INSTITUTO
UNIVERSITÁRIO
DE LISBOA

UNIVERSIDADE
DE LISBOA

The Kristensen and Mele (2011) approach to approximating derivative prices in continuous-time models

Sandro Marcos Fernandes Gaspar

Master in Financial Mathematics

Supervisors:
PhD, José Carlos Gonçalves Dias, Full Professor, ISCTE Business School

PhD, Fernando Correia da Silva, Integrated Researcher, ISCTE Business School

October, 2025

Department of Finance

Department of Mathematics

The Kristensen and Mele (2011) approach to approximating derivative prices in continuous-time models

Sandro Marcos Fernandes Gaspar

Master in Financial Mathematics

Supervisors:
PhD, José Carlos Gonçalves Dias, Full Professor, ISCTE Business School

PhD, Fernando Correia da Silva, Integrated Researcher, ISCTE Business School

October, 2025

Acknowledgments

I would like to sincerely thank both my supervisors, Professor Doutor José Carlos Gonçalves Dias and Doutor Fernando Correia da Silva, for all their knowledge, guidance, and constant support throughout this journey. Their expertise, availability, and assistance, especially in the moments when I had more doubts, were crucial for the completion of this work.

I am deeply grateful to my parents, for all their love and support in every step throughout my life. Their patience, encouragement, and belief in me have been without a doubt the foundation of my achievements, and I am sure I would not be where I am today without them.

I am also very thankful to all my friends and my girlfriend, for being with me and giving me strength in the most challenging times. Their companionship and encouragement made the most challenging and stressful moments more manageable and allowed me to move forward.

Finally, I would also like to thank all those I have met along the way who, in one way or another, had a positive impact on my life and helped me reach this milestone.

Resumo

Em 2011, *Kristensen e Mele* desenvolveram um método de aproximação para modelos de avaliação de opções europeias que não admitem soluções em forma fechada. Este método baseia-se na utilização de um modelo auxiliar mais simples, em torno do qual o modelo principal, que não apresenta solução fechada, é expandido e, através desta expansão e eventuais calibrações, obtém-se uma solução aproximada para o preço da opção, cuja precisão melhora progressivamente e dependendo do modelo auxiliar escolhido.

Este método distingue-se por ser simples de implementar e por permitir obter aproximações rápidas e de elevada precisão, para além da sua transversalidade, uma vez que esta metodologia pode ser aplicada em diferentes contextos, tais como no cálculo dos *Gregos* ou na avaliação de opções assumindo volatilidade constante ou estocástica.

Esta dissertação tem como objetivo estudar detalhadamente o método de aproximação desenvolvido por *Kristensen e Mele* (2011), fazendo uma breve comparação literária com outros métodos bastante utilizados e, de seguida, realizar uma análise dos resultados obtidos da sua implementação face aos benchmarks escolhidos, de forma a avaliar o seu desempenho e a compreender como este método se destaca relativamente às suas alternativas.

Adicionalmente, foi também implementada uma função de calibração sugerida pelos autores para aumentar a precisão dos resultados e avaliar se esta permite efetivamente melhorar a performance relativamente aos resultados obtidos na sua publicação original e, contrariamente aos autores, este método foi aplicado ao modelo padrão de variância com elasticidade constante, visto que originalmente foi aplicado com um processo de reversão à média.

Palavras-chave: Avaliação de Opções, Método de *Kristensen e Mele*, Modelo de *Black, Scholes e Merton*, Modelo de *Heston*, Modelo *CEV*, Volatilidade Estocástica.

Classificação JEL: C63, G13

Abstract

In 2011, *Kristensen and Mele* developed an approximation method for European-style option pricing models that do not admit closed-form solutions. For that, this method is based on the use of a simpler auxiliary model, around which the main model, which does not offer these closed-form solutions, is expanded and, through this expansion and eventual calibrations, a closed-form approximate solution for the option price is obtained, whose accuracy progressively improves depending on the chosen auxiliary model.

This method stands out for being simple to implement and for allowing fast and highly accurate approximations, in addition to its versatility, since the methodology can also be applied to various contexts, such as the calculation of *Greeks* or in option pricing under constant or stochastic volatility settings.

This dissertation aims to study in detail the approximation method developed by *Kristensen and Mele* (2011), doing a brief literary comparison of other widely used methods and, afterwards, performing an analytical review of the results obtained from its implementation against the chosen benchmarks, in order to evaluate its performance and understand how this method distinguishes itself from existing alternatives.

Additionally, it was also implemented a calibration function suggested by the authors in order to increase its precision and evaluate if this effectively improves the performance compared to the results obtained in their original paper and, unlike the authors, this method was applied to the standard *Constant Elasticity of Variance* model, since in the original work it was applied with a mean-reverting process.

Keywords: Options pricing, *Kristensen and Mele* method, *Black, Scholes and Merton* model, *Heston* model, *CEV* model, Stochastic Volatility.

JEL Classification: C63, G13

Contents

List of Tables

List of Figures

Introduction

In the last decades, there has been an increasing demand for new models related to pricing and hedging that cover multiple derivatives contracts and, due to the intrinsic complexity of the financial markets, the continuous-time models have risen in popularity. These models offer a more realistic modeling of the financial frameworks, as the parameters change continuously.

To take advantage of these models, which, due to their high complexity, do not offer closed-form solutions, practitioners can approach these problems via two typical applications. The first application uses a numerical solution to partial differential equations (*PDEs*), e.g., obtained via finite-differences, which work by discretizing the domain of a *PDE* into a finite number of intervals, where the terms are then approximated using finite-difference formulas, which express derivatives as algebraic differences between function values at neighboring points, and, in turn this method transforms the original *PDE* into a set of equations easier to solve numerically. Another possible numerical approach is the *Fourier* inversion theorem, which states that a *PDE* can be converted into the frequency domain using the *Fourier* transform, to solve it more easily, and then use the inverse transformation to recover the solution back to the original domain. This method can provide fast and accurate calculations of vanilla option prices in models with analytically available characteristic functions, however, it has limited model flexibility and can create some numerical instabilities, as documented by *Lord and Kahl* (2007).

Additionally, another way to solve numerical solutions is an approach designated by the tree method, in which the time and space of each option is discretized, and throughout each time step, the stochastic processes evolve to a finite number of states, similar to the branches of a tree. Some famous models have adopted variations of this method, such as *Schwartz* (1977), *Hull and White* (1990), *Scott* (1997), and *Figlewski and Gao* (1999). Each model adapts the tree framework to accommodate different financial products and market dynamics, which makes it more suitable for specific types of derivatives, as *Schwartz* (1977) historically proposed one of the earliest approaches resembling a binomial tree method for valuing American-style options on dividend paying stocks, which do not have closed form solutions, and also analyzed the optimal strategies for exercising these derivatives; *Hull and White* (1990) suggested an explicit finite difference method to value derivative securities, which is related to tree structures and can be used to value derivative securities, e.g. bonds and bond options, depending on single state variables or, by extension, for several state variables; *Scott* (1997) developed a jump-diffusion model with stochastic volatility and interest rates for pricing European-style stock options and, while using mainly *Fourier* inversion methods, the author also addressed approaches using tree-methods; and *Figlewski and Gao* (1999) proposed the adaptive mesh model for option pricing, which greatly improves the accuracy and computational costs by constructing a lattice-based valuation model, this way refining the traditional tree method and standard finite difference approaches by adjusting dynamically the different parts of the tree.

The second typical application is the use of *Monte Carlo* simulations, as suggested by *Boyle* (1977). Although both applications are widely used, both have limitations, since finite-differences method require solving large system of equations, which may introduce rounding and discretization errors, which will be more demanding for high-accuracy solutions, and *Monte Carlo* simulations, when simulating a vast number of paths over time, ultimately, will also create an exchange between accuracy and computational power, especially in more complex situations.

The main objective of this thesis is to review the *Kristensen and Mele* (2011) approach (from now on designated as *KM*) to approximating derivative prices in these continuous-time models and compare it to the *Heston* (1993) model (hereafter designated as only *Heston*). The purpose of this approach is to be able to price models without a closed-form solution and, for that, the authors chose an auxiliary model, which is simpler than the main model (also known as the model of interest), such as *Black and Scholes* (1973) and *Merton* (1973) (hereafter designated as *BSM*), and that is available in closed-form, and then, expand the unknown price of the main model around the auxiliary one, with form of a conditional expectation under the risk-neutral probability, through a *Taylor* series expansion. Additionally, this dissertation also attempts to extend the *KM* approximation method to the *Constant Elasticity of Variance* model (also known as *CEV*) introduced by *Cox* (1975), however, because the original paper did not clearly describe how it was implemented, since it appears to suggest that it follows a *CEV* model with a mean-reverting process, the method expanded here is fundamentally different from the authors, even though it follows the same general methodology.

It is also worth noting that one of the major strengths of *KM's* approximation method is its flexibility and high adaptability to different model conditions. As the framework relies on choosing an auxiliary model that admits a closed-form solution, in case that the model of interest is more complex, the auxiliary model chosen can also be more sophisticated, assuming that it still allows for a closed-form solution. In this way, the choice of the auxiliary model should directly correlate to the accuracy and general capacity of the approximation. For example, when analyzing a main model with dividend paying assets, the auxiliary model must also incorporate dividends in order to mimic the key features of the true model.

A simpler auxiliary model such as *BSM* keeps the method computationally light but may limit its accuracy, while a richer auxiliary model, e.g., with stochastic volatility or jumps, allows the method to handle more advanced dynamics without sacrificing accuracy. This adaptability explains why the *KM* method can be applied to such a wide range of continuous-time models for which exact solutions are otherwise unavailable.

To implement this approach, the *Python* programming language was used with the *Sympy* library, due to its vast application in symbolic differentiation and algebraic manipulation. As *KM's* approximation method becomes increasingly more complex for higher-order expansions, an entirely numerical implementation was considered less suitable in this case, hence, the use of symbolic computation was chosen, as it allows for the derivation of analytical expressions for each term in the expansion, which in turn improves both its efficiency and accuracy.

Moreover, to compare the accuracy of the *KM* approximation applied to the *CEV* model, in the percentage error of the *CEV* analysis, the call prices results were compared against a benchmark obtained from the standard *CEV* model, as this is a method that offers closed-form solutions, making it appropriate for this situation. In order to implement this benchmark, it was used the definition provided by *Schroder* (1989) as the elasticity parameter γ was set to satisfy the condition $\gamma < 1$.

Additionally, in order to compare the results obtained from the implementation of *KM* method applied to the *Heston* model, the *Madan and Carr* (1999) fast *Fourier* transform method was used as a benchmark to evaluate the performance of *KM's* method under a stochastic volatility model of interest. The *Madan and Carr* (1999) fast *Fourier* transform is a numerical approach that assumes the characteristic function of the risk-neutral density is known analytically, which makes this assumption highly reliable and an accurate alternative for pricing options.

CHAPTER 1

Literature review

1.1 Processes definitions

In order to review the *KM* model, it is essential to first have an understanding on the foundations of stochastic processes and concepts for building such models. This includes the *Martingale* processes, the *Wiener or Brownian* motion processes, the *Feynman-Kac* theorem, and the *Taylor* series, since these elements provide some of the necessary background for interpreting and evaluating the model's assumptions, behavior, and their implications.

1.1.1 Martingale processes

A *Martingale* process is a stochastic process, i.e., a random process, in which the expected future value is equal to its present value, given all current available information.

The stochastic process of a *Martingale* can be defined as M_k for all $k \geq 0$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k\geq 0}, \mathbb{P})$, with respect to the probability measure $\mathbb{P}$ and the filtration $(\mathcal{F}_k)_{k\geq 0}$, assuming it follows the below conditions:

1) $\mathbb{E}[|M_k|] < \infty$ for all $k \geq 0$, i.e., M is integrable;
2) $\mathbb{E}[M_{k+1}|\mathcal{F}_k] = M_k$, where $\mathcal{F}_k$ is the filtration and represents all the current information known in the present. Which means that M_{k+1} only depends on the information known until the moment k.

This implies that a *Martingale* process is an adapted stochastic process whose conditional expected value does not change over time with future observations, meaning that, on average, the process will not be influenced, neither positively nor negatively, so it shows no predictable trend.

These properties make *Martingales* especially valuable in financial modeling since, in efficient and arbitrage-free markets, asset price changes are assumed to reflect all current and available information, which behave like a "fair game" over time, resembling the key concept of a *Martingale* process.

1.1.2 Wiener processes

The *Wiener* processes, also commonly known as standard *Brownian* motions, are continuous-time stochastic processes that models random and unpredictable movement over time, which is a property that can be observed in the randomness present in financial markets.

The *Wiener* processes can be defined as W_t for all $t \geq 0$, defined in a filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$ and satisfying the following properties:

1) $W_0 = 0$, almost surely, i.e., $\mathbb{P}(W_0 = 0) = 1$;
2) The paths of W are also almost surely continuous over time t;
3) W has independent and stationary increments, which means that for any $0 \leq s < t$, the increment $W_t - W_s$ is independent of the process history up to time s;
4) The distribution of $W_t - W_s$ depends only on the length of the time interval $t - s$, rather than the individual value of t and s;
5) The increments follow a normal distribution, i.e., $W_t - W_s \sim \mathcal{N}(0, t - s)$, meaning that each increment will have a mean equal to 0 and variance equal to $t - s$.

Due to these properties, *Wiener* processes are widely used in the modelling of financial markets, however, they have the quite strict limitation of assuming that their paths are always smooth and continuous over time, therefore, when new dynamics are introduced in the markets, such as sudden jumps or discontinuities, the standard *Wiener* processes are no longer suitable for capturing the observed randomness, and instead, other approaches must be used, such as jump-diffusion models, which are designed to model paths with sharp and instantaneous shifts more accurately.

1.1.3 Feynman-Kac theorem

The *Feynman-Kac* theorem provides a fundamental connection between stochastic processes and *PDEs*, since it shows that the solution to linear parabolic *PDEs* can be represented as the expected value of a stochastic process, which essentially allows the use of probabilistic methods, such as simulating random paths of an *Itô* process in order to solve *PDEs*.

Let $u(x,t): \mathbb{R} \times [0,T] \to \mathbb{R}$ be the solution to a linear parabolic *PDE* with the form:

$$\frac{\partial u}{\partial t} + \mu(x,t)\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 u}{\partial x^2} - V(x,t)u + f(x,t) = 0, \tag{1.1}$$

for all $x \in \mathbb{R},\ t \in [0,T]$, and subject to the terminal condition $u(x,t) = \psi(x)$, where μ, σ, ψ, V, and f are known functions.

Then, under these conditions, the solution $u(x,t)$ can be represented as the following expected value:

$$u(x,t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T V(X_r,r)dr}\psi(X_T) + \int_t^T e^{-\int_t^u V(X_r,r)dr} f(X_u,u)du \middle| X_t = x\right], \tag{1.2}$$

where X_u follows an *Itô* diffusion process given by the stochastic differential equation:

$$dX_u = \mu(X_u,u)du + \sigma(X_u,u)dW_u^{\mathbb{Q}}. \tag{1.3}$$

This means that, instead of solving the *PDE* analytically, the *Feynman-Kac* theorem is able to simulate the associated stochastic process X_s and compute the expected value numerically. This definition makes this method especially useful when pricing options, since the value of these derivatives can be expressed as the expected discounted payoff under a risk-neutral measure, which interestingly is exactly the expectation that the *Feynman-Kac* theorem describes in equation (1.2).

1.1.4 Taylor series

Lastly, the *Taylor* series, also known as *Taylor* expansions, is another fundamental method that works by approximating functions as an infinite sum of polynomial terms, centered around a point, which in turn allows the chosen function to be represented by derivatives at that specific point.

Therefore, for a function $f(x)$ that is infinitely differentiable at $x = a$, the *Taylor* series expansion around a is defined as:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{1.4}$$

This expansion is quite important, since it allows to study complex functions through simpler polynomial approximations and, assuming that the series converges, when $n \to \infty$ the approximation becomes increasingly accurate, particularly near the point $x = a$. This makes this method especially useful when modelling financial markets, as *Taylor* expansions can be used for approximating non-linear payoffs and for analyzing the model sensitivities, such as for computing the *Greeks* of options.

1.2 Black-Scholes and Merton assumptions

The *BSM* model, developed by *Black and Scholes* (1973) and later extended by *Merton* (1973), is one of the most fundamental concepts in financial mathematics, which serves as one of the foundations for pricing derivatives such as options. Although today it is considered a simpler model which has several well-known limitations, it still remains as a base for more advanced and complex models currently in use.

The model relies on several fundamental assumptions:

1) The underlying asset price must follow a geometric *Brownian* motion, which is modeled using a *Wiener* process with constant drift and volatility;
2) The risk-free interest rate and volatility of the underlying asset are both known and constant over time;
3) The original model defined by *Black and Scholes* (1973) (hereafter designated only as *BS*) assumes that the underlying asset does not pay any dividends throughout the life of the

option, however, *Merton (1973)* later introduced an extension to this model that incorporates a continuous dividend yield;

4) The model does not account for any transaction costs for buying or selling the derivative;
5) Options are only allowed to be exercised at the maturity, meaning that *BSM* is only applicable to European-style options, so it cannot be used to price American-style options.

Based on these assumptions, *BSM* defined the behavior of a dividend-paying underlying asset under the risk-neutral measure as the following stochastic differential equation (hereafter designated as *SDE*):

$$dS(t) = (r - q)\, S(t)dt + \sigma S(t)dW(t), \tag{1.5}$$

where $S(t)$ is the price of a dividend-paying asset at the time t, r is the constant risk-free interest rate, and q is the continuous dividend yield. Additionally, σ is the constant volatility of the asset and $W(t)$ is the standard *Brownian* motion under the risk-neutral probability measure $\mathbb{Q}$.

1.3 Heston model assumptions

While the *BSM* model laid the foundations for modern option pricing, it has the quite strict limitation of always assuming constant volatility, as this assumption fails to reflect the complexity of the real-world financial markets. Over the years, it has been demonstrated through empirical evidence that the asset returns display skewness and kurtosis, deviating from a *Gaussian* distribution, i.e., normal distribution, and, moreover, the assets' volatilities are neither constant nor deterministic, but rather time-varying, and often inversely correlated to the underlying asset price.

To address this limitation various models have been proposed, including one of the most famous developed by *Heston*, where this model extends the classical *BSM* framework by allowing the asset's volatility to oscillate and evolve as a stochastic process, rather than remaining fixed. This means that the volatility is a random and mean-reverting process, reflecting the dynamic behavior observed in financial markets, and, by capturing its randomness, the model captures the market behavior more accurately, especially in pricing options with longer maturities.

It is also worth to note that a notable strength of the *Heston* model, is that the model leverages characteristic functions for option pricing, which by allowing to focus on the characteristic function of the terminal price distribution, rather than the distribution itself, the model is capable of a more efficient approach in the pricing.

Apart from these innovations and capabilities, the model relies on these fundamental assumptions:

1) The underlying asset follows a geometric *Brownian* motion, which implies that the prices evolve randomly according to a *Wiener* process;

2) Unlike the *BSM* model, where the volatility is constant, *Heston* model assumes that it is a stochastic process and fluctuates over time and, under the *Feller* condition, it always remains positive. The instantaneous volatility also follows a *Cox et al.* (1985) (also designated as *CIR*) mean-reverting square-root process, meaning that, even if the volatility can drift from its average value, it will eventually return to the equilibrium;
3) Similar to *BSM*, this model also assumes a constant risk-free interest rate for the duration of the option;
4) The model assumes a constant correlation between the asset price and its volatility, which gives the ability to generate implied volatility smiles and skews, resulting from its capacity to incorporate skewness and kurtosis into option prices. These volatility smiles and skews capture the inverse relationship between the volatility and asset prices, for a practical and straightforward way to analyze;
5) Alongside with *BSM* the model also assumes frictionless market conditions and does not account for sudden jumps or discontinuities in the price process.

Knowing these assumptions, *Heston* defines the stock price $S(t)$ as a stochastic volatility model, with the following *SDE*:

$$\frac{dS(t)}{S(t)} = (r-q)\,dt + \sqrt{v(t)}dW_1(t) \Leftrightarrow dS(t) = (r-q)S(t)\,dt + S(t)\sqrt{v(t)}dW_1(t), \tag{1.6}$$

where the variables remain the same as defined in *BSM*, but, since it is now assumed that the volatility is not a constant, it is implemented as $v(t)$, being the instantaneous return variance.

As mentioned before, it is assumed that the variance $v(t)$ follows a mean-reverting square-root process and has constant elasticity of variance, similarly to the *CEV* model. Under this assumption, the corresponding *SDE* is defined by:

$$dv(t) = \kappa\big(\alpha - v(t)\big)dt + \omega|v(t)|^{\xi}dW_2(t), \tag{1.7}$$

where $W_2(t)$ corresponds to the standard *Brownian* motion under the risk-neutral measure $\mathbb{Q}$, and correlated to $W(t)$, with an instantaneous correlation factor of ρ, and assuming a *CEV* parameter of $\xi > 0$ for the elasticity/diffusion of $v(t)$, which serves to adjust the sensitivity of volatility in relation to the current variance level. Additionally, in regard to the elasticity parameter ξ, in the *Heston* model it is generally assumed a $\xi = 0.5$ and, in the more general *CEV* framework, an elasticity factor of $\xi > 0.5$ usually tends to provide a greater flexibility when capturing market behavior.

