq 1\}$ .

The finiteness of  $h^*(\sigma)$  requires that  $\langle u, f \rangle = \langle \nabla^2 u, \sigma \rangle$  for every  $u \in \mathcal{D}(\mathbb{R}^d)$ , meaning that  $\text{div}^2 \sigma = f$  in  $\mathcal{D}'(\Omega)$ . This proves the validity of (2.6) and concludes Step 2. The proof of Proposition 2.1 is finished.  $\square$

**2.2. Convex order and martingale transport.** Stochastic ordering plays an important role in probability theory as a tool for comparing random variables through their probability laws. Here, we are concerned specifically with the convex order between measures in  $\mathcal{M}_p(\mathbb{R}^d; \mathbb{R}_+)$  ( $p \in \{1, 2\}$ ). To any measure  $\mu \in \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}_+)$ , we associate its total mass  $\|\mu\|$  and its barycentre  $[\mu]$  given by:

$$\|\mu\| = \int \mu, \quad [\mu] = \frac{1}{\|\mu\|} \int x \mu(dx).$$

<sup>1</sup>unfortunately, we were unable to find a direct approximation of an admissible  $u$  by a sequence  $(u_n)$  of compactly supported functions such that  $\text{lip}(\nabla u_n) \leq 1$ .

**Definition 2.2.** Given two non-negative measures  $\mu, \nu$  in  $\mathcal{M}_1(\mathbb{R}^d)$ , we say that  $\nu$  dominates  $\mu$  in the sense of convex order, in short  $\nu \succeq_c \mu$ , if for every convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  there holds the inequality

$$\int \varphi d\nu \geq \int \varphi d\mu. \quad (2.7)$$

By the Moreau-Yosida infimal convolution procedure, we know that any convex lower semi-continuous function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup +\infty$  is the non-decreasing limit of a sequence of convex Lipschitz functions. Therefore, in order to show that  $\nu \succeq_c \mu$ , the inequality (2.7) needs to be checked only for those  $\varphi$  that are Lipschitz. If it is the case, then (2.7) extends to any convex lower semi-continuous  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ .

The following properties are straightforward:

- By testing (2.7) with affine functions (which are integrable), we see that

$$\nu \succeq_c \mu \quad \Rightarrow \quad \|\mu\| = \|\nu\|, \quad \text{and} \quad [\mu] = [\nu].$$

- $\mu \succeq_c \delta_{[\mu]}$  (Jensen inequality).

The next characterization of convex order is crucial.

**Theorem 2.3** (Strassen). *The convex order  $\nu \succeq_c \mu$  holds true if and only if there exists a  $\mu$ -measurable map  $x \mapsto p^x \in \mathcal{P}(\mathbb{R}^d)$  such that:*

- (i)  $[p^x] = x$   $\mu$ -a.e.,
- (ii)  $\nu(B) = \int p^x(B) \mu(dx)$  for any Borel set  $B \subset \mathbb{R}^d$ .

By  $MT(\mu, \nu)$  we denote the set of martingale transports from  $\mu$  to  $\nu$ , i.e. the family of coupling measures  $\gamma \in \Gamma(\mu, \nu)$  whose disintegration with respect to  $\mu$ , given by  $\langle \varphi, \gamma \rangle = \int (\int \varphi(x, y) \gamma^x(dy)) \mu(dx)$ , satisfies the condition  $[p^x] = x$   $\mu$ -a.e. By virtue of Strassen theorem, this family is non-empty if and only if  $\nu \succeq_c \mu$ .

Three straightforward consequences of Strassen theorem for measures  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  are listed below:

- (p1) Assume that  $\nu \succeq_c \mu$ . Then,  $\text{var}(\nu) \geq \text{var}(\mu)$ , while strict inequality holds unless  $\mu = \nu$ . Indeed, assuming that  $[\mu] = 0$ , for  $p^x$  such that  $\nu = \int p^x \mu(dx)$  with  $[p^x] = x$   $\mu$ -a.e., we have

$$\text{var}(\nu) - \text{var}(\mu) = \int (\langle |\cdot|^2, p^x \rangle - |x|^2) \mu(dx),$$

which is positive unless the Jensen inequality  $\langle |\cdot|^2, p^x \rangle \geq |x|^2$  is an equality for  $\mu$ -a.e.  $x$ . By the strict convexity of  $|\cdot|^2$ , this is possible only if  $p^x = \delta_x$ , hence, if  $\nu = \mu$ .

- (p2) Assume that  $[\nu] = 0$ , and take the convolution  $\rho = \mu \star \nu$ . Then, it holds that  $\rho \succeq_c \mu$ . Indeed,  $\rho = \int p^x \mu(dx)$  where  $p^x := (x + \text{id})\# \nu$  satisfies the condition  $[p^x] = x$ . Thanks to this property, one checks easily (see [19]) that, for centred Gaussian distributions  $\mu, \rho$  on  $\mathbb{R}^d$  with the respective covariance matrices  $R, M \in \mathcal{S}_+^{d \times d}$ , the condition  $\rho \succeq_c \mu$  reduces to the order relation  $R \geq M$  (in the sense of quadratic forms).

Finally, we point out that optimal transport problems under the martingale constraint of the kind  $\inf \{ \iint c(x, y) \gamma(dx dy) : \gamma \in MT(\mu, \nu) \}$  are often considered in the literature, most often for the cost  $c(x, y) = |x - y|^p$ ,  $p \geq 1$  (cf. for instance [1], [24], [14], [12]). A particularity of the quadratic cost  $p = 2$  is that, for  $\nu \succeq_c \mu$ , the infimum above is reached by any  $\gamma \in MT(\mu, \nu)$  since the total cost remains constant (equal to  $\text{var}(\nu) - \text{var}(\mu)$ ) on this subset. This fact will be exploited in the proof of Theorem 1.3.

## 3. PROOFS

A quite technical direct proof of Theorem 1.1 could be derived directly, by leveraging the Le Gruyer's three-point characterization (1.13) of the feasible set  $\{u \in C^{1,1} : \text{lip}(\nabla u) \leq 1\}$  (see [15] for more details on this characterization). However, as we aim to emphasize the important link between our initial problem and stochastic optimization under convex order dominance, we choose here to deal first with the proof of Theorem 1.3. After that, our main result in Theorem 1.1 and its Corollary 1.2 will follow nicely.

In the whole section we assume that  $\mu, \nu$  are centred probability measures in  $\mathcal{P}_2(\mathbb{R}^d)$ , that is, in particular,  $[\mu] = [\nu] = 0$ . These conditions are not restrictive since the equalities  $\int \mu = \int \nu$  and  $[\mu] = [\nu]$  are necessary to ensure that  $\mathcal{I}(\nu - \mu) < +\infty$ .

**3.1. Dualization of the minimal variance problem and optimality conditions.** Let us rewrite  $\mathcal{V}(\mu, \nu)$  defined in (1.20) in the form  $\mathcal{V}(\mu, \nu) = \inf \{\text{var}(\rho) : \rho \in \mathcal{A}(\mu, \nu)\}$  where

$$\mathcal{A}(\mu, \nu) := \left\{ \rho \in \mathcal{P}_2(\mathbb{R}^d) : \rho \succeq_c \mu, \rho \succeq_c \nu \right\}. \quad (3.1)$$

By the properties (p1), (p2) that conclude Section 2.2, we know that  $\rho = \mu \star \nu$  belongs to  $\mathcal{A}(\mu, \nu)$ , whence:

$$\max \{\text{var}(\mu), \text{var}(\nu)\} \leq \mathcal{V}(\mu, \nu) \leq \text{var}(\mu) + \text{var}(\nu). \quad (3.2)$$

Next, we consider a new variational problem involving pairs  $(\varphi, \psi)$  of convex functions. Let  $\mathcal{K}$  be the set of convex functions that are in  $C^{1,1}(\mathbb{R}^d)$  (clearly,  $\mathcal{K} \subset X_2(\mathbb{R}^d)$ ). Then, we set

$$\mathcal{V}'(\mu, \nu) := \sup \left\{ \int \varphi d\mu + \int \psi d\nu : (\varphi, \psi) \in \mathcal{F} \right\} \quad (3.3)$$

where  $\mathcal{F} := \{(\varphi, \psi) \in \mathcal{K}^2 : \varphi + \psi \leq |\cdot|^2\}$ .

**Proposition 3.1.** *There exists an optimal  $\rho$  for (1.20), and we have the no-gap equality*

$$\mathcal{V}(\mu, \nu) = \mathcal{V}'(\mu, \nu).$$

Furthermore,  $\rho \in \mathcal{A}(\mu, \nu)$  and  $(\varphi, \psi) \in \mathcal{F}$  are optimal for (1.20) and (3.3), respectively, if and only if the following optimality conditions are fulfilled:

$$\begin{cases} (i) & \varphi + \psi = |\cdot|^2 \quad \rho\text{-a.e.}, \\ (ii) & \int \varphi d\rho = \int \varphi d\mu, \quad \int \psi d\rho = \int \psi d\nu. \end{cases} \quad (3.4)$$

**Remark 3.2.** The existence issue for  $\mathcal{V}'(\mu, \nu)$  is not straightforward. In fact, optimal pairs  $(\varphi, \psi) \in \mathcal{F}$  will be deduced from the solutions to (1.6) by means of Proposition 3.5 in Section 3.2.

*Proof.* We start by proving that (1.20) admits solutions. By (3.2), there exists a maximization sequence  $(\rho_n)$  in  $\mathcal{A}(\mu, \nu)$  such that  $\text{var}(\rho_n) \rightarrow \mathcal{V}(\mu, \nu) < +\infty$ . Then,  $(\rho_n)$  is bounded in  $\mathcal{M}_2(\mathbb{R}^d)$  and, up to extracting a subsequence, we have  $\rho_n \xrightarrow{*} \rho$  in the duality between  $X_{2,0}(\mathbb{R}^d)$  and  $\mathcal{M}_2(\mathbb{R}^d)$ . Therefore, as  $\rho_n \in \mathcal{A}(\mu, \nu)$ , by passing to the limit  $n \rightarrow \infty$ , the convex order relations  $\int f d\rho \geq \sup\{\int f d\mu, \int f d\nu\}$  are deduced for every convex Lipschitz  $f$  (such  $f$  belongs to  $X_{2,0}(\mathbb{R}^d)$ ). As pointed out after Definition 2.2, this is enough to ensure that  $\rho \in \mathcal{A}(\mu, \nu)$ . The optimality of  $\rho$  follows since  $\mathcal{V}(\mu, \nu) = \liminf_n \text{var}(\rho_n) \geq \text{var}(\rho)$ .

Next, we prove the equality  $\mathcal{V}(\mu, \nu) = \mathcal{V}'(\mu, \nu)$ . Notice that the inequality  $\mathcal{V}(\mu, \nu) \geq \mathcal{V}'(\mu, \nu)$  is straightforward since for every admissible  $(\rho; \varphi, \psi)$  we have

$$\int \varphi d\mu + \int \psi d\nu \leq \int \varphi d\rho + \int \psi d\rho \leq \int |\cdot|^2 d\rho. \quad (3.5)$$

To show the opposite inequality, we introduce the perturbation function  $h : X_{2,0}(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  defined by

$$h(\chi) := \inf \left\{ - \left( \int \varphi d\mu + \int \psi d\nu \right) : (\varphi, \psi) \in \mathcal{K}^2, \varphi + \psi + \chi \leq |\cdot|^2 \right\}.$$

We see that  $h(0) = -\mathcal{V}'(\mu, \nu)$  is finite, while the function  $h$  is convex. Moreover, by taking  $\varphi = \psi = -\frac{1}{2}$  as a competitor, we have  $h(\chi) \leq -1$  whenever  $\chi \leq 1 + |\cdot|^2$ . Thus,  $h$  has a finite upper bound on the unit ball of the Banach space  $X_{2,0}(\mathbb{R}^d)$ . Therefore it is continuous at 0 and, by Appendix A, it holds that  $h(0) = h^{**}(0) = -\min h^*$ , where  $h^*$  denotes the Fenchel conjugate of  $h$  on the dual space  $\mathcal{M}_2(\mathbb{R}^d)$ . The asserted equality will follow if we can prove that

$$h^*(\rho) = \text{var}(\rho) \quad \text{if } \rho \in \mathcal{A}(\mu, \nu), \quad h^*(\rho) = +\infty \quad \text{otherwise.} \quad (3.6)$$

Let us compute

$$h^*(\rho) = \sup \left\{ \int \chi d\rho + \int \varphi d\mu + \int \psi d\nu : \chi \in X_{2,0}(\mathbb{R}^d), (\varphi, \psi) \in \mathcal{K}^2, \varphi + \psi + \chi \leq |\cdot|^2 \right\}.$$

Clearly, one has  $h^*(\rho) \leq \text{var}(\rho)$  if  $\rho \in \mathcal{A}(\mu, \nu)$ . To find a lower bound for  $h^*$ , we may restrict the supremum above to pairs  $(\varphi, \psi) \in \mathcal{K}^2$  which are Lipschitz. Fixing such a pair, we see that the function  $\bar{\chi}(z) := |z|^2 - \varphi(z) - \psi(z)$  belongs to  $X_2(\mathbb{R}^d)$  and it is positive for large  $|z|$ . By truncation, it can be approximated by a sequence  $\chi_n \in C_0(\mathbb{R}^d)$  such that  $\chi_n \rightarrow \bar{\chi}$  increasingly, and  $\sup_n \|\chi_n\|_{X_2(\mathbb{R}^d)} < +\infty$ . Since  $\chi_n + \varphi + \psi \leq |\cdot|^2$ , after certain manipulations we are led to<sup>2</sup>

$$h^*(\rho) \geq \int (\chi_n + \varphi + \psi) d\rho + \left( \int \varphi d\mu - \int \varphi d\rho \right) + \left( \int \psi d\nu - \int \psi d\rho \right).$$

Then, passing to the limit as  $n \rightarrow \infty$  (see (2.1)), we get the inequality

$$h^*(\rho) \geq \int |z|^2 d\rho + \left( \int \varphi d\mu - \int \varphi d\rho \right) + \left( \int \psi d\nu - \int \psi d\rho \right),$$

which holds true for every pair of convex Lipschitz functions  $(\varphi, \psi)$ . Therefore, the finiteness of  $h^*(\rho)$  implies that  $\rho$  dominates  $\mu$  and  $\nu$  in the convex order. In this case, we infer that  $\rho \in \mathcal{A}(\mu, \nu)$ , while  $h^*(\rho) \geq \text{var}(\rho)$ . This proves our claim (3.6), hence the equality  $\mathcal{V}(\mu, \nu) = \mathcal{V}'(\mu, \nu)$ .

We see now that a pair  $(\rho, (\varphi, \psi))$  in  $\mathcal{A}(\mu, \nu) \times \mathcal{F}$  is optimal if and only if the inequalities in (3.5) are equalities. In turn, these equalities are equivalent to the conditions (i), (ii) stated in Proposition 3.1.  $\square$

**3.2. Proving the equality  $\mathcal{V}'(\mu, \nu) = \mathcal{I}(\nu - \mu) - \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu))$ .** First, we recall a classical result that establishes a connection between the Hessian constraint and the convexity properties.

**Lemma 3.3.** *For any continuous function  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  the following conditions are equivalent:*

- (i)  $u \in C^{1,1}(\mathbb{R}^d)$  and  $\text{lip}(\nabla u) \leq 1$ ;
- (ii)  $u \in W_{\text{loc}}^{2,\infty}(\mathbb{R}^d)$  and  $-\text{Id} \leq \nabla^2 u \leq \text{Id}$  a.e. in  $\mathbb{R}^d$ ;
- (iii) both functions  $\frac{1}{2}|\cdot|^2 + u$  and  $\frac{1}{2}|\cdot|^2 - u$  are convex.

