## **A NOVEL COSINE NETWORK FOR PRICING EUROPEAN AND BARRIER OPTIONS UNDER STOCHASTIC VOLATILITY MODELS**

MSC THESIS



## **A NOVEL COSINE NETWORK FOR PRICING EUROPEAN AND BARRIER OPTIONS UNDER STOCHASTIC VOLATILITY MODELS**

# MSC THESIS

by

# **Felix BANGERTER**

to obtain the degree of Master of Science at the Delft University of Technology, to be defended publicly on November 24, 2023 at 10:00.



An electronic version of this dissertation is available at <http://repository.tudelft.nl/>.





# **ABSTRACT**

<span id="page-2-0"></span>This thesis presents a comprehensive exploration of the rough Heston model as a means to enhance financial derivative pricing and calibration in the context of the complex behavior of market volatility. Recognizing the limitations of classical models, such as the Black-Scholes and the standard Heston model, which assume constant or meanreverting volatility, this research delves into the application of rough volatility models that account for the empirical 'memory' effect observed in financial markets. These models, inspired by the fractional Brownian motion with a Hurst parameter less than 0.5, offer a more accurate representation of the volatility surface.

A significant portion of the thesis is dedicated to the development of a novel cosine tensor network that expedites the supervised learning of the characteristic function of the lifted Heston model. This advancement is pivotal for the rapid pricing and calibration of European options under the rough Heston framework. The cosine tensor network, leveraging the characteristic function's availability, enables the efficient application of Fourier-based methods, such as the COS method, for option pricing. This approach is further extended to the pricing of path-dependent options like barrier and Bermudan options through the 2-dimensional COS method.

The thesis is methodically structured, beginning with a foundational overview of option pricing and volatility, followed by a literature review that situates rough volatility within the broader context of quantitative finance. Subsequent chapters detail the mathematical framework of the rough Heston model, its derivation from market microstructures, and the introduction of the lifted Heston model as a multi-factor approximation.

Empirical analysis is provided to validate the lifted Heston model against traditional methods, demonstrating its superior performance and accuracy. The thesis culminates in the presentation of a new calibration method, supported by real market data, and a novel benchmark for pricing complex derivatives under the rough Heston model.

The research encapsulated in this thesis not only sets a new standard for computational speed and precision in the pricing and calibration of financial derivatives under the rough Heston model but also opens avenues for future research, particularly in the application of rough volatility models to other areas of financial mathematics.

# **PREFACE**

<span id="page-4-0"></span>This thesis not only ends my time at TU Delft but my whole academic career. Hence, I would like to take a moment and express my gratitude the most important persons along the way. Everything started in October 2017 when I made the decision to pursue a bachelors degree in mathematics at the LMU in Munich. Young and naive, I quickly realized studying mathematics is more than Bernoulli experiments and taking derivatives. Needless to say this degree wouldn't have been possible without my incredible friends Jan-Jürgen, Leo and Josel and our memorable carnival trips. At this moment, special Kudus to Lukas, who I know since high school! Thanks for your incredible friendship.

Coming to the Netherlands to pursue my master degree at TU Delft and leaving Munich behind has been a challenge, but I cannot express how grateful I am for the past two years. First, I would really like to thank my supervisors, Dr. Fang and Dr. Shen, for the possibility to work on this project at FF Quant. Without your help this wouldn't have been possible and I'm really grateful to have been able to work at FFQuant for the past 9 months. It's been a fantastic work environment and even though the thesis project had its challenging periods, I was always able to count on my supervisors and Asian Fantasy. To Marnix, I'm thankful for our lovely Friday afterwork sessions and your help at the COS-CPD method, sometimes even at 9pm on a Sunday. To Sofia and Levi, who were part of my Delft journey from day 1. Without you guys life in Delft wouldn't have been as fun. Honorable mention to Zoé for our CC&C lunches and Gijs for letting me have a proper smile every Tuesday with a good "Morge Gijs".

