Competitive Monte Carlo methods for the pricing of Asian options

B. Lapeyre
CERMICS, Ecole Nationale des Ponts et Chaussées, 77455 Marne La Vallée, France

E. Temam
Université Paris VI, Laboratoire de Probabilité et Modèles Aléatoires and CERMICS, Ecole Nationale des Ponts et Chaussées, 77455 Marne La Vallée, France

We explain how a carefully chosen scheme can lead to competitive Monte Carlo algorithms for the computation of the price of Asian options. We give evidence of the efficiency of these algorithms with a mathematical study of the rate of convergence and a numerical comparison with some existing methods.

1. Introduction

Monte Carlo methods are known to be useful when the state dimension is large. This is widely true, but we will give here an example of a small dimension problem coming from finance where a Monte Carlo method (helped by a variance reduction technique) can be more efficient than other known methods.

This example is based on the price of an Asian option (see subsection 2.1). This problem is known to be computationally hard and a lot of literature deals with this problem, using either analytic methods (Geman and Yor, 1993; Fu *et al.*, 1996), numerical methods based on the associated partial differential equation (Dewynne and Wilmott, 1995, Forsyth *et al.*, 1996, Ingersoll, 1987, Rogers and Shi, 1995) or Monte Carlo methods (Kemna and Vorst, 1990). The originality of this work is to propose (and to give precise results of) new time schemes approximation for the integral of the Black and Scholes model: this point is really a source of concern when using Monte Carlo methods as noted by Fu *et al.* (1996, page 14). We will show that, when we use a suitable scheme and variance reduction, a Monte Carlo method can be more competitive than other methods under some circumstances.

In order to get precise and complete numerical results, we have undertaken extensive comparisons with most of the other known methods: the Forward Shooting Grid (Barraquand and Pudet, 1996), Hull and White (1993), Finite Difference and the already quoted Monte Carlo methods. For moderate precision (10^{-2}), tree methods (FSG and Hull and White) are the most efficient. But, surprisingly, really high precision (10^{-4}) can only be reached by a Monte Carlo method with a relatively large time step (approximately one month).

This article is organized in the following way: we first introduce the mathematical context, then we present the Monte Carlo schemes in sections 1 and 2, section 3 contains the proof of their convergence in the L^p spaces and section 4 briefly describes the variance reduction method used in this article. The end of the paper is devoted to numerical tests and comparisons.

2. Mathematical context

2.1. Financial background

To describe the price of an asset at time t, we use the Black and Scholes model with a risky asset (a share of price S_t at time t) and a no-risk asset (whose price is S_t^0 at time t). The price S_t is given by the following stochastic differential equation

$$\mathrm{d}S_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t)$$

where μ and σ are two positive constants and (B_t) is a standard Brownian motion. The price of the no-risk asset satisfies the ordinary differential equation

$$\mathrm{d}S_t^0 = rS_t^0 \mathrm{d}t$$

As usual, we introduce the process $W_t = B_t + ((\mu - \mathrm{r})/\sigma)t$ which is a Brownian motion under an adapted probability, called the neutral risk probability and denoted by **P**. Thus, the risky asset satisfies a new stochastic differential equation

$$\mathrm{d}S_t = S_t(r\mathrm{d}t + \sigma \mathrm{d}W_t)$$

whose solution is

$$S_t = S_{T_0} \exp\left(\sigma W_t - \frac{\sigma^2}{2}t + rt\right)$$

S_{T0} is the price of the asset at the beginning of the modeling.

Asian options (or options on average) is the general name for a class of options whose payoff depends on the mean of the price of the risky asset in a given period. Thus, the price of an Asian option with maturity T can be written:

$$V(t, S, A) = \mathrm{e}^{-r(T-t)}\mathbf{E}f(S_t, A_S(T_0, t))$$

where

$$A_S(T_0, t) = \frac{1}{t - T_0}\int_{T_0}^{t} S_u \mathrm{d}u$$

From now on, we choose $T_0 = 0$, with no lost of generality. The function f depends on the type of the options

- ❑ For a call with fixed strike: $f(s, a) = (a - K)_+$
- ❑ For a put with fixed strike: $f(s, a) = (K - a)_+$
- ❑ For a call with floating strike: $f(s, a) = (s - a)_+$
- ❑ For a put with floating strike: $f(s, a) = (a - s)_+$

2.2. Practical schemes

If we want to use Monte Carlo methods to compute a price, we have to simulate the average of S_t, therefore we need to approximate an integral. Here, it is not necessary to approximate S_t because it can be exactly simulated. With this aim in view, the interval $[0, T]$ will be divided into N steps. The step size will be noted $h = T/N$, and we define the times $t_k = kT/N = kh$.

The standard scheme

We introduce three schemes to estimate $Y_T = \int_0^T S_u \mathrm{d}u$. Since we are able to simulate S_t at a given t, the integral can be approached by Riemann sums:

$$Y_T^{r,n} = h \sum_{k=0}^{n-1} S_{t_k} \tag{2.1}$$

For example, if M denotes the number of drawings in the Monte Carlo method, an approximation of the price at maturity of a fixed strike Asian call is given by

$$\frac{\mathrm{e}^{-rT}}{M} = \sum_{j=1}^{M} \left(\frac{h}{T} \sum_{k=0}^{n-1} S_{t_k} - K \right)_+$$

Note that the time complexity of this algorithm is $O(1/NM)$ (this is true for every kind of Monte Carlo method) and that it involves two kinds of errors: the Monte Carlo error and the time step error. The Monte Carlo error is of the order $\sigma/\sqrt{M}$ and the time step error is harder to evaluate (see proposition 3.3).

