Monte Carlo Call Options

Abhishek Ranjan Dey

1 Vanilla Call Option

1.1 Introduction & Problem Formulation.

The task is to value a financial contract written on an underlying asset S. It is known that the asset at the current time has value S_0, the trader has found, under a risk-neutral process on the underlying asset, that the asset at time T follows the distribution:

$$S_T \sim \mathcal{N}(f(S_0, T), v^2(S_0, T)T). \tag{1}$$

Where $f(S_0, T)$ is the risk-neutral mean, and $v(S_0, T)$ is the volatility. These functions are known under the risk-neutral distribution to be:

$$f(S_0, T) = S_0(e^{\alpha T} - e^{\beta T}) + \theta(1 + \tanh(\alpha T) - \beta T), \qquad v(S_0, T) = \sigma e^{\beta T} S_0^2 \theta^{-\gamma}. \tag{2}$$

For market fit parameters $\theta = 600.954$, $\alpha = 0.01$, $\beta = 0.02$, $\gamma = 1.1$. The option has a known payoff at time T (Expiry).

$$C(S_T, T) = g(S_T) = \begin{cases} X_2 & \text{if } S_T < X_1 \\ S_T - X_1 & \text{if } X_1 \leq S_T < X_2 \\ S_T & \text{if } S_T \geq X_2 \end{cases} \tag{3}$$

At current time $t = 0$, $S_0 = 601.14$, and the risk-free interest rate is $r = 0.03$. The option expires at $T = 1$, and has strike prices $X_1 = 600$ and $X_2 = 650$. The option itself also has a volatility of $\sigma = 0.18$. The task is to find the value of this option at the current time $t = 0$.

This value $V \equiv V_0 = e^{-rT}\mathbb{E}[g(S_T)]$ is the discounted future payoff and can be found analytically and also through a Monte Carlo valuation. We know that under the risk-neutral measure, S_T follows equation 1, which follows the Gaussian probability density function, meaning the expectation can be written explicitly.

$$\begin{aligned} V = e^{-rT}\mathbb{E}[g(S_T)] &= e^{-rT}\int_{-\infty}^{\infty} g(z)p(z)dz = e^{-rT}\int_{-\infty}^{\infty} g(z)\left[\frac{1}{\sqrt{2\pi v^2 T}}\exp\left(-\frac{(z-f)^2}{2v^2T}\right)\right]dz \\ &= e^{-rT}\frac{1}{\sqrt{2\pi v^2 T}}\int_{-\infty}^{\infty} g(z)\exp\left(-\frac{(z-f)^2}{2v^2T}\right)dz. \end{aligned} \tag{4}$$

The resulting integral is analytic and can be well-approximated through basic quadrature methods. Monte Carlo valuation can also be used, using the standardisation property of normal distributions. Since $S_T \sim \mathcal{N}(f, v^2T)$ under the risk-neutral measure, it can be expressed in terms of the standard normal variable as

$$S_T = f(S_0, T) + v(S_0, T)\sqrt{T}\cdot\phi \tag{5}$$

By generating independent samples $\phi_i \sim \mathcal{N}(0, 1)$, the option value can be approximated by a sum over N of the standardised values 5. Due to the Law of Large Numbers, this sum converges in the expectation [1], [2].

$$V = e^{-rT}\mathbb{E}[g(S_T)] = e^{-rT}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} g(S_T^i) \tag{6}$$

This allows for an approximation to this expectation through a discrete number N of samples of S_T, provided N is large enough.

$$S_T^i = f(S_0, T) + v(S_0, T)\sqrt{T}\phi_i, \qquad V \approx \hat{V} = e^{-rT}\frac{1}{N}\sum_{i=1}^{N} g(S_T^i) \tag{7}$$

1.2 Standard Monte Carlo

Equation 7 is implemented for varying values of N in Figure 1, and the integral in Equation 4 is plotted as a horizontal red line. At $N = 100,000$ samples (paths), the Monte Carlo valuation of the option is $V_{mc} = 465.863741$. In this case, the integral can be computed through quadrature, which gives $V_{int} = 466.347811$. The analysis is repeated 5 times per N, showing how the simulation converges as N increases.

