Electron. Commun. Probab. **27** (2022), article no. 16, 1–12.
https://doi.org/10.1214/22-ECP457
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

The potential of the shadow measure

Mathias Beiglböck* David Hobson† Dominykas Norgilas‡

Abstract

It is well known that given two probability measures μ and ν on $\mathbb{R}$ in convex order there exists a discrete-time martingale with these marginals. Several solutions are known (for example from the literature on the Skorokhod embedding problem in Brownian motion). But, if we add a requirement that the martingale should minimise the expected value of some functional of its starting and finishing positions then the problem becomes more difficult. Beiglböck and Juillet (Ann. Probab. 44 (2016) 42–106) introduced the *shadow* measure which induces a family of martingale couplings, and solves the optimal martingale transport problem for a class of bivariate objective functions. In this article we extend their (existence and uniqueness) results by providing an explicit construction of the shadow measure and, as an application, give a simple proof of its associativity.

Keywords: couplings; martingales; peacocks; convex order; optimal transport.
MSC2020 subject classifications: 60G42.
Submitted to ECP on April 20, 2021, final version accepted on February 22, 2022.
Supersedes arXiv:2008.09936v2.

1 Introduction

Given two probability measures μ, ν on $\mathbb{R}$, a *transport plan* or *coupling* between μ and ν is a probability measure π on $\mathbb{R}^2$ such that $\pi(A\times\mathbb{R}) = \mu(A)$ and $\pi(\mathbb{R}\times B) = \nu(B)$ for all Borel sets A, B of $\mathbb{R}$. It is often convenient to express a coupling π via its disintegration with respect to the first marginal μ, $\pi(dx, dy) = \mu(dx)\pi_x(dy)$ where $x \mapsto \pi_x$ is a μ-almost surely unique probability kernel. In the language of the classical optimal transport, each transport plan π corresponds to a joint distribution of $X \sim \mu$ and $Y \sim \nu$ and then it is natural to ask for couplings which, (not only possess nice structures or properties but also) for a given cost function $c : \mathbb{R} \to \mathbb{R}$, minimise the expected cost $\mathbb{E}^\pi[c(X, Y)]$. Brenier's Theorem (see Brenier [5], and Rüschendorf and Rachev [23]) considers the problem in $\mathbb{R}^d$ with an Euclidean cost $c(x, y) = |y - x|^2$. Then, under some regularity assumptions, the optimal coupling is a push forward measure induced by the gradient of a convex function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$. In other words, the optimal coupling $\hat{\pi}$ is *deterministic* and of the form

$$\hat{\pi}(dx, dy) = \mu(dx)\delta_{\nabla\phi(x)}(dy).$$

In one dimension, this says that $\hat{\pi}$ is concentrated on a graph of an increasing function. Furthermore, it is optimal for (at least) costs $c(x, y) = h(y - x)$, where $h : \mathbb{R} \mapsto \mathbb{R}$ is

*Fakultät für Mathematik, Universität Wien. E-mail: mathias.beiglboeck@univie.ac.at
†Department of Statistics, University of Warwick. E-mail: d.hobson@warwick.ac.uk
‡Department of Mathematics, University of Michigan. E-mail: dnorgila@umich.edu

strictly convex, and coincides with the monotone (Hoeffding-Fréchet) coupling π^{HF} (which is often called the *quantile* coupling).

In the martingale version of optimal transport, introduced in the context of a specific cost function by Hobson and Neuberger [14] and more generally by Beiglböck et al. [2] and Galichon et al. [9], the goal is to construct a martingale $(M_i)_{i=1,2}$ with $M_1 \sim \mu$ and $M_2 \sim \nu$, and such that $\mathbb{E}[c(M_1, M_2)]$ is minimised. The martingale requirement places a non-trivial constraint on the possible joint distributions of M_1 and M_2 and the martingale transports correspond to measures π on $\mathbb{R}^2$ with univariate marginals μ and ν such that the kernel in the disintegration of π satisfies the barycenter property

$$\int y\pi_x(dy) = x, \quad \text{for } \mu\text{-almost every } x \in \mathbb{R}.$$

For arbitrary probability measures μ and ν the existence of a martingale transport is not guaranteed and requires that μ is smaller than ν in convex order, which we write as $\mu \leq_{cx} \nu$. That this condition is necessary and sufficient for there to exist a martingale with given marginals was proved by Strassen [24] (and extended to the level of continuous time processes by Kellerer [18, 19]).

Provided that $\mu \leq_{cx} \nu$, one then seeks to find optimal martingale couplings. For quadratic costs the martingale transport problem is trivial: any martingale coupling is an optimiser. Solutions are known for several other specific but important costs. Hobson and Neuberger [14] and Hobson and Klimmek [13], in the context of mathematical finance, provide the (non-explicit and explicit, respectively) constructions of the optimal martingale couplings π^{HN} and π^{HK} for the cost functions $c(x,y) = -|x-y|$ and $c(x,y) = |x-y|$, respectively. (Hobson and Klimmek [13] work under the dispersion assumption whereby $\mu \geq \nu$ on an interval I and $\mu \leq \nu$ outside I.) Beiglböck and Juillet [4] introduced the so-called left-curtain coupling π^{lc} (a martingale counterpart of the quantile coupling in the classical optimal transport) and proved its optimality for costs of the form $c(x,y) = h(y-x)$ for some differentiable function h with strictly convex derivative.

All of the aforementioned martingale couplings have nice structural properties: if μ is atom-free then card(spt$\{\pi_x^{HN}\}$)≤ 2, card(spt$\{\pi_x^{HK}\}$)≤ 3 with card(spt$\{\pi_x^{HK}\} \setminus \{x\}$)$\leq 2$ and card(spt$\{\pi_x^{lc}\}$)≤ 2 for μ-almost every $x \in \mathbb{R}$. In particular, in each case there exist lower and upper functions on which the couplings are concentrated. Under the dispersion assumption, Hobson and Klimmek [13] constructed the upper and lower functions for π^{HK}, while for π^{HN} only the existence is known. When the initial law μ is atomic, an explicit construction of the characteristic functions of the left-curtain coupling is provided in Beiglböck and Juillet [4]. Another construction of π^{lc} (using ordinary differential equations) is given by Henry-Labordère and Touzi [10] for atomless initial measures μ. For general initial and target laws Hobson and Norgilas [15] constructed the upper and lower functions that characterise the generalised (or *lifted*) left-curtain martingale coupling using weak approximation of measures. Several other authors further investigate the properties and extensions of the left-curtain coupling, see Henry-Labordère et al. [11], Beiglböck et al. [3, 1], Juillet [16, 17], Nutz et al. [20, 21] and Brückerhoff at al. [6].

To study the transport plans in the martingale setting, Beiglböck and Juillet [4] introduced the *extended* convex order of two measures. Its significance lies in the fact that, for each pair of measures μ, ν with μ less than ν in the extended convex order, there exists a martingale that transports μ into ν (without necessarily covering all of ν). In particular, the set of measures η with $\mu \leq_{cx} \eta \leq \nu$, is non-empty. Each such η corresponds to a terminal law of a martingale that embeds μ into ν. Among these terminal laws there are at least two canonical choices, namely, the smallest and the largest element with respect to the convex order, which correspond to the most concentrated and the most disperse attainable terminal law of the transporting martingale, respectively. Beiglböck and Juillet [4] proved the existence and uniqueness of both of these extreme measures and baptised the minimal measure, which is the main object of interest in this paper, the *shadow* (of μ in ν). One of the main achievements of this paper is Theorem 4.7, which provides an explicit construction of the shadow measure. This allows us to give a simple proof of the existence and uniqueness which is considerably more direct than that given in [4].

