Valuing Employee Stock Options

Cian O Brien

Supervisor: Dr. Cónall Kelly

School of Mathematical Sciences

University College Cork

Ireland

27 March 2019

Contents

Introduction

The employee stock option (ESO) can be viewed as a complex American style call option, which is generally offered to employees of firms as part of their remuneration package. The valuing of these stock options is the subject of this project - in particular we will look at taking a Monte Carlo approach to valuing these options. There are many models currently used to value ESOs; binomial and other lattice models are frequently cited, [6] and the standard Black-Scholes option pricing formula, [3], can be applied using a modified time to expiry to capture the early exercise feature. An advantage of the Monte Carlo approach is that it offers a great deal of flexibility when specifying the underlying asset price model, as well as incorporating the various features of the ESO quite easily.

Section 1 will formually introduce the employee stock option and the features of such a contract. In Section 2, we will examine some numerical methods for approximating solutions of stochastic differential equations - these can be used later to approximate asset prices should a more complex asset price model be introduced. We also examine convergence and perform some stability analysis of the methods. Section 3 will discuss the Monte Carlo valuation of these ESO prices, as well as some variance reduction techniques.

1 Features of the ESO

Unlike a standard call option written on the stock of a company, the ESO is a private contract between an employer and an employee. They are not traded on any exchange, cannot be bought or sold and therefore there does not exist a standard set of features which applies to all such options. For the purposes of this project, we will consider a number of features which hope to capture a very broad range of these ESOs.

We will denote the price of a share by $S(t)$, the time to expiry of the option by T, and the strike price of the option by K. Futhermore let r and σ^2 denote the risk-free interest rate and the volatility of the share price respectively. These parameters are enough to calculate the price at time t of a standard European call option where $t \in [0, T]$. We now introduce some additional option features which pertain to the ESO:

- **Vesting Period**. This is a period after the option grant during which the employee cannot exercise the option. The purpose of this period is to incentivise an employee to stay with the

company after the options have been granted. We will denote the time until vesting by T_0, so in general the vesting period will be written as $[0, T_0]$ with $T_0 < T$. Once the vesting period has ended, the option is said to be 'vested'.

- **Departure from company.** We also consider the possibility of the employee leaving the company after being granted an ESO. In general, if an employee leaves a company during the vesting period, the option is forfeited. If the employee leaves after the vesting period, they are usually obliged to exercise within a short period of time - for simplicity we will assume that employee exercises immediately after departure from the company. We will model this departure by a Poisson process with intensity λ. In particular, we allow for the intensity during the vesting period λ_0 to differ from λ.

- **Early exercise.** To capture the possibility of early exercise of the ESO, we will introduce a 'barrier' L which will prompt early exercise if struck. The reasoning behind this barrier is as follows: we assume that the average employee will happy to exercise early and forgo the possibility of higher profits if the share price is currently at a level sufficiently above the strike price K.

These features are discussed in [2] - the authors propose an analytic pricing formula for four variants of the ESO. We will be using these formulae to benchmark our Monte Carlo estimations under the Black-Scholes asset model assumptions in Section 3.

In a general sense, we are trying to estimate the value of

$$C_t = \mathbb{E}[\, e^{-r(T'-t)} \big(S(T') - K\big)^+],$$

the discounted expected payoff of the option at time t. However, in the case of the ESO, the time T' is no longer a pre-determined time as in the case of a standard call option, but a random variable that can be thought of as

$$T' = min\{\tau, T\}$$

where T is the time to expiry and τ is a random time before expiry, $\tau < T$. The value of τ is determined by the various early exercise features of the ESO as discussed above. The fact that the price of the ESO reduces to an expectation allows us to take a Monte Carlo approach to valuing the ESO by simulating a large number of asset trajectories and taking an average of the corresponding payoffs.

2 Numerical methods

This section will examine two particular numerical methods for solving stochastic differential equations, namely the *Euler-Maruyama method* and the *Milstein method*. We will give a derivation of both methods using stochastic Taylor expansions, and look at both strong and weak convergence of the methods. Furthermore, an in-depth look at the stability of both methods will be presented in the context of solving a particular type of SDE, as well as the stability of the solution of this equation.

2.1 The Euler-Maruyama method

Consider the following stochastic differential equation

$$dX(t) = a\big(X(t)\big)dt + b\big(X(t)\big)\,dW(t), \tag{2.1}$$

with $X(t_0) = X_{t_0}$, where $W(t)$ is a Wiener process and $t \in [t_0, T]$. The coefficient functions a and b are assumed to be sufficiently smooth real-valued functions. The equation can be written in integral form as follows:

$$X(t) = X_{t_0} + \int_{t_0}^{t} a\big(X(s)\big)ds + \int_{t_0}^{t} b\big(X(s)\big)\,dW(s). \tag{2.2}$$

Then, for any twice continuously-differentiable function $f : \mathbb{R} \to \mathbb{R}$, the Itô formula gives

$$\begin{aligned} f(X(t)) = \; & f(X_{t_0}) + \int_{t_0}^{t} \left(a\big(X(s)\big)\frac{\partial}{\partial x} f\big(X(s)\big) + \frac{1}{2} b^2\big(X(s)\big)\frac{\partial^2}{\partial^2 x} f\big(X(s)\big) \right) ds \\ & + \int_{t_0}^{t} b\big(X(s)\big)\frac{\partial}{\partial x} f\big(X(s)\big)\,dW(s) \\ & = f(X_{t_0}) + \int_{t_0}^{t} L^0 f\big(X(s)\big)\,ds + \int_{t_0}^{t} L^1 f\big(X(s)\big)\,dW(s) \end{aligned}$$

for $t \in [t_0, T]$. The operators

$$L^0 f(x) = a(x)\frac{\partial}{\partial x} f(x) + \frac{1}{2} b(x)^2 \frac{\partial^2}{\partial^2 x} f(x),$$

$$L^1 f(x) = b(x)\frac{\partial}{\partial x} f(x)$$

have been introduced here. For $f(x) \equiv x$, we have $L^0 f(x) = a(x)$ and $L^1 f(x) = b(x)$, in which case

the above equation reduces to (2.2), that is to

$$X(t) = X_{t_0} + \int_{t_0}^{t} a\big(X(s)\big)ds + \int_{t_0}^{t} b\big(X(s)\big)\, dW(s), \quad t \in [t_0,\, T].$$

If the Itô formula is then applied to the functions $f = a$ and $f = b$, we have that

$$X(t) = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW(s) + R, \quad t \in [t_0,\, T], \tag{2.3}$$

where

$$\begin{aligned} R = &\int_{t_0}^{t}\int_{t_0}^{s} L^0 a\big(X(z)\big)\, dz\, ds + \int_{t_0}^{t}\int_{t_0}^{s} L^1 a\big(X(z)\big)\, dW(z)\, ds \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0 b\big(X(z)\big)\, dz\, dW(s) + \int_{t_0}^{t}\int_{t_0}^{s} L^1 b\big(X(z)\big)\, dW(z)\, dW(s) \end{aligned} \tag{2.4}$$

contains the the final two terms in the Itô expansion of a and b respectively.

This is the simplest non-trivial Itô-Taylor expansion. If the interval $[t_0,\, T]$ is divided into N sub-intervals of length Δt, we can use (2.33) to generate a stochastic process in discrete time which will approximate the values of $X(t)$: by replacing t_0 and T with $n\Delta t$ and $(n+1)\Delta t$ respectively, which are the endpoints of the n^{th} sub-interval, we have

$$\begin{aligned} X\big((n+1)\Delta t\big) &= X(n\Delta t) + \int_{n\Delta t}^{(n+1)\Delta t} a\big(X(n\Delta)\big)\, ds + \int_{n\Delta t}^{(n+1)\Delta t} b\big(X(n\Delta t)\big)\, dW(s) \\ &= X(n\Delta t) + a\big(X(n\Delta t)\big)\Delta t + b\big(X(n\Delta t)\big)\big(W\big((n+1)\Delta t\big) - W(n\Delta t)\big) \end{aligned}$$

By letting $j = (n+1)\Delta t$, we define a stochastic process $\{X_j\}_{j\geq 0}$ by

$$X_j = X_{j-1} + a(X_{j-1})\Delta t + b(X_{j-1})\big(W(j) - W(j-1)\big) \tag{2.5}$$

for $j = 1, 2, \ldots N$. This is the *Euler-Maruyama method*, which can be used to numerically approximate solutions to (2.1).

2.2 The Milstein method

We can use the same analysis to produce another method, namely the *Milstein method*. The Milstein method can be viewed as the Euler-Maruyama method with a correction added to the stochastic term.

We begin with (2.33) which was derived in the previous section, where $X(t)$, a and b are as before:

$$X(t) = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW(s) + R.$$

We are interested in the last remainder term in (2.4), namely

$$I = \int_{t_0}^{t} \int_{t_0}^{s} L^1 b\big(X(z)\big)\, dW(z)\, dW(s)$$

with $L^1 = b\frac{\partial}{\partial x}$ as before. If the Itô formula is applied to $L^1 b$ in the integrand, we arrive at

$$\begin{aligned} X(t) = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW(s) \\ + L^1 b(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW(z)\, dW(s) + R', \quad t \in [t_0, T], \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} R' = R - I + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{z} L^0\, L^1 b\big(X(u)\big)\, du\, dW(z)\, dW(s) \\ + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{z} L^1\, L^1 b\big(X(u)\big)\, dW(u)\, dW(z)\, dW(s). \end{aligned}$$

We are now interested in evaluating the double stochastic integral

$$\int_{t_0}^{t} \int_{t_0}^{s} dW(z)\, dW(s).$$

First note that

$$\begin{aligned} \int_{t_0}^{t} \int_{t_0}^{s} dW(z)\, dW(s) &= \int_{t_0}^{t} \big(W(s) - W(t_0)\big)\, dW(s) \\ &= \int_{t_0}^{t} W(s)\, dW(s) - \int_{t_0}^{t} W(t_0)\, dW(s) = \int_{t_0}^{t} W(s)\, dW(s) - W(t_0)\big(W(t) - W(t_0)\big). \end{aligned}$$

Let $[t_0, t]$ be divided into n sub-intervals $[t_{i-1}, t_i]$, with $t_i - t_{i-1} = \delta t$ for $i = 0, 1, 2, \ldots n$. Then

$$\int_{t_0}^{t} W(s)\, dW(s) \;\approx\; \sum_{i=0}^{n-1} W(t_i)\big(W(t_{i+1}) - W(t_i)\big).$$

Note that $W(t_i) = \frac{1}{2}\big(W(t_{i+1}) + W(t_i)\big) - \frac{1}{2}\big(W(t_{i+1}) - W(t_i)\big)$, so

$$\sum_{i=0}^{n-1} W(t_i)\big(W(t_{i+1}) - W(t_i)\big) = \sum_{i=0}^{n-1} \Big[\big(\frac{1}{2}(W(t_{i+1}) + W(t_i)\big)\big(W(t_{i+1}) - W(t_i)\big) - \frac{1}{2}\big(W(t_{i+1}) - W(t_i)\big)^2\Big]$$

$$= \frac{1}{2}\sum_{i=0}^{n-1}\Big(W(t_{i+1})^2 - W(t_i)^2\Big) - \frac{1}{2}\sum_{i=0}^{n-1}\Big(W(t_{i+1}) - W(t_i)\Big)^2. \tag{2.7}$$

