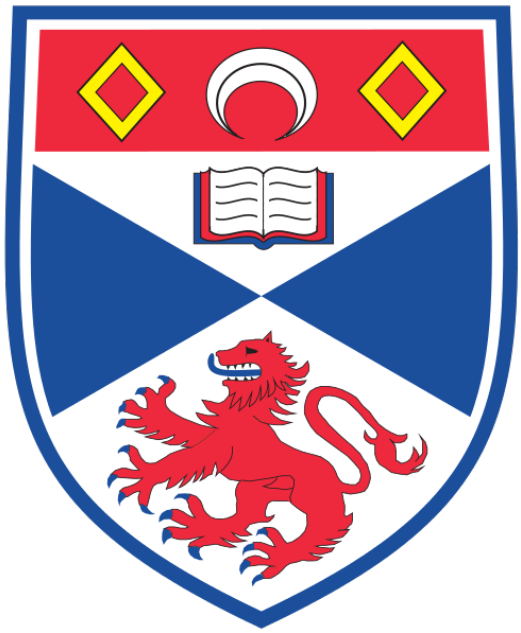
Oscillations Begone: Tackling Weakly Damped Crank-Nicolson Errors in the Valuation of European Call Options

Oliver Löthgren

April 2023



Module code: MT5599

This dissertation is submitted in partial fulfillment for the degree of MPhys (Hons) in Mathematics and Theoretical Physics at the University of St Andrews.

I certify that this project report has been written by me, is a record of work carried out by me, and is essentially different from work undertaken for any other purpose or assessment.

Abstract

Non-smooth initial conditions precipitate spurious oscillations for the Crank-Nicolson method; weakly damped high frequency error terms do not adhere to the scheme's supposed second order convergence properties. Instead, first order convergence is observed. This investigation explores the application of the Crank-Nicolson method in valuing European call options. These are known to have a kink in their initial conditions. Numerous methods attempt to reestablish second order convergence in the Crank-Nicolson scheme. We explore two such modifications, namely Rannacher timestepping and Richardson extrapolation. Rannacher timestepping leverages the superior stability of the backward Euler method; the first R time steps are replaced with aR backward Euler evaluations. Effectively, this serves to prematurely dampen high frequency errors. Two parameterisations, $R = 1$ and $R = 2$, are considered with half step ($a = 2$) partitioning. The Richardson extrapolation, on the other hand, nullifies leading order error terms. It does this via linear combinations of solutions for differing step sizes. As the computational grid is refined for constant $\lambda = \frac{k}{h} = \frac{\Delta t}{\Delta x}$, results confirm the suspected first order convergence of the Crank-Nicolson method. Rannacher timestepping is found to reestablish second order convergence across all λ. Richardson extrapolations are found to improve convergence when dominant leading orders are nullified. This is, however, only the case for small λ when extrapolating Crank-Nicolson solutions. In general, the procedure results in increased accuracy.

Contents

1 Introduction

Appropriate pricing of options is of utmost importance in the financial industry. It quenches arbitrage opportunities and leads to more stable markets for everyone to enjoy. The Black-Scholes equation, or modifications thereof, are often implemented for this purpose [5, 22]. The framework, albeit presumptuous, provides analytical results for valuing European options via the solution of a partial differential equation (PDE). However, a vast range of options do not have such a luxury; American options (and most Exotic options), for example, do not have analytical solutions [3, 12]. As a result, numerical methods must be implemented to approximate the solution of the PDE. Putting the veracity of the Black-Scholes equation aside for a moment, the problem then becomes to judge the accuracy of its numerical solutions. Without analytical solutions, this becomes a challenging task. Results in scenarios where analytical solutions *do* exist must therefore be extrapolated to similar problems where they do not. This, in combination with simulation methods and numerical incorporation of real world qualities, is how numerical schemes are tested [32]. There is therefore value in exploring numerical solutions of European call options, both to validate numerical schemes and to extract implications for other applications.

Finite difference methods provide one particular numerical framework of approximation, and involves discretizing solutions to particular points in share price and time. The implicit Crank-Nicolson method is a particular approach within this larger branch, and is the most accurate of the commonly implemented schemes [22]. It does, however, induce weakly damped high frequency error terms for initial conditions $\notin C^1$. This decreases the rate of convergence as the discretization is refined. Improving the rate of convergence and general accuracy are the aim of this investigation. The Rannacher timestepping procedure, as well as the Richardson extrapolation method will be explored to this end. Considerations of computational is also of particular importance. After all, the faster the option price is determined, the faster it can be sold. Fundamental mathematical concepts are first introduced to provide apt background knowledge for the task at hand. We begin with a synopsis of partial differential equations and the requisite Black-Scholes theory. With an understanding of the underlying problem, finite difference methods are then conceptualised, developing towards an account of the Crank-Nicolson scheme. The theories underpinning our select modification procedures are then detailed, after which numerical results are explored.

1.1 Partial Differential Equations

We begin with an introduction of partial differential equations (PDEs): a PDE is an equation involving an unknown multivariate function and its partial derivatives [1]. The solution of a partial differential equation is the fundamental function which satisfies the PDE and its boundary conditions. A kth order partial differential equation is an equation of the form

$$F(D^k u(\mathbf{x}),\, D^{k-1}\mathbf{u}(\mathbf{x}),\, ...,\, D\mathbf{u}(\mathbf{x}),\, \mathbf{u}(\mathbf{x}),\, \mathbf{x}) = 0,$$

where $F : \mathbb{R}^{d^k} \times \mathbb{R}^{d^{k-1}} \times ... \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{G} \to \mathbb{R}$, with $\mathbf{x} \in G \subset \mathbb{R}^d$ such that $\mathbf{x} = (x_1, ..., x_d) \in G$ are the variables of our function, and $k \in \mathbb{N}$. Here, D^k is a vector ($k = 1$) or array ($k > 1$) containing all combinations of partial derivatives of $\mathbf{u}(\mathbf{x})$ of order k such that $D^k \in \mathbb{R}^{d^k}$. We will be exploring the case of *linear second order partial differential equations*, which can be expressed in the form

$$F(D^2\mathbf{u}, D\mathbf{u}, \mathbf{u}, \mathbf{x}) = -\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\partial_{x_i}\partial_{x_j}\mathbf{u} + \sum_{i=1}^{d} b_i(\mathbf{x})\partial_{x_i}\mathbf{u} + c(\mathbf{x})\mathbf{u} - f(\mathbf{x}) = 0.$$

PDEs are *linear* when there are no products of u or any of its derivatives, i.e. the coefficients of the partial derivatives only involve the independent variables $\mathbf{x}$. Due to the linearity property of partial derivatives, any linear combination of solutions to the PDE will remain a solution. Clearly, this is violated in the *non-linear* case. PDEs can also be *quasi-linear*, where the PDE is linear with respect to the highest order partial derivative. We will further simplify our analysis to the case of $d = 2$, where u is a function of two variables $(x_1, x_2) = (x, y)$, with a PDE of the form

$$\alpha u_{xx} + \beta u_{yx} + \zeta u_{yy} + \kappa u_x + \nu u_y + \gamma u + G = 0, \tag{1}$$

where the coefficients are all functions of (x, y).

PDEs are further distinguished by classification; there are parabolic, elliptic, and hyperbolic PDEs. The coefficients of the PDE determine its class. Second order partial differential equations are classified as follows:

1. Parabolic: $\beta^2 - 4\alpha\zeta = 0$
2. Elliptic: $\beta^2 - 4\alpha\zeta => 0$
3. Hyperbolic: $\beta^2 - 4\alpha\zeta < 0$.

Classification for general second order linear PDEs is detailed in Appendix 1. Nevertheless, the two variable case will suffice for the purposes of this paper.

Parabolic second order PDEs occur most frequently in quantitative finance, particularly in option pricing [1]. The renowned Black-Scholes equation is an example of such a PDE. Its discovery revolutionised the financial industry, and it, or generally modifications thereof, are used in practice to this day [2]. The equation provides an unequivocal formula for asset pricing, proving vastly superior to the heuristic financial analyses conducted prior. Elliptic PDEs may appear in infinite horizon problems, with hyperbolic PDEs arising for pure jump processes with dominating drift in Lévy models. See [1] for further details on this. Nevertheless, we now narrow our focus to that of the parabolic Black-Scholes equation.

1.2 Black-Scholes Equation

Options are contracts giving the holder the right to buy or sell an underlying financial asset, at a certain price, before or within an expiry time T. *European* options, which will be the focus of this investigation, can only be exercised at the expiry time, whereas *American* options may be exercised at any time prior. *Exotic* options follow other criteria, and occasionally require certain conditions to be met in order for them to be exercised. Due to the complicated nature of exotic options, many cannot be valued by use of the Black-Scholes equation [3]. In the case of American options, the moving exercise boundary also makes the problem analytically intractable. Nevertheless, numerical methods and extensions of the Black-Scholes Equation (Section 1.2.1) can be implemented to provide solutions in these cases.

Prior to the 1970s, no rigorous framework was in place for the pricing of options [4]. Instead, valuation was conducted based on analyses of markets, employing economic, financial, and risk management theory in an attempt to predict a reasonable price. Given the vast scale and interpretability of these concepts, outcomes could vary drastically; pricing was a skill, and it required an astounding understanding of market dynamics. However, with markets constantly fluctuating and previous methods becoming obsolete, finding a robust method proved to be a challenge. That is, until the Black-Scholes equation entered the scene. In 1973, Fischer Black and Myron Scholes proved and published what is now known as the Black-Scholes equation [5], a parabolic second order linear PDE governing the price of a financial derivative. Its discovery triggered a paradigm shift in the world of finance; the seemingly impossible problem had now been boiled down to the solution of a PDE. Despite the model being contingent on a large set of restrictive assumptions, it was found to consistently outperform previous methods [6]. We will now delve into these assumptions, after which the Black-Scholes equation will be derived. A discussion of the model's veracity and recent extensions will also be presented.

1.2.1 Model Assumptions and Extensions

Black and Scholes stated the following assumptions for their proposed model [5]:

1. The yearly risk free interest rate, r, is known and taken to be constant.
2. The underlying stock price, S, follows geometric brownian motion in continuous time. This leads to a lognormal distribution for the stock price.

3. The volatility of the stock, σ, defined to be the annualised standard deviation of log returns, is constant.
4. The stock has no dividends or other forms of monetary distribution.
5. There are no transaction costs and the market is frictionless.
6. It is possible to buy any fraction of the underlying stock or to hold it at the short-term risk free interest rate.

Thus, in the Black-Scholes model, the value of the option depends entirely on the current price of the stock and the time until the option's expiry. Though treatment of r and σ as constants is clearly unrepresentative of reality, it drastically simplifies the problem. Hence, an accurate choice of r and σ are vital for the model's performance. The risk-free interest rate is often taken to be the discount rate on US Treasury Bills with 30 day maturities. Generally, these are quite stable, but in times of market turmoil these are subject to fluctuation. Volatility estimation, on the other hand, is a research topic in its own right. The two most common estimates are historical or implied volatility: Historical volatility is estimated by analysis of the stock's historical time series, whereas implied volatility is obtained by reverse engineering the market price of an underlying stock's options to retrieve the corresponding volatility value. See [7, 8] for more on this.

One downfall of the Black-Scholes model is that the behaviour of implied volatilities are not in agreement with what is predicted in theory. Under the model's assumptions, volatility is taken to be constant. Therefore, implied volatilities are expected to be the same for options (on the same underlying stock) of different strike prices. However, in reality, we observe a *volatility smile*, where for significantly larger and smaller strike prices the implied volatilities are greater than expected [9]. Modelling the volatility smile remains an active area of research. *Stochastic volatility* models [10], such as the Heston model [11], are extensions of the Black-Scholes that implements a more realistic treatment of volatility. The exclusion of transaction costs is another inherent flaw of the model, as these are virtually unavoidable in practice. Furthermore, stocks with dividends are completely excluded from treatment, which elucidates the model's incompleteness. Other flaws include underestimation of the tails of the log returns distribution, making *black swan events* theoretically implausible despite their relatively frequent occurences. These are events characterised by large and unpredictable shifts in the market. However, there are adaptations of the Black-Scholes model capable of dealing with these scenarios [1, 12, 13]. Overall, the Black-Scholes model is incredibly robust: Variables r and σ can be adjusted to take into account market behaviour, or additional terms can be implemented for better performance [6]. The assumptions of frictionless trade and fractional pruchases are reasonable for a commonly traded stocks.