Not to forget the two most important people throughout my Delft-life. I can't thank my roommates Robin and Arya enough for the time here in Delft. The two best roommates one could ask for and I will always keep the time at VB24 in mind. I know we grew a friendship for life!

In addition, I would like to express my gratitude to two very special people. First, my very best friend Adrian, You were not only part of my academic journey but throughout my whole life! I know you for 26 years and certainly for the rest of my life. Our friendship was and will always be special! Secondly, thanks to my incredible girlfriend Lucie. Your support in the past year has been more than special and I cannot thank you enough for being the way you are!

Finally, I would like to thank my family, whose unconditionally support and caring certainly helped me through my academic career, and especially through this thesis project. Words - and certainly not these few lines - are not enough to describe how important your support was (and is) to me. One thing is for sure, I wouldn't be here without you having my back all the time! To my brother, you're just the best ever! Last but not least, to cite the great band Migos, "Mama, we made it!".

> *Felix Bangerter Delft, November 2023*

# **LIST OF FIGURES**

<span id="page-6-0"></span>





# **LIST OF TABLES**

<span id="page-10-0"></span>

# **CONTENTS**







# **1**

# **INTRODUCTION**

<span id="page-16-0"></span>Financial markets have always been the epicenter of global economic dynamics, with their behavior influencing decisions ranging from individual investments to national fiscal policies. Understanding, predicting, and quantifying market volatility stands as one of the paramount challenges in the realm of quantitative finance. Over the decades, an assortment of models have been proposed to capture the intricacies of this volatility, each with its own strengths and weaknesses.

The classical Black-Scholes model, which paved the way for modern financial mathematics, assumed constant volatility—a simplification that would soon be challenged by real market behaviors. Subsequent models like the Heston model and stochastic volatility models brought randomness to the volatility itself, offering a more realistic portrayal of market dynamics. However, while these models were more sophisticated, they still often failed to completely capture the rough and irregular nature of actual market volatility.

Enter the concept of rough volatility. Pioneered by researchers like [\[1\]](#page-112-0),[\[2\]](#page-112-1) or [\[3\]](#page-112-2), rough volatility models build upon the observation that the volatility surface in financial markets exhibits behavior akin to fractional Brownian motion, particularly with a Hurst parameter less than 0.5.  $H < 0.5$  suggests that volatility exhibits a sort of "memory" effect. This means that large moves tend to be followed by small moves and small moves by large moves, with this dependence decaying very slowly. This notion of 'roughness' in volatility aligns with the empirical features seen in high-frequency data, leading to models that not only better fit historical data but also offer more precise tools for derivative pricing. However, due to the memory effect rough volatility models do not enjoy the Markov property and classical pricing methods like the PDE approach are not applicable while the well-known Monte-Carlo simulation is converging even slower due to the fractional nature of these models.

With this groundwork of rough volatility, it becomes intuitive to reassess and adapt classical models to better capture market dynamics. Among the myriad of models in the financial literature, the Heston model, known for its portrayal of volatility as a meanreverting process, becomes a prime candidate for revision. Enter the rough Heston model, an innovative fusion that leverages the empirical consistency of rough volatility while retaining the structural advantages of the Heston framework, paving a new path for robust derivative pricing and risk management.

A pivotal advantage of the rough Heston model is the availability of its characteristic function [\[3\]](#page-112-2). This facilitates efficient pricing and calibration techniques, especially when combined with trans-formative mathematical tools. One approach to leverage this is through Fourier-based methods. One of the most popular methods of this type is the COS method, introduced by Fang and Oosterlee in [\[4\]](#page-112-3). The essence of the method lies in approximating the probability density function, which appears in the expectation, through the use of its Fourier-cosine series expansion. It has been established that the Fourier-cosine series coefficients have a closed-form solution with the characteristic function. This allows the COS method to be applied to a broad class of asset price processes for which the characteristic function is available [\[4\]](#page-112-3). For smooth density functions, the COS method is optimal in terms of error convergence and computational complexity for European options. In case the bi-variate characteristic function between two processes is known, we can make use of the 2-dimensional COS method introduced by [\[5\]](#page-112-4). This enables us to price path-dependent options like barrier and Bermudan options efficiently.

