Stochastic Simulations
Mini-project: Multilevel Monte Carlo methods for option pricing

Bruno Rodriguez Carrillo
EPFL

January 19, 2021

We would like to apply the MLMC algorithm for option pricing. We consider the underlying SDE to follow the simple geometric Brownian motion:

$$dS(t) = rS(t)dt + \sigma S(t)dW_t,\ 0 < t < 1$$

For all results we use the following simulation parameters: $S(0) = 1$, $r = 0.05$, $\sigma = 0.20$, $K = 1$ and W_t a standard Brownian motion. We compute the expected payoff $\mu_{CMC} = \mathbb{E}[Y]$ for three different types of options given by:

1. European call option:
$$Y = \exp(-r)\max(0, S(1) - K) \tag{1}$$
2. Asian option:
$$Y = \exp(-r)\max\left(0, \bar{S} - K\right) \tag{2}$$
And
$$\bar{S} = \int_0^1 S(t)dt$$
3. Digital option:
$$Y = \exp(-r) H(S(1) - K); \tag{3}$$
$H(x)$ represents the Heaviside function.

We compute $\bar{S}$ as an average.
For point f only, we use $K = 25$.
Point a)

We know that the total cost and total variance of $\mathbb{E}[P_L]$ are respectively. Then at level l, we can compute the minimum value of variance given a fixed cost C_l as:
We know that total cost is given as $\sum_{l=0}^{L} N_l C_l$. And the total variance as $\sum_{l=0}^{L} V_l N_l^{-1}$. Then, in order to get the equation for minimizing the variance given a fixed cost, we use the following expression:

$$f(N_l, \lambda) = \sum_{l=0}^{L}\left(\lambda N_l C_l + V_l N_l^{-1}\right) \tag{4}$$

with λ a Lagrange multiplier.
Taking the derivative of with respect to $N_j, j = 1, ..., L$, and equaling to zero, we have:

$$\frac{\partial}{N_j} f(N_l, \lambda) = 0 \Rightarrow \frac{\partial}{N_j}\left(\sum_{l=0}^{L}\lambda N_l C_l + V_l N_l^{-1}\right) = 0 \tag{5}$$

At level l, we have: $\frac{V_l}{N_l^2} - \lambda C_l = 0$. Thus, at level 0 and 1, we have: $\frac{V_0}{N_0^2} = \lambda C_0$ and $\frac{V_1}{N_1^1} = \lambda C_1$. Then, eliminating λ, for the 2-level estimator, we have:

$$N_1 \cdot \sqrt{\frac{V_0}{C_0}} = N_0 \cdot \sqrt{\frac{V_1}{C_1}} \tag{6}$$

which is the equation to compute the optimal allocation of N_0 and N_1 that minimizes the variance of the estimator $\widehat{\mu}^{2-level}$.
Similarly, to compute the optimal number of samples at level l, N_l that minimizes the total cost given a variance $V_{MLMC} \leq \epsilon^2$, we use a second function to minimize:

$$f_2(N_l, \lambda) = \sum_{l=0}^{L}\left(N_l C_l + \lambda V_l N_l^{-1}\right) \tag{7}$$

with λ a Lagrange multiplier. As we did before, we take the derivative of with respect to $N_j, j = 1, ..., L$, and equaling to zero, then we have:

$$\frac{\partial}{N_j} f_2(N_l, \lambda) = 0 \Rightarrow \frac{\partial}{N_j}\left(\sum_{l=0}^{L} N_l C_l + \lambda V_l N_l^{-1}\right) = 0 \Rightarrow C_l = \lambda \frac{V_l}{N_l^2} \Rightarrow N_l = \sqrt{\lambda}\sqrt{\frac{V_l}{C_l}} \tag{8}$$

Then, since the total variance $\epsilon^2 = \mathbb{V}[Y] = \sum_{l=0}^{L} V_l N_l^{-1}$, we have:

$$\epsilon^2 = \sum_{l=0}^{L} V_l \left(\sqrt{\lambda}\sqrt{\frac{V_l}{C_l}}\right)^{-1} \Rightarrow \sqrt{\lambda} = \frac{1}{\epsilon^2}\sum_{l=0}^{L}\sqrt{C_l V_l} \tag{9}$$