Although the assumptions are restrictive, in many occasions they are sensible. It is, however, important to verify the veracity of the model given the state of the market. For example, in times of turmoil, the model may not be as accurate due to varying volatilities and interest rates. Extensions should be implemented to capture any expected market effects. A good understanding of markets, for proper implementation, is therefore still a necessity. Nevertheless, the assumption of stocks evolving according to geometric brownian motion is perhaps the most significant. To elucidate why this is an apt assumption, we now explore the underlying theory along with the derivation Black-Scholes PDE.

1.2.2 Derivation

There are several derivations of the Black-Scholes PDE. The original derivation by Black and Scholes involves constructing a *delta hedged* portfolio with the option and the underlying stock,

$$\Pi_t = V_t - \Delta_t S_t,$$

where Π_t is the value of the portfolio, V_t the value of the option, Δ_t is the hedging ratio, and S_t is the share price, all of which are evaluated at time t. *Itô's lemma* (see Appendix A.1) is then used to derive the stochastic differential equation (SDE) governing the evolution of the portfolio's value. Hedging the portfolio removes the stochastic effects in the model, making it inherently risk-less. As a result, it can be assumed to grow at the risk-free interest rate, else there would be an arbitrage opportunity. The remaining differential equation can then be solved, yielding the Black-Scholes equation.

The assumption of *no arbitrage* is key in financial mathematics, and states, simply, that there should be no riskless profit. The *Fundamental Theorem of Asset Pricing* (FTAP) then extends this notion to the equivalence of a market compatible 'absence of arbitrage' property with the existence of a risk-neutral measure. Strictly speaking, the law of

no arbitrage only holds in discrete markets. For more general markets, the notion is extended to that of *no free lunch with vanishing risk*. See (Teichmann, 2020)[14] for a more mathematical treatment and proofs of the FTAP.

Other derivations of the Black-Scholes equation exist. In fact, Black and Scholes include another derivation via the Capital Asset Pricing Model in their paper. Other examples include derivations by portfolio replication, which generalises the original approach, and by the limiting case of the continuous time Cox, Ross, Rubinstein binomial model[15]. We will be exploring a variant of the original derivation.

Geometric Brownian Motion The driving assumption of the Black-Scholes model is that the underlying stock evolves according to a geometric brownian motion (GBM) process. That is, the evolution of the stock price is modelled by the SDE,

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{2}$$

where μ is the natural growth rate of the stock and W_t is a Wiener process (also called Brownian motion). The t subscript denotes the time at which the S evolution is being considered. In the Ito framework, this has an associated analytical t-continuous solution,

$$S_t = S_0 \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\right],$$

which can be shown by application of Ito's formula to calculate $d(\ln S) = \frac{dS_t}{S_t} - \frac{1}{2}\frac{1}{S_t^2}(dS_t)^2 = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t$, retrieving the above result upon integration. In the calculation, dS_t is known from equation (2), and the quadratic variation $(dS_t)^2 = \sigma^2 S_t^2 dt$ because upon expansion, all higher order terms in dt can be neglected. After all, $dt^{m>1} \to 0$ faster than dt as $dt \to 0$. A more formal proof of this result considers the quadratic variation of S_t under a GBM process [16],

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t \\ \Rightarrow S_t &= S_0 + \int_0^t \mu S_u du + \int_0^t \sigma S_u dW_u. \end{aligned}$$

The quadratic variation of S_t is then given by

$$\begin{aligned} [S]_t &= \int_0^t (\mu S_u)^2 d[u] + \int_0^t (\sigma S_u)^2 d[W]_u \\ &= \int_0^t \sigma^2 S_u^2 du \end{aligned}$$

because the quadratic variation of any smooth process is zero, and so $d[S]_t = \sigma^2 S_t^2 dt = (dS_t)^2$ in differential form, as required.

It was previously mentioned that an assumption of the Black-Scholes model is that the share price is lognormally distributed, indeed this is the case for the share price under GBM. The *Fokker-Plank equation* [17],

$$\frac{\partial}{\partial t} f(S,t) + \frac{\partial}{\partial S}\left[a(S,t) f(S,t)\right] = \frac{1}{2}\frac{\partial^2}{\partial S^2}\left[b^2(S,t) f(S,t)\right], \quad f(S,0) = \delta(S),$$

governs the time evolution of the probability density of S. Note that time subscripts have been omitted. Here, $f(S,t)$ is the probability density, at time t, of stochastic variable S, where $\delta(S)$ is the Dirac delta function. The given condition

implies that the initial value of S at time $t = 0$ is known. This statement holds for all Ito processes, i.e. equations of the form

$$dS = a(S,t)dt + b(S,t)dW.$$

In our case, $a(S,t) = \mu S$ and $b(S,t) = \sigma S$, and so the solution of the PDE [17] is the log normal distribution,

$$f(S_t,t) = \frac{1}{S_t}\frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\left(-\frac{[\log\left(\frac{S_t}{S_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)t]^2}{2\sigma^2 t}\right), \tag{3}$$

where S_0 is the known starting price.

Intuitively, GBM seems to be a good model for share price evolution. It incorporates relative changes in magnitude of share price (by scaling terms with the current share price), the notion of inherent randomness via the Wiener process, and the idea of a stock experiencing an adjustable natural rate of growth or decline. A sample of four different GBM paths, starting from an initial share price value of \$15 is shown in Figure 1.

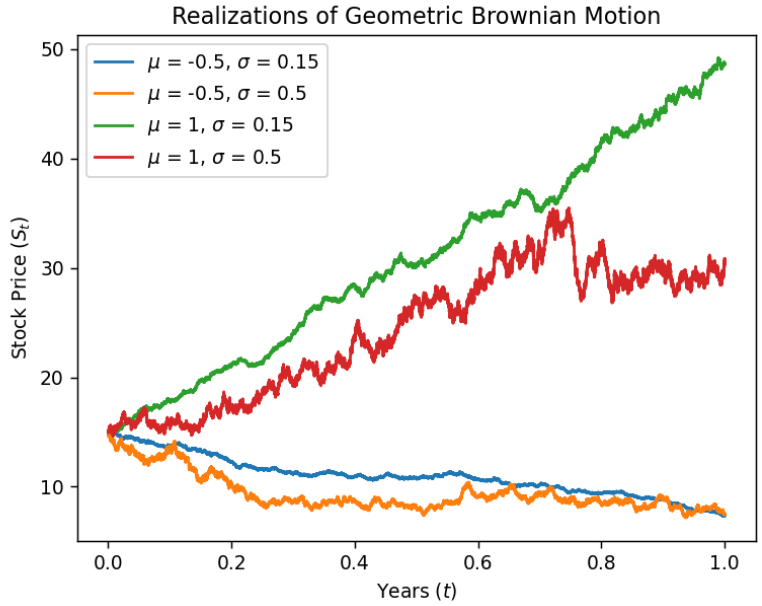


Figure 1: GBM simulations for different drift rates and volatilities, starting from an initial share price of \$15. These were simulated in python using the analytical solution of equation (2).

The inherent properties and behaviour of the GBM process has the potential to closely replicate the behaviour of stock price evolution. However, as previously mentioned, the choice of parameters is paramount. Now that the model's assumptions have been investigated we will delve into the derivation of the Black-Scholes equation. The notions discussed and derived above will come of use.

Obtaining the PDE. As previously stated, we create a delta-hedged portfolio, at time t, of the option and the underlying stock,

$$\Pi(S_t,t) = V(S_t,t) - \Delta(S_t,t)S_t,$$

where Δ is the hedging ratio. We then find that

$$d\Pi = dV - \Delta dS.$$

The hedge is self-financing, hence its differential does not appear. Applying Ito's lemma to calculate dV yields the following expression when collecting terms of dt and dW,

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \frac{\partial V}{\partial S} dS - \Delta dS,$$

where dt is deterministic, and the dS term contains stochastic effects precipitated by dW. The previously derived result for the quadratic variation of S was also used. Thus, the choice of $\Delta = \frac{\partial V}{\partial S}$ removes all stochastic terms, creating a riskless portfolio. As a result, the portfolio will increase in value at the risk-free rate, such that

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt \\ &= \Pi r dt = (V - \frac{\partial V}{\partial S} S) r dt, \end{aligned}$$

which is easily solved and yields a particular convection-diffusion equation,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{4}$$

more commonly known as the Black-Scholes equation. For now, we omit the discussion of initial and boundary conditions. These are explored in Section 1.3.1.

1.2.3 Analytic Solution for a European Call Option

First, we consider a change of measure from original measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$ via *Girsanov's theorem* [18]. Our GBM process for share price is then given by

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t^{\mathbb{P}} \\ dS_t &= r S_t dt + \sigma S_t dW_t^{\mathbb{Q}}, \end{aligned}$$

via $dW_t^{\mathbb{P}} = dW_t^{\mathbb{Q}} - \frac{\mu - r}{\sigma} dt$. We now denote S_t as the known share price at the time we wish to value the option. Applying the *Feynman-Kac Formula* to the Black-Scholes PDE, we find that the solution $V(S_t, t)$ of equation (4), with final condition $V(S_T, T) = h(S_T)$, has the stochastic-integral form

$$V(S_T, T) = V(S_t, t) + \int_t^T rV(S_u, u)du + \int_t^T \sigma S_u dW_u.$$

Taking the expectation at time t for a stock with current price S_t and rearranging for the current value $V(S, t)$ yields

$$V(S_t, t) = E^{\mathbb{Q}}_{t,S_t}[h(S_T)] - E^{\mathbb{Q}}_{t,S_t}\left[\int_t^T rV(S_u, u)du\right],$$

where the SDE of V is given by

$$dV = rV du + \sigma \frac{\partial V}{\partial S} dW_u^{\mathbb{Q}}, \quad V(S_T, T) = h(S_T)$$

for $t \leq u \leq T$. Discounting V at a rate r, such that $G(S_u, u) = e^{-ru}V(S_u, u)$, eliminates the deterministic term from the SDE. Thus,

$$dG = -re^{-ru}Vdu + e^{-ru}dV = e^{-ru}\sigma\frac{\partial V}{\partial S}dW_u^{\mathbb{Q}},$$

and so the discounted price process G is a martingale under the measure $\mathbb{Q}$. As a result, the stochastic representation for G yields

$$\begin{aligned} E_{t,S_t}^{\mathbb{Q}}\left[G(S_t,t)\right] = e^{-rt}V(S_t,t) &= E_{T,S_t}^{\mathbb{Q}}\left[G(S_T,T)\right] \\ &= E_{T,S_t}^{\mathbb{Q}}\left[e^{-rT}V(S_T,T)\right] \\ &= e^{-rT}h(S_T) \end{aligned}$$

Thus, with risk-neutral valuation under the measure $\mathbb{Q}$, we obtain the *risk-neutral valuation formula*

$$V(S,t) = e^{-r(T-t)}E_{t,S}^{\mathbb{Q}}\left[h(S_T)\right]. \tag{5}$$

Given that the model assumptions are satisfied, the PDE can be solved both analytically (via (5)) and numerically for European options. See [3] and [12] for treatments of Exotic and American options. Now, we further narrow our focus to that of the European Call option, which gives the holder the right to purchase an underlying stock for strike price K at time of expiry T. The value of the European call option, at expiry, is therefore

$$V(S_T,T) = h(S_T) = \max(S_T - K, 0) = (S_T - K)^+,$$

and so the analytical value of a European Call option is given by

$$\begin{aligned} V(S,t) &= e^{-r(T-t)}E_{t,S}^{\mathbb{Q}}\left[(S_T-K)^+\right] \\ &= e^{-r(T-t)}\int_0^\infty (S_T-K)^+ f^{\mathbb{Q}}(S_T,T)dS_T \\ &= S\phi(d_1) - Ke^{-r(T-t)}\phi(d_2), \end{aligned} \tag{6}$$

where $f^{\mathbb{Q}}$ simply replaces occurrences of μ with the risk-free rate r in equation (3), ϕ is the cumulative standard normal distribution, and the d_i are given by

$$\begin{aligned} d_1 &= \frac{\ln\left(\frac{S}{K}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \\ d_2 &= \frac{\ln\left(\frac{S}{K}\right) + (r - \frac{1}{2}\sigma^2)(T-t}{\sigma\sqrt{(T-t)}}. \end{aligned}$$

The final result is obtained easily by transformations of variables to express equation (6) in terms of cumulative standard normal distributions.

