Pricing Israeli options: a pathwise approach

C. Kühn*, A. E. Kyprianou† and K. van Schaik*

Abstract

An Israeli option (also referred to as game option or recall option) generalizes an American option by also allowing the seller to cancel the option prematurely, but at the expense of some penalty. Kifer (2000) shows that in the classical Black-Scholes market such contracts have unique no-arbitrage prices. In Kyprianou (2004) and Kühn and Kyprianou (2006) characterizations were obtained for the price of two classes of Israeli options. For the general case, we give a dual resp. pathwise pricing formula similar to Rogers (2002) and investigate this approach numerically.

Keywords: Israeli options, convertible bonds, Dynkin games, hedging strategies, Monte Carlo simulation.

JEL Classification: G13, C73.

Mathematics Subject Classification (2000): 91A15, 60G40, 60G48.

1 Introduction

There are few examples of derivatives having the feature that they can be both exercised by the holder prematurely and recalled by the writer. The most prominent example are convertible bonds. The holder can convert them into a predetermined number of stocks of the issuing firm, and the issuer can recall them, paying some compensation to the holder.

*Address: Frankfurt MathFinance Institute, Johann Wolfgang Goethe-Universität, 60054 Frankfurt a.M., Germany, e-mail: {ckuehn, schaik}@math.uni-frankfurt.de

†Address: School of Mathematical and Computer Sciences, Heriot Watt University, Edinburgh EH14 4AS, Scotland, UK, e-mail: kypriano@ma.hw.ac.uk

These contracts were approached in the economic literature for the first time by Brennan and Schwartz (1977) and Ingersoll (1977a), (1977b). Optimal conversion and call policies were derived. These approaches are limited however to Markovian claim structures. The first general analysis of such kind of derivatives, using no-arbitrage arguments in conjunction with game theory, was made by Kifer (2000). For a practical example of a convertible callable bond see McConnell and Schwartz (1986), for an overview of recent progress in handling convertible bonds see Sirbu, Pikovsky, and Shreve (2004), Kallsen and Kühn (2005) and the references therein.

Let us introduce Kifer's model. Fix some finite time horizon $T \in (0, \infty)$. Suppose that $X = \{X_t : t \in [0,T]\}$ and $Y = \{Y_t : t \in [0,T]\}$ are two stochastic processes, defined on some filtered probability space $\big(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, P\big)$, with values in $\mathbb{R}_+ \cup \{+\infty\}$ and càdlàg paths (right continuous with left limits). We assume that $Y_t \geq X_t$ for all $t \in [0,T)$ and $Y_T = X_T$. The filtered probability space satisfies the usual conditions of right-continuity and completeness.

The Israeli option, as introduced by Kifer (2000), is a contract between a writer and holder at time $t = 0$ such that both have the right to exercise at any time before the expiry date T. If the holder exercises, then (s)he may claim the value of X at the exercise date and if the writer exercise, (s)he is obliged to pay to the holder the value of Y at the time of exercise. If neither have exercised at time T then the writer pays the holder the amount $X_T = Y_T$. If both decide to claim at the same time then the lesser of the two claims is paid. But, it turns out that this marginal case has no impact on the option price as long as the payoff lies in the interval $[X_t, Y_t]$. In short, if the holder will exercise with stopping time σ and the writer with stopping time τ we can conclude that the holder receives at time $\sigma \wedge \tau$ the amount

$$Z_{\sigma,\tau} = X_\sigma \mathbf{1}_{(\sigma \leq \tau)} + Y_\tau \mathbf{1}_{(\tau < \sigma)}. \tag{1.1}$$

In a complete market, with a *unique* risk-neutral measure $\mathbb{P} \sim P$, Kifer obtained a unique no-arbitrage price for such contracts.

Suppose that the financial market consists of one riskless and d risky assets, i.e. $S = (S^0, S^1, \dots, S^d)$, where S^i, $i = 0, \dots, d$, are positive semimartingales. To simplify notations we assume w.l.o.g. that $S^0 = 1$. Put differently, we work with discounted values with respect to the numéraire S^0.

Assume that the market is complete, i.e. there is a *unique* equivalent measure $\mathbb{P}$ under which S is a local martingale. This is the situation in the Black-Scholes model. We shall denote $\mathbb{E}$ to be expectation under $\mathbb{P}$. The following theorem is a slight generalization of Kifer's pricing result. Kifer stated it under the Black-Scholes model and a slightly stronger integrability assumption. However, the arguments are based on martingale representation which holds for complete markets in general, see Theorem 2.1 resp. Theorem 3.3. in Kramkov (1996) applied to the case that the set of equivalent martingale measures is a singleton. For a proof of Theorem 1.2 see Kifer (2000) in conjunction with Step 1 of the proof of Theorem 2.4 in this paper.

Definition 1.1 *A stochastic process $U = (U_t)_{t\in[0,T]}$ is said to be of class (D) if $\{U_\tau \mid \tau \in \mathcal{T}_{0,T}\}$ is uniformly integrable.*

Theorem 1.2 *Suppose that X is of class (D). Let $\mathcal{T}_{t,T}$ be the class of $\mathbb{F}$-stopping times valued in $[t,T]$. There is a unique no-arbitrage price process of the Israeli option. It can be represented by the right continuous process $V = \{V_t : t \in [0,T]\}$ where*

$$\begin{aligned} V_t &= \textit{ess-inf}_{\tau\in\mathcal{T}_{t,T}}\, \textit{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\, \mathbb{E}\left(Z_{\sigma,\tau} \,|\, \mathcal{F}_t \right) \\ &= \textit{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\, \textit{ess-inf}_{\tau\in\mathcal{T}_{t,T}}\, \mathbb{E}\left(Z_{\sigma,\tau} \,|\, \mathcal{F}_t \right), \end{aligned} \tag{1.2}$$

i.e. V_t is the dynamic value of a Dynkin game. Further, if Y has no positive jumps and X has no negative jumps then optimal stopping strategies exist and are given by

$$\sigma_t^* = \inf\{s \in [t,T] : V_s \le X_s\} \ \text{ and } \ \tau_t^* = \inf\{s \in [t,T] : V_s \ge Y_s\} \tag{1.3}$$

for all $t \in [0,T]$.

Remark 1.3 *In incomplete markets no-arbitrage arguments alone are not sufficient to determine unique derivative prices. An established approach to price derivatives in incomplete markets is by utility (indifference) arguments. For American and Israeli options this was done in Kallsen and Kühn (2004). It turns out that the "fair price" of an Israeli option is again the value of a Dynkin game. The unique equivalent martingale measure*

$\mathbb{P}$ in (1.2) is replaced by a well-chosen so-called neutral pricing measure $P^* \sim P$ which plays a crucial role in utility maximization, see Kallsen and Kühn (2004). Therefore, the results of this paper can be used in the same manner to simulate (utility based) option prices in incomplete markets.

Although in special cases the optimal stopping problem in (1.2) can be solved explicitly (see Kühn and Kyprianou (2006) for the Israeli put option), in general, analytical methods are virtually out of the question. Clearly one should not expect to be in any a better position than when posing the same question for American claims. Consider for example, for large $d \in \mathbb{N}$, an American option with contingent claim of the form

$$\left(K - \min\left\{S_t^1, ..., S_t^d\right\}\right)^+ .$$

Trying to characterize free boundary problems for such an option can also be a problem because of the high dimensionality. For such cases, Rogers proposes to work via a dual pricing formula which inspires a different outlook when it comes to simulation.

The goal of this paper is to produce a dual representation of the price of an Israeli option which could in the same way be used for Monte Carlo simulation as Rogers' contribution on American options. American claims are covered in this paper as long as we only assume that the lower bound X is integrable. For it set $Y_t = \infty$, $t \in [0, T)$. We made the following observation which can help when it comes to simulation. Let us interpret the stopping game in (1.2) no longer as a stochastic but as a *deterministic* stopping game, by choosing the optimal exercise strategies for each path $\omega \in \Omega$ separately. This represents the hypothetical situation where all information about future price movements of the underlyings are available at the very beginning. Instead of the payoff function $Z_{s,t}$ as defined in (1.1) we consider $Z_{s,t}(\omega) - M^*_{s\wedge t}(\omega)$ where M^* is a martingale which will be characterized later on. In a complete market M^* corresponds to the gain process generated by a hedging strategy. It turns out that the values of these stopping games coincide $\mathbb{P}$-a.s. with V_0, the price of the Israeli option, see Theorem 2.7. Therefore, with the right martingale M^* it is possible to attain the exact option price by the simulation of a single sample path. Theorem 2.10 provides estimations in the case of a simulation with a wrong martingale $M \neq M^*$. Theorem 2.8 suggests to consider the variance of the pathwise value as a minimization criterion in order to approach the right martingale M^*. This idea is implemented in Section 3.

Cvitanić and Karatzas (1996) established a connection between the value of a Dynkin game, i.e. a stochastic stopping game as it arises in Theorem 1.2, and the solution of a backward stochastic differential equation with reflection. In this framework they also construced a corresponding deterministic stopping game, i.e. a game played path by path. However, their pathwise obtained values of the deterministic stopping games are still stochastic and coincides only in expectation with the value of the stochastic stopping game.

2 Representations of the Israeli option price

Throughout the paper we assume that X is of class (D). Firstly we need a couple of notations. They are rather voluminous. Primarily, this is the price for capturing the case of discontinuous payoff processes.

Definition 2.1 *Denote by $\mathcal{M}_0$ the set of martingales starting in zero.*

In continuation of the usual notations for stopped processes we make the following definition.

