Computational Finance - Hand-in 2

Clara E. Tørsløv (cnp777)

14th November 2021

Simulation

S.1

Replicate figures 2-4 in Frandsen, Pedersen & Poulsen (2021).

We consider the Bachelier model, meaning that (using the same notation as in the referenced article)

$$\begin{aligned} dS(t) &= \sigma dW(t) \\ Call(t) &= \mathbb{E}_t^{\mathbb{Q}}\left[(S(T)-K)^+\right] = F(S(t),t) \\ F(x,t) &= (x-K)\Phi\left(\frac{x-K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{x-K}{\sigma\sqrt{T-t}}\right) \\ \Delta^{true}(x,t) &= \Phi\left(\frac{x-K}{\sigma\sqrt{T-t}}\right) \end{aligned}$$

From a Bachelier simulation we perform a least squares regression of call option payoffs on powers of the starting values.

$$f_\theta(x) = \sum_{i=0}^{p} \theta_i x^i \quad \Rightarrow \quad f'_\theta(x) = \sum_{i=1}^{p} \theta_i i x^{i-1} \tag{1}$$

The left-hand panel of Figure 1 shows the true pricing function (black), an estimated pricing function using a seventh degree polynomial in the regression (red) and the grey dots representing a simulated (initial stock price, call payoff)-pair. The intuition in this figure is that for a given initial stock price a point on the red or black curve represent an expected call payoff, whereas a grey dot represent a realized call payoff.

The right-hand panel of Figure 1 shows the true delta (black) and the delta which we get from the regression function (red), see Eq. 1.

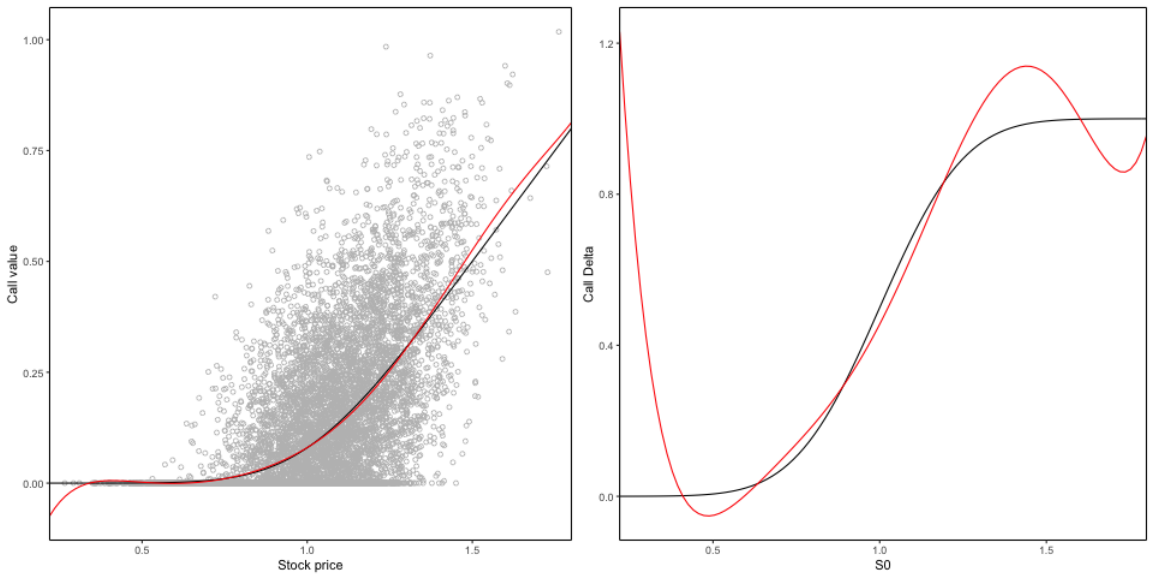

Figure 1: Left: True (black) and estimated (red) call pricing function. Right: True (black) and estimated (red) Δ.

We then expand the regression framework to further examine the estimated Δ functions. To do so we differentiate the call price, interchange expectation and differentiation and apply the chain rule

$$\Delta^{Pathwise} := \frac{\partial Call(0)}{\partial S(0)} = E^{\mathbb{Q}}\left[\frac{\partial}{\partial S(0)}(S(T)-K)^{+}\right] = E^{\mathbb{Q}}\left[\frac{\partial S(T)}{\partial S(0)}\frac{1}{\partial S(T)}(S(T)-K)^{+}\right] = E^{\mathbb{Q}}\left[\mathbb{1}_{S(T)\geq K}\right]$$

where we in the last step have used that the generalized derivative of the $()^{+}$-function is an indicator and

$$\frac{\partial S(T)}{\partial S(0)} = \frac{\partial}{\partial S(0)}\left(S(0)+\sigma W(T)\right) = 1$$

We include this term into the least squares regression criterion (Eq. (2)). The introduction of the right-hand term is what they refer to as Delta-regularization. Figure (2) shows the results of applying different regularization schemes, i.e. using different values of w in Equation 2.

$$\text{crit}(\theta) = w \cdot \sum_j \left(C_j - f_\theta(S_j(0))\right)^2 + (1-w)\cdot \sum_j \left(D_j - f'_\theta(S_j(0))\right)^2 \tag{2}$$

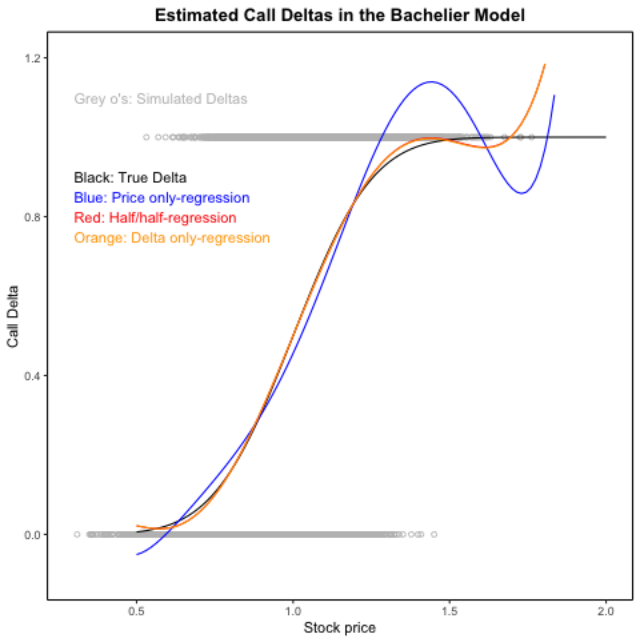

Figure 2: Estimated Δ functions using different regularization schemes

We further test the importance of the choice of regularization scheme in a discrete Delta-hedge experiment. The result is shown in Figure 3. Here we

- Choose a polynomial degree to use in the regression.
- Choose a number of simulations to use in the regression.
- Run separate regressions for expiries $T = 1, \frac{51}{52}, \ldots, \frac{1}{52}$, thus obtaining the respective regression coefficients, $\hat{\theta}$, to be used in the approximated functions, $f_\theta(x)$ and $f'_\theta(x)$
- (Repeat 1000 times) Run a standard discrete Delta-hedge experiment (as we have done hundreds of times in various finance courses), where we use the estimated Δ functions and register the Profit&Loss.
- Estimate the hedge error for the given polynomial order and number of simulations. What we refer to as 'the hedge error' is obtained by dividing the standard deviation of the P&L vector with the initial hedge portfolio price.
- We repeat the above 10 times and average (called 'batching') in order to get one circle in, what they in the article refer to as Figure 4, but here is called Figure 3.

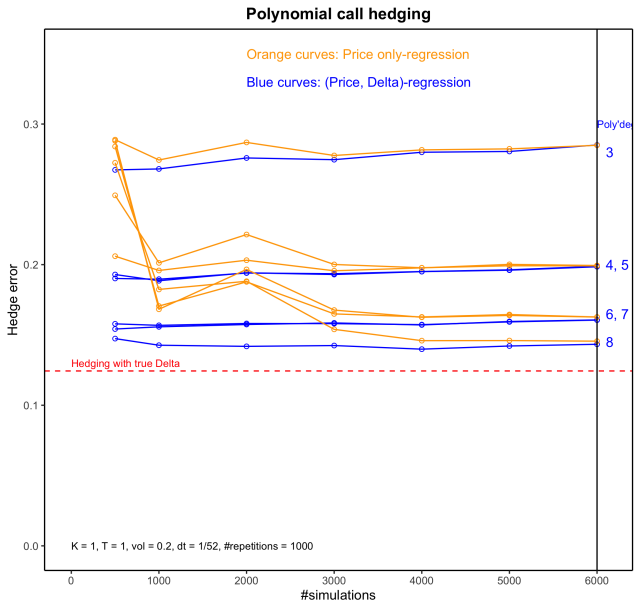

Figure 3: Standard deviation of relative hedge error for different polynomial regressions in the Bachelier model.

S.2

Draw the same graphs but use the Black-Scholes model instead (with the same σ) and comment on any potential differences between the quality of the approximation in the two models.

We now consider the Black-Scholes model, meaning that

$$
\begin{aligned}
dS(t) &= \sigma S(t) dW(t) \\
S(T) &= S(0) e^{-\frac{1}{2}\sigma^2 T + \sigma W(T)} \\
Call(t) &= \mathbb{E}_t^{\mathbb{Q}} \left[(S(T) - K)^+ \right] = F(S(t), t) \\
F(x,t) &= x \cdot \Phi \left(\frac{\log\left(\frac{x}{K}\right) + \frac{1}{2}\sigma^2 (T-t)}{\sigma\sqrt{T-t}} \right) - K \cdot \Phi \left(\frac{\log\left(\frac{x}{K}\right) + \frac{1}{2}\sigma^2 (T-t)}{\sigma\sqrt{T-t}} - \sigma\sqrt{T-t} \right) \\
\Delta^{true} &= \Phi \left(\frac{\log\left(\frac{x}{K}\right) + \frac{1}{2}\sigma^2 (T-t)}{\sigma\sqrt{T-t}} \right) \\
\Delta^{B-S} &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{S(T) \geq K} \frac{S(T)}{S(0)} \right]
\end{aligned}
$$

The last equation follows from the fact that

$$
\frac{\partial S(T)}{\partial S(0)} = \frac{\partial}{\partial S(0)} S(0) e^{-\frac{1}{2}\sigma^2 T + \sigma W(T)} = e^{-\frac{1}{2}\sigma^2 T + \sigma W(T)} = \frac{S(T)}{S(0)}
$$

Using above observations we can follow the same course of action as in S.1 in order to obtain Figures 2-4 from the article. The results are stated in Figure 4, 5 and 6.