Inserting this value of λ into (8), we obtain:

$$N_l = \left\lceil \frac{1}{\epsilon^2} \cdot \sqrt{\frac{V_l}{C_l}}\left(\sum_{l=0}^{L}\sqrt{C_l V_l}\right)\right\rceil \tag{10}$$

Where we must make sure to take the integer part as N_l is the number of samples at level $l = 0, ..., L$. The mean squared error of our estimator is:

$$MSE = \mathbb{E}\left[(Y_L - \mathbb{E}[P])^2\right] = \mathbb{V}[Y_L] + (\mathbb{E}[Y_L] - \mathbb{E}[P])^2 \tag{11}$$

where:

$$\widehat{\mu}^{MLMC} = Y_L = \sum_{0}^{L} Y_l \text{ and } Y_l = \frac{1}{N_l}\sum_{n=0}^{N_l}\left(Y_l^{(n,l)} - Y_{l-1}^{(n,l)}\right) \ , Y_{-1} = 0 \tag{12}$$

$$\mathbb{V}[Y_L] = \sum_{l=0}^{L} V_l N_l^{-1} \text{ and } V_l = \mathbb{V}[Y_l - Y_{l-1}] \tag{13}$$

for each level we generate N_l i.i.d Brownian paths.
To ensure that $MSE \leq 2\epsilon^2$, we must make sure $\mathbb{V}[Y_L] \leq \epsilon^2$ and $(\mathbb{E}[Y_L] - \mathbb{E}[P])^2 \leq \epsilon^2$.
The test for weak convergence tries to ensure that $(\mathbb{E}[Y_L] - \mathbb{E}[P])^2 \leq \epsilon^2$ is bounded by ϵ^2, that is:

$$|\mathbb{E}[Y_L] - \mathbb{E}[P]| \leq \epsilon \tag{14}$$

A before, theorem (XX), states at each level the weak error should converges as $\mathbb{E}[Y_l - Y_{l-1}] \sim 2^{-\alpha l}$. Then, the remaining error should be:

$$\mathbb{E}[Y_L - Y] = \textstyle\sum_{l=L+1}^{\infty} \mathbb{E}[Y_l - Y_{l-1}] = \displaystyle\frac{1}{2^{\alpha} - 1}\mathbb{E}[Y_L - Y_{L-1}] \leq \epsilon$$

L is chosen such that $(\mathbb{E}[Y_L] - \mathbb{E}[P])^2 \leq \epsilon^2$ and the constant of proportionality for N_l is chose such that $\mathbb{V}[Y_L] \leq \epsilon^2$.

As stated in theorem (1) of the problem sheet, we have that if there exist positive constants α, β, γ, c_1, c_2, c_3 and c_4, where $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$, such that:

1. $|\mathbb{E}[Y_l - Y_{l-1}]| \leq c_1 2^{-\alpha l}$
2. $\mathbb{E}[\hat{\mu}_l] = \mathbb{E}[Y_l - Y_{l-1}]$, $Y_{-1} = 0$
3. $V_l \leq c_2 2^{-\beta l}$

4. $C_l \leq c_3 2^{\gamma l}$

It is important to note that the MLMC theorem ensures that $MSE \leq \epsilon^2$ as long as the values of the constants are known, namely, the values of c_1 and c_2, which bound the weak convergence and the variance.

Taking the expression for the optimal N_l and total cost (C), we have that the total cost as a function of variance and cost at level l is given by:

$$C = \frac{1}{\epsilon^2}\left(\sum_{l=0}^{L}\sqrt{V_l C_l}\right)^2 \tag{15}$$

If $\sqrt{V_l C_l}$ gets larger with increasing level l, the most expensive part comes from the very last level $l = L$. Then $C \sim \dfrac{1}{\epsilon^2} V_L C_L$. Now, if $\sqrt{V_l C_l}$ decreases with increasing level l, the most expensive part comes from the very first level $l = 0$. As a result, $C \sim \dfrac{1}{\epsilon^2} V_0 C_0$. Lastly, if $\sqrt{V_l C_l}$ is constant from level to level,

When $\beta > \gamma$, the dominant computational cost is on the coarsest levels where $C_l \sim O(1)$ and $O(\epsilon^{-2})$ samples are required to achieve the desired accuracy; this is the result when we use the CMC approach.

