



Robust bounds for the American put

David Hobson¹ · Dominykas Norgilas¹

Received: 10 November 2017 / Accepted: 02 January 2019 / Published online: 5 March 2019
© The Author(s) 2019

Abstract We consider the problem of finding a model-free upper bound on the price of an American put given the prices of a family of European puts on the same underlying asset. Specifically, we assume that the American put must be exercised at either T_1 or T_2 and that we know the prices of all vanilla European puts with these maturities. In this setting, we find a model which is consistent with European put prices, together with an associated exercise time, for which the price of the American put is maximal. Moreover, we derive the cheapest superhedge. The model associated with the highest price of the American put is constructed from the left-curtain martingale coupling of Beiglböck and Juillet (Ann. Probab. 44:42–106, 2016).

Keywords Model-independent pricing · American put · Martingale optimal transport · Left-curtain coupling · Optimal stopping

Mathematics Subject Classification (2010) 60G40 · 60G42 · 91G20

JEL Classification G13

1 Introduction

This article is motivated by an attempt to understand the range of possible prices of an American put in a robust, or model-independent, framework. In our interpretation, this means that we assume we are given today's prices of a family of European-style vanilla puts (for a continuum of strikes and for a discrete set of maturities). The

✉ D. Hobson
d.hobson@warwick.ac.uk
D. Norgilas
d.norgilas@warwick.ac.uk

¹ Department of Statistics, University of Warwick, Coventry, UK

goal is to find a consistent model for the underlying for which the American put has the highest price, where by definition a model is consistent if the discounted price process is a martingale and the model-based discounted expected values of European put payoffs match the given prices of European puts.

This notion of model-independent, or robust, bounds on the prices of exotic options was introduced in Hobson [17] in the context of lookback options, and has been applied several times since; see Brown et al. [8] (barrier options), Cox and Oblój [11] (no-touch options), Hobson and Neuberger [19] and Hobson and Klimmek [18] (forward-start straddles), Carr and Lee [9] and Cox and Wang [12] (variance options), Stebegg [27] (Asian options) and the survey article Hobson [24]. The principal idea is that the prices of the vanilla European puts determine the marginal distributions of the price process at the traded maturities (but not the joint distributions) and that these distributional requirements, coupled with the martingale property, place meaningful and useful restrictions on the class of consistent models. These restrictions lead to bounds on the expected payoffs of path-dependent functionals, or equivalently, to bounds on the prices of exotic options.

In addition to the pricing problem, there is a related dual or hedging problem. In the dual problem, the aim is to construct a static portfolio of European put options and a dynamic discrete-time hedge in the underlying which combine to form a superhedge (pathwise over a suitable class of candidate price paths) for the exotic option. The value of the dual problem is the cost of the cheapest superhedge. There is a growing literature, beginning with Beiglböck et al. [5] for discrete-time problems, and Galichon et al. [15] in continuous time, which aims to explain how to formulate the problem in such a way that there is no duality gap, i.e., the highest model-based price is equal to the cheapest superhedge, either for specific derivatives or in general.

Many of the early papers on robust hedging exploited a link with the Skorokhod embedding problem (Skorokhod [26, Sect. 7]). For example, in the study of the lookback option in Hobson [17], the consistent model which achieves the highest lookback price is constructed from the Azéma and Yor [2] solution of the Skorokhod embedding problem. More recently, Beiglböck et al. [5] (see also Dolinsky and Soner [14] and Touzi [29]) have championed the connection between robust hedging problems and martingale optimal transport. In this paper, we make use of the left-curtain martingale coupling introduced by Beiglböck and Juillet [4] and developed by Henry-Labordère and Touzi [16] and Beiglböck et al. [6].

The study of American-style claims in a robust framework was initiated by Neuberger [25]; see also Hobson and Neuberger [20], Bayraktar and Zhou [3] and Aksamit et al. [1]. (There is also a paper by Cox and Hoeggerl [10] which asks about the possible shapes of the price of an American put, considered as a function of strike, given the prices of co-maturing European puts.) The main innovation of the present paper is that rather than focusing on general American payoffs and proving that the pricing (primal) problem and the dual (hedging) problem have the same value, we focus explicitly on American puts and try to say as much as possible about the structure of the consistent price process for which the model-based American put price is maximised, and the structure of the cheapest superhedge.

Our problem can be cast as follows. Let $M = (M_0 = \bar{\mu}, M_1 = X, M_2 = Y)$ represent the discounted price of an underlying asset, where $\bar{\mu}$ is a known constant. The

laws of X and Y are presumed to be given and $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$, where μ and ν are (integrable) probability measures on \mathbb{R} with mean $\bar{\mu}$. Given a martingale model (a filtered probability space supporting a stochastic process M which is a martingale), we consider an American put on M with strike K . The option may only be exercised at time 1 or time 2; if the put is exercised at time 1, the payoff is $(K_1 - X)^+$, if the put is exercised at time 2, the payoff is $(K_2 - Y)^+$. Here K_1 and K_2 represent the discounted strikes of the put. For any martingale model, the model-based price of the American put is then given by the expected value of the payoff calculated under the best available stopping time (defined with respect to the filtration associated to the given model, and taking values in time 1 or time 2). Our primal problem is to find the highest possible model-based price of the American put, i.e., the highest expected payoff, where expectations are calculated under the probability measure of a consistent model (a model under which M is a martingale and has the given laws at times 1 and 2).

There is a corresponding dual or hedging problem of finding the cheapest superhedge based on static portfolios of European puts and a piecewise constant holding of the underlying asset; see Sect. 2.2.

Our main achievement is as follows:

Main result 1.1 *Suppose μ is continuous. The highest model-based expected payoff of the American put is equal to the cheapest superhedging price. Moreover, the highest model-based expected payoff is attained by the model associated with the left-curtain martingale coupling of Beiglböck and Juillet [4] (and a judiciously chosen stopping rule). Further, we can characterise the cheapest superhedging strategy, and it is one of four possible types.*

For fixed μ, ν and $K_1 > K_2$, there is typically a family of optimal models. Fixing μ and ν but varying K_1 and K_2 , it turns out that there is a model which is optimal for all K_1 and K_2 simultaneously. This model is related to the left-curtain coupling of Beiglböck and Juillet [4].

The remainder of the paper is structured as follows. In the next section, we formulate precisely our problem of finding the robust, model-independent price of an American put. We also explain how the pricing problem is related to the dual problem of constructing the cheapest superhedge. In Sect. 3, we assume that μ is continuous, transform the primal pricing problem into a martingale optimal transport (MOT) problem, and show by studying a series of ever more complicated setups how to determine the best model and hedge.

By weak duality, the highest model price is bounded above by the cost of the cheapest superhedge. Hence, if, on the one hand, we can identify a consistent model and stopping rule, and, on the other hand, a superhedge such that the expected payoff in that model with that stopping rule is equal to the cost of the superhedge, then we must have identified an optimal model and an optimal stopping rule together with an optimal hedging strategy. Moreover, there is no duality gap. This is the strategy of our proofs. One feature of our analysis is that wherever possible, we provide pictorial explanations and derivations of our results. In our view, this helps bring insights which may be hidden under calculus-based approaches.

2 Preliminaries and setup

2.1 The financial model and model-based prices for American puts

Suppose time is discrete and interest rates are non-stochastic. Without loss of generality, we may identify trading times with the nonnegative integers $t = 0, 1, 2$. Let $M = (M_t)_{t=0,1,2}$ be the discounted asset price which we expect to be a martingale under a pricing measure. We assume M_0 is known at time 0.

We are interested in pricing an American put with strike K and maturity 2. Under the bond price numeraire, this corresponds to a put with strike K_1 at time 1 and K_2 at time 2 (and $K_1 > K_2$ provided interest rates are strictly positive, which we now assume without further comment). Note that we do not allow exercise at $t = 0$.

We suppose we are given European put prices for maturities 1 and 2 for a continuum of strikes. By classical arguments (e.g. Breeden and Litzenberger [7]), it is possible to infer the laws of the price of the asset, and hence the laws of the discounted asset price. Denote the law of $X = M_1$ by μ and the law of $Y = M_2$ by ν . It follows from Jensen’s inequality that if μ and ν have arisen from sets of European put options in this way, μ and ν are in convex order, and we write $\mu \leq_{cx} \nu$ (see Sect. 2.4 for a further discussion of the properties of μ and ν).

Definition 2.1 (Hobson and Neuberger [21]) Suppose we have $\mu \leq_{cx} \nu$ and 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 that $M = (M_0, M_1, M_2) = (\bar{\mu}, X, Y)$ is an (\mathcal{S}, μ, ν) -consistent stochastic process and write $M \in \mathcal{M}(\mathcal{S}, \mu, \nu)$ if

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

We say that (\mathcal{S}, M) is a (μ, ν) -consistent model if \mathcal{S} is a filtered probability space and M is an (\mathcal{S}, μ, ν) -consistent stochastic process. Where μ and ν are clear from the context, this is sometimes abbreviated to a *consistent model*.

Let $B_1 \in \mathcal{F}_1$. Define the stopping time τ_{B_1} by $\tau_{B_1} = 1$ on B_1 and $\tau_{B_1} = 2$ on B_1^c . (Conversely, any stopping rule taking values in $\{1, 2\}$ has a representation of this form.) Suppose (\mathcal{S}, M) is a (μ, ν) -consistent model. The (\mathcal{S}, M) -model-based expected payoff (MBEP) of the American put under the stopping rule τ_{B_1} is

$$\mathcal{A}(B_1, M, \mathcal{S}) = \mathbb{E}[(K_{\tau_{B_1}} - M_{\tau_{B_1}})^+].$$

Then, optimising over stopping rules under the model (\mathcal{S}, M) , the model-based price of the American put is $\mathcal{A}(M, \mathcal{S}) = \sup_{B_1 \in \mathcal{F}_1} \mathcal{A}(B_1, M, \mathcal{S})$. The highest model-based expected payoff for the American put is

$$\mathcal{P} = \mathcal{P}(\mu, \nu) = \sup_{\mathcal{S}} \sup_{M \in \mathcal{M}(\mathcal{S}, \mu, \nu)} \mathcal{A}(M, \mathcal{S}).$$

Amongst the class of consistent models, there is a natural and important class of models which we call the class of *canonical models*. Although we typically expect nonnegative prices in the finance context, in this definition and the mathematical analysis which follows, we allow measures μ and ν supported on \mathbb{R} .

Definition 2.2 Suppose $\mu \leq_{cx} \nu$. We call $(\hat{S} = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{\mathbb{P}} = (\hat{\mathcal{F}}_0, \hat{\mathcal{F}}_1, \hat{\mathcal{F}}_2)), \hat{M})$ a canonical (μ, ν) -consistent model if (\hat{S}, \hat{M}) is a (μ, ν) -consistent model such that $\hat{\Omega} = \mathbb{R} \times \mathbb{R}$, $\hat{\mathbb{P}} = \mathcal{B}(\hat{\Omega})$, $\hat{M}_1(\omega_1, \omega_2) = \omega_1$, $\hat{M}_2(\omega_1, \omega_2) = \omega_2$ and such that $\hat{\mathcal{F}}_0$ is trivial, $\hat{\mathcal{F}}_1 = \sigma(\hat{M}_1)$ and $\hat{\mathcal{F}}_2 = \sigma(\hat{M}_1, \hat{M}_2)$. Then $\hat{B}_1 \in \hat{\mathcal{F}}_1$ can be identified with an element \hat{B} in $\mathcal{B}(\mathbb{R})$ via $\hat{B}_1 = \hat{B} \times \mathbb{R}$.

In the canonical setting, different models (consistent or not) can be parametrised by a probability measure π on \mathbb{R}^2 . To simplify the notation, we write \hat{M}_π for the canonical model $(\hat{S}_\pi = (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}_\pi, \hat{\mathbb{P}}), \hat{M})$, where $\hat{\mathbb{P}}_\pi$ is the probability measure such that $\hat{\mathbb{P}}_\pi[X \in dx, Y \in dy] = \pi(dx, dy)$.

Define $\hat{\mathcal{P}} = \sup \sup_{\hat{B} \in \mathcal{B}(\mathbb{R})} \mathbb{E}[(K_{\tau_{\hat{B}}} - M_{\tau_{\hat{B}}})^+]$, where the first supremum is taken over canonical (μ, ν) -consistent models and $\tau_{\hat{B}}(\omega) = 1$ if $X(\omega) \in \hat{B}$ and $\tau_{\hat{B}}(\omega) = 2$ otherwise. Clearly, since the set of canonical consistent models is a subset of the set of all consistent models, we have $\hat{\mathcal{P}} \leq \mathcal{P}$.

In this paper, we concentrate on the case where μ is continuous. In that case, we show that $\hat{\mathcal{P}} = \mathcal{P}$. However, if μ has atoms, the situation becomes more delicate, as pointed out in Hobson and Neuberger [20]; see also Hobson and Neuberger [21], Bayraktar and Zhou [3] and Aksamit et al. [1]. On the one hand, we must allow a wider range of possible candidates for exercise-determining sets B_1 . On atoms of X , we may want to sometimes stop and sometimes continue, although we must still take stopping decisions which do not violate the martingale property of future price movements. On the other hand, the functions T_d, T_u that characterise the left-curtain coupling (see Sect. 2.5) become ill-defined on the points where μ has atoms. Then it is not clear how the optimal model can be identified. For these reasons, we must extend our notion of a martingale coupling and generalise, in a useful fashion, the left-curtain martingale coupling of Beiglböck and Juillet [4] to the case with atoms. The appropriate extension of the left-curtain coupling to the case with atoms in μ is discussed in a companion paper [22]; in the present paper, we focus on the financial aspects of our results, namely the application to the robust hedging of American puts.

2.2 Superhedging

The following notion of a robust superhedge for an American option was first introduced by Neuberger [25]; see also Bayraktar and Zhou [3] and Hobson and Neuberger [20].

We work in discounted units over two time-points. Consider a general American-style option with payoff a if exercised at time 1, and payoff b if exercised at time 2, where $a : \mathbb{R} \rightarrow \mathbb{R}_+$ and $b : \mathbb{R} \rightarrow \mathbb{R}_+$ are positive functions.

Definition 2.3 $(\phi, \psi, (\theta_i)_{i=1,2})$ is a superhedge for (a, b) if for all $x, y \in \mathbb{R}$,

$$a(x) \leq \phi(x) + \psi(y) + \theta_1(x)(y - x), \tag{2.1}$$

$$b(y) \leq \phi(x) + \psi(y) + \theta_2(x)(y - x). \tag{2.2}$$

The hedging cost (HC) associated with the superhedge $(\phi, \psi, (\theta_i)_{i=1,2})$ is given by

$$\mathcal{C} = \mathcal{C}(\phi, \psi, (\theta_i)_{i=1,2}; \mu, \nu) = \int \phi(x)\mu(dx) + \int \psi(y)\nu(dy),$$

where we set $\mathcal{C} = \infty$ if $\int \phi(x)^+ \mu(dx) + \int \psi(y)^+ \nu(dy) = \infty$. We let $\mathcal{H}(a, b)$ be the set of superhedging strategies $(\phi, \psi, (\theta_i)_{i=1,2})$.

The idea behind the definition is that the hedger purchases a portfolio of maturity-1 European puts (and calls) with payoff ϕ and a portfolio of maturity-2 European puts (and calls) with payoff ψ . (The fact that this can be done and has cost \mathcal{C} follows from arguments of Breeden and Litzenberger [7].) In addition, if the American option is exercised at time 1, the hedger holds θ_1 units of the underlying between times 1 and 2; otherwise the hedger holds θ_2 units of the underlying over this time-period. In the former case, (2.1) implies that the strategy superhedges the American option payout; in the latter case, (2.2) implies the same.

The dual (superhedging) problem is to find

$$\mathcal{D} = \mathcal{D}(a, b; \mu, \nu) = \inf_{(\phi, \psi, (\theta_i)_{i=1,2}) \in \mathcal{H}(a,b)} \mathcal{C}(\phi, \psi, (\theta_i)_{i=1,2}; \mu, \nu).$$

Potentially, the space $\mathcal{H} = \mathcal{H}(a, b)$ could be very large and it is extremely useful to be able to search over a smaller space. The next lemma shows that any convex ψ with $\psi \geq b$ can be used to generate a superhedge $(\phi, \psi, (\theta_i)_{i=1,2})$.

