

Monte Carlo evaluation of Greeks for multidimensional barrier and lookback options.

Guillaume BERNIS*, Emmanuel GOBET†, Arturo KOHATSU-HIGA‡

This version: May 14, 2002

Abstract

In this paper, we consider the problem of the numerical computation of Greeks for a multidimensional barrier and lookback style option: the payoff function depends in a rather general way on the minima and maxima of the coordinates of the d -dimensional underlying asset process. Using Malliavin calculus techniques, we derive additional weights which enable one to compute the Greeks using Monte Carlo simulations. Numerical experiments confirm the efficiency of the method. This work is a multidimensional extension of previous results (see Gobet et al. (2001)).

Forthcoming in Mathematical Finance

JEL Classification : G13

Mathematics Subject Classification (1991): 60H07, 60J60, 65C05

Keywords: barrier and lookback options, option sensitivities, Malliavin calculus.

Introduction

In a frictionless market, we consider a d -dimensional risky asset $(S_t := (S_t^1, \dots, S_t^d))_{t \geq 0}$, whose dynamics, under the risk neutral probability \mathbf{P} , is log-normal with constant volatility $\sigma = (\sigma_{i,j})_{1 \leq i,j \leq d}$:

$$\frac{dS_t^i}{S_t^i} = rdt + \sum_{j=1}^d \sigma_{i,j} dW_t^j. \quad (1)$$

The process $(W_t = (W_t^1, \dots, W_t^d))_{t \geq 0}$ is a standard Brownian motion in \mathbf{R}^d and r is the interest rate. We focus our attention on barrier and lookback European style options, the payoff of which is a

*CAI - Fixed Income, Banque CPR - 30, rue Saint-Georges - 75312 Paris Cedex 09 - FRANCE, e-mail : gbernis@cpr.fr

†Ecole Polytechnique - Centre de Mathématiques Appliquées - 91128 Palaiseau Cedex - FRANCE, e-mail : emmanuel.gobet@polytechnique.fr

‡Universitat Pompeu Fabra. Department of Economics and Business. Ramón Trias Fargas 25-27 08005- Barcelona - SPAIN, email : arturo.kohatsu@econ.upf.es

general function of the extrema of each S^i and of their values S_T^i at maturity T . The payoff function Φ is hence of the form

$$\Phi \left(\max_{s \in I} S_s^1, \dots, \max_{s \in I} S_s^d, \min_{s \in I} S_s^1, \dots, \min_{s \in I} S_s^d, S_T^1, \dots, S_T^d \right). \quad (2)$$

The set $I \subset [0, T]$ is a set of times when the extrema are monitored: $I = [0, T]$ corresponds to the continuous time case, and $I = \{0 = t_0 < t_1 < \dots < t_i < \dots < t_N = T\}$ to the discrete time one. This setting includes obviously usual barrier and lookback options, but also other exotic options. Let us give two illustrative examples:

- *Best-Off* (or *Worst-Off*) option: the payoff function may be of the form

$$\max \left(\max_{s \in I} S_s^1, \dots, \max_{s \in I} S_s^d, K \right).$$

- *BLAC* (Basket Lock Active Coupon) *Down & Out* option: this option pays 1 Euro if at most one of the underlying assets has touched a lower barrier L before the expiration. Hence, the payoff is

$$\prod_{i \neq j} \mathbf{1}_{\min_{s \in I} S_s^i \vee \min_{s \in I} S_s^j > L}.$$

In this Black & Scholes model, the price at time 0 of the option, whose payoff is given by (2), is equal to

$$P(S_0^1, \dots, S_0^d) = \mathbf{E}(e^{-rT} \Phi),$$

and can be computed by Monte Carlo methods. For the discrete time case this calculation is straightforward, but for the continuous time case this may sometimes be difficult, because one may need to simulate local multidimensional minima and maxima (for multidimensional barrier options, such kinds of problems have been handled by one of the authors in Gobet (2001)). Rather than considering the pricing issue, here we focus on the hedging problem, that is, to numerically compute the so-called Greeks, in particular the Delta $\Delta = (\partial_{S_0^i} P(S_0^1, \dots, S_0^d))_{1 \leq i \leq d}$ and the Gamma $\Gamma = (\partial_{S_0^i, S_0^j}^2 P(S_0^1, \dots, S_0^d))_{1 \leq i, j \leq d}$. The idea of this paper is to rewrite these sensitivities as expectations, so that the same simulations used to compute the price can also be used to evaluate the Greeks. For this we will derive an integration by parts formula of Malliavin calculus, which will give that $\Delta = \mathbf{E}(e^{-rT} \Phi H_\Delta)$. This formula will provide an alternative method of unbiased simulation if one simulates also the random weight H_Δ . A similar formula will be also obtained for the Gamma.

Recently the interest of Malliavin calculus for Monte Carlo computations has increased because it has led to new efficient algorithms to evaluate the Greeks for example; we refer to the articles by Fournié et al. (1999), Fournié et al. (2001), Benhamou (2000) and Gobet et al. (2001). The evaluation of sensitivities using an integration by parts formula is not new and has been already considered in other problems; it has been known as the *likelihood ratio* method or *score* method, and this essentially requires to know the joint density of the random variables involved in the problem. For general sensitivities, see Glynn (1986,1987) and Reiman et al. (1986); for an application to the Greeks, see

Broadie et al. (1996); for an application to stochastic control problems, see Kushner et al. (1991). The real advantage of Malliavin calculus is to generalize the previous approach even if the density is unknown, which is the most usual situation. In the framework of this paper, we have to face this nasty situation since the law of $(\max_{s \in I} S_s^1, \dots, \max_{s \in I} S_s^d, \min_{s \in I} S_s^1, \dots, \min_{s \in I} S_s^d, S_T^1, \dots, S_T^d)$ is not explicitly known, especially because of the possible correlation between the assets.

Compared to the Finite Difference method (FD in short) (see Glasserman et al. (1992), L'Ecuyer et al. (1994)), this approach using the integration by parts formula of Malliavin calculus in order to obtain additional random weights has several advantages: for example, the number of parameters to be chosen is smaller and the sensitivity estimation is unbiased. The remaining key point is to know how large is the variance of the simulated random variables for both methods. Numerical evidence (see Fournié et al. (1999)) tends to prove that if the payoff function is irregular, the FD method leads to high variance and the integration by parts approach is more accurate, while for smooth payoff functions the converse is true. Hence, in our framework of barrier style options where indicator functions may be involved in the payoff, finding additional weights H is promising for an efficient computation of the Greeks, and some numerical illustrations at the end of the paper will illustrate this fact.

The case of vanilla and Asian options has been previously handled in Fournié et al. (2001) and in Benhamou (2000). We consider here the context of barrier and lookback style options on multidimensional assets and we present new results, extending those obtained in the case of one risky asset by the two last authors in Gobet et al. (2001). Actually, in the cited paper, the dynamic of the asset is supposed to be a generalized Black-Scholes model with non-constant volatility; our contribution is to consider the multidimensional case, with a restriction to the usual Black-Scholes setting (1). The way to derive the appropriate weights involves Malliavin calculus techniques, which are not so easy in our context because the maxima and minima of stochastic processes are usually not smooth, even if the underlying processes are smooth (see the paper by Bermin (2000), where such questions are discussed through the application of Clark-Ocone's formula). The derivation of weights will require one to localize the usual integration by parts formula as is done in Nualart (1995), Section 2.1.4.

