Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff

**Michael B. Giles · Desmond J. Higham ·
Xuerong Mao**

Abstract Giles (Oper. Res. 56:607–617, 2008) introduced a multi-level Monte Carlo method for approximating the expected value of a function of a stochastic differential equation solution. A key application is to compute the expected payoff of a financial option. This new method improves on the computational complexity of standard Monte Carlo. Giles analysed globally Lipschitz payoffs, but also found good performance in practice for non-globally Lipschitz cases. In this work, we show that the multi-level Monte Carlo method can be rigorously justified for non-globally Lipschitz payoffs. In particular, we consider digital, lookback and barrier options. This requires non-standard strong convergence analysis of the Euler–Maruyama method.

M.B. Giles
Mathematical Institute and Oxford-Man Institute of Quantitative Finance, University of Oxford, Oxford OX1 3LB, UK

D.J. Higham (✉)
Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, UK
e-mail: aas96106@maths.strath.ac.uk

X. Mao
Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH, UK

1 Background

Many models in mathematical finance take the form of the Itô stochastic differential equations (SDEs)

$$dS(t) = a\big(S(t)\big)\,dt + b\big(S(t)\big)\,dW(t), \quad 0 \le t \le T, \; S(0) \text{ given}, \qquad (1.1)$$

and the expected value of some functional of S is typically of interest [5, 8]. We focus here on the case where $S(t)$ represents the price of an asset at time t and the expected payoff of an option is required.

We assume that the SDE is scalar, with drift and diffusion coefficients satisfying global Lipschitz bounds; so, for some constant L,

$$\big|a(x) - a(y)\big| \vee \big|b(x) - b(y)\big| \le L|x - y|, \quad \forall x, y \in \mathbb{R}, \qquad (1.2)$$

and note that this implies the linear growth bounds

$$\big|a(x)\big| \vee \big|b(x)\big| \le R + L|x|, \quad \forall x \in \mathbb{R},$$

where $R = |a(0)| \vee |b(0)|$. We also suppose that the initial data has bounded moments; that is, for any $m \ge 0$ there is a constant Q_m such that $\mathbb{E}[|S(0)|^m] \le Q_m$. Approximating (1.1) with the Euler–Maruyama (EM) method produces the recurrence $\{X_k\}_{k=0}^M$, with $X_0 = S(0)$ and

$$X_{k+1} = X_k + a(X_k)h + b(X_k)\Delta W_k, \qquad (1.3)$$

where $h = T/M$ is a fixed stepsize and $\Delta W_k = W((k+1)h) - W(kh)$ is a Brownian increment [9, 10].

For the purpose of illustration, consider now the problem of estimating the expected final time solution value, $\mathbb{E}[S(T)]$. Using the EM method (1.3) to compute N approximate samples produces the standard Monte Carlo estimate

$$\mu = \frac{1}{N} \sum_{i=1}^N X_M^{[i]},$$

where $X_M^{[i]}$ is the numerical approximation to $X(T)$ on the ith sample path. The overall error may then be decomposed as

$$\begin{aligned}\mathbb{E}\big[S(T)\big] - \mu &= \mathbb{E}\big[S(T) - X_M + X_M - \mu\big] \\ &= \mathbb{E}\big[S(T) - X_M\big] + \big(\mathbb{E}[X_M] - \mu\big). \end{aligned} \qquad (1.4)$$

The first term on the right-hand side in (1.4), $\mathbb{E}[S(T) - X_M]$, is the *bias*. This is a property of the numerical method. For EM (1.3), classical weak convergence theory shows that this bias is $O(h)$ [2]. The second term, $\mathbb{E}[X_M] - \mu$, is the *statistical error* that arises in any Monte Carlo experiment. From the perspective of a *confidence interval width*, this error is $O(1/\sqrt{N})$ [5]. Overall, we may regard μ as giving an accuracy of $O(h) + O(1/\sqrt{N})$. To achieve a target accuracy of $O(\epsilon)$ we should therefore use the scaling $h = 1/\sqrt{N} = \epsilon$.

