Research Article

Shady Ahmed Nagy, Mohamed A. El-Beltagy* and Mohamed Wafa

Multilevel Monte Carlo by using the Halton sequence

https://doi.org/10.1515/mcma-2020-2065
Received October 14, 2018; accepted April 1, 2020

Abstract: Monte Carlo (MC) simulation depends on pseudo-random numbers. The generation of these numbers is examined in connection with the Brownian motion. We present the low discrepancy sequence known as Halton sequence that generates different stochastic samples in an equally distributed form. This will increase the convergence and accuracy using the generated different samples in the Multilevel Monte Carlo method (MLMC). We compare algorithms by using a pseudo-random generator and a random generator depending on a Halton sequence. The computational cost for different stochastic differential equations increases in a standard MC technique. It will be highly reduced using a Halton sequence, especially in multiplicative stochastic differential equations.

Keywords: Stochastic differential equation, Monte Carlo techniques, low discrepancy sequences, multilevel Monte Carlo method

MSC 2010: 81T80, 34k28, 60-XX

1 Introduction

Stochastic differential equations (SDEs) are widely used in application fields such as Engineering [4], Physics [5] and Finance [17]. SDEs are probabilistic models that are generated from choosing different samples randomly, so that SDEs depend on different samples. Random samples that represent the noise which is modeled by using Brownian motion. Each sample of Brownian motion is a different expected path for SDE, so it works on different samples [1, 2]. The MC method is useful by generating different random samples that are with a given distribution example: uniform or normal distribution. These samples are used to get solution statistics for a given function like mean, variance, skewness, and kurtosis. For example, consider the stochastic process (see [15])

$$X_t = X_0 + \int \mu(t, X_t)\, dt + \int \sigma(t, X_t)\, dW_t,$$

where W_t is a one-dimensional white noise, $\mu(t, X_t)$ is the drift term and $\sigma(t, X_t)$ is the diffusion term. The solution of the stochastic integral (second term on the right-hand side of the equation) can be simulated by representing the noise term as random samples and calculate the solution statistics of it by using the MC method:

$$I = \int_\Omega f(x, \omega)\, dx,$$

***Corresponding author: Mohamed A. El-Beltagy,** Engineering Mathematics and Physics Department, Engineering Faculty, Cairo University, Giza 12613, Egypt, e-mail: zbeltagy@eng.cu.edu.eg. https://orcid.org/0000-0001-9272-0096
Shady Ahmed Nagy, Mohamed Wafa, Engineering Mathematics and Physics Department, Engineering Faculty, Cairo University, Giza 12613, Egypt, e-mail: shady.nagy@guc.edu.eg, mwafa@nu.edu.eg

where Ω is the sample space dimension of the integration, and ω is the randomness that effects the differential equation. We find the value of the multi-dimension integration by approximating I using the summation

$$\frac{1}{N}\sum_{i=1}^{N} f(x,\omega_i)$$

which approximates $E(f(x))$ where E is the expectation of the function f and N is the number of independent random samples from any distribution in the pseudo-random generator on the unit cube of dimension $[0,1]^{\Omega}$. The pseudo-random vector generated can be used to calculate the stochastic integral.

It is well known that MC converges slowly, the convergence rate is proportional to the number of samples N resulting in a convergence rate $1/\sqrt{N}$; i.e. decreasing the convergence rate by a factor $\frac{1}{2}$, the number of samples should be increased by 4, e.g., $1/\sqrt{4N} \to 1/2\sqrt{N}$, which is expensive. The MC method depends on the initial seed with clustering and non-uniform distribution for samples.

Depending on the variance reduction technique called control variate, Giles invented MLMC path simulation, which uses the control variate where the random variable is a Brownian motion sample [6]. A coarser path is used as control variate for estimating fine paths since its exact expectation is not known. It will be used recursively as a control variate for estimating the fine paths in a multilevel approach. MLMC will greatly decrease the cost if some conditions are satisfied and additive noise-type SDEs. The computational cost will not be decreased when we use MLMC on the multiplicative type of SDEs. Multi-index Monte Carlo (MIMC) [8] is the most significant extension of uses one scalar level l or more than one level as computation stochastic partial differential equations (PDEs) with different levels for time steps and spatial discretization for distance. MIMC will use levels which will be defined in different directions. It is a combination of sparse grid methods and M sampling that will be better in high-dimensional stochastic applications. Quasi-Monte Carlo (QMC) is the equidistant way for generation different stochastic samples. Low discrepancy sequences will be the seed for generating different random numbers in an equidistant manner that will decrease the convergence rate. We use the Halton sequence, which is a type of low discrepancy sequences to generate random samples that are used as a random generator for stochastic samples in SDEs. The paper is organized as follows. In Section 2, we start with the QMC technique. In Section 3, the low discrepancy sequences that we use to generate stochastic samples are detailed. In Section 4, we present the Halton sequence and its definition. Section 5 consists of the MLMC theorem and its explanations. Section 6 consists of some numerical examples for both different simulations with a pseudo-random generator and a low discrepancy generator on different types of SDEs. We compare decreasing the computational cost in both methods. In Section 7, we discuss the benefits of using low discrepancy sequences as random generators in MLMC.

