

# DUAL ATTAINMENT FOR THE MARTINGALE TRANSPORT PROBLEM

MATHIAS BEIGLBÖCK, TONGSEOK LIM, AND JAN OBLÓJ

ABSTRACT. We investigate existence of dual optimizers in one-dimensional martingale optimal transport problems. While [5] established such existence for weak (quasi-sure) duality, [2] showed existence for the natural stronger (pointwise) duality may fail even in regular cases. We establish that (pointwise) dual maximizers exist when  $y \mapsto c(x, y)$  is convex, or equivalent to a convex function. It follows that when marginals are compactly supported, the existence holds when the cost  $c(x, y)$  is twice continuously differentiable in  $y$ . Further, this may not be improved as we give examples with  $c(x, \cdot) \in C^{2-\varepsilon}$ ,  $\varepsilon > 0$ , where dual attainment fails. Finally, when measures are compactly supported, we show that dual optimizers are Lipschitz if  $c$  is Lipschitz.

*Keywords:* martingale optimal transport, Kantorovitch duality, dual attainment, robust mathematical finance.

## 1. INTRODUCTION

In recent years, there has been a significant interest in optimal transport problems where the transport plan is constrained to be a martingale. Referred to as martingale optimal transport (MOT), they were introduced by [2, 8] to study the mathematical finance question of computing model-independent no-arbitrage price bounds, see [11] for a survey, and have been studied in many papers since, e.g. [13, 7, 6, 14]. They are however of much wider mathematical interest. Mirroring classical optimal transport, they have important consequences for the study of martingale inequalities, see e.g. [4, 10, 16]. In continuous time, they are intimately linked with the Skorokhod embedding problem, see [15] for an overview of the latter, and have already led to new contributions to this well established field, see [1].

Most papers on MOT either study the structure and geometry of optimisers or investigate a form of general Kantorovitch duality. Duality is of particular importance for mathematical finance: the primal problem corresponds to option pricing while the dual offers robust hedging strategies. However the latter poses a challenge: as already shown in [2], the dual problem in MOT does not admit an optimiser in general. One way to recover the dual attainment is relaxing the duality and considering not pointwise but weaker, quasi-sure, inequalities, as shown in [5]. Our aim here instead is to identify sufficient conditions on the problem under which a suitably nice dual optimiser exists. This has immediate applications in robust

---

The project has been generously supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 335421. The first and the second author gratefully acknowledge support from the Austrian Science Foundation (FWF) through grant Y782 and the third author also acknowledge the support of St John's College Oxford.

mathematical finance, where pointwise inequalities are more natural. Equally importantly, we believe, this problem is of intrinsic mathematical interest. In fact, answering such questions is an important prerequisite for the future development of the field and understanding geometry of primal optimisers, or existence of Brenier-type MOT plans in multiple dimensions. So far results in this direction are limited to dimension 1 and 2 [13, 12, 3] and more recently [9]. However, the methods and results of [9] would allow to provide a satisfactory answer to this central question, conditionally on the existence of dual maximisers.

To present in more detail the questions we want to study, we need to introduce some notation. Let  $\Omega := \mathbb{R} \times \mathbb{R}$  be the canonical space and  $(X, Y)$  the canonical process, i.e.  $X(x, y) = x$  and  $Y(x, y) = y$  for all  $(x, y) \in \Omega$ . We also denote by  $P_{\mathbb{R}}$  and  $P_{\Omega}$  the collections of all probability measures on  $\mathbb{R}$  and  $\Omega$ , respectively. For fixed  $\mu, \nu \in P_{\mathbb{R}}$  with finite first moments, we consider the following subsets of  $P_{\Omega}$

$$\Pi(\mu, \nu) := \{\mathbb{P} \in P_{\Omega} : X \sim_{\mathbb{P}} \mu, Y \sim_{\mathbb{P}} \nu\}, \quad (1.1)$$

$$\text{MT}(\mu, \nu) := \{\mathbb{P} \in \Pi(\mu, \nu) : \mathbb{E}^{\mathbb{P}}[Y|X] = X, \mathbb{P} - a.s.\}. \quad (1.2)$$

The set  $\Pi(\mu, \nu)$  is non-empty as it contains the product measure  $\mu \otimes \nu$ . By a classical result of Strassen [17],  $\text{MT}(\mu, \nu)$  is non-empty if and only if  $\mu \preceq \nu$  in convex order:

$$\mu(\xi) \leq \nu(\xi) \text{ for all convex function } \xi, \text{ where } \mu(\xi) := \int \xi(x)\mu(dx). \quad (1.3)$$

Throughout we assume that  $c : \Omega \rightarrow \mathbb{R}$  is a lower-semicontinuous cost function with  $c(x, y) \leq a(x) + b(y)$  for some  $a \in L^1(\mu)$  and  $b \in L^1(\nu)$ . Then  $\mathbb{E}^{\mathbb{P}}[c(X, Y)]$  is a well-defined scalar in  $\mathbb{R} \cup \{-\infty\}$ . The martingale optimal transport problem, as introduced in [2] in the present discrete-time case and in [8] in continuous time, is defined by the following primal problem:

$$\mathbf{P} := \mathbf{P}(c) := \inf_{\mathbb{P} \in \text{MT}(\mu, \nu)} \mathbb{E}^{\mathbb{P}}[c(X, Y)]. \quad (1.4)$$

Its dual is given by

$$\mathbf{D} := \mathbf{D}(c) := \sup_{(f, g) \in D_c} \{\nu(g) - \mu(f)\}, \quad (1.5)$$

where

$$D_c := \{(f, g) : f^- \in L^1(\mu), g^+ \in L^1(\nu), \text{ and for some } h \in L^\infty(\mu), \\ g(y) - f(x) - h(x) \cdot (y - x) \leq c(x, y) \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{R}\}.$$

We assume that  $P(c)$  is finite. Lower-semicontinuity<sup>1</sup> implies that the infimum in (1.4) is attained i.e.  $\mathbf{P} = \mathbb{E}^{\mathbb{P}^*}[c(X, Y)]$  for some  $\mathbb{P}^* \in \text{MT}(\mu, \nu)$ , cf. Remark 3.3 below.

In mathematical finance, the cost  $c$  has the interpretation of the payoff of an exotic derivative and  $\mathbf{P}(c)$  gives its lower no-arbitrage price. A triplet  $(f, g, h)$  on the dual side corresponds to a robust sub-hedging strategy for  $c$ : both  $f$  and  $g$  are bought through trading European options and  $h(x)(y - x)$  corresponds to the payoff from buying  $h(x)$  stocks at time zero. The following result establishes the basic duality between the primal and dual problems.

<sup>1</sup>Throughout the paper lower-semicontinuity is only required to guarantee the existence of some  $\mathbb{P}^* \in \text{MT}(\mu, \nu)$  satisfying  $\mathbb{E}^{\mathbb{P}^*}[c(X, Y)] = \mathbf{P} \in \mathbb{R}$ . Hence we might work with a merely measurable cost function and assume existence of a primal minimizer.

**Theorem 1.1** ([2]). *Let  $\mu \preceq \nu \in P_{\mathbb{R}}$ , and assume that  $c$  is bounded from below. Then  $\mathbf{P} = \mathbf{D}$ .*

In the same paper the authors provided a simple example, based on the cost function  $c(x, y) = -|y - x|$ , where the dual problem is not attained. Our aim here is to study fundamental reasons why dual attainment may fail and to provide sufficient conditions for it to hold and for dual optimiser to have further desirable regularity and integrability properties. We note that [5] showed dual attainment may be recovered if we weaken the dual formulation and require inequalities to hold quasi-surely, i.e. almost surely for any  $\mathbb{P} \in \text{MT}(\mu, \nu)$ . However this is not entirely satisfying in view of the financial and other, mentioned above, applications.

