DUALITY FOR BOREL MEASURABLE COST FUNCTIONS

MATHIAS BEIGLBÖCK AND WALTER SCHACHERMAYER

Abstract. We consider the Monge-Kantorovich transport problem in an abstract measure theoretic setting. Our main result states that duality holds if $c : X \times Y \to [0, \infty)$ is an arbitrary Borel measurable cost function on the product of Polish spaces X, Y. In the course of the proof we show how to relate a non - optimal transport plan to the optimal transport costs via a "subsidy" function and how to identify the dual optimizer. We also provide some examples showing the limitations of the duality relations.

1. Introduction

We consider the *Monge-Kantorovich transport problem* for Borel probability measures μ, ν on Polish spaces X, Y. See [RR98, Vil03, Vil05] for a general account of the theory of optimal transportation. The *cost function* $c : X \times Y \to [0, \infty]$ is assumed to be Borel measurable. $\Pi(\mu, \nu)$ is the set of all *transport plans*, that is, Borel probability measures on $X \times Y$ which have X-marginal μ and Y-marginal ν. The *transport costs* associated to a transport plan π are given by

$$I_c[\pi] = \int_{X \times Y} c \, d\pi \tag{1}$$

and we say that π is a *finite* transport plan if $I_c[\pi] < \infty$. The infimum over all possible transport costs $I_c[\pi], \pi \in \Pi(\mu, \nu)$ will be denoted by I_c. We define $\Phi(\mu, \nu)$ as the set of pairs (ϕ, ψ) of integrable functions $\phi : X \to [-\infty, \infty)$ and $\psi : Y \to [-\infty, \infty)$ which satisfy $\phi(x) + \psi(y) \leq c(x, y)$ for all $(x, y) \in X \times Y$. The dual part of the Monge-Kantorovich problem then consists in maximizing

$$J(\phi, \psi) = \int_X \phi \, d\mu + \int_Y \psi \, d\nu \tag{2}$$

for $(\phi, \psi) \in \Phi(\mu, \nu)$. Monge-Kantorovich duality asserts that $\inf\{I_c[\pi] : \pi \in \Pi(\mu, \nu)\} = \sup\{J(\phi, \psi) : (\phi, \psi) \in \Phi(\mu, \nu)\}$. For example, if X and Y consist of N points, each having measure $1/N$, duality holds true, as this reduces to an elementary linear programming problem. In the literature duality has been established under various conditions, see for instance [Vil05, p 98f] for a short overview. In particular it is known that duality holds if c is lower semi-continuous (see [Kel84, Theorem 2.2] or [Vil05, Theorem 5.10]) or just Borel measurable and bounded by the sum of two integrable functions ([Kel84, Theorem 2.14]). In [RR95, RR96] the

2000 *Mathematics Subject Classification.* 49K27, 28A05.

Key words and phrases. Monge-Kantorovich problem, Monge-Kantorovich Duality, c-cyclical monotonicity, measurable cost function.

The first author gratefully acknowledges financial support from the Austrian Science Fund (FWF) under grant S9612. The second author gratefully acknowledges financial support from the Austrian Science Fund (FWF) under grant P19456, from the Vienna Science and Technology Fund (WWTF) under grant MA13 and from the Christian Doppler Research Association (CDG).

problem is investigated beyond the realm of Polish spaces and it is characterized for which spaces duality holds for all bounded measurable cost functions.

Our main result is that Monge-Kantorovich duality holds in the case of a finite but not necessarily bounded Borel measurable cost function.

Theorem 1. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $c : X \times Y \to [0, \infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that there exists a finite transport plan. Then*

$$I_c = \inf\{I_c[\pi] : \pi \in \Pi(\mu, \nu)\} = \sup\{J(\phi, \psi) : (\phi, \psi) \in \Phi(\mu, \nu)\}. \tag{3}$$

In contrast to the situation where c is lower semi-continuous, duality does not hold in general in the Borel setting if c attains the value ∞ on a large set, see Example 4.1.

1.1. Existence of dual maximizers.

In general it is not possible to find dual maximizers $\hat{\phi}$ and $\hat{\psi}$ for (2) which are *integrable*, even if the cost function is the squared distance on $\mathbb{R}$ (see Example 4.5 below). However it is possible to find dual maximizers in a weaker sense for all Borel measurable cost functions which are $\mu \otimes \nu$-a.e. finite. Assume that π is a finite transport plan and that ϕ, ψ are arbitrary functions satisfying $\phi(x) + \psi(y) \leq c(x, y)$. While $\phi(x) + \psi(y)$ is not necessarily integrable with respect to $\mu \otimes \nu$, it can be integrated with respect to π (possibly assuming the value $-\infty$). Thus we may well define

$$J(\phi, \psi) = \int_{X \times Y} \big[\phi(x) + \psi(y)\big] \, d\pi(x, y). \tag{4}$$

The subsequent lemma shows that the notation $J(\phi, \psi)$ is justified in the sense that this definition does not depend on the particular choice of π.

Lemma 1.1. *Let X, Y be Polish spaces equipped with Borel probability measures μ, ν and let $c : X \times Y \to [0, \infty]$ be a Borel measurable cost function. Let $\pi, \tilde{\pi} \in \Pi(\mu, \nu)$ be finite transport plans and assume that $\phi : X \to [-\infty, \infty), \psi : Y \to [-\infty, \infty)$ are such that $\phi(x) + \psi(y) \leq c(x, y)$ holds π-almost surely as well as $\tilde{\pi}$-almost surely. Then*

$$\int_{X \times Y} \big[\phi(x) + \psi(y)\big] \, d\pi(x, y) = \int_{X \times Y} \big[\phi(x) + \psi(y)\big] \, d\tilde{\pi}(x, y). \tag{5}$$

Moreover there is a sequence $(\phi_n, \psi_n)_{n \geq 1}$ of bounded functions in $\Phi(\mu, \nu)$ such that $\lim_{n\to\infty} J(\phi_n, \psi_n) = \int_{X \times Y} \big[\phi(x) + \psi(y)\big] \, d\pi(x, y)$.

Using Lemma 1.1, Theorem 1 follows from Theorem 2 below. Theorem 2 is stronger in the sense that it guarantees that the supremum on the right side of (3) is in fact a maximum if $J(\phi, \psi)$ is defined as in (4).

Theorem 2. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $c : X \times Y \to [0, \infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that there exists a finite transport plan. Then there exist Borel measurable* dual maximizers *$\hat{\phi}, \hat{\psi}$, i.e. functions $\hat{\phi} : X \to [-\infty, \infty), \hat{\psi} : Y \to [-\infty, \infty)$ satisfying $\hat{\phi}(x) + \hat{\psi}(y) \leq c(x, y)$ for all $(x, y) \in X \times Y$ such that*

$$I_c = \inf\{I_c[\pi] : \pi \in \Pi(\mu, \nu)\} = J\big(\hat{\phi}, \hat{\psi}\big). \tag{6}$$

We point out that the hypothesis that c is $\mu\otimes\nu$-a.e. finite is crucial for the existence of dual optimizers (see Example 4.3). Even in the case of a continuous cost function $c: X\times Y\to[0,\infty]$, it is in general not possible to find dual maximizers although duality holds in this case (see Example 4.2).

In [AP03, Theorem 3.2] it is proved that in the case of a finite lower semi-continuous cost function there exist integrable dual maximizers $(\hat\phi,\hat\psi)\in\Phi(\mu,\nu)$, provided that

$$\mu\Big(\Big\{x\in X:\int_Y c(x,y)\,d\nu(y)<\infty\Big\}\Big)>0, \tag{7}$$

$$\nu\Big(\Big\{y\in Y:\int_X c(x,y)\,d\mu(x)<\infty\Big\}\Big)>0. \tag{8}$$

Their argument yields that functions $\hat\phi,\hat\psi$ which are dual maximizers in the sense of Theorem 2 are μ- resp. ν-integrable whenever (7) and (8) are satisfied.

1.2. **Dual maximizers and strong c-cyclical monotonicity.** Theorem 2 is also connected with the notion of strong c-cyclical monotonicity introduced in [ST08]. A transport plan π is *strongly c-cyclically monotone*[1] if there exist Borel measurable functions $\phi: X\to[-\infty,\infty), \psi: Y\to[-\infty,\infty)$ such that $\phi(x)+\psi(y)\le c(x,y)$ for all $(x,y)\in X\times Y$ and $\phi(x)+\psi(y)=c(x,y)$ for π-almost all $(x,y)\in X\times Y$. In [BGMS08] it is proved that in the case of a finite cost function, a transport plan is strongly c-cyclically monotone if and only if it is optimal. We want to point out that this is also a consequence of Theorem 2:

Corollary 1.2. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ,ν, that $c: X\times Y\to[0,\infty]$ is Borel measurable and $\mu\otimes\nu$-a.e. finite and that there exists a finite transport plan.*

1. *Let π be a finite transport plan and assume that $\phi: X\to[-\infty,\infty), \psi: Y\to[-\infty,\infty)$ witness that π is strongly c-cyclically monotone. Then $J(\phi,\psi)=I_c[\pi]$, thus π is an optimal transport plan and ϕ,ψ are dual maximizers.*
2. *Assume that $\hat\pi$ is an optimal transport plan, i.e. $I_c[\hat\pi]=I_c$. Then $\hat\pi$ is strongly c-cyclically monotone. In fact, this is witnessed by every pair $(\hat\phi,\hat\psi)$ of dual maximizers.*

Proof. Given a transport plan π and functions ϕ,ψ witnessing that π is strongly c-cyclically monotone, we have

$$I_c[\pi]=\int_{X\times Y}c(x,y)\,d\pi(x,y)=\int_{X\times Y}\big[\phi(x)+\psi(y)\big]\,d\pi(x,y)=J(\phi,\psi)\le I_c. \tag{9}$$

Thus $I_c[\pi]=I_c$, hence π is an optimal transport and ϕ,ψ are dual maximizers in the sense of Theorem 2.

