

# Optimal Stopping under Uncertainty in Drift and Jump Intensity

Volker Krätschmer

Faculty 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

Faculty of Mathematics and Economics

Ulm University

`mitja.stadje@uni-ulm.de`

This Version: July 28, 2017

## Abstract

This paper studies the optimal stopping problem in the presence of model uncertainty (ambiguity). We develop a numerically implementable method to 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 also provide asymptotically optimal exercise rules. We analyze the limiting 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.

**OR/MS Classification:** Dynamic programming/optimal control: Models; Dynamic programming/optimal control: Applications; Decision analysis: Risk.

**JEL Classification:** D81, G10, G20.

# 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 [95] and Snell [90], and we mention here only an incomplete selection related to the setting of this paper: Brennan and Schwartz [22], McDonald and Siegel [74], Barone-Adesi and Whaley [3], Dixit [42], Dixit and Pindyck [43], Karatzas and Shreve [64], Dayanik and Karatzas [37], Guo and Pham [54], Dasci and Laporte [38], Peskir and Shiryaev [79], Øksendal and Sulem [76], Henderson and Hobson [60], and Dharma Kwon [41]. Prime applications are a manufacturer's market entry decision 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. 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 objectively or subjectively 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.<sup>1</sup> 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 [65]) 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 [48], Frittelli and Rosazza Gianin [50], and Heath and Ku [59], extending Artzner et al. [2]; see also the early Ben-Tal [14] and Ben-Tal and Teboulle [15, 16]). For applications of convex risk measures in the context of decision and optimization, see e.g., Ruszczynski and Shapiro [85], Lesnevski, Nelson and Staum [69], Ben-Tal, Bertsimas and Brown [18], Choi, Ruszczynski and Zhao [31], Tekaya et al. [93], and Laeven and Stadje [67, 68]. By the representation theorem of convex risk measures, a random future reward, say  $H$ , is evaluated according to

$$U(H) = \inf_{Q \in \mathcal{Q}} \{E_Q[H] + c(Q)\}, \quad (1.1)$$

---

<sup>1</sup>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.

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  $Q$ .<sup>2</sup> Models  $Q$  that have ‘high’ plausibility are associated with a low penalty, while models that have ‘low’ plausibility yield a high 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. The case in which the null sets are also ambiguous (meaning that the decision-maker is not sure if these sets really have zero probability) is called the non-dominated case and will not be considered in this paper.

A canonical class of penalty functions is provided by  $\phi$ -divergences; see e.g., Ben-Tal and Teboulle [16, 17]. 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  $\phi$ -divergence measure with respect to the approximation  $P$ . A similar approach was adopted by Hansen and Sargent [56, 57] in macroeconomics, using the specific Kullback-Leibler ( $\phi$ -)divergence (or relative entropy; see also Csiszár [35] and Ben-Tal [14]). Another special case of interest is given by penalty functions of the form

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

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<sup>3</sup>

$$U(H) = \inf_{Q \in M} E_Q[H], \quad (1.3)$$

which attaches the same plausibility to all probabilistic models in  $M$ ; see e.g., Föllmer and Schied [49] for further details. In a dynamic setting, such as considered in this paper, time-consistent versions of convex measures of risk were analyzed by Riedel [82]. They have also been considered more recently in e.g., Ruszczyński and Shapiro [86], Cheridito, Delbaen and Kupper [30], Ruszczyński [84], Philpott, de Matos and Finardi [80], and Laeven and Stadjé [68]; see also Duffie and Epstein [44], Chen and Epstein [27], and Shapiro, Dentcheva and Ruszczyński [89], Chapter 6. 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 not optimal at time  $t = 0$  will not be optimal 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 [47]. It has been noted that incorporating ambiguity may not only be of theoretical and normative interest, but can also play a potential role in explaining empirically important failures of a purely risk-based framework (Chen and Epstein [27]). Popular approaches to decision-making under ambiguity are provided by the multiple priors preferences of Gilboa and Schmeidler [52] (see also

<sup>2</sup>In the literature, a convex risk measure is usually defined as  $-U(H)$  leading however to the same optimization problem.

<sup>3</sup>In this case,  $U$  corresponds to a coherent risk measure given by  $-U(H)$ .

Schmeidler [87]), also referred to as maxmin expected utility, and the significant generalization of variational preferences developed by Maccheroni, Marinacci and Rustichini [71]. 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 (Cerrei-Vioglio et al. [26]). A version of multiple priors was also studied by Huber [62] in robust statistics; see also the early Wald [95].

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 [81], Krättschmer and Schoenmakers [66], Bayraktar, Karatzas and Yao [4], Bayraktar and Yao [5], Cheng and Riedel [29], Øksendal, Sulem and Zhang [77], and Belomestny and Krättschmer [10, 11]. For optimal stopping under ambiguity aversion with non-dominated families of measures, we refer to Bayraktar and Yao [6, 7, 8], Ekren, Touzi and Zhang [45], Matoussi, Possamai and Zhou [73], Matoussi, Piozin and Possamai [72] and Nutz and Zhang [75]. The development of numerical methods to practically solve dominated robust optimal stopping problems in full generality may, however, currently be considered breaking ground.

In this paper, we develop and analyze a numerically implementable 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 suitably extend and exploit two duality theories of different kinds. The ‘first kind’ of duality theory is the martingale duality approach to standard (non-robust) optimal stopping problems, introduced by Rogers [83] and Haugh and Kogan [58] (see also Davis and Karatzas [36]), and exploited by Andersen and Broadie [1] to develop a numerically implementable algorithm which, like our own approach, generates approximations that converge to the true optimal solution and are ‘biased high’. By an appropriate generalization of the notion of a martingale, and with some suitable modifications, we extend their dual representation to encompass general preference functionals beyond plain conditional expectation; see also Krättschmer and Schoenmakers [66] in a discrete-time setting with simple rewards instead of a continuous-time setting with general rewards as in this paper. The ‘second kind’ of duality theory explicates the connection between time-consistent convex measures of risk and backward stochastic differential equations (BSDEs), which we extend to apply to our setting. We note that powerful numerical tools are nowadays available for BSDEs.

Exploiting these duality results, our first main contribution is then to develop and analyze a numerical method composed of three steps. First, by our 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, by our 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 (and use again duality of the second kind) 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 general dominated robust optimal stopping problems in the literature.

Next, we show that our algorithm yields quite naturally exercise rules (i.e., stopping times) with good asymptotic properties. We first prove that in the case that the optimal stopping time is unique the stopping times derived from our algorithm converge in probability to the proper

limit. In the case that the optimal stopping time is not unique, the question of convergence becomes more subtle. One may actually see that then, the exercise rules do not necessarily converge in probability anymore, since they may on a non-zero set continue to switch between continuation and exercising. However, we show next that, even in this case, the stopping times derived from our algorithm are asymptotically optimal in the sense that, with probability arbitrarily close to one, they agree with an optimal stopping time. So the decision maker who stops according to our algorithm will follow an optimal strategy except on a set whose probability converges to zero. Another way to phrase this non-trivial result would be to say that the stopping times derived from our algorithm converge to the set of optimal stopping times. Of course, if the optimal stopping time is unique, this entails the convergence to the unique optimal stopping time. The mathematical details in these convergence results are delicate.

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 [25] and Longstaff and Schwartz [70]; see also Tsitsiklis and Van Roy [94] and Clément, Lamberton and Protter [32]. These methods, which are connected to the stochastic mesh method of Broadie and Glasserman [23] (see Glasserman [53]), 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 [83] and Haugh and Kogan [58]. Many follow-up papers have further refined their method; see e.g., Belomestny, Bender and Schoenmakers [9] and Schoenmakers, Zhang and Huang [88] 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 to 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.

Our numerical method yields an approximation to the true evaluation of the general robust optimal stopping problem with two important features: (i) it converges asymptotically (in a sense that is made precise in the paper); and (ii) it is a biased high estimate at the pre-limiting level. A variety of algorithms in the standard optimal stopping literature do not produce bounds, but merely have the property of asymptotic convergence, or produce lower bounds. Our approximation, like that of Andersen and Broadie [1], converges asymptotically, and hence can serve as a good approximation when taking a fine time grid, a suitable selection of basis functions and a sufficiently large number of Monte Carlo simulations. Moreover, it also has the additional benefit of being biased high, so that with a finite grid and numbers of simulations and basis functions on average some protection is provided. The development of primal biased low algorithms for robust optimal stopping is left to future research.

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 size of the time grid tends to zero. Bender, Schweizer, and Zhuo [12] 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. A detailed step-wise description of our algorithm and its convergence properties are presented in Section 5. In Section 6, we develop the asymptotically optimal exercise rules. Section 7 contains the numerical examples. Details of all 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  $W = (W^1, \dots, W^d)^\top$ .
- (ii) A standard  $k$ -dimensional Poisson process  $N = (N^1, \dots, N^k)^\top$  with intensities  $\lambda_P = (\lambda_P^1, \dots, \lambda_P^k)^\top$ .

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  $\tilde{N} = (\tilde{N}^1, \dots, \tilde{N}^k)^\top$ , where  $\tilde{N}_t^i = N_t^i - \lambda_P^i t$ ,  $i = 1, \dots, k$ . 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  $\tau$  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

$$H_{t_i} = \Pi(t_i, X_{t_i}) + \sum_{j=i+1}^L h(t_j, X_{t_j}), \quad t_i \in \{t_0 = 0, t_1, \dots, t_L = T\}, \quad (2.1)$$

for functions  $\Pi$  and  $h$  mapping from  $\{t_0 = 0, t_1, \dots, t_L = T\} \times \mathbb{R}^n$  to  $\mathbb{R}$ . Furthermore, we assume that  $\Pi(t_i, X_{t_i})$  and  $h(t_j, X_{t_j})$  are in  $L^2(P)$  for all  $i = 0, \dots, L$  and  $j = 1, \dots, L$ . Standard examples that take the form (2.1) include (see the references provided in the Introduction):

- (a) *The optimal entrance problem:* In this case, typically  $\Pi(t, x) = -\exp(-\rho t) \kappa$ , for a fixed irreversible cost  $\kappa$  depreciating at a continuous rate  $\rho$ , 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,  $\xi$ . 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  $t$ .

**Remark 1** *The setting in this paper features general rewards with a finite set of exercise dates and an underlying process that evolves in continuous-time. In the case of simple rewards, such payoffs are also referred to as Bermudan options, which are common in interest rate and energy markets. Furthermore, American options with continuous exercise possibilities are often treated numerically by approximating the exercise time interval by a discrete set of exercise dates, hence by approximation with a Bermudan option.*

In standard optimal stopping problems, the decision-maker maximizes the expected reward under a given probabilistic model  $P$ :  $\max_{\tau \in \mathcal{T}} \mathbb{E}[H_\tau]$ , where  $\mathcal{T} = \{t_0 = 0 < t_1 < \dots < t_L = T\}$  is the set of possible exercise dates. Specifying the model  $P$  in this 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 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 [52]) and variational preferences (Maccheroni, Marinacci and Rustichini [71]). With linear utility, these decision criteria correspond to coherent (Artzner et al. [2]) and convex measures of risk (Föllmer and Schied [48]). Henceforth, we postulate that the decision-maker adopts a convex measure of risk and evaluates his future reward according to

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

with  $\mathcal{Q} = \{Q | Q \sim P\}$  and  $c : \mathcal{Q} \rightarrow \mathbb{R}_0^+ \cup \{\infty\}$ . (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

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

in which the non-negative, convex, and lower semi-continuous function  $Q \rightarrow c_t(Q)$  reflects the esteemed plausibility of the model  $Q$  given the information up to time  $t$ . We will assume that  $P$  is the reference measure which is not penalized, meaning that  $c_t(P) = 0$ . For a general payoff

$H$ , we define  $U_t(H)$  analogously to (2.3) with  $H_\tau$  replaced by  $H$ . In (2.3), and in the rest of this paper, we define for notational convenience  $\sup := \text{ess.sup}$  and  $\inf := \text{ess.inf}$ . The optimal stopping problem at time  $t_i$  is then given by

$$V_{t_i}^* = \sup_{\tau \in \mathcal{T}_i} U_{t_i}(H_\tau) = \sup_{\tau \in \mathcal{T}_i} \inf_{Q \in \mathcal{Q}} \{E_Q[H_\tau | \mathcal{F}_{t_i}] + c_{t_i}(Q)\}, \quad (2.4)$$

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  $U_t(H_2) \geq U_t(H_1) \Rightarrow U_s(H_2) \geq U_s(H_1)$ ,  $t \geq s$ . 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  $D_t = E \left[ \frac{dQ}{dP} | \mathcal{F}_t \right]$ . 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  $\lambda$  such that the Radon-Nikodym derivative can be written as

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

with  $\frac{\lambda_s}{\lambda_P} := \left( \frac{\lambda_s^1}{\lambda_P^1}, \dots, \frac{\lambda_s^k}{\lambda_P^k} \right)^\top$ . In particular,  $Q$  is uniquely characterized by  $q$  and  $\lambda$ . 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  $N_t$  has jumps with intensity  $\lambda_t$ . 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 [92], who generalized a result of Delbaen, Peng and Rosazza Gianin [40] obtained in a Brownian setting to a setting with jumps. (For the equivalence between (i) and (ii), see Lemma 11.11 on p. 466 of Föllmer and Schied [49].)

**Lemma 2 (Tang and Wei [92])** *Let  $U_t(H) = \inf_{\{Q \sim P | Q=P \text{ on } \mathcal{F}_t\}} \{E_Q[H | \mathcal{F}_t] + c_t(Q)\}$  for  $t \in [0, T]$ . The following statements are equivalent:*

- (i)  $U$  is time-consistent over bounded rewards.
- (ii)  $U$  is recursive, that is,  $U$  satisfies Bellman's dynamic programming principle given by  $U_0(U_t(H)I_A) = U_0(HI_A)$  for every  $t \in [0, T]$ ,  $A \in \mathcal{F}_t$  and bounded  $H$ .
- (iii) There exists a function

$$\begin{aligned} r : [0, T] \times \Omega \times \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \dots \times (-\lambda_P^k, \infty) &\rightarrow \mathbb{R} \cup \{\infty\} \\ (t, \omega, q, v) &\mapsto r(t, \omega, q, v), \end{aligned}$$



which is convex and lower semi-continuous in  $(q, v)$  with  $r(t, 0, 0) = 0$ , such that

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

Notice that analogously to the notation used for random variables we write  $r(t, q, \psi)$  instead of  $r(t, \omega, q, \psi)$ .

**Remark 3** 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  $\mathbb{R}^d \times (-\lambda_P^1, \infty) \times \dots \times (-\lambda_P^k, \infty)$  such that  $r(s, q, v) = I_{C_s}(q, v)$ . This is a consequence of the fact that, for coherent risk measures,  $c$  can only take the values zero or infinity; see also Delbaen [39] for a thorough analysis of this case in a continuous-time setting.

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 jump-diffusion 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

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

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

It is straightforward to show that (G1) implies that  $U$  is time-consistent and recursive also over square-integrable rewards.

**Remark 4** We note that for numerical tractability of the optimal stopping problem, we have assumed in (G1) that the functions  $r$  considered in this paper will not depend on  $\omega$ .

**Remark 5** 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  $\lambda_s$  under  $Q$  from the intensity  $\lambda_P$  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  $\lambda$  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  $(q, \lambda)$  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  $\lambda - \lambda_P$ .

We now illustrate the generality of (2.4) and (G1) with some examples of penalty functions satisfying our conditions. All these examples will reappear later in illustrations.

**Examples 6** (1) Kullback-Leibler divergence: A prototypical example of the penalty function in (2.4) is the Kullback-Leibler ( $\phi$ -)divergence given by

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

and  $\alpha > 0$ ; see Csiszár [35], Ben-Tal [14] and Ben-Tal and Teboulle [16, 17]. 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 [56, 57] to generate model robustness in macroeconomics. The parameter  $\alpha$  measures the degree of trust the decision-maker assigns to the reference model  $P$ . The limiting case of  $\alpha = \infty$  ( $\alpha = 0$ ) induces a maximal degree of trust (distrust). One may verify (see, for example, Proposition 9.10 in [34]) that in our continuous-time setting, for every  $Q$  satisfying  $c(Q) < \infty$ ,  $\alpha \text{KL}_t(Q|P)$  is of the form (2.7), where  $r(s, q, v) = \frac{\alpha}{2}|q|^2 + \alpha(\lambda_P + v)^\top \log(\mathbf{1} + \frac{v}{\lambda_P}) - \alpha \mathbf{1}^\top v$  with  $\mathbf{1} = (1, \dots, 1)^\top$ , where  $q$  and  $\lambda$  correspond to the model  $Q$  characterized through (2.5), and where the logarithm should be taken componentwise.

