Master's Thesis
A comparison of the Fourier-Gauss-Laguerre and Fourier cosine series method in option pricing

Benjamin Kraska

Abstract

We describe the Fourier-Gauss-Laguerre and Fourier cosine series method and test them extensively in four models: Black-Scholes, Black-Scholes with discrete dividends, Heston and Bates. While both methods mostly achieve good accuracy and high computational speed, problems may arise with respect to the optimal choice of the method-specific parameters and the extension of the methods to other models and financial products.

Contents

1 Introduction

This thesis should be seen as an extension to the Benchop project[1], which specifies a number of standard test cases as a benchmark for new option valuation methods. The Benchop project also provides testing results of various already existent option pricing methods, specifically Monte Carlo, Fourier, finite difference and radial basis function methods.

Two of the most interesting methods of the tested ones are the Fourier-Gauss-Laguerre (FGL) and the Fourier cosine series (COS) method. This is because in the Benchop test cases, FGL is one of the or the most accurate method while being fairly fast, and COS is almost always the fastest, while being reasonably accurate.[2] Since the test cases covered by the Benchop project are quite limited, our goal is to test FGL and COS more comprehensively by

- testing many more combinations of parameters
- testing in a model not covered by the Benchop project, namely the Bates model.

We are going to focus on the pricing of European call options. The fact that we cover many more test cases permits us to analyze the error in dependence of

- each single model parameter, such as time to maturity,
- the used model, and
- the used method.

To make this thesis as self-contained as possible, a thorough theoretical description of both methods precedes the specification, description and analysis of the test results. Specifically, the thesis is structured as follows: After giving an overview of option valuation methods in 1.1 and stating some preliminary definitions and conventions in Section 2, we proceed to describe the FGL and COS methods in sections 3 and 4, respectively. Some of our test cases use a Monte Carlo method to produce reference prices, which we elaborate on in Section 5. Lastly, we specify our test methodology, test cases and describe the results in Section 6.

A large part of this thesis relies on numerical experiments, the code for which can be found at github.com/Arebjel/ThesisCode.

1.1 Overview of option valuation methods

Option valuation methods can be roughly categorized into three different types:[3]

- Deterministic numerical solution of the pricing PDEs or - for models with jumps - the corresponding PIDEs.
- Quasi Monte Carlo or Monte Carlo methods combined with various variance reduction techniques.
- Methods based on Fourier transform techniques.

An advantage of the PDE/PIDE-based methods is their versatility, allowing them to be employed to price both plain vanilla and exotic products. Their main drawback is related to the fact that they involve approximating prices on a grid. During the calibration of model parameters, we need to calculate many option prices that are not used, leading to an additional amount of work.

Monte-Carlo methods do not have that disadvantage, since we only need to calculate the prices we use. Furthermore, they are easy to implement.[4] Due to their slow convergence rate ($O\left(1/\sqrt{N}\right)$ for standard Monte Carlo)[5], however, achieving high accuracy is computationally expensive. One may argue that calculating option prices with a higher accuracy than the bid-ask spread is unnecessary. However, this does not hold for the calibration of the model parameters, where even small inaccuracies might lead to very different model parameters in the optimization step. Therefore, the aforementioned slow convergence rate also makes the calibration step slower.

Fourier transform based methods are fast and fairly accurate. A downside is that they require the characteristic function of the log stock price to be known in closed form. This, however, is the case for all the models considered in this thesis and several others, specifically Black-Scholes (BS)[6], Merton[7], Heston[8], Bates[9], Normal Inverse Gaussian[10], Variance Gamma[11],

[1] Benchop = The benchmarking project in option pricing - [von Sydow et al., 2015]
[2] See [von Sydow et al., 2015] p. 9f, tables 2 - 5. The only exception of COS being the fastest method is problem 1, American call.
[3] Section based on [Wiktorsson, 2015], p. 1f.
[4] [von Sydow et al., 2015], p. 10.
[5] [von Sydow et al., 2015], p. 10.
[6] [Black and Scholes, 1973]
[7] [Merton, 1976]
[8] [Heston, 1993]
[9] [Bates, 1996]
[10] (see e.g. [Barndorff-Nielsen, 1997], and the references wherein)
[11] [Madan and Seneta, 1990]

CGMY[12], and Finite Moment log-stable[13].

We conclude this section by pointing out some advantages and drawbacks of some specific Fourier methods: The FGL method can easily lead to numerical overflow, as the involved numbers are too large for the data type *double*. This problem is particularly prevalent in the Bates model, because the involved moment-generating function requires nested evaluation of functions such as $\exp, \coth$ etc. To solve this, one needs either to restrict the optimization domain of the parameter α and the number of weights in the Gauss-Laguerre quadrature, both of which may reduce accuracy, or one must make use of classes specialized for dealing with extremely large numbers.[14] This, however, considerably slows down the method.

As for the COS method, we found that the choice of the parameters $[a,b]$ and N suggested by Fang, Oosterlee can lead to highly inaccurate prices in some cases.

Despite the fact that the FFT is not specifically considered in this thesis, we mention one of its drawbacks for the sake of completeness: the FFT allows the computation of prices on a regularly spaced grid of log strike levels. Market-traded options, however, are usually regularly spaced in the strike level, not the log-strike level. This makes an interpolation procedure necessary, and it makes the accuracy of the FFT dependent on the grid spacing, the interval the integrand is to be evaluated at and the properties of the characteristic function of the log stock price. Apart from that, for each time of maturity one FFT needs to be done, leading to additional work if we are to price options with several dates of maturity, but each only with a few strike levels.

[12] [Carr et al., 2002]

[13] [Carr and Wu, 2003]

[14] For example, the high-precision float (HPF) package (http://de.mathworks.com/matlabcentral/fileexchange/36534-hpf-a-big-decimal-class) on Matlab File Exchange allows for much higher numbers than the usual data type *double*.

2 Preliminaries

In this section, we introduce some important definitions relevant to our analyis. Also, we specify the stock price models used in our test cases.

2.1 Definitions

Definition 2.1. (Fourier transform, inverse Fourier transform)

For a continuous function $f : \mathbb{R} \to \mathbb{R}$, let its *Fourier transform* (FT) be denoted by

$$\mathscr{F}[f](z) = \int_{\mathbb{R}} e^{zk} f(k)\, dk,$$

wherever this integral exists. Let the corresponding *inverse Fourier transform* (iFT) be denoted by

$$\overline{\mathscr{F}}[g](k) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-(\alpha+i\omega)k} g(\alpha+i\omega)\, d\omega,$$

for any $\alpha \in \left\{ \left| \int_{\mathbb{R}} e^{-(\alpha+i\omega)k} g(\alpha+i\omega)\, d\omega \right| < \infty \right\} \cap \{|g(\alpha+i\omega)| < \infty\}$.

Remark 2.2. Note that the regions where $\mathscr{F}[f]$ exists, are always vertical strips in the complex plane. This is because for any complex number $z = \alpha + i\omega$ it holds

$$\int_{\mathbb{R}} \left| e^{\alpha k} e^{i\omega k} f(k) \right| dk = \int_{\mathbb{R}} \left| e^{\alpha k} f(k) \right| dk,$$

and thus the finiteness of $\mathscr{F}[f](\alpha+i\omega)$ does not depend on ω.

Definition 2.3. (conditional moment-generating function)

For a stochastic process X_t on $\{\Omega, \mathscr{F}_t, \mathbb{Q}\}_{t \geq 0}$ and $0 \leq t < T$, let the *conditional moment-generating function of X_T given $\mathscr{F}_t$* be denoted by

$$M^{\mathbb{Q}}_{X_T|\mathscr{F}_t}(z) = \mathbb{E}^{\mathbb{Q}}[\exp(zX_T) \mid \mathscr{F}_t],$$

where $z \in \mathbb{C}$.

2.2 Stock price models

In this subsection, we specify the stock price models used in our tests. They largely overlap with the models used in the Benchop project, allowing us to provide a more comprehensive test coverage (see Section 6) than in that project. The exception thereof is the Bates model, which serves as an interesting addition to the models used in the Benchop project. This is especially true because it incorporates both stock price paths with stochastic jumps as in the Merton model as well as stochastic volatility as in the Heston model.

2.2.1 Black-Scholes model

The Black-Scholes model can be defined by the SDE

$$dS_t = rS_t dt + \sigma S_t dW_t$$

with initial condition $S_0 = s_0$. Here, r is the risk-free interest rate and σ the volatility of the underlying.

2.2.2 Black-Scholes model with discrete dividends

$$dS_t = rS_t dt + \sigma S_t dW_t - \delta(t-\tau) D \cdot S_t$$

At time $\tau = \alpha \cdot T$, a dividend of $D \cdot S_\tau$ is paid. As with the Black-Scholes model, the initial condition is $S_0 = s_0$.

2.2.3 Heston model

The risk-neutral dynamics in the Heston model are given by

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^{(1)}, \tag{2.1}$$

$$dV_t = \kappa(\theta - V_t)\, dt + \sigma_V \sqrt{V_t} dW_t^{(2)}, \tag{2.2}$$

where $W^{(1)}, W^{(2)}$ are Brownian motions with correlation ρ. Furthermore, κ is the mean reversion rate, θ the long-term volatility, σ_V the volatility of the volatility. The initial conditions for the underlying price and volatility are $S_0 = s_0, V_0 = v_0$.

2.2.4 Bates model

The Bates model is specified via the SDEs

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^{(1)} + S_t dJ_t \tag{2.3}$$

$$dV_t = \kappa(\theta - V_t)\, dt + \sigma_V \sqrt{V_t} dW_t^{(2)}. \tag{2.4}$$

Like in the Heston model, $W^{(1)}, W^{(2)}$ are Brownian motions with correlation ρ and the initial conditions for S_0, V_0 are the same. Also, the variables κ, θ and σ_V have the same meaning as in the Heston model. J is a compound Poisson process with intensity λ, with jumps ΔJ such that $\log(1+\Delta J) \sim \mathcal{N}\left(\gamma, \delta^2\right)$ and with drift $-\lambda\left(\exp\left(\delta^2/2+\gamma\right)-1\right)$.[15] The Bates model is essentially a combination of the Heston and Merton model, which incorporate stochastic volatility and jumps into the stock price, respectively.

[15] [Wiktorsson, 2015], p.7.

3 Fourier-Gauss-Laguerre method

This section is divided into two parts:[16] Subsection 3.1 covers the theoretical description of the FGL method. Subsection 3.2 deals with numerical aspects that are relevant to ensure high accuracy in a reasonable amount of computation time.

3.1 Method description

By the fundamental theorem of asset pricing, the price of any European-style contract with maturity T at time t is

$$p = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[\Phi\left(S_T, K\right) \mid \mathscr{F}_t\right], \tag{3.1}$$

where Φ is the payoff function and $\tau := T - t$.

We give a rough outline of the FGL method as follows: Firstly, the payoff function is represented as the inverse Fourier transform of its Fourier transform with respect to the log-strike price:

$$\Phi\left(S_T, K\right) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-(\alpha+i\omega)\ln K} \cdot \mathscr{F}\left[\Phi\right]\left(\ln S_T, \alpha + i\omega\right) d\omega.$$

Discounted conditional expectation yields the option price, resulting in a double integral:

$$p = \frac{e^{-r\tau}}{2\pi}\int_{\Omega}\int_{\mathbb{R}} e^{-(\alpha+i\omega)\ln K} \cdot \mathscr{F}\left[\Phi\right]\left(\ln S_T, \alpha + i\omega\right) d\omega d\mathbb{Q}\left(\cdot \mid \mathscr{F}_t\right).$$

The order of integration is then reversed and explicit formulae for the FT of the payoff function, $\mathscr{F}\left[\Phi\right]$, and the conditional moment-generating function of the stock price are inserted. This results in an expression of the type

$$p = \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}} g\left(\alpha + i\omega\right) d\omega, \tag{3.2}$$

where $\alpha \in \mathbb{R}$ is a parameter governing the shape of g. Expression (3.2) can be approximated via Gauss-Laguerre quadrature. To ensure optimal accuracy of the latter, α is optimized before applying Gauss-Laguerre.

3.1.1 Fourier transform of European call options

As mentioned, the FGL method involves the FT of the payoff function. Therefore, we will provide an explicit formula for it in the case of a European call option. In general, European style contracts have payoff functions of the shape

$$\Phi\left(e^s, e^k, p, c\right) = \max\left(c\left(e^s - e^k\right), 0\right)^p, \tag{3.3}$$

where s is the log stock price, k the log strike, c is one for call options and minus one for put type options, and p is one for standard European options, zero for binary options and $p > 1$ for power options. The FT of (3.3) with respect to k reads

$$h\left(s, z, p, c\right) = \int_{\mathbb{R}} e^{zk} \max\left(c\left(e^s - e^k\right), 0\right)^p dk. \tag{3.4}$$

For call type options, it is $c = 1$ and the Fourier transform exists for $\Re z > 0$. Expression (3.4) then becomes

$$\begin{aligned}
h\left(s, z, p, 1\right) &= \int_{\mathbb{R}} e^{zk} \max\left(e^s - e^k, 0\right)^p dk \\
&= \int_{-\infty}^{s} e^{zk}\left(e^s - e^k\right)^p dk && (3.5) \\
&= \int_{-\infty}^{0} e^{s(z+p)} e^{zk}\left(1 - e^k\right)^p dk, \ \left[x = e^k\right] && (3.6) \\
&= e^{s(z+p)}\int_0^1 x^{z-1}\left(1-x\right)^{p+1-1} dx && (3.7) \\
&= e^{s(z+p)}\frac{\Gamma\left(z\right)\Gamma\left(p+1\right)}{\Gamma\left(z+p+1\right)}, && (3.8)
\end{aligned}$$

which, in the case of a European call option ($p = 1$), equals

$$h\left(s, z, 1, 1\right) = \frac{e^{s(z+1)}}{z\left(z+1\right)}. \tag{3.9}$$

[16] Section based on [Wiktorsson, 2015], p. 2 - 8.

