MONTE CARLO CONSTRUCTION OF HEDGING STRATEGIES AGAINST MULTI-ASSET EUROPEAN CLAIMS

G.N. MILSTEIN[1,2] AND J.G.M. SCHOENMAKERS[1]

Abstract. For evaluating a hedging strategy we have to know at every instant the solution of the Cauchy problem for a parabolic equation (the value of the hedging portfolio) and its derivatives (the deltas). We suggest to find these magnitudes by Monte Carlo simulation of the corresponding system of stochastic differential equations using weak solution schemes. It turns out that with one and the same control function a variance reduction can be achieved simultaneously for the claim value as well as for the deltas. We consider a Markovian multi-asset model with an instantaneously riskless saving bond and also some applications to the LIBOR rate model of Brace, Gatarek, Musiela and Jamshidian.

1) Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, D-10117 Berlin, Germany;
2) Ural State University, Lenin Street 51, 620083 Ekaterinburg, Russia

JEL classification: E43, G13
AMS Subject Classification: 60H30, 65U05, 90A09 Key words: Derivative pricing and hedging, probabilistic representations, weak approximation of solutions of stochastic differential equations, Monte Carlo simulation, variance reduction.

1. Introduction

Our goal in this paper is the construction of a probabilistic method for the valuation *and* the replication (hedge) of a European claim maturing at some future time T, contingent on the values $X^i(T)$, of a basket of m stocks (risky assets) with price processes $X^i(t)$, $i = 1, ..., m$. We assume that in the market also a cash bond (riskless asset) B is available and that the system (B, X) satisfies the stochastic differential equations

$$dB = r(t)Bdt, \ B(t_0) = 1, \tag{1.1}$$

$$dX^i = X^i(\mu^i(t,X)dt + \sum_{j=1}^m \nu_{ij}(t,X)dW^j(t)), \ t \geq t_0, \ i = 1, ..., m,$$

where, for the time being, $r(t)$ is a deterministic interest rate, $X = (X^1, ..., X^m)^\top$, $W = (W^1, ..., W^m)^\top$ is an m-dimensional standard Wiener process on a probability space $(\Omega, \mathcal{F}, P)$. As usual the P-augmentation of the filtration generated by W is denoted by $\{\mathcal{F}_t\}$. Further it is assumed that $r(t)$, the vector $(\mu^1(t,x), ..., \mu^m(t,x))^\top$, and the matrix $\nu(t,x) = \{\nu_{ij}(t,x)\}$, $t \in [t_0, T]$, $x \in R^m_+ := \{x : x^1 > 0, ..., x^m > 0\}$, are sufficiently smooth and such that there exists a unique process $X(t) \in R^m_+$, $t \in [t_0, T]$, with $X(t_0) \in R^m_+$ satisfying (1.1) (for example, all the μ^i, ν_{ij} are smooth and bounded). Moreover, we assume that the volatility matrix $\sigma(t,x) = \{\sigma_{ij}(t,x)\} = \{x^i\nu_{ij}(t,x)\}$ has full rank for every (t,x), $t \in [t_0, T]$, $x \in R^m_+$. Under these assumptions the model (B, X) constitutes a complete market and is, in fact, a special version of the system described in Karatzas and Shreve [7]. Henceforth, we will assume that the originating functions are always sufficiently good in analytical sense without stating their analytical properties explicitly.

The central problem now is to determine the price and hedge of a European claim $f(X(T))$ at maturity time T specified by a payoff function f which depends on X, by constructing a self-financing portfolio or trading strategy. In the construction of this portfolio it is assumed that the stocks pay dividends to the share holders at a rate $r^i(t, X(t))$ for the i-th stock and further a consumption process C is assumed which is defined by a consumption rate $c(t, X(t))$, $t_0 \leq t \leq T$,

$$dC = c(t, X(t))dt, \ C(t_0) = 0. \tag{1.2}$$

$>$From a mathematical point of view the incorporation of both continuous dividends and consumption goes without any difficulties, however, the reader should feel free to take them zero if he prefers.

The portfolio value $V(t)$ of a trading strategy $(\varphi_t, \psi_t) = (\varphi_t, \psi^1_t, ..., \psi^m_t)$, i.e. the positions in bond $B(t)$ and stocks $X^j(t)$ respectively, is given by

$$V(t) = \varphi_t B(t) + \sum_{i=1}^m \psi^i_t X^i(t) \tag{1.3}$$

and the trading strategy is said to be (generalized) self-financing if

$$dV = \varphi_t dB + \sum_{i=1}^m \psi^i_t dX^i + \sum_{i=1}^m r^i(t, X(t))\psi^i_t X^i(t)dt - c(t, X(t))dt. \tag{1.4}$$

It is known (see, e.g., [14], [7]) that in our framework the European claim may be hedged with a uniquely determined self-financing portfolio or trading strategy, which value at time $t < T$ is given by

$$V(t) = v(t, X(t)).$$

Here, v is a function of the variables $t, x^1, ..., x^m$ and satisfies the following Cauchy problem for the parabolic partial differential equation:

$$Lv(t,x) + c(t,x) := \frac{\partial v}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(t,x)\frac{\partial^2 v}{\partial x^i \partial x^j} \tag{1.5}$$

$$+\sum_{i=1}^{m} b^i(t,x)\frac{\partial v}{\partial x^i} - r(t)v + c(t,x) = 0,$$

$$v(T,x) = f(x), \tag{1.6}$$

where $a_{ij} = \sum_{k=1}^{m} \sigma_{ik}\sigma_{jk} = x^i x^j \sum_{k=1}^{m} \nu_{ik}\nu_{jk}$, $b^i = (r - r^i)x^i$, $i = 1, ..., m$.

If $v(t,x)$ is the solution of the problem (1.5)-(1.6), the required hedging strategy $(\varphi_t, \psi_t^1, ..., \psi_t^m)$ as a function of $(t, X(t))$ is then given by

$$\varphi_t = \frac{1}{B(t)}(v(t,X(t)) - \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(t,X(t))X^i(t)), \ \psi_t^i = \frac{\partial v}{\partial x^i}(t,X(t)), \ i = 1, ..., m. \tag{1.7}$$

Frequently, works in numerics for finance (see, e.g., [18] and references therein) are devoted to the evaluation of a portfolio value $v(t,x)$. However, for constructing the hedging strategy we also need the derivatives of v, called *deltas*. Of course, in the case where $v(t,x)$ is known, it is possible to find $\partial v(t,x)/\partial x^i$ approximately, for example, as

$$\frac{\partial v(t,x)}{\partial x^i} \simeq \frac{v(t,x^1,...,x^i+\Delta x^i,...,x^m) - v(t,x^1,...,x^i-\Delta x^i,...,x^m)}{2\Delta x^i}, \ i = 1, ..., m,$$

but such an approach requires very accurate calculations for v and cannot be realized in practice especially in many-dimensional cases. Moreover, in many-dimensional cases (in reality for $m \geq 3$) it is usually impossible to find $v(t,x)$ for all (t,x) because of the complexity of problem (1.5)-(1.6). Therefore, in this sequel we give special attention to the evaluation of the deltas

$$u_i(t,x) := \frac{\partial v}{\partial x^i}(t,x), \ i = 1, ..., m,$$

by probabilistic methods. Indeed, as for the construction of a hedging strategy one only needs at any instant t the individual values $v(t, X(t))$ and $\partial v(t, X(t))/\partial x^i$, $i = 1, ..., m$, where $X(t)$ is the (known) state of the market, the Monte Carlo method is most relevant for such a problem.

The probabilistic approach for the evaluation of a particular value $v(t,x)$ of Cauchy problem (1.5)-(1.6) is well known and comes down to Monte Carlo simulation of probabilistic representations for $v(t,x)$ by stochastic differential equations naturally connected to (1.5)-(1.6).

It is not difficult to obtain probabilistic representations for the derivatives $\partial v(t,x)/\partial x^i$ also (see Section 2) and so it is possible to find the values $v(t,x)$ and $\partial v(t,x)/\partial x^i$, $i = 1, ..., m$, for specific (t,x) by Monte Carlo simulation of suitable representations. The present paper is particularly devoted to these Monte Carlo techniques and the outline of the paper is as follows.

Section 2 contains rather known material and is included for the convenience of the reader. First we give a brief derivation of (1.5)-(1.6). Then we present various probabilistic representations for the solution v of Cauchy problem (1.5)-(1.6) from which we derive straightforwardly probabilistic representations for the partial derivatives $\partial v/\partial x^i$ in terms of a variational system of stochastic differential equations. The representation for v is standard (see, e.g., [3], [4]) and, probably, the representations for $\partial v/\partial x^i$ are also known to specialists in stochastic analysis. However, we could not find them in the literature and for this reason the formulas for $\partial v/\partial x^i$ are given without any references. Section 2 is concluded by a brief discussion of weak methods for numerical integration of SDEs (see [8], [10]) which are needed for Monte Carlo simulation of stochastic differential equations.

Our main results are contained in Sections 3 and 4. As well known fact, variance reduction is of crucial importance for the implementation of Monte Carlo procedures. Therefore, all the various probabilistic representations given in this paper are endowed with a (vector) function to control the variance. For the moment, let us denote this function by h. The function h is such that the mathematical expectations for v and for $\partial v/\partial x^i$ due to the chosen representations do not depend on the choice of h. In the meantime, the corresponding variance *does* depend on h. Therefore it is natural to regard h as a control function which may be chosen such that the variance is minimal. We use this approach for two known methods for variance reduction: In Section 3 for the method of important sampling and in Section 4 for the method of control variates. Moreover, in Section 4 we introduce a combined method which contains both the method of importance sampling and the method of control variates as particular cases. It turns out that with one and the same control function a variance reduction can be achieved simultaneously for the claim value v as well as for the deltas $u_i = \partial v/\partial x^i$, $i = 1, ..., m$. These results are extended for gammas, vegas, and thetas in Section 5 and for barrier options in Section 6. In Section 7, we show the application of the presented methods to LIBOR derivatives in a LIBOR (market) model [6], [1]. Finally, in Section 8 we give some concluding remarks.