Definition 2.2 *Let U be a stochastic process, τ a stopping time, and D an $\mathcal{F}_{\tau-}$-measurable subset of Ω ($\mathcal{F}_{\tau-}$ is the σ-algebra generated by $\mathcal{F}_0$ and by the sets of the form $A \cap \{t < \tau\}$, where $t \in (0,T]$ and $A \in \mathcal{F}_t$). We denote by $U^{\tau,D}$ the process U stopped at $\tau-$, if the event D occurs, and at τ, if D does not occur, i.e.*

$$U_t^{\tau,D} = U_t \mathbf{1}_{(t<\tau)} + U_{\tau-}\mathbf{1}_{\left(t \geq \tau \text{ and } \omega \in D\right)} + U_\tau \mathbf{1}_{\left(t \geq \tau \text{ and } \omega \notin D\right)}. \tag{2.1}$$

Remark 2.3 *If U is a martingale or predictable, then the respective property also holds for $U^{\tau,D}$. This can be derived by using e.g. Proposition I.2.10 in Jacod and Shiryayev (2003). By (2.1) we allow for stopping at $t-$ which can be seen as some closure of usual stopping strategies.*

Define the stopping times

$$\tau^{\varepsilon} = \inf\{s \geq 0 \ \mid \ V_s \geq Y_s - \varepsilon\}, \quad \varepsilon > 0, \quad \text{and} \quad \widehat{\tau} = \lim_{\varepsilon \to 0} \tau^{\varepsilon}. \tag{2.2}$$

The limit exists as τ^{ε} is monotone in ε. If Y is lower semicontinuous (i.e. it has no positive jumps), then $\widehat{\tau} = \tau^*$. Analogously we set

$$\sigma^{\varepsilon} = \inf\{s \geq 0 \ \mid \ V_s \leq X_s + \varepsilon\}, \quad \varepsilon > 0, \quad \text{and} \quad \widehat{\sigma} = \lim_{\varepsilon \to 0} \sigma^{\varepsilon}.$$

Let g be the American claim given by the (discounted) payoff process

$$g_s = X_s \mathbf{1}_{(s<\widehat{\tau})} + Y_{\widehat{\tau}-} \mathbf{1}_{\left(\widehat{\tau} \leq s \text{ and } \tau^{\varepsilon} < \widehat{\tau} \ \forall \varepsilon > 0\right)} + Y_{\widehat{\tau}} \mathbf{1}_{\left(\widehat{\tau} \leq s \text{ and } \exists \varepsilon > 0 \text{ s.t. } \tau^{\varepsilon} = \widehat{\tau}\right)}. \tag{2.3}$$

g is the payoff process the buyer is faced with – given the optimal exercise strategy of the seller. Analogously we define the payoff process the seller is faced with by

$$h_t = X_{\widehat{\sigma}-} \mathbf{1}_{\left(\widehat{\sigma} \leq t \text{ and } \sigma^{\varepsilon} < \widehat{\sigma} \ \forall \varepsilon > 0\right)} + X_{\widehat{\sigma}} \mathbf{1}_{\left(\widehat{\sigma} \leq t \text{ and } \exists \varepsilon > 0 \text{ s.t. } \sigma^{\varepsilon} = \widehat{\sigma}\right)} + Y_t \mathbf{1}_{(t<\widehat{\sigma})}. \tag{2.4}$$

Let $D_1 = \{\tau^{\varepsilon} < \widehat{\tau}, \quad \forall \varepsilon > 0\}$ and $D_2 = \{\sigma^{\varepsilon} < \widehat{\sigma}, \quad \forall \varepsilon > 0\}$.

Theorem 2.4 *The process $V^{\widehat{\tau},D_1}$ is a supermartingale, the process $V^{\widehat{\sigma},D_2}$ is a submartingale, and the process $\left(V^{\widehat{\tau},D_1}\right)^{\widehat{\sigma},D_2} = \left(V^{\widehat{\sigma},D_2}\right)^{\widehat{\tau},D_1}$ is a martingale. It follows that $V^{\widehat{\sigma}\vee\widehat{\tau}}$ enjoys a canonical decomposition of the form*

$$V^{\widehat{\sigma}\vee\widehat{\tau}} = V_0 + M^* + A - B, \tag{2.5}$$

where

(i) *the process $M^* = \{M_t^* : t \in [0,T]\}$ belongs to $\mathcal{M}_0$,*

(ii) *$A = \{A_t : t \in [0,T]\}$ is predictable and non-decreasing such that $A^{\widehat{\tau},D_1} = 0$.*

(iii) *$B = \{B_t : t \in [0,T]\}$ is predictable and non-decreasing such that $B^{\widehat{\sigma},D_2} = 0$.*

In addition, the Snell envelope of g

$$V_t' := \textit{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}\left(g_\sigma \,|\, \mathcal{F}_t \right),$$

exists as an element of class (D) so that it possesses a Doob-Meyer decomposition. Analogously, the lower *Snell envelope of* h

$$V''_t := ess\text{-}inf_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}(\, h_\tau \,|\, \mathcal{F}_t) \,.$$

exists as an element of class (D) and possesses a Doob-Meyer decomposition.

We have $V_0 = V'_0 = V''_0$. *The process* $(M^*)^{\widehat{\tau}, D_1}$ *coincides with the martingale part in the Doob-Meyer decomposition of* V' *and* $(M^*)^{\widehat{\sigma}, D_2}$ *coincides with the martingale part in the Doob-Meyer decomposition of* V''.

Proof. *Step 1:* Let us show that under the assumption that X is of class (D) the Dynkin game possesses an equilibrium point, i.e. the process V in (1.2) exists. This was proven by Lepeltier and Maingueneau (1984) (henceforth LM) for bounded payoff processes X and Y. However, the existence of an equilibrium point holds also under the weaker assumption above. To see this assume in the first instance that Y is bounded by a $\mathbb{P}$-martingale M. W.l.o.g. $\mathbb{P}(M_T > 0) = 1$. By applying the statements of LM to the processes $\widetilde{X}_t := \frac{X_t}{M_t}$ and $\widetilde{Y}_t := \frac{Y_t}{M_t}$ (which have values in $[0,1]$), taking the measure $\widetilde{\mathbb{P}}$ defined by $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \frac{M_T}{M_0}$, we obtain the existence of a value process $\widetilde{V}$. On the other hand, we have for $U \in \{X, Y\}$, $t \in [0,T]$, $\sigma \in \mathcal{T}_{t,T}$

$$\widetilde{\mathbb{E}}\left(\widetilde{U}_\sigma \middle| \mathcal{F}_t \right) = \frac{1}{M_t} \mathbb{E}\left(M_T \widetilde{U}_\sigma \middle| \mathcal{F}_t \right) = \frac{1}{M_t} \mathbb{E}(\, U_\sigma \,|\, \mathcal{F}_t) \,.$$

This implies that the value process V for the payoff processes X, Y and the measure $\mathbb{P}$ also exists with $V_t = M_t \widetilde{V}_t$. Thus, the following implication holds

$$Y \text{ is bounded by a martingale} \implies \text{value process } V \text{ for } (X, Y) \text{ exists.} \tag{2.6}$$

Now, assume only that X is of class (D) instead of $Y \leq M$. By Bank and El Karoui (2004), Lemma 3.10, this implies that $X \leq \bar{M}$ for some martingale $\bar{M}$. Instead of Y consider the upper bound $\bar{Y} := Y \wedge (1 + \bar{M})$ and let $\bar{V}$ be the corresponding value process for the payoff processes $\bar{X} = X$ and $\bar{Y}$. As $X \leq \bar{Y}$ and $X_T = \bar{Y}_T$ the value process of this Dynkin game exists by (2.6). Let us show that the value process V for X and Y also exists and coincides with $\bar{V}$. We have

$$\bar{V}_t \leq \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}(\, X_\sigma \,|\, \mathcal{F}_t) \leq \bar{M}_t. \tag{2.7}$$

Consider now the ε-optimal recall time of the option seller given by $\bar{\tau}_t^\varepsilon := \inf\{s \geq t \mid \bar{V}_s \geq \bar{Y}_s - \varepsilon\}$. By right-continuity we have $\bar{V}_{\bar{\tau}_t^\varepsilon} \geq \bar{Y}_{\bar{\tau}_t^\varepsilon} - \varepsilon$. Together with (2.7) this implies that $\bar{Y}_{\bar{\tau}_t^\varepsilon} < 1 + \bar{M}_{\bar{\tau}_t^\varepsilon}$ for all $\varepsilon \in (0,1)$. Therefore, by definition of $\bar{Y}$,

$$\bar{Y}_{\bar{\tau}_t^\varepsilon} = Y_{\bar{\tau}_t^\varepsilon}, \quad \forall \varepsilon \in (0,1). \tag{2.8}$$

With this we obtain

$$\begin{aligned}
&\text{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\text{ess-inf}_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}\left(X_\sigma \mathbf{1}_{(\sigma\leq\tau)} + Y_\tau \mathbf{1}_{(\tau<\sigma)} \middle| \mathcal{F}_t\right) \\
&\geq \bar{V}_t \\
&= \lim_{\varepsilon\to 0} \text{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\mathbb{E}\left(X_\sigma \mathbf{1}_{(\sigma\leq\bar{\tau}_t^\varepsilon)} + \bar{Y}_{\bar{\tau}_t^\varepsilon} \mathbf{1}_{(\bar{\tau}_t^\varepsilon<\sigma)} \middle| \mathcal{F}_t\right) \\
&= \lim_{\varepsilon\to 0} \text{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\mathbb{E}\left(X_\sigma \mathbf{1}_{(\sigma\leq\bar{\tau}_t^\varepsilon)} + Y_{\bar{\tau}_t^\varepsilon} \mathbf{1}_{(\bar{\tau}_t^\varepsilon<\sigma)} \middle| \mathcal{F}_t\right) \\
&\geq \text{ess-inf}_{\tau\in\mathcal{T}_{t,T}}\text{ess-sup}_{\sigma\in\mathcal{T}_{t,T}}\mathbb{E}\left(X_\sigma \mathbf{1}_{(\sigma\leq\tau)} + Y_\tau \mathbf{1}_{(\tau<\sigma)} \middle| \mathcal{F}_t\right).
\end{aligned} \tag{2.9}$$

The first inequality is by $Y \geq \bar{Y}$. The first equality is by Theorem 11 in LM applied to X and $\bar{Y}$ and the second equality by (2.8). By (2.9) V exists and coincides with $\bar{V}$.