Moreover, κ, α, ω are additional variables, with κ representing the mean-reversion rate, which indicates the speed at which $v(t)$ returns to its long-term mean value α, and ω being the volatility of volatility, which indicates the magnitude of random fluctuations that occur in this process.

The payoff of a European-style call option at maturity is given by the following equation:

$$b\big(S(T)\big) = max(S(T) - K, 0), \tag{1.8}$$

which means that the option pays the maximum value between the difference of the underlying asset price $S(t)$ minus the strike price K, or 0.

Let $w(S, v, t)$ be the option price as of time $t \in [0, T]$, with instantaneous variance v. Practically, this means that, once the option reaches its maturity, the option price $w(S, v, t)$ will coincide with its payoff $b\big(S(T)\big)$, independently of v, as the option will either pay or not.

Additionally, in order to better define the processes described in equations (1.6) and (1.7), the *Heston* model also admits the correlation of both *Brownian* motions as:

$$dW_1(t)dW_2(t) = \rho\, dt. \tag{1.9}$$

1.4 Introduction to Kristensen and Mele approach

As previously described, the *KM* approach works by approximating derivative prices in continuous-time models without closed-form solutions, by comparing it to auxiliary models far simpler to compute, essentially, expanding the unknown price of the main model around a base model. Having the auxiliary model defined (in this case *BSM* as it offers closed-form solutions, in order to simplify the introduction to this method), it is now possible to start analyzing the *KM* approach.

To start the expansion, it is possible to use *Itô's* lemma for bivariable functions around the pricing function $w(S, v, t)$, assuming that $S(t) = x$:

$$dw(x, v, t) = \frac{\partial w}{\partial t}dt + \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial v}dv + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(dx)^2 + \frac{1}{2}\frac{\partial^2 w}{\partial v^2}(dv)^2 + \frac{\partial^2 w}{\partial x \partial v}(dx)(dv). \tag{1.10}$$

Therefore, by applying the processes defined in equations (1.6) and (1.7) to the $dw(x, v, t)$ operator in equation (1.10), it yields the infinitesimal generator $\mathcal{L}$[1]:

$$\begin{aligned}\mathcal{L}w(x, v, t) = {} & \frac{\partial w}{\partial t} + (r - q)x\frac{\partial w}{\partial x} + \kappa(\alpha - v)\frac{\partial w}{\partial v} + \frac{1}{2}vx^2\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}\omega^2 v^{2\xi}\frac{\partial^2 w}{\partial v^2} \\ & + \rho\omega\left(v^{\xi + \frac{1}{2}}\right)x\frac{\partial^2 w}{\partial x \partial v},\end{aligned} \tag{1.11}$$

with a drift and diffusion terms of the underlying asset $S(t)$ equal to $(r - q)x\frac{\partial w}{\partial x}$ and $\frac{1}{2}vx^2\frac{\partial^2 w}{\partial x^2}$, respectively. Additionally, the drift and diffusion factor of the volatility $v(t)$ are equal to $\kappa(\alpha - v)\frac{\partial w}{\partial v}$ and $\frac{1}{2}\omega^2 v^{2\xi}\frac{\partial^2 w}{\partial v^2}$, respectively. And lastly, the correlation factor between the dividend paying asset and its volatility equal to $\rho\omega\left(v^{\xi + \frac{1}{2}}\right)x\frac{\partial^2 w}{\partial x \partial v}$.

[1] See the appendix in section A – Calculation of the infinitesimal generator $\mathcal{L}$, for the auxiliary calculations.

Using the *Feynman-Kac* theorem, it is now possible to expand the *PDE* defined in equation (1.11), minus the risk-neutral discounting term $rw(x, v, t)$, subject to the boundary condition $w(S, v, t) = b(S)$, with the aim to prepare the foundations for introducing the base model:

$$0 = \mathcal{L}w(x, v, t) - rw(x, v, t). \tag{1.12}$$

Next, the objective is to develop a closed-form approximation to the price equation $w(x, v, t)$. For that, *KM* proposed using an auxiliar model which can be solved in closed-form, such as *BSM*, since it is the simplest to use in this context.

Since in the *BSM* model the asset price follows a geometric *Brownian* motion, the equation (1.12) can be rewritten as:

$$0 = \mathcal{L}_0 w^{bs}(x, t; \sigma_0) - rw^{bs}(x, t; \sigma_0), \tag{1.13}$$

where the condition $w^{bs}(x, t; \sigma_0) = b(x)$ and the infinitesimal operator $\mathcal{L}_0$ remain equal, but, because the *BSM* model assumes constant volatility, then $v = \sigma_0^2$, and thus, assuming that the price function w does not depend on the volatility v, all the partial derivatives of w with respect to v vanish, namely, $\frac{\partial w}{\partial v}, \frac{\partial^2 w}{\partial v^2}$, and $\frac{\partial^2 w}{\partial x \partial v}$ are zero.

Naturally, it can also be assumed that the approximation error introduced by using *BSM* to price the option under of the model of interest is given by:

$$\Delta w(x, v, t; \sigma_0) = w(x, v, t) - w^{bs}(x, t; \sigma_0), \tag{1.14}$$

which by rearranging terms becomes:

$$w(x, v, t) = \Delta w(x, v, t; \sigma_0) + w^{bs}(x, t; \sigma_0). \tag{1.15}$$

Following this, *KM* proceed by subtracting equation (1.13) from the equation (1.12), which provides the estimated error for introducing *BSM* to the pricing:

$$\begin{aligned} 0 &= \mathcal{L}w(x, v, t) - rw(x, v, t) - [\mathcal{L}_0 w^{bs}(x, t; \sigma_0) - rw^{bs}(x, t; \sigma_0)] \\ &= \mathcal{L}w(x, v, t) - rw(x, v, t) - \mathcal{L}_0 w^{bs}(x, t; \sigma_0) + rw^{bs}(x, t; \sigma_0), \end{aligned} \tag{1.16}$$

in which, by substituting in the equation (1.16) the new definition of $w(x, v, t)$ in equation (1.15), it is easily simplified to the below equation:

$$\begin{aligned} 0 &= \mathcal{L}[\Delta w(x, v, t; \sigma_0) + w^{bs}(x, t; \sigma_0)] - r[\Delta w(x, v, t; \sigma_0) + w^{bs}(x, t; \sigma_0)] \\ &\qquad - \mathcal{L}_0 w^{bs}(x, t; \sigma_0) + rw^{bs}(x, t; \sigma_0) \\ &= \mathcal{L}\Delta w(x, v, t; \sigma_0) + \mathcal{L}w^{bs}(x, t; \sigma_0) - r\Delta w(x, v, t; \sigma_0) - rw^{bs}(x, t; \sigma_0) \\ &\qquad - \mathcal{L}_0 w^{bs}(x, t; \sigma_0) + rw^{bs}(x, t; \sigma_0) \\ &= \mathcal{L}\Delta w(x, v, t; \sigma_0) - r\Delta w(x, v, t; \sigma_0) + (\mathcal{L} - \mathcal{L}_0) w^{bs}(x, t; \sigma_0). \end{aligned} \tag{1.17}$$

It is also useful to define a mispricing function $\delta(x, v, t; \sigma_0)$ as the approximation error under the *BSM* model price $w^{bs}(x, t; \sigma_0)$, since it will be needed to quantify the error incurred for introducing a simpler model in the pricing:

$$\delta(x, v, t; \sigma_0) = (\mathcal{L} - \mathcal{L}_0) w^{bs}(x, t; \sigma_0). \tag{1.18}$$

And, by substituting this newly defined function in equation (1.17), the estimated error for introducing a new model becomes:

$$0 = \mathcal{L}\Delta w(x, v, t; \sigma_0) - r\Delta w(x, v, t; \sigma_0) + \delta(x, v, t; \sigma_0), \tag{1.19}$$

with the boundary condition $\Delta w(x, v, t; \sigma_0) = 0$.

To compute this mispricing function $\delta(x, v, t; \sigma_0)$, it is necessary to define the new infinitesimal generators $\mathcal{L}$ and $\mathcal{L}_0$ applied to the *BSM* pricing function $w^{bs}(x, t; \sigma_0)$, according to equation (1.18). Since *BSM* assumes constant volatility and does not involve the stochastic volatility parameter v, all the partial derivatives related to this variable can be omitted from the full infinitesimal generator $\mathcal{L}$ given in equation (1.11). Furthermore, since the operator $\mathcal{L}_0$ is directly related to *BSM*, expressed in equation (1.13) (in contrast to $\mathcal{L}$, since it concerns the model of interest assuming stochastic volatility), the volatility variable v can also be substituted by the constant volatility σ_0^2:

$$\mathcal{L}w^{bs}(x, t; \sigma_0) = \frac{\partial w}{\partial t} + (r - q)x\frac{\partial w}{\partial x} + \frac{1}{2}vx^2\frac{\partial^2 w}{\partial x^2}, \tag{1.20}$$

$$\mathcal{L}_0 w^{bs}(x, t; \sigma_0) = \frac{\partial w}{\partial t} + (r - q)x\frac{\partial w}{\partial x} + \frac{1}{2}\sigma_0^2 x^2\frac{\partial^2 w}{\partial x^2} \; (as\ v = \sigma_0^2), \tag{1.21}$$

hence, the equation (1.18) can be rewritten as:

$$\delta(x, v, t; \sigma_0) = (\mathcal{L} - \mathcal{L}_0) w^{bs}(x, t; \sigma_0) = \frac{1}{2}(v - \sigma_0^2)x^2\frac{\partial^2}{\partial x^2}w^{bs}(x, t; \sigma_0), \tag{1.22}$$

which captures the instantaneous hedging cost resulting from the discrepancy between the model of interest and auxiliary model, in this case *BSM*.

Having the mispricing equation properly defined, it is now safe to find a way to compute the pricing of the model of interest, and, after analyzing the equation (1.19), it is now starting to resemble the structure of the *PDE* from the equation (1.1) of *Feynman-Kac* theorem. However, to make this connection explicit, some small adjustments are still needed.

Therefore, by adding a new term for $\frac{\partial u}{\partial t}$, and substituting u as $\Delta w(x, v, t; \sigma_0)$; the functions $\mu(x, t)$ and $\sigma^2(x, t)$ become the drift and diffusion terms of the operator $\mathcal{L}$; the function $V(x, t)$ is taken to be the constant risk-free interest rate r; and $f(x, t)$ is substituted by $\delta(x, v, t; \sigma_0)$, it is now possible to closely see the similarities. With this identification, and given the boundary condition $\Delta w(x, v, t; \sigma_0) = 0$, the corresponding *PDE* for $\Delta w(x, v, t; \sigma_0)$ takes the form of:

$$\frac{\partial \Delta w}{\partial t} + \mathcal{L}\Delta w(x, v, t; \sigma_0) - r\Delta w(x, v, t; \sigma_0) + \delta(x, v, t; \sigma_0) = 0. \tag{1.23}$$

Having established equation (1.23), it is now possible to apply the *Feynman-Kac* representation from equation (1.2), assuming that it follows the *Itô* diffusion process defined by the *SDE* equation (1.3). This allows to express the pricing approximation error $\Delta w(x, v, t; \sigma_0)$, as an expected discounted integral of the mispricing term $\delta(x, v, t; \sigma_0)$, which admits the following solution at the terminal value of the underlying asset, i.e., $X(T)$:

$$\begin{aligned} \Delta w(x, v, t; \sigma_0) &= \mathbb{E}_{x,v,t}\left[e^{-\int_t^T V(X_r, r)dr} b\big(X(T)\big) + \int_t^T e^{-\int_t^u V(X_r, r)dr} \delta(X(u), v(u), u; \sigma_0)du\right] \\ &= \mathbb{E}_{x,v,t}\left[e^{-\int_t^T r dr} b\big(X(T)\big) + \int_t^T e^{-\int_t^u r dr} \delta(X(u), v(u), u; \sigma_0)du\right] \\ &= \mathbb{E}_{x,v,t}\left[e^{-r(T-t)} b\big(X(T)\big) + \int_t^T e^{-r(u-t)} \delta(X(u), v(u), u; \sigma_0)du\right], \end{aligned} \tag{1.24}$$

then, given that the terminal payoff function $b\big(X(T)\big)$ is zero, since the pricing approximation error of an option at the maturity will be null, it is possible to omit the first term. Therefore, the expression simplifies to:

$$\Delta w(x, v, t; \sigma_0) = \mathbb{E}_{x,v,t}\left[\int_t^T e^{-r(u-t)} \delta(X(u), v(u), u; \sigma_0)du\right], \tag{1.25}$$

and, using the pricing difference between both models as defined in equation (1.14), the expression can be rewritten to provide an approximation for the option price under stochastic volatility as:

$$w(x, v, t) = w^{bs}(x, t; \sigma_0) + \mathbb{E}_{x,v,t}\left[\int_t^T e^{-r(u-t)} \delta(X(u), v(u), u; \sigma_0)du\right]. \tag{1.26}$$

As shown by *El Karoui et al.* (1998), and further elaborated by *Corielli* (2006), the mispricing term $\delta(x, v, t; \sigma_0)$ can be interpreted as the instantaneous increment in the total hedging cost, resulting from using an incorrect model (in this case *BSM*) to hedge an option against the model of interest defined under the asset price dynamics of equation (1.6). Since the approximation in equation (1.26) is expressed as a conditional expectation under this true model, it is generally not possible to obtain a closed-form expression for its value, especially for auxiliary models more complex than *BSM*.

Due to the obstacle of the conditional expectation generally lacking a closed-form solution, a possible route is to expand this term to an approximation using the *Taylor* series of the mispricing function over time. This approach is formalized in-depth by *KM* in their *Proposition A.3*, however the below exposition presented tries to simplify this process, and, for that matter, it is convenient to temporarily neglect the exponential discounting term $e^{-r(u-t)}$, since this simplification helps focusing only on the expansion of the mispricing term.

Under this assumption, the integral in equation (1.26) becomes:

$$\int_t^T \delta(X(u), v(u), u; \sigma_0)du = \int_t^T \sum_{n=0}^{\infty} \frac{(u-t)^n}{n!} \delta_n(x, v, t; \sigma_0)du, \tag{1.27}$$

rearranging the terms, it is possible to see that:

$$\sum_{n=0}^{\infty} \frac{\delta_n(x, v, t; \sigma_0)}{n!} \int_t^T (u-t)^n du. \tag{1.28}$$

To develop the integral there must be a change of variables, hence, let $u - t = v$ so that $du = dv$:

$$\begin{aligned}\int_t^T (u-t)^n du = \int_t^T v^n dv = \left[\frac{v^{n+1}}{n+1}\right]_t^T &= \left[\frac{(u-t)^{n+1}}{n+1}\right]_t^T \\ &= \frac{(T-t)^{n+1}}{n+1} - \frac{(t-t)^{n+1}}{n+1} \\ &= \frac{(T-t)^{n+1}}{n+1},\end{aligned} \tag{1.29}$$

therefore, the full expansion becomes:

$$\sum_{n=0}^{\infty} \frac{\delta_n(x, v, t; \sigma_0)}{n!} \times \frac{(T-t)^{n+1}}{n+1} = \sum_{n=0}^{\infty} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(x, v, t; \sigma_0). \tag{1.30}$$

Having the previous expansion, it is safe to express the general form of equation (1.26), using a *Taylor* series for the mispricing function $\delta_n(x, v, t; \sigma_0)$ as:

$$w(x, v, t) = w^{bs}(x, t; \sigma_0) + \sum_{n=0}^{\infty} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(x, v, t; \sigma_0), \tag{1.31}$$

where the function δ_n satisfies the recursive relation below:

$$\delta_{n+1}(x, v, t; \sigma_0) = \mathcal{L}\delta_n(x, v, t; \sigma_0) - r\delta_n(x, v, t; \sigma_0), with\ \delta_0 = \delta. \tag{1.32}$$

However, as suggested by *KM*, since in real life simulations there will not be an infinite number of terms, the general approximation expression can be rewritten to a N finite number of terms:

$$w_N(x, v, t; \sigma_0) = w^{bs}(x, t; \sigma_0) + \sum_{n=0}^{N} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(x, v, t; \sigma_0), for\ all\ N \geq 0. \tag{1.33}$$

This approximation is interesting because, even though it starts from the *BSM* model, which assumes constant volatility, and, by default may provide less accurate results, the *KM* method still incorporates information about stochastic volatility through the mispricing function $\delta(x, v, t; \sigma_0)$, which this correction term allows to reduce some of the intrinsic inaccuracies that arise when using a simpler auxiliary model. Additionally, although the authors note that introducing an auxiliary model inevitably brings in a nuisance parameter σ_0, this parameter does not directly affect the pricing function itself, but it does appear in $w_N(x, v, t; \sigma_0)$. In the *Chapter 2*, this nuisance parameter will be further explained and will outline some of the possible ways to handle it, as suggested by the authors.

1.5 Kristensen and Mele model

Having established the foundations for the *KM* approach, it is now possible to derive their general approximation method for more complex asset price approximations, following the same logic but, this time using the full framework, since it is designed to handle auxiliary models that do not offer a closed-form solution, or that offer semi-closed-form solutions, unlike simple models that admit closed-form pricing solutions such as *BSM*.

In their paper, the *KM* method considers a multifactor stochastic model, in which asset prices consist of a d-dimensional vector of state variables, i.e., $\mathscr{f}(t) = \begin{bmatrix} x \\ v(t) \end{bmatrix}$, with $x = S(t)$ still being the underlying asset price and $v(t)$ the instantaneous asset return variance, with the processes already defined previously in equations (1.6) and (1.7).

Under the risk-neutral probability measure, the dynamics of the state vector $\mathscr{f}(t) \in \mathbb{R}^d$ are given by the following *SDE*:

$$d\mathscr{f}(t) = \mu(\mathscr{f}(t), t)dt + \sigma(\mathscr{f}(t), t)dW(t), \tag{1.34}$$

where $\mu(\mathscr{f}(t), t) \in \mathbb{R}^d$ and $\sigma(\mathscr{f}(t), t) \in \mathbb{R}^{d\times m}$ still represents the drift and diffusion functions, respectively, and $W(t)$ also remains the standard *Brownian* motion under the risk-neutral probability measure $\mathbb{Q}$.

To derive the price $w(\mathscr{f}, t)$, it is first necessary to compute the infinitesimal generator $\mathcal{L}$ associated to the *SDE* above, which can be done using the multivariate *Itô's* lemma, since it was already obtained in equation (1.10), which yields the full infinitesimal generator $\mathcal{L}$ present in equation (1.11). However, another possible format for the infinitesimal generator $\mathcal{L}$, which the authors use is the simpler form, that shortens the equation (1.11) to the drift and diffusion factors of the vector $\mathscr{f}(t)$. This can be achieved by obtaining the shorter form of the equation (1.10), using the multivariate *Itô's* lemma:

$$dw(\mathscr{f}, t) = \frac{\partial w}{\partial t}dt + \sum_{i=1}^{d} \frac{\partial w}{\partial \mathscr{f}_i} d\mathscr{f}_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 w}{\partial \mathscr{f}_i \partial \mathscr{f}_j} d\mathscr{f}_i d\mathscr{f}_j, \tag{1.35}$$

and, converting the equation (1.34) to the expanded version of $d\mathscr{f}(t)$:

$$d\not{f}_i(t) = \mu_i(\not{f}(t), t)dt + \sum_{k=1}^{m} \sigma_{i,k}(\not{f}(t), t)dW_k(t), \tag{1.36}$$

and $d\not{f}_i d\not{f}_j$ as:

$$d\not{f}_i d\not{f}_j(t) = \sum_{k=1}^{m} \sigma_{i,k}(\not{f}, t)\sigma_{j,k}(\not{f}, t)dt = \sigma_i(\not{f}, t)\sigma_j(\not{f}, t)^T = \sigma_{i,j}^2(\not{f}, t) \in \mathbb{R}^{dxd}, \tag{1.37}$$

then, adding all the terms associated with dt, this final version of the infinitesimal generator $\mathcal{L}$ becomes:

$$\mathcal{L}w(\not{f}, t) = \frac{\partial w}{\partial t} + \sum_{i=1}^{d} \mu_i(\not{f}, t)\frac{\partial w}{\partial \not{f}_i} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma_{i,j}^2(\not{f}, t)\frac{\partial^2 w}{\partial \not{f}_i \partial \not{f}_j}. \tag{1.38}$$

Since the authors now consider a more complex scenario, with a model that allows for sporadic and stochastic coupon or dividend payments, they propose a variation to the equation (1.12). This new formula not only introduces this term, but also assumes that the constant risk-free interest rate r becomes stochastic. Hence, the solution to the previous *PDE* becomes:

$$\mathcal{L}w(\not{f}, t) + c(\not{f}, t) = R(\not{f}, t)w(\not{f}, t), \tag{1.39}$$

with boundary condition defined by the payoff at the maturity, i.e., $w(\not{f}, T) = b(\not{f})\ \forall\ \not{f}$. Here, $c(\not{f}, t)$ represents the stochastic instantaneous coupon or dividend rate, and $R(\not{f}, t)$ the stochastic instantaneous short-term interest rate, which under the risk-neutral probability measure, the expected payoff and coupon/dividend payments are discounted accordingly.