When $\beta < \gamma$, the most expensive computational cost is at the finest levels. Because of condition (1), we have $2^{-\alpha L} \sim O(\epsilon)$, and hence $C_L = O(\epsilon^{-\gamma/\alpha})$. If $\beta = 2\alpha$, since $|\mathbb{E}[Y_l - Y_{l-1}]| \sim |\mathbb{E}\left[(Y_l - Y_{l-1})^2\right]| > (\mathbb{E}\left[Y_l - Y_{l-1}\right])^2$, then we have $O(1)$ samples at the finest level.

if $\beta = \gamma$, both the computational effort and the contributions to the overall variance, are equally spread across all levels.

We now give a sketch of the proof of theorem (1). We consider the case then the base of log is 2. This is based entirely on [1] and [2]. Here it also important to mention that the code used for this project is based on the python version reported on the website reported on such sources.

The idea is to select the number of levels L such that:

$$N_l = \left\lceil \alpha^{-1}\log_2(\sqrt{2}c_1\epsilon^{-1}) \right\rceil < \alpha^{-1}\log_2(\sqrt{2}c_1\epsilon^{-1}) + 1$$

and $2^{-\alpha}\dfrac{\epsilon}{\sqrt{(2)}} < c_1 2^{-\alpha L} < \dfrac{\epsilon}{\sqrt{(2)}}$. Combining this with points (1) and points (2), the weak error is bounded as: $\left(\mathbb{E}[Y_L] - \mathbb{E}[Y]\right)^2 \leq \dfrac{1}{2}\epsilon^2$. As the variance is bounded too as $\mathbb{V}[Y] \leq \dfrac{1}{2}\epsilon^2$, this gives us that $MSE < \epsilon^2$.

We should consider three cases depending on the value of β.

1. if $\beta = \gamma$, $N_l = \left\lceil 2\epsilon^{-2}(L+1)c_2 2^{-\beta l} \right\rceil$ and $\mathbb{V}[Y] = \sum_{l=0}^{L} V_l N_l^{-1} \leq \sum_{l=0}^{L} N_l^{-1} c_2 2^{-\beta l} \leq \dfrac{1}{2}\epsilon^2$. Then, the cost $C_{MLMC} \leq \epsilon^{-2}(\log\epsilon)^2$.

2. if $\beta = \gamma$, $N_l = \left\lceil 2\epsilon^{-2}c_2(1 - 2^{-(\beta-\gamma)/2})^{-1}2^{-(\beta+\gamma)l/2} \right\rceil$. Then, $\mathbb{V}[Y_L] \leq \dfrac{1}{2}\epsilon^2$ and $C_{MLMC} \leq \epsilon^{-2}$.

3. if $\beta < \gamma$, $N_l = \left\lceil 2\epsilon^{-2}c_2 2^{(\gamma-\beta)L/2}(1 - 2^{-(\gamma-\beta)/2})^{-1}2^{-(\beta+\gamma)l/2} \right\rceil$. Then, $\mathbb{V}[Y_L] \leq \dfrac{1}{2}\epsilon^2$ and $C_{MLMC} \leq \epsilon^{-2-(\gamma-\beta)/\alpha}$.

Point b)

Implement a standard Monte Carlo estimator for payoff (1). To do so, implement an Euler-Maruyama scheme to evolve the asset price $S(t)$ with $h = 0.10$. Investigate your results for different sample sizes N. Investigate the bias and variance of your estimator with respect to the exact solution given by the Black-Scholes formula:

The Euler-Maruyama scheme is: $S_{n+1} = S_n\left(1 + rh + \sigma\Delta W_n\right), n = 1, \ldots$

$$\mathbb{E}\left[Y\right] = \Phi(d_1)S(t) - \Phi(d_2)K\mathrm{e}^{-r(T-t)} \tag{16}$$

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S(t)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right] \tag{17}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{18}$$

with $\Phi(\cdot)$ the is the cumulative distribution function of a standard normal whose parameters are as stated before.

When computing the expected value of payoff(1) using CMC, we obtained that $\mu_{CMC} = 0.1045$. We computed 5 runs with different samples.