Before going further, we should mention that for the discrete time case $I = \{0 = t_0 < t_1 < \dots < t_i < \dots < t_N = T\}$ the payoff Φ depends only on a finite number of dates t_i ($\Phi = f(S_{t_1}, \dots, S_{t_N})$), and formulae from Fournié et al. (1999) can be used. For the delta Δ with only one risky asset, for example, one finds out that $H = W_{t_1} / (S_0^1 \sigma_{1,1} t_1)$. This weight is really easy to simulate, but it has the undesirable drawback of an exploding variance when the frequency of monitoring increases: $\text{Var}(H) = C/t_1 \rightarrow \infty$ when $t_1 \rightarrow 0$. Some numerical experiments below illustrate that the use of this weight for daily monitored option is really not appropriate, even compared to the FD method, while the weights obtained in Theorem 2.1 behave significantly better than the previous alternatives.

The paper is organized as follows: in Section 1 we give the notations and assumptions used throughout the paper. We also recall some Malliavin calculus results which are useful in our setting. In Section 2 we introduce a dominating process, crucial for our Malliavin calculus computations; then we state the main result concerning the derivation of weights for Δ and Γ . Our results are valid if the support

assumption **(S)** is fulfilled. This condition states that the support of the payoff function (modulo constants) should be included in the interior of the support of the density. This hypothesis may seem undesirable at first but we also show that it is satisfied in the examples we consider. We also study the influence of assumption **(S)** in the numerical results; in particular, we show that in an asymptotic case our results coincide with those of Fournié et al. (1999). In Section 3, we show numerical experiments on the discrete time case, that is, cases where path-dependent payoffs are discretely monitored. In this situation our numerical procedure behaves better than other proposed methods. Proofs are in general postponed to the Appendices.

1 Preliminaries

1.1 Notations and assumptions

We consider the d -dimensional asset process defined in (1), for which one has:

$$\log(S_t^i) = \log(S_0^i) + \mu^i t + \sigma^i \cdot W_t \quad (3)$$

with $S_0^i > 0$, where we set $\mu^i = (r - \frac{1}{2}\|\sigma^i\|^2)$ and $\sigma^i = (\sigma_{i,1}, \dots, \sigma_{i,d})^*$ is the transposition of the i -th row of the matrix $\sigma = (\sigma_{i,j})_{1 \leq i, j \leq d}$. Put $\mu = (\mu^1, \dots, \mu^d)^*$. In the sequel, we also assume that the matrix σ is invertible.

Let us denote by $M^i = \max_{s \in I} S_s^i$ and $m^i = \min_{s \in I} S_s^i$ the maximum and minimum of the i -th asset. We consider in this work a payoff of the form $\Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d)$, which is assumed to be square-integrable; sometimes we will denote it simply by Φ . It may be natural to think that no other properties of the payoff function should be required; nevertheless it appears that an integration by parts formula for barrier and lookback style options cannot be obtained in full generality without an additional support type condition on the payoff function. Before stating the appropriate assumption on Φ , we illustrate this fact with a simple probabilistic example.

Example 1.1. Let $X^x = (X_t^x = x + W_t)_{t \geq 0}$ be a linear Brownian motion starting at $x \in \mathbf{R}$. Here, the quantity of interest is $\mathbf{E}(f(\inf_{t \in [0,1]} X_t^x))$, and suppose we are interested in rewriting $\partial_x \mathbf{E}(f(\inf_{t \in [0,1]} X_t^x)) = \mathbf{E}(f(\inf_{t \in [0,1]} X_t^x)H)$ for some square integrable random variable H and for some bounded continuous function f . Using the explicit law of $\inf_{t \in [0,1]} X_t^x$, one has:

$$\begin{aligned} \partial_x \mathbf{E}(f(\inf_{t \in [0,1]} X_t^x)) &= \partial_x \int_{-\infty}^x f(y) \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{(y-x)^2}{2}\right) dy \\ &= f(x) \frac{2}{\sqrt{2\pi}} + \int_{-\infty}^x f(y) \frac{2}{\sqrt{2\pi}} (y-x) \exp\left(-\frac{(y-x)^2}{2}\right) dy \\ &= f(x) \frac{2}{\sqrt{2\pi}} + \mathbf{E}\left(f\left(\inf_{t \in [0,1]} X_t^x\right) \left(\inf_{t \in [0,1]} X_t^x - x\right)\right). \end{aligned}$$

Hence this basic computation shows that finding a weight H if $f(x) \neq 0$ turns out to be hard.

The preceding example illustrates the necessity of a support type condition of Φ . The following assumption **(S)** is not exactly analogous to $f(x) = 0$ in the previous example, this slight modification coming from the fact that the weights will be centered random variables.

We assume, in the sequel, that the condition¹ below is fulfilled.

(S) There exists $a > 0$ such that for any $i \in \{1, \dots, d\}$, the function $\Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d)$ does not depend on M^i (resp. m_i) if $M^i < S_0^i \exp(a)$ (resp. $m^i > S_0^i \exp(-a)$).

In fact, assumption **(S)** is not so restrictive; it is fulfilled for usual options. For instance, it holds for the two examples from the introduction:

- *Best-Off* option with $\Phi = \max(M^1, \dots, M^d, K)$ for $\max_{1 \leq i \leq d} S_0^i < K$. Assumption **(S)** is clearly satisfied with $a = \min_{1 \leq i \leq d} \log(K/S_0^i) > 0$.
- *BLAC Down & Out* option with $\Phi = \prod_{i \neq j} \mathbf{1}_{m^i \vee m^j > L}$ for $\min_{1 \leq i \leq d} S_0^i > L$. Write $\Phi = \prod_{1 \leq i \leq d} \mathbf{1}_{m^i > L} + \sum_{i=1}^d \mathbf{1}_{m^i \leq L} \prod_{j \neq i} \mathbf{1}_{m^j > L}$; it is now easy to see that assumption **(S)** is fulfilled taking $a = \min_{1 \leq i \leq d} \log(S_0^i/L) > 0$.

1.2 Some basic results from Malliavin Calculus

In this section, we introduce the necessary material for our Malliavin calculus computations. We follow standard definitions and notations from Nualart (1995).

