

**Manuscript version: Published Version**

The version presented in WRAP is the published version (Version of Record).

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/114293>

**How to cite:**

The repository item page linked to above, will contain details on accessing citation guidance from the publisher.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

## THE LEFT-CURTAIN MARTINGALE COUPLING IN THE PRESENCE OF ATOMS

BY DAVID G. HOBSON AND DOMINYKAS NORGILAS

*University of Warwick*

Beiglböck and Juillet (*Ann. Probab.* **44** (2016) 42–106) introduced the left-curtain martingale coupling of probability measures  $\mu$  and  $\nu$ , and proved that, when the initial law  $\mu$  is continuous, it is supported by the graphs of two functions. We extend the later result by constructing the generalised left-curtain martingale coupling and show that for an arbitrary starting law  $\mu$  it is characterised by two appropriately defined lower and upper functions.

As an application of this result, we derive the model-independent upper bound of an American put option. This extends recent results of Hobson and Norgilas (2017) on the atom-free case.

**1. Introduction.** Given two probability measures  $\mu$  and  $\nu$  on  $\mathbb{R}$  and a payoff function  $c : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , the classical problem of optimal transport is to construct a joint distribution for random variables  $Z_1 \sim \mu$  and  $Z_2 \sim \nu$  which minimises  $\mathbb{E}[c(Z_1, Z_2)]$ . Beiglböck et al. [3] and Galichon et al. [10] introduced a martingale version of the transportation problem and related it to the problem of finding model-independent bounds of exotic derivatives in mathematical finance. Given  $\mu$  and  $\nu$  in convex order, the basic problem of martingale optimal transport (MOT) is to construct a martingale  $M$ , with  $M_1 \sim \mu$ ,  $M_2 \sim \nu$ , which minimises  $\mathbb{E}[c(M_1, M_2)]$ . In this setting, a martingale transport or coupling can be identified with a measure  $\pi$  on  $\mathbb{R}^2$  with univariate marginals  $\mu$  and  $\nu$ , and such that  $\int_{y \in \mathbb{R}} \int_{x \in A} (y - x)\pi(dx, dy) = 0$  for all Borel sets  $A$ , and the MOT is to find  $\pi$  to minimise  $\iint c(x, y)\pi(dx, dy)$ . In the context of mathematical finance, this problem was first studied in Hobson and Neuberger [13] for the payoff  $c(x, y) = -|y - x|$ .

Beiglböck and Juillet [5] introduced the notion of a left-monotone martingale coupling and established that for (arbitrary) fixed marginals  $\mu$  and  $\nu$  in convex order there exists a unique such coupling (called the left-curtain martingale coupling and denoted by  $\pi_{1c}$ ). The left-curtain martingale coupling may be viewed as a martingale analogue to the monotone Hoeffding–Fréchet coupling in classical optimal transport. The authors also proved the optimality of  $\pi_{1c}$  for a specific class of payoff functions. Henry-Labordère and Touzi [11] extended the results of Beiglböck and Juillet [5] and showed optimality for a wider class of payoff functions.

---

Received March 2018; revised October 2018.

*MSC2010 subject classifications.* Primary 60G42; secondary 60G40.

*Key words and phrases.* Martingale optimal transport, optimal stopping, model-independent pricing, American put.

Beiglböck et al. [4] analysed the left-curtain coupling further and gave a simplified proof of uniqueness under the additional assumption that  $\mu$  is continuous. Juillet [18] proved that  $\pi_{lc}$  is continuous, and thus, for general distributions, it can be approximated by the left-curtain couplings corresponding to “nice” (e.g., finitely supported or continuous) initial and/or target laws. A number of further articles investigate the properties and extensions of  $\pi_{lc}$ ; see Beiglböck et al. [2, 4] and Nutz et al. [20, 21].

Beiglböck and Juillet [5] also established a martingale version of the fundamental Brenier’s [7] result in the classical optimal transport which states that, for a sufficiently regular initial distribution  $\mu$ , the optimal transport map is unique and supported by the graph of the gradient of some convex function (a monotonically increasing function in one dimension). In particular, the authors showed that under the assumption that the initial law  $\mu$  is continuous, the left-curtain martingale coupling is supported by the graphs of lower and upper functions  $T_d$  and  $T_u$ , respectively, so that  $M_2 \in \{T_d(M_1), T_u(M_1)\}$ . Henry-Labordère and Touzi [11] gave an explicit construction of  $T_d$  and  $T_u$  using differential equations. However, when  $\mu$  has an atom at  $x$  the element  $\pi_{lc}^x(\cdot)$  in the disintegration  $\pi_{lc}(dx, dy) = \mu(dx)\pi_{lc}^x(dy)$  becomes a measure with support on nontrivial subsets of  $\mathbb{R}$  and not just on a two-point set. Then we cannot construct functions  $(T_d, T_u)$ , unless we allow them to be multivalued.

Our goal in this paper is to show how by changing our viewpoint we can again recover the property that  $M_2$  takes values in a two-point set. The idea is to write  $M_1 = h(Z)$  for a continuous random variable  $Z$  (in fact we take  $Z \equiv U \sim U(0, 1)$ ) and then to find  $f_{Z,h}$  and  $g_{Z,h}$  such that  $M_2 \in \{f_{Z,h}(Z), g_{Z,h}(Z)\}$ . Then, although there is uniqueness at the level of martingale couplings  $\pi$ , when  $\mu$  contains atoms there are many possible choices of  $(f_{Z,h}, g_{Z,h})$ , even for fixed  $Z$  and monotonic increasing  $h$ . Nonetheless, we show that amongst this set there is an essentially unique choice  $(f_{Z,h}, g_{Z,h})$  with a special monotonicity property.

The motivation for this extension of the left-curtain martingale coupling comes from mathematical finance. The recent study of American put options in Hobson and Norgilas [17] highlights the role of the left-curtain martingale coupling in finding the model-independent upper bound on the price of the American put. When  $\mu$  is continuous the authors show how the optimal martingale coupling and the optimal stopping time can be obtained from the functions  $f = T_d$  and  $g = T_u$  which arise in the construction of the left-curtain coupling. In particular, for the optimal model there is a Borel subset of  $\mathbb{R}$ , say  $B$ , such that it is optimal to stop at time-1 if  $M_1 \in B$ , and at time-2 otherwise. Moreover, the structure of  $f$  and  $g$  allows us to identify the cheapest superhedging strategy that supports the price of the American put.

If  $\mu$  has atoms then the situation becomes more delicate, essentially because we must allow for a wider range of possible candidates for exercise determining sets  $B$ . On atoms of  $\mu$ , we may want to sometimes stop and sometimes continue, although we must still take stopping decisions which do not violate the martingale

property. As the stopping decision in the continuous case is based on the natural filtration of the martingale  $M$ , if  $M_1$  ends up at the atom of  $\mu$ , then it is not clear, using only the structure of  $f$  and  $g$ , what part of mass at time-1 should be stopped and what part should be allowed to continue. This is the reason why we must extend the notion of the left-curtain martingale coupling.

The main effort in this article is in proving Theorem 1 which extends the left-curtain martingale coupling to the presence of atoms in the starting law  $\mu$ . We show that this extended coupling is again characterised by lower and upper functions,  $R$  and  $S$ , respectively. However, while  $f$  and  $g$  are multivalued on the atoms of  $\mu$ ,  $R$  and  $S$  remain well defined. Then our second achievement is to show how the structure of  $R$  and  $S$  can be used to characterise the model and stopping rule which achieves the highest possible price for the American put, and the cheapest superhedge. This generalises results of Hobson and Norgilas [17]: for arbitrary  $\mu$  and  $\nu$ , the highest model based price of the American put is equal to the cost of the cheapest superhedge.

**2. Preliminaries and set-up.** Let  $\mathcal{M}(\mathbb{R}^n)$  (resp.,  $\mathcal{P}(\mathbb{R}^n)$ ) be the set of measures (resp., probability measures) on  $\mathbb{R}^n$ . Given an integrable  $\eta \in \mathcal{M}(\mathbb{R})$ , that is,  $\int_{\mathbb{R}} |x| \eta(dx) < \infty$ , define  $\bar{\eta} = \frac{\int_{\mathbb{R}} x \eta(dx)}{\int_{\mathbb{R}} \eta(dx)}$  to be a barycentre of  $\eta$ . Let  $\mathcal{I}_\eta$  with endpoints  $\{\ell_\eta, r_\eta\}$  be the smallest interval containing the support of  $\eta$  (with the convention that finite endpoints are included). Define  $P_\eta : \mathbb{R} \mapsto \mathbb{R}^+$  by  $P_\eta(k) = \int_{-\infty}^k (k-x)\eta(dx)$ . Then  $P_\eta$  is convex and increasing,  $\lim_{z \downarrow -\infty} P_\eta(z) = 0$ ,  $\lim_{z \uparrow \infty} P_\eta(z) - \eta(\mathbb{R})(z - \bar{\eta})^+ = 0$  and  $\{k : P_\eta(k) > \eta(\mathbb{R})(k - \bar{\eta})^+\} \subseteq \mathcal{I}_\eta$ . Note that  $P_\eta$  is related to the potential  $U_\eta$  defined by  $U_\eta(k) := -\int_{\mathbb{R}} |k-x|\eta(dx)$  by  $P_\eta(k) = \frac{1}{2}(-U_\eta(k) + (k - \bar{\eta})\eta(\mathbb{R}))$ . For  $\eta \in \mathcal{P}(\mathbb{R})$  let  $F_\eta$  be the distribution function of  $\eta$  and let  $G_\eta : (0, 1) \mapsto \mathbb{R}$  be the quantile function of  $\eta$ , which is taken to be left-continuous unless otherwise stated.

Two measures  $\eta$  and  $\chi$  are in convex order, and we write  $\eta \leq_{\text{cx}} \chi$ , if and only if  $\eta(\mathbb{R}) = \chi(\mathbb{R})$ ,  $\bar{\eta} = \bar{\chi}$  and  $P_\eta(k) \leq P_\chi(k)$  on  $\mathbb{R}$ . Necessarily we must have  $\ell_\chi \leq \ell_\eta \leq r_\eta \leq r_\chi$ . For any two probability measures  $\eta, \chi \in \mathcal{P}(\mathbb{R})$ , we write  $\pi \in \Pi(\eta, \chi)$  if  $\pi \in \mathcal{P}(\mathbb{R}^2)$  and has first marginal  $\eta$  and second marginal  $\chi$ . If  $\pi \in \Pi(\eta, \chi)$  is such that the following martingale condition holds:

$$\begin{aligned}
 \int_{x \in B} \int_{y \in \mathbb{R}} y \pi(dx, dy) &= \int_{x \in B} \int_{y \in \mathbb{R}} x \pi(dx, dy) \\
 (1) \qquad \qquad \qquad &= \int_B x \eta(dx) \qquad \forall \text{ Borel } B \subseteq \mathbb{R},
 \end{aligned}$$

we write  $\pi \in \Pi_M(\eta, \chi) \subset \Pi(\eta, \chi)$  and say that  $\pi$  is a martingale coupling of  $\eta$  and  $\chi$ . By a classical result of Strassen [22],  $\Pi_M(\eta, \chi)$  is nonempty if and only if  $\eta \leq_{\text{cx}} \chi$ .

DEFINITION 1 (Hobson and Neuberger [16]). Suppose  $\mu \leq_{\text{cx}} \nu$ .

Let  $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2\})$  be a filtered probability space. We say  $M = (M_0, M_1, M_2) = (\bar{\mu}, X, Y)$  is a  $(\mathcal{S}, \mu, \nu)$  consistent stochastic process and we write  $M \in \mathbb{M}(\mathcal{S}, \mu, \nu)$  if:

1.  $M$  is a  $\mathcal{S}$ -martingale,
2.  $\mathcal{L}(M_1) = \mu$  and  $\mathcal{L}(M_2) = \nu$ .

We say  $(\mathcal{S}, M)$  is a  $(\mu, \nu)$ -consistent model if  $\mathcal{S}$  is a filtered probability space and  $M$  is a  $(\mathcal{S}, \mu, \nu)$  consistent stochastic process.

Let  $(\eta_n)_{n \geq 1}$  be a sequence of probability measures in  $\mathcal{P}(\mathbb{R})$ . For  $\eta \in \mathcal{P}(\mathbb{R})$ , we write  $\eta_n \xrightarrow{w} \eta$ , and say  $\eta_n$  converges weakly to  $\eta$ , if  $\lim_{n \rightarrow \infty} \int f d\eta_n = \int f d\eta$  for all bounded and continuous functions  $f$  on  $\mathbb{R}$  (see Billingsley [6]). If  $\eta_n \xrightarrow{w} \eta$ , if  $\eta_n \leq_{\text{cx}} \eta$  and if  $(\eta_n)_{n \geq 1}$  is increasing in convex order, that is,  $\eta_n \leq_{\text{cx}} \eta_{n+1}$  for each  $n$ , then we write  $\eta_n \uparrow_{\text{cx}} \eta$ .