Step 2: As V is dominated by the Snell-envelope of X which is dominated by $\bar{M}$ defined in Step 1 we have

$$V \leq \bar{M}. \tag{2.10}$$

In addition we have

$$\begin{aligned}
V_t' &\leq \text{ess-sup}_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}\left(X_\sigma \middle| \mathcal{F}_t\right) + \mathbb{E}\left(V_{\hat{\tau}-}\mathbf{1}_{(\tau^\varepsilon<\hat{\tau}\ \forall\varepsilon>0)} + V_{\hat{\tau}}\mathbf{1}_{\left(\exists\varepsilon>0 \text{ s.t. } \tau^\varepsilon=\hat{\tau}\right)} \middle| \mathcal{F}_t\right) \\
&\overset{(2.10)}{\leq} \bar{M}_t + \bar{M}_t^{\hat{\tau},D_1}
\end{aligned}$$

and

$$V_t'' \leq \mathbb{E}(Y_T|\mathcal{F}_t) + \bar{M}_t \leq 2\bar{M}_t,$$

where the second inequality follows because $Y_T = X_T$. As $\bar{M}$ and $\bar{M}^{\hat{\tau},D_1}$ are martingales they are processes of class (D), cf. e.g. Proposition I.1.46 in Jacod and Shiryaev (2003). Therefore, V, V', and V'' are processes of class (D), which is of later use for the Doob-Meyer decomposition.

Step 3: Let us show that $V^{\widehat{\tau},D_1} = V'$. Fix a $t \in [0,T]$. On the event $\{\widehat{\tau} \le t\}$ the assertion is obvious. In the following calculations, we assume that we are working on the event $\{\widehat{\tau} > t\}$. By Theorem 11 in LM, we have for $\varepsilon > 0$

$$\text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}\left(Z_{\sigma,\tau_t^\varepsilon} \,\middle|\, \mathcal{F}_t \right) \le V_t + \varepsilon, \tag{2.11}$$

where $\tau_t^\varepsilon = \inf\{s \ge t \mid V_s \ge Y_s - \varepsilon\}$. Let $\delta > 0$. Again, by Theorem 11 in LM, we can choose an δ-optimal stopping strategy σ^δ such that $\mathbb{E}(g_{\sigma^\delta} | \mathcal{F}_t) \ge \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}(g_\sigma | \mathcal{F}_t) - \delta$. Let $\varepsilon_n \downarrow 0$ as $n \uparrow \infty$. Possibly after re-definition of $Z_{s,t}$ for $s = t$ (which does not change the value process V), we have pointwise convergence of $Z_{\sigma^\delta,\tau_t^{\varepsilon_n}}$ to g_{σ^δ} as $n \to \infty$ (note that $\tau_t^{\varepsilon_n} = \tau^{\varepsilon_n}$ for sufficiently small ε_n). By (2.10) the sequence $\left(Z_{\sigma^\delta,\tau_t^{\varepsilon_n}}\right)_{n \in \mathbb{N}}$ is uniformly integrable so that we have convergence of $\mathbb{E}\left(Z_{\sigma^\delta,\tau_t^{\varepsilon_n}} \,\middle|\, \mathcal{F}_t\right)$ to $\mathbb{E}(g_{\sigma^\delta} | \mathcal{F}_t)$ in $L^1(\mathbb{P})$. This implies $\mathbb{P}$-a.s. convergence on a subsequence $(\varepsilon_{n_k})_{k \in \mathbb{N}}$, w.l.o.g. $n_k = k$. Together with (2.11) applied to $\varepsilon = \varepsilon_{n_k}$ we obtain

$$\begin{aligned} V_t + \varepsilon_k &\ge \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}\left(Z_{\sigma,\tau_t^{\varepsilon_k}} \,\middle|\, \mathcal{F}_t \right) \\ &\ge \mathbb{E}\left(Z_{\sigma^\delta,\tau_t^{\varepsilon_k}} \,\middle|\, \mathcal{F}_t \right) \\ &\stackrel{k\to\infty}{\longrightarrow} \mathbb{E}(g_{\sigma^\delta} | \mathcal{F}_t) \\ &\ge \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}(g_\sigma | \mathcal{F}_t) - \delta. \end{aligned}$$

As $\delta > 0$ was arbitrary this implies $V_t \ge \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}(g_\sigma | \mathcal{F}_t)$. For the opposite estimation we use that, again by Theorem 11 in LM,

$$V_t - \varepsilon \le \text{ess-inf}_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left(Z_{\sigma_t^\varepsilon,\tau} \,\middle|\, \mathcal{F}_t \right), \tag{2.12}$$

where $\sigma_t^\varepsilon = \inf\{s \ge t \mid V_s \le X_s + \varepsilon\}$. On a subsequence $(n_k)_{k \in \mathbb{N}}$ we have again pointwise convergence of $\mathbb{E}\left(Z_{\sigma_t^\varepsilon,\tau^{\varepsilon_{n_k}}} \,\middle|\, \mathcal{F}_t \right)$ to $\mathbb{E}\left(g_{\sigma_t^\varepsilon} \,\middle|\, \mathcal{F}_t\right)$, w.l.o.g. $n_k = k$. We obtain

$$\begin{aligned} V_t - \varepsilon &\le \text{ess-inf}_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left(Z_{\sigma_t^\varepsilon,\tau} \,\middle|\, \mathcal{F}_t \right) \\ &\le \mathbb{E}\left(Z_{\sigma_t^\varepsilon,\tau^{\varepsilon_k}} \,\middle|\, \mathcal{F}_t \right) \\ &\stackrel{k\to\infty}{\longrightarrow} \mathbb{E}\left(g_{\sigma_t^\varepsilon} \,\middle|\, \mathcal{F}_t \right) \\ &\le \text{ess-sup}_{\sigma \in \mathcal{T}_{t,T}} \mathbb{E}(g_\sigma | \mathcal{F}_t), \quad k \to \infty. \end{aligned}$$

Altogether we arrive at

$$V^{\widehat{\tau},D_1} = V'. \tag{2.13}$$

Using similar reasoning one can conclude that

$$V^{\widehat{\sigma}, D_2} = V''. \tag{2.14}$$

As V' is a supermartingale and V'' is a submartingale it follows that $V^{\widetilde{\tau} \vee \widetilde{\sigma}}$ is a special semimartingale, i.e. it possesses a unique decomposition

$$V = V_0 + M^* + A - B,$$

with the properties as stated in the Theorem. The last assertion of the theorem follows then from (2.13), (2.14), and Remark 2.3. ■

Remark 2.5 *Theorem 2.4 provides a canonical decomposition of the process V only up to $\widehat{\sigma} \vee \widehat{\tau}$. For the whole process such a decomposition need not exist, as V is in general not a semimartingale. For example, V can be deterministic and possess infinite variation on $[0, T]$. This is in contrast to American claims where V is a supermartingale and hence a semimartingale.*

Lemma 2.6 *Let $L = (L_t)_{t \in [0,T]}$ be a positive process of class (D) (then its Snell-envelope exists and is also an element of class (D), see El Karoui (1981), Proposition 2.29). Let $V^L = V_0^L + M^L + A^L$ be the Doob-Meyer decomposition of the Snell-envelope. We have*

$$V_0^L \quad = \quad \sup_{s \in [0,T]} \left(L_s - M_s^L \right), \quad \mathbb{P}\text{-}a.s. \tag{2.15}$$

Proof. By $L \leq V^L \leq V_0^L + M^L$, the left-hand side of (2.15) is obviously not smaller than the right-hand side. On the other hand for each $\varepsilon > 0$ define the $[0, T]$-valued stopping time $\tau^\varepsilon = \inf\{s \geq 0 \mid V_s^L \leq L_s + \varepsilon\}$. By equation (24) in Fakeev (1970) we have $A_{\tau^\varepsilon}^L = 0$ and therefore $L_{\tau^\varepsilon} - M_{\tau^\varepsilon}^L \geq V_{\tau^\varepsilon}^L - M_{\tau^\varepsilon}^L - \varepsilon = V_0^L - \varepsilon$. ■

Theorem 2.7 *We have*

$$V_0 = \sup_{s \in [0,T]} \inf_{t \in [0,T]} \left(Z_{s,t} - M_{s \wedge t}^* \right) = \inf_{t \in [0,T]} \sup_{s \in [0,T]} \left(Z_{s,t} - M_{s \wedge t}^* \right), \quad \mathbb{P}\text{-}a.s., \tag{2.16}$$

where M^* is the martingale part in (2.5).