For $h(\cdot) \in H = L_2([0, T], \mathbf{R}^d)$, denote by $W(h)$ the Wiener stochastic integral $\int_0^T h(t) \cdot dW_t$. Let \mathcal{S} denote the class of random variables of the form $F = f(W(h_1), \dots, W(h_N))$ where $f \in C_p^\infty(\mathbf{R}^N)$, $(h_1, \dots, h_N) \in H^N$ and $N \geq 1$. For $F \in \mathcal{S}$, we define its derivative $\mathcal{D}F = (\mathcal{D}_t F)_{t \in [0, T]}$ as the H -valued random variable given by $\mathcal{D}_t F = \sum_{i=1}^N \partial_{x_i} f(W(h_1), \dots, W(h_N)) h_i(t)$. The operator \mathcal{D} is closable as an operator from $L_p(\Omega)$ to $L_p(\Omega, H)$, for any $p \geq 1$. Its domain is denoted by $\mathbf{D}^{1,p}$ with respect to the norm $\|F\|_{1,p} = [\mathbf{E}|F|^p + \mathbf{E}(\|\mathcal{D}F\|_H^p)]^{1/p}$. We can define the iteration of the operator \mathcal{D} in such a way that for a smooth random variable F , the derivative $\mathcal{D}^k F$ is a random variable with values in $H^{\otimes k}$. As in the case $k = 1$, the operator \mathcal{D}^k is closable from $S \subset L_p(\Omega)$ into $L_p(\Omega; H^{\otimes k})$, $p \geq 1$. If we define the norm $\|F\|_{k,p} = [\mathbf{E}|F|^p + \sum_{j=1}^k \mathbf{E}(\|\mathcal{D}^j F\|_{H^{\otimes j}}^p)]^{1/p}$, we denote its domain by $\mathbf{D}^{k,p}$ and we put $\mathbf{D}^{k,\infty} = \bigcap_{p \geq 1} \mathbf{D}^{k,p}$. We also introduce δ , the Skorohod integral, defined as the adjoint operator of \mathcal{D} : this is a linear operator on $L_2([0, T] \times \Omega, \mathbf{R}^d)$ with values in $L_2(\Omega)$ and we denote by $\text{Dom}(\delta)$ its domain.

The following proposition puts together basic properties of these operators:

Proposition 1.1.

1. **Chain rule property.** Fix $p \geq 1$. For $f \in C_b^1(\mathbf{R}^q, \mathbf{R})$ and $F = (F_1, \dots, F_q)$ a random vector whose components belong to $\mathbf{D}^{1,p}$, $f(F) \in \mathbf{D}^{1,p}$ and for $t \in [0, T]$, one has

$$\mathcal{D}_t(f(F)) = \sum_{i=1}^q \partial_{x_i} f(F) \mathcal{D}_t F_i.$$

¹In Gobet et al. (2001), a stronger condition has been stated but the next computations will show that in fact, assumption **(S)** stated here is sufficient.

2. **Integration by parts formula.** If u belongs to $\text{Dom}(\delta)$, then $\delta(u) = \int_0^T u_t \delta W_t$ is the element of $L_2(\Omega)$ characterized by the integration by parts formula

$$\forall F \in \mathbf{D}^{1,2}, \quad \mathbf{E}(F \delta(u)) = \mathbf{E} \left(\int_0^T \mathcal{D}_t F \cdot u_t dt \right). \quad (4)$$

3. **Some elements in $\text{Dom}(\delta)$.** If u is an adapted process belonging to $L_2([0, T] \times \Omega, \mathbf{R}^d)$, then the Skorohod integral and the Ito integral coincide. The space of weakly differentiable H -valued variables $\mathbf{D}^{1,2}(H)$ is a subset of $\text{Dom}(\delta)$.

4. **Skorohod integral of a process multiplied by a random variable.** If $F \in \mathbf{D}^{1,2}$ and $u \in \text{Dom}(\delta)$ such that $\mathbf{E}(F^2 \int_0^T u_t^2 dt) < +\infty$, one has

$$\delta(F u) = F \delta(u) - \int_0^T \mathcal{D}_t F \cdot u_t dt.$$

In particular, if u is moreover adapted, one simply has:

$$\delta(F u) = F \int_0^T u_t dW_t - \int_0^T \mathcal{D}_t F \cdot u_t dt. \quad (5)$$

Actually, the equality (5), which rewrites the Skorohod integral into Ito and Lebesgue integrals, will be used to simulate the weights in the alternative expressions for the Greeks in Theorem 2.1.

The core of the next computations will concern the differentiability of the maximum/minimum of $(S_t^i)_{t \in I}$ or equivalently of $(\log(S_t^i))_{t \in I}$. For this, we denote by τ_M^i and τ_m^i the random times² in I where the maximum M^i and minimum m^i are attained:

$$M^i = S_{\tau_M^i}^i \quad \text{and} \quad m^i = S_{\tau_m^i}^i.$$

A straightforward application of the more general result from Nualart et al. (1988) yields:

$$\mathcal{D}_t \left(\max_{t \in I} [\mu^i t + \sigma^i \cdot W_t] \right) = \mathbf{1}_{t \leq \tau_M^i} \sigma^i \quad \text{and} \quad \mathcal{D}_t \left(\min_{t \in I} [\mu^i t + \sigma^i \cdot W_t] \right) = \mathbf{1}_{t \leq \tau_m^i} \sigma^i.$$

Combining the chain rule property with $M^i = S_0^i \exp(\max_{t \in I} [\mu^i t + \sigma^i \cdot W_t])$ and $m^i = S_0^i \exp(\min_{t \in I} [\mu^i t + \sigma^i \cdot W_t])$, we complete the proof of the following lemma.

Lemma 1.1. For any $i \in \{1, \dots, d\}$, the random variables M^i, m^i, S_T^i belong to $\mathbf{D}^{1,\infty}$ and their first weak derivatives are given for $t \leq T$ by

$$\mathcal{D}_t M^i = M^i \mathbf{1}_{t \leq \tau_M^i} \sigma^i, \quad \mathcal{D}_t m^i = m^i \mathbf{1}_{t \leq \tau_m^i} \sigma^i \quad \text{and} \quad \mathcal{D}_t S_T^i = S_T^i \sigma^i.$$

Since M^i and m^i do not belong to $\mathbf{D}^{2,p}$, the classical integration by parts formula can be performed and one needs to use a localization procedure as in Nualart (1995), Proposition 2.1.5.

²Owing to the non-degeneracy of σ , it is known that these random times are uniquely defined with probability 1.

2 Main results: computations of Δ and Γ

2.1 Dominating processes

We now introduce some notation to formalize the localization technique; this approach has already been developed in the one-dimensional setting (see Gobet et al. (2001)) and the multidimensional extension requires only minor modifications. Hence, we only state the results, omitting their proofs. First, consider the parameter $a > 0$ from assumption **(S)** and choose a C_b^∞ function Ψ such that $\mathbf{1}_{[-\infty, a/2]}(x) \leq \Psi(x) \leq \mathbf{1}_{[-\infty, a]}(x)$; without loss of generality we assume that $\sup_{x \in \mathbf{R}} |\Psi^{(k)}(x)| \leq C_k/a^k$ for some constants $C_k > 0$. Second, we set a definition for a *dominating process*:

Definition 2.1. *The increasing adapted right-continuous process $(Y_t)_{0 \leq t \leq T}$ is a **dominating process** if*

- *it satisfies $|\mu^i t + \sigma^i \cdot W_t| \leq Y_t$ for any $t \in I$ and $i \in \{1, \dots, d\}$.*
- *there exists a positive function $\alpha : \mathbf{N}^+ \mapsto \mathbf{R}^+$, with $\lim_{q \rightarrow \infty} \alpha(q) = \infty$, such that, for any $q \geq 1$, one has: $\forall t \in [0, T] \quad \mathbf{E}(Y_t^q) \leq C_q t^{\alpha(q)}$. Note that $Y_0 = 0$.*

It is natural that $(Y_t)_{0 \leq t \leq T}$ may not depend on the initial values $(S_0^i)_{1 \leq i \leq d}$, and that is what we assume in the following. Furthermore, the process Y will be required to be differentiable enough for Malliavin calculus computations; this is the following assumption stated for some $q \in \mathbf{N}^*$.

