Electron. J. Probab. **27** (2022), article no. 147, 1–46.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP868

A construction of the left-curtain coupling

David Hobson* Dominykas Norgilas†

Abstract

In a martingale optimal transport (MOT) problem mass distributed according to the law μ is transported to the law ν in such a way that the martingale property is respected. Beiglböck and Juillet (On a problem of optimal transport under marginal martingale constraints, Annals of Probability, 44(1):42-106, 2016) introduced a solution to the MOT problem which they baptised the left-curtain coupling. The left-curtain coupling has been widely studied and shown to have many applications, including to martingale inequalities and the model-independent pricing of American options. Beiglböck and Juillet proved existence and uniqueness, proved optimality for a family of cost functions, and proved that when μ is a continuous distribution, mass at x is mapped to one of at most two points, giving lower and upper functions. Henry-Labordère and Touzi (An explicit martingale version of Brenier's theorem, Finance and Stochastics, 20:635-668, 2016) showed that the left-curtain coupling is optimal for an extended family of cost functions and gave a construction of the upper and lower functions under an assumption that μ and ν are continuous, together with further simplifying assumptions of a technical nature.

In this article we construct these upper and lower functions in the general case of arbitrary centred measures in convex order, and thereby give a complete construction of the left-curtain coupling. In the case where μ has atoms these upper and lower functions are to be interpreted in the sense of a lifted martingale.

Keywords: optimal transport; Brenier's theorem; martingales; convex order.
MSC2020 subject classifications: 60G42.
Submitted to EJP on March 15, 2022, final version accepted on October 17, 2022.
Supersedes arXiv:2102.10549.

1 Introduction

In the classical formulation of the optimal transport (OT) problem due to Kantorovich [32] one seeks a joint law π for random variables $X \sim \mu$ and $Y \sim \nu$ which, for a given cost function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, minimises $\mathbb{E}^\pi[c(X, Y)]$. The cornerstone result in $\mathbb{R}^d$ with an

*Department of Statistics, University of Warwick. E-mail: d.hobson@warwick.ac.uk
†Department of Mathematics, University of Michigan. E-mail: dnorgila@umich.edu

Euclidean cost $c(x,y) = |x-y|^2$ is Brenier's Theorem, see Brenier [11] and Rüschendorf and Rachev [38]. Brenier's Theorem states that, under some regularity assumptions, the optimal coupling is deterministic and realised by a map which corresponds to a gradient of a convex function. In the one-dimensional setting this coupling is identified with the Fréchet-Hoeffding (or quantile) coupling π_{HF}. An important feature of this coupling is that it is optimal for the large class of cost functions satisfying the Spence-Mirrlees condition $c_{xy} > 0$.

In recent years, there has been significant interest in optimal transport problems where the transport plan is constrained to be a martingale. The basic problem of martingale optimal transport (MOT) is, given probability measures μ and ν on $\mathbb{R}$ which are in convex order and a cost function c, to construct a joint law π for $X \sim \mu$ and $Y \sim \nu$ satisfying the martingale constraint $\mathbb{E}^\pi[Y|X] = X$ and such that $\mathbb{E}^\pi[c(X,Y)]$ is minimised. Such problems arise naturally in the context of robust (or model-independent) mathematical finance and were first considered by Hobson and Neuberger [25] and Hobson and Klimmek [24] for the specific (financially relevant) cost functions $c(x,y) = -|y-x|$ and $c(x,y) = |y-x|$. Subsequently, Beiglböck et al. [5] (in a discrete time setting) and Galichon et al. [17] (in continuous time) extended the problem to a more general setting, made the connection to optimal transport problems, and proved duality theorems. The class of MOT problems is of wide mathematical interest and, as well as the application to finance, is closely related to, and has important consequences for, the study of martingale inequalities (see Beiglböck and Nutz [10], Henry-Labordère et al. [18], Obłój et al. [36]) and the Skorokhod embedding problem (see Beiglböck et al. [4], Källblad et al. [33]).

Using an extension of the notion of cyclical monotonicity from the classical OT setting, Beiglböck and Juillet [8] introduced the left-monotone martingale coupling, which can be viewed as a martingale analogue of the monotone Fréchet-Hoeffding coupling. The authors then prove the existence and uniqueness of the left-monotone coupling, which they baptise the *left-curtain* martingale coupling, together with the optimality of this joint probability measure for some specific cost functions. Henry-Labordère and Touzi [19] extended their results to show that the left-curtain coupling is optimal for a wide class of payoffs (essentially[1] those satisfying $c_{xyy} < 0$). Several other authors further investigate the properties and extensions of the left-curtain coupling, see Beiglböck et al. [6, 3], Juillet [30], Nutz et al. [34, 35], Campi et al. [13]. In the case of a continuum of marginals which are increasing in convex order, Henry-Labordère et al. [20], Juillet [31] and Brückerhoff at al. [12] recently showed, amongst other things, how the left-curtain coupling can be used to construct a martingale that fits given marginals at any given time and how it solves a continuous-time version of the martingale optimal transport problem.

In the light of the Brenier's Theorem in classical OT theory, a natural question arises as to whether the aforementioned optimal martingale transport plans possess similar nice structural properties and whether they can be explicitly constructed. Due to the martingale constraint, if the marginals are such that $\mu \neq \nu$, then no martingale coupling can be realized by a single map. Instead, the best one can hope for is a binomial map. Hobson and Neuberger [25] showed that if μ is continuous then there exists a pair of increasing functions on which π_{HN}, their optimal coupling, is concentrated. In particular, if we write $\pi_{HN}(dx,dy) = \mu(dx)\pi_x^{HN}(dy)$ we find π_x^{HN} has support in a two-point set $\{g(x), f(x)\}$ and the functions $x \mapsto g(x)$ and $x \mapsto f(x)$ are both increasing. Hobson and

[1] Some authors seek to minimise $\mathbb{E}^\pi[c(X,Y)]$, whereas others seek to maximise, and therefore some care is needed when moving between articles. Henry-Labordère and Touzi [19] maximise, and therefore the condition they state is $c_{xyy} > 0$, but this becomes the condition in the main text if we move to $-c$ and a minimisation problem.

Klimmek [24], on the other hand, worked under the dispersion assumption and showed that their coupling π_{HK} is such that each portion of mass from initial law μ is mapped to ν by splitting it into three points at most. In particular, π_{HK} is supported on the diagonal and the graphs of two explicitly constructed decreasing functions. When μ is atom free, Beiglböck and Juillet [8] showed that for the left-curtain coupling π_{lc} there exist lower and upper functions T_d, T_u with $T_d(x) \leq x \leq T_u(x)$ such that if we write $\pi_{lc}(dx, dy) = \mu(dx)\pi^{lc}_x(dy)$ then π^{lc}_x has support concentrated on the set $\{T_d(x), T_u(x)\}$. Here T_u is non-decreasing while T_d satisfies a particular left-monotonicity property.

There are two methods in the literature used to obtain the upper and lower functions that characterise the left-curtain coupling π_{lc}. The first method is non-constructive and is to approximate μ by a family of discrete measures and to consider the limits of the resulting upper and lower functions. The second method requires additional regularity assumptions on μ and ν and is to characterise T_d and T_u via differential equations (or integral equations). In the case where both marginals μ and ν are continuous (and satisfy a further particular structural property), Henry-Labordère and Touzi [19] (among other things) construct the upper and lower functions T_d, T_u as the solutions of a pair of coupled ordinary differential equations. Indeed, after some further ingenious manipulations, they show that T_d is the root of an integral equation. However, there are several barriers which make it difficult to extend the construction via differential equations to the general case. First, it cannot cope with atoms in the initial or terminal laws, and requires both measures to have positive densities. Second, it requires a starting condition to initialise the differential equations. Third, the method works by solving for T_d and T_u on a family of intervals, but there may be countably many such intervals, and the set of right-endpoints of these intervals may have (countably many) accumulation points beyond each of which it is difficult to extend the solutions to the differential equations. The same sorts of issues apply to the construction via the integral equation of Henry-Labordère and Touzi [19]

The first point is absolutely fundamental. When μ has an atom at x the probability kernel $\pi^x_{lc}(\cdot)$ in the disintegration $\pi_{lc}(dx, dy) = \mu(dx)\pi^x_{lc}(dy)$ becomes a measure with support on non-trivial subsets of $\mathbb{R}$ and not just on a two-point set. In this case T_d, T_u cannot be constructed, unless we allow them to be multi-valued. By changing the viewpoint, Hobson and Norgilas [26] showed how to recover the property that Y takes values in a two-point set. The idea is to write $X = G_\mu(U)$, where G_μ is a quantile function of μ and $U \sim U[0, 1]$, and then to seek functions R, S satisfying certain monotonicity properties such that $Y \in \{R(U), S(U)\}$. While T_d and T_u are multivalued on the atoms of μ, R and S remain well defined. Hobson and Norgilas explicitly constructed such R and S in the case the initial law μ is finitely supported (while ν is arbitrary), and then by approximating general μ with atomic probability measures showed that the limiting functions give rise to a generalised, or *lifted*, left-curtain martingale coupling.

The goal of this paper is *to give a direct, geometric construction of the lifted left-curtain martingale coupling for general measures μ and ν*. In particular we construct R and S (and (T_d, T_u) in the case where μ is continuous). Our methods rely neither on differential equations nor on the delicate approximation of measures, but rather on a representation of the initial and target laws via potentials. The potential of a measure involves integrating the measure against a test function and this has a smoothing effect. It is this extra smoothness which allows us to give a global construction of the key quantities. Nonetheless, some delicate arguments are needed to prove that the quantities we construct have the appropriate monotonicity properties and do indeed yield the left-curtain martingale coupling, especially since we place no assumptions on the initial or terminal laws.

The power of the potential representation is well-recognised in related settings.

There is a close connection between martingale optimal transport and solutions of the Skorokhod embedding problem (SEP) for Brownian motion (especially for non-trivial initial laws), based on the idea of viewing a martingale as a time-change of Brownian motion. One productive source of elegant solutions to the SEP is the potential-theoretic representation of measures and a geometric description due to Chacon and Walsh [15]. Many of the classical solutions of the Skorokhod embedding problem (and therefore many of the martingale optimal transports which have been proposed in the literature) can be described by drawing tangents (or supporting hyperplanes in the atomic case) on a suitable picture, see Chacon and Walsh [15] and Hobson [22]. The constructions of the SEP due to Dubins [16], Azéma and Yor [1], Jacka [29], Vallois [40], Hobson [23], Hobson and Pedersen [28] (at least) have a representation in this form. The geometric approach has a clear advantage in bypassing many of the technical issues which arise in approximation arguments.

The classical result by Strassen [39] states that it is possible to transport μ to ν using a martingale if and only if the marginal laws respect the convex order condition $\mu \leq_{cx} \nu$ (i.e., μ is less than ν in convex order). In order to study transport plans in the martingale setting, Beiglböck and Juillet [8] introduced the notion of *extended* convex order of two measures, denoted by $\leq_E$, which compares measures of possibly different total mass. If a pair of measures μ, ν is such that $\mu \leq_E \nu$, than 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, and each such η corresponds to a terminal law of a martingale that embeds μ into ν. Beiglböck and Juillet [8] proved that there exists a canonical choice of such η with respect to $\leq_{cx}$. In particular there exists the unique measure $S^\nu(\mu)$, the *shadow* of μ in ν, that is the smallest measure with respect to convex order among measures η satisfying $\mu \leq_{cx} \eta \leq \nu$. Our interest in the shadow measure lies in the fact that the left-curtain martingale coupling can be 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 $S^\nu(\mu|_{(-\infty,x]})$. (For other martingale transports arising using the shadow measure, but different parametrisations of μ, see Beiglböck and Juillet [9].) Recently Beiglböck et al. [7] showed how to explicitly construct, via potential-geometric arguments, the shadow measure $S^\nu(\mu)$ for arbitrary μ and ν with $\mu \leq_E \nu$. This turns out to be the main ingredient of our construction of the upper and lower functions characterising the left-curtain coupling.

One feature of our construction is that in the case where μ is continuous (respectively the general case) for each $x \in \mathbb{R}$ (respectively $u \in (0,1)$) we find $\{T_d(x), T_u(x)\}$ (respectively $\{R(u), S(u)\}$) by considering the convex hull of a certain function. Since we can do this for each x individually, the construction for a single x immediately extends to a construction for all x and to pairs of functions $\{T_d(\cdot), T_u(\cdot)\}$ (with a similar conclusion for the lifted martingale). In some applications, for example the model-independent pricing of American options (see Hobson and Norgilas [26]), it is sufficient to find a triple $(x, T_d(x), T_u(x))$ associated with the left-curtain coupling and with one further property, and the full structure of the construction of the left-curtain coupling is not required. Then the direct nature of our construction is particularly useful.

In recent work Bayraktar et al. [2] provide a geometric construction of the so-called *increasing* supermartingale coupling π_I (introduced by Nutz and Stebegg [34]), which builds upon and extends the construction presented in this article. (π_I can be viewed as supermartingale counterpart of the left-curtain coupling π_{lc}). The increasing coupling π_I has a feature that it behaves as π_{lc} on a certain part of the space, and thus the construction of Bayraktar et al. [2] heavily relies on the ability to construct π_{lc} on these 'martingale' intervals. This is achieved in the present paper and in full generality.

The paper is structured as follows. In Section 2 we discuss the relevant notions of

probability measures and (extended) convex order, and discuss some important (for our main theorems) results regarding the convex hull of a function. In Section 3 we introduce the shadow measure and the left-curtain martingale coupling. Sections 4, 6 and 7 are dedicated to our main results. (Section 5 contains preliminary results for Sections 6 and 7.) In Section 4 we construct a candidate pair of functions that characterise the left-curtain coupling (Theorem 4.9), while in Section 6 we study their regularity properties. Finally in Section 7, we show (Theorem 7.9) that our construction yields the (lifted) left-curtain martingale coupling. Some proofs are deferred until the appendix.

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 $\bar{\eta} = \int_{\mathbb{R}} x\eta(dx)$ to be the first moment of η (and then $\bar{\eta}/\eta(\mathbb{R})$ is the barycentre of η). The support of $\eta \in \mathcal{M}$ is denoted by $\mathrm{supp}(\eta)$; it is the smallest closed set $E \subseteq \mathbb{R}$ with $\eta(E) = \eta(\mathbb{R})$. Let $\ell_\eta := \inf\{k \in \mathrm{supp}(\eta)\}$ and $r_\eta := \sup\{k \in \mathrm{supp}(\eta)\}$. Let $\mathcal{I}_\eta$ be the smallest interval containing the support of η, so that $\{\ell_\eta, r_\eta\}$ are 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 $\eta \in \mathcal{M}$, let $F_\eta : \mathbb{R} \to [0, \eta(\mathbb{R})]$ and $G_\eta : (0, \eta(\mathbb{R})) \to \mathbb{R}$ be the distribution and quantile functions of η, respectively. As usual F_η is right-continuous, however we take an arbitrary version of G_η until further notice. (In Sections 6 and 7 we will work with a left-continuous version of G_η.)

For $\alpha \geq 0$ and $\beta \in \mathbb{R}$, let $\mathcal{D}(\alpha, \beta)$ denote the set of non-negative, non-decreasing convex functions $f : \mathbb{R} \to \mathbb{R}_+$ such that

$$\lim_{z \downarrow -\infty} f(z) = 0, \qquad \lim_{z \uparrow \infty} \{f(z) - (\alpha z - \beta)\} = 0.$$

When $\alpha = 0$, $\mathcal{D}(0, \beta)$ is empty unless $\beta = 0$ and then $\mathcal{D}(0, 0)$ contains one element.

For $\eta \in \mathcal{M}$, define the function $P_\eta : \mathbb{R} \to \mathbb{R}_+$ by

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

(The notation $P_\cdot$ arises from the connection with the expected payoff of a *put option*.) The following properties of P_η can be found in Chacon [14], and Chacon and Walsh [15]: $P_\eta \in \mathcal{D}(\eta(\mathbb{R}), \bar{\eta})$ and $\{k : P_\eta(k) > (\eta(\mathbb{R})k - \bar{\eta})^+\} = (\ell_\eta, r_\eta)$. Conversely (see, for example, Hirsch et al. [21, Proposition 2.1]), if $h \in \mathcal{D}(k_m, k_f)$ 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 $\bar{\eta} = k_f$, such that $h = P_\eta$. In particular, η is uniquely identified by the second derivative of h in the sense of distributions. Note that P_η is related to the potential U_η, defined by

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

via $P_\eta(k) = \frac{1}{2}(-U_\eta(k) + (\eta(\mathbb{R})k - \bar{\eta}))$. We will call P_η a modified potential. Finally note that both second derivatives 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

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

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

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

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 ($\bar{\eta} = \bar{\chi}$). Moreover, necessarily we must have $\ell_\chi \leq \ell_\eta \leq r_\eta \leq r_\chi$. From simple approximation arguments (see Hirsch et al. [21]) we also have that, if η and χ have the same total mass and the same first moment, 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 [8] and say $\eta, \chi \in \mathcal{M}$ are in an *extended* convex order, and write $\eta \leq_E \chi$, if

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

The partial order $\leq_E$ generalises both $\leq$ and $\leq_{cx}$ in the sense that it preserves existing orderings and gives rise to new ones. 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 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

$$\Pi_M(\eta, \chi) = \big\{\pi \in \Pi(\eta, \chi) : (2.2) \text{ holds}\big\},$$

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.

If we ignore the martingale requirement (2.2), it is easy to see that the set of probability measures with given marginals is non-empty, i.e. $\Pi(\eta, \chi) \neq \emptyset$ (consider the product measure $\eta \otimes \chi$). However, the fundamental question whether, for given η and χ, the set of martingale couplings $\Pi_M(\eta, \chi)$ is non-empty, is more delicate. For any $\pi \in \Pi_M(\eta, \chi)$ and convex $f : \mathbb{R} \to \mathbb{R}$, 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 [39] 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 a pair of measures $\eta, \chi \in \mathcal{M}$, let the function $D = D_{\eta,\chi} : \mathbb{R} \to \mathbb{R}$ be defined by $D_{\eta,\chi}(k) = P_\chi(k) - P_\eta(k)$. Note that if η, χ have equal mass and equal first moment then $\eta \leq_{cx} \chi$ is equivalent to $D \geq 0$ on $\mathbb{R}$. Let (ℓ_D, r_D) be the smallest interval containing $\{k : D_{\eta,\chi}(k) > 0\}$; let $\mathcal{I}_D$ be the open interval (ℓ_D, r_D) together with $\{\ell_D\}$ if $\ell_D > -\infty$ and $D'(\ell_D+) := \lim_{k \downarrow \ell_D} (D(k) - D(\ell_D))/(k - \ell_D) > 0$ and $\{r_D\}$ if $r_D < \infty$ and $D'(r_D-) :=$

$\lim_{k\uparrow\ell_D}(D(k)-D(\ell_D))/(k-\ell_D)<0$. Note that, if $\eta\leq_c\chi$, then $\ell_\nu\leq\ell_\mu\leq r_\mu\leq r_\nu$ and $\mathcal{I}_D\subseteq[\ell_\nu,r_\nu]$.

The following result (see Hobson [23, page 254] or Beiglböck and Juillet [8, Section A.1]) tells us that, if $D_{\eta,\chi}(x)=0$ for some x, then in any martingale coupling of η and χ no mass can cross x.

Lemma 2.1. *Suppose η and χ are probability measures with $\eta\leq_{cx}\chi$. Suppose that $D(x)=0$. If $\pi\in\Pi_M(\eta,\chi)$, then we have $\pi((-\infty,x),(x,\infty))+\pi((x,\infty),(-\infty,x))=0$.*

It follows from Lemma 2.1 that, if there is a point x in the interior of the interval $\mathcal{I}_\eta$ such that $D_{\eta,\chi}(x)=0$, then we can separate the problem of constructing martingale couplings of η to χ into a pair of subproblems involving mass to the left and right of x, respectively, always taking care to allocate mass of χ at x appropriately. Indeed, if there are multiple $\{x_j\}_{j\geq1}$ with $D_{\eta,\chi}(x_j)=0$, then we can divide the problem into a sequence of 'irreducible' problems, each taking place on an interval $\mathcal{I}_i$ such that $D>0$ on the interior of $\mathcal{I}_i$ and $D=0$ at the endpoints. All mass starting in a given interval is transported to a point in the same interval. Moreover, by the martingale property, any mass starting at a finite endpoint of $\mathcal{I}_i$ must stay there. Putting this together we may restrict attention to intervals I on which $D>0$ (with $\lim_{x\to e_I}D(x)=0$ at endpoints e_I of I), and we may assume that the starting law has support within the interior of I and the target law has support within the closure of I (and I is the smallest set with this last property).

Notation: For $x\in\mathbb{R}$ let δ_x denote the unit point mass at x. For real numbers c,x,d with $c\leq x\leq d$ define the probability measure $\chi_{c,x,d}$ by $\chi_{c,x,d}=\frac{d-x}{d-c}\delta_c+\frac{x-c}{d-c}\delta_d$ with $\chi_{c,x,d}=\delta_x$ if $(d-x)(x-c)=0$. Note that $\chi_{c,x,d}$ has mean x and is the law of a Brownian motion started at x evaluated on the first exit from (c,d).

2.2 Convex hull

Our key results will make extensive use of the convex hull. For $f:\mathbb{R}\to\mathbb{R}$ 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 $x\leq z$, and define $L^f_{x,z}:\mathbb{R}\to\mathbb{R}$ by

$$L^f_{x,z}(y)=\begin{cases}f(x)+\frac{f(z)-f(x)}{z-x}(y-x), & \text{if } x<z,\\ f(x), & \text{if } x=z.\end{cases}\tag{2.3}$$

Then (see Rockafellar [37, Corollary 17.1.5]),

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

(Note that for (2.4), the definition of $L^f_{x,z}$ outside $[x,z]$ is irrelevant and we could restrict the domain of $L^f_{x,z}$ to $[x,z]$. However, in Sections 4, 6 and 7 we will need $L^f_{x,z}$ to be defined on $\mathbb{R}$.)

Moreover, it is not hard to see (at least geometrically, 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 Appendix A.1.)

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

In the sequel, for a given function f and $y\in\mathbb{R}$, we will want to identify the values $x,z\in\mathbb{R}$ with $x<y<z$ which attain the infimum in (2.4). For this, however, we need to

allow x and z to take values in the extended real line. Therefore we extend the definition of $L^f_{x,z}$ to $L^f_{-\infty,z}$ and $L^f_{x,\infty}$ by taking appropriate limits in (2.3). In particular, we define, for each $y \in \mathbb{R}$,

$$L^f_{-\infty,z}(y) = (z-y)\liminf_{x\downarrow-\infty}\frac{f(x)}{|x|} + f(z),$$
$$L^f_{x,\infty}(y) = f(x) + (y-x)\liminf_{z\uparrow\infty}\frac{f(z)}{z}.$$

Let $\phi = \liminf_{z\uparrow\infty}\frac{f(z)}{z} \in [-\infty,\infty]$, $\psi = \liminf_{x\downarrow-\infty}\frac{f(x)}{|x|} \in [-\infty,\infty]$, and if $\phi \in (-\infty,\infty)$, $\gamma = \inf_{w\in\mathbb{R}}(f(w)-\phi w)$. Then we also define $L^f_{-\infty,\infty}$ by

$$L^f_{-\infty,\infty}(y) = \begin{cases} \infty, & \psi+\phi > 0, \\ \gamma+\phi y & \psi+\phi = 0, \\ -\infty, & \psi+\phi < 0, \end{cases} \quad y \in \mathbb{R},$$

with the convention that $-\infty+\infty = \infty+(-\infty) = -\infty$.

Let $\mathcal{B}(y) = \{(x,z) : -\infty \le x < y < z \le \infty\}$ be the set of open intervals containing y.

Definition 2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function and f^c denote its convex hull. For $y \in \mathbb{R}$, define*

$$X^f(y) = X(y) = \sup\{x : x \le y, f^c(x) = f(x)\},$$
$$Z^f(y) = Z(y) = \inf\{z : z \ge y, f^c(z) = f(z)\},$$

with the convention that $\sup\emptyset = -\infty$ and $\inf\emptyset = \infty$.

The following result is a slight extension of (2.4).

Lemma 2.4. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. Then for $y \in \mathbb{R}$ we have $f^c(y) = L^f_{X(y),Z(y)}(y)$.*

We will need one more result regarding the locations of X and Z.

Lemma 2.5. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. If $(x,z) \in \mathcal{B}(y)$ and $f(k) > L^f_{x,z}(k)$ for all $k \in (x,z)$, then*

$$X(y) \le x \quad \text{and} \quad Z(y) \ge z.$$

3 The shadow measure and left-curtain martingale coupling

3.1 The shadow measure

Given two measures $\mu,\nu \in \mathcal{M}$ with $\mu \le_E \nu$ let $\mathcal{T}(\mu,\nu) = \{\theta \in \mathcal{M} : \mu \le_{cx} \theta \le \nu\}$. Then $\mathcal{T}(\mu,\nu)$ represents the set of all possible target measures in ν to which we can transport μ using a martingale. We are interested in the smallest element of $\mathcal{T}(\mu,\nu)$ with respect to convex order.

Definition 3.1 (Shadow measure)**.** *Let $\mu,\nu \in \mathcal{M}$ and assume $\mu \le_E \nu$. The shadow of μ in ν, denoted by $S^\nu(\mu)$, has the following properties*

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

Beiglböck and Juillet [8, Proposition 4.4 and Lemma 4.6] proved the existence and uniqueness of the shadow measure $S^\nu(\mu)$. Furthermore, a result of Beiglböck et al. [7] says that the modified potential of the shadow measure, $P_{S^\nu(\mu)}$, can be explicitly constructed, and then $S^\nu(\mu)$ is identified as the second derivative of $P_{S^\nu(\mu)}$ in the sense of distributions.

Theorem 3.2 (Beiglböck et al. [7, Theorem 1])**.** *Let $\mu, \nu \in \mathcal{M}$ with $\mu \leq_E \nu$. Then*

$$P_{S^\nu(\mu)} = P_\nu - (P_\nu - P_\mu)^c. \tag{3.1}$$

Corollary 3.3. *If $(P_\nu - P_\mu)^c$ is linear on $[a, b]$ then $\nu - S^\nu(\mu)$ does not charge (a, b).*

3.2 The left-curtain coupling π_{lc} for continuous μ

The left-curtain martingale coupling (introduced by Beiglböck and Juillet [8]), and denoted by π_{lc}, is a martingale coupling that arises via the shadow measure, created working from left to right. More specifically (see Beiglböck and Juillet [8, Theorem 4.18]), π_{lc} is the unique measure in $\Pi_M(\mu, \nu)$ which for each $x \in \mathbb{R}$ transports $\mu|_{(-\infty,x]}$ to the shadow $S^\nu(\mu|_{(-\infty,x]})$. In other words, the first and second marginals of $\pi_{lc}|_{(-\infty,x]\times\mathbb{R}}$ are $\mu|_{(-\infty,x]}$ and $S^\nu(\mu|_{(-\infty,x]})$, respectively, for each x. Furthermore, as a consequence of the minimality with respect to convex order, π_{lc} is also the unique martingale coupling which is left-monotone in the sense of Definition 3.4 (see Beiglböck and Juillet [8, Theorem 5.3]):

Definition 3.4. *A transport plan $\pi \in \Pi(\mu, \nu)$ is said to be left-monotone if there exists $\Gamma \in \mathcal{B}(\mathbb{R}^2)$ with $\pi(\Gamma) = 1$ and such that, if $(x, y^-), (x, y^+), (x', y') \in \Gamma$ we cannot have $x < x'$ and $y^- < y' < y^+$.*

When the initial law μ is continuous, the left-curtain coupling has a rather simple representation. In particular, for $x \in \mathbb{R}$, the element $\pi^x_{lc}(\cdot)$ in the disintegration $\pi_{lc}(dx, dy) = \mu(dx)\pi^x_{lc}(dy)$ is a measure supported on a set of at most two points.

Lemma 3.5 (Beiglböck and Juillet [8, Corollary 1.6])**.** *Let μ, ν be probability measures in convex order and assume that μ is continuous. Then there exists a pair of measurable functions $T_d : \mathbb{R} \to \mathbb{R}$ and $T_u : \mathbb{R} \to \mathbb{R}$ such that $T_d(x) \leq x \leq T_u(x)$, such that for all $x < x'$ we have $T_u(x) \leq T_u(x')$ and $T_d(x') \notin (T_d(x), T_u(x))$, and such that, if we define $\bar{\pi}(dx, dy) = \mu(dx)\chi_{T_d(x),x,T_u(x)}(dy)$, then $\bar{\pi} \in \Pi_M(\mu, \nu)$ and $\bar{\pi} = \pi_{lc}$.*

Since $T_d(x) \leq x \leq T_u(x)$ we call T_d a lower function and T_u an upper function.

Lemma 3.5 is expressed in terms of elements of Π_M. We can give an equivalent expression in terms of a martingale. First we give an analogue of Definition 3.4 for functions.

Definition 3.6. *Given an interval I and an increasing function $g : I \to \mathbb{R}$, a pair of functions $f, h : I \to \mathbb{R}$ is said to be left-monotone with respect to g on I if $f \leq g \leq h$ and if for $x < x'$ we have $h(x) \leq h(x')$ and $f(x') \notin (f(x), h(x))$.*

Corollary 3.7. *Let $(\Omega, \mathcal{F}, \mathbb{P}) = (I \times (0, 1), \mathcal{B}(\Omega), \mu \times Leb((0, 1)))$. Let $\omega = (x, v)$ and let the canonical random variable (X, V) on $(\Omega, \mathcal{F}, \mathbb{P})$ be given by $(X(\omega), V(\omega)) = (x, v)$. Then X has law μ, V is a $U(0, 1)$ random variable and X and V are independent. Let $\mathbb{F} = (\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_1 = \sigma(X), \mathcal{F}_2 = \sigma(X, V))$ and set $\mathbf{M} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.*

Suppose μ is continuous. Then there exists $T_d, T_u : I \to \mathbb{R}$ such that (T_d, T_u) is left-monotone with respect to the identity function on I and such that if we define $Y(x, v) \in \{T_d(x), T_u(x)\}$ by $Y(x, v) = x$ on $T_d(x) = x = T_u(x)$ and

$$Y(x, v) = T_d(x) I_{\left\{v \leq \frac{T_u(x)-x}{T_u(x)-T_d((x)}\right\}} + T_u(x) I_{\left\{v > \frac{T_u(x)-x}{T_u(x)-T_d(x)}\right\}} \tag{3.2}$$

otherwise, then $M = (\bar{\mu}, X, Y(X, V))$ is a $\mathbf{M}$-martingale for which $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$.