Measuring the cost of the computation either in terms of

- the number of evaluations of $a(\cdot)$,
- the number of evaluations of $b(\cdot)$, or
- the number of pseudo-random number calls to obtain the Brownian increments ΔW_k,

we have a computational complexity proportional to the product of the number of paths and the number of steps per path, that is, $O(N/h)$. In terms of the target accuracy ϵ, since $h = 1/\sqrt{N} = \epsilon$, we have computational complexity of $O(\epsilon^{-3})$. This type of complexity analysis is standard and the conclusion holds when EM is used to solve a wide class of problems [3].

Recently, however, Giles [4] showed that by using EM more carefully, notably by taking a range of stepsizes with more paths computed with cheaper, large, stepsizes, it is possible to reduce the complexity. For an option with globally Lipschitz payoff, Giles showed how to achieve complexity of $O(\epsilon^{-2}\log(\epsilon)^2)$. Remarkably, both the algorithm and the underlying analysis depend explicitly on the *strong convergence* property of EM, even though the fundamental quantity to be computed is weak—an expected value. Giles also gave numerical results for options where the payoff is not globally Lipschitz, and showed that improvements on the complexity of standard Monte Carlo also arise. Our aim in this work is to fill a gap in [4] by giving a rigorous analysis of multi-level Monte Carlo for the non-globally Lipschitz cases that were tested numerically. Theorem 3.1 of [4] gives a general complexity result for multi-level Monte Carlo. Using that framework, we need only focus attention on the mean-square error between the exact and numerical payoffs. Hence, for specific types of option, our aim is to derive a bound of the form

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) = O\big(h^\beta\big), \quad \text{for some } \beta < 1, \tag{1.5}$$

where P denotes the payoff from the exact SDE solution, S, and $\widehat{P}$ denotes the corresponding payoff from the EM approximation, $\{X_k\}_{k=0}^N$. Under the assumption that there is a corresponding weak error bound of the form

$$\mathbb{E}(P - \widehat{P}) = O\big(h^\alpha\big), \quad \text{for some } \alpha \geq \frac{1}{2}, \tag{1.6}$$

this allows us to conclude that the standard Monte Carlo complexity of $O(\epsilon^{-2-1/\alpha})$ is improved to $O(\epsilon^{-2-(1-\beta)/\alpha})$ when the multi-level version is used. We note that for the options considered here, we are not aware of a rigorously derived optimal weak order α in (1.6), although work in this direction is progressing [6]. However, we emphasize that (a) our results show that the multi-level approach gives an improvement whatever the value of $\alpha \geq \frac{1}{2}$, and (b) in Sect. 2, where we have $\beta = 1 - \delta$ for any $\delta > 0$, the multi-level complexity becomes independent of α. Our results are also of interest in their own right as non-standard strong error bounds in stochastic numerical analysis.

For valuing path-dependent options we consider the basic numerical method that uses the discrete numerical approximation $\{X_k\}_{k=0}^M$. However, we find it convenient

to work in continuous time, so we define the piecewise linear interpolant $\overline{X}(t)$ by

$$\overline{X}(kh + \theta h) = (1 - \theta)X_k + \theta X_{k+1} \quad \text{for } 0 \leq \theta < 1, \ k = 0, 1, 2, \ldots.$$

We note that for the underlying discrete-time approximation there is a classical mean-square, or *strong*, error bound: there is a constant C such that (for all sufficiently small h)

$$\mathbb{E}\left(\sup_{0 \leq kh \leq T} \left| S(kh) - X_k \right|^2 \right) \leq Ch. \tag{1.7}$$

More generally, for any $m \geq 2$ there is a constant C_m such that

$$\mathbb{E}\left(\sup_{0 \leq kh \leq T} \left| S(kh) - X_k \right|^m \right) \leq C_m h^{m/2}. \tag{1.8}$$

Proofs of these results can be found, for example, in [9], where it is also shown that the analogous error moment bounds hold for a non-computable extension of $\{X_k\}$ to continuous time.