## 2. MAIN RESULTS

We start by defining the crucial notion of a solution for the dual problem (1.5).

**Definition 2.1.** *Let  $\mu \preceq \nu$  be in convex order and let  $c(x, y)$  be a cost function. We say that a triple of functions  $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $g : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ ,  $h : \mathbb{R} \rightarrow \mathbb{R}$  is a dual maximizer, or a solution for the dual problem (1.5), if  $f$  is finite  $\mu$ -a.s.,  $g$  is finite  $\nu$ -a.s., and for any minimizer  $\mathbb{P}^* \in \text{MT}(\mu, \nu)$  for the martingale optimal transport problem (1.4), the following holds:*

$$g(y) - f(x) - h(x) \cdot (y - x) \leq c(x, y) \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad (2.1)$$

$$g(y) - f(x) - h(x) \cdot (y - x) = c(x, y) \quad \mathbb{P}^* \text{-a.a. } (x, y). \quad (2.2)$$

In fact, if (2.1), (2.2) hold for some  $\mathbb{P} \in \text{MT}(\mu, \nu)$ , then  $\mathbb{P}$  is a minimizer of the martingale transport problem (1.4) and moreover (2.1), (2.2) hold for all minimizers of (1.4), c.f. [5, Corollary 7.8].

A simple but important observation is that if  $(f, g, h)$  is a dual maximizer, then  $g$  can always be replaced by a “better” candidate  $\tilde{g}$  induced by  $(f, h)$ , as follows:

$$\tilde{g}(y) := \inf_{x \in \mathbb{R}} (f(x) + h(x) \cdot (y - x) + c(x, y)). \quad (2.3)$$

Observe that then  $g \leq \tilde{g}$  while (2.1)–(2.2) still holds with  $\tilde{g}$ . The minus signs on  $f$  and  $h$  in (2.1) and (2.2) were chosen to define  $\tilde{g}$  by (2.3). In this paper, unless stated otherwise we will always assume that  $g = \tilde{g}$ . Now we state our main theorem.

**Theorem 2.2.** *Let  $\mu \preceq \nu$  be in convex order. Suppose that  $c(x, y)$  is semiconvex in  $y$   $\mu$ -uniformly in  $x$ , in the following sense: there exists a Borel function  $u : J \rightarrow \mathbb{R}$  where  $J := \text{conv}(\text{supp}(\nu))$  such that*

$$\text{for } \mu\text{-a.e. } x, y \mapsto c(x, y) + u(y) \text{ is continuous and convex on } J. \quad (2.4)$$

*If  $\nu$  is not compactly supported, then further suppose that  $y \mapsto c(x, y) + u(y)$  is of linear growth to the direction where  $J$  is unbounded. Then there exists a dual maximizer in the sense of Definition 2.1.*

**Corollary 2.3.** *In the setting of Theorem 2.2, if  $c \in C^{0,2}$  – that is  $\frac{\partial^2 c}{\partial y^2}$  exists and is continuous on  $\Omega$  – and if  $\nu$  is compactly supported then there exists a dual maximizer.*

Note that Definition 2.1 is made in a pointwise sense, that is we do not require  $f \in L^1(\mu)$ ,  $g \in L^1(\nu)$ , nor  $h \in L^\infty(\mu)$ . But as already observed in [3, 5], this classical integrability assumption is too restrictive for the existence of dual maximizer. But by using the “extended notion of integrability” introduced in [5], this pointwise dual

maximizer  $(f, g, h)$  may still be viewed as dual maximizer in the generalized sense. To ensure integrability in the classical sense, further assumptions are required, as summarised in the following result.

**Theorem 2.4.** *Let  $\mu \preceq \nu$  be in convex order where  $J = \text{conv}(\text{supp}(\nu))$  is compact. Assume that  $c$  is Lipschitz on  $J \times J$ , and there is a Lipschitz  $u : J \rightarrow \mathbb{R}$  such that  $y \mapsto c(x, y) + u(y)$  is convex on  $J$  for  $\mu$ -a.e.  $x$ . Then there exists a dual maximizer  $(f, g, h)$  such that  $f$  and  $g$  are Lipschitz on  $J$  and  $h$  is bounded on  $J$ . In particular,  $\mu(f) + \nu(g) = \mathbb{E}^{\mathbb{P}^*} [c(X, Y)]$  for any solution  $\mathbb{P}^*$  to the problem (1.4).*

*Remark 2.5.* In Theorem 2.4 if  $c$  and  $u$  are Lipschitz with Lipschitz constant  $L$ , then  $f$  and  $g$  can be taken to be Lipschitz with constant  $7L$  and  $5L$  respectively on  $J$ , while  $|h|$  is bounded by  $6L$  on  $J$ , as shown in the proof.

We close this section with a discussion of how, and in what sense, the above results are sharp. Examples which support this discussion are presented after the proofs in Section 4. First, we note that the linear growth condition in Theorem 2.2 can not be removed. Indeed, Example 4.1 gives a cost function which violates the linear growth condition together with marginals  $\mu \preceq \nu$  for which the dual maximizers fails to exist. Second, the convexity condition (2.4) on  $J$  can not be relaxed to just local convexity around  $x$ , as shown in Example 4.2, and this even for compactly supported marginals. Third, the  $C^{0,2}$  regularity in Corollary 2.3, is optimal in the sense that for any given  $\varepsilon > 0$  we can construct a cost function  $c \in C^{2-\varepsilon}$  and compactly supported, convex-ordered marginals  $\mu \preceq \nu$  for which a dual maximizer satisfying (2.1)–(2.2) does not exist. This is carried out in Example 4.3. Finally, in Example 4.4 we show the necessity of semiconvexity for the regularity in Theorem 2.4 by showing that there exist 1-Lipschitz cost  $c$  for which there is a dual maximizer  $(f, g, h)$  but  $g \notin L^1(\nu)$ , even when  $(\mu, \nu)$  are compactly supported and irreducible (see Definition 3.1 below).

### 3. PROOFS

To establish Theorem 2.2, we prove Propositions 3.4 and 3.5 below. The key idea is to consider the martingale optimal transport problem on its irreducible components (see Definition 3.1 below and [3, Appendix A]). It is known, see Theorem 3.2 below, that on each irreducible component the dual problem admits a maximizer. Using the semiconvexity assumption on the cost function, we can show that these maximizers are appropriately bounded, such that it is possible to glue them together to obtain global maximizers of the dual problem.

Let  $\mu, \nu$  be probability measures on  $\mathbb{R}$  which are in convex order. It was shown in [3] and [5], see Proposition 3.6 below, that there is a canonical decomposition of  $\mu, \nu$  into irreducible pairs  $(\mu_i, \nu_i)_{i \in \mathbb{N}}$  such that for each irreducible pair  $(\mu_i, \nu_i)$ , the dual problem attains a solution. For a probability measure  $\mu$  on  $\mathbb{R}$ , we define its potential function by

$$u_\mu : \mathbb{R} \rightarrow \mathbb{R}, \quad u_\mu(x) := \int |x - y| d\mu(y).$$

**Definition 3.1.** *Let  $\mu \preceq \nu$  be in convex order and let  $I := \{x : u_\mu(x) < u_\nu(x)\}$ . We say that  $(\mu, \nu)$  is irreducible on the domain  $I$  if  $I$  is an open interval and  $\mu$  is concentrated on  $I$ .*

We recall the following result from [5] (which requires our standing assumption that  $P(c) \in \mathbb{R}$  and  $c(x, y) \leq a(x) + b(y)$  for some  $a \in L^1(\mu)$  and  $b \in L^1(\nu)$ ).

