Robust Optimal Stopping*

Volker Krätschmer
Dept. of Mathematics
University of Duisburg-Essen
Volker.Kraetschmer@uni-due.de

Marcel Ladkau
Stochastic Algorithms and Nonparametric Statistics
Weierstrass Institute Berlin
Marcel.Ladkau@wias-berlin.de

Roger J. A. Laeven
Dept. of Quantitative Economics
University of Amsterdam, EURANDOM
and CentER
R.J.A.Laeven@uva.nl

John G. M. Schoenmakers
Stochastic Algorithms and Nonparametric Statistics
Weierstrass Institute Berlin
schoenma@wias-berlin.de

Mitja Stadje
Dept. of Econometrics and Operations Research
Tilburg University
and CentER
M.A.Stadje@uvt.nl

This Version: April 17, 2015

Abstract

This paper studies the optimal stopping problem in the presence of model uncertainty (ambiguity). We develop a method to practically solve this problem in a general setting, allowing for general time-consistent ambiguity averse preferences and general payoff processes driven by jump-diffusions. Our method consists of three steps. First, we construct a suitable Doob martingale associated with the solution to the optimal stopping problem using backward stochastic calculus. Second, we employ this martingale to construct an approximated upper bound to the solution using duality. Third, we introduce backward-forward simulation to obtain a genuine upper bound to the solution, which converges to the true solution asymptotically. We analyze the asymptotic behavior and convergence properties of our method. We illustrate the generality and applicability of our method and the potentially significant impact of ambiguity to optimal stopping in a few examples.

Keywords: Optimal stopping; Model uncertainty; Robustness; Convex risk measures; Ambiguity aversion; Duality; BSDEs; Monte Carlo simulation; Regression; Relative entropy.

AMS 2000 Subject Classification: 49L20, 60G40, 91B06, 91B16.

JEL Classification: D81, G10, G20.

*We are very grateful to Christian Bender, Juan Vera Lizcano, Jianing Zhang, and participants at the EURANDOM-CWI Workshop on BSDEs for their comments and suggestions. This research was funded in part by the Netherlands Organization for Scientific Research under grants NWO VIDI 2009 (Laeven) and NWO VENI 2012 (Stadje) and by the DFG Research Center Matheon "Mathematics for Key Technologies" in Berlin (Schoenmakers).

1 Introduction

The theory of optimal stopping and control has evolved into one of the most important branches of modern probability and optimization and has a wide variety of applications in many areas, perhaps most notably in operations management, statistics, and economics and finance. There exists a vast literature on both theory and applications of optimal stopping and control, going back to Wald [85] and Snell [80], and we mention here only an incomplete selection related to the setting of this paper: Brennan and Schwartz [16], McDonald and Siegel [63], Pindyck [70], Barone-Adesi and Whaley [3], Dixit [35], Dixit and Pindyck [36], Karatzas and Shreve [55], Dayanik and Karatzas [31], Guo and Pham [46], Dasci and Laporte [32], Peskir and Shiryaev [68], Øksendal and Sulem [65], Henderson and Hobson [52], and Dharma Kwon [34]. Prime applications are a manufacturer's market entry decision[1] or ageing plant closing decision in operations management; a real estate agent's decision to accept a bid or search problems in economics; and the valuation of American-style derivatives in finance.[2] These applications naturally lead to an optimal stopping problem.

Since the (future) reward (sequence) is typically uncertain in these applications, it needs to be evaluated using probabilistic methods, and the main target in the above-mentioned literature on standard optimal stopping is the maximization of the expected reward over a family of stopping strategies. That is, the central object is the *expectation* of the reward induced by the problem's payoff process. Such a setting requires that the reward's expectation can be unambiguously determined by the decision-maker, which is the case in particular if the reward's probability law is given to the decision-maker. In reality, however, this is quite a restrictive requirement: in many situations the decision-maker faces uncertainty about the true probabilistic model, meaning that the probability law generating the future reward is (partially) unknown.[3] In these situations, different probabilistic models may be plausible, each of them potentially leading to very different optimal stopping strategies. Such model uncertainty is usually referred to as ambiguity. In decision theory, the more specific term of Knightian uncertainty (after Knight [56]) is also employed, to distinguish from decision under uncertainty problems in which the probabilistic model is objectively given — the specific case of decision under risk. Approaches that explicitly take ambiguity into account are often referred to as *robust* approaches.

In a general probabilistic setting, a robust approach that has recently gained much attention is provided by convex measures of risk (Föllmer and Schied [39], Frittelli and Rosazza Gianin [41], and Heath and Ku [51], extending Artzner et al. [2]; see also the early Ben-Tal [9] and Ben-Tal and Teboulle [10]). For applications of convex risk measures in the context of decision and optimization, see e.g., Ruszczyńsky and Shapiro [75], Lesnevski, Nelson and Staum [60], Ben-Tal, Bertsimas and Brown [12], Choi, Ruszczyńsky and Zhao [25], Tekaya, Shapiro, Soares and da Costa [83], and Laeven and Stadje [58, 59]. By the representation theorem of convex

[1]From the entrance time onwards, the firm will encounter fixed irreversible costs but will at the same time start generating an (uncertain) reward. The goal of the management would be to maximize their present value.

[2]The buyer of such a derivative wants to find the optimal time to exercise the option such that the reward be maximized.

[3]This is, for instance, the case if estimation is unreliable, data are scarce, or if the evaluation necessarily relies on extrapolating past trends, but past patterns are no longer representative for their future counterpart. Furthermore, in financial decision-making (as in the case of American-style derivatives), investors may need to cope with markets that are inherently incomplete, meaning, in particular, that no unique probabilistic pricing operator exists.

risk measures, a random future reward, say H, is evaluated according to

$$U(H) = \inf_{Q \in \mathcal{Q}} \{\mathrm{E}_Q [H] + c(Q)\}, \tag{1.1}$$

where $\mathcal{Q} = \{Q | Q \sim P\}$ is the set of probabilistic models Q that share the same null sets with a base reference model P, with each Q attaching a different probability law to the future reward H, and c is a penalty function specifying the plausibility of the model Q.[4] Models Q that have 'low' plausibility are associated with a high penalty, while models that have 'high' plausibility yield a low penalty, with $c(Q) = \infty$ corresponding to the case in which the model Q is considered fully implausible. By taking the infimum over $\mathcal{Q}$ a conservative worst-case approach occurs, also typical in (deterministic) robust optimization.

A canonical class of penalty functions is provided by ϕ-divergences; see e.g., Ben-Tal and Teboulle [10, 11]. In this case, the decision-maker starts with a reference model P, which is an approximation or 'an educated guess' to the probabilistic model driving the reward H rather than the true model. The decision-maker therefore does not solely rely on the model P but considers instead a collection of models Q, with esteemed plausibility (or trust) decreasing with their ϕ-divergence measure with respect to the approximation P. A similar approach was adopted by Hansen and Sargent [48, 49] in macroeconomics, using the specific Kullback-Leibler (ϕ-)divergence (or relative entropy; see also Csiszár [29] and Ben-Tal [9]). Another special case of interest is given by penalty functions of the form

$$c(Q) = \begin{cases} 0, & \text{if } Q \in M \subset \mathcal{Q}; \\ \infty, & \text{otherwise}; \end{cases} \tag{1.2}$$

for a fixed set of probabilistic models $M \subset \mathcal{Q}$. The subclass of penalty functions given by an indicator function as in (1.2) yields evaluations of the form[5]

$$U(H) = \inf_{Q \in M} \mathrm{E}_Q [H], \tag{1.3}$$

which attaches the same plausibility to all probabilistic models in M; see e.g., Föllmer and Schied [40] for further details. In a dynamic setting, such as considered in this paper, time-consistent versions of convex measures of risk were discussed by Riedel [72] and have also been considered more recently in e.g., Ruszczyńsky and Shapiro [76], Cheridito, Delbaen and Kupper [24], Ruszczyńsky [74], Philpott, de Matos and Finardi [69], and Laeven and Stadje [59]; see also Duffie and Epstein [37], Chen and Epstein [21], Shapiro, Dentcheva and Ruszczyński [79], Chapter 6, and Glasserman and Xu [45]. The usual definition of time-consistency requires that whenever, in each state of nature at time t, a reward H_2 is preferred over H_1, it is also preferred prior to time t. In our context, this implies in particular that a stopping strategy that is optimal at time $t = 0$ will not be reversed at a later point in time. For dynamic versions of evaluations of the form (1.1), time-consistency is equivalent to a dynamic programming principle (recursiveness).

Decision-making under ambiguity, with probabilities of events unknown to the decision-maker, has been extensively studied in economics since the seminal work of Ellsberg [38]. It has been noted that incorporating ambiguity may not only be of theoretical and normative interest,

[4] In the literature, a convex risk measure is usually defined as $-U(H)$ leading however to the same optimization problem.

[5] In this case, U corresponds to a coherent risk measure given by $-U(H)$.

but can also play a potential role in explaining empirically important failures of a purely risk-based framework (Chen and Epstein [21]). Popular approaches to decision-making under ambiguity are provided by the multiple priors preferences of Gilboa and Schmeidler [43] (see also Schmeidler [77]), also referred to as maxmin expected utility, and the significant generalization of variational preferences developed by Maccheroni, Marinacci and Rustichini [62]. With linear utility, multiple priors essentially reduces to the evaluation (1.3) while variational preferences reduces to (1.1). Such preferences induce aversion to ambiguity (Cerreia-Vioglio et al. [20]). A version of multiple priors was also studied by Huber [53] in robust statistics; see also the early Wald [85].

The theory of convex measures of risk and ambiguity averse preferences is well-established and their use in optimal stopping problems has recently been developing; see, in particular, Riedel [71], Krätschmer and Schoenmakers [57], Bayraktar, Karatzas and Yao [4], Bayraktar and Yao [5], Cheng and Riedel [23] and Øksendal, Sulem and Zhang [66]. However, the development of numerical methods to practically solve robust optimal stopping problems may currently be considered breaking ground.

In this paper, we develop a method to practically solve the optimal stopping problem under ambiguity in a general continuous-time setting, allowing for general time-consistent convex measures of risk, i.e., all time-consistent dynamic counterparts of (1.1), and general (sequences of) rewards. As to the payoff process, we allow for a general jump-diffusion model specification. The key to our method is to expand two duality theories of a different kind. The first kind of duality theory is the martingale duality approach to standard optimal stopping problems, dating back to Rogers [73], Haugh and Kogan [50] and Andersen and Broadie [1] (see also Davis and Karatzas [30]). We expand their martingale dual representation to encompass general preference functionals beyond plain conditional expectation. The second kind of duality theory explicates the connection between time-consistent convex measures of risk and backward stochastic differential equations (BSDEs), which we expand to apply to our setting. We note that powerful numerical tools are nowadays available for BSDEs.

Our method is then composed of three steps. First, expanding duality theory of the second kind and using backward stochastic calculus, we construct a suitable Doob martingale from the Snell envelope generated by the optimally stopped and robustly evaluated payoff process. Second, expanding duality theory of the first kind, we employ this martingale to construct an approximated upper bound to the solution of the optimal stopping problem. Third, we introduce the notion of backward-forward simulation to obtain a genuine upper bound to the solution. We analyze the asymptotic behavior of our method by deriving its convergence properties. To the best of our knowledge, we are not aware of other practical solution methods for robust optimal stopping problems in the literature so far. Finally, to illustrate the generality of our approach and the relevance of ambiguity to optimal stopping, we supplement the presentation of our method with a few examples of robust optimal stopping problems, including Kullback-Leibler divergences, worst case scenarios, and good-deal bounds. Our numerical results illustrate that our algorithm is easily implemented for a wide range of robust optimal stopping problems and has good convergence properties, yielding accurate results in realistic settings at the pre-limiting level. They also reveal that ambiguity can have a significant impact on the robust reward evaluations under standard specifications. Thus, ambiguity really matters for optimal stopping.

The development of methods to practically compute the solution to a standard optimal stopping problem (with plain conditional expectations) has a long history, in particular in the

American-style option literature. Seminal contributions based on regression include Carriere [19] and Longstaff and Schwartz [61]; see also Tsitsiklis and Van Roy [84] and Clément, Lamberton and Protter [26]. These methods, which are connected to the stochastic mesh method of Broadie and Glasserman [17] (see Glasserman [44]), can be used to generate lower bounds to the optimal solution and are part of the literature that is referred to as primal. The development of practical dual methods started with Andersen and Broadie [1] who exploited the dual representation obtained by Rogers [73] and Haugh and Kogan [50]. Many follow-up papers have further refined their method; see e.g., Belomestny, Bender and Schoenmakers [6] and Schoenmakers, Zhang and Huang [78] and their references. Employing duality (of the first kind), our method may, in some sense, be viewed as the analogous contribution for robust optimal stopping problems of the original contribution by Andersen and Broadie [1] for standard optimal stopping problems. But we note that we are not even aware of any primal method to practically solve robust optimal stopping problems in the literature to date. Furthermore, we note that we allow for a more general reward specification.

An interesting aspect of our method, which may be of interest as a contribution to the BSDE literature in its own right, is the introduction of backward-forward Monte Carlo simulation to obtain a *genuine* (biased high) upper bound, which will converge to the true solution as the number of Monte Carlo simulations and basis functions increases and the mesh ration of the time-grid tends to zero. Bender, Schweizer, and Zhuo [7] derive upper and lower bounds on the solution to a discrete-time (reflected) BSΔE, rather than a continuous-time BSDE as we consider, using techniques different from ours.

The remainder of this paper is organized as follows. In Section 2, we introduce our setting, specify the robust optimal stopping problem, recall some basic properties of time-consistent ambiguity averse preferences, and provide some illustrative examples. In Section 3, we present the duality results (of the first and second kind) underpinning our approach, and revisit our examples using duality. In Section 4, we provide a general outline of our algorithm and a preview of our convergence results. Section 5 contains the numerical examples. A detailed step-wise description of our algorithm and its convergence properties are presented in Section 6. Details of our proofs are deferred to the Appendix.

2 Problem Description

2.1 Setting, Rewards and Preferences

Consider a decision-maker (economic agent or firm) who has to decide at what time to stop (or exercise) a certain action in order to maximize his future uncertain (sequence of) rewards. For the dynamics of the rewards, we assume a continuous-time jump-diffusion setting with ambiguity. Formally, we consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and assume that the probability space is equipped with two independent processes, which will serve as our stochastic drivers:

(i) A standard d-dimensional Brownian motion $W = (W^1, \ldots, W^d)^\intercal$.

(ii) A standard k-dimensional Poisson process $N = (N^1, \ldots, N^k)^\intercal$ with intensities $\lambda_P = (\lambda_P^1, \ldots, \lambda_P^k)^\intercal$.

Standard in this case means that the components are assumed to be independent, and, in the case of W, to have zero mean and unit variance. We denote the vector of compensated Poisson processes by $\tilde{N} = (\tilde{N}^1, \ldots, \tilde{N}^k)^\intercal$, where

$$\tilde{N}_t^i = N_t^i - \lambda_P^i t, \qquad i = 1, \ldots, k.$$

We assume that these stochastic drivers generate an n-dimensional adapted Markov process $(X_t)_{t\in[0,T]}$ satisfying the strong Markov property. The process X is exogenous and may represent a production process, a capacity process, a stream of net cash flows, or a price process of e.g., a collection of risky assets.

The decision-maker chooses a stopping time τ taking values between time 0 and a fixed maturity time $T < \infty$. We assume that if the decision-maker exercises at time $\tau = t_i$, he receives the reward

$$H_{t_i} = \Pi(t_i, X_{t_i}) + \sum_{j=i}^{L} h(t_j, X_{t_j}), \qquad t_i \in \{t_0 = 0, t_1, \ldots, t_L = T\}, \tag{2.1}$$

for functions Π and h mapping from $\{t_0 = 0, t_1, \ldots, t_L = T\} \times \mathbb{R}^n$ to $\mathbb{R}$. Furthermore, we assume that $h(t_j, X_{t_j}) \in L^2$ for all $j = 0, \ldots, L$. Standard examples that take the form (2.1) include:

(a) The optimal entrance problem: In this case, typically $\Pi(t, x) = -\exp(-\rho t)\,\kappa$, for a fixed irreversible cost κ depreciating at a continuous rate ρ, and $h(t, x) = \exp(-\rho t)\,(h(x) - \xi)$, which measures the present value of the payoff or the production per time unit, $h(x)$, after entering the market, minus the running costs, ξ. Often times $h(x)$ is simply taken to be equal to x. Of course, the fixed costs may also depend on the state of the economy at time t, X_t.

(b) The optimal (simple) reward problem: In this case, $h \equiv 0$ and $\Pi(t, x)$ is the (simple) reward function of exercising at time t. This problem appears abundantly in the American option pricing literature, with X_t a vector of risky asset values at time t.

For further details on these and other examples, see the references provided in the Introduction.

In standard optimal stopping problems, the decision-maker maximizes the expected reward under a given probabilistic model P:

$$\max_{\tau\in\mathcal{T}} \mathrm{E}\,[H_\tau]\,,$$

where $\mathcal{T} = \{t_0 = 0 < t_1 < \ldots < t_L = T\}$ is the set of possible exercise dates. Specifying the model P in our setting means specifying the distribution of the whole path $(X_t)_{t\in[0,T]}$. In reality, however, the probabilities with which future rewards are received are often times subject to model uncertainty. Therefore, it is appealing to consider instead a robust decision criterion, which induces that the optimal stopping strategy accounts for a whole class of probabilistic models and not just a single one. Different approaches to decision-making under ambiguity have emerged in the literature. Among the most popular approaches are multiple priors (Gilboa and Schmeidler [43]) and variational preferences (Maccheroni, Marinacci and Rustichini [62]). With linear utility, these decision criteria correspond to coherent (Artzner et al. [2]) and convex measures of risk (Föllmer and Schied [39]). Henceforth, we postulate that the decision-maker adopts a convex measure of risk and evaluates his future reward according to

$$U(H_\tau) = \inf_{Q\in\mathcal{Q}} \{\mathrm{E}_Q[H_\tau] + c(Q)\}, \tag{2.2}$$

with $\mathcal{Q} = \{Q|Q \sim P\}$ and $c : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$. (We call Q equivalent to P and write $Q \sim P$ if events that have probability zero under P still have probability zero under Q and vice versa.) For our purposes, we have to consider the dynamic version of (2.2), given by

$$U_t(H_\tau) = \inf_{Q \in \mathcal{Q}} \{\mathrm{E}_Q[H_\tau|\mathcal{F}_t] + c_t(Q)\}, \tag{2.3}$$

in which $c_t(Q)$ reflects the esteemed plausibility of the model Q given the information up to time t. In (2.3), and in the rest of this paper, we define for notational convenience sup := ess.sup and inf := ess.inf. The optimal stopping problem at time t_i is then given by

$$V^*_{t_i} = \sup_{\tau \in \mathcal{T}_i} U_{t_i}(H_\tau) = \sup_{\tau \in \mathcal{T}_i} \inf_{Q \in M} \{\mathrm{E}_Q[H_\tau|\mathcal{F}_{t_i}] + c_{t_i}(Q)\}, \tag{2.4}$$

with $\mathcal{T}_i := \{\tau \geq t_i | \tau \in \mathcal{T}\}$.