The scheme (2.1) can be interpreted as the second variable in the Euler approximation for the following stochastic differential equation

$$\mathrm{d}U_t = B(U_t)\mathrm{d}t + \Sigma(U_t)\mathrm{d}W_t$$
$$\text{with } U_t = \begin{bmatrix} S_t \\ Y_t \end{bmatrix},\ B(U_t) = \begin{bmatrix} rS_t \\ S_t \end{bmatrix} \text{ and } \Sigma(U_t) = \begin{bmatrix} \sigma(S_t) \\ 0 \end{bmatrix}$$

Higher accuracy schemes

One way to obtain higher accuracy for the integral approximation is to note that in L^2, the closest random variable to $((1/T)\int_0^T S_s \mathrm{d}s - K)_+$ when the $(S_{t_k}, k = 0,\ldots, N)$ are known is given by:

$$\mathbf{E}\left[\left(\frac{1}{T} \int_0^T S_u \mathrm{d}u - K \right)_+ \middle| \mathcal{B}_h \right] \tag{2.2}$$

where $\mathcal{B}_h$ is the σ-field generated by the $(S_{t_k}, k = 0,\ldots, N)$. Of course, it is theoretically impossible to compute explicitly this conditional expectation (it is certainly harder than obtaining an explicit formula for V). But, since the conditional law of W_u with respect to $\mathcal{B}_h$ for $u \in [t_k, t_{k+1}]$ is given by

$$\mathcal{L}\left(W_u \mid W_{T_k} = x,\ W_{T_{k+1}} = y\right) = \mathcal{N}\left(\frac{t_{k+1} - u}{h} x + \frac{u - t_k}{h} y, \frac{(t_{k+1} - u)(u - t_k)}{h}\right) \tag{2.3}$$

one can in principle compute

$$\left[\mathbf{E}\left(\frac{1}{T}\int_0^T S_u \mathrm{d}u \,\middle|\, \mathcal{B}_h\right) - K\right]_+ = \left[\frac{1}{T}\int_0^T \mathbf{E}\left(S_u \mid \mathcal{B}_h\right) \mathrm{d}u - K\right]_+ \tag{2.4}$$

as a function of $(W_{t_k}, k = 0, \ldots, N)$. Jensen inequality proves that (2.4) is smaller than (2.2), but we will see that (2.2) is already a really good approximation of Y_T (see proposition 3.3). Using the law described by (2.3), we get

$$\begin{aligned}
&\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u \,\middle|\, \mathcal{B}_h\right] - K \\
&= \frac{1}{T}\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{\left(r - \frac{\sigma^2}{2}\right) u} e^{\sigma \frac{t_{k+1} - u}{h} W_{t_k} + \sigma \frac{u - t_k}{h} W_{t_{k+1}} + \frac{\sigma^2}{2} \frac{(t_{k+1} - u)(u - t_k)}{h}} \mathrm{d}u \\
&= \frac{1}{T}\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{\sigma \frac{u - t_k}{h}\left(W_{t_{k+1}} - W_{t_k}\right) - \frac{\sigma^2}{2} \frac{(u - t_k)^2}{h} + ru} e^{\sigma W_{t_k} - \frac{\sigma^2}{2} t_k} \mathrm{d}u
\end{aligned}$$

In Monte Carlo simulation, this approximation is used in a double loop (in times and in the number of simulation), so it is really necessary to simplify this formula. Hence, a formal Taylor expansion can be done with h small, which leads to a more practical scheme:

$$Y_T^{e,n} = \frac{h}{T}\sum_{k=0}^{n-1} S_{t_k}\left(1 + \frac{rh}{2} + \sigma \frac{W_{t_{k+1}} - W_{t_k}}{2}\right) \tag{2.5}$$

Remark 2.1. Note that this scheme is equivalent to the well known trapezoidal method. We will prove that,

$$\mathbf{E}\left(Y_T^{e,n} - \frac{1}{T}\sum_{k=0}^{n-1} h \frac{S_{t_k} + S_{t_{k+1}}}{2}\right)^2 = O\left(\frac{1}{n^3}\right)$$

So, since the rate of convergence of (2.5) is in $1/n$ (see proposition 3.3), the error above is negligible.

PROOF This result can be obtained using a Taylor expansion:

$$\frac{1}{T}\sum_{k=0}^{n-1} h \frac{S_{t_k} + S_{t_{k+1}}}{2} = \frac{1}{T}\sum_{k=0}^{n-1} \frac{h S_{t_k}}{2}\left(e^{\sigma\left(W_{t_{k+1}} - W_{t_k}\right) - \frac{\sigma^2}{2}h + rh} + 1\right)$$

$$= \frac{1}{T}\sum_{k=0}^{n-1} \frac{h S_{t_k}}{2}\left(2 + \sigma\left(W_{t_{k+1}} - W_{t_k}\right) + rh\right) + \ldots$$

which is exactly the scheme (2.5). The term of order $\sigma^2 h$ and the quadratic variation of $\sigma(Wt_{k+1} - Wt_k)$ cancel each other out, so we do not retain it. The mathematical proof of this remark is the same as proposition 3.3. □

Our last scheme is quite similar. As the Brownian motion is a Gaussian process, $\int_0^T W_u \mathrm{d}u$ has a normal density with respect to the Lebesgues measure on R and can easily be simulated. Hence,

$$Y_T = \frac{1}{T}\int_0^T S_u \mathrm{d}u$$

$$= \frac{1}{T}\sum_{k=0}^{n-1} S_{t_k} \int_{t_k}^{t_{k+1}} \mathrm{e}^{\sigma\left(W_u - W_{t_k}\right) - \frac{\sigma^2}{2}(u - t_k) + r(u - t_k)} \mathrm{d}u$$

So, using a Taylor expansion again, we obtain the scheme:

$$Y_T^{p,n} = \frac{1}{T}\sum_{k=0}^{n-1} S_{t_k}\left(h + \frac{rh^2}{2} + \sigma \int_{t_k}^{t_{k+1}} \left(W_u - W_{t_k}\right)\mathrm{d}u \right) \tag{2.6}$$

REMARK 2.2. In practice, to simulate this scheme, we have, at each time step, to simulate Wt_{k+1} knowing Wt_k and $\left(\int_{t_k}^{t_{k+1}} W_u \mathrm{d}u \,|\, Wt_k, Wt_{k+1}\right)$. For the second random variable, we use the law (2.3) and for the first one the fact that $(Wt_{k+1} - Wt_k, k = 0, \ldots, N-1)$ are Gaussian i.i.d. variables.

