Robust Optimal Stopping

Roger J. A. Laeven
Amsterdam School of Economics
University of Amsterdam, EURANDOM,
and ACRM

based on joint work with

Volker Krätschmer, Marcel Ladkau, John Schoenmakers and Mitja Stadje

April 14, 2016

1. Introduction

- The theory of optimal stopping and control has evolved into one of the most important branches of modern probability and optimization.
- It has a wide variety of applications in many areas:
  - operations management;
  - statistics;
  - economics; and
  - financial and insurance mathematics.
- There exists a vast literature on both theory and applications of optimal stopping and control, going back to Wald (1950) and Snell (1952).
- 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 and optimal policyholder behavior in insurance.
- These applications naturally lead to an optimal stopping problem.

Reward's Expectation

- Since the (future) reward (sequence) is typically uncertain in these applications, it needs to be evaluated using probabilistic methods.

- The main target in the literature on standard optimal stopping is the maximization of the expected reward over a family of stopping strategies.

- That is, the central object is the expectation of the reward induced by the problem's payoff process.

- Such a setting requires that the reward's expectation can be unambiguously determined by the decision-maker, which is the case in particular if the reward's probability law is given to the decision-maker.

Ambiguity

- In reality, however, this is quite a restrictive requirement: in many situations the decision-maker faces uncertainty about the true probabilistic model.

- 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 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.

Convex Measures of Risk

- In a general probabilistic setting, a robust approach that has recently gained much attention is provided by convex measures of risk.
- By the representation theorem of convex risk measures, a random future reward, say H, is evaluated according to

$$U(H) = \inf_{Q \in \mathcal{Q}} \{\mathrm{E}_Q[H] + c(Q)\}, \qquad (1)$$

where $\mathcal{Q} = \{Q | Q \sim P\}$ is the set of probabilistic models Q that share the same null sets with a base reference model P, with each Q attaching a different probability law to the future reward H, and c is a penalty function specifying the plausibility of the model Q.

- Models Q that have 'low' plausibility are associated with a high penalty, while models that have 'high' plausibility yield a low penalty, with $c(Q) = \infty$ corresponding to the case in which the model Q is considered fully implausible.
- By taking the infimum over $\mathcal{Q}$ a conservative worst-case approach occurs, also typical in (deterministic) robust optimization.

Variational Preferences

- 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 (1961).

- Popular approaches to decision-making under ambiguity are provided by the multiple priors preferences of Gilboa and Schmeidler (1989), also referred to as maxmin expected utility, and the significant generalization of variational preferences developed by Maccheroni, Marinacci and Rustichini (2006).

- With linear utility, variational preferences reduces to (1).

- Such preferences induce aversion to ambiguity. A version of multiple priors was also studied by Huber (1981) in robust statistics; see also the early Wald (1950).

Time-Consistency

- In a dynamic setting, such as considered in this paper, time-consistent versions of convex measures of risk were discussed by Riedel (2004); see also Duffie and Epstein (1992) and Chen and Epstein (2002).

- (Strong) 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.

- For dynamic versions of evaluations of the form (1), time-consistency is equivalent to a dynamic programming principle (recursiveness).

This Paper's Motivation

- 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 (2009), Krätschmer and Schoenmakers (2010), Bayraktar, Karatzas and Yao (2010), Bayraktar and Yao (2011), Cheng and Riedel (2013) and Øksendal, Sulem and Zhang (2014).

- However, the development of numerical methods to practically solve robust optimal stopping problems may currently be considered breaking ground.

Contribution [1]

- In this paper, we develop a method to practically solve the optimal stopping problem under ambiguity in a general continuous-time setting, allowing for general time-consistent convex measures of risk, i.e., all time-consistent dynamic counterparts of (1), and general (sequences of) rewards.

- As to the payoff process, we allow for a general jump-diffusion model specification.

- The key to our method is to expand two duality theories of a different kind.

- The first kind of duality theory is the martingale duality approach to standard optimal stopping problems, dating back to Rogers (2002), Haugh and Kogan (2004) and Andersen and Broadie (2004) (see also Davis and Karatzas, 1994).

  - We expand their martingale dual representation to encompass general preference functionals beyond plain conditional expectation.

- The second kind of duality theory explicates the connection between time-consistent convex measures of risk and backward stochastic differential equations (BSDEs), which we expand to our setting.