2. Monte Carlo evaluation of a hedging strategy

We give a brief derivation of the problem (1.5)-(1.6) and formulas (1.7). Since the portfolio (φ_t, ψ_t) is (generalized) self-financing and thus satisfies (1.4), it follows that

$$\begin{aligned} dV &= \varphi_t r(t) B dt + \sum_{i=1}^m \psi_t^i X^i (\mu^i(t,X)dt + \sum_{j=1}^m \nu_{ij}(t,X) dW^j(t)) + \\ &\quad + \sum_{i=1}^m r^i(t,X(t)) \psi_t^i X^i(t) dt - c(t,X(t))dt, \end{aligned} \tag{2.1}$$

which is equivalent with

$$B d\varphi_t + \sum_{i=1}^m X^i d\psi_t^i + \sum_{i=1}^m d\psi_t^i dX^i = \sum_{i=1}^m r^i(t,X(t)) \psi_t^i X^i(t) dt - c(t,X(t))dt. \tag{2.2}$$

Let us now consider a European claim $f(X_T)$ at maturity time T, specified by a payoff function f which depends on $X(T)$. Since the market is complete the claim may be replicated (hedged) by a self-financing portfolio with value process $V(t)$, say. As, moreover, the model (B, X) is Markovian we have

$$V(t) = \varphi_t B(t) + \sum_{i=1}^m \psi_t^i X^i(t) = v(t, X(t)),\ V(T) = v(T, X(T)) = f(X(T)), \tag{2.3}$$

where v is a function of the variables $t, x^1, ..., x^m$.

Just as for the standard Black-Scholes model (one risky asset and one riskless bond) we may derive a parabolic differential equation for the function $v(t,x)$ (see, e.g., [7], [14]).

Due to Itô's formula we have

$$dv(t,X(t)) = \frac{\partial v}{\partial t}dt + \sum_{i=1}^{m}\frac{\partial v}{\partial x^i}dX^i + \frac{1}{2}\sum_{i,j=1}^{m}\frac{\partial^2 v}{\partial x^i \partial x^j}dX^i dX^j \tag{2.4}$$

$$= \frac{\partial v}{\partial t}dt + \sum_{i=1}^{m}\frac{\partial v}{\partial x^i}X^i\mu^i dt + \sum_{i=1}^{m}\frac{\partial v}{\partial x^i}\sum_{j=1}^{m}\sigma_{ij}dw^j(t) + \frac{1}{2}\sum_{i,j=1}^{m}a_{ij}\frac{\partial^2 v}{\partial x^i \partial x^j}dt.$$

>From 2.3 and comparing (2.1) with (2.4), we obtain

$$\psi_t^i = \psi^i(t,X(t)) = \frac{\partial v}{\partial x^i}(t,X(t)),\ \psi^i(t,x) = \frac{\partial v}{\partial x^i}(t,x), \tag{2.5}$$

and

$$\frac{\partial v}{\partial t}(t,X(t)) + \frac{1}{2}\sum_{i,j=1}^{m}a_{ij}(t,X(t))\frac{\partial^2 v}{\partial x^i \partial x^j}(t,X(t))$$

$$= \varphi_t r(t)B(t) + \sum_{i=1}^{m}r^i(t,X(t))\psi_t^i X^i(t) - c(t,X(t)). \tag{2.6}$$

Then, substituting (see (2.3) and (2.5))

$$\varphi_t B(t) = v(t,X(t)) - \sum_{i=1}^{m}\psi_t^i X^i(t) = v(t,X(t)) - \sum_{i=1}^{m}\frac{\partial v}{\partial x^i}(t,X(t))X^i(t),$$

in (2.6) and taking into account (2.5), we get the Cauchy problem (1.5)-(1.6) for a parabolic partial differential equation. If $v(t,x)$ is the solution to this problem, the required hedging strategy is given by formulas (1.7). The equality (2.2) for this strategy can be checked directly.

Remark 2.1. Consider the model (1.1) with now r depending on t and X, i.e., (1.1) with as first equation

$$dB = r(t,X)Bdt,\ B(t_0) = 1.$$

Then, in general, $V(t)$ depends on $t,\ X(t),\ B(t)$, i.e., $V(t) = v(t,X(t),B(t))$. Arguing as above, we now obtain that v satisfies the following equation

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{m}a_{ij}(t,x)\frac{\partial^2 v}{\partial x^i \partial x^j} + \sum_{i=1}^{m}b^i(t,x)\frac{\partial v}{\partial x^i} + r(t,x)B\frac{\partial v}{\partial B} - r(t,x)v + c(t,x) = 0.$$

But, since the claim at maturity T depends on $X(T)$ only, the solution of the above equation satisfying condition (1.6) is independent of B. So $\partial v/\partial B = 0$ and we obtain Cauchy problem (1.5)-(1.6) where $r = r(t,x)$. The formulas for the required hedging strategy, (1.7), remain the same.

Remark 2.2. We note that a Cauchy problem is considered in spite of the fact that the variable x belongs to $R_+^m = \{x:\ x^1 > 0,...,x^m > 0\}$. This is possible because every solution $X(t),\ X(t_0) \in R_+^m$, of system (1.1) evolves in R_+^m during the whole time interval $[t_0,T]$. If we consider, for example, a stock model,

$$X^i = (X^i - \pi_1^i)(\pi_2^i - X^i)(\mu^i(t,X)dt + \sum_{j=1}^{m}\nu_{ij}(t,X)dW^j(t)),\ t \geq t_0,\ i = 1,...,m, \tag{2.7}$$

for suitable coefficients μ^i and ν_{ij}, with stock prices evolving in an open parallelepiped $\Pi = \{x:\ 0 \leq \pi_1^1 < x^1 < \pi_2^1, ..., 0 \leq \pi_1^m < x^m < \pi_2^m\}$, where $\pi_1^k,\ \pi_2^k,\ k = 1,...,m$, are constants (it is possible to consider cases when some of π_2 are equal to ∞), then the construction of a hedging strategy leads to a corresponding Cauchy problem again (i.e. not to a boundary value problem).

We now recall the probabilistic representation for the solution v of the Cauchy problem (1.5)-(1.6). For generality we take $r(t,x)$ in (1.5) instead of $r(t)$. In fact, the solution to problem (1.5)-(1.6) has various probabilistic representations:

$$v(t,x) = E[f(X_{t,x}(T))\,Y_{t,x,1}(T) + Z_{t,x,1,0}(T)],\ t \le T,\ x \in R^m_+, \tag{2.8}$$

where $X_{t,x}(s),\ Y_{t,x,y}(s),\ Z_{t,x,y,z}(s),\ s \ge t$, is the solution of the following system of stochastic differential equations:

$$dX = (b(s,X) - \sigma(s,X)h(s,X))ds + \sigma(s,X)dW(s),\ X(t) = x, \tag{2.9}$$

$$dY = -r(s,X)Yds + h^\top(s,X)\,YdW(s),\ Y(t) = y, \tag{2.10}$$

$$dZ = c(s,X)Yds,\ Z(t) = z, \tag{2.11}$$

with Y and Z being scalar processes and $h(t,x) = (h^1(t,x),...,h^m(t,x))^\top$, where h^i are rather arbitrary functions, however, with good analytical properties. The usual probabilistic representation (see, e.g., [3], [4]) follows from (2.8)-(2.11) for $h = 0$. The representation for $h \neq 0$ is, in fact, a consequence of Girsanov's theorem. Other representations are given in Section 4.

Let us now proceed to derive probabilistic representations for $\partial v/\partial x^i$. In what follows we assume that all the coefficients in (1.5)-(1.6) and in (2.9)-(2.11) and the solution of (1.5)-(1.6) are sufficiently smooth and satisfy necessary growth conditions for large $|x|$, so that, in particular, we may apply the theory of weak methods for numerical integration of SDEs.

We introduce the notation

$$u_k(t,x) = \frac{\partial v}{\partial x^k}(t,x),\ k = 1,...,m. \tag{2.12}$$

By differentiating (1.5)-(1.6) with respect to x^k, it follows that the functions v and $u_k,\ k = 1,...,m$, satisfy the Cauchy problem for the following system of $m+1$ linear parabolic equations consisting of (1.5)-(1.6) and

$$\frac{\partial u_k}{\partial t} + \frac{1}{2}\sum_{i,j=1}^m a_{ij}(t,x)\frac{\partial^2 u_k}{\partial x^i \partial x^j} + \sum_{i=1}^m b^i(t,x)\frac{\partial u_k}{\partial x^i} - r(t,x)u_k \tag{2.13}$$

$$+\frac{1}{2}\sum_{i,j=1}^m \frac{\partial a_{ij}}{\partial x^k}(t,x)\frac{\partial u_j}{\partial x^i} + \sum_{i=1}^m \frac{\partial b_i}{\partial x^k}(t,x)\frac{\partial v}{\partial x^i} - \frac{\partial r}{\partial x^k}(t,x)v + \frac{\partial c}{\partial x^k}(t,x) = 0,$$

$$u_k(T,x) = \frac{\partial f}{\partial x^k}(x),\ k = 1,...,m. \tag{2.14}$$

The Cauchy problem (1.5)-(1.6), (2.13)-(2.14) belongs to a class of problems, which solutions have probabilistic representations given in [11] . However, we obtain a representation from (2.8)-(2.11) directly by differentiating (2.8) with respect to x_k. We thus get

$$u_k(t,x) = \frac{\partial v}{\partial x^k}(t,x)$$

$$(2.15) \qquad = E\left[\sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(X_{t,x}(T))\, \delta_k X^i(T)\, Y_{t,x,1}(T) + f(X_{t,x}(T))\, \delta_k Y(T) + \delta_k Z(T)\right],$$

where

$$(2.16) \qquad \delta_k X^i(s) := \delta_k X^i_{t,x}(s) := \frac{\partial X^i_{t,x}(s)}{\partial x^k},\ \delta_k Y(s) := \delta_k Y_{t,x,1}(s) := \frac{\partial Y_{t,x,1}(s)}{\partial x^k},$$

$$\delta_k Z(s) := \delta_k Z_{t,x,1,0}(s) := \frac{\partial Z_{t,x,1,0}(s)}{\partial x^k},\ t \le s \le T.$$

Let $\delta_k X = (\delta_k X^1, ..., \delta_k X^m)^\top$, for fixed k. Then, the vector functions $\delta_k X(s)$ and the scalars $\delta_k Y(s)$ and $\delta_k Z(s)$ satisfy the following system of first order variation associated with (2.9)-(2.11),

$$(2.17) \qquad d\delta_k X = \sum_{l=1}^{m} \frac{\partial (b(s,X) - \sigma(s,X)h(s,X))}{\partial x^l}\, \delta_k X^l ds$$

$$+ \sum_{l=1}^{m} \frac{\partial \sigma(s,X)}{\partial x^l} \cdot \delta_k X^l dW(s),\ \delta_k X^l(t) = 0,\ \text{if } l \ne k,\ \text{and } \delta_k X^k(t) = 1,$$

$$(2.18) \qquad d\delta_k Y = -\sum_{l=1}^{m} \frac{\partial r(s,X)}{\partial x^l}\, \delta_k X^l\, Y ds - r(s,X)\, \delta_k Y ds$$

$$+ \sum_{l=1}^{m} \frac{\partial h^\top(s,X)}{\partial x^l} \cdot \delta_k X^l Y dW(s) + h^\top(s,X) \cdot \delta_k Y dW(s),\ \delta_k Y(t) = 0,$$

$$(2.19) \qquad d\delta_k Z = \sum_{l=1}^{m} \frac{\partial c(s,X)}{\partial x^l}\, \delta_k X^l Y ds + c(s,X)\, \delta_k Y ds,\ \delta_k Z(t) = 0.$$

We underline here that there is an opportunity for parallelizing: One can consider m problems (2.15), (2.9)-(2.11), (2.17)-(2.19) for every fixed $k = 1, ..., m$ separately.

