Table of Results

k-Number	k24010992
Price by Monte Carlo	42.3908
Confidence interval for Price	$[42.3145, 42.4670]$
Delta by Monte Carlo	-0.7427
Confidence interval for Delta	[-0.7502,-0.7352]
Price by Finite Difference	42.4345
Delta by Finite Difference	-0.7499

1 Introduction and problem statement

The purpose of this essay is to calculate the price and delta of an option under a stock price process defined by a stochastic differential equation. The mathematical arguments, numerical methods and code used are based on John Armstrong's lecture notes [Arm24]. We consider a European put option with a strike price of $K = 186.00$ and maturity $T = 1.70$. The payoff function of this option is

$$(S_T - K)^- = \max(K - S_T, 0)$$

where S_T denotes the terminal value of the underlying price process $(S_t)_{t\geq 0}$. By assumption, S_t satisfies the stochastic differential equation:

$$dS_t = S_t\mu\, dt + S_t^\gamma \sigma(t)\, dW_t. \tag{1}$$

Where $S_0 = 140.00$, $\mu = 0.13$ and $\gamma = 0.99$ are constant and:

$$\sigma(t) = \begin{cases} 0.12 + 0.14 \times \frac{t}{0.60} & t < 0.60 \\ 0.26 & \text{otherwise} \end{cases}$$

(W_t) denotes the Brownian motion process. We also assume the existence of a risk-free asset bearing a continuously compounded interest rate of $r = 0.03$.

2 Derivation of a PDE for option price

We denote the option price, $V(S, t)$, as a function of the underlying stock price and time. Our stated aim is to compute $V(S_0, 0)$ and $\partial_S V(S_0, 0)$.

A PDE analogous to the Black-Scholes partial differential equation [BS73], can be derived to aid in pricing this option. The following argument is an adapted form of a derivation of the Black-Scholes PDE found in Martin Forde's lecture notes [For24] which in turn draw from [BK04].

Proposition 1. *$V(S,t)$ satisfies the following PDE:*

$$V_t(S,t) + rSV_S(S,t) + \frac{1}{2}S^{2\gamma}\sigma(t)^2 V_{SS}(S,t) = rV(S,t) \tag{2}$$

with boundary condition $V(S,T) = (S-K)^-$.

Proof. Considering $V(S_t,t)$, we apply Itô's lemma [Arm24, Topic 6] giving:

$$dV = (V_t + \mu S_t V_S + \frac{1}{2}S_t^{2\gamma}\sigma(t)^2 V_{SS})\,dt + V_S S_t^{\gamma}\sigma(t)\,dW_t. \tag{3}$$

Under self financing assumptions, the wealth, X_t, of a trader holding φ_t units of stock at time t evolves according to the differential form:

$$dX_t = \varphi_t dS_t + (X_t - \varphi_t S_t)rdt.$$

The first term models the effect of changes in the stock price, and the second models interest earned on remaining wealth over an infinitesimal time period dt. Expanding dS_t from (1) we rearrange to:

$$dX_t = (rX_t + \varphi_t S_t(\mu - r))dt + \varphi_t S_t^{\gamma}\sigma(t)dW_t.$$

If, as in the Black-Scholes model, X_t replicates $V(S_t,t)$ the process becomes:

$$dV = (rV + \varphi_t S_t(\mu - r))dt + \varphi_t S_t^{\gamma}\sigma(t)dW_t. \tag{4}$$

Equating the dW_t terms in (3) and (4), requires that $\varphi_t = V_S(S_t,t)$. Equating the dt terms in (3) and (4) yields the PDE in (2) with S_t in place of S. W_t which hits all points in finite time, thus in principle, the PDE must hold for any pairing of (S,t) where $S > 0$. The intrinsic value of the put option trivially satisfies the boundary condition. □

3 Monte Carlo Methods

3.1 The Feynman-Kac Theorem

We reproduce below from [Arm24] a key theorem which allows us to express the solution to our PDE as an expectation of a stochastic process.