In this thesis, we introduce a novel cosine tensor network that significantly enhances the supervised learning of the characteristic function (chf) of the lifted Heston model, a cornerstone in our method for fast pricing and calibration of European options under the rough Heston framework. This innovative approach not only establishes a new benchmark for pricing barrier and Bermudan options within this domain but also enables the use of gradient-based optimization methods. Thanks to the efficiency of the cosine tensor network, we can calibrate the lifted Heston model against market quotes in under one minute, marking a substantial improvement in computational speed and efficiency. Broadly speaking, our work provides a comprehensive toolkit for pricing various derivatives under the rough Heston model, catering to the industry's need for speed and precision.

This thesis is structured as follows. In Chapter 1, we give a broad introduction to option pricing and volatility as well as describe the fundamentals of rough volatility models and their advantages opposite of classical stochastic volatility models like the classical Heston model. In the following chapter, we give a broad overview of the fundamentals of option pricing and implied volatility to understand the benefits of rough volatility. Finally, the rough volatility concept which is the base of the rough Heston model is explained by introducing the fractional Brownian motion and the rough fractional stochastic volatility model.

This is followed by a broad literature review of the world of option pricing and implied volatility as well as the mathematical foundations and the beginning of rough volatility and its advantages in comparison to classical stochastic volatility models.

In Chapter 3, we define the mathematical framework in order to solve the problems to be able to efficiently price and calibrate derivatives under the rough Heston model.

After introducing Hawkes Processes at the end of chapter 3, chapter 4 starts with the derivation of the rough Heston model from the microstructural market interactions (and the following facts of the electronic trading market). Moreover, we introduce the multifactor approximation of the rough Heston model, the so-called Lifted Heston model.

In Chapter 5, we introduce the characteristic function of each of the models in order to use Fourier-based methods to price European options as presented in the following chapter.

In Chapter 6, we give a detailed analysis in order to justify the lifted Heston model as our benchmark method for pricing European options. To do so, we give a brief introduction to the Adams scheme to approximate the characteristic function of the rough Heston model, as well as the implicit-explicit Euler in order to get the characteristic function of the lifted Heston model. Moreover, we give some numerical results comparing both methods in terms of performance as well as accuracy.

Chapter 7 is the first major contribution to this thesis. We use our newly developed cosine network based on canonical polyadic decomposition (CPD) as our supervised machine learning technique in order to price European options in a very fast and efficient way, without using any black-box neural networks. We also present some numerical results calculating single option pricing processes as well as the whole implied volatility surface.

In Chapter 8, we introduce a new calibration method based on the characteristics of our novel pricing method derived in Chapter 6. We also give some numerical results based on real market data.

In the next chapter, we introduce a novel benchmark method for pricing Barrier and Bermudan options under the rough Heston model. For this purpose, we present the bivariate characteristic function between log-stock and volatility process driven by a fractional Brownian motion, as well as the 2D COS method to price options which enables us to price path-dependent options under the rough Heston model.

Finally, in chapter 10, the thesis is concluded with a summary of the results and some future research recommendations.

**1**

# **2**

# **LITERATURE REVIEW**

## <span id="page-20-1"></span><span id="page-20-0"></span>**2.1.** OPTION BASICS AND IMPLIED VOLATILITY

An option is a financial derivative that gives the buyer the right to trade an underlying asset at a predetermined price. Generally speaking, an option gives the holder of the option the option to buy or sell an asset in the future without any obligations. There is a key difference between put options and call options.

A call option gives the holder the right to buy an asset, while a put option gives the right to sell the option. In general, there are different types of options. In this thesis, we mainly focus on European style options, Barrier options, and Bermudan options [\[6\]](#page-112-5).

