

# THE DECOMPOSITION OF STRETCHED BROWNIAN MOTION INTO BASS MARTINGALES\*

Walter Schachermayer<sup>†</sup>

Bertram Tschiderer<sup>‡</sup>

ABSTRACT. In previous work J. Backhoff-Veraguas, M. Beiglböck and the present authors showed that the notions of stretched Brownian motion and Bass martingale between two probability measures on Euclidean space coincide if and only if these two measures satisfy an irreducibility condition.

Now we consider the general case, i.e. a pair of measures which are not necessarily irreducible. We show that there is a paving of Euclidean space into relatively open convex sets such that the stretched Brownian motion decomposes into a (possibly uncountable) family of Bass martingales on these sets. This paving coincides with the irreducible convex paving studied in previous work of H. De March and N. Touzi.

*MSC 2020 subject classifications:* Primary 60G42, 60G44; secondary 91G20.

*Keywords and phrases:* optimal transport, Brenier’s theorem, Benamou–Brenier, stretched Brownian motion, Bass martingale

## 1 INTRODUCTION

As in the parallel paper [BBST23] we consider two probability measures  $\mu$  and  $\nu$  on  $\mathbb{R}^d$  with finite second moments, denoted by  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ . It is always assumed that  $\nu$  dominates  $\mu$  in convex order, written  $\mu \preceq_c \nu$ , and meaning that  $\int f d\mu \leq \int f d\nu$  holds for all convex functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ .

Our object of interest are *martingale transports* starting from  $\mu$  and terminating in  $\nu$ , i.e., one-step martingales  $M = (M_0, M_1)$  with initial distribution  $\text{Law}(M_0) = \mu$  and terminal distribution  $\text{Law}(M_1) = \nu$ . A celebrated result by Strassen [Str65] shows that, under the above assumptions, there always exists a martingale transport from  $\mu$  to  $\nu$ . We may and shall identify the one-step martingale  $M = (M_0, M_1)$  with the joint law  $\pi \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$  of the random vector  $(M_0, M_1)$ . The set of all martingale transports  $\pi$  from  $\mu$  to  $\nu$  is denoted by  $\text{MT}(\mu, \nu)$ .

In [BBHK20] and [BBST23] a special martingale transport  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  has been identified: the *stretched Brownian motion* from  $\mu$  to  $\nu$ , which uniquely exists for all  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . It is defined as the optimizer  $\pi^{\text{SBM}}(dx, dy) = \pi_x^{\text{SBM}}(dy) \mu(dx)$  of the martingale optimal transport problem

$$P(\mu, \nu) = \sup_{\pi \in \text{MT}(\mu, \nu)} \int \text{MCov}(\pi_x, \gamma) \mu(dx), \tag{1.1}$$

---

\*This research was funded by the Austrian Science Fund (FWF) [Grant DOIs: 10.55776/P35197 and 10.55776/P35519]. For open access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

We thank Pietro Siorpaes for his valuable feedback and comments during the preparation of this paper.

<sup>†</sup>Faculty of Mathematics, University of Vienna (email: [walter.schachermayer@univie.ac.at](mailto:walter.schachermayer@univie.ac.at)).

<sup>‡</sup>Faculty of Mathematics, University of Vienna (email: [bertram.tschiderer@univie.ac.at](mailto:bertram.tschiderer@univie.ac.at)).

where  $\gamma$  is the standard Gaussian measure on  $\mathbb{R}^d$ . The maximal covariance  $\text{MCov}(\cdot, \cdot)$  between two probability measures  $p, q \in \mathcal{P}_2(\mathbb{R}^d)$  is defined as

$$\text{MCov}(p, q) = \sup_{\tilde{\pi} \in \text{Cpl}(p, q)} \int \langle y, z \rangle \tilde{\pi}(dy, dz),$$

where the set  $\text{Cpl}(p, q)$  consists of all couplings  $\tilde{\pi}$  between  $p$  and  $q$ , i.e., probability measures  $\tilde{\pi}$  on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $p$  and  $q$ . By [BBHK20, Theorem 2.2], the discrete-time formulation (1.1) is equivalent to the continuous-time martingale optimal transport problem

$$P(\mu, \nu) = \sup_{\substack{M_0 \sim \mu, M_1 \sim \nu, \\ M_t = M_0 + \int_0^t \sigma_s dB_s}} \mathbb{E} \left[ \int_0^1 \text{tr}(\sigma_t) dt \right], \quad (1.2)$$

and there exists a unique-in-law optimizer  $M^{\text{SBM}} = (M_t^{\text{SBM}})_{0 \leq t \leq 1}$  of (1.2). Furthermore, the optimizers  $M^{\text{SBM}}$  of (1.2) and  $\pi^{\text{SBM}}$  of (1.1) are related via  $\pi^{\text{SBM}} = \text{Law}(M_0^{\text{SBM}}, M_1^{\text{SBM}})$ , the values  $P(\mu, \nu)$  in (1.1) and (1.2) are equal, and  $M^{\text{SBM}}$  can be explicitly constructed from  $\pi^{\text{SBM}}$  (see Theorem 2.2 in [BBHK20]). Therefore the stretched Brownian motion from  $\mu$  to  $\nu$  can equivalently be viewed as the law  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  of the one step-martingale  $(M_0^{\text{SBM}}, M_1^{\text{SBM}})$ , or as the continuous-time martingale  $(M_t^{\text{SBM}})_{0 \leq t \leq 1}$  with stochastic differential

$$dM_t^{\text{SBM}} = \sigma_t^{\text{SBM}} dB_t, \quad 0 \leq t \leq 1,$$

where  $(B_t)_{0 \leq t \leq 1}$  is a standard Brownian motion on  $\mathbb{R}^d$ . In the following we shall freely switch between these two representations of a stretched Brownian motion.

A particularly appealing subclass of stretched Brownian motions consists of *Bass martingales*.

**Definition 1.1** ([BBST23, Definition 1.1]). Let  $(B_t)_{0 \leq t \leq 1}$  be a  $d$ -dimensional Brownian motion with initial distribution  $B_0 \sim \alpha$ , where  $\alpha$  is a probability measure on  $\mathbb{R}^d$ . Let  $v: \mathbb{R}^d \rightarrow \mathbb{R}$  be a convex function such that  $\nabla v(B_1)$  is square-integrable. The martingale

$$M_t := \mathbb{E}[\nabla v(B_1) | \sigma(B_s: s \leq t)] = \mathbb{E}[\nabla v(B_1) | B_t], \quad 0 \leq t \leq 1, \quad (1.3)$$

is called *Bass martingale* with *Bass measure*  $\alpha \in \mathcal{P}(\mathbb{R}^d)$ , initial law  $\mu := \text{Law}(M_0)$ , and terminal law  $\nu := \text{Law}(M_1)$ .

Martingales of this form were introduced by Bass [Bas83] (in dimension  $d = 1$  and with  $\alpha$  a Dirac measure) in order to derive a solution of the Skorokhod embedding problem. Definition 1.1 generalizes this concept to multiple dimensions and non-degenerate starting laws. A Bass martingale is a stretched Brownian motion [BBHK20, Theorem 1.10], but not vice versa.

## 1.1 THE IRREDUCIBLE CASE

In [BBST23] the following question was answered: under which conditions on  $\mu$  and  $\nu$  is the stretched Brownian motion  $\pi^{\text{SBM}}$  from  $\mu$  to  $\nu$  a Bass martingale? The answer is given by a connectivity assumption on the marginals, called *irreducibility*.

**Definition 1.2** ([BBST23, Definition 1.2]). For probability measures  $\mu, \nu$  on  $\mathbb{R}^d$  we say that the pair  $(\mu, \nu)$  is *irreducible*, if for all measurable sets  $A, B \subseteq \mathbb{R}^d$  with  $\mu(A) > 0$  and  $\nu(B) > 0$ , there exists a martingale transport  $\pi \in \text{MT}(\mu, \nu)$  such that  $\pi(A \times B) > 0$ .

The irreducibility assumption is necessary and sufficient for a stretched Brownian motion to be a Bass martingale, as shown in [BBST23, Theorem 1.3]. The analysis of stretched Brownian motion and Bass martingales relies on the following dual formulation of the optimization problem (1.1).

**Theorem 1.3** ([BBST23, Theorem 1.4]). *Assume that  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  are in convex order. The value  $P(\mu, \nu)$  of the problem (1.1) is equal to*

$$D(\mu, \nu) = \inf_{\substack{\psi \in L^1(\nu), \\ \psi \text{ convex}}} \left( \int \psi d\nu - \int (\psi^* * \gamma)^* d\mu \right). \quad (1.4)$$

*The infimum is attained by a lower semicontinuous convex function  $\psi: \mathbb{R}^d \rightarrow (-\infty, +\infty]$  with  $\mu(\text{ri}(\text{dom } \psi)) = 1$  if and only if  $(\mu, \nu)$  is irreducible. In this case the (unique) optimizer to (1.2) is given by the Bass martingale*

$$M_t := \mathbb{E}[\nabla v(B_1) | \sigma(B_s: s \leq t)] = \mathbb{E}[\nabla v(B_1) | B_t], \quad 0 \leq t \leq 1,$$

where  $v = \psi^*$  and  $B_0 \sim \nabla(\psi^* * \gamma)^*(\mu)$ .

Here, the symbol  $*$  used as a superscript denotes the convex conjugate of a function, otherwise it is the convolution operator;  $\text{ri}$  stands for relative interior, and  $\text{dom}$  for the domain of a function.

A technical caveat seems in order: it turns out that the optimizer  $\psi$  of (1.4) is not necessarily  $\nu$ -integrable. Therefore, in order for the difference of the integrals in (1.4) to be well-defined, attainment of  $D(\mu, \nu)$  has to be understood in a “relaxed” sense frequently encountered in martingale optimal transport problems; see [BJ16, BNT17, BNS22] and Subsection 1.6 below.

The following simple observation will be of crucial importance for the analysis of the dual problem (1.4): the value of the dual function

$$\psi \mapsto \mathcal{D}(\psi) := \int \psi d\nu - \int (\psi^* * \gamma)^* d\mu \quad (1.5)$$

remains unchanged if the convex function  $\psi$  is replaced by  $\psi + \text{aff}$ , where  $\text{aff}: \mathbb{R}^d \rightarrow \mathbb{R}$  is an arbitrary affine function. As a consequence, dual optimizers  $\psi$  — provided they exist — turn out to be unique, modulo adding affine functions. We thus make the following definition.

**Definition 1.4.** Let  $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$  be a convex function. We write  $[\psi]$  for the equivalence class of nonnegative convex functions of the form  $\psi + \text{aff}$ , where  $\text{aff}: \mathbb{R}^d \rightarrow \mathbb{R}$  is an affine function.

We make one more observation which will turn out to be useful: for an optimizing sequence  $(\psi_n)_{n \geq 1}$  of lower semicontinuous convex functions for the dual problem (1.4) only their values on the closed convex hull  $\widehat{\text{supp}}(\nu)$  of the support of  $\nu$  matter. Indeed, defining

$$\check{\psi}_n := \begin{cases} \psi_n, & \text{on } \widehat{\text{supp}}(\nu), \\ +\infty, & \text{on } \mathbb{R}^d \setminus \widehat{\text{supp}}(\nu), \end{cases}$$

we obtain lower semicontinuous convex functions. It follows from the definition (1.5) of the dual function that  $\mathcal{D}(\check{\psi}_n) \leq \mathcal{D}(\psi_n)$ , so that  $(\check{\psi}_n)_{n \geq 1}$  is an optimizing sequence, too.

The following theorem is an extension of [BBST23, Theorem 1.4]. It characterizes the features of the irreducible case.

**Theorem 1.5 (IRREDUCIBLE CASE).** *Let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . We denote by  $C := \widehat{\text{supp}}(\nu)$  the closed convex hull of the support of  $\nu$  and by  $I := \text{ri } C$  its relative interior. Let  $(\psi_n)_{n \geq 1}$  be an optimizing sequence of convex functions for the dual problem (1.4), satisfying the regularity assumption (1.22).*

(1) *If the pair  $(\mu, \nu)$  is irreducible, then the following assertions hold:*

- (i)  *$\mu(I) = 1$  and  $\widehat{\text{supp}}(\nu) = C^{\text{B}}(x)$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , where  $C^{\text{B}}(x) := \overline{I^{\text{B}}(x)}$  and the set  $I^{\text{B}}(x)$  is defined in (1.10) below.*

(ii) There exist representatives  $\bar{\psi}_n \in [\psi_n]$  and there is a lower semicontinuous convex function  $\psi_{\text{lim}}: \mathbb{R}^d \rightarrow [0, +\infty]$ , which is unique modulo adding affine functions, such that

$$\lim_{n \rightarrow \infty} \bar{\psi}_n(y) = \psi_{\text{lim}}(y) < +\infty, \quad y \in I, \quad (1.6)$$

$$\lim_{n \rightarrow \infty} \bar{\psi}_n(y) = \psi_{\text{lim}}(y) = +\infty, \quad y \in \mathbb{R}^d \setminus C. \quad (1.7)$$

(iii) The function  $\psi_{\text{lim}}$  is a dual optimizer for the pair  $(\mu, \nu)$ .

(iv) There is a Bass martingale  $M = (M_t)_{0 \leq t \leq 1}$  from  $\mu$  to  $\nu$  as defined in (1.3), with

$$v := \psi_{\text{lim}}^* \quad \text{and} \quad \alpha := \nabla(\psi_{\text{lim}}^* * \gamma)^*(\mu). \quad (1.8)$$

(2) Conversely, if  $\mu(I) = 1$  and the sequence  $(\psi_n(y))_{n \geq 1}$  is bounded, for all  $y \in I$ , then the pair  $(\mu, \nu)$  is irreducible and again the assertions (i) – (iv) hold.

Most parts of this theorem were already shown in [BBST23]. We will see in Section 5 that Theorem 1.5 can be viewed as a special case of Theorem 1.6 below, which describes the non-irreducible situation.