Three Steps

Our method is then composed of three steps:

1. First, expanding duality theory of the second kind and using backward stochastic calculus, we construct a suitable Doob martingale from the Snell envelope generated by the optimally stopped and robustly evaluated payoff process.

2. Second, expanding duality theory of the first kind, we employ this martingale to construct an approximated upper bound to the solution of the optimal stopping problem.

3. Third, we introduce the notion of backward-forward simulation to obtain a genuine upper bound to the solution.

Contribution [2]

- We analyze the asymptotic behavior of our method by deriving its convergence properties.

- To the best of our knowledge, we are not aware of other practical solution methods for robust optimal stopping problems in the literature so far.

- Finally, to illustrate the generality of our approach and the relevance of ambiguity to optimal stopping, we supplement the presentation of our method with a few examples of robust optimal stopping problems, including:
  - Kullback-Leibler divergences;
  - worst case scenarios; and
  - good-deal bounds.

- Our numerical results illustrate that our algorithm is easily implemented for a wide range of robust optimal stopping problems and has good convergence properties, yielding accurate results in realistic settings at the pre-limiting level.

- They also reveal that ambiguity can have a significant impact on the robust reward evaluations under standard specifications.

Related Literature

- The development of methods to practically compute the solution to a standard optimal stopping problem has a long history, in particular in the American-style option literature.
  - Seminal contributions based on regression include Carriere (1996) and Longstaff and Schwartz (2001); see also Tsitsiklis and Van Roy (2000) and Clément, Lamberton and Protter (2002).
  - These methods 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 (2004) who exploited the dual representation obtained by Rogers (2002) and Haugh and Kogan (2004).
- Employing duality (of the first kind), our method may, in some sense, be viewed as the analogous contribution for robust optimal stopping problems of the original contribution by Andersen and Broadie (2004) 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 allow for a more general reward specification.

BSDEs

- 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.
- It generates a genuine (biased high) upper bound, which will converge to the true solution as the number of Monte Carlo simulations and basis functions increases and the mesh ration of the time-grid tends to zero.

Outline

2. Problem Description [1]

- Consider a decision-maker 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 equipped with two independent processes, which will serve as our stochastic drivers:
  - (i) A standard d-dimensional Brownian motion $W = (W^1, \ldots, W^d)^{\mathsf{T}}$.
  - (ii) A standard k-dimensional Poisson process $N = (N^1, \ldots, N^k)^{\mathsf{T}}$ with intensities $\lambda_P = (\lambda_P^1, \ldots, \lambda_P^k)^{\mathsf{T}}$.
- The process X, driven by W and N, 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.

Problem Description [2]

- The decision-maker chooses a stopping time τ taking values between time 0 and a fixed maturity time $T < \infty$.

- We assume that if the decision-maker exercises at time $\tau = t_i$, he receives the reward

$$H_{t_i} = \Pi(t_i, X_{t_i}) + \sum_{j=i}^{L} h(t_j, X_{t_j}), \qquad t_i \in \{t_0 = 0, t_1, \ldots, t_L = T\}, \quad (2)$$

for functions Π and h mapping from $\{t_0 = 0, t_1, \ldots, t_L = T\} \times \mathbb{R}^n$ to $\mathbb{R}$.

Two Canonical Examples

Standard examples that take the form (2) include:

(a) The optimal entrance problem: In this case, typically,

$$\Pi(t, x) = -\exp(-\rho t)\,\kappa,$$

for a fixed irreversible cost κ depreciating at a continuous rate ρ, and

$$h(t, x) = \exp(-\rho t)\,(h(x) - \xi),$$

which measures the present value of the payoff or the production per time unit, $h(x)$, after entering the market, minus the running costs, ξ.

- ▶ Often times $h(x)$ is simply taken to be equal to x.
- ▶ Of course, the fixed costs may also depend on the state of the economy at time t, X_t.

(b) The optimal (simple) reward problem: In this case, $h \equiv 0$ and $\Pi(t, x)$ is the (simple) reward function of exercising at time t.

- ▶ This problem appears abundantly in the American option pricing literature, with X_t a vector of risky asset values at time t.

Standard Optimal Stopping

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

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

where $\mathcal{T} = \{t_0 = 0 < t_1 < \ldots < t_L = T\}$ 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 times subject to model uncertainty.
  - 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.