Practically, the equation (1.39) implies that the expected instantaneous asset capital gain, under the risk-neutral probability measure $\mathcal{L}w(\not{f}, t)$, plus the instantaneous coupon rate $c(\not{f}, t)$, must equal the instantaneous return of an asset required by the market.

To start the expansion, *KM* introduce an auxiliary model defined as:

$$d\not{f}_0(t) = \mu_0(\not{f}_0(t), t)dt + \sigma_0(\not{f}_0(t), t)dW(t), \tag{1.40}$$

where, $\mu_0(\not{f}_0(t), t)$ and $\sigma_0(\not{f}_0(t), t)$ denote the drift and diffusions functions of the auxiliary process, respectively.

The objective is to expand the dynamics of the model of interest in equation (1.34) around this auxiliary model. For this purpose, *KM* assume that the dimension of this auxiliary model is the same of the actual model, meaning that $\not{f}_0(t)$ is also composed by a d-dimensional vector, i.e., $\not{f}_0(t) \in \mathbb{R}^d$, just as $\not{f}(t)$. This assumption allows the authors to create a simpler base model to analyze potentially more complex frameworks, for example with higher dimensions, i.e., assuming stochastic volatility or even the possibility of jumps.

Moreover, even though there is a possibility for the auxiliary models to have lower dimensions than the model of interest, i.e., $\mathbb{R}^m$ where $m < d$, it is always possible to extend it by adding additional constant components. For example, let $y(t) \in \mathbb{R}^m$ be a process with drift μ_y, diffusion σ_y and with a m-dimensional standard *Brownian* motion $W_1(t)$, satisfying the *SDE*:

$$dy(t) = \mu_y(y(t), t)dt + \sigma_y(y(t), t)dW_1(t). \tag{1.41}$$

It is also possible to define this vector process as $[y^T \quad \not{f}_{m+1} \quad \dots \quad \not{f}_d]^T$ where the last $d - m$ components are constant over time. As reported by the authors, this process only satisfies the equation (1.40) when the below conditions apply:

$$\mu_{0,i}(\not{f}, t) = \begin{cases} \mu_{Y,i}(y, t) & 1 \leq i \leq m \\ 0 & otherwise \end{cases},$$
$$\sigma_{0,ij}(\not{f}, t) = \begin{cases} \sigma_{Y,ij}(y, t) & 1 \leq ij \leq m \\ 0 & otherwise \end{cases}. \tag{1.42}$$

In other words, the first m components of the drift and diffusion will match those of the lower auxiliary process, while the remaining $d - m$ components will remain constant, due to a zero drift and diffusion.

As proposed by the authors, the auxiliar model will have a similar performance as the actual model since normally it will be chosen a similar payoff $b_0(\not{f}) = b(\not{f})$, so that the auxiliary pricing function will mimic the actual pricing, i.e., $w_0(\not{f}, t) = w(\not{f}, t)$, at the maturity time T, assuming that $w_0(\not{f}, t)$ has a closed-form solution when the vector satisfies the equation (1.40).

Using this framework, it is possible to define the *PDEs* corresponding to both model of interest in equation (1.39), and the auxiliary model as:

$$\mathcal{L}w(\not{f}, t) + c(\not{f}, t) = R(\not{f}, t)w(\not{f}, t), \tag{1.43}$$

$$\mathcal{L}_0 w_0(\not{f}, t) + c_0(\not{f}, t) = R(\not{f}, t)w_0(\not{f}, t). \tag{1.44}$$

Recalling that the difference between the model of interest and the auxiliary model is defined as below, when rearranging the terms of the pricing difference between models it becomes:

$$\Delta w(\not{f}, t) = w(\not{f}, t) - w_0(\not{f}, t) \Leftrightarrow w(\not{f}, t) = \Delta w(\not{f}, t) + w_0(\not{f}, t), \tag{1.45}$$

hence, it is possible to expand the infinitesimal generator $\mathcal{L}$ associated with the pricing function $w(\not{f}, t)$ of the model of interest in equation (1.43), by substituting the terms in equation (1.45), such as:

$$\mathcal{L}w(\not{f}, t) = \mathcal{L}[\Delta w(\not{f}, t) + w_0(\not{f}, t)] = \mathcal{L}\Delta w(\not{f}, t) + \mathcal{L}w_0(\not{f}, t), \tag{1.46}$$

and, substituting this in the full model of interest in equation (1.43), it becomes:

$$
\begin{aligned}
&\mathcal{L}w(\not{f},t) + c(\not{f},t) = R(\not{f},t)w(\not{f},t) \Leftrightarrow \\
&\Leftrightarrow \mathcal{L}\Delta w(\not{f},t) + \mathcal{L}w_0(\not{f},t) + c(\not{f},t) = R(\not{f},t)\Delta w(\not{f},t) + R(\not{f},t)w_0(\not{f},t).
\end{aligned}
\tag{1.47}
$$

Now, subtracting the auxiliary function in equation (1.44) from both sides of the full model in equation (1.47), the new expansion becomes:

$$
\begin{aligned}
&\big(\mathcal{L}\Delta w(\not{f},t) + \mathcal{L}w_0(\not{f},t) + c(\not{f},t)\big) - \big(\mathcal{L}_0 w_0(\not{f},t) + c_0(\not{f},t)\big) \\
&\qquad = \big(R(\not{f},t)\Delta w(\not{f},t) + R(\not{f},t)w_0(\not{f},t)\big) - \big(R(\not{f},t)w_0(\not{f},t)\big) \Leftrightarrow \\
&\Leftrightarrow \mathcal{L}\Delta w(\not{f},t) + \big(\mathcal{L}w_0(\not{f},t) - \mathcal{L}_0 w_0(\not{f},t) + c(\not{f},t) - c_0(\not{f},t)\big) = R(\not{f},t)\Delta w(\not{f},t),
\end{aligned}
\tag{1.48}
$$

and as previously introduced in equation (1.18), the term $\delta(\not{f},t)$ still represents the mispricing function between both models, with the difference that now it includes the stochastic coupon or dividend rate, as defined below, however, these terms can be easily omitted to isolate mainly the relation of both models on the rest of the factors, e.g., to focus on the drift and diffusion factors.

Hence, the mispricing function becomes:

$$
\begin{aligned}
\delta(\not{f},t) &= \mathcal{L}w_0(\not{f},t) - \mathcal{L}_0 w_0(\not{f},t) + c(\not{f},t) - c_0(\not{f},t) \\
&= (\mathcal{L} - \mathcal{L}_0)w_0(\not{f},t) + \big(c(\not{f},t) - c_0(\not{f},t)\big) \\
&= (\mathcal{L} - \mathcal{L}_0)w_0(\not{f},t).
\end{aligned}
\tag{1.49}
$$

Therefore, when substituting $\delta(\not{f},t)$ into equation (1.48) it finally shows the similarity from the equation (1.19), when it was originally defined the *PDE* for the estimated error when introducing a new model, as seen below:

$$
\mathcal{L}\Delta w(\not{f},t) + \delta(\not{f},t) = R(\not{f},t)\Delta w(\not{f},t). \tag{1.50}
$$

The authors also propose the boundary condition $\Delta w(\not{f},T) = d(\not{f})$, where $d(\not{f})$ represents the difference in the terminal payoffs between the two models, i.e., $d(\not{f}) = b(\not{f}) - b_0(\not{f})$, meaning that $d(\not{f})$ captures the mispricing at maturity due to the differences of both models.

To compute the mispricing function $\delta(\not{f},t)$ in equation (1.49), a similar approach to define equation (1.22) can be used, but this time focusing only on the drift and diffusion terms. Therefore, using the multivariate *Itô's* lemma expression for $\mathcal{L}w(\not{f},t)$ from equation (1.38), but now applied to calculate the infinitesimal generators $\mathcal{L}$ and $\mathcal{L}_0$ acting on the auxiliary model's pricing function $w_0(\not{f},t)$:

$$
\mathcal{L}w_0(\not{f},t) = \frac{\partial w_0}{\partial t} + \sum_{i=1}^{d} \mu_i(\not{f},t)\frac{\partial w_0}{\partial \not{f}_i} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma_{i,j}^2(\not{f},t)\frac{\partial^2 w_0}{\partial \not{f}_i \partial \not{f}_j}, \tag{1.51}
$$

$$
\mathcal{L}_0 w_0(\not{f},t) = \frac{\partial w_0}{\partial t} + \sum_{i=1}^{d} \mu_{0,i}(\not{f},t)\frac{\partial w_0}{\partial \not{f}_i} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma_{0,i,j}^2(\not{f},t)\frac{\partial^2 w_0}{\partial \not{f}_i \partial \not{f}_j}. \tag{1.52}
$$

Having defined this, is it easily seen that the mispricing function of equation (1.49) becomes:

$$
\begin{aligned}
\delta(\mathit{f}, t) &= (\mathcal{L} - \mathcal{L}_0) w_0(\mathit{f}, t) \\
&= \left(\frac{\partial w_0}{\partial t} + \sum_{i=1}^{d} \mu_i(\mathit{f}, t) \frac{\partial w_0}{\partial \mathit{f}_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sigma_{i,j}^2(\mathit{f}, t) \frac{\partial^2 w_0}{\partial \mathit{f}_i \partial \mathit{f}_j} \right) \\
&\qquad - \left(\frac{\partial w_0}{\partial t} + \sum_{i=1}^{d} \mu_{0,i}(\mathit{f}, t) \frac{\partial w_0}{\partial \mathit{f}_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sigma_{0,i,j}^2(\mathit{f}, t) \frac{\partial^2 w_0}{\partial \mathit{f}_i \partial \mathit{f}_j} \right) \\
&= \sum_{i=1}^{d} \Delta\mu_i(\mathit{f}, t) \frac{\partial w_0}{\partial \mathit{f}_i} + \frac{1}{2} \sum_{i,j=1}^{d} \Delta\sigma_{i,j}^2(\mathit{f}, t) \frac{\partial^2 w_0}{\partial \mathit{f}_i \partial \mathit{f}_j},
\end{aligned}
\tag{1.53}
$$

which only becomes possible when assuming that the variation of the drift and diffusion factors can be written as $\Delta\mu_i(\mathit{f}, t) = \mu_i(\mathit{f}, t) - \mu_{0,i}(\mathit{f}, t)$ and $\Delta\sigma_{i,j}^2(\mathit{f}, t) = \sigma_{i,j}^2(\mathit{f}, t) - \sigma_{0,i,j}^2(\mathit{f}, t)$, respectively.

In *Theorem 1* the authors expressed their *Asset Price Representation*, which provides a general formula to represent the pricing function. Using the propositions explained below, and while applying the *Feynman-Kac* theorem to the equations (1.43) and (1.50), it is possible to obtain the representation of asset price $w(\mathit{f}, t)$, which is a variation of the equation (1.24) that expresses the general formula of the pricing between the model of interest and the auxiliary model:

$$
w(\mathit{f}, t) = w_0(\mathit{f}, t) + \mathbb{E}_{\mathit{f},t}\left[e^{-\int_t^T R(\mathit{f}(s), s) ds} d\big(\mathit{f}(T)\big) + \int_t^T e^{-\int_t^s R(\mathit{f}(u), u) du} \delta(\mathit{f}(s), s) ds \right],
$$

or equivalently, (1.54)

$$
\Delta w(\mathit{f}, t) = \mathbb{E}_{\mathit{f},t}\left[e^{-\int_t^T R(\mathit{f}(s), s) ds} d\big(\mathit{f}(T)\big) + \int_t^T e^{-\int_t^s R(\mathit{f}(u), u) du} \delta(\mathit{f}(s), s) ds \right],
$$

with $\mathit{f}(t)$ defined in equation (1.34), $d(\mathit{f}) = b(\mathit{f}) - b_0(\mathit{f})$ and $\delta(\mathit{f}, t)$ as per equation (1.53).

Another way of interpreting this formula is being the potential hedging costs calculated using incorrect models to price the models of interest, as this precisely shows the relation of the pricing error and the auxiliary model in the right-hand side term for the mispricing error $\delta(\mathit{f}, t)$.

The above steps in the *Asset Price Representation* formula rely on multiple assumptions and conditions, which the authors discuss in detail in their propositions A.1 and A.2, and for which they provide the necessary proofs. The below tries to provide a simplification to the main assumptions:

1) The *Theorem 1*, defined in equation (1.54), is an application of the *Feynman-Kac* theorem as formulated by *Karatzas and Shreve* (1991) in their theorem 5.7.6, and as explained in *KM's* conditions A.3 to A.5. Additionally, the mispricing function $\delta(\mathit{f}, t)$ and the stochastic instantaneous short-term interest rate $R(\mathit{f}, t)$ are assumed to be continuous and to satisfy the growth condition imposed on $\Delta w(\mathit{f}, t)$, and satisfy the condition $|\Delta w(\mathit{f}, t)| \leq$

$C(1+\|\mathscr{f}\|^q)$ for some constants $C, q > 0$, where $\Delta w(\mathscr{f}, t) \in \mathbb{R}^d \times [0, T]$. Moreover, the drift μ and diffusion σ terms are also assumed to satisfy standard continuity and growth conditions, which ensure the existence of the solutions to the model's *SDE*, as proposed by *Kurz* (2018);

2) The authors define a function space $\mathcal{H}$ with norm $\| \,.\, \|_{\mathcal{H}}$ in which the infinitesimal operator $\mathcal{L}$, in equation (1.38), is stable and converges. Additionally, there exists $\bar{\tau} > 0$ and functions $\phi_d \phi_\delta \in \mathcal{H}$ such that $d : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}$ and $\delta : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}$ satisfy $\mathbb{E}\left[\phi_d\left(\mathscr{f}(\bar{\tau})\right) \middle| \mathscr{f}(0) = \mathscr{f}\right] = d(\mathscr{f})$ and $\mathbb{E}\left[\phi_\delta\left(\mathscr{f}(\bar{\tau})\right) \middle| \mathscr{f}(0) = \mathscr{f}\right] = \delta(\mathscr{f}, \bar{\tau})$, where the function $t \mapsto \delta(\mathscr{f}, t)$ is analytic and uniformly in $\| \,.\, \|_{\mathcal{H}}$, and in equation (1.43), $R(\mathscr{f}, t) = R(\mathscr{f})$ with $sup_x |R(\mathscr{f})| < \infty$. As explained by the authors, this serves to impose conditions on the functions $d(\mathscr{f})$ and $\delta(\mathscr{f}, t)$ so to require that each of the two functions can be matched through conditions moments;

3) The infinitesimal generator $\mathcal{L}$ must have a transition density $p_t(y|\mathscr{f})$ with respect to the *Lebesgue measure*, in order to ensure that the expectations above are well defined. Additionally, there must exist a measure π so that $\pi(\mathscr{f})p_t(y|\mathscr{f}) = \pi(y)p_t(\mathscr{f}|y)$ is true. Both of these conditions are true for most diffusion models, which have transitions densities, while the latter condition is known as the time-reversibility, which is always true for univariate and stationary processes.

Additionally, *KM* advise that these assumptions are abstract and not easily verified for specific models, hence the authors do not further verify but show that these remain true for the models considered in their paper.

The next step is to derive an approximation of the error term, in order to easily adjust the asset price of the auxiliary model $w_0(\mathscr{f}, t)$, for the intrinsic error involved. For this, *KM* now propose the *Asset Price Approximation* in their *Definition 1*, a N-th approximation the solution to $w_N(\mathscr{f}, t)$ on equation (1.54) using a series expansion, which is familiar to the equation (1.33), however, this new expansion adds a term for the difference in terminal payoffs defined as $d(\mathscr{f})$:

$$w_N(\mathscr{f}, t) = w_0(\mathscr{f}, t) + \sum_{n=0}^{N} \frac{(T-t)^n}{n!} d_n(\mathscr{f}, t) + \sum_{n=0}^{N} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(\mathscr{f}, t), \tag{1.55}$$

with $d_0(\mathscr{f}, t)$ defined as $d(\mathscr{f}) = b(\mathscr{f}) - b_0(\mathscr{f})$, and $\delta_0(\mathscr{f}, t)$ as the function $\delta(\mathscr{f}, t)$ in equation (1.53), assuming the below relationships:

$$d_n(\mathscr{f}, t) = \mathcal{L} d_{n-1}(\mathscr{f}, t) - R(\mathscr{f}, t) d_{n-1}(\mathscr{f}, t), \tag{1.56}$$

$$\delta_n(\mathscr{f}, t) = \mathcal{L} \delta_{n-1}(\mathscr{f}, t) - R(\mathscr{f}, t) \delta_{n-1}(\mathscr{f}, t). \tag{1.57}$$

With the above definitions, where N represents the order of approximation, the authors also provide the condition $w_N(\oint, t) \rightarrow w(\oint, t)$ as $N \rightarrow \infty$, under which the *Asset Price Approximation* equation (1.55) must adhere in order to be valid. Additionally, *KM* assume that for an order of approximation $N \geq 1$, $d(\oint)$ must be $2N$ times differentiable in relation to $\oint$, and $\delta(\oint, t)$ must also be $2N$ times differentiable in relation to $\oint$ and additionally, N times differentiable in relation to t. The authors also advise that conditions imposed on $d(\oint) = b(\oint) - b_0(\oint)$, in general rule out choosing models with different terminal payoffs, i.e., $b(\oint) \neq b_0(\oint)$, if the payoff function $b(\oint)$ is non-differentiable, e.g., cases when choosing plain vanilla options.

KM also note that the *Asset Price Approximation* in equation (1.55), is only one possible way of approximating the right-hand side of equation (1.54), specifically the functions $d(\oint)$ and $\delta(\oint, t)$, and that other alternative approaches could be used, such as via simulation-based methods, instead of relying on the *Feynman-Kac* theorem. Using these simulation methods could be beneficial, as the auxiliary model $w_0(\oint, t)$ is assumed to improve the accuracy of the price estimation, however, a downside for this alternative is that the power series method, used to obtain equation (1.55), is significantly more computational efficient than via simulations, therefore it should take far less computation time.

1.6 Greeks approximations

While the *KM* approach was originally developed as a means to approximate asset prices in models that may not admit closed-form solutions, the authors also demonstrate that their expansion method can also be used to derive closed-form approximations of partial derivatives of asset prices, which in turn facilitates the estimation of the *Greeks*.

To provide a brief context, the *Greeks* are financial measures that provide a way to analytically quantify the sensitivities of an option's price to changes in various underlying parameters. They are widely used to assess portfolio risk and to understand how options behave in response to market changes and, while there are several *Greeks*, the most commonly used and relevant for this dissertation are the ones defined as follows:

1) $\Delta = \frac{\partial w(\oint,t)}{\partial x}$: known as *Delta*, it measures the rate of change in the option's value with respect to changes in the underlying asset price;
2) $\Gamma = \frac{\partial^2 w(\oint,t)}{\partial x^2} = \frac{\partial \Delta}{\partial x}$: known as *Gamma*, it measures the rate of change of *Delta* with respect to changes in the underlying asset price;
3) $\mathcal{V} = \frac{\partial w(\oint,t)}{\partial v}$: known as *Vega*, it measures the sensitivity of the option's value to changes in the volatility of the underlying asset;

4) $\Theta = \frac{\partial w(\not{f},t)}{\partial t}$: known as *Theta*, it measures the change in the option's value with respect to the passage of time;

5) $\hat{\rho} = \frac{\partial w(\not{f},t)}{\partial r}$: known as *Rho*, it measures the sensitivity of the option's value to changes in the instantaneous short-term interest rate;

6) $\varepsilon = \frac{\partial w(\not{f},t)}{\partial q}$: known as *Epsilon*, it measures the sensitivity of the option's value to changes in the underlying dividend yield.

Analyzing these sensitivities, it seems natural to assume that the calculation of these parameters is straightforward when the pricing function $w(\not{f},t)$ has a closed-form solution, however, in the cases where such a solution is not readily available, *KM* defend that their approximation method, defined in equation (1.55), may be used to obtain approximations of the *PDE* by differentiating the expansion directly.