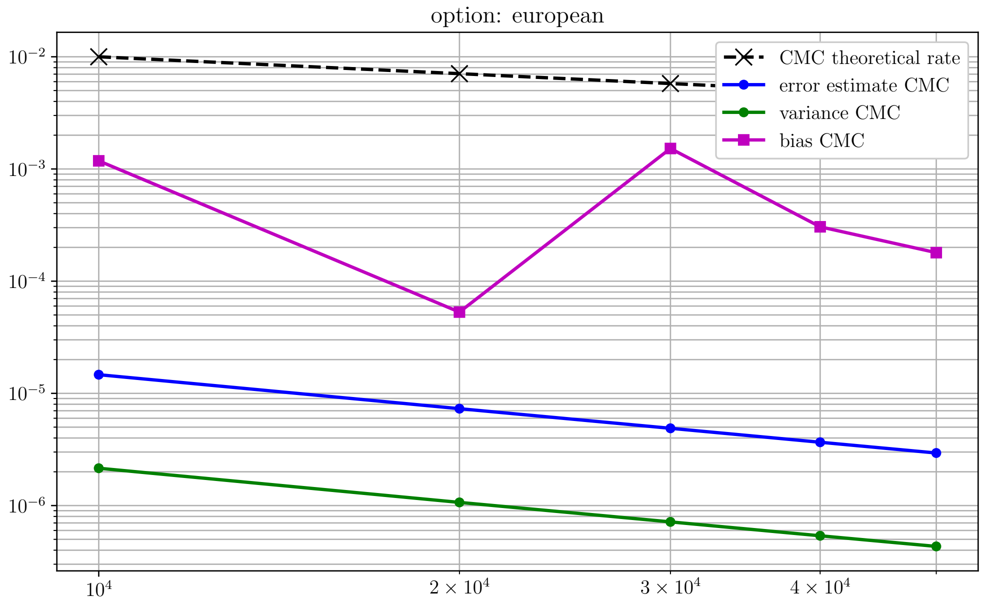


Figure 1: CMC for European option.

And the following values were recorded:

Samples N	μ_{CMC}	$\mathbb{V}\left[\mu_{CMC}\right]$
10000	0.1033	2.1518e-06
20000	0.10455	1.0671e-06
30000	0.1029	7.153e-07
40000	0.1042	5.3798-07
50000	0.1046	4.3254e-07

Table 1: Results for CMC, European option.

The average expected value using CMC is $\mu_{CMC} = 0.1039$. The expected value for the same pricing option using (16) is $\mu_{BS} = 0.1045$. We observe that the values μ_{CMC} are close to the one computed via the Black-Scholes formula. The averaged bias is 0.00064.

Point c)

We consider that the cost of computing a single sample is C, so the total cost in CMC is $N_{CMC}C_{CMC}$. Since we want to estimate the variance reducing when implementing the 2-level MC with the same cost as the CMC, we can write this like:

$$N_{CMC}C = N_0C_0 + N_1C_1 = \frac{C}{2} + N_1C \Rightarrow N_{CMC} = \frac{N_0}{2} + N_1 \tag{19}$$

where we have used the fact that $C_0 = \dfrac{C_1}{2}$. Besides, from (6), assuming the same value for C_0, we obtain: $N_1/N_0 = \sqrt{V_1/2V_0}$. Combining this with (19):

$$N_{CMC}\left(\frac{1}{2} + \sqrt{V_1/2V_0}\right)^{-1} = N_0$$

Now, we first estimate N_1 and N_0 based on a pilot run using two grids $h_1 = 0.1$ and $h_0 = 0.2$. From this pilot run, we compute the quantity $\sqrt{V_1/2V_0}$ and substitute it in the equation above for N_0. Then, we compute the value N_1 according to $N_1/N_0 = \sqrt{V_1/2V_0}$. Finally, these values for N_1 and N_0 are used to perform a two-level MLMC, which yields to the following figure and chart for different values of samples N_{CMC}:

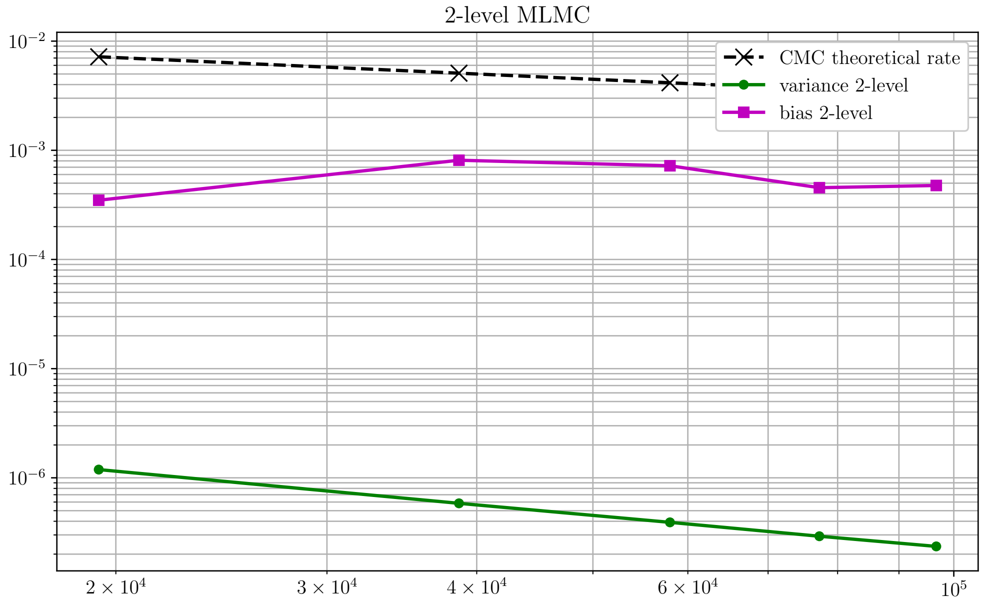


Figure 2: 2-level MLMC for European option.

Samples N_{CMC}	μ_{CMC}	$\mathbb{V}\left[\mu_{CMC}\right]$	$\mu_{CMC}^{2-level}$	$\mathbb{V}\left[\mu_{CMC}^{2-level}\right]$
10000	0.1033	2.1518e-06	0.1048	1.1896e-06
20000	0.10455	1.0671e-06	0.1053	5.8399e-07
30000	0.1029	7.153e-07	0.1037	3.9102e-07
40000	0.1042	5.3798-07	0.1040	2.9171e-07
50000	0.1046	4.3254e-07	0.1040	2.3545e-07

Table 2: Results comparison for CMC and 2-level MLMC.

The expected values is then average $\mu_{CMC}^{2-level} = 0.1044$.

As observed, the variance reduction using the two-level MLMC is: . It is also worth mentioning that the number of samples to estimate the number of samples for the the first estimate of the optimal N_0 and N_1 is one over 5 times the ones using in CMC. The more samples we take, the bigger the variance reduction, as expected. From the same chart, the variance reduction corresponds to. That is, we reduce the variance by using 1/5 of the number of samples for CMC.

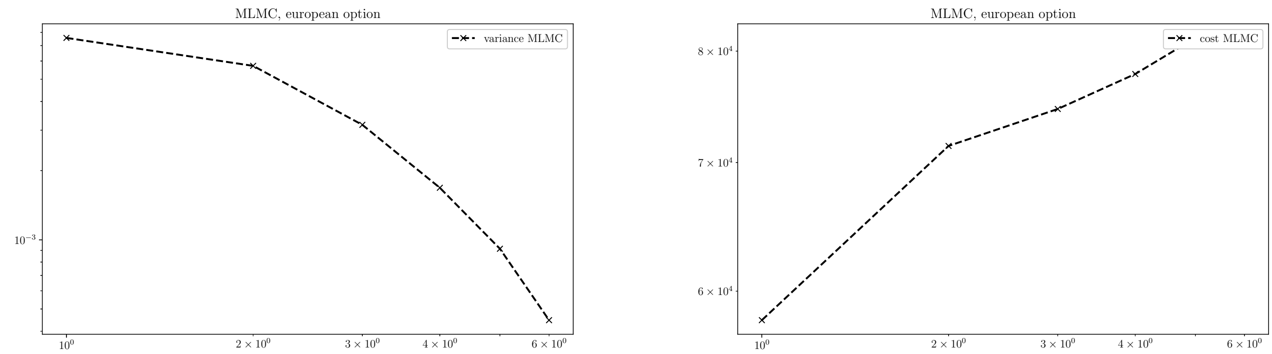
When comparing results in b) and c), we see that the CMC method has a total cost (given as the total number of samples for each run) of 150000 whereas the 2-level estimator has a total cost of 289774. However, for such an estimator the variance is much smaller that the one computed by CMC. The ratio of $N_1/N_0 \simeq 0.037$. This means that almost all computational effort is devoted to perform the estimation at the coarsest level N_0. We obtain a variance reduction using fewer samples at the least expensive level, which reduces the computational effort. It is important to point out that the bias for both CMC and 2-level are quite similar, which on average is 0.00056 for the 2-level estimator. We have a reduction of a half in variance with respect to the CMC.