3.1.2 Representation of the call option price using the Fourier inversion theorem

By applying the Fourier inversion theorem, the payoff function can be represented as

$$\Phi(S_T,K)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{-(\alpha+i\omega)\ln K}h(\ln S_T,\alpha+i\omega,1,1)\,d\omega.$$

Applying conditional expectation and discounting yields the option price at time t:

$$\begin{aligned}
p &= \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}}e^{-(\alpha+i\omega)\ln K}\mathbb{E}^{\mathbb{Q}}\left[h(\ln S_T,\alpha+i\omega,1,1)\mid\mathscr{F}_t\right]d\omega\\
&= \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}}\frac{e^{-(\alpha+i\omega)\ln K}}{(\alpha+i\omega)(\alpha+i\omega+1)}\mathbb{E}^{\mathbb{Q}}\left[e^{\ln S_T(\alpha+i\omega+1)}\mid\mathscr{F}_t\right]d\omega\\
&= \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}}\frac{e^{-(\alpha+i\omega)\ln K}}{(\alpha+i\omega)(\alpha+i\omega+1)}M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(\alpha+i\omega+1)\,d\omega,
\end{aligned}\tag{3.10}$$

where α must be positive and such that the conditional moment-generating function $M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(\alpha+i\omega+1)$ exists. We will cover the range of admissible α's and the optimal choice of α in 3.2.2.

3.1.3 Conditional moment-generating functions

Formula (3.10) requires the evaluation of the conditional MGF of the log stock price.These are given by:

- Black-Scholes model:

$$M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(z)=\exp\left(z(r\tau+\ln S_t)+\tau\frac{\sigma^2}{2}\left(z^2-z\right)\right).$$

- Black-Scholes model with discrete dividends:

$$M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(z)=\exp\left(z(r\tau+\ln S_t)+\tau\frac{\sigma^2}{2}\left(z^2-z\right)+z\log(1-D)\,I(t<\tau<T)\right).$$

- Heston model:

$$\begin{aligned}
d &= \sqrt{(\rho\sigma_V z-\kappa)^2+\sigma_V^2(z-z^2)},\\
C &= \frac{\kappa\theta}{\sigma_V^2}\left((\kappa-\rho\sigma_V z)\tau+2\log\left(\frac{1}{\kappa-\rho\sigma_V z+d\coth(d\tau/2)}\frac{d}{\sinh(d\tau/2)}\right)\right),\\
D &= \frac{z^2-z}{\kappa-\rho\sigma_V z+d\coth(d\tau/2)},\\
M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(z) &= \exp\left(z(r\tau+\ln S_t)+C+DV_t\right).
\end{aligned}$$

- Bates model:

$$\begin{aligned}
d &= \sqrt{(\rho\sigma_V z-\kappa)^2+\sigma_V^2(z-z^2)},\\
C &= \frac{\kappa\theta}{\sigma_V^2}\left((\kappa-\rho\sigma_V z)\tau+2\log\left(\frac{1}{\kappa-\rho\sigma_V z+d\coth(d\tau/2)}\frac{d}{\sinh(d\tau/2)}\right)\right),\\
D &= \frac{z^2-z}{\kappa-\rho\sigma_V z+d\coth(d\tau/2)},\\
M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(z) &= \exp\left(z(r\tau+\ln S_t)+\lambda\tau\left(e^{z^2\delta^2/2+z\gamma}-1-z\left(e^{\delta^2/2+\gamma}-1\right)\right)+C+DV_t\right).
\end{aligned}$$

3.2 Implementation and numerical aspects

3.2.1 Smoothing of the integrand in the inverse Fourier transform

Let the integrand in (3.10) be denoted by $g(\alpha+i\omega)$. As the option price is real, it holds

$$\begin{aligned} p &= \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}} g(\alpha+i\omega)\,d\omega \\ &= \frac{e^{-r\tau}}{2\pi}\int_{\mathbb{R}} \mathfrak{Re}\left(g(\alpha+i\omega)\right)d\omega. \end{aligned} \tag{3.11}$$

Apart from the condition that the integrand and integral have to exist, the choice of α is theoretically irrelevant. For a good numerical approximation of this integral, however, it is important to choose α such that the integrand does not exhibit overpronounced peaks, strong oscillations or extreme values that cause numerical overflow. One possibe choice of α would minimize the maximal absolute change of the integrand, i.e.

$$\alpha_{opt} = \operatorname{argmin}_{\alpha\in A^+_{S_T}} \max_{\omega\in D}\left|\frac{\partial}{\partial\omega}\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right| \tag{3.12}$$

or the mean absolute change, that is

$$\alpha_{opt} = \operatorname{argmin}_{\alpha\in A^+_{S_T}} \int_{\omega\in D}\left|\frac{\partial}{\partial\omega}\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right|d\omega, \tag{3.13}$$

where $D=\mathbb{R}$ and

$$A^+_{S_T} := \left\{\alpha>0 \mid \left|M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(\alpha+i\omega+1)\right|<\infty\right\}.$$

is the set of all α such that the conditional MGF in (3.10) exists. Taking into account that the integral $\int_{\mathbb{R}} g(\alpha+i\omega)\,d\omega$ will be truncated to some interval $[a,b]$ during the numerical approximation, $D=[a,b]$ would be a valid choice, too.

Alternatively, considering that the error term of Gauss-Laguerre quadrature - which we will later employ - is

$$E = \frac{(n!)^2}{(2n)!}\frac{\partial^{2n}}{\partial\omega^{2n}}\left(\omega\mapsto\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right)|_{\omega=\xi}$$

for some $\xi\in(0,\infty)$,[17] it might make sense to replace $\left|\frac{\partial}{\partial\omega}\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right|$ by E in (3.12) and (3.13). Here, n is the number of weights used in the Gauss-Laguerre quadrature.

The problem of these approaches is that they require the evaluation of $\left|\frac{\partial}{\partial\omega}\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right|$ or E at many points ω for each combination of parameters.[18] A computationally less expensive approach is based on the inequality

$$\left|\mathfrak{Re}\left(g(\alpha+i\omega)\right)\right| \le \left|g(\alpha+i\omega)\right| \le g(\alpha),$$

and then choosing

$$\alpha_{opt} = \operatorname{argmin}_{\alpha\in A^+_{S_T}} g(\alpha).$$

As $\alpha\mapsto g(\alpha)$ is a convex function on $A^+_{S_T}$, there exists a unique minimum, which can numerically be approximated by, for example, golden section search.

3.2.2 Range of admissible α's

Note that

$$\begin{aligned} \left|M^{\mathbb{Q}}_{\ln S_T|\mathscr{F}_t}(\alpha+i\omega+1)\right| &\le \mathbb{E}^{\mathbb{Q}}\left[\left|e^{\ln S_T\cdot(\alpha+i\omega+1)}\right| \mid \mathscr{F}_t\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\left|e^{\ln S_T\cdot(\alpha+1)}\right|\cdot\left|e^{\ln S_T\cdot i\omega}\right| \mid \mathscr{F}_t\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[S_T^{1+\alpha} \mid \mathscr{F}_t\right]. \end{aligned}$$

[17] [Abramowitz and Stegun, 1972], p. 890.
[18] [Kahl and Lord, 2006], p.12.

Therefore, a sufficient condition for the existence of the conditional MGF is that $S_T^{1+\alpha}$ be integrable, i.e.

$$A_{S_T}^+ \supseteq \left\{ \alpha > 0 \mid \mathbb{E}^{\mathbb{Q}} \left[S_T^{1+\alpha} \right] < \infty \right\}.$$

Since every $(1+\hat{\alpha})$-integrable variable is $(1+\alpha)$-integrable for $\alpha < \hat{\alpha}$, the set $A_{S_T}^+$ is an entire interval $[0, \alpha_{\max}]$ or $[0, \alpha_{\max})$, where $\alpha_{\max} \in [0, \infty]$. In the Black-Scholes models - with and without discrete dividends -, this set is simply the interval $[0, \infty)$. In the Bates and Heston model, we use the following approximation: A conservative approximation of $A_{S_T}^+$ is the interval between the largest negative zero and the smallest positive zero of the polynomial

$$\begin{aligned} q(x) &= \left(-\rho^3 + \rho\right) \tau^3 \sigma_V^3 x^3 + \left(-\rho \tau^3 \sigma_V^3 + \left(\left(3\kappa\rho^2 - \kappa\right) \tau^3 + \left(6\rho^2 - 6\right) \tau^2\right) \sigma_V^2\right) x^2 \\ &\quad + \left(\left(\kappa\tau^3 + 6\tau^2\right) \sigma_V^2 + \left(-3\kappa^2 \rho \tau^3 - 12\kappa\rho\tau^2 - 24\rho\tau\right) \sigma_V\right) x \\ &\quad + \kappa^3 \tau^3 + 6\kappa^2 \tau^2 + 24\kappa\tau + 48. \end{aligned}$$

Let the said interval be denoted by $[s, r]$. To prevent numerical overflow and following the implementation of FGL in the Benchop project, we further shrink this interval:

$$\check{A}_{S_T}^+ := [0, \min(0.999 \cdot (r-1), 50000)].$$

This last step applies to all models, so that the interval in the Black-Scholes models is $[0, 50000]$.

3.2.3 Gauss-Laguerre method

The Gauss-Laguerre method of order n approximates an exponentially weighted integral via[19]

$$\int_0^\infty e^{-x} f(x)\, dx \approx \sum_{j=1}^n w_j^{(n)} f\left(x_j^{(n)}\right),$$

which can be extended to any integral on $[0, \infty]$ via

$$\int_0^\infty g(x)\, dx = \int_0^\infty e^{-x} \underbrace{e^x g(x)}_{=:f(x)} \approx \sum_{j=1}^n w_j^{(n)} \underbrace{g\left(x_j^{(n)}\right) e^{x_j^{(n)}}}_{f\left(x_j^{(n)}\right)}. \tag{3.14}$$

The abscissa $x_j^{(n)}$ is the j-th zero of the Laguerre polynomial $L_n(x)$, and the j-th weight is

$$w_j^{(n)} = \frac{x_j^{(n)}}{(n+1)^2 \left[L_{n+1}\left(x_j^{(n)}\right)\right]^2}.$$

In this thesis project, we apply this technique to the integral (3.11). Note that the real part of the integrand is symmetric, i.e.

$$\int_{\mathbb{R}} \mathfrak{Re}\left(g(\alpha + i\omega)\right) d\omega = 2 \int_0^\infty \mathfrak{Re}\left(g(\alpha + i\omega)\right) d\omega,$$

making it feasible to use approximation formula (3.14).

[19] [Abramowitz and Stegun, 1972], p. 890.

4 Fourier cosine series method

This section is organized as follows: In Subsection 4.1, we derive the COS formula. The involved Fourier coefficients are presented in 4.2 in the case of European call options. An exception to the general COS formula is the Black-Scholes model with discrete dividends, for which we state a tailored COS formula in Subsection 4.3. We conclude by treating the characteristic function in Subsection 4.4.

4.1 Method description

Our goal is to approximate the price of a European call option by means of the COS method.[20] We denote the price by

$$p(x,t) := e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[p(X_T,T) \mid X_t = x\right], \tag{4.1}$$

where $X_T := \log(S_T/K)$ is the state process, and $p(X_T,T) := \left(K\left(e^{X_T}-1\right)\right)^+$ the payoff function of a European call option. In models with stochastic volatility V_t such as Heston and Bates, the conditional expectation in (4.1) must be replaced by $\mathbb{E}^{\mathbb{Q}}\left[p(X_T,T) \mid X_t = x, V_t = v_t\right]$. The subsequent reasoning, however, remains valid in that case, hence assuming that the option price is of the shape (4.1) leads to no loss of generality.

The expectation in (4.1) can also be expressed as an expectation with respect to the factorized conditional risk-neutral measure of X_T given $X_t = x$, which is $A \mapsto \mathbb{Q}(X_T \in A \mid X_t = x)$ for $A \in \mathscr{F}_T$. Assuming that this measure has a continuous density $q(\cdot \mid x)$ under the risk-neutral measure $\mathbb{Q}$, the price equals

$$\begin{aligned} p(x,t) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[p(X_T,T) \mid X_t = x\right] \\ &= e^{-r\tau}\int_{\mathbb{R}} p(y,T)\, d\mathbb{Q}(X_T = y \mid X_t = x) \\ &= e^{-r\tau}\int_{\mathbb{R}} p(y,T)\, q(y \mid x)\, dy. \end{aligned} \tag{4.2}$$

The derivation of the COS formula can be separated into three steps:

Step 1: The Fourier cosine series expansion allows any function $g : [a,b] \to \mathbb{R}$ to be represented as

$$g(x) = \sum_{k=0}^{\infty}{}'a_k \cos\left(k\cdot\pi\frac{x-a}{b-a}\right), \qquad \text{for } x \in [a,b], \tag{4.3}$$

$$a_k = \frac{2}{b-a}\int_a^b g(y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \tag{4.4}$$

where the $'$ means that the first coefficient should be multiplied by $\frac{1}{2}$. Formula (4.3) requires g to have a bounded interval $[a,b]$ as its domain or support. As we aim to apply it to the payoff function $p(y,T)$ and to the transitional density $q(y \mid x)$, we therefore need to truncate the integral (4.2):

$$p(x,t) = e^{-r\tau}\int_{\mathbb{R}} p(y,T)\, q(y \mid x)\, dy \approx e^{-r\tau}\int_a^b p(y,T)\, q(y \mid x)\, dy =: p_1(x,t;[a,b]) \tag{4.5}$$

We will deal with the choice of a and b in Subsection 6.1.3.

Step 2: Next, consider the Fourier series expansion of $p(y,T)$ and $q(y \mid x)$:

$$\begin{aligned} p(y,T) &= \sum_{k=0}^{\infty}{}'\mathscr{P}_k(T)\cos\left(k\pi\frac{y-a}{b-a}\right), \qquad \text{for } y \in [a,b], \text{ where} \\ \mathscr{P}_k(T) &= \frac{2}{b-a}\int_a^b p(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \text{ and} \\ q(y \mid x) &= \sum_{k=0}^{\infty}{}'\mathscr{Q}_k(x)\cos\left(k\pi\frac{y-a}{b-a}\right), \qquad \text{for } y \in [a,b], \text{ where} \\ \mathscr{Q}_k(x) &= \frac{2}{b-a}\int_a^b q(y \mid x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy. \end{aligned}$$

[20] Section based on [Ruijter and Oosterlee, 2015], p. 1 - 3.