Using their approximation for the k-th order derivative of $w(\not{f},t)$ gives:

$$\frac{\partial^k w_N(\not{f},t)}{\partial \not{f}^k} = \frac{\partial^k w_0(\not{f},t)}{\partial \not{f}^k} + \sum_{n=0}^{N} \frac{(T-t)^n}{n!} d_n^{(k)}(\not{f},t) + \sum_{n=0}^{N} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n^{(k)}(\not{f},t), \tag{1.58}$$

with the equations:

$$d_n^{(k)}(\not{f},t) = \frac{\partial^k d_n(\not{f},t)}{\partial \not{f}^k}, \tag{1.59}$$

$$\delta_n^{(k)}(\not{f},t) = \frac{\partial^k \delta_n(\not{f},t)}{\partial \not{f}^k}. \tag{1.60}$$

KM advise that these functions defined in equations (1.59) and (1.60), can be evaluated either numerically, e.g., via the finite-difference methods, or analytically, but ultimately, the authors demonstrate their expansion for the first and second-order partial derivatives, i.e., for $k=1$ and $k=2$ respectively, via the analytical route.

The exposition provided below summarizes the process for calculating the first-order partial derivates. For $k=1$, the authors proposed the recursions $d_0^{(1)}(\not{f},t) = \frac{\partial d(\not{f},t)}{\partial \not{f}}$ and $\delta_0^{(1)}(\not{f},t) = \frac{\partial \delta(\not{f},t)}{\partial \not{f}}$. Since the functions $d_n^{(k)}(\not{f},t)$ and $\delta_n^{(k)}(\not{f},t)$ have similar forms, it is sufficient to show the process for one of them, so differentiating both sides of the equation (1.56) with respect to $\not{f}$ it becomes:

$$\begin{aligned} &\frac{\partial}{\partial \not{f}}\big(d_n(\not{f},t)\big) = \frac{\partial}{\partial \not{f}}\big(\mathcal{L}d_{n-1}(\not{f},t)\big) - \frac{\partial}{\partial \not{f}}\big(R(\not{f},t)d_{n-1}(\not{f},t)\big) \Leftrightarrow \\ &\Leftrightarrow d_n^{(1)}(\not{f},t) = \frac{\partial}{\partial \not{f}}\big(\mathcal{L}d_{n-1}(\not{f},t)\big) - \frac{\partial}{\partial \not{f}}\big(R(\not{f},t)d_{n-1}(\not{f},t)\big). \end{aligned} \tag{1.61}$$

The next step is to expand the derivatives of both right-hand side terms, so the first becomes:

$$\begin{aligned}\frac{\partial}{\partial f}\left(\mathcal{L}d_{n-1}(f,t)\right) &= \mathcal{L}\frac{\partial d_{n-1}(f,t)}{\partial f}+\mathcal{L}^{(1)}d_{n-1}(f,t)\\ &= \mathcal{L}d_{n-1}^{(1)}(f,t)+\mathcal{L}^{(1)}d_{n-1}(f,t),\end{aligned} \tag{1.62}$$

with $\mathcal{L}^{(1)}$ already defined by the authors as:

$$\mathcal{L}^{(1)}\phi(f,t)=\sum_{i=1}^{d}\frac{\partial\mu_i(f,t)}{\partial f}\frac{\partial\phi(f,t)}{\partial f_i}+\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial\sigma_{i,j}^2(f,t)}{\partial f}\frac{\partial^2\phi(f,t)}{\partial f_i\partial f_j}. \tag{1.63}$$

And the second derivative term of the right-hand side becomes:

$$\begin{aligned}\frac{\partial}{\partial f}\left(R(f,t)d_{n-1}(f,t)\right) &= \frac{\partial R(f,t)}{\partial f}d_{n-1}(f,t)+R(f,t)\frac{\partial d_{n-1}(f,t)}{\partial f}\\ &= \frac{\partial R(f,t)}{\partial f}d_{n-1}(f,t)+R(f,t)d_{n-1}^{(1)}(f,t).\end{aligned} \tag{1.64}$$

Combining these results into equation (1.61), it becomes:

$$\begin{aligned}d_n^{(1)}(f,t) &= \frac{\partial}{\partial f}\left(\mathcal{L}d_{n-1}(f,t)\right)-\frac{\partial}{\partial f}\left(R(f,t)d_{n-1}(f,t)\right)\\ &= \mathcal{L}d_{n-1}^{(1)}(f,t)+\mathcal{L}^{(1)}d_{n-1}(f,t)-\frac{\partial R(f,t)}{\partial f}d_{n-1}(f,t)-R(f,t)d_{n-1}^{(1)}(f,t)\\ &= \mathcal{L}d_{n-1}^{(1)}(f,t)-R(f,t)d_{n-1}^{(1)}(f,t)+\mathcal{L}^{(1)}d_{n-1}(f,t)-\frac{\partial R(f,t)}{\partial f}d_{n-1}(f,t).\end{aligned} \tag{1.65}$$

As the pricing error function $\delta_n^{(1)}(f,t)$ is similar, it is possible to replace the terms related to the function $d(f,t)$ with $\delta(f,t)$, as follows:

$$\delta_n^{(1)}(f,t)=\mathcal{L}\delta_{n-1}^{(1)}(f,t)-R(f,t)\delta_{n-1}^{(1)}(f,t)+\mathcal{L}^{(1)}\delta_{n-1}(f,t)-\frac{\partial R(f,t)}{\partial f}\delta_{n-1}(f,t). \tag{1.66}$$

As mentioned previously, the above process is based on an analytical recursion method, which the authors also compute for the second-order partial derivative, i.e., $k=2$, in their paper. Moreover, for comparison, *Kurz* (2018) discusses in their master's thesis a numerical approach using central differences for each corrective term, which they recommend as it yielded the same results as *KM* approximation but tends to be more efficient, since analytical differentiation becomes increasingly more complex with higher order derivatives.

1.7 Comparison of other approximating methods

To assess the accuracy and relevance of the *KM* approximation, the authors did a literary comparison with other well-established expansion methods, notably against *Yang's* (2006) approximation and the expansion derived from risk-neutral probabilities. This comparison aimed to evaluate possible differences in performance, implementation efficiency and suitability for the various applications.

1.7.1 Yang's asset price expansion

In many ways *Yang's* expansion is similar to *KM* approximation, as both approaches assume that the actual model is too complex and cannot be solved in closed-form, due to admitting features such as stochastic volatilities, jumps, etc., therefore to overcome this constraint, the authors introduce an auxiliary model, such as *BSM*, that does admit a closed-form solution, in which then they combine the two in order to obtain an approximation to the model of interest. However, while *Yang's* base model is similar to *KM's* auxiliary model, the series expansion *Yang* proposes is quite different from that of *KM*.

In equation (1.43) *KM* set the instantaneous short-term rate and the coupon rate to zero, i.e., $R(\fint,t)=c(\fint,t)=0$, and then assumed that the auxiliary model's market would mimic the payoff of the real market, hence, $d(\fint)=b(\fint)-b_0(\fint)=0$.

As in *Yang*, the authors also set up a base model using an infinitesimal generator $\mathcal{L}_0 w^{(0)}=0$, with pricing function $w^{(0)}$, with the same boundary condition $\Delta w(\fint,T)=d(\fint)$. Then, knowing this, the unknown price function $w(\fint,t)$ can be easily rewritten as:

$$\begin{gathered}\mathcal{L}w(\fint,t)=0 \Leftrightarrow \\ \Leftrightarrow \mathcal{L}w(\fint,t)+\big(\mathcal{L}_0 w(\fint,t)-\mathcal{L}_0 w(\fint,t)\big)=0 \Leftrightarrow \\ \Leftrightarrow \mathcal{L}_0 w(\fint,t)+(\mathcal{L}-\mathcal{L}_0)w(\fint,t)=0.\end{gathered} \tag{1.67}$$

These new terms introduced now were already previously defined, so recalling equation (1.21), and noting that $\fint$ is composed by the state variables $\begin{bmatrix} x \\ v \end{bmatrix}$, $\mathcal{L}_0 w(\fint,t)$ becomes:

$$\mathcal{L}_0 w(\fint,t)=\frac{\partial w}{\partial t}+(r-q)x\frac{\partial w}{\partial x}+\frac{1}{2}vx^2\frac{\partial^2 w}{\partial x^2}\ (as\ v=\sigma_0^2), \tag{1.68}$$

since $\mathcal{L}_0$ assumes a constant variance v under the *BSM* assumptions, so the partial derivatives related to v can be omitted.

Additionally, as $(\mathcal{L}-\mathcal{L}_0)w(\fint,t)$ is defined by the difference of the infinitesimal generator $\mathcal{L}$ by $\mathcal{L}_0$, it is possible to subtract the equation (1.68) from the full infinitesimal generator $\mathcal{L}$ already defined in equation (1.11), which yields:

$$(\mathcal{L}-\mathcal{L}_0)w(\fint,t)=\ \kappa(\alpha-v)\frac{\partial w}{\partial v}+\frac{1}{2}\omega^2 v^{2\xi}\frac{\partial^2 w}{\partial v^2}+\rho\omega\left(v^{\xi+\frac{1}{2}}\right)x\frac{\partial^2 w}{\partial x\partial v}, \tag{1.69}$$

which, practically, this equation only keeps the partial derivative terms linked to the stochastic variance, since this is the real difference between an infinitesimal generator $\mathcal{L}$ built under the influence of the model of interest, assuming stochastic variance, and the infinitesimal generator $\mathcal{L}_0$ built under the influence of the auxiliary model, assuming constant variance, with both applied to the pricing function assuming the assumptions of the model of interest $w(\mathscr{f}, t)$.

This then leads to a difference between *KM* and *Yang*, even though the definition of $\delta(\mathscr{f}, t)$ is similar (compared to the approximation error function $\delta(\mathscr{f}, t)$ defined by *KM* in equation (1.49)), in *Yang's* approximation it is applied to the model of interest, so the solution cannot be expressed in closed-form, while in *KM's* model it is applied to the auxiliary model so is known in closed-form, which allows a direct closed-form approximation of the expectation.

The next step is to use *Feynman-Kac* theorem to obtain a similar equation of (1.26):

$$w(\mathscr{f}, t) = w_0(\mathscr{f}, t) + \int_t^T \mathbb{E}^0_{\mathscr{f},t}[(\mathcal{L} - \mathcal{L}_0)w_0(\mathscr{f}(u), u)]du, \tag{1.70}$$

with $\mathbb{E}^0_{x,t}$ being the expectation taken under probability measure of the baseline model, defined by the infinitesimal generator $\mathcal{L}_0$. Hence, as per the above explanation, although this expression is the same in both approximations, it leads to different prices since the unknown price of the mispricing function $\delta(\mathscr{f}, t) = (\mathcal{L} - \mathcal{L}_0)w_0(\mathscr{f}, t)$ differs.

Similar to *KM*, *Yang* also suggests a series expansion, in order to approximate the integral term:

$$w(\mathscr{f}, t) = w_0(\mathscr{f}, t) + \sum_{m=1}^{M} w^{(m)}(\mathscr{f}, t), \tag{1.71}$$

with each term $w^{(m)}(\mathscr{f}, t)$ obtained recursively from:

$$\mathcal{L}_0 w^{(m)} + (\mathcal{L} - \mathcal{L}_0)w^{(m-1)} = 0, \tag{1.72}$$

and, subject to the boundary condition $w^{(m)}(\mathscr{f}, T) = 0 \ \forall\, m = 1, \dots, M$. Additionally, rather than computing an infinite series, *Yang* truncates the expansion of $w^{(m)}(\mathscr{f}, T)$ after M terms, using either numerical or analytical processes for the approximation.

Next, *Yang* proposes to use the *Feynman-Kac* theorem for each *PDE*, as per below:

$$w^{(m)}(\mathscr{f}, t) = \int_t^T \mathbb{E}^0_{\mathscr{f},t}\left[(\mathcal{L} - \mathcal{L}_0)w^{(m-1)}(\mathscr{f}(u), u)\right]du \ \forall\, m = 1, \dots, M. \tag{1.73}$$

Lastly, the authors advise that these recursions could be solved through standard symbolic software packages for the first-order terms, and they do a series of numerical comparisons against *Yang's* approximation, in which their version outperforms *Yang's* for the *Heston* model, against a benchmark computed through *Fourier's* transform proposed by *Madan and Carr* (1999).

1.7.2 Perturbation methods

Another alternative to the asset price approximations are perturbation methods. These approaches involve using expansions to problems that are difficult or even impossible to obtain closed-form solutions around known values of the model's parameters.

To better understand these methods, it is helpful to analyze how they are used in approximating complex asset pricing models, such as *Heston*' stochastic volatility model, explained in equations (1.6) and (1.7), since the model generally does not admit a closed-form solution.

Lewis (2000) proposes a perturbation expansion of the option price $w(\not{f}, t)$ around the volatility of volatility parameter ω, as follows:

$$w(\not{f}, t) = \sum_{m=0}^{\infty} \omega^m w_{(m)}(\not{f}, t). \tag{1.74}$$

This method assumes that when the volatility of volatility parameter ω is small and, for a few terms, the solution is quite accurate, however, if ω has a greater value, the expansion will converge slowly and could become more inaccurate, as it will no longer capture the dynamics of price-volatility for lower m terms. Another important limitation would be for higher m terms, as it will increase the complexity and computational power required to price these functions. Additionally, if assumed a $\omega = 0$, this would mean that the volatility is constant, hence the framework would return to the principles of the *BSM* model.

Fouque et al. (2000) proposed to add perturbations in a stochastic volatility model, characterized by the equation (1.6), however, the constant risk-free interest rate r and the continuous dividend yield q was set to zero, which then replaces the equation as per below:

$$dx = xv\big(y(t)\big)dW(t), \tag{1.75}$$

still assuming that the underlying asset price is $S(t) = x$, and being $W(t)$ the standard *Brownian* motion under the risk-neutral probability measure, $v\big(y(t)\big)$ the instantaneous return volatility function, depending on the stochastic process $y(t)$, which is a mean-reverting *Ornstein-Uhlenbeck* process, defined as:

$$dy(t) = \alpha\big(m - y(t)\big)dt + \sigma_\infty \sqrt{\alpha} dW_y(t), \tag{1.76}$$

where $W_y(t)$ is another standard *Brownian* motion under the risk-neutral probability measure, α is the speed of mean reversion, also called as persistence parameter, which amplifies the diffusion term by an order of $\sqrt{\alpha}$ and, the drift term $\big(m - y(t)\big)$ by α. Moreover, m and $\frac{1}{2}\sigma_\infty^2$ are the ergodic mean and variance parameters, also known as the long-term properties, since, despite the random fluctuations of $y(t)$, the function tends to return to its m and $\frac{1}{2}\sigma_\infty^2$ values over time. To better provide context, it is possible to define $\frac{1}{2}\sigma_\infty^2$ as the variance, since this is concerns an *Ornstein-Uhlenbeck* process, which satisfies the equation:

$$VAR\big(y(t)\big) = \frac{b^2}{2a}, \tag{1.77}$$

where, in this context, b is the diffusion term, so that $b = \sigma_\infty\sqrt{\alpha}$ and a is the speed of mean reversion, such that $a = \alpha$. Hence, substituting these terms in equation (1.77) gives:

$$VAR\big(y(t)\big) = \frac{\left(\sigma_\infty\sqrt{\alpha}\right)^2}{2\alpha} = \frac{\alpha\sigma_\infty^2}{2\alpha} = \frac{1}{2}\sigma_\infty^2. \tag{1.78}$$

Moreover, since m and $\frac{1}{2}\sigma_\infty^2$ are the ergodic parameters of this process, they can also be expressed as:

$$m = \lim_{\tau\to\infty} \mathbb{E}_{y,t}\,[y(t+\tau)], \tag{1.79}$$

and,

$$\frac{1}{2}\sigma_\infty^2 = \lim_{\tau\to\infty} \mathbb{V}_{y,t}\,[y(t+\tau)] \Leftrightarrow \sigma_\infty^2 = 2\lim_{\tau\to\infty} \mathbb{V}_{y,t}\,[y(t+\tau)], \tag{1.80}$$

being $\mathbb{V}$ the ergodic variance operator and $\tau = T - t$.

Under the *Fouque et al.* (2000) approach, the authors suggested that when the mean-reversion speed parameter α is large (or for a slight increase in its inverse, α^{-1}), the volatility process $y(t)$ returns quickly to its equilibrium value. This is quite important for modeling, as it means that over the lifespan of the option, the volatility will linger closely to its average most of the time.

Given these parameters and analysis, the authors advanced with an expansion of option prices as follows:

$$w(\not{f}, t) = \sum_{m=0}^{\infty} \alpha^{-\frac{m}{2}} w_{(m)}(\not{f}, t), \tag{1.81}$$

noticing that the mean-reversion parameter is scaling as $\alpha^{-\frac{m}{2}}$, since this variable affects the diffusion term directly, so it is a key aspect of perturbation methods when applied to fast mean-reverting processes.

On the other hand, *Fouque et al.* (2003), also recommended to use a small mean-reversion parameter α, for “slow” expansions, however, these are more appropriate for relatively short-term options, as it is not expected that the volatility will encounter large fluctuations before its maturity.

While perturbation methods are attractive and offer advantages, such as managing small volatility of volatility parameters, fast mean reversions or small fluctuations, they are not always the most realistic methods, as they critically rely on the assumptions that the parameters are small, e.g., they cannot capture high volatility markets. In contrast, compared to base model expansions, such as *KM* approximation method, these are capable of handling broader parameters, as they do not break for larger fluctuations, which in turn increases their robustness and are able to provide added flexibility even when dealing with far more complex model structures.

1.7.3 Risk-neutral probabilities

In this section, *KM* argue that, since asset prices are defined as conditional expectations under the risk-neutral measure, then, any attempt to approximate asset prices should also account for an approximation of the risk-neutral probabilities themselves. This raises the question of how accurate the authors' approximation method is when compared to other approaches that aim to approximate these probabilities.

To better understand the comparison, the authors propose using the expansion from *Asset Price Representation*, in equation (1.54), to approximate the risk-neutral probability of an asset pricing model by expanding it around the risk-neutral probability of an auxiliary model, which is the same concept that *KM* used in their original approximation method. Additionally, the authors observed that their closed-form approach resembles a *saddlepoint* approximation of risk-neutral densities, which they state is a special case of their approximation framework.

For context, the *saddlepoint approximation method* was initially proposed by *Daniels* (1954) and applied to statistics to increase the accuracy of approximation formulas for density functions, e.g., probability density functions (*PDF*), but was later applied in the context of options pricing, such as by *Aït-Sahalia and Yu* (2006), as a mean to generate closed-form approximations to characteristic functions, when the transition densities or cumulative distribution functions do not offer these solutions.

For simplification purposes, the authors proposed to assume that the instantaneous coupon rate $c(\mathscr{x}, t)$ and the instantaneous short-term interest rate of $R(\mathscr{x}, t)$ are set to zero, as well as the payoff mispricing function $d(\mathscr{x})$, meaning that the payoff at maturity is the same for both the model of interest $b(\mathscr{x})$ and auxiliary model $b_0(\mathscr{x})$.

Then, to better visualize the association, it is possible to recall the equation (1.54) using the *Feynman-Kac* theorem, but this time, only considering the terminal payoff term:

$$w(\mathscr{x}, t) = \mathbb{E}_{x,t}\left[e^{-\int_t^T R(\mathscr{x}(s),s)ds} b\big(\mathscr{x}(T)\big)\right], \tag{1.82}$$

then, the pricing functions for both the model of interest and the auxiliary model in terms of the risk-neutral conditional densities, are defined as:

$$w(\mathscr{x}, t) = \int_{\mathbb{R}^d} b\big(\mathscr{x}(T)\big) p(\mathscr{x}(T), T | \mathscr{x}, t)\, d\mathscr{x}\,, \tag{1.83}$$

$$w_0(\not{f},t) = \int_{\mathbb{R}^d} b_0\big(\not{f}(T)\big)p_0(\not{f}(T),T|\not{f},t)\,d\not{f}.$$

And, again recalling that $b(\not{f}) = b_0(\not{f})$, the above equation can be easily simplified by substituting $b(\not{f})$ in the second equation, so the difference between the two models becomes:

$$\begin{aligned} w(\not{f},t) - w_0(\not{f},t) &= \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)p(\not{f}(T),T|\not{f},t)\,d\not{f} - \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)p_0(\not{f}(T),T|\not{f},t)\,d\not{f} \\ &= \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)[p(\not{f}(T),T|\not{f},t) - p_0(\not{f}(T),T|\not{f},t)]\,d\not{f}, \end{aligned} \tag{1.84}$$

where the difference between both models' risk-neutral conditional densities is defined as $\Delta p = p - p_0$, so the equation (1.84) simplifies to:

$$w(\not{f},t) = w_0(\not{f},t) + \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)\Delta p(\not{f}(T),T|\not{f},t)\,d\not{f},$$

or equivalently, (1.85)

$$\Delta w(\not{f},t) = \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)\Delta p(\not{f}(T),T|\not{f},t)\,d\not{f}.$$

Based on the *Asset Price Representation*, in equation (1.54), the pricing difference between the two models, i.e., $\Delta w(\not{f},t)$, can be expressed as:

$$\begin{aligned} &\int_{\mathbb{R}^d} b\big(\not{f}(T)\big)\Delta p(\not{f}(T),T|\not{f},t)\,d\not{f} \\ &\qquad = \mathbb{E}_{\not{f},t}\left[e^{-\int_t^T R(\not{f}(s),s)ds}d\big(\not{f}(T)\big) + \int_t^T e^{-\int_t^s R(\not{f}(u),u)du}\delta(\not{f}(s),s)ds\right], \end{aligned} \tag{1.86}$$

and, since the payoff mispricing function at the terminal value is zero, i.e., $d\big(\not{f}(T)\big) = 0$, and the instantaneous short-term interest rate of $R(\not{f},t) = R(\not{f},u) = 0$, then:

$$\begin{aligned} \int_{\mathbb{R}^d} b\big(\not{f}(T)\big)\Delta p(\not{f}(T),T|\not{f},t)\,d\not{f} &= \mathbb{E}_{\not{f},t}\left[0 + \int_t^T e^0\delta(\not{f}(s),s)ds\right] \\ &= \mathbb{E}_{\not{f},t}\left[\int_t^T \delta(\not{f}(s),s)ds\right] \\ &= \int_t^T \mathbb{E}_{\not{f},t}[\delta(\not{f}(s),s)]ds, \end{aligned} \tag{1.87}$$

where $\delta(\not{f}(s),s)$ remains the approximation error function between the two models, as defined in equation (1.53).

