Discussion Paper No. 029

Pricing of Path-Dependent American Options by Monte Carlo Simulation

Hajime Fujiwara
and
Masaaki Kijima

May, 2004

21COE
Interfaces for Advanced Economic Analysis
Kyoto University

Pricing of Path-Dependent American Options by Monte Carlo Simulation *

Hajime Fujiwara[a,b] and Masaaki Kijima[a]

[a] *Graduate School of Economics, Kyoto University*

[b] *Quantitative Analysis & IT Department, Mitsubishi Securities Co., Ltd.*

(This Version: May 18, 2004)

Abstract. In this paper, we evaluate anytime Bermudan options, a variant of path-dependent American options, by Monte Carlo simulation. Assuming that the underlying state variable is Markovian, we show that the price of an anytime Bermudan option satisfies a dynamic programming equation. The continuation value in the dynamic programming problem is represented by a conditional expectation. It is shown that the conditional expectation can be transformed to an uncoditional expectation, using the Malliavin calculus, which in turn enables us to evaluate the price of the anytime Bermudan option by Monte Carlo methods. Some numerical examples are given to demonstrate the usefulness of our method.

Keywords: Malliavin calculus, anytime Bermudan option, dynamic programming, callable floating-rate bond.

*Send all correspondence to Masaaki Kijima, Graduate School of Economics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan. TEL:+81-75-753-3456, Email: kijima@econ.kyoto-u.ac.jp

This research is supported in part by a Grant-in-Aid for the 21st Century COE research. One of the authors (M.K.) is supported in part by Daiwa Securities Group Inc.

1 Introduction

This paper considers the pricing of "anytime Bermudan" options, a variant of path-dependent American options, by Monte Carlo simulation.

In order to explain anytime Bermudan options, consider two time epochs, $T_1 < T_2$ say. The cashflow at time T_2 of an anytime Bermudan option depends on the value at time T_1 of the underlying state variable and, unlike ordinary Bermudan options, holders of the option can exercise their rights in any time between T_1 and T_2. Hence, the anytime Bermudan options are American-type options with the "look-back" feature. A typical example of such financial products actually traded in the market is a class of callable, floating-rate bonds.

An American-type option can be formulated by a dynamic programming after an appropriate discretization in time. Any standard method can be applied to solve the dynamic programming problem backwards in time, starting from the maturity of the option.[1] However, such methods do not work if the option has a path-dependent property, simply because the option price depends on the past and, when calculated backwards, the value of the underlying state variable is not yet determined. This forward feature cannot be handled by the standard dynamic programming approach.

On the other hand, Monte Carlo methods are suitable for the evaluation of path-dependent European options, because sample paths of the underlying state variable can be generated forwards in time with ease. Various methods have been proposed to accelarate the speed of convergence in Monte Carlo methods.[2]

Monte Carlo methods cannot be applied to American-type options directly because of the backward feature of the dynamic programming approach. More precisely, in the ordinary American option, the option value in the continuation region can be represented as a conditional expectation of the underlying state variable, $\mathrm{E}[h(X(T))|X(t) = x]$ say, $t < T$. Monte Carlo methods generate sample paths forwards in time and, therefore, we must calculate the conditional expectations for all possible $X(t) = x$.

Monte Carlo methods can be applied to American-type options only if the conditional expectations are evaluated efficiently. Recently, various methods have been proposed for this purpose.[3] Among them, the Malliavin calculus transforms the conditional expectations into unconditional expectations, and Monte Carlo methods can be used directly to evaluate

[1] The standard methods include the finite difference method and the tree method. See, e.g., Rebonato (1998) for details of such standard methods.

[2] Not only various types of variance reduction methods, but also the use of low-discrepancy seuqences can accelarate the speed of convergence. See Jackel (2002) for details.

[3] Such methods include regression-based methods, the duality approach, and stochastic mesh methods. See Glasserman (2003) for details.

the unconditional expectations. See, e.g., Fournie et al. (2001), Bally, Caramellino and Zanette (2003), Bouchard, Ekeland and Touzi (2003), and Mrad, Touzi and Zeghal (2003) for applications of the Malliavin calculus to Monte Carlo methods in finance.

The option value in the continuation region of an anytime Bermudan option can also be represented as a *bivariate* conditional expectation of the underlying state variable, $\mathrm{E}[h(X(T))|X(t_1) = x_1, X(t_2) = x_2]$ say, $t_1 < t_2 < T$. Monte Carlo methods can be used if the conditional expectation can be transformed into the unconditional counterpart. In this paper, we demonstrate that this idea indeed works even for the bivariate case by extending results in Mrad, Touzi and Zeghal (2003).

This paper is organized as follows. In the next section, we formulate the model to evaluate an anytime Bermudan option by the dynamic prigramming approach. The value function in the continuation region is given by a bivariate conditional expectation. It is shown in Section 3 that, using the Malliavin calculus, the bivariate conditional expectation can be represented by an unconditional expectation. These results are then applied to the pricing of anytime Bermudan options by Monte Carlo methods. Section 4 describes the algorithm and numerical examples. Section 5 counclude this paper, and the proof is given in Appendix.

2 The Model

Throughout this paper, we consider a d-dimensional standard Brownian motion $W = \{W_t; 0 \leq t \leq T\}$, $T < \infty$, defined on a complete probability space $(\Omega, \mathcal{F}, P)$ and the filtration $\mathcal{F} = \{\mathcal{F}_t; 0 \leq t \leq T\}$ generated from the Brownian motion W, i.e. $\mathcal{F}_t = \sigma(W_s; 0 \leq s \leq t)$. The underlying state variable is denoted by $X = \{X_t; 0 \leq t \leq T\}$. It is assumed throughout that X satisfies the following stochastic differential equation (SDE for short)

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \qquad X_0 = x_0. \tag{1}$$

The instantaneous interest rate is given by $r_t = r(X_t)$. Hence, both the state variable and the interest rate are adapted to the filtration $\mathcal{F}$.