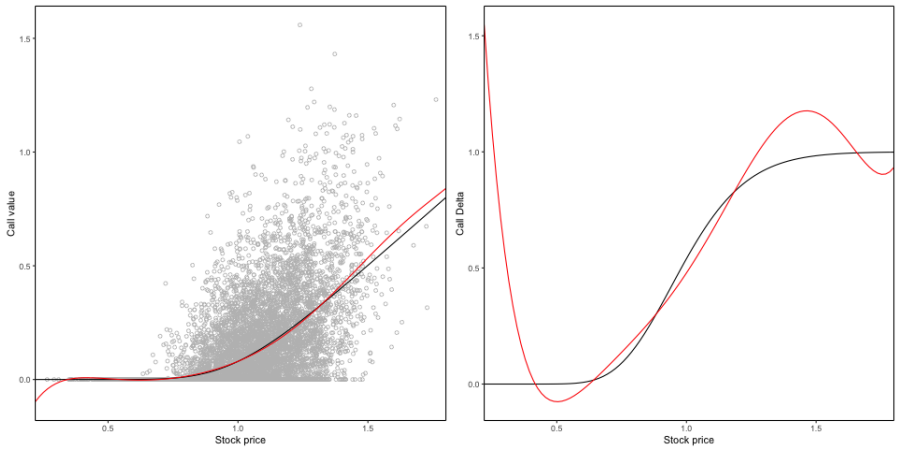

Figure 4: Left: True (black) and estimated (red) call pricing function. Right: True (black) and estimated (red) Δ. In the Black-Scholes model.

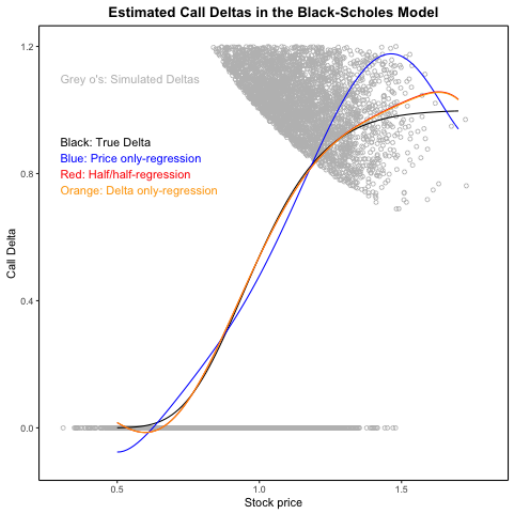

Figure 5: Estimated Δ functions using different regularization schemes in the Black-Scholes model

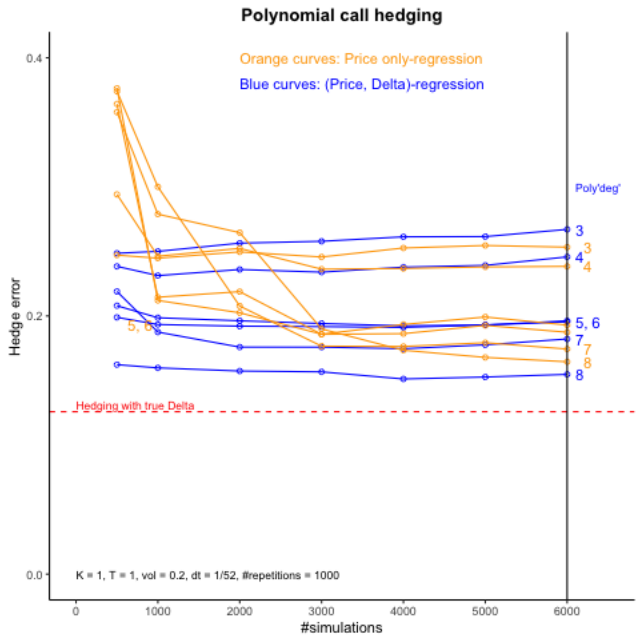

Figure 6: Standard deviation of relative hedge error for different polynomial regressions in the Black-Scholes model.

We see a clear difference in Figure 5, however this is expected as the indicator in the delta regularization term is multiplied with $\frac{S(T)}{S(0)}$ in Δ^{B-S}, which is why we see a 'cloud' of grey dots instead of simply 0's or 1's. Further we notice how the estimated Δ functions (red, orange and blue) seems to be pushed a bit to the right and upwards. This change is more significant in the regression schemes, which use some level of Delta-regularization.

We also see a difference between figures 3 and 6. We first notice that the y-axis is larger, i.e. the largest hedge errors are larger in 6 (Black-Scholes) compared to 3 (Bachelier). This is the case for the 'no delta-regularization' schemes, using polynomial orders 5,6,7 and 8 for low simulations. When #simulations exceeds 3000 the picture looks more or less the same. The only change being the delta-regularization scheme (blue lines) performing noticeable better in the Black-Scholes case when compared to the no-delta-regularization schemes (orange lines) using same polynomial order. We still see a clear increase in hedge performance when the polynomial order is increased.

S.3

Draw a graph similar to that of figure 4, where you compare the quality of the Pathwise and LRM Delta regularization. (You don't need to do it for all polynomial orders, just 3 and 8 is fine) Comment and explain the potential causes of the difference between the two methods.

We consider the Likelihood Ratio Method. We have a random variable of interest, which is a function of the parameter of interest, $Y(\theta)$. The idea in the Likelihood Ratio Method is then to consider this Y to be a function of another random variable, X, which has some density, $g_\theta(x)$, and then place all the dependence of the parameter of interest, θ, into this density.

$$Y(\theta) = f(X)$$

Note that f might be a discontinuous function but not discontinuous in the parameter of interest. Thus we can differentiate the density, as done on slide 62 in the Monte Carlo slides.

$$\frac{\partial}{\partial\theta}\mathbb{E}\left[Y(\theta)\right] = \mathbb{E}\left[f(X)\frac{\partial \log(g_\theta(x))}{\partial\theta}\right]$$

In order for above equation to be useful to us we need an analytical expression for the density. Note that g_θ defines a family of densities and in order for above equation to work they all have to be absolute continuous (i.e. agree on nullsets).

Now consider the Black-Scholes model and a European call option. We wish to get an expression for the delta, thus the parameter of interest is in our case the initial spot, $S(0)$. We then rewrite the model such that the parameter of interest is placed in the density.

$$S(T) = S(0)e^{-\frac{1}{2}\sigma^2 T+\sigma\sqrt{T}Z} = e^{-\frac{1}{2}\sigma^2 T+\sigma\sqrt{T}X}$$

where $Z \sim \mathcal{N}(0,1)$ and $X \sim \mathcal{N}\left(\frac{\log(S(0))}{\sigma\sqrt{T}}, 1\right) = \mathcal{N}(\mu_x, \sigma_x)$. We then have that

$$\begin{aligned}(\theta) = f(X) \quad &\Rightarrow \quad \mathbb{E}[(S(T)-K)^+] = \mathbb{E}[f(X)], \quad \text{where } f(x) = \left(e^{-\frac{1}{2}\sigma^2 T+\sigma\sqrt{T}x} - K\right)^+ \\ &\Rightarrow \quad \frac{\partial}{\partial S(0)}\mathbb{E}[(S(T)-K)^+] = \mathbb{E}\left[f(X)\frac{\partial \log(g_{S(0)}(X))}{\partial S(0)}\right]\end{aligned}$$

We note, as $g_{S(0)}(x) = \frac{1}{\sigma_x}\phi\left(\frac{x-\mu_x}{\sigma_x}\right)$, where ϕ is the std. normal PDF, that

$$\frac{\frac{\partial g_{S(0)}(x)}{\partial S(0)}}{g_{S(0)}(x)} = \frac{1}{g_{S(0)}(x)}\frac{\partial}{\partial S(0)}\phi\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right)$$

Using the chain rule

$$= \frac{1}{g_{S(0)}(x)}\phi'\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right)\cdot\frac{-1}{\sigma\sqrt{T}S(0)}$$

Using that $\phi'(x) = -x\phi(x)$

$$\begin{aligned}
&= \frac{1}{g_{S(0)}(x)}\phi\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right)\cdot\frac{x - \frac{\log(S(0))}{\sigma\sqrt{T}}}{\sigma\sqrt{T}S(0)} \\
&= \frac{1}{\phi\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right)}\phi\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right)\cdot\frac{x - \frac{\log(S(0))}{\sigma\sqrt{T}}}{\sigma\sqrt{T}S(0)} \\
&= \frac{x - \frac{\log(S(0))}{\sigma\sqrt{T}}}{\sigma\sqrt{T}S(0)}
\end{aligned}$$

Thus,

$$\Delta^{LRM} = \mathbb{E}\left[f(X)\frac{\partial\log(g_\theta(X))}{\partial\theta}\right] = \mathbb{E}\left[f(X)\frac{X - \frac{\log(S(0))}{\sigma\sqrt{T}}}{\sigma\sqrt{T}S(0)}\right] = \mathbb{E}\left[f(X)\frac{X-\mu_x}{\sigma\sqrt{T}S(0)}\right] = \mathbb{E}\left[f(X)\frac{Z}{\sigma\sqrt{T}S(0)}\right]$$

What is worth noting here is that we have differentiated the payoff function without actually differentiating the payoff function. We obtain something which has its expectation equal to the Δ. We can use this observation in the same way as in S.2 above. We compare these two in a discrete delta-hedge experiment, where both strategies uses Delta-regularization but different ways of calculating the D_j's used in the regression. The result is stated in Figure 7.