For the computable approximation $\overline{X}(t)$, a continuous-time analogue of (1.8) is available with a slight degradation in the order. Müller-Gronbach [11] established an asymptotic upper bound on $(\mathbb{E}(\sup_{0 \leq t \leq T} |S(t) - \overline{X}(t)|^m))^{1/m}$ that is proportional to $\sqrt{|h \log(h)|}$. For our purposes, it is sufficient to weaken this to the statement that given any $m \geq 2$ and any $\delta > 0$ there is a constant $C_{m,\delta}$ such that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \left| S(t) - \overline{X}(t) \right|^m \right) \leq C_{m,\delta} h^{m/2-\delta}. \tag{1.9}$$

This basic error bound is a key ingredient for our analysis.

The following sections analyse the cases of digital, lookback, up-and-out and down-and-in barrier options. To make the equations less cluttered we have not included the discount factor e^{-rT}, where r is the interest rate (assumed constant). The same complexity results hold, of course, for the discounted expected payoff.

Independent work by Avikainen also provides analysis that is relevant to the multi-level Monte Carlo method. The results in [1] extend material in Sect. 3 by providing a better upper bound for a broader class of final-time payoff functions and an accompanying lower bound.

2 Lookbacks

A floating strike lookback call option differs from the standard European call option in that the strike price is replaced by the smallest asset price observed [5, 7]. So the (undiscounted) lookback payoff is

$$P = S(T) - \inf_{0 \leq t \leq T} S(t).$$

Analogously, our numerical approximation to this payoff is

$$\widehat{P} = \overline{X}(T) - \inf_{0 \leq t \leq T} \overline{X}(t).$$

Hence, using (1.7),

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) \le 2\mathbb{E}\big(\big|S(T) - \overline{X}(T)\big|^2\big) + 2\mathbb{E}\Big(\big|\inf_{0 \le t \le T} S(t) - \inf_{0 \le t \le T} \overline{X}(t)\big|^2\Big)$$

$$= O(h) + 2\mathbb{E}\Big(\big|\inf_{0 \le t \le T} S(t) - \inf_{0 \le t \le T} \overline{X}(t)\big|^2\Big). \tag{2.1}$$

It is straightforward to show that

$$\Big|\inf_{0 \le t \le T} S(t) - \inf_{0 \le t \le T} \overline{X}(t)\Big| \le \sup_{0 \le t \le T} \big|S(t) - \overline{X}(t)\big|. \tag{2.2}$$

Thus, given any $\delta > 0$, using (2.1) and (1.9), we have

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) \le O(h) + 2\mathbb{E}\Big(\sup_{0 \le t \le T} \big|S(t) - \overline{X}(t)\big|^2\Big) = O\big(h^{1-\delta}\big).$$

This shows that (1.5) holds with $\beta = 1 - \delta$ for the floating strike lookback.

Similarly, we have

$$\Big|\sup_{0 \le t \le T} S(t) - \sup_{0 \le t \le T} \overline{X}(t)\Big| \le \sup_{0 \le t \le T} \big|S(t) - \overline{X}(t)\big|, \tag{2.3}$$

which produces the same result for the floating strike lookback put, whose payoff is

$$P = \sup_{0 \le t \le T} S(t) - S(T).$$

A fixed strike lookback call [7] has

$$P = \Big(\sup_{0 \le t \le T} S(t) - K\Big)^+, \quad \text{and so} \quad \widehat{P} = \Big(\sup_{0 \le t \le T} \overline{X}(t) - K\Big)^+.$$

Noting that the function $(.)^+ := \max(., 0)$ has a unit Lipschitz bound, it follows that

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) = \mathbb{E}\Big|\Big(\sup_{0 \le t \le T} S(t) - K\Big)^+ - \Big(\sup_{0 \le t \le T} \overline{X}(t) - K\Big)^+\Big|^2$$

$$\le \mathbb{E}\Big|\sup_{0 \le t \le T} S(t) - \sup_{0 \le t \le T} \overline{X}(t)\Big|^2. \tag{2.4}$$

Using (2.3) and (1.9) we conclude from (2.4) that $\mathbb{E}(|P - \widehat{P}|^2) = O(h^{1-\delta})$.

In this manner we may also obtain the same result for the fixed strike lookback put [7], where $P = (K - \inf_{0 \le t \le T} S(t))^+$.

Hence, all four varieties of lookback allow $\beta = 1 - \delta$ in (1.5), for any $\delta > 0$.