Arguably the most important structural property of the shadow measure is its associativity (see Theorem 4.8). In particular, it is the main ingredient in defining the *left-curtain* martingale coupling, or more generally, any martingale coupling induced by the shadow measure. More specifically, the left-curtain martingale coupling is defined as the unique measure π^{lc} on $\mathbb{R}^2$ such that, for each $x \in \mathbb{R}$, $\pi^{lc}|_{(-\infty,x])\times\mathbb{R}}$ has the first marginal $\mu|_{(-\infty,x]}$ and the second marginal $S^\nu(\mu|_{(-\infty,x]})$, where $S^\nu(\mu|_{(-\infty,x]})$ is the shadow of $\mu|_{(-\infty,x]}$ in ν.

The current paper is structured as follows. In Section 2 we discuss the relevant notions of probability measures and (extended) convex order, and state some crucial (for our main theorems) results regarding the convex hull of a function. In Section 3 we use potential-geometric arguments to explicitly construct the maximal measure with respect to convex order (the opposite of the shadow measure). Section 4 is dedicated to our main results. First, in Theorem 4.7 we provide an explicit construction of the shadow measure in terms of its potential function. Then we use this result to give a simplified proof of the associativity of the shadow, see Theorem 4.8.

2 Preliminaries

2.1 Measures and Convex order

Let $\mathcal{M}$ (respectively $\mathcal{P}$) be the set of measures (respectively probability measures) on $\mathbb{R}$ with finite total mass and finite first moment, i.e., if $\eta \in \mathcal{M}$, then $\eta(\mathbb{R}) < \infty$ and $\int_\mathbb{R} |x|\eta(dx) < \infty$. Given a measure $\eta \in \mathcal{M}$ (not necessarily a probability measure), define $\overline{\eta} = \int_\mathbb{R} x\eta(dx)$ to be the first moment of η (and then $\overline{\eta}/\eta(\mathbb{R})$ is the barycentre of η). Let $\mathcal{I}_\eta$ be the smallest interval containing the support of η, and let $\{\ell_\eta, r_\eta\}$ be the endpoints of $\mathcal{I}_\eta$. If η has an atom at ℓ_η then ℓ_η is included in $\mathcal{I}_\eta$, and otherwise it is excluded, and similarly for r_η.

For $\alpha \geq 0$ and $\beta \in \mathbb{R}$ let $\mathcal{D}(\alpha, \beta)$ denote the set of increasing, convex functions $f : \mathbb{R} \mapsto \mathbb{R}_+$ such that $\lim_{z\downarrow-\infty}\{f(z)\} = 0$ and $\lim_{z\uparrow\infty}\{f(z) - (\alpha z - \beta)\} = 0$. Then, when $\alpha = 0$, $\mathcal{D}(0, \beta)$ is empty unless $\beta = 0$ and then $\mathcal{D}(0, 0)$ contains one element, the zero function.

For $\eta \in \mathcal{M}$, define the functions $P_\eta, C_\eta : \mathbb{R} \mapsto \mathbb{R}^+$ by

$$P_\eta(k) := \int_\mathbb{R} (k-x)^+\eta(dx), \quad k \in \mathbb{R}, \qquad\qquad C_\eta(k) := \int_\mathbb{R} (x-k)^+\eta(dx), \quad k \in \mathbb{R},$$

respectively. Then $P_\eta(k) \geq 0 \vee (\eta(\mathbb{R})k - \overline{\eta})$ and $C_\eta(k) \geq 0 \vee (\overline{\eta} - \eta(\mathbb{R})k)$. Also $C_\eta(k) - P_\eta(k) = (\overline{\eta} - \eta(\mathbb{R})k)$.

The following properties of P_η can be found in Chacon [7], and Chacon and Walsh [8]: $P_\eta \in \mathcal{D}(\eta(\mathbb{R}), \overline{\eta})$ and $\{k : P_\eta(k) > (\eta(\mathbb{R})k - \overline{\eta})^+\} = \{k : C_\eta(k) > (\overline{\eta} - \eta(\mathbb{R})k)^+\} = (\ell_\eta, r_\eta)$. Conversely (see, for example, Proposition 2.1 in Hirsch et al. [12]), if h is a non-negative, non-decreasing and convex function with $h \in \mathcal{D}(k_m, k_f)$ for some numbers $k_m \geq 0$ and $k_f \in \mathbb{R}$ (with $k_f = 0$ if $k_m = 0$), then there exists a unique measure $\eta \in \mathcal{M}$, with total mass $\eta(\mathbb{R}) = k_m$ and first moment $\overline{\eta} = k_f$, such that $h = P_\eta$. In particular, η is uniquely identified by the second derivative of h in the sense of distributions. Furthermore, P_η

and C_η are related to the potential U_η, defined by

$$U_\eta(k) := -\int_{\mathbb{R}} |k-x|\eta(dx), \quad k \in \mathbb{R},$$

via $-U_\eta = C_\eta + P_\eta$. We will call P_η (and C_η) a modified potential. Finally note that all three second derivatives C''_η, P''_η and $-U''_\eta/2$ identify the same underlying measure η.

For $\eta, \chi \in \mathcal{M}$, we write $\eta \leq \chi$ if $\eta(A) \leq \chi(A)$ for all Borel measurable subsets A of $\mathbb{R}$, or equivalently if

$$\int f d\eta \leq \int f d\chi, \quad \text{for all non-negative } f : \mathbb{R} \mapsto \mathbb{R}_+.$$

Since η and χ can be identified as second derivatives of P_χ and P_η respectively, we have $\eta \leq \chi$ if and only if $P_\chi - P_\eta$ is convex, i.e., P_η has a smaller curvature than P_χ.

Two measures $\eta, \chi \in \mathcal{M}$ are in convex order, and we write $\eta \leq_{cx} \chi$, if

$$\int f d\eta \leq \int f d\chi, \quad \text{for all convex } f : \mathbb{R} \mapsto \mathbb{R}. \tag{2.1}$$

Since we can apply (2.1) to all affine functions, including $f(x) = \pm 1$ and $f(x) = \pm x$, we obtain that if $\eta \leq_{cx} \chi$ then η and χ have the same total mass ($\eta(\mathbb{R}) = \chi(\mathbb{R})$) and the same first moment ($\overline{\eta} = \overline{\chi}$). Moreover, necessarily we must have $\ell_\chi \leq \ell_\eta \leq r_\eta \leq r_\chi$. From simple approximation arguments (see Hirsch et al. [12]) we also have that if η and χ have the same total mass and the same barycentre, then $\eta \leq_{cx} \chi$ if and only if $P_\eta(k) \leq P_\chi(k)$, $k \in \mathbb{R}$.

For our purposes in the sequel we need a generalisation of the convex order of two measures. We follow Beiglböck and Juillet [4] and say $\eta, \chi \in \mathcal{M}$ are in an *extended* convex order, and write $\eta \leq_E \chi$, if

$$\int f d\eta \leq \int f d\chi, \quad \text{for all non-negative, convex } f : \mathbb{R} \mapsto \mathbb{R}_+.$$

If $\eta \leq_{cx} \chi$ then also $\eta \leq_E \chi$ (since non-negative convex functions are convex), while if $\eta \leq \chi$, we also have that $\eta \leq_E \chi$ (since non-negative convex functions are non-negative). Note that, if $\eta \leq_E \chi$, then $\eta(\mathbb{R}) \leq \chi(\mathbb{R})$ (apply the non-negative convex function $\phi(x) = 1$ in the definition of $\leq_E$). It is also easy to prove using the above characterisation of convex order via potential functions that if $\eta(\mathbb{R}) = \chi(\mathbb{R})$, then $\eta \leq_E \chi$ is equivalent to $\eta \leq_{cx} \chi$.