LEMMA 1. *Suppose  $\mu \in \mathcal{P}(\mathbb{R})$  is integrable. Then there exists a sequence  $(\mu_n)_{n \geq 1}$  of finitely supported integrable measures in  $\mathcal{P}(\mathbb{R})$  such that  $\mu_n \uparrow_{\text{cx}} \mu$ .*

PROOF. Recall that for any  $\eta \in \mathcal{P}(\mathbb{R})$ ,  $U_\eta$  is concave, linear on each interval  $I \subset \mathbb{R}$  with  $\eta(I) = 0$ ,  $U_\eta(x) \leq -|\bar{\eta} - x| = U_{\delta_{\bar{\eta}}}(x)$  on  $\mathbb{R}$ , and  $\lim_{|x| \rightarrow \infty} U_\eta(x) + |\bar{\eta} - x| = 0$ . Moreover,  $\mu_n \uparrow_{\text{cx}} \mu$  if and only if  $U_{\mu_n} \downarrow U_\mu$  pointwise; see Chacon [8]. Let  $\mathcal{U}_\mu$  be a set of piecewise linear concave functions  $\tilde{U} : \mathbb{R} \rightarrow \mathbb{R}_-$  such that  $U_\mu(x) \leq \tilde{U}(x) \leq U_{\delta_{\bar{\mu}}}(x)$ . Then each  $\tilde{U} \in \mathcal{U}_\mu$  corresponds to a finitely supported integrable probability measure  $\tilde{\mu}$  on  $\mathbb{R}$  such that  $\delta_{\bar{\mu}} \leq_{\text{cx}} \tilde{\mu} \leq_{\text{cx}} \mu$ . Finally, Chacon and Walsh [9] provide a sequence of functions  $(\tilde{U}_n)_{n \geq 1}$  in  $\mathcal{U}_\mu$ , such that  $\tilde{U}_n \downarrow U_\mu$  pointwise, proving our claim.  $\square$

**3. An extension of the left-curtain mapping to the general case.** In this section, we construct a new representation of the left-curtain martingale coupling of Beiglböck and Juillet [5]. Our approach is to construct  $(X, Y)$  from a pair of independent uniform  $U(0, 1)$  random variables  $U$  and  $V$ . The construction of  $X$  is straightforward: we set  $X = G_\mu(U)$ .

It remains to construct  $Y$ . First, we consider the case of a point mass at  $w$ ,  $\mu = \delta_w$ , and show how to construct functions  $R = R_{\mu, \nu}$  and  $S = S_{\mu, \nu}$  with  $R_{\mu, \nu}(u) \leq G_\mu(u) \leq S_{\mu, \nu}(u)$ , such that if  $X = G_{\delta_w}(U) = w$  and  $Y \in \{R(U), S(U)\}$  with  $\mathbb{P}(Y = R(u) | U = u) = \frac{S(u) - G(u)}{S(u) - R(u)}$  then  $Y$  has law  $\nu$ . In particular, conditional on  $U = u$ ,  $Y$  takes values in  $\{R(u), S(u)\}$  and satisfies  $\mathbb{E}[Y | U = u] = G_\mu(u)$ . Second, we show how this result extends to the case of a measure  $\mu$  consisting of finitely many atoms. Third, for the case of general  $\mu$  we construct an approximation  $(\mu_n)_{n \geq 1}$  of  $\mu$  and associated functions  $(R_n, G_n, S_n)_{n \geq 1}$  where each  $\mu_n$  is finitely supported. We show that we can define limits  $(R, G, S)$  such that  $(R, G, S)$  can be used to construct a martingale  $M = (M_0 = \bar{\mu}, M_1 = X, M_2 = Y)$  with the

property that  $X = G(U)$  and  $Y \in \{R(U), S(U)\}$  and such that  $\mathcal{L}(X) = \mu$  and  $\mathcal{L}(Y) = \nu$ . The functions  $R, S : (0, 1) \mapsto \mathbb{R}$  we define have the properties

$$R(u) \leq G(u) \leq S(u);$$

$$(2) \quad S \text{ is increasing; for } 0 < u < v < 1, R(v) \notin (R(u), S(u)).$$

We suppose  $\mu \leq_{cx} \nu$  are fixed and given and we abbreviate the quantile function  $G_\mu$  by  $G$ . The aim of this section is to prove the following theorem.

**THEOREM 1.** *There exist functions  $R, S : (0, 1) \mapsto \mathbb{R}$  satisfying (2) 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  $G(u) = S(u)$  and*

$$(3) \quad Y(u, v) = R(u)I_{\{v \leq \frac{S(u)-G(u)}{S(u)-R(u)}\}} + S(u)I_{\{v > \frac{S(u)-G(u)}{S(u)-R(u)}\}}$$

*otherwise, and if  $U$  and  $V$  are independent  $U(0, 1)$  random variables then  $M = (\bar{\mu}, X(U), Y(U, V))$  is a  $\mathbb{F} = (\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_1 = \sigma(U), \mathcal{F}_2 = \sigma(U, V))$ -martingale for which  $\mathcal{L}(X) = \mu$  and  $\mathcal{L}(Y) = \nu$ .*

*In particular, if  $\Omega = (0, 1) \times (0, 1)$ ,  $\mathcal{F} = \mathcal{B}(\Omega)$ ,  $\mathbb{P} = \text{Leb}(\Omega)$ , if  $\mathbb{F}$  and  $M$  are defined as above and if  $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  then  $(\mathcal{S}, M)$  is a  $(\mu, \nu)$ -consistent model.*

**REMARK 1.** For  $n \geq 1$ , let  $\pi_{lc}^n$  be the left-curtain coupling of the initial law  $\mu_n$  (consisting of  $n$  atoms) and target law  $\nu$ . Juillet [18] proved that if  $(\mu_n)_{n \geq 1}$  converges weakly to  $\mu$  then  $(\pi_{lc}^n)_{n \geq 1}$  converges weakly to the left-curtain coupling of  $\mu$  and  $\nu$ .

Here, we argue differently. We use the fact that  $\pi_{lc}^n$  can be represented by an explicitly constructed triple  $(S_n, G_n, R_n)$ . Then, by sending  $n \rightarrow +\infty$ , we show that the limiting functions give rise to the left-monotone martingale coupling, and thus also to  $\pi_{lc}$ , of  $\mu$  and  $\nu$ .

When  $\mu$  is continuous and  $f$  and  $g$  are well defined the construction of this section is related to that of Beiglböck and Juillet [5] (see also Henry-Labordère and Touzi [11]) via the relationships  $S = g \circ G_\mu$  and  $R = f \circ G_\mu$ . Suppose  $\nu$  is also continuous and fix  $x$ . Then under the left-curtain martingale coupling  $\{f(x), g(x)\}$  with  $f(x) \leq x \leq g(x)$  are solutions to the mass and mean conditions

$$(4) \quad \int_f^x \mu(dz) = \int_f^g \nu(dz),$$

$$(5) \quad \int_f^x z\mu(dz) = \int_f^g zv(dz).$$

When  $\mu$  has atoms,  $G_\mu$  has intervals of constancy and  $f$  and  $g$  are multivalued, but  $R$  and  $S$  remain well defined. See Figure 1. Then, for general  $\mu$  and  $\nu$ , the

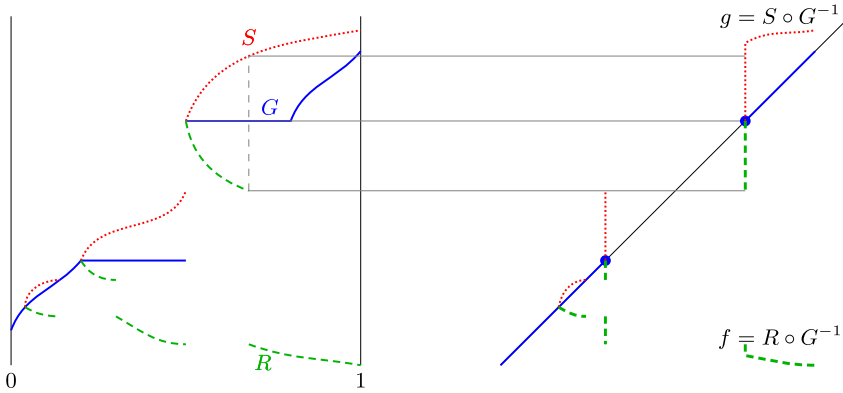


FIG. 1. Sketch of  $R, G, S$  and the corresponding  $f$  and  $g$ . On the atoms of  $\mu$ ,  $G$  is flat, and  $f$  and  $g$  are multivalued, but  $R$  and  $S$  remain well defined.

appropriate generalisations of (4) and (5) are

$$(6) \quad \int_{(R(u), G(u))} \mu(dz) + \bar{\lambda}_u^\mu = \int_{(R(u), S(u))} \nu(dz) + \underline{\lambda}_u^\nu + \bar{\lambda}_u^\nu,$$

$$(7) \quad \int_{(R(u), G(u))} z\mu(dz) + \bar{\lambda}_u^\mu G(u) = \int_{(R(u), S(u))} z\nu(dz) + \underline{\lambda}_u^\nu R(u) + \bar{\lambda}_u^\nu S(u),$$

respectively, where the quantities  $0 \leq \bar{\lambda}_u^\mu \leq \mu(\{G(u)\})$ ,  $0 \leq \underline{\lambda}_u^\nu \leq (\nu - \mu)(\{R(u)\})$ ,  $0 \leq \bar{\lambda}_u^\nu \leq \nu(\{S(u)\})$  are uniquely determined by the triple  $(R, G, S)$ . Essentially, (6) is preservation of mass condition and (7) is preservation of mean condition. Together they give the martingale property.

3.1. *The case where  $\mu$  is a point mass.* The goal in this section is to prove Theorem 1 in the special case where  $\mu$  is a point mass. We assume that  $\mu$  is a unit atom at  $w$  and  $\nu$  is centred at  $w$ . Then  $\mu = \delta_w \leq_{cx} \nu$ .

Let  $P(k) = P_\nu(k) = \int_{-\infty}^\infty (k - z)^+ \nu(dz)$ . Then  $P(k) \geq (k - w)^+$ . For  $p \in [0, P(w)]$ , define  $\alpha : [0, P(w)] \mapsto [w, \infty]$  and  $\beta : [0, P(w)] \mapsto [-\infty, w]$  by

$$(8) \quad \alpha(p) = \operatorname{arginf}_{k > w} \left\{ \frac{P(k) - p}{k - w} \right\}; \quad \beta(p) = \operatorname{argsup}_{k < w} \left\{ \frac{p - P(k)}{w - k} \right\};$$

see Figure 2. Then  $\alpha$  is decreasing and  $\beta$  is increasing. Since the  $\operatorname{arginf}$  and  $\operatorname{argsup}$  may not be uniquely defined (this happens when  $\nu$  has intervals with no mass) we avoid indeterminacy by assuming that  $\alpha$  and  $\beta$  are right continuous. (We also set  $\alpha(P(w)) = \inf\{z > w : F_\nu(z) > F_\nu(w)\}$  and  $\beta(P(w)) = \sup\{z < w : F_\nu(z) < F_\nu(w-)\}$ . Note that  $\alpha(0) = r_\nu$  and  $\beta(0) = \ell_\nu$ .) If  $\nu$  has atoms, then  $\alpha$  and  $\beta$  may fail to be strictly monotonic.

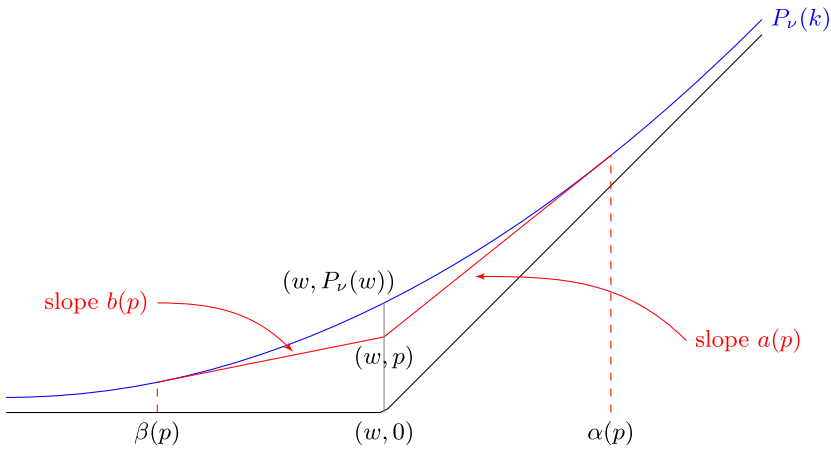


FIG. 2. The definitions of  $\alpha$ ,  $\beta$ ,  $a$  and  $b$ .  $\Upsilon(p)$  is the difference in the slopes of the tangents to  $P_\nu(k)$  which pass through  $(w, p)$ .