Point d)

We performed our computation using $l = 1, ..., L$. This means we do not compute using $h = 0$. Our goal: to have a $MSE \leq \epsilon^2$; $\epsilon = 0.001, 0.0005$. $\alpha = 1$, $\beta = 1$ and $\gamma = 1$. The initial number of samples is 1000 for the three first levels.

European option:

We show the plot for $\epsilon = 0.001$:



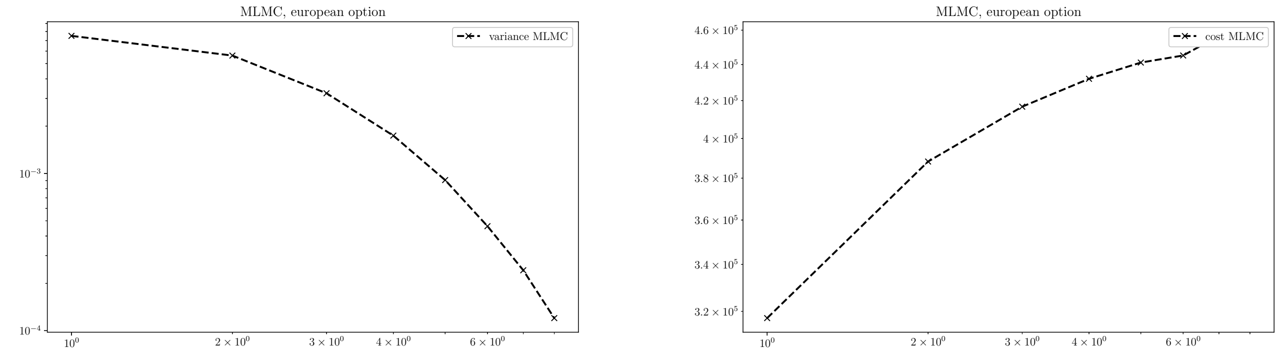
(a) Variance.

(b) Cost.

Figure 3: Cost and variance MLMC, European option.

$\mu_{MLMC} = 0.1027$, variance $V = 9.893e-07$ $L = 6$ and a total cost of 444407 samples including all levels.

We show the plot for $\epsilon = 0.0005$:



(a) Variance.

(b) Cost.

Figure 4: Cost and variance MLMC, European option.

$\mu_{MLMC} = 0.1044$, variance $V = 2.49864e - 07$ $L = 8$ and a total cost of 3351768 samples including all levels.

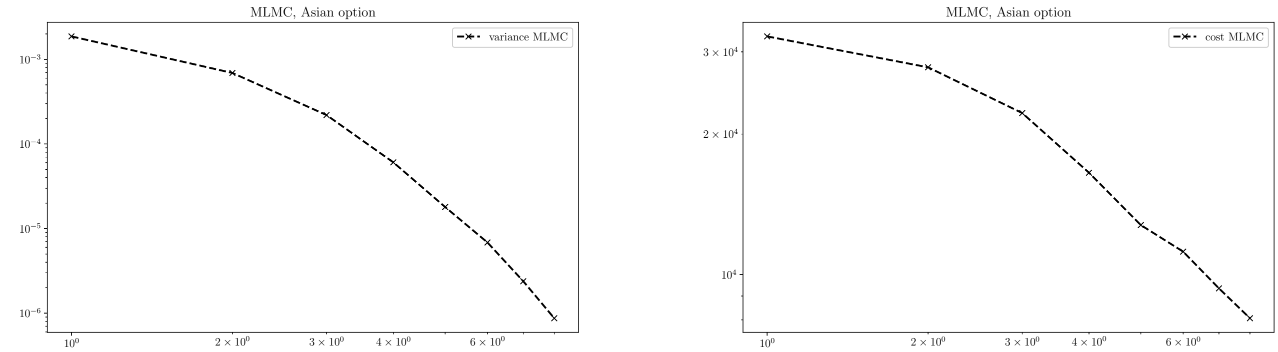
Point e)