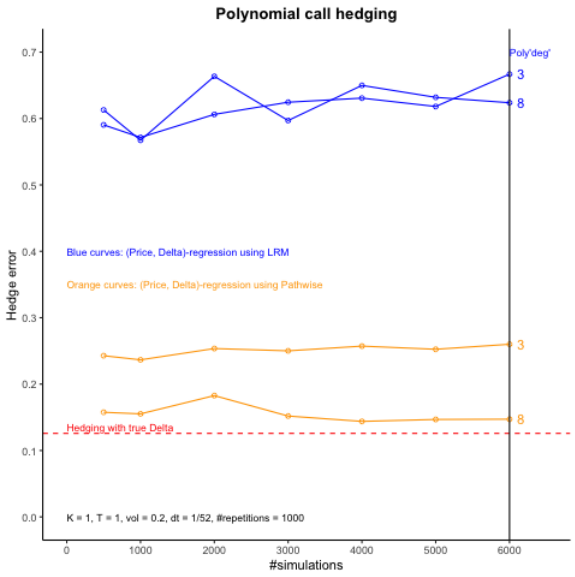

Figure 7: Standard deviation of relative hedge error for different Delta-regularization schemes in the Black-Scholes model

We see from Figure 7 that the Pathwise differentiation method performs significantly better than when the Likelihood Ratio Method has been applied. We examine the LRM approach closer. We repeat the calculations from slide 62, remembering $Y(\theta) = f(X)$, where X is a random variable with density g_θ that depends on θ (think $\theta = S_0$ and $Y = (S_T - K)^+$).

$$\begin{aligned}
\mathbb{E}[Y(\theta)] &= \mathbb{E}[f(X)] \frac{\partial}{\partial \theta} \int_{\mathbb{R}^m} f(x) g_\theta(x) dx \\
&= \int_{\mathbb{R}^m} f(x) \frac{\partial}{\partial \theta} g_\theta(x) dx \\
&= \int_{\mathbb{R}^m} f(x) \frac{\partial g_\theta(x)}{\partial \theta} \frac{g_\theta(x)}{g_\theta(x)} dx \\
&= \int_{\mathbb{R}^m} f(x) \frac{\frac{\partial g_\theta(x)}{\partial \theta}}{g_\theta(x)} g_\theta(x) dx \\
&= \mathbb{E}\left[f(X) \frac{\frac{\partial g_\theta(X)}{\partial \theta}}{g_\theta(X)} \right]
\end{aligned}$$

Notice the red text here - we recognize this trick from Importance Sampling (slide 36). By this slide we know that the Importance Sampling estimator is unbiased iff. the densities in the final expectation agrees on nullsets, i.e. if

$$\frac{\partial g_\theta(x)}{\partial \theta} > 0 \quad \Leftrightarrow \quad g_\theta(x) > 0 \quad \forall x \in \mathbb{R}^m$$

We examine this, as $X \sim \mathcal{N}(\frac{\log(S(0))}{\sigma\sqrt{T}}, 1)$ (by slide 63) we know that

$$\frac{\partial}{\partial S(0)} g_{S(0)}(x) = \phi\left(x - \frac{\log(S(0))}{\sigma\sqrt{T}}\right) \frac{x - \frac{\log(S(0)}{\sigma\sqrt{T}}}{\sigma\sqrt{T}S(0)}$$

Now note that $g_{S(0)}(x) > 0 \, \forall x \in \mathbb{R}^m$, per definition of the standard normal density function, and that

$$\frac{\partial g_{S(0)}(x)}{\partial S(0)} > 0 \quad \Leftrightarrow \quad x > \frac{\log(S(0)}{\sigma\sqrt{T}}$$

Meaning that our Importance Sampling estimator is biased. We remember the example from lectures where we considered a Black Scholes put option. Here we found that the efficiency ratio when the put was OTM were around 4 (i.e. the variance when using IS is smaller than that of the standard Monte Carlo) and when the put was ITM the efficiency ratio were significantly different (around 0.18) thus we were actually increasing the variance by switching from regular Monte Carlo to an Importance Sampling scheme. What we see and our interpretation thereof is in line with the conclusion on slide 73, which states

The LRM method is often not preferable if the pathwise is applicable since $f(X)\frac{\partial \log(g_\theta(X))}{\partial \theta}$ *may often have a significantly larger variance than* $f(X)$.

This additional noise to the Monte Carlo estimator can also be noticed analytically. Assume we are in the same Black-Scholes model but consider a very smooth payoff - namely a constant payoff, $f(x) = c$. Then

$$\mathbb{E}\left[f(S(T))\frac{\frac{\partial g_{S(0)}(S(T))}{\partial S(0)}}{g_{S(0)}(S(T))}\right] = \int_{\mathbb{R}} f(x)\frac{\partial}{\partial S(0)} g_{S(0)}(x)dx$$

Using substitution $x = log(y) \Rightarrow dx = \frac{1}{y}dy$. And, as $\log(S(T) \sim \mathcal{N}(\log(S(0)) - \frac{1}{2}\sigma^2 T, \sigma^2 T)$, then $g_{S(0)}(x) = \frac{1}{\sigma_{ST}}\phi\left(\frac{x-\mu_{ST}}{\sigma_{ST}}\right)$

$$= \int_{\mathbb{R}} \frac{\partial}{\partial S(0)} \frac{1}{\sigma\sqrt{T}} \phi\left(\frac{x - \log(S(0)) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) dy$$

Using the chain rule

$$\begin{aligned}
&= -\frac{1}{\sigma^2 T S(0)} \int_{\mathbb{R}} \frac{\partial}{\partial y} \phi\left(\frac{x - \log(S(0)) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) dy \\
&= -\frac{1}{\sigma^2 T S(0)} \left[\phi\left(\frac{x - \log(S(0)) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right)\right]_{-\infty}^{\infty} \\
&= 0
\end{aligned}$$

Here, the Monte Carlo estimator will be

$$\mathbb{E}\left[f(S(T))\frac{\frac{\partial g_{S(0)}(S(T))}{\partial S(0)}}{g_{S(0)}(S(T))}\right] \approx \frac{1}{n}\sum_{i=1}^{n} c \cdot \frac{\frac{\partial g_{S(0)}(S(\omega_i,T))}{\partial S(0)}}{g_{S(0)}(S(\omega_i,T))}$$

where ω_i indicates the i'th realiztion of a simulated time-T stock price. This Monte Carlo estimator is non-zero. Whereas if we would have used the Pathwise method, we would have instead differentiated the payoff function in the Monte Carlo Estimator, which would then have been zero and thus the correct result.

S.4

Consider now a digital option. Explain why the pathwise method cannot be used and use the LRM method instead to draw graphs comparable to figures 2-4. Do so for the same Black-Scholes model as in question S.2. Why do think the errors are much higher than for the call option?

We still consider the Black-Scholes model, but now consider the Digital call option. This has payoff

$$Call^{Digital}(t) = \mathbb{E}_t^{\mathbb{Q}}[\mathbb{1}_{S(T)>K}] = F^{Digital}(S(t),t)$$

Now, how does $F^{Digital}$ look? $S(T)$ is a random variable where we, by the Black-Scholes model, know

$$\log(S(T)) \sim \mathcal{N}(\underbrace{\log(S(t)) - \frac{1}{2}\sigma^2(T-t)}_{\mu_S}, \underbrace{\sigma^2 T}_{\sigma_S^2})$$

Thus

$$\frac{\log(S(T)) - \mu_S}{\sigma_S} \sim \mathcal{N}(0,1)$$

The digital call option is exercised if $S(T) > K \Leftrightarrow \log(S(T)) > \log(K)$. Thus the probability of the digital call being exercised is

$$\begin{aligned}
\mathbb{P}\left(\log(S(T)) > \log(K)\right) &= \mathbb{P}\left(\frac{\log(S(T)) - \mu_S}{\sigma_S} > \frac{\log(K) - \mu_S}{\sigma_S}\right) \\
&= 1 - \Phi\left(\frac{\log(K) - \mu_S}{\sigma_S}\right) \\
&= \Phi\left(-\frac{\log(K) - \mu_S}{\sigma_S}\right) \\
&= \Phi\left(-\frac{\log(K) - \log(S(t)) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}\right) \\
&= \Phi\left(\frac{\log(\frac{K}{S(t)}) - \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}\right) \\
&= \Phi\left(d_2(S(t),t)\right)
\end{aligned}$$

We recognize the d_2 from the Black-Scholes formula. We can then determine the analytical $\Delta^{Digital}$.

$$\Delta^{Digital}(S(t),t) = \frac{\partial}{\partial S(t)}\Phi\left(d_2(S(t),t)\right) = \phi\left(d_2(S(t),t)\right)\frac{\partial}{\partial S(t)}d_2(S(t),t) = \phi\left(d_2(S(t),t)\right)\frac{1}{\sigma\sqrt{T-t}S(t)}$$

In order to again replicate Figures 2-4 from Frandsen, Pedersen & Poulsen (2021) we need to use Delta-regularization when calculating the regression coefficient. In S.2 we used the Pathwise method to generate the D_j's to be used in the Delta-regularization. However the Pathwise method cannot be used in the case of the digital call option. Remember that the Pathwise estimator is calculated by interchanging the order of differentiation and expectation. Here we just assumed this interchange was justified but this is not always the case. The general rule of thumb is that the interchange is legal when the payoff, Y, is almost surely continuous in the parameter of interest, θ (MC slide 57). In the case of the digital call option, the payoff is clearly discontinuous and we have that $\frac{\partial Y}{\partial S(0)} = 0$ everywhere except in $S(T) = K$ which is a point of non-differentiability, so

$$\mathbb{E}\left[\frac{\partial Y}{\partial S(0)}\right] = 0 \neq \frac{\partial}{\partial S(0)}\mathbb{E}\left[Y\right]$$

Intuitively a change in the initial spot should cause a change in the expected payoff due to the possibility that the change in $S(0)$ could cause $S(T)$ to exceed (or fall below) the strike, K. However this change is not captured by the Pathwise derivative which is zero almost surely.
Instead we use the Likelihood Ratio Method. As we use the same Black-Scholes model as in S.3 and since there is no payoff specific calculations in S.3 we get that the delta for the Digital Call option is

$$\Delta^{LRM} = \mathbb{E}\left[\mathbb{1}_{(S(T)>K)}\frac{Z}{\sigma\sqrt{T}S(0)}\right]$$

The results are stated in Figures 8, 9 and 10.

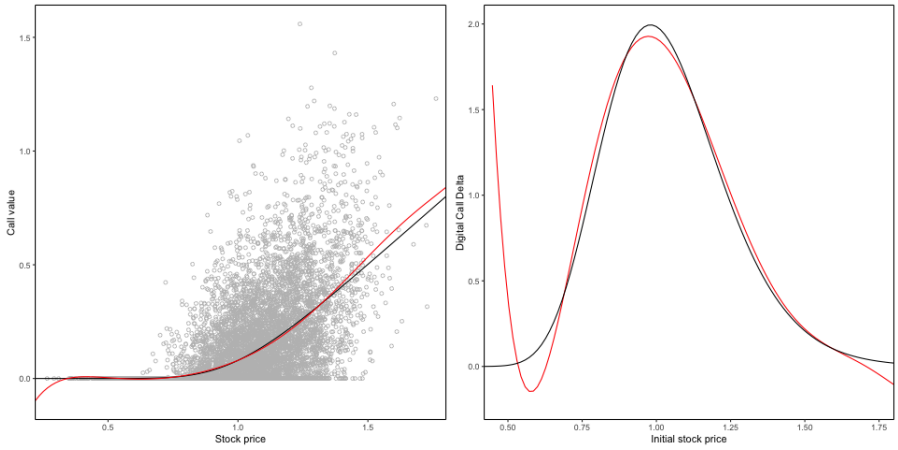

Figure 8: Left: True (black) and estimated (red) digital call pricing function. Right: True (black) and estimated (red) digital call Δ. In the Black-Scholes model.