Consider an anytime Bermudan option with cashflow C_i paid at time T_i, where $0 = T_0 < T_1 < \cdots < T_i < \cdots < T_N = T$. Here, T_0 denotes the current time and we assume without loss of generality that $C_0 = 0$. Recall that, in the anytime Bermudan option, the cashflow C_i is determined at time T_{i-1}. Hence, in our setting, C_i depends on the value of $X_{T_{i-1}}$, i.e. $C_i = C_i(X_{T_{i-1}})$. In the following, we shall denote $X_{i-1} \equiv X_{T_{i-1}}$ for notational simplicity.

Consider the time interval $[T_{i-1}, T_i)$, and denote the class of stopping times in $[T_{i-1}, T_i)$ by $\mathcal{T}_i$. The value of the anytime Bermudan option at time t is denoted by V_t. In particular, we denote $V_i = V_{T_i}$, $i = 0, 1, \ldots, N$.

Suppose $X_{i-1} = x$, and let $h_i(x, X_t)$ denote the payoff of the anytime Bermudan option if exercised at time $t \in [T_{i-1}, T_i)$. If not exercised in the interval, it is given by $h_i(x, X_i) = V_i$. The price of the anytime Bermudan option at time $t \in [T_{i-1}, T_i)$ is then given by

$$V_t = \begin{cases} \sup_{\tau \in \mathcal{T}_i} \mathrm{E}\left[\mathrm{e}^{-\int_t^\tau r_s \mathrm{d}s} h_i(X_{i-1}, X_\tau) \right], & T_{i-1} < t < T_i, \\ C_{i-1} + \sup_{\tau \in \mathcal{T}_i} \mathrm{E}\left[\mathrm{e}^{-\int_t^\tau r_s \mathrm{d}s} h_i(X_{i-1}, X_\tau) \right], & t = T_{i-1}. \end{cases} \tag{2}$$

Since $V_N = C_N(X_{N-1})$, the option price (2) for the interval $[T_{N-1}, T_N)$ can be obtained backwards by the ordinary dynamic programming approach if the value of X_{N-1} is known. In general, the option price (2) for the interval $[T_{i-1}, T_i)$ can be obtained by the ordinary dynamic programming approach if the value X_{i-1} is known. In order to deal with this forward feature, we believe that the use of Monte Carlo methods is most efficient. The backward feature arising from the dynamic programming approach can be resolved by the use of the Malliavin calculus.

More specifically, we discretize the time interval $[T_{i-1}, T_i]$ as $T_{i-1} = t_0^i < t_1^i < \cdots < t_{M-1}^i < t_M^i = T_i$.[4] Suppose $X_{i-1} = x$, and let $DV_j^i(x, X_j^i)$ denote the value of the anytime Bermudan option at time t_j^i in the discrete-time setting, where $X_j^i \equiv X_{t_j^i}$ for notational convenience. Let us define

$$CV_j^i(x, y) = \mathrm{E}\left[\mathrm{e}^{-r(X_j^i)(t_{j+1}^i - t_j^i)} DV_{j+1}^i(x, X_{j+1}^i) \middle| X_{i-1} = x,\ X_j^i = y \right]. \tag{3}$$

The quantity $CV_j^i(x, y)$ represents the time t_j^i value of the anytime Bermudan option if not exercised at that time. It follows that

$$DV_j^i(x, X_j^i) = \max\{h_i(x, X_j^i),\ CV_j^i(x, X_j^i)\}, \tag{4}$$

where $X_{i-1} = x$. The boundary condition at $j = M$, i.e. at the coupon payment date T_i, is given by

$$DV_M^i(x, X_M^i) = DV_0^{i+1}(X_i, X_i) + C_i(x), \tag{5}$$

where $X_{i-1} = x$. Starting from $DV_M^N = C_N(X_{N-1})$ and repeating this procedure, one can finally arrive at the current price DV_0^0 of the anytime Bermudan option.

In the next section, we show that the *bivariate* conditional expectation given in (3) can be represented as an unconditional expectation with the help of the Malliavin calculus. The set of equations (3)–(5) can then be evaluated by Monte Carlo methods.

[4] If the meaning is clear, we denote the time epochs t_j^i simply by t_j.

3 The Malliavin Calculus

In this section, we consider the conditional expectation

$$\mathrm{E}[f(X_{t_3}, x)|X_{t_1} = x, X_{t_2} = y], \qquad 0 < t_1 < t_2 < t_3, \tag{6}$$

where $f : R^d \times R^d \mapsto R$ is assumed to be continuous, for which the conditional expectation exists. In the following, for the sake of notational simplicity, we denote $X_j = X_{t_j}$, $j = 1, 2, 3$.

Let $\delta_x(X)$ denote the delta function, meaning that $\delta_x(X) = 1$ if $X = x$ and $\delta_x(X) = 0$ otherwise. The conditional expectation (6) is represented as

$$\mathrm{E}[f(X_3, x)|X_1 = x, X_2 = y] = \frac{\mathrm{E}[f(X_3, x)\delta_x(X_1)\delta_y(X_2)]}{\mathrm{E}[\delta_x(X_1)\delta_y(X_2)]}. \tag{7}$$

This expression is not useful for Monte Carlo methods, however, because the value in (7) is undefined unless $X_1 = x$ and $X_2 = y$, and the probability that the sample paths in Monte Carlo simulation hit the event is zero, almost surely. In the following, we obtain another expression that are suitable for Monte Carlo methods by extending results in Mrad, Touzi and Zeghal (2003).

Throughout this section, we assume that there exist matrix-valued functions g_t and h_t satisfying

$$\int_0^{t_3} D_t X_1 g_t \mathrm{d}t = 0, \quad \int_0^{t_3} D_t X_2 g_t \mathrm{d}t = I_d, \quad \int_0^{t_3} D_t X_3 g_t \mathrm{d}t = 0 \tag{8}$$

and

$$\int_0^{t_3} D_t X_1 h_t \mathrm{d}t = I_d, \quad \int_0^{t_3} D_t X_2 h_t \mathrm{d}t = 0, \quad \int_0^{t_3} D_t X_3 h_t \mathrm{d}t = 0, \tag{9}$$

respectively, where I_d denotes the d-dimensional identity matirx and where D_t represents the Malliavin derivative.[5]