The effect of the risk-neutral valuation is not to assume that the natural growth rate of the stock, μ, is equal to the risk free interest rate, r. Rather, the fundamental insight from Black and Scholes is that the natural growth rate of a stock becomes irrelevant to its pricing: It is made redundant due to hedging. However, the veracity of the Fundamental Theorem of Asset Pricing is questionable. Arbitrage opportunities will ultimately still appear in imperfect markets, meaning that a risk neutral measure does not exist. Nevertheless, the argument is that as soon as these arbitrage opportunities occur they will be depleted by arbitrageurs, thereby return the market to a state absent of arbitrage.

1.3 Finite Difference Methods

Although an analytical solution exists in the case of European Call options, we aim to explore the effectiveness of numerical methods in estimating solutions to the Black-Scholes equation. The most commonly applied FDMs for this purpose are the *explicit*, *implicit*, and *Crank-Nicolson* schemes [22]. Attention is focused on the Crank-Nicolson method, as it is both unconditionally stable in the L^2 norm [23] and the most accurate of the three methods. A discussion of the *Von Neumann Stability Analysis* for this case can be seen in [22, 24]. The method will be derived in Section 2.3.1 below. Improvements of the Crank-Nicolson scheme, namely Rannacher timestepping and Richardson extrapolation, will also be explored. First, defining notions of the FDM theory will be presented.

1.3.1 Finite Difference Approximations

Forward, backward, and central difference approximations. Consider a real function $u(x)$ that is both continuous and differentiable. Then its derivative, at point x_0, is defined by the difference in $u(x)$ after taking a step from $x = x_0$ to $x = x_0 + h$ divided by the step size h, as h approaches zero. Thus, the *first derivative* of function $u(x)$ is given by

$$u'(x) = \lim_{h \to 0} \frac{u(x+h) - u(x)}{h}.$$

It is natural, then, to consider approximating the value of the derivative at point x by instead taking a small, non-vanishing $h > 0$, such that

$$u'(x) \approx \frac{u(x+h) - u(x)}{h}.$$

This is the *forward difference approximation* of the first derivative. Expanding and rearranging the Taylor series of $u(x+h)$ about x shows that the method has a truncation error of $O(h)$. That is,

$$u(x+h) = u(x) + hu'(x) + \frac{1}{2}h^2u''(x) + \frac{1}{3!}h^3u'''(c) \tag{7}$$

with $c \in [x,\, x+h]$, and so

$$u'(x) = \frac{u(x+h) - u(x)}{h} + O(h).$$

The approximation of the second derivative is derived by taking the forward difference approximation of this result, namely

$$u''(x) = \frac{u(x+2h) - 2u(x+h) + u(x)}{h^2} + O(h).$$

Equivalently, the derivative can be approximated by instead taking a backward step from x to $x - h$, again with non-vanishing step size h. This is the *backward difference approximation*. It too is a first order approximation, where similarly,

$$u(x-h) = u(x) - hu'(x) + \frac{1}{2}h^2u''(x) - \frac{1}{3!}h^3u'''(c) \tag{8}$$

now with $c \in [x-h,\, x]$, and so

$$u'(x) = \frac{u(x) - u(x-h)}{h} + O(h).$$

The backward difference approximation of the second derivative follows. Namely,

$$u''(x) = \frac{u(x) - 2u(x-h) + u(x-2h)}{h^2} + O(h).$$

Taking the average of the forward and backward difference approximation yields the *central difference approximation*,

$$\begin{aligned} u'(x) &= \frac{1}{2}\left(\frac{u(x+h) - u(x)}{h} + O_f(h)\right) + \frac{1}{2}\left(\frac{u(x) - u(x+h)}{h} + O_b(h)\right) \\ \Rightarrow &= \frac{u(x+h) - u(x-h)}{2h} + O(h^2). \end{aligned} \tag{9}$$

Alternatively, the central difference approximation can be thought of as the finite difference approximation obtained by taking steps both forwards and backwards, dividing through by the net step size $2h$. Note that the truncation errors of the forward and backward difference schemes cancel, and so the central difference scheme is a second order method. This can be proven by subtracting equation (7) by equation (8) and rearranging for $u'(x)$.

Finite difference methods are constructed by selecting combinations of the forward, backward, and central difference approximations for derivatives appearing in the PDE of interest. Notable qualities of these approximations include the greater accuracy of the central difference method, as well as the greater stability provided by implementing backward difference approximations in time. However, backwards approximations in time result in implicit equations, requiring the solution of a linear system of equation. This is more computationally demanding. The forward difference approximation, albeit less stable, can be calculated explicitly at a faster rate. As suggested by the above properties, the explicit scheme is constructed by forward difference in time and central difference in space approximations, whereas the implicit scheme utilizes a backward difference in time. The Crank-Nicolson scheme is also an implicit scheme whose derivation will be given in Section 2.4. The substitution of finite difference approximations into a PDE then results in a finite difference equation (FDE). Given a set of initial and boundary conditions, this FDE can be used to approximate PDE solutions at discrete points in the solution space. In our case, we are approximating the value of an option, $V(S,t)$, and so the space under consideration is (S,t).

Discretization of Solution Space. For the Black-Scholes equation, FDMs aim to approximate the value of the option $V(S,t)$. Thus, the solution space under consideration is (S,t), which is discretized to create a computational grid of points for which values of V will be estimated. An even grid spacing will be used for the investigation, where points are separated by step sizes Δt and ΔS in time and share price respectively. A diagram of the computational grid is shown in Figure 2.

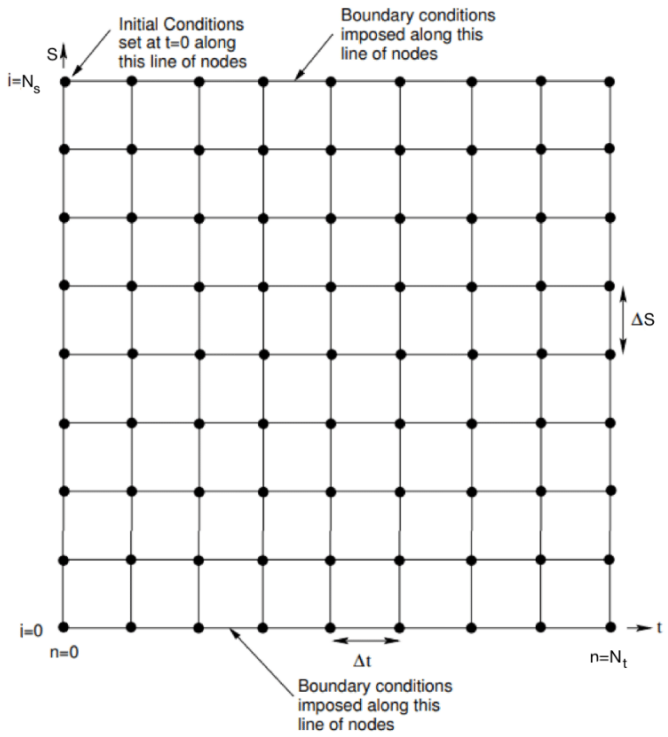


Figure 2: Evenly spaced computational grid for finite difference methods. [33]

Points in share price are indexed by $i = 0, 1, ..., N_S$, whereas points in time will be indexed by $n = 0, 1, ..., N_t$. The current time is taken to be $t = 0$, with the time until option expiry given by $t = T$. Thus, the step size $\Delta t = \frac{T}{N_t}$. For share price, the minimum and maximum bounds will be taken to be $S = 0$ and $S = S_{max}$ respectively, such that $\Delta S = \frac{S_{max}}{N_S}$. The choice of S_{max} will be elaborated on in the discussion below.

Initial and Boundary Conditions. As for the PDE, a unique solution of the FDE can only be obtained given a set of initial and boundary conditions. Both shall be presented here. For a European call option with strike price K, the final time payoff is known to be $V(S_T, T) = (S_T - K)^+$, whereas the options value at the current time is unknown and must be solved for. The Black-Scholes equation must therefore be solved backwards in time, utilizing this payoff function for the initial conditions. The boundary conditions of the Black-Scholes PDE are given by $V(S = 0, t) = 0$ and $V(S, t) = S - Ke^{-r(T-t)}$ as $S \to \infty$. These can be derived from the cumulative standard normal form of the analytical solution (6).

Numerically, considerations of infinite values is intractable, and so $S \to \infty$ is replaced with $S = S_{max}$, an arbitrary number large enough to accurately reflect the nature of the boundary. The choice of S_{max} will therefore be dependent on the price range of the stock under consideration and its strike price K. For instance, the boundary for a stock consistently valued in a range $\$5 - 10$ with a strike price $K = 6$ may be accurately represented by $S_{max} = 500$, whereas a stock within the range $\$50 - 100$ may require a greater $S_{max} = 5000$. The numerical boundary is therefore given by $V(S = S_{max}, t) = S_{max} - Ke^{-r(T-t)}$. As $Ke^{-r(T-t)} << S_{max}$, this will be further simplified to $V(S = S_{max}, t) \approx S_{max}$. To summarise, the initial and boundary conditions of our FDE will be

$$\begin{aligned} &\bullet\ V(S, T) = (S - K)^+ \\ &\bullet\ V(0, t) = 0 \\ &\bullet\ V(S_{max}, t) = S_{max}. \end{aligned} \tag{10}$$

Necessary Properties of an FDE. Consistency, stability, and convergence are essential properties of an FDE, else it will fail to provide accurate solutions to the PDE in question. These are defined as follows:

1. An FDE is said to be *consistent* [19] if, as the grid size tends to zero, the FDE tends to the PDE. Consider a PDE defined by $Fu = f$ in differential operator form for a function $u(x_1, x_2)$. Then, an FDE $F_{\Delta x_1, \Delta x_2} u = f$ of order (r, s), where $F_{\Delta x_1, \Delta x_2}$ is the operator defining the finite difference scheme with step sizes $(\Delta x_1, \Delta x_2)$, is consistent with the PDE if for any smooth function ϕ,

$$F\phi - F_{\Delta x_1, \Delta x_2}\phi = O((\Delta x_1)^r, (\Delta x_2)^s),$$

   which approaches zero as the grid size is decreased. The smoothness of ϕ is required so that derivatives can be computed in the PDE.

2. A finite difference scheme is *stable* [20] if errors introduced in the solution of the FDE are not amplified with the solution of the FDE. That is, the solution of the FDE does not, as a consequence of the method's specification, introduce growing computational errors. Mathematically, a finite difference scheme is stable with respect to the r-norm defined by

$$\|V^{(k)}\|_r = (\sum_{i=1}^{N_S} \Delta S |V_i^k|^r)^{1/r}, \quad 1 \leq r < \infty$$

   if there exists a positive constant C independent of $\Delta S, \Delta t$, such that

$$\|V^{(k)}\|_r \leq C\|V^0\|_r, \quad \forall\, k \leq 0.$$

   Here, $V(k) = (V_0^k, V_1^k, ..., V_{iS}^k)^T$. Stability is commonly classified via the L^2 and L^∞ norms, with $\|V^{(k)}\|_\infty = \sup_{0 \leq i \leq iS} |V_i^k|$.

3. An FDE is *convergent* [19] if, as the grid size approaches zero, the solution of the FDE tends to the real solution of the PDE. That is, if $u(S, t)$ is the real solution of a PDE and $v_i^n = v(i\Delta S, n\Delta t)$ is the solution of a corresponding FDE, then $v_i^n \to u(S, t)$ as $\Delta S, \Delta t \to 0$ with $(i\Delta S, n\Delta t) \to (S, t)$.

Note that the Black-Scholes equation (4) is linear. As a result, we can invoke the *Lax-equivalence theorem*, which states that if an FDE is consistent and stable it will also be convergent in L^2 norm. See [21] for a proof of this result. We now focus our attention on the Crank-Nicolson method.

1.4 Implicit Crank-Nicolson Method

The Crank-Nicolson method is a combination of the explicit and implicit methods that offers second order convergence in both space and time with a truncation error of $O((\Delta t)^2, (\Delta S)^2)$. This outperforms the explicit and implicit schemes in terms of accuracy, both of which have truncation errors of $O(\Delta t, (\Delta S)^2)$. For the Black-Scholes equation, the scheme is consistent [25] and, as previously mentioned, unconditionally stable in L^2 norm. By the Lax-equivalence theorem, it is therefore also L^2-convergent. This is sufficient for our purposes, as the initial condition for the value of a European call option lies in L^2. Higher order methods do exist, and can be derived by iteratively substituting lower order finite difference approximations into those of higher order. However, these methods are both harder to implement and computationally intensive. Ease of implementation and computational efficiency is important in option pricing: The first party to price an option well will also be the first to sell it.