Thus, to find $v(t,x)$ and $\partial v/\partial x^k(t,x)$ we need to evaluate the expectations (2.8) and (2.15). For instance, let us consider (2.8). Usually it is impossible to simulate the random variables $X_{t,x}(T)$, $Y_{t,x,1}(T)$, $Z_{t,x,1,0}(T)$ directly and we are forced to simulate some approximate random variables $\overline{X}_{t,x}(T)$, $\overline{Y}_{t,x,1}(T)$, $\overline{Z}_{t,x,1,0}(T)$, which may be obtained by using weak methods for numerical integration of SDEs (see [8], [10]). The error of such weak approximation is of order $O(h^p)$, where p is the order of weak convergence, depending on the specific method, and h is a time discretization step.

For simplicity we here consider equidistant partitions of the time interval $[t,T] : t = t_0 < t_1 < ... < t_L = T$ with step size $h = (T-t)/L$. Then, for example, the Euler method with simplified simulation of Wiener processes applied to system (2.9)-(2.11) gives

$$(2.20) \qquad \overline{X}(t) = x,\ \overline{X}(t_{l+1}) = \overline{X}(t_l) + (b_l - \sigma_l h_l)h + \sigma_l \cdot \zeta_l \sqrt{h},$$

$$\overline{Y}(t) = 1, \overline{Y}(t_{l+1}) = \overline{Y}(t_l) - r_l \overline{Y}(t_l)h + h_l^\top \cdot \zeta_l \overline{Y}(t_l)\sqrt{h},$$

$$\overline{Z}(t) = 0,\ \overline{Z}(t_{l+1}) = \overline{Z}(t_l) + c_l \overline{Y}(t_l)h,\ l = 0, ..., L-1,$$

where b_l, σ_l, h_l, r_l, and c_l are values of the corresponding functions (scalar, vector or matrix) at $(t_l, \overline{X}(t_l))$ and $\zeta_l = (\zeta_l^1, ..., \zeta_l^m)^\top$ is a vector of two-point random variables ζ_l^j, distributed by the law $P(\zeta_l^j = \pm 1) = 1/2$ and independent in $j = 1, ..., m$; $l = 0, ..., L-1$. We obtain the usual Euler method if ζ_l^j are simulated as $N(0,1)$-distributed random variables. In either case the order of weak convergence is equal to 1, i.e., the following relation

$$|v(t,x) - E[f(\overline{X}(T))\,\overline{Y}(T) + \overline{Z}(T)]| = O(h) \tag{2.21}$$

is fulfilled for a sufficiently large class of functions f. Then, using the Monte Carlo approach, we get

$$E\overline{\xi} \simeq \frac{1}{N}\sum_{n=1}^{N}[f(\overline{X}^{(n)}(T))\,\overline{Y}^{(n)}(T) + \overline{Z}^{(n)}(T)] =: \widehat{v}(t,x), \tag{2.22}$$

where $\overline{\xi} := f(\overline{X}(T))\,\overline{Y}(T) + \overline{Z}(T)$ and $\overline{X}^{(n)}(t_l)$, $\overline{Y}^{(n)}(t_l)$, $\overline{Z}^{(n)}(t_l)$, $n = 1, ..., N$, are independent weak approximate trajectories of the solution of system (2.9)-(2.11).

The statistical error in (2.22) is often defined as $3(D\overline{\xi}(T)/N)^{1/2}$, where $D\overline{\xi}$ is the variance of $\overline{\xi}$, which is close to $D\xi$, $\xi := f(X(T))\,Y(T) + Z(T)$. So the approximation $\widehat{v}(t,x)$ of $v(t,x)$ involves two errors: one error due to the method of numerical integration and a statistical error due to the Monte-Carlo method.

Among the methods with higher order weak convergence let us consider the weak second order Talay-Tubaro extrapolation method [21]. According to the Talay-Tubaro method we have

$$|v(t,x) - 2E[f(\overline{X}_{h/2}(T))\,\overline{Y}_{h/2}(T) + \overline{Z}_{h/2}(T)] + E[f(\overline{X}_h(T))\,\overline{Y}_h(T) + \overline{Z}_h(T)]| = O(h^2),$$

where an approximation (2.20) with step size h is denoted by $\overline{X}_h$, $\overline{Y}_h$, $\overline{Z}_h$. An approximation $\widehat{v}(t,x)$ for $v(t,x)$ is thus obtained by

$$\widehat{v}(t,x) := \frac{2}{N}\sum_{n=1}^{N}[f(\overline{X}_{h/2}^{(n)}(T))\,\overline{Y}_{h/2}^{(n)}(T) + \overline{Z}_{h/2}^{(n)}(T)] - \frac{1}{N}\sum_{n=1}^{N}[f(\overline{X}_h^{(n)}(T))\,\overline{Y}_h^{(n)}(T) + \overline{Z}_h^{(n)}(T)], \tag{2.23}$$

where for $n = 1, ..., N$, $\overline{X}_{h/2}^{(n)}(t_l)$, $\overline{Y}_{h/2}^{(n)}(t_l)$, $\overline{Z}_{h/2}^{(n)}(t_l)$ and $\overline{X}_h^{(n)}(t_l)$, $\overline{Y}_h^{(n)}(t_l)$, $\overline{Z}_h^{(n)}(t_l)$ are independent weak approximate trajectories of the solution of system (2.9)-(2.11), with discretization step $h/2$ and h, respectively.

Now, approximation of $v(t,x)$ with $\widehat{v}(t,x)$ from (2.23) involves an error $\mathcal{O}(h^2) + 3(3D\overline{\xi}(T)/N)^{1/2}$ and so, for reaching the same accuracy, it is possible to take the time step h of numerical integration considerably larger in comparison with the Euler method.

Of course, the same consideration holds with respect to the evaluation of $\partial v/\partial x^k(t,x)$ also.

Concluding, we may say that the error of numerical integration can be reduced by a proper choice of numerical integration scheme and step size h, whereas the statistical error can be reduced (only) by a suitable choice of probabilistic representation for ξ.

We finally remark that if $\overline{X}(t_l)$ in (2.20) is too close to the boundary of R_+^m, it may happen that some of the components $\bar{X}^i(t_{l+1})$, $i = 1, ..., m$, becomes negative. This means that, in fact, one should choose h *not* independently of $\overline{X}$ in this case and, in particular, h should be taken smaller according $\overline{X}$ is closer to the boundary of R_+^m. Such difficulties do not arise if the process evolves in the whole space R^m and we recall that the general theory of numerical integration of SDEs is developed in R^m (for simulation of diffusion processes in bounded domains see [12], [13] and references therein). In our setting, however, we may avoid these difficulties also by using a suitable transformation of R^m_+ into R^m, for instance, by taking $\tilde{x}^i = \ln x^i$, $i = 1, ..., m$.

3. Variance reduction by importance sampling

In this paper we consider variance reduction for the Monte Carlo evaluation of a hedging portfolio as well as for the evaluation of the deltas. In this respect we deal with two methods of variance reduction in connection with the Monte Carlo approach for the linear parabolic Cauchy problem (1.5)-(1.6): the method of importance sampling [5], [10], [15], [16], [22], and the method of control variates [15], [16] (for the initial-boundary value problem see [10], [17]). In this section we explain the method of importance sampling and we show that by this method it is possible to reduce the variance of the estimators corresponding to the probabilistic representations (2.8) and (2.15), simultaneously .

We introduce the variables

$$\xi(s) := v(s, X_{t,x}(s))\, Y_{t,x,1}(s) + Z_{t,x,1,0}(s), \tag{3.1}$$

$$\eta_k(s) := \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(s, X_{t,x}(s))\, \delta_k X^i(s)\, Y_{t,x,1}(s) + v(s, X_{t,x}(s))\, \delta_k Y(s) + \delta_k Z(s). \tag{3.2}$$

Clearly

$$\xi := \xi(T) = f(X_{t,x}(T))\, Y_{t,x,1}(T) + Z_{t,x,1,0}(T), \tag{3.3}$$

$$\eta_k := \eta_k(T) = \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(X_{t,x}(T))\, \delta_k X^i(T)\, Y_{t,x,1}(T) + f(X_{t,x}(T))\, \delta_k Y(T) + \delta_k Z(T). \tag{3.4}$$

Because $D\xi$ $(D\eta_k)$ is close to $D\overline{\xi}$ $(D\overline{\eta}_k)$, the error of a Monte Carlo evaluation of $v(t,x)$ depends on the variance of the random variable ξ, see (2.8), whereas the Monte Carlo error of an evaluation of $u_k(t,x) = \partial v(t,x)/\partial x^k$ depends on the variance of η_k, see (2.15).

The method of evaluating $v(t,x)$ by importance sampling corresponds to the method described in [10]: it is clear that $E\xi$ does not depend on the choice of h. In the meantime, the variance $D\xi = E\xi^2 - (E\xi)^2$ does depend on h. Therefore it is natural to regard $h^1, ..., h^m$ as controls and to choose them such that the variance $D\xi$ is minimal. This problem is solved in [10] and it turns out that, in principle, the variance can be reduced to zero.