2 QMC technique

In order to increase the convergence rate when using MC, we can increase the uniformity of the generated random samples that are substituted in the function to evaluate integration by using QMC. QMC works in the same way as MC, but with more uniformity of the generated samples by choosing an algorithm based on equally distributed random numbers that are generated by using sequences in order to minimize the error. Decreasing gaps and reducing clustering from different random samples lead to an increase in the convergence rate. The convergence rate will be $(\ln N)^{\Omega}/N$, where Ω is the dimension of the problem [13].

QMC depends on low discrepancy sequences which are viewed as deterministic or more uniformly MC, which decreases the error bounds. So all the points are generated in controlled samples. The comparison is shown in Figure 1 with 1000 points that are generated by using the Halton sequence as an example on QMC techniques and a pseudo-random generator that is normally distributed, where the pseudo-random generator is a built-in Matlab function. We can notice that points that are generated in QMC with the Halton sequence are more uniformly distributed by avoidance of clustering than pseudo-random numbers. This is due to correlations between the generated points since different points that are not correlated have a small chance to lie near each other. A simple argument about $\sqrt{N}$ out of N points lies in the clustering as can be seen

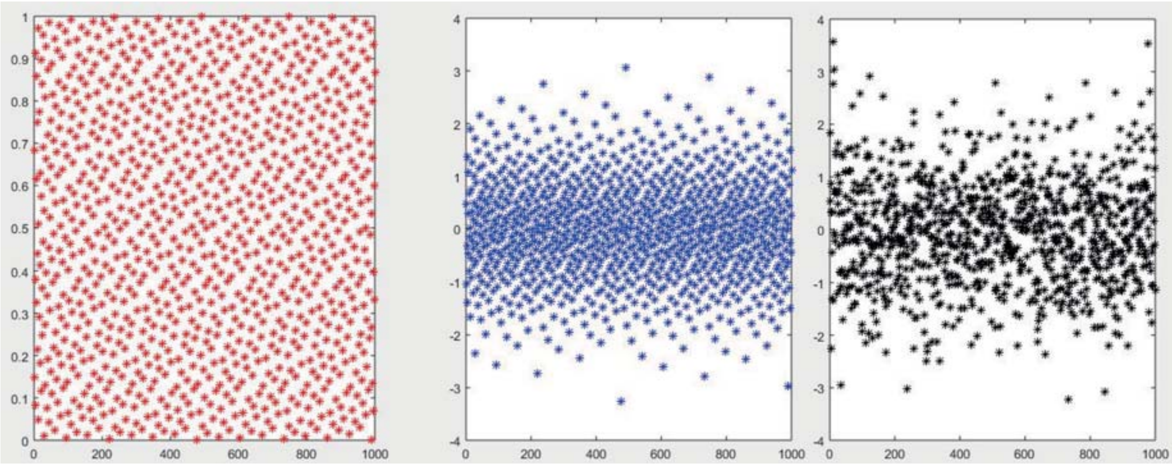

Figure 1: Comparison of three different random numbers, the red points on the left-hand side are Halton points, the blue points in the middle are Halton normalized points, and the black points on the right-hand side are Matlab pseudo-random points.

in Figure 1. They are designed for integration rather than optimization or simulation [12]. When we normalize the points generated from the Halton sequence, the clustering between different points will decrease, as it appears in Figure 1.

Clustering of generated points leads to increased errors in the simulation. Equal distributed numbers are covering more area under the curve in the calculation of the integral, so it will converge more to the exact solution.

3 Low discrepancy sequences

A discrepancy is the difference between the proportion of points in a volume box J compared to the full unit cube I^{Ω}, i.e. the unit cube $[0,1)^{\Omega}$, $\Omega \geq 1$ (see [11]). For N points $x_1, x_2, \ldots, x_N \in I^{\Omega}$ define the discrepancy by

$$D(J, N) = \left|\frac{A(J)}{N} - V(J)\right|,$$

where $A(J)$ is the number of points x_i in J and $V(J)$ is the volume occupied by the dimension of J.

The worst discrepancy is known as the star discrepancy and the goal is to minimize it:

$$D^{*(N)} = \max_{J} |D(J, N)|$$

If we use the low discrepancy sequences, the stochastic integral can be approximated through the sequence points. With the reduction of clustering between the different points, and the substitution of each point in the sequence in a sample of the stochastic integral sample, various sequences have been constructed to achieve the low discrepancy; see, for example, [3, 9, 14, 16].