FDE in (S, t) Space. The method is derived by first considering the half-step central difference approximation for the time derivative, evaluated at the intermediary point $(i, n+\frac{1}{2})$. This yields a truncation that is second order error in time, given by

$$\frac{\partial V}{\partial t}|_{i,n+\frac{1}{2}} = \frac{V_i^{n+1} - V_i^n}{\Delta t} + O((\Delta t)^2).$$

This can be derived by subtracting the Taylor expansion of V_i^{n+1} with that of V_i^n about the intermediary point. The standard central difference approximation (9) is implemented for derivatives with respect to the share price, also evaluated at the intermediary point. To retrieve the full scheme, the terms evaluated at the intermediary time step, $V_i^{n+1/2}$, are replaced with $(V_i^{n+1} - V_i^n)/2$. It can be shown that $V_i^{n+1/2} = (V_i^{n+1} - V_i^n)/2 + O((\Delta t)^2)$ by combining the Taylor expansions of V_i^{n+1} and V_i^n about the point $(i, n+\frac{1}{2})$. The stencil for the Crank-Nicolson method is shown below in Figure 3.

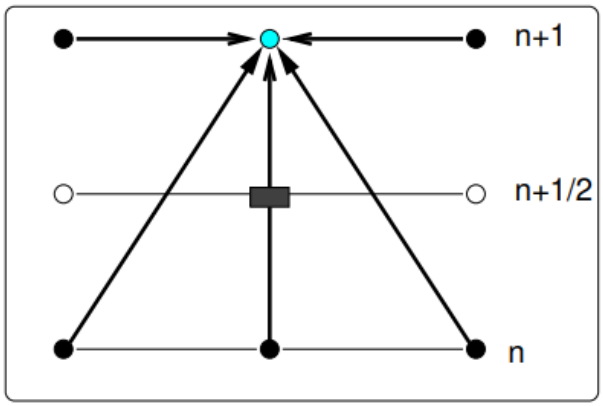


Figure 3: Crank-Nicolson Stencil [33]

Though no information is directly carried from $V_i^{n+\frac{1}{2}}$, the black rectangle demonstrates that the method is derived from this point. Information is carried from both the nth and $(n+1)$th states of the system, i.e. the scheme is implicit.

The final result is obtained by collecting terms in time, yielding the tridiagonal system of equations,

$$-a_i V_{i-1}^n + (1+b)V_i^n - cV_{i+1}^n = a_i V_{i-1}^{n+1} + (1-b)V_i^{n+1} + cV_{i+1}^{n+1} = D_i^{n+1}, \tag{11}$$

for $i = 1, 2, ..., N_S - 1$. It is important to note that solutions are calculated backwards in time, meaning that terms in time $(n+1)$ are known, whereas terms in time (n) are unknown. The coefficients of (11) are defined by

$$\begin{aligned} a_i &= \frac{\Delta t}{4}[\frac{\sigma^2 S_i^2}{(\Delta S)^2} - \frac{rS_i}{\Delta S}] \\ b_i &= \frac{\Delta t}{2}[\frac{\sigma^2 S_i^2}{(\Delta S)^2} + r] \\ c_i &= \frac{\Delta t}{4}[\frac{\sigma^2 S_i^2}{(\Delta S)^2} + \frac{rS_i}{\Delta S}]. \end{aligned}$$

Note that the coefficients depend on the share price at the grid point where the option value is being calculated. The tridiagonal system can be represented in matrix form, $T\mathbf{V}^n = \mathbf{d}^{n+1}$, defined by

$$T = \begin{pmatrix} (1+b_1) & -c_1 & & & \\ -a_2 & (1+b_2) & -c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -a_{N_{S-2}} & (1+b_{N_{S-2}}) & -c_{N_{S-2}} \\ & & & -a_{N_{S-1}} & (1+b_{N_{S-1}}) \end{pmatrix}_{(N_{S-1})\times(N_{S-1})}$$

with

$$\mathbf{V}^n = \begin{pmatrix} V_1^n \\ V_2^n \\ \vdots \\ V_{N_{S-2}}^n \\ V_{N_{S-1}}^n \end{pmatrix}_{(N_{S-1})\times 1} \quad \text{and} \quad \mathbf{d}^{n+1} = \begin{pmatrix} D_1^{n+1} \\ D_2^{n+1} \\ \vdots \\ D_{N_{S-2}}^{n+1} \\ D_{N_{S-1}}^{n+1} + c_{N_{S-1}} S_{max} \end{pmatrix}_{(N_{S-1})\times 1},$$

where boundary conditions must be considered in the $i = 1$ and $i = N_{S-1}$ case. As these values are known, they are incorporated in the constant vector $\mathbf{d}^{n+1}$. The $i = 0$ boundary corresponds to a share price of $S = 0$ with $V(0,t) = 0$, meaning that the term does not explicitly appear. The matrix system can then be solved via the *Tridiagonal Matrix Algorithm* (or *Thomas Algorithm*) which has $O(N)$ complexity [34]. It is also possible to consider bijective transformations of the independent variables of the Black-Scholes equation. These define a new PDE with an alternative solution space. We will now look at the system in $(\ln S, t)$ and (z, τ) space, where we define $z = \ln S - (r - \frac{1}{2}\sigma^2)t$ and $\tau = T - t$. Note that solutions at grid points in these spaces can be inversely transformed for their corresponding (S, t) mapping. Working with these alternative spaces can provide better computational efficiency if the Crank-Nicolson scheme remains stable.

FDE in $(\ln S, t)$ Space. The transformation of variables $x = \ln S$ results in the new Black-Scholes PDE,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + (r - \frac{1}{2}\sigma^2)\frac{\partial V}{\partial x} - rV = 0, \tag{12}$$

as $\frac{\partial}{\partial S} = \frac{1}{S}\frac{\partial}{\partial x}$ and $\frac{\partial^2}{\partial S^2} = \frac{\partial}{\partial S}(\frac{\partial}{\partial S}) = \frac{1}{S^2}\frac{\partial^2}{\partial x^2} - \frac{1}{S^2}\frac{\partial}{\partial x}$. Upon re-deriving the Crank-Nicolson scheme for equation (12), we obtain the new tridiagonal system

$$-aV_{i-1}^n + (1+b)V_i^n - cV_{i+1}^n = aV_{i-1}^{n+1} + (1-b)V_i^{n+1} + cV_{i+1}^{n+1} = D_i^{n+1} \tag{13}$$

for $i = 1, ..., N_x - 1$, where the coefficients are now constants, defined by

$$\begin{aligned} a &= \frac{\Delta t}{4}[\frac{\sigma^2}{(\Delta x)^2} - \frac{r}{\Delta x}] \\ b &= \frac{\Delta t}{2}[\frac{\sigma^2}{(\Delta x)^2} + r] \\ c &= \frac{\Delta t}{4}[\frac{\sigma^2}{(\Delta x)^2} + \frac{r}{\Delta x}]. \end{aligned} \tag{14}$$

The matrix form of the system follows naturally.

In this space, the Crank-Nicolson scheme remains stable and offers greater computational efficiency [23]. Implementation of the Crank-Nicolson method is also simplified, as variable coefficients are no longer involved.

FDE in (z, τ) Space The final space under consideration are the transformed variables that reformulate the Black-Scholes equation as the diffusion equation,

$$\frac{\partial \nu}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 \nu}{\partial z^2},$$

via the transformations $z = \ln S - (r - \frac{1}{2}\sigma^2)t$, $\tau = T - t$, and $\nu(S,t) = e^{-rt}V(S,t)$, where

$$\begin{aligned}
\frac{\partial v}{\partial t} &= -\frac{\partial v}{\partial \tau} - (r - \frac{1}{2}\sigma^2)\frac{\partial v}{\partial z} \\
\frac{\partial v}{\partial S} &= \frac{1}{S}\frac{\partial v}{\partial z} \\
\frac{\partial^2 v}{\partial S^2} &= \frac{\partial}{\partial S}(\frac{\partial v}{\partial S}) = -\frac{1}{S^2}\frac{\partial v}{\partial z} + \frac{1}{S^2}\frac{\partial^2 v}{\partial z^2}.
\end{aligned}$$

Note that the PDE is now solving for $\nu(z, \tau)$ rather than V. Note that with respect to $\tau = T - t$ the FDE is now solved forwards in time. The resulting tridiagonal system is that of the familiar diffusion equation, namely

$$-\alpha V_{i-1}^{n+1} + (1 + 2\alpha)V_i^{n+1} - RV_{i+1}^{n+1} = \alpha V_{i-1}^n + (1 - 2\alpha)V_i^n + RV_{i+1}^n = D_i^n,$$

for $i = 1, ..., N_z - 1$, with $\alpha = \frac{\Delta\tau\sigma^2}{4(\Delta z)^2}$. As grid discretization is carried out in the particular solution space governing the PDE, the initial and boundary conditions (10) must also be transformed to the appropriate variables.

Implementation in the (z, τ) space is more complicated; transformations in share price, $S \leftrightarrow z$, and option value, $V \leftrightarrow \nu$, are now also time dependent. The scheme does, however, remain unconditionally stable (see Appendix A.2 demonstrating this via Von-Neumann stability analysis). Comparing calculation times in solving the tridiagonal system demonstrates that the (z, τ) transformation is more efficient than the (S, t) formulation. However, computation remains the fastest in $(\ln S, t)$ space.

Choice of Solution Space. All spaces considered are unconditionally stable, yet they differ in the complexity of implementation and computational efficiency. Moving forward, we consider the alternative $(x = \ln S, t)$ space, which is both the easiest scheme to implement and the most computationally efficient. Initial and boundary conditions in this new space will therefore be transformed to,

$$\begin{aligned}
&\bullet\, V(x, T) = (e^x - K)^+ \\
&\bullet\, V(x_{min}, t) = 0 \\
&\bullet\, V(x_{max}, t) = S_{max},
\end{aligned} \tag{15}$$

where $S = e^x$, and $x_{max} = \ln(S_{max})$. Note that the treatment of x_{min} is important here, as the logarithm of zero is undefined. Thus, we take an arbitrarily small number $S_{min} \neq 0$, e.g. 1e-10, such that $x_{min} = \ln S_{min}$. Note that this is of course not a completely accurate representation of the boundary, but errors resulting from this will be negligible. Furthermore, the transformation to $x = \ln S$ allows for a larger S_{max} to be considered, thereby providing a better representation of the upper boundary. If S_{max} is increased, however, more grid points must be implemented to sustain accuracy. Nevertheless, the choice of S_{min} and S_{max} has a negligible effect on results as long as they are reasonable for the problem. It is also worth noting that an even grid-spacing in x will result in an un-even grid spacing in S; as S increases so does the spacing in the grid. This must be taken into account for the choices of x_{max} and Δx.

1.4.1 Method Properties

As previously mentioned, the scheme offers second order convergence in space and time, as well as unconditional stability. However, for non-smooth initial conditions, the Crank-Nicolson method has weakly damped high frequency error terms [26]. This is because discrete Fourier decompositions of such initial conditions require the inclusion of

high frequency terms. Unfortunately, errors associated to these frequencies persist over time. Giles and Carter (2006) find that this leads to different rates of convergence for low and high frequency regions. Namely, low frequency oscillations behave as expected, with second order convergence, whereas high frequencies experience a reduced first order convergence. As the initial conditions for the European call option are non-smooth, this presents a problem.

In fact, central difference schemes are not appropriate for solving convection-diffusion equations [27]. To avoid these spurious oscillations, the step size ΔS, must be below a certain critical value [26]. Rannacher time-stepping [28] is a startup procedure that decreases high frequency errors to reestablish second order convergence. Richardson extrapolation [29] can also be implemented to increase method accuracy. These are now discussed as modifications and extensions to the standalone Crank-Nicolson method. Note that alternative modifications may also be implemented for this purpose. See [30] for a review of such techniques. Variable grid spacing appears to be another avenue of interest, however, Giles and Carter (2006) find that this does not resolve the issue.

1.5 Rannacher time-stepping

The Crank-Nicolson modification suggested by Rannacher (1984) decreases high frequency error terms by replacing the first R steps of the Crank-Nicolson scheme with $2R$ half time steps of the implicit method, also referred to as the backward Euler method. Via energy methods, Rannacher proved that a mere replacement of the first Crank-Nicolson step, i.e. $R = 1$, recovers second order convergence. However, Giles and Carter (2006) find that although the choice of $R = 1$ gives second order convergence for the solution of V, the choice of $R = 2$ results in second order convergence for all of V, $\Delta = \frac{\partial V}{\partial S}$, and $\Gamma = \frac{\partial^2 V}{\partial S^2}$. The $R = 1$ parameterisation instead fails to converge for Γ [23]. Note that Δ and Γ, as partial derivatives of V, also satisfy the Black-Scholes equation. These values have many uses in quantitative finance and are referred to as *greeks* [31]. For instance, Δ is the hedging ratio implemented in our derivation of the Black-Scholes equation. Γ therefore measures the rate of change of the hedging ratio with changes in share price. Note that both V and Δ are in L^2, whereas Γ resides in L^∞. As the method is only L^2-convergent, this gives reason the Crank-Nicolson methods failure to converge for Γ.