REMARK 2.3. This scheme can be generalized easily to a higher class of diffusion process. Let S_t be a diffusion with a drift $b(S_t)$ and a diffusion term $\sigma(S_t)$, and $Y_t = \int S_t \mathrm{d}t$. One can use the Euler scheme S_t^n to simulate S_t and

$$Y_T^n = \sum_{k=0}^{n-1} S_{t_k}^n \left(h + \int_{t_k}^{t_{k+1}} \left(S_s^n - S_{t_k}^n\right)\mathrm{d}s \right)$$

Note that when the diffusion S_t can be directly simulated, the following scheme is sufficient

$$Y_T^n = \sum_{k=0}^{n-1} S_{t_k} \left(h + \int_{t_k}^{t_{k+1}} \left(S_s^n - S_{t_k} \right) \mathrm{d}s \right)$$

We can show that the weak convergence holds at the rate $1/n^{3/2}$.

We will now prove some results on the speed of convergence of these three schemes. The weak convergence can be obtained for the scheme (2.1) (at the rate $1/n$) and 2.6) (at the rate $1/n^{3/2}$) under a weak assumption on f using Malliavin calculus and techniques developed by Bally and Talay (1996). We refer to Temam (2001) for a complete proof of these convergence results.

3. Convergence in the L^p spaces

In this section, we consider the strong convergence of the schemes. In L^p, one can only have some upper bound of the rate of convergence, but the precise expansion can only be obtained in the particular case $p = 2$.

Let us start with two well known but important results.

PROPOSITION 3.1. *For a Black and Scholes diffusion,*

$$\mathbf{E}|S_t - S_s|^{2q} \leq C_q |t - s|^q$$

This proposition is available for any diffusion with Lipschitz coefficient (see Revuz and Yor, 1994). The following will also be useful (see Faure, 1992, chapter 3, for a proof):

LEMMA 3.1. *Let* $Z_t = Z_0 + \int_0^T A_s \mathrm{d}W_s + \int_0^t B_s \mathrm{d}_s$ *where* B_s *is a vector in* $\mathbf{R}^n$, A_s *a matrix in* $\mathbf{R}^{n\times d}$, *and* W_t *a d-dimensional Brownian motion. (*Z_t *is an Itô process so A and B are adapted,* $\int |A_s| \mathrm{d}s < +\infty$ and $\mathbf{E}\int B_s{}^2 \mathrm{d}s < +\infty$).

Then, Z_t *satisfies*

$$\mathbf{E}|Z_t|^p \leq \mathbf{E}|Z_0|^p + C\int_0^t \mathbf{E}\left(|Z_s|^p + |A_s|^p + |B_s|^p\right)\mathrm{d}s$$

We can now obtain precise results for the rate of convergence of our schemes.

PROPOSITION 3.2. *With the above notations, there exist three non-decreasing maps* $K_1(T)$, $K_2(T)$, $K_3(T)$ *such that,*

$$\left[\mathbf{E}\left(\sup_{t\in[0,T]} \left| Y_t^{r,n} - Y_t \right|^{2q}\right)\right]^{\frac{1}{2q}} \leq \frac{K_1(T)}{n} \tag{3.1}$$

$$\left[\mathbf{E}\left(\sup_{t\in[0,T]}\left|Y_t^{e,n}-Y_t\right|^{2q}\right)\right]^{\frac{1}{2q}}\leq\frac{K_2(T)}{n}\tag{3.2}$$

$$\left[\mathbf{E}\left(\sup_{t\in[0,T]}\left|Y_t^{p,n}-Y_t\right|^{2q}\right)\right]^{\frac{1}{2q}}\leq\frac{K_3(T)}{n^{3/2}}\tag{3.3}$$

PROOF Let us begin with the inequality (3.1). If we note $\varepsilon_t = Y_t^{r,n} - Y_t$, one has

$$\begin{aligned}\forall t\in[t_k,t_{k+1}]\quad \varepsilon_t &= \varepsilon_{t_k}+\int_{t_k}^{t}\left(S_s-S_{t_k}\right)\mathrm{d}s\\ &= \varepsilon_{t_k}+\int_{t_k}^{t}(t-u)\,rS_u\mathrm{d}u+\int_{t_k}^{t}(t-u)\,\sigma S_u\mathrm{d}W_u\end{aligned}$$

So, lemma 3.1 implies that

$$\forall t\in[t_k,t_{k+1}]\quad \mathbf{E}\,|\varepsilon_t|^{2q}\leq\left(\mathbf{E}\,|\varepsilon_{t_K}|^{2q}+h^{2q+1}\right)+\int_{t_k}^{t}\mathbf{E}\,|\varepsilon_s|^{2q}\mathrm{d}s$$

Applying Gronwall's lemma, we get

$$\mathbf{E}\left|\varepsilon_{t_{k+1}}\right|^{2q}\leq\left(\mathbf{E}\left|\varepsilon_{t_K}\right|^{2q}+h^{2q+1}\right)\mathrm{e}^h$$

Since $x_{n+1} \leq ax_n + b$ implies $x_n \leq a^n x_0 + ne^{n(a-1)}b$ for $a \geq 1$, we have

$$\mathbf{E}\left|\varepsilon_{t_k}\right|^{2q}\leq Ch^{2q}$$

We conclude with the inequality of Burkholder–Davis–Gundy. We use the same kind of arguments for the two other schemes writing them as:

$$\begin{aligned}Y_t-Y_t^{e,n} &= Y_{t_k}-Y_{t_k}^{e,n}+\int_{t_k}^{t}\int_{t_k}^{s}r\left(S_u-S_{t_k}\right)\mathrm{d}u\,\mathrm{d}s+\int_{t_k}^{t}\int_{t_k}^{s}\sigma\left(S_u-\frac{h}{2}S_{t_k}\right)\mathrm{d}W_u\,\mathrm{d}s\\ &= Y_{t_k}-Y_{t_k}^{e,n}+\int_{t_k}^{t}r(t-u)\left(S_u-S_{t_k}\right)\mathrm{d}u+\int_{t_k}^{t}\sigma\left((t-u)\,S_u-\frac{h}{2}S_{t_k}\right)\mathrm{d}W_u\end{aligned}$$

and

$$\begin{aligned}Y_t-Y_t^{p,n} &= Y_{t_k}-Y_{t_k}^{p,n}+\int_{t_k}^{t}\int_{t_k}^{s}r\left(S_u-S_{t_k}\right)\mathrm{d}u\,\mathrm{d}s+\int_{t_k}^{t}\int_{t_k}^{s}\sigma\left(S_u-S_{t_k}\right)\mathrm{d}W_u\,\mathrm{d}s\\ &= Y_{t_k}-Y_{t_k}^{e,n}+\int_{t_k}^{t}r(t-u)\left(S_u-S_{t_k}\right)\mathrm{d}u+\int_{t_k}^{t}\sigma\left((t-u)\,S_u-S_{t_k}\right)\mathrm{d}W_u\end{aligned}$$

From this point, it is easy to get the inequalities (3.2) and (3.3). □

These inequalities give an upper bound for the rate of convergence, but we can obtain more precise results. Moreover, if we study the L^2 convergence, we can obtain an expansion, and this proves the exact rate of convergence.