For any random variable F defined on R, let

$$S_i^h(F) \equiv \int_0^{t_3} F h_t^i \cdot \mathrm{d}W_t, \qquad i = 1, \dots, d, \tag{10}$$

where h_t^i denotes the ith column vector of h_t and where $x \cdot y = \sum_{i=1}^d x_i y_i$ for x, $y \in R^d$. For integers $i_1, i_2, \dots, i_k$, let

$$S_{i_1, i_2, \dots, i_k}^h(F) \equiv S_{i_1}^h \circ S_{i_2}^h \circ \cdots \circ S_{i_k}^h(F), \tag{11}$$

where $f \circ g(x) = f(g(x))$ for f, $g : R \mapsto R$. In particular, we shall write $S^h(F) = S_{1,\dots,d}^h(F)$. Also, $1_A(X)$ denotes the indicator function of A, meaning that $1_A(X) = 1$ if $X \in A$ and $1_A(X) = 0$ otherwise, while $H_x(X)$ represents the Heaviside function, meaning that $H_x(X) = 1$ if $X \geq x$ and $H_x(X) = 0$ otherwise.

We then have the following. The proof is given in the appendix.

[5] See, e.g., Nualart (1995) for the information of the Malliavin calculus.

Proposition 1 *For $a_i, b_i, c_i, d_i \in R$ with $a_i < b_i$, $c_i < d_i$, define $A_i = [a_i, b_i]$ and $B_i = [c_i, d_i]$, $i = 1, \ldots, d$. Let $A \equiv A_1 \times \cdots \times A_d$, $B \equiv B_1 \times \cdots \times B_d$, and let $z \in A$ be any constant. For g_t and h_t satisfying (8) and (9), respectively, we have*

$$\begin{aligned}&\mathrm{E}\left[f(X_3, z)1_A(X_1)1_B(X_2)\right] \\ &\quad = \int_{A\times B} \mathrm{E}\left[H_x(X_1)H_y(X_2)f(X_3,z)S^g \circ S^h(\varphi(X_1 - x)\psi(X_2 - y))\right] \mathrm{d}x\mathrm{d}y,\end{aligned}$$

where $\varphi(x) = \exp(-\eta \cdot x)$ and $\psi(y) = \exp(-\rho \cdot y)$ for some constants η, $\rho \in R^d$.

From Proposition 1, the conditional expectation (7) can be expressed as

$$\begin{aligned}&\mathrm{E}\left[f(X_3, x)|X_1 = x, X_2 = y\right] \\ &\quad = \frac{\mathrm{E}\left[H_x(X_1)H_y(X_2)f(X_3,x)S^g \circ S^h(\varphi(X_1 - x)\psi(X_2 - y))\right]}{\mathrm{E}\left[H_x(X_1)H_y(X_2)S^g \circ S^h(\varphi(X_1 - x)\psi(X_2 - y))\right]}.\end{aligned} \tag{12}$$

The functions $\varphi(x)$ and $\psi(y)$ are called the *localizing functions* and determined so that the variance of the denominator in (12) is minimized. According to Theorem 4.1 in Mrad, Touzi and Zeghal (2003), the parameters η, $\rho \in R^d$ of the localizing functions are given by

$$\eta_i = \sqrt{\frac{\mathrm{E}\left[\int_{R^{2d-1}} H_{x^{-i}}(X_1^{-i})H_y(X_2)S^g \circ S^h(F(x^{-i})G(y))^2 \mathrm{d}x^{-i}\mathrm{d}y\right]}{\mathrm{E}\left[\int_{R^{2d-1}} H_{x^{-i}}(X_1^{-i})H_y(X_2)S^g \circ S^h_{-i}(F(x^{-i})G(y))^2 \mathrm{d}x^{-i}\mathrm{d}y\right]}} \tag{13}$$

and

$$\rho_i = \sqrt{\frac{\mathrm{E}\left[\int_{R^{2d-1}} H_{y^{-i}}(X_2^{-i})H_x(X_1)S^g \circ S^h(G(y^{-i})F(x))^2 \mathrm{d}y^{-i}\mathrm{d}x\right]}{\mathrm{E}\left[\int_{R^{2d-1}} H_{y^{-i}}(X_2^{-i})H_x(X_1)S^g_{-i} \circ S^h(G(y^{-i})F(x))^2 \mathrm{d}y^{-i}\mathrm{d}x\right]}} \tag{14}$$

for $i = 1, \ldots, d$, respectively, where $F(x) \equiv \varphi(X_1 - x)$ and $G(y) \equiv \psi(X_2 - y)$. Here, $x^{-i} \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$ and $S^h_{-i}(F) = S^h_{1,\ldots,i-1,i+1,\ldots,d}(F)$.

It remains to specify the matrix-valued functions g_t, h_t satisfying (8) and (9), respectively. It is readily checked by direct substitution that, when X is a diffusion process satisfying (1), they are given by

$$g_t = (D_t X_3)^{-1} Y_3 Y_2^{-1} \left[\frac{1}{t_2 - t_1} 1_{(t_1,t_2)}(t) - \frac{1}{t_3 - t_2} 1_{[t_2,t_3)}(t)\right] \tag{15}$$

and

$$h_t = (D_t X_3)^{-1} Y_3 Y_1^{-1} \left[\frac{1}{t_1} 1_{[0,t_1)}(t) - \frac{1}{t_2 - t_1} 1_{[t_1,t_2]}(t)\right], \tag{16}$$

where $Y_j \equiv Y_{t_j}$, $j = 1, 2, 3$. The matrix-valued process $Y = \{Y_t; 0 \le t \le T\}$ is called the *first variation process* of X, i.e. $Y_t = \partial X_t / \partial x_0$, and satisfies the SDE

$$\mathrm{d}Y_t = \nabla b(X_t) Y_t \mathrm{d}t + \sum_{i=1}^{d} \nabla \sigma^i(X_t) Y_t \mathrm{d}W_t^i. \tag{17}$$

The Malliavin derivative of X_t can be written as

$$D_s X_t = Y_t Y_s^{-1} \sigma(X_s) 1_{[0,t]}(s). \tag{18}$$

It follows from (15) and (16) that the functions g_t and h_t are expressed as

$$g_t = (\sigma(X_t))^{-1} Y_t Y_2^{-1} \left[\frac{1}{t_2 - t_1} 1_{(t_1,t_2)}(t) - \frac{1}{t_3 - t_2} 1_{[t_2,t_3)}(t) \right] \tag{19}$$

and

$$h_t = (\sigma(X_t))^{-1} Y_t Y_1^{-1} \left[\frac{1}{t_1} 1_{[0,t_1)}(t) - \frac{1}{t_2 - t_1} 1_{[t_1,t_2]}(t) \right], \tag{20}$$

respectively. It should be noted that these functions can be evaluated by Monte Carlo methods, and so are the integrals $S^g \circ S^h(\cdot)$.