## 1.2 THE MAIN THEOREM: GENERAL CASE

We turn to the case of a pair  $(\mu, \nu)$  which is not necessarily irreducible. We refer to Section 6 for several illustrative examples of this case. As indicated by the title of the present paper, our main result is that the stretched Brownian motion  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  decomposes into a family of Bass martingales. For this purpose, we fix an optimizing sequence  $(\psi_n)_{n \geq 1}$  of convex functions for the dual problem (1.4), satisfying the regularity assumption (1.22). We also have to ensure that the sequence  $(\mathcal{D}(\psi_n))_{n \geq 1}$  decreases sufficiently fast to the optimal value  $D(\mu, \nu)$ , in the sense that

$$\sum_{n \geq 1} (\mathcal{D}(\psi_n) - D(\mu, \nu)) < +\infty. \quad (1.9)$$

We derive from the sequence  $(\psi_n)_{n \geq 1}$  the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  of  $\mathbb{R}^d$  into disjoint relatively open convex sets, which are irreducible under any  $\pi \in \text{MT}(\mu, \nu)$ , via

$$I^{\text{B}}(x) := \text{ri} \left\{ y \in \mathbb{R}^d : \sup_{n \geq 1} \sup_{\substack{\bar{\psi}_n \in [\psi_n], \\ \bar{\psi}_n(x) \leq 1}} \bar{\psi}_n(y) < +\infty \right\}, \quad x \in \mathbb{R}^d. \quad (1.10)$$

In Propositions 2.3 and 2.4 below we will show that (1.10) indeed defines a paving of  $\mathbb{R}^d$  enjoying the following properties:

- (i)  $x \in I^{\text{B}}(x)$ ,
- (ii)  $I^{\text{B}}(x) = I^{\text{B}}(x_1)$  for  $x_1 \in I^{\text{B}}(x)$ ,
- (iii)  $I^{\text{B}}(x) \cap I^{\text{B}}(x_1) = \emptyset$  for  $x_1 \notin I^{\text{B}}(x)$ .

We can thus decompose the measure  $\mu$  into its restrictions  $\mu^{I^{\text{B}}}$  to the sets  $I^{\text{B}}$ . For example, if there are at most countably many elements  $\{I_j^{\text{B}}\}_{j \geq 1}$ , we have

$$\mu(\cdot) = \sum_{j=1}^{\infty} \mu(I_j^{\text{B}}) \mu^{I_j^{\text{B}}}(\cdot), \quad \mu^{I_j^{\text{B}}}(\cdot) := \frac{1}{\mu(I_j^{\text{B}})} \mu|_{I_j^{\text{B}}}(\cdot).$$

In general, the Bass paving  $\{I^B(x)\}_{x \in \mathbb{R}^d}$  consists of uncountably many elements, see Example 6.2 below. Following [DMT19], we can still decompose the probability measure  $\mu$  as

$$\mu(\cdot) = \int_{\widehat{\mathcal{K}}} \mu^{I^B}(\cdot) \kappa^B(dI^B), \quad (1.11)$$

where  $\kappa^B$  is a Borel probability measure on the Polish space  $\widehat{\mathcal{K}}$  of closed convex subsets of  $\mathbb{R}^d$ , equipped with the Wijsman topology [Bee91]. Here, the collection  $\widehat{\mathcal{K}}$  of closed convex sets is isomorphically identified with the collection  $\text{ri } \widehat{\mathcal{K}} := \{\text{ri } C : C \in \widehat{\mathcal{K}}\}$  of their relative interiors. For each  $x \in \mathbb{R}^d$ , we denote by  $C^B(x)$  the closure of  $I^B(x)$ .

Using the stretched Brownian motion  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$ , we define

$$\nu^{I^B}(\cdot) := \int_{I^B} \pi_x^{\text{SBM}}(\cdot) \mu^{I^B}(dx). \quad (1.12)$$

We thus obtain a decomposition of the probability measure  $\nu$  via

$$\nu(\cdot) = \int_{\widehat{\mathcal{K}}} \nu^{I^B}(\cdot) \kappa^B(dI^B). \quad (1.13)$$

We now can formulate our main result.

**Theorem 1.6 (GENERAL CASE).** *Let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . Let  $(\psi_n)_{n \geq 1}$  be a sequence of convex functions satisfying the regularity assumption (1.22) and the “fast convergence condition” (1.9). Then there is a Bass paving  $\{I^B(x)\}_{x \in \mathbb{R}^d}$ , a Borel probability measure  $\kappa^B$  on  $\widehat{\mathcal{K}}$ , and  $\kappa^B$ -a.s. unique families of probability measures  $\{\mu^{I^B}\}_{I^B \in \widehat{\mathcal{K}}}$  and  $\{\nu^{I^B}\}_{I^B \in \widehat{\mathcal{K}}}$  in  $\mathcal{P}_2(\mathbb{R}^d)$ , as in (1.11) and (1.12) above, such that the following assertions hold for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ :*

(i)  $\mu^{I^B} \preceq_c \nu^{I^B}$ ,  $\mu^{I^B}(I^B) = 1$ , and  $C^B := \bar{I}^B = \widehat{\text{supp}}(\nu^{I^B})$ .

(ii) *There exist representatives  $\bar{\psi}_n^{I^B} \in [\psi_n]$  and there is a lower semicontinuous convex function  $\psi_{\text{lim}}^{I^B} : \mathbb{R}^d \rightarrow [0, +\infty]$ , which is unique modulo adding affine functions, such that*

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^{I^B}(y) = \psi_{\text{lim}}^{I^B}(y) < +\infty, \quad y \in I^B, \quad (1.14)$$

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^{I^B}(y) = \psi_{\text{lim}}^{I^B}(y) = +\infty, \quad y \in \mathbb{R}^d \setminus C^B. \quad (1.15)$$

(iii) *The function  $\psi_{\text{lim}}^{I^B}$  is a dual optimizer for the pair  $(\mu^{I^B}, \nu^{I^B})$ .*

(iv) *There is a Bass martingale  $M^{I^B} = (M_t^{I^B})_{0 \leq t \leq 1}$  from  $\mu^{I^B}$  to  $\nu^{I^B}$  as defined in (1.3), with*

$$\nu^{I^B} := (\psi_{\text{lim}}^{I^B})^* \quad \text{and} \quad \alpha^{I^B} := \nabla((\psi_{\text{lim}}^{I^B})^* * \gamma)^*(\mu^{I^B}). \quad (1.16)$$

*The Bass martingale  $M^{I^B}$  coincides with the stretched Brownian motion  $\pi^{\text{SBM}}$  restricted to  $(\mu^{I^B}, \nu^{I^B})$ .*

(v) *The pair  $(\mu^{I^B}, \nu^{I^B})$  is irreducible.*

A remarkable message of Theorem 1.6 is that an *arbitrary* sequence  $(\psi_n)_{n \geq 1}$  of convex functions of class (1.22), such that  $(\mathcal{D}(\psi_n))_{n \geq 1}$  converges sufficiently fast (see (1.9)) to the optimal value  $D(\mu, \nu)$ , already contains the *full information* about the structure (1.16) of the family of Bass martingales  $M^{I^B}$  from  $\mu^{I^B}$  to  $\nu^{I^B}$ . Note that in (1.14) and (1.15) we did *not* have to pass to subsequences or convex combinations of the sequence  $(\psi_n)_{n \geq 1}$  in order to obtain pointwise convergence.

We can also show that the Bass paving  $\{I^B(x)\}_{x \in \mathbb{R}^d}$  satisfies the maximality condition (1.17) below, as defined in [DMT19, Theorem 2.1]. In particular, Corollary 1.7 states that the martingale transport  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  is an element of the class of maximal transports  $\pi \in \text{MT}(\mu, \nu)$ , considered by H. De March and N. Touzi in [DMT19] (compare also [OS17]).

**Corollary 1.7.** *Let  $\pi \in \text{MT}(\mu, \nu)$  be an arbitrary martingale transport. For  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , we have that*

$$\widehat{\text{supp}}(\pi_x) \subseteq \widehat{\text{supp}}(\pi_x^{\text{SBM}}) = C^{\text{B}}(x). \quad (1.17)$$

For the convenience of the reader we will present a heuristic outline of the proof of Theorem 1.6 in the next subsection. The rigorous proofs of Theorem 1.6 and Corollary 1.7 are provided in Section 5.

### 1.3 A GUIDED TOUR THROUGH THE PROOF OF THEOREM 1.6

Let  $(\psi_n)_{n \geq 1}$  be a sequence of convex functions satisfying the regularity assumption (1.22) and the “fast convergence condition” (1.9). We define the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  as in (1.10) above. We will show in Propositions 2.3 and 2.4 below that this indeed defines a paving of  $\mathbb{R}^d$  into relatively open convex sets.

Next, we show that the function  $C^{\text{B}}: \mathbb{R}^d \rightarrow \widehat{\mathcal{K}}$  given by

$$x \mapsto C^{\text{B}}(x) = \overline{I^{\text{B}}(x)}, \quad (1.18)$$

where  $\widehat{\mathcal{K}}$  is equipped with the Wijsman topology, is Borel measurable. We then define  $\kappa^{\text{B}} := C^{\text{B}}(\mu)$  to be the pushforward measure of  $\mu$  under the function (1.18), which gives a Borel probability measure on  $\widehat{\mathcal{K}}$ . Applying the disintegration theorem yields the  $\kappa^{\text{B}}$ -a.s. unique family of probability measures  $\{\mu^{I^{\text{B}}}\}_{I^{\text{B}} \in \widehat{\mathcal{K}}}$  in  $\mathcal{P}_2(\mathbb{R}^d)$ , as in (1.11) above. In the following, all statements have to be understood for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , or for  $\kappa^{\text{B}}$ -a.e.  $I^{\text{B}} \in \widehat{\mathcal{K}}$ .

For fixed  $I^{\text{B}} \in \widehat{\mathcal{K}}$ , we use  $\mu^{I^{\text{B}}}$  and the stretched Brownian motion  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  to define  $\nu^{I^{\text{B}}}$  as in (1.12) above. Restricting the stretched Brownian motion  $\pi^{\text{SBM}}$  to  $(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$  yields the martingale transport

$$\pi^{I^{\text{B}}}(dx, dy) := \pi_x^{\text{SBM}}(dy) \mu^{I^{\text{B}}}(dx) \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}}) \quad (1.19)$$

from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ , so that

$$\pi^{\text{SBM}}(\cdot) = \int_{\widehat{\mathcal{K}}} \pi^{I^{\text{B}}}(\cdot) \kappa^{\text{B}}(dI^{\text{B}}). \quad (1.20)$$

It is not hard to verify (see Lemma 3.1 below) that  $\pi^{I^{\text{B}}}$  actually is the stretched Brownian motion from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ . The crucial result is that it is, in fact, a Bass martingale from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ . Let us sketch the argument how to find the dual optimizers corresponding to this family of Bass martingales.

For  $x \in \mathbb{R}^d$ , we consider representatives  $\psi_n^x \in [\psi_n]$  such that the sequence  $(\psi_n^x(x))_{n \geq 1}$  is bounded. For example, by Hahn–Banach, we can choose  $\psi_n^x \in [\psi_n]$  so that  $\psi_n^x(x) = 0$ . By definition (1.10), for  $x \in \mathbb{R}^d$ , the representatives  $(\psi_n^x)_{n \geq 1}$  are pointwise bounded on  $I^{\text{B}}(x)$ . We can even obtain pointwise convergence on  $I^{\text{B}}(x)$  to a limiting function  $\psi_{\text{lim}}^{I^{\text{B}}(x)}: I^{\text{B}}(x) \rightarrow [0, +\infty)$  by choosing (possibly different) representatives  $\bar{\psi}_n^x \in [\psi_n]$ . A possible way to find these representatives  $\bar{\psi}_n^x$  is to choose a maximal set  $\{\xi_1, \dots, \xi_m\}$  of affinely independent points in  $I^{\text{B}}(x)$  and prescribe arbitrary real values  $\{a_1, \dots, a_m\}$  at these points. Then there are affine functions  $\text{aff}_n: \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\bar{\psi}_n^x := \psi_n^x + \text{aff}_n$  attains these values on  $\xi_1, \dots, \xi_m$ , i.e.,

$$\bar{\psi}_n^x(\xi_j) = a_j, \quad j = 1, \dots, m. \quad (1.21)$$

The functions  $\text{aff}_n$  are uniquely determined by (1.21) on the affine span of  $I^{\text{B}}(x)$ . By properly choosing  $\{a_1, \dots, a_m\}$ , we can also make sure that  $\bar{\psi}_n^x \geq 0$ , so that  $\bar{\psi}_n^x \in [\psi_n]$ . Relying on [BBST23, Proposition 7.20], we come to a crucial point: it turns out that this sequence of representatives  $(\bar{\psi}_n^x)_{n \geq 1}$  indeed converges pointwise to some  $\psi_{\text{lim}}^{I^{\text{B}}(x)}$  as in (1.14) and (1.15).

Finally, we show that the limiting function  $\psi_{\text{lim}}^{I^{\text{B}}} := \psi_{\text{lim}}^{I^{\text{B}}(x)}$  is a dual optimizer for the pair  $(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$ . Theorem 1.3 above then implies that there is a Bass martingale  $(M_t^{I^{\text{B}}})_{0 \leq t \leq 1}$  from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ , given by (1.16). The resulting family of Bass martingales then decomposes the stretched Brownian motion  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  via (1.19), (1.20).

## 1.4 LITERATURE

Pavings  $\{I(x)\}_{x \in \mathbb{R}^d}$  associated to martingale transports  $\pi \in \text{MT}(\mu, \nu)$  into relatively open convex sets, which are invariant under  $\pi$  — such as the above Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  — have been studied, notably in [OS17, BJ16, DMT19, GKL19, Cio23a, Cio23b].