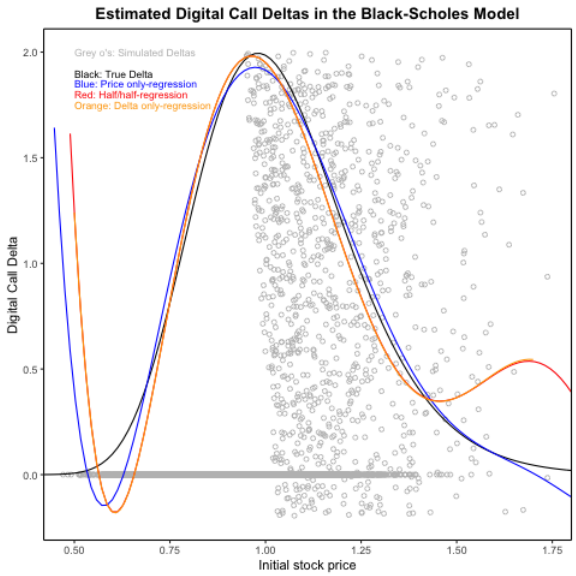

Figure 9: Estimated Δ functions using different regularization schemes in the Black-Scholes model

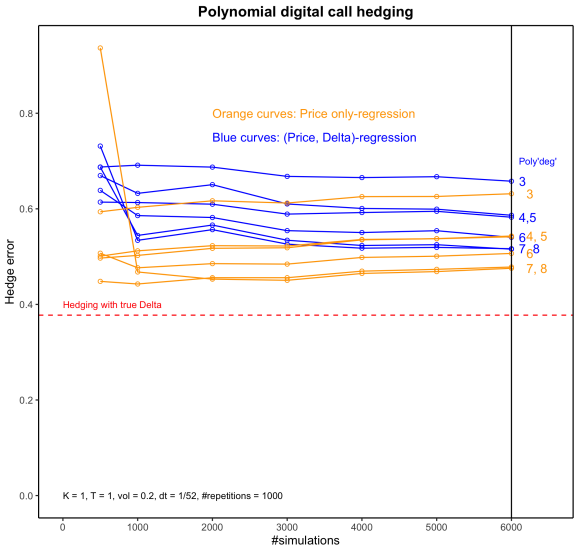

Figure 10: Standard deviation of relative hedge error for different polynomial regressions in the Black-Scholes model.

What we first note here is that the hedge errors are significantly larger than those for the call option, even when hedging with the true Delta. The higher hedge errors could be a product of us using the Likelihood Ratio Method, as we know from S.3 that this often will entail a significantly larger error. In the case of the Call option we were able to use the Pathwise method to obtain a smaller hedge error. In the case of the Digital Call option we know that we are not able to use the Pathwise method. We saw in Figure 7 that when using the LRM approach entails a significantly higher hedge error compared to the Delta regularization scheme using the Pathwise approach.

S.5

Do the same as in the previous question but apply the Mixed method estimator (see MC slides p. 69) with an appropriately chosen ϵ. Comment on the potential differences between the pure LRM and the Mixed method.

Instead of using the pure LRM approach as in S.4 one could instead combine the Pathwise and Likelihood Ratio estimators in order to leverage the strengths of each approach. We consider the Digital Call option, thus the Pathwise approach cannot be used directly, however there is another trick we can use. Consider

$$\mathbb{1}_{x>K} = f_\epsilon(x) + (\mathbb{1}_x > K) - f_\epsilon(x)) = f_\epsilon(x) + h_\epsilon(x)$$

where

$$f_\epsilon(x) = \min\left(1, \frac{\max(0, x - K + \epsilon)}{2\epsilon}\right)$$

The intuition here is that $f_\epsilon(x)$ is a piece-wise linear approximation to the payoff function, $\mathbb{1}_{x>K}$, and $h_\epsilon(x)$ corrects the approximation. Then, as $f_\epsilon(S(T))$ is almost surely continuous in $S(0)$, we can apply the Pathwise approach here and the Likelihood Ratio approach to $h_\epsilon(S(T))$.

$$\frac{\partial}{\partial S(0)}\mathbb{E}\left[f_\epsilon(S(T))\right] = \mathbb{E}\left[f'_\epsilon(S(T))\frac{\partial S(T)}{\partial S(0)}\right]$$
$$\frac{\partial}{\partial S(0)}\mathbb{E}\left[h_\epsilon(S(T))\right] = \mathbb{E}\left[h_\epsilon(S(T))\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))\right]$$

where $g_{S(0)}$ is the density of $S(T)$. Thus the general formula for the Mixed estimator for the delta of a Digital Call option is given by

$$\begin{aligned}\Delta^{mixed} &= \frac{\partial}{\partial S(0)}\mathbb{E}\left[\mathbb{1}_{x>K}\right] = \mathbb{E}\left[f'_\epsilon(S(T))\frac{\partial S(T)}{\partial S(0)} + h_\epsilon(S(T))\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))\right] \\ &= \mathbb{E}\left[\underbrace{f'_\epsilon(S(T))\frac{\partial S(T)}{\partial S(0)} - f_\epsilon(S(T))\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))}_{\text{mean zero}} + \underbrace{\mathbb{1}_{S(T)>K}\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))}_{\text{LR estimator (mean delta)}}\right]\end{aligned}$$

We recognize the two left hand terms as a control variable (CV) to the Likelihood Ratio Estimator. We now insert the Black-Scholes specific expressions. We remember

$$\frac{\partial S(T)}{\partial S(0)} = \frac{S(T)}{S(0)}$$
$$\frac{\partial \log(g_{S(0)})(S(T))}{\partial S(0)} = \frac{Z}{\sigma\sqrt{T}S(0)}$$

We further note that

$$f'_\epsilon(x) = \frac{1}{2\epsilon}\mathbb{1}_{|x-K|<\epsilon}$$

Now what is left is to choose the ϵ. We choose the ϵ which minimizes the variance of Δ^{mixed}. We define a function, which takes in ϵ and returns the weighted variance of Δ^{mixed}. We examine this and the result is stated in Figure 11.

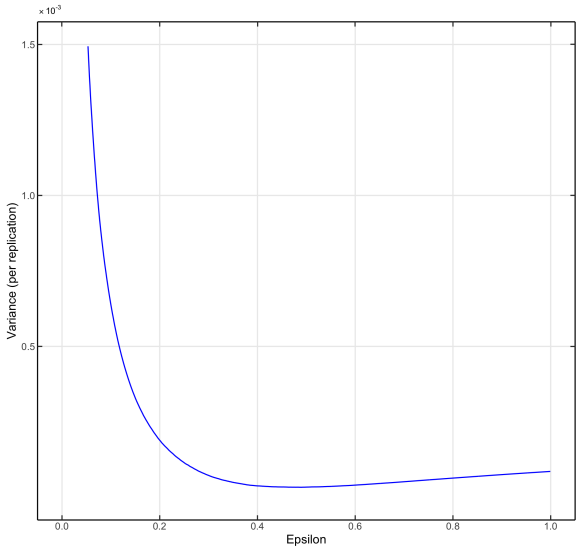

Figure 11: Variance (per replication) of the Mixed Estimator of the delta of a digital call as a function of ϵ. Parameters used is $K = 1$, $S(0) = 1$, $T = 1$, $\sigma = 0.2$, $r = 0$, and $n = 10.000$.

We conclude that the variance minimizing ϵ is 0.48[1]. We proceed using this value. One could argue that it would not be 'correct' to use the same ϵ for different times to maturity, as, if it were possible to calculate the analytical expression of
$mathbbV[\Delta^{mixed}]$, one could imagine this would depend on T (and S(0)). We see a clear difference when changing T to 0.5 rather than 1 (Figure 12). Now the variance minimizing ϵ is equal to 0.33.

[1]For the precise number we have used the R-function optimize().

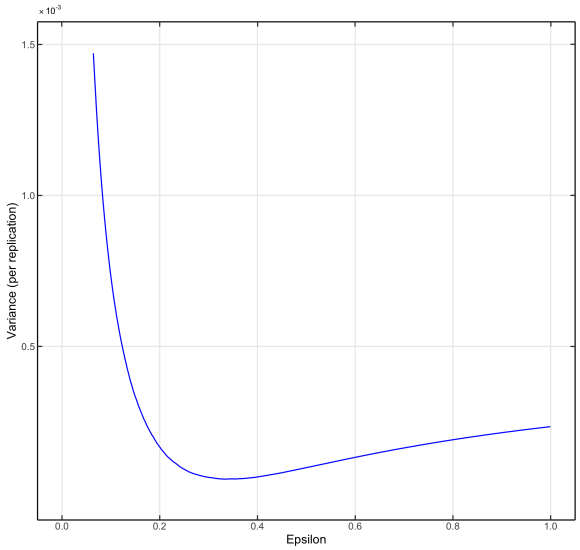

Figure 12: Variance (per replication) of the Mixed Estimator of the delta of a digital call as a function of ϵ. Parameters used is $K = 1$, $S(0) = 1$, $T = 0.5$, $\sigma = 0.2$, $r = 0$, and $n = 10.000$.

One could incorporate this observation into the code for replication of Figure 4 by including an optimization of the variance of the Mixed estimator (Eq. (116) in MC slides) in each timestep. However, with our eyes wide open, we choose to use a constant $\epsilon = 0.48$. The results are stated in Figures 13, 14 and 15.

