

## Mean Exit Times and the Multilevel Monte Carlo Method\*

Desmond J. Higham<sup>†</sup>, Xuerong Mao<sup>†</sup>, Mikolaj Roj<sup>†</sup>, Qingshuo Song<sup>‡</sup>, and George Yin<sup>§</sup>

**Abstract.** Numerical methods for stochastic differential equations are relatively inefficient when used to approximate mean exit times. In particular, although the basic Euler–Maruyama method has weak order equal to one for approximating the expected value of the solution, the order reduces to one half when it is used in a straightforward manner to approximate the mean value of a (stopped) exit time. Consequently, the widely used standard approach of combining an Euler–Maruyama discretization with a Monte Carlo simulation leads to a computationally expensive procedure. In this work, we show that the multilevel approach developed by Giles [*Oper. Res.*, 56 (2008), pp. 607–617] can be adapted to the mean exit time context. In order to justify the algorithm, we analyze the strong error of the discretization method in terms of its ability to approximate the exit time. We then show that the resulting multilevel algorithm improves the expected computational complexity by an order of magnitude, in terms of the required accuracy. Numerical results are provided to illustrate the analysis.

**Key words.** Euler–Maruyama approximation, expected computational cost, expected hitting time, mean exit time, Monte Carlo, multilevel Monte Carlo, stochastic differential equation, stochastic simulation, strong and weak convergence

**AMS subject classifications.** 65C30, 65C20, 65C05

**DOI.** 10.1137/120883803

**1. Background and notation.** We begin with the system of stochastic differential equations (SDEs),

$$(1.1) \quad dX(s) = b(X(s))ds + \sigma(X(s))dW(s),$$

with deterministic initial condition  $X(0) = x$ , over a finite time interval  $[0, T]$ . We assume that  $X$  takes values in  $\mathbb{R}^d$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d_1}$ . Here  $W = \{W(t) : t \geq 0\}$  is a standard  $d_1$ -dimensional Brownian motion, and we let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$  be a complete, filtered probability space satisfying the usual conditions. For a specified open set  $O \subset \mathbb{R}^d$ , the *stopped exit time* is the first time at which  $X(s)$  leaves the open set  $O$ , or  $T$  if this is smaller. Our quantity of interest is the expected value of this random variable.

\*Received by the editors July 6, 2012; accepted for publication July 27, 2012; published electronically March 27, 2013.

<http://www.siam.org/journals/juq/1/88380.html>

<sup>†</sup>Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, UK ([djh@maths.strath.ac.uk](mailto:djh@maths.strath.ac.uk), [x.mao@strath.ac.uk](mailto:x.mao@strath.ac.uk), [mikolaj.roj@strath.ac.uk](mailto:mikolaj.roj@strath.ac.uk)). The first author was supported by a Fellowship from the Leverhulme Trust. The third author was supported by the Engineering and Physical Research Council and by Numerical Algorithms Group (NAG) Ltd.

<sup>‡</sup>Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong ([song.qingshuo@cityu.edu.hk](mailto:song.qingshuo@cityu.edu.hk)). This author's work was supported in part by the Research Grants Council of Hong Kong (CityU 103310) and by CityU Strategic Research grant 7002677.

<sup>§</sup>Department of Mathematics, Wayne State University, Detroit, MI 48202 ([gyin@math.wayne.edu](mailto:gyin@math.wayne.edu)). This author's work was supported in part by the National Science Foundation under DMS-1207667.

Such exit times are important in many applications, including air traffic management (using the Markov chain approach) [23], manufacturing flexibility [37], quantum electrodynamics [4], the adoption of technological innovation [20], electronic systems [22, 30], optimal decision making [27, 33], and finance, insurance, and economics [8, 25, 36, 38].

The approach of approximating a mean exit time by directly simulating paths of the SDE and applying a Monte Carlo technique has been adopted by many authors [2, 5, 26, 28], and, relative to the alternative of solving an associated deterministic partial differential equation, it has the advantages of (a) being straightforward to implement and (b) coping naturally with high dimension and complicated boundaries. However, in this context the inherently high cost of controlling the Monte Carlo sampling error is exacerbated by the large biases in the numerical method—exit time samples are less accurate than the corresponding samples of the solution process itself. In section 4.1 we show that it takes an  $O(\epsilon^{-4})$  amount of computational effort to compute an  $O(\epsilon)$  confidence interval for the mean exit time with the standard Monte Carlo approach.

Our aim in this work is to show that the multilevel idea that Giles [13] introduced for the problem of approximating mean values of the form  $\mathbb{E}[f(X(T))]$  may be adapted to the mean exit time context, reducing the computational cost to  $O(\epsilon^{-3}|\log(\epsilon)|^{1/2})$ . A proper analysis of the multilevel approach requires an understanding of both the weak and strong convergence rates of the underlying discretization method, and for this reason we present what appears to be the first strong convergence result for the mean exit time problem (Theorem 3.1).

The work is organized as follows. Section 2 outlines the existing approximation results that are relevant to our analysis. In section 3, we state and prove a new strong convergence result. Section 4 derives the expected computational cost of standard Monte Carlo, defines a multilevel algorithm, and shows that it offers improved complexity. In section 5 we provide illustrative computational results, and concluding remarks appear in section 6.

We finish this section by introducing some further notation. For the purposes of analysis, we let  $X^{t,z}(s)$  denote the general solution of the SDE (1.1) at time  $s$  with initial condition  $X(t) = z$ . So the specific solution of interest,  $X(s)$ , is shorthand for  $X^{0,x}(s)$ .

The Euler–Maruyama approximation  $Y_n^{t,z} \approx X^{t,z}(s_n)$  has  $Y_0 = z$  and

$$(1.2) \quad Y_{n+1}^{t,z} = Y_n^{t,z} + hb(Y_n^{t,z}) + \sigma(Y_n^{t,z})(W(s_{n+1}) - W(s_n)),$$

where  $s_n = t + nh$ , and  $h$  is the stepsize [24]. A straightforward, continuous-time extension  $Y^{t,z}(s) \approx X^{t,z}(s)$  may then be defined as

$$(1.3) \quad Y^{t,z}(s) = Y_n^{t,z} \quad \text{for } s \in [s_n, s_{n+1}).$$

For consistency, we use  $Y(s)$  as shorthand for our approximation to the given problem; that is,  $Y(s) = Y^{0,x}(s)$ , and we use  $Y_n$  to denote  $Y_n^{0,x}$ .

Next, we introduce notation for the stopped exit time. Having specified the open set  $O$  with boundary  $\partial O$  and a fixed future time  $T$ , we denote the stopped exit time for the SDE by

$$\tau^{t,z} := (\inf\{s > t : X^{t,z}(s) \notin O\}) \wedge T,$$

with  $\tau$  as shorthand for  $\tau^{0,x}$ . Here  $a \wedge b$  denotes  $\min(a, b)$ .

Similarly, for the continuously extended Euler–Maruyama approximation we let

$$\nu^{t,z} := (\inf\{s > t : Y^{t,z}(s) \notin O\}) \wedge T,$$

with  $\nu$  as shorthand for  $\nu^{0,x}$ . We note that this exit time arises when the natural approximation algorithm is used: record the first discrete time point at which the Euler–Maruyama path exits the set  $O$ , or  $T$  if this is smaller. We also use  $\mathbf{1}$  to denote the indicator function.