Replacing the density $q(y \mid x)$ by its Fourier cosine series in definition (4.5), interchanging summation and integration, using the definition of $\mathscr{P}_k(T)$ and truncating the series gives

$$\begin{aligned} p_1(x,t;[a,b]) &= e^{-r\tau}\int_a^b p(y,T)\sum_{k=0}^{\infty}{}'\mathscr{Q}_k(x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= e^{-r\tau}\sum_{k=0}^{\infty}{}'\mathscr{Q}_k(x)\int_a^b p(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= e^{-r\tau}\frac{b-a}{2}\sum_{k=0}^{\infty}{}'\mathscr{Q}_k(x)\,\mathscr{P}_k(T) \\ &\approx e^{-r\tau}\frac{b-a}{2}\sum_{k=0}^{N-1}{}'\mathscr{Q}_k(x)\,\mathscr{P}_k(T) \\ &=: p_2(x,t;[a,b],N). \end{aligned} \tag{4.6}$$

Step 3: Lastly, we can approximate the coefficients $\mathscr{Q}_k(x)$ by means of the conditional characteristic function $\phi(\omega \mid x) := \int_{\mathbb{R}} e^{i\omega y} q(y \mid x)\,dy$ as follows:

$$\mathfrak{Re}\left(\phi\left(\frac{k\pi}{b-a} \mid x\right)e^{-ik\pi\frac{a}{b-a}}\right) = \mathfrak{Re}\left(\int_{\mathbb{R}} e^{ik\pi\frac{y-a}{b-a}} q(y \mid x)\,dy\right) = \int_{\mathbb{R}} \cos\left(k\pi\frac{y-a}{b-a}\right) q(y \mid x)\,dy,$$

hence

$$\begin{aligned} \mathscr{Q}_k(x) &= \frac{2}{b-a}\int_a^b q(y \mid x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &\approx \frac{2}{b-a}\int_{\mathbb{R}} q(y \mid x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= \frac{2}{b-a}\mathfrak{Re}\left(\phi\left(\frac{k\pi}{b-a} \mid x\right)e^{-ik\pi\frac{a}{b-a}}\right). \end{aligned}$$

The advantage of approximating the coefficients $\mathscr{Q}_k(x)$ by the conditional characteristic function $\phi(y \mid x)$ is that the latter is known in many models, while the transitional density $q(y \mid x)$ is not. Plugging this approximation into (4.6) yields the COS formula:

$$\hat{p}(x,t;[a,b],N) := e^{-r\tau}\sum_{k=0}^{N-1}{}'\mathfrak{Re}\left(\phi\left(\frac{k\pi}{b-a} \mid x\right)e^{-ik\pi\frac{a}{b-a}}\right)\mathscr{P}_k(T). \tag{4.7}$$

The fact that the coefficients $\mathscr{P}_k(T)$ do not depend on x allows pre-calculating them and using them for many different values of the state process x. Evaluation of the COS formula (4.7) requires evaluation of $\mathscr{P}_k(T)$ and $\phi(\cdot \mid x)$, for which we will subsequently give analytical formulae.

4.2 Fourier cosine coefficients of European call options

Since the state process $X_t = \log(S_t/K)$ is defined as the scaled log-asset price, the payoff function $p(X_T,T)$ must be expressed in terms of the scaled log-asset price, too.[21] Therefore, denoting $y := \log(S_T/K)$, the payoff function of a European call option reads

$$p(y,T) = (K(e^y - 1))^+.$$

To get an analytical formula for $P_k(T)$, we need the following results: Firstly, the cosine series coefficients χ_k of $g(y) = e^y$ on $[c,d] \subset [a,b]$

$$\chi_k(c,d) := \int_c^d e^y \cos\left(k\pi\frac{y-a}{b-a}\right)dy,$$

are known analytically as

$$\begin{aligned} \chi_k(c,d) : &= \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\left[\cos\left(k\pi\frac{d-a}{b-a}\right)e^d - \cos\left(k\pi\frac{c-a}{b-a}\right)e^c\right. \\ &\left. + \frac{k\pi}{b-a}\sin\left(k\pi\frac{d-a}{b-a}\right)e^d - \frac{k\pi}{b-a}\sin\left(k\pi\frac{c-a}{b-a}\right)e^c\right]. \end{aligned}$$

[21] Section based on [Fang and Oosterlee, 2008], p. 832.

Secondly, the cosine series coefficients ψ_k of $g(y) = 1$ on $[c,d] \subset [a,b]$

$$\psi_k(c,d) := \int_c^d \cos\left(k\pi \frac{y-a}{b-a}\right) dy$$

are known as

$$\psi_k(c,d) = \begin{cases} \left[\sin\left(k\pi \frac{d-a}{b-a}\right) - \sin\left(k\pi \frac{c-a}{b-a}\right)\right] \frac{b-a}{k\pi}, & k \neq 0 \\ d-c, & k = 0. \end{cases}$$

Hence

$$\mathscr{P}_k(T) = \frac{2}{b-a} \int_0^b K(e^y - 1) \cos\left(k\pi \frac{y-a}{b-a}\right) dy = \frac{2}{b-a} K(\chi_k(0,b) - \psi_k(0,b)).$$

4.3 COS formula in Black-Scholes model with discrete dividends

The Black-Scholes model with discrete dividends poses an exception to the COS formula (4.7).[22] We use the following formula to approximate the option price at time t:

$$\hat{p}(x,t) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \Re\mathfrak{e}\left\{\varphi\left(\frac{k\pi}{b-a} \mid x-a\right)\right\} \mathscr{P}_k(\tau^-)$$

with Fourier cosine coefficients

$$\mathscr{P}_k(\tau^-) = \frac{2}{b-a} \int_a^b p(y,\tau^-) \cos\left(k\pi \frac{y-a}{b-a}\right) dy.$$

There holds $p(y,\tau^-) = p(y + \ln(1-D), \tau^+)$.

We use discrete Fourier cosine transforms (DCT) to approximate the Fourier cosine coefficients $\mathscr{P}_k(\tau^-)$. For this, we take N grid points and define an equidistant y-grid

$$y_n := a + \left(n + \frac{1}{2}\right) \frac{b-a}{N} \text{ and } \Delta y := \frac{b-a}{N}.$$

We determine the value of function $p(y,\tau^-) = p(y + \ln(1-D), \tau^+)$ on the N grid points. The midpoint integration rule gives us

$$\begin{aligned} \mathscr{P}_k(\tau^-) &\approx \sum_{n=0}^{N-1} \frac{2}{b-a} p(y_n,\tau^-) \cos\left(k\pi \frac{y_n-a}{b-a}\right) \Delta y \\ &= \sum_{n=0}^{N-1} p(y_n,\tau^-) \cos\left(k\pi \frac{2n+1}{2N}\right) \frac{2}{N} \\ &= \sum_{n=0}^{N-1} p(y_n + \ln(1-D), \tau^+) \cos\left(k\pi \frac{2n+1}{2N}\right) \frac{2}{N}. \end{aligned}$$

The appearing DCT (Type II) can be efficiently calculated by, for example, the function dct in Matlab.

4.4 Characteristic function

In general, it holds for the characteristic function

$$\begin{aligned} \phi(\omega \mid \mathscr{F}_t) &= \mathbb{E}^{\mathbb{Q}}\left[e^{i\omega X_T} \mid \mathscr{F}_t\right] \\ &= e^{-i\omega \ln K} \mathbb{E}^{\mathbb{Q}}\left[e^{i\omega \ln S_T} \mid \mathscr{F}_t\right] \\ &= e^{-i\omega \ln K} M^{\mathbb{Q}}_{\ln S_T \mid \mathscr{F}_t}(i\omega). \end{aligned}$$

In each of the used models, we already provided an explicit formula for $M^{\mathbb{Q}}_{\ln S_T \mid \mathscr{F}_t}(z)$, in Subsection 3.1.3.

[22] Section based on [Ruijter and Oosterlee, 2015], p. 7.

5 Monte-Carlo method

In this section, we describe the Monte-Carlo method used to compute reference prices, which are used to measure the error of FGL and COS. This only applies to the Heston and Bates model, since in the Black-Scholes models we use closed-form formulae to compute reference prices. The used Monte-Carlo method comprises two steps: Obtaining samples of the terminal stock price S_T and using variance reduction techniques to reduce the variance of the Monte-Carlo estimator. The first step is semi-sequential in the sense that paths of the variance V_t are generated, but not of the stock price. It is described in detail in Subsection 5.1 for the Heston model. The Monte-Carlo method can easily be extended to incorporate stochastic jumps in the stock price, as they appear in the Bates model. We shortly describe this extension in Subsection 5.2. The mentioned variance reduction techniques are described in 5.3.

5.1 Heston model

The Heston model, as formulated in the SDEs (2.1) and (2.2), relies on two Brownian motions $W^{(1)}, W^{(2)}$ with correlation ρ.[23] Those SDEs are equivalent to the following:

$$dS_t = rS_t dt + \sqrt{V_t} S_t \left[\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)} \right], \tag{5.1}$$

$$dV_t = \kappa(\theta - V_t)\, dt + \sigma_V \sqrt{V_t} dW_t^{(1)}, \tag{5.2}$$

with two independent Brownian motions $W^{(1)}, W^{(2)}$. The stock price and variance at time t, given their values at $u < t$, permit the following expression:

$$S_t = S_u \exp\left[r(t-u) - \frac{1}{2}\int_u^t V_s ds + \rho \int_u^t \sqrt{V_s} dW_s^{(1)} + \sqrt{1-\rho^2} \int_u^t \sqrt{V_s} dW_s^{(2)} \right] \tag{5.3}$$

$$V_t = V_u + \kappa\theta(t-u) - \kappa \int_u^t V_s ds + \sigma_V \int_u^t \sqrt{V_s} dW_s^{(1)}. \tag{5.4}$$

Our goal is to obtain a sample of S_T so that we can compute a Monte Carlo simulation of the option price at $u = 0$, for instance via

$$p_{MC} = e^{-rT} \sum_{i=0}^{N} \frac{1}{N} \Phi\left(S_T^i\right), \tag{5.5}$$

with Φ being the payoff function and $S_T^1, \ldots, S_T^N$ the individual samples. Actually, our Monte Carlo estimator is slightly more complicated than (5.5) due to various variance reduction techniques, but we will elaborate on this detail in 5.3. As becomes evident from (5.3) - (5.4), for sampling S_T given S_0, V_0, it suffices to follow the following steps:

1. Generate a sample of $V_{t_0}, \ldots V_{t_n}$ on an equidistant time grid $[0 = t_0 < t_1 < \ldots < t_n = T]$ of grid width $t_{i+1} - t_i = \frac{1}{100}$.
2. Using the samples $V_{t_0}, \ldots V_{t_n}$, generate a sample of $\int_0^T V_s ds$.
3. Recover $\int_0^T \sqrt{V_s} dW_s^{(1)}$ from (5.4).
4. Generate a sample of $\int_0^T \sqrt{V_s} dW_s^{(2)}$.
5. Combine all previous steps via (5.3), resulting in a sample of S_T.

In steps 1 and 2, we diverge from ([24]), which does not use a time grid and uses characteristic functions to sample $\int_0^T V_s ds$.

5.1.1 Simulation of V_T

According to ([25]), the variance V_t permits the following representation:

$$V_t = \frac{\sigma_V^2 \left(1 - e^{-\kappa(t-u)}\right)}{4\kappa} \chi_d'^2 \left(\frac{4\kappa e^{-\kappa(t-u)}}{\sigma_V^2 \left(1 - e^{-\kappa(t-u)}\right)} V_u \right), u < t.$$

[23] Section based on [Broadie and Kaya, 2004]
[24] [Broadie and Kaya, 2004]
[25] [Cox et al., 1985]

Here, $\chi_d'^2(\lambda)$ denotes a non-central chi-squared random variable with d degrees of freedom, where

$$\begin{aligned}\lambda &= \frac{4\kappa e^{-\kappa(t-u)}}{\sigma_V^2\left(1-e^{-\kappa(t-u)}\right)}V_u,\\ d &= \frac{4\theta\kappa}{\sigma_V^2}.\end{aligned}$$

Therefore, for sampling from V_t, it suffices to be able to sample from a non-central chi-squared distribution. For the latter, the following two facts are useful: According to [Johnson et al., 1994], if $d>1$, then $\chi_d'^2(\lambda)$ permits the following representation:

$$\chi_d'^2(\lambda) = \chi_1'^2(\lambda) + \chi_{d-1}^2, \tag{5.6}$$

where χ_{d-1}^2 is an ordinary chi-squared random variable. To sample from the latter, we use the Matlab routine chi2rnd. $\chi_1'^2(\lambda)$, in turn, can easily be sampled based on a standard normal random variable Z independent of χ_{d-1}^2, so that (5.6) is equivalent to

$$\chi_d'^2(\lambda) = \left(Z+\sqrt{\lambda}\right)^2 + \chi_{d-1}^2.$$

In the more general case $d>0$, a noncentral chi-squared random variable can be represented as an ordinary chi-squared random variable with a random degrees of freedom parameter. More specifically, let N be a Poisson random variable with mean $\frac{1}{2}\lambda$. Then χ_{d+2N}^2 has the same distribution as $\chi_d'^2(\lambda)$. So to sample from the latter, we first sample from N and then from an ordinary chi-squared random variable χ_{d+2N}^2 with $d+2N$ degrees of freedom.

We apply the aforementioned procedure to all $u=t_i$ and $t=t_{i+1}$, for $i=0,\ldots,n-1$, resulting in samples V_{t_i} for $i=1,\ldots,n$.

5.1.2 Simulation of $\int_0^T V_s ds$

Deviating from the procedure described in ([26]), we use a simple trapezoidal rule to sample $\int_{t_{i-1}}^{t_i} V_s ds$:

$$\int_{t_{i-1}}^{t_i} V_s ds \approx \left(V_{t_i}+V_{t_{i-1}}\right)/2\cdot\left(t_i-t_{i-1}\right).$$

The integral over the entire time interval $[0,T]$ is then simply the sum of all intermediate intervals:

$$\int_0^T V_s ds = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} V_s ds.$$

5.1.3 Simulation of $\int_0^T \sqrt{V_s}dW_s^{(1)}$, $\int_0^T \sqrt{V_s}dW_s^{(2)}$ and S_T

A sample of $\int_0^T \sqrt{V_s}dW_s^{(1)}$ can easily be obtained by solving equation (5.4) for it. This yields

$$\int_0^T \sqrt{V_s}dW_s^{(1)} = \frac{1}{\sigma_V}\left(V_T - V_0 - \kappa\theta T + \kappa\int_0^T V_s ds\right),$$

in which we can simply insert the samples of $\int_0^T V_s ds$ and V_T.