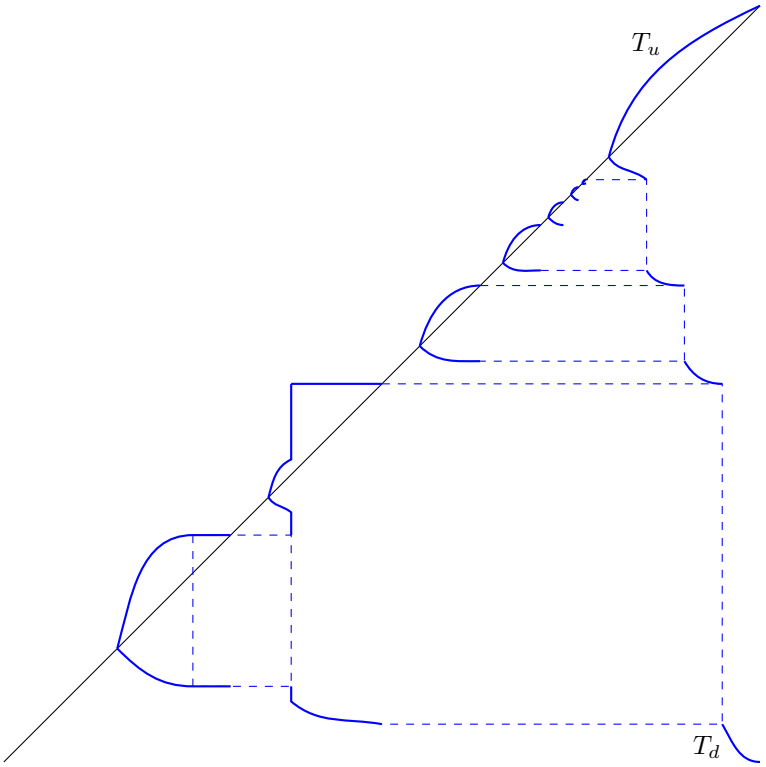

Figure 1: Stylized plot of the functions T_d and T_u in the general case. T_d and T_u are given by the solid lines in the figure. Note that on the set $\{x : T_u(x) = x\}$ we have $T_d(x) = x$. In the figure the set $\{x : T_u(x) > T_d(x)\}$ is a finite union of intervals whereas in general it may be a countable union of intervals (and the set of endpoints of these intervals may have accumulation points). Similarly, in the figure T_d has finitely many downward jumps, whereas in general it may have countably many jumps. Atoms of ν lead to horizontal sections of T_d and T_u. Atoms of μ, which are excluded in Lemma 3.5 and Corollary 3.7, but included in Theorem 3.8 below, lead to vertical (multi-valued) sections of T_d and T_u.

Suppose ν is also continuous and fix $x \in \mathbb{R}$. Under the left-curtain martingale coupling, $\mu|_{(T_d(x),x)}$ is mapped to $\nu|_{(T_d(x),T_u(x))}$. Thus $\{T_d(x), T_u(x)\}$ with $T_d(x) \le x \le T_u(x)$ are solutions to

$$\int_{T_d(x)}^{x} \mu(dz) = \int_{T_d(x)}^{T_u(x)} \nu(dz), \tag{3.3}$$

$$\int_{T_d(x)}^{x} z\mu(dz) = \int_{T_d(x)}^{T_u(x)} z\nu(dz). \tag{3.4}$$

Essentially, (3.3) is preservation of mass condition and (3.4) is preservation of mean and the martingale property.

In general, there can be multiple solutions to (3.3) and (3.4) although under the additional left-monotonicity properties of Definition 3.6, for almost all $x \in \mathbb{R}$ there is a unique solution. (However, even then there may be exceptional x at which T_d jumps and at which there are multiple solutions.)

As observed by Henry-Labordère and Touzi [19], when μ and ν admit continuous densities ρ_μ and ρ_ν respectively, and T_u and T_d are smooth, we find that they satisfy the

pair of coupled differential equations

$$\begin{aligned}\rho_\mu(x) - T_d'(x)\rho_\mu(T_d(x)) &= T_u'(x)\rho_\nu(T_u(x)) - T_d'(x)\rho_\nu(T_d(x))\\ x\rho_\mu(x) - T_d'(x)T_d(x)\rho_\mu(T_d(x)) &= T_u(x)T_u'(x)\rho_\nu(T_u(x)) - T_d'(x)T_d(x)\rho_\nu(T_d(x)).\end{aligned}$$

However, it remains to specify the initial conditions of the differential equations and even in the case of smooth densities, T_d may have downward jumps at locations which depend on the global properties of μ and ν. In the special case where μ and ν satisfy the dispersion assumption, see Hobson and Klimmek [24] or Hobson and Norgilas [27] (see also [19, Section 3.4] for a closely related condition), in the sense that there exists $I_> = (e_-, e_+)$ such that $\rho_\mu > \rho_\nu > 0$ on $I_>$ and $0 \le \rho_\mu \le \rho_\nu$ otherwise, then the situation simplifies. In particular, $T_d(x) = x = T_u(x)$ for $x \le e_-$ and on $[e_-, \infty)$, T_d is strictly decreasing, T_u is strictly increasing and together they solve

$$T_d'(x) = \frac{T_u(x) - T_d(x)}{T_u(x) - x}\frac{(\rho_\nu - \rho_\mu)(T_d(x))}{\rho_\mu(x)} \quad \text{and} \quad T_u'(x) = \frac{T_u(x) - T_d(x)}{x - T_d(x)}\frac{\rho_\nu(T_u(x))}{\rho_\mu(x)}$$

subject to $T_d(e_-) = e_- = T_u(e_-)$ (see also [19, Equations (3.9) and (3.10)]).

Henry-Labordère and Touzi [19, Equations (3.15) and (3.16)]) are able to go further and write $T_d(x)$ as the root of an integral equation, (and then $T_u(x)$ can be deduced from $T_d(x)$). But the integral equation depends on the curve $\{T_d(z); e_- \le z \le x\}$. They are also able to extend beyond the dispersion assumption case by defining T_d on intervals $[m_i, n_i]$ where m_i is an element of a certain set $\mathbb{M}_0 = \mathbb{M}_0(\mu, \nu)$ which, in the case where the densities are well defined, continuously differentiable and not identically equal on an interval, is the set of points where $\rho_\mu = \rho_\nu$ and $\rho_\mu' > \rho_\nu'$. But, the assumption that μ and ν are atom free is essential, and Henry-Labordère and Touzi also assume that $\mathbb{M}_0$ is finite, and to move beyond the case where $\mathbb{M}_0$ can be written as $\mathbb{M}_0 = \{m_j, j \in \mathbb{N} : i < j \Leftrightarrow m_i < m_j\}$ (for example, to allow for accumulation points) would require further arguments. The fundamental question of how to determine T_d and T_u remains, especially since Lemma 3.5 and Corollary 3.7 are purely existence statements.

3.3 The left-curtain coupling in the presence of atoms

In the case with atoms previous work of the authors gives an existence result similar in form to Corollary 3.7.

Theorem 3.8 (Hobson and Norgilas [27, Theorem 1])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1)\times(0,1), \mathcal{B}(\Omega), Leb(\Omega))$. Let $\omega = (u, v)$ and let (U, V) be the canonical random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ given by $(U(\omega), V(\omega)) = (u, v)$ so that U and V are independent $U(0,1)$ random variables. Let $\mathbb{F} = (\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_1 = \sigma(U), \mathcal{F}_2 = \sigma(U, V)\})$ and set $\mathbf{M} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.*

Fix $\mu \le_{cx} \nu$ and let $G = G_\mu$ be a quantile function of μ.

Then there exists $R, S : (0,1) \to \mathbb{R}$ such that the pair (R, S) is left monotone with respect to G on $\mathcal{I} = (0,1)$ and such that if we define $X(u,v) = X(u) = G(u)$ and $Y(u,v) \in \{R(u), S(u)\}$ by $Y(u,v) = G(u)$ on $R(u) = S(u)$ and

$$Y(u,v) = R(u)I_{\left\{v \le \frac{S(u)-G(u)}{S(u)-R(u)}\right\}} + S(u)I_{\left\{v > \frac{S(u)-G(u)}{S(u)-R(u)}\right\}} \tag{3.5}$$

otherwise, then $M = (\bar{\mu}, X(U), Y(U,V))$ is a $\mathbf{M}$-martingale for which $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$.

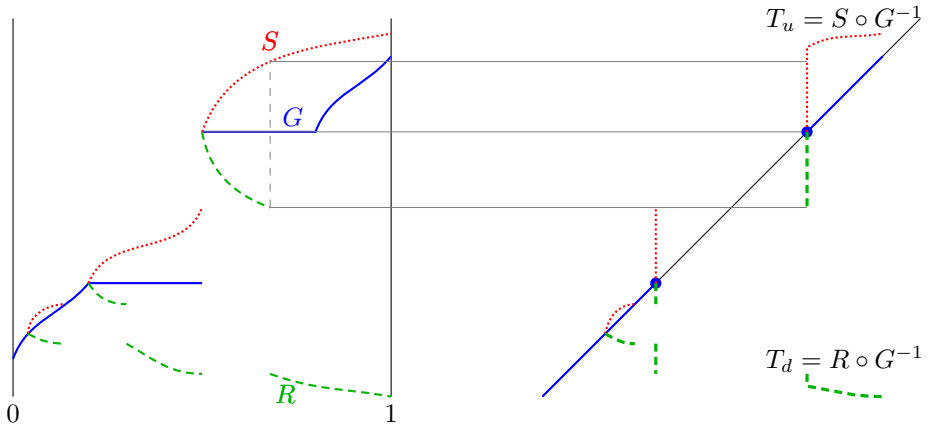

Figure 2: Sketch of R, G, S and the corresponding T_u and T_d. On the atoms of μ, G is flat, and T_d and T_u are multi-valued, but R and S remain well-defined.

Our goals in later sections are first to construct suitable candidate functions R and S satisfying left-monotonicity with respect to G, and second to show that they do indeed lead to a pair (X, Y) with $X \sim \mu$ and $Y \sim \nu$.

3.4 Lifted martingale transport plans

Just as Corollary 3.7 has an equivalent expression via Lemma 3.5, Theorem 3.8 has an equivalent expression in terms of transport plans, provided we generalise the notion of a martingale transport plan. Let $(\mu_u)_{0\le u\le 1}$ be a family of measures with $\mu_u(\mathbb{R}) = u$, $\mu_1 = \mu$ and $\mu_u \le \mu_v$ for $0 \le u \le v \le 1$, and let λ denote Lebesgue measure on the unit interval. Then a *lift* (Beiglböck and Juillet [8, 9]) of μ with respect to $(\mu_u)_{0\le u\le 1}$ is a probability measure $\hat{\mu} \in \Pi(\lambda, \mu)$ such that, for all $u \in [0,1]$ and Borel $A \subseteq \mathbb{R}$, $\hat{\mu}([0,u] \times A) = \mu_u(A)$. A *lifted martingale transport plan* is a probability measure $\hat{\pi} \in \Pi(\hat{\mu}, \nu)$ such that $\int_{\mathbb{R}} y \hat{\pi}_{u,x}(dy) = x$, $\hat{\mu}$-a.e. (u, x), where $\hat{\pi}_{u,x}$ denotes the disintegration of $\hat{\pi} \in \Pi(\hat{\mu}, \nu)$ with respect to $\hat{\mu}$: $\hat{\pi}(du, dx, dy) = \hat{\mu}(du, dx)\hat{\pi}_{u,x}(dy)$.

One of the insights of Beiglböck and Juillet [8, 9] is that, for $(\mu_u)_{0\le u\le 1}$ as above, the shadow measure induces a family of martingale couplings. In particular the idea is that for all $u \in [0,1]$, μ_u is mapped to $S^\nu(\mu_u)$. A crucial result making this possible is the fact that if $0 < u < v < 1$ and $\mu_u \le \mu_v$ then $S^\nu(\mu_u) \le S^\nu(\mu_v)$.

A natural choice for the lift $\hat{\mu}$ of μ is the *quantile lift* $\hat{\mu}^Q$ whose support is of the form $\{(v, G(v)) : 0 < v < 1\}$ where G is a quantile function of μ. Then $\hat{\mu}^Q(du, dx) = du\delta_{G(u)}(dx)$ and for a Borel set A, $\hat{\mu}^Q([0,w] \times A) = \int_0^w du I_{\{G(u)\in A\}}$. Then, by Beiglböck and Juillet [9, Theorem 1.1], there exists a unique lifted martingale transport plan $\hat{\pi}^Q$ such that for all $u \in [0,1]$ and Borel $A, B \subseteq \mathbb{R}$, $\hat{\pi}^Q([0,u]\times A\times\mathbb{R}) = \mu_u(A)$ and $\hat{\pi}^Q([0,u]\times\mathbb{R}\times B) = S^\nu(\mu_u)(B)$. This is the left-curtain martingale coupling.

By analogy with the correspondence between Lemma 3.5 and Corollary 3.7 we have the following equivalent restatement of Theorem 3.8:

Corollary 3.9. *Let μ, ν be probability measures in convex order and let $\hat{\mu}^Q$ be the quantile lift of μ. Then there exists a pair of measurable functions $R : \mathbb{R} \to \mathbb{R}$ and $S : \mathbb{R} \to \mathbb{R}$ such that (R, S) is left-monotone with respect to $G = G_\mu$ and such that if $\hat{\pi}^Q(du, dx, dy) = du\delta_{G(u)}(dx)\hat{\pi}^Q_{u,x}(dy)$ (recall $\hat{\mu}$ has support on $\{(u, G(u)) : 0 < u < 1\}$) then $\hat{\pi}^Q_{u,x}(dy) = \hat{\pi}^Q_{u,G(u)}(dy) = \chi_{R(u),G(u),S(u)}(dy)$ and $\hat{\pi}^Q$ is the lifted-left-curtain martingale transport plan which transports a second marginal μ to third marginal ν.*

4 The geometric construction

Fix $\mu \leq_{cx} \nu$. The goal of this section is to construct candidates for the functions R and S of Theorem 3.8. Then, in Section 7 we will prove that they can indeed be used to define a (left-monotone) martingale coupling of μ and ν.

Recall the definition of $D(k) = P_\nu(k) - P_\mu(k)$, $k \in \mathbb{R}$. In what follows (and in the light of Lemma 2.1 and the subsequent discussion) we assume that $\{k \in \mathbb{R} : D(k) > 0\}$ is an (open) interval, and μ is supported on this set. Moreover, $\{k : D(k) > 0\} = (\ell_\nu, r_\nu)$.

Recall also the definition of the sub-differential $\partial h(x)$ of a convex function $h : \mathbb{R} \to \mathbb{R}$ at x:

$$\partial h(x) = \{\phi \in \mathbb{R} : h(y) \geq h(x) + \phi(y - x) \text{ for all } y \in \mathbb{R}\}.$$

We extend this definition to non-convex functions f so that the subdifferential of f at x is given by

$$\partial f(x) = \{\phi \in \mathbb{R} : f(y) \geq f(x) + \phi(y - x) \text{ for all } y \in \mathbb{R}\}.$$

If h is convex then ∂h is non-empty everywhere, but this is not the case for non-convex functions. Instead we have that $\partial f(x)$ is non-empty if and only if $f(x) = f^c(x)$ and then $\partial f^c(x) = \partial f(x)$.

Let $G = G_\mu$ be a quantile function of μ. (In Sections 6 and 7 we will take G to be the left-continuous quantile function, but for now we let G be *any* quantile function.) For each $u \in (0, 1)$, define $\mu_u \in \mathcal{M}$ by

$$\mu_u(A) = \mu\Big(A \cap \big(-\infty, G(u)\big)\Big) + \left(u - \mu\Big(\big(-\infty, G(u)\big)\Big)\right)\delta_{G(u)}(A), \quad \text{for all Borel } A \subseteq \mathbb{R}.$$

Then for $u \in (0, 1)$, $\mu_u \leq \mu$ and $\mu_u(\mathbb{R}) = u$. Note that μ_u does not depend on the choice of quantile function G.

We have $P_{\mu_u}(k) = P_\mu(k)$ for $k \leq G(u)$, while $P_{\mu_u}(k) \leq P_\mu(k)$ for $k > G(u)$. In particular,

$$P_{\mu_u}(k) = P_\mu(k \wedge G(u)) + u(k - G(u))^+, \quad k \in \mathbb{R},$$

and thus, $P_{\mu_u}(\cdot)$ is linear on $[G(u), \infty)$ and $u \in \partial P_\mu(G(u))$, so that $P'_\mu(G(u)-) \leq u \leq P'_\mu(G(u)+)$.

For each $u \in (0, 1)$ define $\mathcal{E}_u : \mathbb{R} \to \mathbb{R}_+$ by $\mathcal{E}_u = P_\nu - P_{\mu_u}$, so that for $k \in \mathbb{R}$,

$$\mathcal{E}_u(k) = P_\nu(k) - P_{\mu_u}(k) = D(k) + P_\mu(k) - P_{\mu_u}(k).$$

Then, by Theorem 3.2, we have that

$$P_{S^\nu(\mu_u)}(k) = P_\nu(k) - \mathcal{E}_u^c(k), \quad k \in \mathbb{R}.$$

The idea underlying this section is that we can hope to determine the functions characterising the left-curtain coupling by considering the properties of $\mathcal{E}_u$ and $\mathcal{E}_u^c$.

Note that $\mathcal{E}_u(k) = D(k)$ for $k \leq G(u)$. Since $P_\mu - P_{\mu_u}$ is non-negative on $\mathbb{R}$, we have that $\mathcal{E}_u(k) \geq D(k)$ for $k > G(u)$. Moreover, since P_{μ_u} is linear on $[G(u), +\infty)$, $\mathcal{E}_u$ is convex on $(G(u), +\infty)$. It is also easy to see that $k \mapsto \mathcal{E}_u(k) - D(k)$ is non-decreasing.

We now define candidate lower and upper functions. In fact we define two lower functions, which differ on a set of measure zero, either of which could be used in Theorem 3.8. One of the lower functions is in keeping with the definition of X in the study of convex hulls, but it turns out that the other is more convenient in the proof of Theorem 3.8.

The idea is that typically $\mathcal{E}_u$ is not convex, but we can define its convex hull. Moreover, commonly $\mathcal{E}_u(G(u)) > \mathcal{E}_u^c(G(u))$ in which case we can define $Q(u)$ to be the largest point

to the left of $G(u)$ at which $\mathcal{E}_u$ and its convex hull agree, and $S(u)$ to be the smallest point to the right of $G(u)$ where $\mathcal{E}_u$ and $\mathcal{E}_u^c$ agree. Then, when $Q(u) < S(u)$ we can define $\phi(u)$ to be the slope of $\mathcal{E}_u^c$ over this interval. Typically, $Q(u)$ will be the only point k below $G(u)$ such that $\mathcal{E}_u(k) = \mathcal{E}_u(S(u)) + (k - S(u))\phi(u)$, but in exceptional cases there may be other points with this property. In that case we let $R(u)$ be the smallest such point. See Figure 3. Our first goal is to formulate the descriptions in this paragraph precisely, and in such a way that they apply to all situations, including pathological ones. The second goal is prove that the various quantities have certain properties, especially as u varies.

Our motivation to study $R(u)$ and $S(u)$ stems from the fact that, at least in regular cases, by replacing $(T_d(x), x, T_u(x))$ with $(R(u), G(u), S(u))$, the mass and mean preservation conditions (3.3) and (3.4) hold. Indeed, suppose μ and ν are atomless with positive densities everywhere, so that, for each $u \in (0,1)$, $\mathcal{E}_u$ is differentiable. Then if $R(u) < G(u) \leq S(u)$, by construction (see Figure 3) we have that

$$\mathcal{E}_u'(R(u)) = \mathcal{E}_u'(S(u)) \quad \text{and} \quad \mathcal{E}_u(R(u)) + \mathcal{E}_u'(R(u))(S(u) - R(u)) = \mathcal{E}_u(S(u)),$$

which can be easily shown to be equivalent to (3.3) and (3.4). Indeed, we have both

$$\mathcal{E}_u'(R(u)) = P_\nu'(R(u)) - P_{\mu_u}'(R(u)) = P_\nu'(R(u)) - P_\mu'(R(u)) = \int_{-\infty}^{R(u)} \nu(dy) - \int_{-\infty}^{R(u)} \mu(dy),$$
$$\mathcal{E}_u'(S(u)) = P_\nu'(S(u)) - P_{\mu_u}'(S(u)) = P_\nu'(S(u)) - P_\mu'(G(u)) = \int_{-\infty}^{S(u)} \nu(dy) - \int_{-\infty}^{G(u)} \mu(dy),$$

and therefore $\mathcal{E}_u'(R(u)) = \mathcal{E}_u'(S(u))$ is equivalent to the mass preservation condition (3.3). Similarly, by writing

$$\begin{aligned}\mathcal{E}_u(R(u)) = P_\nu(R(u)) - P_{\mu_u}(R(u)) &= P_\nu(R(u)) - P_\mu(R(u)) \\ &= \int_{-\infty}^{R(u)} (R(u) - y)\nu(dy) - \int_{-\infty}^{R(u)} (R(u) - y)\mu(dy)\end{aligned}$$

and

$$\begin{aligned}\mathcal{E}_u(S(u)) &= P_\nu(S(u)) - P_{\mu_u}(S(u)) \\ &= P_\nu(S(u)) - \{P_\mu(G(u)) + P_{\mu_u}'(G(u))(S(u) - G(u))\} \\ &= \int_{-\infty}^{S(u)} (S(u) - y)\nu(dy) - \left\{ \int_{-\infty}^{G(u)} (G(u) - y)\mu(dy) + (S(u) - G(u)) \int_{-\infty}^{G(u)} \mu(dy) \right\} \\ &= \int_{-\infty}^{S(u)} (S(u) - y)\nu(dy) - \int_{-\infty}^{G(u)} (S(u) - y)\mu(dy),\end{aligned}$$

and using the mean preservation condition (3.3), we have that that

$$\mathcal{E}_u(R(u)) + \mathcal{E}_u'(R(u))(S(u) - R(u)) = \mathcal{E}_u(S(u))$$

is equivalent to the mean preservation condition (3.4).

Define $Q, S : (0,1) \to \mathbb{R}$ by

$$Q(u) := X^{\mathcal{E}_u}(G(u)) \tag{4.1}$$
$$S(u) := Z^{\mathcal{E}_u}(G(u)) \tag{4.2}$$

Lemma 4.1. *$Q(u) = G(u)$ if and only if $S(u) = G(u)$.*

Proof. Fix $u \in (0,1)$. By continuity of $\mathcal{E}_u(\cdot)$ we have $\mathcal{E}_u(Q(u)) = \mathcal{E}_u^c(Q(u))$, and $\mathcal{E}_u(S(u)) = \mathcal{E}_u^c(S(u))$.

Suppose $Q(u) = G(u)$. Then $G(u) \in \{z : z \geq G(u), \mathcal{E}_u^c(z) = \mathcal{E}_u(z)\}$ and hence $S(u) = \inf\{z : z \geq G(u), \mathcal{E}_u^c(z) = \mathcal{E}_u(z)\} = G(u)$.

The reverse implication follows by symmetry. □

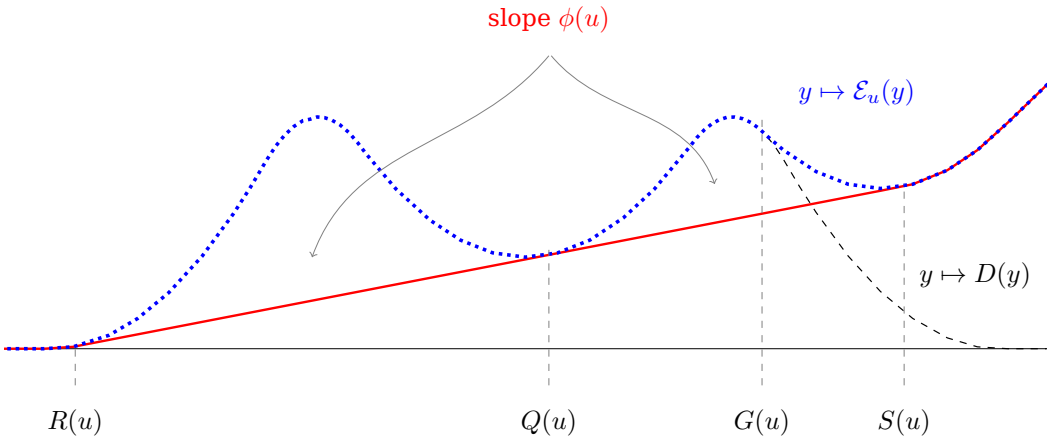

Figure 3: Plot of locations of $R(u)$, $Q(u)$, $G(u)$ and $S(u)$ in the case where $R(u) < Q(u) < G(u) < S(u)$. The dashed curve represents D. The dotted curve corresponds to the graph of $\mathcal{E}_u$. Note that $D = \mathcal{E}_u$ on $(-\infty, G(u)]$, while $\mathcal{E}_u$ is convex and $D \leq \mathcal{E}_u$ on $(G(u), \infty)$. The solid curve below $\mathcal{E}_u$ represents $\mathcal{E}_u^c$. The convex hull $\mathcal{E}_u^c$ is linear on $[R(u), S(u)]$, and its slope is given by $\phi(u)$.

We want to introduce a function $\phi : (0,1) \to \mathbb{R}$ which represents the slope of $\mathcal{E}_u^c(\cdot)$ at $G(u)$. If $Q(u) < G(u) < S(u)$, then this slope is well defined. If $Q(u) = G(u)$ or $G(u) = S(u)$ then the slope of $\mathcal{E}_u^c$ may not be well defined at $G(u)$. To cover all cases we define:

Definition 4.2. *$\phi : (0,1) \to \mathbb{R}$ is given by* $\phi(u) = \inf\{\psi : \psi \in \partial\mathcal{E}_u^c(G(u))\}$.

Lemma 4.3. $\phi(u) = (\mathcal{E}_u^c)'(G(u)-) = (\mathcal{E}_u^c)'(S(u)-)$.

Proof. The first equality is immediate from the definition of ϕ, as is the second one provided $G(u) = S(u)$. On the other hand, if $G(u) < S(u)$, then $\mathcal{E}_u^c = L^{\mathcal{E}_u}_{Q(u),S(u)}$ on $[Q(u), S(u)] \supset \{G(u)\}$ and the second equality follows. □

Now we can introduce our second candidate lower function.

Recall the definition of $L^f_{a,b}$ for any $f : \mathbb{R} \to \mathbb{R}$ (see (2.3)), so that (in the case $a < b$) $L^f_{a,b}$ is the line passing through $(a, f(a))$ and $(b, f(b))$. Define also $L^{f,\psi}_a$ by $L^{f,\psi}_a(y) = f(a) + \psi(y-a)$ so that $L^{f,\psi}_a$ is the line passing through $(a, f(a))$ with slope ψ. (Note that, in the case $a = b$, $L^f_{a,a} = L^{f,0}_a$.) Define $R : (0,1) \to \mathbb{R}$ by

$$R(u) := \inf\{k : k \leq G(u), D(k) = L^{\mathcal{E}_u^c, \phi(u)}_{G(u)}(k)\}. \tag{4.3}$$

If $Q(u) < G(u)$ then the definition of R can be rewritten as $R(u) = \inf\{k : k \leq G(u), D(k) = L^{\mathcal{E}_u}_{Q(u),S(u)}(k)\}$. Note that $Q(u) \in \{k : k \leq G(u), D(k) = L^{\mathcal{E}_u^c, \phi(u)}_{G(u)}(k)\}$ so that $R(u)$ exists in all cases and satisfies $R(u) \leq Q(u)$. See Figure 3.

Recall that ℓ_ν and r_ν are the left- and right-hand endpoints of the interval $\{k : D(k) > 0\}$. The next lemma, the proof of which is postponed until Appendix A.2, shows that R and S are finite on $(0,1)$.

Lemma 4.4. *Fix $u \in (0,1)$. Either $-\infty < \ell_\nu \le R(u)$ or $-\infty = \ell_\nu < R(u)$. Similarly, either $S(u) \le r_\nu < -\infty$ or $S(u) < r_\nu = \infty$.*

If $Q(u) < S(u)$, then by construction, $\mathcal{E}_u^c < \mathcal{E}_u$ on $(Q(u), S(u))$ and $\mathcal{E}_u^c \le \mathcal{E}_u$ on $[R(u), S(u)]$. In particular, $\mathcal{E}_u^c$ is linear on $(R(u), S(u))$, whilst $\mathcal{E}_u^c(S(u)) = \mathcal{E}_u(S(u))$, $\mathcal{E}_u^c(Q(u)) = \mathcal{E}_u(Q(u)) = D(Q(u))$ and $\mathcal{E}_u^c(R(u)) = \mathcal{E}_u(R(u)) = D(R(u))$. It follows that if $Q(u) < S(u)$ then

$$\phi(u) = \frac{\mathcal{E}_u(S(u)) - D(Q(u))}{S(u) - Q(u)} = \frac{\mathcal{E}_u(S(u)) - D(R(u))}{S(u) - R(u)}.$$

Further, $\phi(u)$ is an element of each of $\partial\mathcal{E}_u(R(u))$, $\partial\mathcal{E}_u(Q(u))$ and $\partial\mathcal{E}_u(S(u))$ together with $\partial\mathcal{E}_u^c(R(u))$, $\partial\mathcal{E}_u^c(Q(u))$ and $\partial\mathcal{E}_u^c(S(u))$.

Our goal is to prove first that (R, S) is left-monotone with respect to G on $(0,1)$ in the sense of Definition 3.6 (Theorem 4.9 below) and second that they define a martingale coupling of μ and ν (Theorem 7.9 below). Together, these results give an explicit construction of a pair (R, S) which solve the problem in Theorem 3.8 above.