**Theorem 3.2.** [5, Theorem 6.2] *Let  $\mu \preceq \nu$  be irreducible on the domain  $I$ . Then a dual maximizer exists.*

*Remark 3.3.* [Existence of a primal minimizer] If  $c$  is bounded from below, then standard arguments imply the existence of a primal minimizer of the martingale transport problem (1.4) i.e.  $\mathbf{P} = \mathbb{E}^{\mathbb{P}^*} [c(X, Y)]$  for some  $\mathbb{P}^* \in \text{MT}(\mu, \nu)$ , see [2]. We note that this remains valid in the present setup where  $c$  is lower-semicontinuous,  $c$  is bounded from above in the sense  $c(x, y) \leq a(x) + b(y)$  for some  $a \in L^1(\mu)$  and  $b \in L^1(\nu)$  and  $P(c) \in \mathbb{R}$ , but  $c$  is not necessarily bounded from below. To establish this it is sufficient to argue on irreducible components and on each such component the result is a consequence of [5, Theorem 6.2] and [5, Remark 7.9]

The next proposition claims that for any dual maximizer  $(f, g, h)$  in Theorem 3.2, if the function  $y \mapsto c(x, y)$  is convex for each  $x \in I$ , then the lower envelope function  $g$  has a “desirable shape” modulo an affine function. We first deal with the bounded domain case.

**Proposition 3.4.** *Let  $\mu \preceq \nu$  be irreducible on a bounded domain  $I = ]a, b[$  and assume that  $(f, g, h)$  is a dual maximizer in Theorem 3.2. Suppose that there exist  $A \in \mathbb{R} \cup \{-\infty\}$ ,  $B \in \mathbb{R} \cup \{\infty\}$  such that  $A \leq a < b \leq B$  and that for  $\mu$ -a.e.  $x$ ,*

$$y \mapsto c(x, y) \text{ is continuous and convex on } [A, B]. \quad (3.1)$$

*Then we can find an affine function  $L(y) = L(x) + \nabla L \cdot (y - x)$  such that  $(\tilde{f}(x), \tilde{g}(y), \tilde{h}(x)) := (f(x) - L(x), g(y) - L(y), h(x) - \nabla L)$  is a dual maximizer, and furthermore*

$$\tilde{g}(y) \leq 0 \text{ on } ]a, b[, \quad (3.2)$$

$$\tilde{g}(y) \geq 0 \text{ on } [A, a] \cup [b, B], \quad (3.3)$$

$$\text{If } \nu(a) > 0 \text{ then } \tilde{g}(a) = 0. \text{ If } \nu(b) > 0 \text{ then } \tilde{g}(b) = 0. \quad (3.4)$$

*Proof. Step 1.* Let us begin by recalling some terminology from [3]: for a set  $\Gamma \subseteq \mathbb{R} \times \mathbb{R}$ , denote  $X_\Gamma$  as its projection to the first coordinate space  $\mathbb{R}$ , and  $Y_G$  to the second. We will also write  $\Gamma_x = \{y : (x, y) \in \Gamma\}$ .

Now let  $\Gamma \subseteq \mathbb{R} \times \mathbb{R}$  be the “contact set” induced by the dual optimizer  $(f, g, h)$ , that is

$$\Gamma := \{(x, y) : g(y) - f(x) - h(x) \cdot (y - x) = c(x, y)\}. \quad (3.5)$$

Then by definition every solution  $\mathbb{P}^*$  to the primal problem (1.4) is concentrated on  $\Gamma$ , i.e.  $\mathbb{P}^*(\Gamma) = 1$ . Since every such  $\mathbb{P}^*$  is a martingale measure, we can find a subset  $G$  of  $\Gamma$  such that every solution to (1.4) is still concentrated on  $G$ ,  $Y_G \subseteq [a, b]$  by the irreducibility of  $(\mu, \nu)$ , and furthermore  $G$  is *regular* in the following sense:

$$\text{For each } x \in X_G, \text{ either } x \in \text{int}(\text{conv}(G_x)) \text{ or } G_x = \{x\}. \quad (3.6)$$

Hence, we have

$$g(y) - f(x) - h(x) \cdot (y - x) \leq c(x, y) \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad (3.7)$$

$$g(y) - f(x) - h(x) \cdot (y - x) = c(x, y) \quad \text{on } G \quad (3.8)$$

where  $f, g$  are real-valued on  $X_G, Y_G$  respectively,  $Y_G \subseteq [a, b]$ ,  $G$  is regular,  $\mathbb{P}^*(G) = 1$  for every solution  $\mathbb{P}^*$  to (1.4), and  $y \mapsto c(x, y)$  is convex on  $[A, B]$  for every  $x \in X_G$ . Now for each  $x \in X_G$ , let  $I(x) = \text{int}(\text{conv}(G_x))$  which is an open interval or a point  $\{x\}$  if  $G_x = \{x\}$ . As  $\mathbb{P}^*(G) = 1$  and  $(\mu, \nu)$  is irreducible, we can find a sequence  $\{x_n\} \subseteq X_G$  such that each  $I(x_n)$  is an open interval (and we write  $I(x_n) = ]a_n, b_n[$ ), and defining  $l_n = \min_{i \leq n} a_i$ ,  $r_n = \max_{i \leq n} b_i$  and  $J_n = ]l_n, r_n[$ , we have

$$J_n \cap I(x_{n+1}) \neq \emptyset \quad \forall n, \quad \text{and} \quad (3.9)$$

$$l_n \searrow a, r_n \nearrow b \quad \text{as } n \rightarrow \infty. \quad (3.10)$$

(For more details about the existence of such a sequence, we refer to [3, Appendix A].) We may assume that the case  $l_n < a_{n+1} < b_{n+1} < r_n$  does not occur, since if it occurs then we may simply discard the respective  $x_{n+1}$  from the sequence.

Note that as  $y \mapsto c(x, y)$  is continuous and thus  $g(y)$  is upper semicontinuous (recall (2.3)), (3.8) holds for  $(x_n, a_n)$  and  $(x_n, b_n)$  for every  $n$ . Hence

$$\text{For each } n, \text{ there exist } 1 \leq i, j \leq n \text{ such that (3.8) holds for } (x_i, l_n) \text{ and } (x_j, r_n). \quad (3.11)$$

**Step 2.** For each  $n$ , define

$$v_{x_n}(y) = c(x_n, y) + f(x_n) + h(x_n) \cdot (y - x_n), \quad (3.12)$$

$$g_n(y) = \inf_{i \leq n} (v_{x_i}(y)), \quad (3.13)$$

$$L_n(y) = \frac{g_n(r_n) - g_n(l_n)}{r_n - l_n} \cdot (y - l_n) + g_n(l_n). \quad (3.14)$$

We claim that, for each  $n$ ,

$$g_n(y) \leq L_n(y) \quad \forall y \in [l_n, r_n], \quad (3.15)$$

$$g_n(y) \geq L_n(y) \quad \forall y \in [A, l_n] \cup [r_n, B]. \quad (3.16)$$

The claim is obvious for  $n = 1$  since  $v_{x_1}(y)$  is convex. Suppose that the claim is true for  $n$ . Then there are two cases.

*Case 1 :*  $a_{n+1} \leq l_n < r_n \leq b_{n+1}$ .

First, since  $g_{n+1}(y) = \min(g_n(y), v_{x_{n+1}}(y))$  and since  $g_{n+1}(a_{n+1}) = v_{x_{n+1}}(a_{n+1})$ ,  $g_{n+1}(b_{n+1}) = v_{x_{n+1}}(b_{n+1})$  by (3.7), (3.8), by convexity of  $v_{x_{n+1}}(y)$  we see that

$$g_{n+1}(y) \leq v_{x_{n+1}}(y) \leq L_{n+1}(y) \quad \forall y \in [a_{n+1}, b_{n+1}]. \quad (3.17)$$

This establishes the claim in (3.15) for  $n + 1$ .