At low values of N, the simulation exhibits very high variance. However, the dispersion of simulated values gets lower as N increases, with the lowest variance seen at N =10 million. However, it should be noted that even at N =100,000 paths, there is still a large dispersion in the simulation, showing that basic Monte Carlo methods require a large number N of paths to see a result with good variance. The basic Monte Carlo simulation has a very slow rate of convergence, with its variance scaling as $\text{Var(V)} \propto \frac{1}{N}$[2]. This means that reducing the standard deviation of estimates by a factor of 10 requires the simulation of 100 more paths.

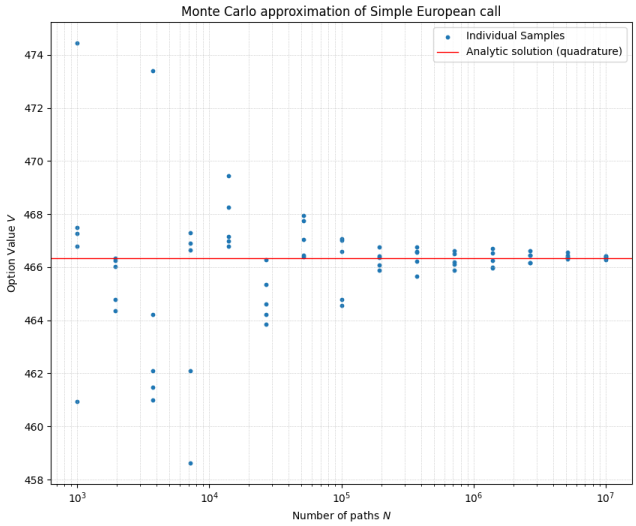

Figure 1: Monte Carlo approximation of the contract value for $N = 1000, .., 10,000,000$ paths. Each N is repeated 5 times

1.3 Variance Reduction Techniques

Monte Carlo, as outlined in equation 7, relies on generating i.i.d samples S_T^i. The standard method leads to variance that scales inversely to the number of paths N. This motivates generating another sequence of S_T^i, but with more minor variance in the estimator[3].

If the new technique reduces the variance by a factor R, the same level of accuracy can be achieved with only $\frac{1}{R}$ as many simulated paths.

1.3.1 Antithetic Variates

Instead of generating i.i.d samples, antithetic variates pair each sample with its negative (antithetic) counterpart, exploiting the symmetry in the normal distribution. This can be interpreted as generating only $N/2$ samples of ϕ_i, generating S_T^i regularly, and calculating the rest of the samples $S_T^{i'}$ with $-\phi_i$. The Monte Carlo estimator, therefore, can be written as:

$$\hat{V}_{antithetic} = \frac{1}{N}\sum_{i=0}^{N/2}\frac{g(S_T^i) + g(S_T^{i'})}{2}. \tag{8}$$

The symmetry of the normal distributions means that $\phi_i \sim -\phi_i$, meaning that the marginal distributions for both are the same. This implies that the variance of S_T^i and $S_T^{i'}$ are equal, and so is the variance of the payoff dependent on them. This leads to writing the variance of the fractional term within the estimator [3]:

$$\mathrm{Var}\left(\frac{g(S_T^i) + g(S_T^{i'})}{2}\right) = \frac{1}{2}\left(\mathrm{Var}(g(S_T^i)) + \mathrm{Cov}(g(S_T^i), g(S_T^{i'}))\right). \tag{9}$$

In the case that g is monotonic, the covariance will be exclusively negative, leading to a maximal variance factor $R = \frac{1}{2}$[3], meaning that if g is monotonic, the same variance can be achieved with antithetic variates using only half of the simulated paths.

1.3.2 Moment Matching

Moment matching introduces non-i.i.d antithetic variates, which ensures for the constructed set $\phi = (\phi_1, -\phi_1, ..., \phi_{N/2}, -\phi_{N/2})$, that the first moment, or expectation, is preserved, i.e. $\mathbb{E}[\phi] = 0$. It is also possible to scale for the second moment of ϕ, since the first moment is zero, the variance of ϕ is:

$$\mathrm{Var}(\phi) = \frac{1}{N}\sum_{i=1}^{N}\phi_i^2. \tag{10}$$

The second moment can be forced by normalising ϕ by scaling it with its standard deviation $\frac{\phi}{\sqrt{\mathrm{Var}(\phi)}} = \tilde{\phi}$, where $\tilde{\phi}$ are now normalised to be the $\sim \mathcal{N}(0,1)$ random variables sampled to generate paths of the underlying asset. This method preserves the statistical properties of the standard normal distribution.

