A LAND OF MONOTONE PLENTY

MATHIAS BEIGLBÖCK AND CLAUS GRIESSLER

Abstract. A fundamental concept in optimal transport is c-cyclical monotonicity: it allows to link the optimality of transport plans to the geometry of their support sets. Recently, related concepts have been successfully applied in the multi-marginal version of the transport problem as well as in the martingale transport problem which arises from model-independent finance.

We establish a unifying concept of *c-monotonicity / finitistic optimality* which describes the geometric structure of optimizers of a generalized moment problem (GMP). This allows us to strengthen known results in martingale optimal transport and for a mass transport problem with a continuum of marginals.

If the optimization problem can be formulated as a multi-marginal transport problem, potentially with additional linear constraints, our contribution is parallel to a recent result of Zaev.

Keywords: cyclical monotonicity, mass transport, moment problem.
MSC (2010): Primary 60G42, 60G44; Secondary 91G20.

1. Introduction

1.1. **Motivation from optimal transport.** Consider the Monge-Kantorovich transport problem for probabilities μ and ν on Polish spaces X and Y, cf. [Vil03, Vil09]. The set $\Pi(\mu,\nu)$ of transport plans consists of all measures on $X \times Y$ with X-marginal μ and Y-marginal ν. Associated to a cost function $c : X \times Y \to \mathbb{R}_+$ and $\gamma \in \Pi(\mu,\nu)$ are the transport costs $\int c\, d\gamma$. The Monge-Kantorovich problem is to determine the value

$$\inf\left\{\int c\, d\gamma : \gamma \in \Pi(\mu,\nu)\right\} \tag{OT}$$

and to identify an *optimal* transport plan $\gamma^* \in \Pi(\mu,\nu)$, minimizing of (OT).

A fundamental concept in the theory of optimal transport is *c-cyclical monotonicity* which leads to a *geometric characterization* of optimal couplings. Its relevance for the theory of optimal transport has been fully recognized by Gangbo and McCann [GM96], based on earlier work of Knott and Smith [KS84] and Rüschendorf [Rüs91, Rüs95] among others.

We postpone precise definitions and just mention that, heuristically, a transport plan is c-cyclically monotone if it cannot be improved by means of cyclical rerouting, i.e. by replacing the transfers

$$x_1 \to y_1, x_2 \to y_2, \dots, x_n \to y_n$$

with

$$x_1 \to y_2, x_2 \to y_3, \dots, x_n \to y_1.$$

Date: February 27, 2015.
We acknowledge financial support through FWF-projects P21209 and P26736.

Connecting optimality and c-cyclical monotonicity is technically intricate. A series of contributions ([AP03, Pra08, ST08, BGMS09, BC10] among others) led to the following clear cut characterization:

Theorem 1.1. *Let $c : X \times Y \to [0, \infty)$ be Borel measurable and assume that $\gamma \in \Pi(\mu, \nu)$ is a transport plan with finite costs $\int c \, d\gamma \in \mathbb{R}_+$. Then γ is optimal if and only if γ is c-cyclically monotone.*

To see whether a transport behaves optimally on a finite number of points is often an elementary and feasible task. However, it is difficult to relate this to the transport between diffuse distributions since single points do not carry positive mass. Theorem 1.1 provides the required remedy to this obstacle: it connects the optimization problem for measures with optimality on a "pointwise" level.

1.2. **Aims of this article.** Several modifications of the classical optimal transport problem have received interest in the literature. We mention a few to which we will come back in more detail in the following section. First, a natural extension is the multi-marginal transport problem where not just two but finitely many marginals are prescribed, see e.g. [Kel84, Car03, Pas11, Pas12, KP13]. Pass [Pas13a, Pas13b] considered the problem with a continuum of marginals prescribed. Furthermore, martingale versions of the transport problem have recently been attracting considerable amounts of research ([BHP13, GHLT14, BJ14, HT13, DM13, DS13] among others). The latter development has been motivated by applications in model-independent finance.

Given the importance of c-cyclical monotonicity it is natural to look for a related concept applicable in these variants of the transport problem. Kim and Pass [KP13] introduced a notion of c-monotonicity, necessary for optimality in the context of the multi-marginal transport problem ([KP13, Proposition 2.3]). They use it to develop a general condition on the cost function which is sufficient to imply existence of a Monge solution and uniqueness results in the multi-marginal optimal transport problem. In [BJ14], the authors introduced a concept of "finitistic optimality" which mimics c-cyclical monotonicity in the case of the 2-period martingale transport problem. A monotonicity principle in the spirit of Theorem 1.1 (cf. [BJ14, Lemma 2.1]) then links finitistic optimality with optimality overall. This allows to determine optimal martingale transport plans in a number of instances.

The main goal of this article is to unify these notions and to make them applicable to the above mentioned variants of the transport problem. The framework we use is the one of a generalized moment problem (GMP). The generalized moment problem constitutes a classical problem in probability, see for instance [Kem68] and [Las10]. We introduce a version of finitistic optimality / c-monotonicity for this problem and establish a "monotonicity principle" (Theorem 2.4) saying that finitistic optimality is necessary for optimality overall. Whereas it has long been known that (OT) is an instance of (GMP), the optimality criterion of c-cyclical monotonicity given in Theorem 2.4 is new to the best of our knowledge.

We highlight a particular novelty of the approach in this article: in all the instances where the monotonicity principle was previously known, the minimization problem (GMP) admits a well understood dual problem and it is known that there is no duality gap (*strong duality*). In the transport literature it is well known that the absence of a duality gap can be used to show that optimal transport plans are cyclically monotone, see e.g. [Vil03, Exercise 2.38]. In fact, assuming certain regularity assumptions, this argument could be used to establish Theorem 2.4 whenever there is no duality gap. The advantage of the approach presented below is twofold. On the one hand it allows to derive the desired implication virtually without regularity assumptions. More importantly, it is applicable also in situations where duality is either unknown or known to fail (cf. [AN87, Section 3.4] for such cases).

Theorem 2.4 allows to obtain improved versions of the results from [BJ14] and [KP13, Proposition 2.3], and it includes one implicaton of the classical result stated in Theorem 1.1. To exemplify the result's applicability beyond optimization on finite products of spaces, we prove a strengthened version of Pass' Monge-type result for a continuum of marginals [Pas13a]. Pass uses his result to derive an infinite dimensional rearrangement inequality and upper bounds on solutions to parabolic PDE among other applications. In contrast to Pass' original result we do not require additional assumptions on the payoff functional resp. the prescribed marginals which seems convenient for these applications.