<sup>2</sup>all integrals involved below are finite since Lipschitz functions belong to  $X_2(\mathbb{R}^d)$

Next, we consider the subclass  $\mathcal{G} \subset C^0(\mathbb{R}^d)$  consisting of continuous functions  $\varphi$  such that  $|\cdot|^2 - \varphi$  admits an affine minorant. Note that  $\varphi \in \mathcal{G}$  if and only if  $\varphi(x) \leq |x - x_0|^2 + b$  for a suitable pair  $(x_0, b) \in \mathbb{R}^d \times \mathbb{R}_+$ . Then, we introduce the transform  $\mathcal{L} : \varphi \in \mathcal{G} \rightarrow \hat{\varphi} \in \mathcal{G}$  defined by:

$$\mathcal{L}\varphi = \hat{\varphi} \quad \text{where} \quad \hat{\varphi}(x) := |x|^2 - (|\cdot|^2 - \varphi)^{**}(x). \quad (3.7)$$

A crucial property of  $\mathcal{L}$  is that it preserves convexity (see the assertion (iii) below).

**Lemma 3.4.** *The transform  $\mathcal{L}$  enjoys the following properties:*

- (i)  $\mathcal{L}\varphi \geq \varphi$ , while  $\mathcal{L}\varphi \equiv \varphi$  if and only if  $|\cdot|^2 - \varphi$  is convex;
- (ii)  $\mathcal{L} \circ \mathcal{L} = \text{Id}$  (idempotence);
- (iii) If  $\varphi$  is convex, then  $\hat{\varphi} := \mathcal{L}\varphi$  is convex and  $C^{1,1}$ , whilst  $u := \frac{1}{2}|\cdot|^2 - \hat{\varphi}$  satisfies  $\text{lip}(\nabla u) \leq 1$ .

*Proof.* The first two properties are straightforward. In order to show that  $\hat{\varphi}$  is convex, we need only to check the Jensen's inequality  $\int \int \hat{\varphi}(z + \xi) p_0(d\xi) \geq \hat{\varphi}(z)$  for every centred finitely supported probability  $p_0$  and for any  $z \in \mathbb{R}^d$ . In view of the particular form of  $\hat{\varphi}$  given in (3.7), this amounts to showing that

$$\int (|\cdot|^2 - \varphi)^{**}(z + \xi) p_0(d\xi) \leq (|\cdot|^2 - \varphi)^{**}(z) + \text{var}(p_0). \quad (3.8)$$

To prove (3.8), we fix  $\varepsilon > 0$  and choose a finitely supported probability  $p^z$  such that  $[p^z] = z$  and

$$(|\cdot|^2 - \varphi)^{**}(z) \geq \int (|\zeta|^2 - \varphi(\zeta)) p^z(d\zeta) - \varepsilon.$$

Then, by applying Jensen inequality to  $(|\cdot|^2 - \varphi)^{**}$  (which is majorized by  $|\cdot|^2 - \varphi$ ), we infer that for every  $\xi \in \mathbb{R}^d$  we have

$$\begin{aligned} (|\cdot|^2 - \varphi)^{**}(z + \xi) - (|\cdot|^2 - \varphi)^{**}(z) &\leq \int \left( (|\zeta + \xi|^2 - \varphi(\zeta + \xi)) - (|\zeta|^2 - \varphi(\zeta)) \right) p^z(d\zeta) + \varepsilon \\ &= |\xi|^2 - 2\langle z, \xi \rangle - \int (\varphi(\zeta + \xi) - \varphi(\zeta)) p^z(d\zeta) + \varepsilon. \end{aligned}$$

By integrating with respect to the centred measure  $p_0(d\xi)$  and by Fubini theorem, we deduce that:

$$\begin{aligned} \int (|\cdot|^2 - \varphi)^{**}(z + \xi) p_0(d\xi) - (|\cdot|^2 - \varphi)^{**}(z) &\leq \text{var}(p_0) + \varepsilon - \iint (\varphi(\zeta + \xi) - \varphi(\zeta)) p^z(d\zeta) \otimes p_0(d\xi) \\ &= \text{var}(p_0) + \varepsilon - \int \left( \int (\varphi(\zeta + \xi) - \varphi(\zeta)) p_0(d\xi) \right) p^z(d\zeta) \\ &\leq \text{var}(p_0) + \varepsilon. \end{aligned}$$

Let us point out that, in order to reach the last line above, we used the convexity of  $\varphi$  which, by Jensen inequality, renders the integral with respect to  $p_0(d\xi)$  non-negative. Since  $\varepsilon$  can be chosen arbitrarily small, we get our claim (3.8), hence the convexity of  $\hat{\varphi}$ .

To complete the proof of the assertion (iii), we observe that the function  $u = \frac{1}{2}|\cdot|^2 - \hat{\varphi}$  is such that  $\frac{1}{2}|\cdot|^2 - u = \hat{\varphi}$  and  $\frac{1}{2}|\cdot|^2 + u = (|\cdot|^2 - \varphi)^{**}$  are convex functions. By virtue of Lemma 3.3, it follows that  $u$  (hence also  $\hat{\varphi}$ ) is  $C^{1,1}$ , and there also holds  $\text{lip}(\nabla u) \leq 1$ .  $\square$

**Proposition 3.5.** *Let  $\bar{u}$  be a solution to (1.6). Then, the pair of convex function  $(\bar{\varphi}, \bar{\psi})$  given by*

$$\bar{\varphi} = \frac{1}{2}|\cdot|^2 - \bar{u}, \quad \bar{\psi} = \frac{1}{2}|\cdot|^2 + \bar{u} \quad (3.9)$$

solves the maximization problem (3.3). Accordingly, we have the equality

$$\mathcal{I}(\nu - \mu) + \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) = \mathcal{V}'(\mu, \nu) .$$

*Proof.* Since  $\bar{u}$  is  $C^{1,1}$  with  $\text{lip}(\nabla \bar{u}) \leq 1$ , the pair of functions  $(\bar{\varphi}, \bar{\psi})$  given by (3.9) belongs to the class  $\mathcal{F}$  of admissible competitors for (3.3), thanks to the equivalence stated in Lemma 3.3. Therefore,

$$\mathcal{I}(\nu - \mu) = \int \bar{u} d\nu - \int \bar{u} d\mu = \int \varphi d\mu + \int \psi d\nu - \frac{\text{var}(\mu) + \text{var}(\nu)}{2} \leq \mathcal{V}'(\mu, \nu) - \frac{\text{var}(\mu) + \text{var}(\nu)}{2} .$$

Thus, we are done if we can prove the converse inequality, namely

$$\mathcal{V}'(\mu, \nu) \leq \mathcal{I}(\nu - \mu) + \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) . \quad (3.10)$$

Let  $(\varphi, \psi) \in \mathcal{F}$  be any admissible pair for (3.3). Since the convex continuous function  $\psi$  admits an affine minorant, the inequality  $\psi \leq |\cdot|^2 - \varphi$  implies that  $\psi = \psi^{**} \leq (|\cdot|^2 - \varphi)^{**}$ , while  $\varphi$  belongs to the subclass  $\mathcal{G}$  on which the  $\mathcal{L}$ -transform is well defined. By virtue of Lemma 3.4,  $\hat{\varphi} := \mathcal{L}\varphi$  is convex and satisfies  $\hat{\varphi} \geq \varphi$ . Therefore, it holds that

$$\int \varphi d\mu + \int \psi d\nu \leq \int \hat{\varphi} d\mu + \int (|\cdot|^2 - \varphi)^{**} d\nu = \int \hat{\varphi} d\mu + \int (|\cdot|^2 - \hat{\varphi}) d\nu .$$

In terms of  $u := \frac{1}{2}|\cdot|^2 - \hat{\varphi}$ , the latter inequality can be rewritten as follows:

$$\int \varphi d\mu + \int \psi d\nu \leq \int u d\nu - \int u d\mu + \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) .$$

By the assertion (iii) of Lemma 3.4,  $u$  is an admissible competitor for (1.6), hence  $\int u d\nu - \int u d\mu \leq \mathcal{I}(\nu - \mu)$ . This gives the following upper bound:

$$\int \varphi d\mu + \int \psi d\nu \leq \mathcal{I}(\nu - \mu) + \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) .$$

The desired inequality (3.10) is obtained by taking the supremum with respect to all pairs  $(\varphi, \psi) \in \mathcal{F}$ .  $\square$

**Remark 3.6.** Given an admissible pair  $(\varphi, \psi) \in \mathcal{F}$ , we can define two bivariate functions  $\tilde{\varphi}(x, z) := \varphi(x) + \langle \nabla \varphi(x), z - x \rangle$  and  $\tilde{\psi}(y, z) := \psi(y) + \langle \nabla \psi(y), z - y \rangle$ . By the convexity assumptions, we have  $\varphi(z) \geq \tilde{\varphi}(x, z)$  and  $\psi(z) \geq \tilde{\psi}(y, z)$ , hence the inequality

$$\varphi(x) + \langle \nabla \varphi(x), z - x \rangle + \psi(y) + \langle \nabla \psi(y), z - y \rangle \leq |z|^2 \quad \forall (x, y, z) \in (\mathbb{R}^d)^3 . \quad (3.11)$$

Take  $u \in C^{1,1}(\mathbb{R}^d)$  such that  $\text{lip}(\nabla u) \leq 1$ . By applying (3.11) to  $(\varphi, \psi) = (\frac{1}{2}|\cdot|^2 - u, \frac{1}{2}|\cdot|^2 + u)$  which belongs to  $\mathcal{F}$  (see Lemma 3.3), we recover the following three-point inequality

$$[u(y) + \langle \nabla u(y), z - y \rangle] - [u(x) + \langle \nabla u(x), z - x \rangle] \leq c(x, y, z) . \quad (3.12)$$

As pointed out in the introduction, this inequality characterizes the admissibility of  $u$  for (1.6).



**3.3. Relation with the three-marginal OT problem.** A key issue is the relation between the admissible subset  $\Sigma(\mu, \nu)$  for the optimal transport problem (1.9) and the admissible subset  $\mathcal{A}(\mu, \nu)$  for (1.20). This relation is illuminated by the following result:

**Lemma 3.7.** *Let  $\pi \in \mathcal{P}_2((\mathbb{R}^d)^3)$  be a 3-plan with marginals  $(\mu, \nu, \rho)$ . Define the marginals  $\pi_{1,3} := \Pi_{1,3}^\#(\pi)$  and  $\pi_{2,3} := \Pi_{2,3}^\#(\pi)$ , which are the push forwards of  $\pi(dx dy dz)$  through the projection maps  $(x, y, z) \rightarrow (x, z)$  and  $(x, y, z) \rightarrow (y, z)$ , respectively. Then,*

$$\pi \in \Sigma(\mu, \nu) \quad \Leftrightarrow \quad \begin{cases} \pi_{1,3} \in MT(\mu, \rho), \\ \pi_{2,3} \in MT(\nu, \rho). \end{cases} \quad (3.13)$$

Accordingly, we obtain the equality

$$\mathcal{A}(\mu, \nu) = \left\{ \rho \in \mathcal{P}_2(\mathbb{R}^d) : \exists \pi \in \Sigma(\mu, \nu), \quad \Pi_3^\#(\pi) = \rho \right\}. \quad (3.14)$$

*Proof.* Recalling the equilibrium conditions (1.10) which characterize the convex subset  $\Sigma(\mu, \nu)$ , checking the equivalence (3.13) amounts to verifying the two equivalences:

$$\begin{aligned} \text{(i)} \quad & \iiint \langle z - x, \Phi(x) \rangle \pi(dx dy dz) = 0 \quad \forall \Phi \in C_0(\mathbb{R}^d; \mathbb{R}^d) \quad \Leftrightarrow \quad \pi_{1,3} \in MT(\mu, \rho); \\ \text{(ii)} \quad & \iiint \langle z - y, \Psi(y) \rangle \pi(dx dy dz) = 0 \quad \forall \Psi \in C_0(\mathbb{R}^d; \mathbb{R}^d) \quad \Leftrightarrow \quad \pi_{2,3} \in MT(\nu, \rho). \end{aligned}$$

Let us prove (i); the proof of (ii) is similar and will be skipped. We consider the disintegration of the measure  $\pi_{1,3}$  with respect to its first marginal  $\mu$  which provides a  $\mu$ -measurable family  $\{p^x\}$  in  $\mathcal{P}(\mathbb{R}^d)$  such that  $\iint \theta(x, z) \pi_{1,3}(dx dz) = \int \left( \int \theta(x, z) p^x(dz) \right) \mu(dx)$  for every  $\theta \in C_0(\mathbb{R}^d \times \mathbb{R}^d)$ . Then, we observe that:

$$\begin{aligned} \iiint \langle z - x, \Phi(x) \rangle \pi(dx dy dz) &= \iint \langle z - x, \Phi(x) \rangle \pi_{1,3}(dx dz) \\ &= \int \left( \int \langle z - x, \Phi(x) \rangle p^x(dz) \right) \mu(dx) \\ &= \int \langle [p^x] - x, \Phi(x) \rangle \mu(dx). \end{aligned}$$

Clearly, these integrals vanish for every  $\Phi \in C_0(\mathbb{R}^d; \mathbb{R}^d)$  if and only if  $[p^x] = x$  holds  $\mu$ -a.e. This is exactly the martingale condition that characterizes  $\pi_{1,3} \in MT(\mu, \rho)$ .

Let us now prove the equality (3.14). By (3.13), the condition  $\pi \in \Sigma(\mu, \nu)$  implies that  $\rho \in \mathcal{A}(\mu, \nu)$ . Conversely, if  $\rho \succeq \mu$  and  $\rho \succeq \nu$ , Strassen theorem ensures the existence of martingale transports  $\gamma_{1,3} \in MT(\mu, \rho)$  and  $\gamma_{2,3} \in MT(\nu, \rho)$ . Then, we can recover an element  $\pi \in \Sigma(\mu, \nu)$  with  $\rho$  for the third marginal by using a gluing construction between  $\gamma_{1,3}$  and  $\gamma_{2,3}$ . A simple one (it is not unique) is as follows: let us consider the disintegrations of the measures  $\gamma_{i,3}$  ( $i \in \{1, 2\}$ ) with respect to their second marginal  $\rho$ . This gives  $\rho$ -measurable families  $\{p_i^z\}$  in  $\mathcal{P}(\mathbb{R}^d)$  such that

$$\gamma_{1,3}(dx dz) = \int (p_1^z(dx) \otimes \delta_\xi(dz)) \rho(d\xi), \quad \gamma_{2,3}(dy dz) = \int (p_2^z(dy) \otimes \delta_\xi(dz)) \rho(d\xi).$$

Then, it is easy to check that the measure  $\pi(dx dy dz) = \int (p_1^z(dx) \otimes p_2^z(dy)) \otimes \delta_\xi(dz) \rho(d\xi)$  has  $(\mu, \nu, \rho)$  for its marginals, and it satisfies  $\pi_{i,3} := \Pi_{i,3}^\#(\pi) = \gamma_{i,3}$ .  $\square$

**3.4. Proof of Theorem 1.3.** By Proposition 3.1 and Proposition 3.5, we already know that

$$\mathcal{I}(\nu - \mu) + \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) = \mathcal{V}'(\mu, \nu) = \mathcal{V}(\mu, \nu).$$

Accordingly, we still have to check that the infimum  $\mathcal{J}(\mu, \nu)$  in the three-marginal problem (1.9) satisfies the equality:

$$\mathcal{J}(\mu, \nu) = \mathcal{V}(\mu, \nu) - \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)). \quad (3.15)$$

Let  $\pi \in \Sigma(\mu, \nu)$  be a competitor for (1.9), and let  $\rho$  be its third marginal. Then, by Lemma 3.13, we know that  $\rho \in \mathcal{A}(\mu, \nu)$ , while  $\iiint \langle z - x, x \rangle \pi(dx dy dz) = \iiint \langle z - y, y \rangle \pi(dx dy dz) = 0$  by particularizing the equilibrium condition (1.10) for  $\Phi = \Psi = \text{id}$ . Thus, recalling the formula for the cost  $c(x, y, z) = \frac{1}{2}(|x - z|^2 + |y - z|^2)$ , we have:

$$\begin{aligned} \iiint c(x, y, z) \pi(dx dy dz) &= \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)) + \text{var}(\rho) - \iiint \langle z, x + y \rangle \pi(dx dy dz) \\ &= \text{var}(\rho) - \frac{1}{2}(\text{var}(\mu) + \text{var}(\nu)). \end{aligned}$$