Theorem 1. *(Feynman-Kac)*

"Suppose that X_t satisfies: $dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dW_t$, and that

$$v(x,t) := \mathbb{E}_t\left(e^{-c(T-t)}f(X_T) \mid X_t = x\right) < \infty,$$

for some constant c. Then, v solves the PDE

$$\frac{\partial v}{\partial t} + \mu(x,t)\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 v}{\partial x^2} - cv = 0,$$

with the final condition $v(x,T) = f(x)$. With appropriate growth conditions on v the converse is also true."

Applying this theorem to the PDE in proposition 1, we see:

$$V(S,t) = \mathbb{E}_t\left(e^{-r(T-t)}(\tilde{S}_T - K)^- \mid \tilde{S}_t = S\right) \tag{5}$$

where the process $\tilde{S}_t$ is such that $\tilde{S}_0 = S_0$, and evolves according to SDE:

$$d\tilde{S}_t = \tilde{S}_t r\,dt + \tilde{S}_t^\gamma \sigma(t)\,dW_t. \tag{6}$$

The strong law of large numbers tells that this expectation can be approximated by a sample average of payoff scenarios.

3.2 Naive Monte Carlo Simulation

We divide the time interval $[0,T]$ into $N+1$ uniformly spaced points:

$$t \in \mathcal{T} = \{0, \frac{T}{N}, \frac{2T}{N}, \ldots, T\}, \text{ where } N \in \mathbb{N} \text{ and } \delta t = \frac{T}{N}.$$

The Euler-Maruyama scheme to simulate a path of process (6) over $\mathcal{T}$, is:

$$\tilde{S}^{\mathcal{T}}_{t+\delta t} = \tilde{S}^{\mathcal{T}}_t + r\tilde{S}^{\mathcal{T}}_t \delta t + (\tilde{S}^{\mathcal{T}}_t)^\gamma \sigma(t)(W^{\mathcal{T}}_{t+\delta t} - W^{\mathcal{T}}_t), \quad \text{with } \tilde{S}^{\mathcal{T}}_0 = S_0.$$

The increment $(W_{t+\delta t} - W_t)$ is normally distributed with variance δt. Thus, given a sequence (Z_i) of i.i.d. standard normal variables the N-step path is:

$$\tilde{S}_{(i+1)\delta t} = \tilde{S}_{i\delta t} + r\tilde{S}_{i\delta t}\delta t + \tilde{S}^\gamma_{i\delta t}\sigma(i\delta t)\sqrt{\delta t}Z_i, \quad \text{for } i \in \{0,1,2,\cdots,N-1\}. \tag{7}$$

This may be repeated to produce M discrete sample paths for $\tilde{S}_t$, provided that all $M \times N$ instances of (Z_i) are i.i.d. Since put options are not path-dependent we need only consider the terminal stock prices which we will denote $\tilde{S}^{(j)}_T$ for $j \in \{1,2,\cdots,M\}$. The "naive" estimator is then given by:

$$V(S_0,0) \approx \mu_M := \frac{1}{M}\sum_{j=1}^{M} e^{-rT}(\tilde{S}^{(j)}_T - K)^- \tag{8}$$

For an $h \in \mathbb{R}$ small with respect to S_0, we can approximate $V(S_0+h,0)$ and $V(S_0-h,0)$ via the technique above. A central difference estimate then gives an approximation for the option delta:

$$\partial_S V(S_0,0) \approx \frac{V(S_0+h,0) - V(S_0-h,0)}{2h}. \tag{9}$$

Importantly for stability, the two estimates must use the same sample paths.