Robust Optimal Stopping

▶ Henceforth, we postulate that the decision-maker adopts a convex measure of risk and evaluates his future reward according to

$$U(H_\tau) = \inf_{Q \in \mathcal{Q}} \{ \mathrm{E}_Q[H_\tau] + c(Q) \}, \qquad (3)$$

with $\mathcal{Q} = \{ Q | Q \sim P \}$ and $c : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$.

▶ For our purposes, we have to consider the dynamic version of (3), given by

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

in which $c_t(Q)$ reflects the esteemed plausibility of the model Q given the information up to time t.

▶ The robust optimal stopping problem at time t_i is then given by

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

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

Time-Consistency and Dynamic Programming [1]

- 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.

- 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 taking a certain form.

Change of Measure

- First, we explain what a change of measure from P to Q implies in our setting.

- If $Q \sim P$, we denote by D_t the Radon-Nikodym derivative $D_t = \mathrm{E}\left[\frac{dQ}{dP}|\mathcal{F}_t\right]$.

- For every model $Q \sim P$, there exist a predictable, $\mathbb{R}^d$-valued, stochastic drift q and a positive, predictable, $\mathbb{R}^k$-valued process λ such that the Radon-Nikodym derivative can be written as

$$D_t = \exp\left\{\int_0^t q_s dW_s + \int_0^t \log\left(\frac{\lambda_s}{\lambda_P}\right) dN_s - \int_0^t \left(\frac{|q_s|^2}{2} + \lambda_s - \lambda_P\right) ds\right\},$$

$t \in [0, T]$, with $\frac{\lambda_s}{\lambda_P} := (\frac{\lambda_s^1}{\lambda_P^1}, \ldots, \frac{\lambda_s^k}{\lambda_P^k})^\mathsf{T}$.

- In particular, Q is uniquely characterized by q and λ.

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

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

(ii) U is *recursive*, that is, U satisfies *Bellman's dynamic programming principle*: $U_0(U_t(H)I_A) = U_0(HI_A)$ for every $t \in [0, T]$, $A \in \mathcal{F}_t$ and square-integrable H.

(iii) There exists a function

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

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

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

Assumptions [1]

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

$$c_t(Q) = \mathbb{E}_Q \left[\int_t^T r(s, q_s, \lambda_s - \lambda_P) ds \middle| \mathcal{F}_t \right], \qquad (4)$$

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

Examples

- We now illustrate the generality of our robust optimal stopping problem and Assumption (G1) with some examples of penalty functions satisfying our conditions.
- All these examples appear in numerical illustrations.

    1. Kullback-Leibler divergence: $c_t(Q) = \alpha \mathrm{KL}_t(Q|P)$, with

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

    and $\alpha > 0$; see Csiszár (1975), Ben-Tal (1985) and Ben-Tal and Teboulle (1987, 2007).

    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.

    3. Worst case with mean (partially) known.

    4. Good-deal bounds (Cochrane and Saá-Requejo, 2000, Hansen and Jagannathan, 1991).

Assumptions [2]

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

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

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

for a compact set

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

▶ Loosely speaking, condition (G3) states that, if the additional drift q or jump intensity $\lambda - \lambda_P$ that the model Q adds to the Brownian motion or the Poisson process when compared to P is 'too large', then the model Q should not be considered.

▶ Condition (G3) may be generalized substantially.

3. Duality Theory

- We show that there exists an optimal stopping family $\left(\tau_{t_i}^*\right)_{t_i \in \{t_0=0, t_1, \ldots, t_L=T\}}$ satisfying

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

- Furthermore, we show that Bellman's principle

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

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

$$U_{t_i}^h := U_{t_i} \left(\sum_{j=i}^{L} h(t_j, X_{t_j}) \right).$$

U-Martingales and the Doob Decomposition

▶ To compute the solution V^*, the (generalized) Snell envelope, to the optimal stopping problem, we will rely on the Doob decomposition of V^* into a martingale and a predictable process.

  ▶ We first need to generalize the notion of a (standard) martingale (with respect to an ordinary conditional expectation) to martingales with respect to classes of functionals.

▶ We will say that M is a U-martingale if $M_s = U_s(M_t)$, $s, t \in \{t_0 = 0, t_1, \ldots, t_L = T\}$ and $s \leq t$.

▶ The class of U-martingales M with $M_0 = 0$ is denoted by $\mathcal{M}_0^U$.