In [GKL19], N. Ghoussoub, Y.-H. Kim and T. Lim focused on such decompositions, for a fixed martingale transport  $\pi \in \text{MT}(\mu, \nu)$ . They obtained descriptions for the minimizers and maximizers for the cost function  $c(x, y) = |x - y|$ , when marginals are supported on  $\mathbb{R}^2$ , as well as for marginals on higher-dimensional state spaces that are in subharmonic order. Given a specific martingale  $M$ , Ghoussoub–Kim–Lim also defined a finest paving of the source space into cells that are invariant under the martingale  $M$ .

While in dimension  $d = 1$  the decomposition constructed in [GKL19] does not depend on the choice of  $\pi \in \text{MT}(\mu, \nu)$  (see the paper [BJ16] by M. Beiglböck and N. Juillet), it was noted by H. De March and N. Touzi in [DMT19] that in dimension  $d \geq 2$  this decomposition does depend on this choice. In [DMT19] and, independently, in the work [OS17] by J. Oblój and P. Siorpaes, the *finest* such decomposition which works *for all*  $\pi \in \text{MT}(\mu, \nu)$  was analyzed. In [DMT19] it was shown that such a universal decomposition exists uniquely (in an almost sure sense). In addition, a martingale transport  $\pi^{\text{DMT}} \in \text{MT}(\mu, \nu)$  has been constructed in [DMT19], which generates this universal decomposition into irreducible sets. While there is no uniqueness in the construction of such *De March–Touzi martingale transports*  $\pi^{\text{DMT}} \in \text{MT}(\mu, \nu)$  in [DMT19], Theorem 1.6 above shows, among other features, that the (unique) stretched Brownian motion  $\pi^{\text{SBM}} \in \text{MT}(\mu, \nu)$  is such a De March–Touzi martingale transport. In particular, the irreducible pavings into relatively open convex sets induced by  $\pi^{\text{DMT}}$  and  $\pi^{\text{SBM}}$  coincide (in an almost sure sense made precise below).

## 1.5 DEFINITIONS AND NOTATION

- We write  $\mathcal{P}(\mathbb{R}^d)$  for the probability measures on  $\mathbb{R}^d$  and  $\mathcal{P}_p(\mathbb{R}^d)$  for the subset of probability measures satisfying  $\int |x|^p d\mu < +\infty$ , for  $p \in [1, +\infty)$ .
- For  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ , we denote by  $\text{Cpl}(\mu, \nu)$  the set of all couplings  $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  between  $\mu$  and  $\nu$ , i.e. probability measures  $\pi$  on  $\mathbb{R}^d \times \mathbb{R}^d$  with first marginal  $\mu$  and second marginal  $\nu$ .
- We say that  $\mu \in \mathcal{P}_1(\mathbb{R}^d)$  is dominated by  $\nu \in \mathcal{P}_1(\mathbb{R}^d)$  in convex order and write  $\mu \preceq_c \nu$ , if for all convex functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  we have  $\int f d\mu \leq \int f d\nu$ .
- For  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$  we define the collection of martingale transports  $\text{MT}(\mu, \nu)$  as those couplings  $\pi \in \text{Cpl}(\mu, \nu)$  with barycenter  $\text{bary}(\pi_x) := \int y \pi_x(dy) = x$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ . Here, the family of probability measures  $\{\pi_x\}_{x \in \mathbb{R}^d} \subseteq \mathcal{P}(\mathbb{R}^d)$  is obtained by disintegrating the coupling  $\pi$  with respect to its first marginal  $\mu$ , i.e.,  $\pi(dx, dy) = \pi_x(dy) \mu(dx)$ .
- The  $d$ -dimensional standard Gaussian distribution is denoted by  $\gamma$ .
- We denote by  $C_q(\mathbb{R}^d)$  the set of continuous functions  $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$  with quadratic growth, meaning that there are constants  $a, k, \ell \in \mathbb{R}$  with

$$\ell + \frac{|\cdot|^2}{2} \leq \psi(\cdot) \leq a + k|\cdot|^2.$$

We also introduce the set

$$C_q^{\text{aff}}(\mathbb{R}^d) := \{\psi(\cdot) + \text{aff}(\cdot) : \psi \in C_q(\mathbb{R}^d), \text{aff}: \mathbb{R}^d \rightarrow \mathbb{R} \text{ is affine}\}. \quad (1.22)$$

- For two measures  $\rho$  and  $q$  on  $\mathbb{R}^d$  we write  $\rho * q$  for their convolution. If  $f$  is a function, the convolution of  $f$  and  $q$  is defined as

$$(f * q)(x) := \int f(x - y) q(dy), \quad x \in \mathbb{R}^d,$$

provided  $f(x - \cdot) \in L^1(q)$ . In particular,  $(f * \gamma)(x) = \int f(x + y) \gamma(dy)$ .

- For a function  $f: \mathbb{R}^d \rightarrow (-\infty, +\infty]$ , its convex conjugate is given by

$$f^*(y) := \sup_{x \in \mathbb{R}^d} (\langle x, y \rangle - f(x)), \quad y \in \mathbb{R}^d,$$

and we write

$$\text{dom } f := \{x \in \mathbb{R}^d: f(x) < +\infty\}$$

for the domain of  $f$ .

- We write  $\text{ri}(A)$  and  $\bar{A}$  for the relative interior and the closure of a set  $A \subseteq \mathbb{R}^d$ , respectively.
- The support and the closed convex hull of the support of a measure  $\rho$  are denoted by  $\text{supp}(\rho)$  and  $\widehat{\text{supp}}(\rho)$ , respectively.
- The Polish space  $\widehat{\mathcal{K}}$  of closed convex subsets of  $\mathbb{R}^d$  is equipped with the Wijsman topology [Bee91] and is isomorphically identified with  $\text{ri } \widehat{\mathcal{K}} = \{\text{ri } C: C \in \widehat{\mathcal{K}}\}$ .

## 1.6 RELAXED FORMULATION OF THE DUAL FUNCTION

Let us return to the technical caveat we made after the statement of Theorem 1.3. For  $\psi \in L^1(\nu)$ , we recall the definition (1.5) of the dual function  $\mathcal{D}(\cdot)$ , which we rewrite as

$$\mathcal{D}(\psi) = \int \left( \int \psi(y) \pi_x(dy) - (\psi^* * \gamma)^*(x) \right) \mu(dx), \quad (1.23)$$

with  $\pi$  being an arbitrary fixed element of  $\text{MT}(\mu, \nu)$ . As shown in [BBST23, Proposition 4.2], the dual function written in the “relaxed” form (1.23) is well-defined and takes values in the interval  $[0, +\infty]$ , for a general convex function  $\psi: \mathbb{R}^d \rightarrow (-\infty, +\infty]$  which is  $\mu$ -a.s. finite. Furthermore, by [BBST23, Lemma 3.5] and [BBST23, Proposition 4.1], the value  $D(\mu, \nu)$  of the dual problem (1.4) is equal to

$$D(\mu, \nu) = \inf_{\substack{\psi \in C_q^{\text{aff}}(\mathbb{R}^d), \\ \psi \text{ convex}}} \mathcal{D}(\psi) = \inf_{\substack{\mu(\text{dom } \psi) = 1, \\ \psi \text{ convex}}} \mathcal{D}(\psi). \quad (1.24)$$

Thus, whenever convenient, by (1.24) we are free to choose optimizing sequences  $(\psi_n)_{n \geq 1}$  for the dual problem (1.4) of class  $C_q^{\text{aff}}(\mathbb{R}^d)$ , as defined in (1.22) above.

**Definition 1.8.** We say that a lower semicontinuous convex function  $\psi: \mathbb{R}^d \rightarrow (-\infty, +\infty]$  satisfying  $\mu(\text{ri}(\text{dom } \psi)) = 1$  is an optimizer of the dual problem (1.24) if  $D(\mu, \nu) = \mathcal{D}(\psi)$ , for the dual function  $\mathcal{D}(\cdot)$  as defined in (1.23). In short, we say that  $\psi$  is a dual optimizer.

We refer to Example 6.7 in [BBST23] for a case of a dual optimizer  $\psi$  in the sense of Definition 1.8 which is not integrable with respect to  $\nu$ .



## 2 EQUIVALENCE CLASSES OF CONVEX FUNCTIONS ON $\mathbb{R}^d$

In this section we will analyze some phenomena arising for the *equivalence classes*  $[\psi]$  of convex functions  $\psi$  modulo adding affine functions, as in Definition 1.4.

**Definition 2.1.** Let  $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$  be a convex function. For  $x \in \mathbb{R}^d$  and  $\varepsilon > 0$ , we denote by  $[\psi]^{x,\varepsilon}$  the equivalence class of convex functions of the form

$$\psi^{x,\varepsilon}(\cdot) = \psi_n(\cdot) + \text{aff}(\cdot), \quad (2.1)$$

where  $\text{aff}: \mathbb{R}^d \rightarrow \mathbb{R}$  runs through all affine functions on  $\mathbb{R}^d$  such that

$$\psi^{x,\varepsilon}(\cdot) \geq 0 \quad \text{and} \quad \psi^{x,\varepsilon}(x) < \varepsilon. \quad (2.2)$$

**Definition 2.2.** Let  $(\psi_n)_{n \geq 1}$  be a sequence of real-valued convex functions on  $\mathbb{R}^d$ . For  $x \in \mathbb{R}^d$ , we define the sets

$$I^b(x) := \left\{ y \in \mathbb{R}^d : \limsup_{\varepsilon \downarrow 0} \sup_{n \geq 1} \sup_{\psi_n^{x,\varepsilon} \in [\psi_n]^{x,\varepsilon}} \psi_n^{x,\varepsilon}(y) < +\infty \right\}, \quad (2.3)$$

$$I^B(x) := \text{ri } I^b(x), \quad (2.4)$$

$$C^B(x) := \overline{I^b(x)}. \quad (2.5)$$

Here, the superscripts “b” and “B” stand for “bounded” and “Bass”, respectively. The operator  $\text{ri}$  denotes the relative interior of a set. We will prove in Proposition 2.4 below that definition (1.10) of  $I^B(x)$  coincides with Definition 2.2, thus justifying this abuse of notation.

**Proposition 2.3.** Let  $(\psi_n)_{n \geq 1}$  be a sequence of real-valued convex functions on  $\mathbb{R}^d$ . The collection  $\{I^B(x)\}_{x \in \mathbb{R}^d}$ , as introduced in Definition 2.2 above, has the following properties:

- (i)  $x \in I^B(x)$ ,
- (ii) for  $x_1 \in I^B(x)$  we have  $I^B(x) = I^B(x_1)$ ,
- (iii) for  $x_1 \notin I^B(x)$  we have  $I^B(x) \cap I^B(x_1) = \emptyset$ .

In conclusion, the collection  $\{I^B(x)\}_{x \in \mathbb{R}^d}$  is a paving of  $\mathbb{R}^d$  into relatively open convex sets.

*Proof of Proposition 2.3. (i):* Let  $x \in \mathbb{R}^d$ . We first note that both  $I^b(x)$  and  $I^B(x)$  are convex sets and by definition the latter set is relatively open. Moreover, it is clear that  $x \in I^b(x)$ . To show that  $x$  is in the relative interior of  $I^b(x)$ , we suppose that  $x \notin I^B(x)$ , i.e., that  $x$  is a relative boundary point of  $I^b(x)$ . By [Roc70, Corollary 11.6.2] there exists a functional  $y^* \in \mathbb{R}^d$  and some  $\bar{x} \in I^b(x)$  such that

$$\langle \bar{x}, y^* \rangle < \max_{y \in I^b(x)} \langle y, y^* \rangle = \langle x, y^* \rangle. \quad (2.6)$$

For notational convenience and without loss of generality we assume that  $|x - \bar{x}| = 1$ . We will show that

$$\limsup_{\varepsilon \downarrow 0} \sup_{n \geq 1} \sup_{\psi_n^{x,\varepsilon} \in [\psi_n]^{x,\varepsilon}} \psi_n^{x,\varepsilon}(\bar{x}) = +\infty, \quad (2.7)$$

so that  $\bar{x} \notin I^b(x)$ , which will yield the desired contradiction.

We draw a straight line through  $\bar{x}$  and  $x$ , i.e., we consider the function

$$t \mapsto x + t(x - \bar{x}), \quad t \in \mathbb{R}. \quad (2.8)$$

For  $\varepsilon > 0$ , let us consider the point  $y = y(\varepsilon) = x + \varepsilon^2(x - \bar{x})$  on this line. By (2.6) the point  $y$  cannot be an element of  $I^b(x)$ . Thus, for  $\varepsilon \in (0, 1)$  we can find  $n \geq 1$  and a function  $\psi_n^{x,\varepsilon} \in [\psi_n]^{x,\varepsilon}$  such that

$$\psi_n^{x,\varepsilon}(x + \varepsilon^2(x - \bar{x})) - \psi_n^{x,\varepsilon}(x) \geq 1. \quad (2.9)$$

By elementary geometry we observe that the convex function  $\psi_n^{x,\varepsilon}$  in (2.9) has the following property: the affine function  $\text{af}(\cdot)$  on the real line given by

$$t \mapsto \text{af}(t) := \frac{1}{\varepsilon}t - \varepsilon, \quad t \in \mathbb{R} \quad (2.10)$$

is less than or equal to the restriction of  $\psi_n^{x,\varepsilon}$  to the line (2.8), i.e., the function

$$t \mapsto \psi_n^{x,\varepsilon}(x + t(x - \bar{x})), \quad t \in \mathbb{R}.$$

By Hahn–Banach we can extend the function  $x + t(x - \bar{x}) \mapsto \text{af}(t)$  to an affine function  $\text{aff}(\cdot)$  defined on  $\mathbb{R}^d$  such that  $\text{aff}(y) \leq \psi_n^{x,\varepsilon}(y)$  for all  $y \in \mathbb{R}^d$ . Note that we have  $\text{aff}(x) = \text{af}(0) = -\varepsilon$ . Therefore the function

$$y \mapsto \psi_n^{x,2\varepsilon}(y) := \psi_n^{x,\varepsilon}(y) - \text{aff}(y), \quad y \in \mathbb{R}^d$$

is a representative of  $[\psi_n]^{x,2\varepsilon}$ . At the same time we have  $\text{aff}(\bar{x}) = \text{af}(-1) = -\frac{1}{\varepsilon} - \varepsilon$  and thus

$$\psi_n^{x,2\varepsilon}(\bar{x}) = \psi_n^{x,\varepsilon}(\bar{x}) - \text{aff}(\bar{x}) \geq \frac{1}{\varepsilon} + \varepsilon \geq \frac{1}{\varepsilon}.$$

This contradicts the assumption  $\bar{x} \in I^b(x)$  and finishes the proof of (i).