For the second claim, fix  $i \in \{1, 2, \dots, n + 1\}$ . Then there exists  $y_i \in [a_{n+1}, b_{n+1}]$  such that  $(x_i, y_i) \in G$ . Then (3.17) implies

$$v_{x_i}(y_i) = g_{n+1}(y_i) \leq L_{n+1}(y_i). \quad (3.18)$$

Also note that by (3.7), (3.8),

$$v_{x_i}(a_{n+1}) \geq v_{x_{n+1}}(a_{n+1}) = g_{n+1}(a_{n+1}), \quad v_{x_i}(b_{n+1}) \geq v_{x_{n+1}}(b_{n+1}) = g_{n+1}(b_{n+1}). \quad (3.19)$$

Now by (3.18), (3.19) and convexity of  $v_{x_i}(y)$ , we deduce that

$$v_{x_i}(y) \geq L_{n+1}(y) \quad \forall y \in [A, a_{n+1}] \cup [b_{n+1}, B]. \quad (3.20)$$

As (3.20) holds for every  $i$ , this verifies the claim (3.16) for  $n + 1$ , completing the inductive step.

Case 2 :  $l_n \leq a_{n+1} \leq r_n \leq b_{n+1}$ .

First, since  $v_{x_{n+1}}(a_{n+1}) = g_{n+1}(a_{n+1}) \leq g_n(a_{n+1})$ , by the induction hypothesis

$$v_{x_{n+1}}(a_{n+1}) \leq L_n(a_{n+1}). \quad (3.21)$$

Also note that by (3.11), we have

$$v_{x_{n+1}}(r_n) \geq g_n(r_n) = L_n(r_n). \quad (3.22)$$

Now the convexity of  $v_{x_{n+1}}$  implies

$$v_{x_{n+1}}(b_{n+1}) \geq L_n(b_{n+1}). \quad (3.23)$$

Note that  $L_n(l_n) = g_n(l_n) = g_{n+1}(l_n)$  and  $g_{n+1}(b_{n+1}) = v_{x_{n+1}}(b_{n+1})$ . Hence by (3.23),

$$\lambda g_{n+1}(l_n) + (1 - \lambda)g_{n+1}(b_{n+1}) \geq \lambda L_n(l_n) + (1 - \lambda)L_n(b_{n+1}) \quad \forall \lambda \in [0, 1]. \quad (3.24)$$

As  $g_{n+1}(y) = \min(g_n(y), v_{x_{n+1}}(y))$ , the induction hypothesis and convexity of  $v_{x_{n+1}}(y)$  along with (3.21), (3.24) imply the first claim (3.15) for  $n + 1$ .

For the second claim, fix  $i \in \{1, 2, \dots, n + 1\}$ . Then there exists  $y_i \in [l_n, b_{n+1}]$  such that  $(x_i, y_i) \in G$ , and the first claim (3.15) for  $n + 1$  gives  $v_{x_i}(y_i) \leq L_{n+1}(y_i)$ . On the other hand,  $v_{x_i}(l_n) \geq g_n(l_n) = g_{n+1}(l_n)$  and  $v_{x_i}(b_{n+1}) \geq v_{x_{n+1}}(b_{n+1}) = g_{n+1}(b_{n+1})$ . Hence by convexity of  $v_{x_i}(y)$ , we deduce that

$$v_{x_i}(y) \geq L_{n+1}(y) \quad \forall y \in [A, l_n] \cup [b_{n+1}, B]. \quad (3.25)$$

As (3.25) holds for every  $i$ , this verifies the claim (3.16) for  $n + 1$ , completing the induction. The case  $a_{n+1} \leq l_n \leq b_{n+1} \leq r_n$  can be treated in the same way.

**Step 3.** We claim that

$$L_n(y)\text{'s are uniformly bounded on } [a, b] \text{ for all } n. \quad (3.26)$$

To prove (3.26), choose  $M > 0$  such that  $|v_{x_1}(y)| \leq M$  on  $[a, b]$ . Then  $L_n(l_n) = g_n(l_n) \leq v_{x_1}(l_n) \leq M$  and  $L_n(r_n) = g_n(r_n) \leq v_{x_1}(r_n) \leq M$ . Hence, as  $L$  is linear,  $L_n(a_1) \leq M$  and  $L_n(b_1) \leq M$ . On the other hand, by Step 2,  $-M \leq v_{x_1}(a_1) = g_n(a_1) \leq L_n(a_1)$  and  $-M \leq v_{x_1}(b_1) = g_n(b_1) \leq L_n(b_1)$ . This implies (3.26). In particular, there exists a subsequence of  $L_n$  (which we denote as  $L_k$ ) such that  $L_k(y)$  uniformly converges to an affine function as  $k \rightarrow \infty$ , say  $L(y)$  on every compact interval in  $\mathbb{R}$ . Now we claim that, for  $v_x(y) := c(x, y) + f(x) + h(x) \cdot (y - x)$  and  $g(y) := \inf_{x \in X_G} (v_x(y))$ ,

$$g(y) \leq L(y) \text{ on } ]a, b[, \quad (3.27)$$

$$g(y) \geq L(y) \text{ on } [A, a] \cup [b, B]. \quad (3.28)$$

First it is easy to see (3.27) as follows: if  $y \in ]a, b[$  then for all large  $k$  we have  $y \in ]l_k, r_k[$ , thus by Step 2,  $g(y) \leq g_k(y) \leq L_k(y)$ . By taking  $k \rightarrow \infty$ , we see that  $g(y) \leq L(y)$ , proving (3.27).

Next, suppose that there exists  $(x, y) \in G$  with  $a < y < b$ . Then again for all large  $k$  we have  $v_x(y) = g(y) \leq g_k(y) \leq L_k(y)$ , thus  $v_x(y) \leq L(y)$ . On the other hand, by (3.7), (3.8) we have  $v_x(l_k) \geq g_k(l_k)$  and  $v_x(r_k) \geq g_k(r_k)$ , thus  $v_x(a) \geq L(a)$  and  $v_x(b) \geq L(b)$  by letting  $k \rightarrow \infty$ . By convexity of  $v_x$ , this implies that  $v_x(y) \geq L(y)$  on  $[A, a] \cup [b, B]$ .

If there is no  $y$  such that  $a < y < b$  and  $(x, y) \in G$ , this means that  $G_x = \{a, b\}$ , i.e.  $(x, a), (x, b) \in G$ . Then without loss of generality we may simply include this  $x$  in the sequence  $\{x_n\}$  defined in Step 1, say we put  $x = x_1$ . This implies that  $l_n = a$

and  $r_n = b$  for all  $n$ . Thus  $v_x(a) = L(a)$  and  $v_x(b) = L(b)$ . By convexity of  $v_x$ , this implies that  $v_x(y) \geq L(y)$  on  $[A, a] \cup [b, B]$ . Hence, for any  $x \in X_G$  we deduce that  $v_x(y) \geq L(y)$  on  $[A, a] \cup [b, B]$ , therefore (3.28) follows.

If  $\nu(a) > 0$  then there exists  $x_a \in I$  such that  $(x_a, a) \in G$ . Then again we may include  $x_a$  in the sequence  $\{x_n\}$  defined in Step 1. This implies that  $l_n = a$  for all large  $n$ , thus  $g(a) = v_{x_a}(a) = L(a)$ . Similarly if  $\nu(b) > 0$  then  $g(b) = L(b)$ .

Finally, we see that  $(\tilde{f}(x), \tilde{g}(y), \tilde{h}(x)) := (f(x) - L(x), g(y) - L(y), h(x) - \nabla L)$  satisfies (3.7), (3.8), (3.2), (3.3) and (3.4), concluding the proof.  $\square$

Next, we deal with the half-infinite domain case.