4 Halton sequences

The Halton sequence is the most basic low discrepancy sequence in multiple dimensions. It is mainly depending on the van der Corput sequence in dimension Ω. If it is of dimension one, then the van der Corput sequence of base 2, the two-dimensional Halton sequence, is depending on the van der Corput sequence of base 3. The Halton sequence of dimension Ω is a van der Corput sequence of prime base Ω. The van der Corput sequence is a one-dimensional low discrepancy sequence. For a given integer n, we want to find its Quasi-Monte Carlo random number in prime base p:

$$n = \sum_{i=0}^{I} c_i p^i = c_0 p^0 + c_1 p^1 + c_2 p^2 + c_3 p^3 + \cdots + c_I p^I, \quad c_i < p^i.$$

So the random number x_n will equal $y_p(n)$, which is computed as $\sum_{i=0}^{I}(c_i/(p^{i+1}))$. As an example, if $n = 22$ and the base of the sequence is $p = 3$, then write

$$22 = c_0 3^0 + c_1 3^1 + c_2 3^2.$$

If $c_2 = 9$, $c_1 = 3$ and $c_0 = 1$, then the quasi-random number x_{22} is

$$x_{22} = y_3(22) = \frac{c_0}{p^1} + \frac{c_1}{p^2} + \frac{c_2}{p^3} = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} = \frac{7}{9},$$

so the generated quasi-random number from the sequence is already in the interval $[0, 1]$. Each number will be included in this according to the seed, for example:

$$\begin{aligned} n = 1: &\quad 1 = 1 * 3^0 = 1, & x_1 = y_3(1) = \frac{1}{3}, \\ n = 2: &\quad 2 = 2 * 3^0 = 2, & x_1 = y_3(2) = \frac{2}{3}, \\ n = 3: &\quad 3 = 0 * 3^0 + 1 * 3^1 = 3, & x_3 = y_3(3) = \frac{0}{3} + \frac{1}{9} = \frac{1}{9}. \end{aligned}$$

The van der Corput sequence in this way with the prime base 3 will be $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \ldots$; note that all numbers are included in the interval $[0, 1]$. The generalization of van der Corput on a multi-dimensional sequence is a Halton sequence on a unit hyper cube $[0, 1]^\Omega$. According to the dimension of the cube, we determine the base of the van der Corput sequence. If the first dimension of the Halton sequence will be base 2 in van der Corput, the second dimension will be base 3, and so on. The s-th dimension of the unit hyper cube is a van der Corput sequence of the same s-th prime base.

	dim = 1 (base 2)	**dim = 2 (base 3)**
$n = 1$	$\frac{1}{2}$	$\frac{1}{3}$
$n = 2$	$\frac{1}{4}$	$\frac{2}{3}$
$n = 3$	$\frac{3}{4}$	$\frac{1}{9}$
$n = 4$	$\frac{1}{8}$	$\frac{4}{9}$
$n = 5$	$\frac{5}{8}$	$\frac{7}{9}$
$n = 6$	$\frac{3}{8}$	$\frac{2}{9}$

Table 1: Halton sequence points and dimensions.

For the number $(17)_3 = 122$, in order to find the Halton sequence number for it, reverse the digits and point a radix point in front of the sequence number so it will be $(0.211)_3$. Therefore, $H_{17} = 0.122$, but in base 3, as it is different if we change the base of the van der Corput sequence. Table 1 shows the relation of the Halton sequence points and the dimension of the sequence based on a prime number. The usage of the Halton sequence in the Quasi-Monte Carlo integration will improve the convergence rate due to the uniform distribution of the different random variables.

5 Multi-level Monte Carlo

Theorem 5.1 (MLMC theorem [7]). *Let P denote a random variable, and let P_l denote the corresponding level l numerical approximation. Assume there exist independent estimators $\widehat{Y}_l$ based on N_l Monte Carlo samples each* *costing C_l, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geqslant \min(\beta, \gamma)$ and*

$$|E[\widehat{P}_l - P]| \leq c_1 * 2^{-\alpha l},$$
$$E[\widehat{Y}_l] = \begin{cases} E[\widehat{P}_0], & l = 0, \\ E[\widehat{P}_l - \widehat{P}_{l-1}], & l > 0, \end{cases}$$
$$V[\widehat{Y}_l] \leq c_2 * N_l^{-1} * 2^{-\beta l},$$
$$E[C_l] \leq c_3 * 2^{\gamma l}.$$