In addition, the SDEs (5.1) - (5.2) of the Heston model make it obvious that V_t is independent of the Brownian motion $W^{(2)}$. Hence, $\int_0^T \sqrt{V_s}dW_s^{(2)}$ is normally distributed with mean 0 and variance $\int_0^T V_s ds$, so it can be simulated via the variable

$$\sqrt{\int_0^T \sqrt{V_s}ds}\cdot G,$$

with G being a standard normal variable.

Combining these results, we generate a sample of S_T by inserting the samples of $\int_0^T V_s ds$, $\int_0^T \sqrt{V_s}dW_s^{(1)}$ and $\int_0^T \sqrt{V_s}dW_s^{(2)}$ into (5.3):

$$S_T = S_0\exp\left[rT - \frac{1}{2}\int_0^T V_s ds + \rho\int_0^T \sqrt{V_s}dW_s^{(1)} + \sqrt{(1-\rho^2)\int_0^T V_s ds}\cdot G\right] \tag{5.7}$$

[26] [Broadie and Kaya, 2004]

5.2 Bates model

Generating samples of S_T in the Bates model is a simple extension of the aforementioned procedure in the Heston model, because the jumps can be treated separately.[27] The procedure is as follows:

1. Ignore the jump part in (2.3) and simulate S_T in the same way as in the Heston model.
2. Generate a sample from a Poisson random variable N_J with mean λT, representing the number of jumps that occured in the time horizon $[0,T]$. Here, λ refers to the jump intensity of the Poisson process in (2.3).
3. Simulate the log-jumps as a normal random variable J with mean $-\lambda T\left(e^{\delta^2/2+\gamma}-1\right)+N_J\gamma$ and variance $\delta^2 N_J$.
4. Compute the S_T sample as

$$S_T = S_0 \exp\left[rT - \frac{1}{2}\int_0^T V_s ds + \rho\int_0^T \sqrt{V_s}dW_s^{(1)} + \sqrt{(1-\rho^2)\int_0^T \sqrt{V_s}ds} + J\right].$$

5.3 Variance reduction techniques

[28]The previously described procedure yields samples of S_T, which can be used to form a crude Monte Carlo estimator. We improve this estimator by applying three variance reduction techniques: Importance sampling, antithetic sampling and control variates. In each of the paragraphs 5.3.1 - 5.3.3, we describe these techniques first in general, and then specify how we apply them to our test setting.

5.3.1 Importance sampling

A problem we encounter in the Heston and Bates models for some test parameters is that many of the Monte Carlo samples are zero. This can occur when the initial stock price S_0 is significantly lower than the strike price K, so that many of the stock samples S_T^i end up being lower than K, too, resulting in a zero payoff at time T. To address this problem, we make use of importance sampling, which is based around the idea of using a measure change to force the payoff $\Phi(S_T)$ to be non-zero more often. More specifically, let $\mathbb{Q}$ be the risk-neutral measure implicitly used all along, so that the option price can be expressed as

$$p = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T) \mid \mathscr{F}_0\right].$$

Now, find a measure $\mathbb{Q}' \sim \mathbb{Q}$ and a $\mathbb{Q}'$-density $L := \frac{d\mathbb{Q}}{d\mathbb{Q}'}$. Then the option price also permits the representation

$$p = e^{-rT}\mathbb{E}^{\mathbb{Q}'}\left[\Phi(S_T)L\right].$$

We want to choose $\mathbb{Q}'$ such that the probability of a non-zero outcome $\mathbb{Q}'(\Phi(S_T) > 0)$ is reasonably high. Simulating S_T and L under $\mathbb{Q}'$ yields a Monte Carlo estimator of

$$p_{MC,IS} = \frac{1}{N}\sum_{i=1}^{N}\Phi\left(S_T^i\right)L_i.$$

The SDEs (2.1), (2.2) and (2.3), (2.4) characterising the Heston and Bates models are special cases of the following SDE system:

$$\begin{aligned} dS_t &= rS_t dt + \sigma(t,S_t,V_t)\,S_t dW_t^{\mathbb{Q}} + S_{t-}dZ_t, \\ dV_t &= m(t,V_t)\,dt + l(t,V_t)\,dW_t^{\mathbb{Q}}. \end{aligned}$$

We choose $\mathbb{Q}'$ such that

$$\begin{aligned} dS_t &= aS_t dt + \sigma(t,S_t,V_t)\,S_t dW_t^{\mathbb{Q}'} + S_{t-}dZ_t, \\ dV_t &= m(t,V_t)\,dt + l(t,V_t)\,dW_t^{\mathbb{Q}'}, \end{aligned} \tag{5.8}$$

[27]Section based on [Broadie and Kaya, 2004], p. 20ff.
[28]Section based on [Boyle, 1997]

where $W^{\mathbb{Q}'}$ is a 2-dimensional standard Brownian motion under $\mathbb{Q}'$, and Z is a compound Poisson process like in the Bates model. The Girsanov kernel g in the measure change should then satisfy $a - \sigma(t, S_t, V_t) g(t) = r$, yielding

$$L = \exp\left(-\frac{1}{2}\int_0^T |g(u)|^2 du - \int_0^T g(u) * dW^{\mathbb{Q}'}(u)\right).$$

In the Heston and Bates models, we have

$$\begin{aligned}
\sigma(t, S_t, V_t) &= \left[\rho\sqrt{V_t}, \sqrt{1-\rho^2}\sqrt{V_t}\right], \\
m(t, V_t) &= \kappa(\theta - V_t), \\
l(t, V_t) &= \left[\sigma_V\sqrt{V_t}, 0\right], \\
g(t) * &= \left[0, \frac{a-r}{\sqrt{1-\rho^2}\sqrt{V_t}}\right],
\end{aligned}$$

and hence

$$L = \exp\left(-\frac{1}{2}\int_0^T \frac{(a-r)^2}{(1-\rho^2)V_u} du - \int_0^T \frac{a-r}{\sqrt{1-\rho^2}\sqrt{V_u}} dW_2^{\mathbb{Q}'}(u)\right).$$

Using that

$$\begin{aligned}
S_T &\stackrel{d}{=} S_0 \exp\left(rT - \frac{1}{2}\int_0^T V_u du + \rho\left(V_T - V_0 - \kappa\theta T + \kappa\int_0^T V_u du\right)/\sigma_V + \sqrt{(1-\rho^2)\int_0^T V_s ds} G + J\right), \qquad (5.9) \\
J &\stackrel{d}{=} -\lambda T\left(\exp\left(\gamma^2/2 + \delta\right) - 1\right) + N_J\delta + \gamma\sqrt{N_J} G_J, \\
N_J &\sim Poi(\lambda T), G \sim \mathcal{N}(0,1), G_J \in \mathcal{N}(0,1), G \text{ and } G_J \text{ independent},
\end{aligned}$$

we can further reduce variance by conditioning L on $\{V_u \mid 0 \le u \le T\}$ and S_T, yielding

$$\tilde{L} = \mathbb{E}^{\mathbb{Q}'}\left[L \mid \{V_u \mid 0 \le u \le T\}, S_T\right] \stackrel{d}{=} \exp\left(-\frac{(a-r)^2 T^2}{2(1-\rho^2)\int_0^T V_u du} - \frac{(a-r)T}{\sqrt{1-\rho^2}\sqrt{\int_0^T V_u du}} G\right), \qquad (5.10)$$

where G is the same standard normally distributed G as above. Numerical experiments suggest to use

$$a = \frac{\ln(K/S_t)}{2T}$$

in (5.8). We use

$$a = \max\left(\frac{\ln(K/S_t)}{2T}, -2r\right),$$

so that we do not change measure when the option is deep in the money.

5.3.2 Antithetic sampling

For simplicity, let $V := \Phi(S_T)$. Ignoring the discounting factor, our goal is to estimate $\mathbb{E}[V]$. Assume there is a variable V' such that

1. $\mathbb{E}[V'] = \mathbb{E}[V]$
2. $\mathbb{V}[V'] = \mathbb{V}[V]$
3. V' can be simulated at the same complexity as V.

Then the variable

$$W := \frac{V + V'}{2}$$

has the same expectation as V and has variance

$$\begin{aligned}\mathbb{V}[W] &= \frac{1}{4}\left(\mathbb{V}[V]+2\mathbb{C}\left[V,V'\right]+\mathbb{V}\left[V'\right]\right)\\ &= \frac{1}{2}\left(\mathbb{V}[V]+\mathbb{C}\left[V,V'\right]\right).\end{aligned}$$

The number of samples of V needed to achieve an error less than any given $\varepsilon > 0$ is

$$N_V > \lambda_{\alpha/2}^2 \frac{\mathbb{V}[V]}{\varepsilon^2}.$$

Likewise, there holds

$$N_W > \lambda_{\alpha/2}^2 \frac{\mathbb{V}[W]}{\varepsilon^2}.$$

Taking into account that simulating W has double the complexity of simulating V (because both V and V' must be simulated, at the same complexity), it is better to use W if

$$2\lambda_{\alpha/2}^2 \frac{\mathbb{V}[W]}{\varepsilon^2} < \lambda_{\alpha/2}^2 \frac{\mathbb{V}[V]}{\varepsilon^2},$$

which is equivalent to

$$\begin{aligned}\mathbb{V}[V]+\mathbb{C}\left[V,V'\right] &< \mathbb{V}[V]\\ \Leftrightarrow \mathbb{C}\left[V,V'\right] &< 0.\end{aligned}$$

Therefore, finding a V' that is negatively correlated to V improves accuracy. The following theorem proves useful in this context:

Theorem 5.1. *Let $V = \varphi(U)$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a monotone function. Moreover, assume that there exists a non-increasing transform $T : \mathbb{R} \to \mathbb{R}$ such that $U \overset{d}{=} T(U)$. Then $V = \varphi(U)$ and $V' = \varphi(T(U))$ are identically distributed and*

$$\mathbb{C}\left[V,V'\right] = \mathbb{C}\left[\varphi(U), \varphi(T(U))\right] \leq 0.$$

In our case, we apply this to the standard normal variable $U = G$ in (5.10) and (5.9), and the transform we use is $T(U) = -G$. Combined with importance sampling, V will be

$$\begin{aligned}V &= \Phi(S_T)\tilde{L}\\ &= \left(S_0 \exp\left(rT - \frac{1}{2}\int_0^T V_u du + \rho\left(V_T - V_0 - \kappa\theta T + \kappa\int_0^T V_u du\right)/\sigma_V + \sqrt{(1-\rho^2)\int_0^T \sqrt{V_s} ds}\, G + d_J\right) - K\right)^+\\ &\quad \cdot \exp\left(-\frac{(a-r)^2 T^2}{2(1-\rho^2)\int_0^T V_u du} - \frac{(a-r)T}{\sqrt{1-\rho^2}\sqrt{\int_0^T V_u du}} G\right).\end{aligned}$$

Simplifying the notation, this is really the same as

$$V = \varphi(G) = \left(S_0 \exp(\mu + \sigma G) - K\right)^+ \cdot \exp\left(\mu' + \sigma' G\right),$$

with $\mu, \sigma, \mu', \sigma'$ defined in the obvious way. Likewise, the antithetic variable is

$$V' = \varphi(T(G)) = \left(S_0 \exp(\mu - \sigma G) - K\right)^+ \cdot \exp\left(\mu' - \sigma' G\right).$$

The transformation φ is a composition of several transformations, namely

$$\begin{aligned}x &\mapsto \mu + \sigma x,\\ y &\mapsto S_0 \exp(y),\\ z &\mapsto (z-K)^+,\\ x' &\mapsto \mu' + \sigma' x',\\ (v,w) &\mapsto v \cdot w.\end{aligned}$$

They are all monotonously increasing except for the last one, which however is also increasing in each argument on the used domain: The arguments $v = \left(S_0 \exp(\mu - \sigma G) - K\right)^+$ and $w = \exp(\mu' - \sigma' G)$ are always positive. Utilizing Theorem 5.1, the variables V, V' have therefore negative correlation, making this a useful improvement. It is also evident that the transformation T does not increase computational complexity.

5.3.3 Control variates

The aforementioned procedures produce samples of S_T and L, which we refer to by S_T^i and L_i for $i = 1, \ldots, N$, respectively. As their antithetic variables, it also yields samples $\tilde{S_T^i}$ and $\tilde{L}_i$ which are negatively correlated with the former samples. Combining importance sampling and antithetic sampling, an unbiased Monte Carlo estimator would be:

$$p_{MC} = e^{-rT} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \left(\Phi\left(S_T^i\right) L_i + \Phi\left(\tilde{S_T^i}\right) \tilde{L}_i \right).$$

For simplicity, we define $X_i := \frac{1}{2} \left(\Phi\left(S_T^i\right) L_i + \Phi\left(\tilde{S_T^i}\right) \tilde{L}_i \right)$ as samples of a random variable X. We further optimize our Monte Carlo estimate by employing a control variate. In general, this technique works as follows: Assume there is a random variable Y such that

1. $m := \mathbb{E}[Y]$ is known, and
2. $X - Y$ can be computed at the same complexity as $\Phi(S_T)$.

Then the random variable

$$Z := X + \alpha (Y - m) \tag{5.11}$$

for any α has the same expectation as X. Hence, if $X_1, \ldots, X_N$ and $Y_1, \ldots, Y_N$ are i.i.d. samples of X and Y, respectively, then

$$\tilde{p}_{MC} = e^{-rT} \frac{1}{N} \sum_{i=1}^{N} (X_i + \alpha (Y_i - m)) \tag{5.12}$$

is an unbiased Monte Carlo estimator of the option price. The trick of the control variate technique lies in choosing α such that the variance of (5.12) or, equivalently, of (5.11) is minimal. Finding the optimal α is simple:

$$\begin{aligned} \mathbb{V}[Z] &= \mathbb{V}[X + \alpha (Y - m)] \\ &= \mathbb{V}[X] + 2\alpha \mathbb{C}[X, Y] + \alpha^2 \mathbb{V}[Y] \end{aligned} \tag{5.13}$$

Differentiating this expression by α, equating to 0 and solving for α yields

$$\alpha^* = -\frac{\mathbb{C}[X, Y]}{\mathbb{V}[Y]},$$

which we can estimate from the sample of S_T and Y. Inserting α^* into (5.13) yields

$$\mathbb{V}[Z] = \mathbb{V}[X] \left(1 - \rho(X, Y)^2\right), \tag{5.14}$$

where $\rho(X, Y)$ denotes the correlation of X and Y. Equation (5.14) thus shows that we should choose Y such that its correlation is close to 1 or -1 in order to achieve a high reduction in variance.