3.3 Confidence Intervals & Variance Reduction

The rate of convergence of a sample mean to the true expectation is governed by the central limit theorem, which gives a 2-sided p-% confidence interval:

$$\mu_M \pm \frac{1}{\sqrt{M}} \Phi^{-1} \left(\frac{1 - \frac{p}{100}}{2} \right) \sigma_M \tag{10}$$

Φ is the standard normal CDF, and σ_M is the sample variance of the payoff scenarios comprising μ_M. We discuss two methods of shortening this interval.

3.3.1 Antithetic Sampling

Antithetic sampling generates "thetic" stock price paths using Euler scheme (7), and a corresponding "antithetic" path for each, constructed by flipping the sign of each random increment Z_i. Let $\tilde{A}_t^{(j)}$ be the antithetic path of $\tilde{S}_t^{(j)}$.

$$\mu_{\text{antithetic}} = \frac{1}{2M} \sum_{j=1}^{M} e^{-rT} [(\tilde{S}_T^{(j)} - K)^- + (\tilde{A}_T^{(j)} - K)^-] \tag{11}$$

For a thetic estimator, X_1, and its antithetic counterpart X_2:

$$\text{Var}(\mu_{\text{antithetic}}) = \text{Var}\left(\frac{X_1 + X_2}{2} \right) = \frac{1}{4} \left(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2) \right).$$

Thus, if $\text{Cov}(X_1, X_2)$ is sufficiently negative then: $\text{Var}(\mu_{\text{antithetic}}) < \text{Var}(\mu_{\text{naive}})$.

3.3.2 Control Variate Technique

A control variate method to compute $\mathbb{E}[V]$ involves instead computing the expectation of $V^* = V + c(T - \tau)$ where T is a random variable with a known closed form expectation $\mathbb{E}[T] = \tau$. Trivially, $\mathbb{E}[V^*] = \mathbb{E}[V]$, but for $c = \frac{-\text{Cov}(V,T)}{\text{Var(T)}}$, it can also be shown that:

$$\text{Var}(V^*) = (1 - \rho^2)\text{Var}(V) \quad \text{where } \rho = \text{Corr}(V, T). \tag{12}$$

This can represent a significant reduction in variance compared to a naive estimate. In our case, a good candidate for T is simulating the discounted payoff under a modified version of $\tilde{S}_t$ which is defined by:

$$X_{(i+1)\delta t} = X_{i\delta t} + X_{i\delta t}(r\delta t + \sigma_{avg}\sqrt{\delta t} Z_i), \quad \text{for } i \in \{0, 1, 2, \cdots, N-1\}.$$

Where $X_0 = S_0$ and the $\sigma_{avg} = \sum_{i=0}^{N-1} \sigma(i\delta t)$. X_t is a geometric Brownian motion, thus $\mathbb{E}[T] = \tau$ is given by the Black-Scholes formula for a put option. Our control variate estimate for V is then:

$$\mu_{\mathrm{CV}} = \frac{1}{M} \sum_{j=1}^{M} [e^{-rT}(\tilde{S}_T^{(j)} - K)^- + e^{-rT} c(X_T^{(j)} - K)^- - c\tau] \tag{13}$$

If there is sufficient correlation between the first and second terms of the summand we should see a significantly lower variance in our estimator.

3.4 Monte Carlo Results

Table 1 gives a representative output for each of the algorithms above, including 95% confidence intervals. Each method was performed using N=100 time steps and M=10000 sample paths.

Method	Result	Confidence Interval	Interval Size
Naive Price	42.8778	[42.2565, 43.4990]	1.2425
Antithetic Price	**42.3908**	**[42.3145, 42.4670]**	**0.1525**
Variate Control Price	42.3580	[42.2398, 42.4762]	0.2363
Delta Monte Carlo	-0.7427	[-0.7502, -0.7352]	0.0150

Table 1: Summary of Prices, Confidence Intervals, and Interval Sizes for Monte Carlo Methods

Despite ρ in equation (12) being observed as high as 0.98, a strong negative covariance between thetic and antithetic payoffs led to even faster convergence with antithetic Monte Carlo. This negative covariance is a product of the option being deep "in-the-money". When thetic path drifts upward, the payoff can drop by up to $K - S_0$, while the corresponding antithetic path's payoff increases without limit. Simulations which start out-of-the-money, perform better under a control-variate method, however, for this essay, we submit the antithetic price estimate as the best Monte Carlo technique.