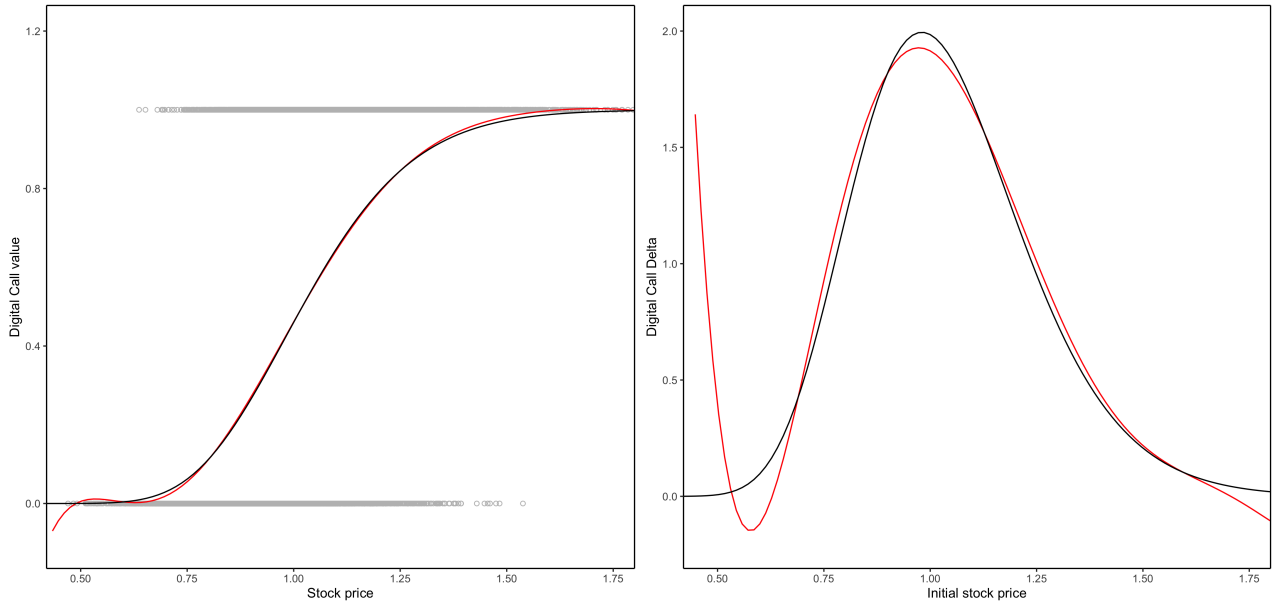

Figure 13: Left: True (black) and estimated (red) digital call pricing function. Right: True (black) and estimated (red) digital call Δ. In the Black-Scholes model.

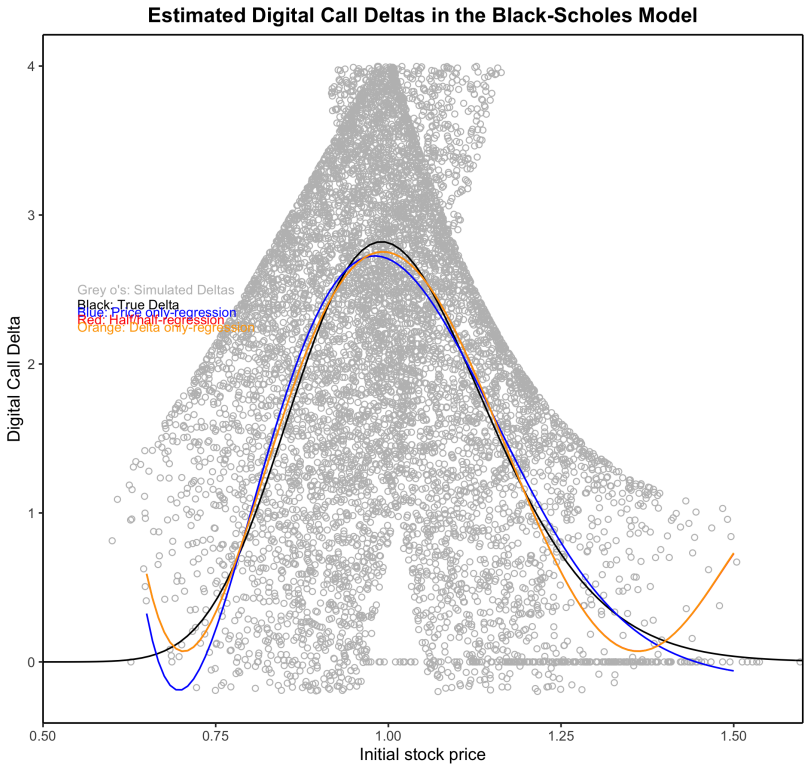

Figure 14: Estimated Δ functions using different regularization schemes in the Black-Scholes model

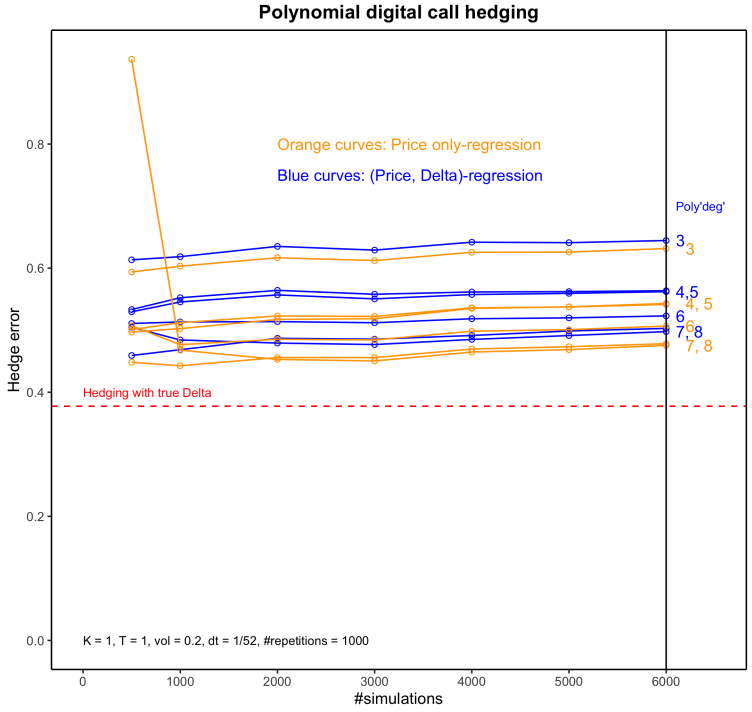

Figure 15: Standard deviation of relative hedge error for different polynomial regressions in the Black-Scholes model.

When comparing Figures 10 (LRM) and 15 (Mixed) we note significantly lower hedge errors. We further note less difference between price only and delta-regularization schemes. In Figure 10 the Delta-regularization schemes performed noticeable worse than the price-only regression schemes, whereas in Figure 15 we see that the Delta-regularization schemes performs better than the price-only regression schemes. To better see the direct difference between the Mixed and the Pure Likelihood Ratio method note that

$$\begin{aligned}\lim_{\epsilon\to 0}\Delta^{mixed} &= \lim_{\epsilon\to 0}\mathbb{E}\left[f'_\epsilon(S(T))\frac{\partial S(T)}{\partial S(0)} - f_\epsilon(S(T))\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T)) + \mathbb{1}_{x>K}\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))\right] \\ &= \mathbb{E}\left[\mathbb{1}_{S(T)>K}\frac{\partial \log(g_{S(0)})}{\partial S(0)}(S(T))\right] \\ &= \Delta^{LRM}\end{aligned}$$

As we saw in Figure 11 when ϵ decreases below a certain value the variance of Δ^{mixed} increases significantly. This is in line with what we have seen throughout this paper - the variance when using Δ^{LRM} is higher when compared to the other methods considered.

Finite Difference

FD.1

Replicate the numbers in column 5 in the table using a suitable implementation of a suitable finite difference method.

In the following we use the Crank-Nicholson method and notations as in Munk note

We consider the problem of determining some function $f(x,t)$, $f : \mathcal{S} \times [0,T] \to \mathbb{R}$, which solves the PDE

$$\frac{\partial f}{\partial t}(x,t) + \mu(x,t)\frac{\partial f}{\partial x}(x,t) + \frac{1}{2}\sigma(x,t)^2\frac{\partial^2 f}{\partial x^2}(x,t) - r(x,t)f(x,t) = 0, \quad (x,t) \in \mathcal{S} \times [0,T) \tag{3}$$

with some terminal condition

$$f(x,T) = F(x)$$

where $F : \mathcal{S} \to \mathbb{R}$ is a known function. As we consider a put option then, for some strike value K,

$$f(x,T) = \max(K - x)$$

According to the Munk note we know that we can do a discretization of the problem.

The main equation in the Crank-Nicholson approach is the following (equation (19) in the Munk note).

$$A_{j,n}f_{j-1,n} + B_{j,n}f_{j,n} + C_{j,n}f_{j+1,n} = -A_{j,n+1}f_{j-1,n+1} + B^*_{j,n+1}f_{j,n+1} - C_{j,n+1}f_{j+1,n+1} \tag{4}$$

where

$$\begin{aligned}
A_{j,n} &= \frac{1}{4}\Delta t\left(\frac{\sigma^2_{j,n}}{(\Delta x)^2} - \frac{\mu_{j,n}}{\Delta x}\right)\\
B_{j,n} &= -1 - \frac{1}{2}\Delta t\left(\frac{\sigma^2_{j,n}}{(\Delta x)^2} + r_{j,n}\right)\\
C_{j,n} &= \frac{1}{4}\Delta t\left(\frac{\sigma^2_{j,n}}{(\Delta x)^2} + \frac{\mu_{j,n}}{\Delta x}\right)\\
B^*_{j,n} &= -1 + \frac{1}{2}\Delta t\left(\frac{\sigma^2_{j,n}}{(\Delta x)^2} + r_{j,n}\right)
\end{aligned}$$

Equation (4) in matrix/vector notation reads

$$Af_n = Bf_{n+1}$$

Where A is a tridiagonal matrix, with $B_{j,n}$ as the main diagonal, $C_{j,n}$ as the first diagonal above the main diagonal and $A_{j,n}$ as the first diagonal below the main diagonal. B is a tridiagonal matrix, with $B^*_{j,n}$ as the main diagonal, $-A_{j,n}$ as the first diagonal below the main diagonal, and $-C_{j,n}$ as the first diagonal above the main diagonal. Lets write it out so it becomes a bit easier to imagine. The system of equations has the following tridiagonal matrix structure

$$\begin{pmatrix} A_{1,n} & B_{1,n} & C_{1,n} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & A_{2,n} & B_{2,n} & C_{2,n} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & & 0 & 0 & 0 \\ \vdots & \vdots & & \ddots & \ddots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & A_{J-2,n} & B_{J-2,n} & C_{J-2,n} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & A_{J-1,n} & B_{J-1,n} & C_{J-1,n} \end{pmatrix} \begin{pmatrix} f_{0,n} \\ f_{1,n} \\ \vdots \\ f_{J-1,n} \\ F_{J,n} \end{pmatrix} =$$

$$\begin{pmatrix} -A_{1,n} & B^*_{1,n} & -C_{1,n} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -A_{2,n} & B^*_{2,n} & -C_{2,n} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & & 0 & 0 & 0 \\ \vdots & \vdots & & \ddots & \ddots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -A_{J-2,n} & B^*_{J-2,n} & -C_{J-2,n} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & -A_{J-1,n} & B^*_{J-1,n} & -C_{J-1,n} \end{pmatrix} \begin{pmatrix} f_{0,n+1} \\ f_{1,n+1} \\ \vdots \\ f_{J-1,n+1} \\ F_{J,n+1} \end{pmatrix}$$