For $\eta, \chi \in \mathcal{P}$, let $\Pi(\eta, \chi)$ be the set of probability measures on $\mathbb{R}^2$ with the first marginal η and second marginal χ. Let $\Pi_M(\eta, \chi)$ be the set of martingale couplings of η and χ. Then $\Pi_M(\eta, \chi) = \{\pi \in \Pi(\eta, \chi) : \text{(2.2) holds}\}$, where (2.2) is the martingale condition

$$\int_{x \in B} \int_{y \in \mathbb{R}} y \pi(dx, dy) = \int_{x \in B} \int_{y \in \mathbb{R}} x \pi(dx, dy) = \int_B x \eta(dx), \quad \forall \text{ Borel } B \subseteq \mathbb{R}. \tag{2.2}$$

Equivalently, $\Pi_M(\eta, \chi)$ consists of all transport plans π (i.e., elements of $\Pi(\eta, \chi)$) such that the disintegration in probability measures $(\pi_x)_{x \in \mathbb{R}}$ with respect to η satisfies $\int_{\mathbb{R}} y \pi_x(dy) = x$ for η-almost every x. For any $\pi \in \Pi_M(\eta, \chi)$ and convex $f : \mathbb{R} \mapsto \mathbb{R}$, by (conditional) Jensen's inequality we have that

$$\int_{\mathbb{R}} f(x)\eta(dx) \leq \int_{\mathbb{R}} \int_{\mathbb{R}} f(y)\pi_x(dy)\eta(dx) = \int_{\mathbb{R}} f(y)\pi(\mathbb{R}, dy) = \int_{\mathbb{R}} f(y)\chi(dy),$$

so that $\eta \leq_{cx} \chi$. On the other hand, Strassen [24] showed that a converse is also true (i.e., $\eta \leq_{cx} \chi$ implies that $\Pi_M(\eta, \chi) \neq \emptyset$), so that $\Pi_M(\eta, \chi)$ is non-empty if and only if $\eta \leq_{cx} \chi$.

For $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$ we define $\mathcal{M}_\mu^\nu = \{\eta \in \mathcal{M} : \mu \leq_{cx} \eta \leq \nu\}$. Then $\mathcal{M}_\mu^\nu$ is a set of terminal laws of a martingale that embeds μ into ν. Note that $\eta \in \mathcal{M}_\mu^\nu$ if and only if $P_\eta \in \mathcal{P}(\mu, \nu)$, where

$$\mathcal{P}(\mu, \nu) := \{\tilde{P} \in \mathcal{D}(\mu(\mathbb{R}), \overline{\mu}) : P_\nu - \tilde{P} \text{ is convex and } P_\mu \leq \tilde{P}\}.$$

Remark 2.1. If $\mu \leq_{cx} \nu$ then $\mathcal{M}_\mu^\nu$ is the singleton $\{\nu\}$.

2.2 Convex hull

Our key results will be expressed in terms of the convex hull. For $f : \mathbb{R} \mapsto (-\infty, \infty)$ let f^c be the largest convex function which lies below f. In our typical application f will be non-negative and this property will be inherited by f^c. However, in general we may have f^c equal to $-\infty$ on $\mathbb{R}$, and the results of this section are stated in a way which includes this case. Note that if a function g is equal to $-\infty$ (or ∞) everywhere, then we deem it to be both linear and convex, and set g^c equal to g.

Fix $x, z \in \mathbb{R}$ with $x \leq z$. For $x < z$, define $L^f_{x,z} : [x, z] \mapsto \mathbb{R}$ by

$$L^f_{x,z}(y) = f(x)\frac{z-y}{z-x} + f(z)\frac{y-x}{z-x}, \quad y \in [x, z] \tag{2.3}$$

and for $x = z$, define $L^f_{x,x} : \{x\} \mapsto \mathbb{R}$ by $L^f_{x,x}(x) = f(x)$. Then, see Rockafellar [22, Corollary 17.1.5],

$$f^c(y) = \inf_{x \leq y \leq z} L^f_{x,z}(y), \quad y \in \mathbb{R}. \tag{2.4}$$

Moreover, it is not hard to see (for example, by drawing the graphs of f and f^c) that f^c replaces the non-convex segments of f by straight lines. (Proofs of lemmas in this section are given in Section 5.)

Lemma 2.2. *Let $f : \mathbb{R} \mapsto \mathbb{R}$ be lower semi-continuous. Suppose $f > f^c$ on $(a, b) \subseteq \mathbb{R}$. Then f^c is linear on (a, b).*

The following lemmas are the main ingredients in the proofs of Theorem 4.7 and Theorem 4.8.

Lemma 2.3. *Let $f, g : \mathbb{R} \mapsto \mathbb{R}$ be two convex functions. Define $\psi : \mathbb{R} \mapsto (-\infty, \infty]$ by $\psi = g - (g - f)^c$. Then ψ is convex.*

Lemma 2.4. *Let $f : \mathbb{R} \mapsto \mathbb{R}$ be any measurable function and let $g : \mathbb{R} \mapsto \mathbb{R}$ be a convex function. Then*

$$(f - g)^c = (f^c - g)^c.$$

Lemma 2.5. *Assume that $f \in \mathcal{D}(\alpha, \beta)$ and $g \in \mathcal{D}(a, b)$ for some $\alpha, a \geq 0$, $\beta, b \in \mathbb{R}$. Let $h : \mathbb{R} \mapsto \mathbb{R}$ be defined by $h(k) := (a - \alpha)k - (b - \beta)$.*

Suppose that $g \geq f$. Then $\alpha \leq a$. If $\alpha = a$ then $\beta \geq b$.

Suppose that $g \geq f$ and $g - f \geq h$. Then $(g - f)^c \in \mathcal{D}(a - \alpha, b - \beta)$.

Note that in the above lemma if $\alpha = a$ and $\beta > b$ then there are no pairs of functions (f, g) with $f \in \mathcal{D}(a, \beta)$, $g \in \mathcal{D}(a, b)$ and $g(z) - f(z) \geq h(z) = \beta - b$. We cannot have both $\lim_{z \downarrow -\infty}(g(z) - f(z)) = 0$ and $g(z) - f(z) \geq h(z) = \beta - b > 0$. In this case the final statement of the lemma is vacuous.

3 The maximal element

In this section we find the largest measure (w.r.t. convex order) in $\mathcal{M}_\mu^\nu$ (see Theorem 3.3). It follows that $\mathcal{M}_\mu^\nu \neq \emptyset$. The reader who is only interested in the shadow measure can omit this section and move directly to Section 4. The only result we will use in the study of the shadow measure is the existence of an element $\eta \in \mathcal{M}_\mu^\nu$ which also follows from Proposition 4.3, although then this paper is no longer self-contained.

The key result of this section is that there is a simple, explicit expression for the maximal element of $\mathcal{M}^{\nu}_{\mu}$ in terms of its potential.

Definition 3.1 (Counter-shadow measure)**.** *Let $\mu, \nu \in \mathcal{M}$ and assume $\mu \leq_E \nu$. The counter-shadow of μ in ν, denoted by $T^{\nu}(\mu)$, has the following properties:*

1. $\mu \leq_{cx} T^{\nu}(\mu) \leq \nu$,
2. *If η is another measure satisfying $\mu \leq_{cx} \eta \leq \nu$ then $\eta \leq_{cx} T^{\nu}(\mu)$.*

By Remark 2.1, if $\mu \leq_{cx} \nu$ then necessarily $T^{\nu}(\mu) = \nu$.