2.2 Time-Consistency, Dynamic Programming and Assumptions

We now consider the question of which class of plausibility indices (penalty functions) to employ in (2.3)–(2.4). To this end, we first recall the notion of time-consistency in dynamic choice problems under uncertainty. We say that a dynamic evaluation $(U_t(H))_{t\in[0,T]}$ is time-consistent if

$$U_t(H_2) \geq U_t(H_1) \qquad \Rightarrow \qquad U_s(H_2) \geq U_s(H_1), \qquad t \geq s.$$

This means that if, in each state of nature at time t, the reward H_2 is preferred over the reward H_1, then H_2 should also have been preferred over H_1 prior to time t. It turns out that requiring time-consistency of U is equivalent to requiring that U satisfies a dynamic programming principle, which, in turn, is equivalent in our setting to the penalty function associated with U taking a certain form, specified later.

Next, we explain what a change of measure from P to Q implies in our setting. If $Q \sim P$, we denote by D_t the Radon-Nikodym derivative $D_t = \mathrm{E}\left[\frac{dQ}{dP}|\mathcal{F}_t\right]$. In our jump-diffusion setting it is known that, for every model $Q \sim P$, there exist a predictable, $\mathbb{R}^d$-valued, stochastic drift q and a positive, predictable, $\mathbb{R}^k$-valued process λ such that the Radon-Nikodym derivative can be written as

$$D_t = \exp\left\{\int_0^t q_s dW_s + \int_0^t \log\left(\frac{\lambda_s}{\lambda_P}\right) dN_s - \int_0^t \left(\frac{|q_s|^2}{2} + \lambda_s - \lambda_P\right) ds\right\}, \qquad t \in [0,T], \tag{2.5}$$

with $\frac{\lambda_s}{\lambda_P} := (\frac{\lambda_s^1}{\lambda_P^1}, \ldots, \frac{\lambda_s^k}{\lambda_P^k})^\intercal$. In particular, Q is uniquely characterized by q and λ. The stochastic exponential on the right-hand side of (2.5) is also referred to as the Doléans-Dade exponential. By Girsanov's theorem, under Q, $W_t^Q := W_t - \int_0^t q_s ds$ is a Brownian motion and the process $\tilde{N}_t$ has jumps with intensity λ_t. The probabilistic model P occurs when $q = 0$ and $\lambda = \lambda_P$.

We then state the form of a penalty function induced by requiring a dynamic evaluation to be time-consistent (or, equivalently, by requiring recursiveness or Bellman's dynamic programming principle). The result is due to Tang and Wei [82], who generalized a result of Delbaen, Peng and Rosazza Gianin [33] obtained in a Brownian setting to a setting with jumps.

Lemma 1 (Tang and Wei [82]) *Let $U_t(H) = \inf_{\{Q\sim P \text{ on } \mathcal{F}_t\}}\{\mathrm{E}_Q[H|\mathcal{F}_t] + c_t(Q)\}$ for $t \in [0,T]$. The following statements are equivalent:*

(i) U is time-consistent on square-integrable rewards.

(ii) U is recursive (i.e., satisfies Bellman's principle): for every $t \in [0,T]$, $A \in \mathcal{F}_t$ and square-integrable H,

$$U_0(U_t(H)I_A) = U_0(HI_A).$$

(iii) There exists a function

$$\begin{array}{lllllllll} r: & [0,T] & \times & \Omega & \times & \mathbb{R}^d & \times & (-\lambda_P^1,\infty)\times\ldots\times(-\lambda_P^k,\infty) & \to & \mathbb{R}\cup\{\infty\} \\ & (t, & & \omega, & & q, & & v) & \longmapsto & r(t,\omega,q,v), \end{array}$$

which is convex and lower semi-continuous in (q,v), such that

$$c_t(Q) = \mathrm{E}_Q\left[\int_t^T r(s,q_s,\lambda_s-\lambda_P)ds\Big|\mathcal{F}_t\right], \qquad t\in[0,T]. \tag{2.6}$$

Remark 2 In the case of a coherent risk measure, (2.6) corresponds to the existence of a convex, closed, set-valued predictable mapping, say C, taking values in $\mathbb{R}^d\times(-\lambda_P^1,\infty)\times\ldots\times(-\lambda_P^k,\infty)$ such that $r(s,q,v) = I_{C_s}(q,v)$.

Violation of time-consistency would lead to situations in which the decision-maker takes decisions that he knows he will regret in every future state of nature. We rule out such situations. Because in our continuous-time setting time-consistency is equivalent to a penalty function of the form (2.6), we henceforth assume:

(G1) $(c_t(Q))_{t\in[0,T]}$ is of the form

$$c_t(Q) = \mathrm{E}_Q\left[\int_t^T r(s,q_s,\lambda_s-\lambda_P)ds\Big|\mathcal{F}_t\right], \tag{2.7}$$

for a function $r:[0,T]\times\mathbb{R}^d\times(-\lambda_P^1,\infty)\times\ldots\times(-\lambda_P^k,\infty)\to\mathbb{R}_0^+\cup\{\infty\}$ mapping $(t,q,v)\mapsto r(t,q,v)$ that is lower semi-continuous and convex in (q,v) with $r(t,0,0)=0$.

Remark 3 We note that for numerical tractability of the optimal stopping problem, we have postulated in (G1) that r does not depend on ω.

Remark 4 Since by (G1) in particular $c_t\geq 0$ and $c_t(P)=0$, we have $U_t(H)=H$ if H is $\mathcal{F}_t$-measurable. That is, if a reward is known, then there is no uncertainty, and therefore the evaluation returns the reward itself.

We note that q may be viewed as an additional drift in the Brownian motion that the reference model P fails to detect, while $\lambda_s-\lambda_P$ is the deviation of the new jump intensity λ_s under Q from the intensity λ_P under P. Since r is non-negative and $r(s,0,0)=0$, r is minimal in $(0,0)$ with $q=0$ and $\lambda=\lambda_P$. These values of q and λ render the probabilistic model P itself. Therefore, the reference model P is associated with the highest plausibility. (Note that, if we would not make the assumption that $r(s,0,0)=0$, we could redefine the reference model P to correspond to a (q,λ) for which the minimum is attained.) The fact that $(q,\lambda-\lambda_P)\mapsto r(t,q,\lambda-\lambda_P)$ is convex in $(q,\lambda-\lambda_P)$ (with minimum assumed to be in $(0,0)$) explicates that penalty functions giving rise to time-consistent evaluations in our setting may be interpreted as penalty functions for which the divergence penalty function r is directly applied to the additional stochastic drift q affecting the Brownian motion and the deviation of the jump intensity $\lambda - \lambda_P$, instead of to the composition of q and $\lambda - \lambda_P$ appearing in the Radon-Nikodym derivative process (2.5).

We now illustrate the generality of (2.4) and (G1) with some examples of penalty functions satisfying our conditions:

Examples 5 (1) Kullback-Leibler divergence: A prototypical example of the penalty function in (2.4) is the Kullback-Leibler (ϕ-)divergence given by

$$c_t(Q) = \alpha \mathrm{KL}_t(Q|P), \quad \alpha > 0, \quad \text{with} \quad \mathrm{KL}_t(Q|P) = \begin{cases} \mathrm{E}_Q\left[\log\left(\frac{dQ}{dP}\right)\Big|\mathcal{F}_t\right], & \text{if } Q \in \mathcal{Q}; \\ \infty, \text{ otherwise;} & \end{cases}$$

see Csiszár [29], Ben-Tal [9] and Ben-Tal and Teboulle [10, 11]. The Kullback-Leibler divergence is also referred to as the relative entropy and measures the distance between the probabilistic models Q and P; it is used e.g., by Hansen and Sargent [48, 49] to generate model robustness in macroeconomics. The parameter α measures the degree of trust the decision-maker assigns to the reference model P. The limiting case $\alpha \uparrow \infty$ ($\alpha \downarrow 0$) induces a maximal degree of trust (distrust). One may verify (see, for example, Proposition 9.10 in [28]) that in our continuous-time setting, for every Q satisfying $c(Q) < \infty$, $\alpha\mathrm{KL}_t(Q|P)$ is of the form (2.7), where $r(s,q,v) = \frac{\alpha}{2}|q|^2 + \alpha(\lambda_P + v)^\intercal \log(\mathbf{1} + \frac{v}{\lambda_P}) - \mathbf{1}^\intercal v$ with $\mathbf{1} = (1,\ldots,1)^\intercal$, where q and λ correspond to the model Q characterized through (2.5), and where the logarithm should be taken componentwise.

(2) Worst case with discrete scenarios: The decision-maker considers a family of finitely many values $q_{1,s},\ldots,q_{L,s}$ and $\lambda_{1,s},\ldots,\lambda_{L,s}$ for the future drift, q_s, and jump intensity, λ_s, that characterize the model Q through (2.5), with $s > t$. Ex ante these L 'scenarios' are equally plausible and the decision-maker adopts a worst case approach. Consider

$$M = \Big\{ Q \in \mathcal{Q} \Big| \text{ for Lebesgue-a.s. all } s: \ (q_s, \lambda_s) \in \{(q_{i,s}, \lambda_{j,s}) |\, i,j \in \{1,\ldots,L\}\} \Big\}.$$

This corresponds to a penalty function of the form (2.7), with

$$r(s,q,\lambda - \lambda_P) = \begin{cases} 0, & \text{if} \quad (q,\lambda) \in \mathrm{conv}\Big(\{(q_{i,s},\lambda_{j,s})|i,j \in \{1,\ldots,L\}\}\Big); \\ \infty, & \text{otherwise;} \end{cases}$$

where $\mathrm{conv}(\cdot)$ is given by its convex hull. (By redefining the reference model, one may ensure (without loss of generality) that $0 \in \mathrm{conv}\Big(\{(q_{i,s},\lambda_{j,s})|i,j \in \{1,\ldots,L\}\}\Big)$.)

(3) Worst case with ball scenarios: The decision-maker considers alternative and equally plausible probabilistic models Q in a small ball around the reference model P and adopts a worst case approach:

$$M = \Big\{ Q \in \mathcal{Q} \Big| |q_t| \leq \delta_1, \quad |\lambda_t| \leq \delta_2, \text{ for Lebesgue-a.s. all } t \Big\},$$

for $\delta_1, \delta_2 > 0$. This corresponds to a penalty function of the form (2.7), with

$$r(s,q,\lambda - \lambda_P) = \begin{cases} 0, & \text{if} \quad |q| \leq \delta_1, \quad |\lambda - \lambda_P| \leq \delta_2; \\ \infty, & \text{otherwise.} \end{cases}$$

For our next examples we will assume that the n-dimensional Markovian process $(X_t)_{t\in[0,T]}$ is either a geometric Brownian motion with jumps and drift, or a Brownian-Poisson process with drift. In the first case,

$$\frac{dX_t^i}{X_t^i} = \mu^i dt + \sigma^i dW_t + J^i d\tilde{N}_t, \qquad i = 1, \ldots, n, \tag{2.8}$$

while in the second case

$$dX_t^i = \mu^i dt + \sigma^i dW_t + J^i d\tilde{N}_t, \qquad i = 1, \ldots, n, \tag{2.9}$$

for $\mu^i \in \mathbb{R}$, $\sigma^i \in \mathbb{R}^{1\times d}$, and $J^i \in (-1,\infty)^{1\times k}$ (former) or $J^i \in \mathbb{R}^{1\times k}$ (latter). We set $\mu = (\mu^1, \ldots, \mu^n)^\intercal \in \mathbb{R}^n$, $\sigma = (\sigma^1, \ldots, \sigma^n)^\intercal \in \mathbb{R}^{n\times d}$ and $J = (J^1, \ldots, J^n)^\intercal \in (-1,\infty)^{n\times k}$ (former) or $J = (J^1, \ldots, J^n)^\intercal \in \mathbb{R}^{n\times k}$ (latter). In optimal entrance/exit decision problems, such as those provided in the Introduction, X often times satisfies either (2.8) or (2.9) (with or without jumps). In finance, μ^i is commonly referred to as the excess return and represents the compensation for bearing the risky asset i. Now let us continue with some examples of penalty functions that induce time-consistent evaluations, i.e., satisfy (G1), and may be considered in the general problem (2.4), assuming dynamics as in (2.8) or (2.9).

Examples 5 (Continued; with (2.8) or (2.9) valid)

(4) Worst case with mean (partially) known: The decision-maker is certain that the (instantaneous or logarithmic instantaneous) mean return μ^Q lies between a known lower and upper bound, (μ^-) and (μ^+), respectively. As a special case, (μ^-) and (μ^+) coincide (mean fully known). By Girsanov's theorem, under Q,

$$\mu_t^Q = \mu + \sigma q_t + J(\lambda_t - \lambda_P).$$

The resulting models are considered equally plausible and the decision-maker adopts a worst case approach:

$$\begin{aligned} M &= \Big\{Q \in \mathcal{Q} | \mu^- \leq \mu_t^Q \leq \mu^+, \quad \text{for Lebesgue-a.s. all } t\} \\ &= \Big\{Q \in \mathcal{Q} | \mu^- - \mu \leq \sigma q_t + J(\lambda_t - \lambda_P) \leq \mu^+ - \mu, \text{ for Lebesgue-a.s. all } t\Big\}. \end{aligned}$$

We assume $B^- \leq q \leq B^+$ for certain vectors $B^+, B^- \in \mathbb{R}^n$ and $d^- \leq \lambda - \lambda_P \leq d^+$ for vectors $d^+, d^- > -\lambda_P$, to ensure well-posedness. This corresponds to a penalty function of the form (2.7) with

$$r(s, q, \lambda - \lambda_P) = \begin{cases} 0, & \text{if} \quad \mu^- - \mu \leq \sigma q + J(\lambda - \lambda_P) \leq \mu^+ - \mu; \\ \infty, & \text{otherwise.} \end{cases}$$

(5) Pricing with Good-Deal Bounds: A fundamental approach to price financial derivatives that are liquidly traded on the financial market is by replicating the derivatives using other (base) assets and applying no-arbitrage arguments. However, if the financial market is incomplete, a full-blown replication is infeasible, and no-arbitrage arguments only yield price bounds. In general, these price bounds are typically too wide to be practically useful.

One approach to narrowing these bounds is provided by the good-deal pricing approach introduced by Cochrane and Saá-Requejo [27]. Under this approach, only pricing kernels that are sufficiently 'close' to the physical measure are considered. Here, 'close' means that only pricing kernels with a variance below a certain bound are considered. By duality results derived in a celebrated paper by Hansen and Jagannathan [47], this corresponds to ruling out portfolios with a too high Sharpe ratio. The intuition is that portfolios with a very high Sharpe ratio, although strictly speaking not providing arbitrage opportunities, are 'too good to be true' and will be eliminated in a competitive market. In a continuous-time setting, such as ours, the bound for the variance of the pricing kernel is equal to the highest (local) Sharpe ratio in the market, say Λ. In this case, the good-deal bound evaluation $U_t(H_\tau)$ is given by

$$U_t(H_\tau) = \inf_{(q,\lambda)\in C} \mathrm{E}_Q\left[H_\tau\right],$$

with $C = (C_t)_{t\in[0,T]}$ given by (see Björk and Slinko [14])

$$C_t = \Big\{(q, \lambda-\lambda_P) \in \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \ldots \times (-\lambda_P^k, \infty) \Big| \mu + \sigma q + J(\lambda - \lambda_P) = 0 \text{ and } |q|^2 + \sum_{i=1}^k \frac{(\lambda^i - \lambda_P^i)^2}{\lambda_P^i} \leq \Lambda\Big\}.$$

This corresponds to a penalty function of the form (2.7) with

$$r(s, q, \lambda - \lambda_P) = \begin{cases} 0, & \text{if} \quad (q,\lambda) \in C; \\ \infty, & \text{otherwise.} \end{cases}$$

For numerical tractability in what follows, we need the following additional assumption:

(G2) We can simulate i.i.d. copies of $(X_t)_{t\in[0,T]}$.

Assumption (G2) is satisfied in particular if X follows a linear SDE, which holds e.g., in the case of a Brownian motion with drift, a Poisson process with drift, an Ornstein-Uhlenbeck process, or a geometric Brownian motion with drift (with or without Poisson type jumps). But note there are by now also very general results available on exact sampling of more general diffusions and jump-diffusions; see, e.g., Beskos and Roberts [13], Broadie and Kaya [18], Chen and Huang [22], or Giesecke and Smelov [42].

In principle, we would only need assumptions (G1)–(G2). However, if the sublevel sets of the penalty function are non-compact (meaning that models that are 'far away' from the reference model may still yield high plausibility), then the associated optimal stopping problem (2.4) would be ill-posed. To verify, consider, for example, the case that $c = 0$ so that $U_0(H_\tau) = \inf_\omega H_{\tau(\omega)}(\omega)$, which leads to a degenerate (and non-semimartingale) evaluation. Therefore, we will assume additionally to (G1)–(G2) that:

(G3) The domain of r is included in a compact set: for every s,

$$\Big\{(q,\lambda) \in \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \ldots \times (-\lambda_P^k, \infty) \Big| \; r(s, q, \lambda - \lambda_P) < \infty\Big\} \subset C_s,$$

for a compact set $C = (C_s)_{s\in[0,T]} \subset [0,T] \times \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \ldots \times (-\lambda_P^k, \infty)$.

Loosely speaking, condition (G3) states that, if the additional drift q or jump intensity $\lambda - \lambda_P$ that the model Q adds to the Brownian motion or the Poisson process when compared to P is 'too large', then the model Q should not be considered. Condition (G3) may be generalized substantially. In fact, it would be sufficient for our purposes to impose a condition on the penalty function that guarantees that the sublevel sets are (weakly) compact. However, in order to keep the exposition simple, we will impose the somewhat stronger condition (G3).

3 Duality Theory

3.1 Duality Theory of the First Kind

Reconsider the optimal stopping problem (2.4). We show in the Appendix that there exists an optimal stopping family $\left(\tau^*_{t_i}\right)_{t_i\in\{t_0=0,t_1,\ldots,t_L=T\}}$ satisfying

$$V^*_{t_i} = \sup_{\tau\in\mathcal{T}_i} U_{t_i}(H_\tau) = U_{t_i}(H_{\tau^*_{t_i}}), \quad t_i \in \{0,\ldots,T\}. \tag{3.1}$$

Furthermore, we show that Bellman's principle

$$V^*_{t_i} = \max\left(\Pi(t_i, X_{t_i}) + U^h_{t_i}, U_{t_i}(V^*_{t_{i+1}})\right), \quad t_i \in \{0,\ldots,t_{L-1}\}, \tag{3.2}$$

holds, with $U^h_{t_i}$ defined as

$$\begin{aligned} U^h_{t_i} :&= U_{t_i}\left(\sum_{j=i}^{L} h(t_j, X_{t_j})\right) \\ &= \inf_{Q\in\mathcal{Q}}\left\{\mathrm{E}_Q\left[\sum_{j=i}^{L} h(t_j, X_{t_j}) + \int_{t_i}^{T} r(s, q_s, \lambda_s - \lambda_P)ds|\mathcal{F}_{t_i}\right]\right\}; \end{aligned} \tag{3.3}$$

see the Appendix for technical details. Recall that in the absence of model uncertainty, $U_{t_i}(H)$ reduces simply to an ordinary conditional expectation (corresponding to the case in which $c_{t_i}(Q) = \infty$ for $Q \neq P$ and $c_{t_i}(P) = 0$ in (2.3)).