Then there exists a positive constant c_4 such that for any $\varepsilon < e^{-1}$ there are values L and N_l for which the multi-level estimator

$$\widehat{Y} = \sum_{l=0}^{L} Y_l$$

has a mean square error (MSE) with bound

$$E[(\widehat{Y} - E[P])^2] < \varepsilon^2,$$

and with computational complexity C with bound

$$E[C] \leq \begin{cases} c_4 \varepsilon^{-2}, & \beta > \gamma, \\ c_4 \varepsilon^{-2} (\log \varepsilon)^2, & \beta = \gamma, \\ c_4 \varepsilon^{-2-(\gamma-\beta)/\alpha}, & 0 < \beta < \gamma. \end{cases}$$

The total variance will be

$$V = \sum_{l=0}^{L} N_l^{-1} V_l,$$

the total cost will be

$$C = \sum_{l=0}^{L} N_l C_l$$

and the mean square error will be

$$E\big[(\widehat{Y} - E[P])^2\big] = N^{-1} V[\widehat{P}] + (E[\widehat{P}] - E[P])^2.$$

Great accuracy occurs with larger N and a small value of the weak error $E[\widehat{P}] - E[P]$. The main idea is that we need to estimate $E[P]$, but we do not have enough information about it, so we can use the control variate variance reduction method that can estimate it through another estimator that is highly correlated to it:

$$N^{-1} \sum_{n=1}^{N} \big[h(\omega^n) - \lambda(g(\omega^n) - E[g])\big].$$

The term P represents the quantity of interest in a given SDE that we want to estimate. The simulation by Matlab is depending on the MC integration for the stochastic integral of a differential equation. It will replace the conventional MLMC that depend on the *randn (Matlab random generator)* pseudo-generator to generate the normal distributed random variables by the new quasi-random generator (Halton sequence). It will increase the convergence rate and also improve the overall cost better than the conventional MLMC. The variance of each level equals $O(2^{-\beta l})$ and the cost of each level equals $O(2^{\gamma l})$, where β is the exponent of variance and γ is the exponent of cost. The optimal number of samples N_l is proportional to $2^{(-(\gamma+\beta)l/2)}$, $l = 0, \dots, L$. Therefore, the cost on level l is proportional to $2^{((\gamma-\beta)l/2)}$. The relation between β and γ determines the behavior of the cost in coarsest levels or finest levels. By plotting the mean, variance and kurtosis, consistency check versus the different levels. We are only interested in the computational cost that has clear change when applying the Quasi-Monte Carlo algorithm on the MLMC routine. Successive approximations of quantity of interest P are $P_0, P_1, P_2, \dots, P_L$ with increasing accuracy, but also increasing cost and convergence to P as $L \to \infty$. The computational cost is approximately proportional to $C \propto \varepsilon^{-2}$ when $\beta > \gamma$, so the multiplication between $C\varepsilon^2$ would be approximately constant; thus we consider decreasing the cost in spite of increasing the number of levels or accuracy. This is the standard result for an MC approach using i.i.d. samples; to do better would require a different approach than QMC methods. When $\beta < \gamma$, the dominant computational cost will be on the finest levels only on L as $C_L = O(\varepsilon^{-\gamma/\alpha})$. The comparison is between the computational costs that will be taken in simulation with different number of stochastic samples for different kinds of equations: Decreasing in computational cost between the conventional Std MC and MLMC in case of using a pseudo-random generator and decreasing in the computational cost between Std MC using Halton sequence points with MLMC.

6 Examples

Consider the Vasciek interest model as SDE with additive noise [18]:

$$r_t = a(b - r_t)dr_t + \sigma\, dw_t,$$

where $r_0 = 2$, $a = 1$ and $b = 0.04$. It gives an explicit characterization of the term structure of interest rates in an efficient market. This SDE depends on the Black–Scholes model for option pricing. This Vasciek model depends on an Ornstein–Uhlenbeck process where the interest rate should fluctuate around a mean equilibrium rate which is between the economic equilibrium of demand and supply. As in Figure 2, we shall consider that MLMC using the Halton sequence has higher cost than MLMC using the same points. The case of pseudo-random generator points Std MC cost is lower than MLMC when using the same sample points. The model with Halton estimates of key MLMC theorem parameters based on linear regression Table 2 are:
- $\alpha = 1.075512$ (exponent for MLMC weak convergence).
- $\beta = 1.151036$ (exponent for MLMC variance).
- $\gamma = 0.469231$ (exponent for MLMC cost).

Accuracy	MLMC Cost	Std MC Cost	Savings
$\varepsilon = 0.005$	40224	83372630	2072.71
$\varepsilon = 0.01$	15024	10380710	690.94
$\varepsilon = 0.02$	11028	1287371	116.75
$\varepsilon = 0.05$	10560	101355	9.59
$\varepsilon = 0.1$	10536	25338.74	2.40

Table 2: Vasciek Interest model values based on Halton points.