### <span id="page-20-2"></span>**2.1.1.** EUROPEAN OPTIONS

European-style options can only be exercised at a pre-specified date, which is known as the maturity of the option. Moreover, the option will only be exercised if the holder of the option makes a profit. For example, for a European call option, a profit is made when the asset price *S* is bigger than the pre-determined price, called strike price *K*, at maturity *T* . Equivalently for the put option a profit is made once the asset price is smaller than the strike price at maturity. This results in the following payoff functions [\[6\]](#page-112-5).

$$
V_{call,T} = \max(S_T - K, 0),
$$
  
\n
$$
V_{put,T} = \max(K - S_T, 0).
$$
\n(2.1)

### <span id="page-20-3"></span>**2.1.2.** BERMUDAN OPTIONS

Bermudan style options can be exercised at multiple pre-specified dates before maturity called the observation dates. A European option is a Bermudan option with only one possible exercise date [\[7\]](#page-112-6). Analogously, a Bermudan call option gives the option holder the right to buy the underlying asset *S* at any of the observation dates at strike price *K*, while a Bermudan put option gives the holder the right to sell *S* at strike *K* at any of the

<span id="page-21-0"></span>

Figure 2.1: Payoff function for a European call and put option with  $K = 50$ .

exercise dates. Hence the payoff profile for Bermudan call and put options is the same as the one for European options at any arbitrary observation date [\[8\]](#page-112-7).

#### <span id="page-21-1"></span>**2.1.3.** BARRIER OPTIONS

Last but not least, we want to introduce Barrier options. Discretely monitored "out" barrier options are options that cease to exist if the asset price hits a certain barrier level, *B*, at one of the pre-specified observation dates [\[9\]](#page-112-8). If  $B > S_0$ , they are called "up-andout" options, and "down-and-out" otherwise. The payoff for an up-and-out option reads

$$
V_{call,T} = \max((S_T - K), 0) \mathbf{1}_{\{S_{t_i} < B\}},
$$
\n
$$
V_{put,T} = \max((K - S_T), 0) \mathbf{1}_{\{S_{t_i} < B\}},
$$
\n
$$
(2.2)
$$

where  $\mathbf{1}_A$  is the indicator function defined as

$$
1_A = \begin{cases} 1 & \text{if } A \text{ is not empty} \\ 0 & \text{otherwise.} \end{cases}
$$

#### <span id="page-21-2"></span>**2.1.4.** OPTION PRICING

**Definition 2.1.1.** *(Arbitrage). An investment strategy is called an arbitrage if the value process V of the strategy satisfies the following two properties:*  $\cdot \mathbb{P}(V_T > (1+r)V_0) > 0$ *.*  $\mathbb{P}(V_T \geq (1+r)V_0) = 1$ , where r denotes the risk-free interest rate in a money-savings ac*count and T is the maturity time.*

In other words, the no-arbitrage principle assures that there is never an opportunity to make a risk-free profit that gives a greater return than that provided by the interest from a bank deposit.

One of the most important results in option value theory is the derivation of the partial differential equation (PDE) by Fischer Black and Myron Scholes to price European

<span id="page-22-0"></span>

Figure 2.2: Payoff function for a Barrier call and put option with  $K = 50$  ,  $B_{call} = 80$  and  $B_{put} = 20$ .

options, published in 1973 [\[10\]](#page-112-9). Assuming that the asset price process *S* is modeled according to a geometric Brownian motion, the value of a European option can be represented as the solution of the following PDE,

$$
\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0
$$
\n(2.3)

The Black-Scholes Partial Differential Equation (PDE) is derived through a series of steps, one of which includes creating what's known as a replicating portfolio. This portfolio is a mix of the underlying asset, like a stock, and a risk-free asset, typically bonds or cash, designed to mimic the returns of an option. The idea is to adjust the amounts of the stock and the risk-free asset so that the portfolio's value moves in tandem with the option's value over time.

According to the principle of no-arbitrage, which states that you can't have a free lunch in the market, two portfolios that offer the same payoff must be priced equally. Therefore, the value of the option should always match the value of the replicating portfolio.