Before proceeding, we provide two examples. These examples will be used later for numerical examples.

Example 1 Suppose $d = 1$, and consider the Ornstein–Uhlenbeck process satisfying the SDE

$$\mathrm{d}r_t = (\theta - a r_t)\mathrm{d}t + \sigma \mathrm{d}W_t,$$

where $a, \theta > 0$. The solution to the SDE is given by

$$r_t = \frac{\theta}{a} + \mathrm{e}^{-at}\left(r_0 - \frac{\theta}{a}\right) + \sigma \mathrm{e}^{-at} \int_0^t \mathrm{e}^{as} \mathrm{d}W_s.$$

From (17), the first variation process Y satisfies the SDE

$$\frac{\mathrm{d}Y_t}{Y_t} = \frac{\partial}{\partial r}(\theta - ar)\mathrm{d}t + \frac{\partial}{\partial r}\sigma \mathrm{d}W_t, \qquad Y_0 = 1.$$

It follows that $Y_t = \mathrm{e}^{-at}$, $t \geq 0$. Hence, from (19) and (20), we obtain

$$g_t = \frac{1}{\sigma}\mathrm{e}^{a(t_2 - t)} \left[\frac{1}{t_2 - t_1} 1_{(t_1,t_2)}(t) - \frac{1}{t_3 - t_2} 1_{[t_2,t_3)}(t) \right]$$

and

$$h_t = \frac{1}{\sigma}\mathrm{e}^{a(t_1 - t)} \left[\frac{1}{t_1} 1_{[0,t_1)}(t) - \frac{1}{t_2 - t_1} 1_{[t_1,t_2]}(t) \right],$$

respectively. The integration by parts formula in the Malliavin calculus together with (8) and (9) then yields

$$\begin{aligned} &S^h(\varphi(r_{t_1} - x)\psi(r_{t_2} - y)) \\ &\quad = \varphi(r_{t_1} - x)\psi(r_{t_2} - y) \left[S^g(1) S^h(1) + \rho S^h(1) + \eta S^g(1) + \eta\rho \right], \end{aligned}$$

where

$$S^h(1) = \frac{\mathrm{e}^{at_1}}{\sigma t_1} \int_0^{t_1} \mathrm{e}^{-at} \mathrm{d}W_t - \frac{\mathrm{e}^{at_1}}{\sigma(t_2 - t_1)} \int_{t_1}^{t2} \mathrm{e}^{-at} \mathrm{d}W_t$$

and

$$S^g(1) = \frac{\mathrm{e}^{at_2}}{\sigma(t_2 - t_1)} \int_{t_1}^{t_2} \mathrm{e}^{-at} \mathrm{d}W_t - \frac{\mathrm{e}^{at_2}}{\sigma(t_3 - t_2)} \int_{t_2}^{t_3} \mathrm{e}^{-at} \mathrm{d}W_t.$$

Example 2 In this example, we consider the two-dimensional case

$$\begin{aligned}
\mathrm{d}x_t^1 &= (\theta - a_1 x_t^1)\mathrm{d}t + \sigma_1 \mathrm{d}W_t^1, \\
\mathrm{d}x_t^2 &= \left(-\frac{\theta}{a_2 - a_1} - a_2 x_t^2\right)\mathrm{d}t - \frac{\sigma_1}{a_2 - a_1}\mathrm{d}W_t^1 + \sigma_2 \mathrm{d}W_t^2,
\end{aligned}$$

where $a_1, a_2, \theta > 0$ and $a_1 \neq a_2$. The solution to the SDE is given by

$$x_t^1 = \frac{\theta}{a_1} + \mathrm{e}^{-a_1 t}\left(x_0^1 - \frac{\theta}{a_1}\right) + \sigma_1 \mathrm{e}^{-a_1 t}\int_0^t \mathrm{e}^{a_1 s}\mathrm{d}W_s^1$$

and

$$\begin{aligned}
x_t^2 &= -\frac{\theta}{a_2(a_2 - a_1)} + \mathrm{e}^{-a_2 t}\left(x_0^2 + \frac{\theta}{a_2(a_2 - a_1)}\right) \\
&\quad - \frac{\sigma_1}{a_2 - a_1}\mathrm{e}^{-a_2 t}\int_0^t \mathrm{e}^{a_2 s}\mathrm{d}W_s^1 + \sigma_2 \mathrm{e}^{-a_2 t}\int_0^t \mathrm{e}^{a_2 s}\mathrm{d}W_s^2.
\end{aligned}$$

As in Example 1, the matrix-valued functions g_t, h_t are calculated as

$$g_t = G_t\left[\frac{1}{t_2 - t_1}1_{(t_1,t_2)}(t) - \frac{1}{t_3 - t_2}1_{[t_2,t_3)}(t)\right]$$

and

$$h_t = H_t\left[\frac{1}{t_1}1_{[0,t_1)}(t) - \frac{1}{t_2 - t_1}1_{[t_1,t_2]}(t)\right],$$

where G_t and H_t are the 2×2 matrices given by

$$G_t = \begin{pmatrix} \frac{1}{\sigma_1}\mathrm{e}^{a_1(t_2 - t)} & 0 \\ \frac{1}{(a_2 - a_1)\sigma_2}\mathrm{e}^{a_1(t_2 - t)} & \frac{1}{\sigma_2}\mathrm{e}^{a_2(t_2 - t)} \end{pmatrix}$$

and

$$H_t = \begin{pmatrix} \frac{1}{\sigma_1}\mathrm{e}^{a_1(t_1 - t)} & 0 \\ \frac{1}{(a_2 - a_1)\sigma_2}\mathrm{e}^{a_1(t_1 - t)} & \frac{1}{\sigma_2}\mathrm{e}^{a_2(t_1 - t)} \end{pmatrix},$$