The equality (3.15) then follows from (3.14) by noticing that taking the infimum with respect to  $\pi \in \Sigma(\mu, \nu)$  on the left hand side above amounts to taking the infimum with respect to  $\rho \in \mathcal{A}(\mu, \nu)$  in the last line. As a consequence, we see that  $\iiint c(x, y, z) \pi(dx dy dz) = \mathcal{J}(\mu, \nu)$  if and only if  $\text{var}(\rho) = \mathcal{V}(\mu, \nu)$ . That proves the assertion (i) of Theorem 1.3. The assertion (ii) is a direct consequence of Proposition 3.1 (existence of optimal  $\rho$ ) and of (3.14) (existence of  $\pi \in \Sigma(\mu, \nu)$  with the third marginal  $\rho$ ).  $\square$

**3.5. Proof of Theorem 1.1.** The existence of an optimal  $u$  solving (1.6) follows from Proposition 2.3 that we apply to the source term of the form  $f = \nu - \mu$ . To prove the central equality  $\mathcal{I}(f) = \mathcal{J}(\mu, \nu)$ , it is now enough to combine the equalities stated in Theorem 1.3, Proposition 3.1, and Proposition 3.5. The existence of an optimal  $\pi \in \Sigma(\mu, \nu)$  has been already established (see Theorem 1.3). Before proving the assertion (ii), we recall that if  $u$  is admissible for (1.6), then by integrating (3.12) with respect to any  $\pi \in \Sigma(\mu, \nu)$  and by taking into account the relations (1.10), we get

$$\begin{aligned} \iiint c(x, y, z) \pi(dx dy dz) &\geq \iiint \left( [u(y) + \langle \nabla u(y), z - y \rangle] - [u(x) + \langle \nabla u(x), z - x \rangle] \right) \pi(dx dy dz) \\ &= \int u d\nu - \int u d\mu. \end{aligned}$$

Therefore, since  $\mathcal{I}(\nu - \mu) = \mathcal{J}(\mu, \nu)$ , the optimality of  $(u, \pi)$  is equivalent to the fact that the above inequality is an equality. In view of (3.12), this happens if and only if (1.12) holds true.  $\square$

**3.6. Proof of Corollary 1.2.** Let  $\pi \in \Sigma(\mu, \nu)$  be an admissible 3-plan for (1.9), and let us consider the associated tensor valued measure, namely  $\sigma = \iiint \sigma^{x, y, z} \gamma(dx dy dz)$ . We claim that:

$$\int \varrho^0(\sigma) \leq \iiint c(x, y, z) \pi(dx dy dz) \quad \text{and} \quad \text{div}^2 \sigma = \nu - \mu \quad \text{in } \mathcal{D}'(\mathbb{R}^d). \quad (3.16)$$

Then, if  $\pi$  is optimal for (1.9), we will deduce that

$$\mathcal{I}'(\nu - \mu) (= \min(1.7)) \leq \int \varrho^0(\sigma) \leq \iiint c(x, y, z) \pi(dx dy dz) = \mathcal{J}(\mu, \nu),$$

hence the optimality of  $\sigma$  since we have  $\mathcal{I}'(\nu - \mu) = \mathcal{I}(\nu - \mu) = \mathcal{J}(\mu, \nu)$  by virtue of Proposition 2.3 and Theorem 1.1.

Let us now prove (3.16). By the subadditivity property of the convex one-homogenous functional  $\mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}) \ni \sigma \mapsto \int \varrho^0(\sigma)$ , we have:

$$\int \varrho^0(\sigma) \leq \iiint \left( \int \varrho^0(\sigma^{x,y,z}) \right) \pi(dx dy dz) \leq \iiint c(x, y, z) \pi(dx dy dz).$$

Indeed, recalling the definition of the rank-one measure  $\sigma^{x,y,z}$  given in (1.14), we have:

$$\int \varrho^0(\sigma^{x,y,z}) \leq \int_{[z,x]} |\xi - z| \mathcal{H}^1(d\xi) + \int_{[z,y]} |\xi - z| \mathcal{H}^1(d\xi) = \frac{1}{2}(|x - z|^2 + |y - z|^2),$$

with the inequality being an equality if the segments  $[x, z]$  and  $[y, z]$  do not overlap. Eventually, let us show that  $\sigma$  satisfies the distributional constraint  $\operatorname{div}^2 \sigma = \nu - \mu$ . Recalling that  $\operatorname{div}^2 \sigma^{x,y,z} = f^{x,y,z}$ , where  $f^{x,y,z} := \delta_y - \delta_x - \operatorname{div}((z - y)\delta_y - (z - x)\delta_x)$  (see (1.16)), for each test function  $\varphi \in \mathcal{D}(\mathbb{R}^d)$  we have

$$\begin{aligned} \langle \operatorname{div}^2 \sigma, \varphi \rangle &= \iiint \langle \operatorname{div}^2 \sigma^{x,y,z}, \varphi \rangle \pi(dx dy dz) = \iiint \langle f^{x,y,z}, \varphi \rangle \pi(dx dy dz) \\ &= \iiint \left( \varphi(y) - \varphi(x) + \langle \nabla \varphi(y), z - y \rangle - \langle \nabla \varphi(x), z - x \rangle \right) \pi(dx dy dz) \\ &= \int \varphi d\nu - \int \varphi d\mu, \end{aligned}$$

where the last equality relied on the relations (1.10). This proves our claim (3.16), hence the first assertion of Corollary 1.2. Let us now consider the marginal  $\gamma = \pi_{1,2}$  of an admissible  $\pi \in \Sigma(\mu, \nu)$  with respect to the first two coordinates. There is no loss of generality in assuming that  $\pi \in \mathcal{P}_2((\mathbb{R}^d)^3)$  (i.e.  $\iiint c d\pi < +\infty$ ). Then, there exists a  $\gamma$ -measurable family  $\{\pi^{x,y}\}$  in  $\mathcal{P}_2(\mathbb{R}^d)$  satisfying the disintegration formula  $\pi(dx dy dz) = \gamma(dx dy) \otimes \pi^{x,y}(dz)$ , see the convention (1.21). It yields

$$\iiint \alpha(x, y, z) \pi(dx dy dz) = \iint \langle \pi^{x,y}, \alpha(x, y, \cdot) \rangle \gamma(dx dy) \quad \forall \alpha \in X_2((\mathbb{R}^d)^3).$$

Let us apply this formula to the following element of  $X_2((\mathbb{R}^d)^3)$ :

$$\alpha_u(x, y, z) := [u(y) + \langle \nabla u(y), z - y \rangle] - [u(x) + \langle \nabla u(x), z - x \rangle] - c(x, y, z),$$

where  $u$  is admissible for (1.6). By (3.12), we have  $\alpha_u \leq 0$  while, by virtue of the second assertion of Theorem 1.1, we have  $\alpha_u = 0$  holding  $\pi$ -a.e. whenever the pair  $(u, \pi)$  is optimal. In this case, we get:

$$0 = \iiint \alpha_u(x, y, z) \pi(dx dy dz) = \iint \langle \pi^{x,y}, \alpha_u(x, y, \cdot) \rangle \gamma(dx dy),$$

yielding that  $\operatorname{sp}(\pi^{x,y}) \subset \{z : \alpha_u(x, y, z) = 0\}$  for  $\gamma$ -almost all  $(x, y) \in (\mathbb{R}^d)^2$ . Next, we show that the subset  $\{\alpha_u(x, y, \cdot) = 0\}$  reduces to the singleton  $\{z_u(x, y)\}$  where

$$z_u(x, y) = \frac{x + y}{2} + \frac{\nabla u(y) - \nabla u(x)}{2}. \quad (3.17)$$

For  $(x, y)$  being fixed, the function  $z \rightarrow \alpha_u(x, y, z)$  is strictly concave; hence, it reaches its maximum on  $\mathbb{R}^d$  at the unique point  $z_u(x, y)$  where  $\partial_z \alpha_u(x, y, z) = \nabla u(y) - \nabla u(x) - (2z - (x + y))$  vanishes. This furnishes (3.17). Since  $\nabla u$  is 1-Lipschitz,  $z_u(x, y)$  belongs to the ball  $B(\frac{x+y}{2}, \frac{|x-y|}{2})$ . Accordingly, any optimal transport plan  $\bar{\pi}$  is supported on  $(\mathcal{B}(\operatorname{sp} \mu, \operatorname{sp} \nu))^3$ , while the associated tensor measure  $\bar{\sigma}$  solving (1.7) satisfies (1.18). The proof of the assertion (ii) is now complete.  $\square$

## 4. EXAMPLES

In this section we give exact solutions for some classes of data  $\mu$  and  $\nu$ . In each case we propose a pair  $(u, \pi)$  and prove its optimality by checking the optimality condition (ii) in Theorem 1.1. It turns out that, after checking the three-point equality (1.12), the main concern is to check the admissibility conditions  $-\text{Id} \leq \nabla^2 u \leq \text{Id}$  and  $\pi \in \Sigma(\mu, \nu)$ . Once the optimality of  $(u, \pi)$  is proved, an optimal convex dominant  $\rho$  is computed as the third marginal of  $\pi$  (see Theorem 1.3). Meanwhile, according to the Corollary 1.2, a solution of the second-order Beckmann problem (1.7) of the form  $\sigma = \iiint \sigma^{x,y,z} \pi(dx dy dz)$  is derived.

**4.1. Ordered measures.** The simplest class of data is the one of  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  that are in convex order. Let us assume that

$$\mu \preceq_c \nu.$$

Then, for any any martingale transport plan  $\gamma \in MT(\mu, \nu)$ , an optimal pair  $(u, \pi)$  is given by:

$$u(x) = \frac{1}{2}|x|^2, \quad \pi(dx dy dz) = \gamma(dx dy) \otimes \delta_y(dz), \quad (4.1)$$

see the convention (1.21). Recall that  $MT(\mu, \nu)$  is non-empty by virtue of Strassen theorem.

Admissibility of  $u$  is clear, and  $\pi \in \Sigma(\mu, \nu)$  follows easily from Lemma 3.7. Due to the form of  $\pi$ , the three-point optimality condition (1.12) has to be checked merely for the triples  $(x, y, z)$  for which  $z = y$ . This is the case since  $u$  satisfies the identity<sup>3</sup>:

$$u(y) - [u(x) + \langle \nabla u(x), y - x \rangle] = \frac{1}{2}|x - y|^2.$$

With the validated optimality of the pair  $(u, \pi)$ , we can deduce the minimal energy:

$$\mathcal{I}(\nu - \mu) = \int u d(\nu - \mu) = \frac{1}{2}(\text{var}(\nu) - \text{var}(\mu)).$$

Moreover, the solution  $\sigma$  provided by Corollary 1.2 takes the form  $\iint \sigma^{x,y,y} \gamma(dx dy)$  where, by (1.14),  $\sigma^{x,y,y}$  is positive semi-definite, thus  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}_+^{d \times d})$ . Eventually, in view of the property (p2) (in Section 2.2), we see that  $\rho = \nu$  is the unique minimizer of the optimal convex dominance problem  $\mathcal{V}(\mu, \nu)$ . In contrast, the solution  $\sigma$  to (1.7) is not unique as it is shown in the forthcoming remark. Our argument will be based on following simple criterium:

**Proposition 4.1.** *Assume that  $\mu \preceq_c \nu$ . Then, a measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$  satisfying the constraint  $\text{div}^2 \sigma = \nu - \mu$  solves the second-order Beckmann problem (1.7) if and only if it is positive semi-definite.*

*Proof.* Using the integration by parts formula (B.1), for any  $\sigma$  satisfying  $\text{div}^2 \sigma = f = \nu - \mu$ , we have

$$\int \varrho^0(\sigma) \geq \int \langle \text{Id}, \sigma \rangle = \int \langle \nabla^2 u, \sigma \rangle = \int u df = \mathcal{I}(f).$$

By (2.3), the tensor measure  $\sigma$  is optimal for (1.7) if and only if  $\int \varrho^0(\sigma) = \mathcal{I}(f)$ . This means that the above inequality is an equality. Noticing that  $\varrho^0(A) = \text{Tr}(A)$  for  $A \in \mathcal{S}^{d \times d}$  implies that all the eigenvalues of  $A$  are non-negative, we infer that an admissible  $\sigma$  is optimal if and only if it is an element of  $\mathcal{M}(\mathbb{R}^d; \mathcal{S}_+^{d \times d})$ .  $\square$

<sup>3</sup>the left hand side is nothing else but the Bregman divergence of  $u$  at  $y$  around  $x$

**Remark 4.2** (*the non-uniqueness issue*). In general, even if  $\rho$  is unique, one can expect that  $\pi$  given in (4.1) is not unique since there may exist multiple martingale transports  $\gamma \in MT(\mu, \nu)$ . In turn, this translates to possibly multiple optimal tensor measures  $\sigma$ . In fact, we can exploit Proposition 4.1 to see that non-uniqueness of optimal  $\sigma$  goes beyond the one induced by the non-uniqueness of  $\pi$ .

Let us consider the simple example when  $\mu = \delta_0$  and  $\nu = \sum_{i=1}^4 \frac{1}{4} \delta_{y_i}$  where  $y_i$  are corners of the square centred at the origin. Clearly  $\mu \preceq_c \nu$ , and  $\gamma = \sum_{i=1}^4 \frac{1}{4} \delta_{(0, y_i)}$  is the unique element of  $MT(\mu, \nu)$ . It follows that  $\Sigma(\mu, \nu)$  is a singleton, which gives uniqueness of optimal  $\pi$ . The induced  $\sigma$  is the rank-one tensor measure defined as follows:

$$\sigma(d\xi) = \iiint \sigma^{x,y,z}(d\xi) \pi(dxdydz) = \sum_{i=1}^4 \frac{|\xi - y_i|}{4} \frac{y_i}{|y_i|} \otimes \frac{y_i}{|y_i|} \mathcal{H}^1(d\xi) \llcorner [0, y_i].$$

Such  $\sigma$  is demonstrated in Fig. 2(a). More accurately, the figure displays the density of  $\varrho^0(\sigma)$  with respect to  $\mathcal{H}^1$  measure restricted to the four segments.

Meanwhile, the set of  $\sigma \geq 0$  for which  $\text{div}^2 \sigma = \nu - \mu$  is very rich. Figs 2(b,c) give examples of such measures. After Proposition 4.1, they are also optimal for the second-order Beckmann problem (1.7). It is even possible to find optimal  $\sigma$  that has an absolutely continuous part. This example not only shows that we may experience great flexibility in the choice of optimal  $\sigma$  but also that not every such optimal measure can be decomposed with respect to a three-point measure  $\pi$  as in Corollary 1.2. This is a significant difference with respect to the classical first-order Beckmann problem where all minimizers can be decomposed along transport rays by virtue of Smirnov theorem (see [23] and Proposition 2.3 in [11]).

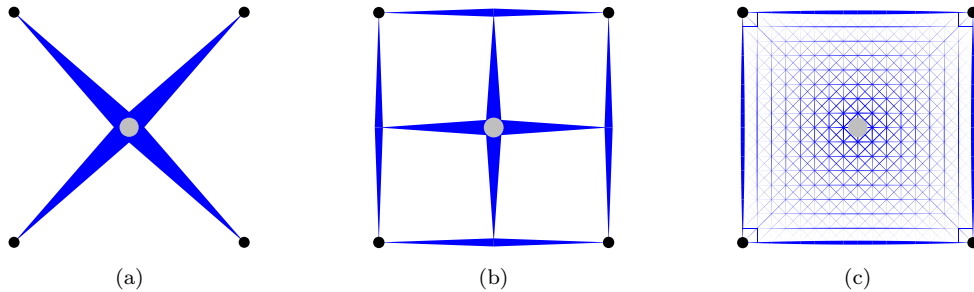


FIGURE 2. Various optimal  $\sigma$  (blue) for the data  $\mu = \delta_0$  (gray) and  $\nu = \sum_{i=1}^4 \frac{1}{4} \delta_{y_i}$  (black). Only the density of the 1D measure  $\sigma$  is displayed.