The Vasciek SDE values with using pseudo-random generator estimates of key MLMC theorem parameters based on linear regression Table 3 are:
- $\alpha = 1.020323$ (exponent for MLMC weak convergence).
- $\beta = 2.028532$ (exponent for MLMC variance).
- $\gamma = 0.878928$ (exponent for MLMC cost).

Accuracy	MLMC Cost	Std MC Cost	Savings
$\varepsilon = 0.005$	15286.50	81703.06	5.34
$\varepsilon = 0.01$	10572	9883.435	0.93
$\varepsilon = 0.02$	10524	2470.859	0.23
$\varepsilon = 0.05$	10512	395.3374	0.03
$\varepsilon = 0.1$	10512	98.83435	0.01

Table 3: Vasciek Interest model values based on pseudo-random points.

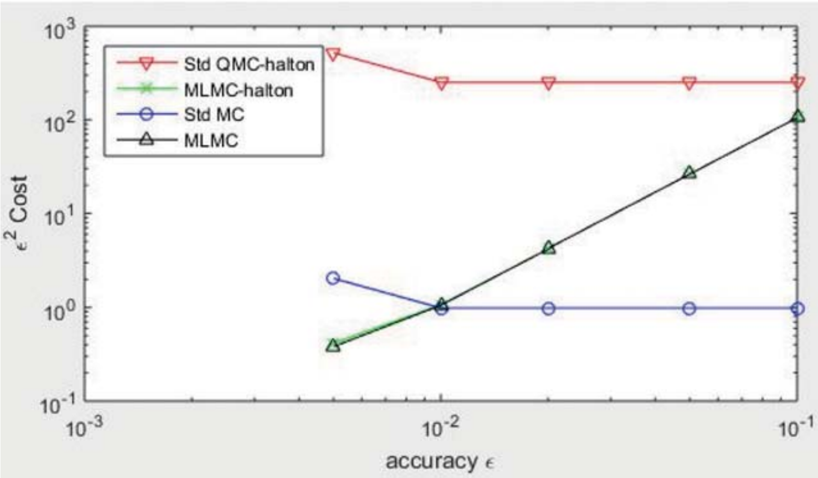

Figure 2: Cost for different MLMC-Halton, Std QMC-Halton, Std MC, and MLMC on the Vasciek interest model.

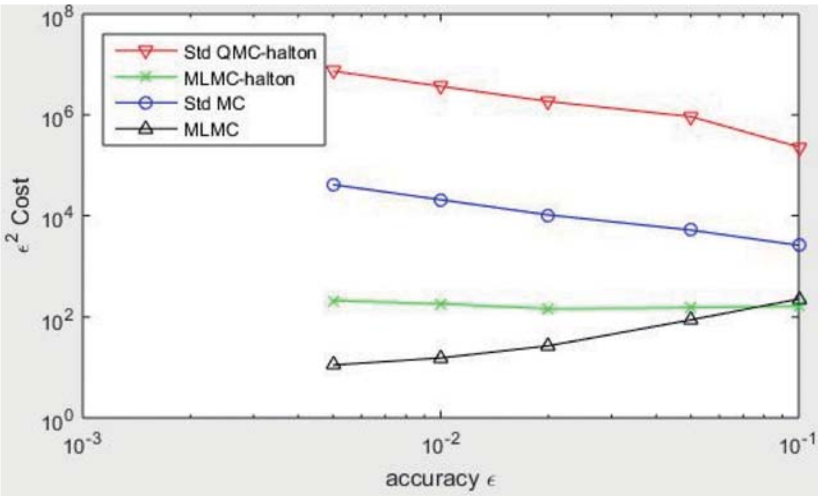

Figure 3: Cost for different MLMC-Halton, Std QMC-Halton, Std MC, and MLMC on an inhomogeneous SDE.

By using Vasciek–Halton points $\beta > \gamma$, we obtain $C = O(\varepsilon^{-2})$, but with more savings than in the ordinary Vasciek model regarding that standard MC using Halton sequence points has higher cost than pseudo-random number generators. Another example of an inhomogeneous SDEs is (see [10])

$$dX_t = (aX_t + b)\,dt + \sigma\,dW_t,$$

where $X_0 = 2$, $a = 2$, $b = 0.4$ and $\sigma = 0.2$.

The same holds for inhomogeneous-type SDEs but the cost remains the same in both algorithms which are decreased in the Halton sequence by increasing the accuracy, which appears in Figure 3. Inhomogeneous based on Halton estimates of key MLMC theorem parameters based on linear regression Table 4 are:

- $\alpha = 0.386942$ (exponent for MLMC weak convergence).
- $\beta = -0.226111$ (exponent for MLMC variance).
- $\gamma = 0.534620$ (exponent for MLMC cost).