respectively. Also, the integration by parts formula together with (8) and (9) yields

$$S^g \circ S^h(\varphi(X_1 - x)\psi(X_2 - y)) = \varphi(X_1 - x)\psi(X_2 - y) \times SH \times SG,$$

where

$$SH = S_{1,2}^h(1) + \eta_1 S_2^h(1) + \eta_2 S_1^h(1) + \eta_1\eta_2$$

and

$$SG = S_{1,2}^g(1) + \rho_1 S_2^g(1) + \rho_2 S_1^g(1) + \rho_1\rho_2.$$

Using g_t and h_t above, the integrals $S(1)$ defined in (10) can be obtained as follows:

$$
\begin{aligned}
S_1^h(1) &= \frac{\mathrm{e}^{a_1 t_1}}{\sigma_1 t_1}\int_0^{t_1}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^1 + \frac{\mathrm{e}^{a_1 t_1}}{\sigma_2(a_2-a_1)t_1}\int_0^{t_1}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^2 \\
&\quad - \frac{\mathrm{e}^{a_1 t_1}}{\sigma_1(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^1 - \frac{\mathrm{e}^{a_1 t_1}}{\sigma_2(a_2-a_1)(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^2, \\
S_1^g(1) &= \frac{\mathrm{e}^{a_1 t_2}}{\sigma_1(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^1 + \frac{\mathrm{e}^{a_1 t_2}}{\sigma_2(a_2-a_1)(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^2 \\
&\quad - \frac{\mathrm{e}^{a_1 t_2}}{\sigma_1(t_3-t_2)}\int_{t_2}^{t_3}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^1 - \frac{\mathrm{e}^{a_1 t_2}}{\sigma_2(a_2-a_1)(t_3-t_2)}\int_{t_2}^{t_3}\mathrm{e}^{-a_1 t}\mathrm{d}W_t^2, \\
S_2^h(1) &= \frac{\mathrm{e}^{a_2 t_1}}{\sigma_2 t_1}\int_0^{t_1}\mathrm{e}^{-a_2 t}\mathrm{d}W_t^2 - \frac{\mathrm{e}^{a_2 t_1}}{\sigma_2(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_2 t}\mathrm{d}W_t^2, \\
S_2^g(1) &= \frac{\mathrm{e}^{a_2 t_2}}{\sigma_2(t_2-t_1)}\int_{t_1}^{t_2}\mathrm{e}^{-a_2 t}\mathrm{d}W_t^2 - \frac{\mathrm{e}^{a_2 t_2}}{\sigma_2(t_3-t_2)}\int_{t_2}^{t_3}\mathrm{e}^{-a_2 t}\mathrm{d}W_t^2, \\
S_{1,2}^h(1) &= S_1^h(1)S_2^h(1) - \frac{\mathrm{e}^{(a_1+a_2)t_1}-1}{\sigma_2^2(a_2^2-a_1^2)t_1^2} - \frac{1-\mathrm{e}^{-(a_1+a_2)(t_2-t_1)}}{\sigma_2^2(a_2^2-a_1^2)(t_2-t_1)^2},
\end{aligned}
$$

and

$$
S_{1,2}^g(1) = S_1^g(1)S_2^g(1) - \frac{\mathrm{e}^{(a_1+a_2)(t_2-t_1)}-1}{\sigma_2^2(a_2^2-a_1^2)(t_2-t_1)^2} - \frac{1-\mathrm{e}^{-(a_1+a_2)(t_3-t_2)}}{\sigma_2^2(a_2^2-a_1^2)(t_3-t_2)^2}.
$$

4 Monte Carlo Simulation

This section applies the results obtained in the previous section to the pricing of anytime Bermudan options by Monte Carlo methods.

Recall that each time interval $[T_{i-1}, T_i]$ is discretized as $T_{i-1} = t_0^i < t_1^i < \cdots < t_{M-1}^i < t_M^i = T_i$. The value of the anytime Bermudan option at time t_j^i in the discrete-time setting is denoted by $DV_j^i(X_{i-1}, X_j^i)$, where $X_j^i = X_{t_j^i}$ and $X_i = X_{T_i}$. In the dynamic programming problem (3)–(5), the conditional expectation (3) should be transformed to an unconditional expectation using (12).

4.1 The Algorithm

Our algorithm to evaluate the anytime Bermudan options by Monte Carlo simulation consists of the following 7 steps.

Step 1. Generate N_{sim} sample paths for the process X in the time interval $[0, T]$, where $T = T_N$. These sample paths are stored in the memory and used throughout the simulation repeatedly. Let $i = N$.

Step 2. At time epoch T_{i-1}, select (and fix) mesh points $x^k \in R^d$, $k = 1, 2, \ldots, N_p$. Let $k = 1$.

Step 3. Let $X_{i-1} = x^k$, and select (and fix) mesh points $x_j^k \in R^d$, $k = 1, 2, \ldots, N_p$, at each time epoch t_j. Let $j = M - 1$. Use the sample paths to solve the dynamic programming (3)–(5) by Monte Carlo simulation for time epoch t_j. If the sample paths do not pass on the mesh points, apply any interpolation to calculate the value $DV_j^i(x^1, x_j^k)$, $k = 1, 2, \ldots, N_p$. Repeat this procedure until time epoch t_2.

Step 4. Starting from $X_{i-1} = x^k$, generate *new* sample points x_2^m, $m = 1, 2, \ldots, N_s$, which correspond to X_2^i and are independent of the original sample paths. The value of the option at time T_{i-1} is estimated by

$$DV_1^i(x^k, x^k) = \frac{1}{N_s} \sum_{m=1}^{N_s} DV_2^i(x^k, x_2^m).$$

If the sample points x_2^m are not on the mesh points, use any interpolation.

Step 5. Repeat Steps 3 and 4 for $k = 2, \ldots, N_p$ to determine the values $DV_1^i(x^k, x^k)$ for all the mesh points x^k.

Step 6. Repeat Steps 2–5 for $i = N - 2, \ldots, 1$ backwards to determine the values $DV_1^1(x^k, x^k)$ for all the mesh points x^k.