2. Formulation of the problem and the optimality criterion

2.1. **The basic optimization problem.** Throughout this article we assume that E is a Polish space and $c : E \to \mathbb{R}$ a Borel measurable cost function. Examples are $E = M^2$, where M is a Riemannian manifold, $E = \left(\mathbb{R}^d\right)^n$, or $E = C[0,T]$, the space of continuous functions $[0,T] \to \mathbb{R}$ with the topology of uniform convergence.

We fix a set $\mathcal{F}$ of Borel-measurable functions on E and write $\Pi_{\mathcal{F}}$ for the set of probability measures γ on E for which $\int f \, d\gamma = 0$ for all $f \in \mathcal{F}$.[1] We deal with the following generalized moment problem: minimizing the total cost choosing from $\Pi_{\mathcal{F}}$, i.e.

$$\min_{\gamma \in \Pi_{\mathcal{F}}} \int c \, d\gamma. \tag{GMP}$$

We give a list of some specific problems that can be posed this way. For a product of spaces $\prod_{i \in I} X_i$ we write p_{X_i} or in short p_i for the projection onto X_i.

2.2. **Classical optimal transport and its multi-marginal version.** To fit the classical Monge-Kantorovich problem (OT) into (GMP), take $E = X \times Y$. To test whether a measure γ is a transport plan in $\Pi(\mu, \nu)$, it is sufficient to verify that

$$\int \varphi(x) \, d\gamma(x,y) = \int \varphi(x) \, d\mu(x), \quad \int \psi(y) \, d\gamma(x,y) = \int \psi(y) \, d\nu(y)$$

[1]By asserting that $\int f \, d\gamma = 0$ we implicitly understand that this integral exists.

for all continuous bounded functions $\varphi : X \to \mathbb{R}, \psi : Y \to \mathbb{R}$. Hence, with

$$\mathcal{F}_1 = \left\{ \varphi \circ p_X - \int \varphi \, d\mu, \psi \circ p_Y - \int \psi \, d\nu : \varphi \in C_b(X), \psi \in C_b(Y) \right\}$$

problem (GMP) is equivalent to (OT).

Analogously for the multi-marginal optimal transport problem: here one considers

$$\inf \left\{ \int c \, d\gamma : \gamma \in \Pi(\mu_1, \dots, \mu_n) \right\}, \tag{1}$$

where $\mu_1, \dots, \mu_n$ are probability measures on Polish spaces $X_1, \dots, X_n$, the set $\Pi(\mu_1, \dots, \mu_n)$ consists of the probability measures γ on $E = X_1 \times \ldots \times X_n$ with $p_i(\gamma) = \mu_i$ for $i = 1, \dots, n$, and we can set

$$\mathcal{F}_2 = \left\{ \varphi \circ p_i - \int \varphi \, d\mu_i : \varphi \in C_b(X_i), 1 \le i \le n \right\}. \tag{2}$$

2.3. **Optimal transport in the continuum marginal case.** In [Pas13a, Pas13b], Pass considers an optimal transport problem with a continuum of marginals prescribed. Specifically, in [Pas13a] the following problem was posed: for $I = [0, T]$, given a family $(\mu_t)_{t \in I}$ of probability measures on $\mathbb{R}$ and a strictly concave function $h : \mathbb{R} \to \mathbb{R}$, determine

$$\inf_{\gamma \in \Pi_C(\mu_t)} \int h\Big(\int_0^T f(t) \, dt \Big) \, d\gamma(f), \tag{B}$$

where $\Pi_C(\mu_t)$ denotes the set of probabilities on $C[0, T]$ with marginals $(\mu_t)_{t \in I}$. The family $(\mu_t)_{t \in I}$ can be assumed to be weakly continuous, as otherwise there is no measure on $C(I)$ with these marginals. Hence, denoting by $q_t : (0, 1) \to \mathbb{R}$ the quantile function[2] of μ_t, the path of an x-quantile evolves continuously over time t. Therefore the map $q : (0, 1) \to C[0, T]$, $x \mapsto q_.(x)$ pushes forward Lebesgue measure λ from $(0, 1)$ to a measure π^* uniformly distributed on the quantile paths of (μ_t).

Pass establishes that π^* is the unique minimizer of (B). He then lists several surprising applications from parabolic equations to mathematical finance and quantum physics. Among other conditions, Pass assumes that the quantile functions satisfy a property of uniform Riemann-integrability which may be difficult to verify in practice. We will see below that it is possible to dispose of this restriction.

To view the problem as an instance of (GMP), set $E = C[0, T]$ and

$$\mathcal{F}_3 = \left\{ \varphi \circ p_t - \int \varphi \, d\mu_t : \varphi \in C_b(\mathbb{R}), t \in [0, T] \right\}. \tag{3}$$

We use Theorem 2.4 to establish the following strengthened version of Pass' main result:

Theorem 2.1. *Let $h : \mathbb{R} \to \mathbb{R}$ be concave and $(\mu_t)_{t \in I}$ a family of probability measures on $\mathbb{R}$, weakly continuous in t and such that*

$$\int_0^T \int |x| \, d\mu_t(x) \, dt < \infty, \quad \textit{and} \quad \int |h| \, d\mu_t < \infty \quad \textit{for all } t \in [0, T].$$

[2]I.e. q_t is the generalized inverse of the cumulative distribution function of μ_t: $q_t(x) = \inf\{y : \mu_t\big((-\infty, y]\big) \ge x\}$

Then π^* *is a minimizer of* (B)*. If the infimum in* (B) *is finite and* h *is strictly concave, then* π^* *is the unique minimizer.*

2.4. **Model-independent finance – Martingale Transport.** For a general overview we refer to the survey of Hobson [Hob11]. Recent contributions on the general theory in discrete time include [ABPS13, HT13, BN13]. Here $E = \mathbb{R}_+^n$ or $\mathbb{R}^n$, and any n-tuple $(x_1, \ldots, x_n)$ is interpreted as a possible evolution of the stock price at future dates $t_1 < t_2 < \ldots < t_n$. A possible price of a "path-dependent option" with payoff $c : E \to \mathbb{R}$ is then calculated as an integral

$$\int c \, d\gamma. \tag{4}$$

Model-independent finance is about determining the minimal (or maximal) possible prices subject to appropriate constraints, i.e. about optimizing (4) over a suitable class of probabilities γ.