Proof. Consider the American contingent claim g defined in (2.3). By Theorem 2.4 its Snell envelope V' possesses a Doob-Meyer decomposition $V' = V_0' + M' + A'$ with $V_0' = V_0$ and $M' = (M^*)^{\widehat{\tau},D_1}$. By this and Lemma 2.6 we obtain

$$V_0 = \sup_{\sigma\in\mathcal{T}_{0,T}} \mathbb{E}(g_\sigma) = \sup_{s\in[0,T]} (g_s - M_s') \geq \inf_{t\in[0,T]} \sup_{s\in[0,T]} (Z_{s,t} - M^*_{s\wedge t}), \quad \mathbb{P}\text{-a.s.} \tag{2.17}$$

To assure oneself of the last inequality choose $t(\omega) = \tau^\varepsilon(\omega)$ for all $\varepsilon > 0$.
Again, by Theorem 2.4 we know that the lower Snell envelope V'' of the process h defined in (2.4) has a Doob-Meyer decomposition $V'' = U_0'' + M'' - B''$ with $V_0'' = V_0$ and $M'' = (M^*)^{\widetilde{\sigma},D_2}$. We obtain as in (2.17)

$$V_0 = \inf_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}(h_\tau) = \inf_{t\in[0,T]} (h_t - M_t'') \leq \sup_{s\in[0,T]} \inf_{t\in[0,T]} (Z_{s,t} - M^*_{s\wedge t}), \quad \mathbb{P}\text{-a.s.} \tag{2.18}$$

(2.17) and (2.18) imply the assertion. ■

Theorem 2.8 *Assume that $\mathbb{E}\left(\sup_{t\in[0,T]} Y_t\right) < \infty$. Let $(M^{(n)})_{n\geq 1} \subset \mathcal{M}_0$ such that the sequence $(\sup_{t\in[0,T]} |M_t^{(n)}|)_{n\in\mathbb{N}}$ is uniformly integrable. Consider the pathwise games with values*

$$V^n := \sup_{s\in[0,T]} \inf_{t\in[0,T]} (Z_{s,t} - M^{(n)}_{s\wedge t}) = \inf_{t\in[0,T]} \sup_{s\in[0,T]} (Z_{s,t} - M^{(n)}_{s\wedge t}). \tag{2.19}$$

Suppose that $(V^n - \mathbb{E}(V^n))_{n\in\mathbb{N}}$ tends to 0 in probability as $n\to\infty$. Then, $\mathbb{E}(V^n)$ converges with $n\to\infty$ to the value of the stochastic game (1.2).

Proof. Let $\varepsilon > 0$ and assume that $\mathbb{P}(|V^n - \mathbb{E}(V^n)| > \varepsilon) \leq \varepsilon$. Define the stopping times $\sigma^n = \inf\{t \geq 0 \mid X_t - M_t^{(n)} \geq \mathbb{E}(V^n) - 2\varepsilon\} \wedge T$ and $\tau^n = \sup\{t \geq 0 \mid Y_t - M_t^{(n)} \leq \mathbb{E}(V^n) + 2\varepsilon\} \wedge T$. On the set $A := \{|V^n - \mathbb{E}(V^n)| \leq \varepsilon\}$ the optimal times for the deterministic game (2.19) are not missed which yields $X_{\sigma^n\wedge\tau^n} - M_{\sigma^n\wedge\tau^n} \leq V^n \leq Y_{\sigma^n\wedge\tau^n} - M_{\sigma^n\wedge\tau^n}$. Together with the right-continuity of the payoff processes we obtain that

$$|Z_{\sigma^n,\tau^n} - M^{(n)}_{\sigma^n\wedge\tau^n} - \mathbb{E}(V^n)| \leq 2\varepsilon \text{ on } A.$$

With $\mathbb{E}(M^{(n)}_{\sigma^n\wedge\tau^n}) = 0$, the fact that $X_0 \leq V^n \leq Y_0$, $\mathbb{P}$-a.s., and some rough estimations on the set $\Omega\setminus A$ it follows that

$$|\mathbb{E}(Z_{\sigma^n,\tau^n}) - E(V^n)| \quad \leq \quad 2\varepsilon + 2\mathbb{E}(1_{\Omega\setminus A} \sup_{t\in[0,T]} Y_t) =: \delta(\varepsilon). \tag{2.20}$$

As $\mathbb{P}(\Omega \setminus A) \leq \varepsilon$, $\delta(\varepsilon)$ tends to 0 for $\varepsilon \to 0$. Analogously it follows that

$$|\mathbb{E}(\inf_{t\in[0,T]}(Z_{\sigma^n,t} - M_{\sigma^n\wedge t})) - \mathbb{E}(V^n)| \vee |\mathbb{E}(\sup_{s\in[0,T]}(Z_{s,\tau^n} - M_{s\wedge\tau^n})) - \mathbb{E}(V^n)| \leq \delta'(\varepsilon) \quad (2.21)$$

where $\delta'(\varepsilon) := \delta(\varepsilon) + \sup_{n\in\mathbb{N}} \sup_{B\in\mathcal{F},\ \mathbb{P}(B)\leq\varepsilon} \mathbb{E}\left(1_B \sup_{t\in[0,T]} |M_t^{(n)}|\right)$. By the uniformly integrability of $(\sup_{t\in[0,T]} |M_t^{(n)}|)_{n\in\mathbb{N}}$ we have $\delta'(\varepsilon) \to 0$ for $\varepsilon \to 0$. Combining (2.20) and (2.21) we obtain

$$\begin{aligned}
\mathbb{E}(Z_{\sigma^n,\tau^n}) &\geq \mathbb{E}(V^n) - \delta(\varepsilon) \\
&\geq \mathbb{E}\left(\sup_{s\in[0,T]}(Z_{s,\tau^n} - M^{(n)}_{s\wedge\tau^n})\right) - \delta(\varepsilon) - \delta'(\varepsilon) \\
&\geq \inf_{\tau\in\mathcal{T}_{0,T}} \sup_{\sigma\in\mathcal{T}_{0,T}} \mathbb{E}\left(Z_{\sigma,\tau} - M^{(n)}_{\sigma\wedge\tau}\right) - \delta(\varepsilon) - \delta'(\varepsilon) \\
&= \inf_{\tau\in\mathcal{T}_{0,T}} \sup_{\sigma\in\mathcal{T}_{0,T}} \mathbb{E}(Z_{\sigma,\tau}) - \delta(\varepsilon) - \delta'(\varepsilon).
\end{aligned}$$

By symmetry $(\sigma^n, \tau^n)_{n\in\mathbb{N}} \subset \mathcal{T}_{0,T} \times \mathcal{T}_{0,T}$ is a sequence of approximate saddle points for the stochastic game (1.2). This implies that $\mathbb{E}(V^n)$ converges to the value V_0 of the game. ■

Remark 2.9 *Theorem 2.7 removes the involvement with stopping strategies and instead transfers the essence of the pricing problem to a clever choice of a martingale M. In complete markets the choice of a $M \in \mathcal{M}_0$ corresponds to a hedging strategy. With the right martingale $M = M^*$ the random variable*

$$\sup_{s\in[0,T]} \inf_{t\in[0,T]} (Z_{s,t} - M_{s\wedge t}) \quad (2.22)$$

degenerates to a real number and the option price can be attained by the simulation of a single sample path.

From a practical point of view it remains of course the problem to find M^*. Arguably, one has not made things any easier. Theorem 2.8 helps at least to approximate and identify the right martingale. One could choose a finite basis of martingales in $\mathcal{M}_0$ and minimize the empirical variance of the random variable (2.22) over linear combinations of this basis. Rogers argues that with sensible choices of basis one can do quite well in this respect. For an effective computation of the dual upper bound we refer the reader to Kolodko and Schoenmakers (2004). For a general overview on Monte Carlo methods for American *options we refer to Glasserman (2004) and the references therein. For the case of an Israeli option the approach is even more robust w.r.t. the choice of the martingale as both players could profit from the wrong choice of the martingale and effects may counter-balance (see Section 3).*

Theorem 2.10 *We have*

$$\begin{aligned} V_0 &= \inf_{M\in\mathcal{M}_0}\inf_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}\left(\sup_{s\in[0,T]}(Z_{s,\tau}-M_{s\wedge\tau})\right) \\ &= \sup_{M\in\mathcal{M}_0}\sup_{\sigma\in\mathcal{T}_{0,T}} \mathbb{E}\left(\inf_{t\in[0,T]}(Z_{\sigma,t}-M_{\sigma\wedge t})\right). \end{aligned} \tag{2.23}$$

Further, the infimum as well as the supremum are achieved when M is chosen to be M^*.

Proof. By symmetry we have only to show (2.23). For any $(\tau, M)\in\mathcal{T}_{0,T}\times\mathcal{M}_0$ we have that

$$\mathbb{E}\left(\sup_{s\in[0,T]}(Z_{s,\tau}-M_{s\wedge\tau})\right) \geq \sup_{\sigma\in\mathcal{T}_{0,T}}\mathbb{E}(Z_{\sigma,\tau}-M_{\sigma\wedge\tau}) = \sup_{\sigma\in\mathcal{T}_{0,T}}\mathbb{E}(Z_{\sigma,\tau}).$$

Taking the infimum over all $\tau\in\mathcal{T}_{0,T}$ this implies that the right-hand side of (2.23) is at least as big as V_0. To prove the other direction we take for each $\varepsilon>0$ the stopping time τ^ε as defined in (2.2) and the martingale part M^ε of the Snell envelope for the American claim with buying back time τ^ε. We obtain

$$\mathbb{E}\left(\sup_{s\in[0,T]}(Z_{s,\tau^\varepsilon}-M^\varepsilon_{s\wedge\tau^\varepsilon})\right) = \sup_{\sigma\in\mathcal{T}_{0,T}}\mathbb{E}(Z_{\sigma,\tau^\varepsilon}) \leq V_0+\varepsilon,$$

where the equality holds by Lemma 2.6 and the inequality by Theorem 11 of LM. ■

Remark 2.11 *By choosing an arbitrary pair $(\tau, M)\in\mathcal{T}_{0,T}\times\mathcal{M}_0$ and simulating $\mathbb{E}\left(\sup_{s\in[0,T]}(Z_{s,\tau}-M_{s\wedge\tau})\right)$ we can get an upper bound for V_0 and by simulating $\mathbb{E}\left(\inf_{t\in[0,T]}(Z_{\sigma,t}-M_{\sigma\wedge t})\right)$ with an arbitrary pair $(\sigma, M)\in\mathcal{T}_{0,T}\times\mathcal{M}_0$ we obtain a lower bound for V_0. Of course, to obtain tight estimations we have to find both a "good" stopping time and a "good" martingale. This corresponds to the fact that for Israeli options a hedging strategy consists of a trading strategy (in the underlyings) and a stopping time.*

In finite discrete time the value of the pathwise game (2.22) can be effectively computed by an algorithm described in the following Proposition.