▶ Define, for $i = 0, \ldots, L$,

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

▶ One may verify that M^{*g} is a U-martingale, A^{*g} is non-decreasing and predictable, $M_0^{*g} = A_0^{*g} = 0$, and that

$$V_{t_i}^* = V_0^* + M_{t_i}^{*g} + A_{t_i}^{*g}, \qquad i = 0, \ldots, L,$$

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

Duality Theory of the First Kind: A Proposition

▶ To construct genuine upper bounds to the optimal solution to our robust optimal stopping problem, which will converge asymptotically to the true value, our method will exploit an additive dual representation of the robust optimal stopping problem.

▶ We expand the well-known dual representation for the standard setting, in which U is just the ordinary conditional expectation (Rogers (2002) and Haugh and Kogan (2004)).

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

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

$$= U_{t_i} \left(\max_{t_j \in \{t_i, \dots, T\}} \left(\Pi(t_j, X_{t_j}) + U_{t_j}^h + M_T^{*g} - M_{t_j}^{*g} \right) \right),$$

$$t_i \in \{t_0 = 0, \dots, T\}.$$

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 (4), we define a function g by Fenchel's duality as follows:

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

with C_t induced by assumption (G3).

- By assumption (G3), g thus defined is Lipschitz continuous.

Computing M^{*g}

- Suppose that, for every exercise date t_j, $j = 0, \ldots, L$, we have a fine time grid $\pi_j = \{s_{j0} = t_j < s_{j1} < \ldots < s_{jP} = t_{j+1}\}$.
- Denote the corresponding overall time grid by $\pi = \{s_{00}, s_{01}, \ldots, s_{LP}\}$.
- The following theorem provides a way to practically compute M^{*g} by connecting it to specific semi-martingale dynamics that can be dealt with numerically efficiently.

Computing M^{*g}: A Theorem

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

$$dU_t^h = -g(t, Z_t^h, \tilde{Z}_t^h)dt + Z_t^h dW_t + \tilde{Z}_t^h d\tilde{N}_t, \quad \text{for } t \in (t_j, t_{j+1}], \qquad (5)$$

and $U_{t_j}^h = U_{t_{j+}}^h + h(t_j, X_{t_j})$, *for each* $j \in \{0, \ldots, L-1\}$.

Furthermore, there exists a unique square-integrable predictable $(Z^*, \tilde{Z}^*)$ *such that*

$$dU_t(V_{t_{j+1}}^*) = -g(t, Z_t^*, \tilde{Z}_t^*)dt + Z_t^* dW_t + \tilde{Z}_t^* d\tilde{N}_t,$$

$$\text{for } t \in [t_j, t_{j+1}], j \in \{0, \ldots, L-1\}. \qquad (6)$$

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

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

BDSEs

- Equations (5)–(6) are also referred to as backward stochastic differential equations (BSDEs) 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 (2004).

Examples Revisited

- To conclude the exposition of the duality theory of the second kind, we employ the penalty functions of our previous examples and compute the corresponding g's.

  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:
     $g(t, z, \tilde{z}) = -\delta_1 |z| - \delta_2 |\tilde{z}|$.
  3. Worst case with mean (partially) known.
  4. Good-deal bounds.

4. Algorithm: General Outline

Our method is composed of three steps.

1. 'Duality theory of the second kind' jointly with Bellman's principle will serve as a *first* stepping stone for our approach, by providing a practical way to find U-martingales

2. ... to be employed in the dual representation, which is our *second* stepping stone ('duality theory of the first kind').

While this bound will be seen to converge to the true optimal solution asymptotically and is an approximated upper bound at the pre-limiting level, it is not a genuine upper bound estimate to the true optimal solution as it is not 'biased high'.

3. 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.

Three Steps Again

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

    Step (1.a.) Compute an approximation to $(U_{t_j}^h)_{t_j \in \{0, \ldots, T\}}$ through backward recursion. 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 $U_s(V_{t_{j+1}}^*)_{t_j < s \leq t_{j+1}}$. This involves least squares Monte Carlo regression.

    Step (1.c.) Given the whole path of $(Z_s^*, \tilde{Z}_s^*)_{s \in [0, T]}$, compute an approximation to $(M_{t_j}^{*g})_{t_j \in \{t_1, \ldots, T\}}$.

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

Step (3.) Introducing backward-forward simulation:

    Step (3.a.) Compute a genuine (biased high) upper bound to $(U_{t_j}^h)_{t_j \in \{0, \ldots, t_{L-1}\}}$ 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.