The first sum on the RHS of the last line of (2.7) is a telescoping sum, which equals $\frac{1}{2}(W(t_{i+1})^2 - W(t_0)^2)$. Write

$$\Delta W(t_i) = W(t_{i+1}) - W(t_i) \sim N(0,\, t_{i+1} - t_i) = N(0,\, \delta t).$$

Now

$$\mathbb{E}\Big[\sum_{i=0}^{n-1}\Big(W(t_{i+1}) - W(t_i)\Big)^2\Big] = \sum_{i=0}^{n-1}\mathbb{E}\Big(\Delta W(t_i)^2\Big)$$

$$= \sum_{i=0}^{n-1}\Delta t = (n-1)\delta t \;\approx\; t - t_0$$

for n large. Also, due to the independence of the Brownian increments,

$$\mathrm{Var}\Big[\sum_{i=0}^{n-1}\Big(\Delta W(t_i)^2\Big)\Big] = \sum_{i=0}^{n-1}\mathrm{Var}\Big(\Delta W(t_i)^2\Big) = \sum_{i=0}^{n-1}\Big[\mathbb{E}\Big(\Delta W(t_i)^4\Big) - \mathbb{E}\Big(\Delta W(t_i)^2\Big)^2\Big]$$

$$= \sum_{i=0}^{n-1}\Big(3\delta t^2 - \delta t^2\Big) = (n-1)2\delta t^2 \;\approx\; 2n\delta t^2$$

for large n. Now $\delta t = \frac{t - t_0}{n}$, so

$$\mathrm{Var}\Big[\sum_{i=0}^{n-1}\Big(\Delta W(t_i)^2\Big)\Big] \;\approx\; \frac{2(t-t_0)^2}{n} \;\longrightarrow\; 0$$

as $n \to \infty$. Hence the second sum in (2.7) is centred at $(t - t_0)$ and its variance tends to 0 as $n \to \infty$. Thus

$$\int_{t_0}^{t} W(s)\, dW(s) = \lim_{n\to\infty}\sum_{i=0}^{n-1} W(t_i)\big(W(t_{i+1}) - W(t_i)\big)$$

$$= \frac{1}{2}\big(W(t)^2 - W(t_0)^2\big) - \frac{1}{2}(t - t_0).$$

So disregarding the remainder term R', (2.6) becomes:

$$X(t) = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW(s) + L^1 b(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW(z)\, dW(s)$$

$$= X_{t_0} + a(X_{t_0})\,(t - t_0) + b(X_{t_0})\big(W(t) - W(t_0)\big)$$

$$+ b(X_{t_0})\, b'(X_{t_0}) \Big(\Big(\frac{1}{2}\big(W(t)^2 - W(t_0)^2\big) - \frac{1}{2}(t - t_0) \Big) - W(t_0)\Big(W(t) - W(t_0)\Big) \Big)$$

$$= X_{t_0} + a(X_{t_0})\,(t - t_0) + b(X_{t_0})\big(W(t) - W(t_0)\big)$$

$$+ \frac{1}{2}\, b(X_{t_0})\, b'(X_{t_0}) \Big(W(t)^2 - W(t_0)^2 - (t - t_0) - 2W(t_0)W(t) + 2W(t_0)^2 \Big)$$

$$= X_{t_0} + a(X_{t_0})\,(t - t_0) + b(X_{t_0})\big(W(t) - W(t_0)\big) + \frac{1}{2}\, b(X_{t_0})\, b'(X_{t_0}) \Big(\big(W(t) - W(t_0)\big)^2 - (t - t_0) \Big).$$

Similarly to above, if $[t_0,\, T]$ is discretized into N sub-intervals of length Δt and by letting $j = (n+1)\Delta t$, we can define $\{X_j\}_{j \geq 0}$ by

$$\begin{aligned} X_j = X_{j-1} + a(X_{j-1})\Delta t + b(X_{j-1})\big(W(j) - W(j-1)\big) \\ + \frac{1}{2} b(X_{j-1}) b'(X_{j-1}) \big((W(j) - W(j-1))^2 - \Delta t \big). \end{aligned} \tag{2.8}$$

This stochastic process is known as the Milstein method. We will see in the following section that the Milstein method outperforms the Euler-Maruyama method in terms of order of strong covergence, due to the stochastic correction.

2.3 Convergence of numerical methods

In this section, we will be discussing both the Milstein and the Euler-Maruyama methods in the context of solving the following stochastic differential equation:

$$dX(t) = \lambda X(t) dt + \mu X(t) dW(t), \quad X(0) = X_0, \tag{2.9}$$

where $t \in [0, T]$ and $\lambda, \mu \in \mathbb{R}$. This equation appears in financial mathematics as an equation governing the price of a risky asset within the Black-Scholes framework. It has solution given by

$$X(t) = X_0 \, e^{(\lambda - \frac{1}{2}\mu^2)t + \mu W(t)}. \tag{2.10}$$

Both these methods can be used to numerically approximate this solution. The method matches the true solution more closely as $\Delta t = T/N$ is decreased, where N is the number of intervals that $[0, T]$ is subdivided into. Therefore, convergence seems to take place. This is illustrated in Fig. 1 - the solution (2.10) plotted over $[0, 1]$, with stepsize δt, along with the Milstein approximation for decreasing stepsizes $12\delta t$, $6\delta t$ and $2\delta t$. It is clear from the plot that the smallest of these stepsizes corresponds to the best approximation. We now give some formal definitions of various types of convergence.

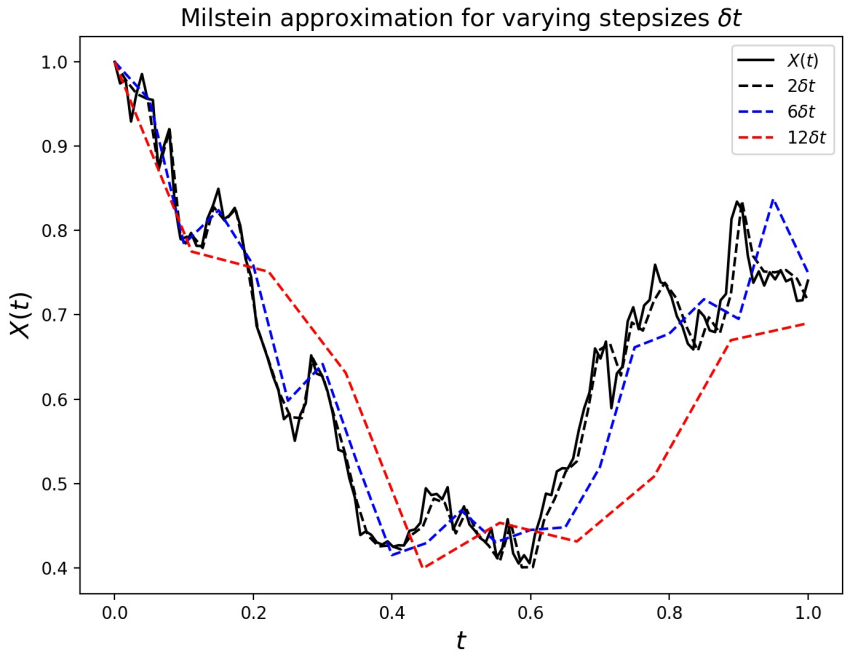


Fig. 1: Visualising the convergence of the Milstein method as δt is decreased. The bold black line represents the trajectory of the solution $X(t)$ for a stepsize of δt as given by (2.10) with parameters $\lambda = -1$ and $\mu = 0.5$, while the black, blue and red dashed lines represent the Milstein approximation to $X(t)$ for stepsizes $2\delta t$, $6\delta t$ and $12\delta t$ respectively.

Strong Convergence. A numerical method is said to have *strong order of convergence in* $\mathcal{L}_1$ equal

to γ if there exists some constant C such that

$$\mathbb{E}\big[\,|X_n - X(h)|\,\big] \leq C\Delta t^{\gamma}, \tag{2.11}$$

for any fixed $h = n\Delta t$.

The strong order of convergence measures the rate at which the "mean of the error" decays as $\Delta t \longrightarrow 0$. An alternative type of convergence is to measure the decay of the "error of the means", which leads to the concept of *weak convergence*.

Weak Convergence. A method is said to have weak order of convergence equal to γ if there exists a constant C such that for all functions p in some class,

$$\big|\mathbb{E}[\,p(X_n)] \,- \mathbb{E}[\,p(X(\tau))]\big| \leq C\Delta t^{\gamma}, \tag{2.12}$$

for any fixed $\tau = n\Delta t \in [0, T]$ and Δt sufficiently small. The class of functions p usually satisfy certain smoothness and polynomial growth conditions, but for the following analysis we will consider the case where p is the identity function.

2.3.1 Convergence of the Euler-Maruyama method

To test the strong convergence of the Euler-Maruyama method in the context of solving equation (2.21), consider the error of the method at the endpoint $t = T$, that is let

$$e_{\Delta t}^{strong} \coloneqq \mathbb{E}\big[\,|X_N - X(T)|\,\big]\,, \tag{2.13}$$

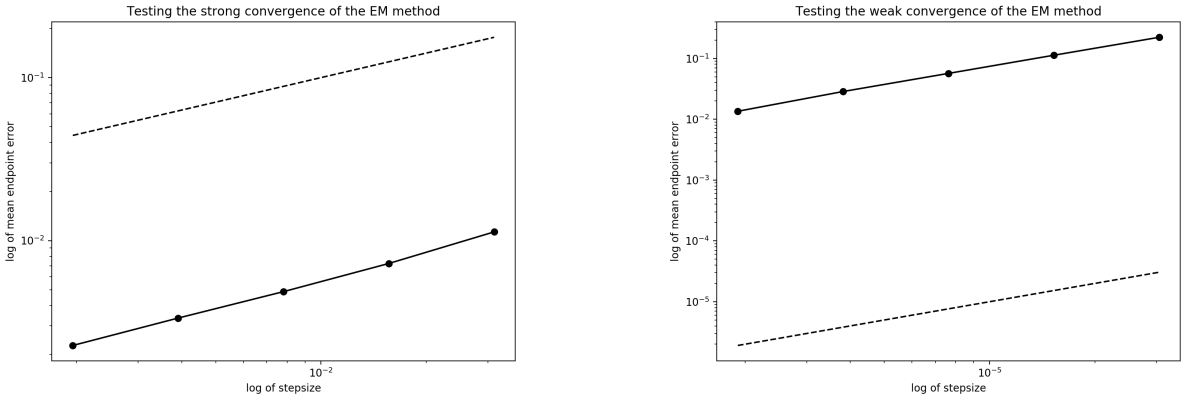
where $N\Delta t = T$, denote the endpoint error in the strong sense. If the bound given in (2.11) holds for all values of $n\Delta t$, then it must hold at the endpoint. Hence we have that

$$e_{\Delta t}^{strong} \leq C\Delta t^{\gamma} \tag{2.14}$$

for sufficiently small Δt. Assuming that (2.14) holds with approximate equality and taking logs, have that

$$\log e_{\Delta t}^{strong} \approx \log C + \gamma \log \Delta t. \tag{2.15}$$

It is known that the EM method has strong order of convergence of $1/2$, see [1]. In order to demonstrate this, $10,000$ Brownian paths are computed over the interval $[0,\, 1]$ with stepsize $\delta t = 2^{-9}$. The true solution to (2.21) is calculated using values of $\lambda = -1$, $\mu = 0.5$ and $X_0 = 1$. For each path, the Euler-Maruyama method is applied using 5 different stepsizes: $\Delta t = 2^{p-1}\delta t$ for $1 \leq p \leq 5$. The endpoint error $|X_N - X(T)|$ for each path is stored, and the mean of this error is taken over all $10,000$ paths for each value of p. The endpoint error $e_{\Delta t}^{strong}$ is plotted against the stepsize Δt on a log-log scale in Fig. 2 (a). A dashed line of slope $1/2$ is also plotted. The two lines appear to be parallel, indicating that