1.3.3 Halton Sequences

The Halton sequences are numbers that cover [0,1] with low discrepancy. This means that numbers generated from the Halton sequence cover the interval very close to uniformly [4]. This is in contrast to sampling the uniform distribution $U([0,1])$ using a random number generator, which can result in clustering [5], and therefore some regions on the interval may be under-represented, and equally, others may be oversampled.

The low discrepancy Halton sequences ensure a closer to uniform coverage of [0,1]. This space is also the space of probabilities, meaning that Halton sequences can be used to map directly back to the normal distribution:

$$\phi_i = \Phi^{-1}(x_i). \tag{11}$$

Where x_i is the $i-$th element of the Halton sequence. This method means that the generated ϕ_i are a much better representation of the normal distribution. This distribution is called a quasi-random distribution, as the Halton sequences are deterministic. When using these quasi-random ϕ_i to generate sample paths for the asset, the resulting Monte Carlo is referred to as a Quasi-Monte Carlo simulation (QMC). Using the Quasi-Monte Carlo simulation leads to the variance scaling $\mathrm{Var}(\hat{V}) \propto \frac{1}{N^2}$, which is a much faster rate of convergence[4].

1.3.4 Results

Figure 2 displays how each variance reduction technique approximates the option value V as N varies. The Halton sequence quickly approaches the integral solution with only $N = 10^4$ paths. The other methods converge to the solution at $N = 10^5$ paths or higher. Otherwise, the variance reduction techniques approach the analytic solution at the same rate as the standard Monte Carlo.

Figure 4 displays how the variance changes with N; it is seen that the Monte Carlo methods all follow the $O(N^{-1})$ trend in the variance, with moment marching and antithetic variates reaching a low variance faster than the standard Monte Carlo. The Halton sequence immediately has a very low variance (<1): the low discrepancy coverage of [0,1] means that the generated samples ϕ_i also have very low variance, which is reflected when generating Monte Carlo sample paths. Figure 3 shows that on closer inspection, the QMC still follows the power law $O(N^{-1})$.

Table 1 displays a metric called efficiency. Here, the efficiency for a method is defined to be:

$$\text{Efficiency} = \frac{\text{MSE}}{\text{Simulation time}} = \frac{\text{Var}(\hat{V}) + (\mathbb{E}[\hat{V}] - V_{true})^2}{\text{Simulation time}}. \tag{12}$$

The definition of MSE used can be found in [6], and V_{true} is the high accuracy quadrature integral found from equation 4. As seen in equation 12, this formulation takes into account the variance of each model, how close to the true solution each model gets (bias), and the length of time each simulation takes. On inspecting table 1, at low N, moment matching outperforms both standard and antithetic variates. This behaviour is maintained throughout most values of N. However, throughout all values of N, it is seen that Halton has the greatest efficiency. Figures 4 and 2 tell us why this behaviour is seen, due to moment matching having a lower bias per for each N, and also having a strong power-law behaviour in its convergence, it on-paper outperforms each model.

However, it is clear that due to the extremely fast convergence for low N, Halton is the most efficient method. It achieves the lowest variance and, therefore, an efficiency much lower than any other Monte Carlo method.