PROPOSITION 3.3. *With the above notations and assumptions, it holds that*

$$\sqrt{\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u - \mathbf{E}\left(\frac{1}{T}\int_0^T S_u \mathrm{d}u \,\middle|\, \mathcal{B}_h\right)\right]^2} = \frac{\sigma}{n}\sqrt{\frac{\mathrm{e}^{(\sigma^2+2r)T}-1}{12\left(\sigma^2+2r\right)}} + O\left(\frac{1}{n\sqrt{n}}\right) \tag{3.4}$$

$$\sqrt{\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u - Y_T^{e,n}\right]^2} = \frac{\sigma}{n}\sqrt{\frac{\mathrm{e}^{(\sigma^2+2r)T}-1}{12\left(\sigma^2+2r\right)}} + O\left(\frac{1}{n\sqrt{n}}\right) \tag{3.5}$$

$$\sqrt{\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u - Y_T^{r,n}\right]^2} = \frac{K_2}{n} + O\left(\frac{1}{n\sqrt{n}}\right) \text{ where } K_2 \geq \sigma\sqrt{\frac{\mathrm{e}^{(\sigma^2+2r)T}-1}{12\left(\sigma^2+2r\right)}} \tag{3.6}$$

$$\sqrt{\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u - Y_T^{p,n}\right]^2} = \frac{1}{n\sqrt{n}}\sqrt{T\sigma^4\frac{\mathrm{e}^{(\sigma^2+2r)T}-1}{12\left(\sigma^2+2r\right)}} + O\left(\frac{1}{n^2}\right) \tag{3.7}$$

Note that the same rate of convergence occurs in equations (3.4) and (3.5), and futhermore the first term in the expansion is the same. It implies that we did not lose anything by approximating the conditional expectation with the scheme (2.5).

If we consider the strong convergence, it seems that the most efficient scheme is the last one. We also remark that the schemes (2.1) and (2.5) have the same rate of convergence and that the first terms in the expansions are quite similar.

The proofs of these four equalities are similar so we will only consider the last one.

PROOF OF (3.7) Let us note

$$A_{u,t_k} = \mathrm{e}^{\sigma\left(W_u - W_{t_k}\right) - \frac{\sigma^2}{2}(u-t_k) + r(u-t_k)} - 1 - r(u-t_k) - \sigma\left(W_u - W_{t_k}\right)$$

Then we get

$$
\begin{aligned}
&\mathbf{E}\left[\frac{1}{T}\int_0^T S_u \mathrm{d}u - Y_T^{p,n}\right]^2 \\
&= \frac{2}{T^2}\sum_{0\le j<i\le N-1}\int_{t_i}^{t_{i+1}}\mathrm{d}v\int_{t_j}^{t_{j+1}}\mathrm{d}u\,\mathbf{E}\left(A_{v,t_i}A_{u,t_j}S_{t_i}S_{t_j}\right) \\
&\qquad + \frac{2}{T^2}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\mathrm{d}v\int_{t_k}^{u}\mathrm{d}u\,\mathbf{E}\left(A_{u,t_k}A_{v,t_k}S_{t_k}^2\right)
\end{aligned}
$$

Since A_{v,t_i} is independent of $\mathcal{F}_{t_i}$, it follows that

$$
\begin{aligned}
\mathbf{E}\left(A_{v,t_i}A_{u,t_j}S_{t_i}S_{t_j}\right) &= \mathrm{e}^{(\sigma^2+r)t_j+rt_i}\left(\mathrm{e}^{r(v-t_i)}-1-r(v-t_i)\right)\times \\
&\qquad \left(\mathrm{e}^{(r+\sigma^2)(u-t_i)}-1-(r+\sigma^2)(u-t_i)\right)
\end{aligned}
$$

With a normalization of the variable in the integrand we get,

$$
\begin{aligned}
&\frac{2}{T^2}\sum_{0\le j<i\le N-1}\int_{t_i}^{t_{i+1}}\mathrm{d}v\int_{t_j}^{t_{j+1}}\mathrm{d}u\,\mathbf{E}\left(A_{v,t_i}A_{u,t_j}S_{t_i}S_{t_j}\right) \\
&= \frac{2}{T^2}\left(\sum_{0\le j<i\le N-1}\mathrm{e}^{(\sigma^2+r)t_j+rt_i}\right)h^2\int_0^1\mathrm{d}v\int_0^1\mathrm{d}u\left(\mathrm{e}^{rhv}-1-rhv\right) \\
&\times\left(\mathrm{e}^{(r+\sigma^2)(hu)}-1-(r+\sigma^2)hu\right)
\end{aligned}
$$

It is clear that

$$
\left(\sum_{0\le j<i\le N-1}\mathrm{e}^{(\sigma^2+r)t_j+rt_i}\right) = O\left(\frac{1}{h^2}\right)
$$

which leads to

$$
\frac{2}{T^2}\sum_{0\le j<i\le N-1}\int_{t_i}^{t_{i+1}}\mathrm{d}r\int_{t_j}^{t_{j+1}}\mathrm{d}u\,\mathbf{E}\left(A_{v,t_i}A_{u,t_j}S_{t_i}S_{t_j}\right) = O\left(h^4\right)
$$