(R(q)) The random variable $\Psi(Y_t)$ belongs to $\mathbf{D}^{q, \infty}$ for each t . Moreover, for $j = 1, \dots, q$, one has

$$\forall p \geq 1 \quad \sup_{r_1, \dots, r_j \in [0, T]} \mathbf{E} \left(\sup_{r_1 \vee \dots \vee r_j \leq t \leq T} \|\mathcal{D}_{r_1, \dots, r_j} \Psi(Y_t)\|^p \right) \leq C_p.$$

We now give some examples of dominating processes, depending on the set I .

Proposition 2.1. *Consider the **discrete time case**: $I = \{0 = t_0 < \dots < t_i < \dots < t_N = T\}$.*

1. *The **extreme process***

$$Y_t = \max_{1 \leq i \leq d} \left(\max_{0 \leq j \leq N: t_j \leq t} [\mu^i t_j + \sigma^i \cdot W_{t_j}] - \min_{0 \leq j \leq N: t_j \leq t} [\mu^i t_j + \sigma^i \cdot W_{t_j}] \right) \quad (6)$$

*is a dominating process satisfying **(R(1))**.*

2. *The **averaged quadratic increments process***

$$Y_t = \sqrt{N \sum_{1 \leq j \leq N: t_j \leq t} \|\mu(t_j - t_{j-1}) + \sigma(W_{t_j} - W_{t_{j-1}})\|^2} \quad (7)$$

*is a dominating process satisfying **(R(q))** for any $q \geq 1$.*

Proposition 2.2. *Consider the **continuous time case**: $I = [0, T]$.*

1. *The extreme process*

$$Y_t = \max_{1 \leq i \leq d} \left(\max_{0 \leq s \leq t} [\mu^i s + \sigma^i \cdot W_s] - \min_{0 \leq s \leq t} [\mu^i s + \sigma^i \cdot W_s] \right) \quad (8)$$

is a dominating process satisfying **(R(1))**.

2. *The averaged modulus continuity process*

$$Y_t = 8 \left(4 \int_0^t \int_0^t \frac{\|\mu(s-u) + \sigma(W_s - W_u)\|^\gamma}{|s-u|^{m+2}} ds du \right)^{1/\gamma} \frac{m+2}{m} t^{m/\gamma}, \quad (9)$$

for an even integer γ and for $0 < m < \frac{\gamma}{2} - 2$, is a dominating process satisfying **(R(q))** under the condition $1 \leq q \leq \gamma - 2(m+2)$.

For the proofs of these results, we refer to Gobet et al. (2001): the arguments are quite basic, except for the last case (*the averaged modulus continuity process*) for which we need Garsia, Rodemich and Rumsey's Lemma (see Garsia et al. (1970)). A dominating process for the case $I = [0, T]$ obviously works for any subset $I \subset [0, T]$, but in general, it will not lead to more efficient numerical results. Hence, in the discrete time case, it is recommended to use dominating processes from Proposition 2.1.

2.2 Statement of the results

Put $e^i = (0 \cdots 0 \ 1 \ 0 \cdots 0)^*$, where 1 is the i -th coordinate. Our main result is the following theorem.

Theorem 2.1. *Suppose (S) is fulfilled.*

1) **Delta.** *If Y is a dominating process satisfying **(R(1))**, then one has*

$$\begin{aligned} 1 \leq i \leq d: \quad \Delta^i &:= \partial_{S_0^i} \mathbf{E} \left(e^{-rT} \Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d) \right) \\ &= \mathbf{E} \left(e^{-rT} \Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d) H_{\Delta^i} \right), \end{aligned} \quad (10)$$

with $H_{\Delta^i} = \frac{1}{S_0^i} \delta \left(\frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^i \right)$.

2) **Gamma.** *If Y is a dominating process satisfying **(R(2))**, then one has*

$$\begin{aligned} 1 \leq i, j \leq d: \quad \Gamma^{i,j} &:= \partial_{S_0^i, S_0^j}^2 \mathbf{E} \left(e^{-rT} \Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d) \right) \\ &= \mathbf{E} \left(e^{-rT} \Phi(M^1, \dots, M^d, m^1, \dots, m^d, S_T^1, \dots, S_T^d) H_{\Gamma^{i,j}} \right), \end{aligned} \quad (11)$$

with $H_{\Gamma^{i,j}} = \frac{1}{S_0^i S_0^j} \delta \left(\delta \left(\frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^i \right) \frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^j \right) - \frac{1}{(S_0^i)^2} \delta \left(\frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^i \right) \mathbf{1}_{i=j}$.

Proof. See Appendix A. □

2.3 About the hypothesis (S)

In this paragraph, we discuss the sensitivity of our weights to the parameter a from condition (S).

We first consider the case when \mathbf{a} **shrinks to 0**. This corresponds to asymptotically removing the assumption (S), or in other words to consider payoffs which depend on extrema at time T close to their initial values. To discuss this case, remark first that using standard computations combined with estimate (16), we can easily obtain that $\|H_{\Delta^i}\|_{L_p} + \|H_{\Gamma^{i,j}}\|_{L_p} \leq C_p/(1 \wedge a^{\beta(p)})$ for any $p \geq 1$ and for some positive function β , which depends on the dominating process. Hence, in general, we cannot ensure that our weights keep a bounded variance when a goes to 0. Actually, this even seems to be false in the continuous time case $I = [0, T]$, in the light of Example 1.1 where an appropriate weight would involve a Dirac measure at the initial value x . Nevertheless, for the discrete time case, in the limit case $a = 0$ our weights coincide with those of Fournié et al. (1999). This is the statement of the following proposition.

Proposition 2.3. *Consider the discrete time case: $I = \{0 = t_0 < \dots < t_i < \dots < t_N = T\}$.*

Take for the dominating process Y either the extreme process (6) or the averaged quadratic increments process (7). Then one has

$$\lim_{a \downarrow 0} H_{\Delta^i}(a) \stackrel{a.s.}{=} \frac{(\sigma^{-1}e^i) \cdot W_{t_1}}{S_0^i t_1},$$

and the convergence also holds in L_p with $p < d$.

Proof. See Appendix B.1 □

Consider now the opposite situation, i.e. \mathbf{a} **is large**. Regarding the maxima and minima, the payoff depends only on their extreme values, and asymptotically only the values of S_T are involved in the payoff. Hence, the weights should correspond to the case of payoffs of the form $\Phi(S_T)$ and Proposition 2.4 below justifies this fact.

Proposition 2.4. *Consider the discrete time case: $I = \{0 = t_0 < \dots < t_i < \dots < t_N = T\}$.*

One has

$$\lim_{a \uparrow \infty} H_{\Delta^i}(a) \stackrel{a.s.}{=} \frac{(\sigma^{-1}e^i) \cdot W_T}{S_0^i T},$$

and the convergence also holds in L_p for any $p \geq 1$.