For  $p \in (0, P(w))$ , define also

$$(9) \quad \begin{aligned} a(p) &= \inf_{k > w} \frac{P(k) - p}{k - w} = \frac{P(\alpha(p)) - p}{\alpha(p) - w}; \\ b(p) &= \sup_{k < w} \frac{p - P(k)}{w - k} = \frac{p - P(\beta(p))}{w - \beta(p)}. \end{aligned}$$

Extend the representations to  $[0, P(w)]$  by taking limits. Then  $a : [0, P(w)] \mapsto [P'(w+), 1]$  is decreasing and  $b : [0, P(w)] \mapsto [0, P'(w-)]$  is increasing. We have the representations

$$a(p) = 1 - \int_0^p \frac{dq}{\alpha(q) - w}; \quad b(p) = \int_0^p \frac{dq}{w - \beta(q)}.$$

Let  $\Upsilon : [0, P(w)] \mapsto [0, 1]$  be given by  $\Upsilon(p) = a(p) - b(p)$ . Then  $\Upsilon(0) = 1$  and  $\Upsilon(P(w)) = \nu(\{w\})$ .  $\Upsilon$  is a decreasing, concave function which is absolutely continuous on  $[0, P(w))$ . We can define an inverse  $\Upsilon^{-1} : [0, 1] \rightarrow [0, P(w)]$  provided we set  $\Upsilon^{-1}(q) = 1$  for  $q \leq \nu(\{w\})$ . Where  $\alpha$  and  $\beta$  are continuous, we have  $\Upsilon'(p) = -\frac{1}{\alpha(p)-w} - \frac{1}{w-\beta(p)}$ . See Figure 3.

Define  $S : (0, 1) \mapsto \mathbb{R}$  by  $S(u) = (\alpha \circ \Upsilon^{-1})(u)$  and  $R : (0, 1) \mapsto \mathbb{R}$  by  $R(u) = (\beta \circ \Upsilon^{-1})(u)$ .

REMARK 2. If  $\nu$  does not charge an open interval  $A \subset (w, \infty)$ , then  $P$  is linear on  $A$ . Then  $\alpha$  jumps over this set and  $S$  does not take values in  $A$ . Similarly, if  $\nu$  does not charge an open interval  $B \subset (-\infty, w)$ , then  $R$  jumps over this interval.



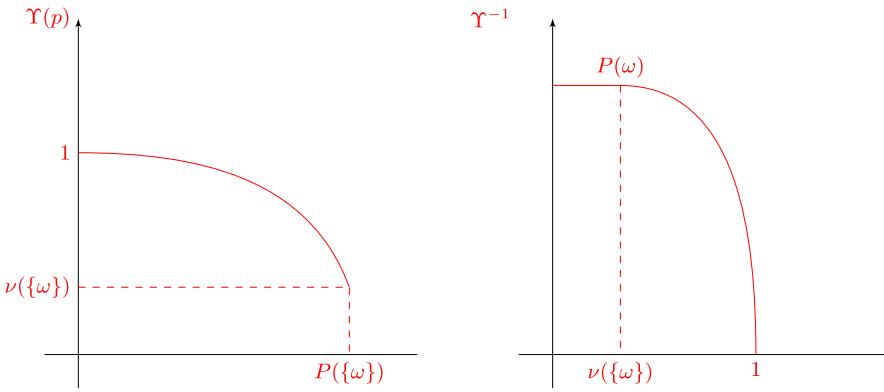


FIG. 3. Sketch of  $\Upsilon$  and  $\Upsilon^{-1}$ .

REMARK 3. By construction,  $\alpha$  and  $\beta$  are both right continuous. Since  $\Upsilon^{-1}$  is continuous and decreasing, it follows that  $R$  and  $S$  are left continuous. Moreover,  $\lim_{u \rightarrow 1} R(u) = \ell_\nu$  and  $\lim_{u \rightarrow 1} S(u) = r_\nu$ .

Let  $Y$  be defined by (3) in Theorem 1. Note that since  $\mu$  is a point mass  $G(u) = w$  for all  $u \in (0, 1)$ .

LEMMA 2. Suppose  $U, V$  are independent uniform random variables. Then  $Y(U, V)$  has law  $\nu$ .

PROOF. Let  $\phi$  be a test function: a continuously differentiable function with support contained in  $[w + \varepsilon, w + \varepsilon^{-1}]$  for some  $\varepsilon \in (0, 1)$ . We will show that  $\mathbb{E}[\phi(Y)] = \int \phi(y)\nu(dy)$ . We can prove a similar result for test functions  $\psi$  with support in  $[w - \varepsilon^{-1}, w - \varepsilon]$ . It follows that  $\mathcal{L}(Y) = \nu$ .

By construction,

$$\begin{aligned} \mathbb{E}[\phi(Y)] &= \int_0^1 du \frac{w - R(u)}{S(u) - R(u)} \phi(S(u)) \\ &= \int_0^1 du \frac{w - \beta \circ \Upsilon^{-1}(u)}{\alpha \circ \Upsilon^{-1}(u) - \beta \circ \Upsilon^{-1}(u)} \phi(\alpha \circ \Upsilon^{-1}(u)) \\ &= \int_0^{P(w)} dp |\Upsilon'(p)| \frac{w - \beta(p)}{\alpha(p) - \beta(p)} \phi(\alpha(p)). \end{aligned}$$

But  $\Upsilon'(p) = -\frac{\alpha(p) - \beta(p)}{(\alpha(p) - w)(w - \beta(p))}$ . Thus, writing  $\psi(y) = \frac{\phi(y)}{(y - w)}$  and using the fact that  $\alpha^{-1}(y) = P(y) - (y - w)P'(y)$  except at the countably many points where

$\alpha^{-1}$  is multivalued,

$$\begin{aligned} \mathbb{E}[\phi(Y)] &= \int_0^{P(w)} dp \frac{\phi(\alpha(p))}{\alpha(p) - w} = - \int_w^\infty d(\alpha^{-1}(y))\psi(y) \\ &= \int_w^\infty [P(y) - (y - w)P'(y)]\psi'(y) dy \\ &= - \int_w^\infty P'(y)[\psi(y) + (y - w)\psi'(y)] dy \\ &= - \int_w^\infty P'(y)\phi'(y) dy = \int \phi(y)\nu(dy). \end{aligned}$$

Hence  $\mathbb{E}[\phi(y)] = \int \phi(y)\nu(dy)$ .  $\square$

REMARK 4. If  $\alpha$  and  $\beta$  are strictly monotonic at  $\Upsilon^{-1}(u)$ , then conditional on  $U \leq u$ ,  $Y$  has law  $\nu$  conditioned to take values in  $[\beta \circ \Upsilon^{-1}(u), \alpha \circ \Upsilon^{-1}(u)]$ . Necessarily,  $\nu([\beta \circ \Upsilon^{-1}(u), \alpha \circ \Upsilon^{-1}(u)]) = u$ .

If there is an atom of  $\nu$  at  $\beta \circ \Upsilon^{-1}(u)$  or  $\alpha \circ \Upsilon^{-1}(u)$ , then we can choose appropriate masses  $\underline{\lambda}_u$  and  $\bar{\lambda}_u$  such that  $\nu((\beta \circ \Upsilon^{-1}(u), \alpha \circ \Upsilon^{-1}(u))) + \underline{\lambda}_u \delta_{\beta \circ \Upsilon^{-1}(u)} + \bar{\lambda}_u \delta_{\alpha \circ \Upsilon^{-1}(u)}$  has total mass  $u$  and mean  $w$ . We must have  $0 \leq \underline{\lambda}_u \leq \nu(\{\beta \circ \Upsilon^{-1}(u)\})$  and  $0 \leq \bar{\lambda}_u \leq \nu(\{\alpha \circ \Upsilon^{-1}(u)\})$ .

On  $U \leq u_1$  let  $Y^{u_1} = Y^{u_1}(U, V)$  be constructed as in (3). On  $U > u_1$ , let  $Y^{u_1}$  be in a graveyard state  $\Delta$ . Then  $\mathcal{L}(Y^{u_1}) = \nu_{u_1} + (1 - u_1)\delta_\Delta$  where  $\nu_{u_1}$  is a measure on  $[R(u_1), S(u_1)]$  with total mass  $u_1$  and mean  $w$ . In particular,  $\nu_{u_1} = \nu$  on  $(R(u_1), S(u_1))$ ,  $\nu_{u_1} \leq \nu$  on  $\{R(u_1), S(u_1)\}$  and  $\nu_{u_1} = 0$  on  $[R(u_1), S(u_1)]^C$ .

3.2. *The case where  $\mu$  consists of a finite number of atoms.* Suppose  $\mu = \sum_{i=1}^N \lambda_i \delta_{x_i}$  where  $x_1 < x_2 < \dots < x_N$  with  $\lambda_i > 0$  and  $\sum_{i=1}^N \lambda_i = 1$ . Suppose  $\nu$  is an arbitrary probability measure satisfying the convex order condition  $\mu \leq_{cx} \nu$ .

For  $0 \leq p \leq P_\nu(x_1)$ , we can construct  $\alpha, \beta, a$  and  $b$  as in (8) and (9) (but relative to  $x_1$  rather than the mean  $w$ ) and set  $\Upsilon = a - b$ . For example,  $\alpha(p) = \operatorname{arginf}_{k > x_1} \frac{P_\nu(k) - p}{k - x_1}$  and  $a(p) = \inf_{k > x_1} \frac{P_\nu(k) - p}{k - x_1}$ . See Figure 4. Note that  $\Upsilon(0) = \Lambda_1 := \inf_{x > x_1} \frac{P_\nu(x)}{x - x_1}$  and since  $P_\nu(x) \geq P_\mu(x) \geq \lambda_1(x - x_1)$  we have  $\Lambda_1 \geq \lambda_1$ . The inverse  $\Upsilon^{-1}$  can be defined on  $[0, \Lambda_1]$ , but we are only interested in  $\Upsilon^{-1}$  over the interval  $[0, \lambda_1]$ . Using  $\Upsilon^{-1}$  and the construction of the previous section we can define  $S = \alpha \circ \Upsilon^{-1} : (0, \lambda_1] \mapsto [x_1, \infty)$  and  $R = \beta \circ \Upsilon^{-1} : (0, \lambda_1] \mapsto (-\infty, x_1]$  with  $S$  increasing and  $R$  decreasing.

By the final comments in Remark 4, the construction of  $R$  and  $S$  on  $(0, \lambda_1]$  is such that if  $Y$  is constructed as in (3), then on  $U \leq \lambda_1$  we find  $Y$  has law  $\nu_{\lambda_1}$ , where  $\nu_{\lambda_1} = \nu$  on  $(R(\lambda_1), S(\lambda_1))$  and  $\nu_{\lambda_1} \leq \nu$  on  $\{R(\lambda_1), S(\lambda_1)\}$ .

We now claim that  $\tilde{\mu}_1 := \mu - \lambda_1 \delta_{x_1} = \sum_{i=2}^N \lambda_i \delta_{x_i}$  and  $\tilde{\nu}_1 = \nu - \nu_{\lambda_1}$  are in convex order. By construction  $\nu_{\lambda_1}$  has mass  $\lambda_1$  and barycentre  $x_1$ . Hence  $\tilde{\mu}_1$  and  $\tilde{\nu}_1$  also have the same total mass and barycentre.

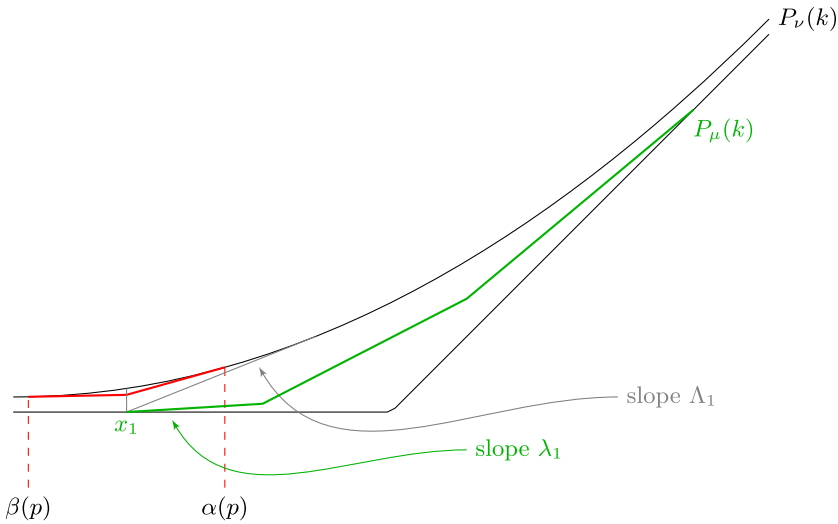


FIG. 4. Calculation of  $\alpha$ ,  $\beta$ ,  $a$  and  $b$  in this case.

LEMMA 3.  $\tilde{\mu}_1 \leq_{cx} \tilde{\nu}_1$ .