- (2) 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. Thus, he considers:

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

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

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

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,

$$\frac{dX_t^i}{X_t^i} = \mu^i dt + \sigma^i dW_t + J^i d\tilde{N}_t, \quad i = 1, \dots, n, \quad (2.8)$$

while in the second case

$$dX_t^i = \mu^i dt + \sigma^i dW_t + J^i d\tilde{N}_t, \quad i = 1, \dots, n, \quad (2.9)$$

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  $\mu = (\mu^1, \dots, \mu^n)^\top \in \mathbb{R}^n$ ,  $\sigma = (\sigma^1, \dots, \sigma^n)^\top \in \mathbb{R}^{n \times d}$  and  $J = (J^1, \dots, J^n)^\top \in (-1, \infty)^{n \times k}$  (former) or  $J = (J^1, \dots, J^n)^\top \in \mathbb{R}^{n \times k}$  (latter). In optimal entrance/exit decision problems, such as those provided in the Introduction,  $X$  often satisfies either (2.8) or (2.9) (with or without jumps). In finance,  $\mu^i$  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 6 (Continued; with (2.8) or (2.9) valid)**

- (3) Worst case with mean (partially) known: The decision-maker is certain that the (instantaneous or logarithmic instantaneous) mean return  $\mu^Q$  lies between a known lower and upper bound,  $(\mu^-)$  and  $(\mu^+)$ , respectively. As a special case,  $(\mu^-)$  and  $(\mu^+)$  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. Thus, he considers:

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

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

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

- (4) 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. 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 [33]. 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 by Hansen and Jagannathan [55], 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  $\Lambda$ . In this case, the good-deal bound evaluation  $U_t(H_\tau)$  is given by  $U_t(H_\tau) = \inf_{(q, \lambda) \in C} \mathbb{E}_Q [H_\tau]$ , with  $C = (C_t)_{t \in [0, T]}$  given by (see Björk and Slinko [20])

$$\begin{aligned} C_t &= \left\{ (q, \lambda - \lambda_P) \in \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \dots \times (-\lambda_P^k, \infty) \mid \mu + \sigma q + J(\lambda - \lambda_P) = 0 \right. \\ & \quad \left. \text{and } |q|^2 + \sum_{i=1}^k \frac{(\lambda^i - \lambda_P^i)^2}{\lambda_P^i} \leq \Lambda \right\}. \end{aligned}$$

(Note that in this example  $C_t$  does not depend on  $t$ .) This corresponds to a penalty function of the form (2.7) with

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

In order to obtain a biased high estimate of the optimal solution to the stopping problem (2.4), we will need the following additional assumption:

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

Assumption (G2) is needed (only) to obtain a genuine upper bound. (An Euler scheme, for example, may induce a bias and therefore violate the biased high property.) 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 [19], Broadie and Kaya [24], Chen and Huang [28], Giesecke and Smelov [51], and Henry-Labordère, Tan and Touzi [61].

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  $s$ ,

$$\left\{ (q, \lambda) \in \mathbb{R}^d \times (-\lambda_P^1, \infty) \times \dots \times (-\lambda_P^k, \infty) \mid r(s, q, \lambda - \lambda_P) < \infty \right\} \subset C_s,$$

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

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  $(\tau_{t_i}^*)_{t_i \in \{t_0=0, t_1, \dots, t_L=T\}}$  satisfying

$$V_{t_i}^* = \sup_{\tau \in \mathcal{T}_{t_i}} U_{t_i}(H_\tau) = U_{t_i}(H_{\tau_{t_i}^*}), \quad t_i \in \{0, \dots, T\}. \quad (3.1)$$

Furthermore, we show that Bellman’s principle

$$V_{t_i}^* = \max \left( \Pi(t_i, X_{t_i}) - h(t_i, X_{t_i}) + U_{t_i}^h, U_{t_i}(V_{t_{i+1}}^*) \right), \quad t_i \in \{0, \dots, t_{L-1}\}, \quad (3.2)$$

holds, with  $U_{t_i}^h$  defined as

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

see the Appendix for the technical details. For the intermediate time instances between the exercises times we set  $U_t^h := \inf_{(q, \lambda) \in C} \left\{ \mathbb{E}_Q \left[ \sum_{t \leq t_j} h(t_j, X_{t_j}) \middle| \mathcal{F}_t \right] + c_t(Q) \right\}$ . 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  $(M_t)_{t \in \{t_0=0, t_1, \dots, t_L=T\}}$  is a  $U$ -martingale if, for each  $t$ , we have  $M_t \in L^2(P)$  and furthermore  $M_s = U_s(M_t)$ , for  $s \leq t$ . By time-consistency, the last equation 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  $\mathcal{M}_0^U$ . Define, for  $i = 0, \dots, L$ ,

$$A_{t_i}^* := \sum_{j=1}^i (U_{t_{j-1}}(V_{t_j}^*) - V_{t_{j-1}}^*), \quad M_{t_i}^* := \sum_{j=1}^i (V_{t_j}^* - U_{t_{j-1}}(V_{t_j}^*)). \quad (3.4)$$

One may verify that  $M^*$  is a  $U$ -martingale,  $A^*$  is non-increasing and predictable,  $M_0^* = A_0^* = 0$ , and that

$$V_{t_i}^* = V_0^* + M_{t_i}^* + A_{t_i}^*, \quad i = 0, \dots, L, \quad (3.5)$$

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

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 extending the well-known dual representation for the standard setting, in which  $U$  is just the ordinary conditional expectation (Rogers [83] and Haugh and Kogan [58]). This generalized additive dual representation, the proof of which uses results obtained by Krättschmer and Schoenmakers [66] in a discrete-time setting with  $h = 0$ , reads as follows:

**Proposition 7** *Let  $M^* \in \mathcal{M}_0^U$  be the (unique)  $U$ -martingale in the  $U$ -Doob decomposition (3.5). Then the optimal stopping problem (2.4) has a dual representation*

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

**Remark 8** 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

$$V_{t_i}^* = \inf_{M \in \mathcal{M}_0} U_{t_i} \left( \max_{t_j \in \{t_i, \dots, T\}} \left( \Pi(t_j, X_{t_j}) - h(t_j, X_{t_j}) + U_{t_j}^h + M_{t_i} - M_{t_j} \right) \right), \quad t_i \in \{t_0 = 0, \dots, T\}, \quad (3.7)$$

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:

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

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 6 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  $t_j$ ,  $j = 0, \dots, L-1$ , we have a finer time grid  $\pi_j = \{s_{j0} = t_j < s_{j1} < \dots < s_{jP} = t_{j+1}\}$ . Denote the corresponding overall time grid by  $\pi = \{s_{00}, s_{01}, \dots, s_{LP}\}$ . The following theorem provides a way to practically compute  $M^*$  in (3.4) by connecting it to specific semi-martingale dynamics that can be dealt with numerically in an efficient way.

**Theorem 9** (a) *There exists a unique square integrable predictable  $(Z^h, \tilde{Z}^h)$  such that*

$$dU_t^h = -g(t, Z_t^h, \tilde{Z}_t^h)dt + Z_t^h dW_t + \tilde{Z}_t^h d\tilde{N}_t, \quad \text{for } t \in (t_j, t_{j+1}], \quad (3.9)$$

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

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

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

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

**Remark 10** Note that by Remark 5 and (3.3), we have terminal conditions  $U_T^h = h(T, X_T)$  and  $U_{t_{j+1}}(V_{t_{j+1}}^*) = V_{t_{j+1}}^*$ , for  $j = 0, \dots, L-1$ , in (3.9) and (3.10). Hence, given  $U_{t_{j+1}}^h$  and  $V_{t_{j+1}}^*$ , we may compute  $U_{t_j}^h$  and  $U_{t_j}(V_{t_{j+1}}^*)$  through the relationships given in Theorem 9(a);  $V_{t_j}^*$  can then be obtained by Bellman's principle (3.2).

**Remark 11** As  $U_{t_{j+1}}(V_{t_{j+1}}^*) = V_{t_{j+1}}^*$ , we can write, by Theorem 9(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  $t \in (t_j, t_{j+1}]$ ,

$$U_t^h = U_{t_{j+1}}^h + \int_t^{t_{j+1}} g(s, Z_s^h, \tilde{Z}_s^h) ds - \int_t^{t_{j+1}} Z_s^h dW_s - \int_t^{t_{j+1}} \tilde{Z}_s^h d\tilde{N}_s. \quad (3.13)$$

**Remark 12** 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  $U_t(H) = \mathbb{E}[H|\mathcal{F}_t]$ . 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 smaller evaluation.

Equations (3.9)–(3.10) are also referred to as backward stochastic differential equations (BSDEs)<sup>4</sup> 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 [78].

To conclude the exposition of the duality theory of the second kind, let us, for illustration purposes, employ the penalty functions of Examples 6 and compute the corresponding  $g$ ’s using (3.8). These  $g$  functions will later be used in numerical illustrations.

**Examples 13** (1) Kullback-Leibler divergence:  $g(t, z, \tilde{z}) = -\frac{|z|^2}{2\alpha} - \alpha \sum_{i=1}^k \lambda^i (e^{-\tilde{z}^i/\alpha} + \frac{\tilde{z}^i}{\alpha} - 1)$ .

(2) Worst case with ball scenarios: Suppose without loss of generality that  $|\lambda_P| \geq \delta_2$ . Then,  $g(t, z, \tilde{z}) = -\delta_1|z| - \delta_2|\tilde{z}|$ .

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

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

with

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

<sup>4</sup>Formally, given a terminal payoff  $H \in L^2$  and a function  $g : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}$ , the solution to the corresponding BSDE is a triple of square-integrable and suitably measurable processes  $(Y, Z, \tilde{Z})$  satisfying

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

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.

- (4) 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  $|(q, v)|_* := \sqrt{|q|^2 + \left| \frac{v}{\lambda_P} \right|^2}$ , 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

$$C = \left\{ (q, \lambda - \lambda_P) \mid A(q, \lambda - \lambda_P)^\top = b \quad \text{and} \quad |(q, \lambda - \lambda_P)|_* \leq \sqrt{\Lambda} \right\}.$$

(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  $P_B(0)$  be the projection of 0 onto the set  $B := \{x \mid Ax = b\}$  in the  $|\cdot|_*$  norm, and define  $P_{\text{Kernel}(A)}((z, \tilde{z})^\top)$  accordingly as the projection of  $(z, \tilde{z})^\top$  in the  $|\cdot|_*$  norm onto the space given by the kernel of the matrix  $A$ . One can prove (see the Appendix) that

$$g(t, z, \tilde{z}) = -\sqrt{\Lambda - |P_B(0)|_*^2} \left| P_{\text{Kernel}(A)}((z, \tilde{z})^\top) \right|_{**} + \langle (z, \tilde{z}), P_B(0) \rangle, \quad (3.15)$$

where  $|(z, \tilde{z})|_{**} := \sqrt{|z|^2 + \left| \sum_{i=1}^k \tilde{z}^i \lambda_P^i \right|^2}$ . This concludes our examples.

## 4 The Algorithm: General Outline

Our method is composed of three steps. Theorem 9 ('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 9(a) yields that, to construct the  $U$ -martingale  $M^*$  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  $(V_s^*)_{t_j < s \leq t_{j+1}}$ . 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 7. While this bound will be seen to converge to the true optimal solution asymptotically and is an approximated upper bound for a finite number of simulations, 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  $U_T^h = h(T, X_T)$ . 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  $U_s(V_{t_{j+1}}^*)_{t_j < s \leq t_{j+1}}$  through (3.10). This involves least squares Monte Carlo regression. We can then set

$$V_{t_j}^* = \max \left( \Pi(t_j, X_{t_j}) - h(t_j, X_{t_j}) + U_{t_j}^h, U_{t_j} \left( V_{t_{j+1}}^* \right) \right),$$

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  $(M_{t_j}^*)_{t_j \in \{t_1, \dots, T\}}$  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 5, but already preview the following results. Since our optimal stopping problem is Markovian, there exists a function  $v^* : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  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 16 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 tight 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 Algorithm: Step-Wise Description

### Preliminaries

Since the approximation scheme adopted in Step (1.a.) below will also be used in the steps that follow, it will be useful to first employ slightly more general notation. Recall the  $n$ -dimensional adapted process  $(X_t)_{t \in [0, T]}$ , satisfying the strong Markov property. Fix a  $t_{j+1}$  and consider  $t_j = s_{j0} < s_{j1} < \dots < s_{jP} = t_{j+1}$ . We start with a function  $w : \mathbb{R}^n \rightarrow \mathbb{R}$  (such that  $w(X_{t_{j+1}})$  is square-integrable) and the function  $g(t, z, \tilde{z})$ ;  $w$  will either be a known payoff function (at initialization  $t_{j+1} = t_L$ ) or a function obtained in the period before (for  $j+1 < L$ ). 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  $|\pi| := \max_{j,p} \Delta_{jp}$ ,  $j = 0, \dots, L-1$ ,  $p = 0, \dots, P-1$ . We will approximate the processes  $U_t^h$  or  $U_t(V_{t_{j+1}}^*)$  given by (3.12) and (3.13), respectively, with a (here, generic) process  $Y^\pi$ .

In general, suppose we are given  $Y_{t_{j+1}}^\pi = y_{t_{j+1}}^\pi(X_{t_{j+1}}) =: w(X_{t_{j+1}})$ . We then do a backward recursion over the  $s_{jp}$  initializing  $p = P$  and  $s_{jP} = t_{j+1}$ . Suppose we have an approximation  $Y_{s_{j(p+1)}}^\pi$  and we want to compute  $Y_{s_{jp}}^\pi$ . Theorem 9 then yields:

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

see (3.9). Taking conditional expectations,

$$Y_{s_{jp}}^\pi \approx \mathbf{E}_{jp} \left[ Y_{s_{j(p+1)}}^\pi \right] + g(s_{jp}, Z_{s_{jp}}^\pi, \tilde{Z}_{s_{jp}}^\pi) \Delta_{jp}, \quad (5.1)$$

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

$$(Z_{s_{jp}}^\pi, \tilde{Z}_{s_{jp}}^\pi) = \operatorname{argmin}_{(Z, \tilde{Z}) \in L_{d+k}^2(\sigma(X_{s_{jp}}))} \mathbf{E}_{jp} \left[ \left( Y_{s_{j(p+1)}}^\pi - Z \Delta W_{jp} - \tilde{Z} \Delta \tilde{N}_{jp} \right)^2 \right].$$

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

$$m^M(s_{jp}, x) = (m_1(s_{jp}, x), \dots, m_M(s_{jp}, x))^\top, \quad \text{for } x \in \operatorname{support}(X_{s_{jp}}),$$

and define  $\psi^M$  and  $\tilde{\psi}^M$  similarly. (In all our numerical implementations  $m$  will actually be defined on the whole  $\mathbb{R}^n$ .) Furthermore, define by  $P_{s_{jp}}^{\pi, M}(Y_{s_{j(p+1)}}^{\pi, M}) := \alpha_{s_{jp}}^{\pi, M} m^M(s_{jp}, X_{s_{jp}})$ , and

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

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

$$\alpha_{s_{jp}}^{\pi, M} = (A_{s_{jp}}^{\pi, M})^{-1} \mathbf{E} \left[ Y_{s_{j(p+1)}}^{\pi, M} m^M(s_{jp}, X_{s_{jp}}) \right],$$

$$\gamma_{s_{jp}}^{\pi, M} = (\bar{A}_{s_{jp}}^{\pi, M})^{-1} \mathbf{E} \left[ Y_{s_{j(p+1)}}^{\pi, M} \psi^M(s_{jp}, X_{s_{jp}}) \Delta W_{jp} \right], \quad (5.2)$$

$$\tilde{\gamma}_{s_{jp}}^{\pi, M} = (\tilde{\bar{A}}_{s_{jp}}^{\pi, M})^{-1} \mathbf{E} \left[ Y_{s_{j(p+1)}}^{\pi, M} \tilde{\psi}^M(s_{jp}, X_{s_{jp}}) \Delta \tilde{N}_{jp} \right], \quad (5.3)$$

with coefficients given by

$$\begin{aligned} (A_{s_{jp}}^{\pi,M})_{1 \leq k, l \leq M} &= \mathbb{E} [m_k^M(s_{jp}, X_{s_{jp}}) m_l^M(s_{jp}, X_{s_{jp}})], \\ (\bar{A}_{s_{jp}}^{\pi,M})_{1 \leq k, l \leq M} &= \mathbb{E} [\psi_k^M(s_{jp}, X_{s_{jp}}) \psi_l^M(s_{jp}, X_{s_{jp}})] \mathbb{E} [\Delta^2 W_{jp}], \end{aligned} \quad (5.4)$$

$$(\tilde{A}_{s_{jp}}^{\pi,M})_{1 \leq k, l \leq M} = \mathbb{E} [\tilde{\psi}_k^M(s_{jp}, X_{s_{jp}}) \tilde{\psi}_l^M(s_{jp}, X_{s_{jp}})] \mathbb{E} [\Delta^2 \tilde{N}_{jp}]. \quad (5.5)$$