Lemma 3.2 (Beiglböck and Juillet [4], Lemma 4.5)**.** *For $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$, $T^{\nu}(\mu)$ exists and is unique.*

Beiglböck and Juillet [4] not only prove the existence and uniqueness of $T^{\nu}(\mu)$, but also show how to construct $T^{\nu}(\mu)$. Our first result provides an alternative, simple and explicit construction of $T^{\nu}(\mu)$ via potential functions. Let $\tilde{P}^{\nu}_{\mu} : \mathbb{R} \mapsto \mathbb{R}$ be given by $\tilde{P}^{\nu}_{\mu}(k) = \min\{P_{\nu}(k), C_{\nu}(k) + (\mu(\mathbb{R})k - \overline{\mu})\}$. See Fig. 1. Recall put-call parity: $C_{\nu}(k) = P_{\nu}(k) + \overline{\nu} - k\nu(\mathbb{R})$. Then an alternative expression for $\tilde{P}^{\nu}_{\mu}$ is

$$\tilde{P}^{\nu}_{\mu}(k) = P_{\nu}(k) - ((\nu(\mathbb{R}) - \mu(\mathbb{R}))k - (\bar{\nu} - \bar{\mu}))^{+}, \quad k \in \mathbb{R}. \tag{3.1}$$

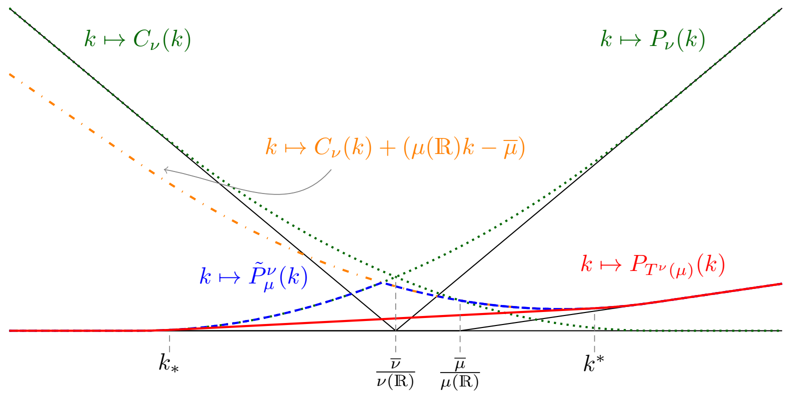

Figure 1: Construction of $P_{T^{\nu}(\mu)}$. Dotted curves correspond to the graphs of C_{ν} and P_{μ}, while the dash-dotted curve corresponds to $k \mapsto C_{\nu}(k) + (\mu(\mathbb{R})k - \overline{\mu})$. The dashed curve represents $\tilde{P}^{\nu}_{\mu}$. Note that $\tilde{P}^{\nu}_{\mu}$ is continuous. The solid curve below $\tilde{P}^{\nu}_{\mu}$ corresponds to $P_{T^{\nu}(\mu)}$, the (modified) potential of the counter-shadow. Note that, on (k_*, k^*), $P_{T^{\nu}(\mu)}$ is linear and strictly below $\tilde{P}^{\nu}_{\mu}$, while on $\mathbb{R} \setminus (k_*, k^*)$ it is equal to $\tilde{P}^{\nu}_{\mu}$.

Theorem 3.3. *Suppose $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$. Then $P_{T^{\nu}(\mu)} = (\tilde{P}^{\nu}_{\mu})^c$.*

Proof. For $\nu(\mathbb{R}) = \mu(\mathbb{R})$ we must have $\bar{\nu} = \bar{\mu}$ and $T^{\nu}(\mu) = \nu$. The result can be verified directly in this case and we exclude it from this point onwards.

The proof in the general case makes extensive use of the definitions and results of Section 2. Note first that since $\mu \leq_E \nu$, $P_{\nu} \geq P_{\mu}$ and $C_{\nu}(k) + (\mu(\mathbb{R})k - \overline{\mu}) \geq C_{\mu}(k) + (\mu(\mathbb{R})k - \overline{\mu}) = P_{\mu}(k)$. Then $\tilde{P}^{\nu}_{\mu} \geq P_{\mu} \geq 0$.

Let $P_{\chi} = (\tilde{P}^{\nu}_{\mu})^c$ be our candidate function. Since P_{χ} is convex, its second derivative (in the sense of distributions) corresponds to a measure, which we denote by χ.

We first show that $P_{T^\nu(\mu)} \leq P_\chi$. Since $P_{T^\nu(\mu)}$ is convex and P_χ is the largest convex function below $\tilde{P}^\nu_\mu$, it is enough to show that $P_{T^\nu(\mu)} \leq \tilde{P}^\nu_\mu$. Since $T^\nu(\mu) \leq \nu$, $P_{T^\nu(\mu)} \leq P_\nu$. On the other hand, since $\mu \leq_{cx} T^\nu(\mu)$, $\mu(\mathbb{R}) = T^\nu(\mu)(\mathbb{R})$ and $\overline{\mu} = \overline{T^\nu(\mu)}$, and therefore, for all $k \in \mathbb{R}$, $P_{T^\nu(\mu)}(k) = C_{T^\nu(\mu)}(k) + (\mu(\mathbb{R})k - \overline{\mu}) \leq C_\nu(k) + (\mu(\mathbb{R})k - \overline{\mu})$, where we again used that $T^\nu(\mu) \leq \nu$. Combining both cases we conclude that $P_{T^\nu(\mu)} \leq \tilde{P}^\nu_\mu$ and thus $P_{T^\nu(\mu)} \leq P_\chi$.

To finish the proof we will show that $\chi \in \mathcal{M}^\nu_\mu$, or equivalently, that $P_\chi \in \mathcal{P}(\mu, \nu)$. Then by the maximality of $T^\nu(\mu)$ we have that $\chi \leq_{cx} T^\nu(\mu)$, which implies that $P_\chi \leq P_{T^\nu(\mu)}$.

First, we already saw that $P_\mu \leq \tilde{P}^\nu_\mu$. Then, since P_μ is convex, $P_\mu \leq (\tilde{P}^\nu_\mu)^c = P_\chi \leq \tilde{P}^\nu_\mu$. Further, since P_μ is an element of $\mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$ and $\tilde{P}^\nu_\mu$ has the same limiting behaviour as an element of $\mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$, it follows that $P_\chi \in \mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$, and therefore $\mu \leq_{cx} \chi$.

It is left to show that $\chi \leq \nu$, or equivalently, that $P_\nu - P_\chi$ is convex. We have that p given by $p(k) = ((\nu(\mathbb{R}) - \mu(\mathbb{R}))k - (\bar{\nu} - \bar{\mu}))^+$ is a convex function and therefore by Lemma 2.3, with $g = P_\nu$ and $f = p$, $P_\nu - (P_\nu - p)^c$ is convex. But, using the expression in (3.1), $P_\nu - (P_\nu - p)^c = P_\nu - P_\chi$. □

Let $\mu = \mu_1 + \mu_2$ for some $\mu_1, \mu_2 \in \mathcal{M}$ and $\nu \in \mathcal{M}$ with $\mu \leq_E \nu$. Then $\mathcal{M}^\nu_{\mu_1} \neq \emptyset$ and, in particular, we can embed μ_1 into ν using any martingale coupling $\pi \in \Pi_M(\mu_1, T^\nu(\mu_1))$. A natural question is then whether $\mathcal{M}^{\nu - T^\nu(\mu_1)}_{\mu_2}$ is non-empty, so that the remaining mass μ_2 can also be embedded in what remains of ν.