According to the martingale pricing paradigm in mathematical finance the measures of interest are *martingale measures*, i.e. probability measures γ such that the coordinate process on $\mathbb{R}^n$ is a martingale (in its own filtration) with respect to γ. Equivalently, γ is a martingale measure iff for each $l < n$, and each continuous bounded function $\varphi : \mathbb{R}^l \to \mathbb{R}$ one has equality and real values in

$$\int x_{l+1} \, \varphi(x_1, \ldots, x_l) \, d\gamma = \int x_l \, \varphi(x_1, \ldots, x_l) \, d\gamma.$$

In applications one has information about the current value $\xi \in \mathbb{R}$ of the stock price, so the expectation of all the marginals of any martingale representing the stock price should equal ξ. This leads us to consider the family of functions

$$\mathcal{F}^{(m)} = \{p_1 - \xi\} \cup \left\{ (p_{l+1} - p_l) \, \left(\varphi \circ p_{\{1,\ldots,l\}}\right) : \varphi \in C_b\left(\mathbb{R}^l\right), 1 \le l < n \right\}. \tag{5}$$

The martingale condition (with expectation ξ) then corresponds to $\int f \, d\gamma = 0$ for all $f \in \mathcal{F}^{(m)}$.

There usually is additional information derived from market-data, again corresponding to $\int f \, d\gamma = 0$ for f in some family of functions $\mathcal{H}$. For $\mathcal{H}$ we list some choices of particular interest: $\mathcal{H} = \emptyset$ is not relevant for mathematical finance but more so in probability through its connection to martingale inequalities: we refer to [ABP$^+$13, BS13, BN13, BN14] for recent developments in this direction. A noteworthy result of Bouchard and Nutz [BN13] is that *every* martingale inequality in finite discrete time can be derived from a "dual", elementary and deterministic inequality.

Provided that European call options on the underlying stock are liquidly traded, it is a reasonable idealization to assume that the marginal distributions of the stock price at particular time instances are known from market data. The case of a given marginal distribution at terminal time t_n has been particularly intriguing.[3] In the present context this corresponds to

[3] This case is naturally connected to the Skorokhod embedding problem, we refer to the survey of Obłój [Obł04].

$p_n(\gamma) = \mu$ for some probability μ, i.e. specifying

$$\mathcal{H} = \left\{ \varphi \circ p_n - \int \varphi \, d\mu : \varphi \in C_b(\mathbb{R}) \right\}. \tag{6}$$

More recently also the case with all intermediate marginals given has been considered. This corresponds to $\mathcal{H} = \mathcal{F}_2$ (where $X_1 = \ldots = X_n = \mathbb{R}$).

Summing up, the principal problem of model independent finance boils down to (GMP) with $\mathcal{F}_4 = \mathcal{F}^{(m)} \cup \mathcal{H}$.

2.5. **A monotonicity principle for martingale optimal transport.** Our motivation for a generalized optimality criterion stems from a characterization of optimizers in the martingale transport problem in [BJ14]. A short discussion follows below, but first recall the definition of *c-cyclical monotonicity*: a set $\Gamma \subseteq X \times Y$ is called c-cyclically monotone if for $(x_1, y_1), \ldots, (x_l, y_l) \in \Gamma$, one always has, setting $y_{l+1} = y_1$,

$$\sum_{i=1}^{l} c(x_i, y_i) \leq \sum_{i=1}^{l} c(x_i, y_{i+1}).$$

A transport plan is called c-cyclically monotone if it is concentrated on a c-cyclically monotone set. An equivalent way of defining cyclical monotonicity of Γ is: each measure α, which is finite and concentrated on finitely many elements of Γ, is a cost minimizing transport between its marginals. I.e., for each such measure α, if α' has the same marginals then

$$\int c \, d\alpha \leq \int c \, d\alpha'.$$

The equivalence of the definitions follows easily from e.g. [AGS08, Thm. 6.1.4]

In [BJ14] this notion was adapted by adding a martingale component[4]: for a measure α on $\mathbb{R}^2$, a measure α' is called a *competitor* if

(1) α and α' have the same marginals, and
(2) $\int x_2 \, d\alpha_{x_1}(x_2) = \int x_2 \, d\alpha'_{x_1}(x_2)$ holds $p_1(\alpha)$-almost surely (i.e., the difference $\alpha - \alpha'$ has the martingale property).

A set $\Gamma \subseteq \mathbb{R}^2$ is then called *finitely optimal* if each measure α, which is finite and concentrated on finitely many elements of Γ, is cost-minimizing amongst its competitors. And a measure is called finitely optimal if it is concentrated on a finitely optimal set. The Variational Principle in [BJ14] states that optimality implies finite optimality provided that c satisfies certain moment conditions, and that the converse holds if c is additionally continuous and bounded.

Note that for finitely supported finite measures, the conditions (1) and (2) above are equivalent to (1) and (2'):

(2') for each bounded Borel-measurable φ, we have

$$\int (x_2 - x_1) \, f(x_1) \, d\alpha(x_1, x_2) = \int (x_2 - x_1) \, f(x_1) \, d\alpha'(x_1, x_2).$$

[4]The article [BJ14] is concerned with the case $E = \mathbb{R}^2$ where the minimization is taken over all transport plans which are martingale measures, i.e. the setup described in the last part of Section 2.4, resp. $\mathcal{F} = \mathcal{F}^{(m)} \cup \mathcal{F}_2$ in the optimization problem (GMP).

Now (1) and (2') can be written in short as

$$\int f\,d\alpha = \int f\,d\alpha',\ \forall f \in \mathcal{F}^{(m)} \cup \mathcal{F}_2.$$

This condition leads us to:

2.6. **A general concept of finitistic optimality and main result.**

Definition 2.2. For a measure α on the Polish space E and a set $\mathcal{F}$ of measurable functions $E \to \mathbb{R}$, a competitor of α is a measure α' on E such that $\alpha(E) = \alpha'(E)$, and for all $f \in \mathcal{F}$ one has

$$\int f\,d\alpha = \int f\,d\alpha'. \tag{7}$$

A set $\Gamma \subseteq E$ is called *finitely minimal / c-monotone* if each measure α, which is finite and concentrated on finitely many atoms in Γ, is cost minimizing amongst its competitors. A measure γ is called *finitely minimal/ c-monotone* if it is concentrated on a finitely minimal / c-monotone set.

Our goal is to establish that optimizers of the problem (GMP) are finitely minimal. To this end we require the following assumption on the family $\mathcal{F}$:

Assumption 2.3. *(1) There exists a function $g : E \to [0, \infty)$ such that each element of $\mathcal{F}$ is bounded by some multiple of g. I.e., for each $f \in \mathcal{F}$ there is a constant $a_f \in \mathbb{R}_+$ such that $|f| \leq a_f g$.*
(2) All functions in $\mathcal{F}$ are continuous, or $\mathcal{F}$ is at most countable.