The continuous-time extension (1.3) takes the form of a step process, so  $\nu$  corresponds to the first grid point where the numerical solution exits the region of interest, or  $T$  if this is smaller. A natural alternative is to use a piecewise linear continuous-time Euler–Maruyama extension, and the analysis that we give here would apply equally well to this case, with the same resulting complexity.

**2. Current results.** Our analysis makes use of existing results concerning the finite-time strong convergence of the Euler–Maruyama method when used to simulate paths of the SDE, and also the weak error arising when this method is used to approximate the stopped exit time.

We impose the following conditions on the drift and diffusion coefficients of the SDE.

*Assumption 1.*

1.  *$C^2$  continuity:*  $b$  and  $\sigma$  have two continuous bounded derivatives on  $O$ .
2. *Strong ellipticity:* for some  $c > 0$ ,  $\sum_{ij}(\sigma(x)\sigma^*(x))_{ij}\xi_i\xi_j > c|\xi|^2$  for all  $x \in O$ ,  $\xi \in \mathbb{R}^d$ .
3. *Regularity of the boundary:* for  $d > 1$ ,  $O \subset \mathbb{R}^d$  is a bounded open set with its boundary  $\partial O$  being  $C^2$  smooth.

In the above,  $C^2$  continuity is imposed on  $b$  and  $\sigma$  to ensure existence and uniqueness of strong solutions of the SDE and to provide extra smoothness needed for (2.3). Strong ellipticity and regularity of  $\partial O$  are needed in Lemma 2.1. We emphasize that the stopped exit time problem restricts the solution to a compact domain. It follows that we may redefine  $b$  and  $\sigma$  outside this domain, if necessary, in order to ensure that they are globally Lipschitz. This allows us to derive results that apply to a wide range of nonlinear SDE models, including those arising in finance and chemical kinetics that pose analytical difficulties through nondifferentiability of the drift at the origin [1, 34] or superlinear growth of the diffusion at infinity [7, 10, 35].

We then have the classical finite-time strong error result for the discrete Euler–Maruyama approximation,

$$(2.1) \quad \left( \mathbb{E} \left[ \sup_{0 \leq nh \leq T} |X(nh) - Y(nh)|^q \right] \right)^{\frac{1}{q}} = O(h^{\frac{1}{2}}), \quad 1 \leq q < \infty;$$

see, for example, [24, 29]. The continuously extended Euler–Maruyama approximation has a slight degradation in order,

$$(2.2) \quad \left( \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X(s) - Y(s)|^q \right] \right)^{\frac{1}{q}} = O(|h \log(h)|^{\frac{1}{2}}), \quad 1 \leq q < \infty.$$

This inequality was proved in [32] for the case where piecewise linear interpolation is used, and a result for the step process then follows via the triangle inequality.

The work of Gobet and Menozzi [19, Theorem 17] (see also [17] and [18]) gives an optimal rate for the weak convergence of the Euler–Maruyama stopped exit time:

$$(2.3) \quad \mathbb{E}[\tau] - \mathbb{E}[\nu] = O(h^{\frac{1}{2}}).$$

Finally, we give a result that follows directly from the smoothness of the mean exit time as a function of the initial condition.

**Lemma 2.1.** *Under Assumption 1 there exists a constant  $K$  such that for any  $z \in O$  we have*

$$\mathbb{E}[\tau^{0,z}] \leq K(d_z \wedge T),$$

where  $d_z = \inf\{|z - y| : y \in \partial O\}$ .

*Proof.* Due to the strong ellipticity in Assumption 1, the function

$$u(z) := \mathbb{E}[\inf\{s > 0 : X^{0,z}(s) \notin O\}]$$

is known to be  $u \in C^2(O) \cap C(\bar{O})$ , satisfying the Dirichlet problem

$$(2.4) \quad Lu + 1 = 0 \text{ on } O, \quad u = 0 \text{ on } \partial O,$$

where  $L$  is a strongly elliptic operator given by

$$Lu = \frac{1}{2} \text{tr}\{\sigma\sigma^* D^2 u\} + b \cdot Du,$$

with  $\text{tr}(A)$  denoting the trace of  $A$  [9]. Since the boundary is  $C^2$  smooth,  $u \in C^2(\bar{O})$ , thanks to [11, Theorem 6.14]. It follows that  $u$  is Lipschitz on  $\bar{O}$ , and we let  $K$  denote an appropriate Lipschitz constant. For any given  $z \in O$  and  $y \in \partial O$  we then have  $u(z) = |u(z) - u(y)| \leq K|z - y|$ , and hence

$$u(z) \leq Kd_z.$$

By construction,  $\mathbb{E}[\tau^{0,z}] \leq u(z)$  and  $\mathbb{E}[\tau^{0,z}] \leq T$ , and the result follows. ■

**3. Main theorem.** In order to justify a multilevel Monte Carlo approach we first establish a rate of strong convergence for the exit time approximation.

**Theorem 3.1.** *Under Assumption 1 we have*

$$\mathbb{E}[|\tau - \nu|^p] = O(|h \log(h)|^{\frac{1}{2}}) \quad \forall p \geq 1.$$

*Proof.* The proof deals separately with two cases. We first consider the event that the Euler–Maruyama approximation exits before the exact process.

*Case I.*  $\nu < \tau$ . We have

$$(3.1) \quad \begin{aligned} \mathbb{E}[(\tau - \nu) \mathbf{1}_{\{\nu < \tau\}}] &= \mathbb{E}[\mathbb{E}[(\tau - \nu) \mathbf{1}_{\{\nu < \tau\}} | \mathcal{F}_\nu]] \\ &= \mathbb{E}[\mathbb{E}[\tau^{\nu, X(\nu)} | \mathcal{F}_\nu]]. \end{aligned}$$

On the right-hand side, we have the expected stopped exit time for the exact process, starting from the stopped exit time of the numerical approximation. In order to bound this quantity, we use the following two properties:

- Lemma 2.1 tells us that the stopped exit time  $\tau^{\nu, X(\nu)}$  will be small in mean if the process starts close to a boundary,
- strong convergence (2.2) tells us that the exact solution is close to the numerical approximation, which has already reached the boundary or been stopped.

More precisely, applying Lemma 2.1 and (2.2) to (3.1), we find

$$(3.2) \quad \mathbb{E} [(\tau - \nu) \mathbf{1}_{\{\nu < \tau\}}] \leq K \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X(s) - Y(s)| \right] = O(|h \log(h)|^{\frac{1}{2}}).$$

Next we consider the event that the exact process exits before the Euler–Maruyama approximation.