Additionally, *KM* argue that their approximation of $\int_t^T \mathbb{E}_{\not{f},t}[\delta(\not{f}(s),s)]ds$, present in *Asset Price Approximation* in their *Definition 1*, is identical to calculating an asset price through an approximation of the risk-neutral transition discrepancy Δp, which the authors provide proof of the derivation in their *Appendix B*, which highlights the fact that their *Asset Price Representation and Approximation* rely on equivalent expressions of the risk-neutral conditional densities.

Even though the expressions use the before equivalency, *KM* conclude that their methods still are more easily to implement, since they offer closed-form solutions for easy to compute pricing errors and, proposed that the left-hand side of the equation (1.87), represented by the pricing error between the model of interest and auxiliary model's risk-neutral conditional densities, will require the use of a *Riemann* integral, which will be computationally demanding, especially with higher model dimensions. However, on the other hand, the right-hand side term, represented by the mispricing function between the model of interest and the auxiliary model, is effectively easier to implement with the use of a power series expansion, similar to the one shown in *KM's Asset Price Approximation* in *Definition 1*.

CHAPTER 2

Implementation and results

The objective of this chapter is to implement, in *Python*, the equations defined in the previous chapter for both models, assuming *CEV* and stochastic volatility. Additionally, to test whether the *KM* approximation method holds, the results will then be compared against the "true" market prices obtained by their benchmarks, also computed in *Python* via the *Schroder* (1989) standard *CEV* model, for the *CEV* analysis, and inspired by *Rouah* (2013) in their *Heston* model extension code utilizing the *Madan and Carr* (1999) fast *Fourier* transform, for the *Heston* analysis.

2.1 Approximating with the CEV model

To test *KM* approximation method under the *CEV* assumptions, it must be used an auxiliary model with constant volatility and equal to σ_0, hence it is assumed a variation of the *BSM* model defined in equation (1.5) and presented as follows:

$$dx(t) = (r - q)\, x(t)dt + \sigma(x,t)x(t)dW(t), \tag{2.1}$$

with a local volatility function $\sigma(x,t)$ as proposed by *KM* of:

$$\sigma(x,t) = \sigma_{CEV} x_t^{\gamma-1}, \tag{2.2}$$

and where σ_{CEV} is the constant volatility in the *CEV* model, also known as the scale parameter which fixes the initial instantaneous volatility at $t = 0$ so that $\sigma(x,0)$ is equal to:

$$\sigma(x,0) = \sigma_{CEV} x_0^{\gamma-1}, \tag{2.3}$$

as explained by *Dias et al.* (2024) and, being $x_t^{\gamma-1}$ the term that allows for the scaling of the volatility parameter, based on the elasticity variable γ. This factor scales the volatility according to the current asset price x, with $\gamma > 0$ controlling the strength of this relationship, such that higher γ values imply a greater volatility. Additionally, it is also worth noting that when this parameter becomes equal to $\gamma = 0.5$, it corresponds to the *Cox and Ross* (1976) square root model, a $\gamma = 1$ displays a zero elasticity scenario, as seen in the *BS* model and, a $\gamma = 0$ corresponds to the price independent volatility as seen in *Merton* (1973).

Moreover, since the *CEV* model is known in closed form, it is possible to benchmark the accuracy of the approximation via the following European-style option call price equation below, as discussed by *Schroder* (1989):

$$c_t(x_t, K, T) = x_t e^{-q\tau} Q(2\mathscr{y}; 2 + 2\mathscr{v}, 2\mathscr{w}) - Ke^{-r\tau}[1 - Q(2\mathscr{w}; 2\mathscr{v}, 2\mathscr{y})] \; for \; \gamma < 1, \tag{2.4}$$

subjected to the following density of non-central chi-square Q parameters equations:

$$\mathscr{y} = \hat{k} K^{2-2\gamma}, \tag{2.5}$$

$$\hat{k} = \begin{cases} \dfrac{(r-q)}{\sigma_{CEV}^2 (1-\gamma)[e^{(r-q)(2-2\gamma)\tau} - 1]} & for \;\; r \neq q \\ \dfrac{2}{\sigma_{CEV}^2 (2-2\gamma)^2 \tau} & for \;\; r = q \end{cases}, \tag{2.6}$$

$$\mathscr{v} = \frac{1}{2-2\gamma} \; for \; \gamma < 1, \tag{2.7}$$

and,

$$\mathscr{w} = \hat{k} {x_t}^{2-2\gamma} e^{(r-q)(2-2\gamma)\tau}. \tag{2.8}$$

In this setting, the approximation of the unknown option price takes a similar form to *KM's* general approximation expression, defined in equation (1.33). However, because the model assumes constant variance, the parameter v does not apply and can therefore be omitted, yielding the following equation:

$$w_N(x, t; \sigma_0) = w_0(x, t; \sigma_0) + \sum_{n=0}^{N} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(x, t; \sigma_0), for \; all \; N \geq 0, \tag{2.9}$$

where $w_0(x, t; \sigma) = w^{bs}(x, t; \sigma_0)$ and assuming an approximation error function $\delta_n(x, t; \sigma_0)$ such as:

$$\delta_{n+1}(x, t; \sigma_0) = \mathcal{L}\delta_n(x, t; \sigma_0) - r\delta_n(x, t; \sigma_0), with \; \delta_0 = \delta. \tag{2.10}$$

The infinitesimal generator $\mathcal{L}$ and the mispricing function are still the same as in equation (1.11), however, the stochastic volatility factor v is replaced by the deterministic variance of the model of interest $\sigma^2(x, t)$, since the *CEV* model does not feature a stochastic volatility variable. Therefore, as in the section 1.4, all the partial derivatives of the price function w with respect to the volatility v, i.e., $\frac{\partial w}{\partial v}, \frac{\partial^2 w}{\partial v^2}$, and $\frac{\partial^2 w}{\partial x \partial v}$ will also be equal to zero:

$$\mathcal{L}w(x, t; \sigma) = \frac{\partial w}{\partial t} + (r-q)x\frac{\partial w}{\partial x} + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} \frac{\partial^2 w}{\partial x^2}, \tag{2.11}$$

and, recalling equation (1.13), it is possible to apply the difference between the infinitesimal generator $\mathcal{L}$ in equation (1.21) and the instantaneous risk-free interest rate r to the pricing equation $w_0(x, t; \sigma)$, such that:

$$\frac{\partial w_0}{\partial t} + (r-q)x\frac{\partial w_0}{\partial x} + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^2 w_0}{\partial x^2} - rw_0 = 0, \tag{2.12}$$

in which, similarly to equation (1.22), the zero-order approximation of the mispricing function is defined as:

$$\delta(x,t;\sigma_0) = \frac{1}{2}(\sigma^2(x_t) - \sigma_0^2)x^2 \frac{\partial^2}{\partial x^2} w_0(x,t;\sigma_0), \tag{2.13}$$

assuming $\sigma^2(x,t)$ is defined in equation (2.2), which can also be represented as $\sigma^2(x_t)$ and is equal to $\sigma^2(x_t) = \sigma_{CEV}^2 x^{2\gamma-2}$. Moreover, $\delta(x,t;\sigma_0)$ is still interpreted as the instantaneous hedging costs, resulting from the discrepancy between the model of interest and auxiliary model, in this case *BSM*, and can also be translated as using the wrong auxiliary model to approximate the model of interest.

Based on the above, it is possible to obtain the *CEV* model iterations of the pricing error function $\delta(x,t;\sigma_0)$, defined in equation (2.10), and assuming the infinitesimal generator $\mathcal{L}$ in equation (2.11) up to the N-th approximation order.

Table 2.1 – *CEV* model iterations of the pricing error function.

n	Pricing error function δ_n iterations
0	$\frac{1}{2}(\sigma_{CEV}^2 x^{2\gamma-2} - \sigma_0^2)x^2 \frac{\partial^2 w_0(x,t;\sigma_0)}{\partial x^2}$
1	$\frac{\partial \delta_0}{\partial t} + (r-q)x\frac{\partial \delta_0}{\partial x} + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}\frac{\partial^2 \delta_0}{\partial x^2} - r\delta_0$
2	$\frac{\partial \delta_1}{\partial t} + (r-q)x\frac{\partial \delta_1}{\partial x} + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}\frac{\partial^2 \delta_1}{\partial x^2} - r\delta_1$
⋮	⋮
N	$\frac{\partial \delta_{N-1}}{\partial t} + (r-q)x\frac{\partial \delta_{N-1}}{\partial x} + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}\frac{\partial^2 \delta_{N-1}}{\partial x^2} - r\delta_{N-1}$

Additionally, it is possible to expand the first-order approximation of the mispricing function $\delta_1(x,t;\sigma_0)$[2] using the infinitesimal generator $\mathcal{L}$ in equation (2.11) and, assuming the *CEV* volatility function in equation (2.2), the first-order approximation becomes:

$$\begin{aligned}\delta_1(x,t;\sigma_0) = \frac{1}{2}\left((2q - 2r - \sigma_0^2)a(x) + (r-q)xa'(x) + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} a''(x)\right)\frac{\partial^2 w_0}{\partial x^2} \\ -\frac{1}{2}\left(2\sigma_0^2 x a(x) - \sigma_{CEV}^2 x^{2\gamma} a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}a^2(x)\frac{\partial^4 w_0}{\partial x^4}.\end{aligned} \tag{2.14}$$

[2] See the appendix in section B – Calculations of the first-order mispricing function in CEV, for the auxiliary calculations.

assuming that

$$a(x) = \sigma_{CEV}^2 x^{2\gamma} - \sigma_0^2 x^2, \tag{2.15}$$

$$a'(x) = 2\gamma\sigma_{CEV}^2 x^{2\gamma-1} - 2\sigma_0^2 x, \tag{2.16}$$

and

$$a''(x) = 2\gamma(2\gamma-1)\sigma_{CEV}^2 x^{2\gamma-2} - 2\sigma_0^2. \tag{2.17}$$

As briefly noted in the introduction of the *KM* approximation method, using *BSM* as the auxiliary model inevitably introduces a nuisance parameter, in this case σ_0. The authors propose that there are multiple ways to handle this parameter, such as assuming that $\hat{\sigma}_0$ is an estimation of σ_0, and then considering $w_N(x,t;\hat{\sigma}_0)$ as an approximation of $w(x,t)$. Another way is defining the estimate through the calibration process:

$$\hat{\sigma}_N(x,t) = arg\ \min_{\sigma}\big(w_N(x,t;\sigma) - w_0(x,t;\sigma)\big)^2,$$

or equivalently,

$$\hat{\sigma}_N(x,t) = arg\ \min_{\sigma}\left(\sum_{n=0}^{N}\frac{(T-t)^{n+1}}{(n+1)!}\delta_n(x,t;\sigma)\right)^2. \tag{2.18}$$

This argument works by selecting the optimal value of the volatility variable σ that minimizes the squared difference between the N-th order approximation of the model of interest and the auxiliary model's prices, this way aligning the two pricing functions as closely as possible.

As N tends to infinity, the estimate $\hat{\sigma}_N(x,t)$ will converge to the implied volatility of the *BSM* model:

$$\lim_{N\to\infty}\hat{\sigma}_N(x,t) = IV(x,t), \tag{2.19}$$

defined as $w(x,t) = w_0\big(x,t;IV(x,t)\big)$. Under these assumptions, the pricing function that depends on the nuisance parameter becomes $w_N\big(x,t;\hat{\sigma}_N(x,t)\big)$ for a fixed N. More generally, the authors represent this as $w_N\big(x,t;\sigma_M(x,t)\big)$, with $M \leq N$, to allow for situations where the parameter is estimated using an approximation of order M while the pricing function itself is computed to order N.

With all of the base functions defined, it is now possible to implement the *KM's* approximation method for the *CEV* framework in *Python*. Due to the complexity of this approximation method, especially for higher-order approximations, which becomes increasingly more difficult to manage and computationally demanding, the *Sympy* library was used, as it provides symbolic differentiation and algebraic manipulation, which is quite helpful in simplifying the construction of recursive approximation terms and makes the overall implementation both clearer and more flexible.

In order to evaluate the accuracy of the implementation, the results obtained in this dissertation were compared against a benchmark composed by the standard *CEV* model defined in equation (2.4), since it already offers closed form solutions.

In this framework, the results of the standard *CEV* model were treated as the "true" market prices, against the results computed for *KM's CEV* method. Then, the accuracy of the approximations was analyzed using the percentage error for each stock price, calculated via the expression:

$$Difference\ (\%) = \left(\frac{KM\ Prices}{CEV\ Prices} - 1\right) \times 100. \tag{2.20}$$

This equation provides a useful performance metric for analyzing the accuracy of the approximation relative to the benchmark, while also giving a way to visualize the error percentage as figures.

In order to try to obtain the most accurate prices, it was also implemented the volatility minimizer function defined in equation (2.18), as proposed by *KM*, however, due to the computational constraints of minimizing the volatility for each approximation order, it was ultimately chosen to optimize the volatility parameter only once using the best approximation order, which in this case was generally $N = 3$. Although this approach is not strictly correct, since each approximation order will have a different optimal volatility, it was considered the most reasonable compromise that provides high accuracy by not sacrificing a lot of computational power.

Unfortunately, due to the fact that it was not explicitly clear how the authors implemented their method in the *CEV* analysis, the implementation in this dissertation differs in a way that is not possible to directly compare their results with the ones obtained here, hence, when trying to reproduce the figures from the paper using similar parameters, they become inherently different.

Taking this into consideration, when attempting to generate similar figures from the paper, these differences become quite evident. To start, this implementation could not obtain the plots for higher approximation orders, e.g., $N > 4$, which the original paper presents up to $N = 6$, due to heavy computational power required, however, even for the fourth-approximation order it is possible to observe some systemic discrepancies, since the curves are far different from *KM's* paper.

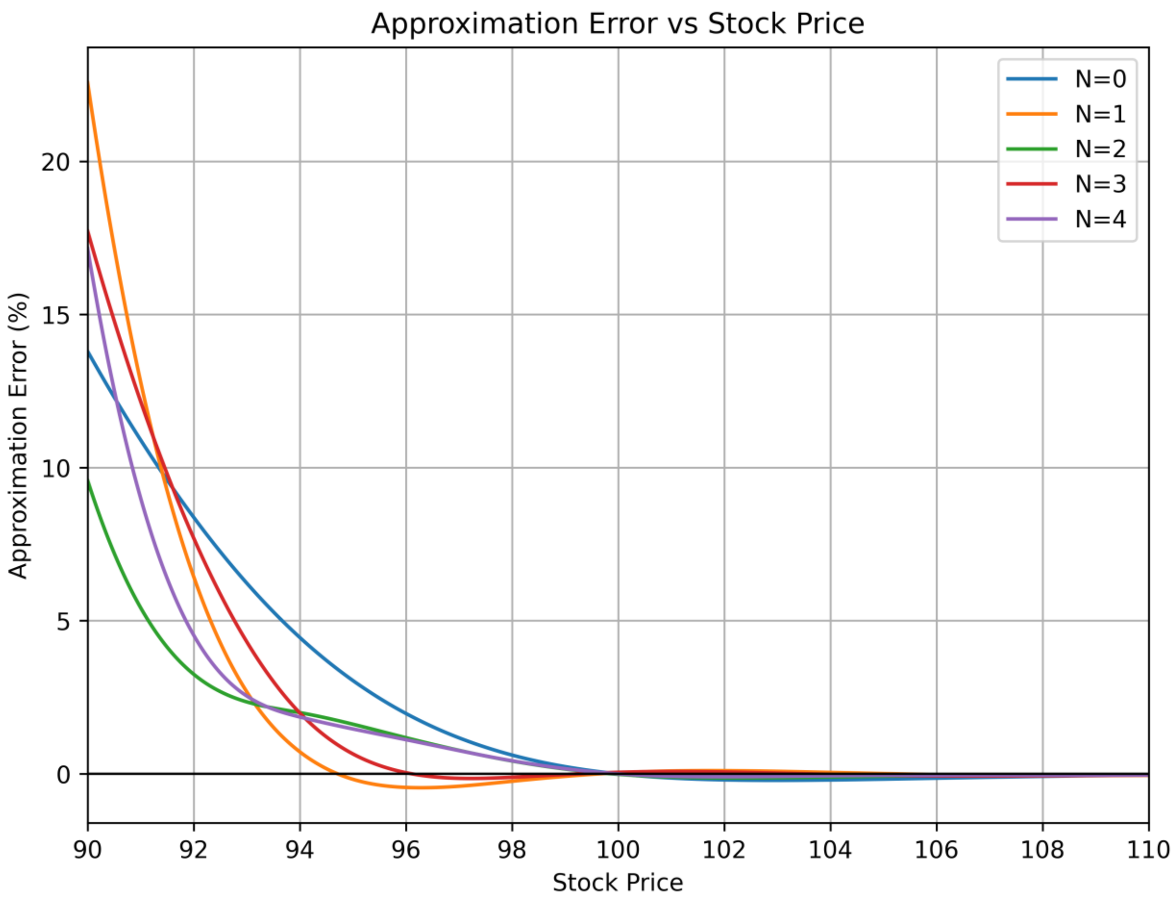

Figure 2.1 – *KM* approximation percentage errors following the *CEV* model defined in equations (2.1) and (2.2).

The parameters were set to $K = 100$, $r = 5\%$, $q = 0\%$, $\tau = 3$ months, $\sigma(x,t) = 10\%$, $\gamma = 0.5$ and $\sigma_{CEV} = \frac{\sigma(x,t)}{x^{\gamma-1}} = \frac{10\%}{x^{0.5-1}} = 0.1x^{0.5}$.

When generating the results with a time to maturity of $\tau = 3$ months, an elasticity parameter of $\gamma = 0.5$ and an optimized volatility of $\sigma(x,t) = 0.10007341$, it is easily seen that all approximation curves have a quite high positive error percentage, meaning that the *KM* approach in this dissertation is overestimating the call prices, particularly for out-of-the-money prices, i.e., $x < K$, but also because with higher approximations orders the error percentage does not decrease as expected, which is trait that *KM* observed in the method they implemented. Nevertheless, the method in this dissertation still achieves a null error near the at-the-money price region, i.e., $x = K$, and it is also possible to see that for in-the-money prices, i.e., $x > K$, the percentage error is extremely small and continues to converge towards zero for far in-the-money prices.