**Proposition 3.5.** *Let  $\mu \preceq \nu$  be irreducible on a half-infinite domain  $I = ]0, \infty[$  and assume that  $(f, g, h)$  is a dual maximizer in Theorem 3.2. Suppose that there exists  $A \in [-\infty, 0]$  such that for  $\mu$ -a.e.  $x$ ,*

*$y \mapsto c(x, y)$  is continuous and convex on  $[A, \infty[$ , and*

*there exists an affine function  $L_x$  such that  $c(x, y) \leq L_x(y)$  for all  $y \geq 0$ .*

*Then we can find an affine function  $L$  such that  $(\tilde{f}(x), \tilde{g}(y), \tilde{h}(x)) := (f(x) - L(x), g(y) - L(y), h(x) - \nabla L)$  is a dual maximizer, and furthermore*

$$\tilde{g}(y) \leq 0 \text{ on } ]0, \infty[, \quad (3.29)$$

$$\tilde{g}(y) \geq 0 \text{ on } [A, 0], \quad (3.30)$$

$$\text{if } \nu(0) > 0 \text{ then } \tilde{g}(0) = 0. \quad (3.31)$$

*Proof. Step 1.* Let  $G \subseteq \mathbb{R} \times \mathbb{R}$  be chosen as in the Step 1 in Proposition 3.4, i.e.  $Y_G \subseteq [0, \infty[$ ,  $G$  is regular,  $\mathbb{P}^*(G) = 1$  for every solution  $\mathbb{P}^*$  to (1.4),  $y \mapsto c(x, y)$  is convex and bounded above by an affine function on  $[0, \infty[$  for every  $x \in X_G$ , and (3.7), (3.8) holds. For each  $x \in X_G$ , let  $I(x) = \text{int}(\text{conv}(G_x))$  as before. We can find a sequence  $\{x_n\} \subseteq X_G$  such that each  $I(x_n)$  is an open interval (and we write  $I(x_n) = ]a_n, b_n[$ , where  $b_n$  can be  $+\infty$ ), and if we define  $l_n = \min_{i \leq n} a_i$ ,  $r_n = \max_{i \leq n} b_i$  and also define  $J_n = ]l_n, r_n[$ , then

$$J_n \cap I(x_{n+1}) \neq \emptyset \quad \forall n, \quad \text{and} \quad (3.32)$$

$$l_n \searrow 0, r_n \nearrow +\infty \quad \text{as } n \rightarrow \infty. \quad (3.33)$$

**Step 2.** Recall definitions (3.12) – (3.14). Altering the triple  $(f, g, h)$  by an appropriate affine function and using the condition of linear growth and convexity satisfied by the cost, we can assume that

$$v_{x_1}(y) \text{ is decreasing on } [A, \infty[, v(0) = 0 \text{ and } \lim_{y \rightarrow \infty} v_{x_1}(y) = b > -\infty. \quad (3.34)$$

Now we claim that, for each  $n$ ,

$$g_n(y) \text{ is decreasing on } [A, l_n], \text{ and } g_n(y) \leq g_n(l_n) \text{ on } [l_n, \infty[. \quad (3.35)$$

Note that the claim (3.35) is obviously true for  $n = 1$  by the assumption (3.34). Suppose the claim is true for  $n$ . We will show that the claim is also true for  $n + 1$ . To see this, note that as  $b_{n+1} \geq l_n$ , using (3.7), (3.8) and the induction hypothesis (3.35), we see that

$$v_{x_{n+1}}(b_{n+1}) \leq g_n(b_{n+1}) \leq g_n(l_n), \text{ while } v_{x_{n+1}}(l_n) \geq g_n(l_n). \quad (3.36)$$

(If  $b_{n+1} = \infty$ , then instead of  $b_{n+1}$  we may argue with arbitrarily large  $c_{n+1}$  satisfying  $(x_{n+1}, c_{n+1}) \in G$ .) By (3.36) and convexity of  $v_{x_{n+1}}(y)$ , we see that  $v_{x_{n+1}}(y)$  is decreasing on  $[A, l_n]$ . As  $g_{n+1}(y) = \min(g_n(y), v_{x_{n+1}}(y))$  and  $g_n(y)$  is decreasing on  $[A, l_n]$ , we see that

$$g_{n+1}(y) \text{ is decreasing on } [A, l_n]. \quad (3.37)$$

In particular, for any  $y \in [l_{n+1}, l_n]$  we have  $g_{n+1}(l_{n+1}) \geq g_{n+1}(y)$ . For  $y \geq l_n$ , we see that  $g_{n+1}(l_{n+1}) \geq g_{n+1}(l_n) = g_n(l_n)$  by (3.11), and  $g_n(l_n) \geq g_n(y) \geq g_{n+1}(y)$  by (3.35). Hence

$$g_{n+1}(y) \leq g_{n+1}(l_{n+1}) \text{ on } [l_{n+1}, \infty[. \quad (3.38)$$

Therefore, (3.35) is proved for all  $n$ .

We have observed that the sequence  $\{g_n(l_n)\}$  is increasing. Note that  $g_n(l_n) \leq v_{x_1}(l_n) \leq v_{x_1}(0) = 0$  for all  $n$ , thus  $\{g_n(l_n)\}$  converges to, say,  $L$ . Then we claim

$$g(y) \geq L \text{ on } [A, 0], \quad (3.39)$$

$$g(y) \leq L \text{ on } ]0, +\infty[, \quad (3.40)$$

where, as before,  $v_x(y) := c(x, y) + f(x) + h(x) \cdot (y - x)$  and  $g(y) := \inf_{x \in X_G} (v_x(y))$ .

To see this, fix  $x > 0$ . Then there exists  $n$  such that  $l_n < x$ . Arguing as above, we see that  $v_x(y)$  is decreasing on  $[A, l_n]$  and  $v_x(l_n) \geq g_n(l_n)$  for all  $n$ . Hence for any  $y \leq 0$  we see that  $v_x(y) \geq g_n(l_n)$ . Letting  $n \rightarrow \infty$  we conclude

$$g(y) \geq L \text{ for all } y \in [A, 0]. \quad (3.41)$$

Now for  $y > 0$ , there exists  $n$  such that  $l_n < y$ . Then by (3.35), we see that  $g(y) \leq g_n(y) \leq g_n(l_n)$  for all large  $n$ , thus by taking  $n \rightarrow \infty$  we conclude

$$g(y) \leq L \text{ for all } y > 0. \quad (3.42)$$

If  $\nu(0) > 0$  then there is  $x \in X_G$  with  $(x, 0) \in G$ , and we may simply put this  $x$  into the sequence  $\{x_n\}$  by letting  $x = x_1$ . Then every  $l_n$  simply becomes 0 and  $\{g_n(l_n)\}$  becomes the constant sequence  $L$ . Hence,  $g(0) = L$ . Finally, altering the triple  $(f, g, h)$  by the constant function  $-L$ , we can assume that  $L = 0$ . This proves the proposition.  $\square$

We are now ready to show the existence of dual optimizers for martingale optimal transport problem in Theorem 2.2. In particular we no longer assume the irreducibility of  $(\mu, \nu)$ . Note that If  $(\mu, \nu)$  is irreducible on the domain  $I = \mathbb{R}$  then Theorem 2.2 simply follows from Theorem 3.2. Otherwise,  $(\mu, \nu)$  can be decomposed into at most countably many irreducible components, and any martingale  $\mathbb{P} \in \text{MT}(\mu, \nu)$  is decomposed accordingly. More precisely we recall:

**Proposition 3.6.** [3, Theorem A.4] *Let  $\mu \preceq \nu$  and let  $(I_k)_{k \geq 1}$  be the open connected components of the set  $\{x : u_\mu(x) < u_\nu(x)\}$ . Set  $I_0 = \mathbb{R} \setminus \cup_{k \geq 1} I_k$  and  $\mu_k = \mu|_{I_k}$  for  $k \geq 0$ , so that  $\mu = \sum_{k \geq 0} \mu_k$ . Then, there exists a unique decomposition  $\nu = \sum_{k \geq 0} \nu_k$  such that*

$$\mu_0 = \nu_0 \quad \text{and} \quad \mu_k \preceq \nu_k \quad \text{for all } k \geq 1,$$

*and this decomposition satisfies  $I_k = \{x : u_{\mu_k}(x) < u_{\nu_k}(x)\}$  for all  $k \geq 1$ . Moreover, any  $\mathbb{P} \in \text{MT}(\mu, \nu)$  admits a unique decomposition  $\mathbb{P} = \sum_{k \geq 0} \mathbb{P}_k$  such that  $\mathbb{P}_k \in \text{MT}(\mu_k, \nu_k)$  for all  $k \geq 0$ .*

Note that as  $\mu_0 = \nu_0$ ,  $\mathbb{P}_0$  must be the identity martingale. We can now give the proof of our first main result.