3 Digitals

We now consider a cash-or-nothing call option, which pays one unit if the final time asset price exceeds the fixed strike price K, and pays zero otherwise [7, 8]. Thus, the

discontinuous payoff function has the form

$$P = \mathbf{1}_{\{S(T) > K\}},$$

with the corresponding EM value

$$\widehat{P} = \mathbf{1}_{\{\overline{X}(T) > K\}}.$$

Here $\mathbf{1}_A$ denotes the indicator function for the set A. For a particular path, the exact and numerical payoffs differ only when one solution exceeds K at expiry and the other does not. In particular, we have

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) = \mathbb{P}\big(\{S(T) > K\} \cap \{\overline{X}(T) \leq K\}\big)$$
$$+ \mathbb{P}\big(\{S(T) \leq K\} \cap \{\overline{X}(T) > K\}\big). \tag{3.1}$$

Now, for any given $\delta \in (0, \frac{1}{2})$, we may choose m sufficiently large for

$$\frac{1}{2m+2} < \delta \quad \text{and set} \quad \widehat{\beta} := \frac{1}{2} - \frac{1}{2m+2} > \frac{1}{2} - \delta. \tag{3.2}$$

We consider first the event where only the exact solution path exceeds the strike price. Then, for $h > 0$,

$$\mathbb{P}\big(\{S(T) > K\} \cap \{\overline{X}(T) \leq K\}\big) = \mathbb{P}\big(\{K + h^{\widehat{\beta}} \geq S(T) > K\} \cap \{\overline{X}(T) \leq K\}\big)$$
$$+ \mathbb{P}\big(\{S(T) > K + h^{\widehat{\beta}}\} \cap \{\overline{X}(T) \leq K\}\big)$$
$$\leq \mathbb{P}\big(\{K + h^{\widehat{\beta}} \geq S(T) > K\}\big)$$
$$+ \mathbb{P}\big(\{S(T) - \overline{X}(T) > h^{\widehat{\beta}}\}\big). \tag{3.3}$$

Here, we have introduced a threshold $h^{\widehat{\beta}}$. On the one hand we should like $h^{\widehat{\beta}}$ to be small so that $K + h^{\widehat{\beta}} \geq S(T) > K$ is a low probability event. On the other hand we should like $h^{\widehat{\beta}}$ to be large so that, by appealing to the strong convergence behavior of EM, $S(T) - \overline{X}(T) > h^{\widehat{\beta}}$ is a low probability event. The power $\widehat{\beta}$ in (3.2) has been chosen to balance these two aims. Under the global Lipschitz assumption (1.2) it follows from the Picard iteration used to establish existence and uniqueness [10] that $S(T)$ is a continuously distributed random variable with a bounded density. So

$$\mathbb{P}\big(\{K + h^{\widehat{\beta}} \geq S(T) > K\}\big) = O\big(h^{\widehat{\beta}}\big). \tag{3.4}$$

But from the Markov inequality [9] and the error moment bound (1.8), we have

$$\mathbb{P}\big(\{S(T) - \overline{X}(T) > h^{\widehat{\beta}}\}\big) \leq \frac{\mathbb{E}|S(T) - \overline{X}(T)|^m}{h^{\widehat{\beta}m}}$$
$$\leq \frac{C_m h^{m/2}}{h^{\widehat{\beta}m}} = O\big(h^{\frac{1}{2} - \frac{1}{2m+2}}\big) = O\big(h^{\widehat{\beta}}\big). \tag{3.5}$$

Combining (3.3)–(3.5) we find that

$$\mathbb{P}\big(\{S(T) > K\} \cap \{\overline{X}(T) \leq K\}\big) = O\big(h^{\widehat{\beta}}\big).$$

Similar arguments give

$$\mathbb{P}\big(\{S(T) \leq K\} \cap \{\overline{X}(T) > K\}\big) = O\big(h^{\widehat{\beta}}\big),$$

and so, in (3.1),

$$\mathbb{E}\big(|P - \widehat{P}|^2\big) = O\big(h^{\widehat{\beta}}\big) = O\big(h^{\frac{1}{2} - \delta}\big).$$

Overall, we have shown that for any $\delta > 0$, (1.5) holds with $\beta = \frac{1}{2} - \delta$.