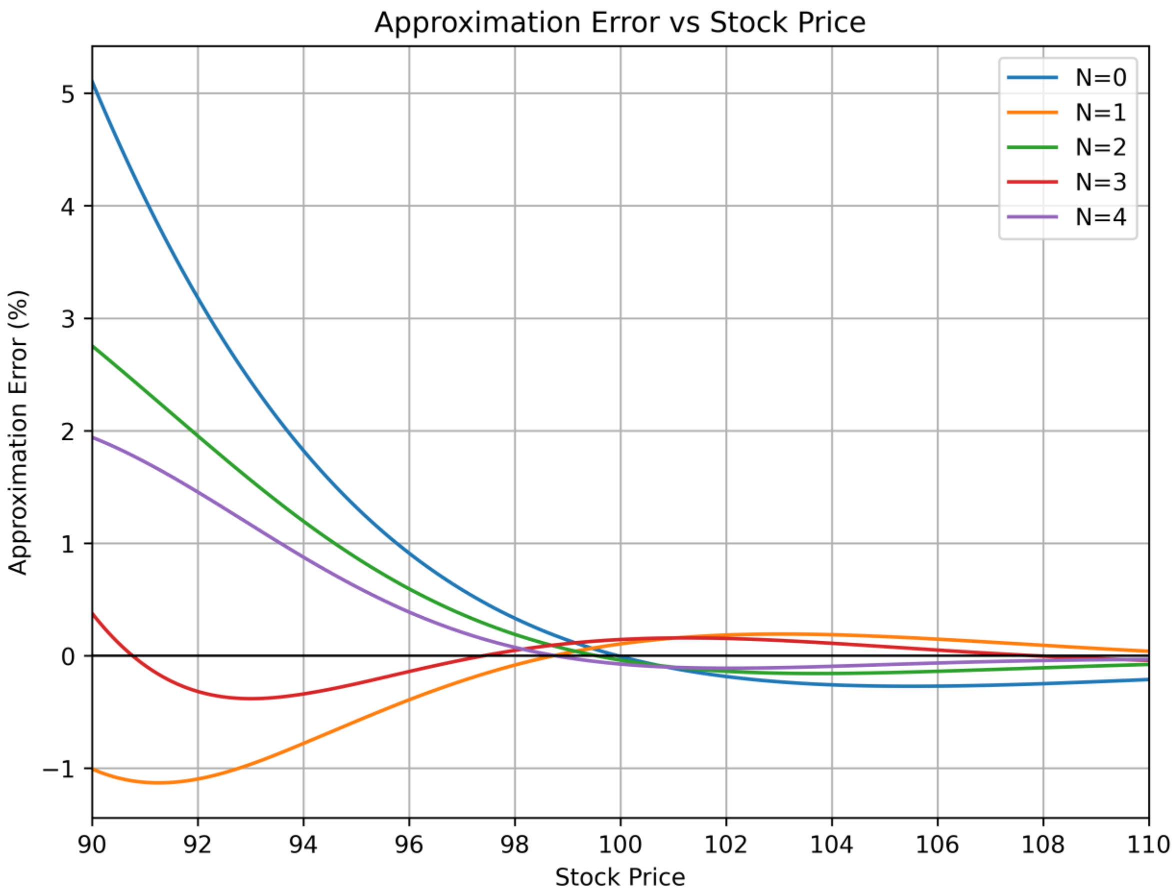

Figure 2.2 – *KM* approximation percentage errors following the *CEV* model defined in equations (2.1) and (2.2).

The parameters were set to $K = 100$, $r = 5\%$, $q = 0\%$, $\tau = 1$ year, $\sigma(x,t) = 10\%$, $\gamma = 0.5$ and

$$\sigma_{CEV} = \frac{\sigma(x,t)}{x^{\gamma-1}} = \frac{10\%}{x^{0.5-1}} = 0.1x^{0.5}.$$

Moreover, in contrast to the figure 2.1, displaying a $\tau = 3$ months, it is easily observed that a longer maturity of $\tau = 1$ year, and an optimal volatility of $\sigma(x,t) = 0.10030050$, provides more accurate results, as this time the maximum error percentage for this range of stock prices does not exceed nearly five percent, and now, each approximation order seems to capture the expected dynamic more effectively. For example, when analyzing $N = 1$, it is possible to see that this curve starts with a negative error, which means that for the first-order approximation the *KM* implementation here is underestimating the call prices. Additionally, as seen in the previous figure, the trend of converging to zero when near at-the-money prices is still showing, with the approximations continuing to converge to zero for far in-the-money option prices. Moreover, it may appear that for in-the-money prices, the curves for each order are much farther from the null approximation error line, however this can be explained as just a matter of perspective, since the figure is now more zoomed in, due to the fact that the maximum error percentage decreased from around 23% to now slightly higher than 5%, therefore as a result, the figure is now highlighting more clearly how each approximation curve behave with each increase of stock price.

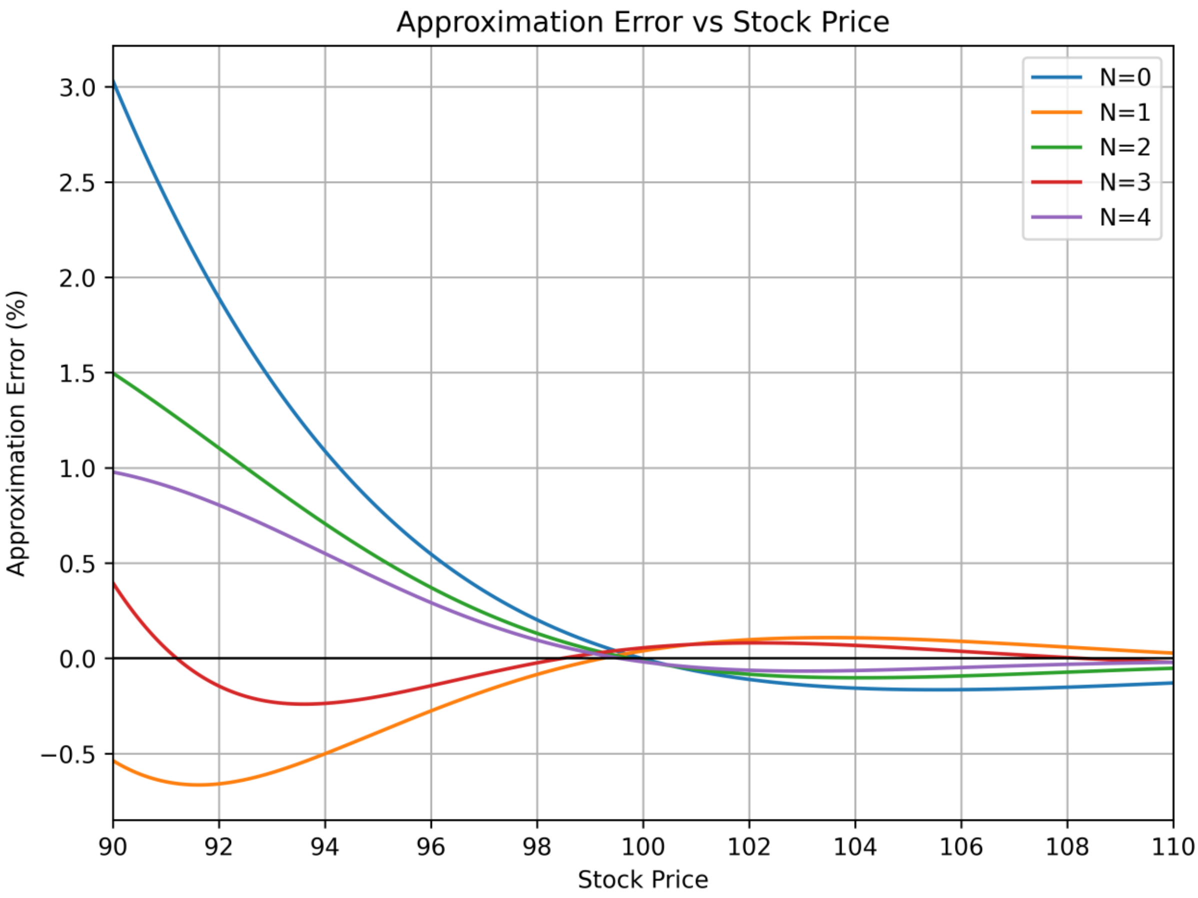

Figure 2.3 – *KM* approximation percentage errors following the *CEV* model defined in equations (2.1) and (2.2).

The parameters were set to $K = 100$, $r = 5\%$, $q = 0\%$, $\tau = 1$ year, $\sigma(x,t) = 10\%$, $\gamma = 0.7$ and

$$\sigma_{CEV} = \frac{\sigma(x,t)}{x^{\gamma-1}} = \frac{10\%}{x^{0.7-1}} = 0.1x^{0.3}.$$

Moreover, the figure 2.3 was generated using the same parameters as figure 2.2, however, this time with a different γ in order to assess the impact of the elasticity parameter in the *KM CEV* method implemented here, and, interestingly, increasing this variable, in this case to $\gamma = 0.7$, leads to a noticeable improvement in the accuracy across all approximation orders, which was a behavior also reported by *Kurz* (2018) in their dissertation. This result highlights the fact that lower γ values imply a stronger dependence of volatility on the underlying asset price x, while γ values closer to one correspond to the *BS* model, where the volatility is assumed to be constant, therefore showing zero elasticity.

Overall, these results seem to suggest that the *KM* approximation method applied to the *CEV* model tends to perform better for higher γ values, i.e., higher elasticity, as the approximation error percentages become more consistent and closer to the benchmark used. This can be explained since in this case, higher γ values cause the volatility of the model of interest to approximate to the one used in the auxiliary model defined using the *BS* model, in turn decreasing the error percentage. However, further analysis would be needed to better understand how this approximation method reacts to other factors.

2.2 Approximating with the Heston model

After examining the method under constant volatility, the analysis in this subchapter serves to compare the stochastic volatility model approximated by *KM*.

In the previous section, it was shown that the original approximation method in equation (1.33) is closely related to the generalized *BSM* model presented in equation (2.1), however, the authors argue that the mispricing function $\delta(x, v, t; \sigma_0)$ defined in equation (1.22) should be able to provide greater accuracy and outperform the generalized *BSM* framework, as it incorporates the stochastic volatility variable v, which the former does not capture. Hence, for approximating a model with stochastic volatility, it is preferable to use the pricing method of equation (1.33) defined as:

$$w_N(x, v, t; \sigma_0) = w_0(x, t; \sigma_0) + \sum_{n=0}^{N} \frac{(T-t)^{n+1}}{(n+1)!} \delta_n(x, v, t; \sigma_0), for\ all\ N \geq 0, \tag{2.21}$$

recalling that $\delta_n(x, v, t; \sigma_0)$ still satisfies the recursive relation below:

$$\delta_{n+1}(x, v, t; \sigma_0) = \mathcal{L}\delta_n(x, v, t; \sigma_0) - r\delta_n(x, v, t; \sigma_0), with\ \delta_0 = \delta. \tag{2.22}$$

and, with the zero-order mispricing function $\delta(x, v, t; \sigma_0)$ as:

$$\delta(x, v, t; \sigma_0) = \frac{1}{2}(v - \sigma_0^2)x^2 \frac{\partial^2}{\partial x^2} w_0(x, t; \sigma_0), \tag{2.23}$$

The following presents the analytical logic for approximating the first-order terms, but for practical purposes, this will later be implemented in *Python* using symbolic computation, since the analytical approach becomes impractical for higher-order approximations.

Having this in mind, the approximation of the equation (1.33) for the first-order, i.e., $N = 1$ is:

$$w_1(x, v, t; \sigma_0) = w_0(x, t; \sigma_0) + \delta(x, v, t; \sigma_0)(T-t) + \delta_1(x, v, t; \sigma_0)\frac{(T-t)^2}{2}, \tag{2.24}$$

since it was expressed earlier that $\delta_0 = \delta$.

The authors advise that the function $\delta(x, v, t; \sigma_0)$ is only based on the last term of the equation (1.54), since the payoffs difference is zero, i.e., $d(x, v) = b(x, v) - b_0(x, v) = 0$, as both *Heston* and

BSM models share the same payoff structure. Additionally, when approximating that last term, some information is inevitably lost, since some stochastic volatility variables do not enter the function, such as κ, α, ρ and ω. This causes the function $\delta(x, v, t; \sigma_0)$ to be quite inaccurate, however, these parameters start appearing in the first-order correction $\delta_1(x, v, t; \sigma_0)$ and, each additional N-th approximation order added incorporates more information about the stochastic volatility process, therefore improving the robustness and gradually increasing the accuracy of the method.

The first-order approximation of the mispricing function $\delta_1(x, v, t; \sigma_0)$[3] can be written as:

$$\begin{aligned}\delta_1(x, v, t; \sigma_0) = \frac{1}{2}(v - \sigma_0^2)\left[(v - \sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2} + 2(v - \sigma_0^2)x^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}(v - \sigma_0^2)x^4\frac{\partial^4 w_0}{\partial x^4}\right] \\ + \frac{1}{2}\kappa(\alpha - v)x^2\frac{\partial^2 w_0}{\partial x^2} + \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right].\end{aligned} \tag{2.25}$$

This expression highlights how the stochastic volatility parameters enter the approximation for the first time, particularly, the drift variables κ and α, as well as the volatility of volatility parameter ω, which only contributes through its interaction with the correlation factor ρ. It is also important to note that this formula differs from the one in the original paper, in the sense that $\delta_1(x, v, t; \sigma_0)$ does not depend explicitly on the time variable t, as the third-order mixed partial derivative $\frac{\partial^3 w_0}{\partial x^2 \partial t}$, which appears in the original derivation, has been simplified.

Additionally, such as table 2.1 from the *CEV* section, it is also possible to obtain the pricing error iterations up to the N-th order using the function $\delta_{n+1}(x, v, t; \sigma_0)$ from the equation (2.22).

[3] See the appendix in section C – Calculations of the first-order mispricing function in Heston, for the auxiliary calculations.

Table 2.2 – *Heston* model iterations of the pricing error function.

n	Pricing error function δ_n iterations
0	$\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^2 w_0(x,t;\sigma_0)}{\partial x^2}$
1	$\frac{\partial\delta_0}{\partial t}+(r-q)x\frac{\partial\delta_0}{\partial x}+\kappa(\alpha-v)\frac{\partial\delta_0}{\partial v}+\frac{1}{2}vx^2\frac{\partial^2\delta_0}{\partial x^2}+\rho\omega\left(v^{\xi+\frac{1}{2}}\right)x\frac{\partial^2\delta_0}{\partial x\partial v}-r\delta_0$
2	$\frac{\partial\delta_1}{\partial t}+(r-q)x\frac{\partial\delta_1}{\partial x}+\kappa(\alpha-v)\frac{\partial\delta_1}{\partial v}+\frac{1}{2}vx^2\frac{\partial^2\delta_1}{\partial x^2}+\frac{1}{2}\omega^2v^{2\xi}\frac{\partial^2\delta_1}{\partial v^2}+\rho\omega\left(v^{\xi+\frac{1}{2}}\right)x\frac{\partial^2\delta_1}{\partial x\partial v}-r\delta_1$
⋮	⋮
N	$\frac{\partial\delta_{N-1}}{\partial t}+(r-q)x\frac{\partial\delta_{N-1}}{\partial x}+\kappa(\alpha-v)\frac{\partial\delta_{N-1}}{\partial v}+\frac{1}{2}vx^2\frac{\partial^2\delta_{N-1}}{\partial x^2}+\frac{1}{2}\omega^2v^{2\xi}\frac{\partial^2\delta_{N-1}}{\partial v^2}+\rho\omega\left(v^{\xi+\frac{1}{2}}\right)x\frac{\partial^2\delta_{N-1}}{\partial x\partial v}-r\delta_{N-1}$

Additionally, it is also worth noting that the term $\frac{1}{2}\omega^2v^{2\xi}\frac{\partial^2\delta_0}{\partial v^2}$ does not exist for $\delta_1(x,v,t;\sigma_0)$ since there is not any partial derivative of $\mathcal{L}$ in equation (1.11), with respect to v^2 that is applied to the zero order pricing error equation $\delta_0(x,v,t;\sigma_0)$. This can also be verified in the equation (C.6) of the auxiliary calculations, as $\frac{\partial^2}{\partial v^2}\delta=0$.

Moreover, as per the *CEV* analysis, in order to achieve more precise results, it is also possible to implement a calibration process to obtain a more accurate volatility parameter, σ. The authors propose determining this parameter by minimizing the squared difference between the prices of the model of interest and the auxiliary model, as shown below, however, they ultimately adopt the simpler approach, such that $\sigma_0=\sqrt{v}$. This assumption is reasonable, but as seen below, the results with this implementation produce an even more precise outcome than those reported in the paper, closely matching the benchmark.

$$\hat{\sigma}_N(x,v,t)=arg\ \min_{\sigma}\left(w_N(x,v,t;\sigma)-w_0(x,t;\sigma)\right)^2.$$

or equivalently,

$$\hat{\sigma}_N(x,v,t)=arg\ \min_{\sigma}\left(\sum_{n=0}^{N}\frac{(T-t)^{n+1}}{(n+1)!}\delta_n(x,v,t;\sigma)\right)^2. \tag{2.26}$$

Having the functions defined, it is now possible to implement *KM's* approximation in *Python*. As with the *CEV* method, since higher-order approximations become increasingly more difficult to handle and are more computationally demanding, the *Sympy* library was also used, as it provides the necessary tools to handle complex equations with multiple approximation orders.

In the context of the *Heston* model, where the stock price x follows the equations (1.6) and (1.7), and with the *CEV* parameter naturally satisfying *Heston's* assumption $\xi = 0.5$, *KM's* approximation method was computed up to the fourth-order, i.e., $N = 4$. Beyond this point, higher-order terms not only increased computational demand but also introduced numerical instability, leading to a loss of precision.

To validate these results, the *KM* approximations were compared against a benchmark constructed using the *Madan and Carr* (1999) fast *Fourier* transform method (also known as *FFT*), which is widely regarded as a reliable and accurate approach for pricing options under the *Heston* framework. Therefore, *FFT* price values obtained were treated as the "true" market prices for the purpose of this comparison.

Furthermore, the relative accuracy of the *KM's* approximation method was also assessed by computing the percentage error at each stock price, according to the following expression:

$$Difference\ (\%) = \left(\frac{KM\ Prices}{FFT\ Prices} - 1\right) \times 100. \tag{2.27}$$

This metric is quite useful as it allows to quantify whether the approximation overestimates or underestimates the benchmark values and by how much, making it possible to directly compare the method implemented in this dissertation to the results obtained by the authors in their paper.

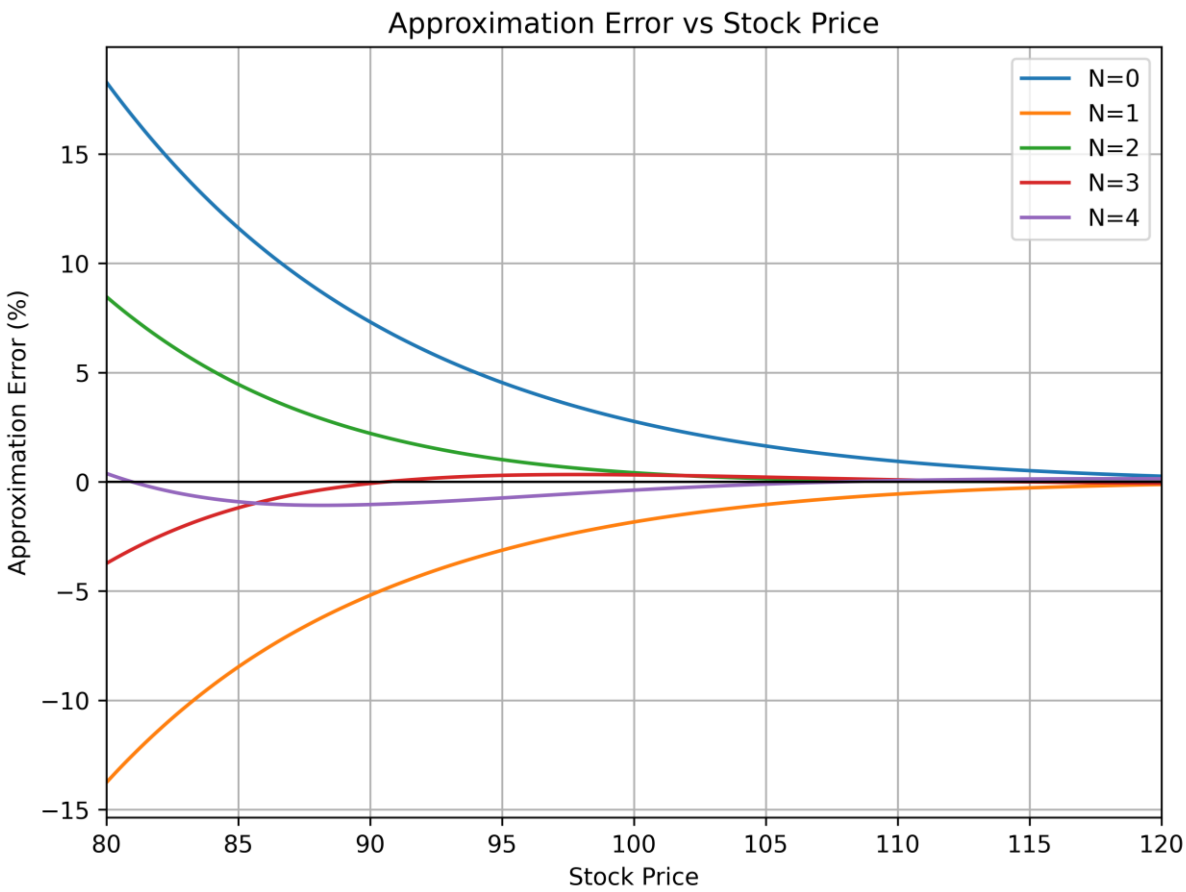

Figure 2.4 – *KM* approximation percentage errors following the *Heston* model defined in equations (1.6) and (1.7).

The parameters were set to $K = 100$, $r = 10\%$, $q = 0\%$, $\tau = 1$ year, $v = 5\%$, $\kappa = 2$, $\alpha = 4\%$, $\omega = 10\%$, $\rho = -0.5$, $\xi = 0.5$ and $\sigma_0 = \sqrt{v} = \sqrt{5\%}$.