4 Finite Difference Methods

4.1 Implementation of Crank-Nicolson Method

We adapt Armstrong's implementation of the Crank-Nicolson method to approximate a solution to PDE (2) on the rectangle $[0, T] \times [S_{\min}, S_{\max}]$ modeled by a discrete grid:

$$\{(t_i, S_j) : i \in \{0, 1, \cdots, N\} \quad \text{for } j \in \{0, 1, \cdots, M\}\}.$$

Here we choose $S_{min} = 0$ and $S_{max} = K \exp(8\sigma_{max}\sqrt{T})$ as an interval containing a large majority of the sample paths of (S_t). The PDE states $\forall i, j$:

$$V_t(t_i, S_j) + rS_jV_S(t_i, S_j) + \frac{1}{2}S_j^{2\gamma}\sigma(t)^2V_{SS}(t_i, S_j) = rV(t_i, S_j). \tag{14}$$

The boundary conditions imposed by the put option imply that for all i, j:

(a) $V(T, S_j) = (S_j - K)^-$, (b) $V(t_i, S_{\max}) = 0$, (c) $V(t_i, S_{\min}) = e^{-r(T-t_i)}K$

We infer the value of V on the interior by using estimates for V_S and V_{SS} to reduce condition (14) to a system of $(M-1)$ first-order ODEs in t. Let:

$$S_j := S_{\min} + j\delta S, \quad \delta S := \frac{S_{\max} - S_{\min}}{M}, \quad V_t^j := V(t, S_j)$$

Then the *central difference estimates* are given by:

$$\frac{\partial V}{\partial S}(t, S_j) \approx \frac{V_t^{j+1} - V_t^{j-1}}{2\delta S}, \quad \frac{\partial^2 V}{\partial S^2}(t, S_j) \approx \frac{V_t^{j+1} - 2V_t^j + V_t^{j-1}}{(\delta S)^2}.$$

Substitution into (14) and regrouping terms, gives the system of ODEs:

$$\frac{dV_t^j}{dt} \approx a_jV_t^{j-1} + b_jV_t^j + c_jV_t^{j+1}$$

where $a_j := -\frac{\sigma(t)^2S_j^{2\gamma}}{2(\delta S)^2} + \frac{rS_j}{2\delta S}$, $b_j := r + \frac{\sigma(t)^2S_j^{2\gamma}}{(\delta S)^2}$, $c_j := -\frac{\sigma(t)^2S_j^{2\gamma}}{2(\delta S)^2} - \frac{rS_j}{2\delta S}$.

We can express this system in matrix form as:

$$\frac{d\mathbf{V}}{dt} \approx \Lambda\mathbf{V}_t + \mathbf{W}_t$$

where $\mathbf{V}_t$ is a vector with components V_t^j, and Λ and $\mathbf{W}_t$ are defined as:

$$\Lambda := \begin{pmatrix} b_1 & c_1 & 0 & 0 & \dots & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 & 0 \\ 0 & a_3 & b_3 & c_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & b_{M-2} & c_{M-2} \\ 0 & 0 & 0 & 0 & \dots & a_{M-1} & b_{M-1} \end{pmatrix}, \quad \mathbf{W}_t := \begin{pmatrix} a_1V_t^0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ c_{M-1}V_t^M \end{pmatrix}$$