N	Standard	Antithetic	Moment Matching	Halton
1000	3.865327e+05	7.442357e+05	3.353812e+05	2.556821e+02
2310	4.218840e+05	8.946775e+04	1.535982e+05	1.423954e+02
5336	2.944735e+05	2.638501e+05	9.285005e+04	1.420814e+01
12328	2.635316e+04	5.067401e+04	2.490806e+04	4.463153e-01
28480	1.323647e+04	4.162091e+03	2.440246e+03	8.141439e-02
65793	1.149507e+03	9.580172e+02	9.781337e+02	3.803697e-03
151991	1.910314e+02	1.188583e+02	2.394442e+01	5.964322e-04
351119	3.336536e+01	2.302379e+01	4.195861e+00	5.449647e-05
811130	4.681051e+00	5.116595e+00	3.320961e+00	1.062476e-05
1873817	8.619332e-01	7.009624e-01	2.709181e-01	2.617514e-07
4328761	1.883382e-01	1.629537e-01	7.426380e-02	2.708328e-08
10000000	2.111067e-02	1.763038e-02	5.062609e-03	6.580462e-09

Table 1: Efficiencies for each Monte Carlo method.

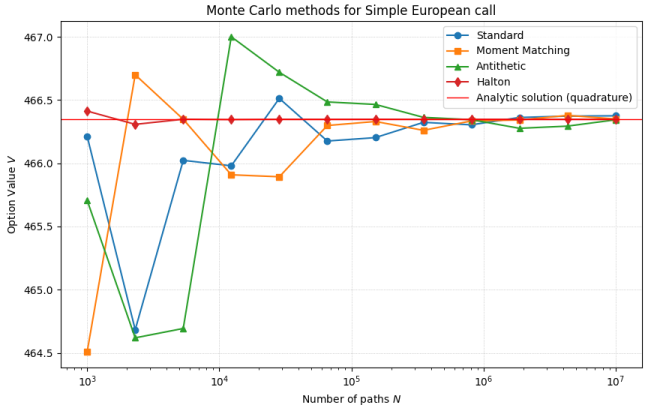

Figure 2: Variance Reduction techniques and their approach to the analytic integral. Each point is an average of 10 paths for each N.

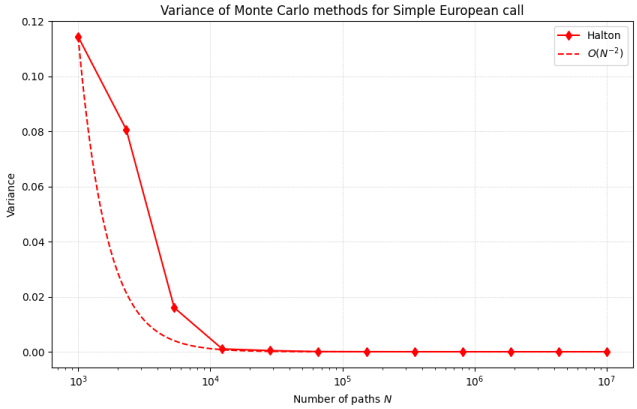

Figure 3: The variance of the Halton sequence with N for the QMC.

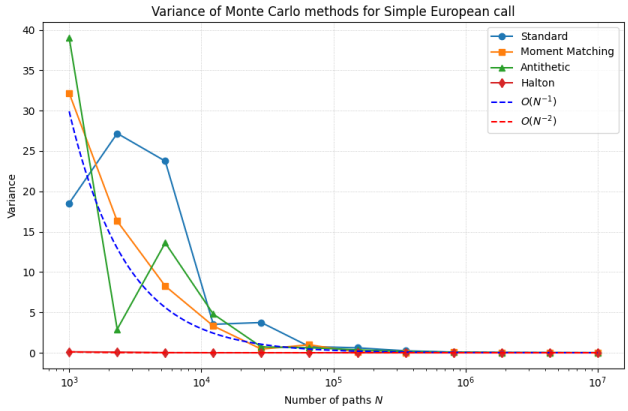

Figure 4: The variance dependency on the number of paths N.