(a) Strong convergence

(b) Weak convergence

Fig. 2: Convergence of the Euler-Maruyama method

(2.15) holds with $\gamma = 1/2$, that is the *strong* order of convergence of the Euler-Maruyama method is equal to $1/2$. A similar analysis can be performed to assess the weak convergence of the EM method. Let

$$e_{\Delta t}^{weak} \coloneqq |\, \mathbb{E}[X_L] - \mathbb{E}[X(T)]\, | \tag{2.16}$$

denote the endpoint error of the Euler-Maruyama method in the *weak* sense. If the bound given in (2.12) holds with approximate equality, again we have

$$\log e_{\Delta t}^{weak} \approx \log C + \gamma \log \Delta t. \tag{2.17}$$

It is known that the EM method has weak order of convergence equal to 1. In order to verify this, $100,000$ Brownian paths are sampled and the EM method is applied with parameters $\lambda = 2$, $\mu = 0.1$ and $X_0 = 1$ for 5 different stepsizes: $\Delta t = 2^{p-10}$ for $1 \leq p \leq 5$. It follows from (2.22) that $\mathbb{E}[X(T)] = e^{\lambda T}$ for the true solution. The corresponding weak endpoint errors $e^{weak}_{\Delta t}$ are plotted against Δt on a log-log scale in Fig. 2 (b). A dashed line of slope 1 is added for reference. Again we see that both lines appear to be parallel, indicating that (2.17) holds with $\gamma = 1$.

2.3.2 Strong Convergence of the Milstein method

We now investigate the strong order of convergence of the Milstein method applied to (2.21). The method is applied to $10,000$ Brownian paths with $\delta t = 2^{-11}$ for stepsizes of $128\delta t$, $64\delta t$, $32\delta t$ and $16\delta t$ over the interval [0, 1]. Similarly to the Euler-Maruyama method, the mean of the endpoint error between the true solution (2.22) and the approximations is plotted against the varying stepsizes Δt on a log-log scale in Fig. 3. As expected, the slope of this line is much sharper than the corresponding

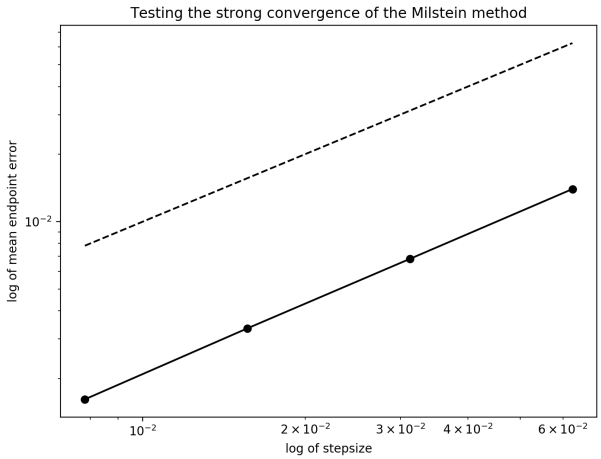


Fig. 3: Strong convergence of the Milstein method

line for the Euler-Maruyama method. A dashed line of slope 1 is also plotted, and the two appear to be parallel, indicating the bound in (2.15) holds with $\gamma = 1$ for the Milstein method.

2.4 Stability of numerical methods

We are also interested in the long-term behaviour or *stability* of an equilibrium solution to a stochastic differential equation given by

$$dX(t) = f\big(X(t)\big)dt + g\big(X(t)\big)\,dW(t), \quad f(0) = g(0) = 0,\ X(0) = X_0, \tag{2.18}$$

and consequently its approximation by a particular numerical method. This stability analysis is important, as a method which is unstable may not converge over time, meaning any application in a Monte Carlo context will be unreliable. We will consider two types of stability, *mean-square stability* and *almost-sure stability*.

1. Mean-square stability. The equilibrium solution $X(t) \equiv 0$ to (2.18) is said to stable in mean-square if

$$\lim_{t\to\infty} \mathbb{E}[\,|X(t)|^2\,] = 0. \tag{2.19}$$

2. Almost-sure stability. The equilibrium solution $X(t) \equiv 0$ to (2.18) is said to be stable almost-surely if

$$\lim_{t\to\infty} |X(t)| = 0 \ \text{ with probability 1.} \tag{2.20}$$

We will be discussing the stability of these methods in the context of solving the following stochastic differential equation:

$$dX(t) = \lambda X(t)dt + \mu X(t)dW(t), \quad X(0) = X_0, \tag{2.21}$$

where $t \in [0,\, T]$ and $\lambda,\, \mu \in \mathbb{C}$. Note that the drift and diffusion coefficients are now allowed to be complex. The solution $X(t)$ is given by

$$X(t) = X_0\, e^{(\lambda - \frac{1}{2}\mu^2)t + \mu W(t)}. \tag{2.22}$$

We first examine the mean-square stability of the equilibrium of zero of (2.18). Letting $W(t) = \sqrt{t}Z$, where $Z \sim N(0,\,1)$, we can write

$$\begin{aligned} X(t)^2 &= X_0^2\, e^{2(\lambda - \frac{1}{2}\mu^2)t + 2\mu\sqrt{t}Z} \\ &= X_0^2\, e^{2\left(\lambda - \frac{1}{2}(\Re(\mu)^2 - 2\Re(\mu)\Im(\mu)i + \Im(\mu)^2)\right)t + 2\mu\sqrt{t}Z}. \end{aligned}$$

It follows from $|e^z| = e^{\Re(z)}$ and the properties of the complex modulus that

$$\begin{aligned} |X(t)^2| &= |X_0|^2\, e^{\Re\left(2\left(\lambda - \frac{1}{2}(\Re(\mu)^2 - 2\Re(\mu)\Im(\mu)i + \mu(\mu)^2)\right)t + 2\mu\sqrt{t}Z\right)} \\ &= |X_0|^2\, e^{\left(2\Re(\lambda) - \Re(\mu)^2 + \Im(\mu)^2\right)t + 2\Re(\mu)\sqrt{t}Z}. \end{aligned}$$

Note that $\left(2\Re(\lambda) - \Re(\mu)^2 + \Im(\mu)^2\right)t + 2\Re(\mu)\sqrt{t}Z$ is normally distributed, hence

$$e^{(2\Re(\lambda) - \Re(\mu)^2 + \Im(\mu)^2)t + 2\Re(\mu)\sqrt{t}Z} \sim LN\big((2\Re(\lambda) - \Re(\mu)^2 + \Im(\mu)^2)t,\ 4\Re(\mu)^2 t\big).$$

If $X \sim LN(\mu,\, \sigma^2)$, then $\mathbb{E}[\,X\,] = e^{\mu + \frac{1}{2}\sigma^2}$. Therefore, we have that

$$\begin{aligned} \mathbb{E}[\,|X(t)^2|\,] &= |X_0|^2\, e^{(2\Re(\lambda) - \Re(\mu)^2 + \Im(\mu)^2)t + \frac{1}{2}4\Re(\mu)^2 t} \\ &= |X_0|^2\, e^{(2\Re(\lambda) + \Re(\mu)^2 + \Im(\mu)^2)t} \\ &= |X_0|^2\, e^{(2\Re(\lambda) + |\mu|^2)t}. \end{aligned}$$

Hence, it follows that

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}[\,|X(t)^2|\,] = 0 &\iff 2\Re(\lambda) + |\mu|^2 < 0 \\ &\iff \Re(\lambda) + \frac{1}{2}|\mu|^2 < 0. \end{aligned}$$

Now the almost sure stability of the equilibrium at zero of (2.21) can be studied by considering the

the behaviour of its Lyapunov exponent as $t \to \infty$ as follows:

$$\lim_{t\to\infty} \frac{1}{t} \log|X(t)| = \lim_{t\to\infty} \frac{1}{t} \log\left(|X_0\, e^{(\lambda - \frac{1}{2}\mu^2)t + \mu W(t)}| \right)$$

$$= \lim_{t\to\infty} \frac{\log|X_0|}{t} + \lim_{t\to\infty} \frac{1}{t} \Re(\lambda - \frac{1}{2}\mu^2)t + \lim_{t\to\infty} \Re(\mu)\frac{W(t)}{t}.$$

By the strong law of large numbers,

$$\lim_{t\to\infty} \frac{W(t)}{t} = 0 \text{ a.s.}$$

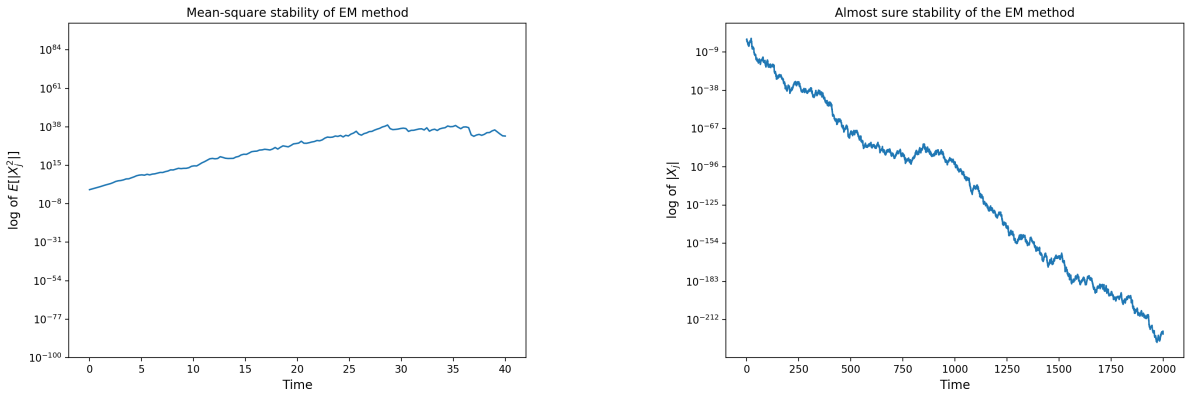
Hence,

$$\lim_{t\to\infty} \frac{1}{t} \log|X(t)| = \Re(\lambda - \frac{1}{2}\mu^2).$$

The solution $X(t)$ will be stable almost surely if the Lyapunov exponent is negative, that is

$$\lim_{t\to\infty} |X(t)| = 0 \iff \lim_{t\to\infty} \frac{1}{t} \log|X(t)| < 0 \iff \Re(\lambda - \frac{1}{2}\mu^2) < 0.$$

Hence it is clear that an equilibrium that is mean-square stable will automatically be stable almost-surely, but the converse is not necessarily true - an equilibrium can be mean-square unstable but stable almost-surely. This is illustrated in Fig. 4.