Proposition 2.12 *Consider the discrete deterministic game where both the maximizer and minimizer choose a strategy s, t resp. from the set $\{0, \dots, N\}$, with corresponding payoff $\mathbf{1}_{\{s \le t\}} l_s + \mathbf{1}_{\{t < s\}} u_t$. Furthermore the sequences $(l_i)_{i \in \{0,\dots,N\}}$ and $(u_i)_{i \in \{0,\dots,N\}}$ satisfy $l_i \le u_i$ for $i = 0, \dots, N-1$ and $l_N = u_N$.*

We denote the running maximum over the l_i's by $\bar{l}_i$ and the running minimum over the u_i's by $\underline{u}_i$. Set

$$m := \inf\{k \in \{1, \dots, N\} \,|\, \bar{l}_{k-1} \ge u_k \text{ or } \underline{u}_{k-1} \le l_k\}$$

(where we understand $\inf \emptyset = \infty$). If $m = \infty$ the value is $\bar{l}_N = l_N = u_N = \underline{u}_N$. Otherwise, the value is $\bar{l}_{m-1}$ if $\bar{l}_{m-1} \ge u_m$ and $\underline{u}_{m-1}$ if $\underline{u}_{m-1} \le l_m$. If both conditions above are satisfied then also $\bar{l}_{m-1} = \underline{u}_{m-1}$.

Proof. First, by induction over k we show that

$$k < m \Longrightarrow \bar{l}_k \le \underline{u}_k. \tag{2.24}$$

The case $k = 0$ is trivial. Suppose that $\bar{l}_{k-1} \le \underline{u}_{k-1}$. By writing $\bar{l}_k = \bar{l}_{k-1} \vee l_k$ and $\underline{u}_k = \underline{u}_{k-1} \wedge u_k$, checking that $\bar{l}_k \le \underline{u}_k$ boils down to considering four cases, each of which can readily be covered by using the induction hypothesis, that $l_k \le u_k$ and that $k < m$ implies $\bar{l}_{k-1} < u_k$ and $\underline{u}_{k-1} > l_k$.

The value if $m = \infty$ follows directly from (2.24). Furthermore if both conditions in the definition of m are satisfied at the same time then using (2.24) together with $l_m \le u_m$ indeed implies $\bar{l}_{m-1} = \underline{u}_{m-1}$. Finally, suppose that $m < \infty$ is triggered by the first condition $\bar{l}_{m-1} \ge u_m$. This inequality together with (2.24) implies that if the maximizer chooses the strategy $\operatorname{argmax}\{l_i \,|\, i \in \{0, \dots, m-1\}\}$ and the minimzer m this is indeed a saddle point with payoff $\bar{l}_{m-1}$. The case $\underline{u}_{m-1} \le l_m$ is dealt with similarly. ■

3 Numerics

In this section we consider two prominent examples of game options. We discuss how the results of the previous sections give rise to a numerical implementation of the pathwise pricing of the games. This is mainly meant to illustrate the results obtained so far and try out how good an approximation we can get by using martingales that are easily derived from the payoff processes involved.

In the first subsection we look at the callable put under geometric Brownian motion and in the second at the convertible bond under a jump-diffusion. Since no completely explicit formulae are known for the value of these game options we use the technique of Canadization to be able to analyze how good the pathwise approximation is. For an introduction to the Canadization approach we refer the reader to Carr (1998) (but this approach is of no importance for the understanding of the current paper).

Theorem 2.7 suggests the following implementation for the pathwise pricing. As in Rogers paper, we consider a set of hedging martingales $M^{(1)}, \dots, M^{(n)}$ from $\mathcal{M}_0$ for some $n \in \mathbb{N}$. We enlarge this set by considering all linear combinations with coefficients $\lambda^{(1)}, \dots, \lambda^{(n)}$ of those martingales and obviously aim at choosing the weights such that we are "closest" to the optimal martingale M^*. As a criterion we use, as indicated by Theorem 2.8, the variance of the value of the pathwise game. The effort of computing the value of the game — given a path sampled at discrete time points – is of linear order in the number of grid points, see Proposition 2.12.

First we sample 300 paths of the underlying price process (in all the upcoming sampling we will take 50 time steps with a step size of 0.01) and determine the value over each of the paths as a function of $\lambda^{(1)}, \dots, \lambda^{(n)}$, whose variance is then minimized. With its minimizing arguments $\lambda_*^{(1)}, \dots, \lambda_*^{(n)}$ we then sample another 5000 paths and determine the average value over the paths and the variance. This average value is then compared to the true value as found by the Canadization. Note that these numbers are the same as the ones used in Rogers paper.

The actual implementation of the computational work and the sampling is done in the software package "Mathematica"*.

---

*Mathematica has the possibility to generate pseudo-random uniformly distributed numbers by the command Random[]. Yet the standard implementation of this function is flawed and the generated numbers are biased. To avoid this problem we used the implementation of Random[] as e.g. discussed in http://forums.wolfram.com/mathgroup/archive/2004/May/msg00002.html

3.1 The callable put

Consider, as in the Black-Scholes model, a geometric Brownian motion S with strictly positive parameters $\sigma, r, S_0 = s$, i.e.

$$S_t = s \exp\left(\sigma W_t + (r - \sigma^2/2)t\right), \quad t \geq 0$$

where, under the martingale pricing measure $\mathbb{P}$, W is a Brownian motion starting in 0. $\mathbb{P}_s$ indicates that S starts in s and $\mathbb{E}_s$ denotes the expectation under $\mathbb{P}_s$. The payoff processes X and Y belonging to the callable put are given by

$$X_t = e^{-rt}(K - S_t)^+ \qquad \text{and} \qquad Y_t = e^{-rt}((K - S_t)^+ + \delta)$$

for $t < T$ and $X_T = Y_T = \exp(-rT)(K - S_T)^+$. $\delta > 0$ is a penalty the minimizer has to pay additionally if (s)he decides to terminate the contract prematurely. The structure of the corresponding value function was studied in Kühn and Kyprianou (2006). The optimal stopping strategies are given by

$$\tau^* = \inf\{t \in [0, t^*] \mid S_t = K\} \wedge T \text{ where } t^* \text{ is given by } v^A(K, T - t^*) = \delta$$

where $v^A(x, u)$ is the value function for the corresponding American put with time to maturity u and

$$\sigma^* = \inf\{t \in [0, T] \mid S_t \leq \varphi(T - t)\} \wedge T$$

where $\varphi : [0, T] \to [0, K]$ is the exercise boundary, a decreasing function in the time to maturity. Note that there is no explicit formula for t^*.

For calculating V^n as in (2.19) we begin with no martingale at all (or, put differently, only the trivial martingale that's 0 everywhere), so the payoffs are the same as in the original stochastic game, but the game is now played with both players knowing the complete path while determining their strategies.

Numerical results for this case are provided in Table 1.

The deviation of the empirical mean $\widehat{E}(V^n)$ from the true value V_0 can only be caused by the randomness of V^n (see Theorem 2.8). We think that the quotient $\frac{|\widehat{E}(V^n) - V_0|}{\sqrt{\widehat{\mathrm{Var}}(V^n)}}$ is a quantitity of interest as it measures the impact of the randomness of the pathwise value on the actual error in the approximation. Therefore it is quoted in all tables. Notice that in the special case of an American option a highly stochastic V^n automatically leads to an overestimation of the option price. By contrast in the game case things are symmetric. Both players could profit from the wrong choice of the martingale and effects could counter-balance.

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation
80	20.6	22.4	4.15	0.88
90	12.4	12.7	4.37	0.14
100	5.00	4.03	3.83	0.50
110	3.64	2.82	6.02	0.33
120	2.54	1.93	5.79	0.25

Table 1: Numerical values with no martingale and parameter values $\sigma = 0.4$, $r = 0.06$, $K = 100$, $T = 0.5$ and $\delta = 5$

Next we take only one martingale M in our set of hedging martingales, namely the discounted value process of the European option that pays out $X_T = Y_T$ to the maximizer at T, so M is given by $M_T = \exp(-rT)(K - S_T)^+$. In figure 1 a plot is shown of the variance of the value of the pathwise games against the weighing coefficient λ, showing that a stable global minimum exists for a value of λ close to 1 which indicates that minimizing over λ is indeed robust. Further numerical values are given in table 2. We see that this martingale already gives a much better approximation than with no martingale at all.