For a convex function χ , let χ'_+ denote the right derivative of χ .

Lemma 2.4 *Suppose $\psi \geq b$ with ψ convex. Define $\phi = (a - \psi)^+$ and set $\theta_2 = 0$ and $\theta_1 = -\psi'_+$. Then $(\phi, \psi, (\theta_i)_{i=1,2})$ is a superhedge.*

Proof We have

$$b(y) \leq \psi(y) \leq \phi(x) + \psi(y) = \phi(x) + \psi(y) + \theta_2(x)(y - x)$$

and (2.2) follows. Also, by the convexity of ψ , $\psi(x) \leq \psi(y) - \psi'_+(x)(y - x)$ and

$$a(x) \leq (a(x) - \psi(x))^+ + \psi(x) \leq \phi(x) + \psi(y) + \theta_1(x)(y - x).$$

Hence (2.1) follows. □

Let $\check{\mathcal{H}} = \check{\mathcal{H}}(b)$ be the set of convex functions ψ with $\psi \geq b$. For $\psi \in \check{\mathcal{H}}$, we can define the associated hedging cost $\check{\mathcal{C}}(\psi; \mu, \nu)$ by

$$\begin{aligned} \check{\mathcal{C}}(\psi; \mu, \nu) &= \mathcal{C}((a - \psi)^+, \psi, \theta_1 = -\psi'_+, \theta_2 = 0; \mu, \nu) \\ &= \int (a(x) - \psi(x))^+ \mu(dx) + \int \psi(y) \nu(dy). \end{aligned}$$

The reduced dual hedging problem restricts attention to superhedges generated from $\psi \in \check{\mathcal{H}}$ and is to find

$$\check{\mathcal{D}} = \check{\mathcal{D}}(a, b; \mu, \nu) = \inf_{\psi \in \check{\mathcal{H}}(b)} \check{\mathcal{C}}(\psi; \mu, \nu).$$

Clearly, we have $\mathcal{D} \leq \check{\mathcal{D}}$; we shall show that $\mathcal{D} = \check{\mathcal{D}}$ for the American put.

2.3 Weak and strong duality

Let (\mathcal{S}, M) be a (μ, ν) -consistent model and let τ be an arbitrary stopping time in this framework. The expected payoff of the American put under this stopping rule is $\mathbb{E}[(K_\tau - M_\tau)^+]$. Conversely, let ψ be any convex function with $\psi(y) \geq (K_2 - y)^+$ and let $\phi(x) = ((K_1 - x)^+ - \psi(x))^+$ and $\theta_i(x) = -\psi'_+(x)I_{\{i=1\}}$. Then for any $i \in \{1, 2\}$, we have $(K_i - M_i)^+ \leq \psi(M_2) + \phi(M_1) + \theta_i(M_1)(M_2 - M_1)$ and hence for any random time τ taking values in $\{1, 2\}$, $(K_\tau - M_\tau)^+ \leq \psi(M_2) + \phi(M_1) + \theta_\tau(M_1)(M_2 - M_1)$. Then $\mathbb{E}[(K_\tau - M_\tau)^+] \leq \mathbb{E}^{X \sim \mu, Y \sim \nu}[\phi(X) + \psi(Y)]$ and we have the weak duality $\mathcal{P} \leq \mathcal{D}$.

In Sect. 3, we show that we can find $(\hat{\mathcal{S}}^*, \hat{M}^*, \hat{B}^*)$, with $(\hat{\mathcal{S}}^*, \hat{M}^*)$ a canonical (μ, ν) -consistent model and $\hat{B}^* \in \mathcal{B}(\mathbb{R})$, and $\psi^* \in \check{H}$ such that

$$\mathcal{A}(\hat{B}^* \times \mathbb{R}, \hat{M}^*, \hat{\mathcal{S}}^*) = \check{\mathcal{C}}(\psi^*; \mu, \nu).$$

Then $\mathcal{A}(\hat{B}^* \times \mathbb{R}, \hat{M}^*, \hat{\mathcal{S}}^*) \leq \hat{\mathcal{P}} \leq \mathcal{P} \leq \mathcal{D} \leq \check{\mathcal{D}} \leq \check{\mathcal{C}}(\psi^*; \mu, \nu)$; but since the two outer terms are equal, we have $\mathcal{P} = \mathcal{D}$ and strong duality. Moreover, $(\hat{\mathcal{S}}^*, \hat{M}^*)$ is a canonical, consistent model which generates the highest price for the American put (and τ^* given by $\tau^* = 1$ if and only if $X \in \hat{B}^*$ is the optimal exercise rule), and ψ^* generates the cheapest superhedge.

2.4 Measures and convex order

Given a finite and integrable measure η (not necessarily a probability measure) on \mathbb{R} , define $\bar{\eta} = \frac{\int_{\mathbb{R}} x\eta(dx)}{\int_{\mathbb{R}} \eta(dx)}$ to be the barycentre of η . Let \mathcal{I}_η be the smallest interval containing the support of η , and let $\{\ell_\eta, r_\eta\}$ be the endpoints of \mathcal{I}_η . Define the function $P_\eta : \mathbb{R} \rightarrow \mathbb{R}_+$ by $P_\eta(k) = \int_{-\infty}^k (k - x)\eta(dx)$. Then P_η is convex and increasing and represents the discounted European put price, expressed as a function of the strike, if the discounted underlying has law η at maturity. In addition, we have $\{k : P_\eta(k) > \eta(\mathbb{R})(k - \bar{\eta})^+\} \subseteq \mathcal{I}_\eta$. Note that P_η is related to the potential U_η 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 any real numbers $c < d$ and a measure η , let $\eta_{c,d}$ be the measure given by $\eta_{c,d}(A) = \eta(A \cap (c, d))$, $A \in \mathcal{B}(\mathbb{R})$. Let $\tilde{\eta}_{c,d} = \eta - \eta_{c,d}$.

Two measures η and χ are in convex order, and we write $\eta \leq_{cx} \chi$, 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} (or equivalently, if $\int_{\mathbb{R}} f d\eta \leq \int_{\mathbb{R}} f d\chi$ for any convex $f : \mathbb{R} \rightarrow \mathbb{R}$). Necessarily, we then have $\ell_\chi \leq \ell_\eta \leq r_\eta \leq r_\chi$. Let $\Pi(\eta, \chi)$ be the set of martingale couplings of η and χ , i.e.,

$$\Pi(\eta, \chi) = \{\pi \in \mathcal{P}(\mathbb{R}^2) : \pi \text{ has first marginal } \eta \text{ and second marginal } \chi, \text{ and (2.3) holds}\},$$

where $\mathcal{P}(\mathbb{R}^2)$ is the set of probability measures on \mathbb{R}^2 and (2.3) is the martingale condition

$$\begin{aligned} \int_{x \in B} \int_{y \in \mathbb{R}} y \pi(dx, dy) &= \int_{x \in B} \int_{y \in \mathbb{R}} x \pi(dx, dy) \\ &= \int_B x \eta(dx), \quad \forall B \in \mathcal{B}(\mathbb{R}). \end{aligned} \tag{2.3}$$

By a classical result of Strassen [28], $\Pi(\mu, \nu)$ is nonempty if and only if $\mu \leq_{cx} \nu$. Note that there is a 1–1 correspondence between canonical, (μ, ν) -consistent models and elements $\pi \in \Pi(\mu, \nu)$, given by $\hat{\mathbb{P}}[\hat{M}_1 \in dx, \hat{M}_2 \in dy] = \pi(dx, dy)$.

Given $\pi \in \Pi(\mu, \nu)$, we say π maps $A \subseteq \mathbb{R}$ to $B \subseteq \mathbb{R}$ if $\pi(A \times \mathbb{R}) = \pi(A \times B)$ (or equivalently, if under the canonical model, $\hat{M}_1 \in A$ implies $\hat{M}_2 \in B$ almost surely). We say π maps A onto B if $\pi(A \times \mathbb{R}) = \pi(A \times B) = \pi(\mathbb{R} \times B)$ (or equivalently, if under the canonical model, $\hat{M}_1 \in A$ if and only if $\hat{M}_2 \in B$ almost surely). Finally, we say π is constant on A if $\pi(C \times \mathbb{R}) = \pi(C \times C)$ for all $C \subseteq A$ (or equivalently, if $\hat{M}_2 = \hat{M}_1$ almost surely on $\{\hat{M}_1 \in A\}$).

For a pair of measures η, χ on \mathbb{R} , let the function $D = D_{\eta, \chi} : \mathbb{R} \rightarrow \mathbb{R}_+$ be defined by $D_{\eta, \chi}(k) = P_\chi(k) - P_\eta(k)$. Note that if η, χ have equal mass and equal barycentre, then $\eta \leq_{cx} \chi$ is equivalent to $D \geq 0$ on \mathbb{R} . Let $\mathcal{I}_D = [\ell_D, r_D]$ be the smallest closed interval containing $\{k : D_{\eta, \chi}(k) > 0\}$. If \mathcal{I}_D is such that $\mathcal{I}_D \subseteq \mathcal{I}_\chi$, then we must have $\eta = \chi$ on $[\ell_\chi, \ell_D] \cup (r_D, r_\chi]$.

The following lemma 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 . The result is well known and can be traced back (at least) to Hobson [23, Sect. 2].

Lemma 2.5 *Suppose η and χ are probability measures with $\eta \leq_{cx} \chi$. Suppose that $D(x) = 0$. If $\pi \in \Pi(\eta, \chi)$, then*

$$\pi((-\infty, x) \times (x, \infty)) + \pi((x, \infty) \times (-\infty, x)) = 0.$$

It follows from Lemma 2.5 that if there is a point x in the interior of the interval \mathcal{I}_η 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 the mass of χ at x appropriately. Indeed, if there are multiple x_j with $D_{\eta, \chi}(x_j) = 0$, then we can divide the problem into a sequence of ‘irreducible’ problems,¹ 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. However, in our setting, in addition to specifying a model (or equivalently a martingale coupling), we also need to specify a stopping rule, and this needs to be defined across all irreducible components simultaneously. For this reason, we do not insist that $D > 0$ on the interior of \mathcal{I}_χ , although this will be the case in the simple settings in which we build our solution.

¹The terminology ‘irreducible’ is due to Beiglböck and Juillet [4], although the idea of splitting a problem of constructing a martingale coupling into separate components is also present in the earlier papers of Hobson [23] and Cox [13].

2.5 The left-curtain coupling

The left-curtain coupling (or martingale transport) was introduced by Beiglböck and Juillet [4] and further studied by Henry-Labordère and Touzi [16] and Beiglböck et al. [6].

For real numbers c, 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 . $\chi_{c,x,d}$ is the law of a Brownian motion started at x evaluated on the first exit from (c, d) .

Lemma 2.6 (Beiglböck and Juillet [4, Corollary 1.6]) *Let μ, ν be probability measures with $\mu \leq_{cx} \nu$ and assume that μ is continuous. Then there exists a pair of measurable functions $T_d : \mathbb{R} \rightarrow \mathbb{R}$ and $T_u : \mathbb{R} \rightarrow \mathbb{R}$ with the properties that $T_d(x) \leq x \leq T_u(x)$, that $T_u(x) \leq T_u(x')$ and $T_d(x') \notin (T_d(x), T_u(x))$ for all $x < x'$, and that if we define $\pi_{lc}(dx, dy) = \mu(dx)\chi_{T_d(x),x,T_u(x)}(dy)$, then $\pi_{lc} \in \Pi(\mu, \nu)$. π_{lc} is called the left-curtain martingale coupling.*

Note that there is no claim of uniqueness of the functions T_d, T_u in Lemma 2.6. Indeed, T_d, T_u are unique only μ -a.s. (see Beiglböck et al. [6, Proposition 3.4]). For example, the definitions of T_d and T_u are immaterial outside $[\ell_\mu, r_\mu]$. Further, if T_u has a (necessarily upward) jump at x' , then it does not matter what value we take for $T_u(x')$ provided $T_u(x') \in [T_u(x'-), T_u(x'+)]$. (Since we are assuming μ is continuous, the probability that we choose an x -coordinate value of x' is zero.) More importantly, if (T_d, T_u) satisfy the properties of Lemma 2.6 and if $T_u(x) = x$ on an interval $[\underline{x}, \bar{x}]$, then we can modify the definition of T_d on $[\underline{x}, \bar{x}]$ to either $T_d(x) = x$ or $T_d(x) = T_d(\underline{x}-)$ and still satisfy the relevant monotonicity properties. Henry-Labordère and Touzi [16] resolve this indeterminacy by setting $T_d(x) = x$ on the set $\{T_u(x) = x\}$ and also taking T_u and T_d to be right-continuous.

In the sequel, we follow Henry-Labordère and Touzi [16] by taking $T_d(x) = x$ on the set $\{T_u(x) = x\}$, but we do not make right-continuity assumptions on T_d and T_u . Also we write (f, g) in place of (T_d, T_u) . Our functions f and g will eventually be defined on \mathbb{R} (see Sect. 3.3.5), but for now we define them just on $[\ell_\mu, r_\mu]$.

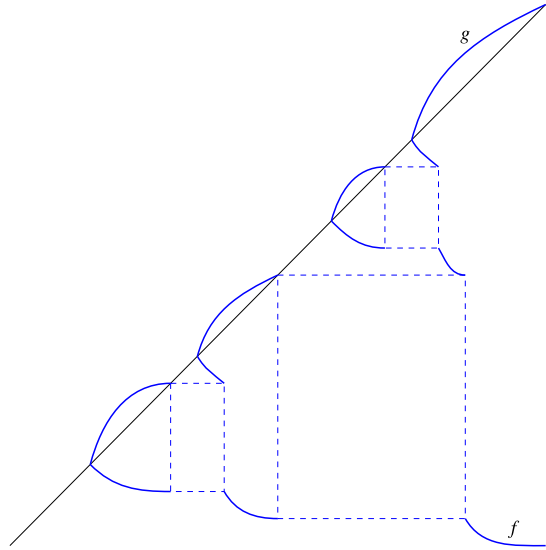
Lemma 2.7 *Let (T_d, T_u) be a pair of functions satisfying the properties in Lemma 2.6. Suppose they lead to a solution $\pi_{lc} \in \Pi(\mu, \nu)$. On $[\ell_\mu, r_\mu]$, set $g(x) = T_u(x)$; on $\{g(x) > x\}$, set $f(x) = T_d(x)$; and on $\{g(x) = x\}$, set $f(x) = x$. Then (f, g) are such that $f(x) \leq x \leq g(x)$ and for all $x' > x$, we have $g(x') \geq g(x)$ and $f(x') \notin (f(x), g(x))$. Moreover,*

$$\mu(dx)\chi_{f(x),x,g(x)}(dy) = \mu(dx)\chi_{T_d(x),x,T_u(x)}(dy).$$

Proof The property $f(x) \leq x \leq g(x)$ is immediate; so we only need to check that for $x' > x$, we have $g(x') \geq g(x)$ and $f(x') \notin (f(x), g(x))$. Monotonicity of g is inherited from monotonicity of T_u . If $g(x) = x$, then $f(x) = x$ and $f(x') \notin (f(x), g(x)) = \emptyset$. If $g(x) > x$ and $g(x') > x'$, then

$$f(x') = T_d(x') \notin (T_d(x), T_u(x)) = (f(x), g(x)).$$

Fig. 2.1 Stylised plot of the functions f and g in the general case (with no atoms). Note that on the set $\{g(x) = x\}$, we have $f(x) = x$



Finally, if $g(x) > x$ and $g(x') = x'$, then

$$f(x') = x' \notin (f(x), x' = g(x')) \supseteq (f(x), g(x)). \quad \square$$

Figure 2.1 gives a stylised representation of f and g in the case where ν has no atoms. (Atoms of ν lead to horizontal sections of f and g ; see Sect. 3.6.) In the figure, the set $\{g(x) > x\}$ is a finite union of intervals whereas in general it may be a countable union of intervals. Similarly, in the figure, f has finitely many downward jumps whereas in general, it may have countably many jumps. Nonetheless Fig. 2.1 captures the essential behaviour of f and g .