Theorem 3.1. *Let the solution $v(t,x)$ of the problem (1.5)-(1.6) be positive. Let*

$$h^j = -\frac{1}{v}\sum_{i=1}^{m} \sigma_{ij} \frac{\partial v}{\partial x^i}. \tag{3.5}$$

Suppose that for any (t,x), $t_0 \le t \le T$, $x \in R^m_+$, there is a solution of the system (2.9)-(2.11), with h^j as in (3.5), for $t \le s \le T$. Then, ξ in (3.3), computed according to (2.9)-(2.11) with h as in (3.5), is deterministic, i.e., $D\xi = 0$.

Proof. By using Itô's formula and taking into account $Lv + c = 0$, we derive

$$d[v(s, X_{t,x}(s))\, Y_{t,x,1}(s) + Z_{t,x,1,0}(s)] = (Lv + c)\, Y ds - \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(\sigma h)^i\, Y ds$$

$$+ \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}\, Y (\sigma dW(s))^i + vYh^\top dW(s) + \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(\sigma dW(s))^i\, Yh^\top dW(s)$$

$$= Y\,(\sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(\sigma dW(s))^i + vh^\top dW(s)) = Y\sum_{j=1}^{m}(\sum_{i=1}^{m}\sigma_{ij}\frac{\partial v}{\partial x^i} + vh^j)dW^j(s),$$

whence

$$v(s, X_{t,x}(s))\, Y_{t,x,1}(s) + Z_{t,x,1,0}(s) = v(t,x) + \int_t^s Y\sum_{j=1}^{m}(\sum_{i=1}^{m}\sigma_{ij}\frac{\partial v}{\partial x^i} + vh^j)dW^j. \tag{3.6}$$

Now, with h as in (3.5), equation (3.6) yields the following identity in s, x with probability1:

$$\xi(s) := v(s, X_{t,x}(s))\, Y_{t,x,1}(s) + Z_{t,x,1,0}(s) \equiv v(t,x), \tag{3.7}$$

i.e., $\xi(s)$ is deterministic. Moreover, $\xi(s)$ is independent of $t \le s \le T$. In particular, by (1.6), we get for $s = T$,

$$\xi(T) = \xi = f(X_{t,x}(T))\, Y_{t,x,1}(T) + Z_{t,x,1,0}(T) = v(t,x), \tag{3.8}$$

hence the theorem is proved.

>From the proof of the theorem above we obtain the following corollary.

Corollary 3.1. For an arbitrary h (of course, the usual conditions of smoothness and boundedness are supposed) the variance $D\xi(T)$ is equal to

$$D\xi(T) = E\int_t^T Y_{t,x,1}^2(s)\cdot\sum_{j=1}^{m}(\sum_{i=1}^{m}\sigma_{ij}\frac{\partial v}{\partial x^i} + vh^j)^2 ds,$$

where the functions $\sigma_{ij},\ \partial v/\partial x^i,\ v,\ h^j$ have $s, X_{t,x}(s)$ as their arguments.

Remark 3.1. Of course, the $h^j,\ j = 1, ..., m$, cannot be constructed without knowing the function v. Nevertheless, the obtained result establishes that, in principle, it is possible to reduce the variance $D\xi$ arbitrarily by properly choosing the functions h^j. The results can be used, e.g., in the following situation. Let all the parameters of the considered problem be close to those one for which the solution is known and equal to v_0. By taking h^j as in (3.5) equal to

$$h^j = -\frac{1}{v_0}\sum_{i=1}^{m}\sigma_{ij}\frac{\partial v_0}{\partial x^i}, \tag{3.9}$$

the variance $D\xi$, although not zero, will be small. As another possibility, it is shown in [20] that under certain circumstances it is optimal to pre-compute a rough approximation for the solution of the Cauchy problem by some finite difference method and then to proceed with variance reduced Monte Carlo simulation, where the controls h^j are computed from the rough approximation.

Remark 3.2. If the condition $v > 0$ in Theorem 3.1 is not satisfied, but e.g., if $v > -K,\ K > 0$, then we consider $\widetilde{v} = v + K$ as a solution of the problem

$$L\widetilde{v} + Kr + c = 0,\ \widetilde{v}(T,x) = f(x) + K$$

and consider instead of (2.11),

$$d\widetilde{Z} = (Kr(s,X) + c(s,X))Y ds,\ \widetilde{Z}(t) = z.$$

Next, taking

$$\widetilde{h}^j = -\frac{1}{v+K}\sum_{i=1}^{m}\sigma_{ij}\frac{\partial v}{\partial x^i} = -\frac{1}{\widetilde{v}}\sum_{i=1}^{m}\sigma_{ij}\frac{\partial \widetilde{v}}{\partial x^i}$$

in (2.9)-(2.10) leads to $\widetilde{\xi} = (f(X_{t,x}(T)) + K)\, Y_{t,x,1}(T) + \widetilde{Z}_{t,x,1,0}(T)$, as being a deterministic variable.

A remarkable fact now is that the variables $\eta_k,\ k = 1, ..., m$, for h^j as in (3.5) are deterministic as well.

Theorem 3.2. *Under the assumptions of Proposition 3.1 the variables $\eta_k = \eta_k(T),\ k = 1, ..., m$, in (3.4), computed according to (2.9)-(2.11) and (2.17)-(2.19) with h as in (3.5) are deterministic.*

Proof. By differentiating (3.7) with respect to x^k we get

$$\frac{\partial v}{\partial x^k}(t,x) = \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(s, X_{t,x}(s))\, \delta_k X^i_{t,x}(s)\, Y_{t,x,1}(s) + v(s, X_{t,x}(s))\, \delta_k Y_{t,x,1}(s) + \delta_k Z_{t,x,1,0}(s).$$

Thus, we have proved that the variables $\eta_k(s)$ (see (3.2)) are deterministic (moreover they do not depend on $s,\ t \le s \le T$). Therefore all $\eta_k(T)$ are deterministic. Theorem 3.2 is proved.

4. Variance reduction by control variates

We now proceed to the method of control variates. In (2.9)-(2.11), we consider h to be fixed and introduce the new random variable

$$\xi_F(T) = \xi(T) + \int_t^T Y_{t,x,1}(s) \sum_{j=1}^{m} F_j(s, X_{t,x}(s)) dW^j(s), \tag{4.1}$$

where $F_j(s,x)$ are functions depending on (s,x) with good analytical properties but further arbitrary.

Clearly, the expectation $E\xi_F(T)$ is equal to $E\xi(T)$ and does not depend on the choice of F. In the meantime, the variance $D\xi_F(T)$ does depend on F. Also in this situation it turns out that the variance can be reduced to zero.

Theorem 4.1. *Let h in (2.9)-(2.11) be a fixed function. Then for*

$$F_j(s,x) = -(\sum_{i=1}^{m} \sigma_{ij}(s,x) \frac{\partial v}{\partial x^i}(s,x) + v(s,x) h^j(s,x)),\ j = 1, ..., m, \tag{4.2}$$

the variable $\xi_F(T)$ is deterministic, i.e., $D\xi_F(T) = 0$.

Proof. The theorem is a consequence of the following equality (see (3.6)),

$$\xi_F(T) = f(X_{t,x}(T))\, Y_{t,x,1}(T) + Z_{t,x,1,0}(T) + \int_t^T Y_{t,x,1}(s) \sum_{j=1}^{m} F_j(s, X_{t,x}(s)) dW^j(s)$$

$$= v(t,x) + \int_t^T Y_{t,x,1}(s) \sum_{j=1}^{m} (\sum_{i=1}^{m} \sigma_{ij} \frac{\partial v}{\partial x^i} + v h^j) dW^j(s) + \int_t^T Y_{t,x,1}(s) \sum_{j=1}^{m} F_j dW^j(s),$$

where the functions $\sigma_{ij},\ \partial v / \partial x^i,\ v,\ h^j,\ F^j$ have $s, X_{t,x}(s)$ as their arguments.

Clearly,

$$D\xi_F(T) = E \int_t^T Y^2_{t,x,1}(s) \sum_{j=1}^{m} (\sum_{i=1}^{m} \sigma_{ij} \frac{\partial v}{\partial x^i} + v h^j + F_j)^2 ds \tag{4.3}$$

which is equal to zero for F_j according to (4.2). Theorem 4.1 is proved.

Of course, a remark similar to Remark 3.1 applies here as well.

The method of control variates in the case $h = 0$ was first considered by N.J. Newton [15]. Following [15], let us look for $F = (F_1, ..., F_m)$ of the form

$$F_j(s,x) = \sum_{i=1}^{m} \sigma_{ij}(s,x) \sum_{r=1}^{l} c_r \gamma_r^i(s,x), \tag{4.4}$$

where $\gamma_r = (\gamma_r^1, ..., \gamma_r^m)$, $r = 1, ..., l$, are known row vectors and c_r are constants. According to (4.3) we then have

$$D\xi_F(T) = E\int_t^T Y_{t,x,1}^2(s) \sum_{j=1}^{m}(\sum_{i=1}^{m} \sigma_{ij}[\frac{\partial v}{\partial x^i} + \sum_{r=1}^{l} c_r \gamma_r^i] + vh^j)^2 ds. \tag{4.5}$$

However, determination of c_r directly by minimization of the right-hand-side of (4.5) is impossible because the functions v and $\partial v/\partial x^i$ are unknown. But, since we have

$$\begin{aligned} v(s, X_{t,x}(s)) &= E(\xi(T; s, X_{t,x}(s)) \mid \mathcal{F}_s), \\ \frac{\partial v}{\partial x^i}(s, X_{t,x}(s)) &= E(\eta_i(T; s, X_{t,x}(s)) \mid \mathcal{F}_s), \end{aligned} \tag{4.6}$$

where

$$\xi(T; s, X_{t,x}(s)) = f(X_{s,X_{t,x}(s)}(T))\, Y_{s,X_{t,x}(s),1}(T) + Z_{s,X_{t,x}(s),1,0}(T),$$

$$\eta_i(T; s, X_{t,x}(s)) = \sum_{k=1}^{m} \frac{\partial f}{\partial x^k}(X_{s,X_{t,x}(s)}(T))\, \delta_i X^k_{s,X_{t,x}(s)}(T)\, Y_{s,X_{t,x}(s),1}(T)$$

$$+ f(X_{s,X_{t,x}(s)}(T))\, \delta_i Y_{s,X_{t,x}(s),1}(T) + \delta_i Z_{s,X_{t,x}(s),1,0}(T),$$

it is not difficult to see that the mentioned minimization problem is equivalent to the following one

$$E\int_t^T Y_{t,x,1}^2(s) \sum_{j=1}^{m}(\sum_{i=1}^{m} \sigma_{ij}[\eta_i(T;\cdot) + \sum_{r=1}^{l} c_r \gamma_r^i] + \xi(T;\cdot)h^j)^2 ds \to \min_{c_1,\dots,c_l}, \tag{4.7}$$

where the functions σ_{ij}, γ_r^i, $\xi(T;\cdot)$, $\eta_i(T;\cdot)$ have $s, X_{t,x}(s)$ as their arguments. Indeed, let us first take the mathematical expectation inside the integral in (4.7). Next, by setting the derivatives of (4.7) with respect to c_r, $r = 1, \dots, l$, equal to zero we get a system of linear equations in c_r, where the coefficients are expressed as integrals of mathematical expectations. By pre-conditioning these expectations on $\mathcal{F}_s$ and using (4.6), it then follows that this system of linear equations coincides with the system yielded by (formally) solving the minimization problem (4.5) directly.