PROOF. Let  $\hat{\nu} = \lambda_1 \delta_{x_1} + \tilde{\nu}_1$ . Since  $\lambda_1 \delta_{x_1} \leq_{cx} \nu_{\lambda_1}$  we have  $\hat{\nu} \leq_{cx} \nu$ . Also  $P_\mu(k) \leq P_{\hat{\nu}}(k)$ . To see this, note that  $P_{\hat{\nu}}$  is continuous everywhere and linear on intervals  $[R(\lambda_1), x_1]$  and  $[x_1, S(\lambda_1)]$ , whereas  $P_\mu$  is continuous and convex on  $[R(\lambda_1), S(\lambda_1)]$ . Moreover,  $P_\mu(R(\lambda_1)) = 0 \leq P_{\hat{\nu}}(R(\lambda_1))$ ,  $P_\mu(x_1) = 0 \leq P_{\hat{\nu}}(x_1)$  and  $P_\mu(S(\lambda_1)) \leq P_\nu(S(\lambda_1)) = P_{\hat{\nu}}(S(\lambda_1))$ . Hence  $P_{\tilde{\mu}}(k) + \lambda_1(x_1 - k)^+ = P_\mu(k) \leq P_{\hat{\nu}}(k) = P_{\tilde{\nu}_1}(k) + \lambda_1(x_1 - k)^+$  and it follows that  $P_{\tilde{\mu}}(k) \leq P_{\tilde{\nu}_1}(k)$  as required.  $\square$

We have constructed  $(R, S)$  on  $(0, \lambda_1]$  with  $S$  increasing and  $R$  decreasing in such a way that the point mass at  $x_1$  is mapped to  $\nu_{\lambda_1}$ . It remains to embed  $\tilde{\nu}_1$  starting from  $\tilde{\mu}_1$ . Note that by Remark 4,  $\tilde{\nu}_1$  places no mass on  $(R(\lambda_1), S(\lambda_1))$ .

As a next step, we embed the atom  $\lambda_2 \delta_{x_2}$  of  $\tilde{\mu}_1$  in  $\tilde{\nu}_1$ .  $x_2$  is the lowest location of an atom in  $\tilde{\mu}_1$  so we can use the same algorithm as before. In this way, for  $\lambda_1 < u \leq \lambda_1 + \lambda_2$  we construct  $S$  increasing with  $S(\lambda_1+) \geq S(\lambda_1-) \vee x_2$  and  $R$  decreasing with  $R(\lambda_1+) \leq x_2$ . By Remark 2,  $R$  jumps over the interval  $(R(\lambda_1), S(\lambda_1))$ . We conclude that for  $0 < u < v < \lambda_1 + \lambda_2$ ,  $R(v) \notin (R(u), S(u))$ .

Thereafter, we proceed inductively on the number of atoms which have been embedded. The initial law is a subprobability  $\tilde{\mu}_k = \sum_{k+1}^N \lambda_i \delta_{x_i}$  which we want to map to a target law  $\tilde{\nu}_k$  where  $\tilde{\mu}_k \leq_{cx} \tilde{\nu}_k$  and  $\tilde{\nu}_k \leq \nu$ . Since  $\mu$  consists of a finite number of atoms the construction terminates. Moreover, the random variable  $Y$  we construct in this way has law  $\nu$  and  $R$  and  $S$  have the properties in (2). It follows that we have proved Theorem 1 in the case where  $\mu$  consists of a finite number of atoms.

3.3. *The martingale coupling and its inverse as maps.* Given  $\nu$  centred, (and  $\mu = \delta_0$ ) we saw in Section 3.1 how to construct  $R : (0, 1) \mapsto \mathbb{R}_-$  and  $S : (0, 1) \mapsto \mathbb{R}_+$  such that  $Y = Y(U, V)$  has law  $\nu$  where  $Y$  is given by  $Y(u, v) = 0$  if  $S(u) = 0$  and

$$(10) \quad Y(u, v) = R(u)I_{\{v \leq \frac{S(u)}{S(u)-R(u)}\}} + S(u)I_{\{v > \frac{S(u)}{S(u)-R(u)}\}}$$

otherwise.

Let  $\mathcal{P}^0(\mathbb{R})$  denote the set of centred probability measures on  $\mathbb{R}$ . Let  $\mathcal{V}^1$  denote the set of pairs of functions  $R, S$  with  $R : (0, 1) \rightarrow \mathbb{R}_-$  and  $S : (0, 1) \rightarrow \mathbb{R}_+$ , let  $\mathcal{V}_{\text{Mon}}^1$  denote the subset of  $\mathcal{V}^1$  for which  $R$  is decreasing and  $S$  is increasing and let  $\mathcal{V}_{\text{Int}}^1$  denote the subset of  $\mathcal{V}^1$  such that  $I(R, S) < \infty$  where

$$I(f, g) = \int_0^1 du \frac{|f(u)|g(u)}{g(u) - f(u)} I_{\{g(u) > 0\}}.$$

Finally, let  $\mathcal{V}_{\text{Mon,Int}}^1 = \mathcal{V}_{\text{Mon}}^1 \cap \mathcal{V}_{\text{Int}}^1$ .

The construction in Section 3.1 can be considered as a pair of maps

$$\begin{aligned} \mathcal{Q}^1 : \mathcal{P}^0(\mathbb{R}) &\mapsto \mathcal{V}_{\text{Mon,Int}}^1, \\ \mathcal{R}^1 : \mathcal{V}_{\text{Mon,Int}}^1 &\mapsto \mathcal{P}^0(\mathbb{R}). \end{aligned}$$

Note that  $\mathbb{E}[|Y|] = 2I(R, S)$  which can be shown using the ideas in the proof of Lemma 2 to be equal to  $2P_\nu(0)$ . Moreover, under  $I(R, S) < \infty$  we have  $\mathbb{E}[Y] = 0$ .

Note that if we take  $(R, S) \in \mathcal{V}_{\text{Mon}}^1 \setminus \mathcal{V}_{\text{Mon,Int}}^1$  then we can still define  $Y$  via (10) but  $\mathcal{L}(Y)$  will not be integrable. Then  $M$  given by  $M_1 = 0, M_2 = Y$  is a local martingale, but not a martingale.

Section 3.2 extends these results from initial laws which consist of a single atom to finite combinations of atoms. Let  $\mathcal{P}_F^0(\mathbb{R})$  be the subset of  $\mathcal{P}^0(\mathbb{R})$  for which the measure consists of a finite set of atoms and let  $\mathcal{C}_F = \{(\zeta, \chi) : \zeta \in \mathcal{P}_F^0(\mathbb{R}), \chi \in \mathcal{P}^0(\mathbb{R}); \zeta \leq_{\text{cx}} \chi\}$ . Let

$$\begin{aligned} \mathcal{V} = \Big\{ &(R, G, S); R : (0, 1) \rightarrow \mathbb{R}, G : (0, 1) \rightarrow \mathbb{R}, \\ &S : (0, 1) \rightarrow \mathbb{R}; R(u) \leq G(u) \leq S(u); \\ &\int_0^1 |G(u)| du < \infty, \int_0^1 G(u) du = 0 \Big\}. \end{aligned}$$

Consider now the subsets

$$\begin{aligned} \mathcal{V}_F &= \{(R, G, S) \in \mathcal{V} : G \text{ nondecreasing and takes only finitely many values}\}, \\ \mathcal{V}_{\text{Mon}} &= \{(R, G, S) \in \mathcal{V} : (2) \text{ holds}\}, \\ \mathcal{V}_{\text{Int}} &= \{(R, G, S) \in \mathcal{V} : I(R, G, S) < \infty\}, \end{aligned}$$

where  $I(R, G, S) = \int_0^1 du \frac{(S(u)-G(u))(G(u)-R(u))}{S(u)-R(u)} I_{\{S(u)>G(u)\}}$ , and consider also intersections of these subsets, for example,  $\mathcal{V}_{\text{Mon,Int}} = \mathcal{V}_{\text{Mon}} \cap \mathcal{V}_{\text{Int}}$ . In Section 3.2, we constructed a map  $Q : \mathcal{C}_F \rightarrow \mathcal{V}_{F,\text{Mon}}$  which we write as  $Q(\zeta, \chi) = (R_{(\zeta, \chi)}, G_\zeta, S_{(\zeta, \chi)})$ . Indeed, since  $\chi \in L^1$  and since  $\mathbb{E}[|Y - X|] \leq \mathbb{E}[|X|] + \mathbb{E}[|Y|] < \infty$  we have that  $\mathbb{E}[|Y - X|] = 2I(R_{(\zeta, \chi)}, G_\zeta, S_{(\zeta, \chi)})$ , so that we actually have a map  $Q : \mathcal{C}_F \rightarrow \mathcal{V}_{F,\text{Mon,Int}}$ . Conversely, the arguments after Lemma 3 show that (3) defines an inverse map  $\mathcal{R} : \mathcal{V}_{F,\text{Mon,Int}} \rightarrow \mathcal{C}_F$ .

Note that given any element  $(R, G, S)$  of  $\mathcal{V}$  we can define the map  $\mathcal{R} : \mathcal{V} \rightarrow \mathcal{P}^0(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$  via  $\mathcal{R}(R, G, S) = (\mathcal{L}(X(U)), \mathcal{L}(Y(U, V)))$  where  $Y(U, V)$  is as given in the statement of Theorem 1. We will make no further use of this idea, but different properties of  $(R, G, S)$  will lead to different (local)-martingale couplings. The embedding of Hobson and Neuberger [13] is of this type. In the Hobson and Neuberger embedding,  $R$  and  $S$  are both increasing.

3.4. *The case of general integrable  $\mu$ .* We assume  $\mu$  is centred at zero, but the general case follows by translation.

Our goal in this section is to extend the map  $Q : \mathcal{C}_F \rightarrow \mathcal{V}_{F,\text{Mon,Int}}$  with inverse  $\mathcal{R}$  to a map  $Q : \mathcal{C} \rightarrow \mathcal{V}_{\text{Mon,Int}}$  where  $\mathcal{C} = \{(\zeta, \chi) : \zeta \in \mathcal{P}^0(\mathbb{R}), \chi \in \mathcal{P}^0(\mathbb{R}); \zeta \leq_{\text{cx}} \chi\}$ . For  $\mu$  a general centred probability measure and  $\nu$  a centred target measure with  $\mu \leq_{\text{cx}} \nu$ , we construct a sequence  $(\mu_n)_{n \geq 1}$  of approximations of  $\mu$  by elements of  $\mathcal{P}_F^0(\mathbb{R})$ . For each  $\mu_n$ , we can construct a triple  $(R_n, G_n, S_n)$ . We show that  $(R_n, G_n, S_n)_{n \geq 1}$  converge to a limit  $(R, G, S)$  first on the rationals and then (almost surely) on  $(0, 1)$ . Convergence of  $G_n$  and  $S_n$  is straightforward, but convergence of  $R_n$  is more subtle, and indeed we only have convergence on  $\{u : S(u) > G(u)\}$ . Finally, we show that  $\mathcal{R}(R, G, S) = (\mu, \nu)$  so that the trio  $(R, G, S)$  defines a martingale coupling between  $\mu$  and  $\nu$ .

Let  $\{q_1, q_2, \dots\}$  be an enumeration of  $\mathbb{Q} \cap (0, 1)$ . Then  $\{S_n(q_1)\}_{n \geq 1}$  converges down a subsequence  $n_{k_1}$  to a limit  $S_\infty(q_1) := \lim_{k_1 \uparrow \infty} S_{n_{k_1}}(q_1)$ . Down a further subsequence if necessary, we have that  $S_{n_{k_2}}(q_2)$  converges to  $S_\infty(q_2)$ . Proceeding inductively, we have by a diagonal argument (see, e.g., Billingsley [6]) that there is a subsequence  $(m_1, m_2, \dots)$  such that  $\{S_{m_k}\}_{k \geq 1}$  converges to  $S_\infty$  at every rational  $q \in \mathbb{Q} \cap (0, 1)$ . This limit is nondecreasing.

Our first result shows that any limit of  $S_n$  is finite valued. Since the ideas behind the proof are not relevant to the arguments of this section the proof is postponed to the Appendix.

LEMMA 4. *Let  $\mu_n \uparrow_{\text{cx}} \mu$ . Then  $\limsup S_n(u) \leq J(u)$  for some function  $J = J_{\mu, \nu} : (0, 1) \mapsto (-\infty, \infty)$ .*

We want to extend the domain from the rationals to  $(0, 1)$ . To this end, define  $S(u) = \lim_{q_j \uparrow u} S_\infty(q_j)$ . This limit is well defined (and nondecreasing) by the monotonicity of  $S_\infty$ . Then from the monotonicity of  $S$ , we conclude that  $S$  has only countably many discontinuities. Note that, by definition,  $S$  is left continuous.

We can construct  $G$  from  $\{G_n\}$  in an identical fashion. In this case, the finiteness of the limit follows from the tightness of the singleton  $\{G\}$ . Moreover, since  $G_n \leq S_n$  by construction, we have  $G_\infty \leq S_\infty$  and  $G \leq S$ . Again, the increasing limit  $G$  has at most countably many discontinuities and is left continuous.