Very similar arguments can be used to obtain the same $\beta = \frac{1}{2} - \delta$ result for a cash-or-nothing put option, where $P = \mathbf{1}_{\{S(T) < K\}}$. Alternatively, the same complexity arises immediately by valuing the put in terms of the call, since their sum is riskless [7, Exercise 17.1]. Similarly, an asset-or-nothing call option, where $P = S(T)\mathbf{1}_{\{S(T) > K\}}$, could be valued via the relation "European call equals asset-or-nothing call minus cash-or-nothing call" and an asset-or-nothing put option, where $P = S(T)\mathbf{1}_{\{S(T) < K\}}$, could be valued in terms of an asset-or-nothing call [7, Exercise 17.8].

4 Barriers

4.1 Ups

An up-and-out call gives a European payoff if the asset never exceeds the barrier, B, where $B > K$; otherwise it pays zero [5, 7, 8]. So, for the exact solution we have

$$P = \big(S(T) - K\big)^+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} S(t) \leq B\}},$$

and for the EM approximation

$$\widehat{P} = \big(\overline{X}(T) - K\big)^+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} \overline{X}(t) \leq B\}}.$$

Consider the events

$$F = \Big\{ \sup_{0 \leq t \leq T} S(t) \leq B \Big\} \quad \text{and} \quad G = \Big\{ \sup_{0 \leq t \leq T} \overline{X}(t) \leq B \Big\}.$$

We have

$$
\begin{aligned}
\mathbb{E}\big(|P - \widehat{P}|^2\big) &= \mathbb{E}\big(\big|(S(T) - K)^+ \mathbf{1}_F - (\overline{X}(T) - K)^+ \mathbf{1}_G\big|^2\big) \\
&= \mathbb{E}\big(\big|(S(T) - K)^+ - (\overline{X}(T) - K)^+\big|^2 \mathbf{1}_{F \cap G}\big) \\
&\quad + \mathbb{E}\big(\big|(S(T) - K)^+\big|^2 \mathbf{1}_{F \cap G^c}\big) + \mathbb{E}\big(\big|(\overline{X}(T) - K)^+\big|^2 \mathbf{1}_{G \cap F^c}\big) \\
&\leq \mathbb{E}\big(\big|S(T) - \overline{X}(T)\big|^2 \mathbf{1}_{F \cap G}\big) \\
&\quad + (B - K)^2 \mathbb{P}\big(F \cap G^c\big) + (B - K)^2 \mathbb{P}\big(G \cap F^c\big)
\end{aligned}
$$

$$\leq \mathbb{E}\big(\big|S(T) - \overline{X}(T)\big|^2\big) + (B - K)^2\big[\mathbb{P}\big(F \cap G^c\big) + \mathbb{P}\big(G \cap F^c\big)\big]$$

$$\leq O(h) + (B - K)^2\big[\mathbb{P}\big(F \cap G^c\big) + \mathbb{P}\big(G \cap F^c\big)\big], \qquad (4.1)$$

where the final step used (1.7).

Now, for any $\delta \in (0, \frac{1}{2})$, choose m sufficiently large and $\gamma > 0$ sufficiently small for

$$\frac{2\gamma + 1}{2m + 2} < \delta \quad \text{so that} \quad \widehat{\beta} := \frac{1}{2} - \frac{2\gamma + 1}{2m + 2} > \frac{1}{2} - \delta.$$

Then

$$\mathbb{P}\big(F \cap G^c\big) = \mathbb{P}\bigg(\bigg\{B - h^{\widehat{\beta}} < \sup_{0 \leq t \leq T} S(t) \leq B\bigg\} \cap G^c\bigg)$$

$$+ \mathbb{P}\bigg(\bigg\{\sup_{0 \leq t \leq T} S(t) \leq B - h^{\widehat{\beta}}\bigg\} \cap G^c\bigg)$$

$$\leq \mathbb{P}\bigg(B - h^{\widehat{\beta}} < \sup_{0 \leq t \leq T} S(t) \leq B\bigg)$$

$$+ \mathbb{P}\bigg(\sup_{0 \leq t \leq T} \overline{X}(t) - \sup_{0 \leq t \leq T} S(t) \geq h^{\widehat{\beta}}\bigg). \qquad (4.2)$$