Note that these properties are satisfied in all examples listed above.

Theorem 2.4. *Let E be a Polish space and $c : E \to \mathbb{R}$ a Borel measurable function. Let $\mathcal{F}$ be a family of Borel-measurable functions on E satisfying Assumption 2.3 and assume that γ^* is a minimizer of the problem*

$$\min_{\gamma \in \Pi_{\mathcal{F}}} \int c\,d\gamma$$

and that $\int c\,d\gamma^* \in \mathbb{R}$. Then γ^* is finitely minimal / c-monotone.

In applications one usually works with continuous or lower semi-continuous cost functions. In that case the existence of an optimizer γ^* can often be established by compactness arguments. However, (semi-)continuity does not simplify our arguments nor does it lead to a more specific result. For instance, in classical optimal transport one obtains a nicer result for the most relevant case in which c is continuous: the support of an optimal transport plan is c-cyclically monotone. But this assertion need not be true in our setup. Juillet [Jui14] gives an example of a two-period martingale transport problem in which the marginals μ, ν are compactly supported, the cost function $c(x, y) = (y - x)^3$ is continuous, the minimizer is unique and its support is not finitely optimal. We have therefore chosen to go with the general formulation above.

2.7. **A counterexample to sufficiency.** It is natural to ask whether the converse of Theorem 2.4 holds true as well, i.e. if finite optimality is also sufficient for optimality overall, at least under additional regularity assumptions on the function c and the underlying spaces. This is not the case as shown by the following counterexample in the context of transport plans which are invariant under group actions (see e.g. [KZ13]).

Example 2.5. Let $X = Y = (0,1)$, and $\mu = \nu = \lambda$. For some irrational number $\xi > 0$, let $T : (0,1) \to (0,1)$ denote the operator $x \mapsto x \oplus \xi$ (addition modulo 1). We want to minimize the cost $c(x,y) = (y-x)^2$ among the transport plans π that are $T \otimes T$-invariant, i.e. the transport plans π for which $\pi = T \otimes T(\pi)$. These transport plans are characterized as those for which

$$\int h\big(T \otimes T\big)\, d\pi = \int h\, d\pi \quad \text{for all } h \in C_b(X \times Y).$$

The unique minimizer here is the uniform distribution on the diagonal, but each other transport plan is also concentrated on a finitely minimal set, as each subset of $X \times Y$ is finitely minimal: every finite and finitely supported α is its only competitor. For a competitor α', the signed measure $\alpha - \alpha'$ is $T \otimes T$-invariant, and hence a continuous measure. The only finitely supported such measure is zero, hence $\alpha = \alpha'$.

2.8. **Connection with** [Zae14]**.** In independent work, Zaev [Zae14] obtains (among a number of further developments) a result which is related to Theorem 2.4. His article is concerned with the multi-marginal transport problem described in Section 2.2, allowing for additional linear constraints. In our notation this corresponds to problem (GMP) on a set E which is a product $X_1 \times \ldots \times X_n$ of Polish probability spaces and where $\mathcal{F}$ is a superset of the set $\mathcal{F}_2$ defined in (6); several important extensions of the transport problem can be phrased in this form. Under continuity and (weak) integrability assumptions Zaev establishes the existence of an optimizer, a version of the classical Monge-Kantorovich duality as well as a necessary geometric condition for optimizers. The latter statement is equivalent to the assertion of Theorem 2.4 (applied to the setup of [Zae14]). The proof given in [Zae14] is based on his duality result and different from the approach pursued in this article.

3. Proof of Theorem 2.4

In the proof of Theorem 2.4 we will make use of the following result from [BGMS09], which is a consequence of a duality result by Kellerer [Kel84]:

Lemma 3.1 ([BGMS09, Proposition 2.1])**.** *Let (E, m) be a Polish probability space, and M an analytic*[5] *subset of E^l, then one of the following holds true:*

(i) there exist m-null sets $M_1, \ldots, M_l \subseteq E$ such that $M \subseteq \bigcup_{i=1}^l p_i^{-1}(M_i)$, or

(ii) there is a measure η on E^l such that $\eta(M) > 0$ and $p_i(\eta) \leq m$ for $i = 1, \ldots, l$.

Proof of Theorem 2.4. Without loss of generality we assume that $|c| \leq g$. We want to find a finitely minimal set Γ with $\gamma^*(\Gamma) = 1$. To obtain this, it is sufficient to show that for each $l \in \mathbb{N}$ there is a set Γ_l with $\gamma^*(\Gamma_l) = 1$ such that: for any finite measure α concentrated on at most l points in Γ_l and

[5] [BGMS09, Proposition 2.1] is stated only for Borel sets, however the same proof applies in the case where M is analytic.

satisfying $\alpha(E) \leq 1$ as well as $\int g\,d\alpha \leq l$, there is no c-better competitor α' on at most l points and satisfying $\int g\,d\alpha' \leq l$. For then $\Gamma := \bigcap_{l\in\mathbb{N}} \Gamma_l$ is finitely minimal.

Hence, fix l and define a subset of E^l,

$$\begin{aligned}M = \{&(z_1, \dots, z_l) \in E^l : \\ &\exists \text{ a measure } \alpha \text{ on } E, \alpha(E) \leq 1, \textstyle\int g\,d\alpha \leq l, \operatorname{supp} \alpha \subseteq \{z_1, \dots, z_l\}, \\ &\text{s.t. there is a } c\text{-better competitor } \alpha', \alpha'(E) \leq 1, \textstyle\int g\,d\alpha' \leq l, |\operatorname{supp} \alpha'| \leq l\}.\end{aligned}$$

Note that M is the projection of the set

$$\begin{aligned}\hat{M} =\Big\{&(z_1, \dots, z_l, \alpha_1, \dots, \alpha_l, z'_1, \dots, z'_l, \alpha'_1, \dots, \alpha'_l,) \in E^l \times \mathbb{R}^l_+ \times E^l \times \mathbb{R}^l_+ : \\ &\sum \alpha_i \leq 1, \sum \alpha_i g(z_i) \leq l, \sum \alpha'_i \leq 1, \sum \alpha'_i g(z'_i) \leq l, \sum \alpha_i = \sum \alpha'_i, \\ &\sum \alpha_i f(z_i) = \sum \alpha'_i f(z'_i) \text{ for all } f \in \mathcal{F}, \sum \alpha_i c(z_i) > \sum \alpha'_i c(z'_i)\Big\}.\end{aligned}$$

onto the first l coordinates. By our Assumption 2.3, the set $\hat{M}$ is Borel, hence M is analytic.