To compute the solution V^* — referred to as the (generalized) Snell envelope — to the optimal stopping problem (2.4), we will rely on the Doob decomposition of V^* into a martingale and a predictable process. However, to do so, we first need to generalize the notion of a (standard) martingale (with respect to an ordinary conditional expectation) to martingales with respect to classes of functionals: We will say that M is a U-martingale if $M_s = U_s(M_t)$, $s, t \in \{t_0 = 0, t_1, \ldots, t_L = T\}$ and $s \leq t$. By time-consistency, this is equivalent to $M_s = U_s(M_T)$ for any s. The class of U-martingales M with $M_0 = 0$ is denoted by $\mathcal{M}^U_0$. Define, for $i = 0,\ldots,L$,

$$A^{*g}_{t_i} := \sum_{j=1}^{i}(U_{t_{j-1}}(V^*_{t_j}) - V^*_{t_{j-1}}), \quad M^{*g}_{t_i} := \sum_{j=1}^{i}(V^*_{t_j} - U_{t_{j-1}}(V^*_{t_j})). \tag{3.4}$$

One may verify that M^{*g} is a U-martingale, A^{*g} is non-decreasing and predictable, $M^{*g}_0 = A^{*g}_0 = 0$, and that

$$V^*_{t_i} = V^*_0 + M^{*g}_{t_i} + A^{*g}_{t_i}, \qquad i = 0,\ldots,L, \tag{3.5}$$

provides a U-Doob decomposition of $V^* = \left(V^*_{t_i}\right)_{t_i\in\{t_0=0,...,T\}}$.

To construct genuine upper bounds to the optimal solution to the stopping problem (2.4), which will converge asymptotically to the true value, our method will exploit an additive dual representation of the optimal stopping problem (2.4), by expanding the well-known dual representation for the standard setting, in which U is just the ordinary conditional expectation (Rogers [73] and Haugh and Kogan [50]). This generalized additive dual representation, the proof of which uses results obtained by Krätschmer and Schoenmakers [57] in a discrete-time setting with $h = 0$, reads as follows:

Proposition 6 *Let $M^{*g} \in \mathcal{M}_0^U$ be the (unique) U-martingale in the U-Doob decomposition (3.5). Then the optimal stopping problem (2.4) has a dual representation*

$$\begin{aligned} V^*_{t_i} &= \inf_{M\in\mathcal{M}_0^U} U_{t_i}\left(\max_{t_j\in\{t_i,...,T\}}\left(\Pi(t_j, X_{t_j}) + U^h_{t_j} + M_T - M_{t_j}\right)\right) \\ &= U_{t_i}\left(\max_{t_j\in\{t_i,...,T\}}\left(\Pi(t_j, X_{t_j}) + U^h_{t_j} + M^{*g}_T - M^{*g}_{t_j}\right)\right), \quad t_i \in \{t_0 = 0, ..., T\}. \end{aligned} \tag{3.6}$$

Remark 7 In the absence of model uncertainty, so that U is a regular conditional expectation, $\mathcal{M}_0^U = \mathcal{M}_0$ is the class of martingales in the usual sense. In this case, interestingly, also

$$V^*_{t_i} = \inf_{M\in\mathcal{M}_0} U_{t_i}\left(\max_{t_j\in\{t_i,...,T\}}\left(\Pi(t_j, X_{t_j}) + U^h_{t_j} + M_{t_i} - M_{t_j}\right)\right), \quad t_i \in \{t_0 = 0, ..., T\}, \tag{3.7}$$

is true. So, for regular conditional expectations, in fact two dual representations hold, namely (3.6) and (3.7). However, (3.7) breaks down in general if U is not a conditional expectation, and only (3.6) is preserved.

3.2 Duality Theory of the Second Kind

Next, we describe the second kind of duality theory on which our method is based. For $t \in [0, T]$, $z \in \mathbb{R}^{1\times d}$ and $\tilde{z} \in \mathbb{R}^{1\times k}$, given a function r specifying the penalty function c through (2.7), we define a function g by Fenchel's duality as follows:

$$g(t, z, \tilde{z}) := \inf_{(q,\lambda-\lambda_P)\in C_t}\{zq + \tilde{z}(\lambda - \lambda_P) + r(t, q, \lambda - \lambda_P)\}, \tag{3.8}$$

with C_t induced by assumption (G3). Note that by assumption (G3), g thus defined is Lipschitz continuous. Note furthermore that (G3) is satisfied in all our Examples 5 above, except for the Kullback-Leibler divergence. In this case, however, we will restrict our analysis to terminal conditions that are Lipschitz continuous in the Brownian motion and the Poisson process, so that the domains of z and $\tilde{z}$ are bounded, and g may be considered to be Lipschitz continuous as well. Furthermore, suppose that, for every exercise date t_j, $j = 0, \ldots, L$, we have a fine time grid $\pi_j = \{s_{j0} = t_j < s_{j1} < \ldots < s_{jP} = t_{j+1}\}$. Denote the corresponding overall time grid by $\pi = \{s_{00}, s_{01}, \ldots, s_{LP}\}$. The following theorem provides a way to practically compute M^{*g} in (3.4) by connecting it to specific semi-martingale dynamics that can be dealt with numerically efficiently.

Theorem 8 *(a) There exists a unique square integrable predictable $(Z^h, \tilde{Z}^h)$ such that*

$$dU^h_t = -g(t, Z^h_t, \tilde{Z}^h_t)dt + Z^h_t dW_t + \tilde{Z}^h_t d\tilde{N}_t, \quad \textit{for } t \in (t_j, t_{j+1}], \tag{3.9}$$

and $U^h_{t_j} = U^h_{t_j+} + h(t_j, X_{t_j})$, *for each* $j \in \{0, \ldots, L-1\}$. *Furthermore, there exists a unique square-integrable predictable* $(Z^*, \tilde{Z}^*)$ *such that*

$$dU_t(V^*_{t_{j+1}}) = -g(t, Z^*_t, \tilde{Z}^*_t)dt + Z^*_t dW_t + \tilde{Z}^*_t d\tilde{N}_t, \text{ for } t \in [t_j, t_{j+1}], j \in \{0, ..., L-1\}. \quad (3.10)$$

(b) For $t \in [0, T]$, $(Z^*, \tilde{Z}^*)$ *from part (a) satisfy*

$$M^{*g}_t = U_t(M^{*g}_T) = -\int_0^t g(s, Z^*_s, \tilde{Z}^*_s)ds + \int_0^t Z^*_s \, dW_s + \int_0^t \tilde{Z}^*_s \, d\tilde{N}_s. \quad (3.11)$$

Remark 9 Note that by Remark 4 and (3.3), we have terminal conditions $U^h_T = h(T, X_T)$ and $U_{t_{j+1}}(V^*_{t_{j+1}}) = V^*_{t_{j+1}}$, for $j = 0, \ldots, L-1$, in (3.9) and (3.10). Hence, given $U^h_{t_{j+1}}$ and $V^*_{t_{j+1}}$, we may compute $U^h_{t_j}$ and $U_{t_j}(V^*_{t_{j+1}})$ through the relationships given in Theorem 8(a); $V^*_{t_j}$ can then be obtained by Bellman's principle (3.2).

Remark 10 As $U_{t_{j+1}}(V^*_{t_{j+1}}) = V^*_{t_{j+1}}$, we can write, by Theorem 8(a), for $t \in [t_j, t_{j+1}]$,

$$U_t(V^*_{t_{j+1}}) = V^*_{t_{j+1}} + \int_t^{t_{j+1}} g(s, Z^*_s, \tilde{Z}^*_s)ds - \int_t^{t_{j+1}} Z^*_s dW_s - \int_t^{t_{j+1}} \tilde{Z}^*_s d\tilde{N}_s. \quad (3.12)$$

Similarly, it follows that, for $t \in (t_j, t_{j+1}]$,

$$U^h_t = U^h_{t_{j+1}} + \int_t^{t_{j+1}} g(s, Z^h_s, \tilde{Z}^h_s)ds - \int_t^{t_{j+1}} Z^h_s dW_s - \int_t^{t_{j+1}} \tilde{Z}^h_s d\tilde{N}_s. \quad (3.13)$$

Remark 11 Note that if $g \equiv 0$ would hold in (3.10), then the increments of the evaluation U were increments of a (standard) martingale. In that case, $U_t(H)$ would simply be a (standard) martingale, and, because $U_T(H) = H$, correspond to the (regular) conditional expectation $U_t(H) = \mathrm{E}[H|\mathcal{F}_t]$. However, our decision-maker is ambiguity averse and considers alternative probabilistic models with potentially different degrees of esteemed plausibility. This leads to $g \leq 0$, which by (3.12)–(3.13) decreases the evaluation. Note furthermore that the couple Z^* and $\tilde{Z}^*$ may be viewed as a measurement of the degree of 'variability' underlying the evaluation — in the same way as the volatility in standard asset pricing models in finance — due to the Brownian motion and the jump component, respectively: The larger $|Z^*|$ ($|\tilde{Z}^*|$), the more variability comes from the local Gaussian part (the jump component) of the model. Because $g(t, \cdot) \leq 0$ is concave in $(z, \tilde{z})$, with maximum in $(0, 0)$, greater variability will lead to a larger 'penalty' term.

Equations (3.9)–(3.10) are also referred to as backward stochastic differential equations (BSDEs)[6] and their solution is often referred to as a (conditional) g-expectation. A g-expectation inherits many properties from a regular (conditional) expectation, such as monotonicity, translation invariance, and the tower property, but not linearity; for further details, see, for instance, the survey of Peng [67].

[6]Formally, given a terminal payoff $H \in L^2$ and a function $g : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$, the solution to the corresponding BSDE is a triple of square-integrable and suitably measurable processes $(Y, Z, \tilde{Z})$ satisfying

$$dY_t = -g(t, Z_t, \tilde{Z}_t)dt + Z_t dW_t + \tilde{Z}_t d\tilde{N}_t, \quad \text{and} \quad Y_T = H.$$

To conclude the exposition of the duality theory of the second kind, let us, for illustration purposes, employ the penalty functions of Examples 5 and compute the corresponding g's using (3.8). These g functions will later be used in numerical illustrations.

Examples 12 (1) Kullback-Leibler divergence:

$$g(t,z,\tilde{z}) = -\frac{|z|^2}{2\alpha} - \alpha \sum_{i=1}^{k} \lambda^i \left(e^{-\tilde{z}^i/\alpha} + \frac{\tilde{z}^i}{\alpha} - 1 \right).$$

(2) Worst case with discrete scenarios:

$$g(t,z,\tilde{z}) = \min_{i=1,\dots,L} z q_{i,t} + \min_{j=1,\dots,L} \tilde{z}(\lambda_{j,t} - \lambda_P).$$

(3) Worst case with ball scenarios: Suppose without loss of generality that $|\lambda_P| \geq \delta_2$. Then,

$$g(t,z,\tilde{z}) = -\delta_1 |z| - \delta_2 |\tilde{z}|.$$

(4) Worst case with mean (partially) known and (2.8) or (2.9): From (3.8),

$$g(t,z,\tilde{z}) = \inf_{(q,\lambda-\lambda_P)\in C_t} \{zq + \tilde{z}(\lambda - \lambda_P)\}, \tag{3.14}$$

with

$$C_t = \Big\{ (q, \lambda - \lambda_P) \in \mathbb{R}^d \times \mathbb{R}^k \Big| \mu^- - \mu \leq \sigma q + J(\lambda - \lambda_P) \leq \mu^+ - \mu, \\ B^- \leq q \leq B^+, \quad d^- \leq \lambda - \lambda_P \leq d^+ \Big\}.$$

In general, g cannot be simplified further, although it can in specific cases, such as $(\mu^-) = (\mu^+)$ (mean fully known). However, in view of (3.14), for fixed $(t, z, \tilde{z})$, g can be obtained as the solution to a linear programming problem.

(5) Good-Deal Bounds and (2.8) or (2.9): Let $b = -\mu$ and let $A = (\sigma, J)$ be a matrix mapping from $\mathbb{R}^d \times \mathbb{R}^k$ to $\mathbb{R}^n$. Define $\langle (z,\tilde{z}), (q, \lambda - \lambda_P) \rangle := qz + \tilde{z}(\lambda - \lambda_P)$. Furthermore, for $q \in \mathbb{R}^d$ and $v \in \mathbb{R}^k$, define $|\langle q, v \rangle|_* := \sqrt{|q|^2 + \left|\frac{v}{\lambda_P}\right|^2}$, where the division is defined componentwise and $|\cdot|$ denotes the Euclidean norm. Then,

$$g(t,z,\tilde{z}) = \inf_{(q,\lambda-\lambda_P)\in C} \langle (z,\tilde{z}), (q, \lambda - \lambda_P) \rangle,$$

with C given by

$$C = \Big\{ (q, \lambda - \lambda_P) | A(q, \lambda - \lambda_P)^\intercal = b \quad \text{and} \quad |\langle q, \lambda - \lambda_P \rangle|_* \leq \sqrt{\Lambda} \Big\}.$$

(Note that the case of no-arbitrage pricing corresponds to $\Lambda = \infty$.) If the set C is non-empty, this optimization problem has an explicit solution: Let $P_W(0)$ be the projection of 0 onto the set $W := \{x | Ax = b\}$ in the $|\cdot|_*$ norm. Using Lagrangian duality techniques, it is not hard to verify that

$$g(t,z,\tilde{z}) = -\left(\sqrt{\Lambda} - |P_W(0)|_*\right) \sqrt{|z|^2 + \Big| \sum_{i=1}^{k} \tilde{z}^i \lambda_P^i \Big|^2} + \langle (z,\tilde{z}), P_W(0) \rangle.$$

This concludes our examples.

4 The Algorithm: General Outline

Our method is composed of three steps. Theorem 8 ('duality theory of the second kind') jointly with Bellman's principle (3.2) will serve as a *first* stepping stone for our approach, by providing a practical way to find U-martingales, to be employed in the dual representation (3.6), which is our *second* stepping stone ('duality theory of the first kind'). In particular, Theorem 8(a) yields that, to construct the U-martingale M^{*g} in the U-Doob decomposition (3.5) of the (generalized) Snell envelope V^* solving our optimal stopping problem, we only have to find $(Z^*, \tilde{Z}^*)$ for every $(V_s^*)_{t_j < s \le t_{j+1}}$. And this can be achieved either by solving a PDE (or PIDE in the presence of jumps) or by least squares Monte Carlo regression and backward stochastic calculus. We will adopt the latter approach. It will provide an approximated upper bound on the solution V^* to the optimal stopping problem, in view of the dual representation (3.6) in Proposition 6. While this bound will be seen to converge to the true optimal solution asymptotically and is an approximated upper bound at the pre-limiting level, it is not a genuine upper bound estimate to the true optimal solution as it is not 'biased high', that is, biased above the Snell envelope V^*. This means that on average this upper bound may not provide enough protection. Our *third* stepping stone, then, is the introduction of backward-forward simulation in the context of BSDEs to obtain a genuine (biased high) upper bound on the solution V^* to our stopping problem (see Step (3.) below).

Therefore, we will:

Step (1.) Exploiting duality theory of the second kind:

- Step (1.a.) Compute an approximation to $(U_{t_j}^h)_{t_j \in \{0,\dots,T\}}$ in (3.3) through backward recursion, using (3.9) and $U_T^h = 0$. This involves least squares Monte Carlo regression.
- Step (1.b.) Set $V_T^* = H_T = \Pi(T, X_T)$ and do a backward recursion over t_j: Given $V_{t_{j+1}}^*$, compute $(Z_s^*, \tilde{Z}_s^*)_{s \in [t_j, t_{j+1}]}$ and $U_s(V_{t_{j+1}}^*)_{t_j < s \le t_{j+1}}$ through (3.10). This involves least squares Monte Carlo regression. We can then set $V_{t_j}^* = \max\left(\Pi(t_j, X_{t_j}) + U_{t_j}^h, U_{t_j}\left(V_{t_{j+1}}^*\right)\right)$, by (3.2). If (and as long as) $t_j > 0$, set $j = j-1$, and repeat the same computation. Otherwise, go to Step (1.c.) below.
- Step (1.c.) Given the whole path of $(Z_s^*, \tilde{Z}_s^*)_{s \in [0,T]}$, compute an approximation to $(M_{t_j}^{*g})_{t_j \in \{t_1,\dots,T\}}$ through (3.11).

Step (2.) Exploiting duality theory of the first kind, obtain an approximated upper bound to V_0^* through (3.6). This involves least squares Monte Carlo regression.

Step (3.) Introducing backward-forward simulation:

- Step (3.a.) Compute a genuine (biased high) upper bound to $(U_{t_j}^h)_{t_j \in \{0,\dots,t_{L-1}\}}$ by using the least squares Monte Carlo results obtained under Step (1.a.) as input in Monte Carlo forward simulations.
- Step (3.b.) Compute a genuine (biased high) upper bound to the Snell envelope V_0^* by using the least squares Monte Carlo results obtained under Steps (1.) and (2.) as input in Monte Carlo forward simulations.

We describe our algorithm (in particular, Steps (1.)–(3.) above) in detail in Section 6, but already preview the following results. Since our optimal stopping problem is Markovian, there exists a function $v^*:[0,T]\times\mathbb{R}^n\to\mathbb{R}$ such that $V_t^*=v^*(t,X_t)$. In particular, $V_0^*=v^*(0,X_0)$. Our method will ultimately provide an approximation to the function v^*, using Monte Carlo simulation techniques that are standard in e.g., the (no-ambiguity) American option literature. This entails that, for a finite number of Monte Carlo simulations, our approximation will inherently be random, as it depends on the stochastic nature of simulations. Our method, then, will be proven to have the following two appealing properties (see Theorem 15 below for the formal results):

(i) Our approximation converges to the true value as the mesh size of the time grid tends to zero and the numbers of Monte Carlo simulations and basis functions tend to infinity.

(ii) For every finite time grid and finite numbers of Monte Carlo simulations and basis functions, our approximation provides a genuine (biased high) upper bound to the true value.

Our numerical examples provided below illustrate that, already after a limited number of Monte Carlo simulations, our method yields rather close estimates in realistic settings. Moreover, by property (ii) above, for a finite time grid and a finite number of simulations, the genuine upper bound will also provide a safety buffer, i.e., a maximal amount the decision-maker (firm or buyer) should be willing to pay or reserve for the action or undertaking. The examples also illustrate the generality of our approach and the relevance of ambiguity to optimal stopping.

5 Numerical Examples

In this section, we present numerical results obtained by applying our algorithm to a few examples of robust optimal stopping problems. We consider two stochastic processes, X_i, $i=1,2$, with dynamics (cf. (2.8))

$$\frac{dX_t^i}{X_t^i}=\mu^i dt+\sigma^i dW_t^i+J^i d\tilde{N}_t^i,\quad X_0^i=x_0^i,$$

where W_t^i is a one-dimensional standard Brownian motion, $\sigma^i\geq 0$ denotes the diffusion coefficient (volatility), $\tilde{N}_t^i$ is a one-dimensional compensated Poisson process with intensity $\lambda_P^i\geq 0$, and $J^i\in(-1,\infty)$ denotes the jump size. The processes W_t^i and $\tilde{N}_t^i$ are assumed to be mutually independent.