We begin with some preliminary lemmas and other results. We are interested in properties of $u \mapsto \mathcal{E}_u(k)$, for fixed $k \in \mathbb{R}$. Let $0 < u < v < 1$. Then, since $\mu_u \le \mu_v$, $P_{\mu_u} \le P_{\mu_v}$ and

$$\mathcal{E}_u(k) \ge \mathcal{E}_v(k), \quad k \in \mathbb{R}. \tag{4.4}$$

Indeed, for $k \in \mathbb{R}$,

$$\begin{aligned}\mathcal{E}_u(k) - \mathcal{E}_v(k) &= P_{\mu_v}(k) - P_{\mu_u}(k)\\ &= \begin{cases} 0, & \text{if } k \le G(u);\\ P_\mu(k) - P_\mu(G(u)) - u(k - G(u)), & \text{if } G(u) < k \le G(v);\\ P_\mu(G(v)) - P_\mu(G(u)) - v(G(v) - G(u)) + (v-u)(k - G(u)), & \text{if } k > G(v).\end{cases}\end{aligned} \tag{4.5}$$

Note that, for $0 < u < v < 1$ and $k \ge G(v)$, (4.5) can be written as

$$\mathcal{E}_u(k) - \mathcal{E}_v(k) = (v-u)\,(k - R(u)) + \Gamma_{u,v}, \tag{4.6}$$

where

$$\Gamma_{u,v} = P_\mu(G(v)) - \{P_\mu(G(u)) + v(G(v) - G(u))\} - (v-u)(G(u) - R(u))$$

does not depend on k. Convexity of P_μ ensures that $\Gamma_{u,v} \le 0$, and if $R(u) < G(u)$ then $\Gamma_{u,v} < 0$. Further, from (4.4) we also have that $\mathcal{E}_v^c \le \mathcal{E}_v \le \mathcal{E}_u$, so that $\mathcal{E}_v^c$ is a convex minorant of $\mathcal{E}_u$ and

$$\mathcal{E}_u^c(k) \ge \mathcal{E}_v^c(k), \quad k \in \mathbb{R}. \tag{4.7}$$

Finally, for any $u \in (0,1)$, $\mathcal{E}_u$ is defined as a difference of two convex functions, and thus its left and right derivatives exist. It follows that for $v > u$ and $k \ge G(v)$,

$$\mathcal{E}_u'(k+) - \mathcal{E}_v'(k+) = v - u. \tag{4.8}$$

Again for $v > u$,

$$\mathcal{E}_u'(k-) - \mathcal{E}_v'(k-) = \begin{cases} v-u, & \text{if } k > G(v);\\ P_\mu'(k-) - u, & \text{if } k \in (G(u), G(v)];\\ 0, & \text{if } k \le G(u).\end{cases} \tag{4.9}$$

Lemma 4.5. *Suppose $k \in [G(v-), G(v+)]$, $u < v$ and $G(u+) < k$. Suppose $\mathcal{E}'_v(k-) \leq \mathcal{E}'_v(k+)$ and $\psi \in [\mathcal{E}'_v(k-), \mathcal{E}'_v(k+)]$. Then $\mathcal{E}'_u(k-) \leq \psi + (v-u) \leq \mathcal{E}'_u(k+)$.*

Proof. The result follows easily from the fact that $\mathcal{E}'_v(k-) = P'_\nu(k-) - P'_{\mu_v}(k-) \geq P'_\nu(k-) - v$ together with $\mathcal{E}'_v(k+) = P'_\nu(k+) - v$ and $\mathcal{E}'_u(k\pm) = P'_\nu(k\pm) - u$. □

The proof of the following lemma and its corollary are deferred until Appendix A.2.

Lemma 4.6. *Fix $v > u$. Suppose $\mathcal{E}^c_u(z) = \mathcal{E}^c_v(z) = \mathcal{E}_u(z)$ for some z. Then $\mathcal{E}^c_u = \mathcal{E}^c_v$ on $(-\infty, z]$.*

Corollary 4.7. *Suppose $v > u$ and $S(u) = S(v) \leq G(u+)$. Then $\mathcal{E}^c_u = \mathcal{E}^c_v$ on $(-\infty, S(v)]$. Further $\phi(u) = \phi(v)$ and $R(u) = R(v)$.*

For $0 < u < v < 1$ let $\xi_{u,v}$ be given by

$$(v-u)\xi_{u,v} = \int_{(G(u),G(v))} x\mu(dx) + G(u)\{\mu(-\infty, G(u)] - u\} + G(v)\{v - \mu(-\infty, G(v))\}.$$

Then $\xi_{u,v}$ is the conditional mean of μ between the quantiles at u and v. It is easily checked that $\xi_{u,v}$ does not depend on the choice of quantile function G.

Lemma 4.8. *We have*

(i) Suppose $k \in [G(v-), G(v+)]$ and $u < v$. Then the line passing through $(k, \mathcal{E}_v(k))$ with slope $P'_\nu(k-) - v$ meets the line passing through $(k, \mathcal{E}_u(k))$ with slope $P'_\nu(k-) - u$ at a point with x-coordinate $\xi_{u,v}$, which does not depend on k.

(ii) Take $k > G(v+)$ and suppose $u < v$. Then the line passing through $(k, \mathcal{E}_v(k))$ with slope $\mathcal{E}'_v(k-)$ meets the line passing through $(k, \mathcal{E}_u(k))$ with slope $\mathcal{E}'_u(k-)$ at a point with x-coordinate $\xi_{u,v}$, which does not depend on k.

(iii) Suppose $k \in [G(w-), G(w+)]$, $u < w$, $G(u+) < k$, $\mathcal{E}'_w(k-) \leq \mathcal{E}'_w(k+)$ and $\psi \in [\mathcal{E}'_w(k-), \mathcal{E}'_w(k+)]$. Then the line passing through $(k, \mathcal{E}_w(k))$ with slope ψ meets the line passing through $(k, \mathcal{E}_u(k))$ with slope $\psi + (w-u)$ at a point with x-coordinate $\xi_{u,v}$, which does not depend on k.

Proof. (i) The line $L_k^{\mathcal{E}_v, P'_\nu(k-)-v}$ is given by

$$L_k^{\mathcal{E}_v, P'_\nu(k-)-v}(z) = \mathcal{E}_v(k) + (P'_\nu(k-) - v)(z-k) = P_\nu(k) + P'_\nu(k-)(z-k) - (P_{\mu_v}(k) + v(z-k)).$$

Similarly the line $L_k^{\mathcal{E}_u, P'_\nu(k-)-u}$ is given by

$$L_k^{\mathcal{E}_u, P'_\nu(k-)-u}(z) = \mathcal{E}_u(k) + (P'_\nu(k-) - u)(z-k) = P_\nu(k) + P'_\nu(k-)(z-k) - (P_{\mu_u}(k) + u(z-k)).$$

These lines intersect at the point where the lines $P_{\mu_v}(k) + v(z-k)$ and $P_{\mu_u}(k) + u(z-k)$ meet. But for $k \geq G(v-)$, $P_{\mu_v}(k) + v(z-k) = P_{\mu_v}(G(v)) + v(z - G(v))$ and $P_{\mu_u}(k) + u(z-k) = P_{\mu_u}(G(u)) + u(z - G(u))$.

It is easily checked that the lines $P_{\mu_v}(G(v)) + v(z - G(v))$ and $P_{\mu_u}(G(v)) + u(z - G(u))$ meet at $z = \xi_{u,v}$.

(ii) If $k > G(v+)$ then from the definition of P_{μ_v} it follows that $P'_{\mu_v}(k-) = v$ and therefore $\mathcal{E}'_v(k-) = P'_\nu(k-) - v$. Since $k > G(v+) \geq G(u)$, the same argument shows that $\mathcal{E}'_u(k-) = P'_\nu(k-) - u$. Then (ii) follows exactly as in the proof of (i).

(iii) This follows similarly to (i). □

Theorem 4.9. *The pair $R, S : (0,1) \to \mathbb{R}$, defined by* (4.3) *and* (4.2) *is left-monotone with respect to G on $(0,1)$ in the sense of Definition 3.6.*

Proof. That $R(u) \le Q(u) \le G(u) \le S(u)$, $u \in (0,1)$, follows by definition.

Fix $0 < u < v < 1$.

If $R(u) = G(u)$ then necessarily $S(u) = G(u)$ (see Lemma 4.1). Then $S(v) \ge G(v) \ge G(u) = S(u)$. Further, $(R(u), S(u)) = \emptyset$ so that $R(v) \notin (R(u), S(u))$ by default.

There are two remaining cases, when $R(u) < G(u) \le G(v) < S(u)$ and $R(u) < G(u) \le S(u) \le G(v)$.

Case 1: Suppose $R(u) < G(u) \le G(v) < S(u)$. We show that

$$\mathcal{E}_v(k) > L^{\mathcal{E}_v}_{R(u),S(u)}(k)$$

for $k \in (R(u), S(u))$. Then it follows from Lemma 2.5 that $R(v) \le Q(v) = X^{\mathcal{E}_v}(G(v)) \le R(u) < S(u) \le Z^{\mathcal{E}_v}(G(v)) = S(v)$ as required.

First, for $k \in (R(u), G(u)]$ and since $\mathcal{E}_u \ge \mathcal{E}_v$ everywhere (with equality to the left of $G(u)$),

$$\mathcal{E}_v(k) = \mathcal{E}_u(k) \ge L^{\mathcal{E}_u}_{R(u),S(u)}(k) > L^{\mathcal{E}_v}_{R(u),S(u)}(k). \tag{4.10}$$

with the strict inequality in (4.10) following from the fact that $S(u) > G(v)$ and $\mathcal{E}_v < \mathcal{E}_u$ on $(G(v), \infty)$.

Second, from (4.9) we have that

$$\begin{aligned}
\mathcal{E}_v'(S(u)-) &= \mathcal{E}_u'(S(u)-) - (v-u) \\
&\le \frac{\mathcal{E}_u(S(u)) - D(R(u))}{S(u) - R(u)} - (v-u) \\
&= \frac{\Gamma_{u,v} + (v-u)(S(u) - R(u)) + \mathcal{E}_v(S(u)) - D(R(u))}{S(u) - R(u)} - (v-u) \\
&< \frac{\mathcal{E}_v(S(u)) - D(R(u))}{S(u) - R(u)}
\end{aligned} \tag{4.11}$$

where we use the fact that $\Gamma_{u,v} < 0$ for $R(u) < G(u)$. We conclude that $\mathcal{E}_v'(S(u-)) < \frac{\mathcal{E}_v(S(u)) - D(R(u))}{S(u) - R(u)}$ which is the slope of $L^{\mathcal{E}_v}_{R(u),S(u)}$. Then from the convexity of $\mathcal{E}_v$ on $[G(v), S(u)]$ we have that $\mathcal{E}_v(k) > L^{\mathcal{E}_v}_{R(u),S(u)}(k)$ for $k \in [G(v), S(u))$.

It remains to show that $\mathcal{E}_v(k) > L^{\mathcal{E}_v}_{R(u),S(u)}(k)$ for $k \in (G(u), G(v))$. This will follow from the following pair of inequalities which are valid on $(G(u), G(v))$:

$$D = \mathcal{E}_v \ge L^{\mathcal{E}_v}_{R(u),G(v)} > L^{\mathcal{E}_v}_{R(u),S(u)}. \tag{4.12}$$

The second of these inequalities is valid on $(R(u), S(u))$ and follows from the fact that $\mathcal{E}_v$ is convex on $(G(v), S(u))$ and (4.11).

Consider, therefore, the first inequality in (4.12). We consider two subcases, namely when $D(G(v)) \ge L^{\mathcal{E}_u}_{R(u),S(u)}(G(v))$ and when $D(G(v)) < L^{\mathcal{E}_u}_{R(u),S(u)}(G(v))$.

Case 1a: $D(G(v)) \ge L^{\mathcal{E}_u}_{R(u),S(u)}(G(v))$. Either $D \ge L^{\mathcal{E}_v}_{R(u),G(v)}$ on $(G(u), G(v))$ in which case we are done, or there exists y with $G(u) < y \le G(v)$ with $D(y) = L^{\mathcal{E}_v}_{R(u),G(v)}(y)$ and $D'(y-) \ge (L^{\mathcal{E}_v}_{R(u),G(v)})'$ whence

$$D'(y-) \ge (L^{\mathcal{E}_v}_{R(u),G(v)})' \ge (L^{\mathcal{E}_u}_{R(u),S(u)})' \ge \mathcal{E}_u'(S(u)-) \ge \mathcal{E}_u'(y+) \ge \mathcal{E}_u'(y-).$$

Indeed by minimality of $S(u)$, $(L^{\mathcal{E}_u}_{R(u),S(u)})' > \mathcal{E}_u'(y+)$. Then, $D'(y-) > \mathcal{E}_u'(y-)$. But, this would mean that $P_\nu'(y-) - P_\mu'(y-) > P_\nu'(y-) - P_{\mu_u}'(y-)$ or equivalently $u > P_\mu'(y-)$. Since $y > G(u)$ we have $P_\mu'(y-) \ge P_\mu'(G(u)+) \ge u$, yielding a contradiction. Hence in *Case 1a:* we have $D \ge L^{\mathcal{E}_v}_{R(u),G(v)}$ on $(G(u), G(v))$ as required.

Case 1b: $D(G(v)) < L^{\mathcal{E}_u}_{R(u),S(u)}(G(v))$. Again, either $D \geq L^{\mathcal{E}_v}_{R(u),G(v)}$ on $(G(u), G(v))$ in which case we are done, or there exists y with $G(u) < y \leq G(v)$ with $D(y) = L^{\mathcal{E}_v}_{R(u),G(v)}(y)$ and $D'(y-) \geq \frac{\mathcal{E}_v(G(v))-D(R(u))}{G(v)-R(u)}$ which is the slope of $L^{\mathcal{E}_v}_{R(u),G(v)}$. Let $w \in (u, v]$ be such that $G(w-) \leq y \leq G(w+)$ and $P'_\mu(y-) = w$. Then $\mathcal{E}'_w(y-) = P'_\nu(y-) - w$. Further, since $G(u) < y$, we have $P'_{\mu_u}(y-) = u$ and therefore $\mathcal{E}'_u(y-) = P'_\nu(y-) - u$.

Now consider the lines $L^{D,P'_\nu(y-)-w}_y \equiv L^{\mathcal{E}_w,P'_\nu(y-)-w}_y$ and $L^{\mathcal{E}_u,P'_\nu(y-)-u}_y$, and note that $D = \mathcal{E}_w$ on $(-\infty, y]$. Since $D'(y-) \geq (L^{\mathcal{E}_v}_{R(u),G(v)})'$, $L^{D,P'_\nu(y-)-w}_y$ lies on or below $L^{\mathcal{E}_v}_{R(u),G(v)}$ to the left of y and $L^{\mathcal{E}_v}_{R(u),G(v)}$ lies on or below $L^{\mathcal{E}_u}_{R(u),S(u)}$ on $(R(u), G(v))$. In contrast, $L^{\mathcal{E}_u,P'_\nu(y-)-u}_y$ lies above $L^{\mathcal{E}_u}_{R(u),S(u)}$ to the left of y by convexity of $\mathcal{E}_u$ on $[G(u), \infty)$. Hence, if $L^{\mathcal{E}_w,P'_\nu(y-)-w}_y$ and $L^{\mathcal{E}_u,P'_\nu(y-)-u}_y$ meet then they must do so outside $(R(u), y]$. But, by Lemma 4.8(i) they meet at $\xi_{u,w} \in [G(u), G(w)]$, and since $\xi_{u,w}$ does not depend on the choice of quantile function G we must have that $\xi_{u,w} \in [G(u), G(w-)]$; a contradiction. Hence $D \geq L^{\mathcal{E}_v}_{R(u),G(v)}$ on $(G(u), G(v))$ as required.

Case 2: Now suppose $R(u) < G(u) \leq S(u) \leq G(v)$. Then $S(v) \geq G(v) \geq S(u)$ so that all that remains to be shown is that $R(v) \notin (R(u), S(u))$.

Suppose to the contrary that $R(u) < R(v) < S(u)$. Consider the lines $L^{\mathcal{E}_w}_{R(w),S(w)}$ for $w \in \{u, v\}$. Recall that $\phi(w)$ is the slope of $L^{\mathcal{E}_w}_{R(w),S(w)}$. Note that, since $\mathcal{E}^c_w = L^{\mathcal{E}_w}_{R(w),S(w)}$ on $[R(w), S(w)]$, for $w \in \{u, v\}$, and $\mathcal{E}^c_u \geq \mathcal{E}^c_v$ everywhere, we have that $L^{\mathcal{E}_u}_{R(u),S(u)} \geq L^{\mathcal{E}_v}_{R(v),S(v)}$ on $[R(v), S(u)]$.

Suppose first that $D(R(v)) \geq L^{\mathcal{E}_u}_{R(u),S(u)}(R(v))$. (This will follow if, for instance, $R(v) \leq G(u+)$ for then since $R(v) > R(u)$ by hypothesis, $D(R(v)) = \mathcal{E}_u(R(v)) \geq L^{\mathcal{E}_u}_{R(u),S(u)}(R(v))$.) Then $D(R(v)) = \mathcal{E}_v(R(v)) = L^{\mathcal{E}_v}_{R(v),S(v)}(R(v)) \leq L^{\mathcal{E}_u}_{R(u),S(u)}(R(v)) \leq D(R(v))$. Therefore $L^{\mathcal{E}_u}_{R(u),S(u)}(R(v)) = L^{\mathcal{E}_v}_{R(v),S(v)}(R(v))$, and since $L^{\mathcal{E}_u}_{R(u),S(u)} \geq L^{\mathcal{E}_v}_{R(v),S(v)}$ to the right of $R(v)$, we must have that $\phi(v) \leq \phi(u)$. If $\phi(v) < \phi(u)$ then $D(R(u)) = L^{\mathcal{E}_u}_{R(u),S(u)}(R(u)) < L^{\mathcal{E}_v}_{R(v),S(v)}(R(u)) \leq \mathcal{E}^c_v(R(u)) \leq \mathcal{E}_v(R(u)) = D(R(u))$, a contradiction. On the other hand, if $\phi(v) = \phi(v)$ then the lines $L^{\mathcal{E}_u}_{R(u),S(u)}$ and $L^{\mathcal{E}_v}_{R(v),S(v)}$ are identical. But then $R(u) = R(v)$, a contradiction.

Now suppose $D(R(v)) < L^{\mathcal{E}_u}_{R(u),S(u)}(R(v))$. Since $L^{\mathcal{E}_u}_{R(u),S(u)}(R(v)) \leq \mathcal{E}_u(R(v))$ we must have that $G(u+) < R(v)$, and then by the minimality of $S(u)$ it follows that $L^{\mathcal{E}_u}_{R(u),S(u)}(R(v)) < \mathcal{E}_u(R(v))$. Since $v > u$ and $R(v) < S(u)$ there exists $w \in (u, v]$ with $G(w-) \leq R(v) \leq G(w+)$ and then we have from the twin facts that $\mathcal{E}_w(R(v)) = D(R(v))$ and $\mathcal{E}^c_w \geq \mathcal{E}^c_v \geq L^{D,\phi(v)}_{R(v)}$ that $\mathcal{E}'_w(R(v)-) \leq \phi(v) \leq \mathcal{E}'_w(R(v)+)$. It follows from Lemma 4.8(iii) with $k = R(v)$ and $\psi = \phi(v)$ that the line passing through $(R(v), \mathcal{E}_w(R(v)))$ with slope $\phi(v)$ meets the line passing through $(R(v), \mathcal{E}_u(R(v)))$ with slope $\phi(v) + (w - u)$ at a point with x-coordinate $\xi_{u,w} \in [G(u), G(w)]$. But since $\xi_{u,w}$ does not depend on the choice of quantile function G we have that $\xi_{u,w} \in [G(u), G(w-)]$, and then by noting that $\xi_{u,w} \neq R(v)$ it follows that $\xi_{u,w} \in [G(u), R(v))$. But, from the convexity of $\mathcal{E}_u$ on $[G(u), \infty)$ and the fact that by Lemma 4.5 $\mathcal{E}'_u(R(v)-) \leq \phi(v) + w - u \leq \mathcal{E}'_u(R(v)+)$, we have that $L^{\mathcal{E}_u,\phi(v)+w-u}_{R(v)} > L^{\mathcal{E}_u}_{R(u),S(u)}$ on $(-\infty, R(v)]$ and hence $L^{D,\phi(v)}_{R(v)} \equiv L^{\mathcal{E}_v}_{R(v),S(v)}$ crosses $L^{\mathcal{E}_u}_{R(u),S(u)}$ in $[G(u), R(v))$. Then

$$\mathcal{E}_u(R(u)) = \mathcal{E}_v(R(u)) \geq L^{\mathcal{E}_v}_{R(v),S(v)}(R(u)) > L^{\mathcal{E}_u}_{R(u),S(u)}(R(u)) = \mathcal{E}_u(R(u)),$$

a contradiction.

We conclude that $R(v) \notin (R(u), S(u))$. □

It is possible to draw pairs of curves $\mathcal{E}_u$ and $\mathcal{E}_v$ which have most of the correct properties (for example, $\mathcal{E}_v \leq \mathcal{E}_u$ with equality to the left of $G(u+)$) and for which

$R(v) \in (R(u), S(u))$. See Figure 4. But crucially $\mathcal{E}_v$ and $\mathcal{E}_u$ in Figure 4 do not satisfy $\mathcal{E}_u - \mathcal{E}_v$ is convex. The extra structure described in (4.9) and Lemma 4.8 makes counterexamples such as the one in the figure infeasible.

By definition G is increasing and by Theorem 4.9 the same property holds for S. Further, for $v > u$, $R(v) \notin (R(u), S(u))$, and so, except at places where $R = G = S$, and except at points where R jumps upwards, we expect R to be decreasing. We now argue that the set where $G < S$ can be divided into a union of disjoint intervals on which R is decreasing.

The functions G and S are monotonic, so we can define left and right limits. If $G(u+) < S(u-)$ then there exists v with $v > u$ such that $G(v+) < S(u-) \le S(v-)$. Conversely, there exists w with $w < u$ such that $S(w-) > G(u+) \ge G(w+)$. Then $\{u : G(u+) < S(u-)\}$ is open, and since each such interval contains a rational we can write $A_< := \{u : G(u+) < S(u-)\}$ as a countable union of disjoint open sets:

$$A_< := \{u \in (0,1) : G(u+) < S(u-)\} = \bigcup_{n \ge 1} A^n_<. \tag{4.13}$$

As the next lemma shows, R is decreasing on each of these sets. Moreover, for $u, v \in A^n_<$ with $u < v$, $R(v) < G(u)$. The proof of Lemma 4.10 is given in Appendix A.2.

Lemma 4.10. *For each $n \ge 1$, $R(\cdot)$ is decreasing on $A^n_<$.*

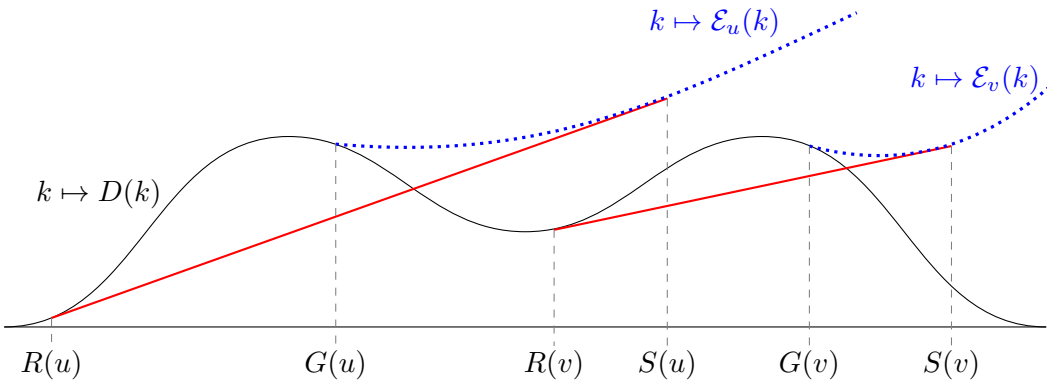

Figure 4: A potential counterexample to the left-monotonicity of R and S. In the figure, for a pair of arbitrary curves $(\mathcal{E}_u, \mathcal{E}_v)$, which are convex beyond $G(u)$ and $G(v)$ respectively, we have $R(v) \in (R(u), S(u))$. But this is not a feasible pair since $\mathcal{E}_u$ and $\mathcal{E}_v$ are not generated from a pair of distributions $\mu \le_{cx} \nu$, and, in particular, $\mathcal{E}_u - \mathcal{E}_v$ is not convex.

The results of this section (especially Theorem 4.9) give functions R, S which are left-monotone with respect to G. The remaining task is to show that they define a martingale transport for μ to ν, see Theorem 3.8. The next two sections give further characterisations and regularity results on the functions ϕ, R and S.

Remark 4.11. The monotonicity of S (see Theorem 4.9) implies that S is (Borel) measurable. On other hand, by Lemma 4.10, the restriction of R to $A_<$ is also measurable. It turns out that, in order to prove Theorem 3.8, global measurability of R is not necessary (see Section 7).

5 Properties of ϕ

The goal in this section is to give some further representations and properties of $\phi(\cdot)$. In particular, although ϕ can jump upwards on $(0,1)$ it is decreasing and Lipschitz continuous on $A_<$ and has a derivative almost everywhere on $A_<$ which we can identify in terms of a rational function of R, G and S.

Lemma 5.1. *We have*

$$\phi(u) = \sup_{k<G(u)} \frac{\mathcal{E}_u(S(u)) - D(k)}{S(u) - k}, \quad u \in (0,1), \tag{5.1}$$

and

$$\phi(v) \le \inf_{k>G(v)} \frac{\mathcal{E}_v(k) - D(R(v))}{k - R(v)}, \quad v \in (0,1), \tag{5.2}$$

with equality in (5.2) *if* $R(v) < G(v)$. *Moreover, if* $u, v \in A^n_<$ *for some* n, *with* $u < v$, *then in* (5.2) *the infimum over* $k > G(v)$ *can be extended to an infimum over* $k > G(u)$ *and we have*

$$\phi(v) = \inf_{k>G(v)} \frac{\mathcal{E}_v(k) - D(R(v))}{k - R(v)} = \inf_{k>G(u)} \frac{\mathcal{E}_v(k) - D(R(v))}{k - R(v)}. \tag{5.3}$$

Proof. First consider (5.1).

If $R(u) < G(u)$ then $\mathcal{E}_u \ge \mathcal{E}^c_u \ge L^{\mathcal{E}_u}_{R(u),S(u)}$ everywhere and $\mathcal{E}^c_u = L^{\mathcal{E}_u}_{R(u),S(u)}$ on $[R(u), S(u)]$. Then, for $k < G(u)$,

$$\frac{\mathcal{E}_u(S(u)) - \mathcal{E}_u(k)}{S(u) - k} \le \frac{L^{\mathcal{E}_u}_{R(u),S(u)}(S(u)) - L^{\mathcal{E}_u}_{R(u),S(u)}(k)}{S(u) - k} = \frac{\mathcal{E}_u(S(u)) - \mathcal{E}_u(R(u))}{S(u) - R(u)} = \phi(u) \tag{5.4}$$

so that (5.1) holds.

Now suppose that $R(u) = G(u)$, and hence that $\mathcal{E}^c_u(G(u)) = \mathcal{E}_u(G(u))$ and $G(u) = S(u)$. Let $\phi^* = \mathcal{E}'_u(G(u)-)$. We show $\phi(u) = \phi^* = \mathcal{E}'_u(G(u)-)$ and that (5.1) holds. Clearly if $\psi \in \partial\mathcal{E}^c_u(G(u))$ then $\psi \ge \frac{\mathcal{E}^c_u(S(u)) - \mathcal{E}_u(k)}{S(u)-k}$ for $k < G(u)$ and then

$$\psi \ge \sup_{k<G(u)} \frac{\mathcal{E}_u(G(u)) - D(k)}{G(u) - k} \ge \lim_{k\uparrow G(u)} \frac{\mathcal{E}_u(G(u)) - \mathcal{E}_u(k)}{G(u) - k} = \mathcal{E}'_u(G(u)-) = \phi^*.$$

Hence, since $\phi(u) \in \partial\mathcal{E}^c_u(G(u))$ by definition, to show that $\phi(u) = \phi^*$ it is sufficient to show that $\phi^* = \mathcal{E}'_u(G(u)-) \in \partial\mathcal{E}^c_u(G(u))$. We show

$$\mathcal{E}_u(y) \ge \mathcal{E}_u(G(u)) + \phi^*(y - G(u)), \tag{5.5}$$

for $y < G(u)$ and $y > G(u)$ separately.

Suppose (5.5) fails for some $r_0 < G(u)$. Then there exists $\epsilon > 0$ such that

$$\mathcal{E}_u(r_0) < \mathcal{E}_u(G(u)) + (\phi^* + \epsilon)(r_0 - G(u)).$$

Let r_1 be the largest solution (with $r < G(u)$) of $\mathcal{E}_u(r) = \mathcal{E}_u(G(u)) + (\phi^* + \epsilon)(r - G(u))$. Then from consideration of the slope of $\mathcal{E}_u$ near $G(u)$, $r_1 < G(u)$; further, from the property of r_1 as the largest solution we have $\mathcal{E}_u > L^{\mathcal{E}_u}_{r_1,G(u)}$ on $(r_1, G(u))$. Then $Q(u) = Y^{\mathcal{E}_u}(G(u)) \le r_1$ by Lemma 2.5 and *a fortiori* $R(u) < G(u)$. But, this contradicts our assumption that $R(u) = G(u)$.

Now we show (5.5) for $y > G(u)$ in the case $R(u) = G(u) = S(u)$. For $y > G(u)$, since $\mathcal{E}_u = \mathcal{E}^c_u$ on $[G(u), \infty)$,

$$\begin{aligned} \mathcal{E}_u(y) = \mathcal{E}^c_u(y) &\ge \mathcal{E}_u(G(u)) + (\mathcal{E}^c_u)'(G(u)+)(y - G(u)) \\ &\ge \mathcal{E}_u(G(u)) + (\mathcal{E}^c_u)'(G(u)-)(y - G(u)) \\ &\ge \mathcal{E}_u(G(u)) + \phi^*(y - G(u)) \end{aligned}$$

and we are done.