Conversely assume that $\hat\pi$ is an optimal transport plan and that $\hat\phi,\hat\psi$ are dual maximizers. Then

$$0=I_c[\hat\pi]-J(\hat\phi,\hat\psi)=\int_{X\times Y}c(x,y)-\big[\hat\phi(x)+\hat\psi(y)\big]\,d\hat\pi(x,y). \tag{10}$$

[1] This notion is called strong c-monotonicity in [ST08, BGMS08]. We find it, however, more consistent with previous notations in the literature to call it strong c-cyclical monotonicity.

Hence $c(x,y) - \big[\hat\phi(x) + \hat\psi(y)\big] = 0$ for $\hat\pi$-almost all $(x,y) \in X \times Y$. Thus, $\hat\phi, \hat\psi$ witness that $\hat\pi$ is strongly c-cyclically monotone. $\square$

1.3. **Continuity of I_c.** The fact that Monge-Kantorovich duality holds for not necessarily finite lower semi-continuous cost functions is related to a certain continuity property of the mapping $c \mapsto I_c$ which is always satisfied in the lower semi-continuous setting.

If the cost function $c : X \times Y \to [0,\infty]$ is lower semi-continuous, there exists a sequence $(c_n)_{n\geq 1}$ of bounded continuous functions such that $c_n \uparrow c$. For each such sequence we have $I_{c_n} \uparrow I_c$. This can easily be derived from the fact that $\Pi(\mu,\nu)$ is weakly compact[2]: Pick for each $n \geq 1$ a transport plan π_n such that $I_{c_n}[\pi_n] \leq I_{c_n} + 1/n$. By passing to a subsequence if necessary we may assume that $(\pi_n)_{n\geq 1}$ converges weakly to some transport plan $\pi \in \Pi(\mu,\nu)$. Then

$$I_c \leq I_c[\pi] = \lim_{m\to\infty} \int c_m\, d\pi = \lim_{m\to\infty} \left(\lim_{n\to\infty} \int c_m\, d\pi_n \right) \tag{11}$$

$$\leq \lim_{m\to\infty} \left(\lim_{n\to\infty} \int c_n\, d\pi_n \right) = \lim_{n\to\infty} I_{c_n}. \tag{12}$$

Since $I_{c_n} \leq I_c$ it follows in fact that $I_{c_n} \uparrow I_c$. Observe that $I_c[\pi] = I_c$, i.e. π is a primal optimizer.

A direct consequence of this simple continuity result is that as soon as we have shown the relatively easy result that duality holds for bounded continuous functions, it already follows for an arbitrary lower semi-continuous function $c : X \times Y \to [0,\infty]$. To see this, pick a sequence of bounded continuous functions $c_n : X \times Y \to [0,\infty)$ and for each $n \geq 1$ a pair of integrable functions (ϕ_n, ψ_n) such that $J(\phi_n,\psi_n) \geq I_{c_n} - 1/n$ and $\phi_n(x) + \psi_n(y) \leq c_n(x,y)$ $\big(\leq c(x,y)\big)$ on $X \times Y$. Then

$$\sup_{n\geq 1} \{J(\phi_n, \psi_n\} \geq \lim_{n\to\infty} (I_{c_n} - 1/n) = I_c, \tag{13}$$

thus duality holds.

A similar continuity property holds in the case of a finite measurable cost function. However we do not know how continuity in this sense can be shown directly, instead we achieve it as a consequence of Theorem 1.

Corollary 1.3. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $c : X \times Y \to [0,\infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that there exists a finite transport plan. Then $I_{c\wedge n} \uparrow I_c$.*

Proof. Given $\varepsilon > 0$ there exist bounded functions $\phi : X \to \mathbb{R}, \psi : Y \to \mathbb{R}$ such that $J(\phi,\psi) > I_c - \varepsilon$ by Theorem 1 and Lemma 1.1. For all large enough n, we have $\phi(x) + \psi(y) \leq (c \wedge n)(x,y)$ on $X \times Y$. Thus $I_{c_n} \geq J(\phi,\psi) \geq I_c - \varepsilon$. $\square$

1.4. **Cost functions with negative values.** For notational convenience we have chosen to consider only non-negative cost functions, but this restriction is somewhat stronger than necessary. Theorem 1 does remain valid (and in fact so do our other results) in the setting of a cost function $c : X \times Y \to [-\infty,\infty]$, provided that c is $\mu\otimes\nu$-a.e. finite and that there exist integrable functions $a : X \to [-\infty,\infty), b : Y \to [-\infty,\infty)$ such that

$$a(x) + b(y) \leq c(x,y) \tag{14}$$

[2]This is a consequence of Prokhorov's Theorem, see [Vil05, p 56].

for all $(x, y) \in X \times Y$. This is an immediate consequence of Theorem 1 applied to the cost function $c(x, y) - a(x) - b(y)$[3].

2. c-cyclical monotonicity

A transport plan π is *c-cyclically monotone* if it is concentrated on a Borel set $\Gamma \subseteq X \times Y$ which is c-cyclically monotone in the sense that

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) - c(x_i, y_i) \geq 0. \tag{15}$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in \Gamma$. (Here we let $x_{n+1} = x_1$.) Heuristically, π is c-cyclically monotone if it cannot be enhanced by means of cyclical rerouting. Hence it is intuitively appealing (and obvious in the finite setting) that optimal transport plans are always c-cyclically monotone. In fact it can be shown that in the case of a Borel measurable cost function $c : X \times Y \to [0, \infty]$ every optimal transport plan is c-cyclically monotone, and that the two notions are equivalent if c is finitely valued ([BGMS08, Theorem 1]). (This equivalence is not true in general, as shown by a beautiful counterexample due to Ambrosio and Pratelli, [AP03, Example 3.5]. The connection between optimality and c-cyclical monotonicity was also studied in [GM96, AP03, Pra08, ST08].)

The concept of c-cyclical monotonicity is crucial for the Monge-Kantorovich duality. We shall review its connection with strong c-cyclical monotonicity. As indicated by the name, it is almost obvious to see that strong c-cyclical monotonicity implies c-cyclical monotonicity: Assume that ϕ, ψ witness that π is strongly c-cyclically monotone. Then π is concentrated on the set $\Gamma = \{(x, y) : c(x, y) = \phi(x) + \psi(y)\}$ and

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) - c(x_i, y_i) = \tag{16}$$

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) - \big[\phi(x_i) + \psi(y_i)\big] \geq \tag{17}$$

$$\sum_{i=1}^{n} \big[\phi(x_{i+1}) + \psi(y_i)\big] - \big[\phi(x_i) + \psi(y_i)\big] = 0 \tag{18}$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in \Gamma$. Less trivially, the subsequent proposition asserts that the strong version can be deduced from the usual one, provided that c is finitely valued. (This is in general not the case if c attains ∞ on a large set, see Example 4.3 below.)

Proposition 2.1. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $c : X \times Y \to [0, \infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that there exists a finite transport plan. Then every c-cyclically monotone transport plan π is strongly c-cyclically monotone.*

Proposition 2.1 can be proved using a well known construction given in [Rüs96], see also [Roc66], [Vil03, Chapter 2] and [AP03, Theorem 3.2]. Assume for notational

[3]Throughout this paper we use the convention $\infty - \infty = \infty$.

convenience that π is concentrated on a c-cyclically monotone set Γ which satisfies $p_X[\Gamma] = X$ and $p_Y[\Gamma] = Y$ and that c is finite on $X \times Y$. Then the definition

$$\phi(x) := \inf\left\{\sum_{i=1}^{n} \big[c(x_{i+1}, y_i) - c(x_i, y_i)\big] : (x_1, y_1), \dots, (x_n, y_n) \in \Gamma, x_{n+1} = x\right\} \tag{19}$$

$$\psi(y) := \inf_{x \in X} c(x, y) - \phi(x), \tag{20}$$

where $x_1 \in X$ is an arbitrary fixed point, yields functions witnessing that π is strongly c-cyclically monotone. Strictly speaking, it might be necessary to alter ϕ and ψ on appropriately chosen null sets to ensure that they are Borel measurable functions, but these are merely technical obstacles which we will not discuss at this point. Instead we shall below derive Proposition 2.1 rigorously from the more general result in Proposition 3.6.

The subsequent statement summarizes how c-cyclical monotonicity connects to the other concepts discussed so far.

Proposition 2.2. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $c : X \times Y \to [0, \infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that there exists a finite transport plan. Let $\hat\phi, \hat\psi$ be dual maximizers in the sense of (4). Then the set $\Gamma := \{(x, y) : \hat\phi(x) + \hat\psi(y) = c(x, y)\}$ is c-cyclically monotone.*

For any finite transport plan π the following conditions are equivalent.

(a) *π is concentrated on Γ.*
(b) *π is c-cyclically monotone.*
(c) *π is strongly c-cyclically monotone.*
(d) *π is optimal.*

Proof. To see that Γ is c-cyclically monotone, argue as in (16). As regards the equivalence of (a) to (d), (a) trivially implies (b). We have seen above that (b) and (c) are equivalent and (c) and (d) are equivalent by Corollary 1.2.