In Sections 5.1 and 5.2, we consider the optimal (simple) reward problem (i.e., $h\equiv 0$). We first analyze in Section 5.1 the setting in which the jump component in X_i is absent (i.e., $J^i\equiv\lambda_P^i\equiv 0$ for $i=1,2$), and next consider in Section 5.2 the general setting with non-trivial jump component. This problem occurs e.g., in American-style derivative pricing in finance, in which case the drift μ^i under the reference model is equal to $\rho-\delta$ (for $i=1,2$), where ρ represents the risk-free rate and δ the dividend rate. In these sections, we deal specifically with simple rewards of the form

$$\Pi(t,X_t)=\exp(-\rho t)\left(X_t-K\right)^+,\qquad\text{or}\qquad \Pi(t,X_t)=\exp(-\rho t)\left(K-X_t\right)^+,$$
$$\text{or}\qquad \Pi(t,X_t)=\exp(-\rho t)\left(\max\left\{X_t^1,X_t^2\right\}-K\right)^+,$$

where we write $X_t=\left(X_t^1,X_t^2\right)$ in the two-dimensional and $X_t=X_t^1$ in the one-dimensional case. Here, $K\geq 0$ is the fixed cost (or reward) associated with exercising. We assume that the agent always has the possibility not to exercise so that his exercising payoff can never become negative. In finance, these rewards resemble the to time 0 discounted payoffs when exercising at time t of Bermudan call, put and max-call options with strike price equal to K, respectively. In Section 5.3, we analyze the optimal entrance problem, with non-trivial h. There, we assume that Π and h are given by

$$\Pi\left(t, X_t\right) = -\exp\left(-\rho t\right)\kappa \qquad \text{and} \qquad h\left(t, X_t\right) = \exp\left(-\rho t\right)\left(X_t - \xi\right),$$

for a fixed irreversible cost $\kappa \geq 0$ and where h measures the payoff, X_t, after entering the market minus the running costs, $\xi \geq 0$, taking into account discounting. An appropriate choice of the basis functions m^M, ψ^M and $\tilde{\psi}^M$, $M \in \mathbb{N}$, that we employ in the least squares Monte Carlo regressions, is crucial to obtain tight upper bounds. We will state them in detail for the various examples that we analyze.

5.1 Optimal Reward Problem with a Geometric Brownian Motion

Let us first consider the situation in which the jump component is absent (i.e., $J^i \equiv \lambda_P^i \equiv 0$ for $i = 1, 2$). We will provide numerical results for the univariate and bivariate cases. Following Andersen and Broadie [1], we take the following parameter set under the reference model:

$$\rho = 0.05,\ \delta = 0.1,\ \sigma = 0.2,\ K = 100,\ T = 3 \text{ years}.$$

Furthermore, we consider exercise dates given by $t_j = \frac{jT}{9}$, $j = 0, \ldots, 9$, and a fine grid $\{s_{jp}\}$ with $\Delta_{jp} = s_{j(p+1)} - s_{jp} = 1/1{,}500$. For the choice of basis functions, we follow Andersen and Broadie [1] by including still-alive European options and corresponding option deltas. Our results are based on 10,000 simulated trajectories for the calculation of the regression coefficients in Step (1.b.) and the U-martingale increments in Step (1.c.), the approximated upper bound to V^* in Step (2.), and the genuine upper bound to V^* in Step (3.b.). For Step (2.), the basis functions ψ_t^M are enlarged by the martingale and maximum processes, as included in the Markov process $\mathcal{X}$ (defined in Step (2.); see Section 6.2). This applies to both the univariate and the bivariate cases.

5.1.1 Univariate Case

In the univariate case, we restrict attention to the simple reward $\Pi(t, X_t) = \exp(-\rho t)\left(X_t - K\right)^+$ (i.e., call options). Let $E_\Pi\left(t, X_t, T\right)$ denote the price at time t of a European call option with maturity time T and let $\frac{\partial E_\Pi(t, X_t, T)}{\partial X_t}$ denote its derivative with respect to the underlying risky asset's price. For m_t^M, $t_j \leq t \leq t_{j+1}$, we take the set of basis functions given by

$$\left\{1,\ \mathrm{Pol}_2\left(X_t\right),\ \mathrm{Pol}_3\left(E_\Pi\left(t, X_t, t_{j+1}\right)\right),\ \mathrm{Pol}_3\left(E_\Pi\left(t, X_t, t_L\right)\right)\right\}. \tag{5.1}$$

Here, $\mathrm{Pol}_n\left(y\right)$ denotes the set of monomials up to degree n of a vector y. Furthermore, for ψ_t^M (corresponding to the Brownian motion driven part of the BSDE), $t_j \leq t \leq t_{j+1}$, we take the set

$$\left\{1,\ X_t \frac{\partial E_\Pi\left(t, X_t, t_{j+1}\right)}{\partial X_t},\ X_t \frac{\partial E_\Pi\left(t, X_t, t_L\right)}{\partial X_t}\right\}. \tag{5.2}$$

Kullback-Leibler divergence: First, we consider the case of the Kullback-Leibler divergence for different values of its parameter α. The results are in Table 1. The last column, with

$\alpha = \infty$, has to be interpreted as $g \equiv 0$. Thus, it corresponds to the (limiting) case of a standard conditional expectation.

	α				
x_0	10	100	10^4	10^6	∞
90	2.4405	4.0546	4.4049	4.4088	4.4088
	2.4932	*4.0673*	*4.4662*	*4.4708*	*4.4708*
	(0.0003)	(0.0009)	(0.0013)	(0.0013)	(0.0013)
100	4.6077	7.4023	7.9848	7.9913	7.9914
	4.8251	*7.3887*	*8.0328*	*8.0402*	*8.0403*
	(0.0005)	(0.0012)	(0.0018)	(0.0019)	(0.0019)
110	10.0281	12.2732	13.1574	13.1661	13.1662
	10.1468	*12.2934*	*13.1656*	*13.1756*	*13.1757*
	(0.0008)	(0.0015)	(0.0024)	(0.0024)	(0.0024)

Table 1: Approximated and genuine (in italics) upper bounds to robust call option prices using the Kullback-Leibler divergence with different values of its parameter α and depending on the initial value of the underlying risky asset's price x_0. Standard errors for the genuine upper bounds are given in parentheses. Univariate case.

Only in the case of $\alpha = \infty$ we have reference values, provided e.g., by Andersen and Broadie [1]. They appear to be very close to our values. For example, for $x_0 = 100$, the *true* value is 7.98, which is to be compared to our approximated and genuine upper bounds equal to 7.99 and 8.04, respectively. With an increase in α we observe an, initially rapid, increase in the robust call option's value. In general, we observe that American call option values may decrease substantially when ambiguity is taken into account. In view of the fact that our approximated and genuine upper bounds turn out to be quite close, we restrict attention henceforth to the approximated upper bounds, when assessing numerically the impact of ambiguity on optimal stopping problems.

Worst case with mean partially known: Next, we consider the example of worst case with mean partially known, where we either take $\mu^- = -0.05$ and vary μ^+ or we take $\mu^+ = -0.05$ and vary μ^-. Furthermore, we choose large values for the parameters B^+ and B^- such that the resulting driver is practically independent of these parameters (specifically, we take $B^+ = 1{,}000$ and $B^- = -1{,}000$). The results are in Table 2.

	μ^+		$\mu^+ = \mu^-$	μ^-	
	0.05	0	-0.05	-0.1	-0.15
x_0	($\mu^- = -0.05$)			($\mu^+ = -0.05$)	
90	4.41	4.41	4.41	2.62	1.64
100	7.99	7.99	7.99	5.57	4.10
110	13.17	13.17	13.17	10.84	10.17

Table 2: Upper bounds to robust call option prices under the worst case with mean partially known example with different values of the parameters μ^+ and μ^- and depending on the initial value of the underlying risky asset's price x_0. Univariate case.

We observe from Table 2 that the robust call option values are insensitive to changes in μ^+ for given μ^-. By contrast, the robust call option values are quite sensitive to changes in μ^- for given μ^+. Of course, the case of $\mu^+ = \mu^-$ yields the case of a standard conditional expectation. It agrees with the last column of Table 1. Note further that, without jumps, the worst case with mean partially known example would agree with the worst case with ball scenarios example (see Examples 5 (3)) whenever $\left|\frac{\mu^- - \mu}{\sigma}\right| = \left|\frac{\mu^+ - \mu}{\sigma}\right| = \delta_1$ subject to $\frac{\mu^+ - \mu}{\sigma} \leq B^+$ and $\frac{\mu^- - \mu}{\sigma} \geq B^-$ holds.

5.1.2 Bivariate case

Next, we consider the bivariate case. In the bivariate case, we analyze the simple reward $\Pi(t, X_t) = \exp(-\rho t)\left(\max\left\{X_t^1, X_t^2\right\} - K\right)^+$ (i.e., a max-call option). We denote the price of a European max-call option at time t with maturity time T by $E_\Pi\left(t, X_t, T\right)$. It is given by the following expression (Johnson [54]):

$$
\begin{aligned}
E_\Pi\left(t, X_t, T\right) \quad = \quad & \sum_{l=1}^{2} X_t^l \frac{e^{-\delta(T-t)}}{\sqrt{2\pi}} \int\limits_{\left(-\infty, d_+^l\right]} \exp\left[-\frac{1}{2}z^2\right] \prod_{l'=1, l' \neq l}^{2} \mathcal{N}\left(\frac{\log\left(\frac{X_t^l}{X_t^{l'}}\right)}{\sigma\sqrt{T-t}} - z + \sigma\sqrt{T-t}\right) dz \\
& - K e^{-\rho(T-t)} + K e^{-\rho(T-t)} \prod_{l=1}^{2}\left(1 - \mathcal{N}\left(d_-^l\right)\right),
\end{aligned}
$$

with

$$
d_-^l := \frac{\log\left(\frac{X_t^l}{K}\right) + \left(\rho - \delta - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_+^l := d_-^l + \sigma\sqrt{T-t}.
$$

Here, $\mathcal{N}$ denotes the standard Gaussian cumulative distribution function. For the corresponding option delta, denoted by $\frac{\partial E_\Pi(t, X_t, T)}{\partial X_t^l}$, it follows that

$$
\frac{\partial E_\Pi\left(t, X_t, T\right)}{\partial X_t^l} \quad = \quad \frac{e^{-\delta(T-t)}}{\sqrt{2\pi}} \int\limits_{\left(-\infty, d_+^l\right]} \exp\left[-\frac{1}{2}z^2\right] \prod_{l'=1, l' \neq l}^{2} \mathcal{N}\left(\frac{\log\left(\frac{X_t^l}{X_t^{l'}}\right)}{\sigma\sqrt{T-t}} - z + \sigma\sqrt{T-t}\right) dz.
$$

In Step (1.b.) of our algorithm, we choose the same set of basis functions for m_t^M as in the univariate case (see (5.1)). For the Brownian motion driven part, we have to adapt to the two-dimensionality of our problem, and for ψ_t^M we now consider the set

$$
\left\{1, \; \left(X_t^l \frac{\partial E_\Pi\left(t, X_t, t_{j+1}\right)}{\partial X_t^l}\right)_{1 \leq l \leq 2}, \; \left(X_t^l \frac{\partial E_\Pi\left(t, X_t, t_L\right)}{\partial X_t^l}\right)_{1 \leq l \leq 2}\right\}.
$$

The parameters are chosen as in the univariate case, with common μ^i, σ^i and J^i for $i = 1, 2$, and assuming independence between W^1 and W^2 and between N^1 and N^2.

Kullback-Leibler divergence: In Table 3, we consider the Kullback-Leibler divergence for different values of its parameter α.

x_0	α: 10	100	10^4	10^6	∞	BBS
90	4.49	7.49	8.12	8.12	8.12	8.09
100	8.12	12.98	13.97	13.98	13.98	13.96
110	13.41	20.12	21.43	21.45	21.45	21.46

Table 3: Upper bounds to robust max-call option prices using the Kullback-Leibler divergence with different values of its parameter α and depending on the common initial value of the underlying risky assets' prices x_0. Bivariate case. The last column displays reference values for the case of $\alpha = \infty$ (or $g \equiv 0$) obtained by Belomestny, Bender and Schoenmakers [6] (BBS).

We observe again that with an increase in α, the robust option value initially increases rapidly. The last column, with $\alpha = \infty$, yields the case of a standard conditional expectation ($g \equiv 0$). Only in this special case do we have reference values given e.g., in Belomestny, Bender and Schoenmakers [6] (BBS). For $g \equiv 0$, our values are very close to the upper bounds obtained by Belomestny, Bender and Schoenmakers [6].

Worst case with mean partially known: Next, we consider the worst case with mean partially known example. Upper bounds on the robust option price are given in Table 4, for different values of μ^+ and μ^-.

	μ^+	$\mu^+ = \mu^-$		μ^-
	0.05	-0.05		-0.15
x_0	$(\mu^- = -0.05)$	KLLSS	BBS	$(\mu^+ = -0.05)$
90	8.12	8.12	8.09	3.02
100	13.98	13.98	13.96	7.08
110	21.45	21.45	21.46	13.58

Table 4: Upper bounds to robust max-call option prices under the worst case with mean partially known example with different values of the parameters μ^+ and μ^- and depending on the common initial value of the underlying risky assets' prices x_0. Bivariate case. In the fifth column we display reference values as obtained by Belomestny, Bender and Schoenmakers [6] (BBS) for the case of $\mu^+ = \mu^-$ (or $g \equiv 0$).

Similar to the univariate case, the robust max-call option values are insensitive to changes in μ^+ for given μ^-, and are quite sensitive to changes in μ^- for given μ^+. The case of $\mu^+ = \mu^-$ yields the case of a standard conditional expectation and agrees with the last two columns of Table 3. Furthermore, as in the univariate case (without a jump component), the worst case with mean partially known example agrees with the worst case with ball scenarios, for specific parameter sets.

5.2 Optimal Reward Problem with a Jump-Diffusion

Let us now consider the situation in which the Poissonian jump component is present, next to the continuous diffusion component. We restrict attention to the univariate case. We take the following parameter set under the reference model:

$$\rho = 0.04,\ \delta = 0,\ \sigma = 0.2,\ J = 0.06,\ K = 100,\ T = 1 \text{ year},$$

and consider different values of λ_P. The exercise dates are given by $t_j = \frac{jT}{10}$, $j = 0, ..., 10$, and the fine grid is given by $\Delta_{jp} = 1/100$.

In this subsection, we analyze the simple reward $\Pi(t, X_t) = \exp(-\rho t)\,(K - X_t)^+$ (i.e., put options). We let $E_\Pi\,(t, X_t, T)$ denote the price at time t of a European put option with maturity time T and we let $\frac{\partial E_\Pi(t,X_t,T)}{\partial X_t}$ denote its derivative with respect to the underlying risky asset's price. This price is given by the following expression (see e.g., Cont and Tankov [28]):

$$E_\Pi\,(t, X_t, T) = e^{-\rho(T-t)} \sum_{n\geq 0} \frac{e^{-\lambda_P(T-t)}\,(\lambda_P(T-t))^n}{n!} \mathcal{BS}\left(T-t, X_t^{(n)}, \sigma\right), \tag{5.3}$$

where

$$X_t^{(n)} = X_t \exp\left(nJ - \lambda_P(T-t)\exp(J) + \lambda_P(T-t)\right),$$

and where $\mathcal{BS}$ denotes the Black-Scholes price of the corresponding European put option.[7]

In Step (1.b.), we choose for m_t^M, $t_j \leq t \leq t_{j+1}$, the set of basis functions given in (5.1), but with X_t now a jump-diffusion and with $E_\Pi(t, X_t, T)$ the price of a European put option. The basis functions for the Brownian motion driven part of the BSDE, ψ_t^M, and the jump part, $\tilde{\psi}_t^M$, are both given by (5.2). Our numerical results are based on 5,000 simulated trajectories for all *relevant* steps of the algorithm. For Step (2.), ψ_t^M and $\tilde{\psi}_t^M$ are enlarged by the martingale and maximum processes, included in the Markov process $\mathcal{X}$, as in the previous subsection.

Kullback-Leibler divergence: In Table 5, we deal with the Kullback-Leibler divergence and present results for different values of its parameter α and of the jump intensity λ_P.

		α				
	x_0	10	100	10^4	10^6	∞
$J = \lambda_P = 0$	90	10.42	11.54	11.83	11.83	11.83
	100	4.63	6.14	6.40	6.41	6.41
	110	2.19	3.06	3.22	3.22	3.22
$J = 0.06,\ \lambda_P = 1$	90	10.54	11.79	12.09	12.10	12.10
	100	4.85	6.46	6.74	6.74	6.74
	110	2.37	3.33	3.50	3.50	3.50
$J = 0.06,\ \lambda_P = 3$	90	11.00	12.52	12.89	12.89	12.89
	100	5.27	7.12	7.45	7.46	7.46
	110	2.74	3.89	4.10	4.10	4.10

Table 5: Upper bounds to robust put option prices using the Kullback-Leibler divergence with different values of its parameter α and of the jump intensity λ_P, and depending on the initial value of the underlying risky asset's price x_0. Univariate case.

We observe from Table 5 that the put options become more valuable if the jump intensity under the reference model increases, and depreciate in the presence of ambiguity, as expected.

Worst case with ball scenarios: In the worst case with ball scenarios example we provide results for different values of δ_1 and δ_2. These are given in Table 6.

[7] The formula given in Cont and Tankov [28] pertains to the case of Gaussian jumps. Here, we face the special case of a fixed degenerate jump size, which can be viewed as a Gaussian jump with mean J and volatility equal to zero. We calculate an approximation to (5.3), which involves an infinite sum, but converges very rapidly.

	x_0	$\delta_1 = \delta_2 = 0.5$	$\delta_1 = 0.5,\ \delta_2 = 1$	$\delta_1 = 1,\ \delta_2 = 0.5$
	90	10.48	10.37	10.20
$\lambda_P = 1$	100	4.22	3.92	3.13
	110	1.73	1.54	1.05
	90	10.90	10.74	10.43
$\lambda_P = 3$	100	4.77	4.44	3.58
	110	2.15	1.92	1.34

Table 6: Upper bounds to robust put option prices under the worst case with ball scenarios example with different values of the parameters δ_1 and δ_2 and of the jump intensity λ_P, and depending on the initial value of the underlying risky asset's price x_0. Univariate case.

Upon comparing the results in Table 6 to the corresponding no-ambiguity results in the last column of Table 5 (with $\alpha = \infty$ hence $g \equiv 0$), we observe that the put options clearly depreciate in the presence of ambiguity with respect to the drift in the Brownian motion (as measured by δ_1) and to the jump intensity (as measured by δ_2).

Worst case with mean partially known: Next, we consider the worst case with mean partially known example. We take $B^+ = 0.5$, $B^- = -0.5$, $d^+ = 0.5$, and $d^- = -0.25$. The results are in Table 7.

		μ^+	$\mu^+ = \mu^-$	μ^-
		0.14	0.04	-0.06
	x_0	$(\mu^- = 0.04)$		$(\mu^+ = 0.04)$
	90	10.48	11.83	11.83
$J = \lambda_P = 0$	100	4.26	6.41	6.41
	110	1.74	3.22	3.22
	90	10.62	12.05	12.05
$J = 0.06,\ \lambda_P = 1$	100	4.57	6.67	6.68
	110	1.95	3.44	3.45
	90	11.10	12.85	12.85
$J = 0.06,\ \lambda_P = 3$	100	5.16	7.40	7.40
	110	2.41	4.04	4.05

Table 7: Upper bounds to robust put option prices under the worst case with mean partially known example with different values of the parameters μ^+ and μ^- and of the jump intensity λ_P, and depending on the initial value of the underlying risky asset's price x_0. Univariate case.