As the backward Euler method exhibits second order local convergence [31], replacing only the first few steps will theoretically maintain second order convergence. Increasing R, however, will result in diminishing accuracy. It is also possible to consider different implementations of Rannacher time-stepping by specifying the partitioning of each Crank-Nicolson step. For instance, quarter time steps may be used instead as in [23]. In this investigation, we shall explore the parameterisations $R = 1$ and $R = 2$, with half step divisions, with focus on the modification's effect on both accuracy and convergence.

1.5.1 The Backward Euler Method

FDE in (x, τ) Space. For our implementation of the backward Euler method, we make an additional transformation in time from $t \to \tau = T - t$. As such, the problem is solved forwards in time with respect to τ. Note that this transformation has no effect on the Crank-Nicolson scheme; the tridiagonal system remains the same. The transformation from t to τ implies that $\frac{\partial}{\partial t} = -\frac{\partial}{\partial \tau}$. That is, it changes the sign of the time derivative in the Black-Scholes equation (12). In turn, the sign of all coefficients a, b, and c in the tridiagonal system will be reversed. At first glance, this appears to change the tridiagonal system, however, the problem is now solved forwards in time. Terms in time (n) are now known, whereas terms in time $(n + 1)$ are unknowns. As a result, the tridiagonal system remains unchanged. Note that implementing the implicit method for a problem solved backwards in time is equivalent to applying the explicit method to a forward time problem. This only offers conditional stability [33]. The transformation is therefore necessary in order for the implicit scheme to be unconditionally stable. In fact, the backward Euler method offers better stability than the Crank-Nicolson method, as the scheme properly dampens high-frequency error terms [30].

The backward Euler method, as previously stated, utilizes a backward approximation in time and central approximation in share price. Implementing the Backward Euler method for the Black-Scholes equation with independent variables (x, τ) results in the tridiagonal system,

$$-aV_{i-1}^{n+1} + (1 + b)V_i^{n+1} - cV_{i+1}^{n+1} = V_i^n, \tag{16}$$

when collecting terms in time. Recall that the truncation error for this scheme is $O(\Delta t, (\Delta x)^2)$, which can be shown by substituting the appropriate finite difference approximations into the PDE. Note also that the coefficients are the same as in the Crank-Nicolson case (14). Initial and boundary conditions follow naturally from (15), with the initial condition now defined at $\tau = 0$ rather than $t = T$, and boundary conditions defined for arbitrary time $0 < \tau < T$.

1.6 Richardson Extrapolation

Next we consider extending both our Crank-Nicolson and Rannacher time-stepping schemes with the Richardson Extrapolation [30]. In essence, the modification combines solutions of two step sizes (differing by a scalar multiple) with appropriate multiplicative factors in order to cancel the leading error term. This accelerates the convergence of the scheme, as the combined result now has a higher order leading error term. Consider a second order approximation for V with a step size h, such that

$$V = V_h + O(h^2) + O(h^3) + O(h^4).$$

Now, consider the same approximation, but with a stepsize $2h$. Then,

$$V = V_{2h} + O((2h)^2) + O((2h)^3)) = V_{2h} + 4O(h^2) + 8O(h^3) + 16O(h^4),$$

where the $O(h^2)$, $O(h^3)$, and $O(h^4)$ terms are equal. Multiplying the first equation by four and subtracting the second yields

$$3V = 4V_h - V_{2h} - 4O(h^3) - 12O(h^4).$$

Thus, an estimate of V can be obtained by dividing this equation by three. Collecting the coefficient into the $O(h^3)$ term then implies that

$$\begin{aligned} V &= \frac{4}{3}V_h - \frac{1}{3}V_{2h} + O(h^3) \\ &= V^{[1]} + O(h^3), \end{aligned} \tag{17}$$

and so we have obtained a third order approximation. Note that this methodology can be iterated, obtaining subsequent $V^{[n]}$ approximations that remove n leading order error terms. For instance, combining results for $V^{[1]}$ with step sizes h and $2h$ implies that the third order leading error of equation (17) can be removed via the combination

$$V = \frac{8}{7}V_h^{[1]} - \frac{1}{8}V_2 h^{[1]} + O(h^4).$$

Thus, a fourth order approximation is obtained. Note, however, that Richardson extrapolation can only be implemented when the leading order error is known. Additionally, the extrapolations are only carried out at grid points defined by step size $2h$; the h step implementation includes additional points in between those specified by step size $2h$. These cannot be compared. Hence, in evaluating solutions at points separated by step size $2h$, the extrapolation makes explicit use of two valuations at the points of interest (from the $2h$ and h specification). Information from intermediary points defined by step size h are also used.

The procedure is simple and computationally efficient. For $V^{[1]}$, the calculation of two step sizes h and $2h$ achieves an accuracy $O(h^3)$ at points defined by stepsize $2h$. To achieve this equivalent order of accuracy, the original approximation with $O(h^2)$ convergence requires corresponding step sizes of $(2h)^{3/2}$. Going back to our original problem, let $2h$ denote the share price step size $2\Delta x$. The Thomas algorithm has $O(N)$ complexity and $N = (x_{max} - x_{min})/h - 1$ for the tridiagonal system (13). Hence, the Richardson extrapolated solution will have $N_h + N_{2h} = 3(x_{max} - x_{min})/2h - 2$ points to evaluate, albeit in two separate tridiagonal systems. The original single step scheme, however, will require $N_{(2h)^{3/2}} = (x_{max} - x_{min})/(2h)^{\frac{3}{2}} - 1$ point evaluations in a single tridiagonal system. As $h \to 0$, it is clear that the Richardson extrapolation is more computationally efficient. Note that due to the coefficients of the leading order error terms, the single step approach may be more efficient down to a finite h.

After this point, however, the extrapolation will surpass the original scheme. This analysis is relevant for the low frequency region, where both the Crank-Nicolson and Rannacher time-stepping methods are observed to have second order convergence. Analysis of the high frequency region shows the same improvement in efficiency, and can be similarly shown.

Note that, due to the contrasting convergences, Richardson extrapolation should be implemented differently to affect the low and high frequency regions. We will be exploring both cases, with procedures detailed below. Richardson extrapolation offers greater accuracy, with the high frequency implementation supposedly diminishing the effect of spurious oscillations.

1.6.1 Crank-Nicolson Implementation

Low-frequency error terms in the Crank-Nicolson scheme are properly damped, and follow the theoretical truncation error, $O((\Delta t)^2, (\Delta x)^2)$. This can be shown by following the derivation of the scheme outlined in Section 1.4. On the other hand, Rannacher (1984) proved that high frequency error terms experience first order convergence, with corresponding truncation error $O(\Delta t, \Delta x)$. We will be applying Richardson extrapolations in two cases, nullifying second and first order leading error terms respectively. The first case will serve to improve the accuracy of low-frequency terms, leaving the high frequency terms unaltered. The second will reestablish second order convergence for high frequencies, whilst second order convergence for low frequencies remains in tact.

As the order of error in both time and share price are the same, we will be calculating errors for step sizes of constant $\lambda = \frac{\Delta t}{\Delta x} = \frac{k}{h}$, where $k = \Delta t$ and $h = \Delta x$. Thus, a value of h will directly imply a value of k and vice versa. Considering only one step size for a given lambda is therefore sufficient, and it keeps both errors in time and space of the same order [23] as the grid is refined for $k, h \to 0$.

In cancelling out second order errors, we implement the Richardson extrapolation,

$$V^{[1]} = \frac{4}{3}V_h - \frac{1}{3}V_{2h}, \tag{18}$$

as derived above. Given that a third order error term follows the leading term in the low frequency case, we expect to now observe third order convergence. To nullify first order errors, the extrapolation becomes,

$$U^{[1]} = 2V_h - V_{2h}, \tag{19}$$

which can be shown similarly to the $V^{[1]}$ case. The resulting order of convergence in the high frequency regime will be investigated numerically.

1.6.2 Rannacher Implementation

For the Rannacher time-stepping modifications of parameterisations $R = 1$ and $R = 2$, the process follows similarly. Giles and Carter (1984) find that the leading error term for the high frequency region, under Rannacher timestepping, now obtains a convergence of $O(h^{2R+1})$. The low-frequency error term sustains a second order convergence. Hence, the low frequency implementation remains the same, with a Richardson extrapolation given by (18), with next order error term Rk^3 [23]. For termination of the leading high-frequency error,

$$W^{[1]} = \frac{2^{2R+1}V_h - V_{2h}}{2^{2R+1} - 1} \tag{20}$$

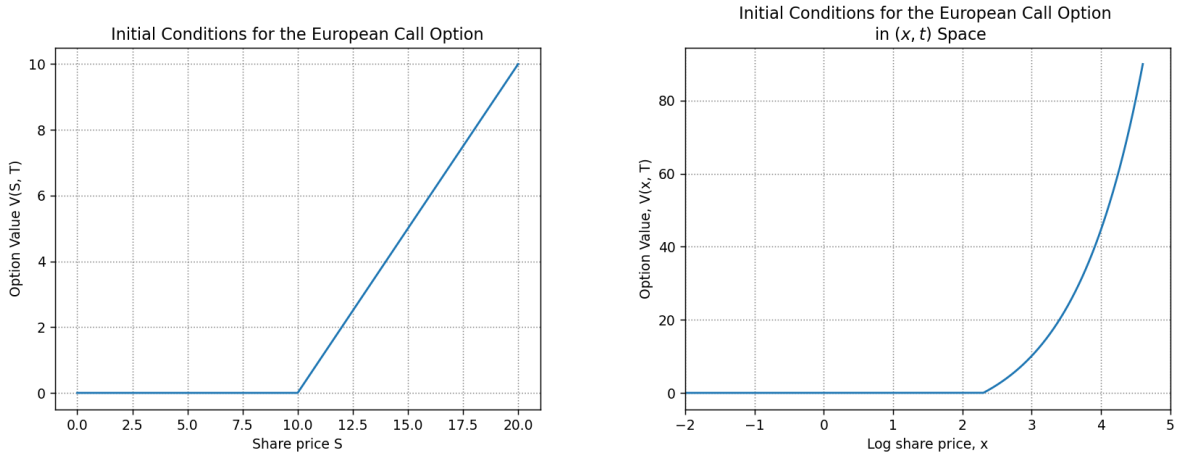
can be shown to be the appropriate extrapolation. Note that the $R = 0$ case retrieves $U^{[1]}$, as expected. As for the Crank-Nicolson implementation, the next order high frequency error term will be determined numerically.

2 Exploration

In our investigation we compare the valuation results obtained for the standard Crank-Nicolson method with both Rannacher timestepping and Richardson extrapolation modifications. The parameterisations $R = 1$ and $R = 2$ of Rannacher time-stepping will first be compared to the Crank-Nicolson scheme. The better of the two parameterisations will then be further modified via Richardson extrapolation. Original Crank-Nicolson results, as well as those obtained via its Richardson extrapolations, will also be compared. Reducing high frequency error terms and increasing the rate of convergence takes precedence in our analysis. We begin with a specification of parameters for the Black-Scholes equation and the European call option that is to be valued.

2.1 Problem Statement

The Black-Scholes equation requires the specification of the risk free interest rate, r, and the volatility of the underlying stock, σ. We will assume a risk-free interest rate of 2%, i.e. $r = 0.02$, and a volatility of $\sigma = 0.15$. The European call option parameters under consideration are a strike price of $K = 10$ and time to expiry of two years, i.e. $T = 2$. Note that all parameters are expressed on a yearly basis, but any other time frame may be used. The initial condition, $V(S, T) = (S - K)^+$, of the specified European call is plotted with respect to both S and x in Figure 4.



(a) (b)

Figure 4: Payoff at expiry for a European Call option with parameters $r = 0.02, \sigma = 0.15, K = 10$, and $T = 2$, with respect to (a) S and (b) $x = \ln S$.