Hence $\phi(u) = \phi^*$. It remains to show (5.1). But

$$\mathcal{E}'_u(G(u)-) = \lim_{k\uparrow G(u)} \frac{\mathcal{E}_u(G(u)) - \mathcal{E}_u(k)}{G(u)-k} \leq \sup_{k<G(u)} \frac{\mathcal{E}_u(S(u)) - \mathcal{E}_u(k)}{S(u)-k} \leq \phi(u) = \phi^*$$

and (5.1) holds.

Now consider (5.2). If $R(v) < G(v)$ then interchanging the roles of R and S, by analogy with (5.4) we obtain for $k \geq G(v)$

$$\frac{\mathcal{E}_v(k) - \mathcal{E}_v(R(v))}{k - R(v)} \geq \phi(v)$$

and (5.2) follows. By considering $k = S(v)$ we see there is equality. If $R(v) = G(v) = S(v)$, then $\mathcal{E}_v(G(v)) = \mathcal{E}^c(G(v))$ and then if $\psi \in \partial\mathcal{E}_v(G(v))$ we have that, for all $k \in \mathbb{R}$,

$$\mathcal{E}_v(k) \geq \mathcal{E}_v(G(v)) + \psi(k - G(v)) = \mathcal{E}_v(R(v)) + \psi(k - R(v)).$$

Hence, for $k > G(v) = R(v)$, we have that $\phi(v) \leq \psi \leq \frac{\mathcal{E}_v(k)-\mathcal{E}_v(R(v))}{(k-R(v))}$ and (5.2) follows.

For the final part, fix $u < v$ with $u, v \in A^n_<$. Then $R(v) \leq Q(v) < G(v) < S(v)$ and we have equality in (5.2) so that $\phi(v) = \inf_{k>G(v)} \frac{\mathcal{E}_v(k)-\mathcal{E}_v(R(v))}{k-R(v)}$. Moreover, $\mathcal{E}_v \geq \mathcal{E}^c_v$ on $[R(v), S(v)]$ and $R(v) \leq R(u) < G(u)$ so that $\mathcal{E}_v \geq \mathcal{E}^c_v$ on $[G(u), G(v)]$ and

$$\inf_{k\in[G(u),G(v)]} \frac{\mathcal{E}_v(k) - \mathcal{E}_v(R(v))}{k - R(v)} \geq \frac{\mathcal{E}_v(S(v)) - \mathcal{E}_v(R(v))}{S(v) - R(v)} = \phi(v).$$

Then the second equality in (5.3) follows. □

Lemma 5.2. *(i)* $\lim_{u\uparrow 1} \phi(u) = 0$.

(ii) For $u, v \in (0,1)$ with $u < v$ we have $\phi(v) \geq \phi(u) - (v-u)$.

(iii) On each $A^n_<$, $n \geq 1$, for $u, v \in A^n_<$ with $u < v$, we have $\phi(v) \leq \phi(u)$ so that ϕ is non-increasing.

In light of (i) above we may extend the domain of ϕ to $(0,1]$ by setting $\phi(1) = 0$, and (ii) still holds, even with $v = 1$.

Corollary 5.3. *On each $A^n_<$ we have that ϕ is Lipschitz continuous. Moreover, ϕ is absolutely continuous on $A^n_<$ and there exists a function $\phi' : A^n_< \to \mathbb{R}$ such that for $u, v \in A^n_<$,*

$$\phi(v) - \phi(u) = \int_u^v \phi'(w)dw.$$

Proof of Lemma 5.2. (i) First consider the limit on ϕ. We have $0 \leq \phi(u) \leq \mathcal{E}'_u(S(u)+) = P'_\nu(S(u)+) - u \leq 1 - u$, and the result follows.

(ii) Suppose $G(u) = S(v)$ and then $G(u) = G(v) = S(u) = S(v)$. Then $L^{\mathcal{E}_v,\phi(v)}_{G(v)}$ lies below $\mathcal{E}_v$, which in turn lies below $\mathcal{E}_u$ (with equality at $G(u) = G(v)$) and so $\phi(v) \in \partial\mathcal{E}_u(G(u))$. Then $\phi(v) \geq \phi(u) \geq \phi(u) - (v-u)$.

Suppose $G(u) < S(v)$ and $R(u) < G(v)$. Then from Lemma 5.1,

$$\begin{aligned}
\phi(v) &= \sup_{k<G(v)} \frac{\mathcal{E}_v(S(v)) - D(k)}{S(v)-k} \\
&\geq \frac{\mathcal{E}_v(S(v)) - D(R(u))}{S(v) - R(u)} \\
&= \frac{\mathcal{E}_u(S(v)) - D(R(u))}{S(v) - R(u)} - \frac{\mathcal{E}_u(S(v)) - \mathcal{E}_v(S(v))}{S(v) - R(u)} \\
&\geq \inf_{k>G(u)} \frac{\mathcal{E}_u(k) - D(R(u))}{k - R(u)} - (v-u) - \frac{\Gamma_{u,v}}{S(v) - R(u)} \\
&\geq \phi(u) - (v-u),
\end{aligned}$$

where we use $R(u) < G(v)$ for the first inequality, (4.6) and $G(u) < S(v)$ for the second and (5.2) and $\Gamma_{u,v} \leq 0$ for the third.

Finally, suppose $G(u) < S(v)$ and $R(u) = G(v)$. Then, $R(u) = Q(u) = G(u) = G(v) = S(u) < S(v)$. Since $L_{G(u)}^{\mathcal{E}_u,\phi(u)}$ lies below $\mathcal{E}_u$ everywhere, while $L_{S(v)}^{\mathcal{E}_v,\phi(v)}$ lies below $\mathcal{E}_v$ everywhere, using (4.9) together with the fact that $S(v) > G(v) = G(u)$ we have that $\phi(v) \geq \mathcal{E}_v'(S(v)-) = \mathcal{E}_u'(S(v)-) - (v-u) \geq \phi(u) - (v-u)$.

(iii) Fix $n \geq 1$ and let $u, v \in A_<^n$ with $u < v$. Then $R(v) < G(u) < S(u)$ and we have

$$\begin{aligned}\phi(u) &= \sup_{k<G(u)} \frac{\mathcal{E}_u(S(u)) - D(k)}{S(u) - k} \\ &\geq \frac{\mathcal{E}_u(S(u)) - D(R(v))}{S(u) - R(v)} \\ &\geq \inf_{k>G(u)} \frac{\mathcal{E}_v(k) - D(R(v))}{k - R(v)} = \phi(v),\end{aligned}$$

where we used that $\mathcal{E}_u \geq \mathcal{E}_v$ everywhere for the second inequality and (5.3) for the last equality. □

Fix $n \geq 1$ and let $u \in A_<^n$. Then

$$\begin{aligned}D(R(u)) - R(u)\phi(u) &= \mathcal{E}_u(S(u)) - \phi(u)S(u) \\ &= \{P_\nu(S(u)) - S(u)(\phi(u)+u)\} - \{P_\mu(G(u)) - G(u)u\}. \qquad (5.6)\end{aligned}$$

Recall that $\phi(u)$ is an element of both $\partial\mathcal{E}_u(R(u))$ and $\partial\mathcal{E}_u(S(u))$. Moreover, $u \in \partial P_{\mu_u}(G(u))$, but since $P_{\mu_u}(\cdot)$ is linear on $[G(u), +\infty)$, we also have that $u \in \partial P_{\mu_u}(S(u))$. Since the subdifferential of the sum of two functions is equal to the sum of individual subdifferentials (at least provided both subdifferentials are non-empty, as in our case) and since $P_\nu(S(u)) = \mathcal{E}_u(S(u)) + P_{\mu_u}(S(u))$, we have that $(\phi(u)+u) \in \partial P_\nu(S(u))$.

Lemma 5.4. *For each $n \geq 1$ and $u, v \in A_<^n$ with $u < v$, we have that*

1. $D(R(u)) - R(u)\phi(u) - \{D(R(v)) - R(v)\phi(v)\} = \int_u^v \phi'(w)R(w)dw$,
2. $P_\mu(G(u)) - G(u)u - \{P_\mu(G(v)) - G(v)v\} = \int_u^v G(w)dw$,
3. $P_\nu(S(u)) - (\phi(u)+u)S(u) - \{P_\nu(S(v)) - (\phi(v)+v)S(v)\} = \int_u^v (\phi'(w)+1)S(w)dw$.

Proof. Let $A_<$ be as in (4.13) and for each $n \geq 1$ write $A_<^n := (u_-^n, u_+^n)$. Recall that R is decreasing on each $A_<^n$. Furthermore, since $R \leq G$ everywhere, $R(u) \leq G(u_-^n+)$ for all $u \in A_<^n$.

Define, for each $n \geq 1$, conjugate functions $C_D^n, C_{P_\nu}^n, C_{P_\mu}^n : (-1, 1) \to \mathbb{R}$ by

$$\begin{aligned}C_D^n(\theta) &= \inf_{-\infty<x\leq G(u_-^n+)} \{D(x) - \theta x\}, \\ C_{P_\nu}^n(\theta) &= \inf_{x\in\mathbb{R}} \{P_\nu(x) - \theta x\}, \\ C_{P_\mu}^n(\theta) &= \inf_{x\in\mathbb{R}} \{P_\mu(x) - \theta x\}.\end{aligned}$$

Note that all three conjugates are concave and thus differentiable almost everywhere.

We prove the first statement, the other two being similar. Fix $u, v \in A_<^n$ with $u < v$. Since $\phi(u) \in \partial\mathcal{E}_u(R(u))$ and $\phi(v) \in \partial\mathcal{E}_v(R(v))$, and $\mathcal{E}_u(R(u)) = D(R(u))$ and $\mathcal{E}_v(R(v)) = D(R(v))$,

$$\begin{aligned}D(R(u)) - R(u)\phi(u) - \{D(R(v)) - R(v)\phi(v)\} &= C_D^n(\phi(u)) - C_D^n(\phi(v)) \\ &= -\int_{\phi(u)}^{\phi(v)} (C_D^n)'(\chi)d\chi \\ &= -\int_u^v \phi'(w)(C_D^n)'(\phi(w))dw,\end{aligned}$$

where we used that, by Corollary 5.3, $\phi(\cdot)$ is differentiable almost everywhere on $A^n_<$.

Moreover, $(C^n_D)'(\phi(w)) = -(D')^{-1}(\phi(w)) = -R(w)$ Lebesgue almost everywhere, and hence

$$D(R(u)) - R(u)\phi(u) - \{D(R(v)) - R(v)\phi(v)\} = \int_u^v \phi'(w)R(w)dw,$$

as required. □

Using (5.6) and Lemma 5.4 we have that for $u < v$ with $u, v \in A^n_<$,

$$\int_u^v \phi'(w)R(w)dw = \int_u^v (\phi'(w)+1)S(w)dw - \int_u^v G(w)dw.$$

Since $n \geq 1$ and $u, v \in A^n_<$ were arbitrary, $\phi'(w)R(w) = (\phi'(w)+1)S(w) - G(w)$ Lebesgue almost everywhere on each $A^n_<$. Hence,

$$\phi'(w) = -\frac{S(w) - G(w)}{S(w) - R(w)} \quad \text{almost everywhere on } A_<.$$

6 Left continuity

In Sections 4 and 5 we allowed G to be any quantile function. To help simplify the analysis going forward, from now on we assume that G is the left-continuous quantile function associated with μ.

Proposition 6.1. *Suppose G is left-continuous.*

(i) *S is left-continuous.*

(ii) *ϕ is left-continuous: $\phi(v-) := \lim_{u\uparrow v}\phi(u) = \phi(v)$. Moreover, for each $v \in (0,1)$ the right limit $\phi(v+)$ exists and $\phi(v+) := \lim_{w\downarrow v}\phi(w) = \lim_{w\downarrow v}(\mathcal{E}^c_v)'(S(w)-)$.*

(iii) *R satisfies $R(v) \leq \liminf_{u\uparrow v} R(u)$.*

Before proving Proposition 6.1 we record a couple of lemmas whose proofs are given in Appendix A.3.

Lemma 6.2. *For each $v \in (0,1)$,*

$$\lim_{u\uparrow v}\mathcal{E}_u(k) = \lim_{u\downarrow v}\mathcal{E}_u(k) = \mathcal{E}_v(k) \quad \textit{and} \quad \lim_{u\uparrow v}\mathcal{E}^c_u(k) = \lim_{u\downarrow v}\mathcal{E}^c_u(k) = \mathcal{E}^c_v(k) \quad \textit{for all } k \in \mathbb{R}. \tag{6.1}$$

Lemma 6.3. *Suppose that $f : (0,1) \to \mathbb{R}$ is non-decreasing. Then*

$$\lim_{u\uparrow v}\mathcal{E}^c_u(f(u)) = \mathcal{E}^c_v(f(v-)).$$

If in addition, for each $u \in (0,1)$, $\mathcal{E}_u$ is non-decreasing on $(f(u), \infty)$, then

$$\lim_{u\uparrow v}\mathcal{E}_u(f(u)) = \mathcal{E}_v(f(v-)).$$

We are now ready to prove Proposition 6.1.

Proof of Proposition 6.1(i). By definition, $G \leq S$, while by Theorem 4.9, S is non-decreasing. Furthermore, $\mathcal{E}_u$ is non-decreasing on $(S(u), \infty)$, and therefore Lemma 6.3 applies with $f = S$.

Fix $u, v \in (0,1)$ with $u < v$. Then $G(u) \leq S(u) \leq S(v)$ and therefore

$$G(v) \leq S(v-) \leq S(v).$$

We will now show that $S(v) \le S(v-)$. Since $\mathcal{E}^c_u(S(u)) = \mathcal{E}_u(S(u))$, letting $u \uparrow v$ on both sides and using Lemma 6.3, we have that

$$\mathcal{E}^c_v(S(v-)) = \lim_{u\uparrow v} \mathcal{E}^c_u(S(u)) = \lim_{u\uparrow v} \mathcal{E}_u(S(u)) = \mathcal{E}_v(S(v-)),$$

which implies that

$$S(v-) \in \{k : k \ge G(v), \mathcal{E}^c_v(k) = \mathcal{E}_v(k)\}.$$

It follows that

$$S(v-) \ge \inf\{k : k \ge G(v), \mathcal{E}^c_v(k) = \mathcal{E}_v(k)\} = S(v). \qquad \square$$

We now turn to the continuity properties of ϕ. The proof of Proposition 6.1(ii) is based on the following lemma.

Lemma 6.4. *For $u, v \in (0,1)$ with $u < v$, and all $k \in \mathbb{R}$,*
(i) $(\mathcal{E}^c_u)'(k-) - (v-u) \le (\mathcal{E}^c_v)'(k-) \le (\mathcal{E}^c_u)'(k-)$,
(ii) $(\mathcal{E}^c_u)'(k+) - (v-u) \le (\mathcal{E}^c_v)'(k+) \le (\mathcal{E}^c_u)'(k+)$.

Proof. (i) First we show that for $u, v \in (0,1)$ with $u < v$, $(\mathcal{E}^c_v)'(k-) \le (\mathcal{E}^c_u)'(k-)$, for all $k \in \mathbb{R}$.

Fix $k \in \mathbb{R}$ and suppose $\theta \in \partial\mathcal{E}^c_u(k)$. We show that $(\mathcal{E}^c_v)'(k-) \le \theta$ and then since $\theta \in \partial\mathcal{E}^c_u(k)$ is arbitrary,

$$(\mathcal{E}^c_v)'(k-) \le \inf\{\theta : \theta \in \partial\mathcal{E}^c_u(k)\} = (\mathcal{E}^c_u)'(k-)$$

as required.

Note that $\theta \in (0, 1-u)$. If $\theta \in [1-v, 1-u)$ then we have $(\mathcal{E}^c_v)'(k-) \le 1-v \le \theta$. Hence in what follows we may assume k and θ are such that $\theta \in \partial\mathcal{E}^c_u(k)$ and $\theta < 1-v$.

Consider the case $k > R(u)$. Given $\theta \in \partial\mathcal{E}^c_u(k)$ there exists $j \ge k$ such that $\mathcal{E}_u(j) = \mathcal{E}^c_u(j)$ and $\theta \in \partial\mathcal{E}^c_u(j)$. (If $R(u) < k \le S(u)$ then $j = S(u)$ is one choice and since $j > R(u)$ we may assume $j \ge S(u)$). Thus

$$\mathcal{E}_u(\ell) \ge \mathcal{E}_u(j) + \theta(\ell - j), \qquad \text{for all } \ell \in \mathbb{R}. \tag{6.2}$$

By the remarks in the previous paragraph we may assume $\theta \in (0, 1-v)$. Then we can find m such that both $\mathcal{E}^c_v(m) = \mathcal{E}_v(m)$ and $\theta \in \partial\mathcal{E}^c_v(m)$, and then

$$\mathcal{E}_v(\ell) \ge \mathcal{E}_v(m) + \theta(\ell - m), \qquad \text{for all } \ell \in \mathbb{R}. \tag{6.3}$$

Then $\mathcal{E}_u(m) \ge \mathcal{E}_u(j) + \theta(m-j)$ and $\mathcal{E}_v(j) \ge \mathcal{E}_v(m) + \theta(j-m)$ and adding

$$\mathcal{E}_v(j) - \mathcal{E}_u(j) \ge \mathcal{E}_v(m) - \mathcal{E}_u(m).$$

But $\mathcal{E}_v(\cdot) - \mathcal{E}_u(\cdot)$ is decreasing, and strictly decreasing on $(G(u+), \infty)$. Therefore, provided $j > G(u+)$, $m \ge j$. Then $(\mathcal{E}^c_v)'(k-) \le (\mathcal{E}^c_v)'(j-) \le (\mathcal{E}^c_v)'(m-) \le \theta$ as required.

If $S(u) \le j \le G(u+)$ we can find m as above with $\mathcal{E}^c_v(m) = \mathcal{E}_v(m)$ and $\theta \in \partial\mathcal{E}^c_v(m)$. If $m \ge j$ then the above argument still works and $(\mathcal{E}^c_v)'(k-) \le (\mathcal{E}^c_u)'(k-)$. So, suppose $m < j$. We begin by arguing that in this case we must have $\mathcal{E}_v(j) = \mathcal{E}_v(m) + \theta(j-m)$. If not then by (6.3) $\mathcal{E}_v(j) > \mathcal{E}_v(m) + \theta(j-m)$. But $m < j \le G(u+)$ and on $(-\infty, G(u+)]$ we have $\mathcal{E}_u = \mathcal{E}_v$. Hence $\mathcal{E}_u(j) > \mathcal{E}_u(m) + \theta(j-m)$, or equivalently $\mathcal{E}_u(m) < \mathcal{E}_u(j) + \theta(m-j)$ contradicting (6.2). It follows that we must have $\mathcal{E}_v(j) = \mathcal{E}_v(m) + \theta(j-m)$ and hence

$$\mathcal{E}_v(\ell) \ge \mathcal{E}_v(m) + \theta(\ell - m) = \mathcal{E}_v(j) + \theta(\ell - j), \qquad \text{for all } \ell \in \mathbb{R}$$

and $\mathcal{E}^c_v(j) = \mathcal{E}_v(j)$ and $\theta \in \partial\mathcal{E}^c_v(j)$. In particular we could take $m = j$, and for this choice of m we have $(\mathcal{E}^c_v)'(k-) \le (\mathcal{E}^c_v)'(j-) \le \theta$ and since $\theta \in \partial\mathcal{E}^c_u(k)$ is arbitrary, $(\mathcal{E}^c_v)'(j-) \le (\mathcal{E}^c_u)'(k-)$.

Now suppose $k \le R(u) \wedge R(v)$. Since $\mathcal{E}_u^c = \mathcal{E}_v^c$ on $(-\infty, R(u) \wedge R(v)]$ we have $(\mathcal{E}_v^c)'(k-) = (\mathcal{E}_u^c)'(k-)$.

Finally, suppose $R(v) < k \le R(u) \le S(v)$. If $\psi = (\mathcal{E}_v^c)'(k-)$ then $\psi \in \partial\mathcal{E}_v^c(R(v))$ and since $\mathcal{E}_v^c(R(v)) = \mathcal{E}_u^c(R(v))$ it follows that $\psi \in \partial\mathcal{E}_u^c(R(v))$. Then, since $k > R(v)$, if $\theta \in \partial\mathcal{E}_u^c(k)$ we must have $\theta \ge \psi$. Then $(\mathcal{E}_u^c)'(k-) \ge (\mathcal{E}_v^c)'(k-)$.

The proof that $(\mathcal{E}_u^c)'(k-) - (v-u) \le (\mathcal{E}_v^c)'(k-)$ is similar, but based on the fact that $\mathcal{E}_u(k) - \mathcal{E}_v(k) - k(v-u)$ is decreasing in k.

(ii) Given that $(\mathcal{E}_w^c)'(k+) = \lim_{j \downarrow k} (\mathcal{E}_w^c)'(j-)$ the result follows from (i) by taking limits from above. □

Proof of Proposition 6.1(ii). Fix $v \in (0,1)$. Consider the left limit $\phi(v-) = \lim_{u \uparrow v} \phi(u)$. We have, using Lemma 6.4 and the last representation of Lemma 4.3,

$$\begin{aligned}
\phi(v) = (\mathcal{E}_v^c)'(S(v)-) &= \lim_{u\uparrow v} (\mathcal{E}_v^c)'(S(u)-) \\
&\le \lim_{u\uparrow v} (\mathcal{E}_u^c)'(S(u)-) = \lim_{u\uparrow v} \phi(u) = \phi(v-) \\
&\le \lim_{u\uparrow v} (\mathcal{E}_u^c)'(S(v)-) \\
&\le \lim_{u\uparrow v} [(\mathcal{E}_v^c)'(S(v)-) + (v-u)] \\
&= (\mathcal{E}_v^c)'(S(v)-) = \phi(v).
\end{aligned}$$

The proof of the result for the right limit is similar. We have

$$\begin{aligned}
\lim_{v\downarrow u} (\mathcal{E}_u^c)'(S(v)-) &\le [\lim_{v\downarrow u} (\mathcal{E}_v^c)'(S(v)-) + (v-u)] \\
&= \lim_{v\downarrow u} (\mathcal{E}_v^c)'(S(v)-) = \lim_{v\downarrow u} \phi(v) = \phi(u+) \\
&\le \lim_{v\downarrow u} (\mathcal{E}_u^c)'(S(v)-). \quad \square
\end{aligned}$$

Proof of Proposition 6.1(iii). We are left to prove the lower semi-continuity (from the left) of R. Fix $v \in (0,1)$. Since $R \le G$ everywhere, $\liminf_{u\uparrow v} R(u) \le G(v)$. Moreover, for $u \in (0,1)$ with $u < v$, by the left-monotonicity we have that $R(v) \notin (R(u), S(u))$, and therefore

$$R(v) \notin (\inf_{u<v} R(u), \sup_{u<v} S(u) = S(v)).$$

Then, if $R(v) < G(v) \le S(v)$, $R(v) \le \inf_{u<v} R(u) \le \liminf_{u\uparrow v} R(u)$, and we are done. So, suppose that $R(v) = G(v) = S(v)$. Let $L = L_{G(v)}^{\mathcal{E}_v, \phi(v)}$. Then $D > L$ to the left of $G(v)$ and $\{k : k < G(v),\ D(k) = L(k)\} = \emptyset$. Furthermore, if $G(w) = G(v)$ for some $w \in (0, v)$, then, for all $u \in (w, v)$, $G(w) = G(u) = S(u) = G(v) = S(v)$, and $\mathcal{E}_u^c = \mathcal{E}_u = \mathcal{E}_v = \mathcal{E}_v^c$ at $G(v)$. Then, by Lemma 4.6, $\mathcal{E}_u^c = \mathcal{E}_v^c$ on $(-\infty, G(v)]$. Then $R(u) = G(u) = S(u) = G(v) = R(v)$ and it follows that $\liminf_{u\uparrow v} R(u) = R(v)$.

Hence suppose that $R(v) = G(v) = S(v)$ and $G(u) < G(v)$ for all $u \in (0, v)$. There are two cases to consider.

First, suppose that $D = \mathcal{E}_v > \mathcal{E}_v^c$ on an interval to left of $G(v)$, and such that this interval cannot be made larger without violating $\mathcal{E}_v > \mathcal{E}_v^c$. Then $\mathcal{E}_v^c$ is linear on this interval, and this interval must be finite since $\mathcal{E}_v \ge \mathcal{E}_v^c \ge 0$ and $\mathcal{E}_v^c$ can only change slope at points where $\mathcal{E}_v = \mathcal{E}_v^c$. But, if the interval is finite, $\{k : k \le G(u), D(k) = L_{G(u)}^{\mathcal{E}_u, \phi(u)}(k)\} \ne \emptyset$, a contradiction.

Second, suppose that, for each $u \in (0, v)$ there exists $k_u \in (G(u), G(v))$ such that $D(k_u) = \mathcal{E}_v^c(k_u)$. We first argue that, for all $u \in (0, v)$, $\mathcal{E}_v^c$ is not linear on $(G(u), G(v))$. Suppose there exists $u_0 \in (0, v)$ such that $\mathcal{E}_v^c$ is linear on $(G(u_0), G(v))$. Let $(k_0, G(v)] \supseteq (G(u_0), G(v)]$ be the largest interval of the form $(k, G(v)]$ on which $\mathcal{E}_v^c$ is linear. As

in the first case we must have that k_0 is finite, but then $R(v) = k_0 < G(v)$, a contradiction. Hence $\mathcal{E}_v^c$ is not linear to the left of $G(v)$. Now, let $\bar{k}_u := \sup\{k \geq k_u : \mathcal{E}_v^c(k) = L_{k_u}^{\mathcal{E}_v^c,(\mathcal{E}_v^c)'(k_u-)}(k)\}$. Then $D \geq \mathcal{E}_v^c > L_{k_u}^{\mathcal{E}_v^c,(\mathcal{E}_v^c)'(k_u-)}$ on $(\bar{k}_u, G(v))$ and $\bar{k}_u \in (G(u), \bar{y}_u)$, where $\bar{y}_u \in (G(u), G(v))$ is the point where $L_{k_u}^{\mathcal{E}_v^c,(\mathcal{E}_v^c)'(k_u-)}$ crosses L. (Note that neither $\bar{k}_u = \bar{y}_u < G(v)$ nor $\bar{k}_u < \bar{y}_u = G(v)$ can happen, since then either $\mathcal{E}_v^c = L$ on $(\bar{y}_u, G(v))$ or $\mathcal{E}_v^c = L_{k_u}^{\mathcal{E}_v^c,(\mathcal{E}_v^c)'(k_u-)}$ on $(k_u, G(v))$, respectively. But this contradicts the fact that $\mathcal{E}_v^c$ is not linear to the left of $G(v)$.) Then, $G(u) < \bar{k}_u \leq R(w) \leq R(v) = G(v)$ for all $w \in (F_\mu(\bar{k}_u), v)$, and therefore $G(u) \leq \liminf_{z\uparrow v} R(z) \leq R(v) = G(v)$. Using the left-continuity of G we conclude that $\liminf_{u\uparrow v} R(u) = R(v)$. □

In the next section we will need further two results.

Corollary 6.5. *If $S(w) > S(v+)$ for all $w > v$ then $\phi(v+) = (\mathcal{E}_v^c)'(S(v+)+)$. Otherwise, if $S(w) = S(v+)$ for some $w > v$ then $\phi(v+) = (\mathcal{E}_v^c)'(S(v+)-)$. In either case $\phi(v+) \geq (\mathcal{E}_v^c)'(S(v+)-)$.*

Proof. If $S(w) > S(v+)$ for all $w > v$ then $\lim_{w\downarrow v}(\mathcal{E}_v^c)'(S(w)-) = (\mathcal{E}_v^c)'(S(v+)+)$ where we use the fact that, for a convex function f, $\lim_{y\downarrow x} f'(y-) = f'(x+)$. Otherwise, if $S(w) = S(v+)$ for some $w > v$ then $\lim_{w\downarrow v}(\mathcal{E}_v^c)'(S(w)-) = (\mathcal{E}_v^c)'(S(v+)-)$. □

Let S^{-1} be the right-continuous inverse to the increasing function S. By our conventions, for $y \in \mathbb{R}$, $S(S^{-1}(y)) \leq y$ with equality whenever S is continuous at $S^{-1}(y)$. Moreover, if $w > S^{-1}(y)$ then $S(w) > y$.

Corollary 6.6. *For $y \in \mathbb{R}$, $\phi(S^{-1}(y)) \leq (\mathcal{E}_{S^{-1}(y)}^c)'(y-) \leq (\mathcal{E}_{S^{-1}(y)}^c)'(y+) \leq \phi(S^{-1}(y)+)$.*

Proof. Since $S(S^{-1}(y)) \leq y \leq S(S^{-1}(y)+)$,

$$\phi(S^{-1}(y)) = (\mathcal{E}_{S^{-1}(y)}^c)'(S(S^{-1}(y))-) \leq (\mathcal{E}_{S^{-1}(y)}^c)'(y-) \leq (\mathcal{E}_{S^{-1}(y)}^c)'(y+),$$

where we use Lemma 4.3 for the equality.

Moreover, if $y = S(S^{-1}(y+))$ then $(\mathcal{E}_{S^{-1}(y)}^c)'(y+) = (\mathcal{E}_{S^{-1}(y)}^c)'(S(S^{-1}(y)+)+)$ and $S(w) > y = S(S^{-1}(y)+)$ for all $w > S^{-1}(y)$, so that by Corollary 6.5, $\phi(S^{-1}(y)+) = (\mathcal{E}_{S^{-1}(y)}^c)'(S(S^{-1}(y)+)+)$ and $(\mathcal{E}_{S^{-1}(y)}^c)'(y+) = \phi(S^{-1}(y)+)$. Otherwise, if $y < S(S^{-1}(y)+)$ then, by Corollary 6.5, $(\mathcal{E}_{S^{-1}(y)}^c)'(y+) \leq (\mathcal{E}_{S^{-1}(y)}^c)'(S(S^{-1}(y)+)-) \leq \phi(S^{-1}(y)+)$. □

7 The candidate coupling is an embedding

We are now almost ready to prove Theorem 3.8. Let $U, V \sim U(0,1)$ be two independent uniform random variables. Then $X = G(U) \sim \mu$. On the other hand, $Y(U,V)$, if defined as in (3.5), may not be a random variable since R may not be measurable (see Remark 4.11). In order to deal with this, we introduce $T : (0,1) \to \mathbb{R}$ given by

$$T(u) = \begin{cases} G(u), & \text{if } S(u) = G(u) \\ R(u), & \text{if } S(u) > G(u) \end{cases} \qquad u \in (0,1). \tag{7.1}$$

The proof of the following lemma is postponed until Appendix A.4.