Here, we define the process  $Y_{t_{j+1}}^{\pi,M}$  by setting  $Y_{t_{j+1}}^{\pi,M} = w(X_{t_{j+1}})$ , and then recursively

$$Y_{s_{jp}}^{\pi,M} = \alpha_{s_{jp}}^{\pi,M} m^M(s_{jp}, X_{s_{jp}}) + g(s_{jp}, \gamma_{s_{jp}}^{\pi,M} \psi^M(s_{jp}, X_{s_{jp}}), \tilde{\gamma}_{s_{jp}}^{\pi,M} \tilde{\psi}^M(s_{jp}, X_{s_{jp}})) \Delta_{jp}. \quad (5.6)$$

To compute the conditional expectations in (5.2)–(5.5) numerically, we simulate  $N_0$  independent paths  $(X_{s_{jp}}^n)_{s_{jp}}$  and  $(W_{s_{jp}}^n)_{s_{jp}}$ , starting with  $X_{t_{j+1}}$  for  $s_{jp} = t_{j+1}$ . Then, for  $n = 1, \dots, N_0$ , we assume that  $w(x) := y_{t_{j+1}}^{\pi,M,N_0}(x)$  is given, and compute

$$\begin{aligned} y_{s_{jp}}^{\pi,M,N_0}(x) &:= \alpha_{s_{jp}}^{\pi,M,N_0} m^M(s_{jp}, x) \\ &\quad + g(s_{jp}, \gamma_{s_{jp}}^{\pi,M,N_0} \psi^M(s_{jp}, x), \tilde{\gamma}_{s_{jp}}^{\pi,M,N_0} \tilde{\psi}^M(s_{jp}, x)) \Delta_{jp}, \end{aligned} \quad (5.7)$$

where

$$\begin{aligned} \alpha_{s_{jp}}^{\pi,M,N_0} &= (A_{s_{jp}}^{\pi,M,N_0})^{-1} \frac{1}{N_0} \sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0,n} m^M(s_{jp}, X_{s_{jp}}^n) \\ \gamma_{s_{jp}}^{\pi,M,N_0} &= (\bar{A}_{s_{jp}}^{\pi,M,N_0})^{-1} \frac{1}{N_0} \sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0,n} \psi^M(s_{jp}, X_{s_{jp}}^n) \Delta W_{jp}^n \\ \tilde{\gamma}_{s_{jp}}^{\pi,M,N_0} &= (\tilde{A}_{s_{jp}}^{\pi,M,N_0})^{-1} \frac{1}{N_0} \sum_{n=1}^{N_0} Y_{s_{j(p+1)}}^{\pi,M,N_0,n} \tilde{\psi}^M(s_{jp}, X_{s_{jp}}^n) \Delta \tilde{N}_{jp}^n, \end{aligned} \quad (5.8)$$

with coefficients given by

$$\begin{aligned} (A_{s_{jp}}^{\pi,M,N_0})_{1 \leq k, l \leq M} &= \frac{1}{N_0} \sum_{n=1}^{N_0} m_k^M(s_{jp}, X_{s_{jp}}^n) m_l^M(s_{jp}, X_{s_{jp}}^n) \\ (\bar{A}_{s_{jp}}^{\pi,M,N_0})_{1 \leq k, l \leq M} &= \frac{1}{N_0} \sum_{n=1}^{N_0} \psi_k^M(s_{jp}, X_{s_{jp}}^n) \psi_l^M(s_{jp}, X_{s_{jp}}^n) \Delta_{jp} \\ (\tilde{A}_{s_{jp}}^{\pi,M,N_0})_{1 \leq k, l \leq M} &= \frac{1}{N_0} \sum_{n=1}^{N_0} \tilde{\psi}_k^M(s_{jp}, X_{s_{jp}}^n) \tilde{\psi}_l^M(s_{jp}, X_{s_{jp}}^n) \lambda_P \Delta_{jp}. \end{aligned} \quad (5.9)$$

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

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

### 5.1.1 Step (1.a.): Construct an Approximation to $U^h$

We use the algorithm constructed above for  $j = L-1, \dots, 0$  and initialize  $w(X_{t_L}) = h(t_L, X_{t_L})$ . Next, in terms of the algorithm above, we define  $u_{s_{jp}}^{h,\pi,M,N_0}(x) := y_{s_{jp}}^{\pi,M,N_0}(x)$ ,  $z_{s_{jp}}^{h,\pi,M,N_0}(x) := \gamma_{s_{jp}}^{\pi,M,N_0} \psi^M(s_{jp}, x)$  and  $\tilde{z}_{s_{jp}}^{h,\pi,M,N_0}(x) := \tilde{\gamma}_{s_{jp}}^{\pi,M,N_0} \tilde{\psi}^M(s_{jp}, x)$ , for  $p = P, \dots, 0$ . After every step, we set (with a slight abuse of notation)  $w(X_{t_j}) = u_{s_{j0}}^{h,\pi,M,N_0}(X_{t_j}) + h(t_j, X_{t_j})$  and  $j = j - 1$ . We continue until  $j = -1$ .

### 5.1.2 Step (1.b.): Construct an Approximation to $V^*$

To do a backward recursion over  $t_j$ , we initialize  $t_j = t_L = T$  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  $U_{s_{jP}}^\pi := U_{t_{j+1}}(V_{t_{j+1}}^{*,\pi,M,N_1}) = V_{t_{j+1}}^{*,\pi,M,N_1}$ . Now, given  $U_{s_{j(p+1)}}^\pi$ , we know from (3.10) that

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

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  $w(\cdot) = v^{*,\pi,M,N_1}(t_{j+1}, \cdot)$  as terminal condition). This yields functions  $u^{\pi,M,N_1}$ ,  $z^{\pi,M,N_1}$  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

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

We stop if  $j = 0$ .

### 5.1.3 Step (1.c.): Construct an Approximation to $M^*$

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

$$\begin{aligned} M_{s_{ip}}^{g,\pi,M,N_1} &:= - \sum_{j=0}^i \sum_{l=0}^{p-1} \int_{s_{jl}}^{s_{j(l+1)}} g(u, z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}), \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}})) du \\ &\quad + \sum_{j=0}^i \sum_{l=0}^{p-1} z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}) \Delta W_{jl} + \sum_{j=0}^i \sum_{l=0}^{p-1} \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}) \Delta \tilde{N}_{jl}, \end{aligned} \quad (5.10)$$

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

$$\begin{aligned} M_{s_{ip}}^{g,\pi,M,N_1,n} &:= - \sum_{j=0}^i \sum_{l=0}^{p-1} \int_{s_{jl}}^{s_{j(l+1)}} g(u, z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n), \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n)) du \\ &\quad + \sum_{j=0}^i \sum_{l=0}^{p-1} z_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n) \Delta W_{jl}^n + \sum_{j=0}^i \sum_{l=0}^{p-1} \tilde{z}_{s_{jl}}^{\pi,M,N_1}(X_{s_{jl}}^n) \Delta \tilde{N}_{jl}^n. \end{aligned} \quad (5.11)$$

Note that (5.10) defines a true discrete-time  $U$ -martingale  $(M_{t_j}^{g,\pi,M,N_1})_{j \in \{0,1,2,\dots,L\}}$ , and that (5.11) gives rise to an exact simulation scheme of it. The reason is that  $M^{g,\pi,M,N_1}$  by definition satisfies a BSDE with terminal condition  $M_T^{g,\pi,M,N_1}$  and driver function  $g$  (details are in the Appendix). 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 time grid  $\pi$ ) will be needed for the numerical approximation.

## 5.2 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 7. To this end, we are faced with the computation of

$$\begin{aligned} V_0^* &= \inf_{M \in \mathcal{M}_0^U} U_0 \left( \max_{t_j \in \{0, t_1, \dots, T\}} \left( \Pi(t_j, X_{t_j}) - h(t_j, X_{t_j}) + U_{t_j}^h + M_T - M_{t_j} \right) \right) \\ &= U_0 \left( \max_{t_j \in \{0, t_1, \dots, T\}} \left( \Pi(t_j, X_{t_j}) - h(t_j, X_{t_j}) + U_{t_j}^h + M_T^* - M_{t_j}^* \right) \right). \end{aligned} \quad (5.12)$$

We set

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

Since we can only compute an approximation to  $M^*$ , we cannot attain the infimum in (5.12). However,  $M_T^{g, \pi, M, N_1}$  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  $N_0 = N_1$ ,

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

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

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

for  $s_{jp} \leq s < s_{j(p+1)}$ . Let us compute  $U_0(F^{\pi, M, N_1})$  numerically. Recall that for a payoff  $H \in L^2(P)$ , by Theorem 9(a),

$$U_t(H) = \inf_{Q \sim P} \left\{ \mathbb{E}_Q \left[ H + \int_t^T r(s, q_s, \lambda_s - \lambda_P) ds \mid \mathcal{F}_t \right] \right\} \quad (5.13)$$

$$= H + \int_t^T g(s, Z_s, \tilde{Z}_s) ds - \int_t^T Z_s dW_s - \int_t^T \tilde{Z}_s d\tilde{N}_s. \quad (5.14)$$

Hence, we can apply the approximation scheme (5.7)–(5.9) (with  $X = \mathcal{X}$  and terminal condition  $\max_{t_l \in \{0, t_1, \dots, t_j\}} \left( \Pi(t_l, X_{t_l}) - h(t_l, X_{t_l}) + u_{t_l}^{h, \pi, M, N_1}(X_{t_l}) - M_{t_l}^{g, \pi, M, N_1} \right)$ ). Simulate  $n = 1, \dots, N_2$  paths

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

Let  $M$  be the number of basis functions in the least squares Monte Carlo regression. We then obtain coefficients, say  $\alpha_j^{\pi, M, N_2}$ ,  $\gamma_j^{\pi, M, N_2}$ ,  $\tilde{\gamma}_j^{\pi, M, N_2}$ , and processes  $(V_t^{\pi, M, N_2}, Z_t^{\pi, M, N_2}, \tilde{Z}_t^{\pi, M, N_2})_{0 \leq t \leq T}$ . Then, by applying Theorem 21 in the Appendix three times, we may conclude that

$$\lim_{\pi \rightarrow 0} \lim_{M \rightarrow \infty} \lim_{N_2, N_1, N_0 \rightarrow \infty} (V^{\pi, M, N_2}, Z^{\pi, M, N_2}, \tilde{Z}^{\pi, M, N_2}) = (V^*, Z^*, \tilde{Z}^*); \quad (5.15)$$

see the technical details provided in the Appendix. In particular,  $V_0^{\pi, M, N_2} \rightarrow V_0^*$  as the mesh size of the grid,  $\pi$ , 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, for finite  $N_0, N_1, N_2$ , our estimates from Step (2.) for the upper bound to  $V^*$  are not biased high (above the Snell envelope), meaning that on average the upper bound may not provide enough protection. For this reason we will subsequently proceed to construct a genuine (biased high) upper bound.

### 5.3 Step (3.): Backward-Forward Simulation

#### 5.3.1 Step (3.a.): Construct a Genuine Upper Bound to $U^h$

By Theorem 9(a), for  $i = 0, \dots, L - 1$ ,

$$U_{t_i}^h = \inf_{Q \sim P} \left\{ \mathbb{E}_Q \left[ \sum_{j=i}^L h(t_j, X_{t_j}) + \int_{t_i}^T r(s, q_s, \lambda_s - \lambda_P) ds \middle| \mathcal{F}_{t_i} \right] \right\}, \quad (5.16)$$

$$= U_{t_{i+1}}^h + \int_{t_i}^{t_{i+1}} g(s, Z_s^h, \tilde{Z}_s^h) ds - \int_{t_i}^{t_{i+1}} Z_s^h dW_s - \int_{t_i}^{t_{i+1}} \tilde{Z}_s^h d\tilde{N}_s + h(t_i, X_{t_i}). \quad (5.17)$$

Denote the  $Q$  that attains the infimum in (5.16) 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 (5.17).

**Proposition 14** *For  $H \in L^2(P)$ , the infimum in (5.16) 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  $(q_s^*, \lambda_s^* - \lambda_P) \in \partial g(s, Z_s^h, \tilde{Z}_s^h)$ , where  $\partial g(s, \cdot, \cdot)$  stands for the subdifferentials of the convex function  $g(s, \cdot, \cdot)$ .<sup>5</sup>

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

- (i) Computing approximations to  $(Z, \tilde{Z})$  by solving (5.17). In view of Proposition 14,  $(Z, \tilde{Z})$  induces an approximation to  $Q^h$ , say  $Q^{h, \text{approx}}$ .
- (ii) Evaluating  $\mathbb{E}_{Q^{h, \text{approx}}} \left[ \sum_{j=i}^L h(t_j, X_{t_j}) + \int_{t_i}^T r(s, q_s, \lambda_s - \lambda_P) ds \middle| X_{t_i} \right]$  and making use of (5.16). This will deliver the desired genuine (biased high) upper bound to  $(U_{t_j}^h)_{t_j \in \{0, \dots, t_{L-1}\}}$ .

So let us carry out our program to compute approximations  $U_{t_j}^{h, n}$ , for  $n = 1, \dots, N_3$ : Simulate  $N_3$  copies of  $(X_{s_{jp}}^n)$  ('outer simulation'). For  $X_{t_j}^n = x$ , let  $N_4 \in \mathbb{N}$  and simulate additional paths  $(X_{s_{jp}}^{t_j, x, n})$  and new  $(W_{s_{jp}}^n)$  for  $n = 1, \dots, N_4$  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

<sup>5</sup>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.

may still be implemented by taking elements in the subgradient.) Define, with  $N_0 = N_1$ ,  $z_{s_{jp}}^{h,\pi,M,N_1}(\bar{x}) := \gamma_{s_{jp}}^{h,\pi,M,N_1} \psi^M(s_{jp}, \bar{x})$ ,  $\tilde{z}_{s_{jp}}^{h,\pi,M,N_1}(\bar{x}) := \tilde{\gamma}_{s_{jp}}^{\pi,M,N_1} \tilde{\psi}^M(s_{jp}, \bar{x})$  and

$$\begin{aligned} q_{s_{jp}}^{h,\pi,t_j,x,n} &:= g_z(s_{jp}, z_{s_{jp}}^{h,\pi,M,N_1}(X_{s_{jp}}^{t_j,x,n}), \tilde{z}_{s_{jp}}^{h,\pi,M,N_1}(X_{s_{jp}}^{t_j,x,n})) \\ \lambda_{s_{jp}}^{h,\pi,t_j,x,n} - \lambda_P &:= g_{\tilde{z}}(s_{jp}, z_{s_{jp}}^{h,\pi,M,N_1}(X_{s_{jp}}^{t_j,x,n}), \tilde{z}_{s_{jp}}^{h,\pi,M,N_1}(X_{s_{jp}}^{t_j,x,n})). \end{aligned}$$

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

$$\begin{aligned} D_{t_i}^{\pi,n}(x) &:= \exp \left( \sum_{t_i \leq s_{jp}} q_{s_{jp}}^{h,\pi,t_j,x,n} \Delta W_{jp}^n + \sum_{t_i \leq s_{jp}} \log \left( \frac{\lambda_{s_{jp}}^{h,\pi,t_j,x,n}}{\lambda_P} \right) \Delta N_{jp}^n \right. \\ &\quad \left. - \sum_{t_i \leq s_{jp}} \left( \frac{1}{2} |q_{s_{jp}}^{h,\pi,t_j,x,n}|^2 + \lambda_{s_{jp}}^{h,\pi,t_j,x,n} - \lambda_P \right) \Delta_{jp} \right); \end{aligned}$$

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

$$\begin{aligned} \tilde{u}_{t_j}^{upp,h,N_4}(x) &:= \frac{1}{N_4} \sum_{n=1}^{N_4} D_{t_j}^{\pi,n}(x) \left[ \sum_{l=j}^L h(t_l, X_{t_l}^{t_j,x,n}) \right. \\ &\quad \left. + \sum_{l=j}^L \sum_{p=0}^{P-1} \int_{s_{lp}}^{s_{l(p+1)}} r(s, q_{s_{lp}}^{h,\pi,t_j,x,n}, \lambda_{s_{lp}}^{h,\pi,t_j,x,n} - \lambda_P) ds \right]. \end{aligned} \quad (5.18)$$

Now  $(D_{t_j}^{\pi,n}(X_{t_j}^n))$ ,  $(q_{s_{jp}}^{h,\pi,t_j,x,n})_{j,p}$  and  $(\lambda_{s_{jp}}^{h,\pi,t_j,x,n})_{j,p}$  are true i.i.d. simulations of  $\frac{dQ^{h,\pi,M,N_4}}{dP}$ , the piecewise constant  $(q_t)$  and  $(\lambda_t)$ , conditioned on  $X_{t_j} = x$ . Therefore, by (5.14),  $\tilde{u}_t^{upp,h,N_4}(X_t^n)$  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 (the proof of which can be found in the Appendix).