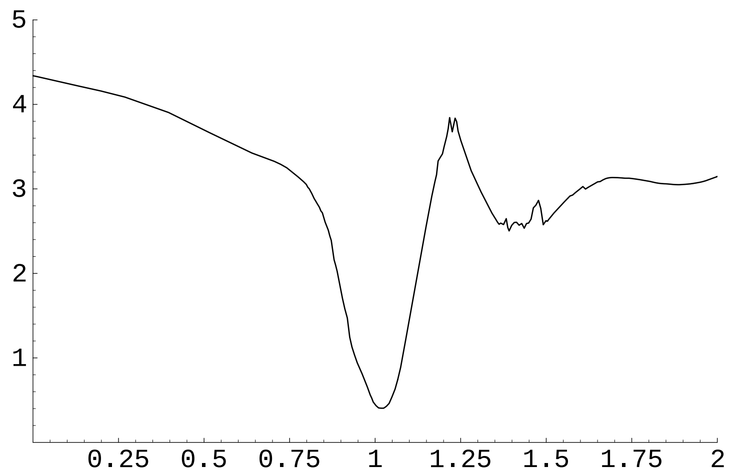

Figure 1: Variance of the value of the pathwise games as a function of λ, parameters as in table 1 and $S_0 = 80$

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation	λ_*
80	20.6	20.8	0.35	0.33	1.00
90	12.4	13.2	1.22	0.72	1.40
100	5.00	5.00	0.00	$--$	1.00
110	3.64	3.77	0.16	0.33	0.58
120	2.54	2.61	0.24	0.14	0.64

Table 2: Numerical values with same parameter values as in table 1

Finally we take in addition to the martingale $M^{(1)} := M$ from the previous setting the martingale $M^{(2)}$ that corresponds to a payoff of δ as soon as S hits K. This is an attempt to bring the influence of the upper payoff into the pathwise game and see if this improves the approximation. Note that the optimal martingale M^* in Theorem 2.7 comes from the Doob-Meyer decomposition of the value process stopped at the *maximum* of both optimal stopping times. But, we take the most practical approach here and keep $M^{(2)}$ constant after S hits K. So, with $\tau := \inf\{t \geq 0 \,|\, S_t = K\}$, we set $M_T^{(2)} := \mathbf{1}_{\{\tau \leq T\}} \exp(-r\tau)\delta$. This gives

$$M_t^{(2)} = \mathbb{E}_s(M_T^{(2)} \,|\, \mathcal{F}_t) = \mathbf{1}_{\{\tau \leq t\}} e^{-r\tau}\delta + \mathbf{1}_{\{\tau > t\}} e^{-rt}\delta F(T-t, S_t), \tag{3.1}$$

where

$$F(u,x) = \mathbb{E}_x\left(\mathbf{1}_{\{\tau \leq u\}} e^{-r\tau}\right) = \frac{\ln(K/x)}{\sigma\sqrt{2\pi}} \int_0^u t^{-3/2} \exp\left(-\frac{(\ln(K/x) - (r - \sigma^2/2)t)^2}{2\sigma^2 t} - rt\right) \mathrm{d}t.$$

Note that since this integral has to be evaluated numerically, including this martingale in our hedging set will cause computations to be considerably slower. And there are more computational issues arising here. The most straightforward way to implement this martingale (which in particular means assigning values to τ) would be to follow the sampled path of S and consider τ to be equal to $\widetilde{\tau}$, which is defined to be the first discrete time point where the sample path crosses the level K. Yet, as it is well-known, the distribution of this $\widetilde{\tau}$ significantly differs from that of τ, even with relatively small time steps, see e.g. Broadie, Glasserman, and Kou (1999) and the references therein. To illustrate this, consider starting from $S_0 = s = 80$ with $K = 100$ as hitting level. With the other parameters the same as in table 1 and taking average values after 5 runs of 5000 paths each, we find the empirical probability of the set $\{\widetilde{\tau} \leq 0.5\}$ to be 0.3720 whereas $\mathbb{P}_s(\tau \leq 0.5) = 0.4182$. So the empirical probability is considerably smaller, which is mainly explained by the fact that there is a significant number of paths of S that actually do cross K when considered in continuous time but stay below K at each of the discrete time points.

The problem with this difference between τ and $\widetilde{\tau}$ is that the process (3.1) is constructed from the law of the time-continuous stopping time τ whereas later on the sample paths are generated for a discrete time skeleton corresponding to the stopping time $\widetilde{\tau}$. This causes the mapping $t \mapsto \widehat{E}(M_t^{(2)})$ (where $\widehat{E}$ denotes the average over all paths) to be significantly non-constant. We take a practical approach in solving this by adjusting our way of sampling as follows. Instead of just sampling from a $\mathcal{N}(\mu\Delta, \sigma^2\Delta)$ distributed random variable to determine the value of the Brownian motion X with volatility σ and drift μ at time point Δ and translating this into the increase of S on a time interval of length Δ, we sample a pair from the joint distribution of X and its running maximum

M at time Δ (or running minimum when starting from above K). Using this pair we can determine the increase of S on a time interval of length Δ together with its running maximum over this time interval, which in turn allows us to redefine $\widetilde{\tau}$ as, say, the middle point of the time interval under consideration if the running maximum indeed indicates that S has actually crossed K somewhere on this time interval (and S has not crossed K before). Of course taking the middle point of the time interval is pretty arbitrary, but the error is at most $\Delta/2 = 0.005$ and the exact time at which S hits K only enters, multiplied by $-r = -0.06$, linearly. Thus, this procedure seems reasonable enough.

The actual sampling from the pair (X_Δ, M_Δ) is pretty straightforward. Making the substitution $Z_1 = 2M_\Delta - X_\Delta$ and $Z_2 = X_\Delta$ the density function of (Z_1, Z_2) becomes

$$f(z_1, z_2) = \mathbf{1}_{\{z_1 \geq |z_2|\}} \frac{\exp\left(-\mu^2\Delta/(2\sigma^2)\right)}{\sqrt{2\pi\sigma^6\Delta^3}} z_1 \exp\left(-\frac{z_1^2}{2\sigma^2\Delta} + \frac{\mu}{\sigma^2} z_2\right).$$

Of course the marginal distribution of Z_2 is still $\mathcal{N}(\mu\Delta, \sigma^2\Delta)$ and as it is clear from the above formula, conditional on $Z_2 = z_2$ the random variable Z_1 follows a Weibull distribution conditioned on being not less than $|z_2|$. Since we can explicitly calculate the inverse of the distribution function of this distribution we can sample $Z_1 | Z_2 = z_2$ by standard means.

Now, if we again run the tests from the previous paragraph, we find that with this improved definition the empirical probability of the set $\{\widetilde{\tau} \leq 0.5\}$ becomes 0.4152. So the deviation from the corresponding probability with respect to τ has improved from 11% to 0.7%. Furthermore, on the issue of the process $M^{(2)}$ being a martingale in the discrete time setting, doing the same tests as above again reveals that on average the difference $|\widehat{E}(M_T^{(2)}) - M_0^{(2)}|$ is even closer to 0 than $|\widehat{E}(M_T^{(1)}) - M_0^{(1)}|$.

Here are the numerics with the martingales $M^{(1)}$ and $M^{(2)}$:

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation	$\lambda_*^{(1)}$	$\lambda_*^{(2)}$
80	20.6	20.7	0.32	0.17	1.00	-0.10
90	12.4	11.7	1.20	9.77	1.00	-0.70
110	3.64	3.72	0.14	0.21	0.58	0.03
120	2.54	2.58	0.24	0.08	0.63	0.05

Table 3: Numerical values with same parameter values as in table 1

There is some improvement over the values from the previous table. Apparently the starting point 90 stays an exception as it has a much larger variance and deviation from the true value than the other ones.

3.2 The convertible bond

A convertible bond is a game contingent claim with payoff processes

$$X_t = \gamma e^{-rt} S_t \qquad \text{and} \qquad Y_t = e^{-rt}(K \vee \gamma S_t)$$

for $t < T$ and $X_T = Y_T = \exp(-rT)(1 \vee \gamma S_T)$, where $K > 1$ and $0 < \gamma < 1$. This game is called a convertible bond because it can be interpreted as the maximizer having the right to convert this contract into γ stocks, with a guaranteed minimum payment of 1 at $t = T$. The minimizer can recall the bond at price K, but then the holder (maximizer) obtains the opportunity to convert the bond immediately.

The stock price S is now being modelled by a jump-diffusion with non-negative, exponentially distributed jumps, including dividend payment at rate δ with $0 < \delta < r$. That is, we have that

$$S_t = \exp\left(\sigma W_t + \mu t + J_t\right), \tag{3.2}$$

where W is again a Brownian motion and J is a compound Poisson process independent of W with jump intensity $\eta > 0$ and increments that follow an exponential distribution with parameter $\theta > 1$ (which ensures finite expectations). Note that if we set the drift $\mu := r - \delta - \sigma^2/2 + \eta/(1-\theta)$, then $\mathbb{P}$ is an equivalent martingale measure, in the sense that the process $(\exp(-rt)S_t + \int_0^t \exp(-ru)\delta S_u\, du)_{t\geq 0}$ is a $\mathbb{P}$-martingale.

Throughout, the optimal exercising strategy of the maximizer is such that he stops as soon as the stock price *exceeds* a time-depending level (called exercise boundary, see the doted line in Figure 2). If time to go is very short, the maximizer will exercise right away if $S_0 > 1/\gamma$ but will wait if $S_0 < 1/\gamma$. Also the minimizer exercises once the stock price *exeeds* an exercise boundary (see the solid line in Figure 2). Coincidence of an exercise boundary with the upper line which is the level K/γ means that there is no "premature" stopping. On the other hand, $S_t \geq K/\gamma$ implies $X_t = Y_t$ and the game is stopped anyway. From Figure 2 it can be seen that never both exercise boundaries lie below K/γ at the same time, but it may happen that there is temporarily no premature exercising. The fact that the minimizer may recall at stock prices strictly smaller than K/γ is caused by the positive jumps in (3.2); see Gapeev and Kühn (2005) for a discussion of this phenomenon in the perpetual case. A plot of the corresponding value function can be seen in Figure 3.