Lemma 7.1. *Let $T, S : (0,1) \to \mathbb{R}$ be defined by* (7.1) *and* (4.2)*. Then T is (Borel) measurable and the pair (T,S) is left-monotone with respect to G on $(0,1)$ in the sense of Definition 3.6.*

Define $Y : (0,1)^2 \to \mathbb{R}$ by $Y(u,v) = G(u)$ on $\{(u,v) \in (0,1)^2 : T(u) = S(u)\}$ and

$$Y(u,v) = T(u) I_{\left\{v \leq \frac{S(u)-G(u)}{S(u)-T(u)}\right\}} + S(u) I_{\left\{v > \frac{S(u)-G(u)}{S(u)-T(u)}\right\}} \tag{7.2}$$

otherwise. Then $Y(U,V)$ is a random variable and the martingale property (see the statement of Theorem 3.8) is a direct consequence of the definition of Y, see (7.2). We are left to show that $Y(U,V)$ has law ν.

Note that with our choice of left-continuous quantile function the definition of $A_<$ becomes $A_< = \{u \in (0,1) : G(u+) < S(u)\} = \cup_{n\geq 1} A^n_<$. Also recall that S^{-1} denotes the right-continuous inverse to the increasing function S.

Let $\chi = \mathcal{L}(Y(U,V))$. We want to show that $\chi = \nu$. We begin by describing the strategy of our proof. Fix $y \in \mathbb{R}$. Then, since $\mathbb{P}(Y(u,V) \leq y) = \frac{S(u)-G(u)}{S(u)-T(u)}$ whenever $T(u) \leq y < S(u)$,

$$\begin{aligned}\chi((-\infty,y]) = \mathbb{P}[Y(U,V)\leq y] &= \mathbb{P}[U \leq S^{-1}(y)] + \mathbb{P}[Y(U,V)\leq y, U > S^{-1}(y)] \\ &= S^{-1}(y) + \int_{S^{-1}(y)}^1 \frac{S(u)-G(u)}{S(u)-T(u)} I_{\{T(u)\leq y<S(u)\}} du \\ &= S^{-1}(y) + \int_{S^{-1}(y)}^1 \frac{S(u)-G(u)}{S(u)-T(u)} I_{\{T(u)\leq y\}} du, \qquad (7.3)\end{aligned}$$

where for the last equality we used that, since S^{-1} is right-continuous, $S(u) > y$ for all $u > S^{-1}(y)$.

We are free to include a multiplicative term $I_{\{S(u)>G(u)\}}$ in the integrand since off this set $\frac{S(u)-G(u)}{S(u)-T(u)} = 0$. (Recall that $T = R$ on $\{S(u) > G(u)\}$.) But $\{S(u) > G(u)\}$ and $A_<$ differ only by a countable (and thus null) set and on $A_<$ we have $\frac{S(u)-G(u)}{S(u)-R(u)} = -\phi'(u)$. Hence, we are interested in integrals of the form $-\int_a^1 \phi'(u) I_{\{R(u)\leq y\}} du$. If $I = (u_-, u_+)$ is an interval over which $R(\cdot) > y$ (and I cannot be made any larger without violating $R(\cdot) > y$) then we expect that $\phi(u_-) = \phi(u_+)$. Adding in such intervals would allow us to replace $-\int_a^1 \phi'(u) I_{\{R(u)\leq y\}} du$ with $-\int_a^1 \phi'(u) du = \phi(a) - \phi(1) = \phi(a)$ since $\phi(1) = 0$. Then

$$\mathbb{P}[Y(U,V) \leq y] = S^{-1}(y) + \phi(S^{-1}(y)). \qquad (7.4)$$

Further, for each $v \in (0,1)$, since $\phi(v) \in \partial\mathcal{E}_v(S(v))$ and $v \in \partial P_{\mu_v}(S(v))$ we have $(v + \phi(v)) \in \partial P_\nu(S(v))$, and then, if $S(v)$ is a continuity point of ν, $v + \phi(v) = P'_\nu(S(v)) = \nu((-\infty, S(v)])$. Hence, provided y is a continuity point of ν and $S(S^{-1}(y)) = y$, we have

$$\chi((-\infty,y]) = \mathbb{P}[Y(U,V)\leq y] = S^{-1}(y) + \phi(S^{-1}(y)) = P'_\nu(S(S^{-1}(y))) = P'_\nu(y) = \nu((-\infty,y])$$

as desired.

There are at least three issues we must overcome to complete this analysis. First, the derivative ϕ' need not exist everywhere. Second, there may be a countably infinite number of intervals $A^m_< = (u^m_-, u^m_+)$ which we must add, on each of which $R(\cdot) > y$. Third, we need a refined argument to cover the case where $S(S^{-1}(y)) < y$.

To deal with the issues about ϕ we introduce a family of modified functions $\psi_{v,x}$, each member of which is monotonically decreasing and Lipschitz, and therefore has a derivative almost everywhere. The introduction of the monotonic function ψ also allows us to easily add the missing intervals since ψ is constant on those intervals by construction; moreover the intervals where ψ is decreasing are precisely the intervals where $R(\cdot) \leq y$. The case where $S(S^{-1}(y)) < y$ requires a careful definition of the initial value of ψ and an additional argument.

Note that it is sufficient to prove $\chi((-\infty,y]) = \nu((-\infty,y])$ at continuity points of ν only. Indeed it is sufficient to prove $\chi((-\infty,y]) = \nu((-\infty,y])$ on a dense set of values of y, so we may also restrict attention to y which are continuity points of μ also.

Lemma 7.2. *Suppose y is a continuity point of ν. Then $\phi(S^{-1}(y)) \leq P'_\nu(y) - S^{-1}(y) \leq \phi(S^{-1}(y)+)$.*

Proof. We have $\mathcal{E}^c_{S^{-1}(y)} = \mathcal{E}_{S^{-1}(y)}$ on $[S(S^{-1}(y)), \infty)$ and $y \geq S(S^{-1}(y))$ so that $\mathcal{E}^c_{S^{-1}(y)}(y) = \mathcal{E}_{S^{-1}(y)}(y)$ and $(\mathcal{E}^c_{S^{-1}(y)})'(y+) = \mathcal{E}'_{S^{-1}(y)}(y+)$. Also, since y is a continuity point of ν, $\mathcal{E}'_{S^{-1}(y)}(y+) = P'_\nu(y+) - S^{-1}(y) = P'_\nu(y) - S^{-1}(y)$. Then the result follows by Corollary 6.6. □

Definition 7.3. *For $v \in (0,1)$ and $x \in [0,1)$, define $\psi_{v,x} : [v,1] \to [0,1]$ by $\psi_{v,x}(w) = x \wedge \inf_{v<u\leq w}\{\phi(u)\}$.*

Here we use the fact that we have extended the domain of ϕ to $(0,1]$ whence also $\psi_{v,x}(1) = 0$. Note that the only relevant cases are when $x \in [\phi(v), \phi(v+)]$. See Figure 5.

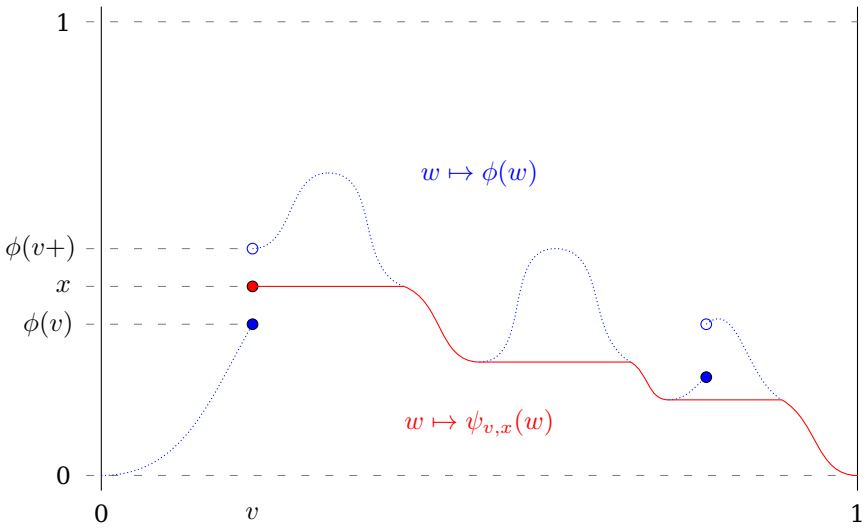

Figure 5: Construction of ψ (solid curve) from ϕ (dotted curve). ϕ can have jumps, but only upwards, and for $x \in [\phi(v), \phi(v+)]$, $w \mapsto \psi_{v,x}(w) = x \wedge \inf_{v<u\leq w}\{\phi(u)\}$ is monotonically decreasing and Lipschitz with Lipschitz constant 1.

The proof of the following lemma is quite straightforward but is deferred to Appendix A.4.

Lemma 7.4. *For $v \in (0,1)$ and $x \in [\phi(v), \phi(v+)]$, $\psi_{v,x}$ is decreasing and absolutely continuous.*

It follows from a combination of Lemma 7.2 and Lemma 7.4 that $\psi_{S^{-1}(y),P'_\nu(y)-S^{-1}(y)}$ is decreasing and absolutely continuous on $(S^{-1}(y), 1)$, and hence $\psi'_{S^{-1}(y),P'_\nu(y)-S^{-1}(y)}$ is defined almost everywhere on $(S^{-1}(y), 1)$.

The next lemma says that the places $\{u : u \in (S^{-1}(y), 1)\}$ where $\psi_{S^{-1}(y),P'_\nu(y)-S^{-1}(y)}$ decreases are essentially the places where $\psi_{S^{-1}(y),P'_\nu(y)-S^{-1}(y)}(u) = \phi(u)$ and $S(u) > G(u+)$.

Lemma 7.5. *For $u \in (0,1)$ and $x \in [\phi(u), \phi(u+)]$ the following set inclusion holds:*

$$\{w : w > u, \psi'_{u,x}(w) < 0\} \subseteq (\{w : w > u, S(w) > G(w+)\} \cap \{w : w > u, \psi_{u,x}(w) = \phi(w)\}) \cup \mathcal{N}_u \quad (7.5)$$

where $\mathcal{N}_u$ is a set of measure zero.

Proof. Let

$$\mathcal{N}_u = \{w \in [u,1) \text{ such that } S(w+) > S(w) \text{ or } G(w+) > G(w) \text{ or } \psi'_{u,x}(w) \text{ does not exist}\}.$$

To prove the lemma we need to show that if $w \in \mathcal{N}_u^c$ and $S(w) = G(w)$ then $\psi'_{u,x}(w) = 0$ and also that if $w \in \mathcal{N}_u^c$ and $\phi(w) > \psi_{u,x}(w)$ then $\psi'_{u,x}(w) = 0$. The second of these is immediate: if $\phi(w) - \psi_{u,x}(w) > \epsilon > 0$, then for $0 < \delta \le \epsilon$ we have $\phi(w+\delta) \ge \phi(w) - \delta > \psi_{u,x}(w)$, and hence $\psi_{u,x}(w+\epsilon) = \psi_{u,x}(w)$. Hence $\psi'_{u,x}(w+) = 0$ and if $\psi'_{u,x}(w)$ is defined — as it must be since $w \in \mathcal{N}_u^c$ — it must take the value 0.

Hence, to complete the proof of the lemma it is sufficient to show that if w is such that S and G are continuous at w and $\psi'_{u,x}(w)$ exists, then $S(w) = G(w)$ implies $\psi'_{u,x}(w) = 0$.

We split into two cases: $R(w) < G(w) = S(w)$ and $R(w) = G(w) = S(w)$.

Case 1: $R(w) < G(w) = S(w)$. Using (5.1) and (5.2) for the inequalities we have that

$$\begin{aligned}
\phi(w+h) &\ge \frac{\mathcal{E}_{w+h}(S(w+h)) - D(R(w))}{S(w+h) - R(w)} \\
&= \frac{\mathcal{E}_w(S(w+h)) - D(R(w))}{S(w+h) - R(w)} + \frac{\mathcal{E}_{w+h}(S(w+h)) - \mathcal{E}_w(S(w+h))}{S(w+h) - R(w)} \\
&\ge \phi(w) + \frac{\mathcal{E}_{w+h}(S(w+h)) - \mathcal{E}_w(S(w+h))}{S(w+h) - R(w)}.
\end{aligned}$$

It is sufficient to show that $\liminf_{h\downarrow 0} \frac{\phi(w+h)-\phi(w)}{h} \ge 0$. This will follow if

$$\lim_{h\downarrow 0} \frac{1}{h} \left(\mathcal{E}_w(S(w+h)) - \mathcal{E}_{w+h}(S(w+h))\right) \le 0.$$

But, by (4.5), $\mathcal{E}_w(S(w+h)) - \mathcal{E}_{w+h}(S(w+h)) = h(S(w+h) - G(w)) + P_\mu(G(w+h)) - P_\mu(G(w)) - (w+h)(G(w+h) - G(w))$. Moreover, $P_\mu(G(w+h)) - P_\mu(G(w)) = \tilde{w}(G(w+h) - G(w))$ for some $\tilde{w} \in [w, w+h]$ and then $|P_\mu(G(w+h)) - P_\mu(G(w)) - (w+h)(G(w+h) - G(w))| \le h(G(w+h) - G(w))$. We conclude $|\frac{1}{h}\left(\mathcal{E}_w(S(w+h)) - \mathcal{E}_{w+h}(S(w+h))\right)| \le S(w+h) - G(w) + G(w+h) - G(w) \to 0$, by the continuity of S and G at w and the fact that $S(w) = G(w)$.

Case 2: $R(w) = G(w) = S(w)$. Then $\phi(w) = D'(G(w)-)$ and $L_{G(w)}^{\mathcal{E}_w,\phi(w)} \le \mathcal{E}_w^c \le \mathcal{E}_w$. Note that $D = \mathcal{E}_w$ to the left of $G(w)$.

Pick $\theta \in (0, \phi(w))$. Define $z_\theta = \sup\{z : \theta \in \partial\mathcal{E}_w^c(z)\}$. Then $z_\theta < G(w)$. Note that $\theta \in \partial\mathcal{E}_w^c(z_\theta)$ and $D(z_\theta) = \mathcal{E}_w^c(z_\theta)$. If $G(u+) = G(w)$ then $\phi(v) = \phi(w)$ for all $v \in (u, w]$, and thus $\psi_{u,x} = x$ on $[u, w]$. It follows that $\psi'_{u,x}(w) = \psi'_{u,x}(w-) = 0$. Otherwise $G(u+) < G(w)$, and by choosing θ large enough we may assume $z_\theta > G(u+)$.

Let $u_\theta = F_\mu(z_\theta)$. Then $u < u_\theta < w$, and either $G(u_\theta) = z_\theta$ or $G(u_\theta) < z_\theta < G(u_\theta+)$. We claim that $\phi(u_\theta) \le \theta$. If $G(u_\theta) = z_\theta$ then by Lemma 4.6 we have that $\mathcal{E}_{u_\theta}^c = \mathcal{E}_w^c$ on $(-\infty, G(u_\theta) = z_\theta]$, and therefore $\phi(u_\theta) = \inf\{z : z \in \partial\mathcal{E}_w^c(z_\theta)\} \le \theta$. Suppose $G(u_\theta) < z_\theta < G(u_\theta+)$. If $\mathcal{E}_{u_\theta}(G(u_\theta)) = D(G(u_\theta)) = \mathcal{E}_w^c(G(u_\theta))$ then using the same argument (i.e., by Lemma 4.6) we have that $\phi(u_\theta) = \inf\{z : z \in \partial\mathcal{E}_w^c(G(u_\theta))\}$. But, by the monotonicity of the subdifferential of $\mathcal{E}_w$, $\inf\{z : z \in \partial\mathcal{E}_w^c(G(u_\theta))\} \le \inf\{z : z \in \partial\mathcal{E}_w^c(z_\theta)\} \le \theta$. Hence $\phi(u_\theta) \le \theta$. Finally, if $\mathcal{E}_{u_\theta}(G(u_\theta)) = D(G(u_\theta)) > \mathcal{E}_w^c(G(u_\theta))$, then since $\mathcal{E}_{u_\theta}$ is convex and coincides with D on $[G(u_\theta), G(u_\theta+)]$, it follows that $S(u_\theta) \le z_\theta$ and $\mathcal{E}_{u_\theta}(S(u_\theta)) = D(S(u_\theta)) = \mathcal{E}_w^c(S(u_\theta))$. Then $\mathcal{E}_{u_\theta}^c = \mathcal{E}_w^c$ on $(-\infty, S(u_\theta)]$ and, by Lemma 4.3, we again have that $\phi(u_\theta) = (\mathcal{E}_u^c)'(S(u_\theta)-) = (\mathcal{E}_w^c)'(S(u_\theta)-) \le (\mathcal{E}_w^c)'(z_\theta-) \le \theta$.

Therefore in all cases $\psi_{u,x}(w) \le \phi(u_\theta) \le \theta < \phi(w)$. By Lemma 5.2(ii) we than have that $\phi(v) > \psi_{u,x}(w)$ for $v \in (w, w + \phi(w) - \psi_{u,x}(w))$ and thus ψ is constant on this interval.

It follows that if $\psi'_{u,x}(w)$ exists, then $\psi'_{u,x}(w) = 0$. □

The final result we need identifies the set where $R(v) \le y$ with the set where $\psi = \phi$. The proof of this result has to cover several cases and consequently is quite long and technical. For this reason the proof is postponed to Appendix A.5.

Proposition 7.6. *Suppose y is a continuity point of μ and ν. Then*

$$\{v : v > S^{-1}(y), R(v) \le y\} = \{v : v > S^{-1}(y), \psi_{S^{-1}(y),P_\nu(y)-S^{-1}(y)}(v) = \phi(v)\}. \tag{7.6}$$

We are now ready to prove the main results.

Proposition 7.7. *Suppose y is a continuity point of both μ and ν. Then*

$$\int_{S^{-1}(y)}^{1} \frac{S(v) - G(v)}{S(v) - T(v)} I_{\{T(v)\le y\}} dv = P'_\nu(y) - S^{-1}(y). \tag{7.7}$$

Proof. From Proposition 7.6 we know that under the hypotheses of the proposition

$$\{u : u > S^{-1}(y), R(u) \le y\} = \{u : u > S^{-1}(y), \psi_{<y>}(u) = \phi(u)\}. \tag{7.8}$$

where $\psi_{<y>}$ is shorthand for $\psi_{S^{-1}(y),P'_\nu(y)-S^{-1}(y)}$.

Then, using the fact that $A_< \subseteq \{v : G(v) < S(v)\}$ but the difference is a set of measure zero,

$$\begin{aligned}
\int_{S^{-1}(y)}^{1} \frac{S(v) - G(v)}{S(v) - T(v)} I_{\{T(v)\le y\}} dv &= \int_{S^{-1}(y)}^{1} \frac{S(v) - G(v)}{S(v) - R(v)} I_{\{R(v)\le y\}} I_{\{G(v)<S(v)\}} dv \\
&= \int_{S^{-1}(y)}^{1} \frac{S(v) - G(v)}{S(v) - R(v)} I_{\{R(v)\le y\}} I_{A_<} dv \\
&= -\int_{S^{-1}(y)}^{1} \phi'(v) I_{\{R(v)\le y\}} I_{A_<} dv \\
&= -\int_{S^{-1}(y)}^{1} \phi'(v) I_{\{\psi_{<y>}(v)=\phi(v)\}} I_{A_<} dv \\
&= -\int_{S^{-1}(y)}^{1} \psi'_{<y>}(v) I_{\{\psi_{<y>}(v)=\phi(v)\}} I_{A_<} dv \\
&= -\int_{S^{-1}(y)}^{1} \psi'_{<y>}(v) \left[I_{\{\psi_{<y>}(v)=\phi(v)\}} I_{A_<} + I_{\mathcal{N}} \right] dv,
\end{aligned}$$

where $\mathcal{N}$ is some set of measure zero.

For any absolutely continuous, decreasing function g we have

$$\int_a^b g'(z) dz = \int_a^b g'(z) I_{\{g'(z)<0\}} dz + \int_a^b g'(z) I_{\{g'(z)=0\}} dz = \int_a^b g'(z) I_A dz$$

for any set $A \supseteq \{g'(z) < 0\}$. We saw in Lemma 7.5 that

$$\begin{aligned}
&\{u : u > S^{-1}(y), \psi'_{<y>}(u) < 0\} \\
&\qquad \subseteq \left(\{u : u > S^{-1}(y), S(u) > G(u+)\} \cap \{\psi_{<y>}(u) = \phi(u)\} \right) \cup \mathcal{N}_{S^{-1}(y)}
\end{aligned} \tag{7.9}$$

for a well chosen set $\mathcal{N}_{S^{-1}(y)}$ of measure zero. Then,

$$\begin{aligned}
&\int_{S^{-1}(y)}^{1} \psi'_{<y>}(v) \left(I_{\{\psi_{<y>}(v)=\phi(v)\}} I_{\{G(v+)<S(v)\}} + I_{\mathcal{N}_{S^{-1}(y)}} \right) dv \\
= &\int_{S^{-1}(y)}^{1} \psi'_{<y>}(v) dv = \psi_{<y>}(1) - \psi_{<y>}(S^{-1}(y)) = -(P'_\nu(y) - S^{-1}(y))
\end{aligned}$$

with the last equality following from Lemma 5.2 and the fact that $\psi_{<y>}(1) = \phi(1) = 0$. □

Proposition 7.8. *Suppose that y is a continuity point of both μ and ν. Then $\chi((-\infty, y]) = \nu((-\infty, y])$.*

Proof. Combining (7.3) and (7.7) we find $\mathbb{P}(Y(U,V) \le y) = P'_\nu(y)$. □

Theorem 7.9. *Define T and S as in* (7.1) *and* (4.2)*. Then they are left-monotone with respect to G and define a construction as in Theorem 3.8 such that $\mathcal{L}(Y) = \nu$, where Y is as in 7.2. In particular, T and S define the (lifted) left-curtain coupling.*

Proof. This follows immediately from Lemma 7.1 and Proposition 7.8, where we note that it is sufficient to show that $\chi((-\infty, y]) = \nu((-\infty, y])$ on a dense subset of the reals and we may exclude non-continuity points of μ and ν. □

A Proofs

A.1 Convex hull

Proof of Lemma 2.2. Without loss of generality we assume that f^c is not equal to $-\infty$ everywhere. Since $f > f^c$, we also cannot have that $f^c = \infty$ everywhere.

We first deal with the case when $a, b \in \mathbb{R}$ with $a < b$. 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

$$f^c(x) < \frac{b-x}{b-a} f^c(a) + \frac{x-a}{b-a} f^c(b) = L^{f^c}_{a,b}(x). \tag{A.1}$$

Let $\eta = \inf_{y\in(a,b)}\{f(y) - L^{f^c}_{a,b}(y)\}$. If $\eta \ge 0$ then $f \ge L^{f^c}_{a,b}$ on (a, b) and $f \ge f^c \vee L^{f^c}_{a,b}$, 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 $z \in \operatorname{arginf}_{y\in[a,b]}\{f(y) - L^{f^c}_{a,b}(y)\}$. Since $f(k) - L^{f^c}_{a,b}(k) = f(k) - f^c(k) \ge 0$ for $k \in \{a, b\}$, $a, b \notin \operatorname{arginf}_{y\in[a,b]}\{f(y) - L^{f^c}_{a,b}(y)\}$, 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. The case when one (or both) of the endpoints of (a, b) are infinite can be reduced to the previous finite case. Indeed, suppose $f > f^c$ on (a, ∞) (resp. $(-\infty, b)$) with $a \in \mathbb{R}$ (resp. $b \in \mathbb{R}$). If f^c is not a straight line on (a, ∞) (resp. $(-\infty, b)$), then there exists $b \in \mathbb{R}$ (resp. $a \in \mathbb{R}$) with $a < b$ such that f^c is not a straight line on (a, b) and (A.1) holds for all $x \in (a, b)$. The rest of the argument remains the same. Finally, if $f > f^c$ on $\mathbb{R}$ but f^c is not linear, then we can find $a, b \in \mathbb{R}$ with $a < b$ such that f^c is not linear on (a, b) and (A.1) holds. We conclude as before. □

Proof of Lemma 2.4. If $f = \infty$ on $\mathbb{R}$ then $f^c = \infty$, $X(y) = y = Z(y)$, and $L^f_{X(y),Z(y)}(y) = \infty = f^c(y)$. Henceforth we exclude this degenerate case.

By continuity of f and necessarily of f^c, if $X(y) > -\infty$ then $f^c(X(y)) = f(X(y))$; similarly, if $Z(y) < \infty$ then $f^c(Z(y)) = f(Z(y))$.

Fix $y \in \mathbb{R}$. If $X(y) = y = Z(y)$ then $f^c(y) = f(y) = L^f_{y,y}(y) = L^f_{X(y),Z(y)}(y)$, as required.

So we may suppose $X(y) < Z(y)$. Suppose for now that $-\infty < X(y) < Z(y) < \infty$. By definition $f > f^c$ on $(X(y), Z(y))$. Then, by Lemma 2.2, f^c is linear on $(X(y), Z(y))$ and, by continuity of f, $f^c(X(y)) = f(X(y))$ and $f^c(Z(y)) = f(Z(y))$. Then $f^c(k) = L^{f^c}_{X(y),Z(y)}(k) = L^f_{X(y),Z(y)}(k)$ on $[X(y), Z(y)]$. Applying this result at $k = y$ we have $f^c(y) = L^f_{X(y),Z(y)}(y)$ as required.

We want to extend this result to the case where one, or both, of $\{X(y), Z(y)\}$ is infinite. Suppose $-\infty < X(y) < Z(y) = \infty$ (the case of $-\infty = X(y) < Z(y) < \infty$ can be treated symmetrically). By Lemma 2.2, f^c is a straight line on $(X(y), \infty)$ and by continuity of f, $f(X(y)) = f^c(X(y))$. Then $f^c(k) = f(X(y)) + \theta(k - X(y))$ on $[X(y), \infty)$, for some θ to be determined. Let $\phi = \liminf_{k\uparrow\infty} f(k)/k$; we show that $\phi = \theta$ for then $f^c(y) = L^f_{X(y),\infty}(y)$. We have

$$\theta = \lim_{k\uparrow\infty}\left[\frac{f^c(k) - f(X(y))}{k - X(y)}\right] \le \liminf_{k\uparrow\infty}\left[\frac{f(k) - f(X(y))}{k - X(y)}\right] = \liminf_{k\uparrow\infty}\frac{f(k)}{k} = \phi.$$

If $\theta < \phi$ then let $J = \inf_k(f(k) - \phi k)$. Then $f^c \vee (J + \phi k)$ is convex, is a minorant of f and is greater than f^c for large enough k, contradicting the maximality of f^c. Hence $\theta = \phi$, and then $f^c(y) = L^f_{X(y),Z(y)}(y)$.

Finally, suppose $-\infty = X(y) < Z(y) = \infty$. Then, by Lemma 2.2, f^c is a straight line, and either $f^c(y) = \alpha + \theta y$ with $\alpha, \theta \in \mathbb{R}$, or $f^c = -\infty$ or $f^c = \infty$. The latter is ruled out since $f^c \le f$ and at least for some x, $f < \infty$ by non-degeneracy.

Suppose $\psi + \phi > 0$. Then both $\psi, \phi > -\infty$. We first show that $f^c \ne -\infty$. Suppose that $\psi, \phi \in \mathbb{R}$. Let ϵ be given by $2\epsilon = \psi + \phi > 0$. Then $f(y) - (\phi - \epsilon)y$ is bounded below on $\mathbb{R}_+$ by $\beta_+ \in \mathbb{R}$ say and $f(y) - (\psi - \epsilon)|y|$ is bounded below on $\mathbb{R}_-$ by $\beta_- \in \mathbb{R}$. Then, if $\beta = \beta_+ \wedge \beta_-$, $f(y) \ge f^c \ge \beta + (\phi - \epsilon)y > -\infty$, for all $y \in \mathbb{R}$. If $\psi = \phi = \infty$, then by the continuity of f and since $f < \infty$ for some $x \in \mathbb{R}$, f is bounded below on $\mathbb{R}$ by some $\alpha \in \mathbb{R}$, and then $f^c \ge \alpha$ everywhere. Finally, suppose that $\psi = \infty$ and $\phi \in \mathbb{R}$ (the case of $\psi \in \mathbb{R}$ and $\phi = \infty$ follows by symmetry). Then, if $\epsilon > 0$, $f(y) - (\phi - \epsilon)y$ is bounded below on $\mathbb{R}$, and it follows that we cannot have $f^c = -\infty$.

So suppose $\psi + \phi > 0$ and $f^c(y) = \alpha + \theta y \le f(y)$, $y \in \mathbb{R}$. We show that this leads to a contradiction (to the fact that $-\infty = X(y) < Z(y) = \infty$). If $\psi + \phi > 0$ then either $\psi > -\theta$ or $\phi > \theta$ (or both). Suppose $\phi > \theta$ (the case of $\psi > -\theta$ can be treated similarly). Then, there exists $y_0 \in \mathbb{R}$ such that for $y \ge y_0$, $f(y) \ge f^c(y_0) + \phi(y - y_0) > f^c(y_0) + \theta(y - y_0) = f^c(y)$. But then $(\alpha + \theta y) \vee (f^c(y_0) + \phi(y - y_0))$ is a convex minorant of f which is greater than f^c.