Moving from the general description of the control variate technique to our specific problem, we chose as a control variate

$$Y := e^{-rT} \left(S_T L + \tilde{S_T} \tilde{L} \right) / 2 - S_0,$$

which has a known expectation of $m = 0$, due to the fact that $e^{-rT} S_T L$ is a martingale under $\mathbb{Q}'$, and thus its antithetic variable $\tilde{S_T} \tilde{L}$, too. It is reasonable to assume that this choice of Y has positive correlation with X.

6 Tests

This section is divided into three parts: In Subsection 6.1, we describe the testing methodology, which includes aspects such as model and method-specific parameters. In Subsection 6.2, we briefly specify the four test cases. Lastly, in Subsection 6.3, we present the test results both in terms of accuracy and computational speed.

6.1 Methodology

6.1.1 Overview

Our goal is to provide a broader test coverage than the Benchop project. Additionally, we aim at discovering dependencies of the option pricing error on the model parameters, the model and the used pricing method. For this purpose, we consider four major test cases: Black-Scholes, Black-Scholes with discrete dividends, Heston and Bates. Each test case encompasses many combinations of model parameters. We refer to ϑ for a combination of parameters and Θ_M for all combinations of model parameters in a model M. For instance, in the Black-Scholes model, we have $\vartheta = (\tau, S_0, r, \sigma, K) \in \Theta_{BS}$. For the choice of Θ_M, we refer to Subsection 6.1.2.

For each single combination of parameters, we compute the reference price of a European call option via analytical formulae, where available, or Monte-Carlo. Such formulae are available for the Black-Scholes model with and without discrete dividends, but not for Heston and Bates. For the latter, we use Monte Carlo because it is a robust and wide-spread technique. As for the choice of the financial product, we chose a European call option because it is very common in financial markets and because this thesis is based on the Benchop project, which also largely used that product.

In addition to the reference price, we compute the price via FGL and COS. The error of both methods are then measured by the relative difference to the reference price: for example, if $p(\vartheta, M, \text{FGL})$ is the FGL option price in model M with model parameters ϑ and $p(\vartheta, M, \text{ref})$ is the reference price in the same model setup, then the raw error is

$$E_{raw}(\vartheta, M, \text{FGL}) := \frac{|p(\vartheta, M, \text{FGL}) - p(\vartheta, M, \text{ref})|}{p(\vartheta, M, \text{ref})}.$$

The COS error is defined analogously. In the Heston and Bates model, the reference method is a non-deterministic one (Monte Carlo) and should therefore not be regarded as an exact option price, but always in conjunction with its $(1-\alpha)$-confidence interval that covers the real option price with probability $1-\alpha$. Consequently, we define the FGL and COS error to be 0 if the respective price falls within this confidence interval. The latter is given by

$$[c_1, c_2] := \left[p_{ref} - u_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{N_S}}, p_{ref} + u_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{N_S}} \right],$$

where p_{ref} is the reference price, $u_{1-\alpha/2}$ the $(1-\alpha/2)$-quantile of the standard normal distribution, N_S the number of simulation steps and $\hat{\sigma}$ is the standard deviation estimated from the Monte Carlo sample. We use $\alpha = 0.05$, so that $u_{1-\alpha/2} \approx 1.96$. To summarise, in the Heston and Bates model we measure the error by

$$E(\vartheta, M, FGL) := \begin{cases} 0, & \text{if } p_{FGL} \in [c_1, c_2], \\ E_{raw}(\vartheta, M, FGL), & \text{else.} \end{cases}$$

and likewise for COS. In the Black-Scholes models, we simply use the raw error.

The described procedure enables us to consider the error a function of the model parameters ϑ, the model M and the pricing method. Since there are generally too many combinations of model parameters to depict or list the error for each one - e.g. 2304 in the Bates test case -, we take the median 1. intra- and 2. inter-model-wise.

By an intra-model median, we mean the following: Let $M \in \{\text{BS1, BS1DD, Heston, Bates}\}$ be any of the tested models, and Θ_M the corresponding set of combinations of model parameters. Now fix a parameter with more than one distinct value in Θ_M, say τ, and for each value of τ take the median of the error with respect to all other parameters. This yields a function

$$\tau \quad \mapsto \quad \text{median} \left\{ E(\vartheta, M, \text{FGL}) \mid \vartheta = \left(\tau^{\vartheta}, S_0^{\vartheta}, \ldots \right) \in \Theta_M, \tau^{\vartheta} = \tau \right\},$$

allowing us to examine the dependence of the error on the parameter τ. We do the above for both FGL and COS and all parameters that have more than one distinct value in Θ_M.

By inter-model median, we simply mean taking the median of all errors within each model and then compare across models. Thus, we are equipped with the function

$$\begin{aligned} \{\text{BS1, BS1DD, Heston, Bates}\} &\to \mathbb{R}_{\geq 0}, \\ M &\mapsto \text{median} \{ E(\vartheta, M, \text{FGL}) \mid \vartheta \in \Theta_M \}, \end{aligned}$$

and its equivalent for the COS method. BS1 stands for Black-Scholes and BS1DD for Black-Scholes with discrete dividends.

Taking the median intra- and inter-model-wise allows us to discover dependencies of the error on each model parameter within a given model and on the model itself. We depict the errors of both methods alongside each other in Subsection 6.3.

6.1.2 Model parameters

In order to provide a more comprehensive test coverage of the FGL and COS methods than in the Benchop project, each model is tested with a set of many different combinations of model parameters, which we refer to by Θ_M for any given model M. This set Θ_M is defined as follows: First, we choose a set of test values for each single model parameter. For instance, in the Black Scholes model, these are $\Theta_M^\tau = \{0.25, 0.5, 0.75, \ldots, 2\}, \Theta_M^{S_0} = \{70, 80, 90, \ldots, 130\}, \Theta_M^r = \{0.03\}, \Theta_M^\sigma = \{0.25, 0.5, 0.75, 1\}, \Theta_M^K = \{100\}$ for time to maturity, current stock price, risk-free interest rate, stock price volatility and strike price, respectively. The set of parameter combinations is simply their Cartesian product:

$$\Theta_{BS} := \Theta_M^\tau \times \Theta_M^{S_0} \times \Theta_M^r \times \Theta_M^\sigma \times \Theta_M^K$$

The exception of this rule is the parameter V_0 (initial volatility) in the Heston and Bates models. To keep the number of simulations at a reasonable level, we set this parameter equal to the long-term volatility θ. Hence, the set of parameter combinations in the Heston model is

$$\Theta_{Heston} := \left\{ (S_0, K, \tau, r, \kappa, \theta, V_0, \sigma_V, \rho) \mid S_0 \in \Theta_M^{S_0}, K \in \Theta_M^K, \tau \in \Theta_M^\tau, r \in \Theta_M^r, \kappa \in \Theta_M^\kappa, \theta \in \Theta_M^\theta, V_0 = \theta, \sigma_V \in \Theta_M^{\sigma_V}, \rho \in \Theta_M^\rho \right\},$$

and the set Θ_{Bates} is defined analogously.

For the sake of consistency, we wanted the different models to share their parameter sets, where possible. For instance, Heston and Bates use the same set of test values for correlation. All parameters are listed in table 6.1.

Due to the high number of parameters of the Bates and Heston models and to keep computational time at a reasonable level, we restricted the size of each parameter test set. The choice of parameters is often motivated by including some small, intermediate and large values.

Parameter	Description	Values	Used in models
τ	time to maturity	0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2	BS1, BS1DD
τ	time to maturity	0.5, 1, 1.5, 2	Heston, Bates
S_0	current stock price	70, 80, 90, 100, 110, 120, 130	BS1, BS1DD
S_0	current stock price	85, 100, 115	Heston, Bates
r	risk-free interest rate	0.03	BS1, BS1DD, Heston, Bates
K	strike price	100	BS1, BS1DD, Heston, Bates
σ	stock price volatility	0.25, 0.5, 0.75, 1	BS1, BS1DD
D	dividend yield	0.01, 0.05, 0.1	BS1DD
α	time of dividend payment (as percentage of τ)	0.25, 0.5, 0.75	BS1DD
σ_V	volatility of volatility	0.5, 0.75, 1	Heston
σ_V	volatility of volatility	0.25, 1	Bates
ρ	correlation of Brownian motions	-0.5, 0, 0.5	Heston, Bates
κ	mean reversion speed of volatility	0.1, 0.5, 1	Heston
κ	mean reversion speed of volatility	0.25, 1	Bates
θ	mean volatility	0.1, 0.5, 1	Heston
θ	mean volatility	0.25, 1	Bates
λ	jump intensity of Poisson process J_t	0.25, 1	Bates
γ	mean jump size	0.25, 1	Bates
δ	jump standard deviation	0.25, 1	Bates

Table 6.1: Table of model parameters. Note that Heston and Bates use a smaller parameter set for time to maturity τ and current stock price S_0. This is because these models have more model parameters than the Black-Scholes type models, and we aimed at keeping the computation time at a reasonable level.

6.1.3 Method-specific parameters

FGL A discretionary parameter of FGL is the number of weights n in the Gauss-Laguerre quadrature. We used $n = 1000$ as a tradeoff between high accuracy and the peril of long computation times and numerical overflow. To understand the latter,

note that the maximum of the abscissae in the Gauss-Laguerre quadrature increase as n increases. As the FGL method involves evaluating the MGF at these abscissae and the MGF in the Heston and Bates model contain exponentials, this may result in numerical overflow.

COS Two of the errors introduced by the COS formula (4.7) stem from the truncation of the integration range in (4.5) to $[a,b]$ and the truncation of the summation in (4.6) to N coefficients. For the choice of $[a,b]$, Fang and Oosterlee recommend:[29]

$$[a,b] = \left[c_1 - L\sqrt{c_2+\sqrt{c_4}}, c_1 + L\sqrt{c_2+\sqrt{c_4}}\right], \quad \text{with } L = 10. \tag{6.1}$$

c_1, c_2, c_4 are the cumulants, i.e. the derivatives of the log-moment generating function of the state process at zero:

$$c_i := \frac{\partial^i}{\partial t^i} \log \mathbb{E}\left[e^{tX_0}\right] \mid_{t=0}.$$

Clearly, the disadvantage of formula (6.1) is that it requires the cumulants to be analytically known, which can be difficult for intricate MGFs as the ones in the Heston and Bates model. In some models, we altered this formula to achieve better accuracy. Overall, the following values for L and N were used:

- Black-Scholes: $[a,b] = \left[c_1 - L\sqrt{c_2}, c_1 + L\sqrt{c_2}\right]$, where $c_1 = \left(r - \sigma^2/2\right)\tau, c_2 = \sigma^2\tau, L = 6.6, N = 26$
- Black-Scholes with discrete dividends: $[a,b] = \left[c_1 - L\sqrt{c_2}, c_1 + L\sqrt{c_2}\right]$, with c_1 and c_2 as above and $L = 6.8, N = 41$
- Heston: $[a,b] = [c_1 - L, c_1 + L]$, where $c_1 = r\tau + \left(1 - e^{-\kappa\tau}\right)(\theta - V_0)/(2\kappa) - \theta\tau/2, L = 7.4, N = 12$
- Bates: $[a,b] = \left[c_1 - L\cdot\left|\sqrt{\bar{c}_2 + c_2 + \sqrt{c_4}}\right|, c_1 + L\cdot\left|\sqrt{\bar{c}_2 + c_2 + \sqrt{c_4}}\right|\right]$, where

$$\begin{aligned}
c_1 &= \tau\left(r + \lambda\left(\gamma - e^{\gamma+\delta^2/2} + 1\right)\right) + \left(1 - e^{-\kappa\tau}\right)(\theta - V_0)/(2\kappa) - \theta\tau/2, \\
\bar{c}_2 &= \theta\tau\left(1 + \sigma_V\right), \\
c_2 &= \lambda\tau\left(\gamma^2 + \delta^2\right), \\
c_4 &= \lambda\tau\left(\gamma^4 + 6\delta^2\gamma^2 + 3\delta^4\lambda\right), \\
L &= 9, \\
N &= 250.
\end{aligned}$$

We obtained the above values for L and N as follows: For each model, we defined a small sets of test cases for calibration, which do not overlap with the test cases described in table 6.1 and adhered to the following procedure:

1. For each model and calibration set, compute the reference prices.
2. For each test case, optimize N and L such that all COS prices match the reference prices well. More precisely, for each $N \in \{2, \ldots, 1000\}$, optimize $L \in (0, 500)$ so that the absolute error of COS price to reference price is minimized. Then, we chose the smallest N and the corresponding L such that the maximal error is below a tolerance of 10^{-4}.
3. With the thusly obtained N, L, repeat computing the COS prices, but this time for the larger set of test cases.

To be clear, in step 3 we used the same L, N for all combinations of parameters. Since the smaller set of test cases for calibration does not coincide with the larger set of test cases, we can argue that the COS prices are not tailored to the reference prices and that we used the calibration set only to get a general idea of which L, N to use. The calibration sets of parameter combinations are defined in table 6.2.

[29] [Fang and Oosterlee, 2008], p. 839.

Parameter	Description	Values	Used in models
τ	time to maturity	0.5, 1, 1.5	BS1, BS1DD
τ	time to maturity	0.5, 1.5	Heston, Bates
S_0	current stock price	85, 115	BS1, BS1DD
S_0	current stock price	90, 110	Heston, Bates
r	risk-free interest rate	0.03	BS1, BS1DD, Heston, Bates
K	strike price	100	BS1, BS1DD, Heston, Bates
σ	stock price volatility	0.33, 0.66, 1	BS1, BS1DD
σ_V	volatility of volatility	0.5	Heston
D	dividend yield	0.05, 0.1	BS1DD
α	time of dividend payment	0.33, 0.66	BS1DD
κ	reversion speed of volatility	0.25, 0.75	Heston, Bates
θ	mean volatility	0.25, 0.75	Heston, Bates
ρ	correlation of Brownian Motions	-0.5, 0.5	Heston
ρ	correlation of Brownian Motions	-0.25, 0.25	Bates
λ	jump intensity of Poisson process J_t	0.5	Bates
γ	mean jump size	0.5	Bates
δ	jump standard deviation	0.5	Bates

Table 6.2: Table of model parameters used for calibration of L,N in the COS method in the BS1, BS1DD, and Heston model. As before, we used all combinations of the respective parameters.

Monte Carlo We set the number of particles to $N_S = 1,000,000$ and the number of time steps N_T proportional to the time to maturity, with 100 per year. So for example, if $\tau = 2$, then $N_T = 200$.