It should be noted further that for simulating $\xi(T;\cdot)$ and $\eta_i(T;\cdot)$ in (4.7) the following relationships are useful:

$$Y_{s,X_{t,x}(s),1}(T) = \frac{Y_{t,x,1}(T)}{Y_{t,x,1}(s)}, \qquad Z_{s,X_{t,x}(s),1,0}(T) = \frac{1}{Y_{t,x,1}(s)}(Z_{t,x,1,0}(T) - Z_{t,x,1,0}(s)),$$

and similar ones for $\delta_i Y_{s,X_{t,x}(s),1}(T)$ and $\delta_i Z_{s,X_{t,x}(s),1,0}(T)$.

The solution of the problem (4.7) thus provides optimal values for c and leads to reduced variance.

To conclude we connect the method of importance sampling and the method of control variates by introduction of the system

$$dX = (b(s,X) - \sigma(s,X)h(s,X))ds + \sigma(s,X)dW(s),\ X(t) = x, \tag{4.8}$$

$$dY = -r(s,X)Y ds + h^\top(s,X)Y dW(s),\ Y(t) = 1, \tag{4.9}$$

$$dZ = c(s,X)Y ds + F^{\top}(s,X)Y dW(s),\ Z(t) = 0, \tag{4.10}$$

and the random variables $\xi(s)$, $\eta_k(s)$ defined according to (3.1), (3.2). Of course, instead of (2.19) the equation for $\delta_k Z$ is now given by

$$d\delta_k Z = \sum_{l=1}^{m} \frac{\partial c(s,X)}{\partial x^l}\,\delta_k X^l\, Y ds + c(s,X)\,\delta_k Y ds \tag{4.11}$$

$$+\sum_{l=1}^{m} \frac{\partial F^{\top}(s,X)}{\partial x^l}\,\delta_k X^l\, Y dW(s) + F^{\top}(s,X)\,\delta_k Y dW(s),\ \delta_k Z(t) = 0.$$

Note that the variables $\xi(s)$ and $\eta_k(s)$ depend on h and F and therefore a more correct notation would be, for example, $\xi_{h,F}(s)$ instead of $\xi(s)$. However, our notation does not lead to any confusion.

The following theorem can be proved analogously to the previous ones.

Theorem 4.2. *Let h and F be such that*

$$\sum_{i=1}^{m} \sigma_{ij}\frac{\partial v}{\partial x^i} + vh^j + F_j = 0,\ j = 1,\dots,m. \tag{4.12}$$

Then $\xi(T)$ from (3.3) computed according to (4.8)-(4.10) and $\eta_k(T)$ in (3.4) computed according to (2.17)-(2.18), (4.11) are deterministic.

Example 4.1. Let all the parameters r, μ^i, ν_{ij}, c, r^i be independent of x, i.e., they are given by deterministic functions of t and let the payoff function f be of the form,

$$f(X(T)) = f_1(X^1(T)) + \dots + f_m(X^m(T)).$$

Then, for $h = 0$ the system (2.9)-(2.11) becomes,

$$dX^i = X^i\,(r(s) - r^i(s))ds + X^i \sum_{j=1}^{m} \nu_{ij}(s)dW^j(s),\ X^i(t) = x^i,\ i = 1,\dots,m,$$

$$dY = -r(s)Y ds,\ Y(t) = 1,$$

$$dZ = c(s)Y ds,\ Z(t) = 0,\ t \le s \le T,$$

and may be solved explicitly:

$$X^i_{t,x}(T) = x^i\, k^i(t)\, \exp\Big(\int_t^T \sum_{j=1}^{m} \nu_{ij}(s)dW^j(s)\Big) = x^i\, k^i(t)\, \exp(\alpha^i\lambda_i(t)),$$

where

$$k^i(t) = \exp\Big(\int_t^T (r(s) - r^i(s))ds - \frac{1}{2}\int_t^T \sum_{j=1}^{m} \nu_{ij}^2(s)ds\Big),$$

$$\lambda_i(t) = \Big(\int_t^T \sum_{j=1}^{m} \nu_{ij}^2(s)ds\Big)^{1/2},$$

and α^i is a normal random variable with zero mean and variance 1.

>From (2.8) we obtain

$$v^0(t,x^1,\dots,x^m) = \sum_{i=1}^{m} E\left[f_i(X^i_{t,x}(T))\, Y_{t,x,1}(T) + Z_{t,x,1,0}(T)\right]$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{m} \int_{-\infty}^{\infty} f_i(x^i k^i(t)\exp(\alpha\lambda_i(t)))\, \exp(-\alpha^2/2)d\alpha\, \exp\Big(-\int_t^T r(s)ds\Big)$$

$$+\int_t^T c(s)\exp(-\int_s^T r(s')ds')ds,$$

whence the derivatives $\partial v^0/\partial x^i$, $i=1,...,m$, can be found explicitly as well.

So, in the case where the parameters of a certain problem do not differ too much from the ones considered above, we can use the recommendation of Remark 3.1 and, for example, take h^j according to (3.9) with $F_j=0$ or $h^j=0$ with $F_j=-\sum_{i=1}^m \sigma_{ij}(s,x)\partial v^0(s,x)/\partial x^i$.

5. Gamma, vega, theta

Clearly, differentiation with respect to x^j in (2.15) gives the probabilistic representation for the gammas $\partial^2 v(t,x)/\partial x^k\partial x^j$, $i,k=1,...,m$. This representation involves, along with the first variations $\delta_k X^i$, $\delta_k Y$, $\delta_k Z$, the second variations

$$\delta_{kj}X^i(s):=\frac{\partial^2 X^i_{t,x}(s)}{\partial x^k\partial x^j},\ \delta_{kj}Y(s):=\frac{\partial^2 Y_{t,x,1}(s)}{\partial x^k\partial x^j},\ \delta_{kj}Z(s):=\frac{\partial^2 Z_{t,x,1,0}(s)}{\partial x^k\partial x^j},\ t\le s\le T.$$

Let us write down the system for these variables where, for notational simplicity, we restrict ourselves to the case $m=1$. In this case X, b, h, σ and W in (2.8)-(2.11) are scalars. For the delta we have

$$u(t,x)=\frac{\partial v}{\partial x}(t,x) \tag{5.1}$$

$$=E\left[\frac{df}{dx}(X_{t,x}(T))\,\delta X(T)\,Y_{t,x,1}(T)+f(X_{t,x}(T))\,\delta Y(T)+\delta Z(T)\right],$$

where (together with (2.9)-(2.11))

$$d\delta X=\frac{\partial(b(s,X)-\sigma(s,X)h(s,X))}{\partial x}\,\delta X\,ds+\frac{\partial\sigma(s,X)}{\partial x}\,\delta X\,dW(s),\ \delta X(t)=1, \tag{5.2}$$

$$d\delta Y=-\frac{\partial r(s,X)}{\partial x}\,\delta X\,Y ds-r(s,X)\,\delta Y ds \tag{5.3}$$

$$+\frac{\partial h(s,X)}{\partial x}\,\delta X\,YdW(s)+h(s,X)\,\delta YdW(s),\ \delta Y(t)=0,$$

$$d\delta Z=\frac{\partial c(s,X)}{\partial x}\,\delta X\,Y ds+c(s,X)\,\delta Y ds,\ \delta Z(t)=0. \tag{5.4}$$

With the notation,

$$\gamma X(s):=\frac{\partial^2 X_{t,x}(s)}{\partial x^2},\ \gamma Y(s):=\frac{\partial^2 Y_{t,x,1}(s)}{\partial x^2},\ \gamma Z(s):=\frac{\partial^2 Z_{t,x,1,0}(s)}{\partial x^2},$$

we thus obtain for the gamma

$$u^\gamma(t,x):=\frac{\partial^2 v}{\partial x^2}(t,x)=E\left[\frac{d^2 f}{dx^2}(X_{t,x}(T))\,[\delta X(T)]^2\,Y_{t,x,1}(T)\right] \tag{5.5}$$

$$+E\left[\frac{df}{dx}(X_{t,x}(T))\,[\gamma X(T)\,Y_{t,x,1}(T)+2\delta X(T)\,\delta Y(T)]+f(X_{t,x}(T))\,\gamma Y(T)+\gamma Z(T)\right],$$

where

$$d\gamma X=\frac{\partial(b(s,X)-\sigma(s,X)h(s,X))}{\partial x}\,\gamma X ds+\frac{\partial\sigma(s,X)}{\partial x}\,\gamma X dW(s) \tag{5.6}$$

$$+\frac{\partial^2(b(s,X)-\sigma(s,X)h(s,X))}{\partial x^2}\,[\delta X]^2ds+\frac{\partial^2\sigma(s,X)}{\partial x^2}\,[\delta X]^2dW(s),\ \gamma X(t)=0,$$

$$(5.7)\qquad d\gamma Y=-\frac{\partial r(s,X)}{\partial x}\gamma X\,Y ds-r(s,X)\,\gamma Y ds+\frac{\partial h(s,X)}{\partial x}\gamma X\,Y dW(s)$$
$$+h(s,X)\,\gamma Y dW(s)-\frac{\partial^2 r(s,X)}{\partial x^2}[\delta X]^2\,Y ds-2\frac{\partial r(s,X)}{\partial x}\delta X\,\delta Y ds$$
$$+\frac{\partial^2 h(s,X)}{\partial x^2}[\delta X]^2\,Y dW(s)+2\frac{\partial h(s,X)}{\partial x}\delta X\,\delta Y dW(s),\ \gamma Y(t)=0,$$

$$(5.8)\qquad d\gamma Z=\frac{\partial c(s,X)}{\partial x}\gamma X\,Y ds+c(s,X)\,\gamma Y ds$$
$$+\frac{\partial^2 c(s,X)}{\partial x^2}[\delta X]^2\,Y ds+2\frac{\partial c(s,X)}{\partial x}\delta X\,\delta Y ds,\ \delta Z(t)=0.$$

Thus, to calculate the gamma one needs to evaluate the expectation (5.5) by virtue of the system consisting of equations (2.9)-(2.11), (5.2)-(5.4)) and (5.6)-(5.8).