We use the same method for the second term,

$$
\begin{aligned}
&\mathbf{E}\left(A_{u,t_k}A_{v,t_k}S_{t_k}^2\right) \\
&= \mathrm{e}^{(\sigma^2+2r)t_k}\left[-\left(1+r(u-t_k)\right)\left(\mathrm{e}^{r(v-t_k)}-1-r(v-t_k)\right)\right. \\
&\left. +\mathrm{e}^{r(u-t_k)}\left(\mathrm{e}^{(\sigma^2+r)(v-t_k)}-1-(r+\sigma^2)(v-t_k)\right)+\sigma^2(v-t_k)\left(\mathrm{e}^{r(v-t_k)}-1\right)\right]
\end{aligned}
$$

Consequently,

$$\frac{2}{T^2}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathrm{d}v \int_{t_k}^{u} \mathrm{d}u\,\mathbf{E}\left(A_{u,t_k} A_{v,t_k} S_{t_k}^2\right)$$

$$= \frac{2}{T^2}\left(\sum_{k=0}^{N-1} \mathrm{e}^{(\sigma^2+2r)t_k}\right) h^2 \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}v \left(\frac{\sigma^4 h^2 v^2}{2} + O(h^3)\right)$$

$$= h^3 \,\frac{\sigma^4 \mathrm{e}^{(\sigma^2+2r)T} - 1}{12(\sigma^2 + 2r)}$$

4. Variance reduction

To increase the efficiency of the Monte Carlo simulation a variance reduction method can be used. We follow the method developed by Kemna and Vorst (1990). It consists in approximating $\frac{1}{T}\int_0^T S_u \mathrm{d}u$ by $\exp\left(\frac{1}{T}\int_0^T \log(S_u)\mathrm{d}u\right)$. We can expect that these two random variables are similar since r and σ are not too large.

The random variable $Z' = \frac{1}{T}\int_0^T \log(S_u)\mathrm{d}u$ obviously has a normal law, and so we can compute explicitly

$$\mathbf{E}\left(\mathrm{e}^{-rT}\left(\exp(Z') - K\right)_+\right)$$

As a consequence, we can choose the random variable Z defined by

$$Z = \mathrm{e}^{-rT}\left(x\mathrm{e}^{\left(r-\frac{\sigma^2}{2}\right)\frac{T}{2}+\frac{\sigma}{T}\int_0^T W_u \mathrm{d}u} - K\right)_+$$

as our control variable.

Note that the control variable has to be computed with the path of the Brownian motion already simulated. Consequently, each control variable has to be adapted to the schemes. So, we retain the same approximation for $\int_0^T W_u \mathrm{d}u$ as for $\int_0^T S_u \mathrm{d}u$:

$$Z_T^{r,n} = \mathrm{e}^{-rT}\left(x\mathrm{e}^{\left(r-\frac{\sigma^2}{2}\right)\frac{T}{2}+\frac{\sigma}{T}\sum_{k=0}^{n-1} hW_{t_k}} - K\right)_+$$

$$Z_T^{e,n} = \mathrm{e}^{-rT}\left(x\mathrm{e}^{\left(r-\frac{\sigma^2}{2}\right)\frac{T}{2}+\frac{\sigma}{T}\sum_{k=0}^{n-1} \frac{h}{2}\left(W_{t_k} + W_{t_{k+1}}\right)} - K\right)_+$$

$$Z_T^{p,n} = \mathrm{e}^{-rT}\left(x\mathrm{e}^{\left(r-\frac{\sigma^2}{2}\right)\frac{T}{2}+\frac{\sigma}{T}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} W_u \mathrm{d}u} - K\right)_+$$

Note that, as it is explained in Remark 2.2, when we use the scheme (2.6), we simulate at each step Wt_{k+1} then the random variables ξ_k defined by $\left(\int_{t_k}^{t_{k+1}} W_u \,\mathrm{d}u \mid Wt_k,\, Wt_{k+1}\right)$. By writing $\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} W_u \,\mathrm{d}u$ in the last equation above, we intend the sum of the generated variables i.e. $\sum_{k=0}^{n-1}\xi_k$.

5. Numerical computation

In this section we compare the accuracy of various schemes to compute the price of an Asian option. Firstly, we will study the convergence of our schemes with and without the variance reduction method. This leads to the foreseeable conclusion that the variance reduction is essential to the efficiency of our schemes. In the second part, we compare the Monte Carlo methods with other methods of pricing, namely Forward Shooting Grid, Hull and White and a finite-difference method.

In the graphs of this section, the letter R represents scheme (2.1), E scheme (2.5) and BB scheme (2.6).

5.1. Convergence of the Monte Carlo methods

Simulation without variance reduction technique

We test the three schemes of subsection 2.2 with the following values: $T = 1$ year, $r = 0.1$, $X_0/K = 1$ and $\sigma = 0.2$.

The first simulation is made without any variance reduction technique and with 100,000 Monte Carlo loops, which is already high. The results presented in Table 1, show that even if the convergence seems to be fast, the confidence intervals are too large for a real financial use (penny accuracy is required).

As we will see later, the real price of this call is not far from 7.04. Table 1 shows that the first scheme is not very efficient compared to the others. The surprise is that (2.5) and (2.6) seem to have the same rate of convergence whereas proposition 3.3 ensures that the second one is more efficient. But, the justification of the numerical simulation can be based on weak convergence, and if we study that kind of convergence, the two schemes will have the same expansion (see Temam, 2001, for details).

Table 1 shows the limit of the Monte Carlo method. In fact, since our objective is to study the convergence of our schemes when the time step number goes to infinity, we want to avoid problems related to the Monte Carlo approximation. With this aim in view, we have to increase the number of paths which also increases the time of computation dramatically. Variance reduction technique is necessary.

TABLE 1 Confidence intervals versus time step number when no variance reduction technique is used

Scheme	Number of time steps	Price	Confidence intervals at 95%
	10	6.396	[6.348,6.444]
(2.1)	50	6.978	[6.926,7.030]
	100	6.930	[6.877,6.982]
	10	6.997	[6.944,7.050]
(2.5)	50	7.034	[6.981,7.087]
	100	7.037	[6.984,7.090]
	10	7.065	[7.012,7.118]
(2.6)	50	7.059	[7.006,7.112]
	100	7.055	[7.002,7.108]

Simulation with variance reduction method

Figure 1 represents the three schemes computed with the variance reduction technique of section 4 and with the same parameters as above except for the number of loops for the Monte Carlo method which is now 20,000 (80% less than before). We observe again that the schemes (2.5) and (2.6) are quite similar and that the Riemann approximation is less efficient.