Suppose ν is also continuous and fix x . Under the left-curtain martingale coupling, mass in the interval $(f(x), x)$ at time 1 is mapped to the interval $(f(x), g(x))$ at time 2. Thus the points $f(x), g(x)$ with $f(x) \leq x \leq g(x)$ are solutions to

$$\int_f^x \mu(dz) = \int_f^g \nu(dz), \tag{2.4}$$

$$\int_f^x z\mu(dz) = \int_f^g z\nu(dz). \tag{2.5}$$

Essentially, (2.4) is a preservation-of-mass condition and (2.5) is preservation of the mean and the martingale property. If ν has atoms, then (2.4) and (2.5) become

$$\int_f^x \mu(dz) = \int_{(f,g)} \nu(dz) + \lambda_f + \lambda_g, \tag{2.6}$$

$$\int_f^x z\mu(dz) = \int_{(f,g)} z\nu(dz) + f\lambda_f + g\lambda_g, \tag{2.7}$$

respectively, where $0 \leq \lambda_f \leq \nu(\{f\})$ and $0 \leq \lambda_g \leq \nu(\{g\})$.

Returning to the case of continuous μ and ν , for fixed x , there can be multiple solutions to (2.4) and (2.5). If, however, we consider f and g as functions of x and impose the additional monotonicity properties of Lemma 2.6 (i.e., $g(x) \leq g(x')$ and $f(x') \notin (f(x), g(x))$ for $x < x'$), then typically, for μ -almost all x , there is a unique solution to (2.4) and (2.5). However, there are exceptional x at which f jumps and at which there are multiple solutions; see Sect. 3.3.

Remark 2.8 In a related problem, Hobson and Klimmek [18] show how under natural simplifying assumptions, including the dispersion assumption below (see Assumption 3.3), upper and lower functions can be characterised as solutions of a pair of coupled differential equations. In our case, (f, g) solve a pair of coupled differential equations on an interval $[e_-, r_\mu)$, obtained from differentiating (2.4) and (2.5), namely

$$\begin{aligned} \frac{df}{dx} &= -\frac{g-x}{g-f} \frac{\rho(x)}{\eta(f) - \rho(f)}, \\ \frac{dg}{dx} &= \frac{x-f}{g-f} \frac{\rho(x)}{\eta(g)}, \end{aligned}$$

with the initial condition $f(e_-) = e_- = g(e_-)$. In addition, see Henry-Labordère and Touzi [16, Eqs. (3.10) and (3.9)], where the construction of T_d and T_u is exactly based on the resolution of (2.4) and (2.5).

3 Robust bounds for American puts when μ is atom-free

3.1 Problem formulation

Our goal in this section is to derive the highest consistent model price for the American put. We begin by giving a concise reformulation of the primal problem (recall Sect. 2.1) as a problem of martingale optimal transport (MOT), and stating the main theorem (Theorem 3.1). Then we first study the problem in a simple special case, second generalise to a case which exhibits all the main features and third present the analysis in the general case.

Recall the definition of the canonical (μ, ν) -consistent model (abbreviated to \hat{M}_π) for which $\hat{\mathbb{P}}[\hat{M}_1 \in dx, \hat{M}_2 \in dy] = \pi(dx, dy)$, where $\pi \in \Pi(\mu, \nu)$. For a pair of fixed constants K_1 and K_2 , the problem we consider is to find

$$\hat{\mathcal{P}} := \sup_{\pi \in \Pi(\mu, \nu)} \sup_{B \in \mathcal{B}(\mathbb{R})} \mathbb{E}^{\mathcal{L}(X, Y) = \pi} [(K_1 - X)^+ I_{\{X \in B\}} + (K_2 - Y)^+ I_{\{X \notin B\}}].$$

Note that $\hat{\mathcal{P}}$ corresponds to the highest model-based price of the American put over the specific subset of consistent models, and therefore $\hat{\mathcal{P}} \leq \mathcal{P}$. By weak duality (recall Sect. 2.3), it follows that $\hat{\mathcal{P}} \leq \mathcal{P} \leq \mathcal{D} \leq \check{\mathcal{D}}$.

Throughout this paper, we assume that μ has no atoms. The same assumption is made in (parts of) Beiglböck and Juillet [4], Henry-Labordère and Touzi [16] and Beiglböck et al. [6]. The extension of the left-curtain martingale coupling to the case where μ has atoms is the subject of Hobson and Norgilas [22].

Theorem 3.1 *Suppose μ has no atoms. Then $\hat{\mathcal{P}} = \mathcal{P} = \mathcal{D} = \check{\mathcal{D}}$.*

We begin by considering a couple of degenerate cases.

We say the put is in the money at time 1 (respectively time 2) if $X < K_1$ (respectively $Y < K_2$). If the inequality is reversed, then the put is out of the money. If $K_1 \leq \ell_\mu$, then the American put is always out of the money at time 1, and the American put is equivalent to the European put with strike K_2 and maturity 2. Since puts with strike K and maturity 1 are costless for $K \leq \ell_\mu$, a simple superhedging strategy is to purchase one European put with strike K_1 and sell one European put with strike K_2 , both with maturity 1, and also purchase one European put with strike K_2 and maturity 2. (This strategy is of the form discussed in Lemma 2.4 and is generated by $\psi(y) = (K_2 - y)^+$.) The cost of this hedge is $P_v(K_2)$; this is also the model-based expected payoff of the American put under any consistent model.

If $K_1 \leq K_2$, then $\mathbb{E}[(K_2 - Y)^+ | X] \geq (K_2 - X)^+ \geq (K_1 - X)^+$ and $\tau = 2$ is optimal. Again, the American put is equivalent to the European put with strike K_2 and maturity 2. In this case, for a superhedge, it is sufficient to purchase one European put with strike K_2 and maturity 2. By Lemma 2.4 (with $\psi(y) = (K_2 - y)^+$ and $\phi = 0$), this generates a superhedge with cost $P_v(K_2)$. Again, this is the model-based expected payoff of the American put under any consistent model.

For the remainder of the paper, we make the

Standing Assumption 3.2 $K_1 > \max\{\ell_\mu, K_2\}$.

3.2 American puts under the dispersion assumption

3.2.1 The left-curtain coupling

The goal in this section is to present the theory in a simple special case, and to illustrate the main features and solution techniques of our approach unencumbered by technical issues or the consideration of exceptional cases. The following assumption is a small modification of one introduced by Hobson and Klimmek [18]; see also Henry-Labordère and Touzi [16]. For illustration, see Fig. 3.1.

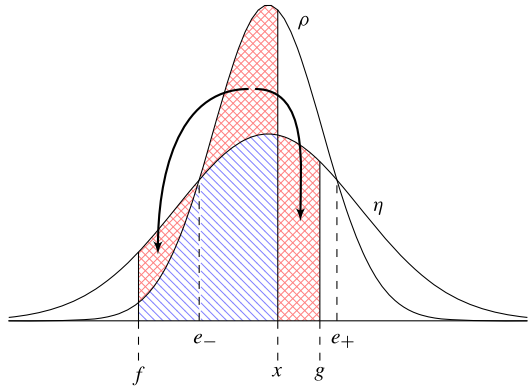
Assumption 3.3 (Dispersion assumption) Laws μ and ν are absolutely continuous with continuous densities ρ and η , respectively; ν has support on $(\ell_\nu, r_\nu) \subseteq (-\infty, \infty)$ and $\eta > 0$ on (ℓ_ν, r_ν) ; μ has support on $(\ell_\mu, r_\mu) \subseteq (\ell_\nu, r_\nu)$ and $\rho > 0$ on (ℓ_μ, r_μ) . In addition:

- $(\mu - \nu)^+$ is concentrated on an interval $E = (e_-, e_+)$ and $\rho > \eta$ on E ;
- $(\nu - \mu)^+$ is concentrated on $(\ell_\nu, r_\nu) \setminus E$ and $\eta > \rho$ on $(\ell_\nu, e_-) \cup (e_+, r_\nu)$.

If $\mu \leq_{\text{cx}} \nu$ are centred normal distributions with different variances or distinct lognormal distribution with common mean, then Assumption 3.3 is satisfied.

Under the dispersion assumption, $\{k : D_{\mu, \nu}(k) > 0\}$ is an interval and $D = D_{\mu, \nu}$ is convex to the left of e_- , concave on (e_-, e_+) and again convex above e_+ .

Fig. 3.1 Sketch of the densities ρ and η and the locations of $f = f(x)$, $g = g(x)$ for given $x > e_-$. Time-1 mass in the interval (f, x) stays in the same place if possible. Mass which cannot stay constant is mapped to (f, e_-) or (x, g) in a way which respects the martingale property



Lemma 3.4 (Henry-Labordère and Touzi [16, Sect. 3.4]) *Suppose that Assumption 3.3 holds. For all $x \in (e_-, r_\mu)$, there exist f, g with $f < e_- < x < g$ such that (2.4) and (2.5) hold. Moreover, if we consider f and g as functions of x on (e_-, r_μ) , then f and g are continuous, f is strictly decreasing and g is strictly increasing, $\lim_{x \downarrow e_-} f(x) = e_- = \lim_{x \downarrow e_-} g(x)$, $\lim_{x \uparrow r_\mu} f(x) = \ell_v$ and $\lim_{x \uparrow r_\mu} g(x) = r_v$.*

The principle behind the left-curtain martingale coupling in Beiglböck and Juillet [4] is that they determine where to map mass at x at time 1 sequentially working from left to right. In our current setting, there is an interval $(\ell_\mu, e_-]$ on which mass can remain unmoved between times 1 and 2. To the right of e_- , we can define f, g in such a way that mass is moved as little as possible. This leads to the ODEs in Remark 2.8.

3.2.2 The American put

Suppose $K_1 \in (e_-, r_\mu]$ and suppose f and g are constructed as in Lemma 3.4. Define $\Lambda : [g^{-1}(K_1), K_1] \rightarrow \mathbb{R}$ by

$$\begin{aligned} \Lambda(x) &= \frac{(K_2 - f(x)) - (K_1 - x)}{x - f(x)} - \frac{K_1 - x}{g(x) - x} \\ &= \frac{g(x) - K_1}{g(x) - x} - \frac{K_1 - K_2}{x - f(x)}. \end{aligned} \tag{3.1}$$

Pictorially Λ is the difference in slope of the two dashed lines in Fig. 3.2.

Lemma 3.5 *Suppose $K_1 \in (e_-, r_\mu]$ and $f(K_1) < K_2$. Then there is a unique scalar $x^* = x^*(\mu, v; K_1, K_2) \in (g^{-1}(K_1), K_1)$ such that $\Lambda(x^*) = 0$. Moreover, $f(x^*) < K_2$ and*

$$\begin{aligned} \frac{K_2 - f(x^*)}{g(x^*) - f(x^*)} &= \frac{K_1 - x^*}{g(x^*) - x^*} = \frac{(x^* - f(x^*)) - (K_1 - K_2)}{x^* - f(x^*)} \\ &= 1 - \frac{K_1 - K_2}{x^* - f(x^*)}. \end{aligned} \tag{3.2}$$

Fig. 3.2 Sketch of put payoffs with points x , f and g marked. $\Lambda(x)$ is the difference in slope of the two dashed lines

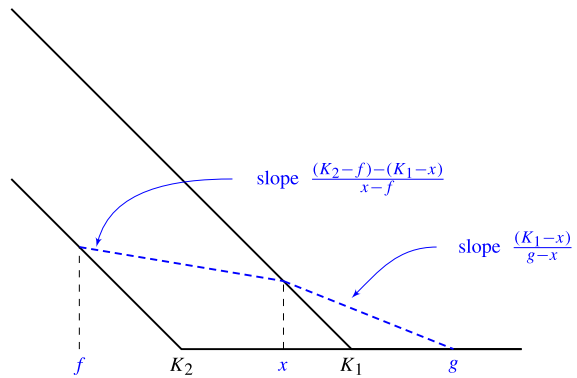
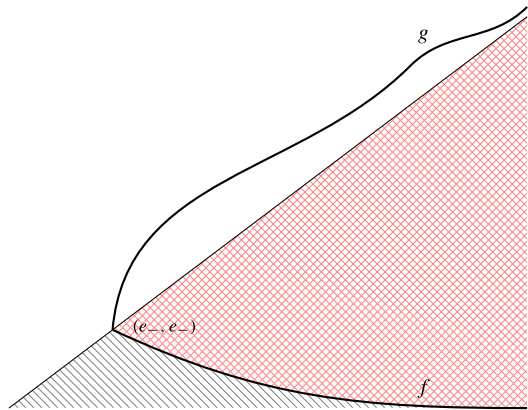


Fig. 3.3 Sketch of functions f and g under the dispersion assumption, with the regions $K_2 < f(K_1)$ and $K_2 > f(K_1)$ shaded. This is a simple special case of Fig. 2.1



Proof First, from the continuity and monotonicity properties of f and g , we have that (see Fig. 3.2) Λ is continuous and strictly increasing. Moreover, $\Lambda(g^{-1}(K_1)) = -\frac{K_1 - K_2}{g^{-1}(K_1) - (f \circ g^{-1})(K_1)} < 0$ and $\Lambda(K_1) = \frac{K_2 - f(K_1)}{K_1 - f(K_1)} > 0$ by hypothesis. Hence there is a unique root to $\Lambda = 0$. At this root, the equalities in (3.2) hold. \square

Suppose $K_1 > e_-$ and $f(K_1) < K_2$ and that $x^* = x^*(\mu, \nu; K_1, K_2) \in (e_-, K_1)$ is such that $\Lambda(x^*) = 0$. It is easy to find a martingale coupling π of μ and ν such that π maps $(f(x^*), x^*)$ onto $(f(x^*), g(x^*))$, and such that π is constant on $(-\infty, f^*)$. For example, we may take $\hat{\pi} = \pi_{lc} = \pi_{lc}(\mu, \nu)$, the left-curtain martingale coupling of Beiglböck and Juillet [4]. More generally, let $\pi_{x^*} \in \Pi(\mu, \nu)$ be any martingale coupling such that π_{x^*} maps $(-\infty, f(x^*))$ to itself, maps $(f(x^*), x^*)$ onto $(f(x^*), g(x^*))$, and maps (x^*, ∞) to $(-\infty, f(x^*)) \cup (g(x^*), \infty)$. The martingale coupling represented in Fig. 3.3 has this property.

Consider a canonical (μ, ν) -consistent model $\hat{M}_{\pi_{x^*}}$, under which the corresponding probability measure $\hat{\mathbb{P}}$ is given by $\hat{\mathbb{P}}[X \in dx, Y \in dy] = \pi_{x^*}(dx, dy)$. Let τ^* be the stopping time such that $\tau^* = 1$ on $(-\infty, x^*)$ and $\tau^* = 2$ otherwise. Our claim in Theorem 3.6 below is that $\hat{M}_{\pi_{x^*}}$ and the stopping time τ^* are such that the model-based price of the American put under this stopping time is the highest possible, over all consistent models.

Continue to suppose $K_1 > e_-$ and $f(K_1) < K_2$. Now we define a superhedge of the American put. Let ψ^* be the function

$$\psi^*(z) = \begin{cases} K_2 - z, & z \leq f(x^*), \\ \frac{(g(x^*)-z)(K_2-f(x^*))}{g(x^*)-f(x^*)}, & f(x^*) < z \leq g(x^*), \\ 0, & z > g(x^*). \end{cases} \tag{3.3}$$

Note that by construction and by (3.2), $\frac{K_2-f(x^*)}{g(x^*)-f(x^*)} = \frac{K_1-x^*}{g(x^*)-x^*}$. Therefore, we have that $\psi^*(x^*) = K_1 - x^*$. Moreover, ψ^* is convex and satisfies $\psi^*(z) \geq (K_2 - z)^+$. Hence by Lemma 2.4, ψ^* can be used to generate a superhedge $(\psi^*, \phi^*, (\theta_i^*)_{i=1,2})$.

In the following theorem, we assume that the American put is not always strictly in the money at time 1 (or equivalently, $K_1 \leq r_\mu$). Discussion of the case $K_1 > r_\mu$ is postponed until Sect. 3.3.5 below.