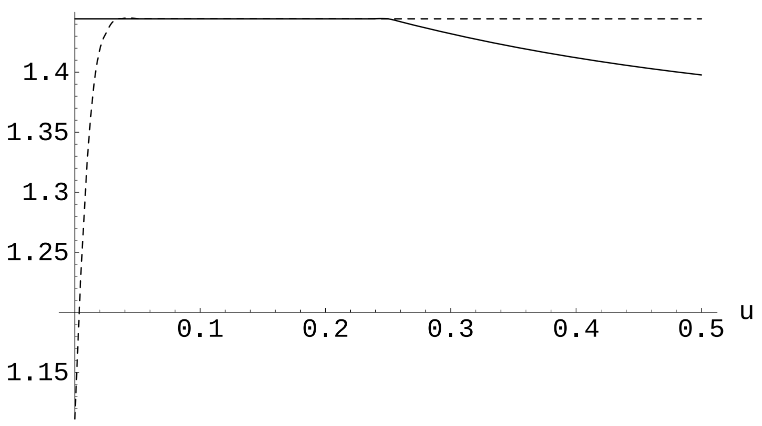

Figure 2: The optimal exercise boundaries (the dotted line is for the maximizer and the solid line for the minimizer). The u-axes represents time to maturity and the upper line is the level K/γ.

Now we turn to the implementation of the pathwise approach. By adding an independent jump process to the geometric Brownian motion the variance increases and con-

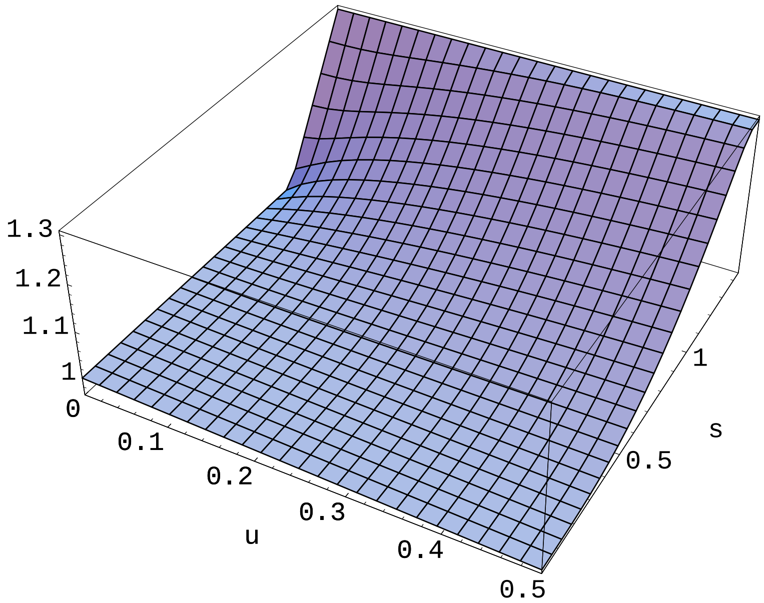

Figure 3: The corresponding value function of a convertible bond, parameters as in table 4. As in Figure 2 the plot is obtained by Canadization.

sequently our sampling variance will also be larger. Tests showed that in our setting doubling the number of sampled paths from 5000 to 10000 pretty much makes up for this. In the procedure of determining the minimizing coefficients with which our martingales are weighted we also doubled the number of sampled paths, from 300 to 600.

As for the callable put we start out with only the trivial 0 martingale, which results in the numerics as shown in table 4.

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation
0.8	1.060	1.031	0.0148	0.24
1	1.133	1.078	0.0225	0.37
1.2	1.224	1.139	0.0250	0.54
1.3	1.272	1.177	0.0231	0.63
1.4	1.300	1.237	0.0149	0.52

Table 4: Numerical values with no martingale and parameter values $\sigma = 0.4$, $r = 0.06$, $K = 1.3$, $T = 0.5$, $\delta = 0.02$, $\gamma = 0.9$, $\eta = 10$ and $\theta = 7$

The form of the lower payoff process X for $t < T$ seems to indicate that it makes sense to try the martingale $M^{(1)}$ defined by $M_T^{(1)} = \gamma \exp(-rT)S_T$. Since $(\exp(-(r-\delta)t)S_t)_{t\geq 0}$ is a $\mathbb{P}_s$-martingale we have that $M_t^{(1)} = \gamma \exp(-rt - \delta(T-t))S_t$ for $t < T$. With this martingale the approximation is indeed already significantly better:

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation	$\lambda_*^{(1)}$
0.8	1.060	1.047	0.0059	0.17	0.3440
1	1.133	1.113	0.0063	0.25	0.4087
1.2	1.224	1.199	0.0052	0.35	0.4834
1.3	1.272	1.241	0.0040	0.49	0.4988
1.4	1.300	1.279	0.0022	0.45	0.4885

Table 5: Numerical values with the same parameters as in table 4

Next it seems to make sense to bring in a martingale that has some connection with the upper payoff Y, therefore we take $M^{(2)}$ to be the martingale defined by $M_T^{(2)} = \exp(-rT)(K \vee \gamma S_T)$. We get $M_t^{(2)} = \exp(-rT)F(T-t, \gamma S_t)$, where $F(u, x) = \mathbb{E}_x[K \vee S_u]$, for $t < T$. Computing F is now of course more involved than its equivalent in the geometric Brownian motion case. A closed form formula is still available, but it turns out to be too complex to use it directly in our simulations in the sense that computing its value takes so long that a simulation of a large number of paths basically becomes unfeasible. Before addressing this issue we derive the formula for F, this can be done along the lines of the calculations in Kou (2002) e.g., see also this paper for more details on the functions defined below.

The density of a $\mathcal{N}(\mu u, \sigma^2 u)$ random variable plus n i.i.d. $\exp(\theta)$ distributed random variables is given by

$$t \mapsto \frac{\exp(\sigma^2\theta^2 u/2 - (t-\mu u)\theta)}{\sigma\sqrt{2\pi u}}(\sigma\theta\sqrt{u})^n \mathrm{Hh}_{n-1}\left(\sigma\theta\sqrt{u} - \frac{t-\mu u}{\sigma\sqrt{u}}\right),$$

where the Hh functions are given by

$$\mathrm{Hh}_n(x) = \int_x^\infty \mathrm{Hh}_{n-1}(y)\,\mathrm{d}y = \frac{1}{n!}\int_x^\infty (t-x)^n e^{-t^2/2}\,\mathrm{d}t \tag{3.3}$$

for $n = 0, 1, \ldots$ and $\mathrm{Hh}_{-1}(x) = \exp(-x^2/2)$.

By conditioning on the number of jumps of S up to time u and using that both the Brownian motion and jump heights in S are independent of this number we can write $F(u,x) = \mathbb{E}_x[K \vee S_u]$ as an infinite sum over integrals of $y \mapsto K \vee (x\exp(y))$ against the density above. This yields

$$\begin{aligned} F(u,x) \;&=\; e^{-\lambda u} K\Phi\left(\frac{\ln(K/x) - \mu u}{\sigma\sqrt{u}}\right) + x e^{(\sigma^2/2+\mu-\lambda)u}\Phi\left(\frac{(\mu+\sigma^2)u - \ln(K/x)}{\sigma\sqrt{u}}\right) \\ &\quad + \frac{\exp((\sigma^2\theta^2/2 + \mu\theta - \lambda)u)}{\sigma\sqrt{2\pi u}} \sum_{n\geq 1} \frac{(\sigma\theta\lambda u^{3/2})^n}{n!}\Bigg[x I_{n-1}\left(\ln(K/x); 1-\theta, \frac{-1}{\sigma\sqrt{u}},\right. \\ &\quad \left. -(\sigma\theta + \frac{\mu}{\sigma})\sqrt{u}\right) + K I_{n-1}\left(\ln(x/K); \theta, \frac{1}{\sigma\sqrt{u}}, -(\sigma\theta + \frac{\mu}{\sigma})\sqrt{u}\right)\Bigg], \end{aligned} \tag{3.4}$$

with Φ the cumulative distribution function of a standard normal and the functions I_n for $n \geq 0$ defined by

$$I_n(y; a, b, c) := \int_y^\infty e^{at}\mathrm{Hh}_n(bt - c)\,\mathrm{d}t.$$

By integration-by-parts we can get rid of the integral in this expression and write

$$I_n(y; a, b, c) = -\frac{e^{ay}}{a}\sum_{i=0}^n \left(\frac{\beta}{a}\right)^{n-i}\mathrm{Hh}_i(by - c) \pm \left(\frac{\beta}{a}\right)^{n+1}\frac{\sqrt{2\pi}}{b}\exp(\frac{ac}{b} + \frac{a^2}{2b^2})\Phi(\mp by \pm c \pm \frac{a}{b}),$$

where the $\pm$ is a $+$ if $b > 0$ and $a \neq 0$, and a $-$ if $b < 0$ and $a < 0$; vice versa for the $\mp$.

The first question when implementing these formulas for calculations is where to truncate the infinite sum involved in the expression of F. Kou suggests that only the first 10 to 15 terms are needed in most cases, we took 20 after doing some numerical tests with our parameters. Furthermore, we see that in order to calculate F for a single pair (u, x) we are required to compute two double sums in which each summand involves a numerical integration because of (3.3). This is perfectly doable for a just a few evaluations, but as remarked above it is too involved for our purpose as this involves 50 (namely, the number of time steps) evaluations of F for a single path and this times 10000 to perform one simulation of a game.

To overcome this difficulty we did the following. Fix some $u > 0$ for a moment. The function $f : x \mapsto F(u, x)$ is smooth, convex and increasing with $f(0+) = K$ and, on account of $(\exp(-(r-\delta)t)S_t)_{t\geq 0}$ being a $\mathbb{P}_s$-martingale, such that $f(x) - x\exp((r-\delta)u) \downarrow 0$ as $x \to \infty$. Furthermore, for $h > 0$ we have that

$$f(x+h) - f(x) = \mathbb{E}_1\left[\mathbf{1}_{\{S_u \in (K/(x+h), K/x]\}}((x+h)S_u - K)\right] + h\mathbb{E}_1\left[\mathbf{1}_{\{S_u > K/x\}} S_u\right]$$

and since the term in the first expectation stays bounded and tends to 0 when $h \downarrow 0$ it follows that $f'(x) = \mathbb{E}_1\left[\mathbf{1}_{\{S_u > K/x\}} S_u\right]$. Given this smooth structure and the fact that 10000 simulations would mean 10000 evaluations of f, it is clear that we can benefit from allowing a small error in evaluating f (in addition to the error that is already caused by the truncation of the sum and numerical integration) and taking a number of base points on the x-axis at which f is approximated by numerically evaluating, after which we interpolate to obtain approximative values for other x.