Note that for put options, the pattern observed is different from (opposite to) what we observed for call options in Tables 2 and 4, in the sense that uncertainty about a potentially lower drift does not impact the put option values, in contrast to the call option values.

Good-deal bounds: In view of the presence of jumps, it is now sensible to also consider good-deal bounds. We provide results for different values of λ_P and Λ, given in Table 8.

x_0	$\lambda_P = 1$, $\Lambda = 0.5$	$\lambda_P = 3$, $\Lambda = 0.5$	$\lambda_P = \Lambda = 1$	$\lambda_P = 3$, $\Lambda = 1$
90	11.50	12.11	10.44	10.70
100	6.06	6.54	4.05	4.20
110	2.98	3.29	1.53	1.47

Table 8: Upper bounds to robust put option prices using good-deal bounds with different values of the parameters λ_P and Λ, and depending on the initial value of the underlying risky asset's price x_0. Univariate case.

Upon comparing the results in Table 8 to the corresponding no-ambiguity results in the last column of Table 5 (with $\alpha = \infty$ hence $g \equiv 0$), we clearly observe that the good-deal bound evaluations of put options can be significantly lower than the corresponding no-ambiguity prices.

5.3 Optimal Entrance Problem

So far, we have considered examples of simple rewards for which $h \equiv 0$. Now we consider the optimal entrance problem, with $\Pi(t, X_t) = -\exp(-\rho t)\,\kappa$ and $h(t, X_t) = \exp(-\rho t)\,(X_t - \xi)$, in a univariate geometric Brownian motion setting (i.e., $J \equiv \lambda_P \equiv 0$). We define the grid of exercise dates by $t_j = j\Delta^c$, $j = 0, ..., T/\Delta^c$, with $1/\Delta^c$ the number of exercise dates in a year. For the fine grid, we take $\Delta_{jp} = \Delta^c/10$, and we vary Δ^c. We use the following parameter set under the reference model:

$$\mu = 0,\ \rho = 0.1,\ \sigma = 0.1,\ \xi = 1,\ \kappa = 1,\ T = 100 \text{ years}.$$

In Steps (1.a.) and (1.b.) of our algorithm, we choose for m_t^M the set of basis functions given by

$$\{1,\ \mathrm{Pol}_3\left(X_t\right),\ \mathrm{Pol}_3\left(h\left(t, X_t\right)\right)\}.$$

The basis functions for the Brownian motion driven part of the BSDE are given by the set

$$\left\{1,\ X_t \frac{\partial h\left(t, X_t\right)}{\partial X_t}\right\}.$$

In Step (2.), we add, as usual, the martingale and maximum processes, included in the Markov process $\mathcal{X}$, to the set of basis functions. We generate 5,000 simulated trajectories in each step of our algorithm.

Standard conditional expectation: First, we consider the case of a standard conditional expectation. In Table 9, we present results for different values of Δ^c and x_0. The results in the second and third columns of Table 9 can be viewed as rough approximations to the continuous-time optimal entrance problem with infinite time horizon as considered, for example, in Dixit [35], where $1/\Delta^c = T = \infty$. The last column in Table 9 displays the corresponding values obtained by Dixit [35].

	$1/\Delta^c$		
x_0	1	10	$1/\Delta^c = T = \infty$ (Dixit)
1	0.79	0.77	0.56
1.375	3.22	3.01	2.75
1.5	4.47	4.26	4.00

Table 9: Upper bounds to robust expected rewards with different values of the number of exercise dates $1/\Delta^c$ and depending on the initial value of the underlying stream of cash flows x_0. Univariate case.

The initial value of 1.375 would be the entrance boundary given by Dixit [35], for the parameter set considered here.

Kullback-Leibler divergence: Next, we consider the Kullback-Leibler divergence for different values of α, taking $1/\Delta^c = 10$. The results are in Table 10.

	α				
x_0	10	100	10^4	10^6	∞
1	0.40	0.70	0.77	0.77	0.77
1.375	1.14	2.64	3.00	3.01	3.01
1.5	1.57	3.83	4.25	4.26	4.26

Table 10: Upper bounds to robustly evaluated rewards using the Kullback-Leibler divergence with different values of its parameter α and depending on the initial value of the underlying stream of cash flows x_0. Univariate case.

Of course, the last column, with $\alpha = \infty$ (or $g \equiv 0$), agrees with the third column in Table 9. Robustly evaluated rewards appear to be fairly sensitive to changes in α, at moderate levels of α, which is in agreement with our observations from Tables 1, 3 and 5.

Worst case with mean partially known: Finally, we consider the worst case with mean partially known example. We take $B^+ = 1{,}000$ and $B^- = -1{,}000$ such that the resulting driver is practically independent of these parameters. The results are in Table 11.

	μ^+	$\mu^+ = \mu^-$	μ^-		
	0.05	0	-0.01	-0.03	-0.05
x_0	$(\mu^- = 0)$			$(\mu^+ = 0)$	
1.375	2.99	3.01	2.03	0.88	0.49
1.5	4.25	4.26	3.15	1.41	0.70

Table 11: Upper bounds to robustly evaluated rewards under the worst case with mean partially known example with different values of the parameters μ^+ and μ^- and depending on the initial value of the underlying stream of cash flows x_0. Univariate case.

We observe from Table 11 that the robustly evaluated rewards are insensitive to changes in μ^+ for given μ^-, and are quite sensitive to changes in μ^- for given μ^+, a pattern consistent with Tables 2 and 4. Again, the case of $\mu^+ = \mu^-$ yields the case of a standard conditional expectation, and agrees with the last column of Table 10 (as well as the third column in Table 9). As explained in Section 5.1.1, the worst case with mean partially known driver coincides with the worst case with ball scenarios driver for certain parameter sets, in the absence of jumps.

In sum, whenever reference values can be obtained by methods that are currently available in the literature, our numerical results confirm that our algorithm has good convergence properties, yielding accurate results at the pre-limiting level. The numerical results also reveal the potentially relevant and significant impact of taking ambiguity into account when evaluating optimal stopping strategies.

6 Algorithm: Step-Wise Description

6.1 Step (1.): Duality Theory of the Second Kind

6.1.1 Step (1.a.): Construct an Approximation to U^h

Since the approximation scheme adopted in Step (1.a.) will also be used in the steps that follow, it will be useful to use slightly more general notation. Recall the n-dimensional adapted process $(X_t)_{t\in[0,T]}$, satisfying the strong Markov property. We start with a function $w:\mathbb{R}^n\to\mathbb{R}$ (such that $w(X_T)$ is square-integrable) and the function $g(t,z,\tilde{z})$. Define $\Delta_{jp}:=s_{j(p+1)}-s_{jp}$, $\Delta W_{jp}:=W_{s_{j(p+1)}}-W_{s_{jp}}$, $\Delta\tilde{N}_{jp}:=\tilde{N}_{s_{j(p+1)}}-\tilde{N}_{s_{jp}}$, and $|\pi|:=\max_{j,p}\Delta_{jp}$, $j=0,\dots,L$, $p=0,\dots,P$. We will approximate U^h in (3.3) with a process Y^π. We initialize $Y^\pi_T=y^\pi(X_T)=w(X_T)$ where (here in Step (1.a.)) $w(X_T)=h(T,X_T)$. We then do a backward recursion over the s_{jp}. Suppose we have an approximation $Y^\pi_{s_{j(p+1)}}$ and we want to compute $Y^\pi_{s_{jp}}$. Theorem 8 then yields:

$$Y^\pi_{s_{jp}}\approx Y^\pi_{s_{j(p+1)}}+g(s_{jp},Z^\pi_{s_{jp}},\tilde{Z}^\pi_{s_{jp}})\Delta_{jp}-Z^\pi_{s_{jp}}\Delta W_{jp}-\tilde{Z}^\pi_{s_{jp}}\Delta\tilde{N}_{jp}\text{ for all }j,p;$$

see (3.9). Taking conditional expectations,

$$Y^\pi_{s_{jp}}\approx \mathrm{E}_{jp}\left[Y^\pi_{s_{j(p+1)}}\right]+g(s_{jp},Z^\pi_{s_{jp}},\tilde{Z}^\pi_{s_{jp}})\Delta_{jp},\tag{6.1}$$

with $\mathrm{E}_{jp}[\cdot]=\mathrm{E}[\cdot|X_{s_{jp}}]$. We take

$$(Z^\pi_{s_{jp}},\tilde{Z}^\pi_{s_{jp}})=\mathrm{argmin}_{(Z,\tilde{Z})\in L^2_{d+k}(\sigma(X_{s_{jp}}))}\mathrm{E}_{jp}\left[\left(Y^\pi_{s_{j(p+1)}}-Z\Delta W_{jp}-\tilde{Z}\Delta\tilde{N}_{jp}\right)^2\right].$$

Suppose that, for all j,p, we have basis functions $(m_k(s_{jp},X_{s_{jp}}))_{k\in\mathbb{N}}$, $(\psi_k(s_{jp},X_{s_{jp}}))_{k\in\mathbb{N}}$ and $(\tilde{\psi}_k(s_{jp},X_{s_{jp}}))_{k\in\mathbb{N}}$ spanning the space $L^2(\sigma(X_{s_{jp}}))$, respectively. Since we can computationally deal only with finitely many basis functions let us fix an $M\in\mathbb{N}$. We write

$$m^M(s_{jp},X_{s_{jp}})=(m_1(s_{jp},X_{s_{jp}}),\dots,m_M(s_{jp},X_{s_{jp}}))^\intercal,$$

and define ψ^M and $\tilde{\psi}^M$ similarly. Furthermore, define by $P^{\pi,M}_{s_{jp}}(Y^{\pi,M}_{s_{j(p+1)}}):=\alpha^{\pi,M}_{s_{jp}}m^M(s_{jp},X_{s_{jp}})$, and

$$Z^{\pi,M}_{s_{jp}}(Y^{\pi,M}_{s_{j(p+1)}}):=\gamma^{\pi,M}_{s_{jp}}\psi^M(s_{jp},X_{s_{jp}})\Delta W_{jp},\quad \tilde{Z}^{\pi,M}_{s_{jp}}(Y^{\pi,M}_{s_{j(p+1)}}):=\tilde{\gamma}^{\pi,M}_{s_{jp}}\tilde{\psi}^M(s_{jp},X_{s_{jp}})\Delta\tilde{N}_{jp},$$

the orthogonal projections on the space spanned by $m^M(s_{jp},X_{s_{jp}})$, $\psi^M(s_{jp},X_{s_{jp}})\Delta W_{jp}$ and $\tilde{\psi}^M(s_{jp},X_{s_{jp}})\Delta\tilde{N}_{jp}$, respectively. (Here and in the remainder of this section, we understand vector multiplication as dot (scalar) product.) Note that

$$
\begin{aligned}
\alpha_{s_{jp}}^{\pi,M} &= (A_{s_{jp}}^{\pi,M})^{-1}\mathrm{E}_{jp}\left[Y_{s_{j(p+1)}}^{\pi,M} m^M(s_{jp}, X_{s_{jp}})\right], \\
\gamma_{s_{jp}}^{\pi,M} &= (\bar{A}_{s_{jp}}^{\pi,M})^{-1}\mathrm{E}_{jp}\left[Y_{s_{j(p+1)}}^{\pi,M} \psi^M(s_{jp}, X_{s_{jp}})\Delta W_{jp}\right], \qquad (6.2)\\
\tilde{\gamma}_{s_{jp}}^{\pi,M} &= (\tilde{A}_{s_{jp}}^{\pi,M})^{-1}\mathrm{E}_{jp}\left[Y_{s_{j(p+1)}}^{\pi,M} \tilde{\psi}^M(s_{jp}, X_{s_{jp}})\Delta \tilde{N}_{jp}\right], \qquad (6.3)
\end{aligned}
$$

with coefficients given by

$$
\begin{aligned}
(A_{s_{jp}}^{\pi,M})_{1\le k,l\le M} &= \mathrm{E}_{jp}\left[m_k^M(s_{jp}, X_{s_{jp}}) m_l^M(s_{jp}, X_{s_{jp}})\right], \\
(\bar{A}_{s_{jp}}^{\pi,M})_{1\le k,l\le M} &= \mathrm{E}_{jp}\left[\psi_k^M(s_{jp}, X_{s_{jp}}) \psi_l^M(s_{jp}, X_{s_{jp}})\right]\mathrm{E}\left[\Delta^2 W_{jp}\right], \qquad (6.4)\\
(\tilde{A}_{s_{jp}}^{\pi,M})_{1\le k,l\le M} &= \mathrm{E}_{jp}\left[\tilde{\psi}_k^M(s_{jp}, X_{s_{jp}}) \tilde{\psi}_l^M(s_{jp}, X_{s_{jp}})\right]\mathrm{E}\left[\Delta^2 \tilde{N}_{jp}\right]. \qquad (6.5)
\end{aligned}
$$

Here, we define the process $Y_T^{\pi,M}$ by setting $Y_T^{\pi,M} = w(X_T)$, and then recursively

$$
Y_{s_{jp}}^{\pi,M} = \alpha_{s_{jp}}^{\pi,M} m^M(s_{jp}, X_{s_{jp}}) + g(s_{jp}, \gamma_{s_{jp}}^{\pi,M}\psi^M(s_{jp}, X_{s_{jp}}), \tilde{\gamma}_{s_{jp}}^{\pi,M}\tilde{\psi}^M(s_{jp}, X_{s_{jp}}))\Delta_{jp}. \qquad (6.6)
$$

To compute the conditional expectations in (6.2)–(6.5) numerically, we simulate N_0 independent paths $(X_{s_{jp}}^n)_{s_{jp}}$, starting with X_T for $s_{jp} = T$. Then, for $n = 1, \ldots, N_0$, we define $y_T^{\pi,M,N_0}(x) := w(x)$, and

$$
\begin{aligned}
y_{s_{jp}}^{\pi,M,N_0}(x) :=& \alpha_{s_{jp}}^{\pi,M,N_0} m^M(s_{jp}, x) \\
&+ g(s_{jp}, \gamma_{s_{jp}}^{\pi,M,N_0}\psi^M(s_{jp}, x), \tilde{\gamma}_{s_{jp}}^{\pi,M,N_0}\tilde{\psi}^M(s_{jp}, x))\Delta_{jp}, \qquad (6.7)
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_{s_{jp}}^{\pi,M,N_0} &= (A_{s_{jp}}^{\pi,M,N_0})^{-1}\frac{1}{N_0}\sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0} m^M(s_{jp}, X_{s_{jp}}^n) \\
\gamma_{s_{jp}}^{\pi,M,N_0} &= (\bar{A}_{s_{jp}}^{\pi,M,N_0})^{-1}\frac{1}{N_0}\sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0} \psi^M(s_{jp}, X_{s_{jp}}^n)\Delta W_{jp}^n \\
\tilde{\gamma}_{s_{jp}}^{\pi,M,N_0} &= (\tilde{A}_{s_{jp}}^{\pi,M,N_0})^{-1}\frac{1}{N_0}\sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0} \tilde{\psi}^M(s_{jp}, X_{s_{jp}}^n)\Delta \tilde{N}_{jp}^n, \qquad (6.8)
\end{aligned}
$$

with coefficients given by

$$
\begin{aligned}
(A_{s_{jp}}^{\pi,M,N_0})_{1\le k,l\le M} &= \frac{1}{N_0}\sum_{n=1}^{N_0} m_k^M(s_{jp}, X_{s_{jp}}^n) m_l^M(s_{jp}, X_{s_{jp}}^n) \\
(\bar{A}_{s_{jp}}^{\pi,M,N_0})_{1\le k,l\le M} &= \frac{1}{N_0}\sum_{n=1}^{N_0} \psi_k^M(s_{jp}, X_{s_{jp}}^n) \psi_l^M(s_{jp}, X_{s_{jp}}^n)\Delta_{jp} \\
(\tilde{A}_{s_{jp}}^{\pi,M,N_0})_{1\le k,l\le M} &= \frac{1}{N_0}\sum_{n=1}^{N_0} \tilde{\psi}_k^M(s_{jp}, X_{s_{jp}}^n) \tilde{\psi}_l^M(s_{jp}, X_{s_{jp}}^n)\lambda_P\Delta_{jp}. \qquad (6.9)
\end{aligned}
$$

We stop if $s_{jp} = 0$.

Finally, we define $u_{s_{jp}}^{h,\pi,M,N_0}(x) := y_{s_{jp}}^{h,\pi,M,N_0}(x)$, $z_{s_{jp}}^{h,\pi,M,N_0}(x) := \gamma_{s_{jp}}^{h,\pi,M,N_0}\psi^M(s_{jp}, x)$ and, similarly, $\tilde{z}_{s_{jp}}^{h,\pi,M,N_0}(x) := \tilde{\gamma}_{s_{jp}}^{h,\pi,M,N_0}\tilde{\psi}^M(s_{jp}, x)$.

6.1.2 Step (1.b.): Construct an Approximation to V^*

To do a backward recursion over t_j, we initialize $t_j = T$ and $V_T^{*,\pi} = v^{*,\pi}(T, X_T) := \Pi(T, X_T)$. Assuming that we are given an approximation $V_{t_{j+1}}^{*,\pi,M,N_1} = v^{*,\pi,M,N_1}(t_{j+1}, X_{t_{j+1}})$, we carry out the following loop: For $p = P$, we initialize $U_{s_{jP}}^\pi := U_{t_{j+1}}(V_{t_{j+1}}^{*,\pi,M,N_1}) = V_{t_{j+1}}^{*,\pi}$. Now, given $U_{s_{j(p+1)}}^\pi$, $U_{s_{jp}}^\pi$ we know from (3.10) that

$$U_{s_{jp}}^\pi \approx U_{s_{j(p+1)}}^\pi + g(s_{jp}, Z_{s_{jp}}, \tilde{Z}_{s_{jp}})(s_{j(p+1)} - s_{jp}) - Z_{s_{jp}}\Delta W_{j(p+1)} - \tilde{Z}_{t_j}\Delta \tilde{N}_{j(p+1)}. \tag{6.10}$$

Therefore, using N_1 simulations we can construct as before the vectors $(u_{s_{jp}}^{\pi,M,N_1})_p, (\alpha_{s_{jp}}^{\pi,M,N_1})_p$, $(\gamma_{s_{jp}}^{\pi,M,N_1})_p$, and $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_1})_p$ (with $T = t_{j+1}$, $t_0 = t_j$, and $w(\cdot) = v^{*,\pi}(t_{j+1}, \cdot)$ as terminal condition). This yields functions u^{π,M,N_1}, z^{π,M,N_1} and $\tilde{z}^{\pi,M,N_1}$. Finally, when we have arrived at $p = 0$, we set $j = j - 1$ and by (3.2) we define

$$v^{*,\pi,M,N_1}(t_j, x) := \max(\Pi(t_j, x) + u_{t_j}^{h,\pi,M,N_1}(x), u_{s_{j0}}^{\pi,M,N_1}(x)).$$

We stop if $j = 0$.