Proof. See Appendix B.2 □

In view of these two results, we can conclude that our weights behave better than those from Fournié et al. (1999), which involve only the first increment W_{t_1} and lead to high variance when $t_1 \rightarrow 0$, as will be seen in the next section about numerical results. In fact, for usual positive values of a , one may expect that

$$\text{Var} \left(\frac{(\sigma^{-1}e^i) \cdot W_T}{S_0^i T} \right) \ll \text{Var} (H_{\Delta^i}(a)) \ll \text{Var} \left(\frac{(\sigma^{-1}e^i) \cdot W_{t_1}}{S_0^i t_1} \right).$$

3 Numerical experiments

We now illustrate the efficiency of formulae from Theorem 2.1. We restrict the presentation to the discrete time case (the continuous time one has already been studied in Gobet et al. (2001)). In the sequel we take $d = 5$ underlying assets; we set $r = 0\%$ for the interest rate and $T = 1$ year for the maturity.

The initial value of each asset is $S_0^i = 100$ Euros. Their volatilities are given by $\|\sigma^1\| = 35\%$, $\|\sigma^2\| = 35\%$, $\|\sigma^3\| = 38\%$, $\|\sigma^4\| = 35\%$ and $\|\sigma^5\| = 40\%$, and the correlation matrix $\rho = (\rho_{i,j})_{1 \leq i,j \leq 5}$ is defined by $\rho_{i,j} = 0.4$ for $i \neq j$. The number of observations is set to $N = 50$ (roughly weekly monitoring).

In the examples below, we will compare the accuracy of three approaches for the computation of Δ with respect the first asset S^1 , for various options:

1. The *FD* method with a centered difference using a perturbation on the parameter of 1%.
2. The Malliavin calculus weight from Fournié et al. (1999) (so called *Classic Malliavin* or *Classic M.*), for option payoffs depending only on a finite number of dates. In this case the weight for Δ^1 is equal to $H_{\Delta^1} = \frac{(\sigma^{-1}e^1) \cdot W_{t_1}}{S_0^1 t_1}$.
3. The Malliavin calculus weight from this paper (so called *Localized Malliavin* or *Local. M.*):

$$H_{\Delta^1} = \frac{1}{S_0^1} \delta \left(\frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^1 \right).$$

The number of simulations has been fixed to $M = 200000$. In the captions of the figures, we give Monte Carlo estimates with the 95%-confidence symmetric interval (i.e. $\pm 1.96 * SD/\sqrt{M}$).

INSERT HERE FIGURE 1

Figure 1: Delta of a *BLAC Down & Out* option. True value $\approx 3.38 \times 10^{-3}$.

FD = $3.15 \times 10^{-3} \pm 3.47 \times 10^{-4}$, *Classic M.* = $3.53 \times 10^{-3} \pm 4.7 \times 10^{-4}$, *Local. M. (e.p.)* = $3.37 \times 10^{-3} \pm 8.34 \times 10^{-5}$, *Local. M. (a.q.i.p.)* = $3.48 \times 10^{-3} \pm 3.26 \times 10^{-4}$.

We first consider a *BLAC Down & Out* option, whose payoff is $\prod_{i \neq j} \mathbf{1}_{\min_{s \in I} S_s^i \vee \min_{s \in I} S_s^j > L}$, with a barrier at $L = 76$ (we take $a = 0.274$). On Figure 1, we plot the Monte Carlo estimates w.r.t. the number of simulations; hence, the range of the fluctuations gives a visual indication of how large is the variance of the simulations for each method. For the *Local. M.* method, we use two different dominating processes, the **extreme process** (*e.p.* in short) given by (6) and the **averaged quadratic increments process** (*a.q.i.p.* in short) defined by (7). It turns out that the weights using *Local. M.* methods yield lower variance than for *FD* or *Classic M.* methods: the mean error is divided by a factor 4 ($\approx 3.47 \times 10^{-4} / 8.34 \times 10^{-5}$), so reaching a given accuracy requires 16 times less simulations if we use the *Local. M.* method. The use of the extreme process as dominating process is more accurate than a.q.i.p: this can be explained by the fact that the former dominates in a closer way than the second, performing a better localization. Hence, in the next simulations, we will only use the extreme process.

INSERT HERE FIGURE 2

Figure 2: Delta of a *BLAC Down & Out* option with $N = 200$. True value $\approx 3.34 \times 10^{-3}$.

$FD = 3.27 \times 10^{-3} \pm 3.54 \times 10^{-4}$, *Classic M.* $= 3.44 \times 10^{-3} \pm 9.09 \times 10^{-4}$, *Local. M.* $= 3.32 \times 10^{-3} \pm 8.53 \times 10^{-5}$.

Still considering the previous example, we now increase the monitoring frequency, taking $N = 200$ (almost daily monitoring). The results on Figure 2 indicate that our localized weights keep working well comparing to *FD* method (still with a gain of accuracy of a factor $4 \approx 3.54 \times 10^{-4} / 8.53 \times 10^{-5}$), while the *Classic M.* performance worsens as the monitoring frequency increases.

INSERT HERE FIGURE 3

Figure 3: Delta of a *BLAC Down & Out* option. High correlation. True value $\approx 3.75 \times 10^{-3}$.

$FD = 3.55 \times 10^{-3} \pm 3.69 \times 10^{-4}$, *Classic M.* $= 3.45 \times 10^{-3} \pm 1.61 \times 10^{-3}$, *Local. M.* $= 3.96 \times 10^{-3} \pm 2.52 \times 10^{-4}$.

An interesting issue concerns the impact of the non-degeneracy condition on the volatility σ ; indeed, the inverse of σ appears explicitly in the Malliavin calculus weights and this presumably leads to a loss a accuracy in the algorithm if σ is nearly degenerate. To see this, we take highly correlated assets, by setting $\rho_{i,j} = 0.9$ for $i \neq j$ (the monitoring frequency is $N = 50$). We see on Figure 3 that *FD* method still behaves well (the range of the half-confidence interval now equals 3.69×10^{-4} instead of 3.47×10^{-4} on Figure 1), whereas the two other methods lead to larger standard deviations (increase of a factor $3 \approx 2.52 \times 10^{-4} / 8.34 \times 10^{-5}$). In this extreme situation, *FD* method and *Local. M.* method roughly lead to same accuracy.

We now consider an *Up & Down barrier* option, whose payoff is given by $\mathbf{1}_{M^1 > U^1} \mathbf{1}_{m^2 < L^2} \mathbf{1}_{m^3 < L^3} \mathbf{1}_{M^4 > U^4} \mathbf{1}_{m^5 < L^5}$, where $U^1 = 123$, $L^2 = 82$, $L^3 = 81$, $U^4 = 126$ and $L^5 = 79$. The correlations are defined by $\rho_{1,2} = \rho_{2,1} = 0.8$, and $\rho_{i,j} = 0.4$ for other $i \neq j$. For this example, $a_0 = 0.198$ is the largest value of a which makes the assumption **(S)** still valid; in the simulations, we have taken $a = a_0$ and $a = a_0/10$. In view of Proposition 2.3, the simulations with $a = a_0/10$ should be close (at least in L_2 norm) to the ones using the *Classic M.* method. Figure 4 confirms this fact, since the two corresponding Monte Carlo estimates are very close. Besides, Figure 4 still illustrates a good accuracy of the method taking $a = a_0$ compared to the other ones.

INSERT HERE FIGURE 4

Figure 4: Delta of an *Up & Down barrier* option. Influence of a . True value $\approx 1.15 \times 10^{-3}$.