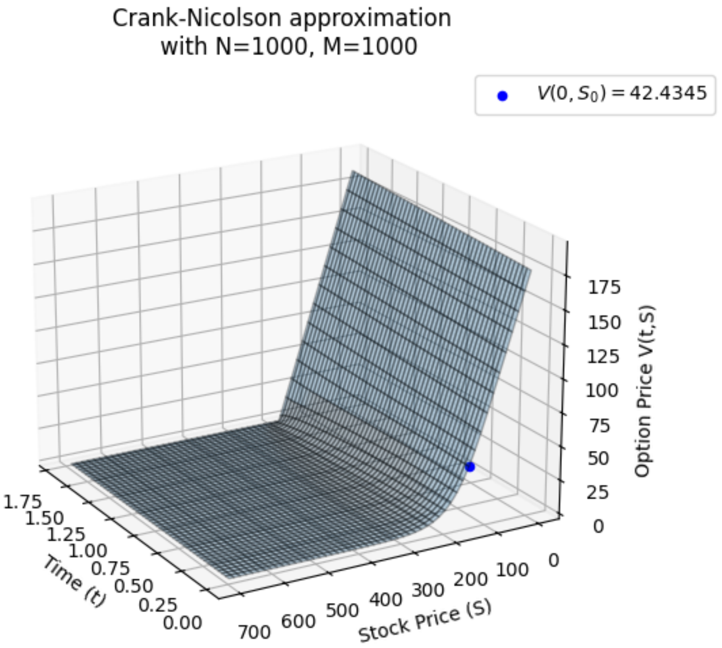

Figure 1: The C-N Approximation for V(t,S) with N=M=1000.The plot omits larger values of S for which V is virtually zero.

The Crank-Nicolson method then makes a central difference approximation:

$$\frac{\mathbf{V}_{t+\delta t}-\mathbf{V}_t}{\delta t} \approx \tfrac{1}{2}\Lambda(\mathbf{V}_{t+\delta t}+\mathbf{V}_t)+\tfrac{1}{2}(\mathbf{W}_{t+\delta t}+\mathbf{W}_t).$$

$\mathbf{V}_t$ can be efficiently solved for given $\mathbf{V}_{t+\delta t}$, using a linear algebra methods. Moving backwards through time from $\mathbf{V}_T$ we can compute V at each point on our grid. To estimate V at an arbitrary point, we linearly interpolate from the values on the surrounding grid points. $\partial_S V(S_0,0)$ is given by a central difference estimate of the form (9), where we interpolate for $V(S_0+h,0)$ and $V(S_0-h,0)$.

4.2 Finite Difference Results

Figure 1 is a plot of the output of the algorithm described above. Given $M=1000$ and $N=1000$, we observed $V(S_0,0)\approx 42.4345$, and $\partial_S V(S_0,0)\approx -0.7498521$. This is broadly in agreement with the Monte Carlo estimates.

5 Comparison of Numerical Methods

Without an analytic solution for V we compare the relative accuracy of the Crank-Nicolson (**CN**) against Monte Carlo (**MC**) by quantifying the order of convergence of their errors. The 95% confidence interval is a useful proxy for the **MC** error. The **CN** error will be quantified using a benchmark computation of $V(S_0, 0)$ on a relatively large grid (50000×50000). We assume that an increasing sequence of lower grid size **CN** computations will converge to the benchmark approximately at the same rate with which errors converge.

The central limit theorem suggests that the **MC** error over M paths will decrease as $\mathcal{O}(M^{-1/2})$, whereas we would expect the overall **CN** error on an $N \times N$ grid to be $\mathcal{O}(N^{-2})$ since estimates of the form (9) have error $\mathcal{O}(h^2)$.

Figure 2 is a log-log plot of the observed errors when running both methods with a wide range of M and N. The gradients of the trend lines match our hypotheses on the order of error closely. These plots do not allow us to directly compare accuracy since on the x-axis M cannot be meaningfully compared with N.