6.1.4 Hard- & software

All code is run on a 2,3 GHz Intel Core i7 with Matlab R2015b and the following toolboxes: Financial Toolbox, Optimization Toolbox, Parallel Computing Toolbox, Signal Processing Toolbox, Statistics and Machine Learning Toolbox and Symbolic Math Toolbox. Additional libraries used: the Benchop files[30], the *allcomb* utility function[31] and a golden section search function from Matlab File Exchange[32].

6.2 Test case specification

6.2.1 Black-Scholes model

- Model: Black-Scholes (see Subsection 2.2.1)
- Reference value: The Black-Scholes formula:

$$\begin{aligned} f(S_0,K) &= N(d_1)S_0 - N(d_2)Ke^{-r\tau}, \\ d_1 &= \frac{1}{\sigma\sqrt{\tau}}\left(\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau\right), \\ d_2 &= d_1 - \sigma\sqrt{\tau}. \end{aligned}$$

- Parameters: All combinations of S_0, τ, K, σ, r in table 6.1.

6.2.2 Black-Scholes model with discrete dividends

- Model: Black-Scholes with discrete dividends (see Subsection 2.2.2)

[30] http://www.it.uu.se/research/project/compfin/benchop
[31] http://www.mathworks.com/matlabcentral/fileexchange/10064-allcomb
[32] http://www.mathworks.com/matlabcentral/fileexchange/25919-golden-section-method-algorithm

- Reference value: By [Björk, 2009], Proposition 16.6, p. 235, the price in this model is the same as the price in the BS model without discrete dividends, but with $S_0(1-D)$ instead of S_0 as the underlying price. Hence, we can employ the BS formula:[33]

$$\begin{aligned} f(S_0,K) &= N(d_1)S_0(1-D) - N(d_2)Ke^{-r\tau}, \\ d_1 &= \frac{1}{\sigma\sqrt{\tau}}\left(\ln\left(\frac{S_0(1-D)}{K}\right) + \left(r+\frac{\sigma^2}{2}\right)\tau\right), \\ d_2 &= d_1 - \sigma\sqrt{\tau}. \end{aligned}$$

- Parameters: All combinations of $S_0, \tau, K, \sigma, r, D, \alpha$ from table 6.1.

6.2.3 Heston model

- Model: Heston (see Subsection 2.2.3)
- Reference value: A Monte-Carlo simulation with $N = 1,000,000$ particles using importance sampling, anithetic sampling, and a control variate.[34]
- Parameters: All combinations of $S_0, \tau, K, \sigma_V, r, \rho, \kappa, \theta$ from table 6.1. The initial volatility V_0 is always equal to σ_V.

6.2.4 Bates model

- Model: Bates (see Subsection 2.2.4)
- Reference value: For each combination of the model parameters, the reference value is a Monte Carlo approximation with $N = 1,000,000$ particles using importance sampling, anithetic sampling, and a control variate.[35]
- Parameters: All combinations of $S_0, K, \tau, r, \kappa, \theta, \sigma_V, \rho, \lambda, \gamma, \delta$ from table 6.1. The initial volatility V_0 is always equal σ_V.

6.3 Results

In this subsection, we present the results of our previously specified test cases. For each test case, we list the median, mean, maximum and standard deviation of the relative error in tables 6.3 - 6.6. The sensitivity of the relative error to the model parameters is depicted in figures 6.1 - 6.5. Furthermore, the mean, median and maximum relative error as a function of the model are shown in figure 6.6. Additionally, the median, mean, maximum and standard deviation of the computation times are listed in tables 6.7 - 6.10. All computation times are given in seconds.

For both methods, we mostly achieve several digits of relative accuracy, with the exception of the maximal COS error in the Bates model. By comparison, the relative Monte Carlo error - as measured by the length of the confidence interval divided by the Monte Carlo price - is typically around $1.4 \cdot 10^{-3}$.

The Black-Scholes models are similar in the sensitivity of both errors to changes in model parameters: The mean relative COS error exhibits a local minimum for values of τ and σ around 0.5. The FGL error is generally smaller than the COS error by several magnitudes and also rather insensitive to changes of model parameters. A noteworthy aspect of the Black-Scholes model with discrete dividends is that the COS error decreases strongly as α increases from 0.25 to 0.5.

In the Heston and Bates model, the FGL error is again generally smaller than the COS error, but to a lesser extent, especially in the Heston model. The FGL error in the Bates model behaves similar to the Black-Scholes models in that it is rather insensitive to changes of the model parameters. This is different in the Heston model, where both errors are similarly sensitive. Furthermore, in the Heston model, both errors decrease for higher values of S_0, τ, ρ, κ. In the Bates model, the COS error increases strongly as $\tau, \lambda, \gamma, \delta$ increase. Since these parameters correspond to the number and size of jumps occuring in the stock price paths, this shows that the COS method performs worse as the jump aspect of that model becomes more prominent.

Comparing the overall error between models, the mean, maximum and standard deviation of both errors are significantly higher in the Heston and Bates model compared to the Black-Scholes models.

The mean and maximum of the COS error increases sharply in the Bates model compared to the other models, while the FGL error is less sensitive.

[33] from [von Sydow et al., 2015], p. 16

[34] See the Matlab routine HestonEuCall_CVISA_ref.m in the thesis code. The Monte Carlo scheme was implemented by Magnus Wiktorsson.

[35] See the Matlab routine BatesEuCall_CVISA_ref.m in the thesis code. The Monte Carlo scheme was implemented by Magnus Wiktorsson.

In terms of computation time, both COS and FGL are by far faster than the Monte Carlo methods, which, on average, take more than a minute for a single price. Unlike FGL and COS, it also becomes slower for high times to maturity τ, because the number of points on each stock price path is proportional to τ. In fact, the slowness of Monte Carlo drove us to restrict the number of parameter combinations to test, so that the computation time could be kept below a reasonable limit. Comparing the two other methods, FGL is generally slower than COS. The difference is, however, small, as even the slowest computation of a FGL price finishes within ≈ 0.7 seconds.

	median	mean	max	stdev
FGL	1.07e-15	1.76e-15	2.76e-14	2.81e-15
COS	2.95e-09	1.68e-08	1.27e-06	8.67e-08

Table 6.3: Black-Scholes model: Relative FGL and COS error. The FGL error is by several magnitudes smaller than the COS error. Furthermore, it varies less, as measured by the lower standard deviation.

	median	mean	max	stdev
FGL	1.13e-15	2.06e-15	3.68e-14	3.53e-15
COS	9.81e-10	5.92e-08	3.06e-05	7.10e-07

Table 6.4: Black-Scholes model with discrete dividends: Relative error of FGL and COS. Similar to the Black-Scholes model, the COS method exhibits a higher error than the FGL method. This is especially true for the maximal error. Additionally, the COS error fluctuates more, as indicated by the higher standard deviation.

	median	mean	max	stdev
FGL	0.00e+00	3.60e-05	2.31e-03	2.01e-04
COS	0.00e+00	4.55e-05	2.44e-03	2.30e-04

Table 6.5: Heston model: Relative error of FGL and COS. In contrast to the Black-Scholes models, FGL and COS perform similarly well here.

	median	mean	max	stdev
FGL	0.00e+00	3.74e-05	2.00e-03	1.83e-04
COS	0.00e+00	4.12e-03	6.83e-01	2.77e-02

Table 6.6: Bates model: Relative FGL and COS error. The FGL error is smaller than the COS error, but the difference is not as big as with the Black-Scholes models. The exception is the maximal error, which is fairly high for the COS method.

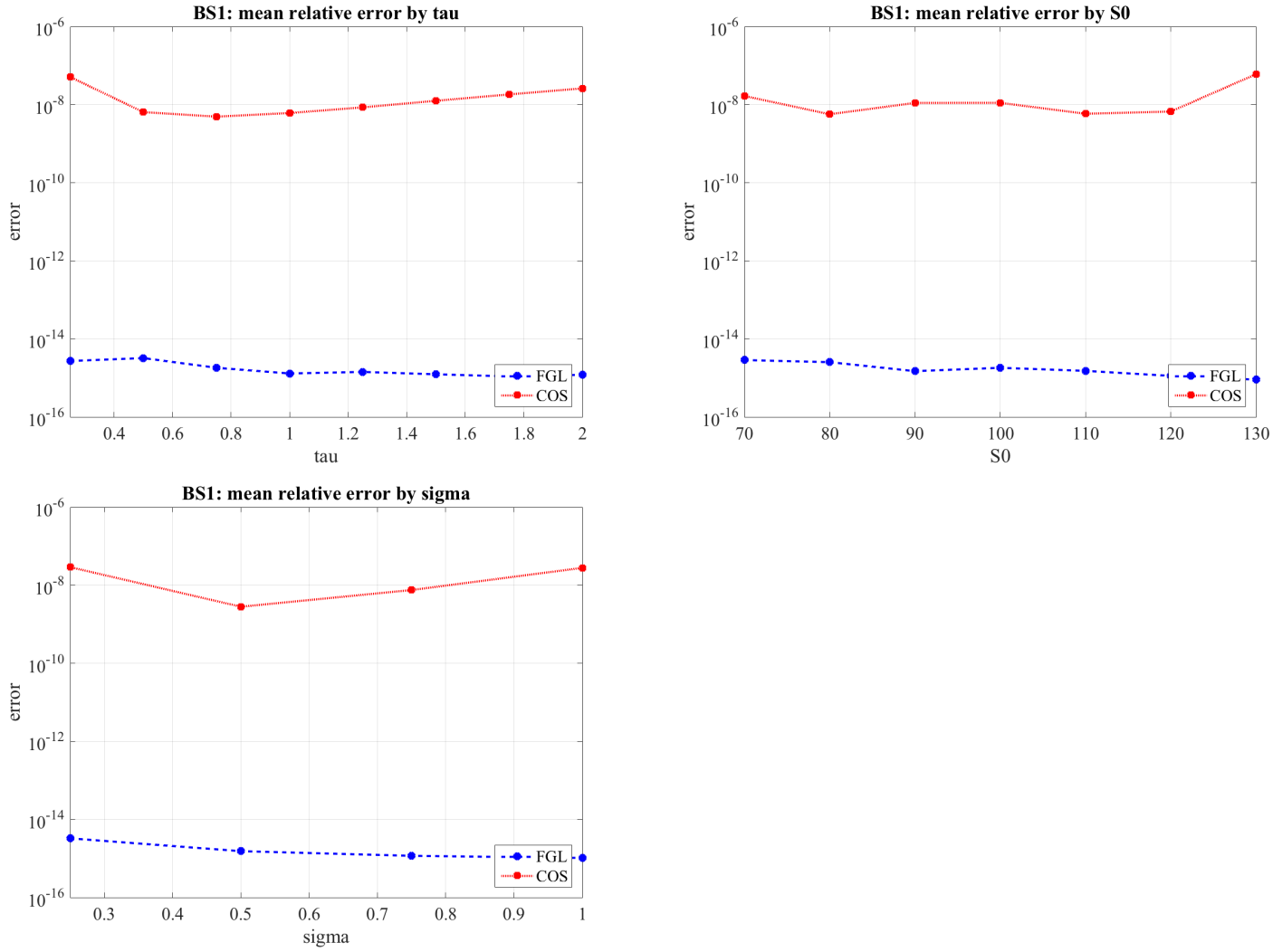
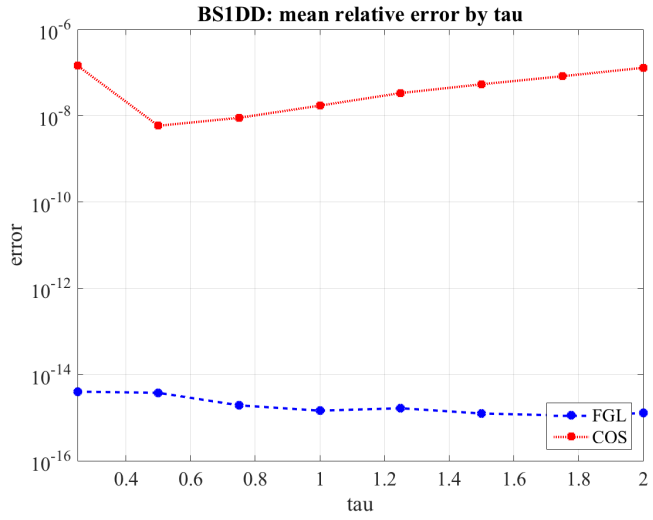
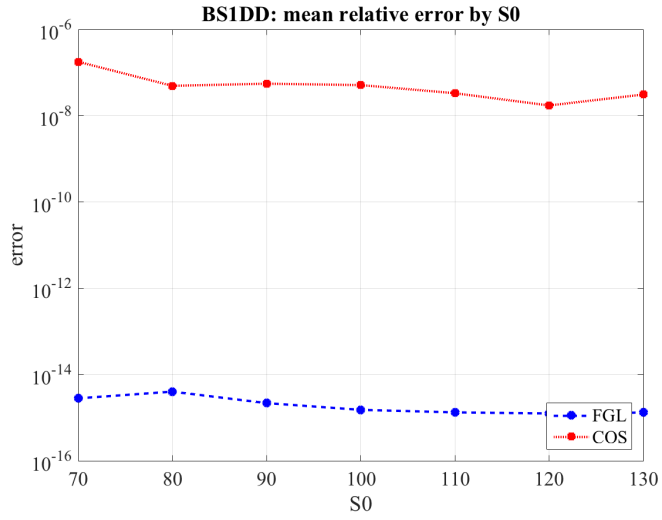
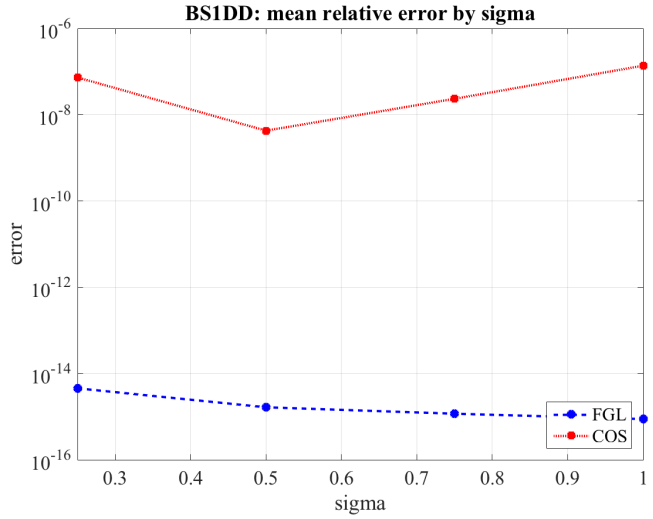
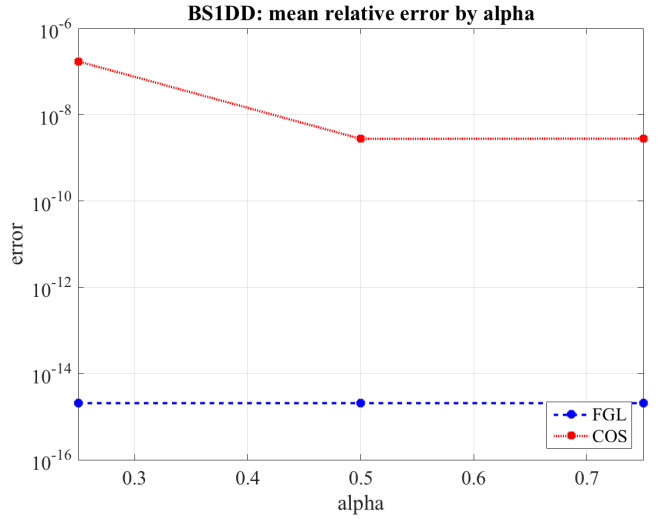


Figure 6.1: Mean relative error of FGL and COS by τ, S_0, σ in the Black-Scholes model with logarithmic scale. The FGL error is always by several magnitudes smaller than the COS error. Furthermore, it is fairly insensitive to the model parameters, whilst the COS error is minimal for low-to-medium maturities and volatilities.