Define  $\mathcal{N}_S = \{u : S_n(u) \not\rightarrow S(u)\}$  and  $\mathcal{N}_G = \{u : G_n(u) \not\rightarrow G(u)\}$  where the subscript  $n$  refers to a subsequence down which  $S_n$  and  $G_n$  converge on rationals. Define also  $\mathcal{N}_S^\Delta = \{u : S(u+) > S(u-)\}$  and  $\mathcal{N}_G^\Delta = \{u : G(+) > G(u-)\}$ .

LEMMA 5.  $\mathcal{N}_S \subseteq \mathcal{N}_S^\Delta$  and  $\mathcal{N}_G \subseteq \mathcal{N}_G^\Delta$ . Moreover,  $\text{Leb}(\mathcal{N}_S \cup \mathcal{N}_G) = 0$ .

PROOF. Suppose  $u$  is a continuity point of  $S$ . Suppose further that there is a subsequence  $(n_j)_{j \geq 1}$  along which  $S_{n_j}(u) > S(u) + \varepsilon$ . Using the continuity of  $S$  at  $u$ , we may pick  $q > u$  such that  $S(q) < S(u) + \varepsilon/2$ . Take  $q_k \in (u, q)$  with  $q_k \downarrow u$ . Then  $S_{n_j}(q_k) \geq S_{n_j}(u) > S(u) + \varepsilon > S(q_k) + \varepsilon/2$ . Letting  $j \uparrow \infty$ ,  $S_\infty(q_k) > S(q_k) + \varepsilon/2$ . Letting  $k \uparrow \infty$ ,  $S(u) \geq S(u) + \varepsilon/2$  which is a contradiction.

A similar argument (without the need of continuity at  $u$ ) shows that down any subsequence  $\lim_j S_{n_j}(u) > S(u) - \varepsilon$ . Hence, if  $S(u) = S(u+)$  then  $S(u) = \lim S_n(u)$ . Since the set of points for which  $S(u+) > S(u)$  is countable we conclude that  $\text{Leb}(\mathcal{N}_S) = 0$ .

An identical argument gives that  $G(u) = \lim_n G_n(u)$  on  $G(u+) = G(u)$  and  $\text{Leb}(\mathcal{N}_G) = 0$ .  $\square$

Now consider  $(R_n)_{n \geq 1}$  and the existence of a possible limit  $R$ . By the same diagonal argument as above, we can define  $R_\infty : \mathbb{Q} \cap (0, 1) \rightarrow \mathbb{R}$  such that on a subsequence  $R_{n_k}(q) \rightarrow R_\infty(q) \in [-\infty, \infty]$  for every  $q$ . (From now on, we work on a subsequence indexed  $n$  such that  $\{S_n\}_n, \{G_n\}_n$  and  $\{R_n\}_n$  converge for every  $q \in \mathbb{Q} \cap (0, 1)$ .) We want to construct  $R$  from  $R_\infty$ , but unlike in the case of  $S$  or  $G$  we do not have monotonicity. Note that for  $q' > q$  we have  $R_n(q') \notin (R_n(q), S_n(q))$  for each  $n$  and this implies  $R_\infty(q') \notin (R_\infty(q), S_\infty(q))$ .

The following lemma shows that  $R_\infty$  is finite valued, at least for  $q$  such that  $G(u+) < S(u)$ .

LEMMA 6. Let  $\mu_n \uparrow_{\text{cx}} \mu$ . Then  $\liminf R_n(u) \geq j(u)$  on  $G(u+) < S(u)$  for some function  $j = j_{\mu, \nu} : (0, 1) \mapsto (-\infty, \infty)$ .

Let  $\mathcal{A} = \{u \in (0, 1) : G(u+) < S(u)\}$ . By the above lemma,  $R_\infty(q) > j(q) > -\infty$  for  $q \in \mathcal{A}$ . If  $u \in \mathcal{A}$  then the left continuity of  $S$  implies that there exists an interval  $(u - \varepsilon, u] \subseteq \mathcal{A}$ ; since every such interval must contain a rational we have that  $\mathcal{A}$  is a countable union of intervals.

We now show that  $R_\infty$  is decreasing on each such interval. Suppose not. Then there exists  $q < q'$  in the same interval  $I$  with  $R_\infty(q') > R_\infty(q)$ . Let  $v = \inf_{q'' \in \mathbb{Q} \cap I} \{q'' : R_\infty(q'') > R_\infty(q)\}$ . Choose  $\tilde{q}_m \uparrow v$  with  $\tilde{q}_m \geq q$  and  $\hat{q}_n \downarrow v$  with  $R_\infty(\hat{q}_n) > R_\infty(q)$ . Then  $R_\infty(\hat{q}_n) \notin (R_\infty(\tilde{q}_m), S_\infty(\tilde{q}_m))$ , and since  $R_\infty(\hat{q}_n) >$

$R_\infty(q) \geq R_\infty(\tilde{q}_m)$  we conclude  $R_\infty(\hat{q}_n) \geq S_\infty(\tilde{q}_m)$ . Letting  $n$  tend to infinity we conclude  $\liminf R_\infty(\hat{q}_n) \geq S_\infty(\tilde{q}_m)$ , and letting  $m$  tend to infinity  $\liminf_{n \uparrow \infty} R_\infty(\hat{q}_n) \geq S(v)$ . However,  $R_\infty(\hat{q}_n) \leq G_\infty(\hat{q}_n)$  and hence  $\limsup_{n \uparrow \infty} R_\infty(\hat{q}_n) \leq G(v+) < S(v)$ . These two statements are inconsistent, and hence  $R_\infty$  must be decreasing on each interval of  $\mathcal{A}$ .

Given that  $R_\infty$  is decreasing on each interval of  $\mathcal{A}$ , we can define  $R$  on  $\mathcal{A}$  by  $R(u) = \lim_{q \uparrow u} R_\infty(q)$ . Then the function  $R$  is decreasing and, therefore, has only countably many discontinuities in any interval of  $\mathcal{A}$ . Away from these discontinuities, we have  $R_n(u) \rightarrow R(u)$  by an argument similar to that in Lemma 5.

Define  $\mathcal{B}_= = \{u \in (0, 1) : G(u) = S(u)\}$  and  $\mathcal{B}_< = \{u \in (0, 1) : G(u) < S(u)\}$ . Then  $\mathcal{B}_< = \mathcal{A} \cup \mathcal{C}$  where  $\mathcal{C} = \{u \in (0, 1) : G(u) < S(u) \leq G(u+)\}$ . Since  $\mathcal{C} \subseteq \mathcal{N}_G^\Delta$ , we have that  $\mathcal{B}_<$  and  $\mathcal{A}$  differ by a set of measure zero and we conclude the following.

LEMMA 7.  $I_{\{u \in \mathcal{B}_<\}}(R_n(u) - R(u)) \rightarrow 0$ , except on a set of measure zero.

Note that we cannot expect  $R_n(u)$  to converge on  $\mathcal{B}_=$ .

It remains to define  $R$  on  $\mathcal{B}_=$  and  $\mathcal{C}$  in such a way that  $R$  satisfies (2). On  $\mathcal{B}_=$ , we set  $R(u) = G(u) = S(u)$ . For  $u \in \mathcal{C}$ , we have by the left continuity of  $S$  that there exists  $\varepsilon > 0$  such that  $I = (u - \varepsilon, u) \subset \mathcal{A}$ . By the same arguments as before, we conclude that  $R_\infty$  is decreasing on  $I$  and we set  $R(u) = \lim_{q \uparrow u} R_\infty(q)$ . Indeed, for  $u \in \mathcal{B}_<$  we have  $R(u) = \lim_{q \uparrow u} R_\infty(q)$ . Note that for  $u \in \mathcal{C}$  we may have that  $R(u+) > R(u)$  and it is not true in general that  $R$  is decreasing on intervals contained in  $\mathcal{B}_<$ .

Fix  $u < v$ . If  $u$  or  $v$  is in  $\mathcal{B}_=$  then since we have defined  $R(w) = G(w) = S(w)$  on  $\mathcal{B}_=$  we trivially have  $R(v) \notin (R(u), S(u))$ . For  $u, v \in \mathcal{B}_<$  choose sequences  $\{q_m\}_m$  with  $q_m < u$  and  $q_m \uparrow u$  and  $\{q_l\}_l$  with  $q_l \in (u, v)$  and  $q_l \uparrow v$ . Then  $R_n(q_l) \notin (R_n(q_m), S_n(q_m))$ , and hence  $R_\infty(q_l) \notin (R_\infty(q_m), S_\infty(q_m))$ . Letting  $l \uparrow \infty$ , we have  $R(v) \notin (R_\infty(q_m), S_\infty(q_m))$ , and letting  $m \uparrow \infty$ , we have  $R(v) \notin (R(u), S(u))$ . Hence,  $(R, G, S)$  satisfy (2).

On the space  $\{(r, g, s); r \leq g \leq s\} \subseteq \mathbb{R}^3$ , define  $\Theta^x = \Theta^x(r, g, s)$  by  $\Theta^x(r, g, s) = I_{\{r \leq x < s\}} \frac{s-g}{s-r}$  with the convention that  $\Theta^x(r, g, s) = 0$  for  $g = s$ . In particular,  $\Theta^x(g, g, g) = 0$ .

PROPOSITION 1. *If  $x$  is such that  $\text{Leb}(\{u : S(u) = x\} \cup \{u : R(u) = x; S(u) > G(u)\}) = 0$ , then we have*

$$\begin{aligned}
 & \int_0^1 du \{I_{\{S_n(u) \leq x\}} + \Theta^x(R_n(u), G_n(u), S_n(u))\} \\
 (11) \quad & \rightarrow \int_0^1 du \{I_{\{S(u) \leq x\}} + \Theta^x(R(u), G(u), S(u))\}.
 \end{aligned}$$

PROOF. Since  $S_n(u) \rightarrow S(u)$  almost surely and since  $\int_0^1 du I_{\{S(u)=x\}} = 0$  by hypothesis, we have  $\int_0^1 du I_{\{S_n(u)\leq x\}} \rightarrow \int_0^1 du I_{\{S(u)\leq x\}}$  by bounded convergence.

Let  $\Omega_< = \{u : S_n(u) \rightarrow S(u), G_n(u) \rightarrow G(u), R_n(u) \rightarrow R(u), G(u) < S(u)\}$  and  $\Omega_ = \{u : S_n(u) \rightarrow S(u), G_n(u) \rightarrow G(u), G(u) = S(u)\}$ . By Lemmas 5 and 7,  $\text{Leb}(\Omega_< \cup \Omega_ ) = 1$ .

Now let  $\Omega_<^x = \{u : S_n(u) \rightarrow S(u) \neq x, G_n(u) \rightarrow G(u), R_n(u) \rightarrow R(u) \neq x, G(u) < S(u)\}$  and  $\Omega_ =^x = \{u : S_n(u) \rightarrow S(u) \neq x, G_n(u) \rightarrow G(u), G(u) = S(u)\}$ . By the hypothesis on  $x$ , we still have that  $\text{Leb}(\Omega_<^x \cup \Omega_ =^x) = 1$ , and by bounded convergence the result of the proposition will follow if we can show that  $\Theta^x(R_n, G_n, S_n) \rightarrow \Theta^x(R, G, S)$  on  $\Omega_<^x \cup \Omega_ =^x$ .

This is immediate on  $\Omega_<^x$ . On  $\Omega_ =^x$  we need only note that

$$\begin{aligned} \Theta^x(R_n, G_n, S_n) &= I_{\{R_n \leq x < S_n\}} \frac{(S_n - G_n)}{(S_n - R_n)} \leq \frac{(S_n - G_n)}{(S_n - x)} I_{\{S_n > x\}} \rightarrow 0 \\ &= \Theta^x(R, G, S). \end{aligned} \quad \square$$

PROOF OF THEOREM 1. All that remains to show is that  $(R, G, S)$  embeds  $\nu$ .

There are at most countably many  $x$  for which  $\text{Leb}(\{u : S(u) = x\}) + \text{Leb}(\{u : R(u) = x; S(u) > G(u)\}) > 0$ . Hence it is sufficient to prove that  $\int_0^1 du \{I_{\{S(u)\leq x\}} + I_{\{R(u)\leq x < S(u)\}} \frac{S(u) - G(u)}{S(u) - R(u)}\} = \nu((-\infty, x])$  outside this set. For such an  $x$ , (11) holds. Then, since  $(R_n, G_n, S_n)$  embeds  $\nu$  from  $\mu_n$ ,

$$\begin{aligned} &\int_0^1 du \left\{ I_{\{S(u)\leq x\}} + I_{\{R(u)\leq x < S(u)\}} \frac{S(u) - G(u)}{S(u) - R(u)} \right\} \\ &= \lim_n \left\{ \int_0^1 du \left\{ I_{\{S_n(u)\leq x\}} + I_{\{R_n(u)\leq x < S_n(u)\}} \frac{S_n(u) - G_n(u)}{S_n(u) - R_n(u)} \right\} \right\} \\ &= \lim_n \nu((-\infty, x]) = \nu((-\infty, x]) \end{aligned}$$

as required.  $\square$

We would like to thank the anonymous referee for the following idea for an alternative proof of Theorem 1.