Example 3.4. Let $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$ and $\nu = \frac{1}{3}(\delta_{-2} + \delta_0 + \delta_2)$. Then $\mu \leq_{cx} \nu$. Consider $\mu_1 = \frac{2}{3}\mu$ and $\mu_2 = \mu - \mu_1 = \frac{1}{3}\mu$. Then $T^\nu(\mu_1) = \frac{1}{3}(\delta_{-2} + \delta_2)$. However, $\mu_2 \leq_{cx} \nu - T^\nu(\mu_1)$ does not hold. Indeed, $\nu - T^\nu(\mu_1) = \frac{1}{3}\delta_0 \leq_{cx} \mu_2$.

As Example 3.4 demonstrates, for $\mu_1, \mu_2, \nu \in \mathcal{M}$ with $\mu_1 + \mu_2 = \mu \leq_E \nu$, if we first transport μ_1 to $T^\nu(\mu_1)$, then we cannot, in general, embed μ_2 in $\nu - T^\nu(\mu_1)$ in a way which respects the martingale property. As a consequence, for arbitrary measures in convex order we cannot expect the maximal element to *induce* a martingale coupling. In the next section we study the minimal element of $\mathcal{M}^\nu_\mu$, namely the shadow measure. The shadow measure has the property that if $\mu_1 + \mu_2 = \mu \leq_E \nu$ and we transport μ_1 to the shadow $S^\nu(\mu_1)$ of μ_1 in ν, then μ_2 is in extended convex order with what remains of ν, i.e., $\mu_2 \leq_E \nu - S^\nu(\mu_1)$.

4 The shadow measure

The goal of this section is to give a simple formula for the shadow of μ in ν via its potential, and hence to give a simple and direct proof of existence of the shadow measure.

Definition 4.1 (Shadow measure)**.** *Let $\mu, \nu \in \mathcal{M}$ and assume $\mu \leq_E \nu$. The shadow of μ in ν, denoted by $S^\nu(\mu)$, has the following properties*

1. *$\mu \leq_{cx} S^\nu(\mu) \leq \nu$,*
2. *If η is another measure satisfying $\mu \leq_{cx} \eta \leq \nu$, then $S^\nu(\mu) \leq_{cx} \eta$.*

Remark 4.2. If $\mu \leq_{cx} \nu$ then, in the light of Remark 2.1, $S^\nu(\mu) = \nu = T^\nu(\mu)$.

Proposition 4.3 (Beiglböck and Juillet [4], Lemma 4.6)**.** *For $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$, $S^\nu(\mu)$ exists and is unique.*

Given μ and ν with $\mu \leq_E \nu$ (and, by Remark 4.2, with $\mu(\mathbb{R}) < \nu(\mathbb{R})$) in this section we construct the shadow measure $S^\nu(\mu)$ via (modified) potential function $P_{S^\nu(\mu)}$. This gives an independent proof of existence; uniqueness is easy.

We begin with some preliminary lemmas.

Lemma 4.4. *Suppose $\mu, \nu \in \mathcal{M}$. The following are equivalent:*

(i) $\mu \leq_E \nu$;

(ii) *there exists* $\eta \in \mathcal{M}$ *such that* $\mu \leq_{cx} \eta \leq \nu$;

(iii) *there exists* $\chi \in \mathcal{M}$ *such that* $\mu \leq \chi \leq_{cx} \nu$.

Proof. That (i) implies (ii) follows from Theorem 3.3 (or from Proposition 4.3, but then this paper is not self-contained).

For (ii) implies (iii), let η be as in part (ii). Then it is easily verified that χ, defined by $\chi = \nu - \eta + \mu$, satisfies $\chi \in \mathcal{M}$ and $\mu \leq \chi \leq_{cx} \nu$. (To prove (iii) implies (ii) reverse the roles of η and χ, i.e., take χ as in part (iii) and define $\eta = \nu - \chi + \mu$.)

Finally we show that (iii) implies (i). Suppose $\mu \leq \chi \leq_{cx} \nu$. Then, for any non-negative and convex $f : \mathbb{R} \mapsto \mathbb{R}_+$,

$$\int f d\mu \leq \int f d\chi \leq \int f d\nu,$$

and thus, from the definition of $\leq_E$ in Section 2, $\mu \leq_E \nu$. (The proof of (ii) implies (i) is identical.) $\square$

Corollary 4.5. *Let $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$. Define $\beta \in \mathcal{M}$ by $\beta := (\nu(\mathbb{R}) - \mu(\mathbb{R}))\delta_{\frac{\overline{\nu}-\overline{\mu}}{\nu(\mathbb{R})-\mu(\mathbb{R})}}$, so that $P_\beta(k) = ((\nu(\mathbb{R}) - \mu(\mathbb{R}))k - (\overline{\nu} - \overline{\mu}))^+$.*

Then $(P_\nu - P_\mu) \geq P_\beta$.

Proof. Let $\eta \in \mathcal{M}$ be as in (ii) of Lemma 4.4. By the remarks in Section 2, since $\mu \leq_{cx} \eta$, we have that $\mu(\mathbb{R}) = \eta(\mathbb{R})$ and $\overline{\mu} = \overline{\eta}$, and therefore $(\nu - \eta)(\mathbb{R}) = \beta(\mathbb{R})$ and $\overline{(\nu - \eta)} = \overline{\nu} - \overline{\mu} = \overline{\beta}$. Since a point mass is smaller in convex order than any other distribution with the same mass and mean, it follows that $\beta \leq_{cx} \nu - \eta$. Then $P_\beta \leq P_{\nu-\eta} = P_\nu - P_\eta \leq P_\nu - P_\mu$. $\square$

The next result follows when Lemma 2.5 with $g = P_\nu$ and $f = P_\mu$ is combined with Corollary 4.5:

Corollary 4.6. *Suppose $\mu \leq_E \nu$. Then $(P_\nu - P_\mu)^c \in \mathcal{D}(\nu(\mathbb{R}) - \mu(\mathbb{R}), \overline{\nu} - \overline{\mu})$.*

Note that if $\mu \leq_E \nu$ and $\mu(\mathbb{R}) = \nu(\mathbb{R})$ then $\mu \leq_{cx} \nu$ and $\overline{\mu} = \overline{\nu}$ so that $(P_\nu - P_\mu)^c$ is the zero function which is the unique element in $\mathcal{D}(0, 0)$.

Theorem 4.7. *Let $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$. Then the shadow of μ in ν is uniquely defined and given by*

$$P_{S^\nu(\mu)} = P_\nu - (P_\nu - P_\mu)^c. \tag{4.1}$$

Proof. Rephrasing Definition 4.1 above for the shadow measure (and splitting the first element into three parts), a function h is the potential of the shadow of μ in ν if

1 $h \in \mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$,

$1'$ $(P_\nu - h) \leq (P_\nu - P_\mu)$,

$1''$ $P_\nu - h$ is a potential function, i.e., $P_\nu - h \in \mathcal{D}(\alpha, \beta)$ for some $\alpha \geq 0, \beta \in \mathbb{R}$,

2 If p is another potential function with properties $1, 1', 1''$ then $(P_\nu - h) \geq (P_\nu - p)$.

By Corollary 4.6 we have $(P_\nu - P_\mu)^c \in \mathcal{D}(\nu(\mathbb{R}) - \mu(\mathbb{R}), \overline{\nu} - \overline{\mu})$. Now set $h = P_\nu - (P_\nu - P_\mu)^c$. First we verify that $h \in \mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$. By applying Lemma 2.3, with $g = P_\nu$ and $f = P_\mu$, we have that h is convex and therefore $h = h^c$. Then, since $h \geq P_{\mu(\mathbb{R})\delta_{\overline{\mu}/\mu(\mathbb{R})}} \geq 0$, applying Lemma 2.5 with $g = P_\nu$ and $f = (P_\nu - P_\mu)^c$ we conclude that $h \in \mathcal{D}(\mu(\mathbb{R}), \overline{\mu})$.