Black and Scholes used this concept to derive their famous pricing equation for European options. However, there's another way to arrive at the Black-Scholes PDE, and that's through the use of martingale theory. This alternative method involves defining a 'risk-neutral' probability measure. Under this measure, the expected growth rate of the underlying asset, when adjusted for risk, is the same as the growth rate of a risk-free asset. The First Fundamental Theorem of Asset Pricing (FTAP1) backs up this approach by confirming that if the market doesn't allow for arbitrage opportunities, such a riskneutral measure must exist.

Both methods ultimately provide a mathematical foundation for pricing options in a way that rules out the possibility of making a riskless profit through arbitrage. By using this measure, the expected value of the option at maturity can be discounted to its value at time  $t < T$ ,

$$
V_t = \mathbb{E}^{\mathbb{Q}} \left[ e^{-r(T-t)} V_T \right]
$$
  
= 
$$
\mathbb{E}^{\mathbb{Q}} \left[ e^{-r(T-t)} H(S, T) \right],
$$

**2**

where  $H(T, S)$  denotes the payoff function of the option. The measure  $\mathbb Q$  indicates that the expectation is taken under the risk-neutral measure. Hence, the value of the option can be derived either by solving the pricing PDE presented or by evaluating the discounted expected payoff. As mentioned above, by the no-arbitrage principle, these approaches should ultimately lead to the same value. The equivalence of both approaches is proved in the famous Feynman-Kac theorem, which connects the discounted expectation as the closed-form solution to the Black-Scholes PDE.

#### <span id="page-23-0"></span>**2.1.5.** IMPLIED VOLATILITY

In order to price options before maturity, quantitative models are chosen. The most famous one is the Black-Scholes model invented by Fischer Black and Scholes [\[10\]](#page-112-9). In this model, the interest rate *r* and volatility  $\sigma$  are supposed to be constant while the stock price follows a geometric Brownian motion, i.e.

$$
dS_t = \mu_t S_t dt + \sigma_t S_t dW_t.
$$

With the above-mentioned Feynman-Kac Theorem, we get the price of a European call option by solving the discounted expectation under the risk-neutral measure Q, i.e.

$$
V_{call,t} = e^{-r(T-t)} \mathbb{E} \left[ \left( S_T - K \right) \mathbf{1}_{\{ S_T > K \}} \mid \mathcal{F}_t \right].
$$

Solving this leads to the famous Black-Scholes price of a European call option [\[8\]](#page-112-7) denoted by

$$
V_{call,t} = S_T \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2).
$$

where Φ is the cumulative distribution function of a standard normal and

$$
d_1 = \frac{\log(S_t/K) + \left(r + \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}},
$$

$$
d_2 = d_1 - \sigma\sqrt{T - t}.
$$

The Black-Scholes model is widely recognized for its simplicity and the powerful framework it provides for option pricing. However, one of its main limitations is the assumption that volatility remains constant over time. Despite this, the model's analytical solution is quite useful in practice. It allows for the calculation of implied volatility through root-finding algorithms. For instance, by observing market prices of options and applying these computational techniques, one can back out the volatility that is implied by current market conditions. Hence, we have to solve

<span id="page-23-1"></span>
$$
V_{BS}(\sigma_{\text{market}}, S_t, r, t, K, T) = V_{\text{market}},\tag{2.4}
$$

where  $V_{BS}(\sigma_{\rm market},S_t,r,t,K,T)$  is the Black-Scholes formula induced with the volatility implied by the market. By knowing this formula we can either calculate the implied volatility when the options price  $V_t$  is given or analogously can calculate the price of the option using the given implied volatility  $[11]$ .

As mentioned, in order to solve [2.4,](#page-23-1) root-finding methods can be applied such as Newton-Raphson.

#### NEWTON-RAPHSON ALGORITHM

We apply the Newton-Raphson algorithm [\[12\]](#page-112-11) in order to find the implied volatility. Let  $\sigma_0$  be a good estimate of *r* and let  $r = \sigma_0 + h$ . Since the true root is *r*, and  $h = r - \sigma_0$ , the number *h* measures how far the estimate  $\sigma_0$  is from the truth.