Since $\sup_{0 \leq t \leq T} S(t)$ is a continuously distributed random variable with bounded density, we have

$$\mathbb{P}\bigg(B - h^{\widehat{\beta}} < \sup_{0 \leq t \leq T} S(t) \leq B\bigg) = O\big(h^{\widehat{\beta}}\big). \qquad (4.3)$$

Also, using (2.3), the Markov inequality and the error moment bound (1.9),

$$\mathbb{P}\bigg(\sup_{0 \leq t \leq T} \overline{X}(t) - \sup_{0 \leq t \leq T} S(t) \geq h^{\widehat{\beta}}\bigg) \leq \mathbb{P}\bigg(\sup_{0 \leq t \leq T} \big(\overline{X}(t) - S(t)\big) \geq h^{\widehat{\beta}}\bigg)$$

$$\leq \frac{1}{h^{\widehat{\beta}m}} \mathbb{E}\bigg(\sup_{0 \leq t \leq T} \big|\overline{X}(t) - S(t)\big|^m\bigg)$$

$$\leq \frac{C_{m,\gamma} h^{m/2 - \gamma}}{h^{\widehat{\beta}m}}$$

$$= O\big(h^{\widehat{\beta}}\big). \qquad (4.4)$$

Using (4.3) and (4.4) in (4.2) we see that $\mathbb{P}(F \cap G^c) = O(h^{\widehat{\beta}})$. Similarly, we can show $\mathbb{P}(G \cap F^c) = O(h^{\widehat{\beta}})$, and hence, in (4.1) we have $\mathbb{E}(|P - \widehat{P}|^2) = O(h^{\widehat{\beta}}) = O(h^{\frac{1}{2} - \delta})$.

An up-and-out put, for which $P = (K - S(T))^+ \mathbf{1}_{\{\sup_{0 \leq t \leq T} S(t) \leq B\}}$, can be analysed in an analogous manner (under the reasonable condition that $S(t) \geq 0$ over $0 \leq t \leq T$ with probability one) giving the same mean-square order.

Up-and-in calls and puts may then be valued with the same complexity using the relation "in plus out equals European" [7, Exercise 19.3].

4.2 Downs

Before analysing options that knock in or out based on a lower bound, we state and prove a lemma.

Lemma 4.1 *Suppose S is a non-negative scalar random variable such that for any integer $q > 0$*

$$\mathbb{E}[S^q] = c_q < \infty.$$

Then for any $x > 0$ and integer $n > 0$

$$\mathbb{E}[S^q \mathbf{1}_{\{S>x\}}] \leq \sqrt{c_{2q+2n}} x^{-n}.$$

Proof Since

$$S^q \mathbf{1}_{\{S>x\}} = S^{q+n}(S^{-n} \mathbf{1}_{\{S>x\}}),$$

the Hölder inequality gives

$$\mathbb{E}[S^q \mathbf{1}_{\{S>x\}}] \leq (\mathbb{E}[S^{2q+2n}])^{1/2}(\mathbb{E}[S^{-2n} \mathbf{1}_{\{S>x\}}])^{1/2} \leq \sqrt{c_{2q+2n}} x^{-n}. \qquad \square$$

A down-and-in call [8] knocks in when the minimum asset price dips below the barrier B, so that

$$P = (S(T) - K)^+ \mathbf{1}_{\{\inf_{0 \leq t \leq T} S(t) \leq B\}},$$

and, accordingly,

$$\widehat{P} = (\overline{X}(T) - K)^+ \mathbf{1}_{\{\inf_{0 \leq t \leq T} \overline{X}(t) \leq B\}}.$$

We can use a similar style of analysis to that in Sect. 4.1 for the up-and-out call. Denote the events

$$H = \left\{ \inf_{0 \leq t \leq T} S(t) \leq B \right\} \quad \text{and} \quad I = \left\{ \inf_{0 \leq t \leq T} \overline{X}(t) \leq B \right\}.$$