Now we claim that $P_\nu - h = (P_\nu - P_\mu)^c$ satisfies the properties $1', 1'', 2$. We already saw that $(P_\nu - P_\mu)^c \in \mathcal{D}((\nu(\mathbb{R}) - \mu(\mathbb{R})), \overline{\nu} - \overline{\mu})$, i.e., $(P_\nu - P_\mu)^c$ is a potential function, and thus property $1''$ is satisfied. On the other hand, properties $1'$ and 2 follow from the definition and the maximality of the convex hull, respectively. $\square$

We now turn to the associativity of the shadow measure. As alluded to in the introduction, it is one of the most important results on the structure of shadows. The proof of the associativity (Theorem 4.8) given in Beiglböck and Juillet [4] is delicate and based on the approximation of μ by atomic measures. Thanks to Theorem 4.7, we are able to provide a simple proof of Theorem 4.8.

Theorem 4.8 (Beiglböck and Juillet [4], Theorem 4.8)**.** *Suppose $\mu = \mu_1 + \mu_2$ for some $\mu_1, \mu_2 \in \mathcal{M}$ and $\mu \leq_E \nu$. Then $\mu_2 \leq_E \nu - S^\nu(\mu_1)$ and*

$$S^\nu(\mu_1 + \mu_2) = S^\nu(\mu_1) + S^{\nu - S^\nu(\mu_1)}(\mu_2). \tag{4.2}$$

Proof. We first prove that $\mu_2 \leq_E \nu - S^\nu(\mu_1)$. Define $P_\theta : \mathbb{R} \to \mathbb{R}_+$ by

$$P_\theta(k) = (P_\nu - P_{\mu_1})^c(k) - ((P_\nu - P_{\mu_1})^c - P_{\mu_2})^c(k), \quad k \in \mathbb{R}.$$

We will show that $P_\theta \in \mathcal{P}(\mu_2, \nu - S^\nu(\mu_1))$. Then the second derivative of P_θ corresponds to a measure $\theta \in \mathcal{M}_{\mu_2}^{\nu - S^\nu(\mu_1)}$, which by Lemma 4.4 is enough to prove the assertion.

Convexity of P_θ is a direct consequence of Lemma 2.3 with $g = (P_\nu - P_{\mu_1})^c$ and $f = P_{\mu_2}$. Moreover, since $P_{\nu - S^\nu(\mu_1)} = P_\nu - P_{S^\nu(\mu_1)} = (P_\nu - P_{\mu_1})^c$, we have that

$$P_{\nu - S^\nu(\mu_1)} - P_\theta = ((P_\nu - P_{\mu_1})^c - P_{\mu_2})^c \leq (P_\nu - P_{\mu_1})^c - P_{\mu_2},$$

and it follows that $(P_{\nu - S^\nu(\mu_1)} - P_\theta)$ is convex and $P_{\mu_2} \leq P_\theta$. To prove that $\mu_2 \leq_E \nu - S^\nu(\mu_1)$ it only remains to show that P_θ has the correct limiting behaviour to ensure that $P_\theta \in \mathcal{D}(\mu_2(\mathbb{R}), \overline{\mu_2})$.

First, since $\mu_1 \leq_E \nu$, by Corollary 4.6 we have that $(P_\nu - P_{\mu_1})^c \in \mathcal{D}(\nu(\mathbb{R}) - \mu_1(\mathbb{R}), \overline{\nu} - \overline{\mu_1})$. Similarly, since $\mu_1 + \mu_2 \leq_E \nu$, $(P_\nu - P_{\mu_1} - P_{\mu_2})^c \in \mathcal{D}(\nu(\mathbb{R}) - \mu_1(\mathbb{R}) - \mu_2(\mathbb{R}), \overline{\nu - \mu_1 - \mu_2})$. But, by Lemma 2.4, with $f = (P_\nu - P_{\mu_1})$ and $g = P_{\mu_2}$, we have that $((P_\nu - P_{\mu_1})^c - P_{\mu_2})^c = (P_\nu - P_{\mu_1} - P_{\mu_2})^c$. Finally, recall that $P_\theta \geq P_{\mu_2} \geq P_{\mu_2(\mathbb{R})\delta_{\overline{\mu_2}/\mu_2(\mathbb{R})}}$ and, since P_θ is convex, $P_\theta = P_\theta^c$. Therefore, by applying Lemma 2.5 with $g = (P_\nu - P_{\mu_1})^c$ and $f = ((P_\nu - P_{\mu_1})^c - P_{\mu_2})^c$, we conclude that $P_\theta \in \mathcal{D}(\mu_2(\mathbb{R}), \overline{\mu_2})$.

We are left to prove the associativity property (4.2). By applying Theorem 4.7 to $S^\nu(\mu_1 + \mu_2)$ we have by Lemma 2.4 that

$$P_{S^\nu(\mu_1+\mu_2)} = P_\nu - ((P_\nu - P_{\mu_1}) - P_{\mu_2})^c = P_\nu - \left((P_\nu - P_{\mu_1})^c - P_{\mu_2}\right)^c,$$

whilst applying Theorem 4.7 to $S^\nu(\mu_1)$ and $S^{\nu - S^\nu(\mu_1)}(\mu_2)$ gives

$$\begin{aligned} P_{S^\nu(\mu_1)} + P_{S^{\nu - S^\nu(\mu_1)}(\mu_2)} &= \{P_\nu - (P_\nu - P_{\mu_1})^c\} + \left\{P_{\nu - S^\nu(\mu_1)} - (P_{\nu - S^\nu(\mu_1)} - P_{\mu_2})^c\right\} \\ &= P_\nu - \left((P_\nu - P_{\mu_1})^c - P_{\mu_2}\right)^c, \end{aligned}$$

where we again used that $P_{\nu - S^\nu(\mu_1)} = P_\nu - P_{S^\nu(\mu_1)} = (P_\nu - P_{\mu_1})^c$. $\square$

We give one further result which is easy to prove using Theorem 4.7 and which describes a structural property of the shadow. The purpose of including this result is to show that the potential method can be used to give new results, as well as to give shorter proofs of existing ones.