Step 7. Repeat Steps 3 and 4 for $X_0 = x^1$ to determine the value $V_0(X_0) = DV_1^0(x^1, x^1)$, the current price of the anytime Bermudan option.

The number of sample points we need for the simulation is $N \times M \times N_{sim}$ for the first step and $N \times N_p \times N_s \times$ for Step 4, whence in total $N(MN_{sim} + N_pN_s)$. Note that, when calculating the continuation value $CV_1^i(x^k, x^k)$ for the first subperiod, it is easier to use the conditional expectation than to evaluate an unconditional expectation. This is the reason why we generate new sample points in Step 4 to evaluate the option value $DV_1^i(x^k, x^k)$.

4.2 Numerical Examples

In this subsection, we consider a floating-rate bond with constant maturity $T = 2$ and face value 100 yen. The coupons are paid semiannually and the coupon rate for the next coupon is set at the current coupon date. The bond is assumed to be callable by the issuer at the face value after a half of the year from the date of issue. Hence, the callable, floating-rate bond can be considered to be an anytime Bermudan option written on the interest rates.

In our numerical experiment, we use the interest-rate models considered in the previous two examples.

Example 1, Continued

The base parameters used in this example are listed in Table 1. Recall that N_{sim} denotes the number of sample paths generated for Monte Carlo simulation, N_p the number of mesh points in each time epoch,[6] N_s the number of sample points generated for Step 4 in the algorithm, M the number of time steps in each period, and N the number of periods.[7] Hence, the number of sample points we need for the base case is $260,000$.

Table 1: Parameter Values for the Base Case

r_0	a	θ	σ	N_{sim}	N_p	N_s	M	N
0.03	0.1	0.01	0.01	10000	50	100	6	4

Note that, in this one-dimensional case, the price of the corresponding European-type bond can be evaluated by numerical integration.[8] We use this price (100.90 yen) as the theoretical value.

Table 2 examines the effect of the localizing functions. Recall that the parameters η, ρ of the localizing functions are obtained in (13) and (14), respectively. The price of the corresponding European-type bond is calculated by Monte Carlo simulation using $m\eta$ and $m\rho$, $m = 0$, 0.5, 1.0, 2.0, as the parameters of the localizing functions. Note that the case $m = 0$ means that the localizing functions are not used in the simulation. The column SDV denotes the standard deviation of the prices calculated by 50 simulation runs. It is explicitly observed that the simulation result become stable when the localizing functions are used appropriately.[9]

Table 2: Effect of Localizing Functions for European-Type Bonds

m	0	0.5	1.0	2.0	Theoretical
Price	83.00	101.82	100.92	100.91	100.90
SDV	26.39	0.12	0.01	0.01	

Table 3 examines the impact of M on the price of the anytime Bermudan-type bonds. The case $M = 1$ corresponds to the ordinary Bermudan-type bonds, i.e. no exercise is allowed between the coupon-payment dates. Since each interval is divided by the equal

[6] In our numerical examples, mesh points are selected to have an equal space.

[7] This means that one year is divided into $2M$ periods in this setting.

[8] See, e.g., Theorem 15.1 in Musiela and Rutkowski (1998).

[9] Since localizing functions are determined so that the variance of the denominator in (12) is minimized, they are considered to be one of the variance reduction methods in Monte Carlo simulation.

length, $M = 2$ means that one can exercise the right in every 3 months, and so on. The column 'Option Value' represents the difference between the prices of the anytime Bermudan-type bond and the corresponding European-type bond. The column 'Time' represents the computational time (measured in seconds) performed by Pentium 4@2GHz. The option value is increasing in M, as expected; but its speed is moderate. The computational time is almost linear in M.

Table 3: Prices of the Anytime Bermudan-Type Bonds

M	1	2	6	12	24
Price	100.24	99.39	98.77	98.61	98.50
Option Value	0.66	1.51	2.13	2.29	2.40
SDV	0.01	0.01	0.02	0.02	0.02
Time		100	490	970	2030

Example 2, Continued

Consider the two-dimensional process (x_t^1, x_t^2) in Example 2, and suppose that the interest rate r_t follows the SDE

$$\mathrm{d}r_t = (x_t^1 - a_2 r_t)\mathrm{d}t + \sigma_2 \mathrm{d}W_t^2,$$

where a_2 is a positive constant. Note that the quantity x_t^1 plays the role of the (stochastic) level for mean reversion in r_t.

The base parameters used in this example are listed in Table 4. Note that a_2 and σ_2 are chosen to be the same as the one-dimensional case (see Table 1). The other parameters for the interest rates are chosen so that the yield curve in the two-dimensional case is the same as that in Example 1. The number of sample points we need for the base case is $520,000$.

Table 4: Parameter Values for the Base Case (2-dim. case)

x_0^1	x_0^2	a_1	θ	σ_1	a_2	σ_2	N_{sim}	N_p	N_s	M	N
0.0087	-0.0766	0.0164	0.0100	0.0089	0.1	0.01	10000	50	100	6	4

Table 5 examines the impact of M on the price of the anytime Bermudan-type bonds in the two-dimensional case. In this example, however, the price of the corresponding European-type bond is not known and so, we calculate it by Monte Carlo simulation. As in the one-dimensional case, the option value is increasing in M, but its speed is moderate. The computational time is again almost linear in M.

Table 5: Prices of the Anytime Bermudan-Type Bonds (2-dim. case)

M	European	1	2	6	12	24
Price	102.12	100.41	99.56	98.96	98.80	98.72
Option Value		1.71	2.56	3.16	3.32	3.40
SDV	0.01	0.01	0.01	0.02	0.03	0.03
Time			235	630	1250	2830

5 Concluding Remarks

In this paper, we propose a Monte Carlo method to evaluate anytime Bermudan options, a variant of path-dependent American options. Assuming that the underlying state variable is Markovian, the price of an anytime Bermudan option satisfies a dynamic programming equation. The continuation value in the dynamic programming problem can be represented by an uncoditional expectation, using the Malliavin calculus, which in turn enables us to evaluate the price of the anytime Bermudan option by Monte Carlo methods. Some numerical examples are given to demonstrate the usefulness of our method.