(a) Mean-square instability

(b) Almost sure stability

Fig. 4: Trajectories of (2.22) with parameters $\lambda = 0.5$, $\mu = \sqrt{6}$ and $\Delta t = 0.25$.

2.4.1 Stability of the Euler-Maruyama method

We now derive mean-square and almost sure stability conditions for the the Euler-Maruyama method applied to (2.21) which is given by

$$X_j = X_{j-1} + \lambda \Delta t X_{j-1} + \mu X_{j-1}\big(W(j) - W(j-1)\big),$$

where $Z_j \sim N(0,1)$ for $j = 1, 2, \ldots N$, and N is the number of sub-intervals of length Δt that $[0,\, T]$ is divided into. Thus $W(j) - W(j-1) \sim N(0, \Delta t)$, so we can write

$$X_j = X_{j-1} + \lambda \Delta t X_{j-1} + \mu \sqrt{\Delta t} X_{j-1} Z_{j-1}.$$

Let $x = 1 + \lambda \Delta t$ and $y = \mu\sqrt{\Delta t}$. Then

$$\begin{aligned}
|X_j| =& |X_{j-1}|\,|\,x + yZ_{j-1}\,| \\
\implies |X_j|^2 = |X_{j-1}|^2\,|\,x + yZ_{j-1}\,|^2 &= |X_{j-1}|^2 (x + yZ_{j-1})(\overline{x + yZ_{j-1}}) \\
=& |X_{j-1}|^2 (x + yZ_{j-1})(\overline{x} + \overline{y} Z_{j-1}) = |X_{j-1}|^2 \big(|x|^2 + (x\overline{y} + \overline{x}y) Z_{j-1} + |y|^2 Z_{j-1}^2\big),
\end{aligned}$$

where $\{Z_j\}_{j \geq 0}$ is a sequence of independent standard normal random variables. Note that X_{j-1} and Z_{j-1} are independent, so we have that

$$\begin{aligned}
\mathbb{E}[\,|X_j|^2\,] &= \mathbb{E}\Big[|X_{j-1}|^2 \big(|x|^2 + (x\overline{y} + \overline{x}y) Z_{j-1} + |y|^2 Z_{j-1}^2\big)\Big] \\
&= \big(|x|^2 + |y|^2\big)\mathbb{E}[\,|X_{j-1}|^2\,].
\end{aligned}$$

Now $\mathbb{E}[\,|X_j|^2\,]$ will satisfy (2.19) if the ratio between $\mathbb{E}[\,|X_j|^2\,]$ and $\mathbb{E}[\,|X_{j-1}|^2\,]$ is less than 1, that is if

$$|x|^2 + |y|^2 < 1 \iff |1 + \lambda \Delta t|^2 + |\mu\sqrt{\Delta t}|^2 < 1.$$

Therefore,

$$\lim_{j\to\infty} \mathbb{E}[\,|X_j|^2\,] = 0 \iff |1+\lambda\Delta t|^2 + |\mu|^2\Delta t < 1. \tag{2.23}$$

For almost-sure stability, let $a = 1+\lambda\Delta t$ and $b = \mu\sqrt{\Delta t}$. Then

$$X_j = X_{j-1}(a+bZ_{j-1}) = X_{j-2}(a+bZ_{j-2})(a+bZ_{j-1})$$

$$\cdots = X_0 \prod_{i=0}^{j-1} (a+bZ_i),$$

where each $Z_i \sim N(0,\,1)$ and are independent. Now taking logs, have

$$\log|X_j| = \log|X_0| + \sum_{i=0}^{j-1} \log|a+bZ_i|.$$

Next, by the strong law of large numbers,

$$\lim_{j\to\infty} \frac{1}{j-1} \sum_{i=0}^{j-1} \log|a+bZ_i| = \mathbb{E}[\,\log|a+bZ|\,]$$

$$\iff \sum_{i=0}^{j-1} \log|a+bZ_i| = \lim_{j\to\infty} (j-1)\mathbb{E}[\,\log|a+bZ|\,].$$

So

$$\lim_{j\to\infty} \log|X_j| = \begin{cases} +\infty, \text{ if } \mathbb{E}[\,\log|a+bZ|\,] > 0, \\ \\ -\infty, \text{ if } \mathbb{E}[\,\log|a+bZ|\,] < 0. \end{cases}$$

We require this limit to tend to $-\infty$ because

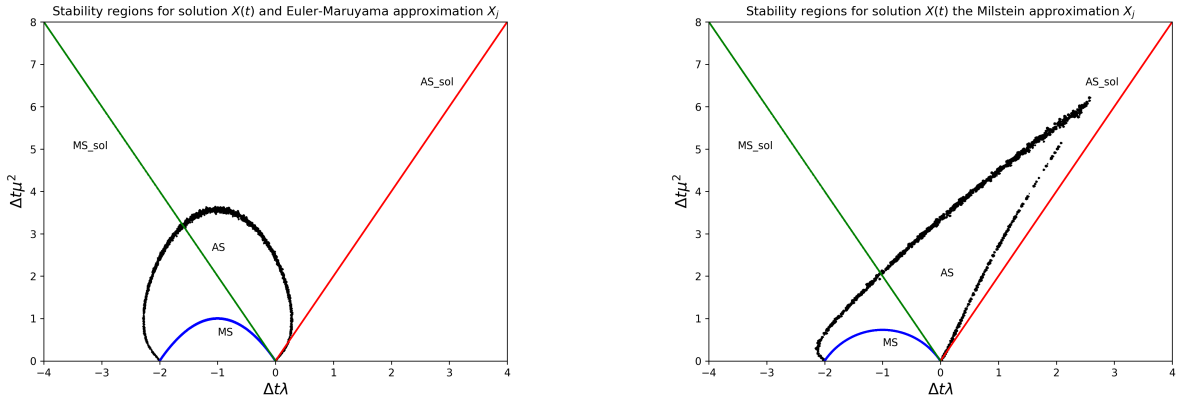
$$\lim_{j\to\infty} \log|X_j| = -\infty \iff \lim_{j\to\infty} |X_j| = 0,$$

as required. Therefore the Euler-Maruyama approximation X_j will be stable *almost surely* if, for $Z \sim N(0,\,1)$, we have

$$\mathbb{E}[\log|1 + \lambda\Delta t + \mu\sqrt{\Delta t}Z|] < 0. \tag{2.24}$$

Fig. 5 (a) aims to visualise the range of values for $\lambda\Delta t$ and $\mu^2\Delta t$ for which the Euler-Maruyama method is stable in both senses. While the stability condition (2.23) corresponds to a simple parabola of the form $-x(x+2)$, the expectation in (2.24) does not appear to have a closed form solution accorrding to current literature. Therefore the almost-sure stability region in Fig. 5 was generated by taking a Monte Carlo approximation of $\mathbb{E}[\log|1 + \lambda\Delta t + \mu\sqrt{\Delta t}Z|]$ for varying values of $\lambda\Delta t$ and $\mu^2\Delta t$. These give a surface in $\mathbb{R}^3$, which was reduced to the plot in Fig. 5 by selecting values of which were within a certain tolerance ϵ of 0. These values were then plotted on a contour plot in Fig. 5 (a), resulting in the egg-shape region labelled 'AS_EM'. The sharpness of the boundary of this area is determined by the sampling error of the Monte Carlo estimate of the expecation, as well as the value of the tolerance ϵ. The mean-square stability region is plotted in blue labelled 'MS_EM'.

It is interesting to note that the mean-square stability region is contained entirely within the almost-sure stability region, meaning that any mean-square method will automatically be almost-surely stable, with the converse being false. This is analogous to the result we saw for the true solution. We have also provided the regions where the equilibrium at 0 of (2.21) is stable in mean-square and almost surely for comparison.



(a) Stability regions of the Euler-Maruyama method

(b) Stability regions of Euler-Milstein method

Fig. 5: Almost sure and mean-square stability regions. The areas to the left of the red and green lines represent the areas where the equilibrium $X(t) \equiv 0$ to (2.21) is stable almost surely and in mean-square respectively.

2.4.2 Stability of the Euler-Milstein method

A similar analysis is now presented for the Euler-Milstein method. Let X_j denote the Euler-Milstein approximation of (2.21) at the j^{th} timestep, that is

$$\begin{aligned} X_j &= X_{j-1} + \lambda\Delta t X_{j-1} + \mu X_{j-1}\big(W(j) - W(j-1)\big) + \frac{1}{2}\mu^2 X_{j-1}\Big(\big(W(j) - W(j-1)\big)^2 - \Delta t\Big) \\ &= X_{j-1} + \lambda\Delta t X_{j-1} + \mu\sqrt{\Delta t} X_{j-1} Z_{j-1} + \frac{1}{2}\mu^2 \Delta t X_{j-1}(Z_{j-1}^2 - 1) \end{aligned} \tag{2.25}$$

where $Z_j \sim N(0,1)$ for $j = 1, 2, \ldots N$. Then it can be shown that

$$\lim_{j\to\infty} \mathbb{E}[\,|X_j|^2\,] = 0 \iff |1 + \lambda\Delta t|^2 + |\mu|^2\Delta t + \frac{1}{2}|\mu|^4\Delta t^2 < 1. \tag{2.26}$$

and

$$\lim_{j\to\infty} |X_j| = 0 \iff \mathbb{E}\Big[\log\big|1 + \lambda\Delta t + \mu\sqrt{\Delta t}Z + \frac{1}{2}\mu^2\Delta t(Z^2 - 1)\big|\Big] < 0. \tag{2.27}$$

We now present a stability region for the Euler-Milstein method in Fig. 5 (b), which was generated in a similar sense to Fig. 5 (a) using (2.27). Here we see that the mean-square stability region for the Euler-Milstein method is contained entirely within the almost sure stability region, similar to what was seen for the EM method. The regions generated in Fig. 5 can be used to determine easily whether a method will be stable in a particular sense by inspecting the values of λ, μ and Δt of the method. This will be useful when applying these methods to asset price models in the Monte Carlo valuation of ESOs in the Section 3.