(ii): Let  $x \in \mathbb{R}^d$ . We first show the following claim: *For every compact subset  $K \subseteq I^B(x)$  there is  $\bar{M} > 0$  and  $\bar{\varepsilon} \in (0, 1]$  such that for each  $n \in \mathbb{N}$  and  $\psi_n^{x,\bar{\varepsilon}} \in [\psi_n]^{x,\bar{\varepsilon}}$  we have*

$$K \subseteq \{\psi_n^{x,\bar{\varepsilon}} \leq \bar{M}\}. \quad (2.11)$$

For the proof of the claim, we define the set

$$B_{M,\varepsilon} := \bigcap_{n \in \mathbb{N}} \bigcap_{\psi_n^{x,\varepsilon} \in [\psi_n]^{x,\varepsilon}} \{\psi_n^{x,\varepsilon} \leq M\}.$$

By Definition 2.2, we have

$$\bigcup_{M>0} \bigcup_{\varepsilon>0} B_{M,\varepsilon} = I^b(x) \supseteq I^B(x).$$

We denote by  $\text{Aff}(x)$  the affine span of  $I^B(x)$  and by  $B_{M,\varepsilon}^\circ$  the relative interior of  $B_{M,\varepsilon}$  in  $\text{Aff}(x)$ . Then for every compact subset  $K \subseteq I^B(x)$  we have

$$\bigcup_{M>0} \bigcup_{\varepsilon>0} B_{M,\varepsilon}^\circ = I^B(x) \supseteq K.$$

By compactness there is some  $\bar{M} > 0$  and  $\bar{\varepsilon} > 0$  such that

$$B_{\bar{M},\bar{\varepsilon}}^\circ \supseteq K,$$

which readily implies (2.11).

Let  $x_1 \in I^B(x)$  and choose  $\delta > 0$  such that

$$K_\delta := \bar{B}_\delta(x_1) \cap \text{Aff}(x) \subseteq I^B(x).$$

Take  $\bar{M} > 0$  and  $\bar{\varepsilon} \in (0, 1]$  such that (2.11) is satisfied for  $K = K_\delta$ . Now we make the following claim: *There exists a constant  $L > 0$  such that, for every  $\varepsilon \in (0, \bar{\varepsilon}]$  and every choice of representatives  $\psi_n^{x,\varepsilon} \in [\psi_n]^{x,\varepsilon}$  and  $\psi_n^{x_1,\varepsilon} \in [\psi_n]^{x_1,\varepsilon}$ , the affine functions*

$$\text{aff}_n := \psi_n^{x,\varepsilon} - \psi_n^{x_1,\varepsilon}$$

satisfy the Lipschitz estimate

$$\forall y, z \in \text{Aff}(x): |\text{aff}_n(y) - \text{aff}_n(z)| \leq \frac{L}{\delta} |y - z|. \quad (2.12)$$

Negating the claim, we suppose that, for every  $L > 0$ , there is a point  $x_2 \in B_\delta(x_1) \cap \text{Aff}(x)$  and  $n \in \mathbb{N}$  such that

$$\text{aff}_n(x_2) - \text{aff}_n(x_1) \geq L.$$

We then obtain

$$\psi_n^{x,\varepsilon}(x_2) - \psi_n^{x,\varepsilon}(x_1) \geq \psi_n^{x_1,\varepsilon}(x_2) - \psi_n^{x_1,\varepsilon}(x_1) + L.$$

Using the estimates  $\psi_n^{x,\varepsilon}(x_1) \geq 0$ ,  $\psi_n^{x_1,\varepsilon}(x_2) \geq 0$ , and  $\psi_n^{x_1,\varepsilon}(x_1) \leq 1$ , we conclude that

$$\psi_n^{x,\varepsilon}(x_2) \geq L - 1,$$

which is a contradiction to  $\psi_n^{x,\varepsilon}(x_2) \leq \bar{M}$ , thus finishing the proof of (2.12).

Finally, it follows from (2.12) that

$$\forall x_1 \in I^{\text{B}}(x): I^{\text{b}}(x) \cap \text{Aff}(x) = I^{\text{b}}(x_1) \cap \text{Aff}(x). \quad (2.13)$$

To obtain the full assertion (ii) from (2.13), we still have to show that  $I^{\text{b}}(x_1)$ , and therefore also  $I^{\text{B}}(x_1)$ , is contained in the affine space  $\text{Aff}(x)$ . If this were not the case, the affine space  $\text{Aff}(x_1)$  spanned by  $I^{\text{B}}(x_1)$  would strictly contain  $\text{Aff}(x)$ . Reversing the roles of  $x$  and  $x_1$ , and applying (2.12) to  $\text{Aff}(x_1)$  rather than  $\text{Aff}(x)$ , we arrive at the desired contradiction, which completes the proof of (ii).

(iii): The final statement (iii) is a consequence of (i) and (ii).  $\square$

We next show a variant of Proposition 2.3, where we do not have to refer to  $\varepsilon > 0$  as in Definitions 2.1 and 2.2 above.

**Proposition 2.4.** *Let  $(\psi_n)_{n \geq 1}$  be a sequence of  $(-\infty, +\infty]$ -valued convex functions on  $\mathbb{R}^d$ . Fix representatives  $\bar{\psi}_n \in [\psi_n]$ ,  $x \in \mathbb{R}^d$ , and  $I^{\text{B}}(x)$  as defined in Definition 2.2. The sequence  $(\bar{\psi}_n(y))_{n \geq 1}$  is bounded for some  $y \in I^{\text{B}}(x)$  if and only if it is bounded for all  $y \in I^{\text{B}}(x)$ . In this case,  $(\bar{\psi}_n)_{n \geq 1}$  is uniformly bounded on compact subsets of  $I^{\text{B}}(x)$ .*

*In particular, the definition (1.10) of  $I^{\text{B}}(x)$  coincides with Definition 2.2.*

*Proof.* Let  $(\varepsilon_n)_{n \geq 1}$  be a null sequence and take  $\psi_n^{x,\varepsilon_n} \in [\psi_n]^{x,\varepsilon_n}$ . By definition (2.3), (2.4) of  $I^{\text{B}}(x)$ , we have that  $(\psi_n^{x,\varepsilon_n}(y))_{n \geq 1}$  is bounded, for all  $y \in I^{\text{B}}(x)$ . For arbitrary representatives  $\bar{\psi}_n \in [\psi_n]$  we have that  $\text{aff}_n := \bar{\psi}_n - \psi_n^{x,\varepsilon_n}$  is an affine function on  $\mathbb{R}^d$ . As  $\bar{\psi}_n \geq 0$ , we have that  $(\text{aff}_n(y))_{n \geq 1}$  is bounded from below, for every  $y \in I^{\text{B}}(x)$ . If we also have that  $(\bar{\psi}_n(y_0))_{n \geq 1}$  is bounded, for some  $y_0 \in I^{\text{B}}(x)$ , we obtain that  $(\text{aff}_n(y_0))_{n \geq 1}$  is bounded. This implies that  $(\text{aff}_n(y))_{n \geq 1}$  is bounded, for all  $y$  in the affine span of  $I^{\text{B}}(x)$ . In particular,  $(\bar{\psi}_n(y))_{n \geq 1}$  is bounded, for all  $y \in I^{\text{B}}(x)$ . By convexity,  $(\bar{\psi}_n)_{n \geq 1}$  is uniformly bounded on compact subsets of  $I^{\text{B}}(x)$ .

As regards the final assertion, the set  $I^{\text{B}}(x)$  defined in (1.10) obviously is contained in the set  $I^{\text{B}}(x)$  defined in (2.3), (2.4). For the reverse inclusion, suppose there is  $y \in \mathbb{R}^d$  and a sequence  $(\psi_n^{x,\varepsilon_n})_{n \geq 1}$  such that  $(\psi_n^{x,\varepsilon_n}(y))_{n \geq 1}$  is unbounded, while, for every choice of representatives  $\bar{\psi}_n \in [\psi_n]$  with  $\bar{\psi}_n(x) \leq 1$ , the sequence  $(\bar{\psi}_n(y))_{n \geq 1}$  is bounded. This is a contradiction to the first part of Proposition 2.4, finishing the proof of the final assertion.  $\square$

So far, the sequence  $(\psi_n)_{n \geq 1}$  of real-valued convex functions on  $\mathbb{R}^d$  had nothing to do with the pair  $(\mu, \nu)$ . We now impose the additional assumption that  $(\psi_n)_{n \geq 1} \subseteq C_{\text{q}}^{\text{aff}}(\mathbb{R}^d)$  is an optimizing sequence of convex functions for the dual problem (1.4). In this way, we can relate the support of  $\nu$  to  $I^{\text{B}}(x)$ .

**Proposition 2.5.** Let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be an optimizing sequence of convex functions for the dual problem (1.4), with induced Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$ . Suppose that there is  $x \in \mathbb{R}^d$  such that, for  $I^{\text{B}} := I^{\text{B}}(x)$ , we have  $\mu(I^{\text{B}}) = 1$ . Then for  $C^{\text{B}} = \overline{I^{\text{B}}}$  we have

$$\widehat{\text{supp}}(\nu) \subseteq C^{\text{B}}. \quad (2.14)$$

*Proof.* Applying [BBST23, Lemma 7.9] gives

$$a := \sup_{n \geq 1} \int_{x \in I^{\text{B}}} \int_{y \in \mathbb{R}^d} (\psi_n(y) - \psi_n(x)) \pi_x^{\text{SBM}}(dy) \mu(dx) < +\infty. \quad (2.15)$$

Arguing by contradiction, we assume that (2.14) is not true. Then we can find an open half space  $H$  with  $H \cap C^{\text{B}} = \emptyset$  such that  $\nu(H) > 0$ . By Proposition 2.4, we can choose representatives  $\bar{\psi}_n \in [\psi_n]$  such that the sequence  $(\bar{\psi}_n)_{n \geq 1}$  is pointwise bounded on  $I^{\text{B}}$ . We denote by  $\tilde{\psi}_n \in L^1(\nu)$  the function which coincides with  $\bar{\psi}_n$  on  $C^{\text{B}}$ , and assumes the value  $+\infty$  outside of  $C^{\text{B}}$ . As observed after Definition 1.4, the sequence  $(\tilde{\psi}_n)_{n \geq 1}$  still is an optimizing sequence of lower semicontinuous convex functions in  $L^1(\nu)$ . It also (trivially) satisfies

$$\lim_{n \rightarrow \infty} \tilde{\psi}_n(y) = +\infty, \quad y \in H. \quad (2.16)$$

Since  $\nu(H) > 0$ , we can find a compact subset  $K \subseteq I^{\text{B}}$  such that

$$\mu(\{x \in K : \pi_x^{\text{SBM}}(H) > 0\}) > 0. \quad (2.17)$$

As a consequence of (2.17), the set  $K \times H$  has positive  $\pi^{\text{SBM}}$ -measure.

Now observe that (2.15) also holds when we replace  $\psi_n$  by  $\tilde{\psi}_n$ . Furthermore, as

$$\int (\tilde{\psi}_n(y) - \tilde{\psi}_n(x)) \pi_x^{\text{SBM}}(dy) \geq 0,$$

for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , we conclude that

$$a \geq \sup_{n \geq 1} \int_{x \in K} \int_{y \in \mathbb{R}^d} (\tilde{\psi}_n(y) - \tilde{\psi}_n(x)) \pi_x^{\text{SBM}}(dy) \mu(dx).$$

As  $(\tilde{\psi}_n)_{n \geq 1}$  is uniformly bounded on compact subsets of  $I^{\text{B}}$ , we have

$$m := \sup_{n \geq 1} \sup_{x \in K} \tilde{\psi}_n(x) < +\infty.$$

Consequently,

$$a \geq \sup_{n \geq 1} \int_{x \in K} \int_{y \in \mathbb{R}^d} \tilde{\psi}_n(y) \pi_x^{\text{SBM}}(dy) \mu(dx) - m\mu(K).$$

Using the fact that  $\tilde{\psi}_n \geq 0$ , we have

$$a \geq \sup_{n \geq 1} \int_{y \in \mathbb{R}^d} \int_{x \in K} \tilde{\psi}_n(y) \pi_x^{\text{SBM}}|_H(dy) \mu(dx) - m\mu(K). \quad (2.18)$$

Next we define a probability measure  $\tilde{\nu}$  on  $\mathbb{R}^d$  by

$$\tilde{\nu}(\cdot) := \frac{1}{c} \int_K \pi_x^{\text{SBM}}|_H(\cdot) \mu(dx),$$

with normalizing constant

$$c := \int_K \pi_x^{\text{SBM}}(H) \mu(dx).$$

Note that  $c > 0$  since  $\pi^{\text{SBM}}(K \times H) > 0$ . In terms of  $\tilde{\nu}$ , we can express the right-hand side of (2.18) as

$$c \sup_{n \geq 1} \int_{\mathbb{R}^d} \tilde{\psi}_n(y) \tilde{\nu}(dy) - m\mu(K).$$

Applying Jensen's inequality yields

$$a \geq c \sup_{n \geq 1} \tilde{\psi}_n(\tilde{y}) - m\mu(K), \quad (2.19)$$

where

$$\tilde{y} := \text{bary}(\tilde{\nu}) = \int_{\mathbb{R}^d} y \tilde{\nu}(dy)$$

denotes the barycenter of the probability measure  $\tilde{\nu}$ . Observing that  $\tilde{y} \in H$  and using (2.16), the right-hand side of (2.19) is equal to  $+\infty$ . This is a contradiction to the finiteness of  $a$  and shows the inclusion (2.14).  $\square$

Combining Proposition 2.5 with [BBST23, Proposition 7.20], we obtain the following result.