In general, however, the matirx-valued processes g_t, h_t are not expressed in closed form, unless the SDE (17) for the first variation process Y has a simple structure. Accordingly, it is difficult to calculate the localizing functions. The importance of the localizing functions in Monte Carlo methods is confirmed in our numerical experience. Hence, the key to success in our method is to calculate numerically the processes g_t, h_t in an efficient way for the general case.

A Appendix: Proofs

In this appendix, we use the following notation. Let N be the set of integers and let N^k denote its k-fold product space. Define

$$\mathcal{J}_k \equiv \left\{ I = (i_1, \ldots, i_k) \in N^k : 1 \le i_1 < \cdots < i_k \le d \right\}.$$

For $I \in \mathcal{J}_k$ and $J \in \mathcal{J}_m$, let $I \vee J \equiv \{i_1, \ldots, i_k\} \cup \{j_1, \ldots, j_m\}$. Also, we denote $\mathcal{J}^{-i} \equiv \{I \in \mathcal{J} : i \notin I\}$ and $x^{-i} \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$. Finally, $|I|$ denotes the cardinality of the set I.

For function $f(x)$, $x \in R^d$, we denote the derivative by

$$\partial_I f(x) = \begin{cases} \dfrac{\partial^{|I|} f(x)}{\partial x_{i_1} \cdots \partial x_{i_{|I|}}}, & I \neq \emptyset, \\ f(x), & I = \emptyset. \end{cases}$$

Also, for $x \in R^d$, we define

$$\pi_i(x) = \begin{cases} (0, \ldots, 0, x_{i+1}, \ldots, x_d), & i = 1, \ldots, d-1, \\ (0, \ldots, 0, 0, \ldots, 0), & i = d. \end{cases}$$

For a function $f(x)$, $x \in R^d$, the function $f_i(x)$ is defined by $f_i(x) \equiv f \circ \pi_i(x)$, $i = 1, \ldots, d$. Finally, we denote $I_i \equiv (i+1, \ldots, d) \in \mathcal{J}_{d-i}$ for $i = 1, \ldots, d-1$ and $I_d \equiv \emptyset$. Recall that we denote $X_{t_i} = X_i$, $i = 1, 2, 3$, for notational simplicity.

Lemma 1 *Suppose that $u : R^d \times R^d \mapsto R$ and $\phi : R^{d-1} \mapsto R$ are bounded. Also, suppose that h_t satisfies (9) and $\varphi(x) = \mathrm{e}^{-\eta \cdot x}$. Then, for any random variable F defined on R with a Malliavin derivative, we have*

$$\begin{aligned} & \mathrm{E}\left[1_{A_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_i(X_1 - x)F)\right] \\ &= \int_{A_i} \mathrm{E}\left[H_{x_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_{i-1}}^h(\varphi_{i-1}(X_1 - x)F)\right] \mathrm{d}x_i \end{aligned} \tag{21}$$

for any $A_i \subset R$.

Proof. We define the random variable G_i as

$$G_i \equiv \int_{-\infty}^{X_1^i} 1_{A_i}(x_i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_{i-1}(X_1 - x)F)\mathrm{d}x_i.$$

The Malliavin derivative of G_i is given by

$$\begin{aligned} D_t G_i &= 1_{A_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_i(X_1 - x)F)D_t X_1^i \\ &\quad + \int_{-\infty}^{X_1^i} 1_{A_i}(x_i)D_t\left\{\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_{i-1}(X_1 - x)F)\right\}\mathrm{d}x_i. \end{aligned}$$

Note that, from (9), we have

$$\int_0^{t_3} D_t X_1^i h_t^i \mathrm{d}t = 1, \qquad \int_0^{t_3} D_t\left\{\phi(X_1^{-i})\right\}h_t^i \mathrm{d}t = \int_0^{t_3} D_t\left\{u(X_3, X_2)\right\}h_t^i \mathrm{d}t = 0.$$

It follows from the chain rule in the Malliavin calculus that

$$\begin{aligned} \int_0^{t_3} D_t G_i h_t^i \mathrm{d}t &= 1_{A_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_i(X_1 - x)F) \\ &\quad + \int_{-\infty}^{X_1^i} 1_{A_i}(x_i)\phi(X_1^{-i})u(X_3, X_2)\mathrm{d}x_i \int_0^{t_3} D_t S_{I_i}^h(\varphi_{i-1}(X_1 - x)F)h_t^i \mathrm{d}t. \end{aligned} \tag{22}$$

On the other hand, from the integration by parts formula, we obtain

$$\int_0^{t_3} G_i h_t^i \mathrm{d}W_t = G_i \int_0^{t_3} h_t^i \mathrm{d}W_t - \int_0^{t_3} D_t G_i h_t^i \mathrm{d}t.$$

Taking the expectation for both sides and substituting the definition of G_i yields

$$\begin{aligned} & \mathrm{E}\left[\int_0^{t_3} D_t G_i h_t^i \mathrm{d}t\right] \\ = \; & \mathrm{E}\left[G_i \int_0^{t_3} h_t^i \mathrm{d}W_t\right] \\ = \; & \int_{A_i} \mathrm{E}\left[H_{x_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_{i-1}(X_1 - x)F) \int_0^{t_3} h_t^i \mathrm{d}W_t\right] \mathrm{d}x_i. \end{aligned} \tag{23}$$

Hence, from (22) and (23), we obtain

$$\begin{aligned} & \mathrm{E}\left[1_{A_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_i}^h(\varphi_i(X_1 - x)F)\right] \\ = \; & \int_{A_i} \mathrm{E}\left[H_{x_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)\left(S_{I_i}^h(\varphi_{i-1}(X_1 - x)F) \int_0^{t_3} h_t^i \mathrm{d}W_t \right.\right. \\ & \left.\left. \qquad - \int_0^{t_3} D_t S_{I_i}^h(\varphi_{i-1}(X_1 - x)F) h_t^i \mathrm{d}t\right)\right] \mathrm{d}x_i \\ = \; & \int_{A_i} \mathrm{E}\left[H_{x_i}(X_1^i)\phi(X_1^{-i})u(X_3, X_2)S_{I_{i-1}}^h(\varphi_{i-1}(X_1 - x)F)\right] \mathrm{d}x_i. \end{aligned}$$

Here, we used the integration by parts formula for the second equality and (9) for the third equality. This proves (21).