2.5 Euler-Milstein method applied to multi-dimensional noise

We now consider the construction of a Milstein-type method for a stochastic process containing more than one stochastic perturbation. Consider the stochastic differential equation given by

$$dX(t) = \lambda X(t)dt + \mu_1 X(t)dW_1(t) + \mu_2 X(t)dW_2(t), \quad X(t_0) = X_{t_0}, \tag{2.28}$$

where λ, μ_1 and μ_2 are constants, $t \in [t_0, T]$. The Wiener processes W_1 and W_2 are assumed to be independent. This equation has solution

$$X(t) = X_{t_0} e^{(\lambda - \frac{\mu^2}{2})t + \mu_1 W_1(t) + \mu_2 W_2(t)} \tag{2.29}$$

for $t \in [t_0, T]$. We will now construct a numerical method X_j, which will be used to approximate $X(t)$ on $[t_0, T]$. We first assume for simplicity that $\mu_1 = \mu_2 = \mu$, so that

$$dX(t) = \lambda X(t)dt + \mu X(t)\big(dW_1(t) + dW_2(t)\big), \quad X(t_0) = X_{t_0},\ t \in [t_0, T], \tag{2.30}$$

which can be rewritten in integral form as

$$X(t) = X_{t_0} + \int_{t_0}^{t} \lambda X(s)\, ds + \int_{t_0}^{t} \mu X(s)\, dW_1(s) + \int_{t_0}^{t} \mu X(s)\, dW_2(s),\ t \in [t_0, T]. \tag{2.31}$$

We now apply the Itô formula for two-dimensional noise - for any continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$, we have that

$$\begin{aligned} f\big(X(t)\big) = \ & f(X_{t_0}) + \int_{t_0}^{t} \left(\lambda X(s) \frac{\partial}{\partial x} f(X(s)) + \frac{\mu^2}{2} X(s)^2 \frac{\partial^2}{\partial^2 x} f(X(s)) \right) ds \\ & + \int_{t_0}^{t} \mu X(s) \frac{\partial}{\partial x} f(X(s)) dW_1(s) + \int_{t_0}^{t} \mu X(s) \frac{\partial}{\partial x} f(X(s)) dW_2(s) \\ & = f(X_{t_0}) + \int_{t_0}^{t} L^0 f(X(s)) ds + \int_{t_0}^{t} L^1 f(X(s)) dW_1(s) + \int_{t_0}^{t} L^1 f(X(s)) dW_2(s) \end{aligned} \tag{2.32}$$

for $t \in [t_0, T]$. The operators

$$L^0 f(x) = \lambda x \frac{\partial}{\partial x} f(x) + \frac{\mu^2}{2} x^2 \frac{\partial^2}{\partial^2 x} f(x),$$

$$L^1 f(x) = \mu x \frac{\partial}{\partial x} f(x)$$

have been introduced here. For $f(x) \equiv x$, have $L^0 f(x) = \lambda x$ and $L^1 f(x) = \mu x$, in which case (2.32)

reduces to (2.31), that is to

$$X(t) = X_{t_0} + \int_{t_0}^{t} \lambda X(s)\, ds + \int_{t_0}^{t} \mu X(s)\, dW_1(s) + \int_{t_0}^{t} \mu X(s)\, dW_2(s), \quad t \in [t_0, T].$$

If the Itô formula is then applied to the functions $f = \lambda X(t)$ and $f = \mu X(t)$, we have that

$$X(t) = X_{t_0} + \lambda X_{t_0} \int_{t_0}^{t} ds + \mu X_{t_0} \int_{t_0}^{t} dW_1(s) + \mu X_{t_0} \int_{t_0}^{t} dW_2(s) + R \tag{2.33}$$

where

$$\begin{aligned} R = &\int_{t_0}^{t}\int_{t_0}^{s} L^0 \lambda X(z)\, dz\, ds + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \lambda X(z)\, dW_1(z)\, ds + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \lambda X(z)\, dW_2(z)\, ds \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0 \mu X(z)\, dz\, dW_1(s) + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \mu X(z)\, dW_1(s)\, dW_1(s) + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \mu X(z)\, dW_2(s)\, dW_1(s) \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0 \mu X(z)\, dz\, dW_2(s) + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \mu X(z)\, dW_1(s)\, dW_2(s) + \int_{t_0}^{t}\int_{t_0}^{s} L^1 \mu X(z)\, dW_2(s)\, dW_2(s). \end{aligned}$$

Now we can expand the four integrals which contain two Brownian differentials in this remainder using L^0 and L^1, resulting in the following expression

$$\begin{aligned} X(t) = &X_{t_0} + \lambda X_{t_0} \int_{t_0}^{t} ds + \mu X_{t_0} \int_{t_0}^{t} dW_1(s) + \mu X_{t_0} \int_{t_0}^{t} dW_2(s) \\ &+ L^1 \mu X_{t_0} \Big(\int_{t_0}^{t}\int_{t_0}^{s} dW_1(s)\, dW_1(s) + \int_{t_0}^{t}\int_{t_0}^{s} dW_2(s)\, dW_1(s) \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} dW_1(s)\, dW_2(s) + \int_{t_0}^{t}\int_{t_0}^{s} dW_2(s)\, dW_2(s) \Big) + R'. \end{aligned} \tag{2.34}$$

Here R' contains the five unused terms in R, as well as the final two terms in each of the four Itô expansions of $L^1 \mu X(z)$. Now we have already shown that

$$\int_{t_0}^{t}\int_{t_0}^{s} dW_i(s)\, dW_i(s) = \frac{1}{2}\big((W_i(t) - W_i(t_0))^2 - (t - t_0)\big) \tag{2.35}$$

for $i = 1, 2$. We now must compute

$$I_{t,\,t_0}(i, j) = \int_{t_0}^{t}\int_{t_0}^{s} dW_i(s)\, dW_j(s)$$

where $i \neq j$. Integrals of this form are difficult to simulate directly - for this particular example, we rely on the following symmetry relation, see [4]

$$I_{t,\,t_0}(i,j) + I_{t,\,t_0}(j,i) = \big(W_i(t) - W_i(t_0)\big)\big(W_j(t) - W_j(t_0)\big).$$

From this and (2.35), we can simplify (2.34) as follows - using $L^1\mu X(t) = \mu^2 X(t)$ and disregarding R', we get

$$\begin{aligned}
X(t) &= X_{t_0} + \lambda X_{t_0}(t-t_0) + \mu X_{t_0}\big(W_1(t) - W_1(t_0)\big) + \mu X_{t_0}\big(W_2(t) - W_2(t_0)\big) \\
&\quad + \frac{\mu^2}{2} X_{t_0}^2 \Big[\big(W_1(t) - W_1(t_0)\big)^2 - (t-t_0) + \big(W_2(t) - W_2(t_0)\big)^2 - (t-t_0) \\
&\quad + 2\big(W_1(t) - W_1(t_0)\big)\big(W_2(t) - W_2(t_0)\big)\Big] \\
&= X_{t_0} + \lambda X_{t_0}(t-t_0) + \mu X_{t_0}\sqrt{t-t_0}Z + \mu X_{t_0}\sqrt{t-t_0}Z \\
&\quad + \frac{\mu^2}{2} X_{t_0}^2 \Big[\Big(\big(W_1(t) - W_1(t_0)\big) + \big(W_2(t) - W_2(t_0)\big)\Big)^2 - 2(t-t_0)\Big] \\
&= X_{t_0} + \lambda X_{t_0}(t-t_0) + \mu X_{t_0}\sqrt{t-t_0}(Z_1+Z_2) + \frac{\mu^2}{2}(t-t_0)\, X_{t_0}^2\,\big((Z_1+Z_2)^2 - 2\big),
\end{aligned}$$

where Z_1 and Z_2 are two independent standard Normal random variables. By discretizing the interval $[t_0,\, T]$ into N sub-intervals $[t_{i-1},\, t_i]$ each of length Δt, and letting $j = i\Delta t$ we can define a stochastic process $\{X_j\}_{j\geq 0}$ for $j = 1, 2, \ldots N$ given by

$$X_j = X_{j-1} + \lambda X_{j-1}\Delta t + \mu X_{j-1}\sqrt{\Delta t}(Z_1+Z_2) + \frac{\mu^2}{2}\Delta t\, X_{j-1}\,\big((Z_1+Z_2)^2 - 2\big) \tag{2.36}$$

which approximates the solution $X(t)$ given in (2.29) at timesteps $j = i\Delta t$, $i = 1, 2, \ldots N$. This is the Euler-Milstein method applied to a system with two Brownian perturbations. It should be noted here that this form is based on the assumption that the diffusion coefficients of *both* perturbations are equal. This assumption allows this method to be easily extended to a system with p stochastic noise componenets. The stability analysis carried out in the previous sections can also be applied to this form of the Milstein method - Fig. 7 shows how the almost sure stability region of the method changes as the number of perturbations is increased.

What we see in Fig. 7 makes intutive sense; as we increase the number of noise terms, the method becomes more sensitive to these perturbations, which is why we see the regions being 'dragged' downwards. As the number of noise terms is increased, we must reduce μ (assuming that Δt remains unchanged), the diffusion coefficient of the process, in order to maintain stability.

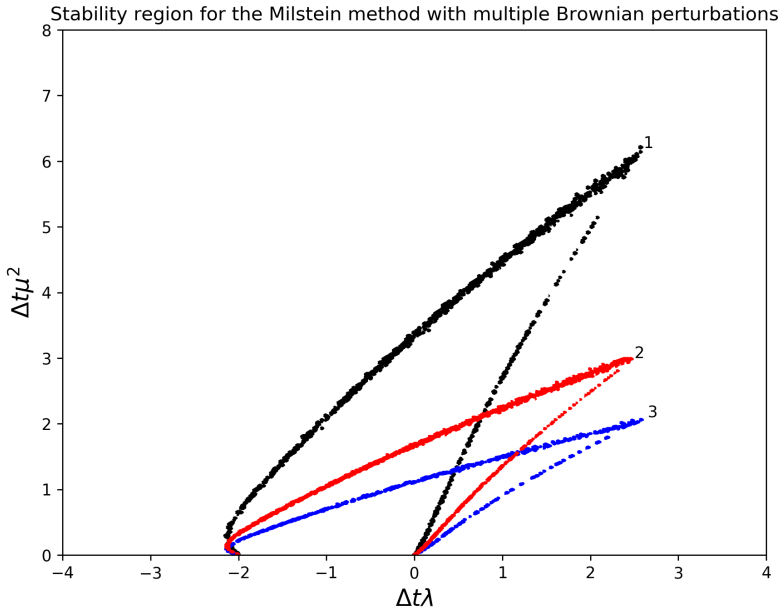


Fig. 6: Stability regions of the Milstein method with 1, 2 and 3 stochastic perturbations

3 Monte Carlo estimation

In this section, we will use Monte Carlo methods to value employee stock options. This approach has some useful advantages - for example, the underlying asset price can be modified from a standard Black-Scholes geometric Brownian motion to a more complicated model that captures more realistic features of asset prices. The Black-Scholes option pricing formula, which is recommended as a means of valuing these options by some professional accounting bodies, relies on the assumption that the price of the underlying share S_t satisfies the following stochastic differential equation:

$$dS(t) = rS(t)dt + \sigma S(t)dB(t), \quad S(0) = S_0.$$

This pricing model is quite limiting. For example, it does not allow for non-constant volatility, that is the parameter σ is a fixed constant for all $t \in [0, T]$. Because the Monte Carlo approach only requires simulated asset trajectories to value these options, the complexity of the asset price model can be increased. One can generate trajectories of such an asset price, or in the case where it is impractical to simulate or if there is no closed form solution, apply one of the numerical methods derived in Section 2 to approximate the asset price.

The Monte Carlo estimate we obtain can be compared to other analytic valuations of the ESO. In particular, we will be comparing the estimate obtained to the corresponding analytic price given in [2], as well as comparing it to a Black-Scholes valuation with a reduced time to expiry. Finally, we will consider some variance reduction techniques in the Monte Carlo approximation - in particular we aim to use the Black-Scholes valuation with reduced time to maturity as a control variate to reduce the sampling error of the estimate.

All the Monte Carlo pricing has been written in Python code, and has been uploaded to [7].