**4.2. Gaussian measures.** In this example we assume the data to be two centred Gaussian distributions on  $\mathbb{R}^d$ :

$$\mu = \mathcal{N}(0, M), \quad \nu = \mathcal{N}(0, N),$$

where  $M, N \in \mathcal{S}_+^{d \times d}$  are two positive semi-definite covariance matrices. Note that, if these matrices are ordered, we find ourselves in the framework of the former example (see the comment after (p2) in Section 2.2). In the general case, at the core of the solution lies the spectral decomposition of the difference of the covariance matrices:

$$N - M = \sum_{i=1}^d \lambda_i a_i \otimes a_i,$$

where  $a_i$  are mutually orthogonal vectors on the unit sphere  $S^{d-1}$ . Let us define the projection matrices

$$P_- := \sum_{\{i: \lambda_i < 0\}} a_i \otimes a_i, \quad P_+ := \sum_{\{i: \lambda_i \geq 0\}} a_i \otimes a_i = \text{Id} - P_-.$$

The following symmetric positive semi-definite matrices will prove to be essential:

$$\begin{aligned} M \vee N &:= M + (N - M)_+ = N + (M - N)_+, \\ M \wedge N &:= M - (M - N)_+ = N - (N - M)_+, \end{aligned}$$

where

$$(N - M)_+ = \sum_{i=1}^d (\lambda_i)_+ a_i \otimes a_i, \quad (M - N)_+ = \sum_{i=1}^d (\lambda_i)_- a_i \otimes a_i.$$

According to Remark 4.3,  $M \vee N$  can be seen as the least majorant of the matrices  $M, N$ , and  $M \wedge N$  as their greatest minorant.

We are going to now show that an optimal pair  $(u, \pi)$  is given by

$$u(x) = \frac{1}{2} \sum_{i=1}^d \text{sgn}(\lambda_i) \langle a_i, x \rangle^2, \quad \pi = \gamma(dx dy) \otimes \delta_{z_u(x, y)}(dz),$$

where we agree to the convention that  $\text{sgn}(0) = 1$ , while

- the transport plan  $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  is a normal distribution

$$\gamma = \mathcal{N}(0, G), \quad G = \begin{bmatrix} M & M \wedge N \\ M \wedge N & N \end{bmatrix}.$$

- the function  $z_u$  is computed according to (1.17), which here leads to

$$z_u(x, y) = P_- x + P_+ y.$$

The positive semi-definiteness of  $G$  is clear since  $M \wedge N$  is a minorant for the both  $M$  and  $N$ . Since  $\nabla^2 u = \sum_{i=1}^d \text{sgn}(\lambda_i) a_i \otimes a_i$ , feasibility of  $u$  is also straightforward. In view of the disintegrated form of  $\pi$ , it is sufficient to show that the equality (1.12) holds for every triple  $(x, y, z_u(x, y))$  where  $(x, y)$  ranges in whole  $(\mathbb{R}^d)^2$ . This reduces to a tedious but elementary computation.

The more involved part is showing the admissibility  $\pi \in \Sigma(\mu, \nu)$ . As the first and second marginals of  $\pi$  coincide with those of  $\gamma$ , they are equal to  $\mu$  and  $\nu$ , respectively. Thus, by virtue of Lemma 3.7, it is enough to show that the marginals  $\pi_{1,3} := \Pi_{1,3}^\#(\pi)$  and  $\pi_{2,3} := \Pi_{2,3}^\#(\pi)$  are martingale plans. Integrating against a test function  $\phi \in C_0(\mathbb{R}^d \times \mathbb{R}^d)$ , we obtain

$$\begin{aligned} \iint \phi(x, z) \pi_{1,3}(dx dz) &= \iiint \phi(x, z) \pi(dx dy dz) = \iint \phi(x, z_u(x, y)) \gamma(dx dy) \\ &= \iint \phi(x, x + P_+(y - x)) \gamma(dx dy) = \iint \phi(x, x + z) \hat{\gamma}(dx dz). \end{aligned}$$

Above  $\hat{\gamma}$  is the push forward of  $\gamma$  through the map  $A(x, y) = (x, z) = (x, P_+(y - x))$ . As  $A$  is linear, it might be identified with a  $2d \times 2d$  matrix. Accordingly,  $\hat{\gamma}$  is another Gaussian given by

$$\hat{\gamma} = \mathcal{N}(0, \hat{G}), \quad \hat{G} = A G A^\top = \begin{bmatrix} M & 0 \\ 0 & (N - M)_+ \end{bmatrix}.$$

Note that the matrix multiplication above is straightforward once we observe that

$$(M \wedge N)P_+ = MP_+, \quad (M \wedge N)P_- = NP_-.$$

The structure of the matrix  $\hat{G}$  shows that  $\hat{\gamma}$  is a product of two Gaussians:  $\hat{\gamma} = \mathcal{N}(0, M) \otimes \mathcal{N}(0, (N - M)_+)$ .

We continue the chain of equalities:

$$\begin{aligned} \iint \phi(x, z) \pi_{1,3}(dx dz) &= \int \left( \int \phi(x, x+z) \mathcal{N}(0, (N - M)_+)(dz) \right) \mathcal{N}(0, M)(dx) \\ &= \int \left( \int \phi(x, z) \mathcal{N}(x, (N - M)_+)(dz) \right) \mathcal{N}(0, M)(dx), \end{aligned}$$

in order to arrive at

$$\pi_{1,3}(dx dz) = \mu(dx) \otimes \mathcal{N}(x, (N - M)_+)(dz). \quad (4.2)$$

It is clear that  $\pi_{1,3}$  is a martingale. In a similar way one shows that  $\pi_{2,3} = \nu \otimes \mathcal{N}(y, (M - N)_+)$ , which is also a martingale. We have thus proved that  $\pi \in \Sigma(\mu, \nu)$  and, ultimately, that  $(u, \pi)$  are optimal. The minimal energy equals:

$$\mathcal{I}(\nu - \mu) = \int u d(\nu - \mu) = \frac{1}{2} \sum_{i=1}^d \operatorname{sgn}(\lambda_i) \langle N - M, a_i \otimes a_i \rangle = \frac{1}{2} \sum_{i=1}^d |\lambda_i| = \frac{1}{2} \varrho^0(N - M).$$

To identify the optimal measure  $\rho$  we compute the third marginal of  $\pi$ . Utilizing the disintegration formula (4.2) for  $\pi_{1,3}$  we find that it is a convolution of two Gaussians:

$$\rho = \pi_3 = \mu \star \mathcal{N}(0, (N - M)_+) = \mathcal{N}(0, M + (N - M)_+) = \mathcal{N}(0, M \vee N)$$

(note that we obtain the same result when computing the second marginal of  $\pi_{2,3}$ ).

**Remark 4.3.** It is possible to show directly that  $\bar{\rho} := \mathcal{N}(0, M \vee N)$  is a solution to the minimal variance problem (1.20). Indeed, since  $\bar{\rho}$  satisfies the dominance constraints (cf. (p2) in Section 2.2), we have  $\mathcal{V}(\mu, \nu) \leq \operatorname{Tr}(M \vee N)$ . In the opposite direction, any admissible  $\rho \in \mathcal{A}(\mu, \nu)$  admits a covariance matrix  $R \in \mathcal{S}_+^{d \times d}$  such that  $R \geq M$ ,  $R \geq N$ . Therefore, since  $\operatorname{var}(\rho) = \operatorname{Tr} R$ , we have:

$$\mathcal{V}(\mu, \nu) \geq \min_{R \in \mathcal{S}_+^{d \times d}} \left\{ \operatorname{Tr} R : R \geq M, R \geq N \right\}.$$

It is not difficult to check that the right hand side above is a semi-definite program which admits a unique solution given by  $R = M \vee N$ . The optimality of  $\bar{\rho}$  follows. Notice that, similarly, the matrix  $M \wedge N$  uniquely solves the analogous maximization problem where the order constraints are reversed. In this sense  $M \vee N$  is the least majorant of the matrices  $M, N$ , whilst  $M \wedge N$  is their greatest minorant.

**4.3. Two-point measures.** The simplest non-trivial data possible is when both measures are supported by two points:

$$\mu = \sum_{i=1}^2 \mu_i \delta_{x_i}, \quad \nu = \sum_{j=1}^2 \nu_j \delta_{y_j}.$$

As the barycentres must coincide, the problem is virtually planar. We can thus *a priori* assume that  $d = 2$ . In addition, we enforce that the four points are not aligned so that 1D scenario is avoided.

As before we assume that the measures are centred, i.e.  $[\mu] = [\nu] = 0$ . In this case  $x_1 = -\frac{\mu_2}{\mu_1} x_2$ ,  $y_1 = -\frac{\nu_2}{\nu_1} y_2$ . Note that the weights follow automatically from the positions:

$$\mu_i = \frac{|x_{i'}|}{|x_1| + |x_2|}, \quad \nu_j = \frac{|y_{j'}|}{|y_1| + |y_2|}, \quad (4.3)$$

where  $i' = 3 - i$ ,  $j' = 3 - j$ .

The main challenge lies in the fact that the type of the solution switches depending on the geometrical property of the convex quadrilateral that the points  $x_1, y_2, x_2, y_1$  form. Indeed, the two cases below must be considered:

- (A) the pairs of opposite edges of the quadrilateral are inclined at an angle non-greater than  $\pi/2$ ;
- (B) the angle between one of the pairs of opposite edges exceeds  $\pi/2$ .

The pairs of lines extending the edges in questions are drawn in Fig. 3(a). In fact, being in the scenario (A) is equivalent to the system of two inequalities:

$$\langle x_2 - y_2, y_1 - x_1 \rangle \geq 0, \quad (4.4a)$$

$$\langle x_1 - y_2, y_1 - x_2 \rangle \geq 0. \quad (4.4b)$$

It is worth emphasizing that at least one of those inequalities is always met.

#### Case (A)

To extent, this case is similar to the Gaussian example as again the spectral decomposition of the difference of the covariance matrices will play the central role. Defining  $M = \int x \otimes x \mu(dx)$  and  $N = \int y \otimes y \nu(dy)$  we can make use of (4.3) to show that

$$M = -x_1 \otimes x_2 = -x_2 \otimes x_1, \quad N = -y_1 \otimes y_2 = -y_2 \otimes y_1. \quad (4.5)$$

Since we assumed that the four points are not collinear, the difference always has two eigenvalues of opposite signs:

$$N - M = \lambda_a a \otimes a + \lambda_b b \otimes b, \quad \lambda_a < 0, \quad \lambda_b > 0,$$

where  $a \perp b$  and  $a, b \in S^1$ . In what follows we prove that in the case (A) the problems  $\mathcal{I}(\nu - \mu)$  and  $\mathcal{J}(\mu, \nu)$  are solved by, respectively,

$$u(x) = \frac{1}{2} (\langle b, x \rangle^2 - \langle a, x \rangle^2), \quad \pi = \sum_{i,j=1}^2 \gamma_{ij} \delta_{(x_i, y_j, z_{ij})}, \quad (4.6)$$

where

$$\gamma_{ij} = \mu_i \frac{\langle b, y_{j'} - x_i \rangle}{\langle b, y_{j'} - y_j \rangle}, \quad z_{ij} = \langle a, x_i \rangle a + \langle b, y_j \rangle b. \quad (4.7)$$

We observe that  $z_{ij} = z_u(x_i, y_j) = P_- x_i + P_+ y_j$  for  $P_- = a \otimes a$ ,  $P_+ = b \otimes b$ . Accordingly, both admissibility of  $u$  and the three-point optimality condition (1.12) can be shown identically as in Example 4.2. The biggest challenge consists in showing that  $\pi \in \Sigma(\mu, \nu)$ . In fact, it is the positivity of  $\gamma_{ij}$  that is the most delicate. The following result shows that it characterizes the case (A):

**Lemma 4.4.** *The inequalities (4.4) hold true if and only if  $\gamma_{ij} \geq 0$  for all  $i, j \in \{1, 2\}$ .*

As the proof is rather long and technical, it is moved to Appendix C. We can readily check that  $\pi \in \Sigma(\mu, \nu)$  relying on Lemma 3.7. The fact that the first marginal of  $\pi$  is  $\mu$  amounts to observing that  $\sum_{j=1}^2 \gamma_{ij} = \mu_i$ . Next, we compute

$$\pi_{1,3} = \sum_{i,j=1}^2 \gamma_{ij} \delta_{(x_i, z_{ij})} = \sum_{i=1}^2 \mu_i \delta_{x_i} \otimes p^i, \quad p^i = \sum_{j=1}^2 \frac{\langle b, y_{j'} - x_i \rangle}{\langle b, y_{j'} - y_j \rangle} \delta_{z_{ij}}.$$

Noting that  $z_{ij} = x_i + \langle b, y_j - x_i \rangle b$ , it is easy to show that  $[p^i] = x_i$ , rendering  $\pi_{1,3}$  a martingale.



To show that  $\pi_2 = \nu$  and that  $\pi_{2,3}$  is martingale as well, we derive an alternative formula for  $\gamma_{ij}$  that is symmetric to (4.7). First, observe that  $\mu_i = \frac{\langle a, x_{i'} \rangle}{\langle a, x_{i'} - x_i \rangle}$  thanks to (4.3). This starts the chain of equalities below in which we exploit the equality  $\langle a \otimes b, M \rangle = \langle a \otimes b, N \rangle$  and formulas (4.5):

$$\begin{aligned} \gamma_{ij} &= \frac{\langle a, x_{i'} \rangle}{\langle a, x_{i'} - x_i \rangle} \frac{\langle b, y_{j'} - x_i \rangle}{\langle b, y_{j'} - y_j \rangle} = \frac{\langle a \otimes b, -x_{i'} \otimes x_i \rangle + \langle a, x_{i'} \rangle \langle b, y_{j'} \rangle}{\langle a, x_{i'} - x_i \rangle \langle b, y_{j'} - y_j \rangle} \\ &= \frac{\langle a \otimes b, -y_j \otimes y_{j'} \rangle + \langle a, x_{i'} \rangle \langle b, y_{j'} \rangle}{\langle a, x_{i'} - x_i \rangle \langle b, y_{j'} - y_j \rangle} = \frac{\langle b, y_{j'} \rangle}{\langle b, y_{j'} - y_j \rangle} \frac{\langle a, x_{i'} - y_j \rangle}{\langle a, x_{i'} - x_i \rangle} = \nu_j \frac{\langle a, x_{i'} - y_j \rangle}{\langle a, x_{i'} - x_i \rangle}. \end{aligned}$$

Readily, arguments put forward above for the marginals  $\pi_1, \pi_{1,3}$  can be now reproduced for  $\pi_2, \pi_{2,3}$ . Admissibility  $\pi \in \Sigma(\mu, \nu)$  is thus established and, hence, also the optimality of the pair  $(u, \pi)$ .

It remains to give the solutions of  $\mathcal{V}(\mu, \nu)$  and of the second-order Beckmann problem (1.7):

$$\rho = \pi_3 = \sum_{i,j=1}^2 \gamma_{ij} \delta_{z_{ij}}, \quad \sigma = \iiint \sigma^{x,y,z} \pi(dx dy dz) = \sum_{i,j=1}^2 \gamma_{ij} \sigma^{x_i, y_j, z_{ij}}.$$

Case (B):

It would be impractical to give a unified solution for all possible positions of the points that fall within the scope of the case (B). Instead, we shall assume that  $\langle x_1, y_1 \rangle \geq 0$  and  $|x_1||y_2| \leq |x_2||y_1|$ . It is not restrictive as one can always relabel the points to guarantee it. Under those assumptions, one can easily observe that the inequality (4.4b) is automatically satisfied. Accordingly, the case (B) is characterized by the strict inequality

$$\langle x_2 - y_2, y_1 - x_1 \rangle < 0. \quad (4.8)$$

We start by defining the point  $z_0 \in \mathbb{R}^2$  as the intersection of the two straight lines that contain segments  $[x_1, y_1]$  and  $[x_2, y_2]$ , see Figs 3(e,f). Let us endow the plane  $\mathbb{R}^2$  with a polar coordinate system  $x \mapsto (\varrho(x), \vartheta(x)) \in [0, \infty) \times [0, 2\pi)$  where the pole and the orientation of the system are fixed by

$$\varrho(z_0) = 0, \quad \vartheta(x_1) = 0, \quad \vartheta(x_2) \in (0, \pi).$$

Next, we define two coefficients:

$$\alpha = \frac{\pi}{2\angle(x_2 - y_2, y_1 - x_1)}, \quad \beta = \frac{\alpha}{4\alpha - 1},$$

where  $\angle$  is the angle between two vectors that *a priori* ranges in  $[0, \pi]$ . Under the assumption (4.8) we have  $\alpha \in (\frac{1}{2}, 1)$  and  $\beta \in (\frac{1}{3}, \frac{1}{2})$ . In particular,  $\alpha \neq \beta$ . In polar coordinates the maximizer of  $\mathcal{I}(\nu - \mu)$  is

$$v(r, \theta) = \frac{1}{2} h(\theta) r^2,$$

where

$$h(\theta) = \begin{cases} h_1(\theta) = \cos(2\alpha\theta) & \text{if } \theta \in [2k\pi, 2k\pi + \pi/(2\alpha)) \text{ for } k \in \mathbb{Z}, \\ h_2(\theta) = \cos(2\beta(2\pi - \theta)) & \text{if } \theta \in [2k\pi + \pi/(2\alpha), 2(k+1)\pi) \text{ for } k \in \mathbb{Z}. \end{cases}$$

Finally, the following pair solves the problems  $\mathcal{I}(\nu - \mu)$  and  $\mathcal{J}(\mu, \nu)$ :

$$u(x) := v(\varrho(x), \vartheta(x)), \quad \pi = \nu_1 \delta_{(x_1, y_1, y_1)} + \nu_2 \delta_{(x_2, y_2, x_2)} + (\mu_1 - \nu_1) \delta_{(x_1, y_2, z_0)}. \quad (4.9)$$

This time, the main difficulty is to prove the admissibility of  $u$ . With the following lemma we see that it holds exactly in the case (B). The proof can be found in the Appendix C.