*Proof of Theorem 2.2.* Notice that by definition of the dual maximizer and the assumption on the cost, we can assume that  $y \mapsto c(x, y)$  is continuous and convex on  $J := \text{conv}(\text{supp}(\nu))$ . Let  $\mathbb{P}^*$  be any minimizer in  $\text{MT}(\mu, \nu)$  for the problem (1.4). Then  $\mathbb{P}_k^*$  is a minimizer in  $\text{MT}(\mu_k, \nu_k)$ . For each  $k \geq 1$ , choose a set  $G_k \subseteq \mathbb{R} \times \mathbb{R}$  and a triple  $(f_k, g_k, h_k)$  provided by Proposition 3.4 if  $I_k$  is bounded, or by Proposition 3.5 if  $I_k$  is half-infinite. We need to define  $G_0$  and  $(f_0, g_0, h_0)$  for  $I_0$ . As  $\mathbb{P}_0^*$  is the identity map, of course we take  $G_0 := \{(x, x) : x \in I_0\}$ . For each  $x \in I_0$  define  $f_0(x) = -c(x, x)$ , and choose  $h_0(x)$  in such a way that the convex function  $v_x(y) := c(x, y) + f_0(x) + h_0(x) \cdot (y - x)$  satisfies  $v(x) = v'(x) = 0$  (more precisely 0 belongs to the subdifferential of  $v_x$  at  $x$ ). Define  $g_0(y) = \inf_{x \in I_0} \{v_x(y)\}$  so in particular  $g_0 = 0$  on  $I_0$ . Finally, define

$$f(x) = f_k(x) \quad \text{if } x \in X_{G_k}, \quad (3.43)$$

$$h(x) = h_k(x) \quad \text{if } x \in X_{G_k}, \quad (3.44)$$

$$g(y) = \inf_{k \geq 0} g_k(y). \quad (3.45)$$

Let  $G = \cup_{k \geq 0} G_k$ . Obviously  $\mathbb{P}^*(G) = 1$ . Now observe that the properties (3.2), (3.3), (3.4), (3.29), (3.30), (3.31) verified in Proposition 3.4, 3.5 clearly indicate that the triples  $(f_k, g_k, h_k)_{k \geq 0}$  are compatible, that is, the duality (3.7), (3.8) holds for  $G$  and  $(f, g, h)$ . This completes the proof.  $\square$

*Remark 3.7.* The linear growth assumption of the function  $y \mapsto c(x, y)$  is required only for those  $x$  in the half-infinite irreducible domain of  $\mu, \nu$  as in Proposition 3.5.

*Proof of Theorem 2.4.* We will say that a function  $f$  is  $L$ -Lipschitz on  $D$  if  $|f(x) - f(y)| \leq L|x - y|$  for all  $x, y \in D$ . Assume  $c$  is  $L_1$ -Lipschitz on  $J \times J$  and  $u$  is  $L_2$ -Lipschitz on  $J$ .

Consider the decomposition of  $(\mu, \nu)$  into irreducible pairs given by Proposition 3.6. Fix  $i \geq 1$  and consider  $(\mu_i, \nu_i)$  which is irreducible on the bounded domain  $I_i = ]a_i, b_i[$ . Let  $\tilde{c}(x, y) := c(x, y) + u(y)$ , and let  $(f_i, g_i, h_i)$  be a dual maximizer for the cost  $\tilde{c}$  and  $(\mu_i, \nu_i)$ , satisfying the conclusion of Proposition 3.4. In the Step 1 of the proof of Proposition 3.4 we explained that there exists a regular set  $G_i \subseteq I_i \times \bar{I}_i$  on which every (decomposed) solution  $\mathbb{P}_i^*$  to the problem (1.4) is concentrated, so that the duality (2.1), (2.2) holds with  $(f_i, g_i, h_i)$ ,  $\tilde{c}$ , and  $G_i$ .

Define  $v_{i,x}(y) := f_i(x) + h_i(x) \cdot (y - x) + \tilde{c}(x, y)$  and note that as  $y \mapsto \tilde{c}(x, y)$  is  $L$ -Lipschitz and convex on  $J = \text{conv}(\text{supp } \nu) = [A, B]$  where  $L = L_1 + L_2$ , we have

$$\frac{dv_{i,x}}{dy}(b_i^-) - \frac{dv_{i,x}}{dy}(a_i^+) \leq 2L. \quad (3.46)$$

Proposition 3.4 tells us that  $g_i(a_i) \geq 0$ ,  $g_i(b_i) \geq 0$  while there is  $y \in [a_i, b_i]$  such that  $g_i(y) \leq 0$ , since  $g_i(y) \leq 0$  whenever  $(x, y) \in G_i$  for some  $x$ . As  $g_i(y) := \inf_{x \in X_{G_i}} \{v_{i,x}(y)\}$ , if  $(x, y) \in G_i$  then we have  $v_{i,x}(a_i) \geq 0$ ,  $v_{i,x}(b_i) \geq 0$  while  $v_{i,x}(y) \leq 0$ . With (3.46) this implies that  $v_{i,x}$  is  $2L$ -Lipschitz for any  $x \in X_{G_i}$ , hence  $g_i$  is also  $2L$ -Lipschitz on  $[a_i, b_i]$ . Proposition 3.4 then tells us that we can replace  $g_i$  with  $\tilde{g}_i := g_i \wedge 0$ , so that  $\tilde{g}_i$  is  $2L$ -Lipschitz on  $\mathbb{R}$ ,  $\tilde{g}_i \leq 0$  on  $[a_i, b_i]$ ,  $\tilde{g}_i = 0$  on  $\mathbb{R} \setminus ]a_i, b_i[$ , and  $(f_i, \tilde{g}_i, h_i)$  satisfies (2.1) – (2.2).

In the proof of Theorem 2.2 we showed that there is a dual maximizer  $(f, g, h)$  to the problem (1.4) where  $g := \inf_{i \geq 0} \tilde{g}_i$  (recall that  $\tilde{g}_0 \equiv 0$  on  $I_0$ ). Hence we get that  $g$  is  $2L$ -Lipschitz on  $\mathbb{R}$ ,  $g \leq 0$  on  $J$ , and  $g = 0$  on  $J^c$ .