3.1 Monte Carlo approximation of a standard European call option

We first present an intuition behind valuing a standard European call option by Monte Carlo methods. The valuing of an ESO can be viewed as an extension of this approach. Consider the case where the price of the underlying asset $S(t)$ is given as the solution to

$$dS(t) = rS(t)dt + \sigma S(t)dB(t), \quad S(0) = S_0,$$

where r is the risk-free rate, σ the volatility of the asset price. Suppose the option expires at time T, therefore the value of the option at time t is given by

$$C_t = e^{-r(T-t)}\mathbb{E}\left[\left(S(T) - K\right)^+\right],$$

where K is the strike price of the option. This expecation can be approximated by generating M realisations of $S(T)$, and taking the average of their payoffs. This is done as follows:

- Assume without loss of generality that $t = 0$. Begin by discretizing the interval $[0, T]$ into N intervals of length $\delta t = T/N$. Therefore the asset price will be calculated at time points $j = n\delta t$, for $n = 1, 2, \ldots N$.
- Generate a Brownian path by sampling N times from a $N(0, \delta t)$ random variable.
- Over each small interval, the asset price will evolve according to

  $$S(j) = S(j-1)e^{(r-\frac{\sigma}{2})\delta t + \sigma Z_{j-1}},$$

  where Z_{j-1} is the $j - 1^{th}$ Brownian increment, and r and σ are the parameters of the asset price model.
- When the asset reaches expiry, i.e. $j = N$, the payoff will be computed as C_T.

This is repeated M times and an average is taken over all M payoffs to give a Monte Carlo approximation of C_0, while quoting the sampling error and a 95% confidence level.

It should be noted that this approach could be viewed as computationally inexpensive, as we only require the share price at expiry to value the standard European call option. Therefore it may seem unecessary to discretize $[0, T]$ and generate the entire path of the asset price - however the value of the ESO is path-dependant, so with this in mind we simulate the entire asset trajectories in order to easily modify this method to value the ESO in the next section.

3.2 Incorporating a departure process

We now consider the simplest variant of the employee stock option, an intensity-based model for departure. Assume there is no vesting period, and that the employee exercises the option either at expiry

T, or in the case that they leave the company at some point during the interval $[0, T]$, i.e. there is no voluntary early exercise feature. We will use a Poisson process with intensity λ to capture this early exercise possibility.

A *Poisson process* is a stochastic process $\{N_t\}_{t\geq 0}$ with *intensity* parameter λ that satisfes

- $N_0 = 0$,
- $\{N_t\}_{t\geq 0}$ has independent increments,
- $N_t - N_s \sim Poi\big(\lambda(t-s)\big)$ for $s < t$.

It is clear from the above properties that the probability of the employee leaving the company during an interval of length δt can be simulated by sampling from a $Poi(\lambda\delta t)$ random variable, and this probability will be independent of the associated probability in the previous time interval.

This can be easily incorporated into the Monte Carlo simulation as follows:

- Discretize the interval $[0, T]$ as above.
- Allow the asset price to evolve from $S(j-1)$ to $S(j)$.
- Simultaneously sample from a $Poi(\lambda\delta t)$ random variable - if this returns a count greater than 0, the employee has left the company during this interval, and so the time of departure j and the asset price $S(j)$ are saved and used to compute C_j.
- This process will continue until either the employee leaves the company, or if the option expires at time T.

The authors in [2] present an analytic formula for pricing such an ESO. We will present the intuition behind their price next.

We seek to value the discounted payoff of the option, which depends on the time of exercise. However the time of exercise of the option is now a random variable given by $min\{\tau, T\}$, where T is the time of expiry and τ is a random time before expiry, corresponding to the time of the first arrival of the Poisson process which models the departure of the employee from the company. Let

$$F(t) = \mathbb{P}_t(\tau \leq t)$$

be the distribution of τ conditional on information available from the asset price up to time t. The discounted option payoff at time 0 is given by

$$C_t = e^{-rt}\big(S(t) - K\big)^+. \tag{3.1}$$

We can write $C_{min\{\tau,T\}}$ as

$$\begin{aligned} C_{min\{\tau,T\}} =& C_\tau \, \mathbf{1}_{\{\tau \le T\}} + C_T \, \mathbf{1}_{\{\tau > T\}} \\ =& C_\tau \, \mathbb{P}(\tau \le T) + C_T \, \mathbb{P}(\tau > T) \\ =& \int_0^T C_u \, dF_u + C_T \, \mathbb{P}(\tau > T). \end{aligned} \tag{3.2}$$

Now Poisson interarrival times are exponentially distributed, so $F(t) = 1 - e^{-\lambda t}$, where λ is the intensity of the process. Therefore we can write

$$C_{min\{\tau,T\}} = \int_0^T \big(S(t) - K\big)^+ \lambda e^{-(r+\lambda)t} dt + \big(S(T) - K\big)^+ e^{-(r+\lambda)T}. \tag{3.3}$$

The value of the option is then obtained by taking the expectation of $C_{min\{\tau,T\}}$ as follows:

$$\begin{aligned} \mathbb{E}[C_{min\{\tau,T\}}] &= \mathbb{E}\left[\int_0^T (S_t - K)^+ \lambda e^{-(r+\lambda)t} dt + (S_T - K)^+ e^{-(r+\lambda)T}\right] \\ &= \lambda \int_0^T e^{-\lambda t} \, \mathbb{E}\big[e^{-rt}(S_t - K)^+\big] dt + \mathbb{E}\big[e^{-(r+\lambda)T}(S_T - K)^+\big]. \end{aligned} \tag{3.4}$$

The expectation inside the integrand of the above expression is equal to the Black-Scholes option pricing formula, and the second term outside the integral is simply the Black-Scholes option pricing formula multiplied by $e^{-\lambda T}$. Therefore the calculation of the price of the option is effectively reduced to evaluating the integral of the Black-Scholes pricing formula with respect to t. A closed-form expression for this expectation is provided in [2], which we have implemented in code in [7].

We now present values obtained for the price of the ESO by this method in Table 1. The values of each of the parameters in the model are $S(0) = 100$, $K = 100$, $r = 0.05$, $\sigma = 0.2$ and $T = 10$. We quote a 95% confidence interval for a sample size of $M = 1,000,000$. For comparison, we also quote the value of the ESO as given by the analytic formula given by (3.4). It should be noted that this

λ	MC price	CI	Analytic price
0.5	14.6118	[14.5802, 14.6435]	14.5941
0.4	16.7536	[16.7174, 16.7898]	16.7236
0.3	19.810	[19.7679, 19.8522]	19.7801
0.2	24.4078	[24.3579, 24.4577]	24.435
0.1	31.985	[31.9254, 32.0446]	32.0064
0.05	37.6487	[37.5845, 37.7129]	37.6401
0.02	41.9038	[41.8375, 41.9701]	41.8961

Tab. 1: Monte Carlo estimates for intensity-based model

valuation lies within the confidence interval quoted for each value of λ, indicating that our Monte Carlo approach closely matches the analytic valuation. We also see a decrease in the value of the option as λ is increased - as we would expect the employee to be more likely to depart and exercise, hence forgoing the remaining time value of the option.

3.2.1 Expected time to expiry

As previously mentioned, in this intensity model the time of expiry is no longer a pre-determined value T as is with a European call option, but a random variable τ, given by

$$\tau = min\{T_\lambda, T\}$$

where T_λ is the time of the first arrival of a Poisson process with intensity λ. It may be of interest to calculate the expected time of expiry for a given set of parameters for this particular variant of ESO. Write τ as

$$\tau = T_\lambda \, \mathbf{1}_{\{T_\lambda \leq T\}} + T \, \mathbf{1}_{\{T_\lambda > T\}}.$$

Now we know that Poisson interarrival times are exponentially distributed, so if F is the distribution function of T_λ, then

$$F(t) = 1 - e^{-\lambda t}$$

as before. Therefore

$$\mathbb{E}[\tau] = \mathbb{E}\big[T_\lambda \, \mathbf{1}_{\{T_\lambda \leq T\}} + T \, \mathbf{1}_{\{T_\lambda > T\}}\big]$$

$$= \mathbb{E}\big[T_\lambda \, \mathbf{1}_{\{T_\lambda \leq T\}}\big] + T \, \mathbb{P}(T_\lambda > T).$$

The expectation can be interpreted as the expected value of a truncated exponential distribution. In general, the expected value of a random variable X, truncated by c is given by

$$\mathbb{E}[X|X < c] = \frac{1}{F(c)} \int_{-\infty}^{c} x f(x)\, dx$$

where F and f are the distribution and density functions respectively of X. Hence we can write

$$\begin{aligned}
\mathbb{E}[\tau] &= \mathbb{E}[\, T_\lambda \,|\, T_\lambda < T\,] + T\,\mathbb{P}(T_\lambda > T) = \frac{1}{1-e^{-\lambda T}} \int_0^T t\lambda e^{-\lambda t}\, dt + Te^{-\lambda T} \\
&= \frac{\lambda}{1-e^{-\lambda T}} \Big(\frac{1}{\lambda} - \frac{1}{\lambda} e^{-\lambda T} - Te^{-\lambda T} \Big) + Te^{-\lambda T} = \frac{1 - e^{-\lambda T} - \lambda T e^{-\lambda T}}{1-e^{-\lambda T}} + Te^{-\lambda T} \qquad (3.5) \\
&= 1 + Te^{-\lambda T} - \frac{\lambda T e^{-\lambda T}}{1-e^{-\lambda T}}.
\end{aligned}$$

It should be noted that, by properties of the limit and applying L'Hôpital's rule,

$$\begin{aligned}
\lim_{\lambda \to 0} \mathbb{E}[\,\tau\,] &= \lim_{\lambda \to 0} \Big(1 + Te^{-\lambda T} - \frac{\lambda T e^{-\lambda T}}{1-e^{-\lambda T}} \Big) \\
&= 1 + T - \lim_{\lambda \to 0} \frac{\lambda T e^{-\lambda T}}{1-e^{-\lambda T}} \\
&= 1 + T - \lim_{\lambda \to 0} \frac{Te^{-\lambda T} + \lambda(-T^2 e^{-\lambda T})}{Te^{-\lambda T}} \\
&= 1 + T - \lim_{\lambda \to 0} (1 - \lambda T) = 1 + T - 1 = T,
\end{aligned}$$

as we would expect, as the payoff of the ESO approaches a standard European call option payoff with expiry T as the intensity parameter is decreased.

3.2.2 Variance reduction by control variate

There are many methods available to reduce the sampling error of a Monte Carlo method - one which may be applicable in this case is a control variate. A control variate is a random variable with known expectation which is correlated with the random variable whose expectation is being approximated. By defining a new random variable that is a linear combination of the two, the variance of the Monte Carlo estimate can be greatly reduced. This reduction is linked to the correlation between the two variables, as we will show below. We plan to use a standard Black-Scholes valuation with a reduced

time to expiry as this second correlated random variable.

Let Λ_E denote the discounted payoff of the ESO in question. The objective is to define a new random variable Z in terms of Λ_E and some correlated random variable Y with known expectation.