2 Maximum Strike Lookback Option

2.1 Introduction and Problem Formulation

The task is to find the value of a financial contract again written on an underlying asset S. This option is path-dependent and, therefore, follows a risk-neutral stochastic differential equation (SDE):

$$dS = f(S,t)dt + v(S,t)dW_t. \tag{13}$$

Path dependent options depend on $S(t_k)$, where t_k is a set of $K+1$ equally spaced samples $\{t_k\}_{k=0}^{k=K}$, where $t_0 = 0, t_K = T$. The function $f(S,t)$ is the deterministic drift under the risk-neutral measure, ensuring proper discounting, while $v(S,t)$ is the diffusion term, which governs the randomness in the asset's price. Under the risk-neutral process, the functions, in this case, are:

$$f(S,t) = \alpha\theta - \beta S, \qquad v(S,t) = \sigma(|S|)^\gamma. \tag{14}$$

The model parameters are $\theta = 6998.53$, $\alpha = 0.02$, $\beta = 0.05$, $\gamma = 0.99$, and volatility $\sigma = 0.43$. The risk-free interest rate is $r = 0.01$.The SDE can, therefore, be rewritten in terms of their increments:

$$\begin{aligned} S(t_k) - S(t_{k-1}) &= f(S(t_{k-1}), t_{k-1})(t_k - t_{k-1}) + v(S(t_{k-1}), t_{k-1})(W_{t_k} - W_{t_{k-1}}) \\ &= f(S(t_{k-1}, t_{k-1})\Delta t + v(S(t_{k-1}), t_{k-1})\sqrt{\Delta t}\phi_{k-1}. \end{aligned} \tag{15}$$

In the final line, a few properties are used and defined. $\Delta t = \frac{T}{K}$ is the time increment used, and is exactly $t_k - t_{k-1} \forall\ t \geq 1$. Further, the stationary increments of brownian motion is used, $W_{t_k} - W_{t_{k-1}} \sim W_{\Delta t} \sim \mathcal{N}(0, \Delta t) \sim \sqrt{\Delta t}\phi_{k-1}$. The scheme outlined in equation 15 is the Euler-Maruyama scheme for solving an SDE[7]. The subscript $k-1$ notation for ϕ emphasises the iterative nature of this scheme, where each step depends on the previously computed value of S. The process constructs a discrete approximation of the stochastic trajectory over the time interval $[0, T]$. In path-dependent options, the entire sequence of asset prices $S(0) = S_0, S(t_1), .., S(t_K) = S_T$ is used to compute the final payoff.

The option being priced is the discrete maximum fixed-strike lookback call option V, it expires at $T = 1.25$, and the paths are discretised over $K = 75$ equally spaced time observations. The fixed strike price is $X = 700$, and the option payoff is:

$$g(S(t_0), ..., S(t_K)) = \max\left(\max_k(S(t_k)) - X, 0\right). \tag{16}$$

This payoff depends on all values of S up to time T; therefore, the Monte Carlo estimator must also depend on this. Taking equation 15 and indexing a simulated path (an instance of $(S_0, ..., S_T)$) with i, the Monte Carlo estimator used with the Euler-Maruyama method can be written, where the average is taken over N paths.

$$S^i(t_k) = S^i(t_{k-1}) + f(S^i(t_{k-1}), t_{k-1})\Delta t + v(S^i(t_{k-1}), t_{k-1})\sqrt{\Delta t}\phi_{k-1}, \tag{17}$$

$$V = e^{-rT}\mathbb{E}[g(S(t_0), ..., S(T), t)] \approx \hat{V} = \frac{1}{N}\sum_{i=1}^{N} g(S^i(t_0), S^i(t_1), ..., S^i(t_K)). \tag{18}$$

2.2 Standard Path-dependent Monte Carlo

Equations 17 and 18 are implemented for varying values of N in Figure 5. The value of the option for $N = 100,000$ paths is $V = 2302.117972$.

Figure 5 displays the behaviour of the Monte Carlo with varying N and how $\text{Var}(\hat{V})$ changes. The interval from 10 to 1,000,000 is evenly logarithmically sampled, and each N is repeated 10 times to create a set of values with variance. The estimates' variance for different N decreases as N^{-1}, as seen in figure 6,

meaning the earlier discussed convergence behaviour of the simulation is still followed. The best estimate for the option value is that with the lowest variance, which is with $N = 1,000,000$ paths, giving a value:

$$\hat{V} = 2311.52316 \pm 2.607824. \tag{19}$$

With a 95% confidence.