Since *h* is 'small,' we can use the linear (tangent line) approximation to conclude that

$$
0 = f(r) = f(\sigma_0 + h) \approx f(\sigma_0) + hf'(\sigma_0),
$$

and therefore, unless  $f'(\sigma_0)$  is close to 0,

$$
h \approx -\frac{f(\sigma_0)}{f'(\sigma_0)}.
$$

It follows that

$$
r = \sigma_0 + h \approx \sigma_0 - \frac{f(\sigma_0)}{f'(\sigma_0)}.
$$

Our new improved estimate  $x_1$  of  $r$  is therefore given by

$$
\sigma_1 = \sigma_0 - \frac{f(\sigma_0)}{f'(\sigma_0)}.
$$

The next estimate  $\sigma_2$  is obtained from  $\sigma_1$  in exactly the same way as  $\sigma_1$  was obtained from  $\sigma_0$  :

$$
\sigma_2 = \sigma_1 - \frac{f(\sigma_1)}{f'(\sigma_1)}
$$

Continue in this way. If  $\sigma_n$  is the current estimate, then the next estimate  $\sigma_{n+1}$  is given by

$$
\sigma_{n+1} = \sigma_n - \frac{f(\sigma_n)}{f'(\sigma_n)}
$$
\n(2.5)

In our context of European call options, we find that

$$
f(\sigma) = V_{market} - V_{BS}(\sigma_{market}, S_t, r, t, K, T) = 0,
$$
\n(2.6)

while

$$
f'(\sigma) = -\frac{d}{d\sigma} V_{BS}(\sigma_{\text{market}}, S_t, r, t, K, T)
$$
 (2.7)

is the so called option vega which is known in closed form solution denoted by

$$
f'(\sigma) = -Ke^{-r(T-t_0)}f_{\mathcal{N}(0,1)}(d_2)\sqrt{T-t_0}
$$
\n(2.8)

<span id="page-24-0"></span>with  $f_{\mathcal{N}(0,1)}(\cdot)$  the standard normal probability density function,  $t_0 = 0$ , and

$$
d_2 = \frac{\log\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T - t_0)}{\sigma\sqrt{T - t_0}}.
$$

## **2.2.** CLASSICAL HESTON MODEL

To address the limitations of the Black-Scholes model, particularly its assumption of constant volatility, numerous alternative models have been proposed. Among the more prominent is the Heston model [\[13\]](#page-112-12). In this paper, this model retains the geometric Brownian motion for asset price dynamics but introduces a key innovation: volatility itself is treated as a stochastic process. Specifically, volatility follows an Ornstein-Uhlenbeck process, which allows it to vary over time in a mean-reverting manner. This ultimately results in the following representation.

$$
dS(t) = \sqrt{\nu(t)}S(t)dW_t,
$$
  
\n
$$
dv(t) = \kappa[\bar{v} - \nu(t)]dt + \nu\sqrt{\nu(t)}dB_t,
$$
\n(2.9)

with initial conditions

 $S(0) = S_0 \ge 0$ ,  $v(0) = v_0 \ge 0.$ 

In this case, the parameters  $\kappa$ ,  $\bar{\nu}$ ,  $\nu_0$  and  $\nu$  are positive describing the speed of mean reversion, the long-term variance, the instantaneous variance, and the volatility of the volatility respectively. *W* and *B* are two Brownian motions with correlation coefficient *ρ*, that is  $\langle dW_t, dB_t \rangle = \rho dt$ .

<span id="page-25-0"></span>Using the Heston model the implied volatility surface can be represented as shown in [2.3.](#page-25-0) Hence, it is more suitable for modeling market option prices, as The Heston model's ability to generate a volatility surface that aligns more closely with observed market data underscores its advantage over the Black-Scholes model, which assumes a flat volatility surface [\[11\]](#page-112-10).