**Proposition 15** *We have  $\mathbb{E} \left[ \tilde{u}_t^{upp,h,N_4}(x) \right] \geq u_t^h(x)$ , for any  $x$ .*

### 5.3.2 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 (5.13) by  $Q^g$ , with corresponding  $(q_s^*, \lambda_s^* - \lambda_P)$ . As in Proposition 14 one may see that  $(q_s^*, \lambda_s^* - \lambda_P) \in \partial g(s, Z_s, \tilde{Z}_s)$  with  $(Z, \tilde{Z})$  from (5.14). 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  $n = 1, \dots, N_3$  and  $j, p$ . Define

$$\begin{aligned} U_{t_j}^{upper,h,N_4,n} &:= \tilde{u}_{t_j}^{upp,h,N_4}(X_{t_j}^n), \\ q_{s_{jp}}^{\pi,M,N_2,n} &:= g_z(s_{jp}, z_{s_{jp}}^{\pi,M,N_2}(X^n), \tilde{z}_{s_{jp}}^{\pi,M,N_2}(X^n)), \\ \lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P &:= g_{\tilde{z}}(s_{jp}, z_{s_{jp}}^{\pi,M,N_2}(X^n), \tilde{z}_{s_{jp}}^{\pi,M,N_2}(X^n)). \end{aligned}$$

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

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

Finally, we set

$$\begin{aligned} \tilde{V}_0^{upp,N_3} := \frac{1}{N_3} \sum_{n=1}^{N_3} \frac{dQ^{g,\pi,M,N_3,n}}{dP} & \left[ \max_{t_j \in \{0, \dots, T\}} \left( \Pi(t_j, X_{t_j}^n) - h(t_j, X_{t_j}^n) + U_{t_j}^{upper,h,N_4,n} + M_T^{g,\pi,M,N_1,n} - M_{t_j}^{g,\pi,M,N_1,n} \right) \right. \\ & \left. + \sum_{j,p} \int_{s_{jp}}^{s_{j(p+1)}} r(s, q_{s_{jp}}^{\pi,M,N_2,n}, \lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P) ds \right], \end{aligned} \quad (5.19)$$

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

## 5.4 Summary and First Main Result

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

- (1.) Run  $N_0$  Monte Carlo simulations to compute  $U^{h,\pi,M,N_0}$ . Run  $N_1$  Monte Carlo simulations to compute  $M^{g,\pi,M,N_1}$ . To fully describe the evolution of these processes, it is sufficient to store the corresponding  $(\alpha_{s_{jp}}^{h,\pi,M,N_0})$ ,  $(\gamma_{s_{jp}}^{h,\pi,M,N_0})$ ,  $(\tilde{\gamma}_{s_{jp}}^{h,\pi,M,N_0})$ ; and  $(\alpha_{s_{jp}}^{\pi,M,N_1})$ ,  $(\gamma_{s_{jp}}^{\pi,M,N_1})$  and  $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_1})$ .
- (2.) With  $N_0 = N_1$ ,  $(\alpha_{s_{jp}}^{h,\pi,M,N_1})$ ,  $(\gamma_{s_{jp}}^{h,\pi,M,N_1})$ ,  $(\tilde{\gamma}_{s_{jp}}^{h,\pi,M,N_1})$  and  $(\alpha_{s_{jp}}^{\pi,M,N_1})$ ,  $(\gamma_{s_{jp}}^{\pi,M,N_1})$ ,  $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_1})$  give rise to a terminal condition  $F^{\pi,M,N_1}$  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  $F^{\pi,M,N_1}$ . Store the corresponding  $(\gamma_{s_{jp}}^{\pi,M,N_2})$  and  $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_2})$ .
- (3.a.) Simulate  $N_3$  (outer simulation) copies of  $(X_{s_{jp}}^n)$ . Simulate, for every  $n, j, p$ ,  $N_4$  additional (inner simulation) copies of  $(X_{s_{jp}}^n)$ , to eventually compute, with  $(\gamma_{s_{jp}}^{h,\pi,M,N_1})$  and  $(\tilde{\gamma}_{s_{jp}}^{h,\pi,M,N_1})$  at hand from the previous Step (1.),  $N_3$  copies of  $U^{upper,h,N_4,n}$ .
- (3.b.) With  $(\gamma_{s_{jp}}^{\pi,M,N_2})$  and  $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_2})$  at hand from the previous Step (2.), simulate  $N_3$  copies of  $\frac{dQ^{g,\pi}}{dP}$ . Furthermore, with  $(\alpha_{s_{jp}}^{\pi,M,N_1})$ ,  $(\gamma_{s_{jp}}^{\pi,M,N_1})$  and  $(\tilde{\gamma}_{s_{jp}}^{\pi,M,N_1})$  at hand from the previous Step (1.), simulate  $N_3$  copies of  $F^{\pi,M,N_1}$ . Using (5.19), 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 first main result, then, reads as follows:



**Theorem 16** *The primal estimator  $V_0^{*,\pi,M,N_1}$  and both the dual estimators  $V_0^{\pi,M,N_2}$  and  $\tilde{V}_0^{upp,N_3}$  converge to the upper Snell envelope, i.e.,*

$$\begin{aligned} \lim_{\pi \rightarrow 0} \lim_{M \rightarrow \infty} \lim_{N_i \rightarrow \infty, i=0,1} V_0^{*,\pi,M,N_1} &= \lim_{\pi \rightarrow 0} \lim_{M \rightarrow \infty} \lim_{N_i \rightarrow \infty, i=0,1,2} V_0^{\pi,M,N_2} \\ &= \lim_{\pi \rightarrow 0} \lim_{M \rightarrow \infty} \lim_{N_i \rightarrow \infty, i=0,\dots,4} \tilde{V}_0^{upp,N_3} = V_0^*. \end{aligned}$$

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.,  $E \left[ \tilde{V}_0^{upp,N_3} \right] \geq V_0^*$ .

**Remark 17** *The limits in Theorem 16 need to be taken consecutively. That is, we first take the limits  $N_i \rightarrow \infty$ , then the limit  $M \rightarrow \infty$ , and finally the limit  $\pi \rightarrow 0$ . This practically complies with the recipe of first choosing a sufficiently fine  $\pi$ , next choosing an appropriate selection of basis functions, and finally running a large number of Monte Carlo simulations. Of course, the ‘optimal trade-off’ between  $\pi$ , the selection of basis functions, and the size of the Monte Carlo sample may be investigated further, but this is beyond the present scope.*

## 6 Optimal Stopping Times

For ease of exposition, we assume throughout this section that  $t_j = j$ . We denote by  $C_j(X_j)$  the robust expected continuation value of the optimal stopping problem at time  $j$  given the Markov process is at  $X_j := X_{t_j}$ . Note that any optimal ‘stop or continue policy’ can be identified with a sequence of stopping times  $(\tau_j)_{j=0,\dots,T}$  which recursively satisfy  $\tau_T = T$  and

$$\tau_j = \begin{cases} j, & \text{if } \pi(j, X_j) > C_j(X_j), \\ \tau_{j+1}, & \text{if } \pi(j, X_j) < C_j(X_j). \end{cases} \quad (6.1)$$

On the set  $\{\pi(j, X_j) = C_j(X_j)\}$ ,  $\tau_j$  is not uniquely defined, but satisfies  $\tau_j \geq j$ . For example, both alternative definitions

$$\tau_j = \begin{cases} j, & \text{if } \pi(j, X_j) = C_j(X_j), \\ \text{or} \\ \tau_{j+1}, & \text{if } \pi(j, X_j) = C_j(X_j), \end{cases}$$

combined with (6.1) defines an optimal strategy. Of course the expected continuation value  $C_j(X_j)$  is typically not known explicitly.

From the functions  $C_j^n(x) := v^{*,\pi,M,N_1}(j, x)$  for  $x \in \text{support}(X_j)$  (which are the approximations known from our algorithm) we can define approximations  $(\tau_j^n)$  to an optimal strategy by setting  $\tau_T^n = T$  and

$$\tau_j^n = \begin{cases} j, & \text{if } \pi(j, X_j) > C_j^n(X_j), \\ \tau_{j+1}^n, & \text{else.} \end{cases} \quad (6.2)$$

By Theorem 16 we have that  $C_j^n(X_j) \xrightarrow{P} C_j(X_j)$ . Note that this convergence takes place on the probability space carrying  $X$  and the *independent* simulations,  $\tilde{X}$  say, needed for the construction of the  $C_j^n(x)$ . We now state the following theorem, which is our second main result:

**Theorem 18** *We have the following:*

- (i): For each  $j \in \{0, \dots, T-1\}$ ,  $\tau_j^n \xrightarrow{P} \tau_j$  on  $\pi(j, X_j) > C_j(X_j)$ .
- (ii): If  $\pi(j, X_j) \neq C_j(X_j)$  a.s. for all  $j \in \{0, \dots, T-1\}$ , then  $\tau_j^n \xrightarrow{P} \tau_j$  for all  $j \in \{0, \dots, T\}$ .
- (iii): For each  $n$  there exists an optimal stopping family  $(\tau_j^{*,n})$  such that

$$P[\tau_j^n \neq \tau_j^{*,n}] \rightarrow 0, \quad \text{for } j = 0, \dots, T, \quad \text{if } n \rightarrow \infty.$$

The definitions (6.1), (6.2), and (A.13) imply the stopping times

$$\begin{aligned} \tau &:= \tau_0, \\ \tau^n &:= \tau_0^n = \min \{j : \tau_j^n = j\}, \\ \tau^{*,n} &:= \tau_0^{*,n} = \min \{j : \tau_j^{*,n} = j\}, \end{aligned}$$

respectively. Thus, by taking  $j = 0$  in the statements of Theorem 18, we obtain the respective statements in terms of the optimal stopping times  $\tau, \tau^{*,n}$ , and the approximate stopping time  $\tau^n$ . Hence, we obtain the following corollary:

**Corollary 19** *There exist optimal stopping times  $(\tau^{*,n})$  such that*

$$P[\tau^n \neq \tau^{*,n}] \rightarrow 0, \quad \text{if } n \rightarrow \infty.$$

*In particular, if the optimal stopping time  $\tau$  is unique, then  $\tau^n \xrightarrow{P} \tau$ .*

## 7 Numerical Examples

In this section, we present numerical results obtained by applying our algorithm to a few examples of robust optimal stopping problems. We have shown in Theorem 16 that our algorithm yields both a genuine upper bound for finite  $N_i$ ,  $i = 0, \dots, 4$ , and estimates that converge asymptotically. We consider two stochastic processes,  $X_i$ ,  $i = 1, 2$ , with dynamics (cf. (2.8))  $\frac{dX_t^i}{X_t^i} = \mu^i dt + \sigma^i dW_t^i + J^i d\tilde{N}_t^i$ ,  $X_0^i = x_0^i$ , 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 7.1 and 7.2, we consider the optimal (simple) reward problem (i.e.,  $h \equiv 0$ ). We first analyze in Section 7.1 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 7.2 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  $\mu^i$  under the reference model is equal to  $\rho - \delta$  (for  $i = 1, 2$ ), where  $\rho$  represents the risk-free rate and  $\delta$  the dividend rate. In these sections, we deal specifically with simple rewards of the form  $\Pi(t, X_t) = \exp(-\rho t) (X_t - K)^+$ , or  $\Pi(t, X_t) = \exp(-\rho t) (K - X_t)^+$ , or  $\Pi(t, X_t) = \exp(-\rho t) (\max \{X_t^1, X_t^2\} - K)^+$ , where we write  $X_t = (X_t^1, X_t^2)$  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 7.3, we analyze the optimal entrance problem, with non-trivial  $h$ . There, we assume that  $\Pi$  and  $h$  are given by  $\Pi(t, X_t) = -\exp(-\rho t)\kappa$  and  $h(t, X_t) = \exp(-\rho t)(X_t - \xi)$ , 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  $m^M, \psi^M$  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.

## 7.1 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:  $\rho = 0.05$ ,  $\delta = 0.1$ ,  $\sigma = 0.2$ ,  $K = 100$ ,  $T = 3$  years. Furthermore, we consider exercise dates given by  $t_j = \frac{jT}{9}$ ,  $j = 0, \dots, 9$ . For the choice of basis functions, we follow Andersen and Broadie [1] by including still-alive European options and corresponding option deltas. For Step (2.), the basis functions  $\psi_t^M$  are enlarged by the martingale and maximum processes, as included in the Markov process  $\mathcal{X}$  (defined in Step (2.); see Section 5.2). This applies to both the univariate and the bivariate cases.

### 7.1.1 Univariate Case

In the univariate case, we restrict attention to the simple reward  $\Pi(t, X_t) = \exp(-\rho t)(X_t - K)^+$ . It is referred to as the discounted payoff of a Bermudan call option, although it may just as well be interpreted as the exercising payoff of other applications of optimal stopping for which the reward function takes this form. Let  $E_\Pi(t, X_t, T)$  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

$$\{1, \text{Pol}_2(X_t), \text{Pol}_3(E_\Pi(t, X_t, t_{j+1})), \text{Pol}_3(E_\Pi(t, X_t, t_L))\}. \quad (7.1)$$

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

$$\left\{1, X_t \frac{\partial E_\Pi(t, X_t, t_{j+1})}{\partial X_t}, X_t \frac{\partial E_\Pi(t, X_t, t_L)}{\partial X_t}\right\}. \quad (7.2)$$

We consider a fine grid  $\{s_{jp}\}$  with  $\Delta_{jp} = s_{j(p+1)} - s_{jp} = 1/1,500$ . 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.).

**Kullback-Leibler divergence:** First, we consider the case of the Kullback-Leibler divergence for different values of its parameter  $\alpha$ . 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.

$x_0$	$\alpha$				
	10	100	$10^4$	$10^6$	$\infty$
90	2.4405	4.0546	4.4049	4.4088	4.4088
	<i>2.4932</i>	<i>4.0673</i>	<i>4.4662</i>	<i>4.4708</i>	<i>4.4708</i>
	(0.0003)	(0.0009)	(0.0013)	(0.0013)	(0.0013)
100	4.6077	7.4023	7.9848	7.9913	7.9914
	<i>4.8251</i>	<i>7.3887</i>	<i>8.0328</i>	<i>8.0402</i>	<i>8.0403</i>
	(0.0005)	(0.0012)	(0.0018)	(0.0019)	(0.0019)
110	10.0281	12.2732	13.1574	13.1661	13.1662
	<i>10.1468</i>	<i>12.2934</i>	<i>13.1656</i>	<i>13.1756</i>	<i>13.1757</i>
	(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  $\alpha$  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 case  $\alpha = \infty$  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  $\alpha$  we observe an, initially rapid, increase in the robust call option's value. In general, we observe that Bermudan call option values may decrease substantially when ambiguity is taken into account. This is also illustrated graphically in Figure 1 where we plot the genuine upper bound  $\tilde{V}_0^{upp, N_3}$  as a function of  $x_0$  and the Kullback-Leibler divergence parameter  $\alpha$ . Occasionally, one may encounter that the genuine upper bound, which is biased high in the sense of expected values and not path-wise, is slightly smaller than the approximated upper bound. 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.