**Proposition 2.6.** *Let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be an optimizing sequence of convex functions for the dual problem (1.4). Suppose that there is  $x \in \mathbb{R}^d$  such that for  $I^{\text{B}} := I^{\text{B}}(x)$  we have  $\mu(I^{\text{B}}) = 1$ . Then there exist representatives  $\bar{\psi}_n \in [\psi_n]$  and a lower semicontinuous convex function  $\psi_{\text{lim}}: \mathbb{R}^d \rightarrow [0, +\infty]$  such that*

$$\lim_{n \rightarrow \infty} \bar{\psi}_n(y) = \psi_{\text{lim}}(y) < +\infty, \quad y \in I^{\text{B}}, \quad (2.20)$$

$$\lim_{n \rightarrow \infty} \bar{\psi}_n(y) = \psi_{\text{lim}}(y) = +\infty, \quad y \in \mathbb{R}^d \setminus C^{\text{B}}. \quad (2.21)$$

Furthermore,

$$\widehat{\text{supp}}(\nu) = C^{\text{B}}, \quad (2.22)$$

and the function  $\psi_{\text{lim}}$  is a dual optimizer for the pair  $(\mu, \nu)$ .

*Proof.* Note that  $\mu(I^{\text{B}}) = 1$  by assumption and  $\widehat{\text{supp}}(\nu) \subseteq C^{\text{B}}$  by Proposition 2.5. Furthermore, by Proposition 2.4, we can choose representatives  $\tilde{\psi}_n \in [\psi_n]$  such that the sequence  $(\tilde{\psi}_n)_{n \geq 1}$  is pointwise bounded on  $I^{\text{B}}$ , i.e.

$$\forall y \in I^{\text{B}}: \sup_{n \geq 1} \tilde{\psi}_n(y) < +\infty.$$

By [BBST23, Proposition 7.20], we conclude that there are representatives  $\bar{\psi}_n \in [\tilde{\psi}_n] = [\psi_n]$  such that (2.20) and (2.21) are satisfied; moreover, we have the identity (2.22), and  $\psi_{\text{lim}}$  is a dual optimizer.  $\square$

In order to apply the disintegration theorem, we will need the following measurability property of the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$ .

**Proposition 2.7.** *Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be a sequence of convex functions. The map  $C^{\text{B}}: \mathbb{R}^d \rightarrow \widehat{\mathcal{K}}$  given by*

$$x \longmapsto C^{\text{B}}(x) = \overline{I^{\text{B}}(x)}, \quad (2.23)$$

where  $\widehat{\mathcal{K}}$  is equipped with the Wijsman topology, is Borel measurable.

*Proof.* As noted in [DMT19], it suffices to show that the set

$$B^V := \{x \in \mathbb{R}^d: C^{\text{B}}(x) \cap V \neq \emptyset\}$$

is Borel measurable, for every open set  $V \subseteq \mathbb{R}^d$ . To see this, we fix an open set  $V \subseteq \mathbb{R}^d$  and note that

$$B^V = \{x \in \mathbb{R}^d: I^{\text{b}}(x) \cap V \neq \emptyset\}.$$

Consequently,

$$B^V = \bigcup_{\varepsilon > 0} \bigcup_{M > 0} \bigcap_{n \in \mathbb{N}} \bigcap_{\substack{\bar{\psi}_n \in [\psi_n], \\ \{\bar{\psi}_n < M\} \cap V \neq \emptyset}} \{x \in \mathbb{R}^d : \bar{\psi}_n(x) < \varepsilon\}.$$

Hence  $B^V$  is Borel measurable if we can show that the last intersection can be replaced by a countable intersection. In order to do so, we fix a basis  $(e_j)_{j=1}^d$  of  $\mathbb{R}^d$  as well as representatives  $\bar{\psi}_n \in [\psi_n]$ , and define the equivalence class  $[\psi_n]^\mathbb{Q}$  to consist of all functions  $\tilde{\psi}_n \in [\psi_n]$  such that the affine function  $\bar{\psi}_n - \tilde{\psi}_n$  has rational coefficients when expressed in the basis  $(e_j)_{j=1}^d$ . We then have to show that

$$U := \bigcap_{\substack{\bar{\psi}_n \in [\psi_n], \\ \{\bar{\psi}_n < M\} \cap V \neq \emptyset}} \{x \in \mathbb{R}^d : \bar{\psi}_n(x) < \varepsilon\} = \bigcap_{\substack{\tilde{\psi}_n \in [\psi_n]^\mathbb{Q}, \\ \{\tilde{\psi}_n < M\} \cap V \neq \emptyset}} \{x \in \mathbb{R}^d : \tilde{\psi}_n(x) < \varepsilon\} =: U^\mathbb{Q}.$$

Clearly  $U \subseteq U^\mathbb{Q}$ , so that it remains to show the inclusion  $\mathbb{R}^d \setminus U \subseteq \mathbb{R}^d \setminus U^\mathbb{Q}$ . Fix  $x \in \mathbb{R}^d$  and  $\bar{\psi}_n \in [\psi_n]$  such that  $\{\bar{\psi}_n < M\} \cap V \neq \emptyset$  and  $\{\bar{\psi}_n(x) \geq \varepsilon\}$ , i.e.,  $x \in \mathbb{R}^d \setminus U$ . Choose  $\eta > 0$  such that we still have

$$\{\bar{\psi}_n + \eta < M\} \cap V \neq \emptyset.$$

Note that  $\bar{\psi}_n + \eta$  is bounded from below by  $\eta$  and that the function  $\bar{\psi}_n$  is in  $C_q^{\text{aff}}(\mathbb{R}^d)$ , so that  $\bar{\psi}_n(y) \geq \frac{|y|^2}{4}$  for  $y$  outside a bounded set. Hence we can find a representative  $\tilde{\psi}_n \in [\psi_n]^\mathbb{Q}$  such that

$$\tilde{\psi}_n \geq 0, \quad \{\tilde{\psi}_n < M\} \cap V \neq \emptyset \quad \text{and} \quad \{\tilde{\psi}_n(x) \geq \varepsilon\}.$$

Therefore  $x \in \mathbb{R}^d \setminus U^\mathbb{Q}$ . This shows the claim  $U = U^\mathbb{Q}$  and we conclude that the set  $B^V$  is Borel measurable.  $\square$

Having established the Borel measurability of the map  $x \mapsto C^{\text{B}}(x)$ , we can disintegrate the probability measure  $\mu$  with respect to this mapping.

**Lemma 2.8.** *Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be a sequence of convex functions. Define  $\kappa^{\text{B}} := C^{\text{B}}(\mu)$  to be the pushforward measure of  $\mu$  under the map  $C^{\text{B}}: \mathbb{R}^d \rightarrow \hat{\mathcal{K}}$  of (2.23), which induces a probability measure on  $\hat{\mathcal{K}}$  (recall that we identify  $\hat{\mathcal{K}}$  with  $\text{ri } \hat{\mathcal{K}}$ ). There exists a  $\kappa^{\text{B}}$ -a.s. unique family of probability measures  $\{\mu^{I^{\text{B}}}\}_{I^{\text{B}} \in \hat{\mathcal{K}}} \subseteq \mathcal{P}_2(\mathbb{R}^d)$  such that  $\mu^{I^{\text{B}}}(I^{\text{B}}) = 1$ , for  $\kappa^{\text{B}}$ -a.e.  $I^{\text{B}} \in \hat{\mathcal{K}}$ , and*

$$\mu(\cdot) = \int_{\hat{\mathcal{K}}} \mu^{I^{\text{B}}}(\cdot) \kappa^{\text{B}}(dI^{\text{B}}). \quad (2.24)$$

*Proof.* The statement follows from applying the disintegration theorem (see, for instance, [AGS08, Theorem 5.3.1] or [DM78, Chapter III, p. 78]), which is possible thanks to the Borel measurability of the function (2.23).  $\square$

Note that Lemma 2.8 rigorously justifies the decomposition (1.11) as already stated in Section 1.

We refer to Examples <sup>1</sup> and <sup>2</sup> below for a concrete visualization of Lemma 2.8. For example, the measure  $\kappa^{\text{B}}$  in <sup>3</sup>, considered as a probability measure on  $\hat{\mathcal{K}}$ , gives probabilities  $\frac{1}{2}$  to  $\mathbb{R}_- = (-\infty, 0)$  and  $\mathbb{R}_+ = (0, +\infty)$ , as well as probability 0 to the singleton  $\{0\}$ . The Bass paving consists of the relatively open sets  $\mathbb{R}_-, \{0\}, \mathbb{R}_+$ , where  $\{0\}$  has measure zero with respect to  $\kappa^{\text{B}}$  and may therefore be neglected. The probability measures  $\mu^{\mathbb{R}_-}$  and  $\mu^{\mathbb{R}_+}$  are the normalized restrictions of  $\mu$  to  $\mathbb{R}_-$  and  $\mathbb{R}_+$ , respectively.

---

<sup>1</sup>Include reference.

<sup>2</sup>Include reference.

<sup>3</sup>Include reference.

### 3 THE DECOMPOSITION OF THE PRIMAL PROBLEM

The disintegration result of Lemma 2.8 allows us to decompose the primal problem (1.1) from  $\mu$  to  $\nu$  into a family of *local* primal problems, defined on  $\kappa^{\text{B}}$ -a.e. set  $I^{\text{B}} \in \widehat{\mathcal{K}}$ . To this end, we define (as in (1.12) above) the  $\kappa^{\text{B}}$ -a.s. unique family of probability measures  $\{\nu^{I^{\text{B}}}\}_{I^{\text{B}} \in \widehat{\mathcal{K}}} \subseteq \mathcal{P}_2(\mathbb{R}^d)$  by

$$\nu^{I^{\text{B}}}(\cdot) := \int_{I^{\text{B}}} \pi_x^{\text{SBM}}(\cdot) \mu^{I^{\text{B}}}(dx), \quad (3.1)$$

where  $\pi_x^{\text{SBM}}(dx, dy) = \pi_x^{\text{SBM}}(dy) \mu(dx) \in \text{MT}(\mu, \nu)$  is the optimizer of the primal problem (1.1). As noted in (1.13), the family  $\{\nu^{I^{\text{B}}}\}_{I^{\text{B}} \in \widehat{\mathcal{K}}}$  then is a decomposition of  $\nu$ , i.e.

$$\nu(\cdot) = \int_{\widehat{\mathcal{K}}} \nu^{I^{\text{B}}}(\cdot) \kappa^{\text{B}}(dI^{\text{B}}). \quad (3.2)$$

By construction we have that  $\mu^{I^{\text{B}}} \preceq_c \nu^{I^{\text{B}}}$ , for  $\kappa^{\text{B}}$ -a.e.  $I^{\text{B}} \in \widehat{\mathcal{K}}$ . We obtain the following decomposition of the primal problem (1.1).

**Lemma 3.1** (DECOMPOSITION OF THE PRIMAL PROBLEM). *For  $\kappa^{\text{B}}$ -a.e.  $I^{\text{B}} \in \widehat{\mathcal{K}}$ , the optimizer of the local primal problem*

$$P(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}}) = \sup_{\pi \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})} \int \text{MCov}(\pi_x, \gamma) \mu^{I^{\text{B}}}(dx)$$

from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$  equals

$$\pi^{I^{\text{B}}}(dx, dy) := \pi_x^{\text{SBM}}(dy) \mu^{I^{\text{B}}}(dx) \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}}). \quad (3.3)$$

In particular,

$$P(\mu, \nu) = \sup_{\pi \in \text{MT}(\mu, \nu)} \int \text{MCov}(\pi_x, \gamma) \mu(dx) = \int \text{MCov}(\pi_x^{\text{SBM}}, \gamma) \mu(dx) \quad (3.4)$$

$$= \int_{\widehat{\mathcal{K}}} \left( \int \text{MCov}(\pi_x^{\text{SBM}}, \gamma) \mu^{I^{\text{B}}}(dx) \right) \kappa^{\text{B}}(dI^{\text{B}}) \quad (3.5)$$

$$= \int_{\widehat{\mathcal{K}}} P(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}}) \kappa^{\text{B}}(dI^{\text{B}}). \quad (3.6)$$

In other words, the martingale transport  $\pi^{I^{\text{B}}} \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$  of (3.3) above is the stretched Brownian motion from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ .

*Proof of Lemma 3.1.* We note that (3.4) and (3.5) follow from the definition of stretched Brownian motion and (2.24), respectively. Therefore we only have to show the equality (3.6). By contradiction, suppose there is a measurable set  $\mathcal{B} \subseteq \widehat{\mathcal{K}}$  with  $\kappa^{\text{B}}(\mathcal{B}) > 0$  and a  $\kappa^{\text{B}}$ -measurable function

$$\mathcal{B} \ni I^{\text{B}} \mapsto \bar{\pi}^{I^{\text{B}}} \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$$

such that

$$\int \text{MCov}(\bar{\pi}_x^{\text{SBM}}, \gamma) \mu^{I^{\text{B}}}(dx) > \int \text{MCov}(\pi_x^{I^{\text{B}}}, \gamma) \mu^{I^{\text{B}}}(dx).$$

Now for  $x \in \mathbb{R}^d$  define

$$\hat{\pi}_x^{I^{\text{B}}} := \begin{cases} \bar{\pi}_x^{I^{\text{B}}}, & I^{\text{B}} \in \mathcal{B}, \\ \pi_x^{\text{SBM}}, & I^{\text{B}} \in \widehat{\mathcal{K}} \setminus \mathcal{B} \end{cases}$$

and

$$\hat{\pi}(dx, dy) := \hat{\pi}_x^{I^B}(dy) \mu^{I^B}(dx) \kappa^B(dI^B) \in \text{MT}(\mu, \nu).$$

But then for the primal problem from  $\mu$  to  $\nu$  we have

$$\int \text{MCov}(\hat{\pi}_x, \gamma) \mu(dx) > \int \text{MCov}(\pi_x^{\text{SBM}}, \gamma) \mu(dx),$$

which is a contradiction to the optimality of  $\pi^{\text{SBM}}$ .  $\square$

## 4 THE DECOMPOSITION OF THE DUAL PROBLEM

Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be an optimizing sequence of convex functions. In the sequel, we have to assume that the values  $(\mathcal{D}(\psi_n))_{n \geq 1}$  converge sufficiently fast to the optimal value  $D(\mu, \nu)$ . To this end, we define the *suboptimality functional*

$$\mathcal{S}(\psi) := \mathcal{D}(\psi) - D(\mu, \nu), \quad \psi \in C_q^{\text{aff}}(\mathbb{R}^d), \quad (4.1)$$

measuring the difference of  $\mathcal{D}(\psi)$  from the optimal value  $D(\mu, \nu)$ . In Lemma 4.1 below we will see that the “fast convergence condition”

$$\sum_{n \geq 1} \mathcal{S}(\psi_n) < +\infty \quad (4.2)$$

on  $(\psi_n)_{n \geq 1}$ , which we have already introduced in (1.9) above, implies that the sequence  $(\psi_n)_{n \geq 1}$  is not only optimal for the pair  $(\mu, \nu)$ , but also for  $\kappa^B$ -a.e. pair  $(\mu^{I^B}, \nu^{I^B})$ .

For a fixed set  $I^B \in \widehat{\mathcal{K}}$  we consider, by analogy with (1.24), the *local* dual problem from  $\mu^{I^B}$  to  $\nu^{I^B}$

$$D(\mu^{I^B}, \nu^{I^B}) = \inf_{\substack{\psi \in C_q^{\text{aff}}(\mathbb{R}^d), \\ \psi \text{ convex}}} \mathcal{D}^{I^B}(\psi),$$

where the *local* dual function  $\mathcal{D}^{I^B}(\cdot)$  now is given by

$$\mathcal{D}^{I^B}(\psi) := \int \psi d\nu^{I^B} - \int (\psi^* * \gamma)^* d\mu^{I^B}, \quad \psi \in C_q^{\text{aff}}(\mathbb{R}^d).$$

Similarly as in (4.1), we define the *local* suboptimality functional

$$\mathcal{S}^{I^B}(\psi) := \mathcal{D}^{I^B}(\psi) - D(\mu^{I^B}, \nu^{I^B}), \quad \psi \in C_q^{\text{aff}}(\mathbb{R}^d).$$

**Lemma 4.1** (DECOMPOSITION OF THE DUAL PROBLEM).

(i) For a convex function  $\psi \in C_q^{\text{aff}}(\mathbb{R}^d)$  we have

$$\mathcal{S}(\psi) = \int_{\widehat{\mathcal{K}}} \mathcal{S}^{I^B}(\psi) \kappa^B(dI^B). \quad (4.3)$$

(ii) Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be a sequence of convex functions satisfying (4.2). Then, for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ , we have that  $(\psi_n)_{n \geq 1}$  is an optimizing sequence for the local dual problem from  $\mu^{I^B}$  to  $\nu^{I^B}$ , i.e.

$$D(\mu^{I^B}, \nu^{I^B}) = \lim_{n \rightarrow \infty} \mathcal{D}^{I^B}(\psi_n). \quad (4.4)$$



*Proof.* (i) We fix a convex function  $\psi \in C_q^{\text{aff}}(\mathbb{R}^d)$  and want to show (4.3). From Lemma 3.1 we know that

$$P(\mu, \nu) = \int_{\widehat{\mathcal{K}}} P(\mu^{I^B}, \nu^{I^B}) \kappa^B(dI^B).$$

Therefore, since by Theorem 1.3 we have  $P(\mu, \nu) = D(\mu, \nu)$  and  $P(\mu^{I^B}, \nu^{I^B}) = D(\mu^{I^B}, \nu^{I^B})$ , for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ , it suffices to show that

$$\mathcal{D}(\psi) = \int_{\widehat{\mathcal{K}}} \mathcal{D}^{I^B}(\psi) \kappa^B(dI^B),$$

which is immediate from (2.24) and (3.2).

(ii) Note that for any  $\psi \in C_q^{\text{aff}}(\mathbb{R}^d)$  and for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ , we have  $\mathcal{S}^{I^B}(\psi) \geq 0$ . By (4.3) and (4.2) we thus have

$$\begin{aligned} \sum_{n \geq 1} \mathcal{S}(\psi_n) &= \sum_{n \geq 1} \int_{\widehat{\mathcal{K}}} \mathcal{S}^{I^B}(\psi_n) \kappa^B(dI^B) \\ &= \int_{\widehat{\mathcal{K}}} \left( \sum_{n \geq 1} \mathcal{S}^{I^B}(\psi_n) \right) \kappa^B(dI^B) < +\infty. \end{aligned}$$

Therefore  $\lim_{n \rightarrow \infty} \mathcal{S}^{I^B}(\psi_n) = 0$ , for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ , which gives (4.4).  $\square$

With the help of Lemma 4.1, we obtain the following consequence from Proposition 2.6.

**Proposition 4.2.** *Let  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  be a sequence of convex functions satisfying (4.2). Then the following assertions hold for  $\kappa^B$ -a.e.  $I^B \in \widehat{\mathcal{K}}$ :*

(i)  $\mu^{I^B} \preceq_c \nu^{I^B}$ ,  $\mu^{I^B}(I^B) = 1$ , and

$$\widehat{\text{supp}}(\nu^{I^B}) = C^B, \quad (4.5)$$

where  $\mu^{I^B}$  and  $\nu^{I^B}$  are defined as in Lemma 2.8 and (3.1), respectively.

(ii) There exist representatives  $\bar{\psi}_n^{I^B} \in [\psi_n]$  and there is a lower semicontinuous convex function  $\psi_{\text{lim}}^{I^B}: \mathbb{R}^d \rightarrow [0, +\infty]$  such that

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^{I^B}(y) = \psi_{\text{lim}}^{I^B}(y) < +\infty, \quad y \in I^B, \quad (4.6)$$

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^{I^B}(y) = \psi_{\text{lim}}^{I^B}(y) = +\infty, \quad y \in \mathbb{R}^d \setminus C^B. \quad (4.7)$$

(iii) The function  $\psi_{\text{lim}}^{I^B}$  is a dual optimizer for the pair  $(\mu^{I^B}, \nu^{I^B})$ .

*Proof.* The following statements are valid for  $\kappa^B$ -a.e.  $I^B$ :

By Lemma 2.8, we have  $\mu^{I^B}(I^B) = 1$ . From the definitions of  $\mu^{I^B}$  and  $\nu^{I^B}$  it follows that  $\mu^{I^B} \preceq_c \nu^{I^B}$ . Moreover, by Lemma 4.1,  $(\psi_n)_{n \geq 1}$  is an optimizing sequence for the local dual problem from  $\mu^{I^B}$  to  $\nu^{I^B}$ . Hence we can apply Proposition 2.6 to the pair  $(\mu^{I^B}, \nu^{I^B})$ , which gives (i) – (iii).  $\square$

## 5 THE PROOFS OF THEOREMS 1.5, 1.6, AND COROLLARY 1.7

Combining the results of Sections 2 – 4 allows us to finally provide the proof of our main result, Theorem 1.6.

*Proof of Theorem 1.6:* As always, let  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . Throughout the proof we fix a sequence  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  of convex functions satisfying (4.2) and consider the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  of Definition 2.2, induced by this sequence. In the following, all statements about a set  $I^{\text{B}} \in \mathcal{K}$  have to be understood in a  $\kappa^{\text{B}}$ -a.s. sense. Once again we emphasize that the “fast convergence condition” (4.2) ensures that  $(\psi_n)_{n \geq 1}$  is an optimizing sequence for the local dual problems from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ . Note that the assertions (i) – (iii) of Theorem 1.6 precisely correspond to the assertions (i) – (iii) of Proposition 4.2.

(iv) As  $\psi_{\text{lim}}^{I^{\text{B}}}$  is a dual optimizer for the pair  $(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$ , by Theorem 1.3, there is a Bass martingale  $M^{I^{\text{B}}} = (M_t^{I^{\text{B}}})_{0 \leq t \leq 1}$  from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$ . Moreover, the pair  $(v^{I^{\text{B}}}, \alpha^{I^{\text{B}}})$  defining the Bass martingale  $M^{I^{\text{B}}}$  is given as claimed in (1.16). The Bass martingale is the stretched Brownian motion  $\pi^{I^{\text{B}}}$  from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$  and

$$\text{Law}(M_0^{I^{\text{B}}}, M_1^{I^{\text{B}}}) = \pi^{I^{\text{B}}} \in \text{MT}(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}}),$$

where the martingale transport  $\pi^{I^{\text{B}}}$  is defined as in (1.19).

(v) The existence of the Bass martingale  $M^{I^{\text{B}}}$  from  $\mu^{I^{\text{B}}}$  to  $\nu^{I^{\text{B}}}$  implies the irreducibility of the pair  $(\mu^{I^{\text{B}}}, \nu^{I^{\text{B}}})$ .  $\square$

Combining the results of [BBST23] with Theorem 1.6, we obtain Theorem 1.5.

*Proof of Theorem 1.5.* (1) Since the pair  $(\mu, \nu)$  is irreducible, by Theorem 1.3 there is a dual optimizer  $\psi_{\text{opt}}$  with  $\mu(\text{ri}(\text{dom } \psi_{\text{opt}})) = 1$ , inducing a Bass martingale  $(M_t)_{0 \leq t \leq 1}$  from  $\mu$  to  $\nu$ . More precisely, the pair  $(v, \alpha)$  defining the Bass martingale is given by  $v = \psi_{\text{opt}}^*$  and  $\alpha = \nabla(\psi_{\text{opt}}^* * \gamma)^*(\mu)$ . In particular,

$$(\nabla \psi_{\text{opt}}^* * \gamma)(\alpha) = \mu \quad \text{and} \quad \nabla \psi_{\text{opt}}^*(\alpha * \gamma) = \nu.$$

As  $I = \text{ri}(\widehat{\text{supp}}(\nu))$ , this implies that  $\text{ri}(\text{dom } \psi_{\text{opt}}) = I$  and thus  $\mu(I) = 1$ .

Applying Theorem 1.6, we conclude that  $C = \widehat{\text{supp}}(\nu) = C^{\text{B}}(x)$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , and obtain the assertions (i) – (iv) of Theorem 1.6. Moreover, we see that  $\psi_{\text{opt}}$  is equal to  $\psi_{\text{lim}}$ , modulo adding an affine function.

(2) Without loss of generality we assume that the optimizing sequence  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  consists of convex functions  $\psi_n$  which are nonnegative. By assumption, the relative interior  $I$  of  $C = \widehat{\text{supp}}(\nu)$  has full  $\mu$ -measure. Furthermore, the sequence  $(\psi_n(y))_{n \geq 1}$  is assumed to be bounded, for all  $y \in I$ . Then, by [BBST23, Proposition 7.20] and [BBST23, Corollary 7.21], the pair  $(\mu, \nu)$  is irreducible, so that again the assertions (ii) – (iv) hold.  $\square$

*Proof of Corollary 1.7:* As in the proof of Theorem 1.6, we fix a sequence  $(\psi_n)_{n \geq 1} \subseteq C_q^{\text{aff}}(\mathbb{R}^d)$  of convex functions satisfying (4.2) and consider the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  of Definition 2.2, induced by this sequence. By Theorem 1.6, the pair  $(\mu^{I^{\text{B}}(x)}, \nu^{I^{\text{B}}(x)})$  is irreducible, for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ . Thus, by [BBST23, Theorem D.1], the probability measure  $\pi_x^{\text{SBM}}$  is equivalent to  $\nu^{I^{\text{B}}(x)}$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ . In particular, we have

$$\text{supp}(\pi_x^{\text{SBM}}) = \text{supp}(\nu^{I^{\text{B}}(x)}),$$

for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ . On the other hand, by (i) of Theorem 1.6, we have that

$$C^{\text{B}}(x) = \widehat{\text{supp}}(\nu^{I^{\text{B}}(x)}),$$

for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ . We conclude that  $\widehat{\text{supp}}(\pi_x^{\text{SBM}}) = C^{\text{B}}(x)$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , i.e. the equality in (1.17).