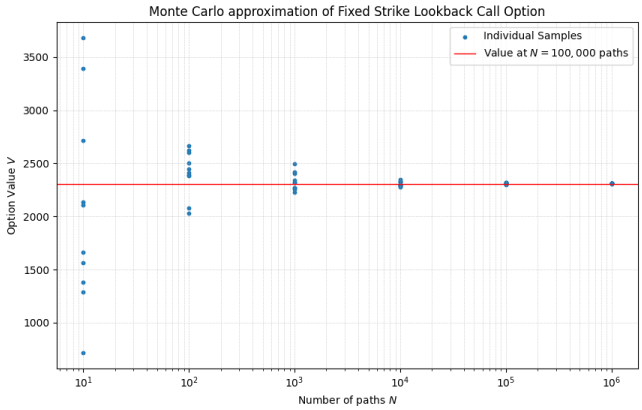

Figure 5: The value of the path-dependent option with varying N, the horizontal line is the calculated $\hat{V}$ for 100,000 paths.

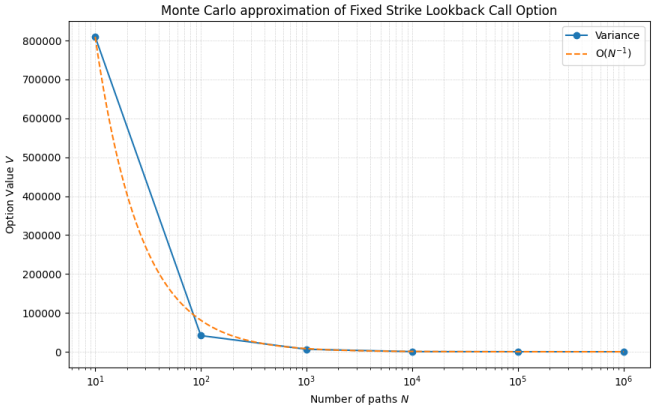

Figure 6: The standard deviation of the path dependent simulation for varying N

2.3 Path Dependent Derivatives

2.3.1 One-Sided & Central differences

The task further is to estimate the partial derivative $\frac{\partial V}{\partial \sigma}$. This can be estimated with a finite difference:

$$\frac{\partial V}{\partial \sigma} = \frac{V(S_0, t = 0; \sigma = 0.43) - V(S_0, t = 0; \sigma = 0.43 + \Delta\sigma)}{\Delta\sigma}. \tag{20}$$

For an increment of $\Delta\sigma$, this method is known as the one-sided forward differences. Its accuracy to the true derivative can be found by expanding $V(S_0, t = 0)$ around $\sigma = 0.43$ in a Taylor series.

$$V(S_0, t = 0; \sigma + \Delta\sigma) = V(S_0, t = 0; \sigma) + \Delta\sigma \frac{\partial V}{\partial \sigma} + \frac{\Delta\sigma^2}{2} \frac{\partial^2 V}{\partial \sigma^2} + O(\Delta\sigma^3)$$
$$\frac{\partial V}{\partial \sigma} = \frac{V(S_0, t = 0; \sigma) - V(S_0, t = 0; \sigma + \Delta\sigma)}{\Delta\sigma} - \frac{\Delta\sigma}{2} \frac{\partial^2 V}{\partial \sigma^2} \tag{21}$$

Meaning that forward differences has a truncation error of $O(\Delta\sigma)$, this is referred to as first order accuracy. When running forward differences on the option with $\Delta\sigma = 1 \times 10^{-12}$, for $N = 100,0000$ paths, the derivative is valued at:

$$\frac{\partial V}{\partial \sigma} = 6303.253031. \tag{22}$$

Figure 7 shows the result of a Monte Carlo estimator for the derivative, indexing an individual valuation of equation 20 with i, The Monte Carlo estimation is given by [8]:

$$\mathbb{E}\left[\frac{\partial V}{\partial \sigma}\right] = \frac{1}{M} \sum_{i=1}^{M} \frac{\partial V_i}{\partial \sigma} \qquad \text{Var}\left(\frac{\partial V}{\partial \sigma}\right) = \frac{1}{M-1} \sum_{i=1}^{M} \left(\mathbb{E}\left[\frac{\partial V}{\partial \sigma}\right] - \frac{\partial V_i}{\partial \sigma}\right)^2. \tag{23}$$

Implementation of this, with the $\pm$ 1 standard deviation interval, is plotted in figure 7. Each N has $M = 100$ valuations of the derivative. The convergence behaviour implies that the value of the derivative does converge to some value.