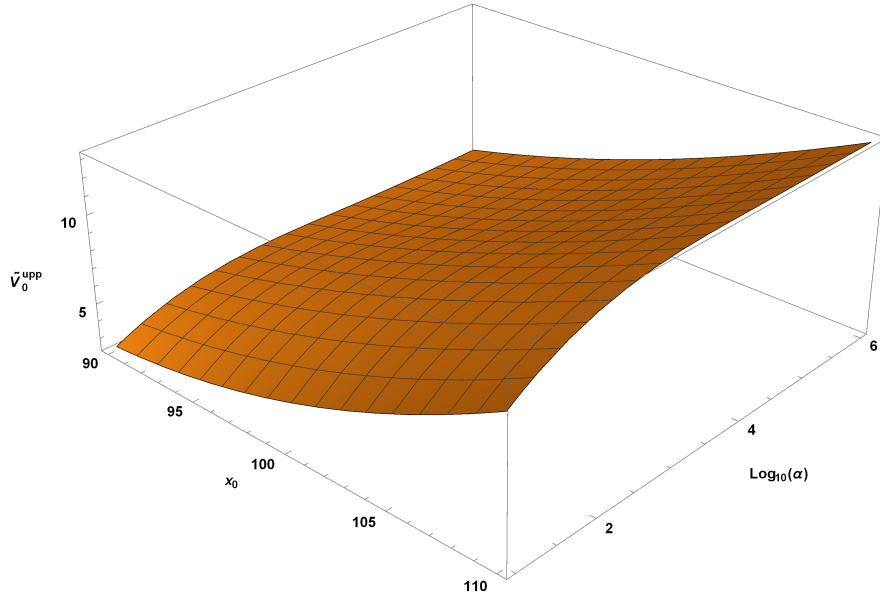
**Worst case with mean partially known:** Next, we consider the example of worst case with mean partially known, where we either take  $\mu^- = -0.05$  and vary  $\mu^+$  or we take  $\mu^+ = -0.05$  and vary  $\mu^-$ . Furthermore, we choose large values for the parameters  $B^+$  and  $B^-$  such that 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.

$x_0$	$\mu^+$		$\mu^+ = \mu^-$	$\mu^-$	
	0.05	0	-0.05	-0.1	-0.15
	$(\mu^- = -0.05)$			$(\mu^+ = -0.05)$	
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  $\mu^+$  and  $\mu^-$  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  $\mu^+$  for given  $\mu^-$ . By contrast, the robust call option values are quite sensitive to changes in  $\mu^-$  for

Figure 1: Genuine upper bound  $\tilde{V}_0^{upp, N_3}$  as a function of  $x_0$  and the Kullback-Leibler divergence parameter  $\alpha$ .



given  $\mu^+$ . 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 6 (2)) 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.

The insensitivity of the results in Table 2 to changes in  $\mu^+$  for given  $\mu^-$  is not surprising because the expected payoff is monotone in  $\mu$ , so that the worst case measure in this situation is simply the one inducing the drift  $\mu^-$  for the wealth process. In the (current) univariate example without jumps this measure is, in fact, unique so that in this specific situation we could also have computed the robust prices using simpler standard techniques. However, we will also consider worst case with mean partially known examples with two assets or with jumps, in which situations, even for an expected payoff that is monotone in  $\mu$ , the measures inducing drifts  $\mu^+$  or  $\mu^-$  are no longer unique, so that a method such as that developed in this paper is really required.

### 7.1.2 Bivariate case

Next, we consider the bivariate case. In the bivariate case, we analyze the simple reward  $\Pi(t, X_t) = \exp(-\rho t) (\max \{X_t^1, X_t^2\} - K)^+$ , referred to as the discounted payoff of a Bermudan max-call option. We denote the price of a European max-call option at time  $t$  with maturity time  $T$  by  $E_\Pi(t, X_t, T)$ . It is given by the following expression (Johnson [63]):

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

$$- K e^{-\rho(T-t)} + K e^{-\rho(T-t)} \prod_{l=1}^2 \left(1 - \mathcal{N}(d_-^l)\right),$$

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

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

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 (7.1)). For the Brownian motion driven part, we have to adapt to the two-dimensionality of our problem, and for  $\psi_t^M$  we now consider the set

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

The parameters are chosen as in the univariate case, with common  $\mu^i$  and  $\sigma^i$  for  $i = 1, 2$ , and assuming independence between  $W^1$  and  $W^2$ . We consider a grid with  $\Delta_{jp} = s_{j(p+1)} - s_{jp} = 1/150$ . Our results are based on 1,000 simulated trajectories for the calculation of the regression coefficients in Step (1.b.) and 5,000 simulated trajectories for the  $U$ -martingale increments in Step (1.c.) and the approximated upper bound to  $V^*$  in Step (2.).

**Kullback-Leibler divergence:** In Table 3, we consider the Kullback-Leibler divergence for different values of its parameter  $\alpha$ .

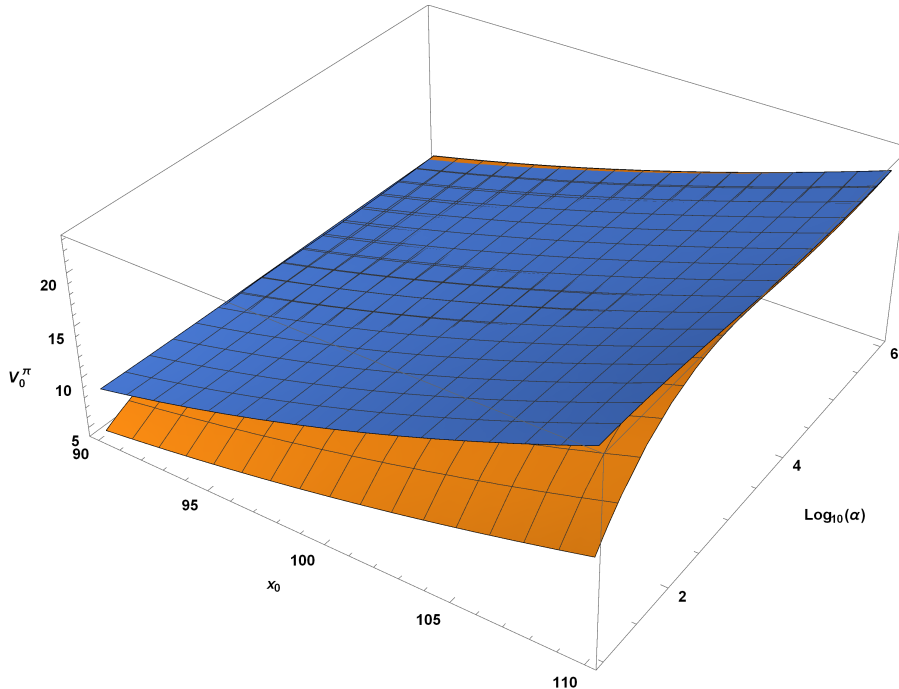
$x_0$	$\alpha$					BBS
	10	100	$10^4$	$10^6$	$\infty$	
90	4.49	7.49	8.12	8.12	8.12	8.09
100	8.12	12.98	13.97	13.98	13.98	13.96
110	13.41	20.12	21.43	21.45	21.45	21.46

Table 3: Upper bounds to robust max-call option prices using the Kullback-Leibler divergence with different values of its parameter  $\alpha$  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 [9] (BBS).

We observe again that with an increase in  $\alpha$ , 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 we have reference values, given e.g., in Belomestny, Bender and Schoenmakers [9] (BBS). For  $g \equiv 0$ , our values are very close to the upper bounds obtained by BBS. We illustrate this graphically in Figure 2, in which we plot the approximated upper bound  $V_0^{\pi, M, N_2}$  as a function of  $x_0$  and the Kullback-Leibler divergence parameter  $\alpha$  along with the surface given by the upper bounds obtained by BBS as a function of  $x_0$  and with  $g \equiv 0$  (and therefore this surface is constant in  $\alpha$ ).

**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  $\mu^+$  and  $\mu^-$ .

Figure 2: Approximated upper bound  $V_0^{\pi, M, N_2}$  as a function of  $x_0$  and the Kullback-Leibler divergence parameter  $\alpha$  (in orange) along with the surface given by the upper bounds obtained by Belomestny, Bender and Schoenmakers [9] as a function of  $x_0$  and with  $g \equiv 0$  (in blue).



	$\mu^+$	$\mu^+ = \mu^-$		$\mu^-$
	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  $\mu^+$  and  $\mu^-$  and depending on the common initial value of the underlying risky assets' prices  $x_0$ . Bivariate case. In the fourth column we display reference values as obtained by Belomestny, Bender and Schoenmakers [9] (BBS) for the case of  $\mu^+ = \mu^-$  (or  $g \equiv 0$ ).

Similar to the univariate case, the robust max-call option values are monotone in  $\mu$  so that the worst case measure is one of the measures inducing  $\mu^-$ . More specifically, the worst case measure corresponds to the stochastic drift  $(q^1, q^2)$  that satisfies  $\sigma^1 q^1 + \sigma^2 q^2 = \mu^- - \mu$  such that the maximum of the optimization problem in (3.14) (without the  $\lambda$  component) is attained for  $(t, \omega)$ . Of course, the optimization problem (3.14) depends in turn on the values of the components of the vector  $z$ . The components of this vector correspond, for each  $(t, \omega)$ , to the local covariance per time unit between the different components of an increment of a Brownian motion and the future values of our upper Snell envelope over one incremental time unit. The case of  $\mu^+ = \mu^-$  pertains to 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.

## 7.2 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:  $\rho = 0.04$ ,  $\delta = 0$ ,  $\sigma = 0.2$ ,  $J = 0.06$ ,  $K = 100$ ,  $T = 1$  year, and consider different values of  $\lambda_P$ . The exercise dates are given by  $t_j = \frac{jT}{10}$ ,  $j = 0, \dots, 10$ , and the fine grid is given by  $\Delta_{jp} = 1/100$ .

In this subsection, we analyze the simple reward  $\Pi(t, X_t) = \exp(-\rho t) (K - X_t)^+$ , referred to as the discounted payoff of a Bermudan put option. 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 [34]):

$$E_{\Pi}(t, X_t, T) = e^{-\rho(T-t)} \sum_{n \geq 0} \frac{e^{-\lambda_P(T-t)} (\lambda_P(T-t))^n}{n!} \mathcal{BS}(T-t, X_t^{(n)}, \sigma), \quad (7.3)$$

where  $X_t^{(n)} = X_t \exp(nJ - \lambda_P(T-t) \exp(J) + \lambda_P(T-t))$ , and where  $\mathcal{BS}$  denotes the Black-Scholes price of the corresponding European put option.<sup>6</sup>

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 (7.1), 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,  $\psi_t^M$ , and the jump part,  $\tilde{\psi}_t^M$ , are both given by (7.2). Our numerical results are based on 5,000 simulated trajectories for all *relevant* steps of the algorithm. For Step (2.),  $\psi_t^M$  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  $\alpha$  and of the jump intensity  $\lambda_P$ .

		$\alpha$					
		$x_0$	10	100	$10^4$	$10^6$	$\infty$
$J = \lambda_P = 0$	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  $\alpha$  and of the jump intensity  $\lambda_P$ , and depending on the initial value of the underlying risky asset's price  $x_0$ . Univariate case.

<sup>6</sup>The formula given in Cont and Tankov [34] 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 (7.3), which involves an infinite sum, but converges very rapidly.



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  $\delta_1$  and  $\delta_2$ . These are given in Table 6.

	$x_0$	$\delta_1 = \delta_2 = 0.5$	$\delta_1 = 0.5, \delta_2 = 1$	$\delta_1 = 1, \delta_2 = 0.5$
$\lambda_P = 1$	90	10.48	10.37	10.20
	100	4.22	3.92	3.13
	110	1.73	1.54	1.05
$\lambda_P = 3$	90	10.90	10.74	10.43
	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  $\delta_1$  and  $\delta_2$  and of the jump intensity  $\lambda_P$ , 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  $\delta_1$ ) and to the jump intensity (as measured by  $\delta_2$ ).

**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.

	$x_0$	$\frac{\mu^+}{0.14}$ ( $\mu^- = 0.04$ )	$\frac{\mu^+ = \mu^-}{0.04}$	$\frac{\mu^-}{-0.06}$ ( $\mu^+ = 0.04$ )
$J = \lambda_P = 0$	90	10.48	11.83	11.83
	100	4.26	6.41	6.41
	110	1.74	3.22	3.22
$J = 0.06, \lambda_P = 1$	90	10.62	12.05	12.05
	100	4.57	6.67	6.68
	110	1.95	3.44	3.45
$J = 0.06, \lambda_P = 3$	90	11.10	12.85	12.85
	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  $\mu^+$  and  $\mu^-$  and of the jump intensity  $\lambda_P$ , 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.

### 7.3 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  $t_j = j\Delta^c$ ,  $j = 0, \dots, T/\Delta^c$ , 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  $\Delta^c$ . We use the following parameter set under the reference model:  $\mu = 0$ ,  $\rho = 0.1$ ,  $\sigma = 0.1$ ,  $\xi = 1$ ,  $\kappa = 1$ ,  $T = 100$  years.

In Steps (1.a.) and (1.b.) of our algorithm, we choose for  $m_t^M$  the set of basis functions given by  $\{1, \text{Pol}_3(X_t), \text{Pol}_3(h(t, X_t))\}$ . The basis functions for the Brownian motion driven part of the BSDE are given by the set

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

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 8, we present results for different values of  $\Delta^c$  and  $x_0$ . The results in the second and third columns of Table 8 can be viewed as rough approximations to the continuous-time optimal entrance problem with infinite time horizon as considered, for example, in Dixit [42], where  $1/\Delta^c = T = \infty$ . The last column in Table 8 displays the corresponding values obtained by Dixit [42].

$x_0$	$1/\Delta^c$		$1/\Delta^c = T = \infty$ (Dixit)
	1	10	
1	0.79	0.77	0.56
1.375	3.22	3.01	2.75
1.5	4.47	4.26	4.25

Table 8: 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 [42], for the parameter set considered here.

**Kullback-Leibler divergence:** Next, we consider the Kullback-Leibler divergence for different values of  $\alpha$ , taking  $1/\Delta^c = 10$ . The results are in Table 9.

$x_0$	$\alpha$				
	10	100	$10^4$	$10^6$	$\infty$
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 9: Upper bounds to robustly evaluated rewards using the Kullback-Leibler divergence with different values of its parameter  $\alpha$  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 8. Robustly evaluated rewards appear to be fairly sensitive to changes in  $\alpha$ , at moderate levels of  $\alpha$ , 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 10.

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

We observe from Table 10 that the robustly evaluated rewards are monotone in  $\mu$  so that the worst case measure corresponds to one of the measures inducing a drift  $\mu^-$  for the wealth process; a pattern consistent with Table 4. Again, the case of  $\mu^+ = \mu^-$  yields the case of a standard conditional expectation, and agrees with the last column of Table 9 (as well as the third column in Table 8). As explained in Section 7.1.1, 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.

## 7.4 CPU Times and Accuracy

To assess the numerical efficiency of our algorithm, we display in Table 11 the CPU times for two of the examples we analyze. We decompose the CPU times per step in the algorithm. The computations have been performed on a machine with the following infrastructure: HP BL680c G7 4xXeon, Ten-Core 2400MHz, and 1024 GB RAM. The SDE's involved are pre-simulated. The associated CPU times are displayed in the second row of Table 11. For the Kullback-Leibler divergence example in Section 7.1.1 (univariate diffusion), we analyze CPU times for both 1,000 and 5,000 simulated trajectories for Step (1.b.), and 5,000 trajectories for all remaining steps, taking  $\Delta_{jp} = 1/150$ . For the Kullback-Leibler divergence example in Section 7.2 (univariate jump-diffusion), we simulate 5,000 trajectories for all steps and take  $\Delta_{jp} = 1/100$  (as in Table 5).

	KL (Section 7.1.1)	KL (Section 7.1.1)	KL (Section 7.2)
SDE simulation	1	1	1
Step 1	32	94	73
Step 1b	14	74	57
Step 1c	18	20	16
Step 2	80	82	55
Total	112	176	128

Table 11: CPU times in seconds. For the Kullback-Leibler divergence example in Section 7.1.1 (univariate diffusion), we analyze CPU times for both 1,000 (second column) and 5,000 (third column) simulated trajectories for Step (1.b.), and 5,000 trajectories for all remaining steps, taking  $\Delta_{jp} = 1/150$ . For the Kullback-Leibler divergence example in Section 7.2 (univariate jump-diffusion), we simulate 5,000 trajectories for all steps and take  $\Delta_{jp} = 1/100$ .