This system of equations has J-1 equations with J+1 unknowns. We fix this by introducing two boundary conditions for respectively high and low values of x. In our case of the put option, we know that for very high values of x (i.e. a high stock price) the put option will be very OTM and reversely for very low values of x the put option will be very ITM. Thus it will be reasonable to assume (as done in the Munk note) that

$$f_{0,n} = F(x_{min}, t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(K-x)^+ | x = x_{min}] = e^{-r(T-n\Delta t)}K$$
$$f_{J,n} = F(x_{max}, t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(K-x)^+ | x = x_{max}] = 0$$

We add the two equations and, using the same matrix notation, we arrive at the following quadratic matrix system of equations

$$\begin{pmatrix} B_{1,n} & C_{1,n} & 0 & 0 & \dots & 0 & 0 \\ A_{2,n} & B_{2,n} & C_{2,n} & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & A_{J-2,n} & B_{J-2,n} & C_{J-2,n} \\ 0 & 0 & \dots & 0 & 0 & A_{J-1,n} & B_{J-1,n} \end{pmatrix} \begin{pmatrix} f_{1,n} \\ \vdots \\ f_{J-1,n} \end{pmatrix} =$$

$$\begin{pmatrix} B*_{1,n} & -C_{1,n} & 0 & 0 & \dots & 0 & 0 \\ -A_{2,n} & B*_{2,n} & -C_{2,n} & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & -A_{J-2,n} & B*_{J-2,n} & -C_{J-2,n} \\ 0 & 0 & \dots & 0 & 0 & -A_{J-1,n} & B*_{J-1,n} \end{pmatrix} \begin{pmatrix} f_{1,n+1} \\ \vdots \\ f_{J-1,n+1} \end{pmatrix} +$$

$$\begin{pmatrix} -A_{1,n+1} - A_{1,n} f_{0,n} \\ 0 \\ \vdots \\ 0 \\ -C_{J-1,n+1} - C_{J-1,n} f_{0,n} \end{pmatrix}$$

We now have $J-1$ equations with $J-1$ unknowns. In order to solve this system we will perform successive iterations, starting with the known values at the date of maturity, working backwards for $t = N-1, ..., 0$. In Section 7 of the Munk Note, an algorithm to solve the above problem is implemented.

The above result is applicable for any PDE, but we are interested in the specific case of the Black Scholes model. The Black Scholes PDE is given by the following, where both the short rate and volatility is constant over time,

$$\frac{\partial f}{\partial t}(x,t) + r \cdot x \frac{\partial f}{\partial x}(x,t) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^f}{\partial x^2}(x,t) - rf(x,t)$$

Comparing to (3) we note that

$$\mu(x,t) = rx$$
$$\sigma(x,t)^2 = \sigma^2 x^2$$

We are now able to implement our chosen finite difference method. The result is given in Table (1). As discussed in the lectures, we know that the Crank-Nicolson approach (which averages the implicit and explicit method) is a lot more precise compared to the implicit method. Therefore less time steps and stock steps are needed in order to get satisfying results. The Crank-Nicolson approach is recommended and as stated in Jesper Andreasen's slideset 1 will almost always be 'the weapon of choice for the pro', which is why we deemed it a 'suitable' finite difference method. Jesper Andreasen further recommends in his slideset 3 that, for Crank-Nicolson, the number of spatial points is 2-4 times bigger than the number of time points. We use 1000 spatial points (x) and 300 time steps (t). From Table (1) we note that, using the above described scheme, we get almost identical results compared to the closed form solution (Table 1, column 5 in the Longstaff & Schwartz paper), while using a significantly smaller number of steps.

Spot	σ	T	Closed form EU	Finite Difference EU
36	0.2	1	3.844	3.844
36	0.2	2	3.763	3.763
36	0.4	1	6.711	6.711
36	0.4	2	7.700	7.699
38	0.2	1	2.852	2.852
38	0.2.	2	2.991	2.991
38	0.4	1	5.834	5.834
38	0.4	2	6.979	6.977
40	0.2	1	2.066	2.066
40	0.2	2	2.356	2.356
40	0.4.	1	5.060	5.058
40	0.4	2	6.326	6.323
42	0.2	1	1.465	1.464
42	0.2	2	1.841	1.841
42	0.4	1	4.379	4.379
42	0.4	2	5.736	5.732
44	0.2	1	1.017	1.017
44	0.2	2	1.429	1.429
44	0.4	1	3.783	3.783
44	0.4	2	5.202	5.197

Table 1: Closed form solution for European put options compared to results calculated with finite difference method. Parameters used are $K = 40$ and $r = 0.06$, with 1000 spatial points and 300 time-steps in the FD scheme.

FD.2

Suppose you want to simulate your way to European option prices of similar accuracy as you get when using a finite difference method from the first part. What would run-times look like? It seems reasonable to define accuracy for a finite difference method via Østerby's equations (10.9), (10.12), and (10.30) and to use the standard errors for simulated values.

Using notation as in Østerby, the mentioned equations are

$$(10.9) \qquad e_1 = v_1 - v_2 - 2h^2 d - 6h^3 f - \dots$$

$$(10.12) \qquad e_2 = \frac{1}{3}(v_1 - v_2) - \frac{4}{3}h^3 f - 4h^4 g - \dots$$

$$(10.30) \qquad v_1 + (v_1 - v_2) + (v_1 - v_4) = u + hke + 2h^2 f + 2k^2 g + \dots$$

As the price function $f(x,t)$ is a function of two independent variables, t and x, the numerical approximation depends on both space discretization steps, h, and time discretization steps, k. This type of accuracy is analyzed in Østerby Section 10.5. As mentioned by Østerby it is important to check the global order of convergence before calculating an estimate of the error, as the formula for the error depends on whether the convergence is of the first or second order. Equation (10.27) and (10.29) in Østerby states the formulas to be used in order to get respectively the space order ratio and the time order ratio

$$(10.27) \qquad \frac{v_2 - v_3}{v_1 - v_2} = 2\frac{c + ke + 6hf + \dots}{c + ke + 3hf + \dots} \quad \text{space (h) order ratio}$$

$$(10.29) \qquad \frac{v_4 - v_5}{v_1 - v_4} = 2\frac{d + he + 6kg + \dots}{d + he + 3kg + \dots} \quad \text{space (k) order ratio}$$

where

v_1 is calculated using step sizes h and k

v_2 is calculated using step sizes 2h and k

v_3 is calculated using step sizes 4h and k

v_4 is calculated using step sizes h and 2k

v_5 is calculated using step sizes h and 4k

The order ratio, we briefly collectively refer to this as q, is examined. If there is first-order convergence then c (or d) is the dominating term, and we will see that

$$q \approx 2\frac{c}{c} = 2$$

If there is second-order convergence then hf (or kg) is the dominating term, and we will see that

$$q \approx 2\frac{6hf}{3hf} = 4$$

When testing the global order of convergence we add values very close to the strike, as we expect to see a spike when $x = K$. We expect this because, when considering the payoff function of a put, this

exact point is non-differentiable and therefore contradict the FD assumption that the true solution is differentiable. This creates what is referred to as 'ringing' when using the Crank-Nicolson approach. As mentioned by Jesper Andreasen (in his slideset 1) there are ways of getting rid of this ringing' however, we will deliberately choose the 'easy' approach by simply adding two points close to the non-differentiable point (98 and 102).

The results are stated in Table 2 and Table 3.

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0											
20	4.00	4.01	4.02	4.05	4.10	4.22	4.75	-0.17	0.10	-0.03	
40	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.02	4.13	0.27	
60	4.00	4.00	4.00	3.99	3.97	4.01	4.00	4.00	3.99	4.03	
80	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.05	
98	1092.03	1050.03	1004.65	954.72	898.35	833.56	762.07	693.83	606.03	388.88	
100	-60.03	-6.05	-2.52	-1.32	-0.76	-0.46	-0.28	-0.16	-0.05	0.26	
102	1009.38	966.52	920.17	869.00	811.12	745.08	673.44	604.04	506.02	301.26	
120	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	3.99	
140	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.02	3.99	
160	4.00	4.00	4.00	4.00	4.00	4.00	3.99	4.00	4.00	4.00	
180	4.00	4.00	4.00	4.00	4.00	4.02	4.00	4.00	4.00	4.03	
200											

Table 2: Examination of global convergence order for space discretization steps for European put options. Crank-Nicolson is used with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

We see from the Table 2 that the global order of convergence is $\mathcal{O}(dx^2)$. Further we note that there is a clear problem around $x = K$ as expected. Missing values are due to boundary conditions, as the values will be the same independent of the step size.

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0											
20	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
40	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.01	4.00	
60	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.01	4.00	3.99	
80	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
99	-5.85	-6.71	-4.76	1.30	9.78	16.85	23.48	58.08	368.41	-10.43	
100	3.71	3.35	3.03	2.74	2.48	2.26	2.07	1.92	1.83	1.82	
101	-6.27	-6.72	-4.14	2.43	10.80	17.49	24.31	69.10	-501.24	-10.20	
120	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
140	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.01	
160	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	3.99	
180	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
200											

Table 3: Examination of global convergence order for time discretization steps for European put options. Crank-Nicolson is used with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

Again we see that the global order of convergence is of the second order, $\mathcal{O}(dt^2)$, and further still seeing the spike around $x = K$ (due to the 'ringing').