$FD = 1.28 \times 10^{-3} \pm 2.17 \times 10^{-4}$, *Classic M.* $= 1.15 \times 10^{-3} \pm 1.82 \times 10^{-4}$, *Local. M.* ($a = a_0$) $= 1.12 \times 10^{-3} \pm 8.73 \times 10^{-5}$, *Local. M.* ($a = a_0/10$) $= 1.17 \times 10^{-3} \pm 1.92 \times 10^{-4}$.

4 Conclusion

In a multidimensional Black-Scholes framework, we have derived, using Malliavin calculus techniques, an integration by parts formula for functions that depend on the maxima and the minima of the

underlying process. We apply this formula to obtain an alternative expression for the Delta and Gamma for general barrier and lookback type options. These formulae lead to new unbiased numerical methods for the calculation of Greeks.

We perform some numerical experiments which show that the method introduced is significantly better than the FD method and the classical Malliavin calculus technique proposed by Fournié et al. (1999). In the discrete time case our approach performs better as the monitoring frequency increases. In cases where the volatility degenerates our method is asymptotically the same as the FD method. When the support condition **(S)** is taken to a limit the behavior of our method is comparable to the results in Fournié et al. (1999).

A Proof of Theorem 2.1

A.1 Proof for Δ^i

The arguments follow those in Gobet et al. (2001). Up to using a density argument (see Fournié et al. (1999)), we consider smooth function $\Phi(x^1, \dots, x^d, x^{d+1}, \dots, x^{2d}, x^{2d+1}, \dots, x^{3d})$ with bounded derivatives. Thus, since M^i, m^i, S_T^i are linear with respect to S_0^i , one clearly has:

$$\Delta^i = \mathbf{E} \left(e^{-rT} \left(\frac{M^i}{S_0^i} \partial_{x^i} \Phi + \frac{m^i}{S_0^i} \partial_{x^{d+i}} \Phi + \frac{S_T^i}{S_0^i} \partial_{x^{2d+i}} \Phi \right) \right) \quad (12)$$

where we have omitted the random arguments of Φ to simplify. Besides, the chain rule property (see Proposition 1.1) combined with Lemma 1.1 yields

$$\mathcal{D}_t \Phi = \sum_{j=1}^d \left(M^j \partial_{x^j} \Phi \mathbf{1}_{t \leq \tau_M^j} + m^j \partial_{x^{d+j}} \Phi \mathbf{1}_{t \leq \tau_m^j} + S_T^j \partial_{x^{2d+j}} \Phi \right) \sigma^j. \quad (13)$$

To apply an integration by parts formula, one needs to remove the sources of non-smoothness (i.e. terms of type $\mathbf{1}_{t \leq \tau_M^j}$ and $\mathbf{1}_{t \leq \tau_m^j}$) and this can be performed using the localizing process $\Psi(Y_t)$. For this, we prove that for any $j \in \{1, \dots, d\}$ and $t \in [0, T]$, one has

$$\partial_{x^j} \Phi \mathbf{1}_{t \leq \tau_M^j} \Psi(Y_t) = \partial_{x^j} \Phi \Psi(Y_t), \quad (14)$$

$$\partial_{x^{d+j}} \Phi \mathbf{1}_{t \leq \tau_m^j} \Psi(Y_t) = \partial_{x^{d+j}} \Phi \Psi(Y_t). \quad (15)$$

Consider first equality (14). This reduces to $0 = 0$ if $M^j < S_0^j \exp(a)$, since Φ satisfies assumption **(S)**. On the contrary if $M^j \geq S_0^j \exp(a)$ and if t is such that $\Psi(Y_t) \neq 0$, one has $\max_{s \in I, s \leq t} S_s^j < S_0^j \exp(a) \leq M^j$: therefore $t \leq \tau_M^j$ and the proof of (14) is completed. Similar arguments apply for (15). From (14) and (15), it successively follows that

$$\begin{aligned} t \in [0, T]: \quad \mathcal{D}_t \Phi \Psi(Y_t) &= \sum_{j=1}^d \left(M^j \partial_{x^j} \Phi + m^j \partial_{x^{d+j}} \Phi + S_T^j \partial_{x^{2d+j}} \Phi \right) \sigma^j \Psi(Y_t), \\ \int_0^T \mathcal{D}_t \Phi \cdot (\Psi(Y_t) \sigma^{-1} e^i) dt &= (M^i \partial_{x^i} \Phi + m^i \partial_{x^{d+i}} \Phi + S_T^i \partial_{x^{2d+i}} \Phi) \int_0^T \Psi(Y_t) dt, \\ \int_0^T \mathcal{D}_t \Phi \cdot \left(\frac{\Psi(Y_t)}{S_0^i \int_0^T \Psi(Y_t) dt} \sigma^{-1} e^i \right) dt &= \frac{M^i}{S_0^i} \partial_{x^i} \Phi + \frac{m^i}{S_0^i} \partial_{x^{d+i}} \Phi + \frac{S_T^i}{S_0^i} \partial_{x^{2d+i}} \Phi. \end{aligned}$$

Plug this equality into (12), take the expectation and apply the integration by parts formula (Proposition 1.1) to finish the proof of (10).

Actually, the entire justification of this last step essentially requires that $(\int_0^T \Psi(Y_t) dt)^{-1} \in \bigcap_{p \geq 1} L_p$, which holds true under the assumption on the L_p estimates in Definition 2.1 (see Gobet et al. (2001) for details). Furthermore, if we focus on the influence of the parameter a , one can even find that for any $p \geq 1$ one has, for some $q = q(p) \geq 0$,

$$\mathbf{E} \left(\int_0^T \Psi(Y_t) dt \right)^{-p} \leq C a^{-q}. \quad (16)$$

A.2 Proof for $\Delta^{i,j}$

We start from the expression (10). If $j \neq i$, note that H_{Δ^i} does not depend on S_0^j , so the differentiation with respect to S_0^j concerns only Φ . Hence the same computations as before have to be repeated and this leads to (11). If $j = i$, one additionally has to differentiate the multiplicative factor $1/S_0^i$ in the term H_{Δ^i} , and we are finished.

B Proof of Propositions 2.3 and 2.4

B.1 Proof for Proposition 2.3

We first focus on the L_p -convergence. Using equality (5), one easily gets

$$\delta \left(\frac{\Psi(Y)}{\int_0^T \Psi(Y_t) dt} \sigma^{-1} e^i \right) = \frac{\int_0^T \Psi(Y_t) \sigma^{-1} e^i . dW_t}{\int_0^T \Psi(Y_t) dt} + \frac{\int_0^T ds \Psi'(Y_s) \sigma^{-1} e^i . \int_0^s \Psi(Y_t) \mathcal{D}_t Y_s dt}{(\int_0^T \Psi(Y_t) dt)^2}, \quad (17)$$

where for the first term of the r.h.s., we have taken into account that $(Y_t)_{t \geq 0}$ is an adapted process, while for the second one, we have interchanged the order of time integrals.