Note the kink in the initial condition at $S = 10$ corresponding to the strike price. This is what provokes spurious high frequency oscillations. Numerical results will be compared to the analytical solution (6), from which error terms $\epsilon = N - A$ can be determined via. Here, N are numerical and A analytical solutions. These will be evaluated at each grid point. Our analysis will focus on the maximal error across the grid, i.e. in L^∞ norm, as this is the most prevalent and widely applicable norm in quantitative finance [23]. In the specification of boundary conditions, we will take $S_{min} =$ 1e-10 and $S_{max} = 500$. The low boundary is close enough to a share price of zero and the upper boundary is sufficient as it represents a practically unattainable value for a stock with strike price $K = 10$. As strike prices are set above the current price of a share, to reach a value of \$500 the share would need to increase in value by more than 5000% in a two year time span. We further restrict our domain in the analysis of maximal error to a more realistic share prices $S < 100$. An increase in value of 1000% in a two year time span is similarly implausible, and so errors after this share value are not crucial to analyse.

2.2 Crank-Nicolson Approach

For the standalone Crank-Nicolson scheme we explore the numerical valuations in relation to the analytical solution, in addition to the convergence of the scheme as $k, h \to 0$. Values of $\lambda = \frac{k}{h} = 3, 4, 5, 6, 10$ will be investigated. As we will see, convergence properties vary based on this ratio. λ is kept constant to ensure that the impact of error terms in time and space is kept constant as the grid is refined. We observe second order convergence for low λ and large step sizes. As suggested by theory, first order convergence begins to be observed with decreasing step size [30]. Despite suffering from weakly damped high frequency errors, the scheme provides an apt approximation of the worth of our European call option. After presenting numerical results, an analysis of convergence is conducted for decreasing step size.

2.2.1 Valuation Results

The Crank-Nicolson scheme, implemented with step sizes $\Delta x = 0.001$ and $\Delta x = 0.1$ with $\lambda = 3$, yields solutions shown in Figure 5. Note that not all numerical grid solutions are shown. Continuous analytical solutions and initial conditions are included for comparison.

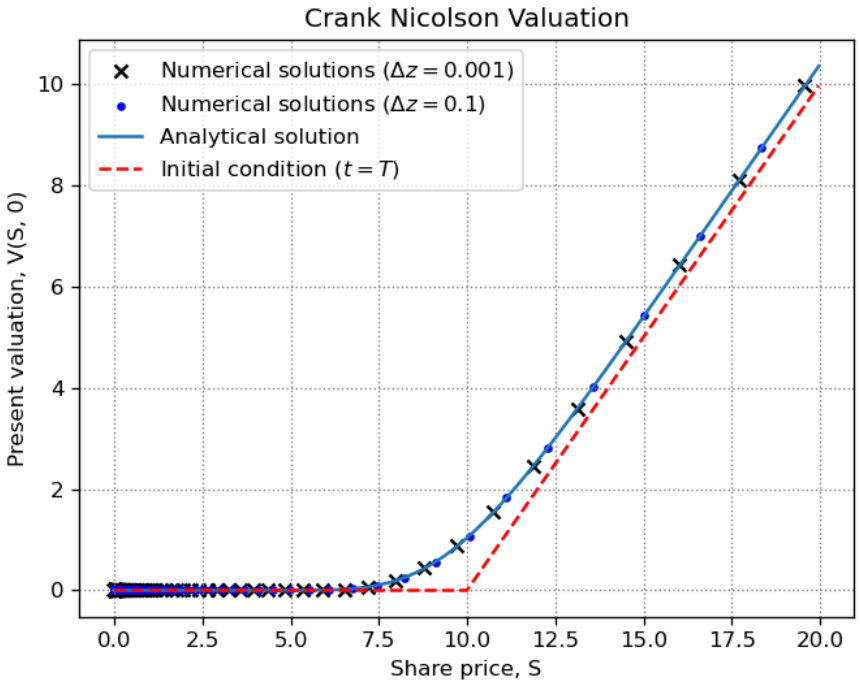


Figure 5: Comparison of numerical and analytical results using the Crank-Nicolson method.

Accuracy of the Crank-Nicolson method appears to be satisfactory, with points only slightly deviating from analytical solutions. The max error in the $\Delta x = 0.1$ case is 0.0166, whereas the largest error for $\Delta x = 0.001$ is only 1.77e-6. A factor of 10^2 decrease in step-size results in a 10^4 factor decrease in error, which appears to follow second order convergence. Errors are almost indiscernible for both step-sizes. Even with $\lambda = 10$ the maximal error is only 0.022, however the lowest observed error is comparatively larger, with a magnitude of 0.00017. Hence, first order convergence is implicated.

Note that valuations, moving backwards in time, slowly evolve from the kinked initial conditions into the smooth analytical solution. Only the final time payoff is non-smooth, any time prior will be differentiable at all points. In

implementing the Black-Scholes equation, it is only the present valuation that is considered accurate; at later times, model parameters may need to be altered.

2.2.2 Convergence Results

Theory suggests that low frequency error terms will experience second order convergence, whereas high frequency terms only converge in first order. Indeed, results show that this appears to be the case. Maximal errors, for varying λ, are plotted in Figure 6 as the computational grid is refined.

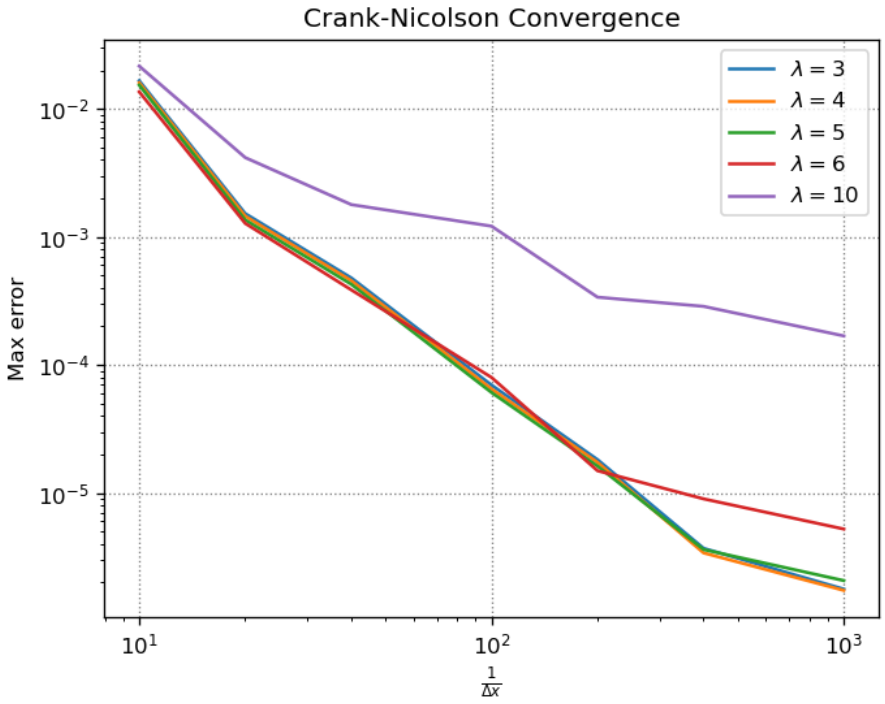


Figure 6: Convergence of Crank-Nicolson scheme for decreasing step size at constant λ.

Values of $\lambda = 3, 4$, and 5 appear to behave similarly, converging in second order for large step sizes, until a step size of $\Delta x = 0.0025$ where first order convergence begins to be observed. For the slightly larger $\lambda = 6$, second order convergence is also observed at greater Δx, however, first order convergence initiates earlier at a step size of $\Delta x = 0.005$. In the $\lambda = 10$ case, there is initially second order convergence down to a step size of $\Delta x = 0.05$, after which first order convergence is observed.

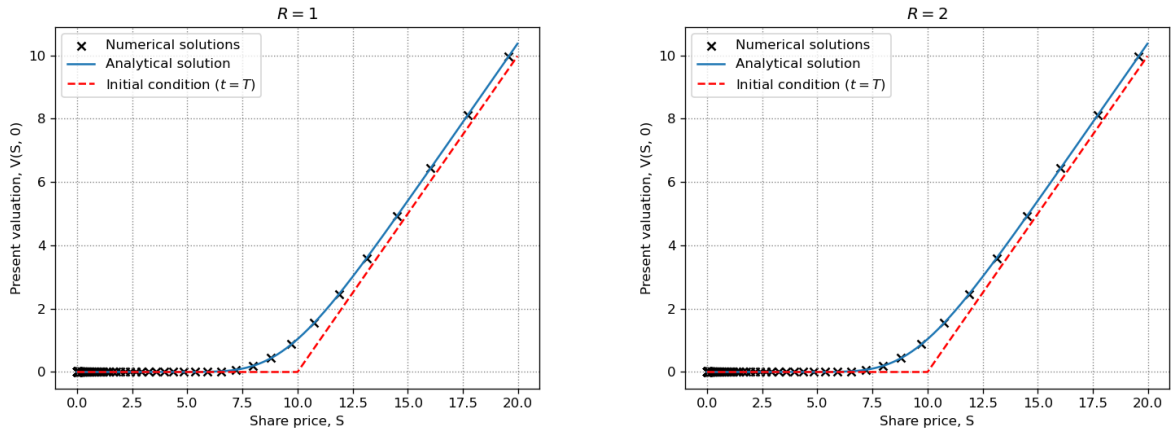
2.3 Crank-Nicolson with Rannacher Timestepping

As discussed, non-smooth initial conditions provoke spurious oscillations in the Crank-Nicolson scheme, leading to first order convergence for high frequency error terms [30]. Numerical results obtained in exploration of the scheme confirm this (see Figure 6). To reestablish the theoretical second order convergence of the Crank-Nicolson method, Rannacher (1984) suggests replacing the R first time steps with solutions from the implicit scheme. Numerical results prove this to provide good approximations of the analytical solutions under both parameterisations. The corresponding present day valuations will first be explored, after which a convergence analysis is conducted for the modified schemes.

2.3.1 Valuation Results

The magnitude of errors, even for the seemingly large step size $\Delta x = 0.1$, remain low. For $R = 1$, we observe maximum errors of 0.0178 and 1.9e-6 in the $\Delta x = 0.1$ and $\Delta x = 0.001$ cases respectively. Errors in the $R = 2$ case are slightly larger. With $\Delta x = 0.1$, the maximal error is 0.019, whilst $\Delta x = 0.001$ has a maximal error of 2.02e-6. As for Crank-Nicolson, this suggests second order convergence.

Figure 7 shows numerical and analytical solutions obtained with Rannacher's modification, considering the case of $\lambda = 3$ with $\Delta x = 0.001$. The $\lambda = 3$ case is shown due to its superior accuracy, but even the $\lambda = 10$ case performs only slightly worse as the grid is refined. Here, the maximal errors are 0.024 and 0.047 for a step size $\Delta x = 0.1$ with the $R = 1$ and $R = 2$ parameterisations, respectively. The lowest observed error for the $\lambda = 10$ case, with a step size of $\Delta x = 0.001$, are also 2.57e-6 and 3.97e-6. See Section 2.3.2 for more on this.

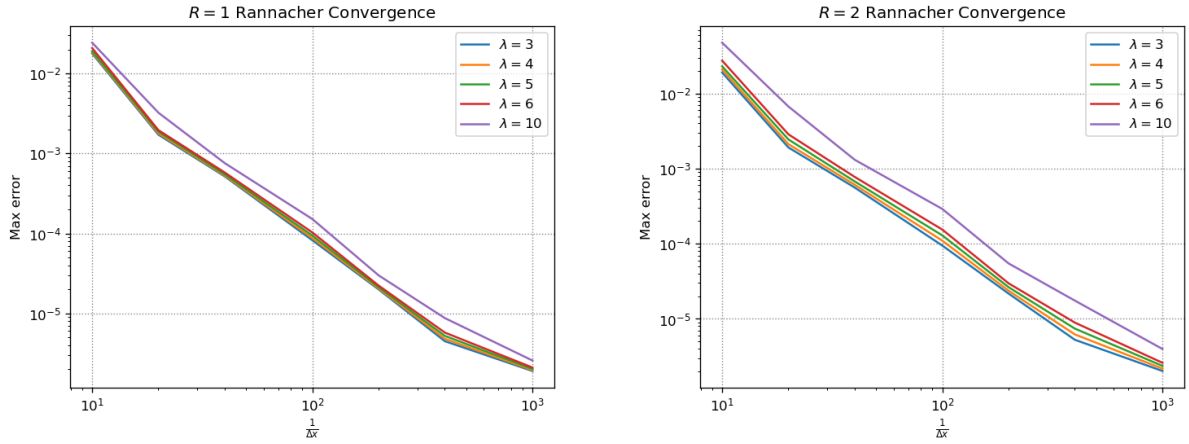


(a) (b)

Figure 7: Comparison of numerical and analytical results using Rannacher timestepping with $R = 1$ and $R = 2$. Stepsizes are $\Delta x = 0.001$ and $\Delta t = 0.003$.