By Østerby chapter 10 we have that the leading error term, in h and k respectively, is given by his equation (10.26) and (10.28). Further, the error for a second order term is given by dividing with 3 (as explained in Østerby eq. (10.12)). Thus the h-component and k-component of the collective error is given by

$$\begin{aligned} (10.26)+\mathcal{O}(dx^3) \quad &\Rightarrow \quad \frac{v_1 - v_2}{3} \quad \text{(h-component)} \\ (10.28)+\mathcal{O}(dt^3) \quad &\Rightarrow \quad \frac{v_1 - v_4}{3} \quad \text{(k-component)} \end{aligned}$$

Now, using Østerby's equation (10.30), we have

$$v_1 + \frac{v_1 - v_2}{3} + \frac{v_1 - v_4}{3}$$

We now run our finite difference algorithm (Crank Nicolson) and pick points where we check the error. The results are given in Table 4. The runtime for the entire calculation of Table 4 is 1.587 seconds.

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	94.18	94.74	95.31	95.89	96.46	97.04	97.63	98.22	98.81	99.40	100.00
20	74.18	74.74	75.31	75.89	76.46	77.04	77.63	78.22	78.81	79.40	80.00
40	54.32	54.83	55.37	55.92	56.48	57.05	57.63	58.22	58.81	59.40	60.00
60	36.11	36.32	36.55	36.81	37.10	37.44	37.83	38.29	38.82	39.40	40.00
80	22.01	21.79	21.54	21.28	21.00	20.69	20.36	20.03	19.73	19.60	20.00
98	13.39	12.93	12.42	11.85	11.22	10.50	9.66	8.67	7.42	5.71	2.00
100	12.65	12.17	11.65	11.05	10.41	9.65	8.81	7.76	6.50	4.67	0.00
102	11.94	11.45	10.91	10.31	9.64	8.89	8.01	6.97	5.67	3.88	0.00
120	7.02	6.50	5.95	5.35	4.69	3.98	3.19	2.32	1.37	0.41	0.00
140	3.79	3.38	2.93	2.47	1.99	1.50	1.02	0.57	0.20	0.02	0.00
160	1.95	1.67	1.38	1.09	0.81	0.54	0.30	0.12	0.02	0.00	0.00
180	0.82	0.69	0.55	0.41	0.28	0.17	0.08	0.02	0.00	0.00	0.00
200	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4: Estimates using Østerby eq. (10.30) and Crank-Nicolson finite difference method, with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

A table with only the errors (i.e. the error of the h-component + the error of the k-component) is stated below (Table 5).

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0
20	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0
40	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001	0.00000	0.00000	0.00000	0.00000	0.00000	0
60	0.00002	0.00002	0.00001	0.00001	0.00001	0.00000	0.00000	-0.00001	-0.00001	0.00000	0
80	0.00004	0.00004	0.00004	0.00004	0.00005	0.00005	0.00005	0.00006	0.00007	0.00007	0
98	0.00005	0.00004	0.00005	0.00004	0.00004	0.00008	-0.00001	0.00013	0.00017	-0.00089	0
100	0.00282	-0.00352	0.00380	-0.00523	0.00511	-0.00841	0.00668	-0.01565	0.00809	-0.04061	0
102	0.00005	0.00004	0.00005	0.00004	0.00004	0.00009	-0.00002	0.00012	0.00020	-0.00091	0
120	0.00003	0.00003	0.00003	0.00004	0.00004	0.00004	0.00004	0.00005	0.00006	0.00012	0
140	0.00002	0.00002	0.00002	0.00002	0.00002	0.00003	0.00003	0.00003	0.00002	-0.00002	0
160	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00000	-0.00001	-0.00001	0
180	0.00001	0.00001	0.00001	0.00001	0.00001	0.00000	0.00000	0.00000	-0.00001	0.00000	0
200	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0

Table 5: Error estimates using Østerby eq. (10.26)+(10.28) and Crank-Nicolson finite difference method, with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

We see that all the errors are very, very small, leading us to the conclusion that we could have used significantly smaller x steps and t steps, thus obtaining a much faster runtime, compared to the 1.587 seconds.

FD.3

Replicate the American option prices in column 4 using Crank-Nicolson. What is the global convergence order (as analyzed in Østerby, chapter 10)? Does that surprise you?

In order to price American options, rather than European, using finite difference methods, we need to be aware of early exercise. This attentiveness is implemented in the boundary conditions and when solving the tridiagonal system.

First, the boundary conditions are adapted to check for early exercise, thus

$$f_{0,n} = \max\left(K - x_{min}, e^{-r(T-n\Delta t)}(K - x_{min})\right)$$
$$f_{J,n} = 0$$

Second, the solver for the tridiagonal system is adapted to check for early exercise. In Section 8 of the Munk Note, this is done. We are now able to replicate the American option prices from column 4 (Table 1, Longstaff & Schwartz). The numbers to be replicated are obtained using an implicit finite difference scheme with 40.000 time steps per year and 1.000 steps for the stock price. The result is stated in Table 6. We note that the prices are almost identical, even though we have used significantly less steps in the replication in both the time- and space-grid, which however is expected when using the Crank-Nicolson approach.

Spot	σ	T	Finite difference American	Crank-Nicolson American
36	0.2	1	4.478	4.484
36	0.2	2	4.840	4.846
36	0.4	1	7.101	7.107
36	0.4	2	8.508	8.501
38	0.2	1	3.250	3.255
38	0.2.	2	3.745	3.749
38	0.4	1	6.148	6.152
38	0.4	2	7.670	7.671
40	0.2	1	2.314	2.318
40	0.2	2	2.885	2.888
40	0.4.	1	5.312	5.314
40	0.4	2	6.920	6.918
42	0.2	1	1.617	1.620
42	0.2	2	2.212	2.215
42	0.4	1	4.582	4.586
42	0.4	2	6.248	6.244
44	0.2	1	1.110	1.112
44	0.2	2	1.690	1.692
44	0.4	1	3.948	3.851
44	0.4	2	5.647	5.639

Table 6: Finite difference results from column 4 compared to the values replicated using Crank Nicolson with 1000 spatial (x) steps and 300 time (t) steps and further $K = 40$, $r = 0.06$, $x_{min} = 0$ and $xmax = 100$.

We now examine the global convergence order. This is done as in Østerby, chapter 10, more specifically Section 10.5, which analyzes the error when considering a function of two independent variables. We

reference FD.2 so as not to repeat ourselves (here the main Østerby results have been restated). We now wish to examine the global convergence order. We will use the values

$$\sigma = 0.4$$
$$T = 1$$
$$K = 100$$
$$x_{min} = 0$$
$$x_{max} = 200$$

Further we will use the following grid points

$$x \in \{0, 20, 40, 60, 80, 90, 98, 100, 102, 120, 140, 160, 180, 200\}$$
$$t \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

We start by examining the global convergence order of the space-component. The result, i.e. the values of the order ratio evaluated in mentioned (x,t) points, is stated in Table 7.

x \ t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0											
20											
40											
60											
80	4.05	4.07	4.11	4.20	4.40	4.87	6.26	16.71	-4.83	8.13	
98	4.03	4.03	4.03	4.03	4.03	4.02	4.02	4.02	4.02	4.00	
100	-1.52	-0.86	-0.51	-0.31	-0.20	-0.12	-0.07	-0.03	0.04	0.32	
102	4.03	4.03	4.03	4.02	4.02	4.02	4.02	4.02	4.01	4.00	
120	4.02	4.02	4.02	4.02	4.02	4.02	4.01	4.01	4.00	3.99	
140	4.02	4.02	4.02	4.02	4.02	4.01	4.01	4.00	4.29	4.01	
160	4.02	4.02	4.02	4.02	4.02	4.01	4.00	4.01	4.01	4.00	
180	4.02	4.02	4.02	4.02	4.02	4.04	4.01	4.01	4.00	4.00	
200											

Table 7: Examination of global convergence order for space discretization steps for American put options. Crank-Nicolson is used with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

We see that almost all the results are very close to 4, except in $x = K$, the reason for which is explained above. Thus the global order of convergence in dx is of the second order, i.e. $\mathcal{O}(dx^2)$. This is of no surprise to us as it is similar to the result in the case of the European option (as seen in Table 2).

What is further worth noting about Table 7 is that there are many missing values. We know from FD.2 that when no values appear it is due to boundary conditions, as values will be the same independent of the step size, meaning that v_1, v_2 and v_3 gives the same price. However, in the case of the American option we have the possibility of early exercise, thus if the spot is low enough, one will always choose early exercise. This makes sense intuitively due to the positive short rate (and zero dividend stock), as one would then, for low enough spot values, be better off by exercising the option and put the money in a bank account, where it can earn interest.

The result of examining the global convergence order in the time-step, is stated in Table 8.

x \ t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0											
20											
40											
60											
80	1.99	1.99	1.98	1.98	1.97	1.96	1.95	1.94	1.94	1.95	
98	1.96	1.96	1.96	1.96	1.95	1.95	1.94	1.95	1.96	1.56	
100	3.31	3.07	2.84	2.62	2.41	2.22	2.06	1.93	1.84	1.83	
102	1.96	1.96	1.96	1.95	1.95	1.94	1.94	1.95	1.95	1.39	
120	1.96	1.95	1.95	1.95	1.94	1.94	1.94	1.94	1.95	2.18	
140	1.95	1.95	1.95	1.95	1.95	1.95	1.96	1.99	2.04	-4.47	
160	1.95	1.96	1.96	1.96	1.97	1.97	1.98	1.96	1.26	4.21	
180	1.96	1.96	1.96	1.97	1.98	1.99	1.96	1.66	8.97	4.05	
200											

Table 8: Examination of global convergence order for time discretization steps for American put options. Crank-Nicolson is used with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0.06$, $\sigma = 0.4$.

We see that the results are approximately 2, meaning that the global order of convergence in dt is of the first order, i.e. $\mathcal{O}(dt)$. This is a surprise to us as the convergence in dt was of second order for European options (as seen in Table 3). Let's consider the difference between the two, namely the possibility of early exercise. We know this difference has an impact when determining the value of the put option. As

discussed just above the positive interest rate makes the choice to exercise early more likely. It therefore makes sense that the error related to the time-steps is more sensitive than the one from valuing European options.

We can examine the interpretation of the impact of the positive interest rate, by examining the global convergence order in the time steps in the case where $r = 0$. The result of such study is stated in Table 9. We see that there are less missing values and that the global convergence is now of second order, $\mathcal{O}(dt^2)$, when all we changed was to go from $r = 0.06$ to $r = 0$. Thus the change in results must be due to the decrease of gain from the early exercise. In contrast to what was discussed above, now the error might not depend as much on the exercise time, which makes the global convergence in the time dimension improve to be of the second order.

x \ t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0											
20	4.00	4.00	4.00	4.00	4.01	4.00	3.31	0.22	-0.43	3.00	
40	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.01	-0.77	
60	4.00	4.00	4.00	4.00	4.00	4.00	4.00	3.98	4.00	3.99	
80	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
98	3.95	3.84	3.76	3.83	4.18	4.72	4.68	2.72	1.79	10.21	
100	3.71	3.35	3.03	2.74	2.48	2.26	2.07	1.92	1.83	1.82	
102	3.91	3.80	3.75	3.89	4.31	4.84	4.53	2.33	2.06	10.72	
120	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
140	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.01	
160	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	3.99	
180	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	
200											

Table 9: Examination of global convergence order for time discretization steps for American put options. Crank-Nicolson is used with 1000 spatial (x) steps and 300 time (t) steps along with $K = 100$, $r = 0$, $\sigma = 0.4$.