Theorem 3.6 *Suppose Assumption 3.3 holds and $K_1 \leq r_\mu$.*

(1) *Suppose $K_1 \in (e_-, r_\mu]$ and $f(K_1) < K_2$. The model $\hat{M}_{\tau_{x^*}}$ described in the previous paragraphs is a canonical (μ, ν) -consistent model for which the price of the American option is the highest. The stopping time τ^* is the optimal exercise time. The function ψ^* defined in (3.3) defines the cheapest superhedge. Moreover, the highest model-based price is equal to the cost of the cheapest superhedge.*

(2) *Suppose that we have either Case A: $K_1 \leq e_-$ or Case B: $K_1 \in (e_-, r_\mu]$, together with $f(K_1) \geq K_2$. Then there exists a canonical, consistent model under which*

$$\{Y < K_2\} = \{X < K_2\} \cup \{X > K_1, Y < K_2\} \quad \text{a.s.}$$

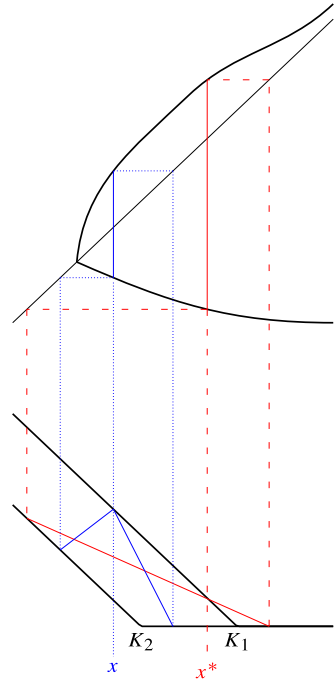
and any model with this property, with the stopping rule $\tau = 1$ if $X < K_1$ and $\tau = 2$ otherwise, attains the highest consistent model price. The cheapest superhedge is generated from $\psi(x) = (K_2 - x)^+$, and the highest model-based price is equal to the cost of the cheapest superhedge.

Remark 3.7 In Part 2 of Theorem 3.6, the left-curtain coupling generates a model for which $\{Y < K_2\} = \{X < K_2\} \cup \{X > K_1, Y < K_2\}$, and hence when associated with the stopping rule of the theorem, this attains the highest consistent model price.

Proof (1) Suppose $K_1 > e_-$ and $f(K_1) < K_2$. Then by Lemma 3.5, there is a unique $x^* \in (g^{-1}(K_1), K_1)$ such that $\Lambda(x^*) = 0$. For this x^* , we can find $f^* = f(x^*)$ and $g^* = g(x^*)$ with $f^* < K_2$ and $K_1 < g^*$ such that $\frac{K_2-f^*}{g^*-f^*} = \frac{K_1-x^*}{g^*-x^*}$; see Fig. 3.4. For typographical reasons, we abbreviate this (x^*, f^*, g^*) to (x, f, g) **for the rest of this proof**.

Since ν is continuous, we have that f, x, g solve (2.4) and (2.5). The elements f, x, g can be used to define a model using the construction after Lemma 3.5 above. For this model, we can calculate the expected payoff of the American put. At the same time, we can use (f, x, g) to define a superhedge. The remaining task is to show that the cost of the superhedge equals that of the model-based expected payoff. Then by the discussion in Sect. 2.3, we have found an optimal model and a cheapest superhedge.

Fig. 3.4 A combination of Figs. 3.3 and 3.2, showing how they jointly define the best model and best hedge. By adjusting x , we can find x^* such that $\Lambda(x^*) = 0$. Together the quantities $(f(x^*), x^*, g(x^*))$ define the optimal model, stopping time and hedge



The model-based expected payoff (MBEP) of the American put (for this model and stopping rule) is

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^x (K_1 - w)\mu(dw) + \int_{-\infty}^f (K_2 - w)(v - \mu)(dw) \\ &= P_\mu(x) + (K_1 - x)P'_\mu(x) + D(f) + (K_2 - f)D'(f). \end{aligned}$$

Now we consider the hedging cost (HC). Set $\Theta = \frac{K_2 - f}{g - f} \in (0, 1)$. Note that since x has $\Lambda(x) = 0$, we have $\Theta = \frac{K_1 - x}{g - x}$. Recall the definition of ψ^* in (3.3). Then

$$\psi^*(y) = \Theta(g - y)^+ + (1 - \Theta)(f - y)^+.$$

Following Lemma 2.4, we can use ψ^* to generate a superhedging strategy. The hedging cost (HC) of this strategy is

$$\text{HC} = \Theta P_v(g) + (1 - \Theta)P_v(f) + (1 - \Theta)(P_\mu(x) - P_\mu(f)), \tag{3.4}$$

where the first two terms arise from the purchase of the static time-2 portfolio ψ^* and the third comes from the purchase of the time-1 portfolio $((K_1 - w)^+ - \psi^*(w))^+$. The expression in (3.4) can be rewritten as

$$P_\mu(x) + D(f) + \Theta(P_v(g) - P_v(f) - P_\mu(x) + P_\mu(f)).$$

Now we consider the difference between the hedging cost and the model-based expected payoff. First recall that $P_\chi(k) = \int_{-\infty}^k (k - x)\chi(dx)$, $\chi \in \{\mu, \nu\}$, and that $D(k) = D_{\mu,\nu}(k) = P_\nu(k) - P_\mu(k)$. Then (2.4) and (2.5) can be rewritten as

$$P'_\mu(x) - P'_\mu(f) = P'_\nu(g) - P'_\nu(f), \tag{3.5}$$

$$\begin{aligned} (xP'_\mu(x) - P_\mu(x)) - (fP'_\mu(f) - P_\mu(f)) &= (gP'_\nu(g) - P_\nu(g)) \\ &\quad - (fP'_\nu(f) - P_\nu(f)). \end{aligned} \tag{3.6}$$

We find

$$\begin{aligned} \text{HC} - \text{MBEP} &= \Theta(P_\nu(g) - P_\nu(f) - P_\mu(x) + P_\mu(f)) \\ &\quad - (K_1 - x)P'_\mu(x) - (K_2 - f)D'(f) \\ &= \Theta(gP'_\nu(g) - xP'_\mu(x) - fD'(f)) \\ &\quad - (K_1 - x)P'_\mu(x) - (K_2 - f)D'(f) \\ &= \Theta((g - x)P'_\mu(x) + (g - f)D'(f)) \\ &\quad - (K_1 - x)P'_\mu(x) - (K_2 - f)D'(f) \\ &= P'_\mu(x)(\Theta(g - x) - (K_1 - x)) + D'(f)(\Theta(g - f) - (K_2 - f)) \\ &= 0, \end{aligned}$$

where we use (3.6), (3.5) and the definition of Θ , respectively. Optimality of the model, stopping rule and hedge now follows.

(2) Now suppose $K_1 \leq e_-$. Consider an exercise rule in which the American put is exercised at time 1 if it is in the money, otherwise it is exercised at time 2, and a model in which mass below K_1 at time 1 stays constant between times 1 and 2. (This is possible since $\mu \leq \nu$ on $(-\infty, e_-)$ and $K_1 \leq e_-$.) The expected payoff of the American put is

$$\int_{-\infty}^{K_1} (K_1 - x)\mu(dx) + \int_{-\infty}^{K_2} (K_2 - y)(\nu - \mu)(dy) = P_\mu(K_1) + P_\nu(K_2) - P_\mu(K_2). \tag{3.7}$$

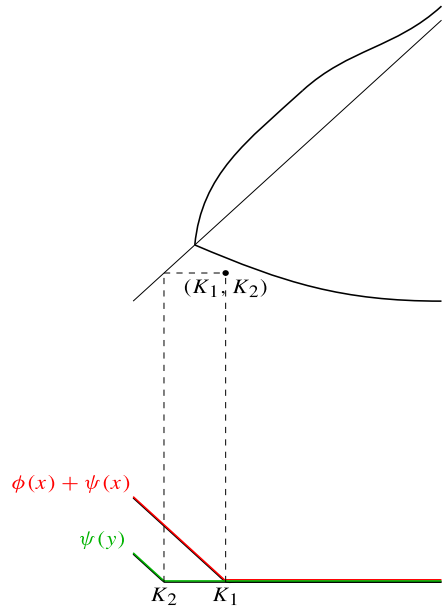
Alternatively, suppose $K_1 > e_-$ and $f(K_1) \geq K_2$. Then under the left-curtain martingale coupling, mass below K_2 at time 1 stays constant between times 1 and 2 (note that $K_2 \leq f(K_1) \leq e_-$), and mass between K_2 and K_1 at time 1 is mapped to (K_2, ∞) . Then, mass which is below K_2 at time 2 was either below K_2 at time 1, or above K_1 at time 1. The expected payoff under this model (using a strategy of exercising at time 1 if the American put is in the money) is again given by (3.7).

Now consider the hedging cost. Let $\psi(y) = (K_2 - y)^+$. If we define ϕ as in Lemma 2.4, we find $\phi(x) = (K_1 - x)^+ - (K_2 - x)^+ = (K_1 - (x \vee K_2))^+$, see Fig. 3.5, and the superhedging cost is

$$\text{HC} = P_\nu(K_2) + P_\mu(K_1) - P_\mu(K_2).$$

Hence the model-based expected payoff equals the hedging cost. □

Fig. 3.5 Sketch of put payoffs with $\psi(y) = (K_2 - y)^+$ and $\phi(x) = (K_1 - x)^+ - (K_2 - x)^+$



3.3 Two intervals of $g > x$ and one downward jump in f

We now relax the dispersion assumption to the case where f is not monotone. The simplest situation when this may arise is when there are two intervals on which $g(x) > x$. We do not contend that there are many natural examples which fall into this situation, but rather that this intermediate case illustrates phenomena which are to be found in the general case, but which were not to be found under the dispersion assumption.

Assumption 3.8 (Single-jump assumption) Laws μ and ν are absolutely continuous with continuous densities ρ and η , respectively; ν has support on $(\ell_\nu, r_\nu) \subseteq (-\infty, \infty)$ and $\eta > 0$ on (ℓ_ν, r_ν) ; μ has support on $(\ell_\mu, r_\mu) \subseteq (\ell_\nu, r_\nu)$ and $\rho > 0$ on (ℓ_μ, r_μ) . In addition:

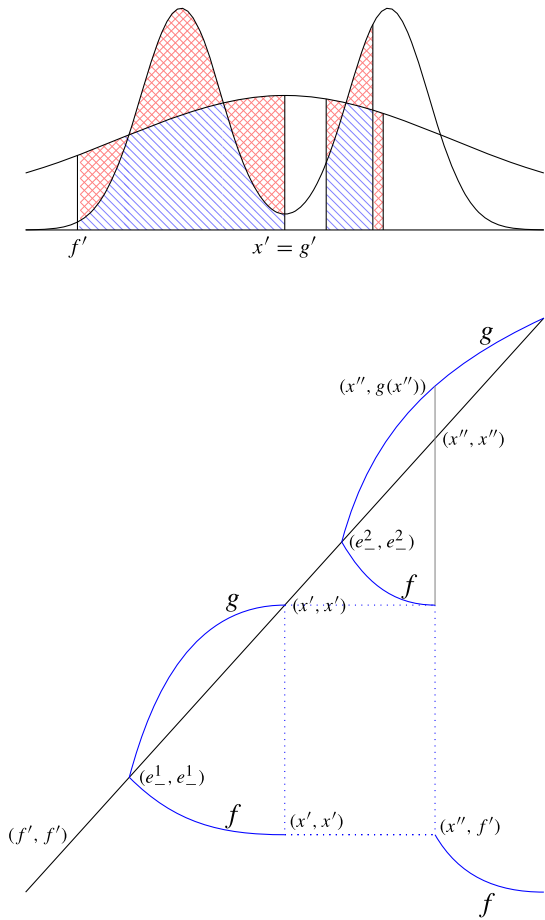
- $(\mu - \nu)^+$ is concentrated on $E = (e_-^1, e_+^1) \cup (e_-^2, e_+^2)$ with $e_+^1 < e_-^2$, and $\rho > \eta$ on E ;
- $(\nu - \mu)^+$ is concentrated on $(\ell_\nu, r_\nu) \setminus E$ and $\eta > \rho$ on $(\ell_\nu, e_-^1) \cup (e_+^1, e_-^2) \cup (e_+^2, r_\nu)$;
- there exist $f' < e_-^1$ and $x' \in (e_+^1, e_-^2)$ such that

$$\int_{f'}^{x'} \mu(dz) = \int_{f'}^{x'} \nu(dz) \quad \text{and} \quad \int_{f'}^{x'} z\mu(dz) = \int_{f'}^{x'} z\nu(dz). \quad (3.8)$$

Under Assumption 3.8, it is possible to find functions $g : (\ell_\mu, r_\mu) \rightarrow (\ell_\nu, r_\nu)$ and $f : (\ell_\mu, r_\mu) \rightarrow (\ell_\nu, r_\nu)$ with the properties (see the lower part of Fig. 3.6)

- (1) $g(x) = x$ on $(\ell_\mu, e_-^1] \cup [x', e_-^2]$;
- (2) $g(x) > x$ on $(e_-^1, x') \cup (e_-^2, r_\mu)$;

Fig. 3.6 Picture of f and g under Assumption 3.8



- (3) g is continuous and strictly increasing;
- (4) $f(x) = x$ on $(\ell_\mu, e_1^1] \cup [x', e_2^2]$;
- (5) $f : (e_1^1, x') \rightarrow (f', x')$ is continuous and strictly decreasing;
- (6) $f : (e_2^2, r_\mu) \rightarrow (\ell_\nu, e_2^2) \setminus (f', x')$ is strictly decreasing;
- (7) there exists $x'' \in (e_2^2, r_\mu)$ with the property that f jumps at x'' and satisfies $f(x''-) = x' > f' = f(x''+)$. Away from x'' , f is continuous on (e_2^2, r_μ) .

By construction, we have that

$$\int_{x'}^{x''} \mu(dz) = \int_{x'}^{g(x'')} \nu(dz) \quad \text{and} \quad \int_{x'}^{x''} z\mu(dz) = \int_{x'}^{g(x'')} z\nu(dz), \quad (3.9)$$

so that if mass in (x', x'') at time 1 is mapped to $(x', g(x''))$ at time 2, then the total mass and mean are preserved. Further, given that (f', x') satisfy (3.8), we also have that $\int_{f'}^{x''} \mu(dz) = \int_{f'}^{g(x'')} \nu(dz)$ and $\int_{f'}^{x''} z\mu(dz) = \int_{f'}^{g(x'')} z\nu(dz)$. In particular, given

(3.8) and (3.9), the pair of equations

$$\int_f^{x''} \mu(dz) = \int_f^{g(x'')} v(dz) \quad \text{and} \quad \int_f^{x''} z\mu(dz) = \int_f^{g(x'')} zv(dz)$$

has two solutions for f , namely $f=x'$ and $f=f'$. Hence, in defining the left-curtain martingale coupling, there are two choices for f at x'' : we may take $f(x'') = x'$ or $f(x'') = f'$. Rather than assuming one of these choices (for example by requiring left-continuity of f), it is convenient to allow f to be multi-valued. Then, for each x such that $g(x) > x$, let $\aleph(x) = \{f : (f, x, g(x)) \text{ solves (2.4) and (2.5)}\}$. Then we have that in the setting of Assumption 3.8, for $x > e_-$, $|\aleph(x)| = 1$ except at x'' and there $\aleph(x'') = \{f(x''+), f(x''-)\} = \{f', x'\}$.

Recall the definition $\Lambda(x) = \frac{g(x)-K_1}{g(x)-x} - \frac{K_1-K_2}{x-f(x)}$. If f is multi-valued, then Λ is also multi-valued. In Sect. 3.2, one of our main steps was to find x such that $\Lambda(x) = 0$, and our aim is similar here.