If π is an optimal transport plan, $\hat\phi, \hat\psi$ witness that π is strongly c-cyclically monotone by Corollary 1.2 (2) such that $\hat\phi(x) + \hat\psi(y) = c(x, y)$ holds for π-almost all $(x, y) \in X \times Y$. Hence (d) implies (a). □

3. Transports with subsidies

In the last section it was described how dual maximizers can be constructed starting from a c-cyclically monotone transport plan. However, in the absence of lower semi-continuity of c, there is no reason why one should find a transport plan which is supported by a c-cyclically monotone set, even in very regular situations as is shown by the subsequent easy example.

Example 3.1. Let $(X, \mu) = (Y, \nu)$ equal the unit interval $[0, 1]$ equipped with Lebesgue measure. Define the cost of moving x to y by $c(x, y) = (x - y)^2$ for $x \neq y$ and let $c(x, x) = 1$. Clearly it is possible find transport plans with arbitrarily small costs, but the infimum 0 is not attained. Thus there exists no optimal and hence no c-cyclically monotone transport plan.

Observe that dual maximizers exist; just set $\hat\phi \equiv \hat\psi \equiv 0$.

Our attempt to overcome this difficulty is to introduce a certain *subsidy function*. To explain this notion we take up the anecdotal interpretation (see [Vil05, Chapter

3]) where (X,μ) models the Parisien bakeries (i.e., croissant sellers) and (Y,ν) the Parisian cafés (i.e., croissant buyers) and $c(x,y)$ denotes the transport cost from bakery x to cafe y. To avoid technicalities we suppose that there are only finitely many cafés and bakeries in Paris and that $c(x,y)$ is finitely valued. (We are not sure to which degree this assumption corresponds to reality.) Suppose that π is the traditional way how the croissants are transported from the bakeries to the cafés and that, for whatever reason, the Parisian authorities want to maintain this transport way also in the future. If the difference α between the present costs $I_c[\pi]$ and the cheapest possible transport costs I_c is strictly positive and the authorities do not intervene, they should expect that market forces will sooner or later cause the transport scheme to switch from π to some other $\tilde{\pi}$ with lower total transport cost. Therefore they might try to introduce a subsidy system, where the transport of each croissant from x to y is subsidized by an amount $f(x,y)\in[0,\infty)$.

The aim of the Parisian authorities is to design the system $f(x,y)$ of subsidies in such a way that the daily total subsidies $F=\int_{X\times Y} f\,d\pi$ effectively paid are minimized under the constraint that the bakers and cafetiers have no rational incentive to change the traditional transport system π by cyclically rerouting their ways of transportation.

A moment's reflection reveals that a lower bound for the cost of subsidy is given by

$$F=\int_{X\times Y} f\,d\pi \geq \alpha = I_c[\pi]-I_c, \tag{LB}$$

and it will turn out that this lower bound is attained which should not be very surprising. In fact, there are (at least) two versions of the "no incentive to change" constraint:

$$\sum_{i=1}^{n} c(x_{i+1},y_i)-\big(c(x_i,y_i)-f(x_i,y_i)\big)\geq 0 \tag{W1}$$

for all $(x_1,y_1),\dots,(x_n,y_n)$ in the support of π or

$$\sum_{i=1}^{n} \big(c(x_{i+1},y_i)-f(x_{i+1},y_i)\big)-\big(c(x_i,y_i)-f(x_i,y_i)\big)\geq 0 \tag{S1}$$

for all $(x_1,y_1),\dots,(x_n,y_n)$ in the support of π.

An interpretation of the two requirements goes as follows. In (W1) the authorities make a "take it or leave it" proposal to the bakers and cafetiers: if you stick to the transport system π we pay the subsidies f, if not we pay nothing. Hence comparing the transport costs from x_i to y_i with the ones from x_{i+1} to y_i, we have to make sure that for every collection $(x_1,y_1),\dots,(x_n,y_n)$ in the support of π the subsidized costs $\sum_{i=1}^n c(x_i,y_i)-f(x_i,y_i)$ are less than or equal to the non-subsidized costs after rerouting $\sum_{i=1}^n c(x_{i+1},y_i)$, which amounts to (W1).

In the interpretation of (S1) the Parisian authorities behave in a less authoritarian way: they promise to pay the subsidies $f(x,y)$ independently of whether the bakers and cafetiers are obedient or not, which amounts to the constraint (S1).

In fact, and this seems somewhat surprising, the validity of (W1) implies that there exists a function f such that $F=\int_{X\times Y} f\,d\pi=\alpha$ and such that the subsequent constraint (S2) which is yet stronger than (S1) is satisfied:

$$\text{(S2)} \qquad \sum_{i=1}^{n} \big(c(x_{i+1}, y_i) - f(x_{i+1}, y_i)\big) - \big(c(x_i, y_i) - f(x_i, y_i)\big) \geq 0$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in X \times Y$ (not necessarily being in the support of π).

Replacing $(x_1, y_1), \ldots, (x_n, y_n)$ by $(x_n, y_{n-1}), (x_{n-1}, y_{n-2}), \ldots, (x_2, y_1), (x_1, y_n)$ one verifies that one must have *equality* in (S2) for all $(x_1, y_1), \ldots, (x_n, y_n) \in X \times Y$. This may be interpreted as follows: there is a subsidy function $f(x, y)$ with total subsidy payment $F = \int_{X\times Y} f\, d\pi = \alpha$ and such that, for the subsidized transport cost we have that $\int_{X\times Y} (c-f)\, d\tilde{\pi}$ is equal for *any* transport plan $\tilde{\pi} \in \Pi(\mu, \nu)$ and if (LB) holds true, one easily verifies that this value must equal I_c. In particular, the bakers and cafetiers have no incentive to change π as they are, in fact, indifferent between all the possible transports $\tilde{\pi} \in \Pi(\mu, \nu)$, if their goal is to minimize the total subsidized transport costs.

In order to prove the existence of a subsidy system $f(x, y)$ satisfying (S2) and (LB) we use the following constraint which strengthens (W1) in a similar way as (S2) strengthens (S1):

$$\text{(W2)} \qquad \sum_{i=1}^{n} c(x_{i+1}, y_i) - \big(c(x_i, y_i) - f(x_i, y_i)\big) \geq 0$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in X \times Y$.

We refrain from giving an intuitive interpretation of (W2). Rather we try to indicate on an intuitive level why all four versions of the constraint are equivalent when minimizing $F = \int_{X\times Y} f\, d\pi$. More precisely, suppose that there is a function $f(x, y)$ such that $\int_{X\times Y} f\, d\pi = \alpha$ satisfying (the weakest form of) constraint (W1) and let us show that there exists a function $\tilde{f}$ with $\int \tilde{f}\, d\tilde{\pi} = I_c[\tilde{\pi}] - I_c$ for any transport plan $\tilde{\pi}$. Hence $\tilde{f}$ satisfies the (strongest form of) constraint (S2).

To pass from (W1) to (W2) is a cheap shot: observe that (W1) as well as $\int_{X\times Y} f\, d\pi$ only pertain to values of f on the support of π. Hence we may alter f outside the support of π to be ∞ (in the case of finite X and Y we clearly may replace ∞ by a sufficiently large real number), which then trivially satisfies (W2). To alleviate notation we still denote by f the function satisfying (W2). To pass from (W2) to (S2) we observe the subsequent "sandwich" type result which seems interesting in its own right.

Denoting $\overline{c}(x, y) = c(x, y)$ and $\underline{c}(x, y) := c(x, y) - f(x, y)$ inequality (W2) may be written as

$$\text{(W3)} \qquad \sum_{i=1}^{n} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) \geq 0$$

for all $(x_1, y_1), \ldots, (x_n, y_n) \in X \times Y$.

We shall show that (W3) implies (at least in our present setting of finite spaces X and Y and a finite cost function $c = \overline{c}$) that we may find functions $\phi(x), \psi(y)$ such that

$$\text{(W3a)} \qquad \overline{c}(x, y) \geq \phi(x) + \psi(y) \geq \underline{c}(x, y),$$

for all $(x, y) \in X$.

To motivate why this result should indeed be considered as a sandwhich theorem note the easy fact that a function $c(x,y)$ on $X \times Y$ may be written as

$$c(x,y) = \phi(x) + \psi(y) \tag{21}$$

if and only if it satisfies

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) - c(x_i, y_i) = 0 \tag{22}$$

for all $(x_1, y_1), \dots, (x_n, y_n) \in X \times Y$. The problem under which conditions a function defined on the product of two sets can be decomposed as the sum of two univariate functions is studied in detail in [BL92].

A precise version of our sandwich theorem under the assumption that $\overline{c}, \underline{c}$ are Borel measurable and $\overline{c}$ is $\mu \otimes \nu$-a.e. finite is given in Proposition 3.6. In our present situation it guarantees the existence of functions ϕ, ψ satisfying

$$c(x,y) \geq \phi(x) + \psi(y) \geq c(x,y) - f(x,y). \tag{W3b}$$

for all $(x,y) \in X \times Y$. This allows us to define

$$\tilde{f}(x,y) := c(x,y) - \big(\phi(x) + \psi(y)\big). \tag{23}$$

The lower bound of (W3b) implies that $f(x,y) \geq c(x,y) - \big(\phi(x) + \psi(y)\big) = \tilde{f}(x,y)$, in particular the desired bound for the total cost of subsidy $\int_{X\times Y} \tilde{f}\, d\pi \leq \int_{X\times Y} f\, d\pi = \alpha$ holds true. The subsidized cost function $(c - \tilde{f})$ is of the form

$$(c - \tilde{f})(x,y) = \phi(x) + \psi(y) \tag{24}$$

and hence satisfies the (strongest form of) constraint (S2).