Inhomogeneous based on pseudo-random generator estimates of key MLMC theorem parameters based on linear regression estimates of key MLMC theorem parameters based on linear regression Table 5 are:

- $\alpha = 0.386948$ (exponent for MLMC weak convergence).
- $\beta = 0.728697$ (exponent for MLMC variance).
- $\gamma = 0.874801$ (exponent for MLMC cost).

Accuracy	MLMC Cost	Std MC cost	Savings
$\varepsilon = 0.005$	1625101×10^4	5790102×10^8	35629.18
$\varepsilon = 0.01$	3609126×10^3	7237626×10^7	20053.68
$\varepsilon = 0.02$	792932500	9047028×10^6	11409.58
$\varepsilon = 0.05$	110701300	7237615×10^5	6537.97
$\varepsilon = 0.1$	23707050	9047002×10^4	3816.17

Table 4: Inhomogeneous SDE values based on Halton sequence points.

Accuracy	MLMC Cost	Std MC cost	Savings
$\varepsilon = 0.005$	95887941	9311257×10^6	97105.61
$\varepsilon = 0.01$	26657640	1163907×10^6	43661.29
$\varepsilon = 0.02$	6530063	7274413×10^4	11139.88
$\varepsilon = 0.05$	1551561	5819525×10^3	3750.75
$\varepsilon = 0.1$	557262	7274393×10^2	1305.38

Table 5: Inhomogeneous SDE values based on pseudo-random points.

By using an inhomogeneous SDE based on Halton points $\beta < \gamma$, we have $C_L = O(\varepsilon^{-\gamma/\alpha}) = O(\varepsilon^{-1.38})$, which is related to the dominant computational cost in the finest level. Ordinary inhomogeneous SDEs with pseudo-random numbers, where $\beta > \gamma$, but the computational cost dominant, has in coarsest levels $C = O(\varepsilon^{-2})$. An example for variable coefficients SDEs is (see [10])

$$dX_t = (a(t)X_t + b(t))\,dt + \sigma b(t)\,dW_t,$$

where $b = 0.4$, $\sigma = 0.2$, $X_0 = 1$, $a(t) = \frac{2}{1+t}$, and $b(t) = 0.4(1 + t^2)^2$ is variable coefficient as a function of time. Decreasing of the cost in the Halton sequence does not grow by increasing the accuracy; it remains the same when changing the accuracies. Variable coefficients SDEs based on Halton sequence estimates of key MLMC theorem parameters based on linear regression Table 6 are:
- $\alpha = 0.542101$ (exponent for MLMC weak convergence).
- $\beta = -0.071720$ (exponent for MLMC variance).
- $\gamma = 0.538967$ (exponent for MLMC cost).

Accuracy	MLMC Cost	Std MC cost	Savings
$\varepsilon = 0.005$	3594948	1730755×10^4	48144.09
$\varepsilon = 0.01$	794838	2163427×10^4	27218.47
$\varepsilon = 0.02$	140728.5	135208×10^4	9607.72
$\varepsilon = 0.05$	28428	108159800	3804.69
$\varepsilon = 0.1$	16668	13518330	811.03

Table 6: Variable coefficients SDE values based on Halton points.

Variable coefficient SDEs with pseudo-random generator estimates of key MLMC theorem parameters based on linear regression Table 7 are:
- $\alpha = 0.542105$ (exponent for MLMC weak convergence).
- $\beta = 1.332413$ (exponent for MLMC variance).
- $\gamma = 0.965883$ (exponent for MLMC cost).

By using variable coefficient SDEs based on Halton points $\beta < \gamma$, we have $C_L = O(\varepsilon^{-\gamma/\alpha}) = O(\varepsilon^{-0.99})$, which relates to the dominant computational cost in the finest level. Ordinary variable SDEs with pseudo-random numbers, where $\beta > \gamma$, but the computational cost is dominant, has in coarsest levels $C = O(\varepsilon^{-2})$. The total savings in cost from Std MC and MLMC by using Halton sequences are higher than using pseudo-random numbers regarding the increasing standard cost in using Halton sequences.

Accuracy	MLMC Cost	Std MC cost	Savings
$\varepsilon = 0.005$	331362	877953700	2649.53
$\varepsilon = 0.01$	134998.5	109743400	812.92
$\varepsilon = 0.02$	62727	13717710	218.69
$\varepsilon = 0.05$	23160	548658.3	23.69
$\varepsilon = 0.1$	16644	68573.91	4.12

Table 7: Variable coefficients SDE values based on pseudo-random points.