Figure 2, which is a zoom of Figure 1 with only the two new schemes, seems to show that these schemes have the same rate of convergence.

Table 2 represents the dependences between the price of the option and the number of time steps.

The results of Table 2 show that the variance reduction method makes the scheme more efficient, both for the convergence and the confidence intervals: Note that

- ❑ The scheme using conditional expectation is a negatively biased approximation.
- ❑ The graph shows that the time step number which is necessary to have a good precision is less with the schemes (2.5) and (2.6) than with (2.1). As a consequence, the use of low-discrepancy sequences can be considered to increase the efficiency of our schemes.

Table 3 gives the number of time steps necessary to have the confidence interval at 95% including in intervals given by the real value of the option and the precision on the second column. To be able to get a real price, we use zero strike Asian option, and the same other parameters as above (in particular 20,000 sample paths).

The results show again how it is efficient to use one of the two new approximations for the integral. Even to get a high precision of 0.01%, the schemes (2.5) and (2.6) do not take more than 1 second against 28 seconds for the Riemann sums.

FIGURE 1 Convergence of the schemes with variance reduction technique; N is the number of time steps; T = 1 year, r = 0.1, X_0/K = 1 and σ = 0.2

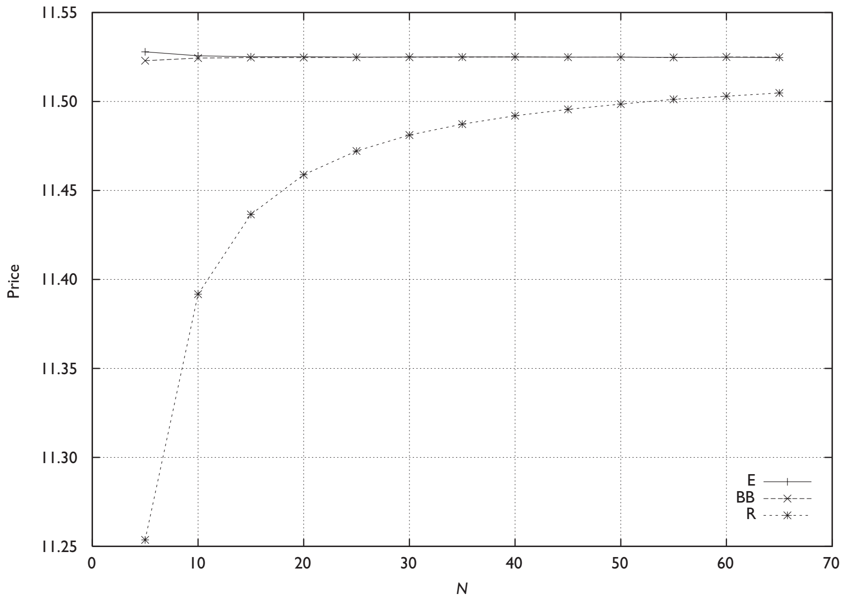


FIGURE 2 Convergence of the schemes with variance reduction technique; N is the number of time steps; T = 1 year, r = 0.1, X_0/K = 1 and σ = 0.2

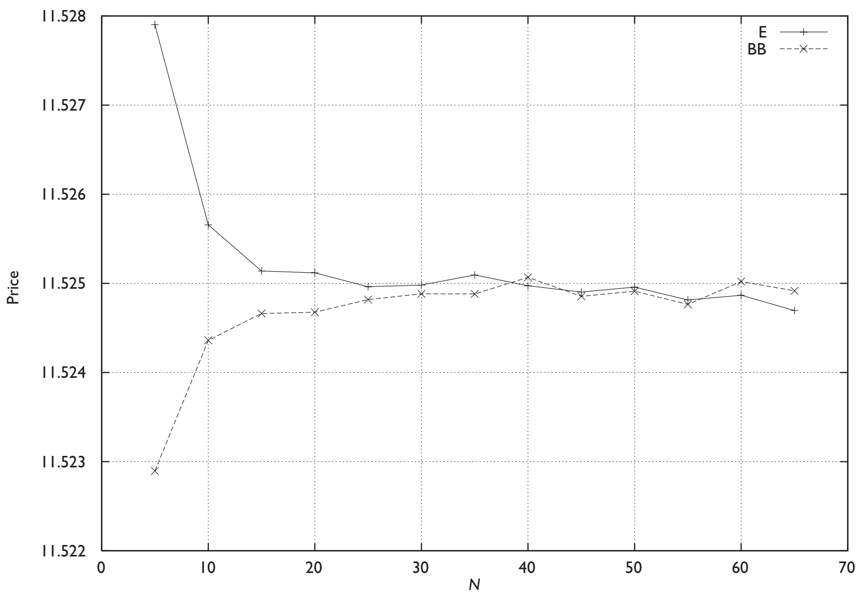


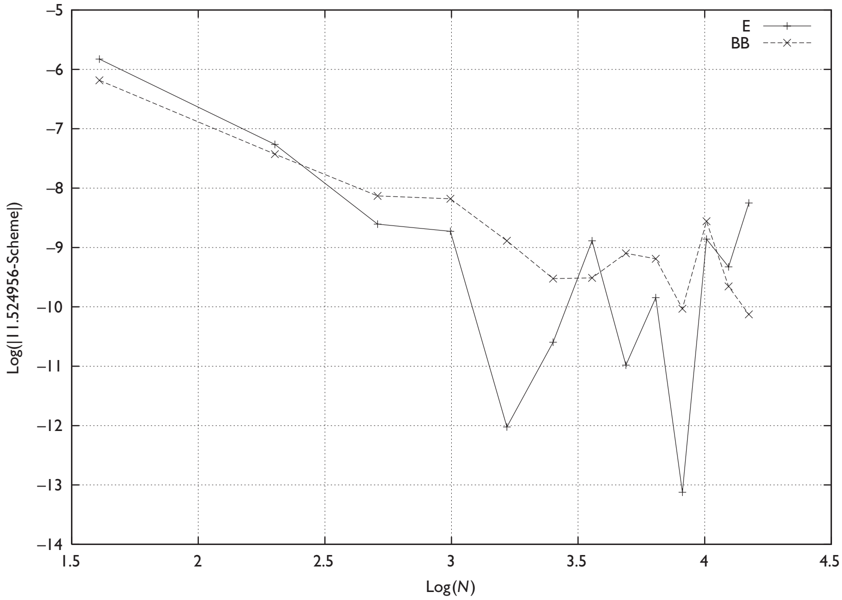
TABLE 2 Confidence intervals crossed with time step number with variance reduction technique