Introduce $\Upsilon = \Upsilon_{K_1, K_2}(f, x, g)$ which is defined for $f \leq K_2, x \leq K_1 \leq g$ by

$$\Upsilon(f, x, g) = \frac{(K_2 - f) - (K_1 - x)}{x - f} - \frac{K_1 - x}{g - x} = \frac{g - K_1}{g - x} - \frac{K_1 - K_2}{x - f}.$$

Instead of seeking x which is a root of $\Lambda(x) = 0$, our goal is to find (f, x, g) with $g = g(x)$ and $f \in \aleph(x)$ such that $\Upsilon(f, x, g) = 0$.

For a fixed K_1 , the value of K_2 such that $\Upsilon(f' = f(x''+), x'', g(x'')) = 0$ is given by $K_2 = f' + (K_1 - x'') \frac{g(x'')-f'}{g(x'')-x''}$. On the other hand, if we define K_2 by $K_2 = x' + (K_1 - x'') \frac{g(x'')-x'}{g(x'')-x''}$, we get $\Upsilon(x' = f(x''-), x'', g(x'')) = 0$. This motivates the introduction of the linear increasing functions $L_u, L_d : [x'', g(x'')] \rightarrow \mathbb{R}$ defined by

$$L_u(x) = x' + (x - x'') \frac{g(x'') - x'}{g(x'') - x''}, \tag{3.10}$$

$$L_d(x) = f' + (x - x'') \frac{g(x'') - f'}{g(x'') - x''}. \tag{3.11}$$

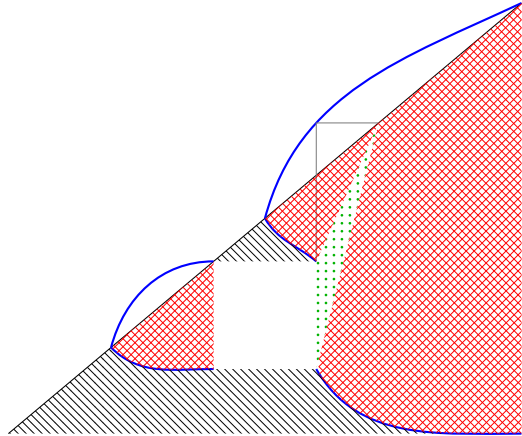
Pictorially, L_d and L_u are the lower and upper boundaries, respectively, of the dotted triangular area \mathcal{G} in Fig. 3.7.

From Fig. 3.7, we identify four regions (and various subregions) on which four different hedging strategies will be needed in order to find the cheapest superhedge for the American put. (Compare this with two regimes under the dispersion assumption in Fig. 3.3.)

Define

$$\mathcal{R}_1 = \{(k_1, k_2) : e_-^1 < k_1 < x', f(k_1) < k_2 < k_1\},$$

Fig. 3.7 Picture of f and g in the single-jump case, now with 4 regions shaded (cross-hatched, diagonally, dotted and blank)



which we write more compactly as $\mathcal{R}_1 = \{e_-^1 < k_1 < x', f(k_1) < k_2 < k_1\}$. Using the same compact notation, define

$$\begin{aligned} \mathcal{R}_2 &= \{e_-^2 < k_1 < x'', f(k_1) < k_2 < k_1\} \cup \{k_1 = x'', x' < k_2 < k_1\}, \\ \mathcal{R}_3 &= \{x'' < k_1 < g(x''), L_u(k_1) \leq k_2 < k_1\}, \\ \mathcal{R}_4 &= \{x'' < k_1 < g(x''), f(k_1) < k_2 \leq L_d(k_1)\}, \\ \mathcal{R}_5 &= \{g(x'') \leq k_1 \leq r_\mu, f(k_1) < k_2 < k_1\}, \\ \mathcal{B}_1 &= \{\ell_\mu \leq k_1 \leq e_-^1, k_2 < k_1\} \cup \{e_-^1 < k_1 < x', k_2 \leq f(k_1)\} \\ &\quad \cup \{x'' < k_1 \leq r_\mu, k_2 \leq f(k_1)\}, \\ \mathcal{B}_2 &= \{x' \leq k_1 \leq x'', k_2 \leq f'\}, \\ \mathcal{B}_3 &= \{x' \leq k_1 \leq e_-^2, x' \leq k_2 < k_1\} \cup \{e_-^2 < k_1 \leq x'', x' \leq k_2 \leq f(k_1)\}, \\ \mathcal{G} &= \{x'' < k_1 < g(x''), L_d(k_1) < k_2 < L_u(k_1)\}, \\ \mathcal{W} &= \{x' \leq k_1 \leq x'', f' < k_2 < x'\}, \end{aligned}$$

and set $\mathcal{R} = \bigcup_{i=1}^5 \mathcal{R}_i$ and $\mathcal{B} = \bigcup_{i=1}^3 \mathcal{B}_i$. In general, on the boundaries between the regions, the boundaries could be allocated to either region. However, we allocate points on the boundary to the region where the hedge is simplest.

Note that $\mathcal{R} \cup \mathcal{B} \cup \mathcal{G} \cup \mathcal{W} = \{(k_1, k_2) : \ell_\mu \leq k_1 \leq r_\mu, k_2 < k_1\}$.

3.3.1 Case $(K_1, K_2) \in \mathcal{R}$

Lemma 3.9 *Suppose that we have $(K_1, K_2) \in \mathcal{R}$. Then there exist a unique $x^* = x^*(\mu, v; K_1, K_2) \in (g^{-1}(K_1), K_1)$ and $f^* \in \mathfrak{N}(x^*)$ such that $\Upsilon(f^*, x^*, g^* = g(x^*)) = 0$.*

Proof Suppose that $(K_1, K_2) \in \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_5$. Consider $\Lambda : [g^{-1}(K_1), K_1] \rightarrow \mathbb{R}$ defined by (3.1). Note that for this choice of (K_1, K_2) , f and g are both continuous

on $[g^{-1}(K_1), K_1]$; see Fig. 3.6. Hence $\Lambda(x) = \Upsilon(f(x), x, g(x))$ is also continuous. Then the same argument as in the proof of Lemma 3.5 shows that there exists a unique $x^* = x^*(\mu, \nu; K_1, K_2) \in (g^{-1}(K_1), K_1)$ such that $\Lambda(x^*) = 0$.

Now suppose $(K_1, K_2) \in \mathcal{R}_3 \cup \mathcal{R}_4$ and consider Λ as before. Recall that Λ is increasing, $\Lambda(g^{-1}(K_1)) < 0$ and $\Lambda(K_1) > 0$. On the other hand, $g^{-1}(K_1) < x''$ and hence Λ has an upward jump at x'' (since f has a downward jump at x''). There are two cases depending on whether $(K_1, K_2) \in \mathcal{R}_3$ or \mathcal{R}_4 .

(1) Suppose that $K_2 > L_u(K_1)$. Then we have $\Lambda(x''-) > 0$. Moreover, since $\Lambda(g^{-1}(K_1)) < 0$, the continuity of Λ on $(g^{-1}(K_1), x'')$ guarantees that there exists a unique scalar $x^* = x^*(\mu, \nu; K_1, K_2) \in (g^{-1}(K_1), x'')$ such that $\Lambda(x^*) = 0$. If $K_2 = L_u(K_1)$, then $\Upsilon(x', x''g(x'')) = 0$ and we take $x^* = x''$, $g^* = g(x'')$ and $f^* = f(x''-) = x'$.

(2) Suppose that $K_2 < L_d(K_1)$. Then $\Lambda(x''+) < 0$. Further, since $\Lambda(K_1) > 0$, there exists a unique $x^* = x^*(\mu, \nu; K_1, K_2) \in (x'', K_1)$ such that $\Lambda(x^*) = 0$. If $K_2 = L_d(K_1)$, then $\Upsilon(f', x'', g(x'')) = 0$ and we take $x^* = x''$, $g^* = g(x'')$ and $f^* = f(x''+) = f'$. □

By Lemma 3.9, for $(K_1, K_2) \in \mathcal{R}$, there exists $(f^* \in \mathfrak{N}(x^*), x^*, g^* = g(x^*))$ such that $\Upsilon(f^*, x^*, g^*) = 0$. Suppose $(K_1, K_2) \in \mathcal{R}_1 \cup \mathcal{R}_4 \cup \mathcal{R}_5$. In this case, we let $\hat{M}_{\pi_{x^*}}$ be a canonical (μ, ν) -consistent model (recall that \hat{M}_π is the abbreviated notation for the canonical model (\hat{S}_π, \hat{M}) for which $\hat{\mathbb{P}}[X \in dx, Y \in dy] = \pi(dx, dy)$). Here $\pi_{x^*} \in \Pi(\mu, \nu)$ is a martingale coupling that is constant on $(-\infty, f^*)$, maps (f^*, x^*) onto (f^*, g^*) and (g^*, ∞) to $(-\infty, f^*) \cup (g^*, \infty)$.

Recall the proof of Theorem 3.6. There, to show that MBEP = HC, we used the fact that for the canonical model $\hat{M}_{\pi_{x^*}}$, π_{x^*} is constant on $(-\infty, f^*)$ and maps (f^*, x^*) onto (f^*, g^*) . In fact, the equality MBEP = HC holds for any canonical model for which the associated martingale coupling has the same property. Then the mass that is ‘unexercised’ at time 1 and is in the money at time 2 has time-2 law given by $(\nu - \mu)|_{(-\infty, f^*)}$, where $f^* < e_-$. When $f(x') < e_-$ (as is the case when $(K_1, K_2) \in \mathcal{R}_1 \cup \mathcal{R}_4 \cup \mathcal{R}_5$), the same proof applies, so that MBEP = HC and we have optimality. On the other hand, if $(K_1, K_2) \in \mathcal{R}_2 \cup \mathcal{R}_3$, then it is not the case that $f^* < e_-$ and thus, in order to specify the optimal model, we need to impose additional structure on the coupling $\tilde{\mu}_{f^*, x^*} \mapsto \tilde{\nu}_{f^*, g^*}$.

Suppose that $(K_1, K_2) \in \mathcal{R}_2 \cup \mathcal{R}_3$. Then $x' < f^*$ so that $(f', x') \cap (f^*, g^*) = \emptyset$. From the defining properties of f' and x' , we see that there exists a martingale coupling, which we term $\pi_{x', x^*} \in \Pi(\mu, \nu)$, which is constant on $(-\infty, f')$ and (x', f^*) , maps (f', x') onto itself and (f^*, x^*) onto (f^*, g^*) , and maps (x^*, ∞) to $(-\infty, f') \cup (x', f^*) \cup (g^*, \infty)$.

If $(K_1, K_2) \in \mathcal{R}_1 \cup \mathcal{R}_4 \cup \mathcal{R}_5$, we have the canonical model $\hat{M}_{\pi_{x^*}}$, and in the case when $(K_1, K_2) \in \mathcal{R}_2 \cap \mathcal{R}_3$, we have $\hat{M}_{\pi_{x', x^*}}$. For both models, we consider a candidate stopping time $\tau^* = 1$ if $X < x^*$ and $\tau^* = 2$ otherwise, and a candidate superhedge $(\psi^*, \phi^*, (\theta_i^*)_{i=1,2})$ generated by the function ψ^* defined in (3.3).

Theorem 3.10 *Suppose Assumption 3.8 holds and $(K_1, K_2) \in \mathcal{R}$. Then, depending on whether (K_1, K_2) is in $\mathcal{R}_1 \cup \mathcal{R}_4 \cup \mathcal{R}_5$ or $\mathcal{R}_2 \cup \mathcal{R}_3$, the models $\hat{M}_{\pi_{x^*}}$ and $\hat{M}_{\pi_{x', x^*}}$*

and the stopping time τ^* are the consistent models for which the price of the American option is the highest. The function ψ^* defined in (3.3) defines the cheapest superhedge. Moreover, the highest model-based price is equal to the cost of the cheapest superhedge.

Proof If $(K_1, K_2) \in \mathcal{R}_1 \cup \mathcal{R}_4 \cup \mathcal{R}_5$, then the proof is essentially the same as the proof of the first case in Theorem 3.6. We repeat the main steps for convenience. First find $x^* \in (g^{-1}(K_1), K_1)$ and $f^* \in \mathfrak{N}(x^*)$ such that $\Upsilon(f^*, x^*, g^* = g(x^*)) = 0$. If $x^* = x''$, we find $f^* = f(x''+) = f'$. Under the candidate model $\hat{M}_{\tau_{x^*}^*}$, mass below f^* at time 1 is mapped to the same point at time 2 (which is possible since $f^* < e_1^1$), and mass in (f^*, x^*) is mapped onto (f^*, g^*) , while mass above x^* is either mapped to below f^* or to above g^* . Then under the candidate stopping rule τ^* , the model-based expected payoff is equal to the cost of the hedging strategy generated by ψ^* , i.e.,

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{x^*} (K_1 - w)^+ \mu(dw) + \int_{-\infty}^{f^*} (K_2 - w)^+ (v - \mu)(dw) \\ &= P_\mu(x^*) + (K_1 - x^*)P'_\mu(x^*) + D(f^*) + (K_2 - f^*)D'(f^*) \\ &= \text{HC}. \end{aligned}$$

Now suppose that $(K_1, K_2) \in \mathcal{R}_2 \cup \mathcal{R}_3$. Then by Lemma 3.9, there exist a unique $x^* \in (g^{-1}(K_1), x'']$ and $f^* \in \mathfrak{N}(x^*)$ such that $\Upsilon(f^*, x^*, g^* = g(x^*)) = 0$. If $x^* = x''$, then we have $f^* = f(x''-) = x'$. Then, since v is continuous, we have that f^*, x^*, g^* solve (2.4) and (2.5). Note, however, that $x' \leq f^* < e_2^-$.

Under the candidate model $\hat{M}_{\tau_{x',x^*}^*}$, mass in (f', x') at time 1 is mapped onto the same interval at time 2. Also, mass below f' and mass in (x', f^*) at time 1 is mapped to the same point at time 2, and mass in (f^*, x^*) is mapped onto (f^*, g^*) . Mass above x^* is either mapped to below f' , to (x', f^*) , or to above g^* . In particular, $(v - \mu)|_{(-\infty, f') \cup (x', f^*)}$ is the law of the mass that was not ‘exercised’ at time 1 and is in the money at time 2. From (3.8), we have $\int_{x'}^{f'} (K_2 - w)(v - \mu)(dw) = 0$. Then

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{x^*} (K_1 - w)\mu(dw) + \int_{-\infty}^{f'} (K_2 - w)(v - \mu)(dw) \\ &\quad + \int_{x'}^{f^*} (K_2 - w)(v - \mu)(dw) \\ &= \int_{-\infty}^{x^*} (K_1 - w)\mu(dw) + \int_{-\infty}^{f^*} (K_2 - w)(v - \mu)(dw) \\ &\quad - \int_{f'}^{x'} (K_2 - w)(v - \mu)(dw) \\ &= \int_{-\infty}^{x^*} (K_1 - w)\mu(dw) + \int_{-\infty}^{f^*} (K_2 - w)(v - \mu)(dw) \\ &= P_\mu(x^*) + (K_1 - x^*)P'_\mu(x^*) + D(f^*) + (K_2 - f^*)D'(f^*) \\ &= \text{HC}. \end{aligned}$$

□

3.3.2 $(K_1, K_2) \in \mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$

Theorem 3.11 *Suppose that Assumption 3.8 holds and $(K_1, K_2) \in \mathcal{B}$. Then there is a consistent model for which $\{Y < K_2\} = \{X < K_2\} \cup \{X > K_1, Y < K_2\}$ a.s. and, if $x' < K_2$, $\{f' < X < x'\} = \{f' < Y < x'\}$. Any model with these properties and with the stopping rule $\tau = 1$ if $X < K_1$ and $\tau = 2$ otherwise attains the highest consistent model price. The cheapest superhedge is generated from $\psi(x) = (K_2 - x)^+$, and the highest model-based price is equal to the cost of the cheapest hedge.*

Proof Let $\psi(y) = (K_2 - y)^+$. As in Lemma 2.4, define a corresponding ϕ . We find that $\phi(x) = (K_1 - x)^+ - (K_2 - x)^+$, and the superhedging cost (which is the same for all the cases) is

$$HC = P_v(K_2) + P_\mu(K_1) - P_\mu(K_2).$$

Suppose $(K_1, K_2) \in \mathcal{B}_1$. Then using the properties of f and g and the left-curtain coupling, we see that the proof that the model-based expected payoff is equal to the hedging cost is the same as in the second case of Theorem 3.6. In particular,