The results in Table 11 reveal that a full-fledged run of the algorithm takes only about two to three minutes. A decomposition per step is instructive. We observe that, with 5,000 trajectories for all steps, Step (1.) takes about 55% of the total CPU time, with (1.b) accounting for 40% to 45% and (1.c.) for 10% to 15%, and Step (2.) accounts for about 45%. Step (1.) can be reduced to about 30% of the total CPU time, if the number of simulations in Step (1.b.) is reduced to 1,000, inducing a reduction in CPU time for this step by a factor 5.

The results for the Kullback-Leibler divergence example in Section 7.1.1 displayed in Table 1 are based on a large number of simulations (10,000 for all steps) and a very fine grid ( $\Delta_{jp} = 1/1,500$ ), to ensure high accuracy. The results associated to the second and third columns of Table 11 are based on lower numbers of simulations and a coarser grid, but differ uniformly by (typically much) less than 1% from the results in Table 1. This illustrates that, already with a quite limited number of Monte Carlo simulations and a reasonably coarse grid, our method yields rather accurate estimates in realistic settings.

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. The numerical results also reveal the potentially relevant and significant impact of taking ambiguity into account when evaluating optimal stopping strategies.

## 8 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 that converges asymptotically to the true solution, yields a safe genuine (biased high) upper bound at the pre-limiting level, and provides asymptotically optimal exercise rules. 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 7:** The proof of existence in principle follows from general existence results in the seminal work by El Karoui [46]. For convenience of the reader we provide a short proof in our specific setting. By time-consistency of  $U$ , a property that is preserved with respect to stopping times, i.e., for any stopping time  $\tau$  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 as a consequence of Theorem 3.2 in Krätschmer and Schoenmakers [66]. While [66] is set in discrete-time, and our jump-diffusion setting is a continuous-time setting (the stochastic drivers generating the Markov process are continuous-time processes), the discretization of our underlying continuous-time processes can simply be seen as the discrete-time process in [66]. Furthermore, (3.2) follows as a consequence of Theorem 3.4 in [66] and Proposition 7 is a consequence of Theorem 5.4 in the same [66]. ■

**Proof of Theorem 9:** For a square-integrable  $H$  that is  $\mathcal{F}_T$ -adapted and  $t \in [0, T]$ , let us consider

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

where we will later, for the first part of the proof, set  $H = 0$  so that  $\tilde{U}^h = U^h$ ; cf. (3.3). By time-consistency, for  $t \in (t_j, t_{j+1}]$ ,

$$\tilde{U}_t^h = \inf_{(q, \lambda) \in C} \left\{ \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h | \mathcal{F}_t \right] + c_t(Q) \right\}. \quad (\text{A.2})$$

Specifically, we have

$$\begin{aligned} \tilde{U}_t^h &= \inf_{(q, \lambda) \in C} \left\{ \mathbb{E}_Q [\tilde{U}_{t_{j+1}}^h | \mathcal{F}_t] + c_t(Q) \right\} \\ &= \inf_{(q, \lambda) \in C} \left\{ \mathbb{E} \left[ \mathcal{E} \left( (q \cdot W)_{t, T} + (\lambda \cdot \tilde{N})_{t, T} \right) \left( \tilde{U}_{t_{j+1}}^h + \int_t^T r(s, q_s, \lambda_s - \lambda_P) ds \right) \middle| \mathcal{F}_t \right] \right\} \\ &= \inf_{(q, \lambda) \in C} \left\{ \mathbb{E} \left[ \mathcal{E} \left( (q \cdot W)_{t, t_{j+1}} + (\lambda \cdot \tilde{N})_{t, t_{j+1}} \right) \left( \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds \right) \middle| \mathcal{F}_t \right] \right\} \\ &= \inf_{(q, \lambda) \in C} \left\{ \mathbb{E}_Q [\tilde{U}_{t_{j+1}}^h | \mathcal{F}_t] + c_{t, t_j}(Q) \right\}, \end{aligned} \quad (\text{A.3})$$

where  $c_{t, t_j}(Q) := \mathbb{E}_Q \left[ \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds | \mathcal{F}_t \right]$  and where  $\mathcal{E} \left( (q \cdot W)_{t, T} + (\lambda \cdot \tilde{N})_{t, T} \right)$  is the Doléans-Dade exponential given by the right-hand side of Eqn. (2.5) with integration bounds from  $t$  to  $T$ . The third equation holds as  $r(s, 0, 0) = 0$  and as the first term in the expectation only depends on  $(q_s, \lambda_s)_{t \leq s \leq t_{j+1}}$ , so that  $(q_s, \lambda_s)$  can be chosen to be  $(0, 0)$  for  $t_{j+1} < s \leq T$ .

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

$$d\tilde{U}_t^h = -g(t, Z_t, \tilde{Z}_t) dt + Z_t dW_t + \tilde{Z}_t d\tilde{N}_t, \quad \text{for } t \in (t_j, t_{j+1}], \quad (\text{A.4})$$

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.4) may be seen as a solution to a BSDE. To be more precise, by Tang and Li [91], 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

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

where we denote by  $\mathcal{S}^2$  the space of all càdlàg processes for which the maximum is square-integrable. Later we shall prove the following auxiliary result:

**Lemma 20** *For every  $t \in [t_j, t_{j+1}]$ ,*

$$Y_t = \inf_{(q, \lambda) \in C} \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds \mid \mathcal{F}_t \right]. \quad (\text{A.5})$$

Furthermore, there exists a measure  $Q^g$  with corresponding  $(q^g, \lambda^g) \in C$  such that, for every  $t \in [t_j, t_{j+1}]$ ,

$$Y_t = \mathbb{E}_{Q^g} \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s^g, \lambda_s^g - \lambda_P) ds \mid \mathcal{F}_t \right].$$

Note that the measure  $Q^g$  can be chosen independently of  $t \in [t_j, t_{j+1}]$ .

Let us first show how to conclude Theorem 9, part (a), from Lemma 20. Note that by (A.3), the right-hand side of (A.5) is equal to  $\tilde{U}_t^h$  for  $t \in (t_j, t_{j+1}]$ . Now setting  $H = 0$  in (A.1) and using Lemma 20 for  $t \in (t_j, t_{j+1}]$  we obtain (3.9). The second part of (a), i.e., (3.10), is seen similarly by setting  $h = 0$  in (A.1). As  $(Y_t)_{t \in [t_j, t_{j+1}]}$  is continuous in  $t$ , we can define  $\tilde{U}_{t_j+}^h = \lim_{t \downarrow t_j} \tilde{U}_t^h = \lim_{t \downarrow t_j} Y_t$ . It is left to show that  $\tilde{U}_{t_j}^h = \tilde{U}_{t_j+}^h + h(t_j, X_{t_j})$ . Given  $(q, \lambda) \in C$ , define  $U_t^{h, Q} := \mathbb{E}_Q \left[ H + \sum_{t \leq t_k} h(t_k, X_{t_k}) \mid \mathcal{F}_t \right] + c_t(Q)$ . Then,

$$U_{t_j}^{h, Q} = \lim_{t \downarrow t_j} U_t^{h, Q} + h(t_j, X_{t_j}),$$

and, taking the infimum over  $(q, \lambda) \in C$ ,

$$\begin{aligned} \tilde{U}_{t_j}^h &= \inf_{(q, \lambda) \in C} U_{t_j}^{h, Q} = \inf_{(q, \lambda) \in C} \lim_{t \downarrow t_j} U_t^{h, Q} + h(t_j, X_{t_j}) \\ &\geq \lim_{t \downarrow t_j} \inf_{(q, \lambda) \in C} U_t^{h, Q} + h(t_j, X_{t_j}) = \tilde{U}_{t_j+}^h + h(t_j, X_{t_j}). \end{aligned} \quad (\text{A.6})$$

Moreover, in view of Lemma 20 we may observe that

$$U_t^{h, Q^g} = Y_t = \tilde{U}_t^h, \quad \text{for } t \in (t_j, t_{j+1}].$$

Hence,

$$\inf_{(q, \lambda) \in C} \lim_{t \downarrow t_j} U_t^{h, Q} + h(t_j, X_{t_j}) \leq \lim_{t \downarrow t_j} U_t^{h, Q^g} + h(t_j, X_{t_j}) = \lim_{t \downarrow t_j} \inf_{(q, \lambda) \in C} U_t^{h, Q} + h(t_j, X_{t_j}). \quad (\text{A.7})$$

Now, putting (A.6) and (A.7) together, we may conclude  $\tilde{U}_{t_j}^h = \tilde{U}_{t_j+}^h + h(t_j, X_{t_j})$ . Thus, it remains to show Lemma 20.

**Proof of Lemma 20:** Let  $Q \in \mathcal{Q}$ . We write

$$\begin{aligned}
Y_t &= \mathbb{E}_Q[Y_t | \mathcal{F}_t] \\
&= \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} g(s, Z_s, \tilde{Z}_s) ds - \int_t^{t_{j+1}} Z_s dW_s - \int_t^{t_{j+1}} \tilde{Z}_s d\tilde{N}_s \mid \mathcal{F}_t \right] \\
&= \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} \left[ -q_s Z_s - \tilde{Z}_s (\lambda_s - \lambda_P) + g(s, Z_s, \tilde{Z}_s) \right] ds \right. \\
&\quad \left. + \int_t^{t_{j+1}} Z_s dW_s^Q + \int_t^{t_{j+1}} \tilde{Z}_s d\tilde{N}_s^Q \mid \mathcal{F}_t \right] \\
&= \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} \left[ -q_s Z_s - \tilde{Z}_s (\lambda_s - \lambda_P) + g(s, Z_s, \tilde{Z}_s) \right] ds \mid \mathcal{F}_t \right] \\
&\leq \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds \mid \mathcal{F}_t \right], \tag{A.8}
\end{aligned}$$

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  $(q, \lambda)$  take values in a compact set and  $Y$  is square-integrable under  $P$ . The third and fourth equalities hold because  $\int_t^t Z_s dW_s^Q$  and  $\int_t^t \tilde{Z}_s d\tilde{N}_s^Q$  are well-defined martingales, since for any  $Q$  with  $(q, \lambda)$  in a compact bounded set we have, again by Cauchy-Schwarz,

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

and a similar argument holds for  $\tilde{Z}$ . It follows from (A.8) and the fact that we can restrict the infimum in (A.2) to  $(q, \lambda) \in C$  that

$$Y_t \leq \inf_{(q, \lambda) \in C} \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds \mid \mathcal{F}_t \right], \text{ for all } t \in [t_j, t_{j+1}].$$

Next, by a measurable selection theorem (see e.g., Beneš [13]), we choose predictable  $(q_s^g, \lambda_s^g - \lambda_P) \in \partial g(s, Z_s, \tilde{Z}_s)$ . Then,  $q^g$  and  $\lambda^g$  induce an equivalent probability measure,  $Q^g$ , with Radon-Nikodym derivative given by (2.5). Proceeding as in (A.8) with  $q^g, \lambda^g$  and  $Q^g$  (where the inequality in (A.8) becomes an equality) yields

$$Y_t = \mathbb{E}_{Q^g} \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s^g, \lambda_s^g - \lambda_P) ds \mid \mathcal{F}_t \right]. \tag{A.9}$$

Therefore, indeed  $Y_t = \inf_{(q, \lambda) \in C} \mathbb{E}_Q \left[ \tilde{U}_{t_{j+1}}^h + \int_t^{t_{j+1}} r(s, q_s, \lambda_s - \lambda_P) ds \mid \mathcal{F}_t \right]$  for all  $t \in [t_j, t_{j+1}]$ , and the infimum is attained in  $Q^g$ . ■

This completes part (a) of Theorem 9.

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

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

From (3.4), part (b) follows. ■

**Proof of Eqn. (3.15):** Throughout this proof, let us set  $C := C_t$ . We want to compute  $g(t, z, \tilde{z}) = \inf_{(q, \lambda - \lambda_P) \in C} \langle (z, \tilde{z}), (q, \lambda - \lambda_P) \rangle$ , with  $C$  given by

$$C = \left\{ (q, \lambda - \lambda_P) \mid A(q, \lambda - \lambda_P)^\top = b \quad \text{and} \quad |(q, \lambda - \lambda_P)|_* \leq \sqrt{\Lambda} \right\}.$$

We assume that the set  $C$  is non-empty. Define  $c := (z, \tilde{z})^\top$ . Then  $g(t, z, \tilde{z})$  is equal to the value of the minimization problem

$$\min c^\top x \text{ subject to } Ax = b, \quad |x|_*^2 \leq \Lambda.$$

Let  $P_B(0)$  be the projection of 0 onto the set  $B := \{x \mid Ax = b\}$  in the  $|\cdot|_*$  norm. Then the minimization problem is equivalent to

$$\min c^\top x' + c^\top P_B(0) \text{ subject to } Ax' = 0, \quad |x'|_*^2 \leq \Lambda - |P_B(0)|_*^2,$$

with  $x' = x - P_B(0)$ . So we have to solve

$$\min_{x' \in \text{Kernel}(A)} c^\top x' = \min_{x' \in \text{Kernel}(A)} P_{\text{Kernel}(A)}^\top(c) x' \text{ subject to } |x'|_* \leq \sqrt{\Lambda - |P_B(0)|_*^2}, \quad (\text{A.10})$$

where  $P_{\text{Kernel}(A)}^\top(c)$  is the projection of  $c$  onto the kernel of  $A$  in the Euclidean norm. Now it is well known that for the Euclidean norm  $|\cdot|$ , we have  $\min_{|x'| \leq \lambda} v^\top x' = -\lambda|v|$ . Since  $P_{\text{Kernel}(A)}^\top(c) \in \text{Kernel}(A)$  and the kernel of  $A$  is a linear vector space, it follows that the solution to the above optimization problem (A.10) is given by  $-\sqrt{\Lambda - |P_B(0)|_*^2} P_{\text{Kernel}(A)}^\top(c)|_{**}$ . Hence, we overall obtain

$$g(t, z, \tilde{z}) = -\sqrt{\Lambda - |P_B(0)|_*^2} \left| P_{\text{Kernel}(A)} \begin{pmatrix} z \\ \tilde{z} \end{pmatrix} \right|_{**} + \langle (z, \tilde{z}), P_B(0) \rangle.$$

■

**Proof of Eqn. (5.10):** Define the processes  $(Z, \tilde{Z})$  given by  $Z_s = z_{s_{jl}}^{\pi, M, N_1}(X_{s_{jl}})$  and  $\tilde{Z}_s = \tilde{z}_{s_{jl}}^{\pi, M, N_1}(X_{s_{jl}})$  if  $s_{jl} \leq s < s_{j(l+1)}$  for  $j = 0, \dots, L-1$ ,  $l = 0, \dots, P-1$ . Then from (5.10) it follows that

$$M_t^{g, \pi, M, N_1} = - \int_0^t g(s, Z_s, \tilde{Z}_s) ds + \int_0^t Z_s dW_s + \int_0^t \tilde{Z}_s d\tilde{N}_s.$$

That  $M^{g, \pi, M, N_1}$  is a  $U$ -martingale with terminal condition  $M_T^{g, \pi, M, N_1}$ , follows now from (3.11) or directly from the arguments given in (A.8) and (A.9). ■

**Proof of Eqn. (5.15):** We now show that our approximation scheme converges. Suppose that equations (5.1)–(5.7) hold with a square-integrable  $p$ -dimensional Markov process,  $\mathcal{X}$ , and an arbitrary function (driver)  $g : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}$  that is uniformly Lipschitz continuous in  $(z, \tilde{z})$ . The following theorem establishes convergence of our approximation scheme:

**Theorem 21** *We have that*

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



**Proof** It follows from Bouchard and Elie [21] that  $Y_t^\pi$  converges to  $Y_t$  in  $L^2$ . From this and Lemma 22 below we may conclude that it is sufficient to prove that  $Y_{T_0}^{\pi,M,N}$  converges to  $Y_{T_0}^{\pi,M}$  in  $L^2$ , which would follow if

$$\lim_{N \rightarrow \infty} Y_{T_0}^{\pi,M,N} \rightarrow Y_{T_0}^{\pi,M} \text{ in } L^2.$$

And this follows from Lemma 23 below. The proof for  $Z^{\pi,M,N}$  and  $\tilde{Z}^{\pi,M,N}$  is similar. ■