Proposition 4.9. *Suppose $\xi, \mu, \nu \in \mathcal{M}$ with $\xi \leq \mu \leq_E \nu$. Then, $\xi \leq_E \nu$, $\xi \leq_E S^\nu(\mu)$ and*

$$S^{S^\nu(\mu)}(\xi) = S^\nu(\xi).$$

Proof. Let f be non-negative and convex. Then, $\int f d\xi \leq \int f d\mu \leq \int f dS^\nu(\mu) \leq \int f d\nu$ and both $\xi \leq_E S^\nu(\mu)$ and $\xi \leq_E \nu$. By Theorem 4.7, we have $P_{S^\nu(\xi)} = P_\nu - (P_\nu - P_\xi)^c$. Applying Theorem 4.7 to $P_{S^{S^\nu(\mu)}(\xi)}$, and writing $\sigma = \mu - \xi$ so that $P_\mu = P_\xi + P_\sigma$,

$$P_{S^{S^\nu(\mu)}(\xi)} = P_{S^\nu(\mu)} - (P_{S^\nu(\mu)} - P_\xi)^c = \Big\{P_\nu - (P_\nu - P_\xi - P_\sigma)^c\Big\} - \Big(P_\nu - (P_\nu - P_\xi - P_\sigma)^c - P_\xi\Big)^c.$$

Three applications of Lemma 2.4 give that

$$P_{S^{S^\nu(\mu)}(\xi)} = \Big\{P_\nu - ((P_\nu - P_\xi)^c - P_\sigma)^c\Big\} - \Big((P_\nu - P_\xi)^c - ((P_\nu - P_\xi)^c - P_\sigma)^c\Big)^c.$$

Finally, by Lemma 2.3 (with $g = (P_\nu - P_\xi)^c$ and $f = P_\sigma$) we have that $(P_\nu - P_\xi)^c - ((P_\nu - P_\xi)^c - P_\sigma)^c$ is convex and then

$$\begin{aligned} P_{S^{S^\nu(\mu)}(\xi)} &= \Big\{P_\nu - ((P_\nu - P_\xi)^c - P_\sigma)^c\Big\} - \Big((P_\nu - P_\xi)^c - ((P_\nu - P_\xi)^c - P_\sigma)^c\Big) \\ &= P_\nu - (P_\nu - P_\xi)^c = P_{S^\nu(\xi)}. \end{aligned}$$

□

Example 4.10. The assertion of Proposition 4.9 does not hold for $\xi, \mu, \nu \in \mathcal{M}$ with $\xi \leq_E \mu \leq_E \nu$. To see this, let $\xi = \frac{1}{3}\delta_0$, $\mu = \frac{1}{3}(\delta_{-2} + \delta_2)$ and $\nu = \frac{1}{3}(\delta_{-2} + \delta_0 + \delta_2)$. Then $S^\nu(\mu) = \mu$ and $S^{S^\nu(\mu)}(\xi) = S^\mu(\xi) = \frac{1}{6}(\delta_{-2} + \delta_2) \neq \xi = S^\nu(\xi)$.

5 Proofs

Proof of Lemma 2.2. If $f^c \equiv -\infty$ then it is linear and we are done. Henceforth we exclude this case.

Suppose f^c is not a straight line on (a, b). Then, by the convexity of f^c, for all $x \in (a, b)$ we have

$$f^c(x) < \frac{b - x}{b - a} f^c(a) + \frac{x - a}{b - a} f^c(b) = L^{f^c}_{a,b}(x).$$

Let $\eta = \inf_{u \in (a,b)}\{f(u) - L^{f^c}_{a,b}(u)\}$. If $\eta \geq 0$ then $f \geq L^{f^c}_{a,b}$ on (a, b) and $f \geq f^c \vee L^{f^c}_{a,b}$, contradicting the maximality of f^c as a convex minorant of f.

Now suppose that $\eta < 0$. Since f is lower semi-continuous, $f - L^{f^c}_{a,b}$ is also lower semi-continuous, and therefore attains its infimum on $[a, b]$. Fix $z \in \operatorname{arginf}_{u \in [a,b]}\{f(u) - L^{f^c}_{a,b}(u)\}$. Since $f(k) - L^{f^c}_{a,b}(k) = f(k) - f^c(k) \geq 0$ for $k \in \{a, b\}$, $a, b \notin \operatorname{arginf}_{u \in [a,b]}\{f(u) - L^{f^c}_{a,b}(u)\}$, and thus $z \in (a, b)$. Then since $f > f^c$ on (a, b) we have $0 > \eta = f(z) - L^{f^c}_{a,b}(z) > f^c(z) - L^{f^c}_{a,b}(z)$. Then $f^c \vee (L^{f^c}_{a,b} + \eta)$ is convex, is a minorant of f and is strictly larger than f^c (in particular at z) again contradicting the maximality of f^c as a convex minorant of f. □

Proof of Lemma 2.3. Let $h = g - f$. We show that $\psi = g - h^c$ is convex.

If $h^c \equiv -\infty$ then $\psi \equiv +\infty$ which is convex. Henceforth we exclude this case.

First note that, since $h^c(y) \leq h(y)$, $\psi(y) = g(y) - h^c(y) \geq f(y)$, $y \in \mathbb{R}$.

Define $\mathcal{A}_= := \{y : h(y) = h^c(y)\}$ and $\mathcal{A}_> := \{y : h(y) > h^c(y)\}$. Then $\psi = f$ on $\mathcal{A}_=$, while $\psi > f$ on $\mathcal{A}_>$.

Recall that $\phi : \mathbb{R} \mapsto \mathbb{R}$ is convex if for all $x, y, z \in \mathbb{R}$, with $x \leq y \leq z$,

$$\phi(y) \leq L^\phi_{x,z}(y).$$

Suppose $y \in \mathcal{A}_=$. Then, for all $x \leq y \leq z$, since f is convex and $\psi \geq f$,

$$\psi(y) = f(y) \leq L^f_{x,z}(y) \leq L^\psi_{x,z}(y).$$

In the rest of the proof we take $y \in \mathcal{A}_>$, and $x, z \in \mathbb{R}$ with $-\infty < x \leq y \leq z < \infty$ and show that $\psi(y) \leq L^\psi_{x,z}(y)$. Let $\mathcal{B}_y$ be the set of open intervals containing y which are subsets of $\mathcal{A}_>$. If $y \in \mathcal{A}_>$ then, by continuity of h and h^c, $\mathcal{B}_y := \{(a, b) \subseteq \mathbb{R} : y \in (a, b) \subseteq \mathcal{A}_>\}$ is non-empty. Moreover $\mathcal{B}_y$ has a largest element: $\hat{\mathcal{B}}_y := \sup \mathcal{B}_y$. Denote by $X(y)$ (resp. $Z(y)$) the left (resp. right) end-point of $\hat{\mathcal{B}}_y$. By Lemma 2.2, we have that h^c is linear on $(X(y), Z(y))$. Moreover, by continuity of h and h^c, if $X(y)$ (resp. $Z(y)$) is finite, then

$X(y) \in \mathcal{A}_=$ (resp. $Z(y) \in \mathcal{A}_=$). If both $X(y)$ and $Z(y)$ are finite then $h^c(y) = L^h_{X(y),Z(y)}(y)$. In general, $\hat{\mathcal{B}} = (X(y), Z(y)) \subseteq \mathcal{A}_>$.

Suppose $-\infty < x \leq X(y) < Z(y) \leq z < \infty$. Then, since g is convex and $h^c(y) = L^h_{X(y),Z(y)}(y)$,

$$\psi(y) = g(y) - h^c(y) \leq L^g_{X(y),Z(y)}(y) - L^h_{X(y),Z(y)}(y) = L^f_{X(y),Z(y)}(y)$$

and then

$$\psi(y) \leq L^f_{X(y),Z(y)}(y) \leq L^{L^f_{x,z}}_{X(y),Z(y)}(y) = L^f_{x,z}(y) \leq L^\psi_{x,z}(y)$$

by the convexity of f (and hence $f \leq L^f_{x,z}$ on $[x,z]$) and the fact that $f \leq \psi$.