We have observed above that (S2) implies that $I_c = \int_{X\times Y} (c - \tilde{f})\, d\tilde{\pi}$ for any finite transport plan $\tilde{\pi} \in \Pi(\mu, \nu)$. Thus $\int_{X\times Y} (c - \tilde{f})\, d\tilde{\pi} = \int_{X\times Y} \phi(x) + \psi(y)\, d\tilde{\pi}(x,y) = J(\phi, \psi)$ yields that ϕ, ψ are dual optimizers.

Finally, note that as a consequence of the fact that the total costs $\int_{X\times Y} \tilde{f}\, d\pi$ of our subsidy system cannot be less than α together with the point-wise inequality $f \geq \tilde{f}$ implies that f and $\tilde{f}$ coincide on the support of π.

3.1. **Existence of subsidy functions.** After the previous heuristic arguments we now pass to a more rigorous analysis. In order to find dual maximizers, we will first prove that there exists a subsidy function f which satisfies (S2) for a given finite transport plan π. Note that in the subsequent Proposition we do not assume that c is $\mu \otimes \nu$-a.e. finitely valued.

Proposition 3.2. *Let X, Y be Polish spaces equipped with Borel probability measures μ, ν. Let $c : X \times Y \to [0, \infty]$ be Borel measurable, assume that π is a finite transport plan and set $\alpha = I_c[\pi] - I_c \geq 0$. Then there exists a function $f : X \times Y \to [0, \infty]$ such that $\int f\, d\pi = \alpha$ and, for all $(x_1, y_1), \dots, (x_n, y_n) \in X \times Y$,*

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) + f(x_i, y_i) - c(x_i, y_i) \geq 0.^{4} \tag{25}$$

[4] We prefer to write $+f(x_i, y_i) - c(x_i, y_i)$ rather then $-(c(x_i, y_i) - f(x_i, y_i))$ in view of our convention $\infty - \infty = \infty$.

The main ingredient in the proof Proposition 3.2 is the following duality theorem due to Kellerer (see [Kel84, Lemma 1.8(a), Corollary 2.18] and [Kel85, p 212]).

Theorem 3 (Kellerer)**.** *Let $Z_1, \dots, Z_n, n \geq 2$ be Polish spaces equipped with Borel probability measures $\pi_1, \dots, \pi_n$ and assume that $C : Z = Z_1 \times \dots \times Z_n \to [-\infty, \infty)$ is Borel measurable and that $b = \sup_Z C$ is finite. Set*

$$I_C := \inf \left\{ \int_Z C \; d\kappa : \kappa \in \Pi(\pi_1, \dots, \pi_n) \right\}, \tag{26}$$

$$S_C := \sup \left\{ \sum_{i=1}^n \int_{Z_i} \phi_i \, d\pi_i : C(z_1, \dots, z_n) \geq \sum_{i=1}^n \phi_i(z_i) \right\}, \tag{27}$$

where $\phi_1, \dots, \phi_n$ are Borel functions taking values in $[-\infty, \infty]$.
Then $I_C = S_C$.

We will use it in the following form:

Corollary 3.3. *Let Z be a Polish space equipped with a Borel probability measure π. Let $e : Z^n \to [0, \infty]$ be a Borel measurable function such that $e(z_1, z_2, \dots, z_{n-1}, z_n) = e(z_2, z_3, \dots, z_n, z_1)$ for all $z_1, \dots, z_n \in Z$ and let $\alpha \geq 0$. Assume that*

$$\sup \left\{ \int_{Z^n} e \, d\kappa : \kappa \in \Pi(\pi, \dots, \pi) \right\} \leq n\alpha. \tag{28}$$

Then, for $\delta > 0$, there exists a function $f : Z \to [0, \infty]$ such that $e(z_1, \dots, z_n) \leq f(z_1) + \dots + f(z_n)$ and $\int f \, d\pi < \alpha + \delta$.

Proof. Applying Kellerer's Theorem to the function $C = -e$, we find functions $f_1, \dots, f_n : Z \to [0, \infty]$ such that $e(z_1, \dots, z_n) \leq f_1(z_1) + \dots + f_n(z_n)$ and $\int f_1 \, d\pi + \dots + \int f_n \, d\pi < n(\alpha + \delta)$. Set $f(z) = (f_1(z) + \dots + f_n(z))/n$. Then

$$e(z_1, \dots, z_n) = \frac{1}{n} \sum_{k=0}^{n-1} e(\sigma^k(z_1, \dots, z_n)) \tag{29}$$

$$\leq \frac{1}{n} \sum_{k=0}^{n-1} f_1(z_{1+k}) + \dots + f_n(z_{n+k}) \tag{30}$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} f_1(z_{1+k}) + \dots + f_n(z_{1+k}) = \sum_{k=0}^{n-1} f(z_{1+k}), \tag{31}$$

where $\sigma(r_1, \dots, r_n) = (r_2, \dots, r_{n+1})$. □

Lemma 3.4. *Let X, Y be Polish spaces equipped with Borel probability measures μ, ν, let $c : X \times Y \to [0, \infty]$ be Borel measurable and assume that π is a finite transport plan satisfying $I_c[\pi] \leq I_c + \alpha$. Set*

$$e(x_1, y_1, \dots x_n, y_n) := \Big(\sum_{i=1}^n c(x_{i+1}, y_i) - c(x_i, y_i) \Big)_- \geq 0. \tag{32}$$

Then

$$\sup \left\{ \int_{(X \times Y)^n} e \, d\kappa : \kappa \in \Pi(\pi, \dots, \pi) \right\} \leq n\alpha. \tag{33}$$

Proof. Denote by $\sigma, \tau : (X \times Y)^n \to (X \times Y)^n$ the mappings

$$\sigma : (x_i, y_i)_{i=1}^n \mapsto (x_{i+1}, y_{i+1})_{i=1}^n \tag{34}$$
$$\tau : (x_i, y_i)_{i=1}^n \mapsto \quad (x_i, y_{i+1})_{i=1}^n. \tag{35}$$

Observe that $\sigma^n = \tau^n = \mathrm{Id}_{(X \times Y)^n}$ and that σ and τ commute. By $p_i : (X \times Y)^n \to X \times Y, (x_1, y_1, \ldots, x_n, y_n) \mapsto (x_i, y_i)$ we denote the projection on the i-th component of the product, while the projections $p_X : X \times Y \to X, (x, y) \mapsto x$ and $p_Y : X \times Y \to Y, (x, y) \mapsto y$ are defined as above.

Pick $\kappa \in \Pi(\pi, \ldots, \pi)$. By replacing κ with

$$\frac{1}{n}\big(\kappa + \sigma_{\#}\kappa + \ldots + (\sigma)^{n-1}{}_{\#}\kappa\big), \tag{36}$$

we may assume that κ is σ-invariant. Set $B = \{e > 0\}$ and consider the restriction of κ to B defined by $\tilde{\kappa}(A) = \kappa(A \cap B)$ for Borel sets $A \subseteq (X \times Y)^n$. $\tilde{\kappa}$ is σ-invariant since both the measure κ and the set B are σ-invariant. Denote the marginal of $\tilde{\kappa}$ in the first coordinate $(X \times Y)$ of $(X \times Y)^n$ by $\tilde{\pi}$. Due to σ-invariance we have

$$p_{i\#}\tilde{\kappa} = p_{i\#}(\sigma_{\#}\tilde{\kappa}) \tag{37}$$
$$= (p_i \circ \sigma)_{\#}\tilde{\kappa} = p_{i+1\#}\tilde{\kappa}, \tag{38}$$

i.e. all marginals coincide and we have $\tilde{\kappa} \in \Pi(\tilde{\pi}, \ldots, \tilde{\pi})$. Furthermore, since $\tilde{\kappa} \leq \kappa$, the same is true for the marginals, i.e. $\tilde{\pi} \leq \pi$. Denote the marginal $p_{1\#}(\tau_{\#}\tilde{\kappa})$ of $\tau_{\#}\tilde{\kappa}$ in the first coordinate $(X \times Y)$ of $(X \times Y)^n$ by $\tilde{\pi}_\beta$. As σ and τ commute, $\tau_{\#}\tilde{\kappa}$ is σ-invariant, so the marginals in the other coordinates coincide with $\tilde{\pi}_\beta$. Note that $\tilde{\pi}_\beta(X \times Y) = \tilde{\kappa}\big((X \times Y)^n\big) = \tilde{\pi}(X \times Y)$. Moreover $\tilde{\pi}$ and $\tilde{\pi}_\beta$ have the same marginals in X resp. Y. Indeed, let $C \subseteq X, D \subseteq Y$ be Borel sets. Then

$$p_{X\#}\tilde{\pi}_\beta(C) = \tau_{\#}\tilde{\kappa}((p_X \circ p_1 \circ \tau)^{-1}[C]) = \tag{39}$$
$$= \tau_{\#}\tilde{\kappa}\{(x_1, \ldots, y_n) : x_1 \in C\} = \tag{40}$$
$$= \tilde{\kappa}\{(x_1, \ldots, y_n) : x_1 \in C\} = p_{X\#}\tilde{\pi}(C) \tag{41}$$
$$p_{Y\#}\tilde{\pi}_\beta(D) = \tau_{\#}\tilde{\kappa}((p_Y \circ p_1 \circ \tau)^{-1}[D]) = \tag{42}$$
$$= \tau_{\#}\tilde{\kappa}\{(x_1, \ldots, y_n) : y_1 \in D\} = \tag{43}$$
$$= \tilde{\kappa}\{(x_1, \ldots, y_n) : y_2 \in D\} = p_{Y\#}\tilde{\pi}(D). \tag{44}$$