We write

$$Z = \Lambda_E + c\,(Y - \mathbb{E}[Y]),$$

where c is some constant. It follows that

$$\mathbb{E}[Z] = \mathbb{E}[\Lambda_E]$$

and

$$\mathrm{Var}(Z) = \mathrm{Var}(\Lambda_E) + c^2\mathrm{Var}(Y) + 2c\,\mathrm{cov}[\Lambda_E, Y],$$

which is minimised by choosing

$$c = -\frac{\mathrm{cov}[\Lambda_E, Y]}{\mathrm{Var}(Y)}, \tag{3.6}$$

giving

$$\begin{aligned}\mathrm{Var}(Z) &= \mathrm{Var}(\Lambda_E) - \frac{\mathrm{cov}[\Lambda_E, Y]}{\mathrm{Var}(Y)} \\ &= (1-\rho^2)\mathrm{Var}(\Lambda_E)\end{aligned} \tag{3.7}$$

where ρ is the correlation coefficient of Λ_E and Y. Thus the greater the correlation between the two, the greater the reduction in variance. A natural choice for Y in this context would be a discounted standard Black-Scholes payoff, denoted Λ_B. It has known expectation, namely the Black-Scholes option pricing formula. It is also reasonable to assume that the standard Black-Scholes payoff with a reduced time to maturity T' (to account for the early exercise possibility) and the payoff of the ESO in question would be correlated, giving a reasonable reduction in variance.

The first point to consider is how to choose a reduced time to maturity T' in a meaningful way. Many professional accounting bodies recommend taking the average of the vesting period and the time to expiry as T'. However this is a rather crude simplification, and as we are not considering a vesting period in this case, we must consider an alternative. The expected time to exercise $\mathbb{E}[\tau]$ as given by (3.5) may give a more realistic result in this particular case. So we aim to use a discounted Black-Scholes payoff with maturity T' equal to $\mathbb{E}[\tau]$ as our control variate.

We must now consider the choice of c. As mentioned above, it is reasonable to assume that these two payoffs will be positively correlated, therefore by condition (3.6), c should at least be strictly negative. However, analytically writing down the optimal valule of c given by (3.6) is more complicated. Note that

$$\text{cov}[\Lambda_E, \Lambda_B] = \mathbb{E}[\Lambda_E \Lambda_B] - \mathbb{E}[\Lambda_E]\mathbb{E}[\Lambda_B].$$

While it is reasonable to assume that $\mathbb{E}[\Lambda_E]$ is equal to the analytic price of the ESO, and we know that $\mathbb{E}[\Lambda_B]$ is the Black-Scholes option pricing formula, the value of $\mathbb{E}[\Lambda_E \Lambda_B]$ is unknown. One could apply the Cauchy-Schwartz inequality to get

$$\mathbb{E}[\Lambda_E \Lambda_B] \leq \sqrt{\mathbb{E}[\Lambda_E^2]\mathbb{E}[\Lambda_B^2]}.$$

There is an expression for $\mathbb{E}[\Lambda_B^2]$ given in [5], so that if an analytic formula were available for $\mathbb{E}[\Lambda_E^2]$, a lower bound for the optimal c could be obtained. This is not provided in [2] - it could however be numerically approximated. We provide some plots for identifying optimal values of c for varying values of λ in Fig. 7. The sampling error for a Monte Carlo estimate of the ESO payoff with Λ_B as a control variate for is plotted against various values of c for a sample size of $M = 500,000$. In choosing values for c to test, we note that due to the expected correlation between the Black-Scholes payoff and the ESO payoff, the ratio of cov[Λ_E, Λ_B] and Var(Λ_B) is unlikely to large, so we choose values of c from the interval [0, 2]. We can see that for these 4 particular values of λ, the optimal value of c, i.e. the value of c which gives the smallest sampling error lies between -1 and -0.75 - in particular, for λ large, the optimal c is closer to -0.75, and approaches -1 as λ is decreased. Table 2 presents a comparison between the brute-force Monte Carlo estimate and the control variate estimate of the price of the ESO for varying values of λ, based on a sample size of $M = 500,000$, and a decreasing value of c from -1 to -0.75. The sampling errors are quoted rather than a confidence interval to show the extent of the variance reduction of the control variate. We see that for all values of λ, the reduction in variance is

λ	Brute-force MC	SE	Control variate	SE
0.1	32.0102	51.9731	32.1105	42.9172
0.075	34.6968	54.9213	34.7634	43.1258
0.05	37.4969	58.0013	37.6873	41.59
0.01	43.3492	63.6601	43.4587	26.1357

Tab. 2: Using a Black-Scholes payoff as control variate

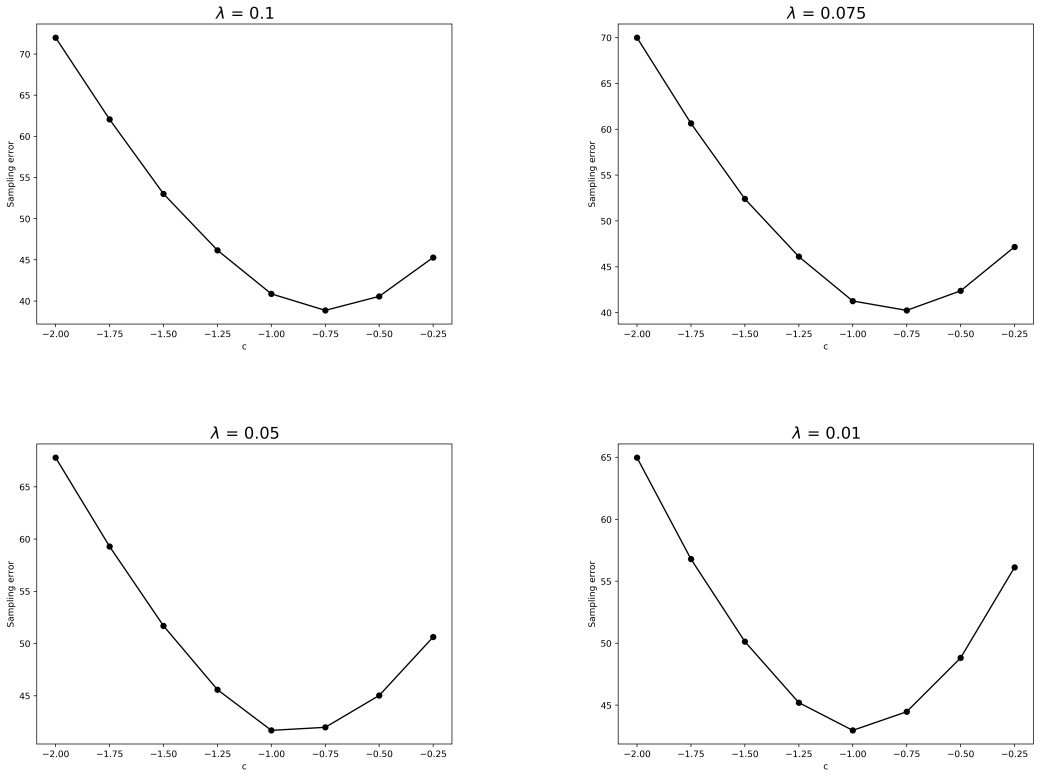


Fig. 7: Finding the optimal value of c for varying λ

significant - in particular, for small values of λ, the reduction is large. A possible reason for this is that when λ is small, the option is more likely to be held until expiry, and the reduced time to expiry T' will approach the time of expiry T, thus increasing the correlation between the two variables. By (3.7), this will lead to an increased reduction in variance, assuming that we have chosen the optimal value of c.

3.3 Capturing the early exercise feature

We now consider an extension to the above model - the early exercise feature of the option. Now, it is difficult to model this feature because the time of early exercise depends on a wide variety of factors, such as the current asset price $S(t)$, an indiviual's risk preferences, individual beliefs of what future share prices will be, etc. So we make a broad assumption about the early exercise patterns of any given employee, that there exists some 'barrier' L, which is higher than the strike K, at which the employee will be happy to exercise their option and surrender the potential for an even higher

profit. This hitting-time feature can be incorporated into the Monte Carlo approximation quite easily by simply checking if the stock price at time t is higher than L - if so, the employee will exercise and the payoff is computed and store easily.

In order to make this barrier more realistic, we allow it to depend on time t. The authors in [2] propose an exponentially decaying barrier L_t given by

$$L_t = Le^{\alpha t}$$

where α is some constant. If we consider the case where α is negative, it can be thought of as the rate at which L_t decays as time progresses. As $t \to T$, the value of the barrier decreases and the employee will exercise at a lower price, representing the decreasing time-value of the option. There are some issues with this barrier which we will discuss in the next section, but we will model this barrier in our Monte Carlo approach first in order to compare with the analytic prices for this ESO.

The implementation of this model is identical to the above case, with the exception of checking that the stock price $S(j)$ does not exceed L_j for each j. If it does, the time and price will be stored and used to compute the payoff, i.e. early exercise has occurred. We present some estimates for varying values of L and α in Table 3. The other parameter values remain unchanged from the previous case with $\lambda = 0.04$ and a sample size of $M = 1,000,000$. While the majority of the Monte Carlo estimates

L	MC price	CI	Analytic price
125	15.9658	[15.9406, 15.991]	15.3372
150	24.4112	[24.3621, 24.4602]	23.9052
200	32.2517	[32.1651, 32.3383]	32.0304

α	MC price	CI	Analytic price
-0.1	16.7212	[16.6818, 16.7605]	12.9485
-0.05	20.7765	[20.733, 20.8201]	20.2291
-0.01	25.4955	[25.4429, 25.548]	25.0745

Tab. 3: Monte Carlo estimates for intensity and hitting-time model

are quite close to their corresponding analytic values, we begin to observe some differences between the two in this case. The most stark of these is the cases of $\alpha = -0.1$. Here we see the analytic pricing formula greatly under-valuing the ESO relative to the Monte Carlo approach. A possible reason for

this is the implementation of the barrier L_t. If $L_t = Le^{\alpha t}$, then we can see that

$$L_t = Le^{\alpha t} < K \iff e^{\alpha t} < \frac{K}{L} \iff \alpha t < \log \frac{K}{L}$$

$$\iff t < \frac{1}{\alpha} \log \frac{K}{L}.$$

Therefore it is possible that, given some α and for some values of t, the barrier will be *below* the strike price. Alternatively we could look at a lower bound for α which ensures the barrier will not drop below the strike by considering the value of the barrier at time T, which will be less than any other time $t < T$:

$$L_T = Le^{\alpha T} > K \iff e^{\alpha T} > \frac{K}{L} \iff \alpha T > \log \frac{K}{L}$$

$$\iff \alpha > \frac{1}{T} \log \frac{K}{L}.$$

Thus for values of $\alpha = -0.1$, $K = 100$ and $L = 150$, the value of the barrier L_t will be *below* the strike price K for $t > 4.055$ - this is very unrealistic of actual early exercise patterns, as the option is worthless when $S(t) < L_t < K$, therefore it should not be exercised. It is unclear whether this anomaly has been taken into account in the analytic pricing formulae, meaning that they may give undervalued prices for the ESO.