*Case II.*  $\tau < \nu$ . We have

$$\begin{aligned} \mathbb{E} [(\nu - \tau) \mathbf{1}_{\{\tau < \nu\}}] &= \mathbb{E} [\mathbb{E} [(\nu - \tau) \mathbf{1}_{\{\tau < \nu\}} | \mathcal{F}_\tau]] \\ &= \mathbb{E} [\mathbb{E} [\nu^{\tau, Y(\tau)} | \mathcal{F}_\tau]]. \end{aligned}$$

Because Lemma 2.1 applies to the exact SDE process, rather than the Euler–Maruyama approximation, the next step is to add and subtract the exact process that runs forward in time from  $t = \tau$  with initial condition  $Y(\tau)$  and apply the triangle inequality to give

$$\begin{aligned} \mathbb{E} [(\nu - \tau) \mathbf{1}_{\{\tau < \nu\}}] &= \mathbb{E} \left[ \mathbb{E} [\nu^{\tau, Y(\tau)} - \tau^{\tau, Y(\tau)} + \tau^{\tau, Y(\tau)} | \mathcal{F}_\tau] \right] \\ &\leq \left| \mathbb{E} [\mathbb{E} [\nu^{\tau, Y(\tau)} | \mathcal{F}_\tau]] - \mathbb{E} [\mathbb{E} [\tau^{\tau, Y(\tau)} | \mathcal{F}_\tau]] \right| \\ &\quad + \mathbb{E} [\mathbb{E} [\tau^{\tau, Y(\tau)} | \mathcal{F}_\tau]]. \end{aligned}$$

The first term on the right-hand side concerns the weak error in the mean exit time algorithm, which is known from (2.3) to be  $O(h^{\frac{1}{2}})$ . The second term on the right-hand side can be bounded using the same arguments that lead from (3.1) to (3.2), showing that it is  $O(|h \log(h)|^{\frac{1}{2}})$ . Hence, we find that

$$(3.3) \quad \mathbb{E} [(\nu - \tau) \mathbf{1}_{\{\tau < \nu\}}] = O(|h \log(h)|^{\frac{1}{2}}).$$

Combining (3.2) and (3.3) provides the result for  $p = 1$ . For a more general  $p \geq 1$ , the required result follows from

$$\mathbb{E} [|\tau - \nu|^p] \leq T^{p-1} \mathbb{E} [|\tau - \nu|]. \quad \blacksquare$$

Strong convergence of Theorem 3.1 shows that  $\nu \rightarrow \tau$  in  $L^1$  with a rate of approximately  $\frac{1}{2}$  as  $h \rightarrow 0$ . However, one may worry about a scenario where the two exit times are close but the exact solution and numerical approximation exit at different areas of phase space. This is illustrated for the scalar case in Figure 1, with  $O := (\alpha, \beta)$ . In what follows, we show that this event is unlikely.

**Corollary 3.2.** *Under Assumption 1 we have*

$$\mathbb{E} [ |X(\tau) - Y(\nu)|^2 ] = O(|h \log(h)|^{\frac{1}{2}}).$$

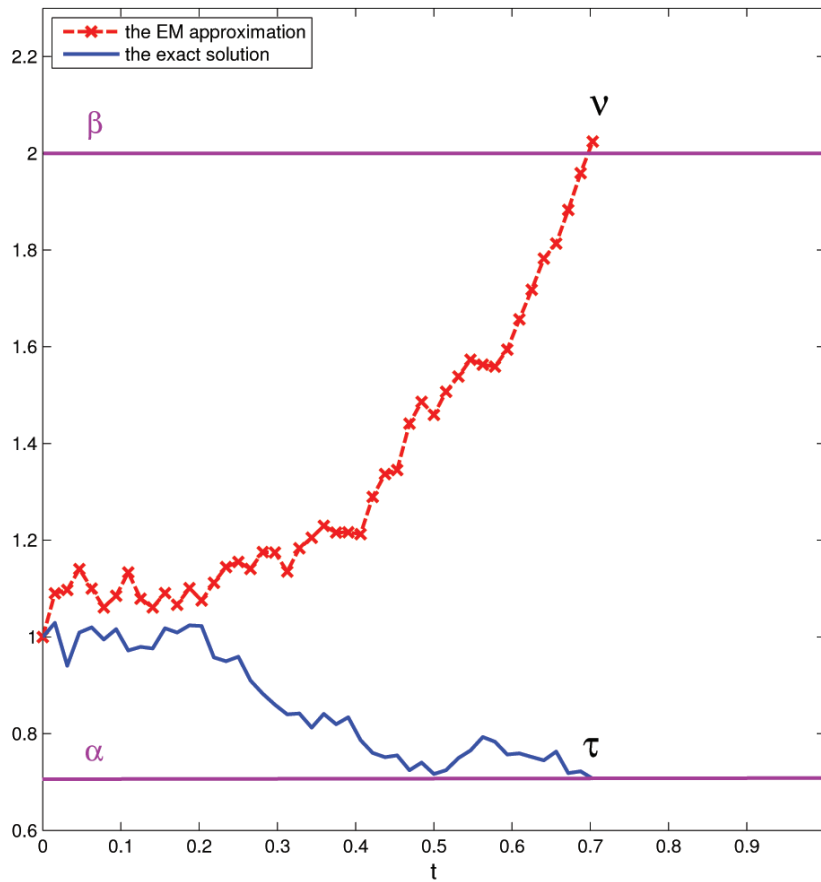


Figure 1. Bad approximation scenario for exit time.

Proof. We have

$$\begin{aligned} \mathbb{E}[|X(\tau) - Y(\nu)|^2] &= \mathbb{E}[|X(\tau) - Y(\nu)|^2 \mathbf{1}_{\{\tau \leq \nu\}}] + \mathbb{E}[|X(\tau) - Y(\nu)|^2 \mathbf{1}_{\{\tau > \nu\}}] \\ &\leq 2 \mathbb{E}[|X(\tau) - Y(\tau)|^2] + 2 \mathbb{E}[|Y(\tau) - Y(\nu)|^2 \mathbf{1}_{\{\tau \leq \nu\}}] \\ &\quad + 2 \mathbb{E}[|X(\nu) - Y(\nu)|^2] + 2 \mathbb{E}[|X(\tau) - X(\nu)|^2 \mathbf{1}_{\{\tau > \nu\}}]. \end{aligned}$$

From (2.2), the first and third terms on the right-hand side here are  $O(|h \log(h)|)$ . Intuitively, the second term is the case where the exact solution is stopped, and we want to show that we cannot move with the numerical approximation too far in a small interval of time. The fourth term is a mirror case of the second term, but this time the numerical approximation is stopped, and we want to show that we cannot move with the exact solution too far in a small interval of time. First, let us focus on bounding the fourth term,

$$\mathbb{E}[|X(\tau) - X(\nu)|^2 \mathbf{1}_{\{\tau > \nu\}}] = \mathbb{E} \left[ \mathbf{1}_{\{\tau > \nu\}} \left( \left| \int_{\nu}^{\tau} b(X(t))dt + \int_{\nu}^{\tau} \sigma(X(t))dW(t) \right|^2 \right) \right]$$

$$\begin{aligned}
&\leq \mathbb{E} \left[ \left| \int_{\nu \wedge \tau}^{\tau} b(X(t))dt + \int_{\nu \wedge \tau}^{\tau} \sigma(X(t))dW(t) \right|^2 \right] \\
&\leq 2 \mathbb{E} \left[ (\tau - \nu \wedge \tau) \int_{\nu \wedge \tau}^{\tau} |b(X(t))|^2 dt + \int_{\nu \wedge \tau}^{\tau} |\sigma(X(t))|^2 dt \right] \\
&\leq 2C_1 \mathbb{E} [ (\tau - \nu \wedge \tau)^2 + (\nu - \nu \wedge \tau) ] \\
&\leq 2C_2 \mathbb{E} [ |\tau - \nu| ] = O(|h \log(h)|^{\frac{1}{2}}),
\end{aligned}$$