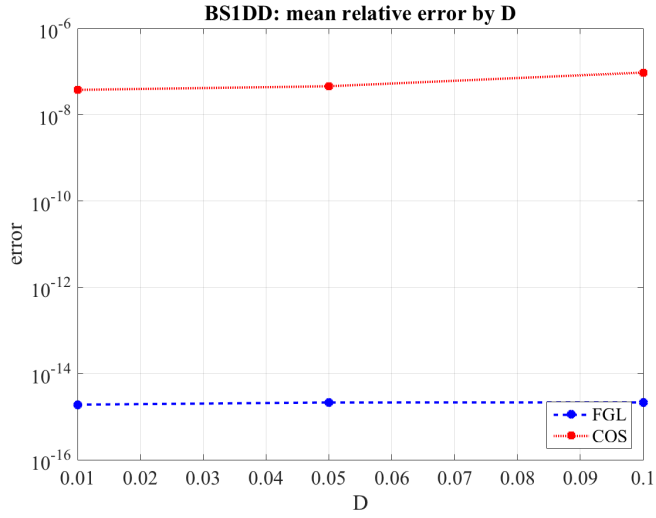


Figure 6.2: Mean relative error of FGL and COS as a function of $\tau, S_0, \sigma, \alpha$ and D in the Black-Scholes model with discrete dividends. As with the Black-Scholes model, the FGL error is far smaller than the COS error and is much less subject to changes in the model parameters. The FGL error decreases slightly as τ, S_0 and σ increase, and is rather insensitive to changes in α and D. The COS error decreases strongly for maturities up to 0.5 and for volatilities of 0.5, then increases again for higher values of each parameter. It also generally decreases as the stock price and α increase.

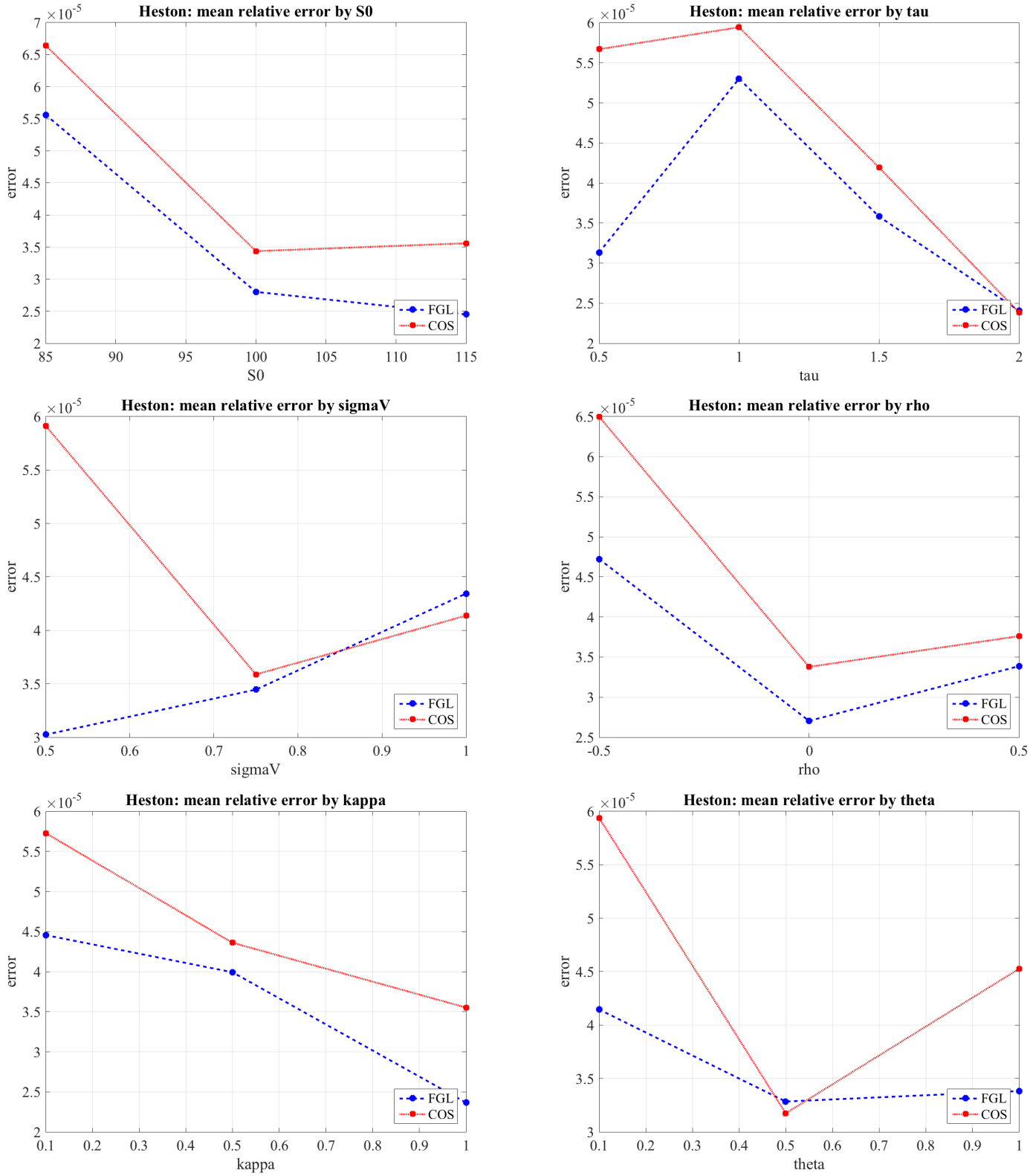


Figure 6.3: Mean relative error in the Heston model by $S_0, \tau, \sigma_V, \rho, \kappa$ and θ, linearly scaled. The dependence on the initial volatility V_0 is omitted because this parameter is always equal to σ_V, as stated in Section 6.1.2. FGL is generally more accurate than COS, although the relative difference is much smaller than in the Black-Scholes models. Both errors decrease as S_0 and κ increase, and are small for medium values of σ_V, ρ and θ. They are maximal for medium time-to-maturity.

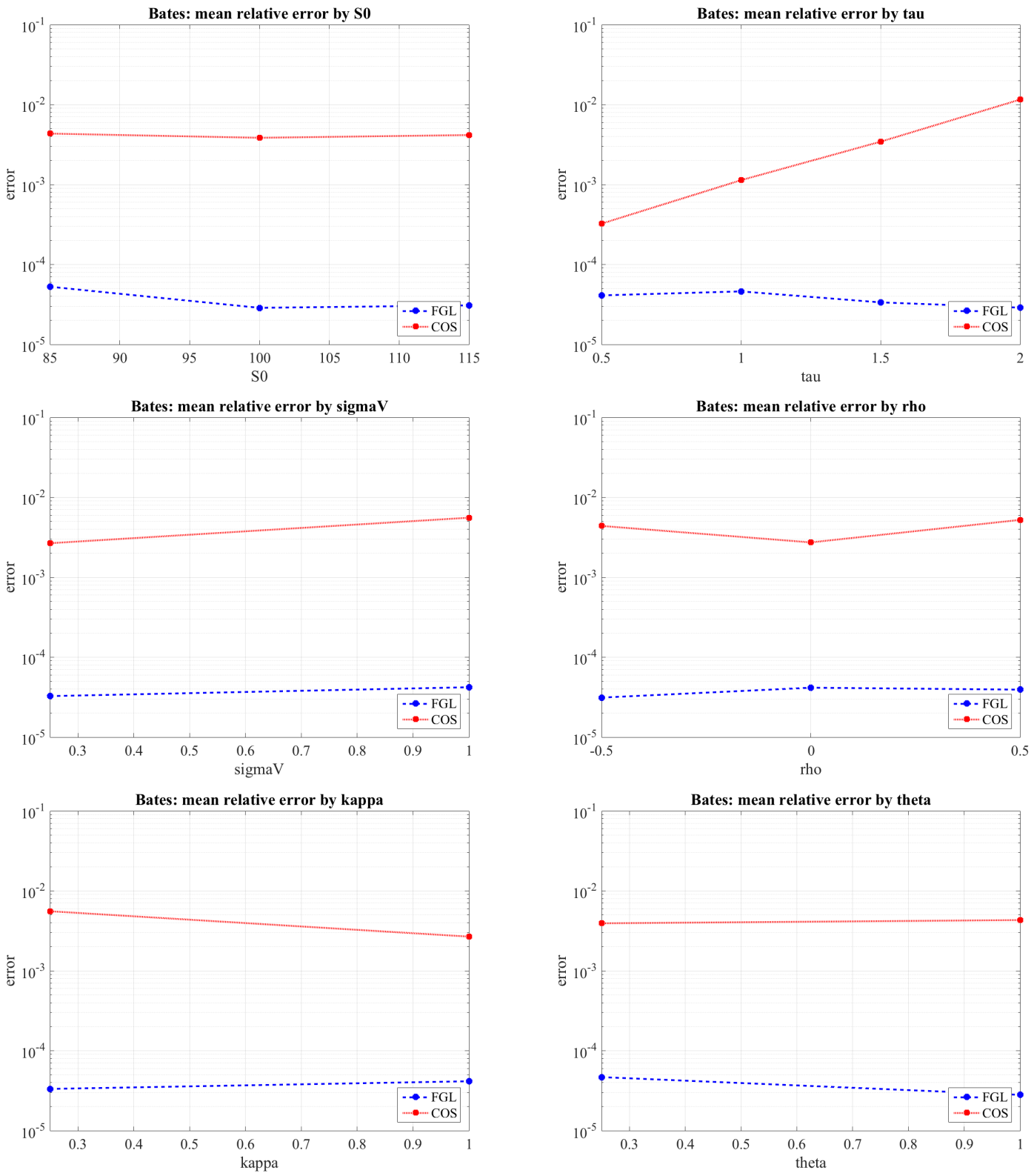


Figure 6.4: Mean relative error of FGL and COS in the Bates model by $S_0, \tau, \sigma_V, \rho, \kappa$ and θ. The FGL error is generally smaller than the COS error, and less sensitive to changes in the model parameters. The COS error increases strongly as τ does.

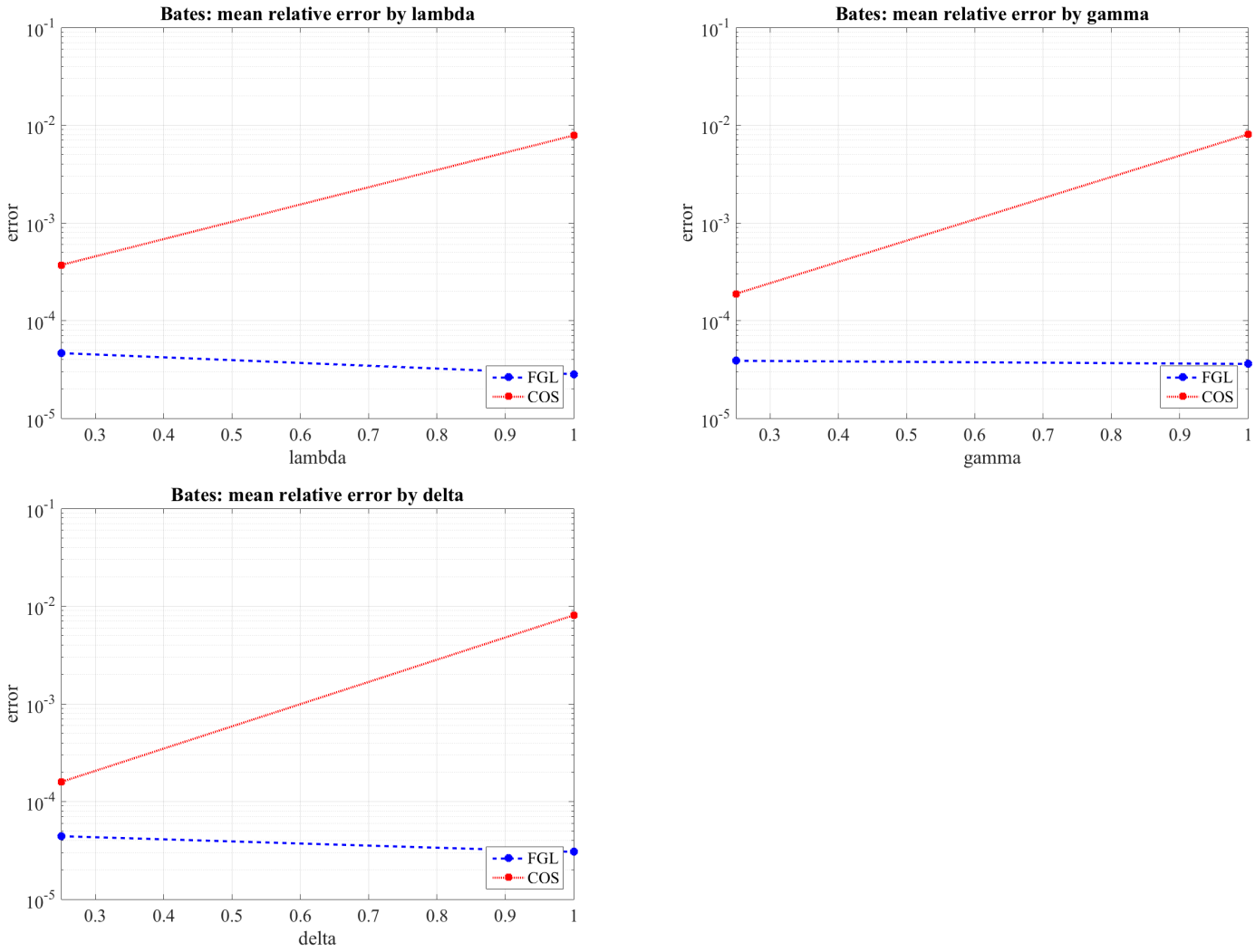


Figure 6.5: Mean relative error of FGL and COS in the Bates model dependent on λ, γ and δ. The COS error increases strongly with all parameters, while the FGL error is rather insensitive.