**Lemma 4.5.** *Assume that  $\angle(x_2 - y_2, y_1 - x_1) \neq 0$ . Then, the function  $u$  in (4.9) is an element of  $W_{\text{loc}}^{2,\infty}(\mathbb{R}^2)$ , whilst  $u \notin C^2(\mathbb{R}^2)$  unless  $\langle x_2 - y_2, y_1 - x_1 \rangle = 0$ . Moreover, the condition*

$$-\text{Id} \leq \nabla^2 u(x) \leq \text{Id} \quad \text{for a.e. } x \in \mathbb{R}^2$$

*holds true if and only if  $\langle x_2 - y_2, y_1 - x_1 \rangle \leq 0$ .*

We move on to check the admissibility  $\pi \in \Sigma(\mu, \nu)$ . First, observe that  $\mu_1 - \nu_1 = \nu_2 - \mu_2$  is non-negative thanks to the assumption  $|x_1||y_2| \leq |x_2||y_1|$ . Then, checking that the first and second marginals of  $\pi$  are equal to  $\mu$  and  $\nu$ , respectively, is straightforward. Prior to showing that  $\pi_{1,3}, \pi_{2,3}$  are martingales we make an observation. By equality of the barycentres there holds  $\int(x - z_0) \mu(dx) = \int(y - z_0) \nu(dy)$ . In this particular case it leads to  $\mu_1(x_1 - z_0) - \nu_1(y_1 - z_0) = \nu_2(y_2 - z_0) - \mu_2(x_2 - z_0)$ . Both triples  $(z_0, x_1, y_1)$  and  $(z_0, y_2, x_2)$  are collinear, and the respective lines are never parallel (cf. Fig. 3(f)), so the vectors on each side of the equality must be zero. In turn, it generates the two equalities:

$$x_1 = \frac{\nu_1}{\mu_1} y_1 + \frac{\mu_1 - \nu_1}{\mu_1} z_0, \quad y_2 = \frac{\mu_2}{\nu_2} x_2 + \frac{\nu_2 - \mu_2}{\nu_2} z_0.$$

By exploiting the first one, we check that  $\pi_{1,3}$  is indeed a martingale:

$$\pi_{1,3} = \nu_1 \delta_{(x_1, y_1)} + \mu_2 \delta_{(x_2, x_2)} + (\mu_1 - \nu_1) \delta_{(x_1, z_0)} = \mu_1 \delta_{x_1} \otimes \left( \frac{\nu_1}{\mu_1} \delta_{y_1} + \frac{\mu_1 - \nu_1}{\mu_1} \delta_{z_0} \right) + \mu_2 \delta_{x_2} \otimes \delta_{x_2}.$$

Handling  $\pi_{2,3}$  is similar. Ultimately,  $\pi \in \Sigma(\mu, \nu)$  is established.

It remains to check the three-point equality (1.12), and, in view of the form of  $\pi$ , it must be tested for the three triples  $(x, y, z)$ . The construction of  $u$  assures that:

$$u(\xi) = \frac{1}{2} |\xi - z_0|^2 \quad \forall \xi \in L_1 \quad \text{and} \quad u(\xi) = -\frac{1}{2} |\xi - z_0|^2 \quad \forall \xi \in L_2, \quad (4.10)$$

where  $L_1$  and  $L_2$  are the lines on which the triples  $(z_0, x_1, y_1)$  and  $(z_0, y_2, x_2)$  lie, respectively. As a result, one arrives at the following identities:

$$u(\xi) - [u(x) + \langle \nabla u(x), \xi - x \rangle] = \frac{1}{2} |\xi - x|^2 \quad \forall \xi, x \in L_1, \quad (4.11)$$

$$[u(y) + \langle \nabla u(y), \xi - y \rangle] - u(\xi) = \frac{1}{2} |\xi - y|^2 \quad \forall \xi, y \in L_2. \quad (4.12)$$

We are ready to verify the condition (1.12). For the triple  $(x_1, y_1, y_1)$  it reduces to (4.11) with  $x = x_1$ ,  $\xi = y_1$ , while for  $(x_2, y_2, x_2)$  to (4.12) with  $y = y_2$ ,  $\xi = x_2$ . Finally, condition (1.12) for the triple  $(x_1, y_2, z_0)$  can be recast by adding equalities (4.11) and (4.12), written for  $x = x_1$ ,  $\xi = z_0$  and, respectively,  $y = y_2$ ,  $\xi = z_0$ .

Optimality of the pair  $(u, \pi)$  is now established. Solutions to  $\mathcal{V}(\mu, \nu)$  and to the second-order Beckmann problem (1.7) read:

$$\rho = \pi_3 = \nu_1 \delta_{y_1} + \mu_2 \delta_{x_2} + (\mu_1 - \nu_1) \delta_{z_0}, \quad \sigma = \nu_1 \sigma^{x_1, y_1, y_1} + \mu_2 \sigma^{x_2, y_2, x_2} + (\mu_1 - \nu_1) \sigma^{x_1, y_2, z_0}.$$

For both cases, the solutions  $\rho$  and  $\sigma$  are displayed in Fig. 3. The blue colour matches the segments where  $\sigma$  is a positive semi-definite rank-one matrix, whilst the red colour matches the negative part. Figs 3(b,c,d) correspond to the case (A) where the two lines form an acute angle. Fig. 3(f) demonstrates case (B) when this angle is obtuse. Finally, Fig. 3(e) shows the limit case for the right angle. In this case, the mass at the point  $z_{21}$  vanishes, and the solution adheres to the formulas given both for the case (A) and (B).

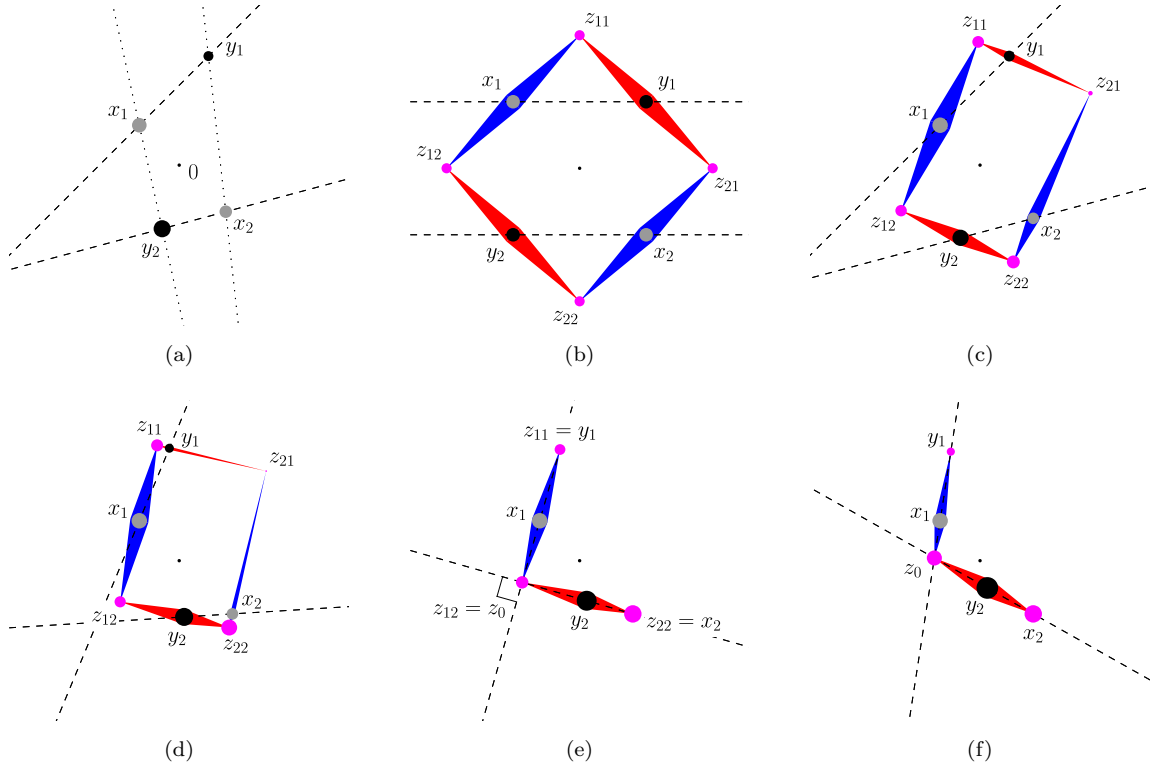


FIGURE 3. Data  $\mu$  (gray) and  $\nu$  (black), optimal  $\rho$  (magenta), and optimal  $\sigma$  (blue and red for the positive and negative part). (a) generic data in the case (A); (b,c,d) solutions for various data in the case (A); (e) solution for the limit case; (f) solution for data in the case (B).

**4.4. The basic first-order distribution data.** Unlike in the previous examples, here we shall consider a source which is not a measure but the first-order distribution  $f^{x,y,z}$  defined in the introduction. It is supported on two the points  $x, y \in \mathbb{R}^2$  and parametrized by a third point  $z$ :

$$f^{x,y,z} = \delta_y - \delta_x - \operatorname{div}((z-y)\delta_y - (z-x)\delta_x).$$

To focus attention we shall assume that the vectors  $x-z$  and  $y-z$  form an angle ranging in  $(0, \pi]$ . That is to say that  $x \neq y$ , while  $z$  cannot lie on the line crossing  $x, y$  except on the open segment  $]x, y[$ .

As announced in the introduction,  $f^{x,y,z} = \operatorname{div}^2 \sigma^{x,y,z}$ . Namely, it is the source term induced by the measure  $\sigma^{x,y,z}$  that serves as an elementary block for building solutions  $\sigma$  of the second-order Beckmann problem (1.7) for sources that are measures. Since  $\sigma^{x,y,z}$  is a competitor in the problem (1.7) for the source  $f = f^{x,y,z}$ , it is natural to ask if it is optimal for such a basic first-order distribution data. This short subsection is to settle this issue.

To that aim we exploit the construction of  $u$  put forth in Example 4.3, case (B). With the polar coordinate system satisfying  $\varrho(z) = 0$ ,  $\vartheta(x) = 0$ ,  $\vartheta(y) \in (0, \pi]$ , we repeat the construction of  $u$  with the parameter  $\alpha = \pi/(2\angle(x-z, y-z))$ . By the property that is analogous to (4.10), one obtains

$$\langle f^{x,y,z}, u \rangle = [u(y) + \langle \nabla u(y), z-y \rangle] - [u(x) + \langle \nabla u(x), z-x \rangle] = \frac{1}{2}|x-z|^2 + \frac{1}{2}|y-z|^2 = c(x, y, z).$$

On the other hand, from the proof of Corollary 1.2 we also know that  $\int \varrho^0(\sigma^{x,y,z}) = c(x, y, z)$ . Owing to the duality result in Proposition 2.1, optimality of the pair  $(u, \sigma^{x,y,z})$  will follow provided that  $u$  is admissible. In view of Lemma 4.5, it is the case only if  $x - z$  and  $y - z$  form an obtuse angle, which is to say that  $z$  lies in the disk of diameter  $[x, y]$ . We have arrived at the following result:

**Proposition 4.6.** *Assume that  $z \in B(\frac{x+y}{2}, \frac{|x-y|}{2})$ , then  $\sigma = \sigma^{x,y,z}$  solves the second-order Beckmann problem (1.7) for the first-order distribution data  $f = f^{x,y,z}$ . Accordingly we have the equality*

$$\mathcal{I}(f^{x,y,z}) = c(x, y, z).$$

**Remark 4.7.** The result above is valid for  $z = x$  (or for  $z = y$ ). In this case,  $\sigma^{x,y,z}$  is negative (or positive) semi-definite, while the optimal potential is given by  $u = -\frac{1}{2}|\cdot|^2$  (or  $u = \frac{1}{2}|\cdot|^2$ ).

On the other hand, we stress the fact that  $\sigma^{x,y,z}$  is no longer optimal if  $z$  is outside of the disc  $B(\frac{x+y}{2}, \frac{|x-y|}{2})$ . Indeed, in this case, we can show that  $\mathcal{I}(f^{x,y,z}) = |z - \frac{x+y}{2}||x - y|$  which is strictly less than  $c(x, y, z)$  for such  $z$ . An exception occurs when  $z$  lies on the extension of the segment  $[x, y]$ . This is due to the cancelling effect between the positive and negative parts of  $\sigma^{x,y,z}$ .

## 5. THE OPTIMAL GRILLAGE

We conclude with a section devoted to an application of the results developed in this paper to optimal design in mechanics. Classically, by a *grillage* one understands a planar multi-junction structure whose components are 1D straight bars. Although geometrically identical to *trusses* [8], a grillage – typically constituting a bearing structure of a ceiling – lies in a horizontally oriented plane and it is loaded vertically at its junctions. The load causes the bars to bend rather than stretch, ultimately resulting in different equilibrium configurations for the two types of structures.

The optimal design of trusses is famously known to be ill-posed, calling for relaxation in the form of the *Michell problem* [8, 16]. We will utilize Corollary 1.2 to prove that, in contrast, optimal grillages do exist provided that the load is a measure. Despite the vast literature on grillage optimization initiated in [22], it seems to be the first result of its kind. Before stating the result we will briefly recall the topic of truss optimization. We will finish with two open problems, including extension of the existence result to data that is a first-order distribution.

**5.1. Review on truss optimization and Michell problem.** A truss is a particular case of a 2D or 3D elastic solid that decomposes to one-dimensional straight bars. In general, the stress tensor in a solid can be described as a matrix valued measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ . It must satisfy the equilibrium equation  $-\operatorname{div} \sigma = F$  in  $(\mathcal{D}'(\mathbb{R}^d))^d$  for a system of forces  $F \in \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$ . For  $\sigma$  to exist, the load  $F$  has to be balanced in the following sense:  $\int \langle v_0, F \rangle = 0$  whenever  $v_0(x) = Ax + b$  for  $b \in \mathbb{R}^d$  and a skew-symmetric  $d \times d$  matrix  $A$ . By a truss we can understand the stress tensors that are of the form:

$$\sigma_\lambda = \iint \sigma^{x,y} \lambda(dxdy), \quad \lambda \in \mathcal{M}((\mathbb{R}^d)^2; \mathbb{R}), \quad (5.1)$$

where  $\sigma^{x,y} = \frac{y-x}{|y-x|} \otimes \frac{y-x}{|y-x|} \mathcal{H}^1 \llcorner [x, y]$  for  $x \neq y$ , and  $\sigma^{x,x} = 0$ . The positive and negative part  $\lambda_+(dxdy)$ ,  $\lambda_-(dxdy)$  represent, respectively, the tensile and compressive forces in the bars  $[x, y]$ .