where in the last line we used Theorem 3.1. In a similar manner it can be shown that

$$\mathbb{E}[ |Y(\tau) - Y(\nu)|^2 \mathbf{1}_{\{\tau \leq \nu\}} ] = O(|h \log(h)|^{\frac{1}{2}}),$$

which finishes the proof.  $\blacksquare$

#### 4. Monte Carlo.

**4.1. Standard Monte Carlo.** A traditional Monte Carlo approach to the estimation of the mean stopped exit time uses the Euler–Maruyama method to compute independent samples  $\nu^{[i]}$  from the distribution of the corresponding random variable  $\nu$ . Let  $\nu^{[i]}$  denote the computed stopped exit time for the  $i$ th simulated path. If we use  $N$  such paths, then the mean value  $\mathbb{E}[\tau]$  is approximated by the sample average

$$\mu = \frac{1}{N} \sum_{i=1}^N \nu^{[i]}.$$

The overall error divides naturally into two parts,

$$(4.1) \quad \mathbb{E}[\tau] - \mu = \mathbb{E}[\tau - \nu + \nu] - \mu = (\mathbb{E}[\tau - \nu]) + (\mathbb{E}[\nu] - \mu).$$

The first term in parentheses is the bias, that is, the weak error of the numerical method in terms of its ability to approximate the mean stopped exit time of the SDE. We know from (2.3) that this term is  $O(h^{\frac{1}{2}})$ . The second term in (4.1) concerns the statistical sampling error. This is known to scale like  $O(1/\sqrt{N})$  from the perspective of confidence interval width (see, for example, [16]).

It is natural to measure the computational cost in terms of either

- the total number of evaluations of the coefficients  $b(\cdot)$  or  $\sigma(\cdot)$  when we use the iteration (1.2), or
- the number of pseudorandom number calls to obtain the Brownian increments in (1.2).

In both cases, the computational cost of each path is proportional to the ratio of the time-span of the numerical approximation,  $\nu^{[i]}$ , and the stepsize,  $h$ . The overall expected computational cost of the standard Monte Carlo method is therefore proportional to  $N\mathbb{E}[\nu]/h$ , which, from (2.3), may be written

$$(4.2) \quad \frac{N \left( \mathbb{E}[\tau] + O(h^{\frac{1}{2}}) \right)}{h}.$$

If we let  $\epsilon$  denote the target level of accuracy, in terms of confidence interval width, then balancing the bias and sampling error in (4.1) gives the scaling  $\epsilon = h^{\frac{1}{2}} = 1/\sqrt{N}$ , whence the complexity measure (4.2) for the method becomes  $O(\epsilon^{-4})$ .

**4.2. Multilevel Monte Carlo.** Giles in [13] considered the problem of computing some function of the final-time SDE solution,  $\mathbb{E}[f(X(T))]$ , where  $f$  is globally Lipschitz, a problem that arises naturally in financial option pricing. He showed the remarkable fact that it is possible to reduce the complexity of the standard Euler–Maruyama/Monte Carlo method from  $O(\epsilon^{-3})$  to  $O(\epsilon^{-2}(\log \epsilon)^2)$ ; see also [21] for related ideas. Further work has extended these results to more general classes of function  $f$  and other discretization methods [12, 14], to quasi-Monte Carlo [15], and also to related problems involving stochastic partial differential equations [3, 6]. Our aim here is to develop the multilevel methodology in the exit time context.

Following [13], we consider a range of different stepsizes of the form  $h_l = M^{-l}T$  for  $l = 0, 1, 2, \dots, L$ . Here  $M$  is a fixed integer. The smallest stepsize,  $h_L = M^{-L}T$ , is chosen so that the bias in the discretization method matches the target accuracy of  $O(\epsilon)$ ; matching  $h_L^{\frac{1}{2}}$  with  $\epsilon$  then gives

$$(4.3) \quad L = \frac{\log \epsilon^{-2}}{\log M}.$$

Intuitively, the multilevel approach exploits the fact that it is not necessary to compute many paths at this high, and expensive, level of resolution. It is sufficient to compute a relatively small number of “high frequency” paths and then pad out the computation with increasingly more information from the increasingly cheaper lower-resolution stepsizes.

To be more precise, we let the random variable  $\nu_l$  denote the stopped exit time arising when the Euler–Maruyama approximation is used with stepsize  $h_l$ . We continue by writing the following trivial identity:

$$(4.4) \quad \mathbb{E}[\nu_L] = \mathbb{E}[\nu_0] + \sum_{l=1}^L \mathbb{E}[\nu_l - \nu_{l-1}].$$

On the left is the exact mean of the high-resolution approximation, which has the required bias. On the right is a telescoping series involving the different levels of resolution. We propose estimating the expected values on the right-hand side of (4.4) as follows. The first term,  $\mathbb{E}[\nu_0]$ , is estimated by a quantity  $Z_0$  that uses the sample mean of  $N_0$  independent paths; so

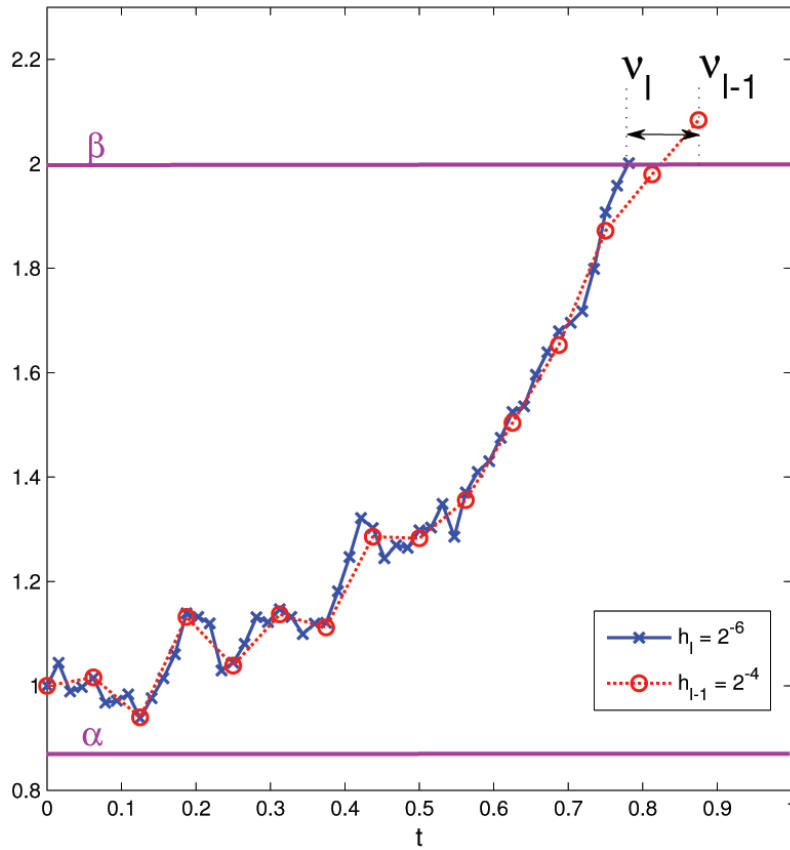
$$Z_0 = \frac{1}{N_0} \sum_{i=1}^{N_0} \nu_0^{[i]}.$$

Each remaining term of the form  $\mathbb{E}[\nu_l - \nu_{l-1}]$  is estimated by a quantity  $Z_l$  based on  $N_l$  independent pairs of paths; thus

$$(4.5) \quad Z_l = \frac{1}{N_l} \sum_{i=1}^{N_l} \left( \nu_l^{[i]} - \nu_{l-1}^{[i]} \right).$$

Here, the two samples  $\nu_l^{[i]}$  and  $\nu_{l-1}^{[i]}$  come from the same Brownian path at the two different levels of resolution. Figure 2 illustrates the case where  $M = 4$  for a scalar SDE, so the two paths are based on stepsizes that differ in size by a factor of 4. Each Brownian increment





**Figure 2.** Illustration of the estimator  $Z_l$ , in the case  $M = 4$  for a scalar SDE. The same path is sampled at two different levels of resolution.

for the lower-resolution computation is given by the sum of the four increments used for the higher-resolution path. Independent paths are used for each  $i$  in (4.5) and for each different level  $l = 1, 2, \dots, L$ .