FD.4

Explain how (stock price, strike)-homogeneity and time-homogeneity of call and put option prices in the Black-Scholes model can be used to price a lot of different options from a single finite difference grid.

By Eulers homogeneity theorem, as the Black-Scholes call option price, $Call()$, is homogeneous of the

first degree in spot and strike, we know that

$$Call(a \cdot S, a \cdot K) = a \cdot Call(S, K)$$

We use this to get option prices when moving along the space-grid (until now we considered the x-grid to be stock price movements, but as we shall use later, this could just as well be a strike grid). Consider a finite difference grid and fix the strike (and time-point), then by scaling up or down the spot price we can price options with different moneyness (relationship between strike and spot). The argument for the Black-Scholes put option price is the same. Now, as an example, consider the put prices we calculated using a finite difference scheme in FD.1, $Put(S = 40, K = 40)$. Then, by (stock price, strike)-homogeneity, we could easily price a put with strike 100 and spot 100,

$$Put(S = 100, K = 100) = Put(S = 2.5 \cdot 40, K = 2.5 \cdot 40) = 2.5 \cdot Put(S = 40, K = 40)$$

We now turn our attention towards the time-homogeneity. This property means that the Black-Scholes call (or put) price at any two time-points depends only on the difference between those time-points (remaining parameters fixed). Imagine that we buy an option at some time-point, t, which has time to maturity, T-t. This option is priced based on a set of parameters (strike, interest rate, volatility etc.). Then buying a call option at the next time-point, t+1, with time to maturity, $(T+1)-(t+1) = T-t$, will yield the same price. Consider now the finite difference time-grid (for $t = 0, ..., T$). Then the time-grid indicates time-to-maturity, such that prices calculated fx in a time-grid with $T = 2$ for $t = 1$ denotes the price of a call (or put) option with 1 year to maturity.

In conclusion, by having a single finite difference grid, we can price call and put options in the Black-Scholes model with varying moneyness (moving along the x-axis) and time to maturity (when moving along the t-axis).

FD.5

If we go beyond basic Black-Scholes, the previous homogeneity tricks do not work. But something else does (for European options)

- *Look up what Dupire's forward equation is (Equation 4)*
- *Adjust your finite difference code to solve Dupire's forward equation for European puts and calls.*
- *Using the CEV model as an example, demonstrate that your code can now calculate many option prices from a single finite difference grid.*

Using notation as in the linked article, for dynamics of the form

$$\frac{\partial S}{S} = \mu(t)dt + \sigma(S, t; S_0)dZ$$

Dupires forward equation holds and is given by

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 K^2}{2}\frac{\partial C}{\partial^2 K^2} + (r(T) - D(T))(C - K\frac{\partial C}{\partial K})$$

Assume no dividends, $D_T = 0$, and rewrite such that

$$0 = \frac{\partial C}{\partial T} + r(T)K\frac{\partial C}{\partial K} - \frac{1}{2}\sigma^2 K^2 \frac{\partial C}{\partial^2 K^2} - r_T C$$

In FD.1 Eq (3) we used Finite Difference methods to solve PDEs of the form

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

The difference is now that we will perform forward successive iterations instead of backwards. Further we will let the space-dimension represent the strike instead of the stock price, thus

$$(x,t) \in K \times [0,T)$$

We therefore also need to adjust the boundary conditions. We know that the payoff of a European Call option is $\max(S_T - K, 0)$, thus (for fixed spot) the lower the strike, the higher the payoff. Using reverse arguments as for the put option in FD.1, the boundary conditions are given by

$$f_{0,n} = e^{-r(T-t)}(S_0 - x_{min})$$
$$f_{J,n} = 0$$

We remember the system of equations, (ref - Eq (4), but we repeat for the reader)

$$A_{j,n}f_{j-1,n} + B_{j,n}f_{j,n} + C_{j,n}f_{j+1,n} = -A_{j,n+1}f_{j-1,n+1} + B^*_{j,n+1}f_{j,n+1} - C_{j,n+1}f_{j+1,n+1}$$

In matrix notation

$$Af_n = Bf_{n+1}$$

Before, when moving backwards, we started at time-of-maturity and the right hand side, f_{n+1}, was the

known values used to calculate the unknown values on the left hand side, $f_{j-1,n}, f_{j,n}, f_{j+1,n}$. Now, as we move forward in time, we start at time $t = 0$ by calculating the immediate payoff and thus the LHS of equation (4) is known and the RHS involves the unknowns, $f_{j-1,n+1}, f_{j,n+1}, f_{j+1,n+1}$. Note that the time-grid now indicates time-to-maturity, such that prices calculated fx for $t = 1$ denotes the price of the call option with 1 year to maturity. By discounting the prices, as the Dupire's forward equation returns undiscounted prices, we are able to create a grid of time-0 prices of call options (and put options by the put-call parity), for different time-to-maturities and strike. Thus by the homogeneity arguments from FD.4 we are able to price a lot of different options from a single finite difference grid.

We now turn our attention to the CEV model and wish to fit this model into our Finite Difference framework. The spot dynamics in the CEV model is given by

$$\frac{\partial S}{S} = \mu dt + \underbrace{\sigma_{cev} S^{\gamma-1}}_{=\sigma^c(S,t)} dW_t$$

Now, making a few adjustments to combine the CEV dynamics, Dupires forward equation and Equation (3), we arrive at the PDE

$$0 = \frac{\partial C}{\partial T} + \mu(T)\frac{\partial C}{\partial K} + \frac{1}{2}\sigma(K,T)\frac{\partial^2 C}{\partial K^2} - r(T)C$$

where

$$\begin{aligned} \mu(T) &= r(T)K \\ \sigma(K,T) &= -\sigma^c(S,t)^2 K^2 \end{aligned}$$

To get an idea of whether our code works or not we wish to, at last, compare our results to the known Black-Scholes pricing function. To do so, we note the following relation between the diffusion terms in the two models

$$\sigma_{BS} \approx \sigma_{CEV} S^{\gamma-1}$$

Thus with $\sigma_{BS} = 0.2$, $\gamma = 0.5$ and $S_0 = 100$, we have $\sigma_{CEV} = 2$.

Now, using the finite difference scheme, prices for call and put options for various strike levels are displayed in the top two plots in Figure 16, plotted against the prices from the Black-Scholes model for comparison. Similarly prices for call and put options for various time-to-maturity are displayed in the bottom two plots of Figure 16. We see that the tendencies in both figures are very much alike, when

comparing CEV prices and Black-Scholes prices and we conclude that our code works.

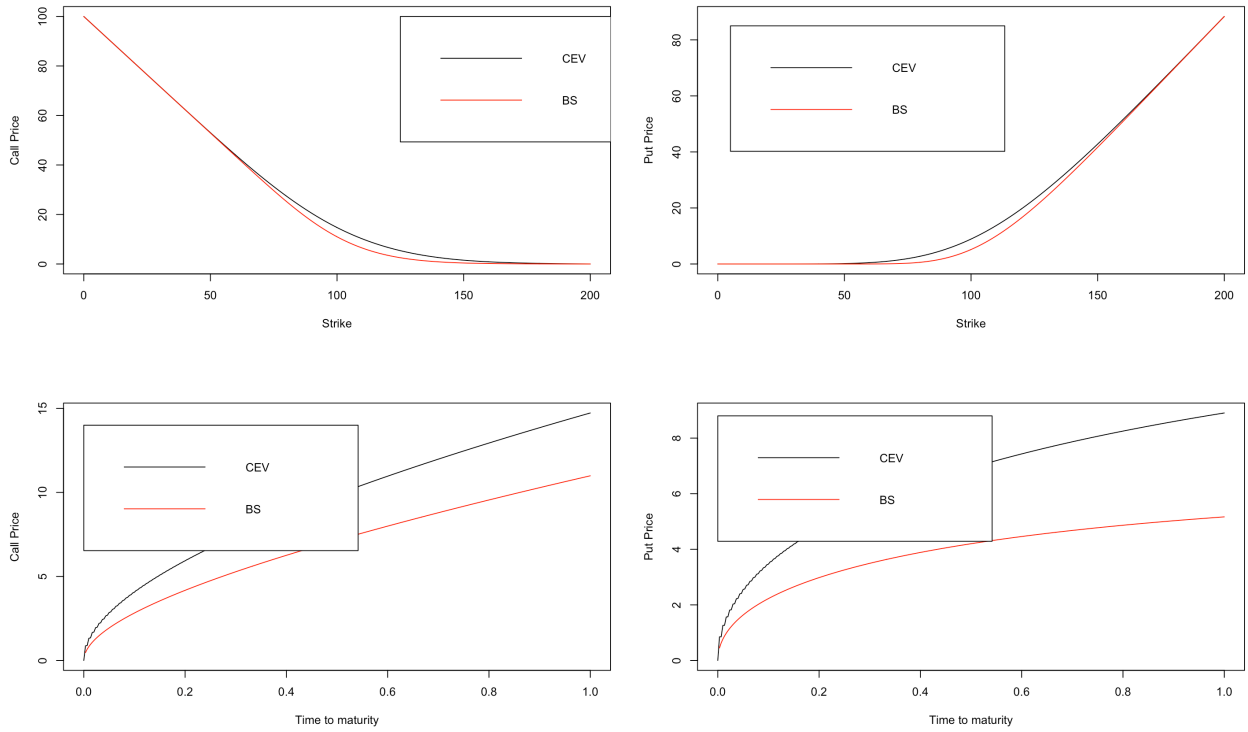

Figure 16: Top panel: (Strike,Price)-plots. Bottom panel: (Time-to-maturity,Price)-plots. Red: Black-Scholes prices. Black: CEV prices calculated using Finite Difference method and Dupire's forward equation. Lines are European Call (left) and Put (right) option prices.

Appendix

All code is on this link: https://github.com/cnp777/CompFinHI2.