Suppose $X(y) \leq x \leq y \leq z \leq Z(y)$. Note that we allow $X(y) = -\infty$ (resp. $Z(y) = \infty$), but in that case $x > X(y) = -\infty$ (resp. $z < Z(y) = \infty$). By Lemma 2.2, h^c is linear on (x,z), and therefore $h^c(y) = L^{h^c}_{x,z}(y)$. Using convexity of g on $\mathbb{R}$ we conclude that

$$\psi(y) = g(y) - h^c(y) \leq L^g_{x,z}(y) - L^{h^c}_{x,z}(y) = L^\psi_{x,z}(y).$$

Suppose $X(y) \leq x \leq y < Z(y) \leq z$. (The case $x \leq X(y) < y \leq z \leq Z(y)$ follows by symmetry.) Since $z < \infty$ we have that $Z(y) < \infty$, but $X(y)$ may be finite or infinite. In this case h^c is also linear on $(x, Z(y))$ and therefore $h^c(y) = L^{h^c}_{x,Z(y)}(y)$. Then

$$\psi(y) = g(y) - h^c(y) \leq L^g_{x,Z(y)}(y) - L^{h^c}_{x,Z(y)}(y) = L^\psi_{x,Z(y)}(y), \tag{5.1}$$

where the inequality follows from the convexity of g on $\mathbb{R}$. Now note that, since f is convex and $f \leq \psi$,

$$\psi(Z(y)) = g(Z(y)) - h(Z(y)) = f(Z(y)) \leq L^f_{x,z}(Z(y)) \leq L^\psi_{x,z}(Z(y)),$$

from which we conclude that $L^\psi_{x,Z(y)}(y) \leq L^\psi_{x,z}(y)$, and then combining with (5.1), $\psi(y) \leq L^\psi_{x,z}(y)$. This finishes the proof. □

Proof of Lemma 2.4. First, since $f \geq f^c$, $(f-g)^c \geq (f^c - g)^c$. On the other hand, we have $(f-g)^c \leq (f-g)$ and therefore $(f-g)^c + g \leq f$. Since the sum of two convex functions is convex, $(f-g)^c + g$ is also convex. Hence, $(f-g)^c + g \leq f^c$, and therefore $(f-g)^c \leq (f^c - g)$. Since $(f^c - g)^c$ is the largest convex function dominated by $(f^c - g)$, $(f-g)^c \leq (f^c - g)^c$. It follows that $(f-g)^c = (f^c - g)^c$. □

Proof of Lemma 2.5. Since $g \in \mathcal{D}(a,b)$ and $f \in \mathcal{D}(\alpha,\beta)$ with $g \geq f$, we have that

$$\begin{aligned} 0 \leq \lim_{k\to\infty} \{g(k) - f(k)\} &= \lim_{k\to\infty} \{g(k) - (ak - b) - f(k) + (\alpha k - \beta) + (a-\alpha)k - (b-\beta)\} \\ &= \lim_{k\to\infty} \{(a-\alpha)k - (b-\beta)\} = \lim_{k\to\infty} h(k), \end{aligned}$$

and therefore $a \geq \alpha$. Also, if $\alpha = a$ then $\beta \geq b$.

Now suppose $f \leq g$ and $g - f \geq h$. Then $g - f \geq h^+$ and since h^+ is convex, we have that $(g-f) \geq (g-f)^c \geq h^+$. Then, $\lim_{|k|\to\infty}\{g(k) - f(k) - h^+(k)\} = 0$, and it follows that $(g-f)^c \in \mathcal{D}(a-\alpha, b-\beta)$. □

References

[1] Beiglböck, M., Cox A.M.G., Huesmann, M.: *The geometry of multi-marginal Skorokhod Embedding.* Probab. Theory Relat. Fields, (2019). MR4087489

[2] Beiglböck M., Henry-Labordère P., Penkner F.: *Model-independent bounds for option prices—mass transport approach.* Finance Stoch., 17(3):477–501, (2013). MR3066985

[3] Beiglböck M., Henry-Labordère P., Touzi N.: *Monotone martingale transport plans and Skorokhod embedding.* Stochastic Process. Appl., 127(9):3005–3013, (2017). MR3682121

[4] Beiglböck M., Juillet N.: *On a problem of optimal transport under marginal martingale constraints.* Ann. Probab., 44(1):42–106, (2016). MR3456332

[5] Brenier Y.: *Décomposition polaire et réarrangement monotone des champs de vecteurs.* C. R. Acad. Sci. Paris Sér. I Math., 305(19): 805-808, (1987). MR0923203

[6] Brückerhoff M., Huesmann M., Juillet N.: *Shadow martingales–a stochastic mass transport approach to the peacock problem.* arXiv preprint, (2020). Available online at: https://arxiv.org/abs/2006.10478

[7] Chacon, R.V.: *Potential processes.* Trans. Amer. Math. Soc., 226:39–58, (1977). MR0501374

[8] Chacon, R.V., Walsh J.B.: *One-dimensional potential embedding.* In Sémin. Probab. X, vol. 511 of Lecture Notes in Mathematics, 19–23, Springer, Berlin, (1976). MR0445598

[9] Galichon A., Henry-Labordère P., Touzi N.: *A stochastic control approach to no-arbitrage bounds given marginals with an application to lookback options.* Ann. Appl. Probab., 24(1):313–336, (2014). MR3161649

[10] Henry-Labordère P., Touzi N.: *An explicit martingale version of the one-dimensional Brenier's theorem.* Finance Stoch., 20(3):635–668, (2016). MR3519164

[11] Henry-Labordère P., Tan X., Touzi N.: *An explicit martingale version of the one-dimensional Brenier's Theorem with full marginals constraint.* Stochastic Process. Appl., 126(9):2800–2834, (2016). MR3522302

[12] Hirsch F., Roynette B.: *A new proof of Kellerer's theorem.* ESAIM Probab. Stat., 16:48–60, (2012). MR2911021

[13] Hobson D.G., Klimmek M.: *Robust price bounds for the forward starting straddle.* Finance Stoch., 19(1):189–214, (2015). MR3292129

[14] Hobson D.G., Neuberger A.: *Robust bounds for forward start options.* Math. Finance, 22(1):31–56, (2012). MR2881879

[15] Hobson D.G., Norgilas D.: *The left-curtain martingale coupling in the presence of atoms.* Ann. Appl. Probab., 29(3):1904–1928, (2019). MR3914560

[16] Juillet N.: *Stability of the shadow projection and the left-curtain coupling.* Ann. Inst. Henri Poincaré Probab. Stat., 52(4):1823–1843, (2016). MR3573297

[17] Juillet N.: *Martingales associated to peacocks using the curtain coupling.* Electron. J. Probab., 23(9):1–29, (2018). MR3771745

[18] Kellerer H.G.: *Markov-Komposition und eine Anwendung auf Martingale.* Math. Ann., 198:99–122, (1972). MR0356250

[19] Kellerer H.G.: *Integraldarstellung von Dilationen.* In Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Tech. Univ., Prague, 1971; dedicated to the memory of Antonín Špaček), pp. 341–374. Academia, Prague, (1973). MR0356249

[20] Nutz M., Stebegg F.: *Canonical supermartingale couplings.* Ann. Probab., 46(6):3351–3398, (2018). MR3857858

[21] Nutz M., Stebegg F., Tan X.: *Multiperiod martingale transport.* Stochastic Process. Appl., 130(3):1568-1615, (2020). MR4058283

[22] Rockafellar R. T.: *Convex analysis.* (No. 28) Princeton Univ. Press, (1970). MR0274683

[23] Rüschendorf L., Rachev S.T.: *A characterization of random variables with minimum L^2-distance.* J. Multivariate Anal., 32:48–54, (1990). MR1035606

[24] Strassen V.: *The existence of probability measures with given marginals.* Ann. Math. Statist., 36(2):423-439, (1965). MR0177430