Also, one can prove that for h^j as in (3.5) the corresponding gamma estimators are deterministic again.

If the problem under consideration depends on some parameter α, hence $v=v(t,x;\alpha)$, it is possible to find $\partial v(t,x;\alpha)/\partial\alpha$, called vega, in the same way. Let us find, for example, vega in the case of one-dimensional model, i.e. (1.1) with $m=1$, where instead of $\sigma(t,x)=x\nu(t,x)$ we now have $\sigma(t,x;\alpha)=\alpha x\nu(t,x)$. We then have

$$(5.9)\qquad v(t,x;\alpha)=E[f(X_{t,x}(T;\alpha))\,Y_{t,x,1}(T;\alpha)+Z_{t,x,1,0}(T;\alpha)],$$

where

$$(5.10)\qquad dX=(b(s,X)-\sigma(s,X;\alpha)h(s,X;\alpha))ds+\sigma(s,X;\alpha)dW(s),\ X(t)=x,$$

$$(5.11)\qquad dY=-r(s,X)Y ds+h(s,X;\alpha)Y dW(s),\ Y(t)=y,$$

$$(5.12)\qquad dZ=c(s,X)Y ds,\ Z(t)=z.$$

Therefore,

$$(5.13)\qquad \frac{\partial v}{\partial\alpha}(t,x;\alpha)=E\left[\frac{df}{dx}(X_{t,x}(T;\alpha))\,\delta_\alpha X(T;\alpha)Y_{t,x,1}(T;\alpha)\right]$$
$$+E[f(X_{t,x}(T;\alpha))\,\delta_\alpha Y(T;\alpha)+\delta_\alpha Z(T;\alpha)],$$

where

$$\delta_\alpha X(s;\alpha)=\frac{\partial X_{t,x}(s;\alpha)}{\partial\alpha},\ \delta_\alpha Y(s;\alpha)=\frac{\partial Y_{t,x,1}(s;\alpha)}{\partial\alpha},\ \delta_\alpha Z(s;\alpha)=\frac{\partial Z_{t,x,1,0}(s;\alpha)}{\partial\alpha}$$

satisfy the following system

$$(5.14)\quad d\delta_\alpha X=\frac{\partial(b-\sigma h)}{\partial x}\delta_\alpha X ds+\frac{\partial\sigma}{\partial x}\delta_\alpha X dW(s)-\frac{\partial(\sigma h)}{\partial\alpha}ds+\frac{\partial\sigma}{\partial\alpha}dW(s),\ \delta_\alpha X(t)=0,$$

$$(5.15)\qquad d\delta_\alpha Y=-\frac{\partial r}{\partial x}\delta_\alpha X\,Y ds-r\,\delta_\alpha Y ds$$
$$+\frac{\partial h}{\partial x}\delta_\alpha X\,Y dW(s)+h\delta_\alpha Y dW(s)+\frac{\partial h}{\partial\alpha}Y dW(s),\ \delta_\alpha Y(t)=0,$$

$$(5.16)\qquad d\delta_\alpha Z=\frac{\partial c}{\partial x}\delta_\alpha X\,Y ds+c\delta_\alpha Y ds,\ \delta_\alpha Z(t)=0.$$

Let us now point out how to find theta: $u_{m+1}(t,x):=\partial v(t,x)/\partial t$. The above way of differentiation under the expectation sign is now impossible because of the non-differentiability of $X_{t,x}(s)$ with respect to t (e.g., the problem $dX = dW(s),\ X(t) = x,\ s \geq t$, has the solution $X_{t,x}(s) = x + W(s) - W(t)$ which is evidently nondifferentiable with respect to t). Of course, one can find theta by the initial equation (1.5) after evaluating the deltas and the gammas. However, if we do not need the gammas actually, this way is rather irrational. It is better to consider the system of $m+2$ parabolic equations consisting of (1.5)-(1.6), (2.13)-(2.14) and

$$\frac{\partial u_{m+1}}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(t,x)\frac{\partial^2 u_{m+1}}{\partial x^i \partial x^j} + \sum_{i=1}^{m} b^i(t,x)\frac{\partial u_{m+1}}{\partial x^i} - r(t,x)\,u_{m+1} \tag{5.17}$$
$$+\frac{1}{2}\sum_{i,j=1}^{m}\frac{\partial a_{ij}}{\partial t}(t,x)\frac{\partial u_j}{\partial x^i} + \sum_{i=1}^{m}\frac{\partial b_i}{\partial t}(t,x)\frac{\partial v}{\partial x^i} - \frac{\partial r}{\partial t}(t,x)\,v + \frac{\partial c}{\partial t}(t,x) = 0,$$

$$u_{m+1}(T,x) = -\frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(T,x)\frac{\partial^2 f}{\partial x^i \partial x^j}(x) - \sum_{i=1}^{m} b^i(T,x)\frac{\partial f}{\partial x^i}(x) \tag{5.18}$$
$$+r(T,x)\,f(x) - c(T,x) := g(x),$$

and to use then consequently the probabilistic representations given in [11].

Let us now consider a model in which the coefficients ν_{ij} (and so a_{ij}) do not depend on t. In such case the parabolic system consists of two equations for v and u_{m+1} only. Namely, (1.5)-(1.6) and the following equation

$$\frac{\partial u_{m+1}}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(t,x)\frac{\partial^2 u_{m+1}}{\partial x^i \partial x^j} + \sum_{i=1}^{m} b^i(t,x)\frac{\partial u_{m+1}}{\partial x^i} - r(t,x)\,u_{m+1} \tag{5.19}$$
$$+\sum_{i=1}^{m}\frac{\partial b_i}{\partial t}(t,x)\frac{\partial v}{\partial x^i} - \frac{\partial r}{\partial t}(t,x)\,v + \frac{\partial c}{\partial t}(t,x) = 0,$$

$$u_{m+1}(T,x) = g(x), \tag{5.20}$$

with $g(x)$ as in (5.18) and $a_{ij}(T,x) = a_{ij}(x)$.

The probabilistic representation for the solution of Cauchy problem (1.5)-(1.6), (5.19)-(5.20) has the following simple form (see [11]). Introduce the system of stochastic differential equations

$$dX = (b(s,X) - \sigma(s,X)h(s,X))ds + \sigma(s,X)dW(s),\ X(t) = x, \tag{5.21}$$

$$dY^1 = -r(s,X)Y^1 ds - \frac{\partial r(s,X)}{\partial s}Y^2 ds + h^\top(s,X)Y^1 dW(s),\ Y^1(t) = y^1, \tag{5.22}$$

$$dY^2 = -r(s,X)Y^2 ds + h^\top(s,X)Y^2 dW(s) \tag{5.23}$$
$$+(\sigma^{-1}(s,X)\frac{\partial b(s,X)}{\partial s})^\top Y^2 dW(s),\ Y^2(t) = y^2,$$

$$dZ = c(s,X)Y^1 ds + \frac{\partial c(s,X)}{\partial s}Y^2 ds,\ Z(t) = 0, \tag{5.24}$$

and the random variable

$$\xi_{t,x,y^1,y^2} = f(X_{t,x}(T))\,Y^1_{t,x,y^1,y^2}(T) + g(X_{t,x}(T))\,Y^2_{t,x,y^1,y^2}(T) + Z_{t,x,y^1,y^2,0}(T), \tag{5.25}$$

where Y^1 and Y^2 are scalars. Then, the required solution $v(t,x)$, $u_{m+1}(t,x)$ can be found from the relations

$$v(t,x) = E\xi_{t,x,1,0}\ ,\ \ u_{m+1}(t,x) = E\xi_{t,x,0,1}\ . \tag{5.26}$$

This fact can be verified in the following way. By Itô's formula we derive

$$\begin{aligned} (5.27) \qquad & v(s, X_{t,x}(s))\, Y^1_{t,x,y^1,y^2}(s) + u_{m+1}(s, X_{t,x}(s))\, Y^2_{t,x,y^1,y^2}(s) + Z_{t,x,y^1,y^2,0}(s) \\ & -v(t,x)\, y^1 - u_{m+1}(t,x)\, y^2 \\ & = \int_t^s \left[F_1^\top(\vartheta, X_{t,x}(\vartheta) \cdot Y^1_{t,x,y^1,y^2}(\vartheta) + F_2^\top(\vartheta, X_{t,x}(\vartheta) \cdot Y^2_{t,x,y^1,y^2}(\vartheta)\right] dW(\vartheta), \end{aligned}$$

where F_1 and F_2 are some known vector-functions and then relations (5.26) follow immediately.