In order to show the inclusion in (1.17), we let  $\pi \in \text{MT}(\mu, \nu)$  be an arbitrary martingale transport. By contradiction, we assume that the set

$$A := \left\{ x \in \mathbb{R}^d : \widehat{\text{supp}}(\pi_x) \cap (\mathbb{R}^d \setminus C^{\text{B}}(x)) \neq \emptyset \right\}$$

has strictly positive  $\mu$ -measure. By (ii) of Theorem 1.5, for  $\mu$ -a.e.  $x \in A$ , we can select representatives  $\bar{\psi}_n^x \in [\psi_n]$  and there is a lower semicontinuous convex function  $\psi_{\text{lim}}^x: \mathbb{R}^d \rightarrow [0, +\infty]$ , such that

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^x(y) = \psi_{\text{lim}}^x(y) < +\infty, \quad y \in I^{\text{B}}(x), \quad (5.1)$$

$$\lim_{n \rightarrow \infty} \bar{\psi}_n^x(y) = \psi_{\text{lim}}^x(y) = +\infty, \quad y \in \mathbb{R}^d \setminus C^{\text{B}}(x). \quad (5.2)$$

This selection can be done in a  $\mu$ -measurable way. In particular, by (5.1) and since  $x \in I^{\text{B}}(x)$ , the sequence  $(\bar{\psi}_n^x(x))_{n \geq 1}$  is bounded, for  $\mu$ -a.e.  $x \in A$ . Next, we choose  $\tilde{A} \subseteq A$  with  $\mu(\tilde{A}) > 0$  such that the sequence  $(\bar{\psi}_n^x(x))_{n \geq 1}$  is uniformly bounded by some  $M > 0$ , for all  $x \in \tilde{A}$ . Since

$$\int (\bar{\psi}_n^x(y) - \bar{\psi}_n^x(x)) \pi_x(dy) \geq 0,$$

for  $\mu$ -a.e.  $x \in \mathbb{R}^d$ , we conclude that

$$\int \int (\bar{\psi}_n^x(y) - \bar{\psi}_n^x(x)) \pi_x(dy) \mu(dx) \geq \int_{x \in \tilde{A}} \int_{y \in \mathbb{R}^d} \bar{\psi}_n^x(y) \pi_x(dy) \mu(dx) - M.$$

Similarly, as  $\bar{\psi}_n^x(y) \geq 0$ , we have that

$$\int_{x \in \tilde{A}} \int_{y \in \mathbb{R}^d} \bar{\psi}_n^x(y) \pi_x(dy) \mu(dx) \geq \int_{x \in \tilde{A}} \int_{y \in \mathbb{R}^d \setminus C^{\text{B}}(x)} \bar{\psi}_n^x(y) \pi_x(dy) \mu(dx).$$

By Fatou's lemma and using (5.2), we obtain that

$$\liminf_{n \rightarrow \infty} \int_{x \in \tilde{A}} \int_{y \in \mathbb{R}^d \setminus C^{\text{B}}(x)} \bar{\psi}_n^x(y) \pi_x(dy) \mu(dx) = +\infty.$$

Altogether, we have

$$\sup_{n \geq 1} \int \int (\bar{\psi}_n^x(y) - \bar{\psi}_n^x(x)) \pi_x(dy) \mu(dx) = +\infty,$$

which is a contradiction to [BBST23, Lemma 7.9].  $\square$

## 6 EXAMPLES

We start with a simple example motivating the treatment of the general case in Theorem 1.6 when the Bass paving  $\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d}$  of the pair  $(\mu, \nu)$  is not reduced to one element so that there does not exist a Bass martingale from  $\mu$  to  $\nu$ .

**Example 6.1.** We first spell out the case of a geometric Brownian motion on  $\mathbb{R}_+ = (0, +\infty)$ , which is an example of a Bass martingale (recall Definition 1.1).

Let  $Z$  be a standard Gaussian random variable on  $\mathbb{R}$  and define

$$\mu_+ := \delta_1, \quad \nu_+ := \text{Law}(\exp(Z - \frac{1}{2})),$$

so that  $\mu_+ \preceq_c \nu_+$ . It is straightforward to calculate all the functions defining the Bass martingale from  $\mu_+$  to  $\nu_+$ . In particular, we have

$$v_+(z) = \frac{d}{dz} v_+(z) = \exp(z - \frac{1}{2}), \quad z \in \mathbb{R},$$

with the property that  $\text{Law}(v_+(Z)) = \nu_+$ . The convex conjugate of  $v_+$  and its derivative are equal to

$$\psi_+(y) = y(\log y - \frac{1}{2}), \quad \frac{d}{dy} \psi_+(y) = \log y + \frac{1}{2}, \quad y > 0.$$

Summing up,  $\pi_+ = \delta_1 \otimes \nu_+$  defines a Bass martingale from  $\mu_+$  to  $\nu_+$ , for which we can explicitly compute all its ingredients.

Next, we mirror this example along the vertical axis, i.e.

$$\mu_- := \delta_{-1}, \quad \nu_- := \text{Law} \left( -\exp\left(Z - \frac{1}{2}\right) \right).$$

Again, we can explicitly calculate the relevant quantities of the Bass martingale joining  $\mu_-$  to  $\nu_-$  and obtain the mirrored functions on  $\mathbb{R}_- = (-\infty, 0)$ . In particular,

$$\psi_-(y) = -y(\log(-y) - \frac{1}{2}), \quad y < 0.$$

Finally, we define the convex combination of these two examples, i.e.

$$\mu := \frac{1}{2}(\mu_+ + \mu_-), \quad \nu := \frac{1}{2}(\nu_+ + \nu_-).$$

It is still possible to pin down the primal optimizer  $\pi^{\text{SBM}}$  for the pair  $(\mu, \nu)$ , namely

$$\pi^{\text{SBM}} = \frac{1}{2}(\delta_1 \otimes \nu_- + \delta_{-1} \otimes \nu_+).$$

However, we do not have a Bass martingale any more but only a stretched Brownian motion from  $\mu$  to  $\nu$ , since we cannot find a function inducing the transport  $\pi^{\text{SBM}}$ . Intuitively speaking, the stretched Brownian motion consists of the above two Bass martingales which “do not talk to each other”. The Bass paving is given by

$$\{I^{\text{B}}(x)\}_{x \in \mathbb{R}^d} = \{(-\infty, 0), \{0\}, (0, +\infty)\},$$

with  $\kappa^{\text{B}}((-\infty, 0)) = \kappa^{\text{B}}((0, +\infty)) = \frac{1}{2}$  and  $\kappa^{\text{B}}(\{0\}) = 0$ . We remark that in the one-dimensional case such invariant decompositions were already studied in [BJ16].

Regarding the dual side of optimization, we know from (1.24) that there is an optimizing sequence  $(\psi_n)_{n \geq 1} \subseteq C_{\text{q}}^{\text{aff}}(\mathbb{R})$  of convex functions for the dual problem (1.4). On  $\mathbb{R}_+$  the functions  $(\psi_n)_{n \geq 1}$  should approximate the dual optimizer  $\psi_+$  for  $(\mu_+, \nu_+)$ , while on  $\mathbb{R}_-$  they should approximate the dual optimizer  $\psi_-$  for  $(\mu_-, \nu_-)$ . A naive guess would be to choose (or approximate) the function

$$\psi(y) := \begin{cases} \psi_+(y), & \text{for } y \geq 0, \\ \psi_-(y), & \text{for } y \leq 0. \end{cases} \quad (6.1)$$

Two problems arise when one tries to approximate  $\psi$  as defined in (6.1) by convex functions  $\psi_n$  in  $C_{\text{q}}^{\text{aff}}(\mathbb{R})$ : firstly at 0 (convexity) and secondly at  $\pm\infty$  (boundedness). The boundedness requirement for  $\psi_n \in C_{\text{q}}^{\text{aff}}(\mathbb{R})$  is easily taken care of and we ignore this issue. The crucial point is that, by trying to paste  $\psi_+$  and  $\psi_-$  together at 0, we grossly violate the convexity property, as  $\psi'(0_+) = -\infty$  while  $\psi'(0_-) = +\infty$ .

Here is the remedy to this problem: let  $\psi_{n,+}$  and  $\psi_{n,-}$  be convex approximations of  $\psi_+$  and  $\psi_-$ , respectively, which both have a finite derivative at  $y = 0$  (which clearly is possible). Defining

$$\tilde{\psi}_n(y) := \begin{cases} \psi_{n,+}(y), & \text{for } y \geq 0, \\ \psi_{n,-}(y), & \text{for } y \leq 0, \end{cases}$$

we still have not found a sequence of convex functions approximating  $\psi_+$  and  $\psi_-$  on  $\mathbb{R}_+$  and  $\mathbb{R}_-$ , respectively. But for large enough  $C_n$ , the function

$$\psi_n(y) := \tilde{\psi}_n(y) + C_n|y|$$

becomes convex. While  $\psi_n$ , of course, neither approximates  $\psi_+$  on  $\mathbb{R}_+$  nor  $\psi_-$  on  $\mathbb{R}_-$ , recall that the dual functions should be considered “modulo adding affine functions”. Passing to

$$\psi_n(y) - C_n y, \quad y \geq 0$$

we get a good approximation of  $\psi_+$  on  $\mathbb{R}_+$ , and passing to

$$\psi_n(y) + C_n y, \quad y \leq 0$$

we get a good approximation of  $\psi_-$  on  $\mathbb{R}_-$ . Hence “locally”, i.e. on  $\mathbb{R}_+$  and  $\mathbb{R}_-$ , the sequence  $(\psi_n)_{n \geq 1}$  converges — after adding proper affine functions — to  $\psi_+$  and  $\psi_-$ , respectively. But, of course, the sequence  $(\psi_n)_{n \geq 1}$  itself does not converge on  $\mathbb{R}$  and it should come as no surprise that there is no “global” optimizer  $\psi$  defined on  $\mathbb{R}$ .  $\diamond$

**Example 6.2.** We thank Mathias Beiglböck and Krzysztof Ciosmak for kindly suggesting this example to us, the former for the case  $R > \frac{3}{2}$ , the latter for the case  $R = \frac{3}{2}$ .

We denote by  $\mathbf{S}_r$  the sphere in  $\mathbb{R}^2$  with radius  $r$ , i.e.

$$\mathbf{S}_r := \{x \in \mathbb{R}^2 : |x| = r\}.$$

We let  $\mu$  be the uniform distribution on  $\mathbf{S}_1$  and define the probability measure

$$\nu^{(R)} := \frac{1}{2}(\nu_{\frac{1}{2}} + \nu_R),$$

where  $\nu_{\frac{1}{2}}$  and  $\nu_R$  are uniform distributions on  $\mathbf{S}_{\frac{1}{2}}$  and  $\mathbf{S}_R$ , respectively. It is straightforward to check — and will follow from the discussion below — that  $\nu^{(R)}$  dominates  $\mu$  in convex order if and only if  $R \geq \frac{3}{2}$ .

$R = \frac{3}{2}$  Let us first examine the case  $R = \frac{3}{2}$ . For  $x \in \mathbf{S}_1$ , we define the probability measure

$$\pi_x^{(\frac{3}{2})} := \frac{1}{2}(\delta_x + \delta_{\frac{3x}{2}})$$

and claim that the unique martingale transport from  $\mu$  to  $\nu^{(\frac{3}{2})}$  is given by

$$\pi^{(\frac{3}{2})}(dx, dy) := \pi_x^{(\frac{3}{2})}(dy) \mu(dx) \in \text{MT}(\mu, \nu^{(\frac{3}{2})}).$$

Clearly,  $\pi^{(\frac{3}{2})}$  is a martingale transport between  $\mu$  and  $\nu^{(\frac{3}{2})}$ . To see that  $\pi^{(\frac{3}{2})}$  is in fact the unique element of  $\text{MT}(\mu, \nu^{(\frac{3}{2})})$ , we consider the convex function  $\psi(x) = |x|$  and note that

$$\int \psi d\mu = \int \psi d\nu^{(\frac{3}{2})} = 1.$$

For any candidate martingale transport  $\pi(dx, dy) = \pi_x(dy) \mu(dx) \in \text{MT}(\mu, \nu^{(\frac{3}{2})})$  we thus obtain from Jensen’s inequality that

$$\int \psi d\pi_x(dy) = \psi(x),$$

which implies that  $\pi_x = \pi_x^{(\frac{3}{2})}$ , for  $\mu$ -a.e.  $x \in \mathbb{R}^2$ .

In particular, the pair  $(\mu, \nu^{(\frac{3}{2})})$  is not irreducible. In fact, identifying  $\mathbb{R}^2$  with  $\mathbb{C}$ , the Bass paving consists ( $\kappa^{\text{B}}$ -almost surely) of the continuum many relatively open line segments

$$I_\varphi := I^{\text{B}}(e^{i\varphi}) = \{re^{i\varphi} : r \in (\frac{1}{2}, \frac{3}{2})\}, \quad \varphi \in [0, 2\pi),$$

as  $\pi^{(\frac{3}{2})}$  transports  $I_\varphi$  into  $C_\varphi := \overline{I_\varphi}$ .

$R > \frac{3}{2}$  We next pass to the case  $R > \frac{3}{2}$ , which behaves quite differently. For  $\alpha \in [-\frac{\pi}{6}, \frac{\pi}{6}]$ , the line

$$t \mapsto \ell(t) := (1, 0) + t(1, \vartheta), \quad (6.2)$$

where  $\vartheta = \tan \alpha$ , intersects the sphere  $S_{\frac{1}{2}}$  at a point  $x^{(\frac{1}{2})}$ , as indicated in Figure 1 below. For  $\alpha \in (-\frac{\pi}{6}, \frac{\pi}{6})$ , there are two intersection points and we choose  $x^{(\frac{1}{2})}$  as the one which is closer to  $(1, 0)$ . Furthermore, we denote by  $x^{(R)}$  the intersection point of the line  $t \mapsto \ell(t)$  with the sphere  $S_R$ , where we choose  $x^{(R)}$  as the intersection point to the right of  $(1, 0)$ . Elementary geometry reveals that, for  $\alpha \in (0, \frac{\pi}{6}]$ , there is a unique  $R = R(\alpha) > \frac{3}{2}$  such that

$$|(1, 0) - x^{(\frac{1}{2})}| = |(1, 0) - x^{(R)}|.$$

In other words,  $R$  is chosen such that  $(1, 0)$  is the midpoint between  $x^{(\frac{1}{2})}$  and  $x^{(R)}$ .