It remains for us to control the statistical sampling error by choosing  $\{N_l\}_{l=0}^L$  in order to give an overall variance of  $O(\epsilon^2)$  for the estimator.

Using Theorem 3.1, we have

$$\text{Var}[\nu_l - \tau] \leq \mathbb{E}[(\nu_l - \tau)^2] = O(|h_l \log(h_l)|^{\frac{1}{2}})$$

and

$$(4.6) \quad \text{Var}[\nu_l - \nu_{l-1}] \leq (\sqrt{\text{Var}[\nu_l - \tau]} + \sqrt{\text{Var}[\nu_{l-1} - \tau]})^2 = O(|h_l \log(h_l)|^{\frac{1}{2}}).$$

So  $\text{Var}[Z_l] = O(|h_l \log(h_l)|^{\frac{1}{2}}/N_l)$ , and, because the computations are independent over different levels, our overall estimator  $Z := Z_0 + \sum_{i=1}^L Z_i$  has variance

$$\text{Var}[Z] = \text{Var}[Z_0] + \sum_{l=1}^L O\left(\frac{|h_l \log(h_l)|^{\frac{1}{2}}}{N_l}\right).$$

Taking

$$(4.7) \quad N_0 = O(\epsilon^{-2}),$$

$$(4.8) \quad N_l = O(\epsilon^{-2} M^{L/4} h_l^{3/4} |\log(h_l)|^{1/2}) \quad \text{for } l = 1, \dots, L,$$

it follows that

$$\begin{aligned} \text{Var}[Z] &= O(\epsilon^2) + \sum_{l=0}^L O\left(\frac{|h_l \log(h_l)|^{1/2}}{N_l}\right) \\ &= O(\epsilon^2) + \sum_{l=0}^L O\left(\frac{|h_l \log(h_l)|^{1/2}}{\epsilon^{-2} M^{L/4} h_l^{3/4} |\log(h_l)|^{1/2}}\right) \\ &= O(\epsilon^2) + O\left(\epsilon^2 M^{-L/4} \sum_{l=0}^L h_l^{-1/4}\right) \\ &= O(\epsilon^2) + O\left(\epsilon^2 M^{-L/4} \sum_{l=0}^L M^{l/4}\right) \\ &= O(\epsilon^2), \end{aligned}$$

as required.

Having specified the algorithm, we may now work out the expected computational cost. Each pair of paths at level  $l$  has a cost proportional to  $v_l^{[i]}/h_l$ . Hence, the expected computational cost of the multilevel Monte Carlo method is

$$\begin{aligned} \sum_{l=0}^L \frac{N_l}{h_l} \mathbb{E}[\nu_l] &= \sum_{l=0}^L \frac{\epsilon^{-2} M^{L/4} h_l^{3/4} |\log(h_l)|^{1/2}}{h_l} \left(\mathbb{E}[\tau] + O(h_l^{1/2})\right) \\ &= O\left(\epsilon^{-2} M^{L/4} |\log(h_L)|^{1/2} \sum_{l=0}^L h_l^{-1/4}\right) \\ &= O\left(\epsilon^{-2} M^{L/2} |\log(\epsilon)|^{1/2}\right). \end{aligned}$$

From (4.3) we see that  $M^{L/2} = O(\epsilon^{-1})$ , and so the expected computational cost may be written

$$O\left(\epsilon^{-3} |\log(\epsilon)|^{1/2}\right).$$

This should be compared with the value  $O(\epsilon^{-4})$  that was derived in section 4.1 for standard Monte Carlo.

**5. Computational results.** We now present computational results, focusing on the scalar case. Our aims are to test the sharpness of the analysis and to check whether the asymptotically valid improvement in complexity can be observed in a real simulation. Here  $O$  takes the form of an open interval, denoted  $(\alpha, \beta)$ .