We apply Lemma 3.1 to the space (E, γ^*) and the set M: if (i) holds, then define $N := \bigcup_{i=1}^{l} M_i$. Then $\Gamma_l := E \setminus N$ has full measure, $\gamma^*(\Gamma_l) = 1$. From the definitions of M and N it can be directly seen that Γ_l is as needed.

If (i) does not hold, (ii) has to. Hence, let us derive a contradiction from it.

Write p_i for the projection of an element of E^l onto its i-th component. We may assume that the measure η in (ii) is concentrated on M, and also fulfills $p_i(\eta) \leq \frac{1}{l}\gamma^*$ for $i = 1, \dots, l$.

We now apply the Jankow – von Neumann selection theorem to the set $\hat{M}$ to define a mapping

$$z \mapsto \big(\alpha_1(z), \dots, \alpha_l(z), z'_1(z), \dots, z'_l(z), \alpha'_1(z), \dots, \alpha'_l(z)\big)$$

such that

$$\big(z, \alpha_1(z), \dots, \alpha_l(z), z'_1(z), \dots, z'_l(z), \alpha'_1(z), \dots, \alpha'_l(z)\big) \in \hat{M}$$

for $z \in M$, and the mapping is measurable with respect to the σ-field generated by the analytic subsets of E^l. Setting

$$\alpha_z := \sum_i \alpha_i(z)\delta_{z_i}, \alpha'_z := \sum_i \alpha'_i(z)\delta_{z'_i(z)}$$

we thus obtain kernels $z \mapsto \alpha_z$, $z \mapsto \alpha'_z$ from E^l with the σ-field generated by its analytic subsets to E with its Borel-sets. We use these kernels to define measures ω, ω' on the Borel-sets on E through

$$\omega(B) = \int \alpha_z(B)\,d\eta(z),\ \omega'(B) = \int \alpha'_z(B)\,d\eta(z).$$

By construction $\omega \leq \gamma^*$. Moreover ω' is a c-better competitor of ω: for each $f \in \mathcal{F}$ we have

$$\int f\ d\omega' = \iint f\ d\alpha'_z d\eta(z) = \iint f\ d\alpha_z d\eta(z) = \int f\ d\omega.$$

Note that the first and last equality are justified since $\int g\,d\alpha_z,\ \int g\,d\alpha'_z \le l$ for all z. Similarly, since $|c| \le g$, we obtain

$$\int c\,d\omega' = \iint c\,d\alpha'_z d\eta(z) < \iint c\,d\alpha_z d\eta(z) = \int c\,d\omega.$$

Summing up, we obtain a probability measure $\gamma' := \gamma^* - \omega + \omega'$ with $\int c\,d\gamma' < \int c\,d\gamma^*$ and $\gamma' \in \Pi_{\mathcal{F}}$. This contradicts the optimality of γ^*. □

4. The continuum marginal transport problem revisited

This section is devoted to establishing Theorem 2.1. W.l.o.g. we work with $I = [0, T] = [0, 1]$ from now on.

For completeness and to fix ideas, we discuss a result that can be seen as a finite-dimensional predecessor to [Pas13a] and has been well-known for at least several decades. We mention the note by [KDV$^+$02] for a simple geometric proof and further references, and for a more general result [Car03]. We denote by π_n^* the n-dimensional analogue of the measure π^* introduced in section 2.3. I.e., given n probability measures $\mu_1, \dots, \mu_n$ on $\mathbb{R}$, π_n^* is the push forward of Lebesgue measure λ on $(0, 1)$ to $\mathbb{R}^n$ via $x \mapsto \big(q_1(x), \dots, q_n(x)\big)$, where, as before, q_i is the quantile function of μ_i.

Theorem 4.1. *Let $h : \mathbb{R} \to \mathbb{R}$ be strictly concave and $\mu_1, \dots, \mu_n$ be probability measures on $\mathbb{R}$ such that*

$$\int |x|\,d\mu_i < \infty, \quad \int |h|\,d\mu_i < \infty, \quad \textit{for } 1 \le i \le n.$$

Then π_n^* is the unique minimizer of

$$\inf_{\gamma \in \Pi_n(\mu_1,\dots,\mu_n)} \int h(x_1 + \cdots + x_n)\,d\gamma(x). \tag{8}$$

It is intuitive to see why the monotonicity principle should come in useful for results as in Theorems 2.1 and 4.1. For in these situations, finite optimality of a set A (in $\mathbb{R}^n$ or $C[0, 1]$, respectively) implies that A must be a monotone set, i.e. $\le$ must be a total order on A: if f and g are both in A, then either $f \le g$ or $g \le f$. Else, set $f' = \max\{f, g\}$ and $g' = \min\{f, g\}$, and let α be the measure $\frac{1}{2}\delta_f + \frac{1}{2}\delta_g$ and α' the measure $\frac{1}{2}\delta_{f'} + \frac{1}{2}\delta_{g'}$. Then α' is a measure with the same marginals as α (on $\mathbb{R}^n$, or $C[0, 1]$, respectively). But due to strict concavity of h, it is easy to see that α' leads to lower costs than α in both cases (8) and (B), contradicting the definition of local optimality. The argument of optimality of π_n^* (or π^*, respectively) is then completed by another well-known fact, a proof of which we include for the convenience of the reader:

Lemma 4.2. *Let γ be a probability measure on $\mathbb{R}^n$ with marginals $\mu_1, \dots, \mu_n$. If there is a monotone Borel set M with $\gamma(M) = 1$, then $\gamma = \pi_n^*$.*
Let γ be a probability measure on $C[0, 1]$ with marginals $\big(\mu_t\big)_{t \in I}$. If there is a monotone Borel set M with $\gamma(M) = 1$, then $\gamma = \pi^*$.