Then, using the strong convergence result (1.7),

$$\begin{aligned}
\mathbb{E}(|P - \widehat{P}|^2) &= \mathbb{E}(|(S(T) - K)^+ \mathbf{1}_H - (\overline{X}(T) - K)^+ \mathbf{1}_I|^2) \\
&= \mathbb{E}(|(S(T) - K)^+ - (\overline{X}(T) - K)^+|^2 \mathbf{1}_{H \cap I}) \\
&\quad + \mathbb{E}(|(S(T) - K)^+|^2 \mathbf{1}_{H \cap I^c}) + \mathbb{E}(|(\overline{X}(T) - K)^+|^2 \mathbf{1}_{I \cap H^c}) \\
&\leq \mathbb{E}(|S(T) - \overline{X}(T)|^2) \\
&\quad + \mathbb{E}(|(S(T) - K)^+|^2 \mathbf{1}_{H \cap I^c}) + \mathbb{E}(|(\overline{X}(T) - K)^+|^2 \mathbf{1}_{I \cap H^c}) \\
&= O(h) + \mathbb{E}(|(S(T) - K)^+|^2 \mathbf{1}_{H \cap I^c}) \\
&\quad + \mathbb{E}(|(\overline{X}(T) - K)^+|^2 \mathbf{1}_{I \cap H^c}).
\end{aligned} \tag{4.5}$$

Now, for any given $\delta \in (0, \frac{1}{2})$, choose $m > 2$ sufficiently large and $\gamma > 0$ sufficiently small for

$$\max\left(\frac{2}{m-2}, \frac{2\gamma+1}{2m+2}\right) < \frac{\delta}{2} \quad \text{so that} \quad \widehat{\beta} := \frac{1}{2} - \frac{2\gamma+1}{2m+2} > \frac{1}{2} - \frac{\delta}{2}.$$

Letting $J = \{S(T) > h^{-1/(m-2)}\}$, we have

$$\mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_{H \cap I^c}\big) = \mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_{H \cap I^c \cap J}\big)$$
$$+ \mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_{H \cap I^c \cap J^c}\big). \qquad (4.6)$$

Using Lemma 4.1 with $q = 2$, $x = h^{-1/(m-2)}$ and $n = m - 2$, we see that there is a constant D_m such that

$$\mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_{H \cap I^c \cap J}\big) \leq \mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_J\big) \leq D_m h. \qquad (4.7)$$

Also,

$$\mathbb{E}\big(\big|(S(T)-K)^+\big|^2 \mathbf{1}_{H \cap I^c \cap J^c}\big) \leq h^{-2/(m-2)} \mathbb{P}\big(H \cap I^c\big). \qquad (4.8)$$

Now

$$\mathbb{P}\big(H \cap I^c\big) = \mathbb{P}\left(\left\{B - h^{\widehat{\beta}} < \inf_{0 \leq t \leq T} S(t) \leq B\right\} \cap I^c\right)$$
$$+ \mathbb{P}\left(\left\{\inf_{0 \leq t \leq T} S(t) \leq B - h^{\widehat{\beta}}\right\} \cap I^c\right)$$
$$\leq \mathbb{P}\left(B - h^{\widehat{\beta}} < \inf_{0 \leq t \leq T} S(t) \leq B\right)$$
$$+ \mathbb{P}\left(\inf_{0 \leq t \leq T} \overline{X}(t) - \inf_{0 \leq t \leq T} S(t) \geq h^{\widehat{\beta}}\right). \qquad (4.9)$$

Since $\inf_{0 \leq t \leq T} S(t)$ is a continuously distributed random variable, we have

$$\mathbb{P}\left(B - h^{\widehat{\beta}} < \inf_{0 \leq t \leq T} S(t) \leq B\right) = O\big(h^{\widehat{\beta}}\big). \qquad (4.10)$$