2.3.2 Convergence Results

As expected, Rannacher timestepping proves to be efficient in sufficiently dampening high frequency error terms. A plot of max error with decreasing step size for both parameterisations is shown in Figure 8.



(a) (b)

Figure 8: Convergence of Rannacher time stepping schemes (with parameters (a) $R = 1$ and (b) $R = 2$) for decreasing step size.

Second order convergence is reestablished for all λ. Most notable is the $\lambda = 10$ case, where improvement is substantial; the previously dominating first order error is completely dissipated, with second order convergence being observed in its place. The impact of the Rannacher start up procedure is also elucidated in the $\lambda = 6$ case, where the previously observed first order convergence after a step size of $\Delta x = 0.005$ is no more. Note that in the plot, convergence for the $\lambda = 3$ case appears to be first order for small step size. This is, however, due to the ill behaved nature of maximal error values. After all, larger λ are found to exhibit better convergence, despite a correspondingly larger time step. Less noticeable increases in convergence are observed for $\lambda = 3$ and 4 as a result. Nevertheless, the improvement in convergence for high frequency terms is noticeable.

Comparing the $R = 1$ and $R = 2$ parameterisations, there are a few noticeable differences. Errors under $R = 2$ are all greater than that of $R = 1$, albeit only by arguably negligible amounts. Additionally, differences in error for contrasting λ are more apparent. In both cases, a larger λ implies slightly larger error. However, in the $R = 2$ case, the increase in error is amplified.

2.4 Richardson Extrapolation

With the low and high frequency terms experiencing differing degrees of convergence, Richardson extrapolations (Section 1.6.1 and 1.6.2) will be implemented to target leading order error terms in both regions. Extrapolations will be conducted on the Crank-Nicolson and $R = 1$ Rannacher time stepping methods. The $R = 2$ case is excluded from analysis, as it provides equivalent results to the $R = 1$ case, but performs slightly worse in terms of accuracy. In targetting error terms of certain frequencies, the aim is to increase accuracy and hasten convergence.

2.4.1 Crank-Nicolson

The Crank-Nicolson method has leading error terms of $O((\Delta t)^2, (\Delta x)^2)$ and $O(\Delta t, \Delta x)$ for low and high frequency terms, respectively. Hence, the $V^{[1]}$ extrapolation (18) is implemented in an attempt to nullify the second order leading error in the low frequency case. Similarly, the $U^{[1]}$ approximation (19) is utilized to target the first order leading error of high frequency error terms. Varying degrees of effectiveness are observed in both cases. Convergence results for $\lambda = 3$ and $\lambda = 10$ are shown in Figure 9.

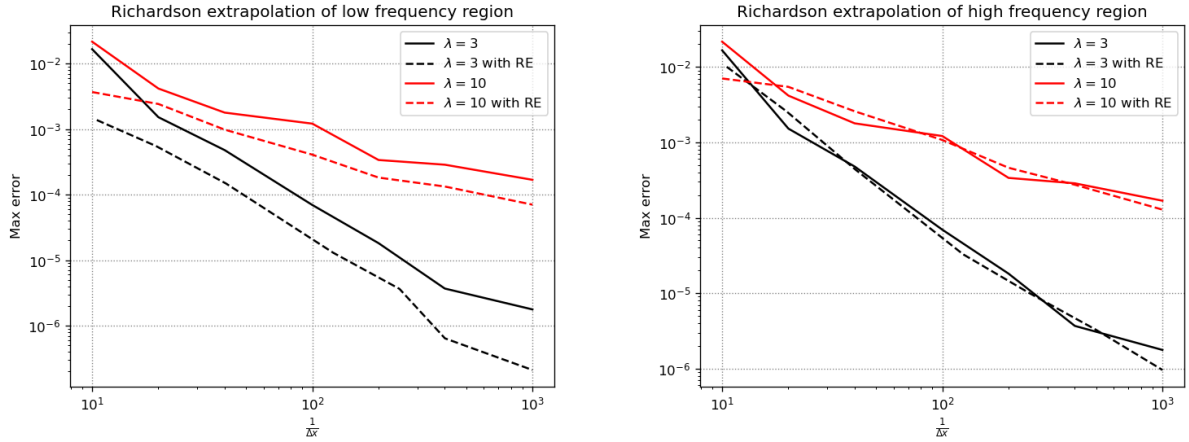


Figure 9: Convergence with Richardson extrapolations (a) $V^{[1]}$ and (b) $U^{[1]}$ targeting the low and high frequency error terms respectively.

Targeting low-frequency error terms in (a) proves to be effective in reducing overall errors for both small and large λ, however, the convergence rate is not altered from the original Crank-Nicolson implementation. In (b), nullifying the first order leading errors of high frequency terms is successful for $\lambda = 3$, where second order convergence is restored past a step size of $\Delta x = 0.0025$. For the larger $\lambda = 10$, the extrapolation is found to be ineffective in increasing the rate of convergence. In both (a) and (b), there is no effect on convergence for relatively large step sizes.

2.4.2 Rannacher timestepping

With $R = 1$ Rannacher timestepping, the truncation error, $O((\Delta t)^2, (\Delta x)^2)$, is unaltered for error terms of low frequency. As a result, the $V^{[1]}$ extrapolation remains appropriate here. Nonetheless, improvement in convergence is observed for the high frequency region; the first order leading error of Crank-Nicolson is superseded. A truncation error of $O((\Delta t)^3, (\Delta x)^3)$ is realised in its place. The appropriate Richardson extrapolation is then $W^{[1]}$ with $R = 1$. Results for the low and high frequency extrapolations are shown in Figure 10.

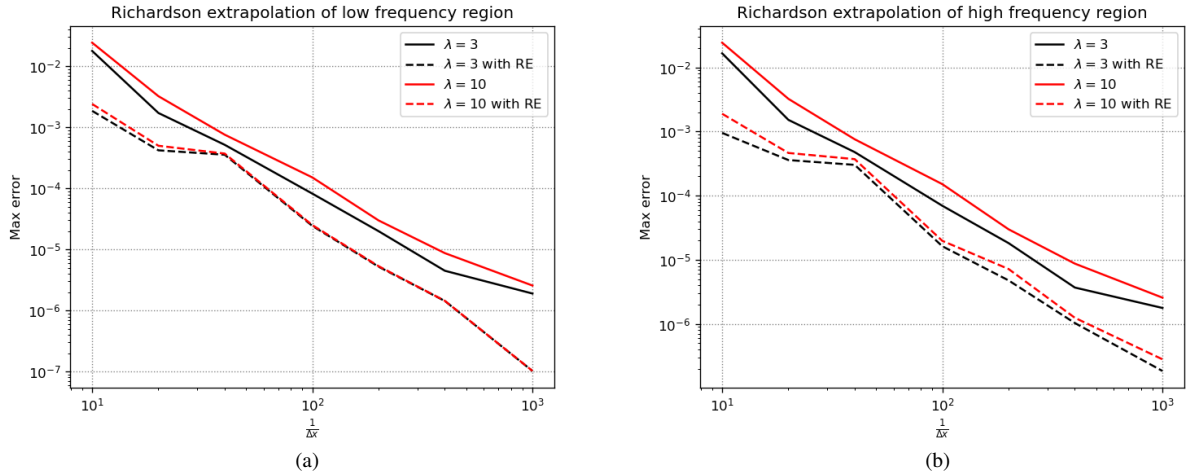


Figure 10: Convergence with Richardson extrapolations (a) $V^{[1]}$ and (b) $W^{[1]}$ targeting the low and high frequency error terms respectively.

For both extrapolations, there is a noticeable increase in accuracy regardless of the λ value. Interestingly, extrapolated convergences are essentially identical for both $\lambda = 3$ and $\lambda = 10$, indicating that higher order terms have less of an effect in this scheme. However, this is more so the case for the targeting of low frequency error terms in (a). Convergence is not improved for large step sizes, despite second order errors being cancelled. There is, however, third order convergence established for step-sizes below $(\Delta x) = 0.0025$. Supplanting the high frequency leading error in (b) is found to not improve the rate of convergence. Agreement of results with theoretical considerations, as well as deviations thereof, will now be discussed in further detail.

3 Discussion

Despite suffering from weakly damped high frequency oscillations, the Crank-Nicolson method performs adequately for all λ, maintaining second order convergence with decreasing step size until a critical Δx is reached. Thereafter, first order convergence begins to be observed (see Figure 6). As the computational grid is refined, altering convergence trends are to be expected; leading order errors only begin to dominate for small enough step sizes. Nevertheless, it is clear that solutions for smaller λ provide better accuracy, and are more resistant to this deterioration in convergence. For larger step size, the contributions of the $O(h^2)$ and $O(h)$ error terms (and those of higher order) to the overall error will be related to their coefficients. In this case, second order convergence is observed. However, as step size decreases the anticipated $O(h)$ error begins to dominate, signifying a transition from low to high-frequency regimes. The critical step size at which a convergence transition is initiated is found to depend on λ; for larger λ the transition from low to high frequency regimes appears to be hastened. Naturally, a larger λ signifies an increase in corresponding step sizes of time. As a result, sizeable values of λ will experience greater effects from persisting higher order error terms. This is what precipitates first order convergence at earlier critical values of Δx. A Fourier analysis of Crank-Nicolson error confirms that high frequency error terms are found to explicitly depend on the value of λ [23].

With Rannacher-timestepping, spurious oscillations are sufficiently dampened and second order convergence is reinstated across all values of λ. This is thanks to the superior stability of the implicit scheme [30]. As suggested by Rannacher (1984), a parameterisation of $R = 1$ is sufficient for reestablishing second order convergence in option value. However, at larger step sizes, the replacement of R Crank-Nicolson steps with $2R$ implicit half steps leads to a slight increase in error. As the implicit scheme exhibits only first order convergence in time, this is to be expected. In theory then, increasing R will result in an increase in error. Indeed, this is observed in comparisons of errors under the $R = 2$ and $R = 1$ parameterisations. Nevertheless, for smaller step sizes where Crank-Nicolson succumbs to first order convergence, the improvement in accuracy is substantial.

Implementing Richardson extrapolation to target leading low frequency errors fails to improve convergence for the Crank-Nicolson scheme. Theoretical considerations suggest this behaviour. The next order low frequency error is $O(h^3)$, but because second order convergence is observed for larger step sizes, no improvement will be observed here; the $O(h^2)$ term, in combination with higher order terms (included due to the larger step size), will still experience second order convergence if the $O(h^2)$ element is removed. Nevertheless, there is a discernible increase in accuracy. This is likely due to the combined effects of decreasing higher order error coefficients as well as cancellation of the leading second order low frequency error. For smaller step sizes, the first order error of high frequency terms will still dominate. The low frequency extrapolation for the Rannacher timestepping scheme proves to be more effective. Giles and Carter (2006) find that the next order error is proportional to $k^3 = (\Delta t)^3$ [23]. For the same reasons, the rate of convergence is initially unaltered for larger step sizes. However, past a step size of $h = 0.0025$, third order convergence is realised. The reason for this is that the leading high frequency term is of order $O(k^3, h^3)$, and because the second order low frequency leading error is cancelled, this begins to dominate as step size is decreased.

Extrapolations targeting high frequency leading errors also behave as expected. In the Crank-Nicolson case, the original leading error is $O(h)$. For small λ, nullification of the first order error does not affect convergence for larger stepsizes. This implies that changes in magnitudes of higher order terms cancel out in the extrapolation. For sizes smaller than $h = 0.0025$, however, there is a noticeable improvement in convergence; the original first order convergence is voided, and succeeded by one of second order. This suggests that the next error term of high frequency oscillations are $O((\Delta t)^2, (\Delta x)^2)$. No discernible changes are noticed in the extrapolation of large λ solutions. The correspondingly larger k for a larger λ means that higher order error terms have a greater effect. Despite cancelling the high frequency leading error, the overall error remains largely impacted by these terms. Hence, the dominating convergence effect is not yet observed. For even smaller stepsizes, it is expected that the second order convergence will be reinstated. However, such calculations will be inefficient. One may as well consider smaller λ. In the Rannacher-timestepping case, the leading high frequency error term is of order $O(h^3)$. Because this is of a greater order than the lower frequencies leading error term, the extrapolation should have no effect on convergence. After all, for large step sizes it can be seen that the involvement of higher order error terms originally results in second order convergence. Nullifying third order errors will therefore increase the influence of the $O(h^2)$ error, as well as altering coefficients of higher order terms. For large step sizes, this appears to result in an unaltered second order convergence with greater accuracy. Naturally, for smaller step size, the $O(h^2)$ error remains dominant, and so there is no discernible change in convergence.