6.1.3 Step (1.c.): Construct an Approximation to M^{*g}

We then obtain a martingale $M_{s_{ip}}^{g,\pi,M,N_1}$ by defining

$$\begin{aligned} M_{s_{ip}}^{g,\pi,M,N_1} :=& -\sum_{j=0}^{i}\sum_{l=0}^{p-1}\int_{s_{jl}}^{s_{j(l+1)}} g(u, z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}), \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}))du \\ &+ \sum_{j=0}^{i}\sum_{l=0}^{p-1} z_{s_{lp}}^{\pi,M,N_1}(X_{s_{jl}})\Delta W_{jl} + \sum_{j=0}^{i}\sum_{l=0}^{p-1} \tilde{z}_{s_{lp}}^{\pi,M,N_1}(X_{s_{jl}})\Delta \tilde{N}_{jl}, \end{aligned} \tag{6.11}$$

see (3.11). Given i.i.d. simulations X^n we can then simulate i.i.d. copies of $M_{s_{ip}}^{g,\pi}$ through

$$\begin{aligned} M_{s_{ip}}^{g,\pi,M,N_1,n} :=& -\sum_{j=0}^{i}\sum_{l=0}^{p-1}\int_{s_{jl}}^{s_{j(l+1)}} g(u, z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n), \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n))du \\ &+ \sum_{j=0}^{i}\sum_{l=0}^{p-1} z_{s_{lp}}^{\pi,M,N_1}(X_{s_{jl}}^n)\Delta W_{jl}^n + \sum_{j=0}^{i}\sum_{l=0}^{p-1} \tilde{z}_{s_{lp}}^{\pi,M,N_1}(X_{s_{jl}}^n)\Delta \tilde{N}_{jl}^n. \end{aligned} \tag{6.12}$$

Note that (6.11) defines a true discrete-time U-martingale $(M_{t_j}^{g,\pi,M,N_1})_{j\in\{0,1,2,\ldots,L\}}$, and that (6.12) gives rise to an exact simulation scheme of it. The simulations $(M^{g,\pi,M,N_1,n})_{t_j}$ will be employed to establish a dual upper bound to the Snell envelope and the simulations $(M^{g,\pi,M,N_1,n})_{s_{jp}}$ (living on the finer grid π) will be needed for the numerical approximation.

6.2 Step (2.): Duality Theory of the First Kind and an Approximated Upper Bound to V^*

Eventually (in Step (3.) below) we will find a genuine (biased high) upper bound for V_0^* according to Proposition 6. To this end, we are faced with the computation of

$$\begin{aligned} V_0^* &= \inf_{M \in \mathcal{M}_0^U} U_0 \left(\max_{t_j \in \{0,t_1,\ldots,T\}} \left(\Pi(t_j, X_{t_j}) + U_{t_j}^h + M_T - M_{t_j} \right) \right) \\ &= U_0 \left(\max_{t_j \in \{0,t_1,\ldots,T\}} \left(\Pi(t_j, X_{t_j}) + U_{t_j}^h + M_T^{*g} - M_{t_j}^{*g} \right) \right). \end{aligned} \tag{6.13}$$

We set

$$F := \max_{t_j} \left(\Pi(t_j, X_{t_j}) + U_{t_j}^h + M_T^{*g} - M_{t_j}^{*g} \right).$$

Since we can only compute an approximation to M^{*g}, we cannot attain the infimum in (6.13). However, M_T^{g,π,M,N_1} obtained in the previous Step (1.c.) is a true U-martingale, which can be used to obtain an approximation to an upper bound. Let us first define, with $N_0 = N_1$,

$$F^{\pi,M,N_1} := \max_{t_j \in \{0,t_1,\ldots,T\}} \left(\Pi(t_j, X_{t_j}) + U_{t_j}^{h,\pi,M,N_1} + M_T^{g,\pi,M,N_1} - M_{t_j}^{g,\pi,M,N_1} \right).$$

Next, define the $(n+2)$-dimensional Markov process

$$\mathcal{X}_s^{\pi,M,N_1} := \left(X_{s_{jp}}, M_{s_{jp}}^{g,\pi,M,N_1}, \max_{t_l \in \{0,t_1,\ldots,t_j\}} \left(\Pi(t_l, X_{t_l}) + U_{t_l}^{h,\pi,M,N_1} - M_{t_l}^{g,\pi,M,N_1} \right) \right)$$

for $s_{jp} \le s < s_{j(p+1)}$. Let us compute $U_0(F^{\pi,M,N_1})$ numerically. Recall that for a payoff H, by Theorem 8(a),

$$\begin{aligned} U_t(H) &= \inf_{Q \sim P} \left\{ \mathrm{E}_Q \left[H + \int_t^T r(s, q_s, \lambda_s - \lambda_P) ds) \big| \mathcal{F}_t \right] \right\} && (6.14) \\ &= H + \int_t^T g(s, Z_s, \tilde{Z}_s) ds - \int_t^T Z_s dW_s - \int_t^T \tilde{Z}_s d\tilde{N}_s. && (6.15) \end{aligned}$$

Hence, we can apply the approximation scheme (6.7)–(6.9) (with $X = \mathcal{X}$ and terminal condition $\max_{t_l \in \{0,t_1,\ldots,t_j\}} \left(\Pi(t_l, X_{t_l}) + U_{t_l}^{h,\pi,M,N_1} - M_{t_l}^{g,\pi,M,N_1} \right)$). Simulate $n = 1, \ldots, N_2$ paths

$$(\mathcal{X}_{s_{jp}}^{\pi,M,N_1,n})_j = \left(X_{s_{jp}}^{\pi,n}, M_{s_{jp}}^{g,\pi,M,N_1,n}, \max_{t_l \in \{0,t_1,\ldots t_j\}} \left(\Pi(t_l, X_{t_l}^n) + U_{t_l}^{h,\pi,M,N_1} - M_{t_l}^{g,\pi,M,N_1,n} \right) \right).$$

Let M be the number of basis functions in the least squares Monte Carlo regression. We then obtain coefficients, say α_j^{π,M,N_2}, γ_j^{π,M,N_2}, $\tilde{\gamma}_j^{\pi,M,N_2}$, and processes $(V_t^{\pi,M,N_2}, Z_t^{\pi,M,N_2}, \tilde{Z}_t^{\pi,M,N_2})_{0 \le t \le T}$. Then, by applying Theorem 16 in the Appendix three times, we may conclude that

$$\lim_{\pi \to 0} \lim_{M \to \infty} \lim_{N_2, N_1, N_0 \to \infty} (V^{\pi,M,N_2}, Z^{\pi,M,N_2}, \tilde{Z}^{\pi,M,N_2}) = (V^*, Z^*, \tilde{Z}^*); \tag{6.16}$$

see the technical details provided in the Appendix. In particular, $V_0^{\pi,M,N_2} \to V_0$ as the mesh ratio of the grid, π, tends to zero, and the number of Monte Carlo simulations and basis functions tend to infinity. Thus, our algorithm will converge to the *true* value of the (U-)Snell envelope V^*.

However, at the pre-limiting level, our estimates from Step (2.) for the upper bound to V^* are not biased high (above the Snell envelope), meaning that in the average the upper bound may not provide enough protection. For this reason we will subsequently proceed to construct a genuine (biased high) upper bound.

6.3 Step (3.): Backward-Forward Simulation

6.3.1 Step (3.a.): Construct a Genuine Upper Bound to U^h

By Theorem 8(a), for $i = 0, \ldots, L-1$,

$$U^h_{t_i} = \inf_{Q\sim P}\left\{\mathrm{E}_Q\left[\sum_{j=i}^{L} h(t_j, X_{t_j}) + \int_{t_i}^{T} r(s, q_s, \lambda_s - \lambda_P)ds\big|\mathcal{F}_{t_i}\right]\right\}, \tag{6.17}$$

$$= U^h_{t_{i+1}} + \int_{t_i}^{t_{i+1}} g(s, Z^h_s, \tilde{Z}^h_s)ds - \int_{t_i}^{t_{i+1}} Z^h_s dW_s - \int_{t_i}^{t_{i+1}} \tilde{Z}^h_s d\tilde{N}_s + h(t_i, X_{t_i}). \tag{6.18}$$

Denote the Q that attains the infimum in (6.17) by Q^h.

The following proposition provides a way to practically obtain the extremal Q^h (leading in the end to an upper bound) by computing $(Z^h, \tilde{Z}^h)$ in (6.18).

Proposition 13 *For $H \in L^2(P)$, the infimum in (6.17) is attained at*

$$\frac{dQ^h}{dP} = \exp\left\{\int_0^t q^*_s dW_s + \int_0^t \log\left(\frac{\lambda^*_s}{\lambda_P}\right) dN_s - \int_0^T \left(\frac{|q^*_s|^2}{2} + \lambda^*_s - \lambda_P\right) ds\right\},$$

for every $(q^*_s, \lambda^*_s - \lambda_P) \in \partial g(s, Z^h_s, \tilde{Z}^h_s)$, where $\partial g(s, \cdot, \cdot)$ stands for the subdifferentials of the convex function $g(s, \cdot, \cdot)$.[8]

We then compute a genuine upper bound to $(U^h_{t_j})_{t_j\in\{0,\ldots,t_{L-1}\}}$ by:

(i) Computing approximations to $(Z, \tilde{Z})$ by solving (6.18). In view of Proposition 13, $(Z, \tilde{Z})$ induces an approximation to Q^h, say $Q^{h,\mathrm{approx}}$.

(ii) Evaluating $\mathrm{E}_{Q^{h,\mathrm{approx}}}\big[\sum_{j=i}^{L} h(t_j, X_{t_j}) + \int_{t_i}^{T} r(s, q_s, \lambda_s - \lambda_P)ds|X_{t_i}\big]$ and making use of (6.17). This will deliver the desired genuine (biased high) upper bound to $(U^h_{t_j})_{t_j\in\{0,\ldots,t_{L-1}\}}$.

So let us carry out our program to compute approximations $U^{h,n}_{t_j} = u^{h,n}_{t_j}(X^n_{t_j})$, for $n = 1, \ldots, N_3$: Simulate N_3 copies of $(X^n_{s_{jp}})$ ('outer simulation'). For $X^n_{t_j} = x$, let $N_4 \in \mathbb{N}$ and simulate additional paths $(X^{t_j,x,n}_{s_{jp}})$ for $n = 1, \ldots, N_4$ and j, p ('inner simulation'). For simplicity, assume that $g(s, \cdot, \cdot)$ is continuously differentiable. (If this is not the case, then our algorithm may still be implemented by taking elements in the subgradient.) Define, with $N_0 = N_1$, $z^{h,\pi,M,N_1}_{s_{jp}}(\bar{x}) := \gamma^{h,\pi,M,N_1}_{s_{jp}}\psi^M(s_{jp}, \bar{x})$, $\tilde{z}^{h,\pi,M,N_1}_{s_{jp}}(\bar{x}) := \tilde{\gamma}^{\pi,M,N_1}_{s_{jp}}\tilde{\psi}^M(s_{jp}, \bar{x})$ and

$$q^{h,\pi,t_j,x,n}_{s_{jp}} := g_z(s_{jp}, z^{h,\pi,M,N_1}_{s_{jp}}(X^{t_j,x,n}_{s_{jp}}), \tilde{z}^{h,\pi,M,N_1}_{s_{jp}}(X^{t_j,x,n}_{s_{jp}}))$$

$$\lambda^{h,\pi,t_j,x,n}_{s_{jp}} - \lambda_P := g_{\tilde{z}}(s_{jp}, z^{h,\pi,M,N_1}_{s_{jp}}(X^{t_j,x,n}_{s_{jp}}), \tilde{z}^{h,\pi,M,N_1}_{s_{jp}}(X^{t_j,x,n}_{s_{jp}})).$$

[8]Formally, $\partial f(x)$ of a convex function is given by the set of all slopes of all tangents at $f(x)$. Of course, in the one-dimensional case, $\partial f(x) = [f_-(x), f'_+(x)]$. Furthermore, $\partial f(x) = \{f'(x)\}$ if f is differentiable.

Next, for each $X_{t_j}^n = x$, define i.i.d. simulations of the measure $\frac{dQ^{g,\pi,M,N_4,t_j,x}}{dP}$ via the Radon-Nikodym derivative

$$D_{t_i}^{\pi,n}(x) := \exp\Bigg(\sum_{t_i \le s_{jp}} q_{s_{jp}}^{\pi,t_j,x,n} \Delta W_{jp}^n + \sum_{t_i \le s_{jp}} \log\left(\frac{\lambda_{s_{jp}}^{\pi,t_j,x,n}}{\lambda_P} \right) \Delta N_{jp}^n \\ - \sum_{t_i \le s_{jp}} \Big(\frac{1}{2} |q_{s_{jp}}^{\pi,t_j,x,n}|^2 + \lambda_{s_{jp}}^{\pi,t_j,x,n} - \lambda_P \Big) \Delta_{jp} \Bigg);$$

for $i = 1, \ldots, L$, see also (2.5). We then set

$$\widetilde{u}_{t_j}^{upp,h,N_4}(x) := \frac{1}{N_4} \sum_{n=1}^{N_4} D_{t_j}^{\pi,n}(x) \Bigg[\sum_{l=j}^{L} h(t_l, X_{t_l}^{t_j,x,n}) \\ + \sum_{l=1}^{L} \sum_{p=1}^{P} \int_{s_{lp}}^{s_{l(p+1)}} r(s, q_{s_{lp}}^{h,\pi,t_l,x,n}, \lambda_{s_{lp}}^{h,\pi,t_l,x,n} - \lambda_P) ds \Bigg]. \tag{6.19}$$

Now $(D_{t_j}^{\pi,n}(X_{t_j}^n))$, $(q_{s_{jp}}^{h,\pi,M,N_4,n})_{j,p}$ and $(\lambda_{s_{jp}}^{h,\pi,M,N_4,n})_{j,p}$ are true i.i.d. simulations of $\frac{dQ^{h,\pi,M,N_4}}{dP}$, the piecewise constant (q_t) and (λ_t), conditioned on $X_{t_j} = x$. Therefore, by (6.15), $\widetilde{u}_t^{upp,h,N_4}(X_t^n)$ can be taken as approximative simulations of U_t^h, yielding a genuine (biased high) upper bound to $U_t^h = u_t^h(X_t)$. Summarizing this step, we obtain the following proposition.

Proposition 14 *We have* $\mathrm{E}\left[\widetilde{u}_t^{upp,h,\pi}(x)\right] \ge u_t^h(x)$, *for any* x.

6.3.2 Step (3.b.): Construct a Genuine Upper Bound to V_0^*

In this final step, we proceed as in Step (3.a.) above, but this time we only need to compute an upper bound at time $t = 0$: Denote the Q that attains the infimum in (6.14) by Q^g, with corresponding $(q_s^*, \lambda_s^* - \lambda_P)$. As in Proposition 13 one may see that $(q_s^*, \lambda_s^* - \lambda_P) \in \partial g(s, Z_s, \tilde{Z}_s)$ with $(Z, \tilde{Z})$ from (6.15). We shall exploit this to practically compute our approximation. Let $N_3 \in \mathbb{N}$ and simulate paths $(W_{s_{jp}}^n)$ and $(X_{s_{jp}}^n)$ for $n = 1, \ldots, N_3$ and j, p. Define

$$U_{t_j}^{upper,h,\pi,n} := \tilde{u}_{t_j}^{upp,h,\pi}(X_{t_j}^n),$$

$$q_{s_{jp}}^{\pi,M,N_2,n} := g_z(s_{jp}, z_{s_{jp}}^{\pi,M,N_2,n}(X^n), \tilde{z}_{s_{jp}}^{\pi,M,N_2,n}(X^n))$$
$$\lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P := g_{\tilde{z}}(s_{jp}, z_{s_{jp}}^{\pi,M,N_2,n}(X^n), \tilde{z}_{s_{jp}}^{\pi,M,N_2,n}(X^n)).$$

Next, define i.i.d. simulations $\frac{dQ^{g,\pi,M,N_3,n}}{dP}$ via

$$\frac{dQ^{g,\pi,M,N_3,n}}{dP} := \exp\Bigg(\sum_{j,p} q_{s_{jp}}^{\pi,M,N_2,n} \Delta W_{jp}^n + \sum_{j,p} \log\left(\frac{\lambda_{s_{jp}}^{\pi,M,N_2,n}}{\lambda_P} \right) \Delta N_{jp}^n \\ - \sum_{j,p} \Big(\frac{1}{2} |q_{s_{jp}}^{\pi,M,N_2,n}|^2 + \lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P \Big) \Delta_{jp} \Bigg),$$

Finally, we set

$$
\begin{aligned}
\widetilde{V}_0^{upp,N_3} := & \frac{1}{N_3}\sum_{n=1}^{N_3}\frac{dQ^{g,\pi,M,N_3,n}}{dP}\Bigg[\\
& \max_{t_j\in\{0,\dots,T\}}\Big(\Pi(t_j,X^n_{t_j})+U^{upper,h,\pi,n}_{t_j}+M^{g,\pi,M,N_1,n}_T-M^{g,\pi,M,N_1,n}_{t_j}\Big) \\
& \qquad +\sum_{j,p}\int_{s_{jp}}^{s_{j(p+1)}} r(s,q^{\pi,M,N_2,n}_{s_{jp}},\lambda^{\pi,M,N_2,n}_{s_{jp}}-\lambda_P)ds\Bigg],
\end{aligned} \tag{6.20}
$$

where $M^{g,\pi,M,N_1,n}_{t_j}$ should be simulated using α^{π,M,N_1}, γ^{π,M,N_1} and $\tilde{\gamma}^{\pi,M,N_1}$ estimated previously (under Step (1.)).

6.4 Summary and Main Result

Let us summarize our algorithm more succinctly. Given a fixed time grid π and M basis functions:

(1.) Run N_0 Monte Carlo simulations to compute U^{h,π,M,N_0}. Run N_1 Monte Carlo simulations to compute M^{g,π,M,N_1}. To fully describe the evolution of these processes, it is sufficient to store the corresponding $(\alpha^{h,\pi,M,N_0}_{s_{jp}})$, $(\gamma^{h,\pi,M,N_0}_{s_{jp}})$, $(\tilde{\gamma}^{h,\pi,M,N_0}_{s_{jp}})$; and $(\alpha^{\pi,M,N_1}_{s_{jp}})$, $(\gamma^{\pi,M,N_1}_{s_{jp}})$ and $(\tilde{\gamma}^{\pi,M,N_1}_{s_{jp}})$.

(2.) With $N_0=N_1$, $(\alpha^{h,\pi,M,N_1}_{s_{jp}})$, $(\gamma^{h,\pi,M,N_1}_{s_{jp}})$, $(\tilde{\gamma}^{h,\pi,M,N_1}_{s_{jp}})$ and $(\alpha^{\pi,M,N_1}_{s_{jp}})$, $(\gamma^{\pi,M,N_1}_{s_{jp}})$, $(\tilde{\gamma}^{\pi,M,N_1}_{s_{jp}})$ give rise to a terminal condition F^{π,M,N_1} and a Markov process $\mathcal{X}^{\pi,M,N_1}$ defined under Step (2.). Run N_2 Monte Carlo simulations to calculate $(V^{\pi,M,N_2},Z^{\pi,M,N_2},\tilde{Z}^{\pi,M,N_2})$ as the solution to corresponding BSΔEs with the Markov process $\mathcal{X}^{\pi,M,N_1}$ and terminal condition F^{π,M,N_1}. Store the corresponding $(\gamma^{\pi,M,N_2}_{s_{jp}})$ and $(\tilde{\gamma}^{\pi,M,N_2}_{s_{jp}})$.