Optimizing trusses amounts to looking for a measure  $\lambda$  that, under the condition of equilibrating  $F$ , minimizes the total energy, cf. [8]. Energy of a single bar  $(x, y)$  that is subject to a unit tensile/compressive force is the total variation  $\int |\sigma^{x,y}| = |y - x|$ . Accordingly, the optimal truss problem reads:

$$\inf \left\{ \iint |y - x| |\lambda|(dxdy) : \lambda \in \mathcal{M}((\mathbb{R}^d)^2; \mathbb{R}), \quad -\operatorname{div} \sigma_\lambda = F \text{ in } (\mathcal{D}'(\mathbb{R}^d))^d \right\}. \quad (5.2)$$

Note that the support of  $\lambda$  can exceed the set  $(\operatorname{sp} F)^2$ , which is to say that we can add junctions that are not loaded.

In (5.2) the total mass of  $\lambda$  is not controlled, raising the issue of existence. Moreover, in practice engineers expect that for a finitely supported load  $F$  there is a solution  $\bar{\lambda}$  that is finitely supported. It means that the structure can be manufactured as a junction of a finite number of bars. Meanwhile, already at the dawn of the 20th century, A.G.M. Michell observed that an optimal truss does not exist even for the simplest of loads. In his celebrated paper [17] he considered the *bridge problem* where the data is the three vertical forces in a plane  $\mathbb{R}^2$ :

$$F = \frac{e_2}{2} \delta_{e_1} + \frac{e_2}{2} \delta_{-e_1} - e_2 \delta_0, \quad (5.3)$$

where  $e_1 = (1, 0)$ ,  $e_2 = (0, 1)$ . Then, looking for finitely supported solutions of (5.2) leads to construction of minimizing sequences  $\lambda_h$  with the number of points in  $\operatorname{sp} \lambda_h$  going to infinity. When taking the weak-\* limit  $\bar{\sigma}$  of the sequence  $\sigma_{\lambda_h}$  one discovers that it is not representable through (5.1), see Fig. 4(a). The measure  $\bar{\sigma}$  is a solution of what today is known as the Michell problem:

$$\min \left\{ \int \varrho^0(\sigma) : \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}), \quad -\operatorname{div} \sigma = F \text{ in } (\mathcal{D}'(\mathbb{R}^d))^d \right\}. \quad (5.4)$$

Recall that  $\varrho^0$  is the Schatten norm:  $\varrho^0(S) = \sum_{i=1}^d |\lambda_i(S)|$ . In the modern measure theoretic setting the Michell problem was first formulated in [8]. Therein, it was proved that  $\inf (5.2) = \min (5.4)$ . Once a compactly supported  $F$  satisfies the balance condition, the minimum in the Michell problem is attained. From Fig. 4(a) one can discern that solutions may charge a curved curve (the thick lines in the figure). It rules out representing solutions through (5.1). To address this, the work [8] put forward another formulation where one seeks a signed measure on the space of regular curves, thus allowing for curved bars. To date, the existence issue remains open.

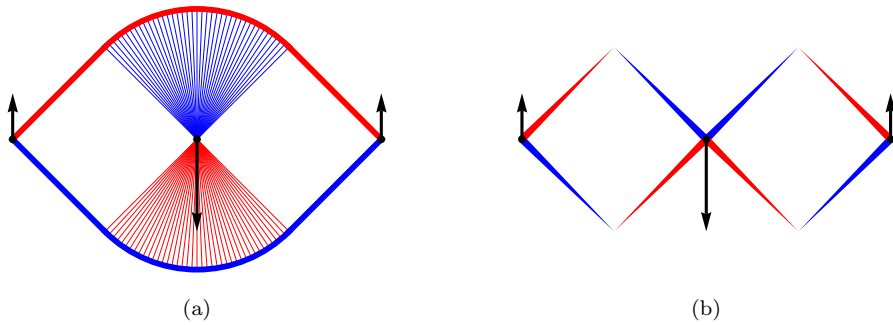


FIGURE 4. (a) Michell structure for a finitely supported system of forces  $F$ ; (b) optimal grillage for a finitely supported torque  $f = -\operatorname{div} F$ .

**5.2. Optimal grillage via the three-marginal optimal transport.** The other example of a structure that is built from 1D bars is a grillage, and it is a special case of a *plate*. Plates are by definition two dimensional bodies occupying a horizontal plane  $\mathbb{R}^d = \mathbb{R}^2$ . In the case of plates, the measure  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$  represents the *bending moment tensor*. The out-of-plane equilibrium of the plates is governed by the equation  $\operatorname{div}^2 \sigma = f$  in  $\mathcal{D}'(\mathbb{R}^d)$ , where  $f = f_0 - \operatorname{div} F$  is a first-order distribution. The measure  $f_0$  models out-of-plane forces. One can think of the positive part  $f_{0,+}$  as of the gravity pull, whilst  $f_{0,-}$  plays the role of upward reaction forces. The term  $F$  represents torques that act about in-plane axes. The balance condition for the load  $f$  reads as in (2.2).

With the second-order equilibrium equation, the decomposition of the measure  $\sigma$  to segments allows for adding affinely varying density. One of the ways of achieving this is through using  $\sigma^{x,y,z}$  as the basic measure. It concentrates on the union of segments  $[x, z] \cup [z, y]$ , see Fig. 1. Thus, by a grillage we will understand the bending moment tensor of the form

$$\sigma_\pi = \iiint \sigma^{x,y,z} \pi(dx dy dz), \quad \pi \in \mathcal{M}_+(\mathbb{R}^d)^3. \quad (5.5)$$

In the case of grillages,  $\pi(dx dy dz)$  enjoys the interpretation of the *transverse shear force* in the two-bar structure. Assuming that the segments  $[x, z]$  and  $[z, y]$  do not overlap, the energy of this structure is  $\int |\sigma^{x,y,z}| = \frac{1}{2}(|x - z|^2 + |y - z|^2) = c(x, y, z)$ . Accordingly, the optimal grillage problem can be formulated as follows:

$$\mathcal{I}_{\text{OG}}(f) = \inf \left\{ \iiint c(x, y, z) \pi(dx dy dz) : \pi \in \mathcal{M}_+(\mathbb{R}^d)^3, \operatorname{div}^2 \sigma_\pi = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\}.$$

Note that  $\pi$  is a positive Borel measure that is not necessarily finite. In fact, the condition  $\iiint c d\pi < \infty$  is sufficient for  $\sigma_\pi$  to be a well defined element of the space  $\mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$ .

*A priori*, the optimal grillage problem shares the issues of non-compactness that are known for truss optimization. A natural candidate for relaxation is the second-order Beckmann problem (1.7):

$$\mathcal{I}(f) = \min \left\{ \int \varrho^0(\sigma) : \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}), \operatorname{div}^2 \sigma = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\}.$$

The solution is guaranteed to exist provided that  $f_0 \in \mathcal{M}_2(\mathbb{R}^d)$ , and  $F \in \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d)$  (see Section 2.1). Utilizing the subadditivity of the functional  $\sigma \mapsto \int \varrho^0(\sigma)$  we can show that  $\int \varrho^0(\sigma_\pi) \leq \iiint c d\pi$ , which furnishes the inequality

$$\mathcal{I}_{\text{OG}}(f) \geq \mathcal{I}(f). \quad (5.6)$$

Historically, the systematic study of optimal grillages was initiated in the engineering paper [22]. Inspired by the theory of Michell structures, the author has tackled the Beckmann problem  $\mathcal{I}(f)$  from the outset. However, in the numerous analytical examples worked out in [22] and subsequent works, e.g. [21], one can discern the grillage-like structure (5.5) for the optimal solutions  $\bar{\sigma}$ . Unlike in the Michell problem, the curved bars are not exhibited at optimality.

The result below lays out a foundation for the foregoing observations. Exploiting the novel three-marginal optimal transport formulation developed in this paper, we show that the optimal grillage problem  $\mathcal{I}_{\text{OG}}(f)$  admits a solution when  $f$  is a measure, i.e. the first-order term  $-\operatorname{div} F$  is absent. On top of that, we prove that the optimal grillage consists of a finite number of bars once the load  $f$  is discrete.

**Theorem 5.1.** *If the load distribution  $f$  is a measure in  $\mathcal{M}_2(\mathbb{R}^d)$ , then the equality  $\mathcal{I}_{\text{OG}}(f) = \mathcal{I}(f)$  holds true together with the following statements.*

- (i) *There exists a solution  $\bar{\pi}$  of the optimal grillage problem  $\mathcal{I}_{\text{OG}}(f)$ , and, for any such solution,  $\sigma_{\bar{\pi}}$  solves the Beckmann problem  $\mathcal{I}(f)$ . Moreover,  $\bar{\pi}$  can be chosen such that  $\bar{\pi}((\mathbb{R}^d)^3) < \infty$ , and*

$$\text{sp } \sigma_{\bar{\pi}} \subset \mathcal{B}(\text{sp } f_+, \text{sp } f_-), \quad (5.7)$$

where  $f = f_+ - f_-$  is the Jordan decomposition to the positive and negative part.

- (ii) *If, in addition, the measure  $f$  is finitely supported, then one can choose a finitely supported solution  $\bar{\pi}$ . In particular,*

$$\sigma_{\bar{\pi}} \ll \mathcal{H}^1 \llcorner G$$

where  $G \subset \mathbb{R}^d$  is a union of at most  $2mn$  segments where  $m, n$  is the cardinality of  $\text{sp } f_+, \text{sp } f_-$ , respectively.

*Proof.* It is not restrictive to assume that  $f = \nu - \mu$ , for the probability distributions  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  that are centred,  $[\mu] = [\nu] = 0$ . Let  $\bar{\pi} \in \mathcal{P}((\mathbb{R}^d)^3)$  be a solution of the problem  $\mathcal{J}(\mu, \nu)$ , see (1.9). By the virtue of Corollary 1.2,  $\sigma_{\bar{\pi}}$  solves the second-order Beckmann problem  $\mathcal{I}(f)$ . In particular, it satisfies the equation  $\text{div}^2 \sigma_{\bar{\pi}} = f$ , so that  $\bar{\pi}$  is a competitor in  $\mathcal{I}_{\text{OG}}(f)$ . Thanks to assertion (i) of Theorem 1.3 and to the inequality (5.6), we obtain

$$\iiint c d\bar{\pi} = \mathcal{J}(\mu, \nu) = \mathcal{I}(f) \leq \mathcal{I}_{\text{OG}}(f) \leq \iiint c d\bar{\pi},$$

which proves optimality of  $\bar{\pi}$ . Finiteness of  $\bar{\pi}$  is trivial as it is a probability, while the inclusion (5.7) is the final assertion of Corollary 1.2. This concludes the proof of the part (i).

To prove the statement (ii), we assume that  $\mu = \sum_{i=1}^m \mu_i \delta_{x_i}$  and  $\nu = \sum_{j=1}^n \nu_j \delta_{y_j}$ . Let  $\bar{u}$  be any solution of the problem (1.6). Then, by assertion (ii) of Corollary 1.2, the 3-plan  $\bar{\pi}$  must be of the form

$$\bar{\pi} = \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \delta_{(x_i, y_j, z_{ij})} \quad \text{where} \quad z_{ij} = \frac{x_i + y_j}{2} + \frac{\nabla \bar{u}(y_j) - \nabla \bar{u}(x_i)}{2},$$

and thus  $\sigma_{\bar{\pi}} = \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \sigma^{x_i, y_j, z_{ij}}$ . The proof is complete since  $\text{sp } \sigma^{x_i, y_j, z_{ij}} \subset [x_i, z_{ij}] \cup [y_j, z_{ij}]$ .  $\square$

Optimal grillages have been already presented in Example 4.3, where  $\mu, \nu$  were two-point measures. The grillages  $\sigma_{\bar{\pi}}$  were showed in Fig. 3, and they consisted of eight or four bars with affinely varying cross section. Handling more complex data  $\mu, \nu$  calls for numerical treatment of the three-marginal optimal transport problem (1.9). For a discrete load  $\mu, \nu$ , it can be rewritten as a finite dimensional second-order conic program. It can be tackled using off-the-shelf convex optimization software.

**Example 5.2 (discrete load).** Here we present an optimal grillage found numerically for the discrete load  $f = \nu - \mu$  as in Fig. 5(a). The measure  $\mu$  is uniformly distributed on a grid of  $29 \times 29$  points, simulating the gravity pull coming from a square concrete slab. The five equal reaction forces in the columns are encoded by  $\nu$ . The numerical simulation of an optimal grillage  $\sigma_{\bar{\pi}}$  is showed in Fig. 5(c). Meanwhile, Fig. 5(b) presents the probability  $\bar{p}$  solving the problem  $\mathcal{V}(\mu, \nu)$ .

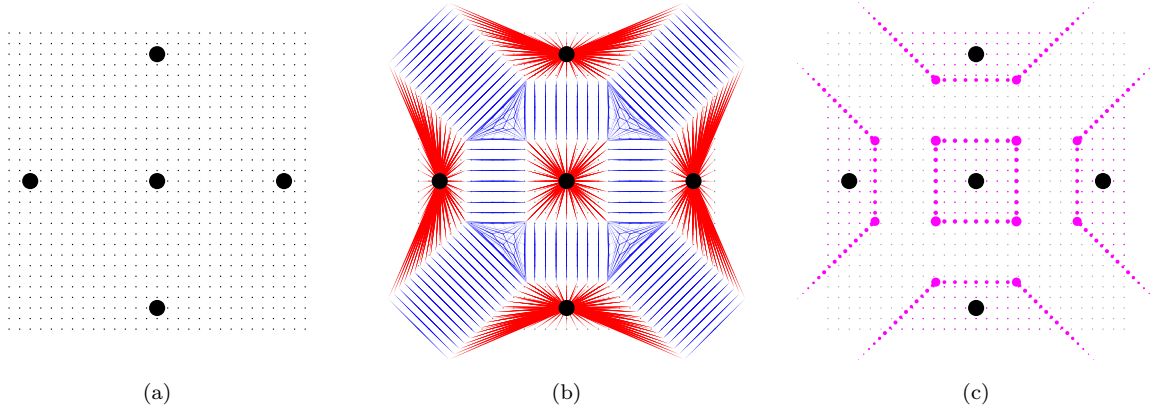


FIGURE 5. Numerical solution of the optimal grillage problem: (a) finitely supported data  $\mu, \nu$ ; (b) optimal grillage  $\sigma_{\bar{\pi}}$  where blue and red indicate, respectively, the positive and the negative part; (c) solution  $\bar{\rho}$  of the optimal dominance problem  $\mathcal{V}(\mu, \nu)$ .

**Example 5.3** (*continuous load*). In engineering practice, it is typical to assume that the weight of a slab is transferred to the grillage through a finite system of point loads, as demonstrated in the previous example. Nonetheless, it is natural to explore the optimal grillage problem also when the load is continuous:  $\mu = \mathcal{L}^2 \llcorner Q$ , where  $Q$  is the unit square, see Fig. 6(a). Numerically, it comes down to a fine discretization of  $\mu$ , here by a  $113 \times 113$  mesh. Fig. 6(c) shows the approximation  $\sigma_{\pi_h}$  of an optimal grillage  $\sigma_{\bar{\pi}}$ . It is clear that the support of  $\sigma_{\bar{\pi}}$  exceeds the square  $Q$ , but is contained within the set  $\mathcal{B}(\text{sp } \mu, \text{sp } \nu)$ . A prediction of the exact solution  $\bar{\rho}$  for the optimal dominance problem  $\mathcal{V}(\mu, \nu)$  is presented in Fig. 6(b). Based on the numerical simulation the authors expect that in the five quadrilateral regions  $\bar{\rho}$  is equal to  $\mu$ , i.e. to the Lebesgue measure. Partially on their boundaries, there is a part of  $\bar{\rho}$  that is absolutely continuous with respect to  $\mathcal{H}^1$ . Finally, at the vertices, there are concentrations in the form of Dirac delta masses.

### 5.3. Open problems.

5.3.1. *Loads that are general first-order distributions.* Generalization of Theorem 5.1 towards general first-order distributions  $f = f_0 - \text{div } F$  is not straightforward. Unlike  $f_0 = \nu - \mu$ , the vector measure  $F$  does not admit a natural decomposition to a pair of measures. It makes it difficult to propose a generalization of the set  $\Sigma(\mu, \nu)$ , and thus to find the right optimal transport formulation like (1.9) whose solution is guaranteed to exist.