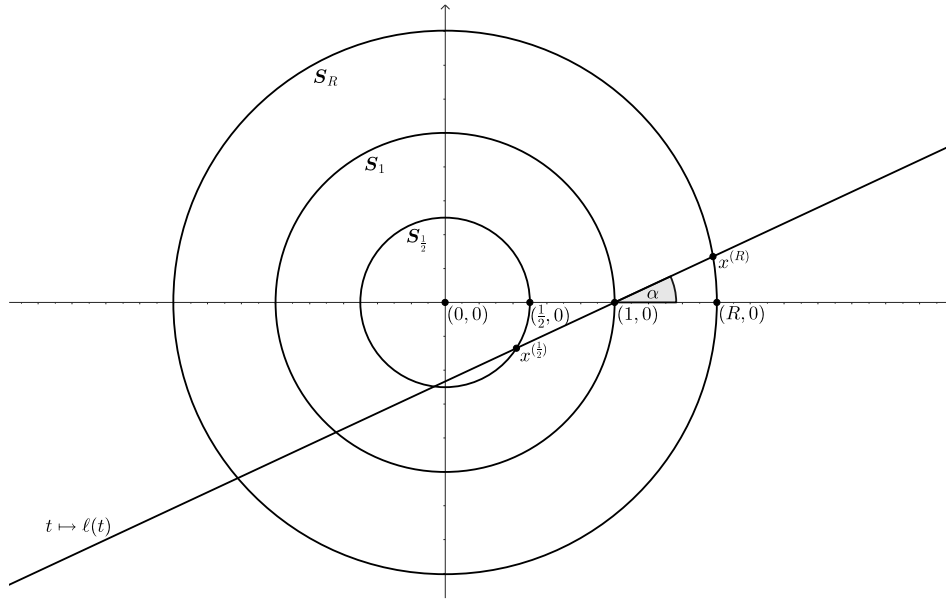


Figure 1: Illustration of Example 6.2 in the case  $\alpha = 25^\circ$  and  $R \approx \frac{8}{5}$ .

After these preparations, we turn to the martingale transports in  $\text{MT}(\mu, \nu^{(R)})$ , for  $R = R(\alpha) > \frac{3}{2}$  as above. Figure 1 indicates a martingale transport analogous to the one constructed above in the case  $R = \frac{3}{2}$ . Identifying again  $\mathbb{R}^2$  with  $\mathbb{C}$ , we define

$$\pi_{e^{i\varphi}}^{(R)} := \frac{1}{2}(\delta_{e^{i\varphi}x^{(\frac{1}{2})}} + \delta_{e^{i\varphi}x^{(R)}}), \quad \varphi \in [0, 2\pi).$$

As in the case  $R = \frac{3}{2}$  it follows from rotational invariance that

$$\pi^{(R)}(dx, dy) := \pi_x^{(R)}(dy) \mu(dx) \in \text{MT}(\mu, \nu^{(R)})$$

defines a martingale transport from  $\mu$  to  $\nu^{(R)}$ . What about the invariant cells attached to the martingale transport  $\pi^{(R)}$ ? Again as in the case  $R = \frac{3}{2}$ , there are continuum many relatively open line segments

$$I_\varphi := I^{\text{B}}(e^{i\varphi}) = (e^{i\varphi}x^{(\frac{1}{2})}, e^{i\varphi}x^{(R)}), \quad \varphi \in [0, 2\pi), \quad (6.3)$$

such that  $\pi^{(R)}$  leaves  $C_\varphi := \overline{I_\varphi}$  invariant.

But now the transport  $\pi^{(R)}$  is not unique any more. A striking way to see this is to replace  $\alpha > 0$  by  $-\alpha$ , i.e. mirroring the line  $t \mapsto \ell(t)$  defined in (6.2) along the  $x$ -axis. In this way we obtain another martingale transport, denoted by  $\overline{\pi}^{(R)}$ , which is different from  $\pi^{(R)}$ . In particular, the invariant cells of  $\overline{\pi}^{(R)}$  are obtained by “flipping” the cells in (6.3). This phenomenon is in sharp contrast to the one-dimensional setting  $d = 1$ , analyzed in [BJ16], where it was shown that the paving of  $\mathbb{R}$  into invariant cells only depends on the pair  $(\mu, \nu)$ , but not on the special choice of  $\pi \in \text{MT}(\mu, \nu)$ . In the case  $d \geq 2$ , we refer to [GKL19] for an analysis of the invariant pavings of  $\mathbb{R}^d$  which do depend on the choice of  $\pi \in \text{MT}(\mu, \nu)$ , for  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ . In particular, our example in dimension  $d = 2$  shows that the invariant cells of a martingale transport  $\pi \in \text{MT}(\mu, \nu^{(R)})$  may depend on the choice of  $\pi$ .

The above discussion motivates the approach of [DMT19] and [OS17], which focused on the *universal* partition of  $\mathbb{R}^d$ , induced by  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\mu \preceq_c \nu$ , as defined in [DMT19, Theorem 2.1] and revisited in Corollary 1.7 above. It turns out that in the above example of the pair  $(\mu, \nu^{(R)})$ , the De March–Touzi paving is in fact trivial, i.e. the pair  $(\mu, \nu^{(R)})$  is irreducible. More precisely, we will see that there is a Bass martingale  $M = (M_t)_{0 \leq t \leq 1}$  from  $\mu$  to  $\nu^{(R)}$ , for which we have that the law of  $M_t$  is equivalent to Lebesgue measure on the closed disk with radius  $R$  in  $\mathbb{R}^2$ , for  $0 < t < 1$ , while, of course,  $\text{Law}(M_0) = \mu$  and  $\text{Law}(M_1) = \nu^{(R)}$ .

We now consider the dual optimization problem for  $(\mu, \nu^{(R)})$  in the case  $R > \frac{3}{2}$ , which can be explicitly solved. Let  $\alpha^{(\rho)}$  denote the uniform distribution on the sphere  $\mathcal{S}_\rho$  with radius  $\rho > 0$  to be specified later. Denoting by  $\gamma$  the standard Gaussian measure on  $\mathbb{R}^2$ , we consider the convolution measure  $\alpha^{(\rho)} * \gamma$  and define the Brenier map  $\nabla v^{(\rho, R)}$  from  $\alpha^{(\rho)} * \gamma$  to  $\nu^{(R)}$ . Here  $v^{(\rho, R)}$  is a convex function on  $\mathbb{R}^2$ , unique up to an additive constant, for which the derivative  $\nabla v^{(\rho, R)}$  exists Lebesgue-almost everywhere and therefore also  $(\alpha^{(\rho)} * \gamma)$ -almost surely. As  $\nu^{(R)}$  is supported by  $\mathcal{S}_{\frac{1}{2}}$  and  $\mathcal{S}_R$ , we have that  $\nabla v^{(\rho, R)}$  takes its values almost surely in  $\mathcal{S}_{\frac{1}{2}} \cup \mathcal{S}_R$ .

It is intuitively rather obvious that on each radial ray of  $\mathbb{R}^2$ , say on  $\{\lambda(1, 0) : \lambda > 0\}$ , the directional derivative along this ray has to take either the value  $\frac{1}{2}$  or  $R$ , each with conditional probability  $\frac{1}{2}$  under the measure  $\alpha^{(\rho)} * \gamma$ , induced on this ray. Restricting  $v^{(\rho, R)}$  to the  $x$ -axis, we obtain the following picture.

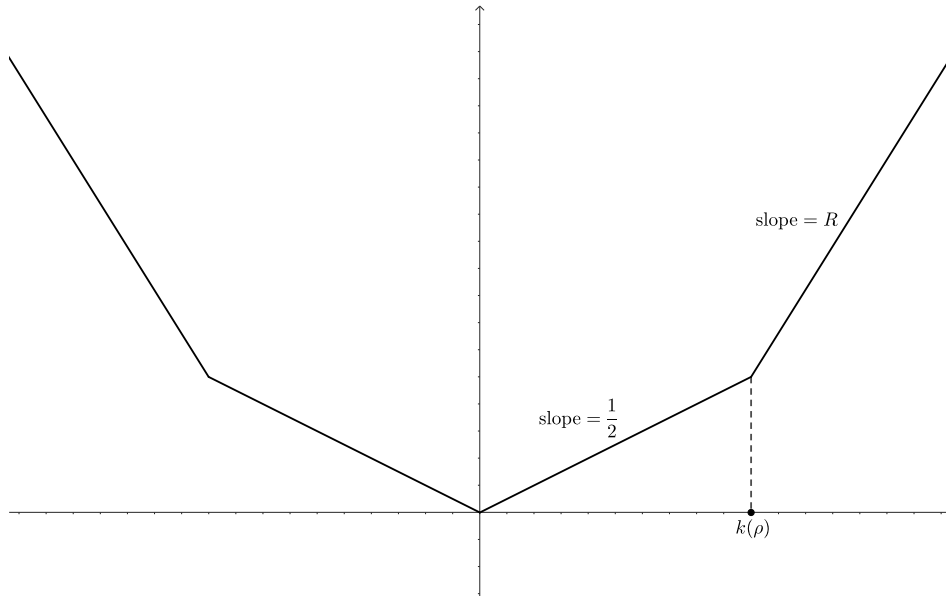


Figure 2: Restriction of the convex function  $v^{(\rho, R)}$  to the  $x$ -axis, with  $R \approx \frac{8}{5}$ .

The function in Figure 2 has a kink at some point  $k(\rho)$ , which can be determined as the median of the conditional law of  $\alpha^{(\rho)} * \gamma$ , induced on the ray  $\{\lambda(1, 0) : \lambda > 0\}$ . Finally we determine, for  $\rho > 0$ , the number  $R = R(\rho) > \frac{3}{2}$  as the unique number such that the rotationally invariant function  $\nabla v^{(\rho, R)} * \gamma$  maps the measure  $\delta_{\rho(0,1)} * \gamma$  on  $\mathbb{R}^2$  to the measure

$$(\nabla v^{(\rho, R)} * \gamma)(\delta_{\rho(0,1)} * \gamma)$$

with barycenter  $(0, 1)$ . We know from the general theory that there must be a unique such  $R > \frac{3}{2}$  for any given  $\rho > 0$ . It also follows from the general theory that, for  $0 < t < 1$ , the law of  $M_t$  is equivalent to Lebesgue measure on the closed disk with radius  $R$  in  $\mathbb{R}^2$ , i.e.

$$\{x \in \mathbb{R}^2 : |x| \leq R\}.$$

Indeed, by [BBST23, Remark 6.3, (iii)] the law of  $M_t$  is the image of  $\alpha * \gamma^t$  under the map  $\nabla v^{(\rho, R)} * \gamma^{1-t}$ . While  $\alpha * \gamma^t$  is equivalent to Lebesgue measure on  $\mathbb{R}^2$ , for  $0 < t \leq 1$ , the range of  $\nabla v^{(\rho, R)} * \gamma^{1-t}$  is the open ball of radius  $R$  on  $\mathbb{R}^2$ , for  $0 \leq t < 1$ . In conclusion, the support of the law of  $M_t$  is the closed disk of radius  $R$  in  $\mathbb{R}^2$ , for  $0 < t < 1$ .

On a last note, we discuss the limiting behaviour of the function  $\rho \mapsto R(\rho)$ , for  $\rho > 0$ . It turns out that  $\lim_{\rho \rightarrow +\infty} R(\rho) = \frac{3}{2}$  while  $\lim_{\rho \rightarrow 0} R(\rho) = +\infty$ .  $\diamond$



## REFERENCES

- [AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lect. Math. ETH Zürich. Birkhäuser, Basel, second edition, 2008.
- [Bas83] R.F. Bass. Skorokhod imbedding via stochastic integrals. In J. Azéma and M. Yor, editors, *Sémin. Probab. XVII 1981/82 — Proceedings*, volume **986** of *Lecture Notes in Math.*, pages 221–224. Springer, Berlin, Heidelberg, 1983.
- [BBHK20] J. Backhoff-Veraguas, M. Beiglböck, M. Huesmann, and S. Källblad. Martingale Benamou–Brenier: A probabilistic perspective. *Ann. Probab.*, **48**(5):2258–2289, 2020.
- [BBST23] J. Backhoff-Veraguas, M. Beiglböck, W. Schachermayer, and B. Tschiderer. The structure of martingale Benamou–Brenier in  $\mathbb{R}^d$ . [arXiv:2306.11019](#), 2023.
- [Bee91] G. Beer. A Polish Topology for the Closed Subsets of a Polish Space. *Proc. Amer. Math. Soc.*, **113**(4):1123–1133, 1991.
- [BJ16] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, **44**(1):42–106, 2016.
- [BNS22] M. Beiglböck, M. Nutz, and F. Stebegg. Fine properties of the optimal Skorokhod embedding problem. *J. Eur. Math. Soc.*, **24**(4):1389–1429, 2022.
- [BNT17] M. Beiglböck, M. Nutz, and N. Touzi. Complete duality for martingale optimal transport on the line. *Ann. Probab.*, **45**(5):3038–3074, 2017.
- [Cio23a] K.J. Ciosmak. General localisation scheme I: theory. [arXiv:2312.12281](#), 2023.
- [Cio23b] K.J. Ciosmak. General localisation scheme II: applications. [arXiv:2312.13167](#), 2023.
- [DM78] C. Dellacherie and P.-A. Meyer. *Probabilities and Potential*, volume **29** of *North-Holland Mathematics Studies*. North-Holland Publishing Company, Amsterdam, 1978.
- [DMT19] H. De March and N. Touzi. Irreducible convex paving for decomposition of multidimensional martingale transport plans. *Ann. Probab.*, **47**(3):1726–1774, 2019.
- [GKL19] N. Ghoussoub, Y.-H. Kim, and T. Lim. Structure of optimal martingale transport plans in general dimensions. *Ann. Probab.*, **47**(1):109–164, 2019.
- [OS17] J. Oblój and P. Siorpaes. Structure of martingale transports in finite dimensions. [arXiv:1702.08433](#), 2017.
- [Roc70] R.T. Rockafellar. *Convex analysis*, volume **28** of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1970.
- [Str65] V. Strassen. The Existence of Probability Measures with Given Marginals. *Ann. Math. Statist.*, **36**(2):423–439, 1965.