Next, observing the duality relation

$$g(y) - \tilde{c}(x, y) \leq f(x) + h(x) \cdot (y - x) \quad \forall x \in J, \forall y \in J, \quad (3.47)$$

note that we can replace  $f, h$  by  $\tilde{f}, \tilde{h}$  respectively, as follows: define  $H : J \times J \rightarrow \mathbb{R}$  by the upper concave envelope in  $y$  variable

$$H(x, y) := \text{conc}[g(\cdot) - \tilde{c}(x, \cdot)](y).$$

Then we define  $\tilde{f}(x) := H(x, x)$  and  $\tilde{h}(x) := \frac{\partial H(x, y)}{\partial y} \Big|_{y=x}$ . More precisely,  $\tilde{h}(x)$  is an element of the superdifferential of the concave function  $y \rightarrow H(x, y)$  at  $x$ , and there exists a measurable choice of such an  $\tilde{h}$ . Now in view of (3.47), it is clear that  $(\tilde{f}, g, \tilde{h})$  is a dual maximizer. Observe that since  $y \mapsto g(y) - \tilde{c}(x, y)$  is  $3L$ -Lipschitz, it is immediate that  $|\tilde{h}| \leq 3L$  on  $J$ . Then, to see that  $\tilde{f}$  is Lipschitz, note that since  $x \mapsto g(y) - \tilde{c}(x, y)$  is  $L_1$ -Lipschitz, by definition of  $H$  we have

$$|H(x, y) - H(x', y)| \leq L_1|x - x'| \quad \forall x, x', y \in J.$$

On the other hand, since the concave envelope of a Lipschitz function is Lipschitz,

$$|H(x, y) - H(x, y')| \leq 3L|y - y'| \quad \forall x, y, y' \in J.$$

These inequalities immediately imply that, for any  $x, x' \in J$ ,

$$|\tilde{f}(x) - \tilde{f}(x')| = |H(x, x) - H(x', x')| \leq (L_1 + 3L)|x - x'|.$$

Finally, recall that  $\tilde{c}(x, y) = c(x, y) + u(y)$  where  $u$  is  $L_2$ -Lipschitz, we replace  $g$  with  $\tilde{g} = g - u$  which is  $2L + L_2$  Lipschitz on  $J$ . Then  $(\tilde{f}, \tilde{g}, \tilde{h})$  is a dual maximizer satisfying the conclusion of Theorem 2.4.  $\square$

#### 4. SOME INSIGHTFUL EXAMPLES

*Example 4.1.* In this example, we show that the linear growth condition on the cost for the half-infinite domain case cannot be dropped in Theorem 2.2.

Let  $x_n = n$ ,  $n = 1, 2, 3, \dots$  and let  $G_n = \{n - 1, n + 1\}$ . For each  $n$ , choose  $c(n, y)$ ,  $f(n)$  and  $h(n)$  appropriately such that  $v_n(y) = c(n, y) + f(n) + h(n) \cdot (y - n)$  becomes

$$v_n(y) = y^2 \quad \text{if } y \geq n - 1, \quad (4.1)$$

$$v_n(y) = (n - 1)y \quad \text{if } y \leq n - 1. \quad (4.2)$$

Then the triple  $(f, g, h)$  supports the set  $G = \{(n, n + 1), (n, n - 1) : n \in \mathbb{N}\}$  in view of (3.7), (3.8), and clearly

$$g(y) = -\infty \quad \text{on } ]-\infty, 0[. \quad (4.3)$$

Hence, Proposition 3.5 cannot hold in this case.

*Example 4.2.* In this example, we show that if the convexity assumption on  $y \mapsto c(x, y)$  holds only locally around  $x$ , then the dual maximizer can fail to exist.

Define the cost function by

$$c(x_\infty, y) = 0, \quad \text{and} \quad c(x_n, y) = \begin{cases} 0 & \text{if } y \in [y_{n-1}, y_n], \\ y - y_{n-1} & \text{if } y \leq y_{n-1}, \\ -y + y_n & \text{if } y \geq y_n. \end{cases} \quad (4.4)$$

Let  $y_0 = 0$ ,  $y_n = \sum_{k=1}^n \frac{1}{n^2}$ ,  $x_n = (y_{n-1} + y_n)/2$ , and  $x_\infty = y_\infty = \sum_{k=1}^\infty \frac{1}{n^2}$ . Let  $\mu$  be any probability measure whose support is  $\{x_n\}_{1 \leq n \leq \infty}$ , and construct a martingale measure  $\mathbb{P}$  whose disintegration  $(\mathbb{P}_x)_x$  w.r.t.  $\mu$  is as follows:

$$\mathbb{P}_{x_n} = \frac{1}{2}(\delta_{y_{n-1}} + \delta_{y_n}) \quad \forall n = 1, 2, \dots \quad \text{and} \quad \mathbb{P}_{x_\infty} = \delta_{x_\infty},$$

and define  $\nu$  as the second marginal of  $\mathbb{P}$ . Notice that then  $\mathbb{P}$  is the unique element in  $\text{MT}(\mu, \nu)$ . Now we will show that this (optimal)  $\mathbb{P}$  does not allow a dual maximizer, that is, there does not exist a triple  $(f, g, h)$  which satisfies the following:

$$g(y) \leq c(x_n, y) + f(x_n) + h(x_n) \cdot (y - x_n) \quad \forall n \in \mathbb{N} \cup \{\infty\}, \quad \forall y \in \mathbb{R}, \quad (4.5)$$

$$g(y_n) = c(x_n, y_n) + f(x_n) + h(x_n) \cdot (y_n - x_n) \quad \forall n \geq 1, \quad (4.6)$$

$$g(y_{n-1}) = c(x_n, y_{n-1}) + f(x_n) + h(x_n) \cdot (y_{n-1} - x_n) \quad \forall n \geq 1, \quad (4.7)$$

$$g(y_\infty) = c(x_\infty, y_\infty) + f(x_\infty) + h(x_\infty) \cdot (y_\infty - x_\infty) = f(x_\infty). \quad (4.8)$$

Recall that once such a  $(f, g, h)$  exists, then we can redefine  $g$  as follows:

$$g(y) := \inf_{n \in \mathbb{N} \cup \{\infty\}} (c(x_n, y) + f(x_n) + h(x_n) \cdot (y - x_n)). \quad (4.9)$$

We claim that, if we have such a  $(f, g, h)$ , then we must have

$$g(y_\infty) = -\infty,$$

which is a contradiction to (4.8). To see this, for convenience let us define

$$v_x(y) := c(x, y) + f(x) + h(x) \cdot (y - x). \quad (4.10)$$

Then by (4.5), (4.6), (4.7) we must have

$$v_{x_n}(y_n) = v_{x_{n+1}}(y_n), \quad \text{and} \quad (4.11)$$

$$v_{x_n}(y_{n-1}) \leq v_{x_{n+1}}(y_{n-1}), \quad \forall n \geq 1. \quad (4.12)$$

Notice that these with the definition of  $c(x_n, y)$  immediately implies

$$h(x_n) \geq h(x_{n+1}) + 1, \quad \forall n \geq 1. \quad (4.13)$$

Also notice that  $g(y)$  is a piecewise linear function on  $[0, y_\infty[$ , and in fact  $g(y) = f(x_n) + h(x_n) \cdot (y - x_n)$  on  $[y_{n-1}, y_n]$ . Hence by (4.13) and the fact  $\sum_n \frac{1}{n} = \infty$  and the concavity of  $g$ , we see that

$$g(y_\infty) = \lim_{y \nearrow y_\infty} g(y) = -\infty,$$

a contradiction to the fact that  $g(y_\infty)$  must be real-valued.