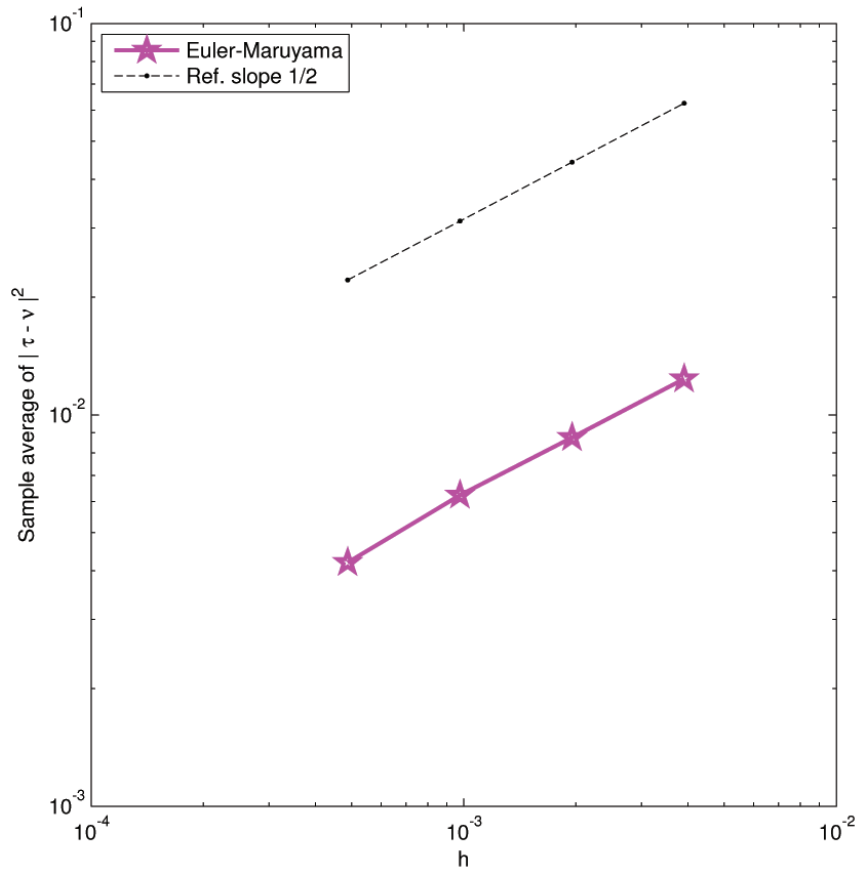


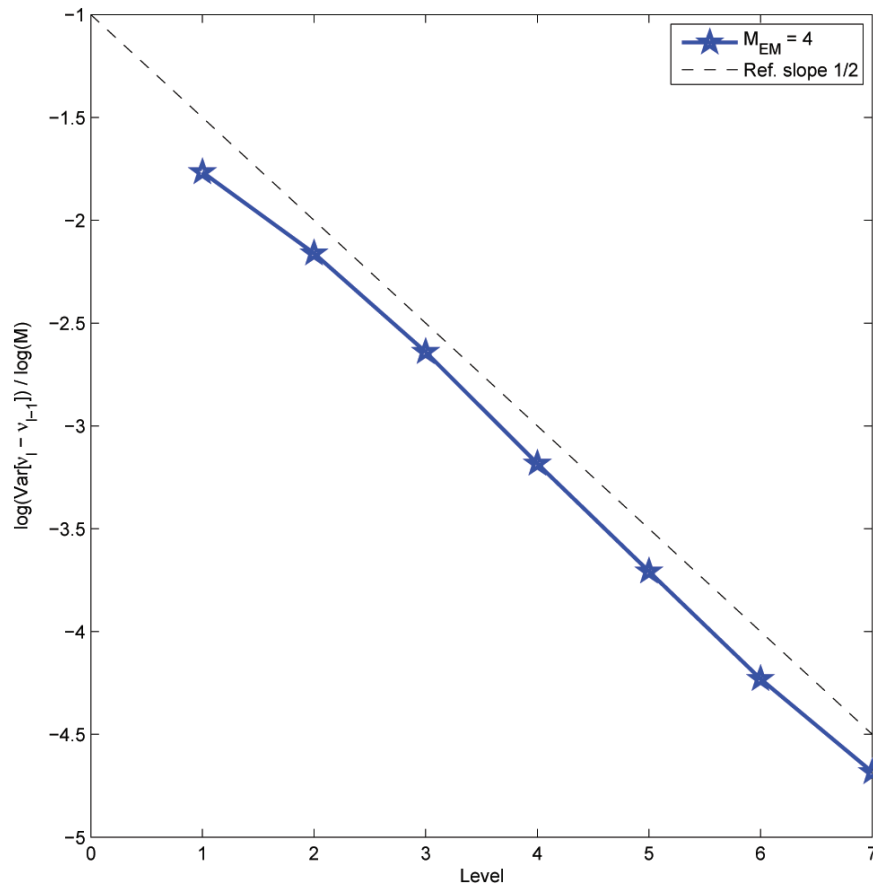
Figure 3. Strong error in stopped exit time.

**5.1. Strong error in exit time.** We begin by checking the sharpness of the strong convergence rate from Theorem 3.1 on a geometric Brownian motion model,

$$(5.1) \quad dX(s) = bX(s)ds + \sigma X(s)dW(s),$$

with constant drift coefficient  $b = 0.05$ , constant volatility  $\sigma = 0.2$ , and initial value  $X(0) = 2$ . We set the boundaries to  $\alpha = 1.7$  and  $\beta = 2.3$ , and the finite cutoff time to  $T = 1$ . We computed the reference solution  $\mathbb{E}[\tau] = 0.4962$  using a numerical solver for the boundary value ordinary differential equation. We approximated the strong exit time error  $\text{err}_h := \mathbb{E}[|\tau - \nu|^2]$  using the sample mean from  $N = 5000$  path simulations, and stepsizes  $h = 2^{-8}, 2^{-9}, 2^{-10}, 2^{-11}$ . The error behavior on a log-log scale is shown in Figure 3. A least squares fit for  $\log C$  and  $q$  in  $\log \text{err}_h = \log C + q \log h$  produced  $q = 0.5167$  with a least squares residual of 0.0289. Error bars representing 95% confidence intervals are small enough to be covered by stars in the figure. Thus, our results suggest that the strong order of convergence equal to approximately one-half in Theorem 3.1 is sharp for  $p = 2$ .

**5.2. Variance behavior.** A key step in our analysis of the multilevel Monte Carlo algorithm was the derivation of the variance estimate (4.6). In Figure 4 we plot the quantity



**Figure 4.** Variance of  $v_l - v_{l-1}$  over different levels.

$\log(\text{Var}[v_l - v_{l-1}]) / \log(M)$  over a sequence of levels, where the most refined level corresponds to a user-specified accuracy of  $\epsilon = 0.001$ . If (4.6) is sharp, then this plot will have a slope of approximately  $-1/2$ . A least squares fit for the slope gives  $q = -0.4984$  with a residual of 0.1040. We also include a line with a slope  $-1/2$  for reference. Here we used the linear SDE (5.1) with boundaries changed to  $\alpha = 0.9$ ,  $\beta = 1.1$  and initial value  $X(0) = 1$ , whence  $\mathbb{E}[\tau] = 0.2480$ .

**5.3. Multilevel Monte Carlo.** Finally, we compare the computational cost versus accuracy for standard and multilevel mean exit time computation. We measure the computational cost of the multilevel method as

$$\text{Cost}_{\text{MLMC}} := \left( N_0 + \sum_{l=1}^L N_l M^l \right) \mathbb{E}[\tau].$$

For the complexity analysis in section 4.2, we specified an order of magnitude for the number of paths per level, (4.8), and showed that this (a) produces the correct variance and (b) has attractive complexity.

In our implementation, we used the values

$$(5.2) \quad N_l = \left\lceil 2\epsilon^{-2} \sqrt{\text{Var}[\nu_l - \nu_{l-1}] M^{-l}} \left( \sum_{l=0}^L \sqrt{\text{Var}[\nu_l - \nu_{l-1}] M^l} \right) \right\rceil, \quad 0 \leq l \leq L.$$

More precisely, we used order-of-magnitude estimates of the variances  $\text{Var}[\nu_l - \nu_{l-1}]$  that were obtained numerically in a preprocessing step of negligible cost. To justify the choice (5.2), we note that

$$\begin{aligned} \text{Var}[Z] &= \sum_{l=0}^L \left( \frac{\text{Var}[\nu_l - \nu_{l-1}]}{N_l} \right) \\ &= \sum_{l=0}^L \frac{\text{Var}[\nu_l - \nu_{l-1}]}{2\epsilon^{-2} \sqrt{\text{Var}[\nu_l - \nu_{l-1}] M^{-l}} \left( \sum_{l=0}^L \sqrt{\text{Var}[\nu_l - \nu_{l-1}] M^l} \right)} \\ &= \frac{1}{2} \epsilon^2, \end{aligned}$$

so we achieve a small enough variance.

We also note that the order of magnitude for  $N_l$  in (4.8) that we used for the complexity analysis satisfies

$$\begin{aligned} \epsilon^{-2} M^{L/4} h_l^{3/4} |\log(h_l)|^{1/2} &\leq \epsilon^{-2} M^{L/4} h_l^{3/4} |\log(h_l)|^{1/4} |\log(h_L)|^{1/4} \\ &\leq \epsilon^{-2} h_l^{3/4} |\log(h_l)|^{1/4} \sum_{l=0}^L M^{l/4} |\log(h_l)|^{1/4}. \end{aligned}$$

Since  $\text{Var}[\nu_l - \nu_{l-1}] = O(h_l^{1/2} |\log(h_l)|^{1/2})$  it follows that our practical choice gives a comparable order of paths per level.