Based on theorem (1), we know that the the weak error, variance and cost from one coarse level to a finer one should decay as $m_l \sim 2^{-\alpha} m_{l-1}$, $V_l \sim 2^{-\beta} V_{l-1}$ and $C_l \sim 2^{l\gamma}$, respectively. Consequently, to compute the values α, β and γ which ensure the expected behavior, we use the least squares method taking into account that we work with log_2. In [1], it is mentioned that for robustness, after performing the least squares method to compute such values, the estimates for m_l and V_l are not allowed to decrease by more than factor 0.50 relative to this anticipated value.

To compute the values of α and β for pricing options 2 and 3, we use the first 3 levels and $\epsilon = 0.001$ and $\epsilon = 0.005$ respectively; initial samples in each level $N = 1000$. We always use $\gamma = 1$.

Asian option

$\alpha = 1.0829$ and $\beta = 1.60408$



(a) Variance.

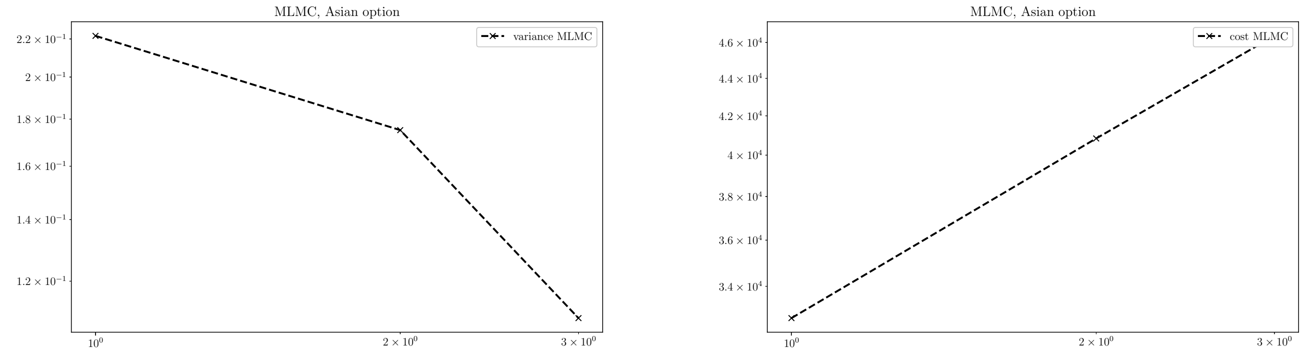
(b) Cost.

Figure 5: Cost and variance MLMC, Asian option.

$\mu_{MLMC} = 0.05726$, variance $V = 2.4870e - 07$, $L = 8$ and a total cost of 140265 samples including all levels.

Digital option

$\alpha = 1.0829$, $\beta = 1.60408$ and $\epsilon = 0.005$.



(a) Variance. (b) Cost.

Figure 6: Cost and variance MLMC, digital option.

$\mu_{MLMC} = 0.5351$, variance $V = 2.4797e - 05$, $L = 3$ and a total cost of 119898 samples including all levels.

For Asian and digital option, α and β sometimes result in values less than 0.5. By taking these during the simulation, we do not keep the rate according to theorem 1.

Point f)

Consider now a higher strike price $K = 25$. For each of the three payoffs, compute a (crude) Monte Carlo estimator of $\mathbb{E}[Y]$. Propose and implement a Variance Reduction Technique (VRT) for such an estimator and report your results. Can this VRT be used in the context of MLMC as well?

Given the three payoffs, it is almost impossible that the value $\mathbb{E}[Y]$ computed by the CMC method gives a value different from 0 since $S(T = 1)$ is mostly unlike to be such that $S(T = 1) \geq 25$ given $S_0 = 1$.

When impleenting the CMC, the expected value is zero during CMC simulation.

As a result, we can try moving the expected value $\mathbb{E}[S(T = 1)]$ toward $K = 25$, which can be achieved via importance sampling method.

Bibliography

[1] Cliffe, K. A., Giles, M. B., Scheichl, R., & Teckentrup, A. L. (2011). Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. Computing and Visualization in Science, 14(1), 3.

[2] Giles, M. B. (2015). Multilevel monte carlo methods. Acta Numerica, 24, 259.