A more informative comparison is given by figure 3 which log-log plots the error of computations using a wide range of paths/grids, against the time taken to run those computations using the same resources. For lower values of M and N, the Monte Carlo estimate was more accurate which is likely due to the Crank-Nicolson method's higher algorithmic complexity (more computation time is required initially to gain a reasonable estimate). As grid size and path numbers increase, we observe that the Crank-Nicolson method becomes substantially more accurate than a Monte Carlo simulation using the same resources. The gradients of the two regression lines express the differing rates of convergence with respect to computation time and clearly show that the finite difference error is converging at a higher order. The computations observed closest to the intersection of the linear regressions are a Monte Carlo simulation with $M = 5700$ and a Crank-Nicolson with $N = 155$, which illustrate the magnitude of M and N at which a Crank-Nicolson method will overtake a Monte Carlo approach.

In conclusion, the data presented suggests that the Crank-Nicolson method converges more quickly to the true price than a comparable Monte Carlo estimate, and so, given adequate resources, its output will almost certainly be more accurate.

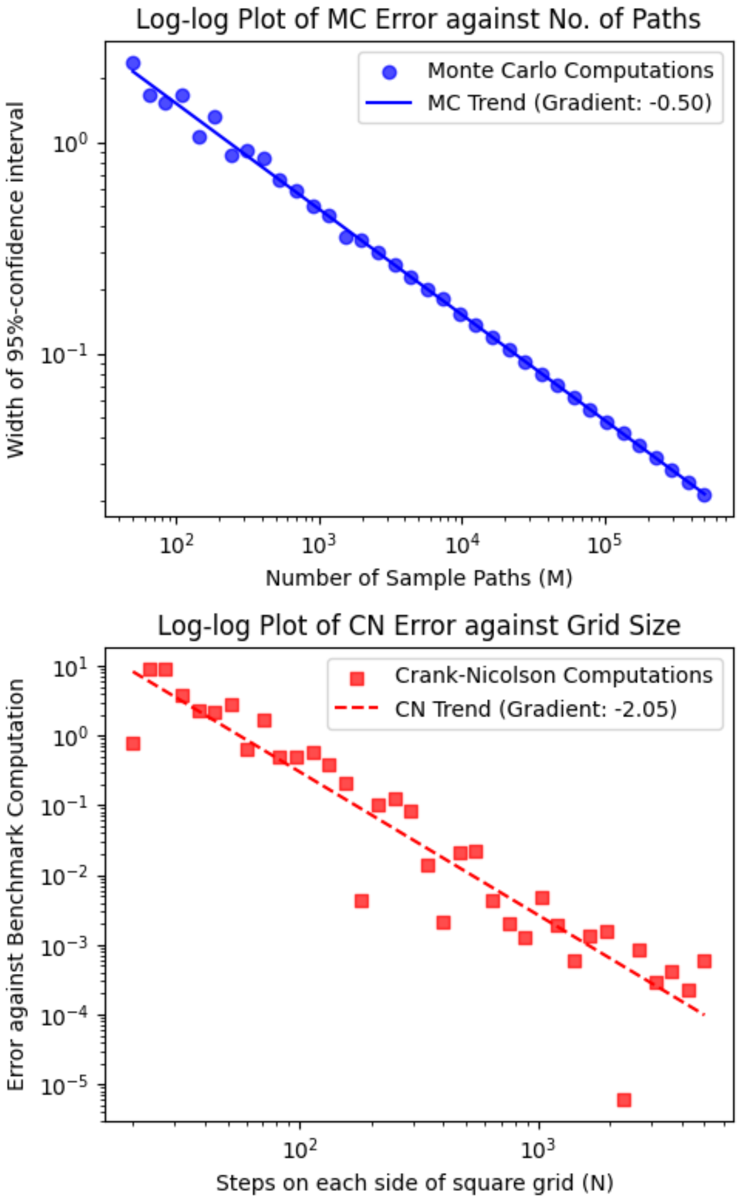

Figure 2: The upper plot has a gradient of -0.49, agreeing with the expected order of convergence $M^{-1/2}$ for the Monte Carlo method. The lower plot has a gradient of -2.05, close to the expected order of N^{-2} for the Crank-Nicolson method. Note the two plots are not directly comparable due to different quantities on the $x-$axis. Best-fit lines were fitted using a standard least squares linear regression.