6. Barrier options

Let us consider a model consisting of a cash bond $B(s)$ and a stock $X(s)$ (we take only one stock for notational simplicity), where the price of the stock satisfies the equation

$$dX = \mu(s, X)ds + \sigma(s, X)dW(s). \tag{6.1}$$

Let $0 \leq \pi_1 < \pi_2,\ \ \pi_1 < x < \pi_2,\ \ \tau = \tau_{t,x} = T \wedge \inf\{s : X_{t,x}(s) \notin [\pi_1, \pi_2],\ t \leq s \leq T\}$ (we put inf to be equal ∞ for an empty set). We now look at an example of a barrier option. The option is specified by a payoff equal to zero if $\tau < T$ and equal to $f(X_{t,x}(T))$ if $\tau = T$, where $f(x)$ is a function defined on $[\pi_1, \pi_2]$. We should note here that a more rigorous notation for (6.1) would be

$$dX = 1_{\{\tau > s\}}\mu(s, X)ds + 1_{\{\tau > s\}}\sigma(s, X)dW(s),$$

where X describes the value of the stock up to time τ, but we use (6.1) as long as it doesn't lead to confusion. In addition, we assume that $f(x)$ is equal to zero in some neighborhood of π_1 and π_2 respectively. Then, it is not difficult to show that the portfolio value $V(t)$ of the hedging strategy is equal to $v(t, X(t))$ where $v(t,x)$ satisfies the following boundary value problem

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 v}{\partial x^2} + r(t)x\frac{\partial v}{\partial x} - r(t)v = 0,\ t_0 \leq t < T,\ \pi_1 < x < \pi_2, \tag{6.2}$$

$$v(T,x) = f(x),\ v(t,\pi_1) = v(t,\pi_2) = 0, \tag{6.3}$$

and as before we have

$$V(t) = v(t, X(t)) = \varphi_t B(t) + \psi_t X(t),$$

with

$$\varphi_t = \frac{1}{B(t)}(v(t, X(t)) - \frac{\partial v}{\partial x}(t, X(t))X(t)),\ \psi_t = \frac{\partial v}{\partial x}(t, X(t)).$$

Note that in contrast to (1.1), in this model we do not use the multiplier X in the stock equation . This can be explained in the following way. The multipliers X^i in (1.1) ensure the positivity property of prices X^i during infinite time. Analogously, the multipliers $(X^i - \pi_1^i)(\pi_2^i - X^i)$ in (2.7) ensure evolving the prices in the open parallelepiped Π. In the here considered problem the stock price process is only relevant in the interval $[\pi_1, \pi_2]$, due to the definition of the barrier option and so we don't need any special multipliers.

The solution of the boundary value problem (6.2)-(6.3) for the barrier option has the following probabilistic representation,

$$v(t,x) = E1_{\{\tau_{t,x} = T\}}[f(X_{t,x}(T))\, Y_{t,x,1}(T)], \tag{6.4}$$

where

$$\begin{aligned} (6.5) \qquad & dX = (r(t)X - \sigma(s,X)h(s,X))ds + \sigma(s,X)dW(s),\ X(t) = x, \\ & dY = -r(t)Y ds + h(s,X)Y dW(s),\ Y(t) = 1, \end{aligned}$$

and

$$(6.6)\qquad \frac{\partial v}{\partial x}(t,x) = E1_{\{\tau_{t,x}=T\}}\left[\frac{\partial f}{\partial x}(X_{t,x}(T))\cdot\delta X(T)\cdot Y_{t,x,1}(T) + f(X_{t,x}(T))\cdot\delta Y(T)\right],$$

where the equations for $\delta X(T)$ and $\delta Y(T)$ are analogous to (2.17), (2.18).

The method of importance sampling can be developed for this model without any serious difficulties (see [10] in connection with variance reduction of $v(t,x)$ due to (6.4)). After introducing a system similar to (4.8)-(4.10) a more general representations for v and $\partial v/\partial x$ can be given and also the results of Section 4 (in particular the method of control variates) can be carried over to the barrier option considered above.

Remark 6.1. The option under consideration is known as nullified barrier option [9]. For more general barrier options the boundary value conditions are nonzero and instead of (6.3) we have

$$(6.7)\qquad v(T,x) = f(x),\ v(t,\pi_1) = v_1(t),\ v(t,\pi_2) = v_2(t).$$

Let Γ denote the set where the condition (6.7) is specified. Then (6.7) can be written as

$$(6.8)\qquad v\mid_\Gamma = g,$$

where $g(T,x) = f(x),\ g(t,\pi_1) = v_1(t),\ g(t,\pi_2) = v_2(t)$.

Instead of (6.4) we may now write

$$(6.9)\qquad v(t,x) = E[g(\tau_{t,x}, X_{t,x}(\tau_{t,x}))\cdot Y_{t,x,1}(\tau_{t,x})].$$

It should be noted that in this case there is no expression for $\partial v(t,x)/\partial x$ such as (6.6) because the dependence on x is now more complicated due to the presence of $\tau_{t,x}$ and the problem of effective numerical construction of a hedging strategy requires special examination.

7. Hedging of European LIBOR derivative claims

In this section we show the application of the presented probabilistic methods to the LIBOR[1] interest rate model in connection with some specific European LIBOR derivative claims. The LIBOR (market) model by Brace, Gatarek, Musiela [1] and Jamshidian [6] is based on an arbitrage free system of zero coupon bonds and in the frame work of Jamshidian it may be represented as below (we now use the notation from [6] which differs a little from the one used in the previous sections).

For a given tenor structure $0 < T_1 < T_2 < \ldots < T_m$ we consider the forward LIBOR process $L_i(t)$, $0 \le t \le T_i$, $1 \le i \le m-1$, as the effective forward rate over the period $[T_i, T_{i+1}]$, which is defined in terms of zero coupon bonds B_i, B_{i+1}, maturing at T_i, T_{i+1} respectively, by

$$L_i(t) := \delta_i^{-1}\left(\frac{B_i(t)}{B_{i+1}(t)} - 1\right),$$

with $\delta_i := T_{i+1} - T_i$. Then, in the so called terminal bond measure P_m, which is a measure such that B_i/B_m are P_m (local) martingales, the dynamics of the LIBOR process is given by

$$(7.1)\qquad dL_i = -\sum_{j=i+1}^{m-1}\frac{\delta_j L_i L_j\,\gamma_i^\top\gamma_j}{(1+\delta_j L_j)}\,dt + L_i\,\gamma_i^\top(t,L)dW^{(m)},$$

[1] LIBOR stands for London Inter Bank Offer Rate

where $W^{(m)}$ is a d-dimensional P_m-Brownian motion, $d \leq m-1$, and for our purposes we assume that the R^d-valued functions γ_i, defined in $[t_0, T_i] \times R_+^{m-1}$, are smooth, bounded, and such that the matrix $\Sigma := (\gamma_i^\top \gamma_j)$ has constant rank d.

We now illustrate how to price and hedge a European LIBOR derivative claim at maturity time T_1 with a typical degree 1 homogeneous payoff structure of the form $V(T_1, B) := B_m(T_1) f(L(T_1))$ for some $f(\cdot) \ : \mathbb{R}^{m-1} \to \mathbb{R}$. By application of Theorems 4.7 and 5.2 in [6] in this setting it follows that the claim value at time $t \leq T_1$ has the form

$$V(t, B(t)) = B_m(t) v(t, L(t)),$$

where $v(\cdot,\cdot): \ \mathbb{R}\times\mathbb{R}^{m-1} \to \mathbb{R}$ satisfies the Cauchy problem

$$\frac{\partial v}{\partial t} - \sum_{k=1}^{m-1} \sum_{p=k+1}^{m-1} \frac{\partial v}{\partial y_k} \frac{\delta_p y_p y_k}{1+\delta_p y_p} \gamma_k^\top \gamma_p + \frac{1}{2} \sum_{k,l=1}^{m-1} \frac{\partial^2 v}{\partial y_k \partial y_l} y_k y_l \, \gamma_k^\top \gamma_l = 0, \tag{7.2}$$

with boundary condition

$$v(T_1, y) = f(y).$$

For convenience we introduce,

$$\lambda_k(t, y) = - \sum_{p=k+1}^{m-1} \frac{\delta_p y_p y_k}{1+\delta_p y_p} \gamma_k^\top \gamma_p,$$

$$\eta_{kl}(t, y) = y_k y_l \, \gamma_k^\top \gamma_l.$$

and then the Cauchy problem for v reads,

$$\frac{\partial v}{\partial t} + \sum_{k=1}^{m-1} \lambda_k(t, y) \frac{\partial v}{\partial y_k} + \frac{1}{2} \sum_{k,l=1}^{m-1} \eta_{kl}(t, y) \frac{\partial^2 v}{\partial y_k \partial y_l} = 0, \quad v(T, y) = f(y). \tag{7.3}$$

Following Jamshidian, [6], Th. 4.7, the European claim may be hedged by a self-financing portfolio in zero coupon bonds (ψ, B),

$$V(t, B(t)) := \sum_{k=1}^{m} \psi_k(t, B(t)) B_k(t) = V(0, B(0)) + \sum_{k=1}^{m} \int_0^t \psi_k(s, B(s)) dB_k(s), \tag{7.4}$$

where

$$\psi_k(t, x) := \frac{\partial V}{\partial x_k}.$$

Since we may write

$$\begin{aligned} V(t, x) = V(t, x_1, ..., x_m) &= x_m v(t, \delta_1^{-1}(\frac{x_1}{x_2} - 1), ..., \delta_{m-1}^{-1}(\frac{x_{m-1}}{x_m} - 1)) \\ &= x_m v(t, y_1, .., y_{m-1}) = x_m v(t, y), \end{aligned}$$

where (y_k corresponds to L_k)

$$y_k := \delta_k^{-1}(\frac{x_k}{x_{k+1}} - 1), \quad k = 1, ..., m-1,$$

V is homogeneous of degree 1 in x and so the hedge quantities ψ_k are homogeneous of degree 0 and may be seen as functions of t and y. We will thus derive a representation for the hedge quantities ψ_i in terms of t and y. For $i < m$ we have

$$\psi_k(t, x) = \frac{\partial V}{\partial x_i} = x_m \sum_{k=1}^{m-1} \frac{\partial v}{\partial y_k} \frac{\partial y_k}{\partial x_i} = -x_m \frac{\partial v}{\partial y_{i-1}} \delta_{i-1}^{-1} \frac{x_{i-1}}{x_i^2} + x_m \frac{\partial v}{\partial y_i} \delta_i^{-1} \frac{1}{x_{i+1}} \tag{7.5}$$

$$= -\frac{\partial v}{\partial y_{i-1}}\delta_{i-1}^{-1}\cdot(1+\delta_{i-1}y_{i-1})\prod_{k=i}^{m-1}\frac{1}{1+\delta_k y_k}+\frac{\partial v}{\partial y_i}\delta_i^{-1}\cdot\prod_{k=i+1}^{m-1}\frac{1}{1+\delta_k y_k} =: \varphi_i(t,y)$$

and

(7.6)

$$\psi_m(t,x)=\frac{\partial V}{\partial x_m}=v(t,y)-\frac{1}{x_m}\sum_{i=1}^{m-1}x_i\frac{\partial V}{\partial x_i}=v(t,y)-\sum_{i=1}^{m-1}\prod_{k=i}^{m-1}(1+\delta_k y_k)\varphi_i(t,y)=:\varphi_m(t,y).$$