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{K_1} (K_1 - x)\mu(dx) + \int_{-\infty}^{K_2} (K_2 - y)(v - \mu)(dy) \\ &= P_\mu(K_1) + P_v(K_2) - P_\mu(K_2). \end{aligned}$$

Now suppose $(K_1, K_2) \in \mathcal{B}_2$. Then under the left-curtain coupling, mass from (f', x') at time 1 is mapped onto the same interval at time 2. Therefore mass which is below K_2 at time 2 was either below K_2 at time 1, or above x' at time 1. Therefore, we again have

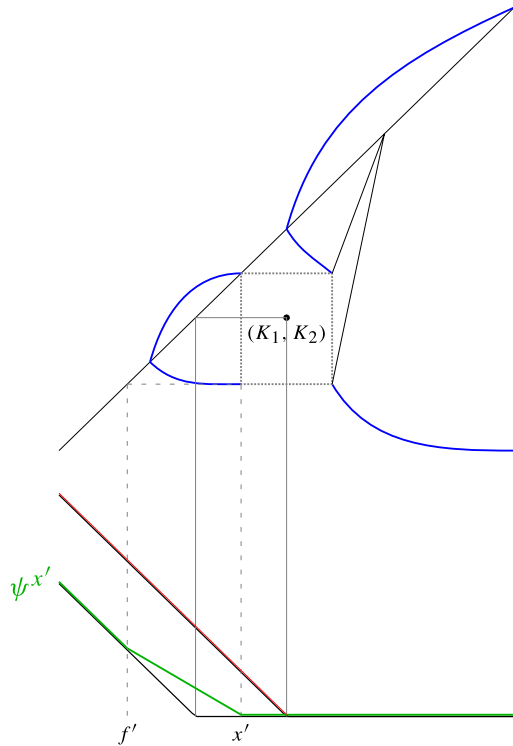
$$\text{MBEP} = \int_{-\infty}^{K_1} (K_1 - x)\mu(dx) + \int_{-\infty}^{K_2} (K_2 - y)(v - \mu)(dy).$$

Finally, suppose $(K_1, K_2) \in \mathcal{B}_3$. We again utilise the fact that under the left-curtain coupling, mass from (f', x') at time 1 is mapped onto the same interval at time 2. In both cases, the mass which is below K_2 at time 2 was either below K_2 at time 1, or above K_1 at time 1. In particular, the law of the mass that can be ‘exercised’ at time 2 is given by $(v - \mu)|_{(-\infty, f') \cup (x', K_2)}$. Then using $\int_{f'}^{x'} (K_2 - z)(v - \mu)(dz) = 0$, we again have

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{K_1} (K_1 - x)\mu(dx) + \int_{-\infty}^{f'} (K_2 - y)(v - \mu)(dy) \\ &\quad + \int_{x'}^{K_2} (K_2 - y)(v - \mu)(dy) \\ &= \int_{-\infty}^{K_1} (K_1 - x)\mu(dx) + \int_{-\infty}^{K_2} (K_2 - y)(v - \mu)(dy), \end{aligned}$$

which ends the proof. □

Fig. 3.8 Picture of f and g along with superhedge for $(K_1, K_2) \in \mathcal{W}$



3.3.3 $(K_1, K_2) \in \mathcal{W}$

Suppose $(K_1, K_2) \in \mathcal{W}$. For this case, we associate the following superhedge: let $\psi^{x'}$ be given by

$$\psi^{x'}(z) = \begin{cases} K_2 - z, & z \leq f', \\ (K_2 - f') - (z - f') \frac{K_2 - f'}{x' - f'}, & f' < z \leq x', \\ 0, & z > x'; \end{cases} \quad (3.12)$$

see Fig. 3.8. Since $\psi^{x'}$ is convex and $\psi^{x'}(z) \geq (K_2 - z)^+$, we can use Lemma 2.4 to generate a corresponding superhedging strategy $(\psi^{x'}, \phi^{x'}, (\theta_i^{x'})_{i=1,2})$.

Theorem 3.12 *Suppose Assumption 3.8 holds and $(K_1, K_2) \in \mathcal{W}$. Then there is a consistent model for which $\{f' < X < x'\} = \{f' < Y < x'\}$, and any model with this property and with the stopping rule $\tau = 1$ if $X < K_1$ and $\tau = 2$ otherwise attains the highest consistent model price. The cheapest superhedge is generated from $\psi^{x'}$ defined in (3.12), and the highest model-based price is equal to the cost of the cheapest hedge.*

Proof First note that

$$\psi^{x'}(z) = \Theta(x' - z)^+ + (1 - \Theta)(f' - z)^+,$$

where $\Theta = \frac{K_2 - f'}{x' - f'}$. Since $x' < K_1$, we have

$$\begin{aligned} \phi^{x'}(w) + \psi^{x'}(z) &= (K_1 - w)^+ - \psi^{x'}(w) + \psi^{x'}(z) \\ &= (K_1 - w)^+ + \Theta((x' - z)^+ - (x' - w)^+) \\ &\quad + (1 - \Theta)((f' - z)^+ - (f' - w)^+). \end{aligned}$$

It follows that $HC = P_\mu(K_1) + \Theta D(x') + (1 - \Theta)D(f')$ is the cost of this strategy (under any consistent model).

Now consider the model-based expected payoff. From (3.8), it follows that $\mu_{f',x'}$ and $\nu_{f',x'}$ have the same mean and mass and are in convex order. Moreover, the same holds for $\tilde{\mu}_{f',x'}$ and $\tilde{\nu}_{f',x'}$. Therefore, there exists a martingale coupling, which we term $\pi_{x'} \in \Pi(\mu, \nu)$, which is constant on $(-\infty, f')$ and maps (f', x') onto itself. It follows that under this model, the law of the mass that can be ‘exercised’ at time 2 is given by $(\nu - \mu)|_{(-\infty, f')}$.

Note that since f' and x' satisfy (3.8), and hence $\int_{f'}^{x'} (x' - w)(\nu - \mu)(dw) = 0$,

$$\begin{aligned} D(x') - D(f') &= \int_{-\infty}^{x'} (x' - w)^+(\nu - \mu)(dw) - \int_{-\infty}^{f'} (f' - w)^+(\nu - \mu)(dw) \\ &= \int_{-\infty}^{f'} (x' - f')(\nu - \mu)(dw) + \int_{f'}^{x'} (x' - w)(\nu - \mu)(dw) \\ &= (x' - f') \int_{-\infty}^{f'} (\nu - \mu)(dw). \end{aligned}$$

Then given that we stop at time 1 if $X < K_1$ and at time 2 otherwise, we have

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{K_1} (K_1 - w)^+ \mu(dw) + \int_{-\infty}^{f'} (K_2 - w)^+(\nu - \mu)(dw) \\ &= \int_{-\infty}^{K_1} (K_1 - w)\mu(dw) + \int_{-\infty}^{f'} (f' - w)(\nu - \mu)(dw) \\ &\quad + (K_2 - f') \int_{-\infty}^{f'} (\nu - \mu)(dw) \\ &= P_\mu(K_1) + D(f') + \Theta(D(x') - D(f')) \\ &= P_\mu(K_1) + \Theta D(x') + (1 - \Theta)D(f') = HC, \end{aligned}$$

as required. □

3.3.4 $(K_1, K_2) \in \mathcal{G}$

Recall from (3.10), (3.11) the construction of L_u and L_d . For $K_1 \in (x'', g(x''))$ and $K_2 \in (L_d(K_1), L_u(K_1))$, there does not exist any $x^* \in (g^{-1}(K_1), K_1)$ such that $\Lambda(x^*) = 0$; instead we have that $\Lambda(x''-) < 0 < \Lambda(x''+)$. On the other hand,

from (3.9) we have that there exists a martingale coupling of $\mu_{x',x''}$ and $\nu_{x',g(x'')}$. Moreover, note that the restrictions of $\tilde{\mu}_{f',x'}$ to (x', x'') and $\tilde{\nu}_{f',x'}$ to $(x', g(x''))$ are equal to $\mu_{x',x''}$ and $\nu_{x',g(x'')}$, respectively. Then we define a martingale coupling $\pi_{x',x''} \in \Pi(\mu, \nu)$ which is constant on $(-\infty, f')$, maps (f', x') onto itself, (x', x'') onto $(x', g(x''))$, and (x'', ∞) to $(-\infty, f') \cup (g(x''), \infty)$. Let $\hat{M}_{\pi_{x',x''}}$ be the canonical model under which $\hat{\mathbb{P}}[X \in dx, Y \in dy] = \pi_{x',x''}(dx, dy)$. Note that the model $\hat{M}_{\pi_{x',x''}}$ is a refinement of $\hat{M}_{\pi_{x'}}$ used in the proof of Theorem 3.12.

Given x' and thus also x'' , we define the superhedge as follows. First define linear functions $\Delta_1 : [f', x'] \rightarrow \mathbb{R}$ and $\Delta_2 : [x', g(x'')] \rightarrow \mathbb{R}$ by

$$\begin{aligned} \Delta_1(x) &= (K_2 - f') - (x - f') \frac{(K_2 - f') - \Delta_2(x')}{x' - f'}, \\ \Delta_2(x) &= (g(x'') - x) \frac{K_1 - x''}{g(x'') - x''}. \end{aligned}$$

Then $\Delta_1(f') = (K_2 - f')$, $\Delta_1(x') = \Delta_2(x')$, $\Delta_2(x'') = K_1 - x''$ and $\Delta_2(g(x'')) = 0$. Moreover, direct calculation shows that $-1 < \Delta'_1(x) < \Delta'_2(x) < 0$. Now define a function $\psi^{x',x''}$ by

$$\psi^{x',x''}(z) = \begin{cases} K_2 - z, & z \leq f', \\ \Delta_1(z), & f' < z \leq x', \\ \Delta_2(z), & x' < z \leq g(x''), \\ 0, & z > g(x''). \end{cases} \tag{3.13}$$

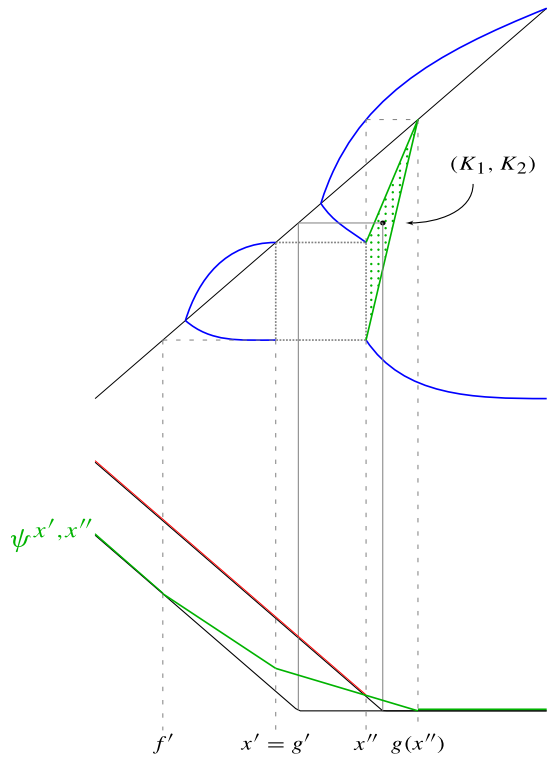
By construction, $\psi^{x',x''}$ is convex and $\psi^{x',x''}(z) \geq (K_2 - z)^+$ (see Fig. 3.9), and thus by Lemma 2.4, it can be used to construct a superhedge $(\psi^{x',x''}, \phi^{x',x''}, \theta_{1,2}^{x',x''})$.

Theorem 3.13 *Suppose Assumption 3.8 holds and $(K_1, K_2) \in \mathcal{G}$. The model $\hat{M}_{\pi_{x',x''}}$ and the stopping time $\tau = 1$ if $X < x''$ and $\tau = 2$ otherwise attain the highest consistent model price. Moreover, $\psi^{x',x''}$ defined in (3.13) generates the cheapest superhedge, and the highest model-based price is equal to the cost of the cheapest superhedge.*

Proof The candidate canonical model is associated with the martingale coupling $\pi_{x',x''}$ which is constant on $(-\infty, f')$, maps (f', x') onto itself, maps (x', x'') onto $(x', g(x''))$, and (x'', ∞) to $(-\infty, f') \cup (g(x''), \infty)$. Then for the candidate stopping time τ (exercise at time 1 if $X < x''$ and at time 2 otherwise), we have that the law of Y (under $\hat{M}_{\pi_{x',x''}}$), on the event that the option was not exercised at time 1, is given by $(\nu - \mu)|_{(-\infty, f')} + \nu|_{(g(x''), \infty)}$. Therefore

$$\begin{aligned} \text{MBEP} &= \int_{-\infty}^{x''} (K_1 - w)^+ \mu(dw) + \int_{-\infty}^{f'} (K_2 - w)^+ (\nu - \mu)(dw) \\ &= P_\mu(x'') + (K_1 - x'')P'_\mu(x'') + D(f') + (K_2 - f')D'(f'). \end{aligned}$$

Fig. 3.9 Picture of f and g along with superhedge for the dotted region \mathcal{G} . The hedge function $\psi^{x',x''}$ has a kink at x'



Now consider the hedging cost generated by $\psi^{x',x''}$. Let us define the quantities $\Theta_1 = \frac{K_2 - f' - \Delta_2(x')}{x' - f'} = -\Delta'_1$ and $\Theta_2 = \frac{K_1 - x''}{g(x'') - x''} = -\Delta'_2$. Note that we can rewrite (3.13) as

$$\psi^{x',x''}(z) = \Theta_2(g(x'') - z)^+ + (\Theta_1 - \Theta_2)(x' - z)^+ + (1 - \Theta_1)(f' - z)^+.$$

Then

$$\phi(z) = (1 - \Theta_1)((x' - z)^+ - (f' - z)^+) + (1 - \Theta_2)((x'' - z)^+ - (x' - z)^+),$$

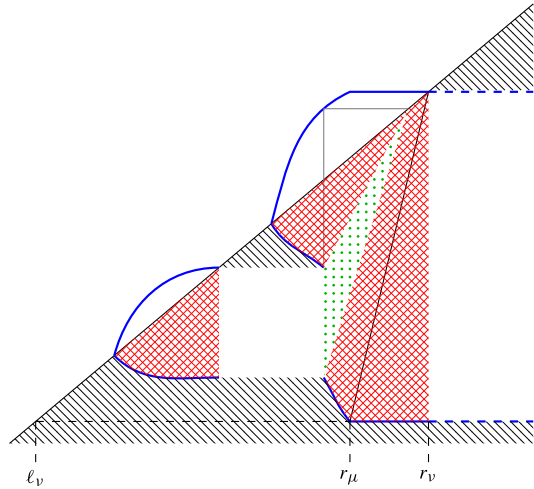
and thus the hedging cost is

$$\begin{aligned} \text{HC} &= \Theta_2 P_v(g(x'')) + (1 - \Theta_1)D(f') + (1 - \Theta_2)P_\mu(x'') + (\Theta_1 - \Theta_2)D(x') \\ &= P_\mu(x'') + D(f') + \Theta_1(D(x') - D(f')) \\ &\quad + \Theta_2(P_v(g(x'')) - P_v(x') - P_\mu(x'') + P_\mu(x')). \end{aligned}$$

Now using (3.8) and the fact that $g(x') = x'$, we have that $D'(f') = D'(x')$ and $f'D'(f') - D(f') = x'D'(x') - D(x')$. Hence

$$\Theta_1(D(x') - D(f')) = (K_2 - f')D'(f') - \Delta_2(x')D'(f'). \tag{3.14}$$

Fig. 3.10 The various cases for $K_1 > r_v$ in the setting of Sect. 3.3



Moreover, (3.9) gives that

$$\begin{aligned}
 & \Theta_2\left(P_v(g(x'')) - P_v(x') - P_\mu(x'') + P_\mu(x')\right) \\
 &= \Theta_2\left(g(x'')P'_v(g(x'')) - x''P'_\mu(x'') - x'D'(x')\right) \\
 &= \Theta_2\left((g(x'') - x'')P'_\mu(x'') + (g(x'') - x')D'(x')\right) \\
 &= (K_1 - x'')P'_\mu(x'') + \Delta_2(x')D'(f').
 \end{aligned}
 \tag{3.15}$$

Then combining (3.14) and (3.15), we conclude that $HC = MBEP$. □

3.3.5 $K_1 > r_\mu$

In Lemma 3.4 and under the dispersion assumption, we have constructed f and g , but only on the interval $(e_-, r_\mu]$. More generally, when μ is continuous, the arguments of Beiglböck and Juillet [4] and Henry-Labordère and Touzi [16] allow us to construct $T_d = f$ and $T_u = g$ on $[\ell_\mu, r_\mu]$ for arbitrary laws $\mu \leq_{cx} \nu$. For their purposes, the definitions of f and g outside the range of μ are not important since they have no impact on the construction of the left-curtain martingale coupling.