Similar to the figure in *KM's* paper, when the stock prices increase, the errors percentage tends to converge to zero, which is the expected behavior as *KM's* method becomes closer to the "true" prices, obtained via the *FFT* benchmark. Additionally, as seen in the figure, when adding more orders, the function tends to improve its precision as the lines become closer to zero. However, in this case, the most precise approximation order is the third-order approximation, as the error percentage is the closest to zero for the majority of the stock prices selected. Therefore, the subsequent tables will present results using $N = 3$, as it delivers the most accurate results for the parameters set.

Proceeding to analyze the approximation prices over a different range of stock values, the first step is to compute the optimal volatility parameter σ_0, in order to achieve the most accurate price against the benchmark, however, for the implementation in this dissertation, it was chosen to use the optimized volatility based on the best approximation order, which in this case is $N = 3$, since recalculating the optimal volatility for each iteration was computationally heavy, and did not provide many benefits. Therefore, using equation (2.26), the optimal volatility was found to be $\sigma_0 \approx 0.71744383$, which differs slightly from the volatility chosen by the authors of $\sigma_0 = \sqrt{v} = \sqrt{0.5172} \approx 0.71916618$. Moreover, as shown in the tables below, this small difference in σ_0 becomes more pronounced in the *KM* method results with varying stock prices, particularly as the approximation order N increases, by improving accuracy and reliability compared to the benchmark.

The table below compares the benchmark prices in the first column, computed using the *FFT* pricing method, followed by the *KM* approximation results and corresponding percentage differences reported in their paper, which also follows the same error formula as equation (2.27) for calculating the difference against the benchmark, and finally, the last column presents this dissertation's *KM* approximation results, together with their percentage differences against the benchmark.

Table 2.3 – Comparison of the *Heston* model prices for varying stock prices.
The methods used were the *FFT* benchmark, *KM* approximation in their paper, and this dissertation's approximation, with parameters set to $K = 1000$, $r = q = 0$, $\tau = 1/12$ years, $v = 0.5172$, $\kappa = 0.1465$, $\alpha = 0.5172$, $\omega = 0.5786$, $\rho = -0.0243$, $\xi = 0.5$ and $\sigma_0 \approx 0.71744383$.

	***FFT* Benchmark**	**Paper approximation**		**This dissertation approximation**	
Stock	Price	Price	% Difference	Price	% Difference
950	57.8425	57.8449	0.0042	57.8434	0.0015
960	62.3711	62.3738	0.0043	62.3719	0.0013
970	67.1005	67.1033	0.0042	67.1012	0.0010
980	72.0291	72.0321	0.0042	72.0297	0.0008
990	77.1553	77.1584	0.0040	77.1557	0.0005
1000	82.4766	82.4797	0.0038	82.4768	0.0003
1010	87.9903	87.9934	0.0036	87.9903	0.0000
1020	93.6933	93.6964	0.0033	93.6931	-0.0002
1030	99.5822	99.5852	0.0030	99.5818	-0.0004
1040	105.6532	105.6560	0.0027	105.6525	-0.0006
1050	111.9021	111.9048	0.0023	111.9013	-0.0008

As observed, the pricing errors obtained in this dissertation are significantly smaller than those reported in *KM's* paper, particularly for lower stock prices. However, contrary to their results, which tend to improve at higher stock prices, the percentage error here starts slightly higher, with an initial overestimation that converges to the benchmark price around $x = 1010$, before gradually shifting to a slight underestimation. This highlights an important aspect of the method implemented in this dissertation, as while *KM's* approximation in the paper systematically overestimates the European-style call prices relative to the benchmark, the implementation in this dissertation overestimates for lower stock prices and then transitions to underestimate for higher call prices.

This difference could be attributed to the calibration method, which shifts the direction of the estimation, or potentially due to implementation differences. This being said, both approaches still produce very close results, with some slight deviations of only fractions of a basis point, which can be economically insignificant for lower-order approximations. However, as previously mentioned, these discrepancies can accumulate when higher-order terms are introduced, which can potentially lead to meaningful differences, particularly in more complex situations, when option prices show greater sensitivity to model parameters.

The next table also follows the same comparison structure however, this time, for at-the-money European-style call prices, since the stock price is the same as the strike, i.e., $x = K$, only changing the variance parameter for each line.

Table 2.4 – Comparison of the *Heston* model prices for varying variances.
The methods used were the *FFT* benchmark, *KM* approximation in their paper, and this dissertation's approximation, with parameters set to $x = K = 1000$, $r = q = 0$, $\tau = 1/12$ years, $\kappa = 0.1465$, $\alpha = 0.5172$, $\omega = 0.5786$, $\rho = -0.0243$, $\xi = 0.5$ and $\sigma_0 = \sqrt{v} = \sqrt{0.5172}$.

	***FFT* Benchmark**	**Paper approximation**		**This dissertation approximation**	
Variance	Price	Price	% Difference	Price	% Difference
0.1	36.4488	36.4854	0.1004	36.4337	-0.0414
0.2	51.4125	51.4255	0.0253	51.4087	-0.0074
0.3	62.8997	62.9068	0.0113	62.8984	-0.0020
0.4	72.5792	72.5838	0.0063	72.5789	-0.0004
0.5	81.1007	81.1040	0.0041	81.1008	0.0002
0.6	88.7981	88.8006	0.0028	88.7985	0.0004
0.7	95.8702	95.8721	0.0021	95.8707	0.0006
0.8	102.4465	102.4481	0.0016	102.4471	0.0006
0.9	108.6171	108.6184	0.0013	108.6178	0.0006
1.0	114.4477	114.4488	0.0010	114.4484	0.0006
1.1	119.9878	119.9888	0.0008	119.9886	0.0006

Similar to table 2.3, the implementation in this dissertation proves to be more accurate than the paper's approximation when the variance levels vary. While *KM's* percentage errors converge smoothly towards zero at higher variances, which is a trend also observed previously with varying stock prices, this implementation shows a slight divergence, though the differences remain extremely small.

An additional observation is the direction of the estimation, as *KM's* approximation method continues to systematically overestimating call prices across all variance levels, this implementation starts underestimating at lower variances but then gradually shifts into a mild overestimation as variance increases, which interestingly is the opposite as seen with varying stock prices. This shift may reflect the effect of the chosen calibration for σ_0, which influences the correction terms differently across variance levels.

As seen previously, from a practical perspective, the absolute errors in both approaches are fractions of a basis point, meaning that both methods are numerically very stable and remain economically insignificant for lower-order approximations. Still, as mentioned before the different convergence patterns may have more pronounced implications for higher-order terms or other parameters, where the approximation errors become more significant.

Conclusion

The objective of this dissertation was to explain the foundations of the *KM* approximation method, including the underlying processes and model definitions, and provide a brief literary review and comparison to other expansion methods, such as *Yang's* asset price expansion, perturbation methods, and risk-neutral probabilities, and later analyze and implement it for both the standard *CEV* and *Heston* models, with the latter being the main focus in this dissertation. Finally, both implementations were compared against the "true" market prices of the chosen benchmarks, in this case using the standard *CEV* model and the *Madan and Carr* (1999) fast *Fourier* transform, respectively, across a wide range of parameters for European-style call options.

Additionally, the method implemented in this dissertation took advantage of a calibration function suggested by the original authors, which minimized the volatility of the pricing equations in order to obtain the most precise computation for the parameters analyzed. This minimizer function was particularly useful for the *Heston* model as it was able to produce better results in comparison to the reported in the original paper.

Another point that makes this dissertation stand out from the original paper is that it managed to derive a further simplification of the mispricing function, which is not dependent of the time variable. This provides a clever solution that could open opportunities for further research on this approximation method.

Concerning the *KM* approximation method applied to the *CEV* model, since the authors did not explicitly describe in their original paper how they implemented their application to *CEV*, although it seems to suggest that their approach follows a *CEV* model with a mean-reverting process, it was not possible to fully replicate and verify their findings, hence the results obtained in this dissertation differ significantly. That being said, it was observed that the method implemented here had some inherent difficulties and was not able to compute approximation orders beyond $N > 4$ since it was too computationally demanding and, additionally, the percentage errors produced were considerably higher. Nonetheless, it was found that the best performing approximation order was generally $N = 3$, where the error percentage decreased significantly for in-the-money call prices, and that this particular implementation benefited from slight adjusting parameters, such as using higher γ values and longer maturity, which further reduced the error levels significantly. Unfortunately, it was not possible to conduct the second planned analysis, in which the *KM* approach would be compared to a *Monte Carlo* benchmark, as in the original paper, in order to evaluate its performance for a range of stock prices and variances.

On the other hand, in contrast to the *CEV* case, the *KM* method applied to the *Heston* model provided much better results, and in some cases could even outperform the original authors' results. In the first analysis, when comparing the approximation errors to the benchmark composed by *Madan and Carr* (1999) *FFT*, it is easily seen that the figure closely matches the one obtained by *KM* in the original paper, where each iteration shows smooth curves that gradually converged to zero for far in-the-money call prices, with the best results obtained being the third-order approximation, i.e., $N = 3$. Additionally, in the second part of the analysis, the method was also compared to the same benchmark, but this time for a different set of parameters, with higher stock and strike prices but displaying a far shorter time-to-maturity, of only one month instead of one year. In this case, all the results were better than those reported by the authors, showing very small deviations from the benchmark, however, while the *KM* results tended to converge smoothly towards zero with higher stock prices, this implementation initially overestimated call prices before reaching its highest accuracy at around $x = 1010$, after which it shifted to negative percentage errors, indicating a slight underestimation. Moreover, a similar pattern was observed when comparing the approximation errors for at-the-money call prices with varying variance levels, as the error results in this implementation still remained quite more accurate than those in the original paper and continued to improve for higher stock prices. However, this time the trend appeared to be reversed, since smaller variances led to a slight underestimation, while higher variances gradually shifted to a mild overestimation, being the most precise values around the variance level of $v = 0.5$, and for $v = 1.1$, the prices were almost identical to the ones reported by the authors, with just a 0.0002 percentual points difference, meaning that possibly for higher variances, the original paper implementation could outperform these results, similarly to what was seen in the analysis across multiple stock prices.

Overall, the *KM's* approach is considered to be a robust and efficient alternative to other popular expansion methods, for both *CEV* and stochastic volatility frameworks, as it is proved to be a reliable and accurate technique, with a high versatility and simple implementation, which makes it particularly appealing when pricing options in multiple frameworks, such as when assuming constant volatility or within multifactor models with stochastic volatility, or even for the computation of the *Greeks* sensitivities. Additionally, since this method directly approximates the price of the model of interest, it also provides a way to analyze the error obtained between the auxiliary and main model, which can be interpreted as the hedging costs associated from the choice of the auxiliary model and could therefore provide valuable information for further analysis.

Nevertheless, this method presents some limitations mainly because it is an approximation method, with its accuracy being strongly dependent on the auxiliary model chosen, which must adequately mimic the key features of the model of interest, such as assuming the same payoff or dividend paying structure, in order to maintain the fidelity to the main model under study. Therefore, if both models are not aligned, the approximation becomes at risk of losing its precision and may eventually diverge from the dynamics of the target model.

Despite these constraints, the *KM* approximation method remains a powerful method that widens the scope of option pricing techniques. Further research could be done to extend the analysis of this approach, as the complexity of the frameworks studied in this dissertation limited a more detailed testing of *CEV* models and the analysis of the term structure of interest rates. Additionally, future research could also explore the application of the *KM* approach to models that admit jumps or barrier options, which are also interesting extensions of *KM's* method that the authors highlighted in their original paper.

References

Aït-Sahalia, Y., & Yu, J. (2006). Saddlepoint approximations for continuous-time Markov processes. *Journal of Econometrics, 134*(2), 507–551. https://doi.org/10.1016/j.jeconom.2005.07.004

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637–654. https://doi.org/10.1086/260062

Boyle, P. P. (1977). Options: A Monte Carlo approach. Journal of Financial Economics, 4(3), 323–338. https://doi.org/10.1016/0304-405x(77)90005-8

Carr, P., & Madan, D. (1999). Option valuation using the fast Fourier transform. *The Journal of Computational Finance, 2*(4), 61–73. https://doi.org/10.21314/jcf.1999.043

Corielli, F. (2006). Hedging with energy. *Mathematical Finance, 16*(3), 495–517. https://doi.org/10.1111/j.1467-9965.2006.00280.x

Cox, J. C. (1975). *Notes on option pricing I: Constant elasticity of variance diffusions* [Unpublished manuscript]. Sloan School of Management, Massachusetts Institute of Technology

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics, 3*(1–2), 145–166. https://doi.org/10.1016/0304-405x(76)90023-4

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53*(2), 385–407. https://doi.org/10.2307/1911242

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics, 25*(4), 631–650. https://doi.org/10.1214/aoms/1177728652

Dias, J. C., Nunes, J. P., & da Silva, F. C. (2024). Finite maturity caps and floors on continuous flows under the constant elasticity of variance process. *European Journal of Operational Research, 316*(1), 361–385. https://doi.org/10.1016/j.ejor.2024.01.039

El Karoui, N., Jeanblanc-Picqué, M., & Shreve, S. E. (1998). Robustness of the Black and Scholes formula. *Mathematical Finance, 8*(2), 93–126. https://doi.org/10.1111/1467-9965.00047

Figlewski, S., & Gao, B. (1999). The adaptive mesh model: A new approach to efficient option pricing. *Journal of Financial Economics, 53*(3), 313–351. https://doi.org/10.1016/s0304-405x(99)00024-0

Fouque, J.-P., Papanicolaou, G., & Sircar, K. R. (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge University Press.

Fouque, J.-P., Papanicolaou, G., Sircar, R., & Solna, K. (2003). Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation, 2*(1), 22–42. https://doi.org/10.1137/030600291

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*(2), 327–343. https://doi.org/10.1093/rfs/6.2.327

Hull, J., & White, A. (1990). Valuing derivative securities using the explicit finite difference method. *The Journal of Financial and Quantitative Analysis, 25*(1), 87–100. https://doi.org/10.2307/2330889

Karatzas, I., & Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer.

Kristensen, D., & Mele, A. (2011). Adding and subtracting Black-Scholes: A new approach to approximating derivative prices in continuous-time models. *Journal of Financial Economics, 102*(2), 390–415. https://doi.org/10.1016/j.jfineco.2011.05.007

Kurz, M. (2018). *Closed-form approximations in derivatives pricing: The Kristensen-Mele approach*. arXiv. https://doi.org/10.48550/arXiv.1804.08904

Lewis, A. L. (2000). *Option valuation under stochastic volatility: With Mathematica Code*. Finance Press.

Lord, R., & Kahl, C. (2007). Optimal Fourier inversion in semi-analytical option pricing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.921336

Rouah, F. D. (2013). *The Heston model and its extensions in MATLAB and C#*. Wiley Finance Series, John Wiley & Sons, Inc.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *The Journal of Finance, 44*(1), 211–219. https://doi.org/10.2307/2328285

Schwartz, E. S. (1977). The valuation of warrants: Implementing a new approach. *Journal of Financial Economics, 4*(1), 79–93. https://doi.org/10.1016/0304-405x(77)90037-x

Scott, L. O. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of Fourier inversion methods. *Mathematical Finance, 7*(4), 413–426. https://doi.org/10.1111/1467-9965.00039

Appendix

A – Calculation of the infinitesimal generator $\mathcal{L}$

In order to create the infinitesimal generator $\mathcal{L}$, which is needed to start the expansion of *KM's* approximation method, it is first necessary to expand the equation (1.10) with the support of the processes defined in equations (1.6) and (1.7).

Assuming that $S(t) = x$ then $dS(t) = dx$, and, hence, it is possible to calculate the derivative terms dx and dv of the pricing function $w(S, v, t)$.

Firstly, expanding the square of the new dx term and applying *Itô's* lemma:

$$\begin{aligned} dx &= (r-q)xdt + x\sqrt{v}dW_1(t) \Leftrightarrow \\ \Leftrightarrow (dx)^2 &= [(r-q)xdt]^2 + 2\big[(r-q)xdt \times x\sqrt{v}dW_1(t)\big] + \Big(x\sqrt{v}dW_1(t)\Big)^2 \\ &= [(r-q)x]^2(dt)^2 + 2(r-q)x^2\sqrt{v}dtdW_1(t) + x^2v\big(dW_1(t)\big)^2. \end{aligned} \tag{A.1}$$

And, according to *Itô's* lemma, as the increment of time become infinitesimally small, i.e., as $dt \to 0$, $(dt)^2$ will become negligible, since $(dt)^2 \approx 0$, and, due to the quadratic variation of a *Wiener* process, the squared term of $dW(t) = \sqrt{dt}$, becomes $\big(dW(t)\big)^2 \approx dt$. Having this in mind, it can also be assumed that $dtdW(t) \approx dt\sqrt{dt} = dt^{\frac{3}{2}} \approx 0$, so the equation (A.1) can be reduced to:

$$(dx)^2 = x^2v\big(dW_1(t)\big)^2 = vx^2dt. \tag{A.2}$$

Next, expanding the square of the dv term and applying *Itô's* lemma, while considering the earlier assumptions:

$$\begin{aligned} dv &= \kappa(\alpha - v)dt + \omega|v|^{\xi}dW_2(t) \Leftrightarrow \\ \Leftrightarrow (dv)^2 &= [\kappa(\alpha - v)dt]^2 + 2\big[\kappa(\alpha - v)dt \times \omega v^{\xi}dW_2(t)\big] + \Big(\omega v^{\xi}dW_2(t)\Big)^2 \\ &= [\kappa(\alpha - v)]^2(dt)^2 + 2\kappa(\alpha - v)\omega v^{\xi}dtdW_2(t) + \big(\omega v^{\xi}\big)^2\big(dW_2(t)\big)^2 \\ &= \omega^2 v^{2\xi}dt. \end{aligned} \tag{A.3}$$

The cross term $(dx)(dv)$ can also be expanded following the same principles:

$$\begin{aligned} (dx)(dv) &= \Big((r-q)xdt + x\sqrt{v}dW_1(t)\Big)\Big(\kappa(\alpha - v)dt + \omega v^{\xi}dW_2(t)\Big) \\ &= \big((r-q)x\kappa(\alpha - v)(dt)^2\big) + \Big((r-q)x\omega v^{\xi}\,dtdW_2(t)\Big) \\ &\quad + \Big(x\sqrt{v}\kappa(\alpha - v)dtdW_1(t)\Big) + \Big(x\omega v^{\xi}\sqrt{v}dW_1(t)dW_2(t)\Big), \end{aligned} \tag{A.4}$$

and, recalling equation (1.9) for the correlation of two *Brownian* motions, the expansion can be developed as:

$$(dx)(dv) = \left(x\omega v^{\xi}\sqrt{v}dW_1(t)dW_2(t)\right) = x\omega v^{\xi}\sqrt{v}\rho dt = \rho\omega\left(v^{\xi+\frac{1}{2}}\right)xdt. \tag{A.5}$$

Lastly, the infinitesimal generator $\mathcal{L}$, which defines the *PDE* satisfied by the pricing function $w(x,v,t)$ under the risk-neutral measure, can be obtained by substituting the equations (A.2) to (A.5) into the *Itô's* expansion of dw defined in equation (1.10):

$$\begin{aligned}\mathcal{L}w(x,v,t) = {} & \frac{\partial w}{\partial t} + (r-q)x\frac{\partial w}{\partial x} + \kappa(\alpha - v)\frac{\partial w}{\partial v} + \frac{1}{2}vx^2\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}\omega^2 v^{2\xi}\frac{\partial^2 w}{\partial v^2} \\ & + \rho\omega\left(v^{\xi+\frac{1}{2}}\right)x\frac{\partial^2 w}{\partial x \partial v}.\end{aligned} \tag{A.6}$$

B – Calculations of the first-order mispricing function in *CEV*

To approximate the pricing error function to the first-order approximation, i.e., $\delta_1(x,t;\sigma_0)$, assuming a *CEV* model, it is first necessary to apply this mispricing function to the infinitesimal generator $\mathcal{L}$ defined in equation (2.11). Since the first iteration, i.e., $n = 0$, is already provided, when substituting σ^2 for the volatility scaling function in equation (2.2) it becomes:

$$\begin{aligned}\delta(x,t;\sigma_0) &= \frac{1}{2}(\sigma^2 - \sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2} \\ &= \frac{1}{2}(\sigma_{CEV}^2 x^{2\gamma-2} - \sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2} \\ &= \frac{1}{2}(\sigma_{CEV}^2 x^{2\gamma} - \sigma_0^2 x^2)\frac{\partial^2 w_0}{\partial x^2} \\ &= \frac{1}{2}a(x)\frac{\partial^2 w_0}{\partial x^2} \; with \; \delta_0 = \delta.\end{aligned} \tag{B.1}$$