Therefore the hedging strategy is constructed by

$$V(t):=\sum_{i=1}^{m}\varphi_i(t,L(t))B_i(t). \tag{7.7}$$

According to (7.5) and (7.6), for calculating $\varphi_i(t,L(t)),\ i=1,...,m$, we have to find v and $\partial v/\partial y_i,\ i=1,...,m-1$. Clearly, in the manner which was outlined in the previous sections, v can be found by Monte Carlo simulation of a suitable probabilistic representation for Cauchy problem (7.3). For its derivatives we set

$$u_i:=\frac{\partial v}{\partial y_i},$$

and differentiate (7.3) with respect to y_i to yield,

$$\frac{\partial u_i}{\partial t}+\sum_{k=1}^{m-1}\frac{\partial\lambda_k}{\partial y_i}u_k+\sum_{k=1}^{m-1}\lambda_k(t,y)\frac{\partial u_i}{\partial y_k}+\frac{1}{2}\sum_{k,l=1}^{m-1}\frac{\partial\eta_{kl}}{\partial y_i}\frac{\partial u_k}{\partial y_l}+\frac{1}{2}\sum_{k,l=1}^{m-1}\eta_{kl}(t,y)\frac{\partial^2 u_i}{\partial y_k y_l}=0. \tag{7.8}$$

So, along with Cauchy problem (7.3) for v, we have the Cauchy problem for u_i : (7.8) with boundary condition

$$u_i(T,y)=\frac{\partial f}{\partial y_i}, \tag{7.9}$$

which can be solved also by the previously presented probabilistic methods.

Remark 7.1. It should be noted that in the case of a low factor LIBOR model ($d << m$) the diffusion matrices in the Cauchy problems (7.3) and (7.8)-(7.9) are highly degenerate, however, the different probabilistic representations in Sections 2-4 still go through. In fact, in Section 2 nondegeneracy was merely assumed to guarantee that the multi-asset system (1.1) is arbitrage free.

The developed general probabilistic method for the price and hedge of a European claim can be applied to various “Over The Counter” European LIBOR derivatives. As an illustration we consider two so-called ”fixed income” instruments: the ”swaption” and the ”callable” reverse floater. Although especially the swaption is a very liquidly traded instrument and therefore will be hedged hardly ever in practice, these examples serve nevertheless as a clear illustration of the methods presented. Besides, in a LIBOR market model, for a fairly large family of LIBOR derivatives one can derive analytical approximation formulas, see e.g. [19], which, in principle, can be used for variance reduction.

Example 7.1. *European swaption.* A swap contract with maturity T_1 and strike κ on a loan of \$1 over the period $[T_1,T_m]$ obliges to pay a fixed coupon κ and receive spot LIBOR at the settlement dates $T_2,...,T_m$. From a standard portfolio argument it is obvious that the present value of this contract is equal to

$$Swap(t) = B_1(t) - B_m(t) - \kappa \sum_{j=1}^{m-1} \delta_j B_{j+1}(t), \ t_0 \le t \le T_1.$$

The swap rate $S(t)$ is now defined as that fixed coupon which sets this contract value to zero:

$$S(t) := \frac{B_1(t) - B_m(t)}{\sum_{j=1}^{m-1} \delta_j B_{j+1}(t)}.$$

A swaption contract with maturity T_1, strike κ and principal \$1 gives the right to contract at T_1 to pay a fixed coupon κ and receive spot LIBOR at the settlement dates $T_2, ..., T_m$. Equivalently, one can contract for receiving the T_1-swaprate and one can show that the payoff of the swaption is equivalent to a T_1-cashflow of

$$Swpn(T_1) = \sum_{j=1}^{m-1} 1_A B_{j+1}(T_1)(L_j(T_1) - \kappa)\delta_j, \tag{7.10}$$

where A denotes the $\mathcal{F}_{T_1}$-measurable event $\{S(T_1) > \kappa\}$ and the swaprate $S(T_1)$ is given by (see [19])

$$S(T_1) := \frac{B_1(T_1) - B_m(T_1)}{\sum_{j=1}^{m-1} \delta_j B_{j+1}(T_1)} = \frac{-1 + \prod_{i=1}^{m-1}(1 + \delta_i L_i(T_1))}{\sum_{j=1}^{m-1} \delta_j \prod_{i=j+1}^{m-1}(1 + \delta_i L_i(T_1))} .$$

>From (7.10) we see that the swaption cashflow is homogeneous of degree one. Therefore we may compute the swaption price and the corresponding hedge by Monte Carlo simulation of the probabilistic representations for (7.3), (7.8)-(7.9), with f given by

$$f(y) := \sum_{j=1}^{m-1} 1_A(y)(y_j - \kappa)\delta_j \prod_{k=j+1}^{m-1} (1 + \delta_k y_k), \tag{7.11}$$

where

$$A = \left\{ y : \ \frac{-1 + \prod_{k=1}^{m-1}(1 + \delta_k y_k)}{\sum_{k=1}^{m-1} \delta_k \prod_{i=k+1}^{m-1}(1 + \delta_i y_i)} > \kappa \right\} ,$$

and obtain variance reduction, for instance, by the "industrial standard" Black market formula for swaptions [6, 19].

Example 7.2. *The callable reverse floater.* Let $K > 0$. A reverse floater (RF) contracts for receiving $L_i(T_i)$ while paying $\max(K - L_i(T_i), 0)$ at time T_{i+1} for $i = 1, \ldots, m-1$ with respect to a unit principal. A callable reverse floater (CRF) is an option to enter into a reverse floater at T_1. In [19] it is shown that in a LIBOR market model the reverse floater can be evaluated analytically and that the contract is equivalent with a T_1-cashflow of

$$RF(T_1) = B_1(T_1) - B_m(T_1) - \sum_{i=1}^{m-1} B_{i+1}(T_1) F_i(T_1, K), \tag{7.12}$$

where $F_i(T_1, K)$ is known explicitly as a Black-type formula, only involving T_1, K, and the deterministic γ_i, $i = 1, ..., m-1$, [19]. So the payoff of the CRF, being

$$CRF(T_1) = \max(RF(T_1), 0),$$

is clearly homogeneous of degree one and the reverse floater price and hedge may be computed by Monte Carlo simulation of the probabilistic representations for the system (7.3), (7.8)-(7.9) and f given by an expression derived from (7.12). Moreover, in [19] a one factor approximation formula is derived which could be used for variance reduction.

Remark 7.2. It should be noted that in practice LIBOR derivatives may not be hedged by zero coupon bonds directly as they are usually not always available in the market. Instead, zero coupon bonds may be constructed, for instance, by simple linear combinations of swap contracts and certain particular Government bonds and so the hedge positions in the "virtual" zero bonds need to be translated to these assembling instruments.

8. Concluding remarks

Although the parabolic Cauchy and boundary value problems associated with pricing and hedging of European multi-asset claims are practically impossible to solve when the number of assets exceeds three, the probabilistic methods presented in this article are still feasible and straightforward to implement even when the asset dimension is relatively high. The generalization of these methods to high dimensional American options, for instance, in the spirit of the work of Broadie and Glasserman [2] is to our opinion an interesting and challenging problem for the future.

References

[1] Brace, A., Gatarek, D., Musiela, M. (1997). The Market Model of Interest Rate Dynamics. *Mathematical Finance,* 7 (2), 127-155.

[2] Broadie, M. and Glasserman, P. (1997). Monte Carlo Methods for Pricing High-Dimensional American Options: An Overview, NetExposure, December, 15-37.

[3] Dynkin, E.B. (1965). *Markov Processes.* Springer: Berlin (engl. transl. from Russian 1963).

[4] Freidlin, M.I. (1996). *Markov Processes and Differential Equations: Asymptotic Problems.* Birkhäuser: Basel.

[5] Gladyshev, S.A. and Milstein, G.N. (1984). A Runge-Kutta method for computing Wiener integrals of functionals of exponential type. *Zh. Vychisl. Mat. i Mat. Fiz.*, **24**, no. 8, 1136-1150.

[6] Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics*, **1**, 293-330.

[7] Karatzas, I., Shreve, S.E. (1998). *Methods of Mathematical Finance.* Springer-Verlag Berlin Heidelberg.

[8] Kloeden, P.E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations.* Springer: Berlin.

[9] Kwok, Y.K. (1998). *Mathematical Models of Financial Derivatives.* Springer: Berlin.

[10] Milstein, G.N. (1995). *Numerical Integration of Stochastic Differential Equations.* Kluwer Academic Publishers, (engl. transl. from Russian 1988).

[11] Milstein, G.N. (1978). On a probabilistic solution of linear systems of elliptic and parabolic equations. *Theory Prob. Appl.*, **23**, 851-855.

[12] Milstein, G.N. (1997). Weak approximation of a diffusion process in a bounded domain. *Stochastics and Stochastics Reports*, **62,** 147-200.

[13] Milstein, G.N. (1998). On the mean-square approximation of a diffusion process in a bounded domain. *Stochastics and Stochastics Reports*, **64,** 211-233.

[14] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling.* Springer: Berlin.

[15] Newton, N.J. (1994). Variance reduction for simulated diffusion. *SIAM J. Appl. Math.*, **54**, 1780-1805.

[16] Newton, N.J. (1997). Continuous-time Monte Carlo methods and variance reduction. In *Numerical Methods in Finance*, edited by L.C.G. Rogers and D. Talay, Cambridge University Press, 22-42.

[17] Newton, N.J. (1996). Variance reduced Monte Carlo methods for PDEs'. *Z. angewandte Math. Mech.*, **76**, Suppl. 3, 327-330.

[18] Numerical Methods in Finance. (1997). Edited by L.C.G. Rogers and D. Talay. Cambridge University Press.

[19] Schoenmakers, J.G.M. and Coffey, B. (1999). LIBOR rate models, related derivatives and model calibration. *Preprint no. 480,* Weierstrass Institute Berlin.

[20] Schoenmakers, J.G.M. and Heemink, A.W. (1997). Fast Valuation of Financial Derivatives. *Journal of Computational Finance*, **1**, no. 1, 47-62.

[21] Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and Appl.*, **8**, 483-509.

[22] Wagner, W. (1988). Monte Carlo evaluation of functionals of solutions of stochastic differential equations. Variance reduction and numerical examples. *Stoch. Anal. and Appl.*, 6, 447-468.