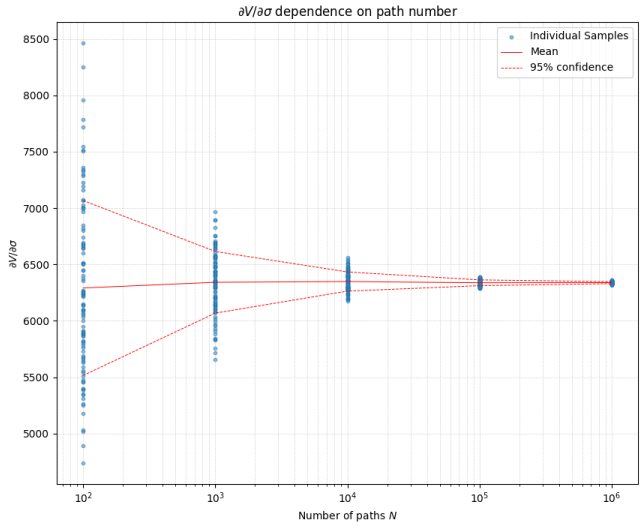

Figure 7: Convergence of the derivative with N for $M = 100$ per N.

A simple extension of the forward differences is to value the derivative with central differences.

$$\frac{\partial V}{\partial \sigma} = \frac{V(S_0, t = 0; \sigma = 0.43 - \Delta\sigma) - V(S_0, t = 0; \sigma = 0.43 + \Delta\sigma)}{2\Delta\sigma}. \tag{24}$$

Through a similar analysis, this method is accurate to $O(\Delta\sigma^2)$, meaning that it should result in a more stable estimate of the derivative for small $\Delta\sigma$, as the error scales quadratically.

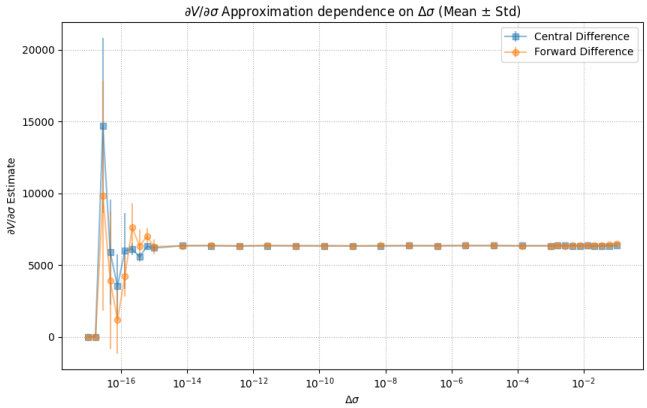

Figure 8: The value of the derivative with varying $\Delta\sigma$, all are at $N = 100,000$, where error bars represent standard deviation over 5 samples

Figure 8 displays the behaviour of both methods over the range 10^{-17} to 1 The value of the derivative is stable at $10^{-17} \leq \Delta\sigma \leq 10^{-1}$. At 10^{-17}, the estimated derivative is heavily influenced by numerical noise, making both methods produce near-zero estimates. At 10^{-1}, both methods start to increase. The value $\Delta\sigma = 10^{-12}$ is chosen as it falls in a range where numerical stability is maintained whilst remaining accurate. While central differences reduce bias and variance in deterministic cases, in Monte Carlo methods, their advantage is less pronounced due to the dominating stochastic variance.

2.3.2 Best Estimate of the derivative

From Section 1, the most efficient and accurate estimate came from using Halton sequences. Figure 9 shows this behaviour again but for path-dependent Monte Carlo valuations of these derivatives. Again, it is seen that the QMC with Halton sequences converges the fastest and consistently outperforms the typical Monte Carlo. The figure also shows the Halton Sequences underperforming compared to its usual power law behaviour, this could be due to the generation of two-dimensional Halton sequences having a slightly worse variance, following $O(\frac{\log(N)^2}{N^2})$ [4].