The new definition of $\sigma_{CEV}^2 x^{2\gamma} - \sigma_0^2 x^2 = a(x)$ is useful in order to simplify the expressions, due to the complexity of the partial derivatives in the next calculations. Therefore, the first and second partial derivatives terms of $a(x)$ with respect to x become:

$$\frac{\partial}{\partial x}a(x) = \frac{\partial}{\partial x}(\sigma_{CEV}^2 x^{2\gamma} - \sigma_0^2 x^2) = 2\gamma\sigma_{CEV}^2 x^{2\gamma-1} - 2\sigma_0^2 x, \tag{B.2}$$

$$\frac{\partial^2}{\partial x^2}a(x) = \frac{\partial}{\partial x}(2\gamma\sigma_{CEV}^2 x^{2\gamma-1} - 2\sigma_0^2 x) = 2\gamma(2\gamma-1)\sigma_{CEV}^2 x^{2\gamma-2} - 2\sigma_0^2. \tag{B.3}$$

Having these defined, it is now possible to start the calculations for the $n = 1$ iteration in equation $\delta_1(x,t;\sigma_0)$:

$$\delta_1(x,t;\sigma_0) = \mathcal{L}\delta_0(x,t;\sigma_0) - r\delta_0(x,t;\sigma_0). \tag{B.4}$$

To start applying $\delta_0(x,t;\sigma_0)$ to the infinitesimal generator $\mathcal{L}$, it is first necessary to expand the function mispricing function to each partial derivate of $\mathcal{L}$, therefore, the partial derivative with respect to t becomes:

$$\frac{\partial}{\partial t}\delta = \frac{\partial}{\partial t}\left(\frac{1}{2}a(x)\frac{\partial^2 w_0}{\partial x^2}\right) = \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^2 \partial t}, \tag{B.5}$$

and the first and second-order partial derivatives with respect to x become:

$$\begin{aligned}\frac{\partial}{\partial x}\delta &= \frac{\partial}{\partial x}\left(\frac{1}{2}a(x)\frac{\partial^2 w_0}{\partial x^2}\right)\\ &= \frac{\partial}{\partial x}\left(\frac{1}{2}a(x)\right)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}a(x)\frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\\ &= \frac{1}{2}a'(x)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^3},\end{aligned} \tag{B.6}$$

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\delta &= \frac{\partial}{\partial x}\left(\frac{1}{2}a'(x)\frac{\partial^2 w_0}{\partial x^2}\right) + \frac{\partial}{\partial x}\left(\frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^3}\right)\\ &= \left[\frac{\partial}{\partial x}\left(\frac{1}{2}a'(x)\right)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}a'(x)\frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\right]\\ &\qquad + \left[\frac{\partial}{\partial x}\left(\frac{1}{2}a(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}a(x)\frac{\partial}{\partial x}\left(\frac{\partial^3 w_0}{\partial x^3}\right)\right]\\ &= \frac{1}{2}a''(x)\frac{\partial^2 w_0}{\partial x^2} + a'(x)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}a(x)\frac{\partial^4 w_0}{\partial x^4}.\end{aligned} \tag{B.7}$$

Therefore, when constructing the first term of the mispricing function, $\mathcal{L}\delta_0(x,t;\sigma_0)$ becomes:

$$\begin{aligned}\mathcal{L}\delta_0(x,t;\sigma_0) &= \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^2 \partial t} + \frac{1}{2}(r-q)a'(x)x\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(r-q)a(x)x\frac{\partial^3 w_0}{\partial x^3}\\ &\quad + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma}a''(x)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}a'(x)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma}a(x)\frac{\partial^4 w_0}{\partial x^4},\end{aligned} \tag{B.8}$$

which can be simplified to:

$$\begin{aligned}\mathcal{L}\delta_0(x,t;\sigma_0) &= \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^2 \partial t} + \left(\frac{1}{2}(r-q)a'(x)x + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma}a''(x)\right)\frac{\partial^2 w_0}{\partial x^2}\\ &\quad + \left(\frac{1}{2}(r-q)a(x)x + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma}a(x)\frac{\partial^4 w_0}{\partial x^4},\end{aligned} \tag{B.9}$$

and now, it is possible to construct the function $\delta_1(x,t;\sigma_0)$ defined in equation (B.4):

$$\delta_1(x,t;\sigma_0) = \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^2 \partial t} + \left(\frac{1}{2}(r-q)a'(x)x + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a''(x)\right)\frac{\partial^2 w_0}{\partial x^2}$$
$$+ \left(\frac{1}{2}(r-q)a(x)x + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a(x)\frac{\partial^4 w_0}{\partial x^4}$$
$$- r\delta_0(x,t;\sigma_0), \tag{B.10}$$

or equivalently,

$$\delta_1(x,t;\sigma_0) = \frac{1}{2}a(x)\frac{\partial^3 w_0}{\partial x^2 \partial t} + \left(\frac{1}{2}(r-q)a'(x)x + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a''(x) - \frac{1}{2}a(x)r\right)\frac{\partial^2 w_0}{\partial x^2}$$
$$+ \left(\frac{1}{2}(r-q)a(x)x + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a(x)\frac{\partial^4 w_0}{\partial x^4}.$$

Another possible simplification of the equation $\delta_1(x,t;\sigma_0)$ in (B.10), could be reducing the partial derivative term dependent of the time variable t, defined previously as $\frac{\partial^3 w_0}{\partial x^2 \partial t}$. For that, it is necessary to recall the *PDE* equation (2.12) which, by rearranging the terms, becomes:

$$\frac{\partial w_0}{\partial t} = -(r-q)x\frac{\partial w_0}{\partial x} - \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^2 w_0}{\partial x^2} + rw_0. \tag{B.11}$$

The left-hand side of this expression corresponds directly to the third-order mixed partial derivative $\frac{\partial^3 w_0}{\partial x^2 \partial t}$, once it is differentiated twice in order of x, which does provide a convenient substitution that helps with the simplification of the time dependent third-order mixed partial derivative in the expansion of the mispricing function.

Then, differentiating twice the term $\frac{\partial w_0}{\partial t}$ in order to match with the term in $\delta_1(x,t;\sigma_0)$ gives:

$$\frac{\partial^2}{\partial x^2}\frac{\partial w_0}{\partial t} = \frac{\partial^3 w_0}{\partial x^2 \partial t}. \tag{B.12}$$

The term $(r-q)x\frac{\partial w_0}{\partial x}$ gives:

$$\frac{\partial^2}{\partial x^2}\left((r-q)x\frac{\partial w_0}{\partial x}\right) = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}\big((r-q)x\big)\frac{\partial w_0}{\partial x} + (r-q)x\frac{\partial}{\partial x}\left(\frac{\partial w_0}{\partial x}\right)\right]$$
$$= \frac{\partial}{\partial x}\left((r-q)\frac{\partial w_0}{\partial x} + (r-q)x\frac{\partial^2 w_0}{\partial x^2}\right) \tag{B.13}$$

$$= \frac{\partial}{\partial x}(r-q)\frac{\partial w_0}{\partial x} + (r-q)\frac{\partial}{\partial x}\left(\frac{\partial w_0}{\partial x}\right) + \frac{\partial}{\partial x}\big((r-q)x\big)\frac{\partial^2 w_0}{\partial x^2} + (r-q)x\frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)$$
$$= (r-q)\frac{\partial^2 w_0}{\partial x^2} + (r-q)\frac{\partial^2 w_0}{\partial x^2} + (r-q)x\frac{\partial^3 w_0}{\partial x^3}$$

$$= 2(r-q)\frac{\partial^2 w_0}{\partial x^2} + (r-q)x\frac{\partial^3 w_0}{\partial x^3}.$$

The term $\frac{1}{2}\sigma_{CEV}^2 x^{2\gamma}\frac{\partial^2 w_0}{\partial x^2}$ gives:

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\left(\frac{1}{2}\sigma_0^2 x^2 \frac{\partial^2 w_0}{\partial x^2}\right) &= \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}\left(\frac{1}{2}\sigma_0^2 x^2\right)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\right] \\
&= \frac{\partial}{\partial x}\left(\sigma_0^2 x \frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^3 w_0}{\partial x^3}\right) \\
&= \frac{\partial}{\partial x}(\sigma_0^2 x)\frac{\partial^2 w_0}{\partial x^2} + \sigma_0^2 x \frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right) + \frac{\partial}{\partial x}\left(\frac{1}{2}\sigma_0^2 x^2\right)\frac{\partial^3 w_0}{\partial x^3} \\
&\quad + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial}{\partial x}\left(\frac{\partial^3 w_0}{\partial x^3}\right) \\
&= \sigma_0^2 \frac{\partial^2 w_0}{\partial x^2} + \sigma_0^2 x \frac{\partial^3 w_0}{\partial x^3} + \sigma_0^2 x \frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^4 w_0}{\partial x^4} \\
&= \sigma_0^2 \frac{\partial^2 w_0}{\partial x^2} + 2\sigma_0^2 x \frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^4 w_0}{\partial x^4}.
\end{aligned} \tag{B.14}$$

And, lastly, the term rw_0 becomes:

$$\frac{\partial^2}{\partial x^2} r w_0 = r\frac{\partial^2 w_0}{\partial x^2}. \tag{B.15}$$

Having obtained these partial derivative terms, it is now possible to insert them in equation (B.11), which gives:

$$\begin{aligned}
\frac{\partial^3 w_0}{\partial x^2 \partial t} &= -2(r-q)\frac{\partial^2 w_0}{\partial x^2} - (r-q)x\frac{\partial^3 w_0}{\partial x^3} - \sigma_0^2 \frac{\partial^2 w_0}{\partial x^2} - 2\sigma_0^2 x \frac{\partial^3 w_0}{\partial x^3} - \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^4 w_0}{\partial x^4} \\
&\quad + r\frac{\partial^2 w_0}{\partial x^2} \\
&= (2q - r - \sigma_0^2)\frac{\partial^2 w_0}{\partial x^2} + (q - r - 2\sigma_0^2)x\frac{\partial^3 w_0}{\partial x^3} - \frac{1}{2}\sigma_0^2 x^2 \frac{\partial^4 w_0}{\partial x^4}.
\end{aligned} \tag{B.16}$$

Hence, substituting this new term in the equation $\delta_1(x, t; \sigma_0)$ defined in (B.10) becomes:

$$\begin{aligned}
\delta_1(x, t; \sigma_0) &= \frac{1}{2}(2q - r - \sigma_0^2)a(x)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(q - r - 2\sigma_0^2)xa(x)\frac{\partial^3 w_0}{\partial x^3} \\
&\quad - \frac{1}{4}\sigma_0^2 x^2 a(x)\frac{\partial^4 w_0}{\partial x^4} \\
&\quad + \left(\frac{1}{2}rxa'(x) - \frac{1}{2}qxa'(x) + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a''(x) - \frac{1}{2}ra(x)\right)\frac{\partial^2 w_0}{\partial x^2} \\
&\quad + \left(\frac{1}{2}rxa(x) - \frac{1}{2}qxa(x) + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}\sigma_{CEV}^2 x^{2\gamma} a(x)\frac{\partial^4 w_0}{\partial x^4},
\end{aligned} \tag{B.17}$$

which can be further simplified to:

$$\delta_1(x,t;\sigma_0) = \frac{1}{2}\left((2q-2r-\sigma_0^2)a(x) + (r-q)xa'(x) + \frac{1}{2}\sigma_{CEV}^2 x^{2\gamma} a''(x)\right)\frac{\partial^2 w_0}{\partial x^2} - \frac{1}{2}\left(2\sigma_0^2 x a(x) - \sigma_{CEV}^2 x^{2\gamma} a'(x)\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}a^2(x)\frac{\partial^4 w_0}{\partial x^4}. \tag{B.18}$$

C – Calculations of the first-order mispricing function in *Heston*

To approximate the mispricing error function to the first-order approximation, i.e., $\delta_1(x,v,t;\sigma_0)$, which is needed in *Chapter 2.2* to approximate with the *Heston* model, it is necessary to apply the infinitesimal generator $\mathcal{L}$ defined in equation (1.11) to the zero-order function $\delta_0(x,v,t;\sigma_0)$, defined in equation (2.23), such that:

$$\delta_1(x,v,t;\sigma_0) = \mathcal{L}\delta_0(x,v,t;\sigma_0) - r\delta_0(x,v,t;\sigma_0), \tag{C.1}$$

since $\mathcal{L}$ is related to the model of interest while δ is dependent on the auxiliary model. Then, expanding the function $\delta_0(x,v,t;\sigma_0)$ to each partial derivative of $\mathcal{L}$, the partial derivative with respect to t becomes:

$$\frac{\partial}{\partial t}\delta = \frac{\partial}{\partial t}\left(\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2}\right) = \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^3 w_0}{\partial x^2 \partial t}. \tag{C.2}$$

The first and second-order partial derivatives with respect to x are:

$$\begin{aligned}\frac{\partial}{\partial x}\delta &= \frac{\partial}{\partial x}\left(\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2}\right)\\ &= \frac{\partial}{\partial x}\left(\frac{1}{2}(v-\sigma_0^2)x^2\right)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\\ &= (v-\sigma_0^2)x\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^3 w_0}{\partial x^3},\end{aligned} \tag{C.3}$$

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\delta &= \frac{\partial}{\partial x}\left((v-\sigma_0^2)x\frac{\partial^2 w_0}{\partial x^2}\right) + \frac{\partial}{\partial x}\left(\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^3 w_0}{\partial x^3}\right)\\ &= \left[\frac{\partial}{\partial x}\left((v-\sigma_0^2)x\right)\frac{\partial^2 w_0}{\partial x^2} + (v-\sigma_0^2)x\frac{\partial}{\partial x}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\right]\\ &\qquad + \left[\frac{\partial}{\partial x}\left(\frac{1}{2}(v-\sigma_0^2)x^2\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial}{\partial x}\left(\frac{\partial^3 w_0}{\partial x^3}\right)\right]\\ &= (v-\sigma_0^2)\frac{\partial^2 w_0}{\partial x^2} + 2(v-\sigma_0^2)x\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^4 w_0}{\partial x^4}.\end{aligned} \tag{C.4}$$

The first and second-order partial derivatives with respect to v are:

$$\frac{\partial}{\partial v}\delta = \frac{\partial}{\partial v}\left(\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2}\right) \tag{C.5}$$

$$= \frac{\partial}{\partial v}\left(\frac{1}{2}(v-\sigma_0^2)x^2\right)\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial}{\partial v}\left(\frac{\partial^2 w_0}{\partial x^2}\right)$$

$$= \frac{1}{2}x^2\frac{\partial^2 w_0}{\partial x^2},$$

$$\frac{\partial^2}{\partial v^2}\delta = \frac{\partial}{\partial v}\left(\frac{1}{2}x^2\frac{\partial^2 w_0}{\partial x^2}\right) = 0. \tag{C.6}$$

Finally, the second-order mixed partial derivative is:

$$\begin{aligned}\frac{\partial^2}{\partial x\partial v}\delta &= \frac{\partial}{\partial v}\left((v-\sigma_0^2)x\frac{\partial^2 w_0}{\partial x^2}\right) + \frac{\partial}{\partial v}\left(\frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^3 w_0}{\partial x^3}\right)\\ &= \left[\frac{\partial}{\partial v}\left((v-\sigma_0^2)x\right)\frac{\partial^2 w_0}{\partial x^2} + (v-\sigma_0^2)x\frac{\partial}{\partial v}\left(\frac{\partial^2 w_0}{\partial x^2}\right)\right]\\ &\quad + \left[\frac{\partial}{\partial v}\left(\frac{1}{2}(v-\sigma_0^2)x^2\right)\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial}{\partial v}\left(\frac{\partial^3 w_0}{\partial x^3}\right)\right]\\ &= x\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^2\frac{\partial^3 w_0}{\partial x^3}.\end{aligned} \tag{C.7}$$

Hence, constructing the first term of the mispricing function, i.e., $\mathcal{L}\delta_0(x,v,t;\sigma_0)$ it becomes:

$$\begin{aligned}\mathcal{L}\delta_0(x,v,t;\sigma_0) &= \frac{1}{2}(v-\sigma_0^2)x^2\frac{\partial^3 w_0}{\partial x^2\partial t} + (r-q)(v-\sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2}\\ &+ \frac{1}{2}(r-q)(v-\sigma_0^2)x^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}\kappa(\alpha-v)x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}(v-\sigma_0^2)vx^2\frac{\partial^2 w_0}{\partial x^2}\\ &+ (v-\sigma_0^2)vx^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{4}(v-\sigma_0^2)vx^4\frac{\partial^4 w_0}{\partial x^4} + \rho\omega\left(v^{\xi+\frac{1}{2}}\right)x^2\frac{\partial^2 w_0}{\partial x^2}\\ &+ \frac{1}{2}\rho\omega\left(v^{\xi+\frac{1}{2}}\right)x^3\frac{\partial^3 w_0}{\partial x^3},\end{aligned} \tag{C.8}$$

which can be simplified as:

$$\begin{aligned}\mathcal{L}\delta_0(x,v,t;\sigma_0) &= \frac{1}{2}(v-\sigma_0^2)\left[x^2\frac{\partial^3 w_0}{\partial x^2\partial t} + 2(r-q)x^2\frac{\partial^2 w_0}{\partial x^2} + (r-q)x^3\frac{\partial^3 w_0}{\partial x^3} + vx^2\frac{\partial^2 w_0}{\partial x^2}\right.\\ &\left. + 2vx^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}vx^4\frac{\partial^4 w_0}{\partial x^4}\right] + \frac{1}{2}\kappa(\alpha-v)x^2\frac{\partial^2 w_0}{\partial x^2}\\ &+ \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right],\end{aligned} \tag{C.9}$$

and then, constructing and simplifying the first-order approximation of the mispricing error function gives:

$$\delta_1(x,v,t;\sigma_0) = \frac{1}{2}(v-\sigma_0^2)\left[x^2\frac{\partial^3 w_0}{\partial x^2 \partial t} + 2(r-q)x^2\frac{\partial^2 w_0}{\partial x^2} + (r-q)x^3\frac{\partial^3 w_0}{\partial x^3} + vx^2\frac{\partial^2 w_0}{\partial x^2}\right.$$
$$\left. + 2vx^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}vx^4\frac{\partial^4 w_0}{\partial x^4}\right] + \frac{1}{2}\kappa(\alpha - v)x^2\frac{\partial^2 w_0}{\partial x^2}$$
$$+ \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right] - r\delta_0(x,v,t;\sigma_0),$$

or equivalently, (C.10)

$$\delta_1(x,v,t;\sigma_0) = \frac{1}{2}(v-\sigma_0^2)\left[x^2\frac{\partial^3 w_0}{\partial x^2 \partial t} + 2(r-q)x^2\frac{\partial^2 w_0}{\partial x^2} + (r-q)x^3\frac{\partial^3 w_0}{\partial x^3} + vx^2\frac{\partial^2 w_0}{\partial x^2}\right.$$
$$\left. + 2vx^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}vx^4\frac{\partial^4 w_0}{\partial x^4} - rx^2\frac{\partial^2 w_0}{\partial x^2}\right] + \frac{1}{2}\kappa(\alpha - v)x^2\frac{\partial^2 w_0}{\partial x^2}$$
$$+ \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right].$$

Once again, it is also possible to do a further simplification step of $\delta_1(x,v,t;\sigma_0)$, which involves reducing the partial derivative term dependent of the time variable t, defined previously as $\frac{\partial^3 w_0}{\partial x^2 \partial t}$. Therefore, recalling the equation (B.16) in the last subsection, and substituting it in the first-order mispricing function $\delta_1(x,v,t;\sigma_0)$, which was defined in equation (C.10), gives:

$$\delta_1(x,v,t;\sigma_0) = \frac{1}{2}(v-\sigma_0^2)\left[(2q - r - \sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2} + (q - r - 2\sigma_0^2)x^3\frac{\partial^3 w_0}{\partial x^3}\right.$$
$$- \frac{1}{2}\sigma_0^2 x^4\frac{\partial^4 w_0}{\partial x^4} + 2(r-q)x^2\frac{\partial^2 w_0}{\partial x^2} + (r-q)x^3\frac{\partial^3 w_0}{\partial x^3} + vx^2\frac{\partial^2 w_0}{\partial x^2}$$
$$\left. + 2vx^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}vx^4\frac{\partial^4 w_0}{\partial x^4} - rx^2\frac{\partial^2 w_0}{\partial x^2}\right] + \frac{1}{2}\kappa(\alpha - v)x^2\frac{\partial^2 w_0}{\partial x^2} \quad (C.11)$$
$$+ \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right],$$

and further simplifying this equation becomes:

$$\delta_1(x,v,t;\sigma_0) = \frac{1}{2}(v-\sigma_0^2)\left[(v-\sigma_0^2)x^2\frac{\partial^2 w_0}{\partial x^2} + 2(v-\sigma_0^2)x^3\frac{\partial^3 w_0}{\partial x^3} + \frac{1}{2}(v-\sigma_0^2)x^4\frac{\partial^4 w_0}{\partial x^4}\right]$$
$$+ \frac{1}{2}\kappa(\alpha - v)x^2\frac{\partial^2 w_0}{\partial x^2} + \rho\omega\left(v^{\xi+\frac{1}{2}}\right)\left[x^2\frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2}x^3\frac{\partial^3 w_0}{\partial x^3}\right]. \quad (C.12)$$