This enables us to define an improved transport plan by

$$\pi_\beta = (\pi - \tilde{\pi}) + \tilde{\pi}_\beta. \tag{45}$$

Since $\tilde{\pi} \leq \pi$, we have that $(\pi - \tilde{\pi})$ is a positive measure, hence (45) defines a positive measure as well. Since $\tilde{\pi}$ and $\tilde{\pi}_\beta$ have the same total mass, π_β is a probability measure. Furthermore $\tilde{\pi}$ and $\tilde{\pi}_\beta$ have the same marginals in X, resp. Y, so π_β is indeed a transport plan. It remains to apply the assumption that the transport costs of π_β cannot be cheaper by more than α than the ones of π:

$$\alpha \geq I_c[\pi] - I_c[\pi_\beta] \tag{46}$$

$$= \int_{X\times Y} c\, d(\tilde\pi - \tilde\pi_\beta) \tag{47}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{(X\times Y)^n} c\circ p_i\, d(\tilde\kappa - \tau_{\#}\tilde\kappa) \tag{48}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{(X\times Y)^n} c(x_i,y_i) - c(x_{i+1},y_i)\, d\tilde\kappa(x_1,\dots y_n) \tag{49}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{B} c(x_i,y_i) - c(x_{i+1},y_i)\, d\kappa(x_1,\dots y_n) \tag{50}$$

$$= \frac{1}{n}\int_B e(x_1,\dots,y_n)\, d\kappa \tag{51}$$

$$= \frac{1}{n}\int_{(X\times Y)^n} e(x_1,\dots,y_n)\, d\kappa. \tag{52}$$

Since $\kappa\in\Pi(\pi,\dots,\pi)$ was arbitrary, this yields (33). □

In the proof of Proposition 3.2 we shall apply a result of Komlós ([Kom67]).

Lemma 3.5. *Let $(f_n)_{n\geq1}$ be a sequence of measurable $[0,\infty]$-valued functions on a probability space (Z,π) such that $\sup_{n\geq1}\|f_n\|_1<\infty$. Then there exists a subsequence $(\tilde f_n)_{n\geq1}$ such that the functions*

$$\frac{1}{n}(\tilde f_1+\ldots+\tilde f_n),\quad n\geq 1 \tag{53}$$

converge π-a.e. to a function taking values in $[0,\infty]$.

In particular, there exist functions $g_n\in\operatorname{conv}(f_n,f_{n+1},\dots)$ such that $(g_n)_{n\geq1}$ converges π-a.e..

Proof. The first part of Lemma 3.5 is Komlós' original result which we will not prove. The assertion that there exist $g_n\in\operatorname{conv}(f_n,f_{n+1},\dots)$ such that $(g_n)_{n\geq1}$ converges almost everywhere[5] is a simple consequence which we will derive for the sake of completeness. Assume that $\frac{1}{n}(\tilde f_1+\ldots+\tilde f_n)$ converges π-a.e. to a function g. Since all functions $f_n, n\geq1$ are π-a.e. finitely valued, for each $k\geq1$ there exists some n_k such that

$$\pi\left(\left\{\frac{1}{n\,_k}(\tilde f_1+\ldots+\tilde f_k)>\frac{1}{k}\right\}\right)<\frac{1}{2^k}. \tag{54}$$

Set $g_k=\frac{1}{n_k}(\tilde f_{k+1}+\ldots+\tilde f_{n_k+k})$. □

Proof of Proposition 3.2. Combining Lemma 3.4 and Corollary 3.3 we achieve that for each $n\geq2$ there exists a function f_n such that $\int f_n\,d\pi\leq\alpha+1/n$ and for all $(x_1,y_1),\dots,(x_n,y_n)\in X\times Y$

$$\sum_{i=1}^{n} c(x_{i+1},y_i)+f_n(x_i,y_i)-c(x_i,y_i)\geq 0. \tag{55}$$

[5] In fact, this result holds true without *any* integrability assumptions, see [DS94, Lemma A1.1].

Observe that for all $p \geq 1$ the function f_{np} satisfies (55) as well since we can run through the cycle

$$(x_1, y_1) \to \ldots \to (x_n, y_n) \to (x_1, y_1) \to \ldots \tag{56}$$

p times. Also note that any convex combination and any pointwise limit of functions which satisfy (55) for some fixed n satisfies (55) (for the same n) as well. Thus we may apply Lemma 3.5 to find functions $g_n \in \operatorname{conv}(f_{n!}, f_{(n+1)!}, \ldots)$ which converge π-almost everywhere. Defining g as the pointwise limit where this limit exists and ∞ elsewhere, yields a function which satisfies (55) for every n. Moreover

$$\int g \, d\pi = \int \liminf_{n\to\infty} g_n \, d\pi \leq \liminf_{n\to\infty} \int g_n \, d\pi = \alpha, \tag{57}$$

hence the desired bound for the total costs of subsidy holds true. □

3.2. **Dual maximizers by subsidized transport plans.** The main goal of this section is to prove the sandwich-type result announced above.

Proposition 3.6. *Assume that X, Y are Polish spaces equipped with Borel probability measures μ, ν, that $\overline{c} : X \times Y \to (-\infty, \infty]$ is Borel measurable and $\mu \otimes \nu$-a.e. finite and that $\underline{c} : X \times Y \to [-\infty, \infty)$ is Borel measurable. If*

$$\sum_{i=1}^{n} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) \geq 0 \tag{58}$$

for all $x_1, \ldots, x_n \in X$, $y_1, \ldots, y_n \in Y$, there exist Borel measurable functions $\phi : X \to [-\infty, \infty), \psi : Y \to [-\infty, \infty)$ and Borel sets $X' \subseteq X, Y' \subseteq Y$ of full measure such that

$$\underline{c}(x, y) \leq \phi(x) + \psi(y) \leq \overline{c}(x, y), \tag{59}$$

where the lower bound holds for $x \in X', y \in Y'$ and the upper bounded is valid for all $x \in X, y \in Y$.

Observe that Proposition 3.6 is in fact a generalization of Proposition 2.1: Let $\pi \in \Pi(\mu, \nu)$ be a finite transport plan which is concentrated on a c-cyclically monotone Borel set Γ. Without loss of generality assume that c is finite on Γ. Set $\overline{c} = c$ and

$$\underline{c}(x, y) := \begin{cases} c(x, y) & \text{if } (x, y) \in \Gamma \\ -\infty & \text{else} \end{cases}. \tag{60}$$

Then $\overline{c}, \underline{c}$ satisfy the assumptions of Proposition 3.6 and thus there exist Borel measurable functions ϕ, ψ which satisfy $\phi(x) + \psi(y) \leq c(x, y)$ for all $x \in X, y \in Y$ and $\phi(x) + \psi(y) \geq c(x, y)$ for π-almost all $(x, y) \in X \times Y$, hence π is strongly c-cyclically monotone.

Before giving the proof of Proposition 3.6 we need some preliminaries, in particular we will recapitulate some facts from the theory of analytic sets. They will be needed to deal with certain measurability issues which arise in the course of the proof.

Let X be a Polish space. A set $A \subseteq X$ is *analytic* if there exist a Polish space Z, a Borel measurable function $f : Z \to X$ and a Borel set $B \subseteq Y$ such that $f(B) = A$.

Lemma 3.7. *Let X, Z be Polish spaces and $g : X \times Z \to [-\infty, \infty]$ a Borel measurable function. Set*

$$\phi(x) = \inf_{z \in Z} g(x, z). \tag{61}$$

Then $\{\phi < \alpha\}$ is analytic for every $\alpha \in [\infty, \infty]$.

Proof. We have

$$\phi(x) < \alpha \iff \exists z, g(x, z) < \alpha. \tag{62}$$

Thus $\{\phi < \alpha\} = p_X[\{g(x, z) < \alpha\}]$. □

Given a Borel measure μ on X, we denote its completion by $\tilde{\mu}$. By a result of Luzin (see for instance [Kec95, Theorem 21.10]) every analytic set $A \subseteq X$ is the disjoint union of a Borel set and a $\tilde{\mu}$-null-set. This will allow us to replace a function which only satisfies that the sets $\{\phi < \alpha\}, \alpha \in [-\infty, \infty]$ are analytic by a Borel measurable function.

Lemma 3.8. *Let X be a Polish space and μ a finite Borel measure on X. If $\phi : X \to [-\infty, \infty]$ satisfies that $\{\phi(x) < \alpha\}$ is analytic for each $\alpha \in [-\infty, \infty]$, then there exists a Borel measurable function $\tilde{\phi} : X \to [-\infty, \infty]$ such that $\tilde{\phi} \leq \phi$ everywhere and $\phi = \tilde{\phi}$ almost everywhere with respect to $\tilde{\mu}$.*

Proof. Let $(I_n)_{n=1}^{\infty}$ be an enumeration of the intervals $[-\infty, \alpha), \alpha \in \mathbb{Q}$. Then for each $n \in \mathbb{N}$, $\phi^{-1}[I_n]$ is $\tilde{\mu}$-measurable and hence the union of a Borel set B_n and a $\tilde{\mu}$-null set N_n. Let N be a Borel null set which covers $\bigcup_{n=1}^{\infty} N_n$. Let $\tilde{\phi}(x) = \phi(x) - \infty \cdot \mathbf{1}_N(x)$. Clearly $\tilde{\phi}(x) \leq \phi(x)$ for all $x \in X$ and $\phi(x) = \tilde{\phi}(x)$ for $\tilde{\mu}$-almost all $x \in X$. Furthermore, $\tilde{\phi}$ is Borel measurable since $(I_n)_{n=1}^{\infty}$ is a generator of the Borel σ-algebra on $[-\infty, \infty)$ and for each $n \in \mathbb{N}$ we have that $\tilde{\phi}^{-1}[I_n] = B_n \setminus N$ is a Borel set. □