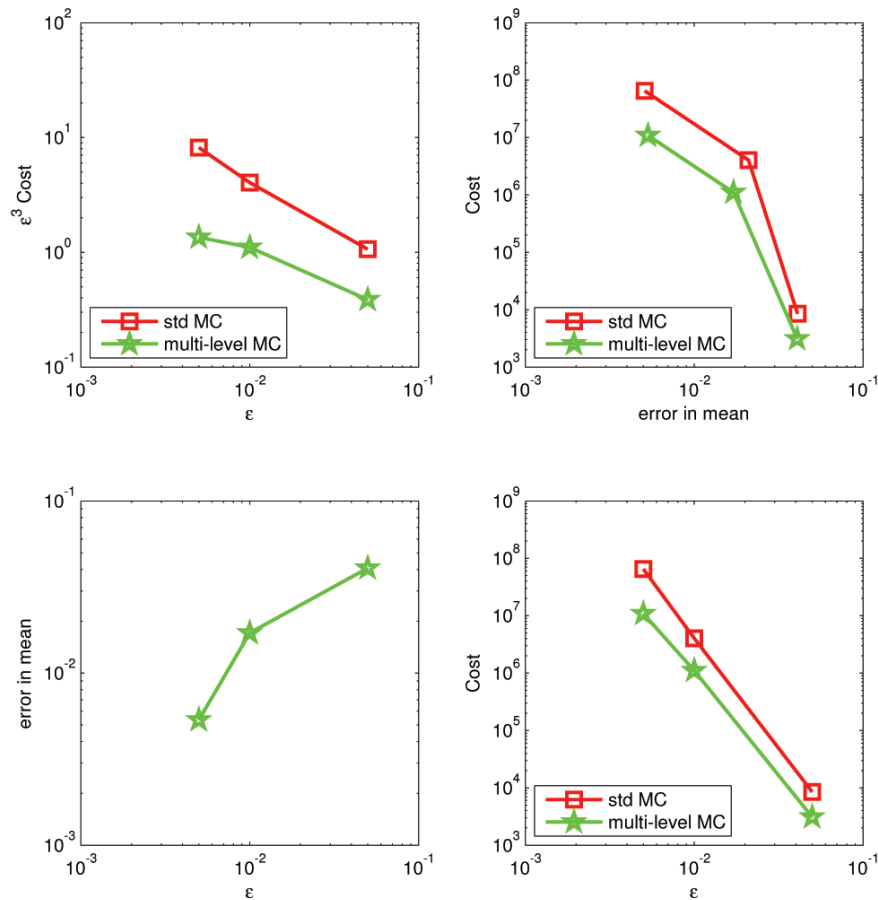
The computational cost of the standard method is measured as

$$\text{Cost}_{\text{stdMC}} := \frac{N}{h} \mathbb{E}[\tau],$$

where  $N$  is the total number of sample paths and  $h$  is the fixed stepsize such that  $h = O(\epsilon^2)$ . Here  $N$  was chosen adaptively to produce the required  $O(\epsilon^2)$  variance. We used three choices of the accuracy parameter,  $\epsilon = 10^{-1.5}, 10^{-2}, 10^{-2.5}$ . For the SDE and parameters used in Figure 4, the lower-left picture in Figure 5 shows the accuracy obtained by the multilevel algorithm as a function of  $\epsilon$ . This confirms that the algorithm produces an error that scales like  $\epsilon$ .

In the top-left picture of Figure 5 we plot  $\epsilon$  against  $\epsilon^3 \text{Cost}$  for the two methods. The results are consistent with the predictions from our analysis: for the multilevel method this scaled complexity increases slowly like  $|\log \epsilon|^{1/2}$ , whereas for the standard Monte Carlo method it increases like  $\epsilon$ .

On the right-hand side of Figure 5, we repeat this information in a slightly different format. The upper-right-hand picture plots the overall cost against the attained error, and the lower-right-hand picture plots the overall cost against the accuracy parameter. We see that for the



**Figure 5.** Cost versus accuracy for the standard Monte Carlo method and the multilevel Monte Carlo algorithm on geometric Brownian motion.

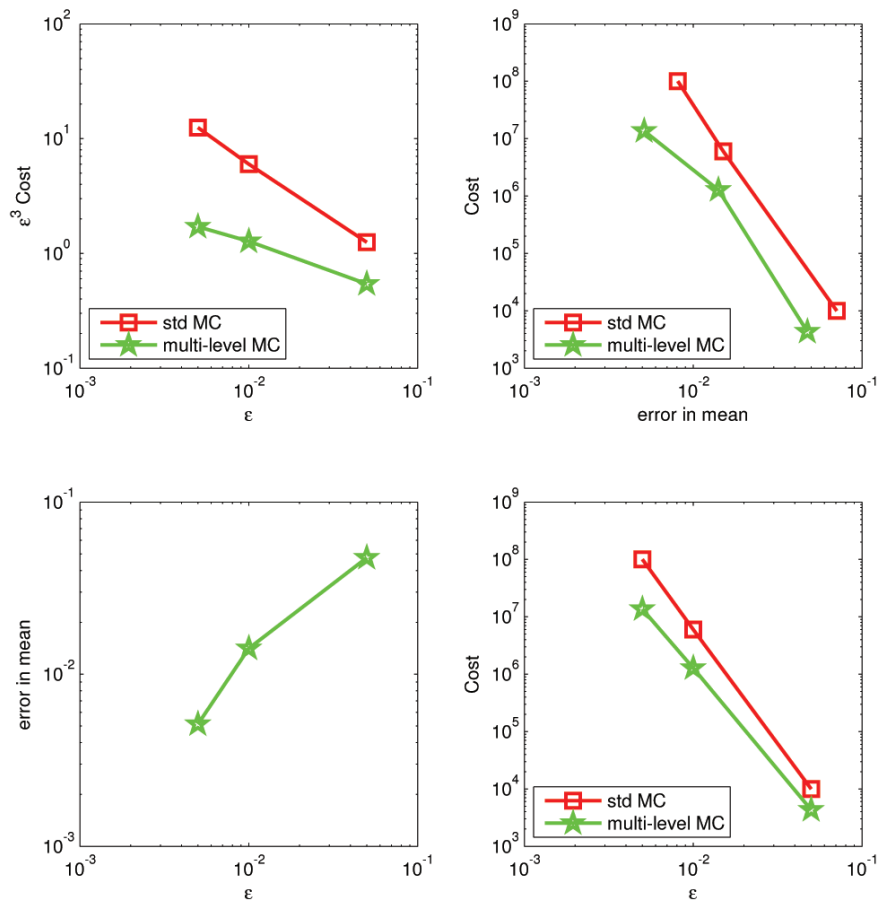
most stringent accuracy requirement the multilevel version is about an order of magnitude cheaper.

As we mentioned in section 2, one of the benefits of a mean exit time analysis is that our attention is restricted to a compact domain, so rigorous results can be derived for nonlinear SDEs with coefficients that do not satisfy a global Lipschitz condition. We consider now a mean reverting square root process commonly used to model interest rates [29],

$$dX(s) = b(\mu - X(s))ds + \sigma\sqrt{X(s)}dW(s),$$

noting that the diffusion coefficient is not Lipschitz at the origin. Keeping all the parameters the same as for the geometric Brownian motion test in Figure 5, additionally we fix  $\mu = 1$ , which causes the stopped mean exit time to increase to 0.2495. The complexity results, shown in Figure 6, are consistent with those for the geometric Brownian motion case.