**Lemma 22** *For every  $t \in [T_0, T_1]$  and for fixed  $\pi$ , we have that  $Y_t^{\pi,M} \rightarrow Y_t^\pi$ ,  $Z_t^{\pi,M} \rightarrow Z_t^\pi$  and  $\tilde{Z}_t^{\pi,M} \rightarrow \tilde{Z}_t^\pi$  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  $Y_{s_{jp}}^{\pi,M} \rightarrow Y_{s_{jp}}^\pi$ ,  $Z_{s_{jp}}^{\pi,M} \rightarrow Z_{s_{jp}}^\pi$  and  $\tilde{Z}_{s_{jp}}^{\pi,M} \rightarrow \tilde{Z}_{s_{jp}}^\pi$  in  $L_1^2$ ,  $L_d^2$  and  $L_k^2$ , respectively. Since our basis functions span the entire space,  $L_1^2(\mathcal{F}_{s_{jp}})$ , the lemma clearly holds for  $s_{jp} = T$ . (Without loss of generality we may set  $Z_{T_1}^{\pi,M} = Z_{T_1}^\pi$  and  $\tilde{Z}_{T_1}^{\pi,M} = \tilde{Z}_{T_1}^\pi$ .) It will be useful to consider the projection onto the span of  $\psi^M(s_{jp}, \mathcal{X}_{s_{jp}}^\pi)$  and  $\tilde{\psi}^M(s_{jp}, \mathcal{X}_{s_{jp}}^\pi)$ , say  $\hat{P}^{\pi,M}$  and  $\tilde{\hat{P}}^{\pi,M}$ , respectively, instead of the projection onto the span of  $\psi^M(s_{jp}, \mathcal{X}_{s_{jp}}^\pi) \Delta W_{jp}$  and  $\tilde{\psi}^M(s_{jp}, \mathcal{X}_{s_{jp}}^\pi) \Delta \tilde{N}_{jp}$ , respectively. We write

$$\begin{aligned} \gamma_{s_{jp}}^{\pi,M} \psi^M(s_{jp}, \mathcal{X}_{s_{jp}}^\pi) &= \hat{P}^{\pi,M} (Y_{s_{j(p+1)}}^M \Delta W_{jp}) / \mathbb{E} [\Delta^2 W_{jp}] \\ &= \hat{P}^{\pi,M} \left( \mathbb{E}_{jp} [Y_{s_{j(p+1)}}^{\pi,M} \Delta W_{jp}] \right) / \mathbb{E} [\Delta^2 W_{jp}] \\ &\xrightarrow{M \rightarrow \infty} \mathbb{E}_{jp} [Y_{s_{j(p+1)}}^\pi \Delta W_{jp}] / \mathbb{E} [\Delta^2 W_{jp}] = Z_{s_{jp}}^\pi, \end{aligned}$$

in  $L^2$ , where we used (5.2) and (5.4) in the first equality. The convergence then follows since, by the induction assumption, we have that  $\mathbb{E}_{jp} [Y_{s_{j(p+1)}}^{\pi,M} \Delta W_{jp}]$  converges in  $L^2$  to  $\mathbb{E}_{jp} [Y_{s_{j(p+1)}}^\pi \Delta W_{jp}]$  as  $M$  tends to infinity. Similarly,

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

in  $L^2$ , where we used (5.3) and (5.5) in the first equality. The lemma is now a consequence of (5.1) and (5.6). ■

**Lemma 23** *For all  $j$ , we have that  $\alpha_{s_{jp}}^{\pi,M,N} \rightarrow \alpha_{s_{jp}}^{\pi,M}$ ,  $\gamma_{s_{jp}}^{\pi,M,N} \rightarrow \gamma_{s_{jp}}^{\pi,M}$  and  $\tilde{\gamma}_{s_{jp}}^{\pi,M,N} \rightarrow \tilde{\gamma}_{s_{jp}}^{\pi,M}$  as  $N$  tends to infinity.*

**Proof** By the Law of Large Numbers (LLN), we have that  $(A_{s_{jp}}^{\pi,M,N})$ ,  $(\bar{A}_{s_{jp}}^{\pi,M,N})$  and  $(\tilde{A}_{s_{jp}}^{\pi,M,N})$  converge to  $(A_{s_{jp}}^{\pi,M})$ ,  $(\bar{A}_{s_{jp}}^{\pi,M})$  and  $(\tilde{A}_{s_{jp}}^{\pi,M})$ , 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

$$\begin{aligned} & \frac{1}{N} \sum_{n=1}^N F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M}, \gamma_{s_{j(p+1)}}^{\pi, M}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M}, \mathcal{X}_{s_{j(p+1)}}^{\pi, n}) m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi, n}) \\ & \xrightarrow{N \rightarrow \infty} \mathbb{E}[F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M}, \gamma_{s_{j(p+1)}}^{\pi, M}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M}, \mathcal{X}_{s_{j(p+1)}}^{\pi}) m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi})]. \end{aligned}$$

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

$$\begin{aligned} & \frac{1}{N} \sum_{n=1}^N (F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M, N}, \gamma_{s_{j(p+1)}}^{\pi, M, N}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M, N}, \mathcal{X}_{s_{j(p+1)}}^{\pi, n}) \\ & \quad - F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M}, \gamma_{s_{j(p+1)}}^{\pi, M}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M}, \mathcal{X}_{s_{j(p+1)}}^{\pi, n})) m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi, n}) \\ & \leq \left( |\alpha_{s_{j(p+1)}}^{\pi, M, N} - \alpha_{s_{j(p+1)}}^{\pi, M}| + |\gamma_{s_{j(p+1)}}^{\pi, M, N} - \gamma_{s_{j(p+1)}}^{\pi, M}| + |\tilde{\gamma}_{s_{j(p+1)}}^{\pi, M, N} - \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M}| \right) \\ & \quad \times \frac{1}{N} \sum_{n=1}^N m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi, n}) \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

it follows that

$$\begin{aligned} & \frac{1}{N} \sum_{n=1}^N F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M, N}, \gamma_{s_{j(p+1)}}^{\pi, M, N}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M, N}, \mathcal{X}_{s_{j(p+1)}}^{\pi, n}) m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi, n}) \\ & \xrightarrow{N \rightarrow \infty} \mathbb{E}[F(s_{j(p+1)}, \alpha_{s_{j(p+1)}}^{\pi, M}, \gamma_{s_{j(p+1)}}^{\pi, M}, \tilde{\gamma}_{s_{j(p+1)}}^{\pi, M}, \mathcal{X}_{s_{j(p+1)}}^{\pi})] m^M(s_{jp}, \mathcal{X}_{s_{jp}}^{\pi}). \end{aligned}$$

Therefore,

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

By replacing  $\alpha_{s_{jp}}^{\pi, M, N}$  by  $\gamma_{s_{jp}}^{\pi, M, N}$ ,  $A_{s_{jp}}^{\pi, M}$  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  $\tilde{\gamma}_{s_{jp}}^{\pi, M, N}$ ,  $\bar{A}_{s_{jp}}^{\pi, M}$  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 (5.15). ■

**Proof of Proposition 14:** This follows from (A.9) in the proof of Theorem 9(a). ■

**Proof of Proposition 15:** We have

$$\begin{aligned}
& \mathbb{E} \left[ \tilde{u}_{t_j}^{upp,h,N_4}(x) \right] \\
&= \mathbb{E} \left[ \frac{1}{N_4} \sum_{n=1}^{N_4} D_{t_j}^{\pi,n}(x) \left[ \sum_{l=j}^L h(t_l, X_{t_l}^{t_j,x,n}) + \sum_{l=1}^L \sum_{p=0}^{P-1} \int_{s_{lp}}^{s_{l(p+1)}} r(s, q_{s_{lp}}^{h,\pi,t_j,x,n}, \lambda_{s_{lp}}^{h,\pi,t_j,x,n} - \lambda_P) ds \right] \right] \\
&= \mathbb{E} \left[ \frac{dQ^{g,\pi,M,N_4,t_j,x}}{dP} \left[ \sum_{l=j}^L h(t_l, X_{t_l}^{t_j,x}) + \sum_{l=1}^L \sum_{p=0}^{P-1} \int_{s_{lp}}^{s_{l(p+1)}} r(s, q_{s_{lp}}^{h,\pi,t_j,x}, \lambda_{s_{lp}}^{h,\pi,t_j,x} - \lambda_P) ds \right] \right] \\
&\geq \inf_{Q \sim P | Q=P \text{ on } \mathcal{F}_{t_j}} \left\{ \mathbb{E}_Q \left[ \sum_{l=j}^L h(t_l, X_l) + \int_{t_j}^T r(s, q_s, \lambda_s - \lambda_P) ds \middle| \mathcal{F}_{t_j} \right] \right\} = u_{t_j}^h(x),
\end{aligned}$$

where the last inequality used the fact that  $Q^{g,\pi,M,N_4,t_j,x}$  is a conditional probability measure that is absolutely continuous with respect to  $P$  and agrees with  $P$  on  $\mathcal{F}_{t_j}$ . ■

**Proof of Theorem 16:** The stated convergence results follow as a consequence of our convergence results for BSΔEs (see the *proof* of (5.15)). Next, choose a fixed  $n \in \{1, \dots, N_3\}$ . To show the biased high property, we then write

$$\begin{aligned}
\mathbb{E} \left[ \tilde{V}_0^{upp,N_3} \right] &= \mathbb{E} \left[ \frac{dQ^{g,\pi,M,N_2,n}}{dP} \mathbb{E} \left[ \max_{t_j \in \{0, \dots, T\}} \left( \Pi(t_j, X_{t_j}^n) - h(t_j, X_{t_j}^n) \right) \right. \right. \\
&\quad \left. \left. + \tilde{u}_{t_j}^{upp,h,N_4}(X_{t_j}^n) + M_T^{g,\pi,M,N_1,n} - M_{t_j}^{g,\pi,M,N_1,n} \right) \right. \\
&\quad \left. \left. + \sum_{j,p} \int_{s_{jp}}^{s_{j(p+1)}} r(s, q_{s_{jp}}^{\pi,M,N_2,n}, \lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P) ds \right. \right. \\
&\quad \left. \left. \left| W_{s_{jp}}^{n'}, N_{s_{jp}}^{n'}, X_{s_{jp}}^{n'}, p = 0, \dots, P, j = 0, \dots, L-1, n' = 1, \dots, N_3 \right] \right] \\
&\geq \mathbb{E} \left[ \frac{dQ^{g,\pi,M,N_2,n}}{dP} \left[ \max_{t_j \in \{0, \dots, T\}} \left( \Pi(t_j, X_{t_j}^n) - h(t_j, X_{t_j}^n) + u_{t_j}^h(X_{t_j}^n) + M_T^{g,\pi,M,N_1,n} - M_{t_j}^{g,\pi,M,N_1,n} \right) \right. \right. \\
&\quad \left. \left. + \sum_{j,p} \int_{s_{jp}}^{s_{j(p+1)}} r(s, q_{s_{jp}}^{\pi,M,N_2,n}, \lambda_{s_{jp}}^{\pi,M,N_2,n} - \lambda_P) ds \right] \right] \\
&\geq U_0 \left( \max_{t_j \in \{0, \dots, T\}} \left( \Pi(t_j, X_{t_j}^n) - h(t_j, X_{t_j}^n) + u_{t_j}^h(X_{t_j}^n) + M_T^{g,\pi,M,N_1,n} - M_{t_j}^{g,\pi,M,N_1,n} \right) \right) \\
&\geq U_0 \left( \max_{t_j \in \{0, \dots, T\}} \left( \Pi(t_j, X_{t_j}^n) - h(t_j, X_{t_j}^n) + u_{t_j}^h(X_{t_j}^n) + M_T^{*,n} - M_{t_j}^{*,n} \right) \right) = V_0^*,
\end{aligned}$$

where we have used Proposition 15 and Jensen's inequality in the first inequality, (5.13) in the second inequality, and Proposition 7 in the last inequality and also in the last equality. ■

**Proof of Theorem 18:** (i): Let us define  $A_j := \{\pi(j, X_j) > C_j(X_j)\}$ . Then, since by Theorem 16  $C_j^n(X_j) \xrightarrow{P} C_j(X_j)$ , it holds by Lemma 24 below that

$$\begin{aligned}
P[\tau_j^n = j, A_j] &= P[C_j^n(X_j) - \pi(j, X_j) < 0, C_j(X_j) - \pi(j, X_j) < 0] \\
&\rightarrow P[C_j(X_j) < \pi(j, X_j)] = P[A_j],
\end{aligned} \tag{A.11}$$

and therefore

$$P[(\tau_j^n > j) \cap A_j] = P[A_j] - P[(\tau_j^n = j) \cap A_j] \rightarrow 0.$$

Thus, obviously, for any  $0 < \epsilon < 1$ ,

$$P[|(\tau_j^n - j) 1_{A_j}| > \epsilon] \rightarrow 0,$$

i.e.,  $(\tau_j^n - j) 1_{A_j} \xrightarrow{P} 0$ , and so

$$\tau_j^n 1_{A_j} = j 1_{A_j} + (\tau_j^n - j) 1_{A_j} \xrightarrow{P} j 1_{A_j} = \tau_j 1_{A_j}. \quad (\text{A.12})$$

(ii): Next, assume that  $\pi(j, X_j) \neq C_j(X_j)$  a.s. for all  $j$ . We will show by backward induction over  $j = T, \dots, j = 0$ , that  $\tau_j^n$  converge in probability to  $\tau_j$  given by (6.1). The case  $j = T$  is clear. Next, define

$$B_j := \{C_j(X_j) > \pi(j, X_j)\}.$$

We have again by Theorem 16 and Lemma 24 below that

$$\lim_{n \rightarrow \infty} P[-C_j^n(X_j) + \pi(j, X_j) < 0, -C_j(X_j) + \pi(j, X_j) < 0] = P[-C_j(X_j) + \pi(j, X_j) < 0] = P[B_j].$$

Hence,

$$\begin{aligned} P[B_j] &\geq \limsup_{n \rightarrow \infty} P[\tau_j^n = \tau_{j+1}^n, B_j] \geq \liminf_{n \rightarrow \infty} P[\tau_j^n = \tau_{j+1}^n, B_j] \\ &\geq \liminf_{n \rightarrow \infty} P[-C_j^n(X_j) + \pi(j, X_j) < 0, -C_j(X_j) + \pi(j, X_j) < 0] \\ &= P[-C_j(X_j) + \pi(j, X_j) < 0] = P[B_j]. \end{aligned}$$

Thus, all inequalities must be equalities and we obtain

$$\lim_{n \rightarrow \infty} P[\tau_j^n = \tau_{j+1}^n, B_j] = P[B_j].$$

Now, by writing

$$\tau_j^n 1_{B_j} = \tau_{j+1}^n 1_{B_j} + (\tau_j^n - \tau_{j+1}^n) 1_{B_j},$$

using that  $\tau_{j+1}^n 1_{B_j} \xrightarrow{P} \tau_{j+1} 1_{B_j}$  since by induction  $\tau_{j+1}^n \xrightarrow{P} \tau_{j+1}$ , and that for any  $0 < \epsilon < 1$  (recalling that  $\tau_j^n$  and  $\tau_{j+1}^n$  are integer-valued)

$$P[|1_{B_j} (\tau_j^n - \tau_{j+1}^n)| > \epsilon] = P[B_j] - P[\tau_j^n = \tau_{j+1}^n, B_j] \rightarrow 0,$$

i.e.,  $(\tau_j^n - \tau_{j+1}^n) 1_{B_j} \xrightarrow{P} 0$ , we have that

$$\tau_j^n 1_{B_j} \xrightarrow{P} \tau_{j+1} 1_{B_j} = \tau_j 1_{B_j} \quad \text{a.s.}$$

Since by assumption  $\pi(j, X_j) \neq C_j(X_j)$  a.s. it holds that  $1_{A_j} + 1_{B_j} = 1$  a.s., so combined with (A.12) we get  $\tau_j^n \xrightarrow{P} \tau_j$ .