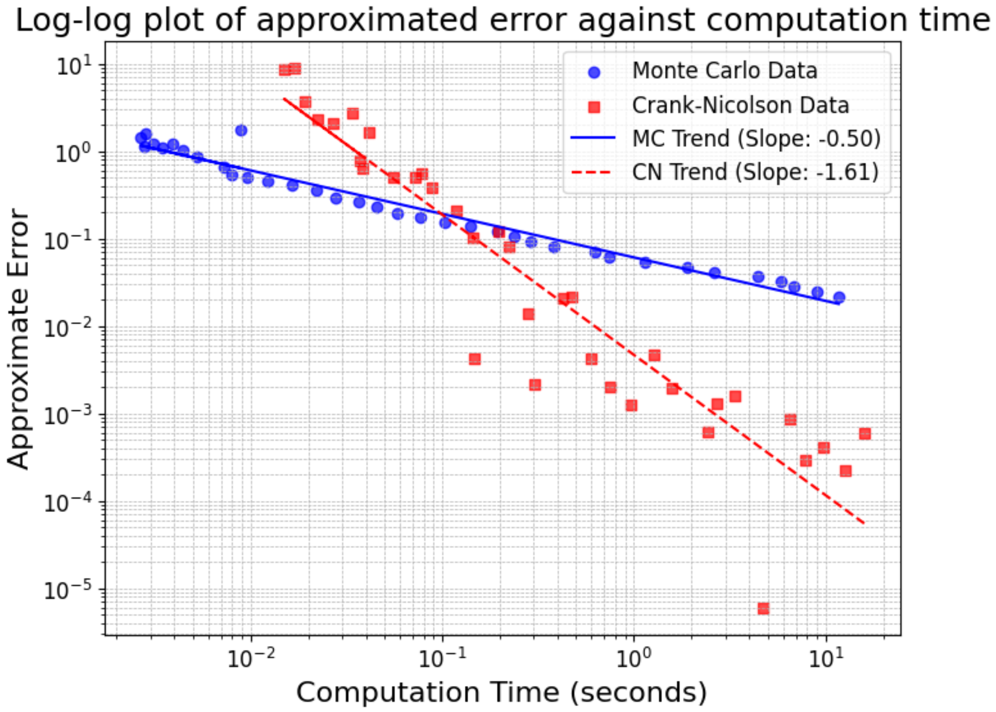

Figure 3: Each point corresponds to a computation using a different choice of paths/grids. The gradients of the trend lines show the relative speed of convergence between the two methods, without the need to directly compare the number of paths with grid sizes.

6 Testing

Multiple tests were made to ensure the accuracy and robustness of all the techniques implemented, most of which can be found at the end of appendix A. The most important indicator of accuracy is that computations from the Crank-Nicolson method agree with each Monte Carlo method, despite these implementations relying on different code and theory. Moreover, since the code was written using very general parameters it is easy to confirm that when S_t is altered to have $\gamma = 1$ and constant σ, all methods agree with an implementation of the analytic Black-Scholes formula for a put option.

References

[Arm24] John Armstrong. Lecture notes for 7CCMFM06, Computational and Numerical Methods in Mathematical Finance. https://keats.kcl.ac.uk/course/view.php?id=119832, 2024.

[BK04] N H Bingham and Rudiger Kiesel. *Risk-neutral valuation*. Springer Finance Textbooks. Springer, 2nd edition, May 2004.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[For24] Matin Forde. Lecture notes for 7CCMFM02, Risk Neutral Valuation. https://keats.kcl.ac.uk/pluginfile.php/11910351/mod_resource/content/9/BlackScholesModel.pdf, 2024.