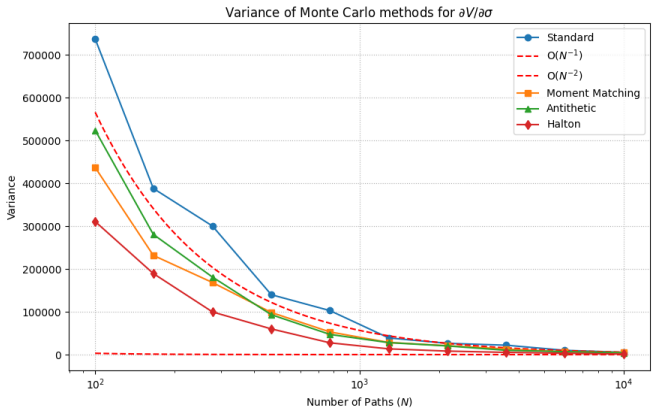

Figure 9: Variance of different Monte Carlo techniques with N.

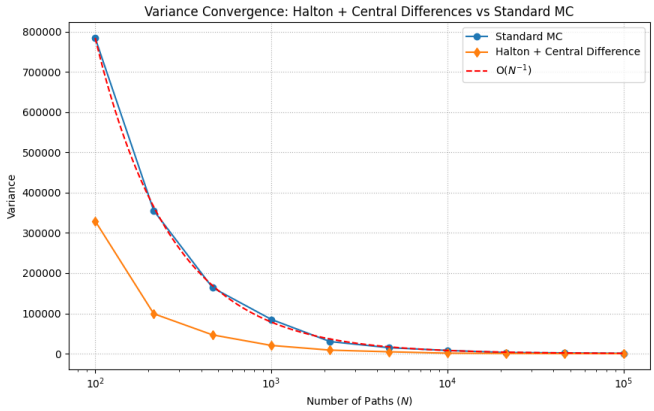

Figure 10: The variance of Standard MC and Halton with central derivatives to find the derivative.

Figure 10 displays the improvement of Halton sequences with central differences; the variance is lower throughout and decreases towards zero faster than the standard method.

The best prediction, therefore, that can be made for the derivative is the Quasi-Monte Carlo valuation using Halton sequences, for $N = 100000$ paths and $M = 100$ valuations of the derivative. This gives the value of the derivative:

$$\frac{\partial V}{\partial \sigma} = 6338.489129 \pm 7.215523 \tag{25}$$

References

[1] T. Budd, *Mathematical concepts in computing*, version 1.0, 2021. [Online]. Available: `https://hef.ru.nl/~tbudd/mct/mct_book.pdf`.

[2] J. C. Watkins, *Introduction to Probability and Statistics*. 2021, Chapter 10.2. [Online]. Available: `https://math.arizona.edu/~jwatkins/statbook.pdf`.

[3] D. J. Higham, *An introduction to financial option valuation*. Cambridge, England: Cambridge University Press, Apr. 2004, Chapter 26.

[4] S. Asmussen and P. W. Glynn, *Stochastic simulation: Algorithms and analysis* (Stochastic Modelling and Applied Probability), en, 2007th ed. New York, NY: Springer, Jul. 2007.

[5] S. Sen, T. Samanta, and A. Reese, "Quasi-versus pseudo-random generators: Discrepancy, complexity and integration-error based comparison," *Int J Innov Comput Info Control*, vol. 2, Jan. 2006.

[6] Z. Yang, Y. Yu, C. You, J. Steinhardt, and Y. Ma, *Rethinking bias-variance trade-off for generalization of neural networks*, 2020. arXiv: `2002.11328 [cs.LG]`. [Online]. Available: `https://arxiv.org/abs/2002.11328`.

[7] "Chapter 8: Euler—maruyama," in *An Introduction to the Numerical Simulation of Stochastic Differential Equations*, pp. 79–89. DOI: `10.1137/1.9781611976434.ch8`. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611976434.ch8`. [Online]. Available: `https://epubs.siam.org/doi/abs/10.1137/1.9781611976434.ch8`.

[8] D. J. Higham, *An introduction to financial option valuation*. Cambridge, England: Cambridge University Press, Apr. 2004.