4 Conclusion

Numerical results show that the Crank-Nicolson method, as well as extensions involving the Rannacher time-stepping and Richardson extrapolation procedures, provide reasonable results for the present value of our European call option. Nevertheless, the Crank-Nicolson scheme offers the worst performance, particularly for larger λ. Significant improvement is observed with Rannacher timestepping, which leverages the superior stability of the implicit method. Its implementation eradicates the first order high frequency errors of the Crank-Nicolson scheme. However, this comes at the cost of a slight decrease in accuracy for larger step sizes. A larger step-size is only implemented for faster computation. As results for smaller step sizes can still be calculated relatively quickly, this is not a significant issue.

Richardson extrapolations provide additional improvements in accuracy, excepting the case of Crank-Nicolson with large λ. In fact, the $U^{[1]}$ extrapolation proves to also be effective in reestablishing second order convergence for the Crank-Nicolson method. In the case of Rannacher timestepping, the $V[1]$ extrapolation leads to third order convergence for smaller step size. Extrapolation is found to only improves convergence when targeting the order of error that dominates at small step-size.

Thus, the superior method in valuing our European call option appears to involve both modifications. Namely, a $V[1]$ Richardson extrapolation of the $R = 1$ Rannacher time-stepping scheme. This provides the best accuracy, and third order convergence with decreasing step size. Additionally, the computational efficiency remains reasonable, albeit slower than if the extrapolation procedure were discarded. A smaller λ is found to be preferential in achieving greater accuracy. However, it is also important to consider the cost in computation time relative to the garnered increase of accuracy. An area of further investigation would therefore be to explore optimal λ and method choice in relation to computational efficiency.

The investigation of our European call option serves as a foundation for further exploration of the Rannacher and Richardson modifications. For instance, an interesting extension would be to investigate their application in the valuation of American or Digital options. Comparisons of these two modifications with alternative solutions, such as the moving average or Pearson methods [30], would also be of interest. Indeed, implications of these findings are not merely restricted to finance: They are applicable to any problem involving the Crank-Nicolson scheme and non-smooth initial conditions. For this purpose, an exploration of alternative Rannacher timestep divisions, such as quarter time step replacements, would be of value. The Richardson extrapolation may also be extended to any numerical setting involving known leading order errors, and will provide significant improvements in accuracy at low computational cost.

5 References

1. Hilber, N., Reichmann, O., Schwab, C., & Winter, C. (2013). *Computational Methods for Quantitative Finance*. Springer-Verlag, Berlin Heidelberg.

2. Yalincak, O. (July 22, 2012). *Criticism of the Black-Scholes Model: But Why is It Still Used?: (The Answer is Simpler than the Formula)*. Journal of Applied Economics and Finance, Vol. 7, No. 4. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2115141#:~:text=However%2C%20while%20the%20formula%20has,%2D%C3%A0%2Dvis%20alternative%20models.

3. Austing, P. (2014). *Exotic Options in Black—Scholes. In: Smile Pricing Explained. Financial Engineering Explained.* Palgrave Macmillan, London. https://doi.org/10.1057/9781137335722_5

4. Bernstein, P. (1993). *Capital Ideas: The Improbable Origins of Modern Wall Street*. The Free Press, New York.

5. Black, F. & Scholes, M. (May, 1973). *The Pricing of Options and Corporate Liabilites*. The Journal of Political Economy Vol. 81, No. Retrieved from https://www.cs.princeton.edu/courses/archive/fall09/cos323/papers/black_scholes73.pdf

6. Wilmott, P. (November, 2019). *Frequently Asked Questions in Quantitative Finance, 2nd Edition*. John Wiley & Sons Ltd, West Sussex, United Kingdom. Retrieved from http://101.34.87.177/read/CQF/book/7.%20frequently%20asked%20questions%20in%20quantitative%20fin.pdf

7. Keenan, R. et al. (NA). *Estimating the Volatility in the Black-Scholes Formula*. Retrieved from https://www.valpo.edu/mathematics-statistics/files/2015/07/Estimating-the-Volatility-in-the-Black-Scholes-Formula.pdf

8. Brous, P. et al. (2010). *Volatility forecasting and liquidity: Evidence from individual stocks*. Journal of Derivatives & Hedge Funds, Vol. 16, 144-159. Retrieved from https://link.springer.com/article/10.1057/jdhf.2009.17

9. Hull, J. (2012). *Options, Futures, and Other Derivatives*. Pearson Education Limited, Edinburgh Gate, England.

10. Lorig, M. & Sircar, R. (February, 2015). *Stochastic Volatility: Modeling and Asymptotic Approaches to Option Pricing & Portfolio Selection*. Financial Signal Processing and Machine Learning. Retrieved from https://sircar.princeton.edu/Public/ARTICLES/LorigSircar-VolReview_revised.pdf

11. Dunn, R. et al. (NA). *Estimating Option Prices with Heston's Stochastic Volatility Model*. Retrieved from https://www.valpo.edu/mathematics-statistics/files/2015/07/Estimating-Option-Prices-with-Heston%E2%80%99s-Stochastic-Volatility-Model.pdf

12. Grossinho, M. et al. (June, 2018). *Pricing American Call Options by the Black-Scholes Equation with a Nonlinear Volatility Function*. Retrieved from https://arxiv.org/pdf/1707.00358.pdf,

13. Sterk, W. (2001). *The Relative Accuracy Of The Modified Black-Scholes Model For Options On Underlying Stocks With Dividends*. Journal of Financial Education, Vol. 27, 1-9. Retrieved from https://www.jstor.org/stable/41948351

14. Teichmann, J. (2020). *Fundamental Theorem of Asset pricing*. ETH Zurich lecture notes. Retrieved from https://people.math.ethz.ch/~jteichma/slides_ftap.pdf

15. Rouah, F. (NA). *Four Derivations of the Black Scholes PDE*. Retrieved from https://www.frouah.com/finance%20notes/Black%20Scholes%20PDE.pdf

16. Possamai, D. (2022). *Lecture notes in Mathematical Finance*. ETH Zurich, Autumn term.

17. Viktor Stojkoski, Zoran Utkovski, Lasko Basnarkov, and Ljupco Kocarev. (2019). *Cooperation dynamics in networked geometric brownian motion*. Physical Review E, 99(6).

18. Lecture Notes in Mathematical Finance. *The Feynman-Kac Formula, Risk-Neutral Valuation and the Continuous Black-Scholes Formula*. Imperial College London. Retrieved from https://www.ma.imperial.ac.uk/~bin06/Math482/2015/liv10.pdf

19. Koev, P. (September 6, 2012). *Numerical Methods for Partial DIfferential Equations*. Massachusetts Institute of Technology. Retrieved from https://math.mit.edu/~plamen/18.336/math336.pdf.

20. Ammari, H. (NA). Lecture 6: Finite difference methods. Department of Mathematics, ETH Zürich. Retrieved from https://people.math.ethz.ch/~grsam/SS21/NAII/resources/slides/ODE-Lecture6.pdf.

21. Tekriwal, M. et al. (NA) . *A formal proof of the Lax equivalence theorem for finite difference schemes*. Retrieved from http://www-personal.umich.edu/~jeannin/papers/tekriwal2021formal.pdf

22. Tavella, D. and Randall, C. (2000). *Pricing Financial Instruments, The Finite Difference Method*. John Wiley & Sons New York.

23. Giles, M. & Carter, R. (2006). *Convergence analysis of Crank–Nicolson and Rannacher time-marching*. Journal of Computational Finance, Vol. 9. Retrieved from https://people.maths.ox.ac.uk/gilesm/files/giles_carter.pdf

24. Thomas, J.W. (1998). *Numerical Partial Differential Equations, Volume I Finite Difference Methods*. Springer New York.

25. Richtmyer, R. & Morton, K. (1967). *Difference Methods for Initial-Value Problems*. Wiley-Interscience, 2nd edition.

26. Duffy, D. (2004). *A Critique of the Crank Nicolson Scheme Strengths and Weaknesses for Financial Instrument Pricing*. Retrieved from https://www.ma.imperial.ac.uk/~ajacquie/IC_Num_Methods/IC_Num_Methods_Docs/Literature/DuffyCN.pdf

27. Duffy, D. (1980). *Uniformly Convergent Difference Schemes for Problems with a Small Parameter in the Leading Derivative*. Trinity College Dublin.

28. Rannacher, R. (1984). *Finite element solution of diffusion problems with irregular data*. Numerische Mathematik 43, 309–327.

29. Burden, R. Faires, D. (2011). *Numerical Analysis*. 9th Edition, Brooks/Cole, Boston MA.

30. Osterby, O. (2003). *Five Ways of Reducing the Crank–Nicolson Oscillations*. BIT Numerical Mathematica Vol. 43(4), 811-822. Retrieved from https://www.researchgate.net/publication/226513679_Five_Ways_of_Reducing_the_Crank-Nicolson_Oscillations

31. Lyuu, Y. et al. (January, 2019). *A systematic and efficient simulation scheme for the Greeks of financial derivatives*. Journal of Quantitative Finance, Vol. 19(7). 1199-1219, DOI: 10.1080/14697688.2018.1562196.

32. Jauregui, R & Silva, F. (2011). *Numerical Validation Methods*. Numerical Analysis, Theory and Applications, Ch 8, 155–174.

33. Stewart, A. (2023). *Advanced Computational Techniques*. Lecture Notes, University of St Andrews.

34. Kumar, S. et al. (1993). *An algorithm for the numerical inversion of a tridiagonal matrix*. Numerical Methods in Biomedical Engineering, Vol. 9(4). Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.1640090409-

A Appendix

A.1 Itô's lemma

Consider an underlying Itô process, S_t, defined by the stochastic differential equation

$$dS_t = a(S,t)dt + b(S,t)dW_t$$

and a twice-differentiable function $G(x,t) \in \mathbb{R}$ with $x, t \in \mathbb{R}$. Then, $G(S_t, t)$ will also be an Itô process, defined by

$$dG(S_t,t) = \left(\frac{\partial G}{\partial t} + a\frac{\partial G}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 G}{\partial S^2}\right)dt + b\frac{\partial G}{\partial S}dW_t.$$

This is derived by considering $dG_t = \frac{\partial G}{\partial t}dt + \frac{\partial G}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 G}{\partial S^2}d[S]_t$, where $[S]_t$ is the quadratic variation of S_t.

A.2 Unconditional Stability in (z, τ) Space

Applying Von-Neumann Stability Analysis [25] on the Crank-Nicolson scheme constructed from the diffusion equation yields

$$\begin{aligned} & -\gamma\epsilon_{j-1}^{n+1} + (1+2\gamma)\epsilon_j^{n+1} - \gamma\epsilon_{j+1}^{n+1} = \gamma\epsilon_{j-1}^n + (1-2\gamma)\epsilon_j^n + \gamma\epsilon_j^n \\ \Rightarrow \quad & -\gamma E_m(t+\Delta t)e^{ik_m(y-\Delta y)} + (1+2\gamma)E_m(t+\Delta t)e^{ik_m y} - \gamma E_m(t+\Delta t)e^{ik_m(y+\Delta y)} = \\ & \gamma E_m(t)e^{ik_m(y-\Delta y)} + (1-2\gamma)E_m(t)e^{ik_m y} + \gamma E_m(t)e^{ik_m(y+\Delta y)}. \end{aligned}$$

where $\epsilon_i^n = N_i^n - D_i^n$ is the discretization error at the point corresponding to the indices (i,n), with N_i^n and D_i^n being the numerical and actual solutions respectively. Rearranging to make g the subject of the equation and taking the absolute value results in

$$\begin{aligned} \left|\frac{E_m(t+\Delta t)}{E_m(t)}\right| &= \left|\frac{\gamma(e^{ik_m\Delta y} + e^{-ik_m\Delta y} - 2) + 1}{-\gamma(e^{ik_m\Delta y} + e^{-ik_m\Delta y} - 2) + 1}\right| \\ &= \frac{|-4\gamma\sin^2(\frac{k_m\Delta y}{2}) + 1|}{|4\gamma\sin^2(\frac{k_m\Delta y}{2}) + 1|} \leq 1. \end{aligned}$$

Therefore, the Crank-Nicolson method is found to be *unconditionally stable* for the diffusion equation.