REMARK 5 (Alternative construction). Let  $(\pi_{lc}^x)_{x \in \mathbb{R}}$  be the disintegration of  $\pi_{lc}$  with respect to  $\mu$ , so that  $\pi_{lc}(dx, dy) = \mu(dx)\pi_{lc}^x(dy)$ . It follows that for any  $\mu' \leq \mu$ ,  $\pi'(dx, dy) := \mu'(dx)\pi_{lc}^x(dy)$  is again a left-curtain coupling. Decompose  $\mu = \mu_c + \sum_n \alpha_n \delta_{x_n}$  into continuous and discrete parts, respectively. The desired representation of  $\pi_{lc}$  through graphs of functions can then be obtained by pasting together the representations of  $\pi_c(dx, dy) := \mu_c(dx)\pi_{lc}^x(dy)$  and  $\pi_d(dx, dy) := \sum_n \alpha_n \delta_{x_n}(dx)\pi_{lc}^{x_n}(dy)$ . Note that in the case of  $\pi_c$ , the result of Theorem 1 follows from the original theorem of Beiglböck and Juillet [5], while the case of  $\pi_d$  follows from the arguments given in Section 3.



**4. Robust bounds for the American put.** Our motivation for the study of the left-curtain mapping came from a connection with the robust pricing of American puts. In robust or model-independent pricing (Hobson [12, 15]), the idea is that instead of writing down a model for the asset price (e.g., geometric Brownian motion or a stochastic volatility model) we consider the class of all models for which the discounted asset price is a martingale and which are consistent with the prices of traded vanilla options. Then, given an exotic option which we would like to price, we search over this class of models to find the range of feasible model-based prices.

Typically, the set of traded vanilla options is taken to be the set of European-style puts and calls. Given a family of European puts and calls for a fixed maturity and a continuum of strikes, we can infer the law of the asset price at that maturity (under the market measure used for pricing). Given the prices of puts and calls for a sequence of maturities, we can infer the marginal distributions of the asset price, but not the joint distributions. Then, working under the bond-price numeraire, the class of asset price processes which are consistent with the prices of traded vanilla options can be identified with the class of martingales with given marginals. The problem of finding the robust upper bound on the price of an American-style option becomes a search over consistent martingale models of the model-based price of the American option; see Neuberger [19], Hobson and Neuberger [14, 16] and Bayraktar et al. [1]. Crucially, the primal pricing problem can be identified with a dual hedging problem.

When the American-style option is an American put and the number of candidate exercise dates is two, Hobson and Norgilas [17] solve for the robust upper bound under an assumption that the law of the underlying at the first exercise date is continuous. It turns out that the consistent model for which the American put has highest price is the model associated with the left-curtain coupling of Beiglböck and Juillet [5]. Here, we briefly explain how the results of Hobson and Norgilas extend to the atomic case, and why the atomic case is important. There is a subtlety in the case with atoms which is not present when there are no atoms, and to deal with this subtlety we need the extension of the left-curtain coupling to the atomic case as constructed in this paper.

We are interested in pricing the American put which, in discounted units has strike  $K_1$  at maturity 1 and strike  $K_2$  at maturity 2, with  $K_2 < K_1$ ; see Hobson and Norgilas [17]. The expected payoff arising from a given joint law  $\pi \in \Pi_M(\mu, \nu)$  and a given stopping rule  $\tau$  taking values in  $\{1, 2\}$  is

$$\phi_\pi(\tau) = \mathbb{E}^{\mathcal{L}(X_1, X_2) \sim \pi} [(K_1 - X_1)^+ I_{\{\tau=1\}} + (K_2 - X_2)^+ I_{\{\tau=2\}}].$$

Here,  $X$  represents the discounted asset price, and is a martingale with joint law  $\pi$ .

For a Borel set  $B$ , we can let  $\tau_B$  be the stopping rule  $\tau_B = 1$  if  $X_1 \in B$  and  $\tau_B = 2$  otherwise. Then the payoff under the stopping rule  $\tau_B$  is  $\Phi_\pi(B) := \phi_\pi(\tau_B)$  and the American put price under the model is  $\overline{\Phi}_\pi = \sup_B \Phi_\pi(B)$ .

Bayraktar et al. [1]<sup>1</sup> define the upper bound on the price of the American put to be

$$\mathcal{P}_{\text{BHZ}} = \sup_{\pi \in \Pi_M(\mu, \nu)} \bar{\Phi}_\pi = \sup_{\pi \in \Pi_M(\mu, \nu)} \sup_B \Phi_\pi(B).$$

The definition of the model-independent upper bound on the price of the American put given by Neuberger [19] and Hobson and Neuberger [16] is different. Suppose  $(\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), X = (X_0, X_1, X_2))$  is a  $(\mu, \nu)$ -consistent model. The model-based price of the American put is

$$\mathcal{A}(\mathcal{S}, X) = \sup_{\tau \in \mathcal{T}_{1,2}(\mathcal{S})} \mathbb{E}^{\mathcal{S}, X}[(K_\tau - X_\tau)^+],$$

where  $\mathcal{T}_{1,2}(\mathcal{S})$  is the set of all  $\mathbb{F}$ -stopping times taking values in  $\{1, 2\}$ . Then (Neuberger [19], Hobson and Neuberger [16]) the highest model-based price is

$$(12) \quad \mathcal{P}_N = \sup_{\mathcal{S}, X} \mathcal{A}(\mathcal{S}, X),$$

where the supremum is taken over  $(\mu, \nu)$ -consistent models.

Set  $\bar{\Omega} = \mathbb{R} \times \mathbb{R} = \{\omega = (\omega_1, \omega_2)\}$ ,  $\bar{\mathcal{F}} = \mathcal{B}(\Omega)$  and  $(X_1(\omega), X_2(\omega)) = (\omega_1, \omega_2)$ , and let  $\bar{\mathbb{P}}$  be such that  $\mathcal{L}(X_1) = \mu$  and  $\mathcal{L}(X_2) = \nu$ . Let  $\bar{\mathcal{F}}_0 = \{\emptyset, \Omega\}$ ,  $\bar{\mathcal{F}}_1 = \sigma(X_1)$  and  $\bar{\mathcal{F}}_2 = \sigma(X_1, X_2)$ . If  $\bar{\mathcal{S}} = (\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{F}}, \bar{\mathbb{P}})$ , then  $(\bar{\mathcal{S}}, \bar{X})$  is a  $(\mu, \nu)$ -consistent model.

Consistent models of the form  $(\bar{\mathcal{S}}, \bar{X})$  can be identified with martingale couplings  $\pi$ . It follows that  $\mathcal{P}_{\text{BHZ}} \leq \mathcal{P}_N$ , the inequality following from the fact that in principle we could work on a richer probability space. It follows from the work of Hobson and Norgilas [17] that if  $\mu$  is continuous then the martingale coupling associated with the optimiser for either  $\mathcal{P}_{\text{BHZ}}$  or  $\mathcal{P}_N$  is the left-curtain coupling and  $\mathcal{P}_{\text{BHZ}} = \mathcal{P}_N$ . Our interest in extending the left-curtain mapping arose from the fact that when  $\mu$  has atoms we may have  $\mathcal{P}_{\text{BHZ}} < \mathcal{P}_N$ . Then, in order to construct the optimiser for  $\mathcal{P}_N$  we need an appropriate extension of the left-curtain coupling.

4.1. *The trivial law for  $\mu$ .* The difference between the modelling approaches of Bayraktar et al. [1] and Hobson and Neuberger [16] can be illustrated most simply when  $\mu = \delta_w$ . Also for simplicity, we assume  $\nu$  has a continuous law with mean  $w$ .

In the framework of Bayraktar et al. [1], since the filtration generated by  $X$  is still trivial at time 1, the only choices facing the holder of the American put are either to always stop at time 1, or to never stop at time 1. The expected payoff of the American put does not depend on the martingale coupling, and thus

$$\begin{aligned} \mathcal{P}_{\text{BHZ}} &= \sup_{\pi \in \Pi_M(\mu, \nu)} \max\{\Phi_\pi(\Omega), \Phi_\pi(\emptyset)\} = \sup_{\pi \in \Pi_M(\mu, \nu)} \max\{\phi_\pi(1), \phi_\pi(2)\} \\ &= \max\left\{(K_1 - w)^+, \int (K_2 - z)^+ \nu(dz)\right\}. \end{aligned}$$

---

<sup>1</sup>[1] contains many interesting and important results and this is just a small element of the paper.

On the other hand, we can construct a richer model which is  $(\delta_w, \nu)$  consistent. Set  $\Omega = (0, 1) \times (0, 1)$  and let  $\mathbb{P}$  be Lebesgue measure on  $\Omega$ . Let  $(U, V)$  be a pair of independent uniform random variables, let  $(\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_1 = \sigma(U), \mathcal{F}_2 = \sigma(U, V))$  and let  $X_0 = X_1 = w$  and  $X_2 = Y$ , where  $Y = Y(U, V)$  is as given in (3) with  $G(u) \equiv w$ . Here,  $(R, S)$  are a pair of monotonic functions with

$$(13) \quad u = \int_{R(u)}^{S(u)} \nu(dz), \quad 0 = \int_{R(u)}^{S(u)} (z - w)\nu(dz).$$

In this way, we construct a  $(\mu, \nu)$ -consistent model.

Under this model, the value  $A(u)$  of the American put under the stopping rule  $\tau_u$  where  $\tau_u = 1$  if  $U < u$  and  $\tau_u = 2$  otherwise is

$$\begin{aligned} A(u) &= \mathbb{E}[(K_1 - X_1)^+ I_{\{\tau_u=1\}} + (K_2 - X_2)^+ I_{\{\tau_u=1\}}] \\ &= (K_1 - w)u + \int_{-\infty}^{R(u)} (K_2 - z)^+ \nu(dz) + \int_{S(u)}^{\infty} (K_2 - z)^+ \nu(dz). \end{aligned}$$

It follows that  $\mathcal{P}_N \geq \sup_{u \in [0,1]} A(u)$ . (In the next section, we will argue that there is equality here.) Note that  $\mathcal{P}_{\text{BHZ}} = A(0) \vee A(1)$ , so that  $\mathcal{P}_N > \mathcal{P}_{\text{BHZ}}$  will follow if  $\sup_{u \in [0,1]} A(u) > A(0) \vee A(1)$ .

For a simple example, suppose  $w = 1$  and  $\nu = U[0, 2]$ ; suppose  $K_1 = \frac{5}{4}$  and  $K_2 = 1$ . Then  $R(u) = 1 - u$  and  $S(u) = 1 + u$ . We have

$$A(u) = \frac{u}{4} + \int_0^{1-u} (1 - z) \frac{dz}{2} = \frac{1 + u - u^2}{4}.$$

Then  $\mathcal{P}_N \geq \max_{u \in [0,1]} A(u) = \frac{5}{16} > \frac{1}{4} = A(0) \vee A(1) = \mathcal{P}_{\text{BHZ}}$ .

REMARK 6. In our set-up, there are two possible exercise times for the American put, denoted 1 and 2, and we construct a martingale  $(X_0 = w, X_1, X_2)$  to match the marginals at these times. But if  $\mathcal{L}(X_1) = \delta_{X_0}$  the problem can be recast as a problem for a stochastic process  $\tilde{X} = (\tilde{X}_0, \tilde{X}_1)$  where  $\tilde{X}_0 = X_0 = X_1$  and  $\tilde{X}_1 = X_2$ . We also set  $\tilde{\tau} = \tau - 1$ ; then  $\tilde{\tau} \in \{0, 1\}$  and  $\tilde{\tau} = 0$  corresponds to immediate exercise. Put another way, one way to allow for immediate exercise of the American put, is to introduce an additional point (labelled 1) into the time-indexing set and to require  $\mathcal{L}(X_1) = \delta_{X_0}$ . For this reason, it is very natural for  $\mu$  to have a trivial law, if we want to allow immediate exercise.

4.2. *Tightness of the bound for a trivial law  $\mu$ .* Our goal in this section is to show that  $\mathcal{P}_N = \sup_{u \in [0,1]} A(u)$ . We do this by finding an upper bound on the American put pricing problem and then showing that this bound is equal to  $\sup_{u \in [0,1]} A(u)$ .