The situation improves when the supports of the measures  $f_0, F$  are finite. It is then possible to prove that there exists a finitely supported solution of the optimal grillage problem  $\mathcal{I}_{\text{OG}}(f)$ . The main argument is using *the minimal extensions of jets* put forth in [2]. We skip the details here, and instead in Fig. 4(b) we show the optimal grillage for  $f_0 = 0$  and  $F$  as in the bridge problem, see (5.3). Note that the mechanical nature of  $F$  differs for trusses ( $F$  are forces) and grillages ( $F$  are torques).

If the measure  $F$  is not finitely supported, the issue of existence is more subtle. The authors found examples of  $F$  that charge a curved curve for which existence of solutions  $\bar{\pi}$  to the optimal grillage problem  $\mathcal{I}_{\text{OG}}(-\text{div } F)$  must imply that  $\bar{\pi}((\mathbb{R}^2)^3) = \infty$ . The infinite mass of  $\bar{\pi}$  makes it possible to construct infinite chains of straight bars whose lengths tend to zero, while their thickness is bounded



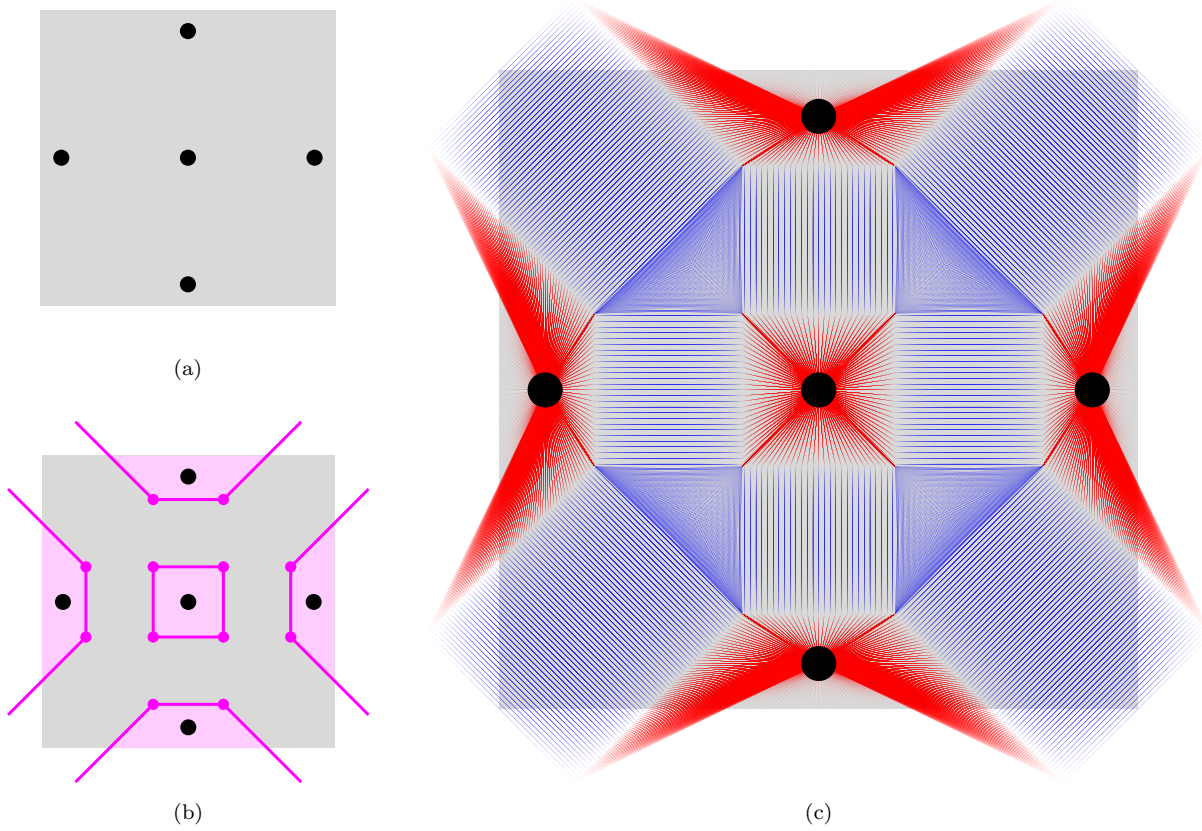


FIGURE 6. (a) Absolutely continuous loading  $\mu$  versus discrete reactions  $\nu$ ; (b) prediction of solution  $\bar{\rho}$  of the problem  $\mathcal{V}(\mu, \nu)$  consisting of 2D, 1D, and atomic parts; (c) numerical approximation of an optimal grillage  $\sigma_{\bar{\pi}}$  via a fine discretization of  $\mu$ .

from below by a positive constant. Such chains seem to open the door to forming solutions  $\bar{\pi}$  for such data  $F$ . Ultimately, the optimal grillage problem for data that are general first-order distributions is not well understood at the moment, and it remains to leave the reader with the following question:

**Problem 5.4.** *Assume that  $f = f_0 - \operatorname{div} F$  where  $(f_0, F) \in \mathcal{M}_2(\mathbb{R}^d) \times \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d)$ , and that the support of  $F$  is infinite. Does the optimal grillage problem  $\mathcal{I}_{\text{OG}}(f)$  admit a solution?*

5.3.2. *Domain confinement.* In practical applications, engineers often work within a prescribed design domain  $\Omega$ , a bounded open and connected subset of  $\mathbb{R}^d$ . For instance, a natural choice for  $\Omega$  in Example (5.3) is the square  $Q = \operatorname{sp} \mu$  being the outline of a ceiling. The domain confinement can be easily accounted for in the optimal grillage problem  $\mathcal{I}_{\text{OG}}(f)$  by adding the constraint  $\operatorname{sp} \sigma_{\pi} \subset \bar{\Omega}$ . Assuming that the load  $f$  is a measure, from assertion (i) of Corollary 5.1 we can see that the whole result holds true provided that

$$\mathcal{B}(\operatorname{sp} f_+, \operatorname{sp} f_-) \subset \bar{\Omega}. \quad (5.8)$$

In this case, the constraint  $\text{sp } \sigma_\pi \subset \bar{\Omega}$  is not binding. If the inclusion (5.8) is not satisfied, then one should work within the framework of the second-order Beckmann problem, whose modification now reads

$$\mathcal{I}(f, \Omega) := \min \left\{ \int \varrho^0(\sigma) : \sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d}), \text{ sp } \sigma \subset \bar{\Omega}, \text{ div}^2 \sigma = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \right\}.$$

Numerical experiments in 2D indicate that, with the condition (5.8) violated, there might not be solutions of  $\mathcal{I}(f, \Omega)$  which take the form  $\sigma_\pi$ . It appears that optimal  $\bar{\sigma}$  may charge subsets of the boundary  $\partial\Omega$  with the density being a full-rank matrix. In terms of mechanics, it corresponds to 1D bars (possibly curved) subject not only to bending moments but also to torsion. In the interior  $\Omega$ , however, the solution seems to decompose to straight bars  $\sigma^{x,y,z}$ . These observations lead to the following open problem:

**Problem 5.5.** *Assume a bounded domain  $\Omega \subset \mathbb{R}^d$  with Lipschitz regular boundary and a load  $f \in \mathcal{M}(\bar{\Omega})$ . Do there exist  $\sigma_{\partial\Omega} \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$  concentrated on  $\partial\Omega$  and  $\pi \in \mathcal{M}_+(\bar{\Omega}^3)$  such that*

$$\bar{\sigma} = \iiint \sigma^{x,y,z} \pi(dx dy dz) + \sigma_{\partial\Omega}$$

*solves the confined optimal grillage problem  $\mathcal{I}(f, \Omega)$ ?*

## APPENDIX A. CONVEX ANALYSIS

Let  $X$  be a normed space and let  $h : X \rightarrow \mathbb{R} \cup \{+\infty\}$  be a convex function. Recall that the Moreau-Fenchel conjugate of  $h$  is defined on the dual space  $X^*$  by:

$$h^*(x^*) := \sup_{x \in X} \{ \langle x, x^* \rangle - h(x) \} \quad \forall x^* \in X^*.$$

Clearly,  $h^*$  is convex and lower semi-continuous with respect to the weak-\* topology on  $X^*$ . Next, we define the biconjugate of  $h$  on  $X$  by:

$$h^{**}(x) := \sup_{x^* \in X^*} \{ \langle x, x^* \rangle - h^*(x^*) \} \quad \forall x \in X.$$

The following classical result (due to J.J. Moreau [18] in the infinite dimensional case) is used several times in this paper.

**Proposition A.1.** *Assume that there exists  $r > 0$  such that  $\sup\{h(x) : \|x\| \leq r\} < +\infty$ . Then:*

- (i)  *$h$  is continuous at 0, while  $h^*$  is coercive and attains its minimum on  $X^*$ ;*
- (ii) *we have the equalities:  $h(0) = h^{**}(0) = -\min h^*$ .*

## APPENDIX B. INTEGRATION BY PARTS

We give here the justification of the integration by parts formula on the whole  $\mathbb{R}^d$  that was required in Section 2 (see (2.5)) and in the proof of Proposition 4.1.

**Lemma B.1.** *Let  $f = f_0 - \text{div } F$  where  $(f_0, F)$  is any pair in  $\mathcal{M}_2(\mathbb{R}^d) \times \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d)$  satisfying (2.2). Let  $\sigma \in \mathcal{M}(\mathbb{R}^d; \mathcal{S}^{d \times d})$  satisfy  $\text{div}^2 \sigma = f$  in  $\mathcal{D}'(\mathbb{R}^d)$ . Then, for every  $u \in C^2(\mathbb{R}^d)$  with  $\text{lip}(\nabla u) < +\infty$ , we have:*

$$\int \langle \nabla^2 u, \sigma \rangle = \langle u, f \rangle = \langle u, f_0 \rangle + \langle \nabla u, F \rangle. \quad (\text{B.1})$$

*Proof.* By the orthogonality conditions (2.2), (B.1) is valid for affine functions. Therefore, it is not restrictive to assume that  $u$  and  $\nabla u$  vanish at 0; we may also assume that  $\text{lip}(\nabla u) \leq 1$ . On the other hand, the equality  $\text{div}^2 \sigma = f$  in the sense of distributions implies that (B.1) holds true if  $u \in \mathcal{D}(\mathbb{R}^d)$ . By using smooth convolution kernels, this can be extended to the case where  $u$  is compactly supported in  $\mathbb{R}^d$ . In order to remove this latter condition, we consider a sequence of radial cut-off functions  $\eta_k(x) := \eta\left(\frac{|x|}{k}\right)$  where

$$\eta \in \mathcal{D}(\mathbb{R}; [0, 1]), \quad \eta(t) = 1 \quad \text{if } |t| \leq k, \quad \eta(t) = 0 \quad \text{if } t \geq 2k.$$

Then we set  $u_k := u \eta_k$ . Since  $u_k$  satisfies (B.1), we have only to check that the sequence  $(v_k)$ , given by  $v_k := u - u_k = (1 - \eta_k)u$ , satisfies:

$$\langle v_k, f_0 \rangle \rightarrow 0, \quad \langle \nabla v_k, F \rangle \rightarrow 0, \quad \langle \nabla^2 v_k, \sigma \rangle \rightarrow 0. \quad (\text{B.2})$$

Since  $v_k(x)$  vanishes for  $|x| \leq k$ , while  $|v_k(x)| \leq |u(x)| \leq \frac{1}{2}|x|^2$  elsewhere, we infer that  $(v_k)$  is bounded in  $X_2(\mathbb{R}^d)$ . Hence,  $\langle v_k, f_0 \rangle \rightarrow 0$  by applying (2.1) with  $\mu = f_0 \in \mathcal{M}_2(\mathbb{R}^d)$ . In the same way,  $\nabla v_k$  is supported on the subset  $\{k \leq |x| \leq 2k\}$  where it satisfies the upper bound

$$|\nabla v_k| = \left| (1 - \eta_k) \nabla u - \frac{1}{k} u \eta' \left( \frac{|x|}{k} \right) \right| \leq |\nabla u| + \frac{\text{lip}(\eta)}{k} |u| \leq (1 + \text{lip}(\eta)) |x|.$$

In the last inequality we used the fact that  $|\nabla u|(x) \leq |x|$ , while  $|u(x)| \leq \frac{1}{2}|x|^2 \leq k|x|$  on  $\text{sp}(\nabla v_k)$ . It follows that  $(\nabla v_k)$  is bounded in  $X_1(\mathbb{R}^d; \mathbb{R}^d)$  and, recalling that  $F \in \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d)$ , we may apply (2.1) to infer that  $\langle \nabla v_k, F \rangle \rightarrow 0$ . Next, we compute:  $\nabla^2 v_k = (1 - \eta_k) \nabla^2 u - (\nabla \eta_k \otimes \nabla u + \nabla u \otimes \nabla \eta_k + u \nabla^2 \eta_k)$  where

$$\nabla \eta_k(x) = \frac{1}{k} \eta' \left( \frac{|x|}{k} \right) \frac{x}{|x|}, \quad \nabla^2 \eta_k(x) = \frac{1}{k^2} \eta'' \left( \frac{|x|}{k} \right) \left[ \frac{x \otimes x}{|x|^2} + \frac{1}{|x|} \eta' \left( \frac{|x|}{k} \right) \left( \text{Id} - \frac{x \otimes x}{|x|^2} \right) \right].$$

We notice that:

- for  $x \geq \text{lip}(\eta)$ , the symmetric tensor appearing inside the large bracket above is non-negative and not larger than the identity. Thus  $\varrho(\nabla^2 \eta_k) \leq \frac{\text{lip}(\eta')}{k^2}$ ;
- $\nabla \eta_k$  and  $\nabla^2 \eta_k$  are supported on  $\{k \leq |x| \leq 2k\}$ , where it holds that  $|u(x)| \leq \frac{1}{2}|x|^2 \leq 2k^2$  and  $|\nabla u(x)| \leq |x| \leq 2k$ ;
- $\varrho(\nabla \eta_k \otimes \nabla u + \nabla u \otimes \nabla \eta_k) \leq \frac{\text{lip}(\eta)}{k} (1 + |\nabla u|)$ .

All in all, we obtain a uniform upper bound for  $\varrho(\nabla^2 v_k)$  whose support is contained in  $\{k \leq |x| \leq 2k\}$

$$\begin{aligned} \varrho(\nabla^2 v_k) &\leq \varrho(\nabla^2 u) + \varrho(\nabla \eta_k \otimes \nabla u + \nabla u \otimes \nabla \eta_k) + |u| \varrho(\nabla^2 \eta_k) \\ &\leq 1 + \frac{\text{lip}(\eta)}{k} (1 + |\nabla u|) + |u| \frac{\text{lip}(\eta')}{k^2} \\ &\leq 1 + \text{lip}(\eta) \frac{1 + |x|}{k} + \frac{1}{2} \text{lip}(\eta') \left( \frac{|x|}{k} \right)^2 \\ &\leq C := 1 + 3 \text{lip}(\eta) + 2 \text{lip}(\eta'). \end{aligned}$$

By virtue of the inequality  $|\langle \nabla^2 v_k, \sigma \rangle| \leq \varrho(\nabla^2 v_k) \varrho^0(\sigma) \leq C \varrho^0(\sigma)$  holding in the sense of measures, it follows that:

$$|\langle \nabla^2 v_k, \sigma \rangle| \leq C \int_{k \leq |x| \leq 2k} \varrho^0(\sigma).$$

Since  $\int_{\mathbb{R}^d} \varrho^0(\sigma) < +\infty$ , we conclude that  $\langle \nabla^2 v_k, \sigma \rangle \rightarrow 0$  for  $k \rightarrow \infty$  as required in (B.2). This concludes the proof.  $\square$

## APPENDIX C. TWO-POINT MEASURES – ADDITIONAL PROOFS

*Proof of Lemma 4.4.* It is not restrictive to assume that  $\langle x_1, y_1 \rangle \geq 0$ . It is then easy to check that inequality (4.4b) is met automatically, thus we can focus on (4.4a) only. Next, we can enforce the orientation of the eigenvector  $b$  so that:

$$\langle b, y_1 \rangle > 0, \quad \langle b, y_2 \rangle < 0, \quad \langle b, x_1 \rangle \geq 0, \quad \langle b, x_2 \rangle \leq 0. \quad (\text{C.1})$$

Accordingly, one can check that  $\gamma_{ij} > 0$  when  $i = j$ , no matter if inequality (4.4a) holds or not. Therefore, we have to show that (4.4a) is equivalent to the two inequalities  $\gamma_{12} \geq 0$ ,  $\gamma_{21} \geq 0$ . What is more, this equivalence is trivial to show when  $\langle x_1, y_1 \rangle = 0$ . In the sequel we thus assume that  $\langle x_1, y_1 \rangle > 0$ .