Now suppose $\psi + \phi < 0$. Choose $z_n \uparrow \infty$ such that $f(z_n)/z_n \to \phi$ and $x_m \downarrow -\infty$ such that $f(x_m)/|x_m| \to \psi$. By (5) (i.e., the result of Rockafellar that $f^c(y) = \inf_{x\le y\le z} L^f_{x,z}(y)$), for big enough m and n, we have $f^c(y) \le L^f_{x_m,z_n}(y)$. Hence, if $-\infty < \phi + \psi < 0$, then $-\infty \le f^c(y) \le \lim_m(\lim_n L^f_{x_m,z_n}(y)) = \lim_m L^f_{x_m,\infty}(y) = L^f_{-\infty,\infty}(y) = -\infty$.

Similarly, if $\phi = -\infty$, then $L^f_{x,\infty}(y) = -\infty$ for $x < y$ and $\lim_m L^f_{x_m,\infty}(y) = -\infty$, irrespective of the value of ψ. If on the other hand, $\psi = -\infty$, then we reverse the order of taking limits and use $f^c(y) \le \lim_n(\lim_m L^f_{x_m,z_n}(y)) = \lim_n L^f_{-\infty,z_n}(y) = -\infty$. Thus, if $\phi + \psi = -\infty$ we have both $L^f_{-\infty,\infty} = -\infty$ and $f^c = -\infty$ and then $f^c = L^f_{X(y),Z(y)}$.

Finally, suppose $\phi + \psi = 0$ and $L^f_{-\infty,\infty}(y) = \gamma + \phi y$. For any $k \in \mathbb{R}$ we have that $\gamma \le f(k) - \phi k$ and hence $L^f_{-\infty,\infty}(k) = \gamma + \phi k \le f(k)$. Hence $L^f_{-\infty,\infty}$ is a convex minorant of f.

Suppose $\gamma = -\infty$, so that $L^f_{-\infty,\infty} = -\infty$. Let g be any convex function with $g \le f$. We show that g is identically equal to $-\infty$. Suppose not. Then there exists k such that $g(k) > -\infty$ and by the convexity of g, $g(y) \ge g(k) + (y - k)g'(k-)$. Dividing by y and letting $y \to \infty$ we conclude $\phi \ge g'(k-)$. Dividing by $|y|$ and letting $y \to -\infty$ we conclude $\psi \ge -g'(k-)$. Then $-\psi = \phi \ge g'(k-) \ge -\psi$ and $f(y) \ge g(y) \ge g(k) + \phi(y - k)$. Then $\inf_w(f(w) - \phi w) \ge g(k) - \phi(k) > -\infty$, contradicting the fact that $\gamma = -\infty$. Finally, since $g = -\infty$ is the only convex minorant of f we have that $f^c = -\infty = L^f_{-\infty,\infty}$.

Now suppose $\gamma \in \mathbb{R}$. We first show that f^c must be linear. Suppose not. Then, since $f^c \ge L^f_{-\infty,\infty}$, either there exists $k_+ \in \mathbb{R}$ such that $f^c > L^f_{-\infty,\infty}$ on (k_+, ∞) or there exists

$k_- \in \mathbb{R}$ such that $f^c > L^f_{-\infty,\infty}$ on $(-\infty, k_-)$ (or both). In either case $f^c(k) > f(k)$ for some (large or small enough) $k \in \mathbb{R}$, a contradiction.

Hence suppose $f^c(y) = \alpha + \theta y$, $y \in \mathbb{R}$, for some $\alpha, \theta \in \mathbb{R}$. We aim to show that $\theta = \phi$ and $\alpha = \gamma$ so that $f^c = L^f_{-\infty,\infty}$. Suppose $\theta < \phi$. Then for large enough y, $f^c(y) < L^f_{-\infty,\infty}(y)$, so that $f^c \vee L^f_{-\infty,\infty}$ is a convex minorant of f which is bigger than f^c, thus contradicting the maximality of f^c. $\theta > \phi$ can be ruled out similarly by considering large negative y; hence $\theta = \phi$. Finally, choose k_n such that $f(k_n) - \phi k_n \downarrow \gamma$. Then $0 \le f(k_n) - f^c(k_n) = f(k_n) - \phi k_n - \alpha \to \gamma - \alpha$. Hence $\gamma \ge \alpha$ and $f^c \le L^f_{-\infty,\infty} \le f$. Since f^c is the largest convex minorant we conclude $\gamma = \alpha$ and $f^c = L^f_{-\infty,\infty}$. □

Proof of Lemma 2.5. Fix $y \in \mathbb{R}$ and suppose $(x, z) \in \mathcal{B}(y)$ are such that $f(k) > L^f_{x,z}(k)$ for all $k \in (x, z)$. If $X(y) = -\infty$, then we trivially have that $X(y) = -\infty \le x$. Now suppose that $X(y)$ is finite. Using again that $f \ge f^c$ and f^c is convex we have that

$$f(k) > L^f_{x,z}(k) \ge L^{f^c}_{x,z}(k) \ge f^c(k), \quad k \in (x, z).$$

Therefore, if $X(y) \in (x, z)$, we have a contradiction since $f(X(y)) = f^c(X(y))$. A symmetric argument shows that $z \le Z(y)$. □

A.2 The Geometric construction

Proof of Lemma 4.4. First note that, since $\{k : D(k) > 0\} = (\ell_\nu, r_\nu) \supseteq (\ell_\mu, r_\mu)$ and by hypothesis $\mu(\{\ell_\nu\}) = 0$ (resp. $\mu(\{r_\mu\}) = 0$) in the case $\ell_\mu = \ell_\nu$ (resp. $r_\mu = r_\nu$), we have that $G(u) \in (\ell_\nu, r_\mu)$ for all $u \in (0, 1)$. Hence $S(u) < r_\nu$ (resp. $\ell_\nu < R(u)$) in the case $G(u) = S(u)$ (resp. $R(u) = G(u)$). Recall that if $R(u) = G(u)$, then $R(u) = Q(u) = G(u) = S(u)$ (see Lemma 4.1).

If $-\infty \le R(u) < \ell_\nu$ (or $-\infty = R(u) = \ell_\nu$), then $\mathcal{E}^c_u(R(u)) = \mathcal{E}_u(R(u)) = D(R(u)) = 0$ (or $\lim_{k \downarrow R(u)} \mathcal{E}_u(R(u)) = \lim_{k \downarrow R(u)} \mathcal{E}^c_u(R(u)) = \lim_{k \downarrow R(u)} D(k) = 0$). In both cases we have that $L^{\mathcal{E}^c_u}_{R(u),S(u)}$ has zero slope. It follows that either $S(u) < -\infty$ and $\mathcal{E}^c_u(S(u)) = \mathcal{E}_u(S(u)) = D(S(u)) = 0$ or $S(u) = \infty$ and then $\mathcal{E}^c_u \equiv 0$. In the former case we must have that $r_\nu \le S(u) < \infty$. But then $\mathcal{E}^c_u$ is linear on $(R(u), S(u))$, and thus (see Corollary 3.3) $\nu - S^\nu(\mu_u)$ does not charge $[\ell_\nu, r_\nu)$. It follows that we have $(\mu - \mu_u) \le_{cx} (\nu - S^\nu(\mu_u)) \le \nu(\{r_\nu\})$, and then $(\mu - \mu_u)$ must be a point mass at r_ν which is excluded by hypothesis, since μ places no mass there. In the latter case, since $\mathcal{E}^c_u \equiv 0$, Corollary 3.3 implies that $\nu = S^\nu(\mu_u)$. But $\nu(\mathbb{R}) = 1 > u = \mu_u(\mathbb{R}) = S^\nu(\mu_u)(\mathbb{R})$, a contradiction. We conclude that either $-\infty < \ell_\nu \le R(u)$ or $-\infty = \ell_\nu < R(u)$.

Now we deal with $S(u)$. Suppose that either $r_\nu < S(u) \le \infty$ or $r_\nu = S(u) = \infty$. Then, by Theorem 3.2, linearity of $\mathcal{E}^c_u$ on $(R(u), S(u))$ implies that $\nu - S^\nu(\mu_u)$ does not charge $(R(u), r_\nu]$, so that $\mathrm{supp}(\nu - S^\nu(\mu_u)) \subseteq (-\infty, R(u)]$. On the other hand, $\mathrm{supp}(\mu - \mu_u) \subseteq [G(u), \infty)$. But, $S(u) > G(u)$, and therefore, by Lemma 4.1, $R(u) \le Q(u) < G(u)$. This contradicts the fact that $(\mu - \mu_u) \le_{cx} (\nu - S^\nu(\mu_u))$. It follows that either $S(u) \le r_\nu < \infty$ or $S(u) < r_\nu = \infty$. □

Proof of Lemma 4.6. We have $\mathcal{E}_v(z) \le \mathcal{E}_u(z) = \mathcal{E}^c_v(z) \le \mathcal{E}_v(z)$ so that $\mathcal{E}_v(z) = \mathcal{E}_u(z)$.

Consider $\tilde{\mathcal{E}} : \mathbb{R} \to \mathbb{R}$ given by $\tilde{\mathcal{E}} = \mathcal{E}^c_u$ on $(-\infty, z]$ and $\tilde{\mathcal{E}} = \mathcal{E}^c_v$ on (z, ∞). If we can show that $\tilde{\mathcal{E}} \le \mathcal{E}_v$ and $\tilde{\mathcal{E}}$ is convex then $\tilde{\mathcal{E}} \le \mathcal{E}^c_v$ everywhere and therefore $\mathcal{E}^c_u \le \mathcal{E}^c_v$ on $(-\infty, z]$. But $\mathcal{E}^c_v \le \mathcal{E}^c_u$ everywhere, and in particular $\mathcal{E}^c_u = \mathcal{E}^c_v$ on $(-\infty, z]$.

To show that $\tilde{\mathcal{E}} \le \mathcal{E}_v$ note that $\mathcal{E}_v < \mathcal{E}_u$ on $(G(u+), \infty)$ and so we must have $z \le G(u+)$. Then, for $x \in (-\infty, z]$, $\tilde{\mathcal{E}}(x) = \mathcal{E}^c_u(x) \le \mathcal{E}_u(x) = D(x) = \mathcal{E}_v(x)$ whereas for $x > z$, $\tilde{\mathcal{E}}(x) = \mathcal{E}^c_v(x) \le \mathcal{E}_v(x)$.

To show that $\tilde{\mathcal{E}}$ is convex note that if $x < y \le z$ then for $\lambda \in (0, 1)$, by convexity of $\mathcal{E}^c_u$,

$$\tilde{\mathcal{E}}(\lambda x + (1-\lambda)y) = \mathcal{E}^c_u(\lambda x + (1-\lambda)y) \le \lambda \mathcal{E}^c_u(x) + (1-\lambda)\mathcal{E}^c_u(y) = \lambda \tilde{\mathcal{E}}(x) + (1-\lambda)\tilde{\mathcal{E}}(y).$$

We obtain a similar inequality for $z \le x < y$ using the convexity of $\mathcal{E}_v^c$.

So suppose $x < z < y$. Suppose λ is such that $\lambda x + (1-\lambda)y \ge z$. Then

$$\begin{aligned}&\tilde{\mathcal{E}}(\lambda x + (1-\lambda)y) = \mathcal{E}_v^c(\lambda x + (1-\lambda)y) \le \lambda \mathcal{E}_v^c(x) + (1-\lambda)\mathcal{E}_v^c(y)\\ &\le \lambda \mathcal{E}_u^c(x) + (1-\lambda)\mathcal{E}_v^c(y) = \lambda \tilde{\mathcal{E}}(x) + (1-\lambda)\tilde{\mathcal{E}}(y).\end{aligned}$$

Finally suppose that $x < z < y$ and $\lambda x + (1-\lambda)y < z$. For $x < r < k$ consider $f = f(k,r)$ given by

$$f(k,r) = \mathcal{E}_u^c(x) + \frac{r-x}{k-x}(\mathcal{E}_v^c(k) - \mathcal{E}_u^c(x)).$$

Since $\mathcal{E}_u^c(x) \ge \mathcal{E}_v^c(x)$ and $\mathcal{E}_v^c$ is convex it is easily seen that $f(k,r)$ is increasing in k. Then

$$\begin{aligned}\lambda\tilde{\mathcal{E}}(x) + (1-\lambda)\tilde{\mathcal{E}}(y) &= \mathcal{E}_u^c(x) + (1-\lambda)(\mathcal{E}_v^c(y) - \mathcal{E}_u^c(x))\\ &= f(y, \lambda x + (1-\lambda)y)\\ &\ge f(z, \lambda x + (1-\lambda)y)\\ &= \mathcal{E}_u^c(x) + \frac{(1-\lambda)(y-x)}{z-x}(\mathcal{E}_v^c(z) - \mathcal{E}_u^c(x))\\ &= \left(1 - \frac{(1-\lambda)(y-x)}{z-x}\right)\mathcal{E}_u^c(x) + \frac{(1-\lambda)(y-x)}{z-x}\mathcal{E}_u^c(z)\\ &\ge \mathcal{E}_u^c\left(\left(1 - \frac{(1-\lambda)(y-x)}{z-x}\right)x + \frac{(1-\lambda)(y-x)}{z-x}z\right)\\ &= \mathcal{E}_u^c(\lambda x + (1-\lambda y)) = \tilde{\mathcal{E}}(\lambda x + (1-\lambda y)). \quad \square\end{aligned}$$

Proof of Corollary 4.7. Set $s = S(u) = S(v)$. Then $G(u+) \le G(v) \le S(v) = s \le G(u+)$ so that $G(u+) = G(v) = s$ also.

By Lemma 4.6, in order to conclude that $\mathcal{E}_u^c = \mathcal{E}_v^c$ on $(-\infty, S(v)]$ it is sufficient to show that $\mathcal{E}_u^c(s) = \mathcal{E}_v^c(s) = \mathcal{E}_v(s)$. But $\mathcal{E}_u^c(s) = \mathcal{E}_u(s)$ since $s = S(u)$, $\mathcal{E}_u(s) = \mathcal{E}_v(s)$ since $s \le G(u+)$ and $\mathcal{E}_v(s) = \mathcal{E}_v^c(s)$ since $s = S(v)$, and we are done.

Inspection of the proof of Lemma 5.1 shows that provided we replace $D(r)$ with $\mathcal{E}_u(r)$ in the expression for $\phi(u)$ in (5.1), the supremum over $k < G(u)$ can be replaced by a supremum over $k < S(u)$. (The two cases can be considered separately, and if $G(w) = S(w)$ there is nothing to prove.) Then, since $S(u) = s = S(v)$ and $\mathcal{E}_u = \mathcal{E}_v$ on $(-\infty, s]$,

$$\phi(u) = \sup_{r < S(u)} \frac{\mathcal{E}_u(S(u)) - \mathcal{E}_u(r)}{S(u) - r} = \sup_{r < S(v)} \frac{\mathcal{E}_v(S(v)) - \mathcal{E}_v(r)}{S(v) - r} = \phi(v).$$

Further, since $\mathcal{E}_u^c = \mathcal{E}_v^c$ on $(-\infty, S(u) = s = S(v)]$ and $R(\cdot)$ depends on $\mathcal{E}_\cdot$ only through $\mathcal{E}_\cdot^c$ on $(-\infty, S(\cdot)]$, we have $R(u) = R(v)$. $\square$

Proof of Lemma 4.10. Fix $n \ge 1$ and $u, v \in A_<^n$ with $u < v$. We have that $R(u) < G(u) \le G(u+) < S(u-) \le S(u)$. Furthermore, by Theorem 4.9, $R(v) \notin (R(u), S(u))$. We will show that $R(v) \le R(u)$.

Suppose not, so that $S(u) \le R(v)$. We aim to find a contradiction. Define $\bar{u} := \inf\{w \in (u,v] : S(u) \le R(w)\}$ and note that, since $v \in \{w \in (u,v] : S(u) \le R(w)\}$, $\bar{u}$ is well-defined in $[u,v]$.

First note that $\bar{u} > u$. Indeed, since $G(u+) < S(u)$, there exists $\bar{\epsilon} > 0$ such that $G(u+) \le G(u+\epsilon) < S(u)$ for all $0 \le \epsilon \le \bar{\epsilon}$. But then $R(u+\epsilon) \le G(u+\epsilon) < S(u)$ for all $0 \le \epsilon \le \bar{\epsilon}$. It follows that $\bar{u} > u$.

We claim that $R(\cdot)$ is decreasing on $[u, \bar{u})$. Let $u_1, u_2 \in [u, \bar{u})$ with $u_1 < u_2$. For $w \in \{u_1, u_2\}$ $R(w) < S(u)$ and then by the results of the previous paragraph and left-monotonicity, $R(w) \le R(u) < S(u) \le S(u_1) \le S(u_2)$. Since $R(u_2) \notin (R(u_1), S(u_1))$ it follows that $R(u_2) \le R(u_1)$, which proves the claim.

Now observe that $R(\bar{u}) \notin (R(w), S(w))$ for all $w \in [u, \bar{u})$, and therefore, by the monotonicity of $R(\cdot)$ and $S(\cdot)$ on $[u, \bar{u})$, we have that $R(\bar{u}) \notin (R(\bar{u}-), S(\bar{u}-))$. But $\bar{u} \in A^n_<$, so that $G(\bar{u}) \le G(\bar{u}+) < S(\bar{u}-)$. Hence, since $R(\bar{u}) \le G(\bar{u})$, we must have that $R(\bar{u}) \le R(\bar{u}-)$ and therefore $R(\bar{u}) < S(u)$. In particular, $\bar{u} < v$.

Finally, we have that $R(\bar{u}) < G(\bar{u}) \le G(\bar{u}+) < S(\bar{u}-) \le S(\bar{u})$ and we can pick $\tilde{\epsilon} > 0$ such that $G(\bar{u} + \epsilon) < S(\bar{u})$ for all $0 \le \epsilon \le \tilde{\epsilon}$. But then $R(\bar{u} + \epsilon) \le G(\bar{u} + \epsilon) < S(\bar{u})$ and therefore $R(\bar{u} + \epsilon) \le R(\bar{u}) < S(u)$ for all $0 \le \epsilon \le \tilde{\epsilon}$, contradicting the definition of $\bar{u}$ as the infimum of points in $\{w \in (u, v] : S(u) \le R(w)\}$. This is the desired contradiction and we conclude that $R(v) \le R(u)$. □

A.3 Left continuity

Proof of Lemma 6.2. By the left-continuity of G we have

$$\begin{aligned}\lim_{u\uparrow v} \mathcal{E}_u(k) &= \lim_{u\uparrow v}\{P_\nu(k) - P_\mu(k \wedge G(u)) + u(k - G(u))^+\} \\ &= P_\nu(k) - P_\mu(k \wedge G(v)) + v(k - G(v))^+ = \mathcal{E}_v(k).\end{aligned}$$

Also, for all $u \in (0, 1)$ with $v < u$, by (4.5) we have that

$$0 \le \mathcal{E}_v(k) - \mathcal{E}_u(k) \le P_\mu(G(u)) - P_\mu(G(v)) - u(G(u) - G(v)) + (u - v)(k - G(v))^+, \quad k \in \mathbb{R}.$$

But, since μ does not charge $(G(v), G(v+))$ (provided $G(v) < G(v+)$), P_μ is linear on $(G(v), G(v+))$ with slope v, and we have that

$$\lim_{u\downarrow v}\{P_\mu(G(u)) - P_\mu(G(v)) - u(G(u) - G(v))\} = P_\mu(G(v+)) - P_\mu(G(v)) - v(G(v+) - G(v)) = 0.$$

It follows that, for each $k \in \mathbb{R}$, $\mathcal{E}_u(k) \uparrow \mathcal{E}_v(k)$ as $u \downarrow v$.

For convergence of convex hulls note that, for $u \in (0, 1)$ with $u < v$, $\mathcal{E}_u(k) \ge \mathcal{E}_v(k) \ge \mathcal{E}^c_v(k)$ and therefore $\mathcal{E}_u(k) \ge \mathcal{E}^c_u(k) \ge \mathcal{E}^c_v(k)$, $k \in \mathbb{R}$. Then

$$\mathcal{E}_v(k) = \lim_{u\uparrow v} \mathcal{E}_u(k) \ge \lim_{u\uparrow v} \mathcal{E}^c_u(k) \ge \mathcal{E}^c_v(k), \quad k \in \mathbb{R}.$$

Since a point-wise limit of convex functions is convex, $\lim_{u\uparrow v} \mathcal{E}^c_u$ is a convex minorant of $\mathcal{E}_v$. Hence $\mathcal{E}^c_v \ge \lim_{u\uparrow v} \mathcal{E}^c_u$ and equality follows. Similarly, for $u \in (0, 1)$ with $v < u$, we have that $\mathcal{E}_v \ge \mathcal{E}_u \ge \mathcal{E}^c_u$, and therefore $\mathcal{E}^c_v \ge \lim_{u\downarrow v} \mathcal{E}^c_u$.

It remains to show that $\mathcal{E}^c_v \le \lim_{u\downarrow v} \mathcal{E}^c_u$. First note that if $s_n \to s$ and $v_n \downarrow v$ then $\mathcal{E}_{v_n}(s_n) \to \mathcal{E}_v(s)$. To see this note that $\mathcal{E}_{v_n}(s_n) \le \mathcal{E}_v(s_n)$ and so $\lim_n \mathcal{E}_{v_n}(s_n) \le \lim_n \mathcal{E}_v(s_n) = \mathcal{E}_v(s)$ by the continuity of $\mathcal{E}_v$. Conversely, since $|\frac{\mathcal{E}_u(k) - \mathcal{E}_u(j)}{j-k}| \le 1$ for all $j \ne k$ and $u \in (0, 1)$, $\mathcal{E}_{v_n}(s_n) \ge \mathcal{E}_{v_n}(s) - |s_n - s|$ and $\lim_n \mathcal{E}_{v_n}(s_n) \ge \lim_n \mathcal{E}_{v_n}(s) = \mathcal{E}_v(s)$.

Fix k such that $D(k) > 0$. Either there exists a sequence $v_n \downarrow v$ such that $\mathcal{E}^c_{v_n}(k) = \mathcal{E}_{v_n}(k)$ or there exists a sequence $v_n \downarrow v$ such that for each n there exists $r_n = r_n(k)$ and $s_n = s_n(k)$ with $r_n < k < s_n$ with

$$\mathcal{E}^c_{v_n}(k) = \mathcal{E}_{v_n}(r_n)\frac{(s_n - k)}{s_n - r_n} + \mathcal{E}_{v_n}(s_n)\frac{(k - r_n)}{s_n - r_n} = \mathcal{E}_{v_n}(s_n) - \frac{\mathcal{E}_{v_n}(s_n) - \mathcal{E}_{v_n}(r_n)}{s_n - r_n}(s_n - k). \quad \text{(A.2)}$$

In the former case we have

$$\mathcal{E}^c_v(k) \le \mathcal{E}_v(k) = \lim_n \mathcal{E}_{v_n}(k) = \lim_{v_n\downarrow v} \mathcal{E}^c_{v_n}(k).$$

In the latter case we can choose a subsequence such that $s_n \to s \ge k$ and $r_n \to r \in [-\infty, k]$. If $s = k$ then from (A.2) and using $|\frac{\mathcal{E}_{v_n}(s_n) - \mathcal{E}_{v_n}(r_n)}{s_n - r_n}| \le 1$,

$$\lim_n \mathcal{E}^c_{v_n}(k) = \lim_n \mathcal{E}_{v_n}(s_n) = \mathcal{E}_v(k) \ge \mathcal{E}^c_v(k).$$

Otherwise, if $s > k$, then taking limits in (A.2),

$$\lim_n \mathcal{E}^c_{v_n}(k) = \mathcal{E}_v(s) - \frac{\mathcal{E}_v(s) - \mathcal{E}_v(r)}{s-r}(s-k) = \mathcal{E}_v(r)\frac{s-k}{s-r} + \mathcal{E}_v(s)\frac{k-r}{s-r} \geq \mathcal{E}^c_v(k). \qquad \square$$

Proof of Lemma 6.3. Fix $u, v \in (0,1)$ with $u < v$. Since $\mathcal{E}_u \geq \mathcal{E}_v$, $\mathcal{E}_u(f(u)) \leq \mathcal{E}_u(k)$ for all $k \geq f(u)$, and f is non-decreasing, using (6.1) we have that

$$\mathcal{E}_v(f(v-)) = \lim_{u \uparrow v} \mathcal{E}_v(f(u)) \leq \lim_{u \uparrow v} \mathcal{E}_u(f(u)) \leq \lim_{u \uparrow v} \mathcal{E}_u(f(v-)) = \mathcal{E}_v(f(v-)). \tag{A.3}$$

Equation (A.3) still holds when $\mathcal{E}$ is replaced by $\mathcal{E}^c$ (note that $\mathcal{E}^c_u$ is non-decreasing everywhere for each $u \in (0,1)$), which concludes the proof. $\square$

A.4 The candidate coupling is an embedding

Proof of Lemma 7.1. First we argue that (T, S) is left-monotone with respect to G. We need to show that if $v > u$ then $T(v) \notin (T(u), S(u))$.

If $S(u) = G(u)$ then $(T(u), S(u)) = \emptyset$ so there is nothing to prove. So take $T(u) = R(u)$. Suppose first that $G(v) = S(v)$. Then $T(v) = G(v) = S(v) \geq S(u)$. Alternatively suppose $G(v) < S(v)$. Then $T(v) = R(v)$. Finally, properties of R (see Theorem 4.9) imply that either $T(v) = R(v) \geq S(u)$ and $T(v) \geq S(u)$ or $T(v) = R(v) \leq R(u) = T(u)$.

Measurability of T now follows easily. We have $(0,1) = (\cup_{n \geq 1} A^n_<) \cup (u : G(u) < G(u+), G(u) \leq S(u) \leq G(u+)) \cup (u : G(u) = S(u) = G(u+))$. Then $(T(u) \leq t) = (\cup_{n \geq 1} A^n_< \cap \{u : T(u) \leq t\}) \cup \{u : G(u) < G(u+), G(u) \leq S(u) \leq G(u+), T(t) \leq t\} \cup \{u : G(u) = S(u) = G(u+), T(u) \leq t\}$.

But T is decreasing on $A^n_<$ and so $A^n_< \cap \{u : T(u) \leq t\}$ is a Borel subset of $(0,1)$; $\{u : G(u) = S(u) = G(u+), T(u) \leq t\} = \{u : G(u) = S(u) = G(u+), G(u) \leq t\}$ is Borel from the measurability of G and S, and $\{u : G(u) < G(u+), G(u) \leq S(u) \leq G(u+), T(t) \leq t\}$ is countable. Hence T is measurable. $\square$

Proof of Lemma 7.4. From Lemma 5.2(ii) we have $\phi(v+) \geq \phi(v)$. Further, for $v \leq w \leq u \leq 1$ we have $\phi(w) - \phi(u) \leq u - w$. Introduce $\tilde{\phi} : [v,1] \to [0,1]$ by $\tilde{\phi}(v) = x$ and $\tilde{\phi}(w) = \phi(w)$ for $w \in (v,1]$. Then, for $v < u \leq 1$, $\tilde{\phi}(v) - \tilde{\phi}(u) \leq \lim_{n \uparrow \infty} \phi(v + 1/n) - \phi(u) \leq \lim_{n \uparrow \infty} \{u - (v + 1/n)\} = u - v$. By checking the easy cases $w = v = u$ and $v < w \leq u \leq 1$ separately we conclude $\tilde{\phi}(w) - \tilde{\phi}(u) \leq u - w$ for all $v \leq w \leq u \leq 1$.

By construction, $\psi_{v,x}(w) = \inf_{v \leq u \leq w} \{\tilde{\phi}(u)\}$. Then $\psi_{v,x}$ is decreasing, $\psi_{v,x}(v) = \tilde{\phi}(v) = x$ and $\psi_{v,x} \leq \tilde{\phi}$ on $[v,1]$. Fix w and u with $v \leq w < u$. Let $(u_m)_{m \geq 1}$ with $v \leq u_m \leq u$ be such that $\lim_{m \uparrow \infty} \tilde{\phi}(u_m) \downarrow \psi_{v,x}(u)$. Taking a convergent subsequence if necessary we may assume $u_m \to \tilde{u} \in [v,u]$. Then, if $w < \tilde{u}$,

$$\begin{aligned} 0 \leq \psi_{v,x}(w) - \psi_{v,x}(u) \leq \tilde{\phi}(w) - \lim_m \tilde{\phi}(u_m) &= \lim_m \left\{ \tilde{\phi}(w) - \tilde{\phi}(u_m) \right\} \\ &\leq \lim_m \{u_m - w\} = \tilde{u} - w \leq u - w. \end{aligned}$$

On the other hand, if $\tilde{u} < w$, $\psi_{v,x}(w) \leq \lim_m \tilde{\phi}(u_m) = \psi_{v,x}(u)$ so that $\psi_{v,x}(w) = \psi_{v,x}(u)$. Finally, if $\tilde{u} = w$ then either there exists a sequence $u_m \to \tilde{u}$ with $\lim_m \tilde{\phi}(u_m) = \psi_{v,x}(u)$ and $u_m \geq \tilde{u}$, or there exists a sequence $u_m \to \tilde{u}$ with $\lim_m \tilde{\phi}(u_m) = \psi_{v,x}(u)$ and $u_m \leq \tilde{u}$ (or both). In either case the corresponding proof shows that $\psi_{v,x}(w) = \psi_{v,x}(u)$. It follows that $0 \leq \psi_{v,x}(w) - \psi_{v,x}(u) \leq u - w$ so that $\psi_{v,x}$ is absolutely continuous on $[v,1]$ (and not just on $[v,1] \cap A_<$) and has a derivative ψ' such that $\int_a^b \psi'_{v,x}(u)du = \psi_{v,x}(b) - \psi_{v,x}(a)$, for all $v \leq a \leq b \leq 1$. $\square$

A.5 Proof of Proposition 7.6

Our goal is to show that $R(\cdot) \le y$ is equivalent to $\psi(\cdot) = \phi(\cdot)$ for a well-chosen element $\psi = \psi_{v,x}$. In particular, we want to show

$$\{v : v > S^{-1}(y), R(v) \le y\} = \{v : v > S^{-1}(y), \psi_{S^{-1}(y),(\mathcal{E}^c_{S^{-1}(y)})'(y-)}(v) = \phi(v)\} \tag{A.4}$$

in a sufficiently rich set of circumstances, with the ultimate aim of proving Proposition 7.6. We begin with a partial result, valid in the case where $w = S^{-1}(y)$ is such that $R(w) < S(w)$. Later the issue will be to show that (A.4) holds also for both y such that $R(S^{-1}(y)) = S(S^{-1}(y))$ and for y which are not of the form $y = S(w)$ for some w.