Proof. The second part is a simple consequence of the first one since the distribution of a continuous process is determined by its finite dimensional marginal distributions. Hence, let γ be as in the first statement. For arbitrary points $a_1, \dots, a_n \in \mathbb{R}$, we show that for $I = (-\infty, a_1] \times \cdots \times (-\infty, a_n]$

we have $\gamma(I) = \pi_n^*(I)$. Set $z = \sup\{x : q_i(x) \leq a_i \text{ for } i = 1, \dots, n\}$. Then we have $\pi_n^*(I) = z$, and for at least one i_0 we have $\mu_{i_0}\big((-\infty, a_{i_0}]\big) = z$. We can hence conclude that $\gamma(I) \leq z$. And, in fact, equality must hold. For observe that from the definition of z we have $\mu_i\big((-\infty, a_i]\big) \geq z$ for all $i = 1, \dots, n$. Hence $\gamma(I) < z$ would imply that for each i there is an element $(b_1^{(i)}, \dots, b_n^{(i)}) \in \Gamma$ such that $b_i^{(i)} \leq a_i$, and $b_{j_i}^{(i)} > a_{j_i}$ for some $j_i \neq i$. This contradicts the monotonicity of Γ. □

Proof of Theorem 4.1. The set $\Pi(\mu_1 \dots, \mu_n)$ is weakly compact. Due to the assumptions on first moments and h-moments of the marginal measures μ_i, the operator to be minimized is lower semi-continuous and bounded. Hence there is a finite minimizer. Strict concavity of h and the above outlined application of the monotonicity principle yield that each finite minimizer must be concentrated on a finitely minimal, hence monotone set. By the preceding lemma, each minimizer must be equal to π_n^*. □

Now we turn to proving Theorem 2.1: here the neat argument for Theorem 4.1 does not work as smoothly as before, as $\Pi_C\big(\mu_t\big)$ need not be compact. This can be seen by easy counterexamples. Hence we have to find a way to establish the existence of an optimizer at all. Here is how we want to proceed: we will solve a problem for a countable index set as an intermediate step, where we also add monotonicity and boundedness (from above) to the assumptions on h. We then use the intermediate result in the proof of Theorem 2.1 at the end of this section. Writing $Q = [0,1] \cap \mathbb{Q}$, we define $\Pi_Q(\mu_q)$ as the set of probability measures on $\mathbb{R}^Q$ with marginals $(\mu_q)_{q \in Q}$. Furthermore, we fix a sequence of finite partitions $(\mathcal{P}_n)$ of $[0,1]$ with $\mathcal{P}_n \subseteq \mathcal{P}_{n+1} \subseteq Q$ and $\bigcup_n \mathcal{P}_n = Q$. We then replace the original problem (B) by

$$\text{(B')} \qquad \inf_{\gamma \in \Pi_Q(\mu_q)} \int h\Big(\limsup_{n\to\infty} \sum_{t_i \in \mathcal{P}_n} f_{t_i}(t_i - t_{i-1})\Big)\, d\gamma(f).$$

Writing π_Q^* for the Q-analogue of π^*, we claim:

Proposition 4.3. *Let $h : \mathbb{R} \to \mathbb{R}$ be concave, increasing, and non-positive. Provided that $\int |x|\, d\mu_q(x) < \infty$, $\int |h|\, d\mu_q < \infty$ for all $q \in Q$, the measure π_Q^* is a minimizer of Problem* (B').

The proof is preceded by Lemmas 4.4, 4.5, and 4.6. The assumptions here on h and the marginals are as in Theorem 4.3.

Lemma 4.4. $\Pi_Q\big(\mu_q\big)$ *is weakly compact.*

Proof. By Prochorov's theorem: let $\varepsilon > 0$ be arbitrary. Then, with $Q = \{q_1, q_2, \dots\}$, for each q_k there exists a compact set $K_k \subseteq \mathbb{R}$ with $\mu_{q_k}(K_k) > 1 - \frac{\varepsilon}{2^k}$. The set $K = \Pi_{k=1}^{\infty} K_k$ is a compact subset of $\mathbb{R}^Q$. For a measure $\gamma \in \Pi_Q\big(\mu_q\big)$ we have

$$\gamma(K) = \lim_{n\to\infty} \gamma\big(p^{-1}_{q_1,q_2,\dots,q_n}(K_1 \times K_2 \times \dots \times K_n)\big).$$

As for each n

$$\gamma\big(p^{-1}_{q_1,q_2,\dots,q_n}(K_1 \times K_2 \times \dots \times K_n)\big) > 1 - \sum_{k=1}^{n} \frac{\varepsilon}{2^k} \geq 1 - \varepsilon$$

we have $\gamma(K) \geq 1-\varepsilon$, and Prochorov's theorem can be applied. $\square$

We introduce some notation:

$$\begin{aligned}
s_n &: \mathbb{R}^Q \to \mathbb{R}, && f \mapsto \sum_{t_i \in \mathcal{P}_n} f_{t_i}(t_i - t_{i-1}),\\
s_n^{(h)} &: \mathbb{R}^Q \to \mathbb{R}, && f \mapsto \sum_{t_i \in \mathcal{P}_n} h(f_{t_i})(t_i - t_{i-1}),\\
\varphi_n &: \mathbb{R}^Q \to \mathbb{R} \cup \{\infty\}, && f \mapsto \sup_{k \geq n} s_k(f),\\
\varphi &: \mathbb{R}^Q \to \mathbb{R} \cup \{-\infty, \infty\}, && f \mapsto \inf_n \varphi_n(f) = \limsup_n s_n(f).
\end{aligned}$$

We continue with

Lemma 4.5. *For each n, the operators defined on $\Pi_Q(\mu_q)$,*

$$S_n : \gamma \mapsto \int h \circ s_n \, d\gamma$$

and

$$\Phi_n : \gamma \mapsto \int h \circ \varphi_n \, d\gamma$$

are lower-semi-continuous (w.r.t. weak convergence) and have minimizers. The values of the minima are finite.

Proof. The existence of minimizers will follow from lower-semi-continuity of the operators and compactness of $\Pi_Q(\mu_q)$. Hence, let $(\gamma_l)_{l \in \mathbb{N}}$ be a sequence in $\Pi_Q(\mu_q)$ converging weakly to some γ_0.
We have

$$\varphi_n \geq s_n$$

and hence, by monotonicity and concavity of h that

$$h \circ \varphi_n \geq h \circ s_n \geq s_n^{(h)}.$$

For each $\gamma \in \Pi_Q(\mu_q)$,

$$\int s_n^{(h)} \, d\gamma = \sum_{t_i \in \mathcal{P}_n} (t_i - t_{i-1}) \int h(f_{t_i}) \, d\gamma(f) = \sum_{t_i \in \mathcal{P}_n} (t_i - t_{i-1}) \int h \, d\mu_{t_i}.$$

Hence in particular

$$\lim_{l \to \infty} \int s_n^{(h)} \, d\gamma_l = \int s_n^{(h)} \, d\gamma_0.$$

As $s_n^{(h)}$ is continuous, the prerequisites of Lemma 4.3. in [Vil09] are met for both S_n and Φ_n, and applying that result we get

$$\liminf_{l \to \infty} S_n(\gamma_l) \geq S_n(\gamma_0)$$

and

$$\liminf_{l \to \infty} \Phi_n(\gamma_l) \geq \Phi_n(\gamma_0).$$

Finally, the finiteness of the minimal values follows from h being bounded from above, the assumption on finite h-moments of the marginals, and
$h \circ \varphi_n \geq h \circ s_n \geq s_n^{(h)}$. $\square$