Scheme	Number of time steps	Price	Confidence intervals at 95%
	10	6.778	[6.776,6.781]
(2.1)	50	6.991	[6.988,6.995]
	100	7.016	[7.013,7.020]
	10	7.035	[7.031,7.038]
(2.5)	50	7.039	[7.035,7.042]
	100	7.042	[7.038,7.045]
	10	7.041	[7.038,7.044]
(2.6)	50	7.041	[7.037,7.044]
	100	7.041	[7.037,7.044]

TABLE 3 Minimal number of time steps and time of computation to obtain a given precision

Scheme	Precision (%)	Number of time steps	Time of computation
	5	2	<1s
	1	5	<1s
	0.5	9	1s
	0.1	42	2s
(2.1)	0.07	65	3s
	0.05	92	3s
	0.03	175	7s
	0.01	800	28s
	5	2	<1s
	1	2	<1s
	0.5	2	<1s
	0.1	2	<1s
(2.5)	0.07	3	<1s
	0.05	3	<1s
	0.03	4	<1s
	0.01	9	1s
	5	2	<1s
	1	2	<1s
	0.5	2	<1s
	0.1	2	<1s
(2.6)	0.07	2	<1s
	0.05	2	<1s
	0.03	3	<1s
	0.01	8	1s

FIGURE 3 Convergence of the schemes with variance reduction technique; we cross the logarithm of the precision (11.524956 is the real price) with the logarithm of the time step number *N*



5.2. A numerical study of the rate of convergence

We attempt, in this section, to obtain numerically the rate of convergence in law of the schemes described in section 2.2. To achieve this, we compute the schemes with a very high number of paths in the Monte Carlo method (10^7) in order to avoid Monte Carlo noise.

The parameters of the option computed are: $r = 0.09$, $\sigma = 0.2$, $T = 0.4$ and $K = S_{T_0} = 100$.

By running the routine FSG (see Barraquand and Pudet, 1996) which will be described in the next section with a low time step we have a very good approximation of the real price (~ 11.524956). In Figure 3, we plot the logarithm of the difference between this price and the result of the schemes (2.5) and (2.6) as a function of $\log(N)$.

The coefficient of correlation for these two sets of data is –1.46 for (2.6) and –1.33 for (2.5), which approximately confirms that the rate of weak convergence of (2.5) and (2.6) is $1/N^{3/2}$.

5.3. Comparison with other numerical methods

In this part, we have implemented some other methods commonly used for the pricing of Asian options:

- A method based on the one-dimensional partial derivative equation of which V is the solution (see Dewynne and Wilmott, 1995, Ingersoll, 1987, or Rogers and Shi, 1995). This equation is

$$\frac{\partial V}{\partial t}(x,t)+\tfrac{1}{2}\sigma^2 x^2\,\frac{\partial^2 V}{\partial x^2}(x,t)-\big(\rho(t)+rx\big)\frac{\partial V}{\partial x}(x,t)=0 \tag{5.1}$$

  with limit condition and price

$$V(x,T) \;=\; x_- \qquad \text{Price} \;=\; S_{T_0}V\left(\frac{K}{S_{T_0}},\,0\right)$$

  for a fixed strike and

$$V(x,T) \;=\; (1+x)_- \qquad \text{Price} \;=\; S_{T_0}V(0,\,0)$$

  for a floating strike. This equation was approximated by a finite-difference method with viscosity arguments to increase the efficiency of the method.
- Two methods (often called "tree" methods) based on approximation of the transition kernel of S with the Cox, Ross and Rubinstein discrete model. These methods have been developed by Barraquand and Pudet (FSG of Forward Shooting Grid, Barraquand and Pudet, 1996) and Hull and White (Hull and White, 1993). We choose for the time step number $N = 12$ (one month), 24, 52 (one week), 100.
- The three schemes for Monte Carlo methods are described here, with and without variance reduction techniques. The number of paths was 20,000 and the time step numbers $N = 2$, 6, 12, 24, 52.

To compare the various schemes, we use the methodology proposed by Broadie and Detemple (1997). It consists in simulating 200 options, whose parameters are randomly chosen as follows:

- The volatility σ follows a uniform law between 0.1 and 0.6.
- The maturity T is, with probability 0.75, uniform in $\{k/12,\ k = 1,\ldots,12\}$ and, with probability 0.25 uniform in $\{0, 1, 2, 3, 4, 5\}$.
- The interest rate r is, with probability 0.8, uniform in [0.0, 0.1], and, with probability 0.2 equal to 0.
- S_0, the initial value of the model (the stock), is uniform between 70 and 130.

The computation was made on a dual-processor Intel Pentium II PC at 450 Mhz. To be able to evaluate the precision, we chose $K = 0$, which is the only case where there exists a closed formula. The results of this test are given in Figure 4, which represents the logarithm of the mean error

$$E = \sqrt{\frac{1}{200}\sum_{i=1}^{200}\left(\frac{\text{True price} \;-\; \text{Simulated price}}{\text{True price}}\right)^2}$$

versus the logarithm of the CPU time of computation.