Lemma 3.9. *Assume that in the setting of Proposition 3.6 we have that*

$$\tilde{\mu}(p_X[\{\underline{c} > -\infty\} \setminus (X \times N)]) = 1 \tag{63}$$

for every ν-null-set $N \subseteq Y$. Then there exist $x_1 \in X$ and a Borel set $X' \subseteq X$ with $\mu(X') = 1$ such that for each $x \in X'$ there are $y_1, y \in Y$ satisfying

$$\underline{c}(x_1, y_1) > -\infty, \quad \overline{c}(x, y_1) < \infty, \tag{64}$$

$$\underline{c}(x, y) > -\infty, \quad \overline{c}(x_1, y) < \infty. \tag{65}$$

Proof. Set

$$X_1 = \{x : \nu(\{y : \overline{c}(x, y) < \infty\}) = 1\} \tag{66}$$

$$Y_1 = \{y : \mu(\{x : \overline{c}(x, y) < \infty\}) = 1\}. \tag{67}$$

By Fubini's Theorem, $\mu(X_1) = \nu(Y_1) = 1$. Since $\tilde{\mu}(p_X[\{\underline{c} > -\infty\} \cap (X \times Y_1)]) = 1$, there exist $x_1 \in X_1, y_1 \in Y_1$ such that $\underline{c}(x_1, y_1) > -\infty$. Since $x_1 \in X_1$,

$$Y' := \{y \in Y_1 : \overline{c}(x_1, y) < \infty\} \tag{68}$$

has ν-measure one. Consequently

$$X' := \{x : \overline{c}(x, y_1) < \infty\} \cap p_X[\{\underline{c} > -\infty\} \cap (X \times Y')] \tag{69}$$

has full $\tilde{\mu}$-measure. It remains to check that the assertions of Lemma 3.9 are satisfied. Choose $x \in X'$. By definition of X', $\overline{c}(x, y_1) < \infty$. Since $x \in p_X[\{\underline{c} >$

$-\infty\} \cap (X \times Y')]$, there exists some $y \in Y'$ such that $\underline{c}(x,y) > -\infty$. Since $y \in Y'$, $\overline{c}(x_1, y) < \infty$. □

Proof of Proposition 3.6. Note that it is sufficient to define ϕ and ψ on Borel sets $X' \subseteq X, Y' \subseteq Y$ with $\mu(X') = \nu(Y') = 1$, since they can then be extended to X and Y by setting them $-\infty$ on the null-sets $X \setminus X', Y \setminus Y'$. This will be used several times in the course of the proof.

Next we will show that it is sufficient to consider the case that $\tilde{\mu}(p_X[\{\underline{c} > -\infty\} \setminus (X \times N)]) = 1$ for every ν-null-set $N \subseteq Y$, such that Lemma 3.9 is applicable. Set

$$\beta := \inf\{\tilde{\mu}(p_X[\{\underline{c} > -\infty\} \setminus (X \times N)]) : N \subseteq Y, \nu(N) = 0\}. \tag{70}$$

Choose for each $k \geq 1$ a ν-null-set $N_k \subseteq Y$ such that

$$\beta + \frac{1}{k} \geq \tilde{\mu}(p_X[\{\underline{c} > -\infty\} \setminus (X \times N_k)]) \tag{71}$$

Set $N := \bigcup_{k=1}^{\infty} N_k, Y' = Y \setminus N$ and let $M \subseteq p_X[\{\underline{c} > -\infty\} \setminus (X \times N)]$ be a Borel set with $\mu(M) = \beta$. Then $\tilde{\mu}(M \cap p_X[\{\underline{c} > -\infty\} \setminus (X \times N')]) = \beta$ for every ν-null-set $N' \subseteq Y'$ and it is sufficient to define ϕ, ψ on M and Y'; as above they can then be extended to X and Y by setting them $-\infty$ on $X \setminus M$ resp. N. Thus we may assume without loss of generality that $M = X$ or, equivalently, that $p_X[\{\underline{c} > -\infty\} \setminus (X \times N)]$ has full measure for every ν-null-set $N \subseteq Y'$.

Choose x_1 and X' according to Lemma 3.9 and set

$$\phi_n(x) = \inf\left\{\sum_{i=1}^{n} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) : x_{n+1} = x\right\} \tag{72}$$

$$\text{and } \phi(x) = \inf_{n\geq 1} \phi_n(x). \tag{73}$$

Fix $x \in X'$. To see that $-\infty < \phi(x) < \infty$, pick $y, y_1 \in Y'$ according to Lemma 3.9 such that $\underline{c}(x_1, y_1) > -\infty, \overline{c}(x, y_1) < \infty, \underline{c}(x, y) > -\infty, \overline{c}(x_1, y) < \infty$. Then

$$\phi(x) \leq \phi_1(x) \leq \overline{c}(x, y_1) - \underline{c}(x_1, y_1) < \infty. \tag{74}$$

To prove the lower bound, we pick $n \geq 1$ and $x_2, \dots, x_n \in X', y_1, \dots, y_n \in Y'$ and set $x_{n+1} = x, y_{n+1} = y, x_{n+2} = x_1$. By (58),

$$\sum_{i=1}^{n+1} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) \geq 0 \tag{75}$$

$$\text{whence } \sum_{i=1}^{n} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) \geq \underbrace{-\big[\overline{c}(x_{n+2}, y_{n+1}) - \underline{c}(x_{n+1}, y_{n+1})\big]}_{=-\left[\overline{c}(x_1,y) - \underline{c}(x,y)\right]}. \tag{76}$$

Taking the infimum over all possible choices of $n \geq 1$ and $x_2, \dots, x_n \in X', y_1, \dots, y_n \in Y'$, we achieve that $\phi(x) \geq -\big[\overline{c}(x_1, y) - \underline{c}(x, y)\big] > -\infty$.

Next observe that for $x, x' \in X', y \in Y'$ and $n \geq 1$

$$\phi_{n+1}(x) \leq \inf\left\{\sum_{i=1}^{n+1} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) : x_{n+2} = x, x_{n+1} = x', y_{n+1} = y\right\} \tag{77}$$

$$= \phi_n(x') + \big[\overline{c}(x, y) - \underline{c}(x', y)\big] \tag{78}$$

Taking the infimum over $n \geq 1$ yields that

$$\phi(x) \leq \inf_{n\geq 1} \phi_{n+1}(x) \leq \phi(x') + \big[\overline{c}(x,y) - \underline{c}(x',y)\big] \tag{79}$$

$$\implies \qquad \underline{c}(x',y) - \phi(x') \leq \overline{c}(x,y) - \phi(x). \tag{80}$$

At this point we will take care about measurability of ϕ. First apply Lemma 3.7 to the spaces X' and $Z = \bigcup_{n=1}^{\infty}(Y' \times X)^n$ to see that $\{\phi < \alpha\}$ is analytic for each $\alpha \in [-\infty, \infty]$. Then we may shrink X' a little bit to achieve, by Lemma 3.8, that ϕ is even Borel measurable. Note that (78) is then still valid for all $x, x' \in X', y \in Y'$.

By Fubini's Theorem, for ν-almost all $y \in Y'$ there exists some $x \in X$ such that $\overline{c}(x,y) < \infty$. By shrinking Y' a little bit if necessary, we may assume that this is the case for all $y \in Y'$. Then the function

$$\psi(y) := \inf_{x\in X'} \overline{c}(x,y) - \phi(x) \tag{81}$$

is finitely valued on Y'. As above, we apply Lemma 3.7 and Lemma 3.8 and shrink Y' a little further to achieve that ψ is Borel measurable. Moreover

$$\underline{c}(x,y) \leq \phi(x) + \psi(y) \leq \overline{c}(x,y) \tag{82}$$

holds for all $x \in X', y \in Y'$ by (78) and (81). $\square$

Proof of Theorem 2. Pick an arbitrary finite transport plan $\pi \in \Pi(\mu,\nu)$. Choose a subsidy function $f : X \times Y \to [0,\infty]$ according to Proposition 3.2 such $\int f\, d\pi \leq I_c[\pi] - I_c$. Set $\overline{c}(x,y) = c(x,y)$ and

$$\underline{c}(x,y) := \begin{cases} c(x,y) - f(x,y) & \text{if } c(x,y) \neq \infty \\ -\infty & \text{if } c(x,y) = \infty \end{cases}, \tag{83}$$

in particular $\underline{c} = c - f$ holds π-a.e. since c is π-a.e. finite. For all $x_1, \ldots, x_n \in X$ and $y_1, \ldots, y_n \in Y$,

$$\sum_{i=1}^{n} \overline{c}(x_{i+1}, y_i) - \underline{c}(x_i, y_i) \geq \tag{84}$$

$$\sum_{i=1}^{n} c(x_{i+1}, y_i) + f(x_i, y_i) - c(x_i, y_i) \geq 0, \tag{85}$$

Thus by Proposition 3.6, there exist functions $\phi : X \to [-\infty, \infty)$, $\psi : Y \to [-\infty, \infty)$ such that $\phi(x) + \psi(y) \leq \overline{c}(x,y) = c(x,y)$ for all $x \in X, y \in Y$ and $\phi(x) + \psi(y) \geq \underline{c}(x,y) = c(x,y) - f(x,y)$ for π-almost all $(x,y) \in X \times Y$. This implies

$$I_c \geq J(\phi,\psi) \geq \int c(x,y) - f(x,y)\, d\pi(x,y) \geq I_c[\pi] - \big(I_c[\pi] - I_c\big) = I_c, \tag{86}$$

thus $I_c = J(\phi,\psi)$ and hence ϕ, ψ are dual maximizers. $\square$

To conclude the proof of Theorem 1 we show Lemma 1.1.