Note now that $0 < t_1 \leq \int_0^T \Psi(Y_t) dt \leq T$ since Y is a piecewise constant process, starting at 0. Besides, one can prove that $\mathbf{E} \left[\int_0^T (\Psi(Y_t) - \mathbf{1}_{[0,t_1]}(t))^2 dt \right]^p \xrightarrow{a \rightarrow 0} 0$ for any $p \geq 1$, using the dominated convergence theorem combined with $Y_{t_1} > 0$ a.s.. Then it easily follows that $\lim_{a \downarrow 0} \int_0^T \Psi(Y_t) \sigma^{-1} e^i . dW_t = \sigma^{-1} e^i . W_{t_1}$ and $\lim_{a \downarrow 0} \int_0^T \Psi(Y_t) dt = t_1$, the convergence holding in L_p for any $p \geq 1$.

It remains to prove that $B_a = \int_0^T ds \Psi'(Y_s) \sigma^{-1} e^i . \int_0^s \Psi(Y_t) \mathcal{D}_t Y_s dt$ converges to 0 in L_p -norm for $p < d$. We denote by $Y^{(-1)}$ the generalized inverse of the increasing process Y . For the dominating process given by (6) or (7), an easy computation yields $\sup_{0 \leq s, t \leq T} \|\mathcal{D}_t Y_s\| \leq C$ for some constant C , and it readily follows that

$$|B_a| \leq C \int_0^T ds |\Psi'(Y_s)| \leq \frac{C}{a} \left(T \wedge Y^{(-1)}(a) - t_1 \right). \quad (18)$$

Here the value of the floating constant C has changed from one inequality to another.

A direct application of estimate (19) below leads to the convergence of B_a to 0 in L_p norm for $p < d$, and the L_p -convergence result of Proposition 2.3 is proved.

To show that

$$\mathbf{P} \left(Y^{(-1)}(a) > t_1 \right) \leq C a^d \tag{19}$$

for some constant C , it is enough to note that $\mathbf{P} \left(Y^{(-1)}(a) > t_1 \right) = \mathbf{P} \left(Y_{t_1} < a \right) \leq \mathbf{P} \left(\|\mu t_1 + \sigma W_{t_1}\|_\infty < a \right) \leq \mathbf{P} \left(\|\sigma^{-1} \mu t_1 + W_{t_1}\|_\infty < \|\sigma^{-1}\|a \right) \leq \left(\frac{2\|\sigma^{-1}\|a}{\sqrt{2\pi t_1}} \right)^d$, using the independence of the d Brownian motion components and an uniform upper bound for the Gaussian density.

The *a.s.*-convergence can be proved directly from (17). We omit details.

B.2 Proof for Proposition 2.4

Analogous arguments as for the proof of Proposition 2.3 can apply to show that the following convergence holds in L_p for any $p \geq 1$: $\lim_{a \rightarrow \infty} \int_0^T \Psi(Y_t) \sigma^{-1} e^i \cdot dW_t = \sigma^{-1} e^i \cdot W_T$ and $\lim_{a \rightarrow \infty} \int_0^T \Psi(Y_t) dt = T$. To derive the L_p -convergence of B_a to 0, we remark that inequality (18) obviously yields $|B_a| \leq \frac{C}{a} T$, which completes the proof. As for $a \rightarrow 0$, the *a.s.*-convergence is easy to check from equality (17).

REFERENCES

- Benhamou, E. (2000): "An application of Malliavin calculus to continuous time Asian Options'Greeks." Technical report, London School of Economics, <http://fmg.lse.ac.uk/~benhamou/>.
- Bermin, H.P. (2000): "Hedging lookback and partial lookback options using Malliavin calculus." *Applied Mathematical Finance*, 7(2), 75-100.
- Broadie, M., and P. Glasserman (1996): "Estimating security price derivatives using simulation." *Management Science*, 42(2), 269-285.
- Fournié, E., J.M. Lasry, J. Lebuchoux, P.L. Lions, and N. Touzi (1999): "Applications of Malliavin calculus to Monte Carlo methods in finance." *Finance and Stochastic*, 3(4), 391-412.
- Fournié, E., J.M. Lasry, J. Lebuchoux, and P.L. Lions (2001): "Applications of Malliavin calculus to Monte Carlo methods in finance, II." *Finance and Stochastic*, 5(2), 201-236.
- Gobet, E., and A. Kohatsu-Higa (2001): "Computation of Greeks for barrier and lookback options using Malliavin calculus." Technical report 464, CMAP, <http://www.cmap.polytechnique.fr/~gobet/>.
- Glynn, P.W. (1986): "Stochastic approximation for Monte Carlo optimization," in *Proceedings of the 1986 Winter Simulation Conference*, eds. J. Wilson, J. Henriksen and S. Roberts, 356-365.
- Glynn, P.W. (1986): "Likelihood ratio gradient estimation: an overview," in *Proceedings of the 1987 Winter Simulation Conference*, eds. A. Thesen, H. Grant and W.D. Kelton, 366-375.
- Gobet, E. (2001): "Euler schemes and half-space approximation for the simulation of diffusions in a domain." *ESAIM: Probability and Statistics*, 5, 261-297.
- Garsia, A.M., E. Rodemich, and H.Jr. Rumsey (1970): "A real variable lemma and the continuity of paths of some Gaussian processes." *Indiana University Mathematics Journal*, 20(6), 565-578.
- Glasserman, P., and D.D. Yao (1992): "Some guidelines and guarantees for common random numbers." *Management Science*, 38(6), 884-908.
- Kushner, H.J., and J. Yang (1991): "A Monte Carlo method for sensitivity analysis and parametric optimization of nonlinear stochastic systems." *SIAM J. Control Optim.*, 29(5), 1216-1249.
- L'Ecuyer, P., and G. Perron (1994): "On the convergence rates of IPA and FDC derivative estimators." *Oper. Res.*, 42(4), 643-656.

Nualart, D. (1995): *Malliavin calculus and related topics*. Springer Verlag.

Nualart, D., and J. Vives (1988): "Absolute continuity of the law of the maximum of a continuous process." *C. R. Acad. Sci., Paris*, 307(7), 349-354.

Reiman, M.I., and A. Weiss (1986): "Sensitivity analysis via likelihood ratios," in *Proceedings of the 1986 Winter Simulation Conference*, eds. J. Wilson, J. Henriksen and S. Roberts, 285–289.