Let  $\psi$  be a convex function with  $\psi(z) \geq (K_2 - z)^+$ . Let  $\phi(z) = ((K_1 - z)^+ - \psi(z))^+$  and let  $\theta(z) = -\psi'_+(z)$ , where  $\psi'_+$  is the right derivative. Then, for all  $x_1$

and  $x_2$ ,

$$\begin{aligned} (K_1 - x_1)^+ &\leq \phi(x_1) + \psi(x_2) + (x_2 - x_1)\theta(x_1), \\ (K_2 - x_2)^+ &\leq \phi(x_1) + \psi(x_2). \end{aligned}$$

It follows that for any set  $B \in \mathcal{F}$  and for every  $\omega$ ,

$$(K_1 - X_1)^+ I_B + (K_2 - X_2)^+ I_{B^c} \leq \phi(X_1) + \psi(X_2) + (X_2 - X_1)\theta(X_1) I_B.$$

In particular, if we think of  $B$  as the set of scenarios on which the put is exercised at time 1 then we have that the payoff of the American put is bounded above by the sum of the European-style payoffs  $\phi$  and  $\psi$  and the gains from trade from a strategy which involves holding  $\theta(X_1)$  units of the underlying over the time-interval  $(1, 2]$ , provided the put was exercised at time 1. Then, for  $B \in \mathcal{F}_1$ ,

$$\begin{aligned} \mathbb{E}[(K_{\tau_B} - X_{\tau_B})^+] &\leq \mathbb{E}[\phi(X_1)] + \mathbb{E}[\psi(X_2)] \\ &= \int ((K_1 - x)^+ - \psi(x))^+ \mu(dx) + \int \psi(y) \nu(dy). \end{aligned}$$

In our context with  $\mu = \delta_w$ , this simplifies to  $((K_1 - w)^+ - \psi(w))^+ + \int \psi(y) \times \nu(dy) =: \mathcal{D}(\psi)$ . Let  $\mathcal{D} = \inf_{\psi} \mathcal{D}(\psi)$  (where the infimum is taken over convex  $\psi$  with  $\psi(z) \geq (K_2 - z)^+$ ).  $\mathcal{D}$  forms an upper bound for the price of the American option under any consistent model, and hence  $\mathcal{P}_N \leq \mathcal{D}$ .

Let  $R$  and  $S$  be defined as in Section 3.1. Let  $P_\nu(z) = \int (z - x)^+ \nu(dx)$ . Then (13) can be rewritten as  $u = P'_\nu(S(u)) - P'_\nu(R(u))$  together with

$$(14) \quad (S(u) - w)P'_\nu(S(u)) - P_\nu(S(u)) = (R - w)P'_\nu(R(u)) - P_\nu(R(u)).$$

Fix  $K_2 < K_1$  with  $K_1 > w$  and define  $\Lambda_w : (-\infty, K_2 \wedge w) \times (K_1, \infty) \mapsto \mathbb{R}$  by

$$\Lambda_w(r, s) = \frac{K_1 - w}{s - w} - \frac{(K_2 - r) - (K_1 - w)}{w - r}.$$

Since  $\nu$  is continuous by assumption,  $R$  and  $S$  are strictly decreasing and strictly increasing, respectively. Define  $u_w = \inf\{u \in (0, 1) : R(u) < K_2 \text{ and } S(u) > K_1\}$ , and for  $u \in (u_w, 1)$  set  $\bar{\Lambda}_w(u) = \Lambda_w(R(u), S(u))$ . It follows that  $\bar{\Lambda}_w$  is strictly decreasing.

Suppose that  $\mathcal{I}_\nu = [\ell_\nu, r_\nu]$  is such that  $\frac{K_1 - w}{r_\nu - w} < \frac{(K_2 - \ell_\nu) - (K_1 - w)}{w - \ell_\nu}$  (e.g., this will follow if  $0 = \ell_\nu < w < r_\nu = \infty$  and  $K_2 > K_1 - w$ ). This assumption is sufficient to guarantee that there exists  $u^* \in (u_w, 1)$  such that  $\bar{\Lambda}_w(u^*) = 0$ . Then  $S^* := S(u^*) > K_1 > K_2 > R(u^*) =: R^*$ . Also  $\bar{\Lambda}_w(u^*) = 0$  implies  $\frac{K_1 - w}{S^* - w} = \frac{K_2 - R^*}{S^* - R^*}$ . For the model constructed in Section 3.1, we have

$$\begin{aligned} \sup_{u \in [0, 1]} A(u) &\geq A(u^*) = (K_1 - w)^+ u^* + \int_{-\infty}^{R^*} (K_2 - z)^+ \nu(dz) \\ &= (K_1 - w)[P'_\nu(S^*) - P'_\nu(R^*)] \\ &\quad + P_\nu(R^*) + (K_2 - R^*)P'_\nu(R^*). \end{aligned}$$

Conversely, let  $\Theta = \frac{K_1-w}{S^*-w} = \frac{K_2-R^*}{S^*-R^*} = \frac{(K_2-R^*)-(K_1-w)}{w-R^*} \in (0, 1)$  and let  $\psi^*(x) = \Theta(S^* - x)^+ + (1 - \Theta)(R^* - x)^+$ . Note that by design  $\psi^*(R^*) = \Theta(S^* - R^*) = (K_2 - R^*)$  so that  $\psi^*(z) \geq (K_2 - z)^+$ . Further,  $\psi^*(w) = \Theta(S^* - w) = (K_1 - w)$  so that  $\phi^*(w) = 0$  where  $\phi^*(z) = ((K_1 - z)^+ - \psi^*(z))^+$ . Then  $\mathcal{D} \leq \Theta P_\nu(S^*) + (1 - \Theta)P_\nu(R^*) = \mathcal{D}(\psi^*)$ .

Now consider  $\mathcal{D}(\psi^*) - A(u^*)$ . Using (14) for the second equality and the alternative characterisations of  $\Theta$  for the third, we have

$$\begin{aligned} \mathcal{D}(\psi^*) - A(u^*) &= \Theta(P_\nu(S^*) - P_\nu(R^*)) - (K_1 - w)[P'_\nu(S^*) - P'_\nu(R^*)] \\ &\quad - (K_2 - R^*)P'_\nu(R^*) \\ &= P'_\nu(S^*)[\Theta(S^* - w) - (K_1 - w)] \\ &\quad - P'_\nu(R^*)[\Theta(w - R^*) - (K_1 - w) + (K_2 - R^*)] \\ &= 0. \end{aligned}$$

Then  $\mathcal{D}(\psi^*) = A(u^*) \leq \sup_{u \in [0,1]} A(u) \leq \mathcal{P}_N \leq \mathcal{D} \leq \mathcal{D}(\psi^*)$ . It follows that this chain of inequalities is in fact a chain of equalities and  $\mathcal{P}_N = \sup_{u \in [0,1]} A(u)$ . Moreover, we have identified an optimal model and an optimal stopping rule. The model which yields the highest price for the American put is our extension of the left-curtain coupling.

4.3. *American puts with a general time-1 law.* We seek to generalise the arguments of the previous section to allow for nontrivial initial laws. Define  $\Lambda = \Lambda(r, g, s)$  via

$$\Lambda(r, g, s) = \frac{K_1 - g}{s - g} - \frac{(K_2 - r) - (K_1 - g)}{g - r}.$$

Suppose we are in the case of continuous  $\mu$ . Define  $\hat{\Lambda}(x) = \Lambda(f(x), x, g(x))$  where  $f$  and  $g$  are the lower and upper functions which arise in the Beiglböck–Juillet [5] characterisation of the left-curtain martingale coupling. In our notation, this can be written as  $\hat{\Lambda}(x) = \Lambda((R \circ G^{-1})(x), x, (S \circ G^{-1})(x))$ . The fundamental insight in Hobson and Norgilas [17] is that, in the case of continuous  $\mu$ , the cheapest superhedge can be described in terms of a simple portfolio of European-style puts whose strikes depend on quantities which arise from looking for the root  $x^*$ , if any, of  $\hat{\Lambda}(\cdot) = 0$ . Moreover, the most expensive model is the model described by the left-curtain coupling, and an optimal exercise rule is to exercise at time-1 if and only if  $X_1 < x^*$ . Hobson and Norgilas [17] identify four archetypes of hedging portfolios. The first two cases correspond to when there is a root to  $\hat{\Lambda} = 0$  and when  $\hat{\Lambda} < 0$  for all  $x$ . (The remaining cases correspond to cases where  $\hat{\Lambda}$  is discontinuous, and jumps downwards over the value 0.)

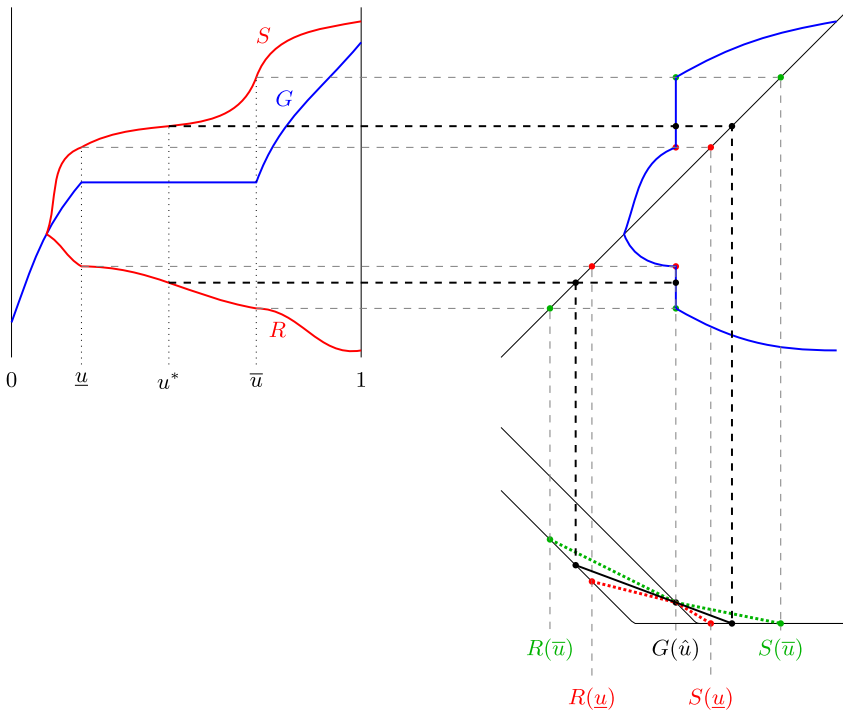


FIG. 5. Finding the optimal hedge for general measures.  $\mu$  has an atom of size  $\bar{u} - \underline{u}$ . Moreover, the piecewise linear curve joining  $(R(\underline{u}), K_2 - R(\underline{u}))$ ,  $(G(\hat{u}), K_1 - G(\hat{u}))$  and  $(S(\underline{u}), 0)$  is concave (where  $\hat{u}$  is any element of  $(\underline{u}, \bar{u}]$ ), whereas the piecewise linear curve joining  $(R(\bar{u}), K_2 - R(\bar{u}))$ ,  $(G(\hat{u}), K_1 - G(\hat{u}))$  and  $(S(\bar{u}), 0)$  is convex. There exists  $u^* \in (\underline{u}, \bar{u}]$  such that  $(R(u^*), K_2 - R(u^*))$ ,  $(G(\hat{u}), K_1 - G(\hat{u}))$  and  $(S(u^*), 0)$  all lie on a straight line. The figure describes the optimal coupling (via  $(U, V)$  and (3)) and the optimal exercise strategy for the American put is to exercise at time-1 if  $U \leq u^*$ .

In the case with atoms in  $\mu$ , we cannot use  $\hat{\Lambda}$  directly since  $G^{-1}$  has jumps. Instead, following the analysis in Section 4.1 we define  $\bar{\Lambda}(u) = \Lambda(R(u), u, S(u))$ , and look for solutions, if any, to  $\bar{\Lambda}(\cdot) = 0$ . We may still have the cases where  $\bar{\Lambda} < 0$  for all  $u \in (0, 1)$  or where  $\bar{\Lambda}(\cdot)$  jumps over zero, but these cases can be dealt with as in [17]. The new case is when the root  $u^*$  of  $\bar{\Lambda} = 0$  occurs in an interval  $(\underline{u}, \bar{u}]$  over which  $G$  is constant. This means that there is an atom of  $\mu$  at  $G(u^*)$ ; see Figure 5. A model which maximises the price of the American put is the extended left-curtain martingale coupling model, and the optimal stopping rule is to exercise at time-1 whenever  $X_1 < G(u)$  and to sometimes exercise when  $X_1 = G(u)$ . When  $X_1 = G(u)$ , the optimal stopping rule is to exercise precisely when  $U \in (\underline{u}, u^*]$  and to wait if  $U \in (u^*, \bar{u}]$ . Because  $R$  and  $S$  are monotonic over  $(\underline{u}, \bar{u}]$  paths with low future variability are exercised at time-1 whereas on paths with high future variability exercise is delayed to time-2.

APPENDIX: PROOFS

PROOF OF LEMMA 4. We begin our study of the upper bound on  $S_n$  by considering the case of a single starting measure  $\mu$  and fixed target law  $\nu$ . First, we assume that  $\mu$  and  $\nu$  are regular (no atoms and no intervals within the support with no mass), before extending to the general case. Then we consider what happens when we consider  $\mu_n \uparrow_{cx} \mu$ .

Suppose  $\mu$  and  $\nu$  have no atoms and no intervals within the support with no mass. Then  $G_\mu$  is continuous and strictly increasing. Fix  $u \in (0, 1)$  and let  $\ell_1 \equiv \ell_1^u$  be the tangent to  $P_\mu$  with slope  $u$ ; see Figure 6. By construction this tangent meets  $P_\mu$  at  $G = G_\mu(u)$ . Let  $H = H(u)$  be the point where the tangent crosses the  $x$ -axis. Let  $\ell_2 \equiv \ell_2^u$  be the tangent to  $P_\nu$  with slope greater than  $u$  which passes through  $(G, P_\mu(G))$ ; this tangent meets  $P_\nu$  at the  $x$ -coordinate  $J = J(u) = J_{\mu,\nu}(u)$ .