*Example 4.3.* In this example, we show that the  $C^2$  regularity required in Theorem 2.2 is optimal in the following sense: for any  $1 < r < 2$ , we construct a cost function  $c \in C^r$  and compactly supported marginals  $\mu \preceq \nu$  for which the dual attainment fails. This example shall be a slight modification of the previous one. First, let

$$c(x, y) = -|x - y|^r, \quad (4.14)$$

and choose  $s$  such that

$$s > 1 \quad \text{and} \quad sr < 2. \quad (4.15)$$

Let  $y_0 = 0$ ,  $y_n = \sum_{k=1}^n \frac{1}{n^s}$ ,  $x_n = (y_{n-1} + y_n)/2$ , and  $x_\infty = y_\infty = \sum_{k=1}^\infty \frac{1}{n^s}$ . Define a martingale measure  $\mathbb{P}$  and its marginals  $\mu, \nu$  as in Example 4.2. Note that  $\mu, \nu$  are compactly supported since  $s > 1$ . Again we will show that this (optimal)  $\mathbb{P}$  does not allow a dual maximizer, that is, there does not exist a triple  $(f, g, h)$  which

satisfies (4.5), (4.6), (4.7), (4.8), where  $g$  is given as in (4.9). Again we will show that  $g(y_\infty) = -\infty$ , which is a contradiction to (4.7). To see this, again define  $v_x$  as in (4.10) so that we have (4.11), (4.12). Next, let us consider the slope

$$b_n = \frac{v_{x_n}(y_n) - v_{x_n}(y_{n-1})}{y_n - y_{n-1}}. \quad (4.16)$$

In order to estimate  $b_n$ , we will first estimate  $b_n - b_{n+1}$ . For this, as we can modify the  $(f, g, h)$  by an affine function, we can assume that  $f(x_{n+1}) = h(x_{n+1}) = 0$ , thus without loss of generality we can assume that  $b_{n+1} = 0$ . Now notice that, because of (4.5), (4.6), (4.7), we must have the following inequality:

$$\begin{aligned} b_n - b_{n+1} &\geq \frac{v_{x_{n+1}}(y_n) - v_{x_{n+1}}(y_{n-1})}{y_n - y_{n-1}} \\ &= n^s [ -|y_n - x_{n+1}|^r + |y_{n-1} - x_{n+1}|^r ] \\ &= n^s \left[ -\left(\frac{1}{2(n+1)^s}\right)^r + \left(\frac{1}{2(n+1)^s} + \frac{1}{n^s}\right)^r \right] \\ &= n^s \left[ \frac{(2(n+1)^s + n^s)^r - n^{sr}}{2n^{sr}(n+1)^{sr}} \right] \\ &\approx Cn^s \cdot n^{sr} \cdot n^{-2sr} = Cn^{s-sr}. \end{aligned} \quad \text{Hence we deduce that}$$

$$-b_n = \sum_{k=0}^{n-1} (b_k - b_{k+1}) \gtrsim Cn^{1+s-sr}. \quad \text{This implies that, since } sr < 2,$$

$$(y_n - y_{n-1})b_n \lesssim -Cn^{1-sr} \quad \implies \quad \sum_{n=1}^{\infty} (y_n - y_{n-1})b_n = -\infty.$$

Again as in Example 4.2, this tells us that  $g(y_\infty) = -\infty$ , a contradiction to (4.8).

*Example 4.4.* In this example we show the necessity of semiconvexity for the regularity in Theorem 2.4 and the Lipschitzness of  $c$  alone is not sufficient, by constructing a 1-Lipschitz cost  $c$  and a compactly supported, irreducible pair  $(\mu, \nu)$  for which  $(f, g, h)$  is a dual maximizer, but  $g \notin L^1(\nu)$ .

To do this, we take  $c(x, y) = -|x - y|$ , and let  $I = ]0, 1[$  and  $\mu = \text{Leb}|_{[0,1]}$ . Choose a smooth and strictly concave function  $\xi : I \rightarrow \mathbb{R}_-$  such that  $\xi \leq 0$ ,  $\xi(\frac{1}{2}) = 0$ ,  $\lim_{x \rightarrow 0^+} \xi(x) = \lim_{x \rightarrow 1^-} \xi(x) = -\infty$ , and  $\int_0^1 \xi(x) \mu(dx) = -\infty$ . Now we will construct a probability measure  $\nu$  where  $\nu(I) = 1$  and  $(\mu, \nu)$  are irreducible, and also find a dual maximizer  $(f, g, h)$  where  $g = \xi$ . Then  $\int g(x) \nu(dx) \leq \int g(x) \mu(dx) = -\infty$ , as claimed.

To construct such  $\nu$  and  $(f, g, h)$ , observe that for each  $x \in I$  there exist unique  $f(x), h(x)$  such that the function  $v_x(y) := f(x) + h(x) \cdot (y - x) - |x - y|$  satisfies

- (1)  $v_x(y) \geq \xi(y) \quad \forall x \in I, \forall y \in I$ , and
- (2) for each  $x \in I$ ,  $v_x$  is tangent to  $\xi$  at two points, say  $y^-(x), y^+(x)$ .

Note that then  $y^-, y^+$  are well-defined on  $I$ , and  $0 < y^-(x) < x < y^+(x) < 1$ . Define a probability measure  $\mathbb{P}_x := \frac{y^+(x) - x}{y^+(x) - y^-(x)} \delta_{y^-(x)} + \frac{x - y^-(x)}{y^+(x) - y^-(x)} \delta_{y^+(x)}$ , and  $\mathbb{P} \in P(\mathbb{R}^2)$  by  $\mathbb{P}(dx, dy) = \mathbb{P}_x(dy) \cdot \mu(dx)$ , i.e.  $(\mathbb{P}_x)_x$  is a disintegration of  $\mathbb{P}$  with respect to  $\mu$ . Define  $\nu$  as the second marginal of  $\mathbb{P}$  and note that by definition of  $\mathbb{P}$ ,  $(\mu, \nu)$  are irreducible and are concentrated on  $I$ . Now observe that the definition of

$f, h$  gives us that  $g(y) := \inf_{x \in I} \{v_x(y)\}$  satisfies  $g = \xi$  so that  $\int g(x)\nu(dx) = -\infty$ , and  $(f, g, h)$  is a dual maximizer with respect to  $\mu, \nu$  and  $c$ .

## REFERENCES

- [1] M. Beiglböck, A. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math.*, 208(2):327–400, 2017.
- [2] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: A mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013.
- [3] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016.
- [4] M. Beiglböck and M. Nutz. Martingale inequalities and deterministic counterparts. *Electron. J. Probab.*, 2014.
- [5] M. Beiglböck, M. Nutz, and N. Touzi. Complete Duality for Martingale Optimal Transport on the Line. *Ann. Probab.*, to appear, 2016.
- [6] L. Campi, I. Laachir, and C. Martini. Change of numeraire in the two-marginals martingale transport problem. *Finance Stoch.*, 21(2):471–486, June 2017.
- [7] Y. Dolinsky and H. M. Soner. Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields*, 160(1-2):391–427, 2014.
- [8] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1):312–336, 2014.
- [9] N. Ghoussoub, Y.-H. Kim, and T. Lim. Structure of optimal martingale transport plans in general dimensions. arXiv:1508.01806, 2016.
- [10] P. Henry-Labordère, J. Oblój, P. Spoida, and N. Touzi. The maximum maximum of a martingale with given  $n$  marginals. *Ann. Appl. Probab.*, 26(1):1–44, 2016.
- [11] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, Berlin, 2011.
- [12] D. Hobson and M. Klimmek. Robust price bounds for the forward starting straddle. *Finance Stoch.*, 9(1):189–214, April 2015.
- [13] D. Hobson and A. Neuberger. Robust bounds for forward start options. *Math. Finance*, 22(1):31–56, 2012.
- [14] M. Nutz, F. Stebegg, and X. Tan. Multiperiod Martingale Transport. *ArXiv e-prints*, March 2017.
- [15] J. Oblój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004.
- [16] J. Oblój, P. Spoida, and N. Touzi. Martingale inequalities for the maximum via pathwise arguments. In *In Memoriam Marc Yor-Séminaire de Probabilités XLVII*, pages 227–247. Springer, 2015.
- [17] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.