(3.a.) Simulate N_3 (outer simulation) copies of $(X^n_{s_{jp}})$. Simulate, for every n,j,p, N_4 additional (inner simulation) copies of $(X^n_{s_{jp}})$, to eventually compute, with $(\gamma^{h,\pi,M,N_1}_{s_{jp}})$ and $(\tilde{\gamma}^{h,\pi,M,N_1}_{s_{jp}})$ at hand from the previous Step (1.), N_3 copies of $U^{upper,h,n}$.

(3.b.) With $(\gamma^{\pi,M,N_2}_{s_{jp}})$ and $(\tilde{\gamma}^{\pi,M,N_2}_{s_{jp}})$ at hand from the previous Step (2.), simulate N_3 copies of $\frac{dQ^{g,\pi}}{dP}$. Furthermore, with $(\alpha^{\pi,M,N_1}_{s_{jp}})$, $(\gamma^{\pi,M,N_1}_{s_{jp}})$ and $(\tilde{\gamma}^{\pi,M,N_1}_{s_{jp}})$ at hand from the previous Step (1.), simulate N_3 copies of F^{π,M,N_1}. Using (6.20), a genuine (biased high) estimate for V^* can then be obtained.

The total computation time is determined by $MLP(N_0+N_1+N_2+N_3(1+LPN_4))$. In case the function h is identical zero so that the optimal stopping problem is a (simple) reward problem, the inner simulation is not needed and N_0 and N_4 may be set equal to zero.

Our main result, then, reads as follows:

Theorem 15 *The primal estimator V^{*,π,M,N_1}_0 and both the dual estimators V^{π,M,N_2}_0 and $\widetilde{V}^{upp,N_3}_0$ converge to the upper Snell envelope, i.e.,*

$$
\begin{aligned}
\lim_{\pi\to 0}\lim_{M\to\infty}\lim_{N_i\to\infty,i=0,1} V^{*,\pi,M,N_1}_0 &= \lim_{\pi\to 0}\lim_{M\to\infty}\lim_{N_i\to\infty,i=0,1,2} V^{\pi,M,N_2}_0 \\
&= \lim_{\pi\to 0}\lim_{M\to\infty}\lim_{N_i\to\infty,i=0,\dots,4} \widetilde{V}^{upp,N_3}_0 = V^*_0.
\end{aligned}
$$

Furthermore, with $(\alpha_{s_{jp}}^{\pi,M,N_1})$, $(\gamma_{s_{jp}}^{\pi,M,N_1})$, $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_1})$, $(\gamma_{s_{jp}}^{\pi,M,N_2})$ *and* $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_2})$ *fixed from the preceding Steps (1.) and (2.), our estimator in Step (3.) gives rise to a genuine (biased high) upper bound, i.e.,* $\mathrm{E}\left[\widetilde{V}_0^{upp,N_3}\right] \geq V_0^*$.

7 Conclusion

We have developed a method to practically compute the solution to the optimal stopping problem in a general continuous-time setting featuring general time-consistent ambiguity averse preferences and general rewards driven by jump-diffusions. The resulting algorithm delivers an approximation to the solution that converges asymptotically to the true solution and yields a safe genuine (biased high) upper bound at the pre-limiting level. Our method is widely applicable, numerically efficient, and eventually requires only simple least squares Monte Carlo regression techniques. Our method may be generalized to encompass multiple stopping problems, which we intend to consider in future research.

A Appendix: Proofs

Proof of Eqns. (3.1)–(3.2) and Proposition 6: By time-consistency of U, a property that is preserved with respect to stopping times, i.e., for any stopping time τ with $0 \leq t \leq \tau \leq T$ (by backward induction),

$$U_t = U_t \circ U_\tau,$$

we have

$$\sup_{\tau \in \mathcal{T}} U_0(H_\tau) = \sup_{\tau \in \mathcal{T}} U_0(U_\tau(H_\tau)) = \sup_{\tau \in \mathcal{T}} U_0(\tilde{H}_\tau),$$

where $\tilde{H}_t := U_t(H_t)$ for $t \in [0, T]$. Hence, the optimal stopping problem (2.4) with non-adapted rewards $(H_t)_{t \in \mathcal{T}}$ can be transformed into an (equivalent) optimal stopping problem with adapted rewards $(\tilde{H}_t)_{t \in \mathcal{T}}$. Therefore, the existence of an optimal stopping time in (3.1) follows, upon continuous embedding, as a consequence of Theorem 3.2 in Krätschmer and Schoenmakers [57]. Furthermore, upon continuous embedding, (3.2) follows as a consequence of Theorem 3.4 in [57] and Proposition 6 is a consequence of Theorem 5.4 in the same [57]. ■

Proof of Theorem 8: For a square-integrable H that is $\mathcal{F}_T$-adapted and $t \in [0, T]$, let us consider

$$U_t^h = \inf_{(q,\lambda) \in C} \left\{ \mathrm{E}_Q \left[H + \sum_{t \leq t_j} h(t_j, X_{t_j}) | \mathcal{F}_t \right] + c_t(Q) \right\}, \tag{A.1}$$

where (for the first part of the proof) $H = 0$. Of course, $U_{t_j}^h = U_{t_j+}^h + h(t_j, X_{t_j})$ and by time-consistency, for $t \in (t_j, t_{j+1}]$,

$$U_t^h = \inf_{(q,\lambda) \in C} \left\{ \mathrm{E}_Q \left[U_{t_{j+1}}^h | \mathcal{F}_t \right] + c_t(Q) \right\}. \tag{A.2}$$

The first part of (a) would follow if we could show that there exists a predictable, square-integrable $(Z, \tilde{Z})$ such that

$$dU_t^h = -g(t, Z_t, \tilde{Z}_t)dt + Z_t dW_t + \tilde{Z}_t d\tilde{N}_t, \quad \text{for } t \in (t_j, t_{j+1}], \tag{A.3}$$

with $j = 0, \dots, L-1$. Let $t \in (t_j, t_{j+1}]$. Notice that an adapted process, say Y, satisfying the RHS of (A.3) may be seen as a solution to a BSDE. To be more precise, by Tang and Li [81], there exists a unique triple of processes, say $(Y_t, Z_t, \tilde{Z}_t)_{t \in [t_j, t_{j+1}]} \in \mathcal{S}^2 \times L^2(dP \times ds) \times L^2(dP \times ds)$, satisfying

$$dY_t = -g(t, Z_t, \tilde{Z}_t)dt + Z_t dW_t + \tilde{Z}_t d\tilde{N}_t, \quad \text{and} \quad Y_{t_{j+1}} = U^h_{t_{j+1}},$$

where we denote by $\mathcal{S}^2$ the space of all processes for which the maximum is square-integrable. We need to show that $U^h_t = Y_t$ for $t \in (t_j, t_{j+1}]$. Let $Q \in \mathcal{Q}$. We write

$$\begin{aligned}
Y_t &= \mathrm{E}_Q[Y_t | \mathcal{F}_t] \\
&= \mathrm{E}_Q\Big[U^h_{t_{j+1}} + \int_t^{t_{j+1}} g(s, Z_s, \tilde{Z}_s)ds - \int_t^{t_{j+1}} Z_s dW_s - \int_t^{t_{j+1}} \tilde{Z}_s d\tilde{N}_s \mid \mathcal{F}_t\Big] \\
&= \mathrm{E}_Q\Big[U^h_{t_{j+1}} + \int_t^{t_{j+1}} \Big[-q_s Z_s - \tilde{Z}_s(\lambda_s - \lambda_P) + g(s, Z_s, \tilde{Z}_s)\Big]ds \\
&\qquad + \int_t^{t_{j+1}} Z_s dW^Q_s + \int_t^{t_{j+1}} \tilde{Z}_s d\tilde{N}^Q_s \mid \mathcal{F}_t\Big] \\
&= \mathrm{E}_Q\Big[U^h_{t_{j+1}} + \int_t^{t_{j+1}} \Big[-q_s Z_s - \tilde{Z}_s(\lambda_s - \lambda_P) + g(s, Z_s, \tilde{Z}_s)\Big]ds \mid \mathcal{F}_t\Big] \\
&\le \mathrm{E}_Q\Big[U^h_{t_{j+1}} + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P)ds \mid \mathcal{F}_t\Big],
\end{aligned} \tag{A.4}$$

where we used in the first equality that Y_t is $\mathcal{F}_t$-measurable. Note that the conditional expectation in the first equality is well-defined by the inequality of Cauchy-Schwarz, as (q, λ) take values in a compact set and Y is square-integrable under P. The third and fourth equalities hold because $\int_{t_j}^t Z_s dW^Q_s$ and $\int_{t_j}^t \tilde{Z}_s \tilde{N}^Q_s$ are well-defined martingales, since for any Q with (q, λ) in a compact bounded set we have, again by Cauchy-Schwarz,

$$\mathrm{E}_Q\left[\sqrt{\int_{t_j}^{t_{j+1}} |Z_s|^2 ds}\right] = \mathrm{E}\left[\frac{dQ}{dP}\sqrt{\int_{t_j}^{t_{j+1}} |Z_s|^2 ds}\right] \le \sqrt{\mathrm{E}\left[\left(\frac{dQ}{dP}\right)^2\right]}\sqrt{\mathrm{E}\left[\int_{t_j}^{t_{j+1}} |Z_s|^2 ds\right]} < \infty,$$

and a similar argument holds for $\tilde{Z}$. It follows from (A.4) and the fact that we can restrict the infimum in (A.2) to $Q \in C$ that $Y_t \le U^h_t$.

Next, by a measurable selection theorem (see e.g., Benes [8]), choose predictable $(q_s, \lambda_s - \lambda_P) \in \partial g(s, Z_s, \tilde{Z}_s)$. Then, q and λ induce an equivalent probability measure, Q^g, with Radon-Nikodym derivative given by (2.5). Proceeding as in (A.4) with q, λ and Q^g (where the inequality in (A.4) becomes an equality) yields

$$Y_t = \mathrm{E}_{Q^g}\Big[U^h_{t_{j+1}} + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P)ds \mid \mathcal{F}_t\Big]. \tag{A.5}$$

Thus, by the definition of U^h_t in (A.2), we get $Y_t \ge U^h_t$. Therefore, indeed $Y_t = U^h_t$ for all $t \in (t_j, t_{j+1}]$. This shows (3.9). (3.10) is seen similarly by setting $h = 0$ in (A.1). This proves part (a) of the theorem.

To see part (b), note that by part (a), there exist square-integrable $(Z^*, \tilde{Z}^*)$ such that (3.11) holds. Hence,

$$
\begin{aligned}
V^*_{t_{j+1}} - U_{t_j}\left(V^*_{t_{j+1}}\right) &= M^{*,g}_{t_{j+1}} - A^{*,g}_{t_{j+1}} - U_{t_j}(M^{*,g}_{t_{j+1}} + A^{*,g}_{t_{j+1}}) \\
&= M^{*,g}_{t_{j+1}} - M^{*,g}_{t_j} \\
&= \int_{t_j}^{t_{j+1}} Z^*_s dW_s + \int_{t_j}^{t_{j+1}} \tilde{Z}^*_s d\tilde{N}_s - \int_{t_j}^{t_{j+1}} g(s, Z^*_s, \tilde{Z}^*_s) ds.
\end{aligned}
$$

From (3.4), part (b) follows. ■

Proof of Eqn. (6.16): We now show that our approximation scheme converges. Suppose that equations (6.1)–(6.7) hold with a square-integrable p-dimensional Markov process, $\mathcal{X}$, and an arbitrary function (driver) $g : [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ that is uniformly Lipschitz continuous in $(z, \tilde{z})$. The following theorem establishes convergence of our approximation scheme:

Theorem 16 *We have that*

$$
\begin{aligned}
&\lim_{\pi\to 0} \lim_{M\to\infty} \lim_{N\to\infty} Y^{\pi,M,N}_{T_0} \to Y_{T_0} \ \textit{in } L^2, \\
&\lim_{\pi\to 0} \lim_{M\to\infty} \lim_{N\to\infty} Z^{\pi,M,N} \to Z \ \textit{in } L^2(dP \times ds, \Omega \times [0,T]), \\
&\lim_{\pi\to 0} \lim_{M\to\infty} \lim_{N\to\infty} \tilde{Z}^{\pi,M,N} \to \tilde{Z} \ \textit{in } L^2(dP \times ds, \Omega \times [0,T]).
\end{aligned}
$$

Proof It follows from Bouchard and Elie [15] that Y^π_t converges to Y_t in L^2. From this and Lemma 17 below we may conclude that it is sufficient to prove that $Y^{\pi,M,N}_{T_0}$ converges to $Y^{\pi,M}_{T_0}$ in L^2, which would follow if

$$
\lim_{N\to\infty} Y^{\pi,M,N}_{T_0} \to Y^{\pi,M}_{T_0} \text{ in } L^2.
$$

And this follows from Lemma 18 below. The proof for $Z^{\pi,M,N}$ and $\tilde{Z}^{\pi,M,N}$ is similar. ■

Lemma 17 *For every $t \in [T_0, T_1]$ and for fixed π, we have that $Y^{\pi,M}_t \to Y^\pi_t$, $Z^{\pi,M}_t \to Z^\pi_t$ and $\tilde{Z}^{\pi,M}_t \to \tilde{Z}^\pi_t$ in L^2 as M tends to infinity.*

Proof The lemma would follow if we could show by a backward induction that, for every s_{jp}, we have $Y^{\pi,M}_{s_{jp}} \to Y^\pi_{s_{jp}}$, $Z^{\pi,M}_{s_{jp}} \to Z^\pi_{s_{jp}}$ and $\tilde{Z}^{\pi,M}_{s_{jp}} \to \tilde{Z}^\pi_{s_{jp}}$ in L^2_1, L^2_d and L^2_k, respectively. Since our basis functions span the entire space, $L^2_1(\mathcal{F}_{s_{jp}})$, the lemma clearly holds for $s_{jp} = T$. (Without loss of generality we may set $Z^{\pi,M}_{T_1} = Z^\pi_{T_1}$ and $\tilde{Z}^{\pi,M}_{T_1} = \tilde{Z}^\pi_{T_1}$.) It will be useful to consider the projection onto the span of $\psi^M(s_{jp}, \mathcal{X}^\pi_{s_{jp}})$ and $\tilde{\psi}^M(s_{jp}, \mathcal{X}^\pi_{s_{jp}})$, say $\hat{P}^{\pi,M}$ and $\tilde{\hat{P}}^{\pi,M}$, respectively, instead of the projection onto the span of $\psi^M(s_{jp}, \mathcal{X}^\pi_{s_{jp}})\Delta W_{jp}$ and $\tilde{\psi}^M(s_{jp}, \mathcal{X}^\pi_{s_{jp}})\Delta\tilde{N}_{jp}$, respectively. We write

$$
\begin{aligned}
\gamma^{\pi,M}_{s_{jp}} \psi^M(s_{jp}, \mathcal{X}^\pi_{s_{jp}}) &= \hat{P}^{\pi,M}(Y^M_{s_{j(p+1)}} \Delta W_{jp}) / \mathrm{E}\left[\Delta^2 W_{jp}\right] \\
&= \hat{P}^{\pi,M}\Big(\mathrm{E}_{jp}[Y^{\pi,M}_{s_{j(p+1)}} \Delta W_{jp}]\Big) / \mathrm{E}\left[\Delta^2 W_{jp}\right] \\
&\overset{M\to\infty}{\to} \mathrm{E}_{jp}[Y^\pi_{s_{j(p+1)}} \Delta W_{jp}] / \mathrm{E}\left[\Delta^2 W_{jp}\right] = Z^\pi_{s_{jp}},
\end{aligned}
$$

in L^2, where we used (6.2) and (6.4) in the first equality. The convergence then follows since, by the induction assumption, we have that $\mathrm{E}_{jp}\left[Y^{\pi,M}_{s_{j(p+1)}}\Delta W_{jp}\right]$ converges in L^2 to $\mathrm{E}_{jp}\left[Y^{\pi}_{s_{j(p+1)}}\Delta W_{jp}\right]$ as M tends to infinity. Similarly,

$$\begin{aligned}
\tilde{\gamma}^{\pi,M}_{s_{jp}}\tilde{\psi}^M(s_{jp},\mathcal{X}^{\pi}_{s_{jp}}) &= \hat{\tilde{P}}^{\pi,M}(\mathrm{Y}^M_{s_{j(p+1)}}\Delta\tilde{N}_{jp})/\mathrm{E}\left[\Delta^2\tilde{N}_{jp}\right] \\
&= \hat{\tilde{P}}^{\pi,M}\Big(\mathrm{E}_{jp}[\mathrm{Y}^{\pi,M}_{s_{j(p+1)}}\Delta\tilde{N}_{jp}]\Big)/\mathrm{E}\left[\Delta^2\tilde{N}_{jp}\right] \\
&\overset{M\to\infty}{\to} \mathrm{E}_{jp}[Y^{\pi}_{s_{j(p+1)}}\Delta\tilde{N}_{jp}]/\mathrm{E}\left[\Delta^2\tilde{N}_{jp}\right] = \tilde{Z}^{\pi}_{s_{jp}},
\end{aligned}$$

in L^2, where we used (6.3) and (6.5) in the first equality. The lemma is now a consequence of (6.1) and (6.6). ■

Lemma 18 *For all j, we have that $\alpha^{\pi,M,N}_{s_{jp}} \to \alpha^{\pi,M}_{s_{jp}}$, $\gamma^{\pi,M,N}_{s_{jp}} \to \gamma^{\pi,M}_{s_{jp}}$ and $\tilde{\gamma}^{\pi,M,N}_{s_{jp}} \to \tilde{\gamma}^{\pi,M}_{s_{jp}}$ as N tends to infinity.*

Proof By the Law of Large Numbers (LLN), we have that $(A^{\pi,M,N}_{s_{jp}})$, $(\bar{A}^{\pi,M,N}_{s_{jp}})$ and $(\tilde{A}^{\pi,M,N}_{s_{jp}})$ converge to $(A^{\pi,M}_{s_{jp}})$, $(\bar{A}^{\pi,M}_{s_{jp}})$ and $(\tilde{A}^{\pi,M}_{s_{jp}})$, respectively. We prove the claim by a backward induction. For $\alpha,\gamma,\tilde{\gamma}\in\mathbb{R}^M$ and $x\in\mathbb{R}^d$ set

$$\begin{aligned}
&F(T_1,\alpha,\gamma,\tilde{\gamma},x) := w(x) \\
&F(s_{jp},\alpha,\gamma,\tilde{\gamma},x) := \alpha m^M(s_{jp},x) + g(s_{jp},\gamma\psi^M(s_{jp},x),\tilde{\gamma}\tilde{\psi}^M(s_{jp},x))\Delta_{jp} \quad \text{ for } s_{jp} < T_1.
\end{aligned}$$

Furthermore, for every j,p, $F(s_{jp},\cdot)$ is continuous in x and Lipschitz continuous in $(\alpha,\gamma,\tilde{\gamma})$. Moreover, by the LLN we have that