Nonetheless, we can extend the definitions of f and g to \mathbb{R} in a way which respects the conditions in Lemma 2.6, by setting

$$\begin{aligned}
 f(x) &= x = g(x), & -\infty < x \leq \ell_\mu, \\
 f(x) &= \ell_v \quad \text{and} \quad g(x) = r_v, & r_\mu < x < r_v, \\
 f(x) &= x = g(x), & r_v \leq x < \infty.
 \end{aligned}$$

We shall show that with these definitions for f and g , the analysis of the previous sections extends to the case $K_1 > r_\mu$; see Fig. 3.10.

Suppose that $r_v > r_\mu$ and $r_\mu < K_1 < r_v$. Then $\Lambda(r_\mu) = \frac{r_v - K_1}{r_v - r_\mu} - \frac{K_1 - K_2}{r_\mu - \ell_v}$ and $\Lambda(r_v -) = \infty$. If $\Lambda(r_\mu) \geq 0$ and Λ is continuous, then there exists $x^* \in [\ell_\mu, r_\mu]$ such that $g(x^*) > x^*$ and $\Lambda(x^*) = 0$. Then, exactly as in Sect. 3.2.2, we can construct a model, stopping time and superhedge such that the model-based expected payoff equals the hedging cost, and hence the model, stopping time and hedge are all optimal. The model could be based on the left-curtain coupling, and the optimal exercise rule is to exercise the American put at time 1 if $X < x^*$. Even if Λ is not continuous, there may exist x^* such that $\Lambda(x^*) = 0$ and the same arguments apply (see Sect. 3.3.1). If not, then we are in the setting of Sect. 3.3.4, but again we can identify the optimal model and hedge. Essentially, the case $\Lambda(r_\mu) \geq 0$ is covered by a direct extension of existing arguments. Note that $\Lambda(r_\mu) \geq 0$ is equivalent to

$$K_2 \geq K_1 - \frac{(r_\mu - \ell_v)(r_v - K_1)}{r_v - r_\mu}.$$

Now suppose $r_\mu < K_1 < r_v$ and $K_2 < K_1 - \frac{(r_\mu - \ell_v)(r_v - K_1)}{r_v - r_\mu}$. Then $\Lambda(r_\mu) < 0$, and since $\Lambda(r_v -) = \infty$ and Λ is continuous on $[r_\mu, r_v]$ (note that we have defined f and g to be constant on this range), there must exist $x^* \in (r_\mu, K_1)$ such that $\Lambda(x^*) = 0$. It is always optimal to exercise at time 1, and any martingale coupling can be used to generate a model which attains the highest model-based price $P_\mu(K_1) = (K_1 - \bar{\mu})$. A cheapest superhedge is generated by

$$\psi(y) = \frac{K_2 - \ell_v}{r_v - \ell_v}(r_v - y)^+ + \frac{r_v - K_2}{r_v - \ell_v}(\ell_v - y)^+. \tag{3.16}$$

The cost of this hedge is

$$\begin{aligned} & \frac{K_2 - \ell_v}{r_v - \ell_v} P_v(r_v) + \frac{r_v - K_2}{r_v - \ell_v} P_v(\ell_v) + P_\mu(K_1) - \frac{K_2 - \ell_v}{r_v - \ell_v} P_\mu(r_v) - \frac{r_v - K_2}{r_v - \ell_v} P_\mu(\ell_v) \\ &= \frac{K_2 - \ell_v}{r_v - \ell_v}(r_v - \bar{\mu}) + (K_1 - \bar{\mu}) - \frac{K_2 - \ell_v}{r_v - \ell_v}(r_v - \bar{\mu}) = K_1 - \bar{\mu}. \end{aligned}$$

Finally, if $K_1 > r_v$, then $Y < K_1$ almost surely under any consistent model and

$$\mathbb{E}[(K_2 - Y)^+ | \mathcal{F}_1] \leq \mathbb{E}[(K_1 - Y)^+ | \mathcal{F}_1] = \mathbb{E}[K_1 - Y | \mathcal{F}_1] = K_1 - X.$$

Therefore, it is always optimal to exercise the American put at time 1. If $K_2 > r_v$ or $K_2 < \ell_v$ then we are in the case studied in Sect. 3.3.2 and the cheapest hedge is generated by a time-2 payoff $\psi(y) = (K_2 - y)^+$. If $K_2 \in [\ell_v, r_v]$, then we are in the case studied in Sect. 3.3.3 and the cheapest superhedge is generated by $\psi = \psi(y)$, where ψ is given by (3.16). In either case, the highest model-based expected payoff is $P_\mu(K_1) = (K_1 - \bar{\mu})$, and this is also the cost of the superhedge; see Fig. 3.10.

3.4 Intervals where ν has no mass, or $\nu = \mu$

The definition of the left-curtain martingale coupling (recall Lemma 3.4) only requires that $g = T_u$ is increasing, and not that it is continuous. In general, g may have jumps; such jumps occur when there is an interval on which ν places no mass.

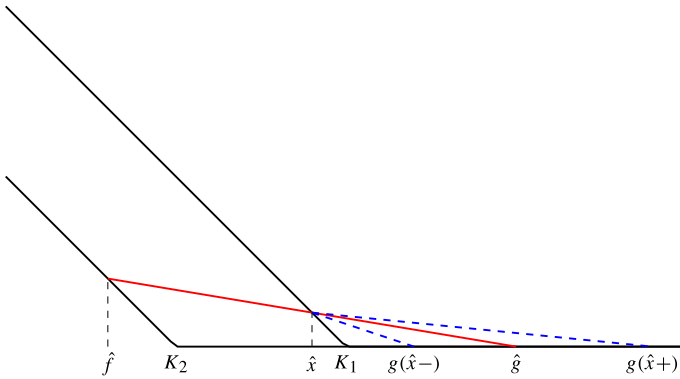


Fig. 3.11 Sketch of put payoffs with points \hat{x} , \hat{f} and \hat{g} marked

If g has a jump, then we need to adapt the superhedge. Suppose g has a jump at \hat{x} (which has to be upwards since g is increasing) and f is continuous at \hat{x} . Suppose further that K_1 is such that $\hat{x} \in (g^{-1}(K_1), K_1)$. Then as before, we should like to find $x^* \in (g^{-1}(K_1), K_1)$ such that $\Lambda(x^*) = 0$. Recall that Λ is increasing and suppose $\Lambda(g^{-1}(K_1)) < 0 < \Lambda(K_1)$. If $\Lambda(\hat{x}-) < 0$ and $\Lambda(\hat{x}+) > 0$, then there will be no solution to $\Lambda = 0$. However, by keeping $x = \hat{x}$, $\hat{f} = f(\hat{x})$ fixed in (3.1) and varying g only, we can find $\hat{g} \in (g(\hat{x}-), g(\hat{x}+))$ such that $(\hat{g} - K_1)/(\hat{g} - \hat{x}) = (K_1 - K_2)/(\hat{x} - \hat{f})$ so that $\Upsilon(f(\hat{x}), \hat{x}, \hat{g}) = 0$. Then the candidate (and indeed optimal) superhedging strategy is generated by ψ^* given in (3.3), with $(f^*, x^*, g^*) = (\hat{f}, \hat{x}, \hat{g})$; see Fig. 3.11. Moreover, since ν does not charge $(g(\hat{x}-), g(\hat{x}+))$, the triple $(\hat{f}, \hat{x}, \hat{g})$ solves the mass and mean equations (2.4) and (2.5). The strong duality between the model-based expected payoff and the hedging cost follows as before.

Alternatively, suppose f has a downward jump at \bar{x} . This can happen if $\nu = \mu$ on $(f(\bar{x}+), f(\bar{x}-))$. Suppose that K_1 is such that $\bar{x} \in (g^{-1}(K_1), K_1)$ and $\Lambda(\bar{x}-) < 0$ and $\Lambda(\bar{x}+) > 0$, so that again we cannot find $x \in (g^{-1}(K_1), K_1)$ with $\Lambda(x) = 0$. We can deal with this similarly as in the case of a discontinuity in g : we first choose $\bar{f} \in (f(\bar{x}+), f(\bar{x}-))$ such that $\Upsilon(\bar{f}, \bar{x}, g(\bar{x})) = 0$, and then consider a hedging strategy generated by ψ^* with $(f^*, x^*, g^*) = (\bar{f}, \bar{x}, g(\bar{x}))$. Note that $\mu = \nu$ on $(f(\bar{x}+), f(\bar{x}-))$; so if (2.4) and (2.5) hold for some $f \in [f(\bar{x}+), f(\bar{x}-)]$ (with \bar{x}, \bar{g}), then they hold for all f in this interval. It follows that we can construct a coupling in which (\bar{f}, \bar{x}) is mapped to (\bar{f}, \bar{g}) , and strong duality holds.

In the case of f and g jumping simultaneously, we have a pictorial representation of the regions of pairs (K_1, K_2) which lead to a hedging strategy which has to be adapted as above; see Fig. 3.12. If g has a jump at \hat{x} , then $\Lambda(\hat{x}-) < 0$ and $\Lambda(\hat{x}+) > 0$ is equivalent to the point (K_1, K_2) lying in the interior of a triangle with vertices $\{(g(\hat{x}-), g(\hat{x}-)), (g(\hat{x}+), g(\hat{x}+)), (\hat{x}, f(\hat{x}))\}$. On the other hand, if f jumps downwards at \bar{x} , then $\Lambda(\bar{x}-) < 0$ and $\Lambda(\bar{x}+) > 0$ is equivalent to the point (K_1, K_2) lying in the interior of a triangle with vertices $\{(\bar{x}, f(\bar{x}-)), (\bar{x}, f(\bar{x}+)), (g(\bar{x}), g(\bar{x}))\}$ (compare this with the region \mathcal{G}).

Exceptionally, we may have simultaneous jumps in g and f at \check{x} . Then the set of (K_1, K_2) for which these arguments are needed is a quadrilateral with ver-

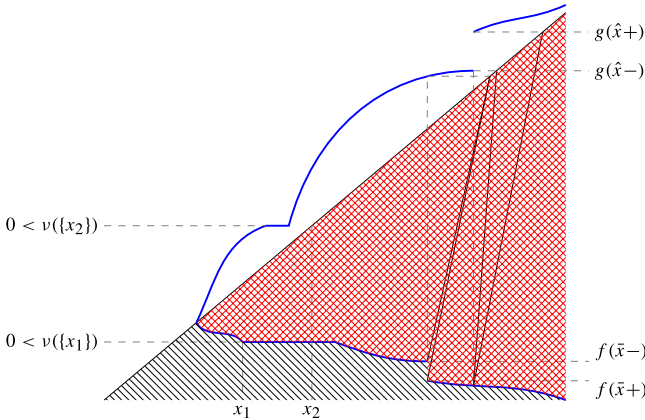


Fig. 3.12 Atoms of ν correspond to flat sections in f and g . Regions of no mass of ν correspond to jumps of f and g

tices $\{(\check{x}, f(\check{x}-)), (\check{x}, f(\check{x}+)), (g(\check{x}+), g(\check{x}+)), (g(\check{x}-), g(\check{x}-))\}$. In particular, then there are multiple pairs (\check{f}, \check{g}) with $\check{f} \in (f(\check{x}+), f(\check{x}-))$ and $\check{g} \in (g(\check{x}-), g(\check{x}+))$ such that $\Upsilon(\check{f}, \check{x}, \check{g}) = 0$, so that an optimal hedging strategy is not unique.

3.5 The general case for continuous ν

In the previous sections, we have shown how the left-curtain coupling can be used to find an optimal model, exercise strategy and a superhedge, under the assumption that both μ and ν are continuous together with further regularity and simplifying assumptions which we labelled the dispersion assumption and the single-jump assumption. Under the latter assumption, the existence of points that solve (3.8) led us to identify two further types of hedging strategy that were not present under the dispersion assumption, making four in total.

If we relax the assumptions further and require only that both μ and ν are continuous, then we expect that in some cases there may exist multiple pairs (f'_i, x'_i) , $i = 1, 2, 3, \dots$, that solve (3.8). Note that from the monotonicity of g , we can write $\{x : g(x) > x\}$ as a countable union of intervals, and on each such interval, f is decreasing. f jumps over the intervals (f'_i, x'_i) identified above (at least those with x' to the left of the current value of x). In particular, f has only countably many downward jumps. Figure 2.1 is a stylised representation of the general left-curtain martingale coupling, not least because in the figure f has only finitely many jumps. Starting from Fig. 2.1 and using the constructions in Sect. 3.3, we can divide the area of all $(K_1, K_2 < K_1)$ into four regions; see Fig. 3.13. The key point is that these four regions are characterised exactly as in the cases described in Sect. 3.3. For given (K_1, K_2) , we can determine which of the types of hedging strategy is a candidate optimal superhedge, and determine a candidate optimal stopping rule. (We can always use the model associated with the left-curtain martingale coupling π_{lc} .) The fact that these candidates are indeed optimal can be proved using exactly analogous techniques to those used in Sect. 3.3.

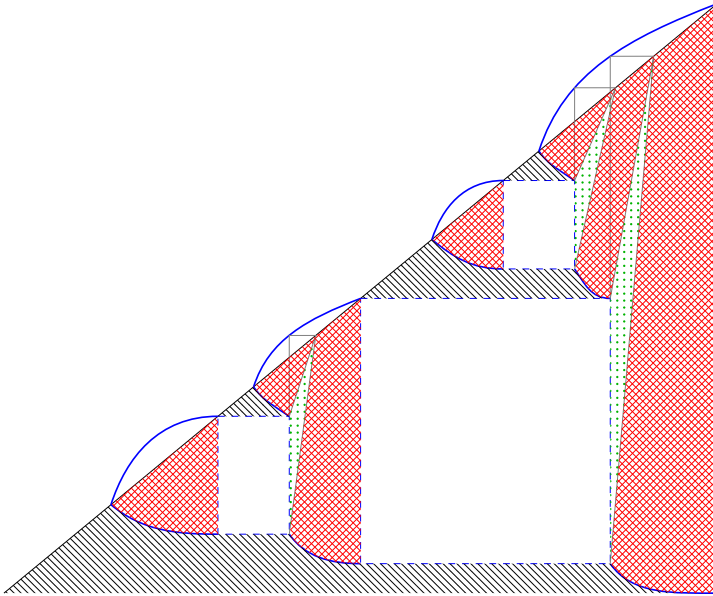


Fig. 3.13 General picture of f, g with shading of regions. There remain four types of shading corresponding to four forms of optimal hedge

More specifically, we can divide $\{(k_1, k_2) : k_2 < k_1\}$ into two disjoint regions, $\{(k_1, k_2) : k_2 \leq f(k_1)\}$ and $\{(k_1, k_2) : f(k_1) < k_2 < k_1\}$. We can divide the former into two further regions $\mathcal{W} = \{(k_1, k_2) : k_2 < k_1, \exists x \leq k_1 \text{ with } f(x) < k_2 < g(x)\}$ and $\mathcal{B} = \{(k_1, k_2) : k_2 \leq f(k_1)\} \setminus \mathcal{W}$. The latter we again divide into two regions \mathcal{G} and $\mathcal{R} = \{(k_1, k_2) : f(k_1) < k_2 < k_1\} \setminus \mathcal{G}$. Here we can write $\mathcal{G} = \bigcup_{x: f(x-) > f(x+)} \Delta(x)$, where $\Delta(x)$ is a triangle with vertices $\{(x, f(x+)), (x, f(x-)), (g(x), g(x))\}$. Then on each of the regions $\mathcal{W}, \mathcal{B}, \mathcal{G}$ and \mathcal{R} , we have a superhedge exactly as described in Sect. 3.3. Moreover, again by the arguments of Sect. 3.3, we can show that the hedging cost associated with the superhedging strategy is precisely the model-based expected payoff of the American put under the martingale coupling π_{1c} (and candidate stopping rule), thus proving the optimality of the hedge and of the model/exercise rule.