For  $t > 0$  we define:

$$\tilde{x}_1(t) = t x_1, \quad \tilde{x}_2(t) = \frac{1}{t} x_2, \quad g(t) = t \langle \tilde{x}_2(t) - y_2, y_1 - \tilde{x}_1(t) \rangle.$$

The function  $g$  is quadratic on  $\mathbb{R}$ . Thanks to  $\langle x_1, y_1 \rangle > 0$ , one can show that  $g$  is concave, and it admits two positive roots:  $0 < t_1 < t_2$ . For each  $k \in \{1, 2\}$  we define two vectors:

$$v_k = \tilde{x}_2(t_k) - y_2, \quad w_k = y_1 - \tilde{x}_1(t_k). \quad (\text{C.2})$$

We shall show that, for both  $k$ ,  $(v_k, w_k)$  are mutually orthogonal eigenvectors of  $N - M$  (not necessarily normalized). Orthogonality follows from the fact that  $t_k$  are roots for  $g$ . The next observation is key:

$$-y_1 \otimes y_2 + \tilde{x}_1(t) \otimes \tilde{x}_2(t) = -y_1 \otimes y_2 + x_1 \otimes x_2 = N - M$$

for any  $t > 0$ . We exploit it to obtain:

$$\begin{aligned} (N - M) v_k &= -\langle y_1, v_k \rangle y_2 + \langle \tilde{x}_1(t_k), v_k \rangle \tilde{x}_2(t_k) = -\langle y_1 - w_k, v_k \rangle y_2 + \langle \tilde{x}_1(t_k), v_k \rangle \tilde{x}_2(t_k) \\ &= -\langle \tilde{x}_1(t_k), v_k \rangle y_2 + \langle \tilde{x}_1(t_k), v_k \rangle \tilde{x}_2(t_k) = \langle \tilde{x}_1(t_k), v_k \rangle (\tilde{x}_2(t_k) - y_2) = \langle \tilde{x}_1(t_k), v_k \rangle v_k. \end{aligned}$$

Similarly, one shows that  $(N - M) w_k = \langle -y_2, w_k \rangle w_k$ . The corresponding eigenvalues are  $\lambda_{v_k} = \langle \tilde{x}_1(t_k), v_k \rangle$  and  $\lambda_{w_k} = \langle -y_2, w_k \rangle$ . Next, we assess which of the four vectors (C.2) are parallel to  $b$ . To that aim we compare the signs; recall that  $\lambda_a < 0$ ,  $\lambda_b > 0$ . We compute the derivative of  $g$  at its roots:

$$g'(t_k) = \langle -y_2, y_1 - t_k x_1 \rangle + \langle x_2 - t_k y_2, -x_1 \rangle = \langle -y_2, w_k \rangle - \langle v_k, \tilde{x}_1(t_k) \rangle = \lambda_{w_k} - \lambda_{v_k}. \quad (\text{C.3})$$

Due to concavity of the quadratic function  $g$ , it must satisfy  $g'(t_1) > 0$  and  $g'(t_2) < 0$ . We conclude that  $\lambda_{v_1} < \lambda_{w_1}$  and  $\lambda_{v_2} > \lambda_{w_2}$ . As a result,  $v_2, w_1$  must be the eigenvectors that are parallel to  $b$ . One can easily check that the three vectors have also the same orientations. To sum up, we have:

$$b = \frac{v_2}{|v_2|} = \frac{w_1}{|w_1|} = \frac{1}{|v_2|} (\tilde{x}_2(t_2) - y_2) = \frac{1}{|w_1|} (y_1 - \tilde{x}_1(t_1)). \quad (\text{C.4})$$

We are ready to prove our assertion. Since  $\langle b, y_1 - y_2 \rangle > 0$  due to (C.1), we deduce that  $\text{sgn}(\gamma_{12}) = \text{sgn}(\langle b, y_1 - x_1 \rangle)$  and  $\text{sgn}(\gamma_{21}) = \text{sgn}(\langle b, x_2 - y_2 \rangle)$ . Defining the two functions:

$$f_1(t) := |v_2| \langle b, y_1 - t x_1 \rangle = \langle \tilde{x}_2(t_2) - y_2, y_1 - \tilde{x}_1(t) \rangle, \quad (\text{C.5})$$

$$f_2(t) := |w_1| \langle \frac{1}{t} x_2 - y_2, b \rangle = \langle \tilde{x}_2(t) - y_2, y_1 - \tilde{x}_1(t_1) \rangle \quad (\text{C.6})$$

we see that  $\text{sgn}(\gamma_{12}) = \text{sgn}(f_1(1))$ , and  $\text{sgn}(\gamma_{21}) = \text{sgn}(f_2(1))$ . Due to (C.1), the function  $f_1$  is strictly decreasing, and  $f_2$  is strictly increasing on  $(0, \infty)$ . The alternative formulas for  $f_1, f_2$  given above follow by (C.4). They provide the equalities  $f_1(t_2) = f_2(t_1) = 0$  since  $g(t_1) = g(t_2) = 0$ .

Let us now assume that the inequality (4.4a) is satisfied or, equivalently,  $g(1) \geq 0$ . By the properties of  $g$ , there holds  $t_1 \leq 1 \leq t_2$ . Since  $f_1$  is decreasing, we have  $f_1(1) \geq f_1(t_2) = 0$ . Similarly, because  $f_2$  is increasing,  $f_2(1) \geq f_2(t_1) = 0$ . This gives  $\gamma_{12} \geq 0$  and  $\gamma_{21} \geq 0$  due to.

Contrarily, assume that (4.4a) does not hold, which gives  $g(1) < 0$ . Then, either  $1 < t_1 < t_2$  or  $t_1 < t_2 < 1$ . In the first case, we have  $f_2(1) < f_2(t_1) = 0$ , which yields  $\gamma_{21} < 0$ . In the second case  $f_1(1) < f_1(t_2) = 0$ , and thus  $\gamma_{12} < 0$  by the same token. The proof is complete.  $\square$

*Proof of Lemma 4.5.* Let us observe that  $h : \mathbb{R} \rightarrow \mathbb{R}$  is  $C^1$  and  $2\pi$ -periodic on  $\mathbb{R}$ . We thus immediately infer that  $u$  is  $C^1$  on  $\mathbb{R}^2 \setminus \{z_0\}$ . However, thanks to the factor  $r^2$  in the definition of  $v$ , it can be showed that  $\nabla u$  is also continuous at  $z_0$  with  $\nabla u(z_0) = 0$ , that is  $u \in C^1(\mathbb{R}^2)$ . Function  $h$  is also piecewise  $C^2$ . More precisely,  $h''$  has discontinuity points  $2k\pi$  and  $2k\pi + \pi/(2\alpha)$  for integer  $k$  if and only if  $\alpha \neq \beta$ . As a result,  $u$  is not of class  $C^2$  except for the case  $\langle x_2 - y_2, y_1 - x_1 \rangle = 0$  which corresponds to the condition  $\alpha = \beta$  exactly. Nonetheless, the piecewise continuity of  $h''$  is enough to deduce that

$$u \in C^2(V_1 \cup V_2), \quad V_i = \{(\varrho, \vartheta)^{-1}(r, \theta) : r > 0, \theta \in A_i\}, \quad A_1 = ]0, \pi/(2\alpha[, \quad A_2 = ]\pi/(2\alpha), 2\pi[.$$

Moreover, on each open set  $V_i$  there holds  $\nabla^2 u(x) = (Q(x))^\top H_i(\varrho(x), \vartheta(x)) Q(x)$ , where  $Q(x)$  is a rotation matrix, and

$$H_i(r, \theta) = \begin{bmatrix} \frac{\partial^2 v}{\partial r^2} & -\frac{1}{r^2} \frac{\partial v}{\partial \theta} + \frac{1}{r} \frac{\partial^2 v}{\partial r \partial \theta} \\ -\frac{1}{r^2} \frac{\partial v}{\partial \theta} + \frac{1}{r} \frac{\partial^2 v}{\partial r \partial \theta} & \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + \frac{1}{r} \frac{\partial v}{\partial r} \end{bmatrix} = \begin{bmatrix} h_i(\theta) & \frac{1}{2} h_i'(\theta) \\ \frac{1}{2} h_i'(\theta) & \frac{1}{2} h_i''(\theta) + h_i(\theta) \end{bmatrix}.$$

Since  $\mathbb{R}^2 \setminus (V_1 \cup V_2)$  is Lebesgue negligible and  $h_i$  are cosine functions, we infer that  $u \in W_{\text{loc}}^{2,\infty}(\mathbb{R}^2)$ , which establishes the first part of the assertion.

To prove the second part, it is enough that we check that for  $i = 1, 2$  the eigenvalues of  $H_i(r, \theta) = H_i(\theta)$  remain in the regime  $[-1, 1]$  if and only if  $\langle x_2 - y_2, y_1 - x_1 \rangle \leq 0$ . Starting from  $i = 1$ , we obtain

$$H_1(\theta) = \begin{bmatrix} \cos(2\alpha\theta) & \alpha \sin(2\alpha\theta) \\ \alpha \sin(2\alpha\theta) & (1 - 2\alpha^2) \cos(2\alpha\theta) \end{bmatrix}$$

and, after using the Pythagorean trigonometric identity, formulas for the eigenvalues  $\lambda_-, \lambda_+$  follow:

$$\lambda_{\pm}(\theta) = (1 - \alpha^2) \cos(2\alpha\theta) \pm \alpha \sqrt{1 - (1 - \alpha^2) \cos^2(2\alpha\theta)}.$$

Assume first that  $\langle x_2 - y_2, y_1 - x_1 \rangle > 0$ , which gives  $\alpha > 1$ . Then, clearly  $\lambda_-(0) = 1 - 2\alpha^2 < -1$ . It remains to check the case when  $\langle x_2 - y_2, y_1 - x_1 \rangle \leq 0$ , for which  $\alpha, \beta \leq 1$ . Thanks to elementary computations we get the estimate:

$$\begin{aligned} \left( \pm 1 - (1 - \alpha^2) \cos(2\alpha\theta) \right)^2 &= (1 - \alpha^2) (1 \mp \cos(2\alpha\theta))^2 + \alpha^2 (1 - (1 - \alpha^2) \cos^2(2\alpha\theta)) \\ &\geq \left( \alpha \sqrt{1 - (1 - \alpha^2) \cos^2(2\alpha\theta)} \right)^2 \end{aligned}$$

where we acknowledged that  $\alpha \leq 1$ . Since the term  $(1 - \alpha^2) \cos(2\alpha\theta)$  ranges in  $[-1, 1]$ , from the estimate above we can deduce that indeed  $\lambda_{\pm}(\theta) \in [-1, 1]$ . Handling the matrix  $H_2(\theta)$  amounts to replacing  $\alpha$  with  $\beta$ . However, since  $\beta \leq 1$  as well, the same reasoning stands.  $\square$

## REFERENCES

- [1] J.-J. Alibert, G. Bouchitté, T. Champion: A new class of costs for optimal transport planning. *Eur. J. Appl. Math.* 30:1229–1263, 2019.
- [2] D. Azagra, E. Le Gruyer, C. Mudarra: Explicit formulas for  $C^{1,1}$  and  $C_{\text{conv}}^{1,\omega}$  extensions of 1-jets in Hilbert and superreflexive spaces. *J. Funct. Anal.* 274(10):3003–3032, 2018.
- [3] K. Bołbotowski, G. Bouchitté: Optimal design versus maximal Monge–Kantorovich metrics. *Arch. Ration. Mech. Anal.*, 243:1449–1524, 2022.
- [4] K. Bołbotowski, L. He, M. Gilbert: Design of optimum grillages using layout optimization. *Struct. Multidiscip. Optim.* 58:851–868, 2018.
- [5] G. Bouchitté, G. Buttazzo: Characterization of optimal shapes and masses through Monge-Kantorovich equation. *J. Eur. Math. Soc.* 3:139–168, 2001.
- [6] G. Bouchitté, G. Buttazzo, P. Seppecher: Shape optimization solutions via Monge-Kantorovich equation. *C. R. Acad. Sci. Paris Sér. I Math.* 324:1185–1191, 1997.
- [7] G. Bouchitté, I. Fragalà: Optimality conditions for mass design problems and applications to thin plates. *Arch. Ration. Mech. Anal.* 184:257–284, 2007.
- [8] G. Bouchitté, W. Gangbo, P. Seppecher: Michell trusses and lines of principal action. *Math. Models Methods Appl. Sci.* 18:1571–1603, 2008.
- [9] S. Conti, F. Maggi, S. Müller: Rigorous derivation of Föppl’s theory for clamped elastic membranes leads to relaxation. *SIAM J. Math. Anal.* 38(2):657–680, 2006.
- [10] D. Dentcheva, A. Ruszczyński: Optimization with stochastic dominance constraints. *SIAM J. Optim.* 14:548–566, 2003.
- [11] S. Dweik, F. Santambrogio:  $L^p$  bounds for boundary-to-boundary transport densities, and  $W^{1,p}$  bounds for the BV least gradient problem in 2D. *Calc. Var. Partial Differ. Equ.*, 58:Article. No. 31, 2019.
- [12] N. Ghoussoub, Y.-H. Kim, T. Lim: Structure of optimal martingale transport plans in general dimensions. *Ann. Probab.* 47:109–164, 2019.
- [13] C. Goffman, J. Serrin: Sublinear functions of measures and variational integrals. *Duke Math. J.* 31:159–178, 1964.
- [14] N. Gozlan, C. Roberto, P.-M. Samson, P. Tetali: Kantorovich duality for general transport costs and applications. *J. Funct. Anal.* 273:3327–3405, 2017.
- [15] E. Le Gruyer: Minimal Lipschitz extensions to differentiable functions defined on a Hilbert space. *Geom. Funct. Anal.* 19:1101–1118, 2009.
- [16] T. Lewiński, T. Sokół, C. Graczykowski: *Michell Structures*. Springer International Publishing, Cham, 2019.
- [17] A.G.M. Michell: The limits of economy of material in frame-structures. *Phil. Mag. S.* 8:589–597, 1904.
- [18] J.-J. Moreau: Fonctionnelles convexes. *Séminaire Jean Leray* 2:1–108, 1966.
- [19] A. Müller: Stochastic ordering of multivariate normal distributions. *Ann. Inst. Statist. Math.* 53:567–575, 2001.
- [20] A. Müller, M. Scarsini: Stochastic order relations and lattices of probability measures. *SIAM J. Optim.* 16:1024–1043, 2006.
- [21] W. Prager, G.I.N. Rozvany: Optimal layout of grillages. *J. Struct. Mech.* 5:1–18, 1977.
- [22] G.I.N. Rozvany: Grillages of maximum strength and maximum stiffness. *Int. J. Mech. Sci.* 14:651–666, 1972.
- [23] S.K. Smirnov: Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. *Algebra i Analiz* 5:206–238, 1993.
- [24] J. Wiesel, E. Zhang: An optimal transport-based characterization of convex order. *Depend. Model.*, 11:Article No. 20230102, 2023.
- [25] V.M. Zolotarev: Ideal metrics in the problem of approximating the distributions of sums of independent random variables. *Teor. Veroyatnost. i Primenen.* 22:449–465, 1977.

Karol Bołbotowski:

Department of Structural Mechanics and Computer Aided Engineering

Faculty of Civil Engineering, Warsaw University of Technology

16 Armii Ludowej Street, 00-637 Warsaw - POLAND

and

Lagrange Mathematics and Computing Research Center

103 rue de Grenelle, Paris 75007 - FRANCE

`karol.bolbotowski@pw.edu.pl`

Guy Bouchitté:

Laboratoire IMATH, Université de Toulon

BP 20132, 83957 La Garde Cedex - FRANCE

`bouchitte@univ-tln.fr`