FIGURE 4 Results of the advanced test; FSG and Hull stand for the tree methods and MC denotes Monte Carlo methods; KV indicates that the variance reduction method was used

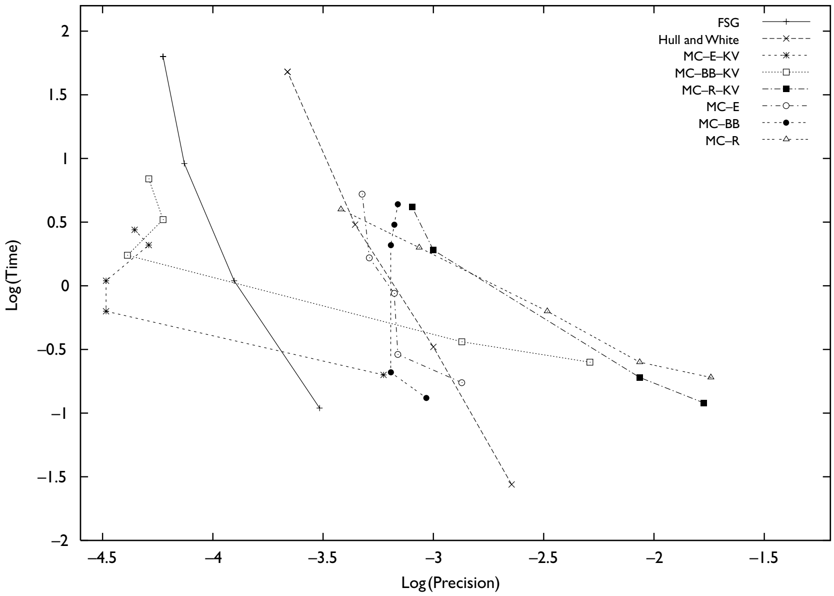


Figure 4 can be read as follows:

- ❑ MC represents Monte Carlo methods.
- ❑ As above, R is the abbreviation for the scheme (2.1), E for (2.5) and BB for (2.6).
- ❑ KV indicates when the variance reduction method described above was used.
- ❑ FSG is the "Forward Shooting Grid Method".
- ❑ FD represents the Finite Difference method implemented here.

The results concerning the analytic methods (developed by Geman and Yor, 1993, and Fu *et al.*, 1996) are unconfirmed. These methods are difficult to implement because of the repetition of numerical inversions of Laplace transform. For the moment their results are disappointing. Using standard algorithms to solve this problem leads to disappointing results. Nevertheless, we are not absolutely sure that we have used the best available method.

A standard finite difference on the one-dimensional equation (5.1) fails if the volatility is low. Since the variance is a decreasing map of the volatility σ, Monte Carlo methods can be very useful in this case.

We still have to consider the Monte Carlo methods and the tree methods. For the first class, variance reduction technique and precise schemes (2.5 and 2.6) represent the only way to obtain correct results.

To obtain moderate accuracy ($\sim 10^{-3}$), the algorithm of Hull and White seems to be the most efficient. With a low step number (say 6 per year), it gives approximations with a precision at 10^{-3} in a time lower than 0.03 seconds whereas for the same precision, the forward shooting grid takes 0.1 seconds and Monte Carlo methods 0.3 seconds (ten times more!).

By contrast, if higher precision is required, the Monte Carlo methods become more efficient. In fact, the time complexity of the tree methods is higher (we can see that on the graph, by comparing the slope of the trees methods and Monte Carlo methods). For a precision at 10^{-4}, the scheme MC–E–KV runs in 0.4 seconds whereas the algorithm FSG takes 1.3 seconds and Hull and White does not reach this precision for a time step number lower than 100.

The ends of the lines representing the schemes MC–BB–KV and MC–E–KV show that with a time step number above 24, the error due to Monte Carlo becomes greater than the error due to the time step on average.

In conclusion, if we use both the new schemes and the variance reduction technique, Monte Carlo methods can be competitive with the trees methods for pricing Asian options and represent the only way to get high accuracy.

REFERENCES

Bally, V., and Talay, D. (1996). The law of the Euler scheme for stochastic differential equations. I: Convergence rate of the distribution function. *Probability Theory and Related Fields* **104**(1), 43–60.

Barraquand, J., and Pudet, T. (1996). Pricing of American path-dependent contingent claims. *Mathematical Finance*, **6**, 17–51.

Broadie, M., and Detemple, J. (1997). Recent advances in numerical methods for pricing derivative securities. In: Rogers, L. C. G., *et al.* (ed.), *Numerical Methods in Finance.* Cambridge University Press, pp. 43–66.

Dewynne, J., and Wilmott, P. (1995). Asian options as linear complementary problems: analysis and finite difference solutions. *Advances in Futures and Operations Research*, **8**, 145–73.

Dufresne, D. (1999). Laguerre series for Asian and other options. Forthcoming in *Mathematical Finance*.

Faure, O. (1992). Simulation du mouvement brownien et des diffusions. Ph.D. thesis, Ecole Nationale des Ponts et Chaussées.

Forsyth, P. A., Vetzal, K., and Zvan, R. (1996). Robust numerical methods for PDE models of Asian options. Technical Report, University of Waterloo, Canada.

Forsyth, P. A., Vetzal, K., and Zvan, R. (1999). Convergence of lattice and PDE methods for pricing Asian options. Technical Report, University of Waterloo, Canada.

Fu, M., Madan, D., and Wang, T. (1996). Pricing continous time Asian options: A comparison of analytical and Monte Carlo methods. Forthcoming in *Journal of Computational Finance*.

Geman, H., and Yor, M. (1993). Bessel processes, Asian options, and perpetuities. *Mathematical Finance*, **3**(4), 349–75.

Hull, J., and White, A. (1993). Efficient procedures for valuing European and American path-dependent options. *Journal of Derivatives*, **1**, 21–31.

Ingersoll, J., Jr. (1987). *Theory of Financial Decision Making*. Roman & Littlefield, Totowa, New Jersey.

Kemna, A. G. Z., and Vorst, A. C. F. (1990). A pricing method for options based on average asset values. *Journal of Banking Finance*, March.

Klassen, T. (2000). Simple, fast, and flexible pricing of Asian options. Forthcoming in *Journal of Computational Finance*.

Revuz, D., and Yor, M. (1994). *Continous Martingales and Brownian Motion*. Springer-Verlag, New York.

Rogers, L. C. G., and Shi, Z. (1995). The value of an Asian option. *Journal of Applied Probability*, **32**(4), 1077–88.

Temam, E. (2001). Couverture approchée des options exotiques. Pricing des options asiatiques. Ph.D. Thesis, University of Paris VI.