Lemma A.1. *Suppose w is such that $R(w) < S(w)$. Then*

$$\{v : v > w, R(v) \le R(w)\} = \{v : v > w, \psi_{w,\phi(w)}(v) = \phi(v)\}.$$

Proof. Let $I^R_w = \{v : v > w, R(v) \le R(w)\}$ and $I^\phi_w = \{v : v > w, \psi_{w,\phi(w)}(v) = \phi(v)\}$.

First we argue that, for all $u, v \in (0,1)$ with $u < v$, $\phi(u) < \phi(v)$ implies $R(v) \ge S(u)$. Suppose $u < v$ and $\phi(u) < \phi(v)$. Suppose $(S(v), \mathcal{E}_v(S(v)))$ lies on or below the line $L^{\mathcal{E}_u,\phi(u)}_{S(u)}$. Then $L^{\mathcal{E}_v,\phi(v)}_{S(v)} < L^{\mathcal{E}_u,\phi(u)}_{S(u)} \le \mathcal{E}_u$ on $(-\infty, S(v)) \supseteq (-\infty, S(u))$ and $L^{\mathcal{E}_v,\phi(v)}_{S(v)} < \mathcal{E}_u = D$ on $(-\infty, G(u+))$. Hence $R(v) \ge G(u+)$ and by left-monotonicity $R(v) \ge S(u)$. Conversely, if $(S(v), \mathcal{E}_v(S(v))$ lies above $L^{\mathcal{E}_u,\phi(u)}_{S(u)}(S(v))$ then $L^{\mathcal{E}_v,\phi(v)}_{S(v)}(S(u)) \le \mathcal{E}_v(S(u)) \le \mathcal{E}_u(S(u)) = L^{\mathcal{E}_u,\phi(u)}_{S(u)}(S(u))$. Then $L^{\mathcal{E}_v,\phi(v)}_{S(v)}(k) < L^{\mathcal{E}_u,\phi(u)}_{S(u)}(k) \le \mathcal{E}_u(k) = \mathcal{E}_v(k)$ for $k < G(u) \le G(v)$ and since $L^{\mathcal{E}_v,\phi(v)}_{S(v)}(R(v)) = \mathcal{E}_v(R(v))$ we must have $R(v) \ge G(u)$. If $R(u) < S(u)$ then $R(v) \ge G(u) > R(u)$ and then by left-monotonicity $R(v) \ge S(u)$. Similarly, if $R(u) = S(u)$ then $R(v) \ge G(u) = R(u) = S(u)$ and again $R(v) \ge S(u)$.

Now we show that $I^R_w \subseteq I^\phi_w$. Suppose $v > w$ but $v \notin I^\phi_w$. Then there exists $u \in [w, v)$ such that $\phi(u) < \phi(v)$. Then by the above argument, $R(v) \ge S(u) \ge S(w) > R(w)$ so that $v \notin I^R_w$. Hence $I^R_w \subseteq I^\phi_w$.

For the converse we show that if $v > w$ but $v \notin I^R_w$ (so that $R(v) > R(w)$) then there exists $u \in [w, v)$ such that $\phi(u) < \phi(v)$ and hence $v \notin I^\phi_w$.

So, suppose $R(v) > R(w)$. By left-monotonicity we deduce that $R(v) \ge S(w)$. Note also that since $R(w) < G(w+)$ and $w < v$, $\mathcal{E}_v(R(w)) = \mathcal{E}_w(R(w)) = D(R(w))$.

If $\mathcal{E}_v(S(v)) > L^{\mathcal{E}_w,\phi(w)}_{S(w)}(S(v))$ then since $R(w) < R(v) \le G(v)$, by Lemma 5.1 we have

$$\phi(v) \ge \frac{\mathcal{E}_v(S(v)) - \mathcal{E}_v(R(w))}{S(v) - R(w)} > \frac{L^{\mathcal{E}_w,\phi(w)}_{S(w)}(S(v)) - L^{\mathcal{E}_w,\phi(w)}_{S(w)}(R(w))}{S(v) - R(w)} = \phi(w).$$

It follows that $\phi(v) > \phi(w)$.

If $\mathcal{E}_v(S(v)) = L^{\mathcal{E}_w,\phi(w)}_{S(w)}(S(v))$ then, since $L^{\mathcal{E}_v,\phi(v)}_{S(v)} \le \mathcal{E}^c_v \le \mathcal{E}^c_w$ everywhere, we must have that $\phi(v) \ge \phi(w)$. But if $\phi(w) = \phi(v)$ then $L^{\mathcal{E}_v,\phi(v)}_{S(v)}$ and $L^{\mathcal{E}_w,\phi(w)}_{S(w)}$ coincide and hence $R(v) = R(w)$, a contradiction to our hypothesis that $R(v) > R(w)$. It follows again that $\phi(v) > \phi(w)$.

Finally consider the case where $\mathcal{E}_v(S(v)) = L^{\mathcal{E}_v,\phi(v)}_{S(v)}(S(v)) < L^{\mathcal{E}_w,\phi(w)}_{S(w)}(S(v))$. If $\mathcal{E}_v(R(v)) = L^{\mathcal{E}_v,\phi(v)}_{S(v)}(R(v)) \ge L^{\mathcal{E}_w,\phi(w)}_{S(w)}(R(v))$ then $R(w) < R(v) < S(v)$ and we must have $L^{\mathcal{E}_v,\phi(v)}_{S(v)}(R(w)) > L^{\mathcal{E}_w,\phi(w)}_{S(w)}(R(w))$. Then $\mathcal{E}_w(R(w)) = \mathcal{E}_v(R(w)) \ge L^{\mathcal{E}_v,\phi(v)}_{S(v)}(R(w)) > L^{\mathcal{E}_w,\phi(w)}_{S(w)}(R(w)) = \mathcal{E}_w(R(w))$, a contradiction. Hence we must have

$$\mathcal{E}_v(R(v)) < L^{\mathcal{E}_w,\phi(w)}_{S(w)}(R(v)).$$

Consider $\mathcal{E}^c_v$. There exists an interval $[x_1, x_2]$ with $x_1 \le R(w) < G(w+) < x_2$ such that $\mathcal{E}^c_v$ is linear on $[x_1, x_2]$, $\mathcal{E}^c_v(x_1) = \mathcal{E}_v(x_1) = D(x_1)$ and $\mathcal{E}^c_v(x_2) = \mathcal{E}_v(x_2)$. Necessarily

$\mathcal{E}_v(x_1) \geq L_{S(w)}^{\mathcal{E}_w,\phi(w)}(x_1)$ and $\mathcal{E}_v(x_2) < L_{S(w)}^{\mathcal{E}_w,\phi(w)}(x_2)$. There exists u such that $G(u) \leq x_2 \leq G(u+)$. Note that since $x_2 > G(w+)$ we have $u > w$. Also $x_2 \leq R(v) \leq G(v)$, else $R(v) \leq x_1 \leq R(w)$ which contradicts our assumption that $R(v) > R(w)$ and hence we may take $u \leq v$. Finally, we must have $u < v$ since if $u = v$ we find $R(v) = R(u) = x_1 \leq R(w)$, again a contradiction. In summary, $u \in (w, v)$.

Since $x_2 \leq G(u+) \leq G(v) \leq S(v)$ by the convexity of $\mathcal{E}_v^c$ we have $(\mathcal{E}_v^c)'(x_2-) \leq (\mathcal{E}_v^c)'(S(v)-)$. If there is equality here then $\mathcal{E}_v^c$ is linear on $(x_1, S(v))$ and $R(v) \leq x_1 \leq R(w)$, a contradiction. Hence there is strict inequality and $\phi(u) = (\mathcal{E}_v^c)'(x_2-) < (\mathcal{E}_v^c)'(S(v)-) = \phi(v)$. Hence $\psi_{w,\phi(w)}(v) \leq \phi(u) < \phi(v)$ and $v \notin I_w^\phi$. □

Now we introduce a condition which will allow us to prove an analogue of (A.4) in a wider set of circumstances.

Definition A.2. *Fix $x \in \mathbb{R}$ and $v \in (S^{-1}(x), 1)$. Define $\mathcal{A}(v, x) = \{w : S^{-1}(x) < w \leq v, R(w) < x\}$. Then Condition $A(x)$ is that $\mathcal{A}(v, x)$ is non-empty for each $v \in (S^{-1}(x), 1)$.*

Lemma A.3. *Suppose $D(y) < \mathcal{E}_{S^{-1}(y)}(y)$. Then Condition $A(y)$ holds.*

Proof. It is clear that $D(y) < \mathcal{E}_{S^{-1}(y)}(y)$ if and only if $G(S^{-1}(y)+) < y$.

Choose $h_0 > 0$ such that $G(S^{-1}(y) + h_0) < y$. Then for all $0 < h \leq h_0$ we have $R(S^{-1}(y) + h) \leq G(S^{-1}(y) + h) \leq G(S^{-1}(y) + h_0) < y$. □

Lemma A.4. *Suppose y is a continuity point for both μ and ν. Suppose further that $D(y) = \mathcal{E}_{S^{-1}(y)}(y)$ and for all $v \in (S^{-1}(y), 1)$, there exists k such that*

$$\mathcal{E}_v(k) < L_y^{\mathcal{E}_{S^{-1}(y)},(\mathcal{E}^c_{S^{-1}(y)})'(y-)}(k).$$

Then, Condition $A(y)$ holds.

Proof. Write L as shorthand for $L_y^{\mathcal{E}_{S^{-1}(y)},(\mathcal{E}^c_{S^{-1}(y)})'(y-)}$. Fix $h_0 > 0$ and choose $k > y$ such that $\mathcal{E}_{S^{-1}(y)+h_0}(k) < L(k)$. Such a k exists under the assumptions of the lemma. See Figure 6.

Since y is a continuity point of both μ and ν we have $(\mathcal{E}^c_{S^{-1}(y)})'(y-) = (\mathcal{E}^c_{S^{-1}(y)})'(y+)$ and $(\mathcal{E}^c_{S^{-1}(y)})'(y) = \mathcal{E}'_{S^{-1}(y)}(y) = D'(y)$. For $j \in (y, k]$ define $f(j) = \frac{\mathcal{E}_{S^{-1}(y)+h_0}(j) - D(y)}{j-y}$. Then f is continuous on $(y, k]$ with $f(y+) = \mathcal{E}'_{S^{-1}(y)}(y)$ and $f(k) < \frac{L(k)-D(y)}{k-y} = (\mathcal{E}^c_{S^{-1}(y)})'(y)$. In particular f attains its minimum value. Let $\underline{f}$ be this minimal value and let $\underline{j}$ be the smallest value at which it is attained. Then $\underline{j} > y$.

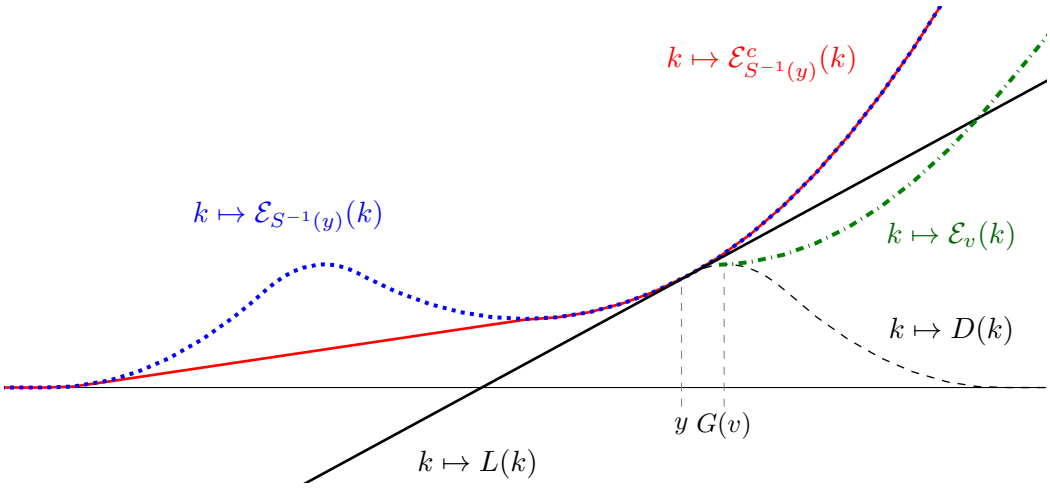

Figure 6: Plot of $\mathcal{E}_{S^{-1}(y)}$ (dotted curve), $\mathcal{E}^c_{S^{-1}(y)}$ (solid curve below $\mathcal{E}_{S^{-1}(y)}$), D (dashed curve), $\mathcal{E}_v$ (dash-dotted curve) and $L \equiv L_y^{\mathcal{E}_{S^{-1}(y)},(\mathcal{E}^c_{S^{-1}(y)})'(y-)}$ (line tangent to $\mathcal{E}^c_{S^{-1}(y)}$ at y) under the assumptions of Lemma A.4. In the figure, $R(S^{-1}(y)) = G(S^{-1}(y)) = S(S^{-1}(y)) = y$ and, for all $v \in (S^{-1}(y), 1)$, $\mathcal{E}_v(k) < L(k)$ for some $k > y$. Furthermore, $\mathcal{E}'_{S^{-1}(y)}(y)$ and $(\mathcal{E}^c_{S^{-1}(y)})'(y)$ both exist, and $\phi(S^{-1}(y)) = \mathcal{E}'_{S^{-1}(y)}(y) = (\mathcal{E}^c_{S^{-1}(y)})'(y)$ is the slope of L.

Suppose $\underline{j} \geq G(S^{-1}(y)+h_0)$. Then the line $\hat{L} = L_{y,\underline{j}}^{\mathcal{E}_{S^{-1}(y)+h_0}}$ joining $(y, \mathcal{E}_{S^{-1}(y)+h_0}(y))$ to $(\underline{j}, \mathcal{E}_{S^{-1}(y)+h_0}(\underline{j}))$ lies below L on $(y, \underline{j})$, but strictly above L on an interval to the left of y. If follows that we can find small enough $\epsilon > 0$ and $\check{y} < y$ such that $L_{\underline{j}}^{\mathcal{E}_{S^{-1}(y)+h_0},\underline{f}+\epsilon}(\check{y}) = D(\check{y})$ and $\mathcal{E}_{S^{-1}(y)+h_0} > \hat{L}$ on $(\check{y}, \underline{j})$. Then $R(S^{-1}(y)+h_0) \leq Q(S^{-1}(y)+h_0) \leq \check{y} < y < \underline{j} \leq S(S^{-1}(y)+h_0)$. In particular, $R(S^{-1}(y)+h_0) < y$.

Now suppose $\underline{j} < G(S^{-1}(y)+h_0)$. Then there exists $h \in (0, h_0]$ such that $G(S^{-1}(y)+h) \leq \underline{j} \leq G((S^{-1}(y)+h)+)$. It follows that $\mathcal{E}_{S^{-1}(y)+h}(\underline{j}) = D(\underline{j})$. Then, exactly as before we can find $\epsilon > 0$ and $\check{y} < y$ such that $L_{\underline{j}}^{\mathcal{E}_{S^{-1}(y)+h},\underline{f}+\epsilon}(\check{y}) = D(\check{y})$ and $D \equiv \mathcal{E}_{S^{-1}(y)+h} > L_{\underline{j}}^{\mathcal{E}_{S^{-1}(y)+h},\underline{f}+\epsilon}$ on $(\check{y}, \underline{j})$. As before it follows that $R(S^{-1}(y)+h) < y$. □

Lemma A.5. *Suppose $x \in \mathbb{R}$ is such that Condition $A(x)$ holds. Then*

$$\{v : v > S^{-1}(x), R(v) \leq x\} = \{v : v > S^{-1}(x), \psi_{S^{-1}(x),(\mathcal{E}^c_{S^{-1}(x)})'(x-)}(v) = \phi(v)\}. \tag{A.5}$$

Remark A.6. It will follow from the proof of Lemma A.5 that if x is such that Condition $A(x)$ holds then also

$$\{v : v > S^{-1}(x), R(v) < x\} = \{v : v > S^{-1}(x), \psi_{S^{-1}(x),(\mathcal{E}^c_{S^{-1}(x)})'(x-)}(v) = \phi(v)\}. \tag{A.6}$$

The equivalence of the left-hand-sides of (A.5) and (A.6) can also be seen directly. It is sufficient to argue that if Condition $A(x)$ holds then $\{v : v > S^{-1}(x), R(v) = x\} = \emptyset$. To see this, given $v > S^{-1}(x)$ choose $w \in (S^{-1}(x), v)$. Then, since $\mathcal{A}(w, x)$ is non-empty there exists $u \in (S^{-1}(x), w]$ such that $R(u) < x < S(u)$. Then $R(v) \notin (R(u), S(u))$ and in particular $R(v) \neq x$.

Proof of Lemma A.5. We begin with a definition which will be useful in both the forward and reverse implication of (A.5).

Fix $v > S^{-1}(x)$. Define $\underline{z} = \sup\{w : w \in \mathcal{A}(v,x)\}$. Note that by assumption $\mathcal{A}(v,x)$ is non-empty and $S^{-1}(x) < \underline{z} \leq v$. We show that $\underline{z} \in \mathcal{A}(v,x)$. If this is not immediately the case then since $\underline{z} > S^{-1}(x)$ there exists $(z_m)_{m\geq 1}$ with $z_m \uparrow \underline{z}$ and $R(z_m) < x$. Then, by Proposition 6.1 we have that $R(\underline{z}) \leq \liminf_{u\uparrow\underline{z}} R(u) \leq x$. Moreover, $R(\underline{z}) \notin (R(z_m), S(z_m)) \supseteq \{x\}$. Hence $R(\underline{z}) < x$ and $\underline{z} \in \mathcal{A}(v,x)$. We also have $R(\underline{z}) < S(\underline{z})$.

Let $H_{S^{-1}(x),z} = \{w \in (S^{-1}(x), z] : \phi(w) = \psi_{S^{-1}(x),\phi^*}(w)\}$ where ϕ^* is shorthand for $(\mathcal{E}^c_{S^{-1}(x)})'(x-)$. We claim that $\phi(\underline{z}) \leq (\mathcal{E}^c_{S^{-1}(x)})'(x-) = \phi^*$ and $\underline{z} \in H_{S^{-1}(x),v}$.

For the first of these claims, suppose to the contrary that $\phi(\underline{z}) > \phi^*$. Then, for $k < x$, $\mathcal{E}_{\underline{z}}(x) + \phi(\underline{z})(k-x) < \mathcal{E}_{S^{-1}(x)}(x) + \phi^*(k-x) \leq D(k)$, since $\phi^* \in \partial\mathcal{E}^c_{S^{-1}(x)}(x)$. Since also $\mathcal{E}_{\underline{z}}(S(\underline{z})) + \phi(\underline{z})(x - S(\underline{z})) \leq \mathcal{E}_{\underline{z}}(x)$, $\mathcal{E}_{\underline{z}}(S(\underline{z})) + \phi(\underline{z})(k - S(\underline{z})) \leq \mathcal{E}_{\underline{z}}(x) + \phi(\underline{z})(k-x) < D(k)$ for all $k \in (-\infty, x)$. In particular, $R(\underline{z}) \geq x$. But we saw above that $R(\underline{z}) < x$, and this is our contradiction.

Now we show $\underline{z} \in H_{S^{-1}(x),v}$. Suppose not, i.e. suppose $\psi_{S^{-1}(x),\phi^*}(\underline{z}) < \phi(\underline{z})$. Let $(z_m)_{m\geq 1}$ be such that $\phi(z_m) \downarrow \psi_{S^{-1}(x),\phi^*}(\underline{z}) < \phi(\underline{z})$. Pick m such that $S^{-1}(y) < z_m < \underline{z}$ with $\phi(z_m) < \phi(\underline{z})$. Then by the left-monotonicity of R (and S) we have that $R(\underline{z}) \notin (R(z_m), S(z_m))$. Since $S(z_m) > x$ and $R(\underline{z}) < x$ we have $R(\underline{z}) \leq R(z_m) \leq S(z_m) \leq S(\underline{z})$. If $R(z_m) = R(\underline{z}) < x < S(z_m)$, then $\mathcal{E}^c_{z_m}(S(z_m)) = \mathcal{E}_{z_m}(S(z_m)) = \mathcal{E}_{z_m}(R(z_m)) + \phi(z_m)(S(z_m) - R(z_m)) < \mathcal{E}_{\underline{z}}(R(\underline{z})) + \phi(\underline{z})(S(z_m) - R(\underline{z})) = L^{\mathcal{E}_{\underline{z}}}_{R(\underline{z}),S(\underline{z})}(S(z_m)) = \mathcal{E}^c_{\underline{z}}(S(z_m))$, a contradiction to $\mathcal{E}^c_{z_m} \geq \mathcal{E}^c_{\underline{z}}$. On the other hand, if $R(\underline{z}) < R(z_m)$, by convexity of $\mathcal{E}^c_{z_m}$ we must have that $D(R(\underline{z})) = \mathcal{E}_{z_m}(R(\underline{z})) \geq \mathcal{E}_{z_m}(S(z_m)) - \phi(z_m)(S(z_m) - R(\underline{z})) > \mathcal{E}^c_{\underline{z}}(S(z_m)) - \phi(\underline{z})(S(z_m) - R(\underline{z})) = L^{\mathcal{E}_{\underline{z}}}_{R(\underline{z}),S(\underline{z})}(R(\underline{z})) = D(R(\underline{z}))$, again a contradiction. We conclude that we cannot have $\phi(\underline{z}) > \psi_{S^{-1}(x),\phi^*}(\underline{z})$ and hence $\underline{z} \in H_{S^{-1}(x),v}$ as claimed.

Forward implication: First we show that $\{v : v > S^{-1}(x), \psi_{S^{-1}(x),\phi^*}(v) = \phi(v)\} \subseteq \{v : v > S^{-1}(x), R(v) \leq x\}$. In fact we show the apparently stronger result that $\{v : v > S^{-1}(x), \psi_{S^{-1}(x),\phi^*}(v) = \phi(v)\} \subseteq \{v : v > S^{-1}(x), R(v) < x\}$, although by Remark A.6 these implications are equivalent.

We wish to show that if $v > S^{-1}(x)$ and $R(v) \geq x$ then $\phi(v) > \psi_{S^{-1}(x),\phi^*}(v)$.

Suppose $R(v) \geq x$. We argue that $\phi(\underline{z}) < \phi(v)$ and therefore that $\psi_{S^{-1}(x),\phi^*}(v) < \phi(v)$.

Temporarily let $L = L^{\mathcal{E}_{\underline{z}}}_{R(\underline{z}),S(\underline{z})}$.

Since $R(v) \geq x$ we have that $\underline{z} < v$. First, we claim that $S(\underline{z}) \leq G(\underline{z}+)$ (and then also $\mathcal{E}_{\underline{z}} = D$ on $[G(\underline{z}), G(\underline{z}+)] \supseteq [G(\underline{z}), S(\underline{z})]$, since P_μ is linear on $(G(\underline{z}), G(\underline{z}+))$ as μ does not charge $(G(\underline{z}), G(\underline{z}+))$). Suppose not: then $S(\underline{z}) > G(\underline{z}+)$ and there exists $h \in (0, v - \underline{z})$ such that $G(\underline{z}+h) < S(\underline{z})$. But $R(\underline{z}+h) \leq G(\underline{z}+h)$ and, by the left-monotonicity result (Theorem 4.9), $R(\underline{z}+h) \notin (R(\underline{z}), S(\underline{z}))$. It follows that $R(\underline{z}+h) \leq R(\underline{z}) < x$ contradicting the supposed maximality of $\underline{z}$. We conclude that $S(\underline{z}) \leq G(\underline{z}+)$.

Second we claim that, for a sufficiently small $h > 0$, $\mathcal{E}_{\underline{z}+h} \geq L$. Suppose to the contrary that there exists $0 < h < v - \underline{z}$ and $\epsilon > 0$ such that $L^{\mathcal{E}_{\underline{z}},\phi(\underline{z})-\epsilon}_{R(\underline{z})}$ intersects $\mathcal{E}_{\underline{z}+h}$ to the right of $S(\underline{z})$. Then, letting $\tilde{k} = \tilde{k}(\underline{z}, h)$ be the x-coordinate of the smallest intersection point to the right of $S(\underline{z})$ we have $\mathcal{E}_{\underline{z}+h} > L^{\mathcal{E}_{\underline{z}+h}}_{R(\underline{z}),\tilde{k}}$ on $(R(\underline{z}), \tilde{k})$. If $G(\underline{z}+h) \leq \tilde{k}$ then $Q(\underline{z}+h) \leq R(\underline{z})$ and so $R(\underline{z}+h) \leq Q(\underline{z}+h) \leq R(\underline{z}) < x$. This contradicts the maximality of $\underline{z}$. Otherwise, if $G(\underline{z}+h) > \tilde{k}$ then by an argument very similar to the last part of the proof of Lemma A.4 we can again conclude that there exists $0 < h_1 \leq h$ for which $R(\underline{z}+h_1) < x$. Again this contradicts the maximality of $\underline{z}$.

Define $\overline{z} = \sup\{z : \mathcal{E}_z \geq L \text{ on } \mathbb{R}\}$. Then since $\lim_{z\uparrow\overline{z}} \mathcal{E}_z(k) = \mathcal{E}_{\overline{z}}(k)$ it follows that $\mathcal{E}_{\overline{z}} \geq L$ and there exists $k > S(\underline{z})$ such that $\mathcal{E}_{\overline{z}}(k) = L(k)$, and then $R(\overline{z}) = R(\underline{z})$ and $\phi(\underline{z}) = \phi(\overline{z})$. Then $R(\overline{z}) < x$ and, by the maximality of $\underline{z}$, we have that $v < \overline{z}$ (note that $v = \overline{z}$ is excluded since $R(v) \geq x > R(\underline{z})$). Since $v < \overline{z}$, $\mathcal{E}_v \geq L$ on $(S(\underline{z}), \infty)$ and $\mathcal{E}_v(S(v)) > L(S(v))$ (note

that if $\mathcal{E}_v(S(v)) = L(S(v))$ then $R(v) \le R(\underline{z}) < x$, a contradiction). Hence

$$\phi(v) \ge \frac{\mathcal{E}_v(S(v)) - D(R(\underline{z}))}{S(v) - R(\underline{z})} > \frac{L(S(v)) - D(R(\underline{z}))}{S(v) - R(\underline{z})} = \phi(\underline{z}) \ge \psi_{S^{-1}(x),\phi^*}(v).$$

Reverse implication:
Now we show that $\{v : v > S^{-1}(x), R(v) \le x\} \subseteq \{v : v > S^{-1}(x), \psi_{S^{-1}(x),\phi^*}(v) = \phi(v)\}$. We suppose that $v > S^{-1}(x)$ and $\psi_{S^{-1}(x),\phi^*}(v) < \phi(v)$ and show that $R(v) > x$.

From the opening comments of the proof of the lemma we have that $S^{-1}(x) < \underline{z} \le v$, $R(\underline{z}) < x$ and $\underline{z} \in H_{S^{-1}(x),v}$. Then the assumption that $\psi_{S^{-1}(x),\phi^*}(v) < \phi(v)$ implies that $\underline{z} < v$ so that $\underline{z} \in (S^{-1}(x), v)$.

Let $\underline{\underline{z}} = \sup\{w : w \in H_{S^{-1}(x),v}\}$. We will see that $\underline{\underline{z}}$ (respectively $\overline{\overline{z}}$ introduced below) plays a very similar role to $\underline{z}$ (respectively $\overline{z}$) from the forward implication.

We have $\underline{\underline{z}} \ge \underline{z} > S^{-1}(x)$. If $\underline{\underline{z}} = v$ then there exists a sequence $(z_n)_{n\ge1}$ with $z_n \uparrow v$ and $\phi(z_n) = \psi_{S^{-1}(y),\phi^*}(z_n)$. Then, by Proposition 6.1 and the left-continuity of ϕ and $\psi_{S^{-1}(x),\phi^*}$, $\phi(v) = \liminf \phi(z_n) = \liminf \psi_{S^{-1}(x),\phi^*}(z_n) = \psi_{S^{-1}(x),\phi^*}(v)$ which is a contradiction. Hence we may conclude $S^{-1}(x) < \underline{z} \le \underline{\underline{z}} < v$ and $\psi_{S^{-1}(x),\phi^*}(v) = \phi(\underline{\underline{z}}) < \phi(v)$.

First we show that $S(\underline{\underline{z}}) \le G(\underline{\underline{z}}+)$. Suppose to the contrary and take $0 < h < v - \underline{\underline{z}}$ such that $G(\underline{\underline{z}} + h) < S(\underline{\underline{z}})$. We claim that $\phi(\underline{\underline{z}} + h) < \phi(\underline{\underline{z}})$, contradicting the maximality of $\underline{\underline{z}}$.