Using the Markov inequality, (2.2) and (1.9), we then have

$$\mathbb{P}\left(\inf_{0 \leq t \leq T} \overline{X}(t) - \inf_{0 \leq t \leq T} S(t) \geq h^{\widehat{\beta}}\right) \leq \mathbb{P}\left(\sup_{0 \leq t \leq T}\big(\overline{X}(t) - S(t)\big) \geq h^{\widehat{\beta}}\right)$$
$$\leq \frac{1}{h^{\widehat{\beta}m}} \mathbb{E}\left(\sup_{0 \leq t \leq T}\big|\overline{X}(t) - S(t)\big|^m\right)$$
$$\leq \frac{C_{m,\gamma} h^{m/2-\gamma}}{h^{\widehat{\beta}m}}$$
$$= O\big(h^{\widehat{\beta}}\big). \qquad (4.11)$$

Using (4.10) and (4.11) in (4.9) we have $\mathbb{P}(H \cap I^c) = O(h^{\widehat{\beta}})$, whence (4.7) and (4.8) in (4.6) give

$$\mathbb{E}\big(\big|\big(S(T) - K\big)^+\big|^2 \mathbf{1}_{H \cap I^c}\big) = O\big(h^{\frac{-2}{m-2} + \widehat{\beta}}\big) = O\big(h^{\frac{1}{2} - \delta}\big). \tag{4.12}$$

Next we note from (1.9) that $\mathbb{E}[|\overline{X}(T)|^m]$ has an upper bound that is uniform over small h, so that Lemma 4.1 remains applicable. So, analogously to (4.12), we can show that

$$\mathbb{E}\big(\big|\big(\overline{X}(T) - K\big)^+\big|^2 \mathbf{1}_{I \cap H^c}\big) = O\big(h^{\frac{1}{2} - \delta}\big),$$

so that, from (4.5), the overall error bound $\mathbb{E}(|P - \widehat{P}|^2) = O(h^{\frac{1}{2} - \delta})$ holds for any $\delta > 0$.

Entirely analogous arguments give the same order for the down-and-in put [7], where $P = (K - S(T))^+ \mathbf{1}_{\{\inf_{0 \le t \le T} S(t) \le B\}}$, and, as mentioned in Sect. 4.1, the "out" versions can be valued via "in plus out equals European" [7, Exercise 19.3].

5 Discussion

Our aim in this work was to develop new mean-square convergence rates for Euler–Maruyama approximations to non-globally Lipschitz option payoffs. Specifically, we have shown that, for any $\delta > 0$, (1.5) holds with $\beta = 1 - \delta$ in the case of lookbacks and with $\beta = \frac{1}{2} - \delta$ in the case of digitals and barriers. As discussed in Sect. 1, these results allow us to quantify an improvement in computational complexity when standard Monte Carlo is replaced by the multi-level version of [4]; they also explain the numerical results presented there. Promising topics for further work in this area include the analysis of (a) the weak error rate α in (1.6) for path-dependent options, (b) methods with higher strong order, and (c) quasi-Monte Carlo.

References

1. Avikainen, R.: On irregular functionals of SDEs and the Euler Scheme. Finance Stoch. **13**, 381–401 (2009)
2. Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations (i): convergence rate of the distribution function. Probab. Theory Relat. Fields **104**, 43–60 (1995)
3. Duffie, D., Glynn, P.: Efficient Monte Carlo simulation of security prices. Ann. Appl. Probab. **7**, 325–348 (1995)
4. Giles, M.B.: Multi-level Monte Carlo path simulation. Oper. Res. **56**, 607–617 (2008)
5. Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer, Berlin (2004)
6. Gobet, E., Menozzi, S.: Discrete sampling of functionals of Itô processes. In: Séminaire de Probabilités XL, pp. 355–375. Springer, Berlin (2007)
7. Higham, D.J.: An Introduction to Financial Option Valuation: Mathematics, Stochastic and Computation. Cambridge University Press, Cambridge (2004)
8. Hull, J.C.: Options, Futures, & Other Derivatives, 4th edn. Prentice Hall, New Jersey (2000)
9. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations, 3rd edn. Springer, Berlin (1999)
10. Mao, X.: Stochastic Differential Equations and Applications, 2nd edn. Horwood, Chichester (2007)
11. Müller-Gronbach, T.: The optimal uniform approximation of systems of stochastic differential equations. Ann. Appl. Probab. **12**, 664–690 (2002)