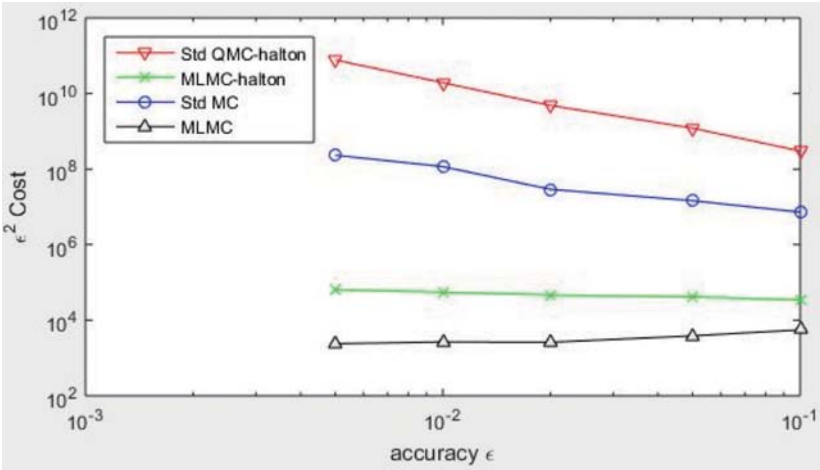

Figure 4: Cost for different MLMC-Halton, Std QMC-Halton, Std MC, and MLMC on a variable coefficient SDE.

The Verhulst stochastic equation (see [19, 20])

$$dX_t = rX(t)\Big(1 - \frac{X(t)}{K}\Big)\,dt + \sigma X(t)\,dW_t$$

with $X_0 = 5$, $\sigma = 1$, $r = 0.001$ and $K = 200$, is known as a logistic equation. It is used to model the growth rate with boundaries related to growing and decaying intervals of an environment. The model is continuous in time where the rate of the reproduction is proportional with the available resources and the existing population, r defines the growth rate of the population and K is the carrying capacity of the required environment. Therefore, $(\frac{K-X(t)}{K})$ population is limited by the carrying capacity of the environment, and the increasing of $X(t)$ is limited by K.

The Verhulst stochastic equation is a representative for multiplicative noise-type SDEs. Computational cost in the MLMC method is more than the standard MC algorithm (see Figure 4), but when we apply Halton sequence on the MLMC it will have lower computational cost than Std MC algorithm, which is an enhancement in the computational cost.

Verhulst based on Halton points estimates of key MLMC theorem parameters based on linear regression Table 8 are:

- $\alpha = -3.010288$ (exponent for MLMC weak convergence).
- $\beta = -5.638101$ (exponent for MLMC variance).
- $\gamma = 0.631639$ (exponent for MLMC cost).

Verhulst based on pseudo-random generator estimates of key MLMC theorem parameters based on linear regression Table 9 are:

- $\alpha = -0.076567$ (exponent for MLMC weak convergence).
- $\beta = -0.156781$ (exponent for MLMC variance).
- $\gamma = 0.981791$ (exponent for MLMC cost).

Accuracy	MLMC Cost	Std MC Cost	Savings
$\varepsilon = 0.005$	340346409	240912381×10^9	707844640.80
$\varepsilon = 0.01$	85086609	60228095×10^9	707844584.65
$\varepsilon = 0.02$	21271660	15057023×10^9	707844409.95
$\varepsilon = 0.05$	271222152	5162408×10^9	19033873.67
$\varepsilon = 0.1$	70946178	1290602×10^9	18191283.60

Table 8: Verhulst SDE values based on Halton points.

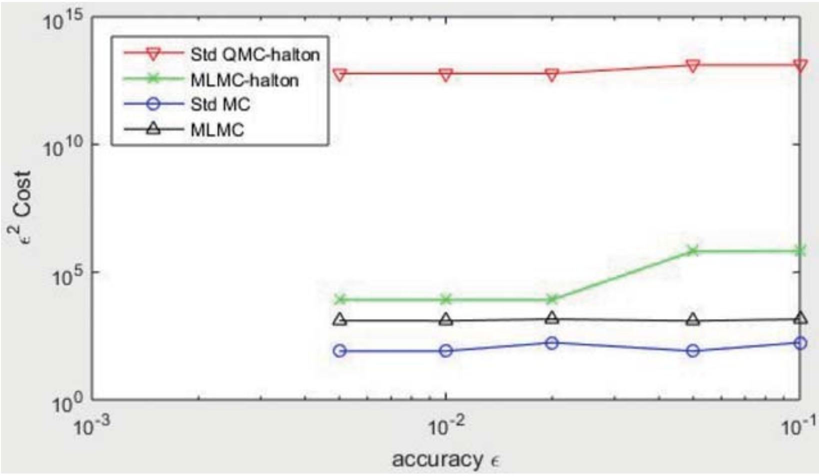

Figure 5: Cost for different MLMC-Halton, Std QMC-Halton, Std MC, and MLMC on an Verhulst SDE.