3.3.1 Modifying the barrier

In light of the problems highlighted previously, we could propose another barrier, given by

$$K_t = Ke^{\alpha(t-T)} \tag{3.8}$$

where K is the strike price, α a constant, taken to be negative, and T is the time to expiry. It is clear that as $t \to T$, $K_t \to K$, meaning that the barrier will decrease towards K as the expiry of the option approaches. This would more accurately reflect the exercise patterns or a risk-averse individual. In order to visualise this barrier, K_t is plotted (dashed line) for various values of α in Fig. 7. We have also plotted the above barrier L_t with $L = 150$ and $\alpha = -0.05$ in black. It is clear that the barrier L_t lies below the strike price as $t \to T$. However, our proposed barrier of K_t lies strictly above the strike K. Therefore we can see from this visualisation that any model with a barrier of the form L_t will severely under-value ESOs when α is relatively high as the barrier will eventually be below and remain below

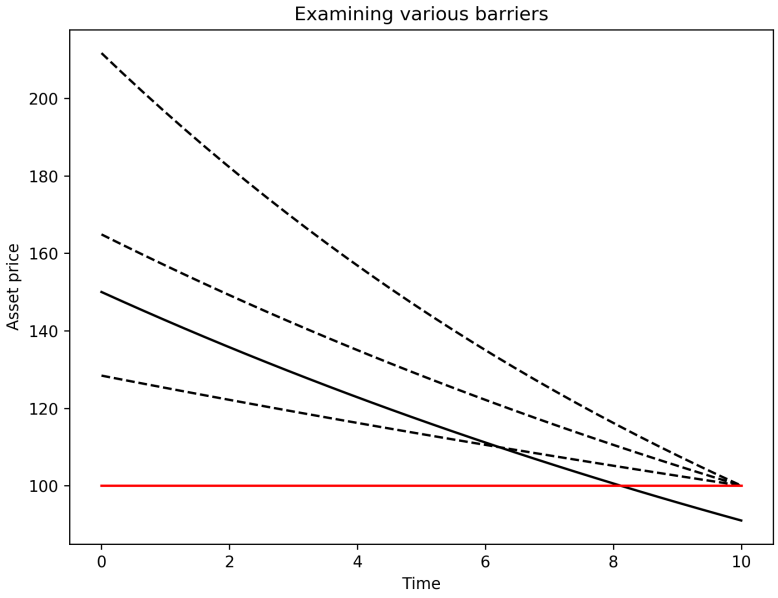


Fig. 8: Proposed barrier K_t with α equal to -0.075, -0.05 and -0.025

the strike price, resulting in the option being voluntarily be exercised when it is out-of-the-money. In light of this, we present some Monte Carlo valuations of the ESO with departure process and barrier K_t in Table 4. We can see that with the barrier K_t, the Monte Carlo approach gives greatly differing

α	MC estimate	CI
-0.1	30.3433	[30.2832, 30.4034]
-0.05	24.0525	[24.015, 24.0901]
-0.01	9.1211	[9.09762, 9.1447]

Tab. 4: Monte Carlo estimates with modified barrier K_t

prices of the ESO compared to with barrier L_t. For the larger two of the three values of α, the K_t barrier gives a higher price of the ESO, whereas for $\alpha = -0.01$, it gives a significantly lower price than the L_t barrier. This is due to the fact that the K_t is quite conservative for small values of α compared to L_t. However, it is clear that $K_t > K$ for all $t \in [0, T]$, which does not hold for L_t as we showed above. Therefore both barriers have advantages - a compromise could be obtained by considering another barrier which is a combination of both.

Consider a barrier given by

$$L_\tau(t) = K + \frac{L - K}{T - \tau}(T - t), \tag{3.9}$$

where K is the strike price, L is an arbitrary barrier, T is the time to expiry and τ is a parameter which determines the behaviour of the barrier. Fig. 9 plots this barrier for values of $\tau = 3$ (dashed line) and $\tau = 7$ (bold line). We can see that the barrier begins at some level $L = 150$, and remains at L until time τ, then decays to K in a straight line. The smaller the value of τ, the sooner this decay begins. An advantage of this barrier is that is does not impose a small barrier early in the life of the option as K_t does for small α. If τ is small, the decline of the barrier begins sooner, but also has a more gentle slope, which will hopefully capture more of the time-value of the option. Table 4 presents some Monte Carlo prices of an ESO for various values of τ. For this barrier, we see the price of the

τ	MC estimate	CI
1	19.1696	[19.1101, 19.2291]
3	23.1425	[23.0947, 23.1903]
5	24.8505	[24.8078, 24.8932]
7	25.7839	[25.7433, 25.8245]
9	26.3504	[26.3102, 26.3906]

Tab. 5: Monte Carlo estimates for barrier $L_\tau(t)$

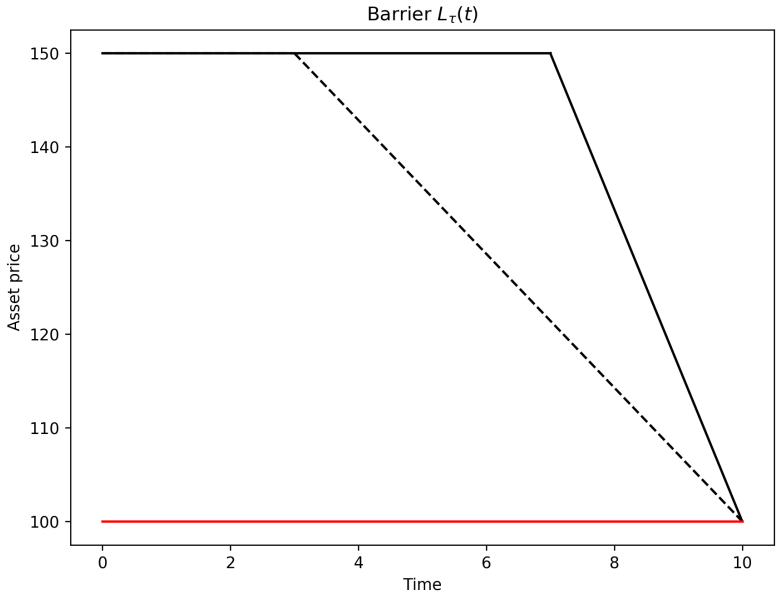


Fig. 9: Barrier $L_\tau(t)$ for $\tau = 3$ and $\tau = 7$.

ESO increasing as τ is increased, as expected. We also see more consistently increasing prices, unlike with the barrier K_t above.

3.4 Addition of a vesting period

The final complete model we consider for the Monte Carlo valuation is the above model with the inclusion of a vesting period. This is a period $[0, T_0]$, $T_0 < T$, during which the option cannot be exercised, and if an employee leaves the company, the option is forfeited. This model is similar to the above model, however we now consider two Poisson process - one for the vesting period, with intensity λ_0 and one for after vesting, with instensity λ. The reasoning behind this is that if an employee knows that their option is worthless if they leave the company during the vesting period, they will be less likely to voluntarily leave the company during this period than after. Therefore it makes sense to consider two separate processes with $\lambda_0 \leq \lambda$. As before, we also consider a barrier L_t given by

$$L_t = Le^{\alpha(t-T_0)}$$

which results in the option being exercised should the underlying asset price exceed L_t. The implementation of this is similar to the previous models - during the vesting period the asset price $S(t)$ evolves as follows:

- Let $\delta t = T/N$ and $j = n\delta t$. For $n = 1, \dots N_0, \quad N_0 = \lfloor T_0/\delta t \rfloor$, the asset price $S(j)$ evolves over intervals of length δt, during which a $Poi(\lambda_0\, \delta t)$ random variable is sampled to determine whether the employee has left the company during this interval. However in this case, because we are in the vesting period, if this random variable gives a count of 1 at some time j, the option becomes worthless, so we can set $C_j = 0$.

- For $n = N_0,\, N_0 + 1, \dots T$, the asset price evolves and a $Poi(\lambda\delta t)$ variable is sampled to determine if a departure has occured. Also, we must ensure that the price $S(j)$ has not exceeded L_j - if it has the option is exercised and, j and $S(j)$ are stored to compute C_j.

- This continues until the option is either exercised, or expires at time T.

We again present Monte Carlo estimates of the price of the ESO in this case with their corresponding analytic prices in Table 4. It should be noted that this model suffers from the same problem as the above case with regards to its barrier, therefore both the Monte Carlo estimate and the analytic price should be considered as under-valuing the ESO in this paricular case. The parameters of the model remain unchanged with $L = 150$, $\alpha = -0.05$ and $M = 1,000,000$. Once again, the Monte Carlo price

T_0	MC price	CI	Analytic price
1	24.5598	[24.5433, 24.5763]	24.603
3	26.5112	[26.4834, 26.5391]	26.7845
5	28.6967	[28.6554, 28.738]	28.5948
7	30.1066	[30.0533, 30.16]	29.6227

Tab. 6: Varying vesting period T_0

closely matches the analytic price here.

4 Conclusions

In Section 1, we give a discription of the basic features of the employee stock option and define some terminology. We have presented a construction of two numerical methods for solving stochastic differential equations, namely the Euler-Maruyama and the Euler-Milstein methods. We also investigate of the convergence of these methods, and examine the almost sure and mean-square stability of both methods in Section 2. The purpose of this section is to provide a rigorous framework in which these numerical methods can be applied to a more complex asset price model than the standard Black-Scholes model and generate trajectories of such an asset price. These trajectories will generally be required when using Monte Carlo methods to approximate the value of an option written on this asset. In particular, the stability regions we have presented can be used to determine whether the parameters of a given asset price model can give rise to a numerically stable method.

In Section 3, we have valued three different variants of the employee stock option by Monte Carlo methods. We have successfully implemented control variate, namely a standard Black-Scholes valuation with a reduced time to expiry, to obtain a significant reduction in the sampling error of our Monte Carlo estimate of the ESO price. We have also examined various methods for incorporating an early exercise feature into the ESO by considering three barriers which aim to model the early exercise patterns of an individual.

We have also compared the Monte Carlo prices obtained with the analytic formulae presented in [2], which are closely matched in two of the three cases. The authors' implementation of the barrier feature seems to have some flaws for particular choices of the parameters of the model, which we have discussed, as well as proposing some alternative barriers. The second barrier we propose appears to have the most consistent results, as well as the modeling the most realistic early exercise patterns.

To conclude, the Monte Carlo approach we present for valuing these ESOs appears to be quite capable of giving meaningful prices in a variety of cases. While being computationally expensive compared to any closed-form expression for the option price, the Monte Carlo approach has proved to be a more flexible method, allowing for various features of the option to be modified. For example, we were able to implement two new barriers to capture the early exercise feature with relative ease - modifying the analytic prices given in [2] for these changes would be extremely difficult. However, the analytic prices are a useful benchmark for the Monte Carlo estimate in the majority of cases presented in this report.

References

[1] J. D. Higham, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Review **43**, no. 3 (2001), 525-46.

[2] Cvitanić, Jakša, et al., *Analytic Pricing of Employee Stock Options*, The Review of Financial Studies, **21,** no. 2, 2008, 683–724.

[3] F. Black, M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Polit. Econ. **81** (1973), no. 3, 637-659.

[4] Y. Alnafisah, *The Implementation of Milstein Scheme in Two-Dimensional SDEs Using the Fourier Method*, Abstract and Applied Analysis, **2018**, Article ID 3805042, 7 pages, 2018.

[5] A. Ben-Meir, J. Schiff, (2012, April 17) , *The Variance of Standard Option Returns*, retrieved from https://arxiv.org/pdf/1204.3452.pdf.

[6] J. Hull, A. White, *How to Value Employee Stock Options*, Financial Analysts Journal, **60**, no. 1, 2004, 114–119.

[7] C. O Brien, (2019, March 21), *Jupyter notebooks for project 'Valuing Employee Stock Options'*, retrieved from https://github.com/Cianob10/ESO-Code.