For example, suppose $(K_1, K_2) \in \mathcal{W}$. (The cases for $(K_1, K_2) \in \mathcal{R} \cup \mathcal{B}$ are generally even simpler, and for $(K_1, K_2) \in \mathcal{G}$, the story is roughly equally involved.) Recall that under the (single-jump) Assumption 3.8, in order to show that MBEP = HC, we used the existence of \bar{x} and \bar{f} satisfying (3.8) together with the fact that

$$\int_y \int_{x > K_1} (K_2 - y)^+ \pi_{1c}(dx, dy) = \int_{-\infty}^{\bar{f}} (K_2 - y)(v - \mu)(dy). \tag{3.17}$$

Then for general probability measures μ and ν , provided we can find \bar{x}, \bar{f} satisfying (3.8) and (3.17), the proof that MBEP = HC and hence of optimality follows exactly as in Theorem 3.12.

Lemma 3.14 *Suppose $(K_1, K_2) \in \mathcal{W}$. Then there exist \bar{x}, \bar{f} such that*

$$\bar{f} < K_2 < \bar{x} \leq K_1 \quad \text{and} \quad \int_{\bar{f}}^{\bar{x}} z^i \mu(dz) = \int_{\bar{f}}^{\bar{x}} z^i \nu(dz), \quad i \in \{1, 2\}, \quad (3.18)$$

and such that under the left-curtain coupling, we have

$$\{X > K_1, Y \leq K_2\} = \{X > K_1, Y \leq \bar{f}\} = \{Y \leq \bar{f}\} \setminus \{X \leq \bar{f}\} \quad \text{a.s.},$$

so that (3.17) holds.

Proof Define $\mathcal{X} = \mathcal{X}_{K_1, K_2} = \{x : x \leq K_1, f(x) < K_2 < g(x)\}$. Since (K_1, K_2) is in \mathcal{W} , \mathcal{X}_{K_1, K_2} is nonempty. Define $\hat{x} = \sup\{x : x \in \mathcal{X}\}$. We show that \hat{x} and a suitably defined \hat{f} are such that (3.17) and (3.18) hold.

First suppose that $\hat{x} < K_1$. Suppose further that $g(\hat{x}) > \hat{x}$. Take $\tilde{x} \in (\hat{x}, g(\hat{x}) \wedge K_1)$. Then $g(\tilde{x}) \geq g(\hat{x}) > \tilde{x}$. Also $f(\tilde{x}) \notin (f(\hat{x}), g(\hat{x}))$ and if $f(\tilde{x}) \geq g(\hat{x})$, then we have $f(\tilde{x}) \geq g(\hat{x}) > \tilde{x}$ which is a contradiction. So $f(\tilde{x}) \leq f(\hat{x}) < K_2 < g(\hat{x}) \leq g(\tilde{x})$, and then $\tilde{x} \in \mathcal{X}$, contradicting the maximality of \hat{x} . Hence $g(\hat{x}) \leq \hat{x}$ (and thus $g(\hat{x}) = \hat{x}$). But then $f(\hat{x}) = \hat{x}$ and $\hat{x} \notin \mathcal{X}$.

Hence there exists $(x_n)_{n \geq 1}$ such that $x_n \in \mathcal{X}$ and $x_n \uparrow \hat{x}$. Let $g(\hat{x}-) = \lim g(x_n)$. By the same argument as above, we cannot have $g(\hat{x}-) > \hat{x}$. Hence $\hat{x} = g(\hat{x}-) > K_2$.

Now suppose $\hat{x} = K_1 > K_2$. Then $K_1 \notin \mathcal{X}$ because we cannot have both $K_2 \leq f(K_1)$ and $f(K_1) < K_2 < g(K_1)$. Hence there exists $(x_n)_{n \geq 1}$ such that $x_n \in \mathcal{X}$ and $x_n \uparrow \hat{x}$. Let $g_n = g(x_n)$ and $f_n = f(x_n)$. If $g(K_1) > K_1$, then there exists n_0 such that for all $n \geq n_0$, $g_n > K_1$. Then $f(K_1) \notin (f_n, g_n)$ and therefore $f(K_1) \leq f_n < K_2$, contradicting $K_2 \leq f(K_1)$. Hence $g(\hat{x}-) = g(K_1-) = K_1$.

In either case, $\hat{x} \notin \mathcal{X}$ and there exists $(x_n)_{n \geq 1}$ such that $x_n \in \mathcal{X}$, $x_n \uparrow \hat{x}$ and $(f_n)_{n \geq 1}$ is a decreasing sequence while $(g_n)_{n \geq 1}$ satisfies $g_n \uparrow \hat{x}$. Let $\hat{f} = \lim_{n \rightarrow \infty} f_n$. Then

$$\int_{\hat{f}}^{\hat{x}} z^i \mu(dz) = \int_{f_n}^{g_n} z^i \nu(dz), \quad i \in \{1, 2\},$$

and by taking limits, we have that \hat{x} and \hat{f} solve (3.18). Note also that $\hat{x} > K_2$.

We are left to show that \hat{x} and \hat{f} solve (3.17). This follows from the fact that $\hat{f} < K_2$, together with the set identifications

$$\{X > K_1, Y \leq K_2\} = \{X > K_1, Y \leq \hat{f}\} = \{X > \hat{f}, Y \leq \hat{f}\} = \{Y \leq \bar{f}\} \setminus \{X \leq \bar{f}\}. \quad \square$$

Remark 3.15 The set $\{x : g(x) > x\}$ is a collection of intervals and we let I_+ denote the set of right endpoints of these intervals. As remarked above, Fig. 3.13 is drawn in the case of ‘finite complexity’ in the sense that the set I_+ contains a finite number of elements. The results extend easily to countable I_+ , provided I_+ contains no accumulation points.

In general, I_+ may contain an accumulation point, and as discussed in Henry-Labordère and Touzi [16], care is needed in the construction of the left-curtain mappings (T_d, T_u) in this case. However, from our perspective, such subtleties do not

cause a problem. The reason is that we do not aim to derive the left-curtain coupling, but rather take the left-curtain coupling as a given and use it to solve the put pricing problem.

Our construction of the best model and the cheapest hedge is local in the sense that when in Fig. 3.13, we examine in which region the point (K_1, K_2) lies, the fine detail of the picture in other parts of (k_1, k_2) -space is not important. So the existence of accumulation points can only be an issue if K_1 is equal to one of those accumulation points.

Let x_∞ be such an accumulation point in I_+ and suppose $K_1 = x_\infty$. Depending on the value of K_2 , then either there exists (x', f') with $f' < K_2 < x'$ such that (3.8) holds or not. In the former case, we can follow the analysis of Sect. 3.3.3, and in the latter, Sect. 3.3.2; in either case, we construct a model and hedge such that the model price and hedging cost agree, thus proving optimality of both.

3.6 Atoms in the target law

When ν has atoms, the preservation-of-mass and mean conditions become (2.6) and (2.7), respectively. In particular, atoms of ν correspond to the flat sections in f or g ; see Fig. 3.12. In this case, we still can find all the optimal quantities as before. In particular, $\Lambda(x) := \frac{g(x)-K_1}{g(x)-x} - \frac{K_1-K_2}{x-f(x)}$ is strictly increasing in x , even if f and/or g is constant. Hence we can find solutions to $\Lambda = 0$ (more generally solutions $x, f \in \mathfrak{N}(x)$ to $\Upsilon(f, x, g = g(x)) = 0$) exactly as before. The superhedge is unchanged. A little care is needed in constructing the optimal model, but under the associated martingale coupling, mass in $(f(x^*), x^*)$ is mapped onto $(f(x^*), g(x^*))$ together with (potentially) atoms at $f(x^*)$ or $g(x^*)$. Specifically, given f^*, x^*, g^* , we can find λ_f^* and λ_g^* such that (2.6) and (2.7) hold. Then in any optimal canonical model \hat{M}_π , π is constant on $(-\infty, f^*)$, and the law of \hat{M}_2 on the event $\{\hat{M}_1 \in (f^*, x^*)\}$ is ν_{x^*} which is defined to be $\nu_{x^*} = \nu|_{(f^*, g^*)} + \lambda_{f^*}^* \delta_{f^*} + \lambda_{g^*}^* \delta_{g^*}$. We also find that the law of \hat{M}_2 on the event $\{\hat{M}_1 \in (x^*, \infty)\}$ is $\nu - \nu_{x^*} - \mu|_{(-\infty, f^*)}$.

4 Discussion and extensions

4.1 The role of the left-curtain coupling

For any pair of strikes (K_1, K_2) , the left-curtain model attains the highest expected payoff for the American put. However, although it optimises simultaneously across all pairs of strikes, it is not (in general) optimal for linear combinations of American puts. For example, if we consider a generalised American option with payoff a if exercised at time 1 and b if exercised at time 2, where $a(x) = \sum_{j=1}^J (K_1^j - x)^+$ and $b(y) = \sum_{j=1}^J (K_2^j - y)^+$ (with $K_2^j \leq K_1^j$ for each j), then the model associated with the left-curtain coupling is typically not optimal. The reason is that a model (\mathcal{S}, M) is only optimal when it is combined with the best stopping rule, and the optimal stopping rule *does* depend on (K_1, K_2) .

Conversely, although the model associated with the left-curtain coupling is optimal (simultaneously across all pairs (K_1, K_2)), we do not need the full power of this coupling when we work with fixed (K_1, K_2) . In the dispersion assumption case, all we need is a coupling in which $(f(x^*), x^*)$ is mapped onto $(f(x^*), g(x^*))$ where x^* is such that $\Lambda(x^*) = 0$, and $(-\infty, f^*)$ is mapped to itself, but not necessarily in a constant fashion. There are many martingale couplings which have this property.

The intuition behind the optimality of the left-curtain coupling is as follows. With American puts, there is a tension between the time-decay of the option payout promoting early exercise, and the convexity of the payoff function promoting delay. If the aim is to maximise the payoff of the option, then any paths which are in the money at time 1 and will remain in the money are best exercised at time 1. However, once a path has been exercised, any further volatility is irrelevant. In particular, when designing a candidate optimal model, we should try to keep paths which are exercised at time 1 constant (or near constant) whenever possible. Thus the probability space should be split into two regions: one region where the put is in the money at time 1 and is exercised, and thereafter paths move little, and a second region where the put is out of the money at time 1 (and sometimes just in the money, but left unexercised at time 1), and then paths move a long way between times 1 and 2. The left-curtain coupling has this property.

4.2 Multiple exercise times

It is natural to ask if it is possible to extend the analysis to American puts which can be exercised at multiple dates (T_1, T_2, \dots, T_N) where $N > 2$, or equivalently to martingales $M = (M_n)_{0 \leq n \leq N}$ with marginals (μ_n) where μ_1 has mean $M_0 = \bar{\mu}$ and $\mu_n \leq_{cx} \mu_{n+1}$ for $1 \leq n \leq N - 1$. It is clear that many of the ideas extend naturally to the multi-marginal case. However, the number of types of hedging strategy may grow exponentially with N . This is left as future work.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Aksamit, A., Deng, S., Oblój, J., Tan, X.: Robust pricing-hedging duality for American options in discrete time financial markets. *Math. Finance* (2019). Forthcoming, available online at <https://doi.org/10.1111/mafi.12199>
2. Azéma, J., Yor, M.: Une solution simple au problème de Skorokhod. In: Dellacherie, C., et al. (eds.) *Séminaire de Probabilités, XIII. Lecture Notes in Math.*, vol. 721, pp. 90–115. Springer, Berlin (1979)
3. Bayraktar, E., Huang, Y.J., Zhou, Z.: On hedging American options under model uncertainty. *SIAM J. Financ. Math.* **6**, 425–447 (2015)
4. Beiglböck, M., Juillet, N.: On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44**, 42–106 (2016)

5. Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17**, 477–501 (2013)
6. Beiglböck, M., Henry-Labordère, P., Touzi, N.: Monotone martingale transport plans and Skorokhod embedding. *Stoch. Process. Appl.* **127**, 3005–3013 (2017)
7. Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in option prices. *J. Bus.* **51**, 621–651 (1978)
8. Brown, H., Hobson, D.G., Rogers, L.C.G.: Robust hedging of barrier options. *Math. Finance* **11**, 285–314 (2001)
9. Carr, P., Lee, R.: Hedging variance options on continuous semimartingales. *Finance Stoch.* **14**, 179–207 (2010)
10. Cox, A.M.G., Hoeggerl, C.: Model-independent no-arbitrage conditions on American put options. *Math. Finance* **26**, 431–458 (2016)
11. Cox, A.M.G., Oblój, J.: Robust pricing and hedging of double no-touch options. *Finance Stoch.* **15**, 573–605 (2011)
12. Cox, A.M.G., Wang, J.: Root’s barrier: construction, optimality and applications to variance options. *Ann. Appl. Probab.* **23**, 859–894 (2013)
13. Cox, A.M.G.: Extending Chacon–Walsh: minimality and generalised starting distributions. In: Donati-Martin, C., et al. (eds.) *Séminaire de Probabilités, XLI. Lecture Notes in Math.*, vol. 1934, pp. 233–264. Springer, Berlin (2008)
14. Dolinsky, Y., Soner, M.: Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields* **160**, 391–427 (2014)
15. Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals with an application to lookback options. *Ann. Appl. Probab.* **24**, 313–336 (2014)
16. Henry-Labordère, P., Touzi, N.: An explicit martingale version of the one-dimensional Brenier theorem. *Finance Stoch.* **20**, 635–668 (2016)
17. Hobson, D.G.: Robust hedging of the lookback option. *Finance Stoch.* **2**, 329–347 (1998)
18. Hobson, D.G., Klimmek, M.: Robust price bounds for the forward starting straddle. *Finance Stoch.* **19**, 189–214 (2015)
19. Hobson, D.G., Neuberger, A.: Robust bounds for forward start options. *Math. Finance* **22**, 31–56 (2012)
20. Hobson, D.G., Neuberger, A.: More on hedging American options under model uncertainty. Preprint (2016). Available online at [arXiv:1604.02274](https://arxiv.org/abs/1604.02274)
21. Hobson, D.G., Neuberger, A.: Model uncertainty and the pricing of American options. *Finance Stoch.* **21**, 285–329 (2017)
22. Hobson, D.G., Norgilas, D.: The left-curtain martingale coupling in the presence of atoms. *Ann. Appl. Prob.* **29**, 1904–1928 (2019)
23. Hobson, D.G.: The maximum maximum of a martingale. In: Azéma, J., et al. (eds.) *Séminaire de Probabilités, XXXII. Lecture Notes in Math.*, vol. 1686, pp. 250–263. Springer, Berlin (1998)
24. Hobson, D.G.: The Skorokhod embedding problem and model-independent bounds for option prices. In: Carmona, R.A., et al. (eds.) *Paris–Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math.*, vol. 2003, pp. 267–318. Springer, Berlin (2011)
25. Neuberger, A.: Bounds on the American option. Preprint (2007). Available online at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=966333
26. Skorokhod, A.V.: *Studies in the Theory of Random Processes*. Addison-Wesley Publishing Co., Inc., Reading (1965)
27. Stebegg, F.: Model-independent pricing of Asian options via optimal martingale transport. Preprint (2014). Available online at [arXiv:1412.1429](https://arxiv.org/abs/1412.1429)
28. Strassen, V.: The existence of probability measures with given marginals. *Ann. Math. Stat.* **36**, 423–439 (1965)
29. Touzi, N.: Martingale inequalities, optimal martingale transport, and robust superhedging. *ESAIM Proc. Surv.* **45**, 32–47 (2014)