Accuracy	MLMC Cost	Std MC Cost	Savings
$\varepsilon = 0.005$	51450846	3272463	0.06
$\varepsilon = 0.01$	12591252	818115.8	0.06
$\varepsilon = 0.02$	3638969	438276.3	0.12
$\varepsilon = 0.05$	499488	32724.63	0.07
$\varepsilon = 0.1$	144475.5	17531.05	0.12

Table 9: Verhulst SDE values based on pseudo-random points.

By using the Verhulst–Halton points $\beta < \gamma$, we have $C_L = O(\varepsilon^{\frac{-\gamma}{\alpha}}) = O(\varepsilon^{0.21})$ which relates with the dominant computational cost in the finest level. Ordinary variable SDEs with pseudo-random numbers, where $\beta < \gamma$, but the computational cost is dominant, is in coarsest levels $C = O(\varepsilon^{12.8})$. The exponent of the accuracy value that appears in Figure 5 shows that MLMC cost is higher than for Std MC with different accuracies.

The total savings in cost from Std MC and MLMC by using Halton sequences is higher than using pseudo-random numbers regarding the increasing standard cost in using Halton sequences.

7 Conclusions

In this paper, We use random numbers generated with different techniques in the MLMC method and apply them on different SDEs. Depending on low discrepancy sequences, multi-level Quasi-Monte Carlo is more efficient than the conventional MLMC method as it decreases the computational cost. However, the computational cost by using the Halton sequence as a random generator for stochastic samples is larger than Std MC by using pseudo-random generators. The computational cost decreases with more savings values with the large difference between the Std MC and MLMC using the Halton sequence. The Halton sequence as a random generator will enhance and decrease MLMC computational cost in several types of SDEs especially in multiplicative noise-type SDEs. We obtain high performance by decreasing the computational cost in case of SDEs with multiplicative noise when applying MLMC using Halton sequences as random generators.

References

[1] W. T. Coffey, Development and application of the theory of Brownian motion, *Adv. Chem. Phys.* (1985), DOI 10.1002/9780470142875.ch2.

[2] A. Einstein, *Investigations on the Theory of the Brownian Movement*, Dutton, New York, 1926.

[3] H. Faure, Discrépance de suites associées à un système de numération (en dimension s), *Acta Arith.* **41** (1982), no. 4, 337–351.

[4] A. Friedman, *Stochastic Differential Equations and Applications. Vol. 2*, Probab. Math. Statist. 28, Academic Press, New York, 1976.

[5] I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations and Their Applications* (in Russian), Naukova Dumka, Kiev, 1982.

[6] M. B. Giles, Multilevel Monte Carlo path simulation, *Oper. Res.* **56** (2008), no. 3, 607–617.

[7] M. B. Giles, Multilevel Monte Carlo methods, *Acta Numer.* **24** (2015), 259–328.

[8] A.-L. Haji-Ali, F. Nobile and R. Tempone, Multi-index Monte Carlo: When sparsity meets sampling, *Numer. Math.* **132** (2016), no. 4, 767–806.

[9] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.* **2** (1960), 84–90.

[10] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. (N. Y.) 23, Springer, Berlin, 1992.

[11] J. Matoušek, On the L_2-discrepancy for anchored boxes, *J. Complexity* **14** (1998), no. 4, 527–556.

[12] W. J. Morokoff and R. E. Caflisch, Quasi-random sequences and their discrepancies, *SIAM J. Sci. Comput.* **15** (1994), no. 6, 1251–1279.

[13] H. Niederreiter, Point sets and sequences with small discrepancy, *Monatsh. Math.* **104** (1987), no. 4, 273–337.

[14] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 63, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1992.

[15] B. Øksendal, *Stochastic Differential Equations*, 6th ed., Universitext, Springer, Berlin, 2003.

[16] I. M. Sobol, The distribution of points in a cube and the approximate evaluation of integrals, *USSR Comput. Math. Math. Phys.* **7** (1964), 86–112.

[17] J. M. Steele, *Stochastic Calculus and Financial Applications*, Appl. Math. (N. Y.) 45, Springer, New York, 2001.

[18] O. Vasciek, An equilibrium characterization of the term structure, *J. Financial Econ.* **5** (1977), no. 2, 177–188.

[19] P. F. Verhulst, Recherché mathématiques sur la loi d’accroissement de la population, *Nouv. Acad. Roy. Sci. Belles-Lettres Bruxelles* **18** (1845), 1–41.

[20] P. F. Verhulst, Deuxiéme mémoire sur la loi d’accroissement de la population, *Mém. Acad. Roy. Sci. Lettres Beaux-Arts Belgique* **20** (1847), 1–32.