We have $R(\underline{\underline{z}}+h) \le R(\underline{\underline{z}}) < G(\underline{\underline{z}}) \le G(\underline{\underline{z}}+h) < S(\underline{\underline{z}}) \le S(\underline{\underline{z}}+h)$, so that $\mathcal{E}^c_w = L^{\mathcal{E}_w}_{R(w),S(w)}$ on $[R(w), S(w)]$ for $w \in \{\underline{\underline{z}}, \underline{\underline{z}} + h\}$. Since $G(\underline{\underline{z}} + h) < S(\underline{\underline{z}})$, $\mathcal{E}_{\underline{\underline{z}}}(S(\underline{\underline{z}})) > \mathcal{E}_{\underline{\underline{z}}+h}(S(\underline{\underline{z}})) \ge \mathcal{E}^c_{\underline{\underline{z}}+h}(S(\underline{\underline{z}}))$, and therefore $L^{\mathcal{E}_{\underline{\underline{z}}}}_{R(\underline{\underline{z}}),S(\underline{\underline{z}})} = L^{\mathcal{E}_{\underline{\underline{z}}},\phi(\underline{\underline{z}})}_{S(\underline{\underline{z}})} > L^{\mathcal{E}^c_{\underline{\underline{z}}+h},\phi(\underline{\underline{z}})}_{S(\underline{\underline{z}})}$ everywhere. Then for $\psi \ge \phi(\underline{\underline{z}})$, $D \ge L^{\mathcal{E}_{\underline{\underline{z}}}}_{R(\underline{\underline{z}}),S(\underline{\underline{z}})} > L^{\mathcal{E}^c_{\underline{\underline{z}}+h},\psi}_{S(\underline{\underline{z}})}$ on $(-\infty, G(\underline{\underline{z}})] \supset (-\infty, R(\underline{\underline{z}} + h)]$. But $D(R(\underline{\underline{z}} + h)) = L^{\mathcal{E}_{\underline{\underline{z}}+h}}_{R(\underline{\underline{z}}+h),S(\underline{\underline{z}}+h)}(R(\underline{\underline{z}} + h)) = L^{\mathcal{E}^c_{\underline{\underline{z}}+h},\phi(\underline{\underline{z}}+h)}_{S(\underline{\underline{z}})}(R(\underline{\underline{z}} + h))$ and therefore we conclude that $\phi(\underline{\underline{z}} + h) < \phi(\underline{\underline{z}})$. It follows that $S(\underline{\underline{z}}) \le G(\underline{\underline{z}}+)$ as claimed.

Second we show that $\mathcal{E}_{\underline{\underline{z}}+h} \ge L^{\mathcal{E}_{\underline{\underline{z}}},\phi(\underline{\underline{z}})}_{R(\underline{\underline{z}})}$ for a sufficiently small $h > 0$. If this is not the case then for any $\epsilon > 0$ there exists some $h \in (0, \epsilon \wedge (v - \underline{\underline{z}}))$ we have that $\delta = \inf_{k > S(\underline{\underline{z}})}\{\mathcal{E}_{\underline{\underline{z}}+h}(k) - L^{\mathcal{E}_{\underline{\underline{z}}},\phi(\underline{\underline{z}})}_{R(\underline{\underline{z}})}(k)\} < 0$. Let L^δ be given by $L^\delta(k) = L^{\mathcal{E}_{\underline{\underline{z}}},\phi(\underline{\underline{z}})}_{R(\underline{\underline{z}})}(k) + \delta$ and let $\underline{\underline{s}} > S(\underline{\underline{z}})$ be such that $L^\delta(\underline{\underline{s}}) = \mathcal{E}_{\underline{\underline{z}}+h}(\underline{\underline{s}})$. We have $\mathcal{E}_{\underline{\underline{z}}+h} \ge \mathcal{E}^c_{\underline{\underline{z}}+h} \ge L^\delta$ and $\mathcal{E}_{\underline{\underline{z}}+h}(\underline{\underline{s}}) = \mathcal{E}^c_{\underline{\underline{z}}+h}(\underline{\underline{s}}) = L^\delta(\underline{\underline{s}})$. It follows that $(\mathcal{E}^c_{\underline{\underline{z}}+h})'(\underline{\underline{s}}-) \le \phi(\underline{\underline{z}})$.

Suppose $G(\underline{\underline{z}} + h) \le \underline{\underline{s}}$. Then $S(\underline{\underline{z}} + h) \le \underline{\underline{s}}$ and

$$\phi(\underline{\underline{z}} + h) = (\mathcal{E}^c_{\underline{\underline{z}}+h})'(S(\underline{\underline{z}} + h)-) \le (\mathcal{E}^c_{\underline{\underline{z}}+h})'(\underline{\underline{s}}-) \le \phi(\underline{\underline{z}}).$$

Then $\underline{\underline{z}}$ is not maximal in $H_{S^{-1}(x),v}$, a contradiction. Now suppose $G(\underline{\underline{z}} + h) > \underline{\underline{s}}$. Then there exists $\underline{\underline{u}} \in (\underline{\underline{z}}, \underline{\underline{z}} + h)$ with $G(\underline{\underline{u}}) \le \underline{\underline{s}} \le G(\underline{\underline{u}}+)$ for which $\mathcal{E}_{\underline{\underline{u}}} = \mathcal{E}_{\underline{\underline{z}}+h}$ on $(-\infty, \underline{\underline{s}}]$. Then by the same argument as in the case $G(\underline{\underline{z}} + h) \le \underline{\underline{s}}$ but with $\underline{\underline{z}} + h$ replaced by $\underline{\underline{u}}$ we find that $\phi(\underline{\underline{u}}) \le \phi(\underline{\underline{z}})$ again contradicting the maximality of $\underline{\underline{z}}$ in $H_{S^{-1}(x),v}$. Indeed $\mathcal{E}_{\underline{\underline{u}}} \ge \mathcal{E}^c_{\underline{\underline{u}}} \ge L^\delta$ and $\mathcal{E}_{\underline{\underline{u}}}(\underline{\underline{s}}) = \mathcal{E}^c_{\underline{\underline{u}}}(\underline{\underline{s}}) = L^\delta(\underline{\underline{s}})$ so that $S(\underline{\underline{u}}) \le \underline{\underline{s}}$ and $(\mathcal{E}^c_{\underline{\underline{u}}})'(\underline{\underline{s}}-) \le \phi(\underline{\underline{z}})$. Then $\phi(\underline{\underline{u}}) = (\mathcal{E}^c_{\underline{\underline{u}}})'(S(\underline{\underline{u}})-) \le (\mathcal{E}^c_{\underline{\underline{u}}})'(\underline{\underline{s}}-) \le \phi(\underline{\underline{z}})$.

Now let $L \equiv L^{\mathcal{E}_{\underline{\underline{z}}},\phi(\underline{\underline{z}})}_{R(\underline{\underline{z}})}$ and introduce

$$\overline{\overline{z}} = \sup\{z : z > \underline{\underline{z}} \text{ such that } \mathcal{E}_z \ge L \text{ on } \mathbb{R}\}.$$

It is clear that $\{z : z > \underline{\underline{z}} \text{ such that } \mathcal{E}_z \ge L \text{ on } \mathbb{R}\}$ is non-empty and thus $\overline{\overline{z}}$ is well-defined.

It follows similarly to the forward implication that $\phi(\underline{\underline{z}}) = \phi(\overline{\overline{z}})$ and $R(\underline{\underline{z}}) = R(\overline{\overline{z}})$. Then we must have $v < \overline{\overline{z}}$ since otherwise $\underline{\underline{z}}$ is not maximal. Then $\mathcal{E}_v > \mathcal{E}_{\overline{\overline{z}}}$ on $(G(\overline{\overline{z}}), \infty)$ and

$\mathcal{E}_v(S(v)) > L(S(v))$ since otherwise $\phi(v) \le \phi(\underline{z})$ which was ruled out above. It follows that the line joining $(S(\underline{z}), \mathcal{E}_v(S(\underline{z})))$ to $(S(v), \mathcal{E}_v(S(v)))$ has slope steeper than $\phi(\underline{z})$. Further, this line lies strictly below $\mathcal{E}_v \equiv D$ to the left of $S(\underline{z})$. Hence $R(v) \ge S(\underline{z})$. Finally, by the right-continuity of S^{-1} we have $S(\underline{z}) > x$ and hence $R(v) > x$ as required. □

We are now ready to prove Proposition 7.6.

Proof of Proposition 7.6. Suppose $D(y) < \mathcal{E}_{S^{-1}(y)}(y)$ then the result follows by combining Lemmas A.3 and A.5.

Suppose $D(y) = \mathcal{E}_{S^{-1}(y)}(y)$ and for all $v > S^{-1}(y)$ there exists k such that $\mathcal{E}_v(k) < L_y^{\mathcal{E}_{S^{-1}(y)}, (\mathcal{E}^c_{S^{-1}(y)})'(y-)}(k)$. Then the result follows from Lemma A.4 and Lemma A.5.

The remaining case is if $D(y) = \mathcal{E}_{S^{-1}(y)}(y)$ and there exists $v_0 > S^{-1}(y)$ such that $\mathcal{E}_{v_0} \ge L$ where $L = L_y^{\mathcal{E}_{S^{-1}(y)}, (\mathcal{E}^c_{S^{-1}(y)})'(y-)}$. Since y is a continuity point of μ and ν, $(\mathcal{E}^c_{S^{-1}(y)})'$ and $\mathcal{E}'_{S^{-1}(y)}$ exist at y and $L = L_y^{\mathcal{E}_{S^{-1}(y)}, (\mathcal{E}^c_{S^{-1}(y)})'(y)} = L_y^{\mathcal{E}_{S^{-1}(y)}, \mathcal{E}'_{S^{-1}(y)}(y)}$.

Define $\tilde{z} = \sup\{u : u > S^{-1}(y), \mathcal{E}_u \ge L\}$ and note that it is well-defined due to the existence of v_0. Then, $R(\tilde{z}) \le y$ and $\phi(\tilde{z}) = \mathcal{E}'_{S^{-1}(y)}(y) = D'(y)$. Further, by the right-continuity of S^{-1} we must have $\tilde{z} \le \overline{z}$ where $\overline{z} := S^{-1}(S(\tilde{z}))$.

Recall that $\psi_{<y>}$ is shorthand for $\psi_{S^{-1}(y), P'_\nu(y) - S^{-1}(y)}$. We will prove (7.6) by proving that

$$\{v : v \in I, R(v) \le y\} = \{v : v \in I, \psi_{<y>}(v) = \phi(v)\} \tag{A.7}$$

for each of $I = (S^{-1}(y), \tilde{z}]$, $I = (\tilde{z}, \overline{z}]$ and $I = (\overline{z}, 1)$ separately.

Case $I = (S^{-1}(y), \tilde{z}]$.

First note that if $v \in (S^{-1}(y), \tilde{z}]$ then since $\mathcal{E}_v \ge L$, we have that $\mathcal{E}_{S^{-1}(y)} \ge \mathcal{E}_v \ge \mathcal{E}^c_v \ge L$ with all inequalities being equalities at y. Since the subdiferential of $\mathcal{E}^c_v$ is non-decreasing and $y < S(v)$, we have that $\phi(v) \ge \mathcal{E}'_{S^{-1}(y)}(y)$ and therefore $\psi_{<y>}(v) = \mathcal{E}'_{S^{-1}(y)}(y)$. It follows that $\psi_{<y>}(v) = \phi(\tilde{z}) = \mathcal{E}'_{S^{-1}(y)}(y) = P'_\nu(y) - S^{-1}(y)$. Further, if $\psi_{<y>}(v) = \phi(v)$ we must have that $\mathcal{E}_v(S(v)) = L(S(v))$ and consequently $R(v) \le y$.

On the other hand, suppose $v \in (S^{-1}(y), \tilde{z}]$ and $R(v) \le y$. As $v > S^{-1}(y)$ we have $y < S(v)$. Further, since $\mathcal{E}^c_{S^{-1}(y)}(y) = \mathcal{E}^c_v(y) = \mathcal{E}_{S^{-1}(y)}(y) = \mathcal{E}_v(y) = D(y)$, by Lemma 4.6 we have that $\mathcal{E}^c_v = \mathcal{E}^c_{S^{-1}(y)}$ on $(-\infty, y]$. If $R(v) = y$ then $\phi(v) = (\mathcal{E}^c_v)'(y) = D'(y) = \psi_{<y>}(v)$. Otherwise, if $R(v) < y$ we have $\phi(v) = (\mathcal{E}^c_v)'(R(v)+) = (\mathcal{E}^c_{S^{-1}(y)})'(R(v)+) \le (\mathcal{E}^c_{S^{-1}(y)})'(y) = \psi_{<y>}(v) \le \phi(v)$. Again we conclude $\phi(v) = \psi_{<y>}(v) = D'(y)$ as required.

For future reference note that $\psi_{<y>}(\tilde{z}) = \phi(\tilde{z})$.

Case $I = (\tilde{z}, \overline{z}]$.

Recall $\overline{z} = S^{-1}(S(\tilde{z}))$. Throughout this section we assume $\tilde{z} < \overline{z}$ else there is nothing to prove.

Let $\overline{s} = S(\overline{z})$. By construction S is constant on $(\tilde{z}, \overline{z}]$ and $\overline{z} = S^{-1}(S(\overline{z}))$. Further, $S(z) > \overline{s}$ if and only if $z > \overline{z}$. We show that both $\{v : v \in (\tilde{z}, \overline{z}], R(v) \le y\} = (\tilde{z}, \overline{z}]$ and $\{v : v \in (\tilde{z}, \overline{z}], \psi_{<y>}(v) = \phi(v)\} = (\tilde{z}, \overline{z}]$.

We have $S(\tilde{z}) \le \overline{s}$ and $L(S(\tilde{z})) = \mathcal{E}_{\tilde{z}}(S(\tilde{z}))$. We show that we must have $\mathcal{E}_{\tilde{z}}(\overline{s}) = L(\overline{s})$ so that if $S(\tilde{z}) < \overline{s}$ then $\mathcal{E}_{\tilde{z}} = L$ on $[S(\tilde{z}), \overline{s}]$. Suppose for a contradiction that $\mathcal{E}_{\tilde{z}}(\overline{s}) > L(\overline{s})$. Then either $\mathcal{E}_{\tilde{z}}(\overline{s}) = D(\overline{s}) > L(\overline{s})$ or both $\mathcal{E}_{\tilde{z}}(\overline{s}) > D(\overline{s})$ and $\mathcal{E}_{\tilde{z}}(\overline{s}) > L(\overline{s})$. In the former case we can find small enough $\breve{z} \in (\tilde{z}, 1)$ such that $\mathcal{E}_{\breve{z}} \ge L$, contradicting the maximality of $\tilde{z}$. In the second case $G(\tilde{z}+) < \overline{s}$ (since $\mathcal{E}_{\tilde{z}}(\overline{s}) > D(\overline{s})$), $S(\tilde{z}) < \overline{s}$ (since $\mathcal{E}_{\tilde{z}}(\overline{s}) > L(\overline{s})$), and there must exist $z_1 \in (\tilde{z}, \overline{z}]$ such that, for all $z \in (\tilde{z}, z_1)$, $G(z) < \overline{s}$ and $\mathcal{E}_{\tilde{z}}(\overline{s}) > \mathcal{E}_z(\overline{s}) > L(\overline{s}) \vee D(\overline{s})$. Then by the maximality of $\tilde{z}$ there must be $k_1 \in (G(\tilde{z}), \overline{s})$ such that $\mathcal{E}_{z_1}(k_1) < L(k_1)$. Fix $z \in (\tilde{z}, z_1)$ and let $\phi_1 = \inf_{x>y} \frac{\mathcal{E}_z(x) - \mathcal{E}_z(y)}{x-y}$ and let the infimum be attained at x_1. Then $\phi_1 < \phi(\tilde{z})$ and $x_1 \in (G(\tilde{z}), \overline{s})$. If $G(z) \le x_1$ then $x_1 \le S(z) \le s_1 = \sup\{s : s > x_1, \mathcal{E}_z(s) \le L(s)\} < \overline{s}$, a contradiction since $S(z) = S(\overline{z}) = \overline{s}$.

Conversely, if $x_1 < G(z)$ then there exists $z_2 \in (\tilde{z}, z)$ such that $G(z_2) \le x_1 \le G(z_2+)$. Then if $s_2 = \sup\{s : s > x_1, \mathcal{E}_{z_2}(s) \le L(s)\}$ we have $s_2 < \overline{s}$ and $S(z_2) \le s_2 < \overline{s}$, again a contradiction. Hence we must have $\mathcal{E}_{\tilde{z}}(\overline{s}) = L(\overline{s})$.

Note that for $v \in (\tilde{z}, \overline{z}]$, $S(v) = S(S^{-1}(S(\tilde{z}))) = S(\overline{z}) = \overline{s}$. Furthermore, we must have that $G(\tilde{z}) \le G(\tilde{z}+) < \overline{s}$. (If $G(\tilde{z}+) \ge \overline{s}$, then, for all $v \in (\tilde{z}, \overline{z}]$, $G(\tilde{z}+) \le G(v) \le S(v) = \overline{s} \le G(\tilde{z}+)$ and hence $G(v) = S(v)$. Then using (4.5) we deduce that $\mathcal{E}_v = \mathcal{E}_{\tilde{z}}$ on $(-\infty, \overline{s} = S(v)]$. Then $\mathcal{E}^c_v(\overline{s}) = \mathcal{E}_v(\overline{s}) = \mathcal{E}_{\tilde{z}}(\overline{s}) = L(\overline{s}) = \mathcal{E}^c_{\tilde{z}}(\overline{s})$ which implies by Lemma 4.6 that $\mathcal{E}^c_v = \mathcal{E}^c_{\tilde{z}}$ on $(-\infty, \overline{s}]$. Then, since $L = \mathcal{E}^c_v$ on an interval to the left of $\overline{s}$ we have $L \le \mathcal{E}^c_v$ everywhere, contradicting the maximality of $\tilde{z}$.) It follows by (4.5) again that $\mathcal{E}_v(\overline{s}) < \mathcal{E}_{\tilde{z}}(\overline{s}) = L(\overline{s})$ for all $v \in (\tilde{z}, \overline{z}]$.

First suppose that $G(\tilde{z}) \le S(\tilde{z}) \le G(\tilde{z}+) < \overline{s}$ and continue to take $v \in (\tilde{z}, \overline{z}]$. Since $\mathcal{E}_{\tilde{z}} \equiv L$ on $[G(\tilde{z}+), \overline{s}]$ and $P_{\mu_{\tilde{z}}}$ is linear on $[G(\tilde{z}+), \infty)$, we have that $P_\nu = \mathcal{E}_{\tilde{z}} + P_{\mu_{\tilde{z}}}$ is also linear on $[G(\tilde{z}+), \overline{s}]$. Then $D = P_\nu - P_\mu$ must be concave and below L on $(G(\tilde{z}+), \overline{s}]$ (recall that $D(\overline{s}) \le \mathcal{E}_v(\overline{s}) < L(\overline{s})$). If $R(v) \in [S(\tilde{z}), G(\tilde{z}+)]$, then $\mathcal{E}_v(R(v)) = D(R(v)) = L(R(v))$ and $L^{\mathcal{E}_v}_{R(v),\overline{s}}$ must cross L at $R(v)$. If $R(v) \in (G(\tilde{z}+), \overline{s})$ then, since D is concave and equal to or below L at $R(v)$, $L^{\mathcal{E}_v}_{R(v),\overline{s}}$ crosses L at some $k \in [G(\tilde{z}+), \overline{s})$. In both cases we have that $\mathcal{E}^c_{\tilde{z}} = L < L^{\mathcal{E}_v}_{R(v),\overline{s}} \le \mathcal{E}^c_v$ on an interval to the left of the corresponding crossing point, a contradiction to $\mathcal{E}^c_v \le \mathcal{E}^c_{\tilde{z}}$. Hence, since $R(v) \notin (R(\tilde{z}), S(\tilde{z}))$, either $R(v) \le R(\tilde{z}) \le y$ or $R(v) = G(v) = S(v) = \overline{s}$. However the latter cannot happen due to concavity of D on $[G(\tilde{z}+), \overline{s}]$. Hence $\{v : v \in (\tilde{z}, \overline{z}], R(v) \le y\} = (\tilde{z}, \overline{z}]$.

We now show that ϕ is non-increasing on $(\tilde{z}, \overline{z}]$, and hence, since $\psi_{<y>}(\tilde{z}) = \phi(\tilde{z})$, that $\psi_{<y>}(v) = \phi(v)$ on $(\tilde{z}, \overline{z}]$. Take $\tilde{z} < v < w \le \overline{z}$. Since $R(w) < S(w) = S(v) = \overline{s}$ we have $R(w) \le R(v) < S(v)$. Then, since $\partial\mathcal{E}^c_{\tilde{z}}$ is non-decreasing, and using Lemma 6.4 for the second inequality,

$$\phi(w) = (\mathcal{E}^c_w)'(R(w)+) \le (\mathcal{E}^c_w)'(R(v)+) \le (\mathcal{E}^c_v)'(R(v)+) = \phi(v)$$

and the result follows.

Now suppose $G(\tilde{z}+) < S(\tilde{z})$. Then $\{v \in (\tilde{z}, \overline{z}] : G(v) < S(\tilde{z})\}$ is non-empty. Define $\hat{z} := \sup\{v \in (\tilde{z}, \overline{z}] : G(v) < S(\tilde{z})\}$ and note that $\tilde{z} < \hat{z} \le \overline{z}$. We now show that

$$\{v \in (\tilde{z}, \hat{z}] : R(v) \le y\} = \{v \in (\tilde{z}, \hat{z}] : \psi_{<y>}(v) = \phi(v)\} = (\tilde{z}, \hat{z}]. \tag{A.8}$$

For $u \in (\tilde{z}, \hat{z})$, $R(u) \notin (R(\tilde{z}), S(\tilde{z}))$ and $R(u) \le G(u) < S(\tilde{z})$, and therefore $R(u) \le R(\tilde{z}) \le y$. Then by Proposition 6.1(iii), $R(\hat{z}) \le \liminf_{u \uparrow \hat{z}} R(u) \le y$. On the other hand, the same argument shows that $R(v) \le R(u) \le y$ for $u, v \in [\tilde{z}, \hat{z})$ with $u < v$. Then we must have that $\mathcal{E}^c_z(R(z)) = \mathcal{E}^c_{\tilde{z}}(R(z))$ for $z \in \{u, v\}$. Since the subdifferential of $\mathcal{E}^c_{\tilde{z}}$ is non-decreasing and $\mathcal{E}^c_u(S(u)) = \mathcal{E}_u(S(u)) > \mathcal{E}_v(S(v)) = \mathcal{E}^c_v(S(v))$, for $\tilde{z} \le u < v < \hat{z}$ we have that $\phi(u) \ge \phi(v)$. Then since $\phi(\tilde{z}) = \psi_{<y>}(\tilde{z})$ we have $\psi_{<y>}(\tilde{z}) \ge \phi(u) \ge \phi(v) \ge \lim_{u \uparrow \hat{z}} \phi(u) = \phi(\hat{z})$, which proves the claim.

Now, if $\hat{z} = \overline{z}$, then (A.7) follows for $I = (\tilde{z}, \overline{z}]$. Therefore, suppose that $\hat{z} < \overline{z}$. We claim that

$$\{v \in (\hat{z}, \overline{z}] : R(v) \le y\} = \{v \in (\hat{z}, \overline{z}] : \psi_{<y>}(v) = \phi(v)\} = (\hat{z}, \overline{z}].$$

Indeed, using the definition of $\hat{z}$, we have that $G(v) = \overline{s}$ for all $v \in (\hat{z}, \overline{z}]$, and therefore $G(\hat{z}+) = \overline{s}$. Then by (4.5), for all $v \in (\hat{z}, \overline{z}]$, $\mathcal{E}_v = \mathcal{E}_{\hat{z}}$ on $(-\infty, \overline{s} = S(v)]$. It follows from Corollary 4.7 with $u = \hat{z}$ that for all $v \in (\hat{z}, \overline{z}]$ we have $R(v) = R(\hat{z}) \le y$ and $\psi_{<y>}(\hat{z}) = \psi_{<y>}(v) = \phi(v)$. We conclude that (A.7) holds for $I = (\hat{z}, \overline{z}]$, and hence, given (A.8), for $I = (\tilde{z}, \overline{z}]$.

Case $v \in (S^{-1}(S(\tilde{z})), 1)$**.**

In the case $\tilde{z} = \overline{z}$ then $\tilde{z} = \overline{z} = S^{-1}(S(\overline{z})) = S^{-1}(S(\tilde{z}))$ and then from the first case $R(\overline{z}) = R(\tilde{z}) \leq y < S(\overline{z})$ and $\psi_{<y>}(\overline{z}) = \psi_{<y>}(\tilde{z}) = \phi(\tilde{z}) = \phi(\overline{z})$. In the case where $\tilde{z} < \overline{z}$ then again we have $\overline{z} = S^{-1}(S(\overline{z}))$ and then from the second case $R(\overline{z}) \leq y < S(\overline{z})$ and $\psi_{<y>}(\overline{z}) = \phi(\overline{z})$.

By the left-monotonicity of R, for $v > \overline{z}$ we have $R(v) \notin (R(\overline{z}), S(\overline{z}))$. Since $R(\overline{z}) \leq y < S(\overline{z})$ we conclude $R(v) \leq y$ is equivalent to $R(v) \leq R(\overline{z})$. Then $\{v : v > \overline{z}, R(v) \leq y\} = \{v : v > \overline{z}, R(v) \leq R(\overline{z})\}$.

Note that for $v > \overline{z}$ we have that $\psi_{<y>}(v) = \psi_{\overline{z},\phi(\overline{z})}(v)$. Hence to show (A.7) for $I = (\overline{z}, 1)$ we need to show that

$$\{v : v > \overline{z}, R(v) \leq R(\overline{z}))\} = \{v : v > \overline{z}, \psi_{\overline{z},\phi(\overline{z})}(v) = \phi(v)\}. \tag{A.9}$$

But, since $R(\overline{z}) < S(\overline{z})$ this is immediate from Lemma A.1. □

References

[1] Azéma J., Yor M.: *Une solution simple au problème de Skorokhod.* In Séminaire de probabilités XIII, vol. 721 of Lecture Notes in Mathematics, 90–115, Springer, Berlin, (1979). MR0544782

[2] Bayraktar E., Deng S., Norgilas D.: *A potential-based construction of the increasing supermartingale coupling.* Available online at: https://arxiv.org/abs/2108.03450, (2021).

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

[4] Beiglböck, M., Cox A.M.G., Huesmann, M.: *Optimal transport and Skorokhod embedding.* Invent. Math., 208(2):327–400, (2017). MR3639595

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

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

[7] Beiglböck M., Hobson D., Norgilas D.: *Potential of the shadow measure.* Electronic Communications in Probability, 27, paper no. 16, 1–12 (2022). MR4389158

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

[9] Beiglböck M., Juillet N.: *Shadow couplings.* arXiv preprint, (2016). Available online at: https://arxiv.org/abs/1609.03340.

[10] Beiglböck M., Nutz M.: *Martingale inequalities and deterministic counterparts.* Electron. J. Probab., 19(95):1–15, (2014). MR3272328

[11] Brenier Y.: *Décomposition polaire et réarrangement monotone des champs de vecteurs.* Comptes Rendus de l'Académie des Sciences, Series I - Mathematics, Paris, 305(19): 805–808, (1987). MR0923203

[12] 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 MR4490406

[13] Campi L., Laachir I., Martini C.: *Change of numeraire in the two-marginals martingale transport problem.* Finance Stoch., 21:471–486, (2017). MR3626622

[14] Chacon, R.V.: *Potential processes.* Transactions of the American Mathematical Society, 226:39–58, (1977). MR0501374

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

[16] Dubins L.E.: *On a theorem of Skorohod.* Ann. Math. Statist., 39:2094–2097, (1968). MR0234520

[17] Galcihon A., Henry-Labordère P., Touzi N.: *A stochastic control approach to no-arbitrage bounds given marginals with an application to lookback options.* The Annals of Applied Probability, 24(1):313–336, (2014). MR3161649

[18] Henry-Labordère P., Obłój J., Spoida P., Touzi N. *The maximum maximum of a martingale with given n marginals.* Ann. Appl. Probab., 26(1):1–44, (2016). MR3449312

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

[20] 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

[21] Hirsch F., Roynette B.: *A new proof of Kellerer's theorem.* ESAIM: Probability and Statistics, 16:48–60, (2012). MR2911021

[22] Hobson D.G.: *Robust hedging of the lookback option.* Finance and Stochastics, 2(4):329–347, (1998).

[23] Hobson D.G.: *The maximum maximum of a martingale.* In Séminaire de probabilités XXXII, vol. 1686 of Lecture notes in Mathematics, 250–263, Springer, Berlin, (1998). MR1655298

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

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

[26] Hobson D.G., Norgilas D.: *Robust bounds for the American Put.* Finance and Stochastics, 23(2):359–395, (2019). MR3933425

[27] Hobson D.G., Norgilas D.: *The left-curtain martingale coupling in the presence of atoms.* The Annals of Applied Probability, 29(3):1904–1928, (2019). MR3914560

[28] Hobson D.G., Pedersen J.: *The minimum maximum of a continuous martingale with given initial and terminal laws.* The Annals of Applied Probability, 30(2):978–999, (2002). MR1906424

[29] Jacka S.: *Doob's inequalities revisited: a maximal H^1-embedding.* Stochastic Process. Appl., 29(2):281–290, (1988). MR0958505

[30] Juillet N.: *Stability of the shadow projection and the left-curtain coupling.* Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 52(4):1823–1843, (2016). MR3573297

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

[32] Kantorovich L. V.: *On the translocation of masses.* In Doklady Akademii Nauk SSSR, 37(7-8):227–229, (1942). MR0009619

[33] Källblad S., Tan X., Touzi, N. *Optimal Skorokhod embedding given full marginals and Azéma-Yor peacocks.* Ann. Appl. Probab., 27(2):686–719, (2017). MR3655851

[34] Nutz M., Stebegg F.: *Canonical supermartingale couplings.* The Annals of Probability, 46(6):3351–3398, (2018). MR3857858

[35] Nutz M., Stebegg F., Tan X.: *Multiperiod martingale transport.* Stochastic Processes and their Applications (in press), available online at: https://doi.org/10.1016/j.spa.2019.05.010, (2019). MR4058283

[36] Obłój J., Spoida P., Touzi N. *Martingale inequalities for the maximum via pathwise arguments.* In Memoriam Marc Yor-Séminaire de Probabilités XLVII, pages 227–247. Springer, (2015). MR3444301

[37] Rockafellar R. T.: *Convex analysis.* Princeton Univ., Princeton, NJ., second printing (1990). MR0274683

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

[39] Strassen V.: *The existence of probability measures with given marginals.* The Annals of Mathematical Statistics, 36(2):423–439, (1965). MR0177430

[40] Vallois P.: *Le problème de Skorokhod sur $\mathbb{R}$: une approche avec le temps local.* In Séminaire de Probabilités, XVII, volume 986 of Lecture Notes in Math., pages 227–239. Springer, Berlin, (1983). MR0770416