(iii): Define the family  $(\tau_j^{*,n})_{j=0, \dots, T}$  recursively by  $\tau_T^{*,n} = T$ , and

$$\tau_j^{*,n} = \begin{cases} j & \text{if } \pi(j, X_j) > C_j(X_j), \\ j & \text{if } \tau_j^n = j \text{ and } \pi(j, X_j) = C_j(X_j), \\ \tau_{j+1}^{*,n} & \text{else.} \end{cases} \quad (\text{A.13})$$

Clearly,  $\tau^{*,n}$  does not stop (i.e.  $\tau_j^{*,n} > j$ ) if the continuation value is greater than the exercise value and does not continue (i.e.  $\tau_j^{*,n} = j$ ) if the exercise value is greater than the continuation value. Hence,  $\tau^{*,n}$  is an optimal strategy. Now we prove the statement by backward induction for  $j = T, \dots, 0$ . For  $j = T$  the statement is clear. Suppose that the statement is true for  $j + 1$ . Let  $A_j$  and  $B_j$  be as in (i) and (ii) above, and  $I_j := \{\pi(j, X_j) = C_j(X_j)\}$ . Then, we may write

$$P[\tau_j^n \neq \tau_j^{*,n}] = P[\{\tau_j^n \neq \tau_j^{*,n}\} \cap A_j] + P[\{\tau_j^n \neq \tau_j^{*,n}\} \cap B_j] + P[\{\tau_j^n \neq \tau_j^{*,n}\} \cap I_j]. \quad (\text{A.14})$$

First, we have

$$P[\{\tau_j^n \neq \tau_j^{*,n}\} \cap A_j] = P[\{\tau_j^n \neq j\} \cap A_j] = P[A_j] - P[\{\tau_j^n = j\} \cap A_j] \rightarrow 0,$$

due to (A.11). Second, in the middle term of (A.14), we have

$$\begin{aligned} \{\tau_j^n \neq \tau_j^{*,n}\} \cap B_j &= \{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \\ &= \left( \{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap A_j^n \right) \cup \left( \{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap (A_j^n)^c \right) \\ &= (B_j \cap A_j^n) \cup \left( \{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap (A_j^n)^c \right), \end{aligned}$$

where  $A_j^n := \{\pi(j, X_j) > C_j^n(X_j)\}$ . Note that

$$\begin{aligned} P[B_j \cap A_j^n] &= P[(\pi(j, X_j) < C_j(X_j)) \cap (\pi(j, X_j) > C_j^n(X_j))] \\ &= P[(\pi(j, X_j) < C_j(X_j))] \\ &\quad - P[(\pi(j, X_j) < C_j(X_j)) \cap (\pi(j, X_j) \leq C_j^n(X_j))] \\ &= P[(\pi(j, X_j) < C_j(X_j))] - P[(\pi(j, X_j) < C_j(X_j)) \cap (\pi(j, X_j) < C_j^n(X_j))] \\ &\quad - P[(\pi(j, X_j) < C_j(X_j)) \cap (\pi(j, X_j) = C_j^n(X_j))] \\ &=: \eta_n - P[(\pi(j, X_j) < C_j(X_j)) \cap (\pi(j, X_j) = C_j^n(X_j))] \leq \eta_n, \end{aligned}$$

with  $\eta_n \geq 0$  and  $\eta_n \rightarrow 0$  by Lemma 24. So we have  $P[B_j \cap A_j^n] \rightarrow 0$ . Furthermore,

$$\{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap (A_j^n)^c = \{\tau_{j+1}^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap (A_j^n)^c,$$

and so, by the induction hypothesis,

$$P[\{\tau_j^n \neq \tau_{j+1}^{*,n}\} \cap B_j \cap (A_j^n)^c] \leq P[\{\tau_{j+1}^n \neq \tau_{j+1}^{*,n}\}] \rightarrow 0.$$

For the third term in (A.14) we have

$$\{\tau_j^n \neq \tau_j^{*,n}\} \cap I_j = \{\tau_{j+1}^n \neq \tau_{j+1}^{*,n}\} \cap I_j,$$

and hence

$$P[\{\tau_j^n \neq \tau_j^{*,n}\} \cap I_j] \leq P[\{\tau_{j+1}^n \neq \tau_{j+1}^{*,n}\}] \rightarrow 0,$$

by induction again. ■

**Lemma 24** *If  $Y_n$  converges in probability to  $Y$ , then, for every  $t$ ,*

$$\lim_n P[Y_n < t, Y < t] = P[Y < t].$$

**Acknowledgements:** We are very grateful to an Area Editor and two anonymous referees for their careful reading, comments and suggestions that have significantly improved the paper. We are also grateful to Christian Bender, Juan Vera Lizcano, Hans Schumacher, Jianing Zhang, and conference participants at the EURANDOM-CWI Workshop on BSDEs, the Lorentz Center Workshop on Models and Numerics in Financial Mathematics, and the First Zurich-Hannover Workshop on Insurance and Financial Mathematics 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).

## References

- [1] ANDERSEN, L., AND BROADIE, M. (2004). A primal-dual simulation algorithm for pricing multi-dimensional 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] BAYRAKTAR, E., AND YAO, S. (2014). On the robust optimal stopping problem. *SIAM Journal on Control and Optimization* 52, 3135–3175.
- [7] BAYRAKTAR, E., AND YAO, S. (2017a). On the robust Dynkin game. *The Annals of Applied Probability* 27, 1702–1755.
- [8] BAYRAKTAR, E., AND YAO, S. (2017b). Optimal stopping with random maturity under nonlinear expectations. *Stochastic Processes and their Applications* 127, 2586–2629.
- [9] BELOMESTNY, D., BENDER, C. AND SCHOENMAKERS, J. (2009). True upper bounds for Bermudan products via non-nested Monte Carlo. *Mathematical Finance* 19, 53–71.
- [10] BELOMESTNY, D., AND KRÄTSCHEMER, V. (2016). Optimal stopping under model uncertainty: A randomized stopping times approach. *The Annals of Applied Probability* 26, 1260–1295.
- [11] BELOMESTNY, D., AND KRÄTSCHEMER, V. (2017). Optimal stopping under probability distortions and law invariant coherent risk measures. *Mathematics of Operations Research*, in press.
- [12] BENDER, C., SCHWEIZER, N., AND ZHUO, J. (2017). A primal-dual algorithm for BSDEs. *Mathematical Finance* 27, 866–901.
- [13] BENEŠ, 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.
- [14] BEN-TAL, A. (1985). The entropic penalty approach to stochastic programming. *Mathematics of Operations Research* 10, 263–279.
- [15] BEN-TAL, A., AND TEBoulLE, M. (1986). Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science* 32, 1445–1466.
- [16] BEN-TAL, A., AND TEBoulLE, M. (1987). Penalty functions and duality in stochastic programming via  $\phi$ -divergence functionals. *Mathematics of Operations Research* 12, 224–240.
- [17] BEN-TAL, A., AND TEBoulLE, M. (2007). An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance* 12, 449–476.

- [18] BEN-TAL, A., BERTSIMAS, D., AND BROWN, D. (2010). A soft robust model for optimization under ambiguity. *Operations Research* 58, 1220–1234.
- [19] BESKOS, A., AND ROBERTS, G. (2005). Exact simulation of diffusions. *The Annals of Applied Probability* 15, 2422–2444.
- [20] BJÖRK, T., AND SLINKO, I. (2006). Towards a general theory of good-deal bounds. *Review of Finance* 10, 221–260.
- [21] BOUCHARD, B., AND ELIE, R. (2008). Discrete-time approximation of decoupled Forward-Backward SDE with jumps. *Stochastic Processes and their Applications* 118, 53–75.
- [22] BRENNAN, M., AND SCHWARTZ, E. (1985). Evaluating natural resource investments. *Journal of Business* 58, 135–157.
- [23] BROADIE, M., AND GLASSERMAN, P. (2004). A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance* 7, 35–72.
- [24] BROADIE, M., AND KAYA, O. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54, 217–231.
- [25] CARRIERE, J. (1996). Valuation of the early-exercise price for options using simulations and non-parametric regression. *Insurance: Mathematics and Economics* 19, 19–30.
- [26] CERREIA-VIOGLIO, S., MACCHERONI, F., MARINACCI, M., AND MONTRUCCHIO, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory* 146, 1275–1330.
- [27] CHEN, Z. AND EPSTEIN, L.G. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70, 1403–1443.
- [28] CHEN, N., AND HUANG, Z. (2013). Localization and exact simulation of Brownian motion driven stochastic differential equations. *Mathematics of Operations Research* 38, 591–616.
- [29] CHENG, X., AND RIEDEL, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics* 7, 29–68.
- [30] CHERIDITO, P., DELBAEN, F., AND KUPPER, M. (2006). Coherent and convex monetary risk measures for unbounded càdlàg processes. *Finance and Stochastics* 10, 427–448.
- [31] 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.
- [32] 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.
- [33] COCHRANE, J. AND SAÁ-REQUEJO, J. (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy* 108, 79–119.
- [34] CONT, R., AND TANKOV, R. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton.
- [35] CSISZÁR, I. (1975).  $I$ -divergence geometry of probability distributions and minimization problems. *Annals of Probability* 3, 146–158.
- [36] 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.
- [37] DAYANIK, S., AND KARATZAS, I. (2003). On the optimal stopping problem for one-dimensional diffusions. *Stochastic Processes and their Applications* 107, 173–212.
- [38] DASI A., AND LAPORTE, G. (2005). A continuous model for multistore competitive location. *Operations Research* 53, 263–280.
- [39] DELBAEN, F. (2006). The structure of  $m$ -stable sets and in particular of the set of risk neutral measures. Émery, M., and Yor, M. (eds.). *In Memoriam Paul-André Meyer*. Springer, Berlin. 215–258.

- [40] DELBAEN, F., PENG, S., AND ROSAZZA GIANIN, E. (2010). Representation of the penalty term of dynamic concave utilities. *Finance and Stochastics* 14, 449–472.
- [41] DHARMA KWON, H. (2010). Invest or exit? Optimal decisions in the face of a declining profit stream. *Operations Research* 58, 638–649.
- [42] DIXIT, A. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy* 97, 620–638.
- [43] DIXIT, A., AND PINDYCK, R.S. (1996). *Investment under uncertainty*. Princeton University Press, Princeton.
- [44] DUFFIE, D. AND EPSTEIN, L.G. (1992). Stochastic differential utility. *Econometrica* 60, 353–394.
- [45] EKREN, I., TOUZI, N., AND ZHANG, J. (2014). Optimal stopping under nonlinear expectation. *Stochastic Processes and their Applications* 124, 3277–3311.
- [46] EL KAROUI, N. (1981). Les aspects probabilistes du contrôle stochastique. Hennequin, P.L. (ed.). *Ecole d’Eté de Probabilités de Saint-Flour IX-1979*, Springer, Berlin. 73–238.
- [47] ELLSBERG, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics* 75, 643–669.
- [48] FÖLLMER, H., AND SCHIED, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics* 6, 429–447.
- [49] FÖLLMER, H., AND SCHIED, A. (2011). *Stochastic Finance*, Walter de Gruyter & Co., Berlin.
- [50] FRITTELLI, M., AND ROSAZZA GIANIN, E. (2002). Putting order in risk measures. *Journal of Banking & Finance* 26, 1473–1486.
- [51] GIESECKE, K., AND SMELOV, D. (2013). Exact sampling of jump-diffusions. *Operations Research* 61, 894–907.
- [52] GILBOA, I., AND SCHMEIDLER, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141–153.
- [53] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York.
- [54] 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.
- [55] HANSEN, L., AND JAGANNATHAN, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99, 225–262.
- [56] HANSEN, L., AND SARGENT, T. (2001). Robust control and model uncertainty. *American Economic Review* 91, 60–66.
- [57] HANSEN, L., AND SARGENT, T. (2007). *Robustness*. Princeton University Press, Princeton.
- [58] HAUGH, M., AND KOGAN, L. (2004). Pricing American options: A duality approach. *Operations Research* 52, 258–270.
- [59] HEATH, D., AND KU, H. (2004). Pareto equilibria with coherent measures of risk. *Mathematical Finance* 14, 163–172.
- [60] HENDERSON, V., AND HOBSON, D. (2013). Risk aversion, indivisible timing options and gambling. *Operations Research* 61, 126–137.
- [61] HENRY-LABORDÈRE, P., TAN, X., AND TOUZI, N. (2015). Exact simulation of multi-dimensional stochastic differential equations. Available at SSRN, id: 2598505.
- [62] HUBER, P. (1981). *Robust Statistics*. Wiley, New York.
- [63] JOHNSON, H. (1987). Options on the maximum or the minimum of several assets. *Journal of Financial and Quantitative Analysis* 22, 277–283.
- [64] KARATZAS, I., AND SHREVE, S. (1998). *Methods of Mathematical Finance*. Vol. 39. Springer, New York.



- [65] KNIGHT, F.H. (1921). *Risk, Uncertainty, and Profit*. Houghton Mifflin Co., Boston.
- [66] KRÄTSCHMER, V., AND SCHOENMAKERS, J. (2010). Representations for optimal stopping under dynamic monetary utility functionals. *SIAM Journal on Financial Mathematics* 1, 811–832.
- [67] LAEVEN, R.J.A., AND STADJE, M. (2013). Entropy coherent and entropy convex measures of risk. *Mathematics of Operations Research* 38, 265–293.
- [68] LAEVEN, R.J.A., AND STADJE, M. (2014). Robust portfolio choice and indifference valuation. *Mathematics of Operations Research* 39, 1109–1141.
- [69] LESNEVSKI, V., NELSON, B. AND STAUM, J. (2007). Simulation of coherent risk measures based on generalized scenarios. *Management Science* 53, 1756–1769.
- [70] LONGSTAFF, F., AND SCHWARTZ, E. (2001). Valuing american options by simulation: A simple least-squares approach. *Review of Financial Studies* 14, 113–147.
- [71] MACCHERONI, F., MARINACCI, M., AND RUSTICHINI, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* 74, 1447–1498.
- [72] MATOUSSI, A., PIOZIN, L., AND POSSAMAÏ, D. (2014). Second-order BSDEs with general reflection and game options under uncertainty. *Stochastic Processes and their Applications* 124, 2281–2321.
- [73] MATOUSSI, A., POSSAMAÏ, D., AND ZHOU, C. (2013). Second order reflected backward stochastic differential equations. *The Annals of Applied Probability* 23, 2420–2457.
- [74] McDONALD, R., AND SIEGEL, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics* 101, 707–727.
- [75] NUTZ, M., AND ZHANG, J. (2015). Optimal stopping under adverse nonlinear expectation and related games. *The Annals of Applied Probability* 25, 2503–2534.
- [76] ØKSENDAL, B., AND SULEM, A. (2007). *Applied stochastic control of jump diffusions*. 2nd Ed. Universitext Springer, Berlin.
- [77] 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.
- [78] PENG, S. (2004). *Nonlinear Expectations, Nonlinear Evaluations and Risk Measures*. Vol. 1856 of *Lecture Notes in Mathematics*. Springer, Berlin.
- [79] PESKIR, G., AND SHIRYAEV, A. N. (2006). *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel.
- [80] PHILPOTT, A., DE MATOS, V., AND FINARDI, E. (2013). On solving multistage stochastic programs with coherent risk measures. *Operations Research* 61, 957–970.
- [81] RIEDEL, F. (2009). Optimal stopping with multiple priors. *Econometrica* 77, 857–908.
- [82] RIEDEL, F. (2004). Dynamic coherent risk measures. *Stochastic processes and their Applications* 112, 185–200.
- [83] ROGERS, L.C.G. (2002). Monte Carlo valuation of American options. *Mathematical Finance* 12, 271–286.
- [84] RUSZCZYŃSKI, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming Series B* 125, 235–261.
- [85] RUSZCZYŃSKI, A., AND SHAPIRO, A. (2006). Optimization of convex risk functions. *Mathematics of Operations Research* 31, 433–452.
- [86] RUSZCZYŃSKI, A., AND SHAPIRO, A. (2006). Conditional risk mappings. *Mathematics of Operations Research* 31, 544–561.
- [87] SCHMEIDLER, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57, 571–587.

- [88] 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.
- [89] SHAPIRO, A., DENTCHEVA, D. AND RUSZCZYŃSKI, A. (2009). *Lectures on stochastic programming: Modeling and theory*. SIAM, Philadelphia.
- [90] SNELL, L.J. (1952). Applications of martingale system theorems. *Trans. Amer. Math. Soc.* 73, 293–312.
- [91] 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.
- [92] TANG, S., AND WEI, W. (2012). Representation of dynamic time-consistent convex risk measures with jumps. *Risk and Decision Analysis* 3, 167–190.
- [93] 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.
- [94] TSITSIKLIS, J. N., AND VAN ROY, B. (2000). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks* 12, 694–703.
- [95] WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York.