Proof of Lemma 1.1. Similarly as in [ST08] we define

$$\begin{aligned} \phi_n(x) &= (-n \vee \phi(x)) \wedge n, \\ \psi_n(y) &= (-n \vee \psi(y)) \wedge n, \\ \xi_n(x,y) &= \phi_n(x) + \psi_n(y), \\ \xi(x,y) &= \phi(x) + \psi(y), \end{aligned}$$

for $(x,y) \in X \times Y$ and $n \geq 1$. Observe that $\xi_n \uparrow \xi$ on $\{\xi \geq 0\}$ and $\xi_n \downarrow \xi$ on $\{\xi \leq 0\}$, as $n \to \infty$. Moreover $\xi_n \leq c$ so that $(\phi_n, \psi_n) \in \Phi(\mu,\nu)$.

Additionally, $\int \xi \, d\pi \leq \int c \, d\pi, \int \xi \, d\tilde{\pi} < \int c \, d\tilde{\pi}$ exist, taking possibly the value $-\infty$, since $\xi \leq c$ holds π-almost surely as well as $\tilde{\pi}$-almost surely and we assume that $\int c \, d\pi, \int c \, d\tilde{\pi} < \infty$. By the assumption on equal marginals of π and $\tilde{\pi}$ we obtain

$$\int \xi_n \, d\pi = \int \phi_n \, d\pi + \int \psi_n \, d\pi \tag{87}$$

$$= \int \phi_n \, d\tilde{\pi} + \int \psi_n \, d\tilde{\pi} \tag{88}$$

$$= \int \xi_n \, d\tilde{\pi}, \tag{89}$$

for $n \geq 0$, hence

$$\int \xi_n 1_{\{\xi \geq 0\}} + \xi_n 1_{\{\xi \leq 0\}} \, d\pi = \int \xi_n 1_{\{\xi \geq 0\}} + \xi_n 1_{\{\xi \leq 0\}} \, d\tilde{\pi}. \tag{90}$$

By our previous considerations we can pass to the limits and obtain $\int \xi \, d\pi = \int \xi \, d\tilde{\pi}$. Indeed the limits are monotone on $\{\xi \geq 0\}$ and on $\{\xi \leq 0\}$ and the convergence is dominated by c on $\{\xi \geq 0\}$. Hence the limits of $\int \xi_n \, d\pi = \int \xi_n \, d\tilde{\pi}$ exist as $n \to \infty$ and are equal. Consequently $\int \xi \, d\pi = \int \xi \, d\tilde{\pi}$. □

4. Examples

We start with a simple example which shows that Monge-Kantorovich duality does not hold in general for a measurable cost function $c : X \times Y \to [0,\infty]$.

Example 4.1. Let $X = Y = [0,1]$, $\mu = \nu$ the Lebesgue-measure and set

$$c(x,y) = \begin{cases} \infty & \text{for } x < y \\ 1 & \text{for } x = y \\ 0 & \text{for } x > y \end{cases} \tag{91}$$

for $(x,y) \in X \times Y$. The optimal (and in fact the only finite) transport plan π is concentrated on the diagonal and yields costs of one. Assume that $\phi, \psi : [0,1] \to [-\infty,\infty)$ are integrable functions satisfying $\phi(x)+\psi(y) \leq c(x,y)$ for all $x,y \in [0,1]$. Then

$$\int \phi \, d\mu + \int \psi \, d\nu = \lim_{\alpha \downarrow 0} \int_0^{1-\alpha} \phi(x+\alpha) + \psi(x) \, dx \tag{92}$$

$$\leq \lim_{\alpha \downarrow 0} \int_0^{1-\alpha} c(x+\alpha, x) \, dx = 0. \tag{93}$$

Thus there is a duality gap, i.e.,

$$1 = I_c > \sup\{J(\phi,\psi) : (\phi,\psi) \in \Phi(\mu,\nu)\} = 0. \tag{94}$$

Note also that I_c fails to be continuous in the sense of Corollary 1.3: For each $n \geq 1$, there exist transport plans assigning arbitrarily small costs to the function $c \wedge n$ such that $\lim_{n\to\infty} I_{c\wedge n} = 0 \neq 1 = I_c$.

In [BGMS08, Example 5.3] a certain variant of Example 4.1 is considered. By setting $c(x,y) = \sqrt{x-y}$ for $x > y$, the cost function becomes lower semi-continuous. In this case duality does hold true, but there are no dual maximizers, that is, the optimal transport plan π is not strongly c-cyclically monotone. We present here yet another variant of Example 4.1 which shows that dual maximizers need not exist, even if the cost function is assumed to be continuous.

Example 4.2. Let $X = Y = \mathbb{N} \cup \{\omega\}$ where we take ω to be a "number" larger than all $n \in \mathbb{N}$. Equip X and Y with the discrete topology and define $\mu = \nu$ such that positive measure is assigned to each point in X, resp. Y. Set

$$c(x,y) = \begin{cases} \infty & \text{for } x < y \\ 1 & \text{for } x = y \\ 0 & \text{for } x > y \end{cases} \tag{95}$$

for $(x,y) \in X \times Y$. As above we find that the only finite transport plan $\hat{\pi}$ is concentrated on the diagonal and yields costs of 1. Since X and Y are discrete spaces, c is continuous with respect to the product topology on $X \times Y$ and hence duality holds true. Striving for a contradiction, we assume that there exist dual maximizers $\hat{\phi}, \hat{\psi} : \mathbb{N} \cup \{\omega\} \to [-\infty, \infty)$. Note that $\hat{\phi}$ and $\hat{\psi}$ are necessarily finitely valued since X and Y have no non trivial null-sets. Since $\hat{\phi}, \hat{\psi}$ witness that $\hat{\pi}$ is strongly c-monotone we have

$$\hat{\phi}(n) + \hat{\psi}(n) = 1, \hat{\phi}(n+1) + \hat{\psi}(n) \le 0, \text{ and } \hat{\phi}(\omega) + \hat{\psi}(n) \le 0 \tag{96}$$

for each $n \in \mathbb{N}$. This yields

$$\hat{\phi}(n+1) \le \hat{\phi}(n) - 1, \text{ and } \hat{\phi}(\omega) \le \hat{\phi}(n) - 1 \tag{97}$$

for all $n \in \mathbb{N}$ which is impossible for a finitely valued function.

One can try to overcome the difficulties encountered in Examples 4.1, 4.2 and [BGMS08, Example 5.3] by admitting dual optimizers from a larger class of mappings: Consider functions $\phi : X \to [-\infty, \infty), \psi : Y \to [-\infty, \infty)$ which do not necessarily satisfy the inequality $\phi(x) + \psi(y) \le c(x,y)$ for *all* $(x,y) \in X \times Y$ but do satisfy the potentially weaker condition

$$\phi(x) + \psi(y) \le c(x,y), \ \pi\text{-a.e. for every } \textit{finite} \text{ transport plan } \pi \in \Pi(\mu,\nu). \tag{98}$$

It follows then that $J(\phi,\psi) = \int_{X\times Y} \phi(x) + \psi(y)\, d\pi(x,y) \le I_c[\pi]$ for each finite transport plan $\pi \in \Pi(\mu,\nu)$, such that

$$J(\phi,\psi) \le \inf_{\pi \in \Pi(\mu,\nu)} I_c[\pi] = I_c. \tag{99}$$

Therefore it is reasonable to consider all pairs of functions (ϕ,ψ) satisfying (98) as admissible solutions of the dual part of the Monge-Kantorovich problem. In particular a transport plan π is optimal provided that there exist measurable (not necessarily integrable) functions $\phi : X \to [-\infty, \infty), \psi : Y \to [-\infty, \infty)$ satisfying (98) and

$$\int_{X\times Y} \big[\phi(x) + \psi(y)\big]\, d\pi(x,y) = \int_{X\times Y} c\, d\pi. \tag{100}$$

Clearly, (100) is tantamount to requiring that $\phi(x) + \psi(y) = c(x,y)$ for π-almost all (x,y). Observe that dual optimizers in this weak sense exist in Examples 4.1 and 4.2. Since the only finite transport plan is the optimal one, we may just take $\phi \equiv \psi \equiv 1/2$. However the subsequent construction (which is a variation of [AP03, Example 3.5]) shows that, in general, dual optimizers do not even exist in this weak sense.

Example 4.3. Let $X = Y = [0, 1]$, equipped with Lebesgue measure $\lambda = \mu = \nu$. Pick $\alpha \in [0, 1)$ irrational. Set

$$\Gamma_0 = \{(x, x) : x \in X\}, \quad \Gamma_1 = \{(x, x \oplus \alpha) : x \in X\},$$

where $\oplus$ is addition modulo 1. Define $c : X \times Y \to [0, \infty]$ by

$$c(x, y) = \begin{cases} 0 & \text{for } (x, y) \in \Gamma_0, x \in [0, 1/2] \\ 2 & \text{for } (x, y) \in \Gamma_0, x \in (1/2, 1) \\ 1 & \text{for } (x, y) \in \Gamma_1 \\ \infty & \text{else} \end{cases}. \tag{101}$$

Note that c is lower semi-continuous whence duality holds (cf. [Kel84, Theorem 2.2]). For $i = 0, 1$, let π_i be the obvious transport plan concentrated on Γ_i. Then all finite transport plans are given by convex combinations of the form $\varrho\pi_0 + (1 - \varrho)\pi_1, \varrho \in [0, 1]$ and each of these transport plans leads to costs of 1. Assume that $\hat{\phi}, \hat{\psi} : [0, 1) \to [-\infty, \infty)$ are measurable functions which satisfy

(1) $J(\hat{\phi}, \hat{\psi}) = 1$,
(2) $\hat{\phi}(x) + \hat{\psi}(y) \leq c(x, y)$ for π_0- and π_1-almost all (x, y).