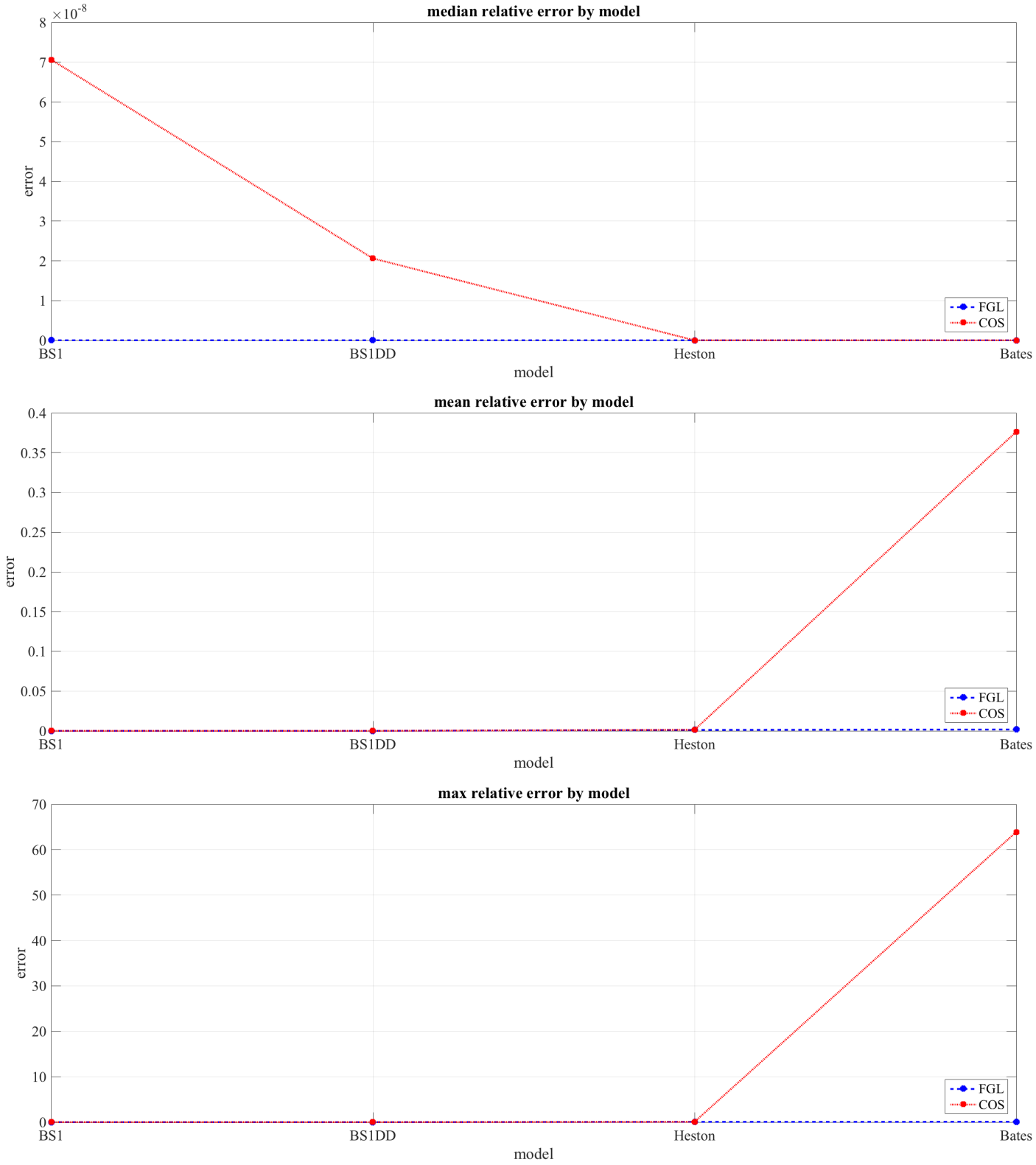


Figure 6.6: Mean relative error of FGL and COS by tested model. The mean and maximal COS error is much greater than the FGL error in the Bates model.

	median	mean	max	stdev
ref	8.54e-04	9.13e-04	2.70e-03	1.70e-04
FGL	1.80e-02	1.82e-02	2.51e-02	1.04e-03
COS	6.69e-05	8.38e-05	1.41e-03	9.20e-05

Table 6.7: Computational time of the reference method (BS formula), FGL and COS in the Black-Scholes model in seconds. The most apparent difference is that FGL is generally slower than COS or the BS formula. Even so, all methods are fairly fast, with the overall longest computation time being less than 3/100 seconds.

	median	mean	max	stdev
ref	8.60e-04	9.28e-04	2.90e-03	2.08e-04
FGL	1.81e-02	1.88e-02	3.60e-02	2.23e-03
COS	3.48e-04	3.88e-04	3.24e-03	1.37e-04

Table 6.8: Computational time of the reference method (BS formula), FGL and COS in the Black-Scholes model with discrete dividends. The results are similar to the ordinary Black-Scholes model, with FGL being the slowest method.

	median	mean	max	stdev
ref	6.19e+01	8.81e+01	2.73e+02	8.31e+01
FGL	1.97e-02	2.07e-02	7.04e-01	2.20e-02
COS	3.32e-04	3.49e-04	1.46e-03	8.32e-05

Table 6.9: Computational time of the reference method (Monte Carlo), FGL and COS in the Heston model. Both FGL and COS are by far faster than Monte Carlo, which takes on average more than a minute per computed price. By contrast, even the slowest computation of a FGL or COS price finishes within ≈ 0.7 seconds. Comparing only FGL and COS with one another, one again notices the same pattern as in the previous models: COS is faster than FGL.

	median	mean	max	stdev
ref	2.16e+01	6.71e+01	1.15e+03	8.45e+01
FGL	5.13e-02	4.29e-02	7.90e-02	1.31e-02
COS	2.87e-04	2.99e-04	1.80e-03	5.50e-05

Table 6.10: Computational time of the reference method (Monte Carlo), FGL and COS in the Bates model. The results are similar to the Heston model: Monte Carlo is by far the slowest method, while COS and FGL are very fast, with COS being slightly faster.

7 Summary

We conclude by discussing some of the advantages and disadvantages of each method, including Monte Carlo.

The main advantage of the FGL and COS methods compared to Monte Carlo is their speed: While the computation of a Monte Carlo price can take several minutes, the slowest computation of a FGL / COS price was ≈ 0.7 seconds. Typically, this time is much shorter, though, with ≈ 0.02 seconds per FGL price and ≈ 0.003 seconds for a COS price. This is clearly a big computational gain, and might be especially useful in areas where option pricing needs to be done very frequently, such as the calibration of model parameters.

Furthermore, the Monte Carlo computational time increases with time to maturity, as longer stock price and volatility paths need to be simulated. FGL and COS do not have this dependency. However, the relative abundance of cheap computational power might render this advantage less important. This is especially true since Monte Carlo methods can be parallelized to a large extent.

A major disadvantage of both Fourier methods - FGL and COS - is that they are not as straightforward to understand and implement as Monte Carlo methods. This can make adjustments of the pricing procedure to new products and models difficult: FGL requires the Fourier transform of the payoff function to be known, and the COS method involves the coefficients in the Fourier cosine series expansion of the payoff function. Additionally, both methods need the moment-generating function of the log-stock price to be known. This is not the case with the Monte Carlo method, where an existing pricing routine can be easily extended to price other products.

The method-specific parameters of Monte Carlo are the size of the time grid for the stock price and volatility paths, and the number of paths N_S. At least with respect to the second parameter, the effect on the error is also easy to predict, as it decreases proportional to $1/\sqrt{N_S}$. By contrast, the method-specific parameters of the Fourier methods caused a lot of tricky issues in our experiments: The number of weights in the Gauss-Laguerre quadrature in FGL, n, should generally be set sufficiently high to allow for an accurate integration. However, when it is too high, numerical overflow occurs, especially in the Bates model due to the nested exponentials in the moment-generating functions. This can be solved by using specialized classes to handle extremely large numbers, but from our experience, this slows down computation enormously.

As for COS, we found that the error behaves rather irregularly depending on the choice of L, $[a,b]$ and N. Furthermore, the recommendation of Fang and Oosterlee to use $[a,b] = \left[c_1 - L\sqrt{c_2 + \sqrt{c_4}}, c_1 + L\sqrt{c_2 + \sqrt{c_4}}\right]$ or $[a,b] = \left[c_1 - L\sqrt{c_2}, c_1 + L\sqrt{c_2}\right]$ caused very inaccurate prices in some cases. Furthermore, it introduces the need to compute the cumulants, which can be quite intricate in the Heston and Bates model. We mostly resolved these problems by appropriate choices of $[a,b]$ and optimization of L, but we remain cautious as to whether this choice holds for model parameters or models other than the tested ones.

In terms of accuracy, both methods generally show good performance, typically with several digits of accuracy. A notable exception thereof is the Bates model, in which the relative COS error gets quite high. We also noted that the mean relative COS error increases as the number and size of jumps do, showing that the COS method does not handle the jump aspect of that model well. For the FGL method, we find overall low sensitivity to this aspect and to changes in model parameters in general.

Overall, we appreciate the swiftness of the Fourier methods, but find that the method-specific parameters and difficulty in extending these methods to other pricing problems are significant caveats. A more thorough investigation of optimal choices of model parameters and an even more comprehensive test coverage would help.

References

[Abramowitz and Stegun, 1972] Abramowitz, M. and Stegun, I. A., editors (1972): Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. Dover. p. 890.

[Barndorff-Nielsen, 1997] Barndorff-Nielsen, O. E. (1997): Processes of Normal Inverse Gaussian Type, Finance and Stochastics, 2, 41–68.

[Bates, 1996] Bates, D. (1996): Jumps and Stochastic Volatility: the Exchange Rate Processes Implicit in Deutschemark Options, Review of Financial Studies, 9, 69–107.

[Björk, 2009] T. Björk (2009): Arbitrage Theory in Continuous Time, 3rd ed., Oxford University Press.

[Black and Scholes, 1973] Black, F. and Scholes, M. (1973): The Pricing of Options and Corporate Liabilities, Journal of Political Economy, 81, 637–659.

[Boyle, 1997] Phelim Boyle, Mark Broadieb, Paul Glasserman (1997): Monte Carlo methods for security pricing. Journal of Economic Dynamics and Control ELSEVIER 21, 1267-1321

[Broadie and Kaya, 2004] Mark Broadie,¨Özgür Kaya (2006): Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes, Operations Research,Volume 54, Issue 2,pages 217-331, http://dx.doi.org/10.1287/opre.1050.0247.

[Carr and Wu, 2003] Carr, P. and Wu, L. (2003): The Finite Moment Log stable Process and Option Pricing, Journal of Finance, 58, 753–778.

[Carr et al., 2002] Carr, P., Geman, H., Madan, D., and Yor, M. (2002): The Fine Structure of Asset Returns: An Empirical Investigation, The Journal of Business, 75, 305–332.

[Cox et al., 1985] Cox, J.C., Ingersoll, J.E., and Ross, S.A., (1985): A theory of the term structure of interest rates, Econometrica, Vol.53, No.2, 385-407.

[Fang and Oosterlee, 2008] F. Fang and C.W. Oosterlee (2008): A novel pricing method for European options based on Fourier-cosine series expansions, SIAM J. Sci. Comput. 31.

[Heston, 1993] Heston, S. L. (1993): A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, Review of Financial Studies, 6, 327–343.

[Johnson et al., 1994] Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994): Continuous Univariate Distributions, Volume 2, Second Edition, Wiley, New York.

[Kahl and Lord, 2006] Lord, Roger; Kahl, Christian (2006): Optimal Fourier Inversion in Semi- analytical Option Pricing, Tinbergen Institute Discussion Paper, No. 06-066/2

[Madan and Seneta, 1990] Madan, D. B. and Seneta, E. (1990): The Variance Gamma (VG) Model for Share Market returns, Journal of Business, 63, 511–524.

[Merton, 1976] Merton, R. C. (1976): Option Pricing When the Underlying Stock Returns are Discontinuous, Journal of Financial Economics, 5, 125–144.

[Ruijter and Oosterlee, 2015] M. J. Ruijter and C. W. Oosterlee (2015): Notes on the Benchop implementations for the COS method, http://www.it.uu.se/research/project/compfin/benchop/COSnotes.pdf

[von Sydow et al., 2015] Lina von Sydow, Lars Josef Höök, Elisabeth Larsson, Erik Lindström, Slobodan Milovanović, Jonas Persson, Victor Shcherbakov, Yuri Shpolyanskiy, Samuel Sirén, Jari Toivanen, Johan Waldén, Magnus Wiktorsson, Jeremy Levesley, Juxi Li, Cornelis W. Oosterlee, Maria J. Ruijter, Alexander Toropov, Yangzhang Zhao (2015): BENCHOP – The BENCHmarking project in option pricing. International Journal of Computer Mathematics Vol. 92, Iss. 12. http://www.tandfonline.com/doi/abs/10.1080/00207160.2015.1072172

[Wiktorsson, 2015] Wiktorsson, Magnus (2015): Notes on the Benchop implementations for the Fourier Gauss Laguerre FGL method, http://www.it.uu.se/research/project/compfin/benchop/FGLnotes.pdf