The next result can be proved similarly.

Lemma 2 *Suppose that $u : R^d \times R^d \mapsto R$ and $\phi : R^{d-1} \mapsto R$ are bounded. Also, suppose that g_t satisfies (8) and $\psi(x) = e^{-\rho \cdot x}$. Then, for any random variable F defined on R with a Malliavin derivative, we have*

$$\begin{aligned} & \mathrm{E}\left[1_{B_i}(X_2^i)\phi(X_2^{-i})u(X_3, X_1)S_{I_i}^g(\psi_i(X_2 - y)F)\right] \\ & \quad = \int_{B_i} \mathrm{E}\left[H_{y_i}(X_2^i)\phi(X_2^{-i})u(X_3, X_1)S_{I_{i-1}}^g(\psi_{i-1}(X_2 - y)F)\right] \mathrm{d}y_i \end{aligned} \tag{24}$$

for any $B_i \subset R$.

Proof of Proposition 1.

Since $S_\emptyset^h(x) = x$, $\phi_d(x) = 1$ and $I_d = \emptyset$, we obtain

$$\mathrm{E}\left[f(X_3, z)1_A(X_1)1_B(X_2)\right] = \mathrm{E}\left[f(X_3, z)1_A(X_1)1_B(X_2)S_{I_d}^h(\phi_d(X_1 - x))\right].$$

Let $\phi(X_1^{-d}) = \prod_{i=1}^{d-1} 1_{A_i}(X_1^i)$. Then, $1_A = 1_{A_d}\phi(X_1^{-d})$. Hence, from (21), we obtain

$$\begin{aligned} & \mathrm{E}\left[f(X_3,z)1_A(X_1)1_B(X_2)S^h_{I_d}(\phi_d(X_1-x))\right] \\ = & \int_{A_d} \mathrm{E}\left[f(X_3,z)H_{x_d}(X_1^d)\phi(X_1^{-d})1_B(X_2)S^h_{I_{d-1}}(\varphi_{d-1}(X_1-x))\right]\mathrm{d}x_d. \end{aligned}$$

Similalrly, let $\phi(X_1^{-(d-1)}) = H_{x_d}(X_1^d)\prod_{i=1}^{d-2} 1_{A_i}(X_1^i)$. Then,

$$H_{x_d}(X_1^d)\phi(X_1^{-d}) = 1_{A_{d-1}}\phi(X_1^{-(d-1)}).$$

It follows from (21) that

$$\begin{aligned} & \int_{A_d} \mathrm{E}\left[f(X_3,z)H_{x_d}(X_1^d)\phi(X_1^{-d})1_B(X_2)S^h_{I_{d-1}}(\varphi_{d-1}(X_1-x))\right]\mathrm{d}x_d \\ = & \int_{A_d}\mathrm{d}x_d\int_{A_{d-1}}\mathrm{d}x_{d-1}\mathrm{E}\left[f(X_3,z)H_{x_{d-1}}(X_1^{d-1})\phi(X_1^{-(d-1)})1_B(X_2)S^h_{I_{d-2}}(\varphi_{d-2}(X_1-x))\right]. \end{aligned}$$

Repeating this argument, one finally obtains

$$\begin{aligned} & \mathrm{E}\left[f(X_3,z)1_A(X_1)1_B(X_2)\right] \\ = & \int_{A_d}\mathrm{d}x_d\cdots\int_{A_1}\mathrm{d}x_1\mathrm{E}\left[f(X_3,z)H_x(X_1)1_B(X_2)S^h(\varphi(X_1-x))\right]. \end{aligned} \tag{25}$$

Now, let $F = S^h(\varphi(X_1-x))$. From (24), we have

$$\begin{aligned} & \mathrm{E}\left[f(X_3,z)H_x(X_1)1_B(X_2)F\right] \\ = & \mathrm{E}\left[f(X_3,z)H_x(X_1)1_B(X_2)S^g_{I_d}(\psi_d(X_2-y)F)\right] \\ = & \int_{B_d}\mathrm{E}\left[f(X_3,z)H_x(X_1)H_{y_d}(X_2^d)\phi(X_2^{-d})S^g_{I_{d-1}}(\psi_{d-1}(X_2-y)F)\right]\mathrm{d}y_d. \end{aligned}$$

Repeated application of (24) then yields

$$\begin{aligned} & \mathrm{E}\left[f(X_3,z)H_x(X_1)1_B(X_2)F\right] \\ = & \int_{B_d}\mathrm{d}y_d\cdots\int_{B_1}\mathrm{d}y_1\mathrm{E}\left[f(X_3,z)H_x(X_1)H_y(X_2)S^g(\psi(X_2-y)F)\right]. \end{aligned} \tag{26}$$

The proposition is now proved by (25) and (26).

References

[1] V. Bally, L. Caramellino and A. Zanette (2003), "Pricing and hedging American options by Monte Carlo methods using a Malliavin calculus approach," Working Paper, INRIA.

[2] B. Bouchard, I. Ekeland and N. Touzi (2003), "On the Malliavin approach to Monte Carlo approximation of condtional expectations," Working Paper, CREST.

[3] E. Fournie, J. Lasry, J. Lebuchoux and P. Lions (2001), "Applications of Malliavin calculus to Monte-Carlo methods in finance, II," *Finance and Stochastics*, **5**, 201–236.

[4] P. Glasserman (2003), *Monte Carlo Methods in Financial Engineering*, Springer, New York.

[5] P. Jackel (2002), *Monte Carlo Methods in Finance*, John Wiley & Sons, New York.

[6] M. Mrad, N. Touzi and A. Zeghal (2003), "Monte Carlo estimation of a joint density using Malliavin Calculus," Working Paper, CREST.

[7] M. Musiela and M. Rutkowski (1998), *Maritngale Methods in Financial Modelling*, Springer, New York.

[8] D. Nualart (1995), *The Malliavin Calculus and Related Topics*, Springer, New York.

[9] R. Rebonato (1998), *Interest-Rate Option Models*, 2nd Ed., John Wiley & Sons, New York.