We now show that  $S(u) \leq J$ .

Choose  $\gamma \in [H, G)$ . Let  $\ell_3^\gamma$  be the tangent to  $P_\mu$  which passes through  $(\gamma, \ell_1(\gamma))$  and has slope less than  $u$ . Suppose this tangent meets  $P_\mu$  at  $r = r(\gamma)$ ;

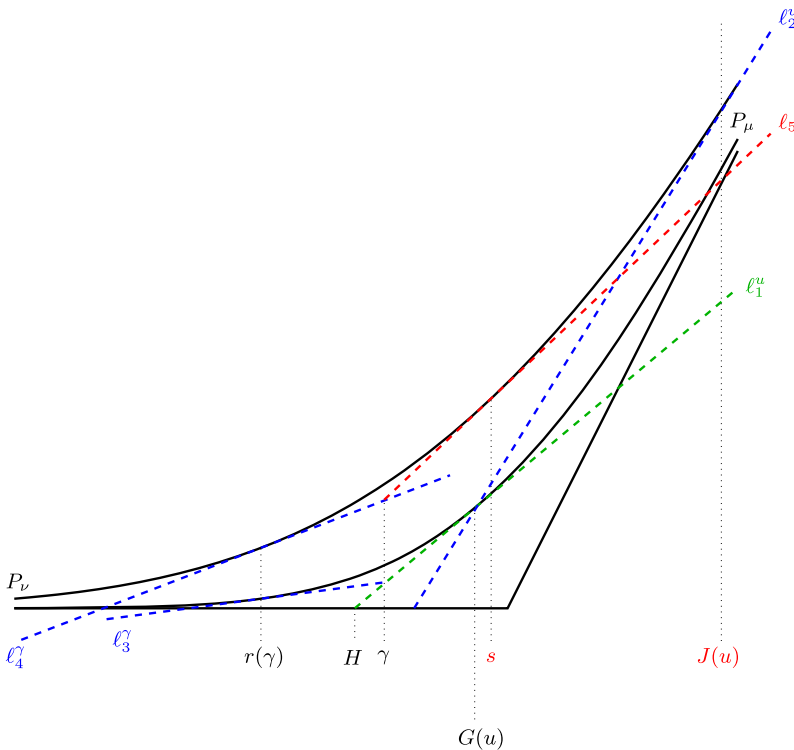


FIG. 6. Construction of function  $J$  that bounds the upper function  $S$  on  $(0, 1)$ .

the slope of the tangent is  $P'_\mu(r)$ . Let  $\ell_4^\gamma$  be the tangent to  $P_\nu$  at  $r$ . Finally, let  $\ell_5^\gamma$  be the line passing through  $(\gamma, \ell_4^r(\gamma))$  with slope  $u + P'_\nu(r) - P'_\mu(r)$ .

If there exists  $\gamma$  such that  $\ell_5^\gamma$  is a tangent to  $P_\nu$  (meeting  $P_\nu$  at  $s$  say), then  $(r, G, s)$  satisfy

$$(15) \quad \int_r^G w^i \mu(dw) = \int_r^s w^i \nu(dw), \quad i = 0, 1$$

(and moreover  $\gamma = \int_r^G w \mu(dw) / \int_r^G \mu(dw) = \int_r^s w \nu(dw) / \int_r^s \nu(dw)$  is the bary-centre of the measures  $\mu|_{(r,G)}$  and  $\nu|_{(r,s)}$ ).

For each  $u$ , there may be multiple  $\gamma$  which lead to a triple  $(r, G, s)$  which satisfies (15). We show that in each case  $s \leq J$ . It follows that  $S(u) \leq J$ .

Suppose  $\ell_4^\gamma(\gamma) \leq \ell_3^\gamma(\gamma) = \ell_1(\gamma)$ . Then necessarily  $P'_\nu(r) < P'_\mu(r)$  and  $\ell_5^\gamma$  lies below  $\ell_1$  to the right of  $\gamma$ ; in particular  $\ell_5^\gamma$  stays below  $P_\mu$  to the right of  $\gamma$  and cannot be a tangent to  $P_\nu$ . Hence if  $(r, G, s)$  satisfies (15) we must have  $\ell_4^\gamma(\gamma) > \ell_1(\gamma)$ . Then, if  $\ell_5^\gamma$  is a tangent to  $P_\nu$  we must have that the point of tangency is below  $J$ .

In the above, we used the regularity assumptions on  $\mu$  and  $\nu$  to conclude that there was a unique tangent to  $P \in \{P_\mu, P_\nu\}$  at a given point, and that there was a unique point at which  $P$  had a given slope. If  $\mu$  or  $\nu$  is not regular then, for fixed  $u$ , there may be multiple quintiles  $G$ , multiple points  $r$  and multiple tangents to  $P_\nu$  at  $r$ . The point is that although there are multiple versions of the construction in this case each candidate triple  $(r, G, s)$  satisfying (15) has  $s \leq J$  where  $J$  is defined using an arbitrary point  $G \in [G_\mu(u), G_\mu(u+)]$ . We define  $J_- = J_-(u)$  to be the smallest  $x$ -coordinate at which the tangent to  $P_\nu$  with slope greater than  $u$  passing through  $(G(u), P_\mu(G(u)))$  meets  $P_\nu$  and  $J_+ = J_+(u)$  to be the largest  $x$ -coordinate at which the tangent to  $P_\nu$  with slope greater than  $u$  passing through  $(G(u+), P_\mu(G(u+)))$  meets  $P_\nu$ . We have  $S(u) \leq J_-(u) \leq J_+(u)$ .

Finally, we want to show that if we approximate  $\mu$  by  $\mu_n$  (with  $\mu_n \uparrow_{cx} \mu$ ) then the bound  $\limsup S_n \leq J_+$  remains valid, where  $J_+$  is constructed from  $\mu$  and  $\nu$ .

Define  $K(k) = \overline{\text{argsup}}_k \frac{P_\nu(k) - P_\mu(k)}{k - k}$ . The notation  $\overline{\text{argsup}}$  is used to indicate that where there are multiple elements in the argsup we choose the largest one. Then  $K$  is increasing and right continuous in  $k$ . Note that  $J_+(u) = K(G(u+))$ . In a similar fashion, we can define  $K_n$  and  $J_n$  using  $P_{\mu_n}$  in place of  $P_\mu$ . (The target law is assumed fixed throughout.) Since  $P_{\mu_n}(k) \uparrow P_\mu(k)$  and  $K$  is right continuous, we have  $K_n(k) \downarrow K(k)$ . Then, for  $\varepsilon > 0$ ,

$$\limsup_n J_n(u) = \limsup_n K_n(G_n(u+)) \leq \limsup_n K_n(G(u+) + \varepsilon) \leq K(G(u+) + \varepsilon).$$

Since  $\varepsilon$  is arbitrary and  $K$  is right continuous,  $\limsup S_n(u) \leq \limsup_n J_n(u) \leq J_+(u)$ .  $\square$

**PROOF OF LEMMA 6.** As for the proof of Lemma 4, we begin by considering a single initial law  $\mu$ , and supposing that  $\mu$  and  $\nu$  are regular.



Fix  $u \in (0, 1)$  and let  $\ell_1$  be the tangent to  $P_\mu$  with slope  $u$ . Let  $H = H(u)$  be the point where this tangent crosses the  $x$ -axis. Suppose that  $\ell_1$  is not a tangent to  $P_\nu$ . Then  $\ell_1$  must lie strictly below  $P_\nu$ . There exists  $\varepsilon = \varepsilon(u) > 0$  such that the line passing through  $(H, \varepsilon)$  with slope  $u + \varepsilon$  lies below  $P_\nu$ . Now choose  $j = j(u)$  such that the tangents to  $P_\mu$  and  $P_\nu$  at  $j$  both have slope less than  $\varepsilon$  and both cross the line  $y = x$  below  $\varepsilon$ . Then  $R(u) \geq j$ .

To see this, let  $\gamma$  be the  $x$ -coordinate of the point where the tangent to  $P_\mu$  at  $j$  crosses  $\ell_1$ . Then if  $\ell_4$  is the tangent to  $P_\nu$  at  $j$  then  $\ell_4(\gamma) < \varepsilon$ ; if  $\ell_5$  is the line passing through  $(\gamma, \ell_4(\gamma))$  with slope  $u + P'_\nu(j) - P'_\mu(j) < u + \varepsilon$ , then by our defining assumption on  $\varepsilon$ ,  $\ell_5$  lies below  $P_\nu$ . Hence  $R(u) > j$ .

We can extend the result to irregular measures, and to  $\liminf R_n(u)$  by similar techniques as for  $S$ . The only extra issue that arises is our assumption that  $\ell^1$  is not a tangent to  $P_\nu$ . But, if for each  $n$ ,  $\ell^1$  is a tangent to  $P_\nu$ , then the same is certainly true in the limit. Then there must exist  $x$  such that  $\ell_1(x) = P_\mu(x) = P_\nu(x)$  and then  $S(u) \leq x \leq G(u+)$ . This case is excluded by hypothesis.  $\square$

## REFERENCES

- [1] BAYRAKTAR, E., HUANG, Y.-J. and ZHOU, Z. (2015). On hedging American options under model uncertainty. *SIAM J. Financial Math.* **6** 425–447. [MR3356981](#)
- [2] BEIGLBÖCK, M., COX, A. M. G. and HUESMANN, M. (2017). The geometry of multi-marginal Skorokhod embedding. Preprint. Available at [arXiv:1705.09505](#).
- [3] BEIGLBÖCK, M., HENRY-LABORDÈRE, P. and PENKNER, F. (2013). Model-independent bounds for option prices—A mass transport approach. *Finance Stoch.* **17** 477–501. [MR3066985](#)
- [4] BEIGLBÖCK, M., HENRY-LABORDÈRE, P. and TOUZI, N. (2017). Monotone martingale transport plans and Skorokhod embedding. *Stochastic Process. Appl.* **127** 3005–3013. [MR3682121](#)
- [5] BEIGLBÖCK, M. and JUILLET, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. [MR3456332](#)
- [6] BILLINGSLEY, P. (2013). *Convergence of Probability Measures*. Wiley, New York. [MR0233396](#)
- [7] BRENIER, Y. (1987). Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.* **305** 805–808. [MR0923203](#)
- [8] CHACON, R. V. (1977). Potential processes. *Trans. Amer. Math. Soc.* **226** 39–58. [MR0501374](#)
- [9] CHACON, R. V. and WALSH, J. B. (1976). One-dimensional potential embedding. In *Séminaire de Probabilités, X. Lecture Notes in Math.* **511** 19–23. Springer, Berlin. [MR0445598](#)
- [10] GALICHON, A., HENRY-LABORDÈRE, P. and TOUZI, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24** 312–336. [MR3161649](#)
- [11] HENRY-LABORDÈRE, P. and TOUZI, N. (2016). An explicit martingale version of the one-dimensional Brenier theorem. *Finance Stoch.* **20** 635–668. [MR3519164](#)
- [12] HOBSON, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris–Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math.* **2003** 267–318. Springer, Berlin. [MR2762363](#)
- [13] HOBSON, D. and NEUBERGER, A. (2012). Robust bounds for forward start options. *Math. Finance* **22** 31–56. [MR2881879](#)

- [14] HOBSON, D. and NEUBERGER, A. (2017). Model uncertainty and the pricing of American options. *Finance Stoch.* **21** 285–329. [MR3590709](#)
- [15] HOBSON, D. G. (1998). Robust hedging of the lookback option. *Finance Stoch.* **2** 329–347.
- [16] HOBSON, D. G. and NEUBERGER, A. (2016). More on hedging American options under model uncertainty. Preprint. Available at [arXiv:1604.02274](#).
- [17] HOBSON, D. G. and NORGILAS, D. (2017). Robust bounds for the American put. Preprint. Available at [arXiv:1711.06466](#).
- [18] JUILLET, N. (2016). Stability of the shadow projection and the left-curtain coupling. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1823–1843. [MR3573297](#)
- [19] NEUBERGER, A. (2007). Bounds on the American option. Preprint. SSRN:966333.
- [20] NUTZ, M. and STEBEGG, F. (2018). Canonical supermartingale couplings. *Ann. Probab.* **46** 3351–3398. [MR3857858](#)
- [21] NUTZ, M., STEBEGG, F. and TAN, X. (2017). Multiperiod martingale transport. Preprint. Available at [arXiv:1703.10588](#).
- [22] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36** 423–439. [MR0177430](#)

DEPARTMENT OF STATISTICS  
UNIVERSITY OF WARWICK  
COVENTRY CV4 7AL  
UNITED KINGDOM  
E-MAIL: [d.hobson@warwick.ac.uk](mailto:d.hobson@warwick.ac.uk)  
[d.norgilas@warwick.ac.uk](mailto:d.norgilas@warwick.ac.uk)