Results

- Since our optimal stopping problem is Markovian, there exists a function $v^* : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ such that $V_t^* = v^*(t, X_t)$. In particular, $V_0^* = v^*(0, X_0)$.
- Our method, then, will be proven to have the following two appealing properties:
  - (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 illustrate that, already after a limited number of Monte Carlo simulations, our method yields rather close estimates in realistic settings.
    - Moreover, by property (ii) above, for a finite time grid and a finite number of simulations, the genuine upper bound will also provide a safety buffer.
    - The examples also illustrate the generality of our approach and the relevance of ambiguity to optimal stopping.

5. Numerical Examples

- We consider two stochastic processes, X_i, $i = 1, 2$, with dynamics

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

  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.

- First, we consider the optimal (simple) reward problem (i.e., $h \equiv 0$). and analyze 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 the general setting with non-trivial jump component.
    - This problem occurs e.g., in American-style derivative pricing in finance, in which case the drift μ^i under the reference model is equal to $\rho - \delta$ (for $i = 1, 2$), where ρ represents the risk-free rate and δ the dividend rate.

- An appropriate choice of the basis functions m^M, ψ^M and $\tilde{\psi}^M$, $M \in \mathbb{N}$, that we employ in the least squares Monte Carlo regressions, is crucial to obtain tight upper bounds.

Optimal Reward Problem with a Geometric Brownian Motion

- Following Andersen and Broadie (2004), 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 $t_j = \frac{jT}{9}$, $j = 0, \ldots, 9$, and a fine grid $\{s_{jp}\}$ with $\Delta_{jp} = s_{j(p+1)} - s_{jp} = 1/1{,}500$.
- For the choice of basis functions, we follow Andersen and Broadie (2004) by including still-alive European options and corresponding option deltas.
- Our results are based on 10,000 simulated trajectories for the calculation of the regression coefficients in Step (1.b.) and the U-martingale increments in Step (1.c.), the approximated upper bound to V^* in Step (2.), and the genuine upper bound to V^* in Step (3.b.).
- In the univariate case, we restrict attention to the simple reward $\Pi(t, X_t) = \exp(-\rho t)(X_t - K)^+$.

Numerical Results [1]

- We consider the case of the Kullback-Leibler divergence for different values of its parameter α.
- The results are in the table below.
- 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	α				
	10	100	10^4	10^6	∞
90	2.4405	4.0546	4.4049	4.4088	4.4088
	2.4932	*4.0673*	*4.4662*	*4.4708*	*4.4708*
	(0.0003)	(0.0009)	(0.0013)	(0.0013)	(0.0013)
100	4.6077	7.4023	7.9848	7.9913	7.9914
	4.8251	*7.3887*	*8.0328*	*8.0402*	*8.0403*
	(0.0005)	(0.0012)	(0.0018)	(0.0019)	(0.0019)
110	10.0281	12.2732	13.1574	13.1661	13.1662
	10.1468	*12.2934*	*13.1656*	*13.1756*	*13.1757*
	(0.0008)	(0.0015)	(0.0024)	(0.0024)	(0.0024)

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

Numerical Results [2]

- Only in the case of $\alpha = \infty$ we have reference values, provided e.g., by Andersen and Broadie (2004).
- They appear to be very close to our values.
- For example, for $x_0 = 100$, the *true* value is 7.98, which is to be compared to our approximated and genuine upper bounds equal to 7.99 and 8.04, respectively.
- With an increase in α we observe an, initially rapid, increase in the robust call option's value.
- In general, we observe that Bermudan call option values may decrease substantially when ambiguity is taken into account.
- For numerical results:
  - in the multivariate case;
  - with jumps;
  - with other forms of ambiguity;
  - with other rewards;

  see the paper.

6. Conclusions

- We have developed a method to practically compute the solution to the optimal stopping problem in a general continuous-time setting featuring general time-consistent ambiguity averse preferences and general rewards driven by jump-diffusions.

- The resulting algorithm delivers an approximation to the solution that converges asymptotically to the true solution and yields a safe genuine (biased high) upper bound at the pre-limiting level.

- Our method is widely applicable, numerically efficient, and eventually requires only simple least squares Monte Carlo regression techniques.

- Extensive numerical illustrations reveal the potential importance of ambiguity to optimal stopping.

- Our method may be generalized to encompass multiple stopping problems, which we intend to consider in future research.