More precisely, we simply take linear interpolation and choose a strategy for determining the base points that allows us to control the maximal error made by interpolating explicitly. In order to do so, we use that due to the fact that S has only positive jumps, f' is bounded below by the function g given by

$$g(x) := \mathbb{E}_1\left[\mathbf{1}_{\{\tilde{S}_u > K/x\}} \tilde{S}_u\right],$$

where $\tilde{S}$ is the continuous part of S, that is $\tilde{S}_t = \exp(\sigma W_t + \mu t)$ for $t \geq 0$. Given two base points $x_k < x_{k+1}$, the maximal difference on the interval $[x_k, x_{k+1}]$ between the straight line l through the points $(x_k, f(x_k))$ and $(x_{k+1}, f(x_{k+1}))$ and the function f is bounded by the minimum of $f(x_{k+1}) - f(x_k)$ and the maximal difference on this interval between l

and G, where G is the antiderivative of g such that $G(x_k) = f(x_k)$. Note that G qualifies as a usable bound because we expect it to have a convex structure comparable with f and its maximal difference with l can be easily computed. Even though this can't be done explicitly, due to G being convex and smooth and having a representation in terms of Φ, this maximal difference can be very quickly and reliably approximated by Mathematica.

The exact implementation looks as follows. We start with an equidistant grid of base points with typically a grid distance of 0.5 and make sure that our largest base point is chosen such that the difference between f and $x \mapsto x\exp((r-\delta)u)$ is smaller than the maximal allowed error, which we chose to be 10^{-3}. After calculating f on these base points we check using the procedure described in the previous paragraph for each interval between two base points whether the bound on the error is already smaller than the maximal allowed error. If not, then we refine the grid by adding the middle point of this interval to it. This we repeat until the grid is fine enough. Obviously, typically f' will be flatter in the regions close to 0 and where x is large while it's more steep in the region around the point K/γ, which translates in the grid being finer around this point than in both other regions. This operation is feasible on an average PC although it still takes several hours to complete it for all $u \in \{0.01, 0.02, \dots, 0.5\}$.

Having all this established, we can return to our actual goal, namely using the martingale $M^{(2)}$ in the simulation of the pathwise game. Since the operation for establishing its values introduces numerical errors to the true value at different stages, one could wonder how much this influences the martingale property of the sampled process. We checked this by running 10 simulations of 10000 paths each from the different starting points also used in tables 4 and 5. As before, we look at how much the quantity $\widehat{E}(M_T^{(2)}) - M_0^{(2)}$ differs from 0 as a measure for $M^{(2)}$ being a martingale. Note in this perspective also that $M_T^{(2)}$ is determined without using any numerical approximation to F while $M_0^{(2)}$ is deterministic and calculating its value does make use of the numerical approximation to $F(T, \cdot)$. Reassuringly, it turns out that for all tests and different starting points, $|\widehat{E}(M_T^{(2)}) - M_0^{(2)}|$ is even smaller than $|\widehat{E}(M_T^{(1)}) - M_0^{(1)}|$. This may seem surprising at first, but it makes more sense if one thinks about it by looking at $M_t^{(1)}$ also as a function of the underlying random variable S_t. This function has a constant derivative $\gamma\exp(-rt - \delta(T-t))$, whereas for $M_t^{(2)}$ this derivative boils down to $\gamma \exp(-rT)(\partial F/\partial x)(T-t, \gamma S_t)$ and using the properties of f' we discussed before we see that this function is small for S_t close to 0 and it's bounded above by $\gamma^2 \exp(-rt - \delta(T-t))$, so in fact this derivative is always smaller. This implies that $M_t^{(2)}$ has a smaller variance than $M_t^{(1)}$ and as the numerical tests indicate, the numerical errors are of smaller order than this effect.

Now we are ready to move on to the actual simulation of the pathwise game using the martingale $M^{(2)}$. The results are as follows:

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation	$\lambda_*^{(2)}$
0.8	1.060	1.059	0.0015	0.03	1.4460
1	1.133	1.131	0.0019	0.05	1.3598
1.2	1.224	1.219	0.0012	0.14	1.3003
1.3	1.272	1.263	0.0006	0.37	1.2637
1.4	1.300	1.294	0.0001	0.60	1.1627

Table 6: Numerical values with the same parameters as in table 4

We see that this martingale gives a very good approximation, considerably better than $M^{(1)}$ did. Some explanation might be found in the fact that the payoff X_T of the maximizer at T looks, due to the face value 1, similar to $M_T^{(2)}$ (recall that $K = 1.3$) while on the other hand $M^{(1)}$ and Y look less similar.

Finally we see whether we can still improve on the values in the previous table by taking linear combinations of $M^{(1)}$ and $M^{(2)}$:

S_0	True value	Average of pathwise values	Variance of pathwise values	Absolute error / standard deviation	$\lambda_*^{(1)}$	$\lambda_*^{(2)}$
0.8	1.060	1.060	0.0012	−−	0.0617	1.2734
1	1.133	1.132	0.0014	0.03	0.0980	1.1233
1.2	1.224	1.220	0.0010	0.13	0.0959	1.0991
1.3	1.272	1.265	0.0006	0.29	0.0299	1.2035
1.4	1.300	1.294	0.0001	0.60	0.0074	1.1454

Table 7: Numerical values with the same parameters as in table 4

We see there is some improvement here but not too much. Note that the two starting values with the least good approximations are the ones closest to the level K/γ. Furthermore it can be seen in the plots (and is also indicated by the value $1.300 = K$ in the tables) that when starting from $S_0 = 1.4$, the minimizer recalls right away. One could interpret this as an indication that what's still mainly missing in our set of hedging martingales is a martingale that is in some sense associated with the payoffs that occur already after short times, when S quickly enters the early exercise region e.g.

References

[1] Bank, P. and El Karoui, N. (2004) A Stochastic Representation Theorem with Applications to Optimization and Obstacle Problems. *Annals of Probability* **32,** 1030–1067.

[2] Brennan, M.J. and Schwartz, E.S. (1977) Convertible bonds: valuation and optimal strategies for call and conversion. *Journal of Finance* **32,** 1699–1715.

[3] Broadie, M. and Glasserman, P. and Kou, S.G. (1999) Connecting discrete and continuous path-dependent options. *Finance and Stochastics* **3,** 55–82.

[4] Carr, P. (1998) Randomization and the American Put. *Review of Financial Studies* **11,** 597–626.

[5] Cvitanić, J. and Karatzas, I. (1996) Backward stochastic differential equations with reflection and Dynkin games. *Annals of Probability* **24,** 2024–2056.

[6] El Karoui, N. (1981) Les aspects probabilistes du contrôle stochastique. *In: Ninth Saint Flour Probability Summer School (Saint Flour 1979)*, 73–238.

[7] Fakeev, A. (1970) Optimal Stopping Rules for Stochastic Processes with Continuous Parameters. *Theory of Probability and its Applications* **15,** 324–331.

[8] Gapeev, P.V. and Kühn, C. (2005) Perpetual convertible bonds in jump-diffusion models. *Statistics & Decisions* **23,** 15–31.

[9] Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering.* Springer.

[10] Ingersoll, J.E. (1977a) A contingent-claims valuation of convertible securities. *Journal of Financial Economics* **4,** 289–322.

[11] Ingersoll, J.E. (1977b) An examination of corporate call policies on convertible securities. *Journal of Finance* **32,** 463–478.

[12] Jacod, J. and Shiryayev, A.N. (2003) *Limit Theorems for Stochastic Processes.* Springer. 2.Auflage.

[13] Kallsen, J. and Kühn, C. (2004) Pricing Derivatives of American and Game Type in Incomplete Markets. *Finance & Stochastics* **8,** 261–284.

[14] Kallsen, J. and Kühn, C. (2005) Convertible bonds: financial derivatives of game type. *In: Exotic Option Pricing and Advanced Lévy Models.* Edited by Kyprianou, A., Schoutens, W., and Wilmott, P., Wiley, Chichester, 277–291.

[15] Kifer, Y. (2000) Game options. *Finance & Stochastics* **4,** 443–463.

[16] Kolodko, A. and Schoenmakers, J. (2004) Upper Bounds for Bermudan Style Derivatives. *Monte Carlo Methods and Applications* **10,** 331–343.

[17] Kou, S.G.(2002) A Jump-Diffusion Model for Option Pricing. *Management Science***48,** 1086–1101.

[18] Kramkov, D. (1996) Optional decomposition of supermartingales and hedging contingent claims in incomplete markets. *Probability Theory and Related Fields* **105,** 459–479.

[19] Kühn, C. and Kyprianou, A.E. (2006) Callable puts as composite exotic options. *Preprint.*

[20] Kyprianou, A.E. (2004) Some calculations for Israeli options. *Finance & Stochastics* **8,** 73–86.

[21] Lepeltier, J. and Maingueneau, M. (1984) Le jeu de Dynkin en théorie générale sans l'hypothèse de Mokobodski. *Stochastics* **13,** 25–44.

[22] McConnell, J.J. and Schwartz, E.S. (1986) LYON Taming. *Journal of Finance* **41,** 561–576.

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

[24] Sirbu, M. and Pikovsky, I. and Shreve, S. (2004) Perpetual Convertible Bonds. *SIAM J. Control Optim.* **43,** 58–85.