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} F(s_{j(p+1)},\alpha^{\pi,M}_{s_{j(p+1)}},\gamma^{\pi,M}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M}_{s_{j(p+1)}},\mathcal{X}^{\pi,n}_{s_{j(p+1)}})m^M(s_{jp},\mathcal{X}^{\pi,n}_{s_{jp}}) \\
&\qquad \overset{N\to\infty}{\to} \mathrm{E}[F(s_{j(p+1)},\alpha^{\pi,M}_{s_{j(p+1)}},\gamma^{\pi,M}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M}_{s_{j(p+1)}},\mathcal{X}^{\pi}_{s_{j(p+1)}})m^M(s_{jp},\mathcal{X}^{\pi}_{s_{jp}})].
\end{aligned}$$

Since, by Lipschitz continuity of g and the induction assumption, we have that

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N}(F(s_{j(p+1)},\alpha^{\pi,M,N}_{s_{j(p+1)}},\gamma^{\pi,M,N}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M,N}_{s_{j(p+1)}},\mathcal{X}^{\pi,n}_{s_{j(p+1)}}) \\
&\qquad - F(s_{j(p+1)},\alpha^{\pi,M}_{s_{j(p+1)}},\gamma^{\pi,M}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M}_{s_{j(p+1)}},\mathcal{X}^{\pi,n}_{s_{j(p+1)}}))m^M(s_{jp},\mathcal{X}^{\pi,n}_{s_{jp}}) \\
&\quad \le \Big(|\alpha^{\pi,M,N}_{s_{j(p+1)}}-\alpha^{\pi,M}_{s_{j(p+1)}}| + |\gamma^{\pi,M,N}_{s_{j(p+1)}}-\gamma^{\pi,M}_{s_{j(p+1)}}| + |\tilde{\gamma}^{\pi,M,N}_{s_{j(p+1)}}-\tilde{\gamma}^{\pi,M}_{s_{j(p+1)}}|\Big) \\
&\qquad \times \frac{1}{N}\sum_{n=1}^{N} m^M(s_{jp},\mathcal{X}^{\pi,n}_{s_{jp}}) \overset{N\to\infty}{\to} 0,
\end{aligned}$$

it follows that

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} F(s_{j(p+1)},\alpha^{\pi,M,N}_{s_{j(p+1)}},\gamma^{\pi,M,N}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M,N}_{s_{j(p+1)}},\mathcal{X}^{\pi,n}_{s_{j(p+1)}})m^M(s_{jp},\mathcal{X}^{\pi,n}_{s_{jp}}) \\
&\qquad \overset{N\to\infty}{\to} \mathrm{E}[F(s_{j(p+1)},\alpha^{\pi,M}_{s_{j(p+1)}},\gamma^{\pi,M}_{s_{j(p+1)}},\tilde{\gamma}^{\pi,M}_{s_{j(p+1)}},\mathcal{X}^{\pi}_{s_{j(p+1)}})]m^M(s_{jp},\mathcal{X}^{\pi}_{s_{jp}}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\alpha_{s_{jp}}^{\pi,M,N} &= (A_{s_{jp}}^{\pi,M,N})^{-1}\frac{1}{N}\sum_{n=1}^{N} Y_{s_{j(p+1)}}^{\pi,M,N} m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n}) \\
&= (A_{s_{jp}}^{\pi,M,N})^{-1}\frac{1}{N}\sum_{n=1}^{N} F(s_{j(p+1)},\alpha_{s_{j(p+1)}}^{\pi,M,N},\gamma_{s_{j(p+1)}}^{\pi,M,N},\tilde{\gamma}_{s_{j(p+1)}}^{\pi,M,N},\mathcal{X}_{s_{j(p+1)}}^{n,\pi}) m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n}) \\
&\to (A_{s_{jp}}^{\pi,M})^{-1}\mathrm{E}\left[F(s_{j(p+1)},\alpha_{s_{j(p+1)}}^{\pi,M},\gamma_{s_{j(p+1)}}^{\pi,M},\tilde{\gamma}_{s_{j(p+1)}}^{\pi,M},\mathcal{X}_{s_{j(p+1)}}^{\pi}) m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi})\right] \\
&= (A_{s_{jp}}^{\pi,M})^{-1}\mathrm{E}\left[Y_{s_{j(p+1)}}^{\pi,M} m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi})\right] = \alpha_{s_{jp}}^{\pi,M}.
\end{aligned}$$

By replacing $\alpha_{s_{jp}}^{\pi,M,N}$ by $\gamma_{s_{jp}}^{\pi,M,N}$, $A_j^{\pi,M}$ by $\bar{A}_{s_{jp}}^{\pi,M}$, and $m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n})$ by $\psi^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n})$, it follows similarly that $\gamma_{s_{jp}}^{\pi,M,N}$ converges to $\gamma_{s_{jp}}^{\pi,M}$. Also, by replacing $\alpha_{s_{jp}}^{\pi,M,N}$ by $\tilde{\gamma}_{s_{jp}}^{\pi,M,N}$, $\tilde{A}_{s_{jp}}^{\pi,M}$ by $\tilde{\bar{A}}_{s_{jp}}^{\pi,M}$, and $m^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n})$ by $\tilde{\psi}^M(s_{jp},\mathcal{X}_{s_{jp}}^{\pi,n})$, it follows similarly that $\tilde{\gamma}_{s_{jp}}^{\pi,M,N}$ converges to $\tilde{\gamma}_{s_{jp}}^{\pi,M}$. This proves the induction. ■

Then, applying Theorem 16 above three times completes the proof of (6.16). ■

Proof of Proposition 13: This follows from (A.5) in the proof of Theorem 8(a). ■

Proof of Theorem 15: The stated convergence results follow as a consequence of our convergence results for BSΔEs (see the *proof* of (6.16)). Next, choose a fixed $n \in \{1,\dots,N_3\}$. To show the biased high property, we then write

$$\begin{aligned}
&\mathrm{E}\left[\widetilde{V}_0^{upp,N_3}\right] \\
&= \mathrm{E}\Bigg[\frac{dQ^{g,\pi,M,N_2,n}}{dP}\mathrm{E}\Bigg[\max_{t_j\in\{0,\dots,T\}}\Big(\Pi(t_j,X_{t_j}^n)+U_{t_j}^{upper,h,\pi,M,N_3,n}+M_T^{g,\pi,M,N_1,n}-M_{t_j}^{g,\pi,M,N_1,n}\Big) \\
&\qquad\qquad +\sum_{j,p}\int_{s_{jp}}^{s_{j(p+1)}} r(s,q_{s_{jp}}^{\pi,M,N_2,n},\lambda_{s_{jp}}^{\pi,M,N_2,n}-\lambda_P)ds \\
&\qquad\qquad\qquad \Bigg| W_{s_{jp}}^{n'},N_{s_{jp}}^{n'},X_{s_{jp}}^{n'},p=1,\dots,P,j=1,\dots,L,n'=1,\dots,N_3\Bigg]\Bigg] \\
&\geq \mathrm{E}\Bigg[\frac{dQ^{g,\pi,M,N_2,n}}{dP}\Bigg[\max_{t_j\in\{0,\dots,T\}}\Big(\Pi(t_j,X_{t_j}^n)+u_{t_j}^h(X_{t_j}^n)+M_T^{g,\pi,M,N_1,n}-M_{t_j}^{g,\pi,M,N_1,n}\Big) \\
&\qquad\qquad +\sum_{j,p}\int_{s_{jp}}^{s_{j(p+1)}} r(s,q_{s_{jp}}^{\pi,M,N_2,n},\lambda_{s_{jp}}^{\pi,M,N_2,n}-\lambda_P)ds\Bigg]\Bigg] \\
&\geq U_0\Big(\max_{t_j\in\{0,\dots,T\}}\Big(\Pi(t_j,X_{t_j}^n)+u_{t_j}^h(X_{t_j}^n)+M_T^{g,\pi,M,N_1,n}-M_{t_j}^{g,\pi,M,N_1,n}\Big)\Big) \\
&\geq U_0\Big(\max_{t_j\in\{0,\dots,T\}}\Big(\Pi(t_j,X_{t_j}^n)+u_{t_j}^h(X_{t_j}^n)+M_T^{g*,n}-M_{t_j}^{g*,n}\Big)\Big) = V_0^*,
\end{aligned}$$

where we have used Proposition 14 and Jensen's inequality in the first inequality, (6.14) in the second inequality, and Proposition 6 in the last inequality and also in the last equality. ■

References

[1] Andersen, L., and Broadie, M. (2004). A primal-dual simulation algorithm for pricing multidimensional American options. *Management Science 50*, 1222–1234.

[2] Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

[3] Barone-Adesi, G., and Whaley, R. (1987). Efficient analytic approximation of American option values. *The Journal of Finance 42*, 301–320.

[4] Bayraktar, E., Karatzas, I., and Yao, S. (2010). Optimal stopping for dynamic convex risk measures. *Illinois Journal of Mathematics 54*, 1025–1067.

[5] Bayraktar, E., and Yao, S. (2011). Optimal stopping for non-linear expectations—Part I. *Stochastic Processes and their Applications 121*, 185–211.

[6] Belomestny, D., Bender, C. and Schoenmakers, J. (2009). True upper bounds for Bermudam products via non-nested Monte Carlo. *Mathematical Finance 19*, 53–71.

[7] Bender, C., Schweizer, N., and Zhuo, J. (2014). A primal-dual algorithm for BSDEs. Mimeo, Saarland University.

[8] Beněs, V. E. (1970). Existence of Optimal Strategies Based on Specified Information, for a Class of Stochastic Decision Problems. *SIAM Journal of Control 8*, 179–188.

[9] Ben-Tal, A. (1985). The entropic penalty approach to stochastic programming. *Mathematics of Operations Research 10*, 224–240.

[10] Ben-Tal, A., and Teboulle, M. (1987). Penalty functions and duality in stochastic programming via ϕ-divergence functionals. *Mathematics of Operations Research 12*, 224–240.

[11] Ben-Tal, A., and Teboulle, M. (2007). An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance 12*, 449–476.

[12] Ben-Tal, A., Bertsimas, D., and Brown, D. (2010). A soft robust model for optimization under ambiguity. *Operations Research 58*, 1220–1234.

[13] Beskos, A., and Roberts, G. (2005). Exact simulation of diffusions. *Annals of Applied Probability 15*, 2422–2444.

[14] Björk, T., and Slinko, I. (2006). Towards a general theory of good-deal bounds. *Review of Finance 10*, 221–260.

[15] Bouchard, B., and Elie, R. (2008). Discrete-time approximation of decoupled Forward-Backward SDE with jumps. *Stochastic Processes and their Applications 118*, 53–75.

[16] Brennan, M., and Schwartz, E. (1985). Evaluating natural resource investments. *Journal of Business 58*, 135–157.

[17] Broadie, M., and Glasserman, P. (2004), A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance 7*, 35–72.

[18] Broadie, M., and Kaya, O. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research 54*, 217–231.

[19] Carriere, J. (1996). Valuation of the early-exercise price for options using simulations and non-parametric regression. *Insurance: Mathematics and Economics 19*, 19–30.

[20] Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Montrucchio, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory 146*, 1275–1330.

[21] Chen, Z. and Epstein, L.G. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica 70*, 1403–1443.

[22] Chen, N., and Huang, Z. (2013). Localization and exact simulation of Brownian motion driven stochastic differential equations. *Mathematics of Operations Research 38*, 591–616.

[23] Cheng, X., and Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics 7*, 29–68.

[24] Cheridito, P., Delbaen, F., and Kupper, M. (2006). Coherent and convex monetary risk measures for unbounded càdlàg processes. *Finance and Stochastics 10*, 427–448.

[25] Choi, S., Ruszczyński, A., and Zhao, Y. (2011). A multiproduct risk-averse newsvendor with law-invariant coherent measures of risk. *Operations Research 59*, 346–364.

[26] Clément, E., Lamberton, D., and Protter, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance and Stochastics 6*, 449–471.

[27] Cochrane, J. and Saá-Requejo, J. (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy 108*, 79–119.

[28] Cont, R., and Tankov, R. (2004). *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, Boca Raton.

[29] Csiszár, I. (1975). *I*-divergence geometry of probability distributions and minimization problems. *Annals of Probability 3*, 146–158.

[30] Davis, M., and Karatzas, I. (1994). A deterministic approach to optimal stopping. Kelly, F.P. (ed.), Probability, statistics and optimisation. A tribute to Peter Whittle. Wiley, Chichester. Wiley Series in Probability and Mathematical Statistics. 455–466.

[31] Dayanik, S., and Karatzas, I. (2003). On the optimal stopping problem for one-dimensional diffusions. *Stochastic Processes and their Applications 107*, 173–212.

[32] Dasci A., and Laporte, G. (2005). A continuous model for multistore competitive location. *Operations Research 53*, 263–280.

[33] Delbaen, F., Peng, S., and Rosazza Gianin, E. (2010). Representation of the penalty term of dynamic concave utilities. *Finance and Stochastics 14*, 449–472.

[34] Dharma Kwon, H. (2010). Invest or exit? Optimal decisions in the face of a declining profit stream. *Operations Research 58*, 638–649.

[35] Dixit, A. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy 97*, 620–638.

[36] Dixit, A., and Pindyck, R.S. (1996). *Investment under uncertainty.* Princeton University Press, Princeton.

[37] Duffie, D. and Epstein, L.G. (1992). Stochastic differential utility. *Econometrica 60*, 353–394.

[38] Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics 75*, 643–669.

[39] Föllmer, H., and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics 6*, 429–447.

[40] Föllmer, H., and Schied, A. (2004). *Stochastic Finance*, Walter de Gruyter & Co., Berlin.

[41] Frittelli, M., and Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking & Finance 26*, 1473–1486.

[42] Giesecke, K., and Smelov, D. (2013). Exact sampling of jump-diffusions. *Operations Research 61*, 894–907.

[43] Gilboa, I., and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*, 141–153.

[44] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer-Verlag, New York.

[45] Glasserman, P., and Xu, X. (2013). Robust portfolio control with stochastic factor dynamics. *Operations Research 61*, 874–893.

[46] Guo, X., and Pham, H. (2005). Optimal partially reversible investment with entry decision and general production function. *Stochastic Processes and their Applications 115*, 705–736.

[47] Hansen, L., and Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy 99*, 225–262.

[48] Hansen, L., and Sargent, T. (2001). Robust control and model uncertainty. *American Economic Review 91*, 60–66.

[49] Hansen, L., and Sargent, T. (2007). *Robustness.* Princeton University Press, Princeton.

[50] Haugh, M., and Kogan, L. (2004). Pricing American options: A duality approach. *Operations Research 52*, 258–270.

[51] Heath, D., and Ku, H. (2004). Pareto equilibria with coherent measures of risk. *Mathematical Finance 14*, 163–172.

[52] Henderson, V., and Hobson, D. (2013). Risk aversion, indivisible timing options and gambling. *Operations Research 61*, 126–137.

[53] Huber, P. (1981). *Robust Statistics.* Wiley, New York.

[54] Johnson, H. (1987). Options on the maximum or the minimum of several assets. *Journal of Financial and Quantitative Analysis 22*, 277–283.

[55] Karatzas, I., and Shreve, S. (1998). *Methods of Mathematical Finance.* Vol. 39. Springer, New York.

[56] Knight, F.H. (1921). *Risk, Uncertainty, and Profit.* Houghton Mifflin Co., Boston.

[57] Krätschmer, V., and Schoenmakers, J. (2010). Representations for optimal stopping under dynamic monetary utility functionals. *SIAM Journal on Financial Mathematics 1*, 811–832.

[58] Laeven, R.J.A., and Stadje, M. (2013). Entropy coherent and entropy convex measures of risk. *Mathematics of Operations Research 38*, 265–293.

[59] Laeven, R.J.A., and Stadje, M. (2014). Robust portfolio choice and indifference valuation. *Mathematics of Operations Research 39*, 1109–1141.

[60] Lesnevski, V., Nelson, B. and Staum, J. (2007). Simulation of coherent risk measures based on generalized scenarios. *Management Science 53*, 1756–1769.

[61] Longstaff, F., and Schwartz, E. (2001). Valuing american options by simulation: A simple least-squares approach. *Review of Financial Studies 14*, 113–147.

[62] Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica 74*, 1447–1498.

[63] McDonald, R., and Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics 101*, 707–727.

[64] Øksendal, B. (2000). Irreversible investment problems. *Finance and Stochastics 4*, 223–250.

[65] Øksendal, B., and Sulem, A. (2007). *Applied stochastic control of jump diffusions*, 2nd Ed. Universitext Springer, Berlin.

[66] Øksendal, B., Sulem, A., and Zhang, T. (2014). Singular control and optimal stopping of SPDEs, and backward SPDEs with reflection. *Mathematics of Operations Research 39*, 464–486.

[67] Peng, S. (2004). *Nonlinear Expectations, Nonlinear Evaluations and Risk Measures*, Vol. 1856 of *Lecture Notes in Mathematics*. Springer, Berlin.

[68] Peskir, G., and Shiryaev, A. N. (2006). *Optimal stopping and free-boundary problems.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel.

[69] Philpott, A., de Matos, V., and Finardi, E. (2013). On solving multistage stochastic programs with coherent risk measures. *Operations Research 61*, 957–970.

[70] Pindyck, R.S. (1986). Irreversible investment, capacity choice, and the value of the firm. *American Economic Review 78*, 969–985.

[71] Riedel, F. (2009). Optimal stopping with multiple priors. *Econometrica 77*, 857–908.

[72] Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic processes and their Applications 112*, 185–200.

[73] Rogers, L.C.G. (2002). Monte Carlo valuation of American options. *Mathematical Finance 12*, 271–286.

[74] Ruszczyński, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming Series B 125*, 235–261.

[75] Ruszczyński, A., and Shapiro, A. (2006). Optimization of convex risk functions. *Mathematics of Operations Research 31*, 433–452.

[76] Ruszczyński, A., and Shapiro, A. (2006). Conditional risk mappings. *Mathematics of Operations Research 31*, 544–561.

[77] Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica 57*, 571–587.

[78] Schoenmakers, J., Zhang, J., and Huang, J. (2013). Optimal dual martingales, their analysis and application to new algorithms for Bermudan products. *SIAM Journal on Financial Mathematics 4*, 86–116.

[79] Shapiro, A., Dentcheva, D. and Ruszczyński, A. (2009). *Lectures on stochastic programming: modeling and theory.* SIAM, Philadelphia.

[80] Snell, L.J. (1952). Applications of martingale system theorems. *Trans. Amer. Math. Soc. 73*, 293–312.

[81] Tang, S., and Li, X. (1994). Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM Journal of Control and Optimization 32*, 1447–1475.

[82] Tang, S., and Wei, W. (2012). Representation of dynamic time-consistent convex risk measures with jumps. *Risk and Decision Analysis 3*, 167–190.

[83] Tekaya, W., Shapiro, A.S., Soares, M.P., and da Costa, J.P., (2013). Worst-case-expectation approach to optimization under uncertainty. *Operations Research 61*, 1435–1449.

[84] Tsitsiklis, J. N., and Van Roy, B. (2000). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks 12*, 694–703.

[85] Wald, A. (1950). *Statistical Decision Functions.* Wiley, New York.

[86] Zhu, H., Ye, F., and Zhou, E. (2013). True martingales for upper bounds on Bermudan option prices under jump-diffusion processes. *Proceedings of the 2013 Winter Simulation Conference.*