Lemma 4.6. *For each $n \in \mathbb{N}$, the measure π_Q^* minimizes Φ_n on $\Pi_Q\big(\mu_q\big)$.*

Proof. We first show that, when h is strictly concave, the following stronger assertion is true: π_Q^* is the unique measure in $\Pi_Q\big(\mu_q\big)$ doing the following:

(0) it minimizes Φ_n,
(1) among the minimizers of Φ_n it minimizes S_1,
(2) among the measures fulfilling (0) and (1), it minimizes S_2,
$\vdots$
(k) among the measures fulfilling (0), (1), ..., ($k-1$), it minimizes S_k
$\vdots$

We show existence of a measure fulfilling all the conditions $(0), (1), \dots$: write K_0 for the set of minimizers of Φ_n. By the previous lemma, $K_0 \neq \emptyset$. Also, K_0 is compact: for it is a closed subset of the compact set $\Pi_Q(\mu_q)$, where closedness is due to the semi-continuity of Φ_n. Hence, among the minimizers of Φ_n, there is a minimizer of the lower-semi-continuous operator S_1. Writing K_1 for the set of these minimizers, by the same argument as above, K_1 is nonempty and compact. Hence, the set K_2 of minimizers of S_2 on K_1 is nonempty and again compact. By induction we obtain a decreasing sequence of compact nonempty sets K_k. Hence the set $K = \bigcap_k K_k$ is nonempty and each of its elements fulfills properties $(0), (1), \dots$ Pick such an element and denote it by π_0. We now apply the monotonicity principle to show that π_0 must indeed be equal to π_Q^*: π_0 is concentrated on a set Γ that is locally optimal for each of the problems (k). Observe first that local optimality of Γ for problem (0) alone does not need to imply that Γ is monotone.[6] However, local optimality of Γ for problem (1) - i.e. the optimization of S_1 on the set K_0 - does imply that Γ must be monotone on $\mathcal{P}_1$, that is, if $f, g \in \Gamma$, then either $f|_{\mathcal{P}_1} \leq g|_{\mathcal{P}_1}$ or $f|_{\mathcal{P}_1} \geq g|_{\mathcal{P}_1}$. For if there were f, g not ordered on $\mathcal{P}_1$, then write $f' = \mathbf{1}_{\mathcal{P}_1} \max(f, g) + \mathbf{1}_{\mathcal{P}_1^c} f$ and $g' = \mathbf{1}_{\mathcal{P}_1} \min(f, g) + \mathbf{1}_{\mathcal{P}_1^c} g$. Set $\alpha = \frac{1}{2}\delta_f + \frac{1}{2}\delta_g$ and $\alpha' = \frac{1}{2}\delta_{f'} + \frac{1}{2}\delta_{g'}$, where δ_f denotes the Dirac-measure on f, etc. Then apparently $S_1(\alpha') < S_1(\alpha)$, but α' is also a competitor of α: it clearly has the same marginals, and we have $\varphi_n(f') = \varphi_n(f)$ and $\varphi_n(g') = \varphi_n(g)$, as manipulating a function $f \in \mathbb{R}^Q$ on finitely many points does not change the value of φ_n. Hence, also $\Phi_n(\alpha') = \int h \circ \varphi_n \, d\alpha' = \int h \circ \varphi_n \, d\alpha = \Phi_n(\alpha)$. The existence of an S_1-better competitor is a contradiction to local optimality, so Γ must indeed be monotone on $\mathcal{P}_1$. Now for problem (2), we also find that Γ must be monotone on $\mathcal{P}_2$: let $f, g \in \Gamma$, and assume, due to monotonicity of Γ on $\mathcal{P}_1$, that $f|_{\mathcal{P}_1} \geq g|_{\mathcal{P}_1}$. If f and g were not ordered on $\mathcal{P}_2$, then the same construction of f', g', α and α' as above (with $\mathcal{P}_2$ in place of $\mathcal{P}_1$) will give a contradiction to local optimality: note that $s_1(f') = s_1(f)$ and $s_1(g') = s_1(g)$, as $f' = f$ and $g' = g$ on $\mathcal{P}_1$. Hence, $\Phi_n(\alpha') = \int h \circ \varphi_n \, d\alpha' = \int h \circ \varphi_n \, d\alpha = \Phi_n(\alpha)$, $S_1(\alpha') = \int h \circ s_1 \, d\alpha' = \int h \circ s_1 \, d\alpha = S_1(\alpha)$, and α' is really a competitor of α.

[6] What local optimality does imply is the following: if f, g are in Γ, and $\varphi_n(f) > \varphi_n(g)$, then one must have $\varphi_n\big((f-g)^+\big) = 0$. This is a weaker condition than $\leq$ being an order on Γ, and explains why one works with the sequence of problems (k) rather than just with problem (0).

Iterating this argument one finds that Γ must indeed be monotone on each $\mathcal{P}_k$, and henceforth monotone. But then π_0 must be π^*_Q, because π^*_Q is the only measure in $\Pi_Q(\mu_q)$ concentrated on a monotone set. This last statement follows easily from Lemma 4.2.

Finally, we discuss the case where h is concave, but not necessarily strictly concave. Then, due to the finiteness of $\int |x|\,d\mu_q$ for all $q \in Q$, there is, for each $k \in \mathbb{N}$, a strictly concave function h_k such that $\int |h_k|\,d\mu_q < \infty$ for all $q \in \mathcal{P}_k$. Then by adapting the above argument, it is easy to see that π^*_Q is the only measure in $\Pi_Q\big(\mu_q\big)$ that

(0) minimizes Φ_n
(1') among the minimizers of Φ_n, it minimizes $\int h_1(s_1)\,d\gamma$,
$\vdots$
(k') among the measures fulfilling (0), ..., (k-1'), it minimizes $\int h_k(s_k)\,d\gamma$,
$\vdots$

$\square$

Proof of Proposition 4.3. Let γ be a measure in $\Pi_Q(\mu_q)$. Then for each n, according to the previous lemma

$$\int h \circ \varphi_n \, d\gamma \geq \int h \circ \varphi_n \, d\pi^*_Q.$$

As h is increasing and non-positive, and φ_n decreases to $\varphi = \limsup_n s_n$, an application of monotone convergence finishes the proof. $\square$

Finally we can prove Theorem 2.1:

Proof of Theorem 2.1. First, note that due to the regularity assumption of $\int_0^1 \int |x| d\mu_t \, dt < \infty$, it is w.l.o.g to assume that h is non-positive. If we further assume for the time being that h is increasing, we can apply Proposition 4.3 to see the optimality of π^* as follows: let p_Q be the projection $\mathbb{R}^I \to \mathbb{R}^Q$, and write p for its restriction on $C[0,1]$. It is easy to see that p is a Borel isomorphism from $C[0,1]$ onto $\mathbb{R}^Q_c$, the set of all elements of $\mathbb{R}^Q$ that are restrictions of continuous functions on $[0,1]$. For an arbitrary $\gamma \in \Pi_C(\mu_t)$, the measure $p(\gamma)$ is in $\Pi_Q(\mu_q)$ and clearly

$$\int h\Big(\int_0^1 f\,dt\Big)\,d\gamma = \int h\Big(\limsup_{n\to\infty} \sum_{t_i \in \mathcal{P}_n} f_{t_i}(t_i - t_{i-1})\Big)\,d\,p(\gamma).$$

But for the right-hand-side one also has, due to Theorem 4.3,

$$\int h\Big(\limsup_{n} \sum_{t_i \in \mathcal{P}_n} f_{t_i}(t_i - t_{i-1})\Big)\,d\,p(\gamma) \geq \int h\Big(\limsup_{n} \sum_{t_i \in \mathcal{P}_n} f_{t_i}(t_i - t_{i-1})\Big)\,d\,\pi^*_Q.$$

As the right-hand-side of this equals $\int h\Big(\int_0^1 f\,dt\Big)\,d\pi^*$ we have

$$\int h\Big(\int_0^1 f\,dt\Big)\,d\gamma \geq \int h\Big(\int_0^1 f\,dt\Big)\,d\pi^*.$$

If h is not increasing, then assume first it is decreasing. If in problem (B') we replace lim sup by lim inf one can show, with the statement and proof of Theorem 4.3 and the above argument suitably adapted, that π^* must be again optimal. Finally, if h is neither increasing nor decreasing, then it can still be written as a sum $h_1 + h_2$, where h_1 is concave, increasing and non-positive, and h_2 is concave, decreasing and non-positive, and again π^* is an optimizer. (h_1 and h_2 will satisfy the regularity assumptions as long as h does.)

If the minimum is finite and h is strictly concave, each other minimizer must be concentrated on a finitely minimal, hence monotone set and thus be equal to π^*. □

References

[ABP+13] B. Acciaio, M. Beiglböck, F. Penkner, W. Schachermayer, and J. Temme. A trajectorial interpretation of doob's martingale inequalities. *The Annals of Applied Probability*, 23(4):1494–1505, 2013.

[ABPS13] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer. A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *ArXiv e-prints*, January 2013.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[AN87] E.J. Anderson and P. Nash. *Linear programming in infinite-dimensional spaces*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Ltd., Chichester, 1987. Theory and applications, A Wiley-Interscience Publication.

[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.

[BC10] S. Bianchini and L. Caravenna. On optimality of c-cyclically monotone transference plans. *C. R. Math. Acad. Sci. Paris*, 348(11-12):613–618, 2010.

[BGMS09] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009.

[BHP13] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: A mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.

[BJ14] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab., to appear*, 2014.

[BN13] B. Bouchard and M. Nutz. Arbitrage and Duality in Nondominated Discrete-Time Models. *Ann. Appl. Probab., to appear*, May 2013.

[BN14] M. Beiglböck and M. Nutz. Martingale inequalities and deterministic counterparts. *Electron. J. Probab.*, 2014.

[BS13] M. Beiglböck and P. Siorpaes. Pathwise versions of the Burkholder-Davis-Gundy-inequality. *Bernoulli, to appear*, 2013.

[Car03] G. Carlier. On a class of multidimensional optimal transportation problems. *Journal of Convex Analysis*, 10(2):517–529, 2003.

[DM13] Y. Dolinsky and H. Mete Soner. Robust Hedging with Proportional Transaction Costs. *Finance Stoch., to appear*, February 2013.

[DS13] Y. Dolinsky and M. H. Soner. Robust Hedging with Proportional Transaction Costs. *Finance Stoch. (to appear)*, February 2013.

[GHLT14] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336, 2014.

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

[Hob11] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, Berlin, 2011.

[HT13] P. Henry-Labordere and N. Touzi. An Explicit Martingale Version of Brenier's Theorem. *ArXiv e-prints*, February 2013.

[Jui14] N. Juillet. On displacement interpolation of measures involved in brenier's theorem. 2014.

[KDV+02] R. Kaas, J. Dhaene, D. Vyncke, M.J. Goovaerts, and M. Denuit. A simple geometric proof that comonotonic risks have the convex-largest sum. *ASTIN Bulletin*, 32(1):71–80, 2002.

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

[Kem68] J. H. B. Kemperman. The general moment problem, a geometric approach. *Annals Math. Stat.*, 39:93–122, 1968.

[KP13] Y.-H. Kim and B. Pass. A general condition for Monge solutions in the multi-marginal optimal transport problem. *SIAM J. Math. Anal., to appear*, July 2013.

[KS84] M. Knott and C. S. Smith. On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43(1):39–49, 1984.

[KZ13] A. V. Kolesnikov and D. A. Zaev. Optimal transportation of processes with infinite Kantorovich distance. Independence and symmetry. *ArXiv e-prints*, March 2013.

[Las10] Jean-Bernard Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2010.

[Obł04] J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004.

[Pas11] B. Pass. Uniqueness and Monge solutions in the multimarginal optimal transportation problem. *SIAM J. Math. Anal.*, 43(6):2758–2775, 2011.

[Pas12] B. Pass. On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calc. Var. Partial Differential Equations*, 43(3-4):529–536, 2012.

[Pas13a] B. Pass. On a class of optimal transportation problems with infinitely many marginals. *SIAM J. Math. Anal.*, 45:2557–2575, 2013.

[Pas13b] B. Pass. Optimal transportation with infinitely many marginals. *J. Funct. Anal.*, 264(4):947–963, 2013.

[Pra08] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 2008. in press.

[Rüs91] L. Rüschendorf. Fréchet-bounds and their applications. In *Advances in probability distributions with given marginals (Rome, 1990)*, volume 67 of *Math. Appl.*, pages 151–187. Kluwer Acad. Publ., Dordrecht, 1991.

[Rüs95] L. Rüschendorf. Optimal solutions of multivariate coupling problems. *Appl. Math. (Warsaw)*, 23(3):325–338, 1995.

[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.

[Vil09] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.

[Zae14] D. Zaev. On the Monge-Kantorovich problem with additional linear constraints. *ArXiv e-prints*, April 2014.