This implies that, in fact, $\hat{\phi}(x) + \hat{\psi}(y) = c(x, y)$ for π_0- and π_1-almost all (x, y). Thus

$$\hat{\phi}(x) + \hat{\psi}(x) = \begin{cases} 0 & \text{for } x \in [0, 1/2] \\ 2 & \text{for } x \in (1/2, 1) \end{cases}, \quad \hat{\phi}(x) + \hat{\psi}(x \oplus \alpha) = 1 \tag{102}$$

$$\text{whence} \quad \hat{\psi}(x \oplus \alpha) = \begin{cases} \hat{\psi}(x) + 1 & \text{for } x \in [0, 1/2] \\ \hat{\psi}(x) - 1 & \text{for } x \in (1/2, 1) \end{cases} \tag{103}$$

for all $x \in I$, where $I \subseteq [0, 1]$ is a Borel set of measure 1. By passing to a subset of full measure, we may additionally assume that $I \oplus \alpha = I$. Pick a set $B \subseteq I$ such that $\lambda(B) > 0$ and $\sup_{x \in B} \hat{\psi}(x) - \inf_{x \in B} \hat{\psi}(x) < 1$. By a classic theorem of Steinhaus, $B - B$ contains a non-empty open set. Since α is irrational, $\{(2n + 1)\alpha : n \in \mathbb{N}\}$ is dense in $(0, 1]$, thus there exists $n \in \mathbb{N}$ such that $x, x \oplus (2n + 1)\alpha \in B$. By (103) this implies that $|\hat{\psi}(x \oplus (2n + 1)\alpha) - \hat{\psi}(x)| \geq 1$, contradicting the definition of B.

We conclude with two examples which show that in general it is not possible to find dual maximizers $\hat{\phi}$ and $\hat{\psi}$ which are integrable even if c is continuous and finitely valued.

Example 4.4. Let $X = Y = (0, 1]$, take μ, ν to be the Lebesgue measure and set $c(x, y) = |1/x - 1/y + 1|$. Define π to be the transport plan concentrated on the diagonal such that $I_c[\pi] = 1$. The functions $\hat{\phi}(x) = 1/x, \hat{\psi}(y) = 1 - 1/y$ witness that π is strongly c-cyclically monotone and hence optimal. Recall that by Corollary 1.2 functions that witness strong c-cyclical monotonicity of a transport plan correspond precisely to dual maximizers. In particular, $\hat{\phi}$ and $\hat{\psi}$ are dual maximizers. Of course $\hat{\phi}$ and $\hat{\psi}$ are not integrable. Let (ϕ, ψ) be another pair of dual maximizers. We will see that there is a constant $\beta \in \mathbb{R}$ such that $\hat{\phi} = \phi + \beta, \hat{\psi} = \psi - \beta$ almost surely, hence there exist no integrable dual maximizers.

Fix $a, b \in (0, 1]$ such that $a < b$ and $(1/x - 1/y + 1)$ is positive on $[a, b]^2$. Let $\pi_{a,b}$ be the transport plan which equals π on $X \times Y \setminus [a, b]^2$ and is $1/(b - a)$ times the product measure on $[a, b]^2$. As above, $\hat{\phi}$ and $\hat{\psi}$ witness that $\pi_{a,b}$ is strongly

c-cyclically monotone and thus optimal. But then also ϕ, ψ witness that $\pi_{a,b}$ is strongly c-cyclically monotone, hence

$$\phi(x) + \psi(y) = c(x, y) = \hat{\phi}(x) + \hat{\psi}(y) \tag{104}$$

for almost all $(x, y) \in [a, b]^2$. This yields

$$(\phi - \hat{\phi})(x) = (\hat{\psi} - \psi)(y), \tag{105}$$

thus both sides are almost everywhere in $[a, b]$ equal up to some constant $\beta_{a,b}$. Since we can cover $(0, 1]$ with sufficiently small overlapping intervals $[a_n, b_n], n \geq 1$, we achieve that this constant $\beta_{a,b} =: \beta$ does in fact not depend on the choice of a and b. It follows that indeed $\phi = \hat{\phi} + \beta$ and $\psi = \hat{\psi} - \beta$ almost everywhere.

In fact, one can find an example displaying the same phenomenon as Example 4.4 where c is just the squared distance on $\mathbb{R}$.

Example 4.5. Let $X = Y = \mathbb{R}$, let $c(x, y) = (x - y)^2$, let μ be a Borel probabilty measure on X, define $T : X \to Y$ by $T(x) = x + 1$ and assume that $\nu = T_{\#}\mu$. Let $\pi \in \Pi(\mu, \nu)$ be the obvious transport plan concentrated on the graph $\Gamma = \{(x, x + 1) : x \in \mathbb{R}\}$ of T. Set $\hat{\phi}(x) = -2x, \hat{\psi}(y) = 2y - 1$. Then

$$c(x, y) - \hat{\phi}(x) - \hat{\psi}(y) = (x - y)^2 - 2(x - y) + 1 = (x - y - 1)^2 \tag{106}$$

is non negative for all $(x, y) \in X \times Y$ and $\phi(x) + \psi(y) = c(x, y)$ holds precisely for $(x, y) \in \Gamma$. Thus π is strongly c-cyclically monotone and the corresponding transport costs $\int_\Gamma 1 \, d\pi = 1$ are minimal.

We claim that the dual optimizers $\hat{\phi}, \hat{\psi}$ are essentially unique if the Lebesgue measure λ is absolutely continuous with respect to μ.

Let ϕ, ψ be dual optimizers. Since $J(\phi, \psi) = \int_{X \times Y} \phi(x) + \psi(y) \, d\pi = 1$, we have $\phi(x) + \psi(y) = c(x, y) = 1$ for π-almost all $(x, y) \in X \times Y$. Thus there is a set $X' \subseteq X$ with $\lambda(X \setminus X') = \mu(X \setminus X') = 0$ such that

$$\phi(x) + \psi(1 + x) = c(x, x + 1) = 1 \implies \psi(1 + x) = 1 - \phi(x) \tag{107}$$

for all $x \in X'$. Since $\phi(x) + \psi(1 + y) \leq c(x, 1 + y) = (x - y - 1)^2$ for all $x, y \in \mathbb{R}$, (107) yields that for all $x \in X, y \in X'$

$$\phi(x) + 1 - \phi(y) \leq (x - y - 1)^2 \tag{108}$$

$$\implies \big(\phi(x) + 2x\big) - \big(\phi(y) + 2y\big) \leq (x - y - 1)^2 - 2(x - y - 1) + 1. \tag{109}$$

Setting $f(x) = \phi(x) + 2x$, (109) is tantamount to

$$f(x) - f(y) \leq (x - y)^2 \tag{110}$$

for all $x \in X, y \in X'$. Thus f is constant on X', whence there exists a constant $\beta \in \mathbb{R}$ such that

$$\phi(x) = -2x + \beta = \hat{\phi}(x), \quad \psi(y) = 2y - 1 + \beta = \hat{\psi}(y). \tag{111}$$

for λ- as well as μ-almost all $x \in X$ and λ- as well as ν-almost all $y \in Y$.

If we pick μ such that λ is absolutely continuous with respect to μ and such that $\int_{\mathbb{R}} x \, d\mu(x)$ does not exist, $\hat{\phi}$ and $\hat{\psi}$ are not integrable and by (111) no other pair of dual maximizers can be integrable either.

Acknowledgment. *The authors are grateful to Martin Goldstern, Gabriel Maresch and Josef Teichmann for many helpful discussions on the topic of this paper.*

References

[AP03] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In *Optimal transportation and applications (Martina Franca, 2001)*, volume 1813 of *Lecture Notes in Math.*, pages 123–160. Springer, Berlin, 2003.

[BGMS08] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport. *arXiv:0802.0646v1*, 2008.

[BL92] J. M. Borwein and A. S. Lewis. Decomposition of multivariate functions. *Canad. J. Math.*, 44(3):463–482, 1992.

[DS94] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300(3):463–520, 1994.

[GM96] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[Kec95] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[Kel84] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984.

[Kel85] H. G. Kellerer. Duality theorems and probability metrics. In *Proceedings of the seventh conference on probability theory (Braşov, 1982)*, pages 211–220, Utrecht, 1985. VNU Sci. Press.

[Kom67] J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.*, 18:217–229, 1967.

[Pra08] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 258(3):677–690, 2008.

[Roc66] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17:497–510, 1966.

[RR95] D. Ramachandran and L. Rüschendorf. A general duality theorem for marginal problems. *Probab. Theory Related Fields*, 101(3):311–319, 1995.

[RR96] D. Ramachandran and L. Rüschendorf. Duality and perfect probability spaces. *Proc. Amer. Math. Soc.*, 124(7):2223–2228, 1996.

[RR98] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[Rüs96] L. Rüschendorf. On c-optimal random variables. *Statist. Probab. Lett.*, 27(3):267–270, 1996.

[ST08] W. Schachermayer and J. Teichmann. Characterization of optimal transport plans for the Monge-Kantorovich problem. *Proceedings of the A.M.S.*, 2008. in press.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil05] C. Villani. Optimal transport, old and new. Lecture Notes, 2005.

Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien
Wiedner Hauptstrasse 8-10/104, 1040 Wien, Austria
E-mail address: `mathias.beiglboeck@tuwien.ac.at`

Institut für Wirtschaftsmathematik, Technische Universität Wien
Wiedner Hauptstrasse 8-10/105, 1040 Wien, Austria
E-mail address: `walter.schachermayer@tuwien.ac.at`