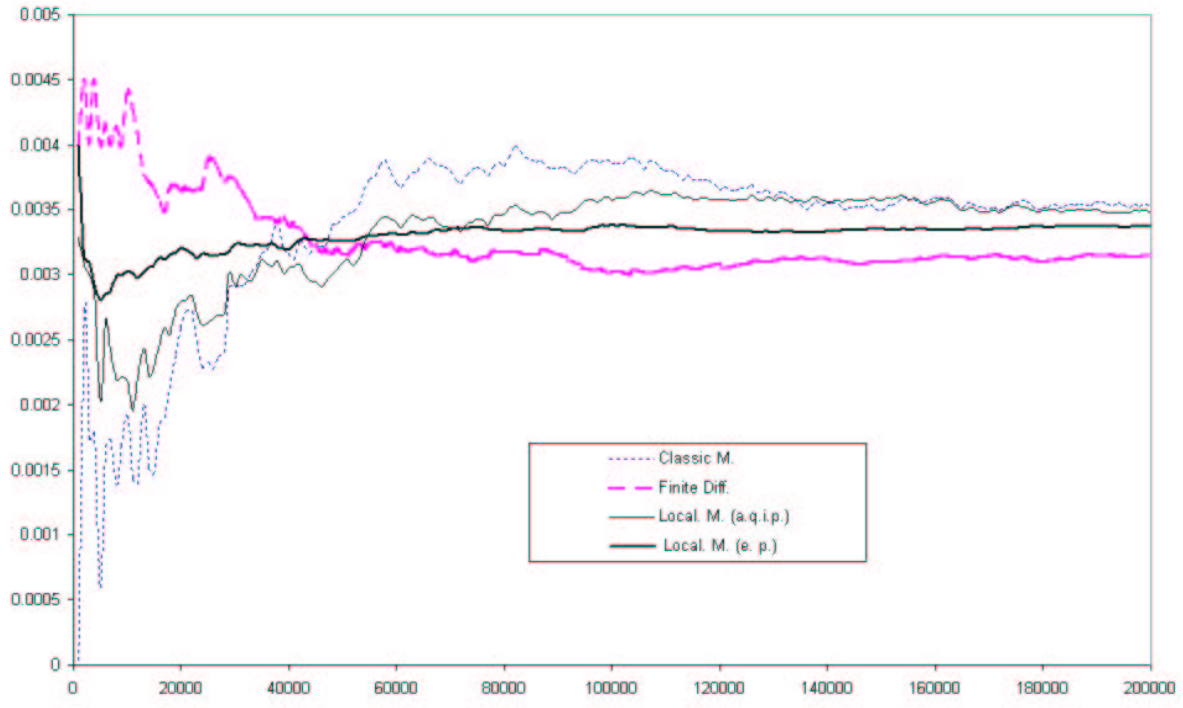


FIGURE 1



FIGURE 2

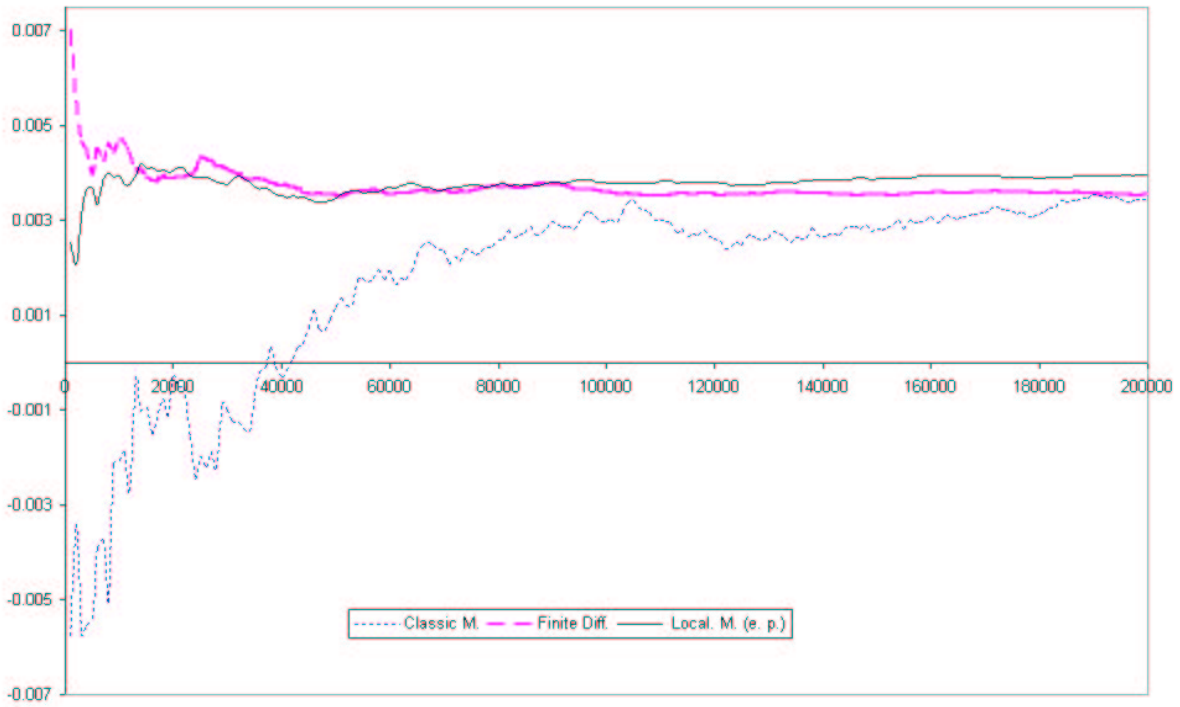


FIGURE 3

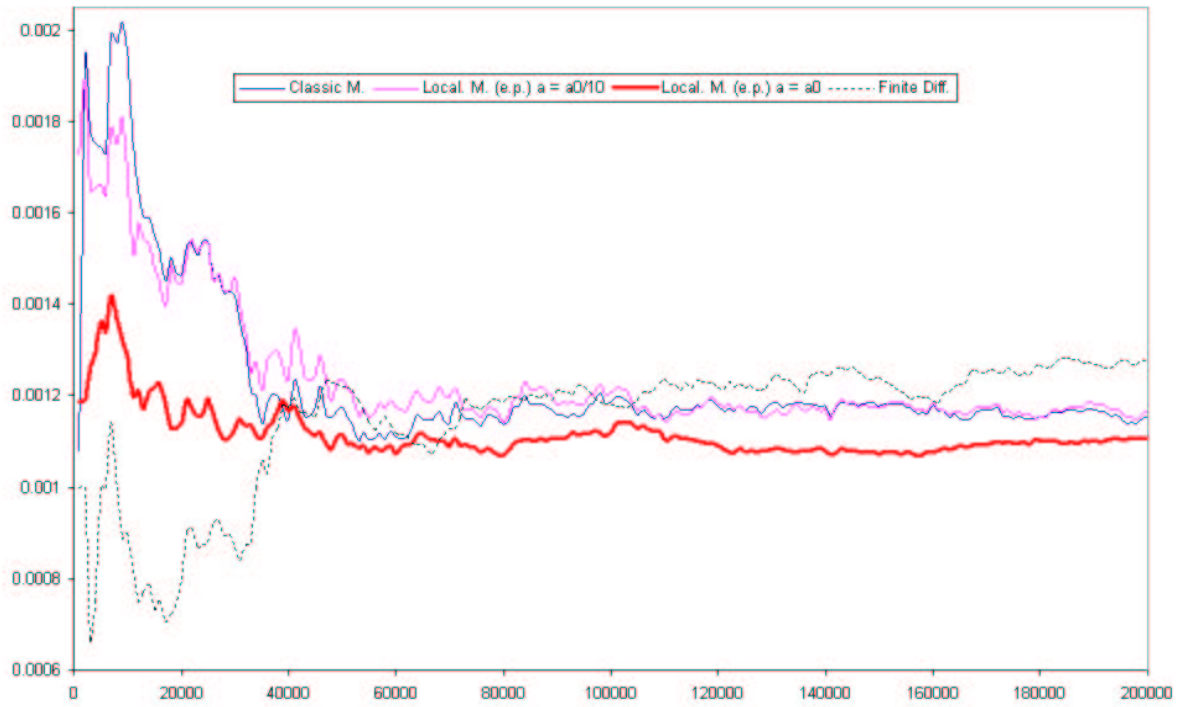


FIGURE 4