**6. Discussion.** Our overall aim here was to design and rigorously analyze a multilevel version of a simple and widely used algorithm: Euler/Monte Carlo simulation for a mean exit



**Figure 6.** Cost versus accuracy for the standard Monte Carlo method and the multilevel Monte Carlo algorithm on the mean reverting square root process.

time. We showed that it is possible to dramatically reduce the computational complexity. A key step in the analysis was to establish a rate of strong convergence for the numerical discretization.

Of course there are several other general approaches to mean exit time computation, each having their own advantages and disadvantages [31], and in future work it would be of great interest to compare this multilevel algorithm with other classes of methods on a range of practical problems. Within the context of basic timestepping methods, it would also be of interest to investigate the possibility of incorporating higher order discretization or adaptive stepsize control into the multilevel setting, as well as considering the use of quasi-Monte Carlo methods and fine-tuning the algorithm to high performance computing architectures.

## REFERENCES

- [1] Y. AIT-SAHALIA, *Testing continuous-time models of the spot interest rate*, Rev. Financial Stud., 9 (1996), pp. 385–426.

- [2] L. ANDERSEN, *A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor market model*, J. Comput. Finance, 3 (2000), pp. 5–32.
- [3] A. BARTH, C. SCHWAB, AND N. ZOLLINGER, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.
- [4] D. BEDINGHAM, *State reduction dynamics in a simplified QED model*, J. Phys. A, 41 (2008), 495205.
- [5] P. BOYLE, M. BROADIE, AND P. GLASSERMAN, *Monte Carlo methods for security pricing*, J. Econom. Dyn. Control, 21 (1997), pp. 1267–1321.
- [6] K. CLIFFE, M. GILES, R. SCHEICHL, AND A. TECKENTRUP, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14 (2011), pp. 3–15.
- [7] T. CONLEY, L. HANSEN, E. LUTTMER, AND J. SCHEINKMAN, *Short-term interest rates as subordinated diffusions*, Rev. Financial Stud., 10 (1997), pp. 525–577.
- [8] J. CVITANIĆ AND I. KARATZAS, *Hedging contingent claims with constrained portfolios*, Ann. Appl. Probab., 3 (1993), pp. 652–681.
- [9] J. DOOB, *Classical Potential Theory and Its Probabilistic Counterpart*, Grundlehren Math. Wiss. 262, Springer, New York, 1984.
- [10] A. GALLANT AND G. TAUCHEN, *Estimation of continuous-time models for stock returns and interest rates*, Macroeconomic Dynamics, 1 (2005), pp. 135–168.
- [11] D. GILBARG AND N. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer, New York, 2001.
- [12] M. GILES, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, in Monte Carlo and Quasi-Monte Carlo Methods 2006, A. Keller, S. Heinrich, and H. Niederreiter, eds., Springer-Verlag, 2007, pp. 343–358.
- [13] M. GILES, *Multi-level Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.
- [14] M. GILES, D. HIGHAM, AND X. MAO, *Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff*, Finance Stoch., 13 (2009), pp. 403–413.
- [15] M. GILES AND B. WATERHOUSE, *Multi-level quasi-Monte Carlo path simulation*, Radon Series Comp. Appl. Math., 8 (2009), pp. 165–181.
- [16] P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag, Berlin, 2004.
- [17] E. GOBET AND S. MENOZZI, *Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme*, Stoch. Process. Appl., 112 (2004), pp. 201–223.
- [18] E. GOBET AND S. MENOZZI, *Discrete sampling of functionals of Itô processes*, in Séminaire de Probabilités XL, Lecture Notes in Math. 1899, Springer, Berlin, 2007, pp. 355–374.
- [19] E. GOBET AND S. MENOZZI, *Stopped diffusion processes: Boundary corrections and overshoot*, Stoch. Process. Appl., 120 (2010), pp. 130–162.
- [20] S. GRENADIER AND A. WEISS, *Investment in technological innovations: An option pricing approach*, J. Financial Econ., 44 (1997), pp. 397–416.
- [21] S. HEINRICH, *Multilevel Monte Carlo methods*, in Lecture Notes in Large Scale Scientific Computing, Springer-Verlag, Berlin, 2001, pp. 58–67.
- [22] A. HUSSAIN, *Multisensor distributed sequential detection*, in Aerospace and Electronic Systems, IEEE Trans. Aerospace Electron. Systems 30, IEEE Press, Piscataway, NJ, 2002, pp. 698–708.
- [23] N. KANTAS, A. LECCHINI-VISINTINI, AND J. MACIEJOWSKI, *Simulation-based Bayesian optimal design of aircraft trajectories for air traffic management*, Internat. J. Adaptive Control Signal Process., 24 (2010), pp. 882–899.
- [24] P. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer, New York, 1999.
- [25] H. KUSHNER AND P. DUPUIS, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, Berlin, 2001.
- [26] F. LONGSTAFF AND E. SCHWARTZ, *Valuing American options by simulation: A simple least-squares approach*, Rev. Financial Stud., 14 (2001), pp. 113–148.
- [27] S. MAJD AND R. PINDYCK, *Time to build, option value, and investment decisions*, J. Financial Econ., 18 (1987), pp. 7–27.
- [28] R. MANNELLA, *Absorbing boundaries and optimal stopping in a stochastic differential equation*, Phys. Lett. A, 254 (1999), pp. 257–262.
- [29] X. MAO, *Stochastic Differential Equations and Applications*, 2nd ed., Horwood, Chichester, UK, 2007.



- [30] H. MEYR AND G. POLZER, *A simple method for evaluating the probability density function of the sample number for the optimum sequential detector*, in Communications, IEEE Trans. Communications 35, IEEE Press, Piscataway, NJ, 2002, pp. 99–103.
- [31] G. N. MILSTEIN AND M. V. TRETYAKOV, *Stochastic Numerics for Mathematical Physics*, Scientific Computation, Springer, Berlin, 2004.
- [32] T. MÜLLER-GRONBACH, *The optimal uniform approximation of systems of stochastic differential equations*, Ann. Appl. Probab., 12 (2002), pp. 664–690.
- [33] R. PINDYCK, *Irreversible investment, capacity choice, and the value of the firm*, Amer. Econom. Rev., 78 (1988), pp. 969–985.
- [34] L. SZPRUCH AND D. J. HIGHAM, *Comparing hitting time behavior of Markov jump processes and their diffusion approximations*, Multiscale Model. Simul., 8 (2010), pp. 605–621.
- [35] L. SZPRUCH, X. MAO, D. HIGHAM, AND J. PAN, *Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model*, BIT, 51 (2011), pp. 405–425.
- [36] M. TAKSAR, *Optimal risk and dividend distribution control models for an insurance company*, Math. Methods Oper. Res., 51 (2000), pp. 1–42.
- [37] A. TRIANTIS AND J. HODDER, *Valuing flexibility as a complex option*, J. Finance, 45 (1990), pp. 549–565.
- [38] G. YIN, C. XU, AND L. WANG, *Optimal remapping in dynamic bulk synchronous computations via a stochastic control approach*, IEEE Trans. Parallel Distrib. Systems, 14 (2003), pp. 51–62.