#### Implied Volatility Surface



Figure 2.3: Implied volatility surface created using the Heston dynamics with parameters  $\rho = -0.4$ ,  $\bar{v} = 0.056$ ,  $v = 0.15$ ,  $\kappa = 2.0$  and  $v_0 = 0.060$ 

## <span id="page-26-0"></span>**2.3.** ROUGH VOLATILITY

#### <span id="page-26-1"></span>**2.3.1.** FRACTIONAL BROWNIAN MOTION

We first give an idea of what it means to have rough volatility. For this purpose, we introduce fractional Brownian motions. In order to define fractional Brownian motions, we first define *α*-Hölder continuity

**Definition 2.3.1.** *A path X* : [0, *T*]  $\rightarrow$  *E is*  $\alpha$ *-Hölder continuous for*  $\alpha$  > 0 *if*  $\|x\|_{\alpha}$  <  $\infty$ *, with* 

$$
||x||_{\alpha} := \sup_{t \neq s} \frac{|X_t - X_s|}{|t - s|^{\alpha}}
$$
 (2.10)

A *α*-Hölder path is said to be a *rough path* when the path itself is *α* ′ -Hölder continous for every  $\alpha' > \alpha$  and  $\alpha \leq \frac{1}{2}$ . The rough path we need to define rough volatility is the fractional Brownian motion introduced by Mandelbrodt and van Ness. Fractional Brownian motion is a generalized version of Brownian motion where the increments are not independent anymore. The fractional Brownian motion can be defined by the following Mandelbrot - van Ness representation [\[14\]](#page-113-1).

**Definition 2.3.2.** *(Fractional Brownian motion (fBM))*

$$
W_t^H = \frac{1}{\Gamma(H+1/2)} \left[ \int_{-\infty}^0 \left( (t-s)^{H-1/2} - (-s)^{H-1/2} \right) dW_s + \int_0^t (t-s)^{H-1/2} dW_s \right], \quad (2.11)
$$

*where*  $H \in (0,1)$  *is called the Hurst parameter and*  $\Gamma(\cdot)$  *is the gamma function.* 

It can easily be derived that for  $H = \frac{1}{2}$  the process coincides with the classical Brownian motion. The fractional Brownian motion has two important propositions on the increments.

**Proposition 2.3.1.** *The covariance function of the fractional Brownian motion is denoted by*

$$
\mathbb{E}\left[W_t^H W_s^H\right] = \frac{1}{2} \left(|t|^{2H} + |s|^{2H} - |t - s|^{2H}\right). \tag{2.12}
$$

From this equation, it can be observed that for  $H < \frac{1}{2}$  the increments of the fBM are negatively correlated and for  $H > \frac{1}{2}$  the increments are positively correlated. Intuitively one could say that *H* < 0.5 signifies anti-persistent behavior, where significant changes tend to be followed by smaller ones, introducing some predictability. *H* > 0.5 indicates persistent behavior, where significant changes are more likely to be followed by further significant changes, making predictions less straightforward.

**Proposition 2.3.2.** *For any t*  $\in \mathbb{R}, \Delta \geq 0, q > 0$ *, we find that* 

$$
\mathbb{E}\left[\left|W_{t+\Delta}^{H} - W_{t}^{H}\right|^{q}\right] = K_{q}\Delta^{qH}.
$$
\n(2.13)

<span id="page-26-2"></span>As a result , for  $H < \frac{1}{2}$  the fBM is an *α*-Hölder rough path. Hence throughout this thesis, we only consider  $H < \frac{1}{2}$ .

<span id="page-27-0"></span>

Figure 2.4: Fractional Brownian motion with different Hurst parameter

#### **2.3.2.** ROUGH VOLATILITY

As we have introduced above volatility can be modeled in terms of very regular paths in the case of the well-known Black Scholes model or with a regularity really close to that of the classical Brownian motion in the case of stochastic volatility like the Heston model. However, there are two main reasons for the introduction of rough volatility models.

#### (a) **Better fitting of stochastic volatility models**

For this purpose, we take a look at the term structure of at-the-money volatility skew, which is defined as

<span id="page-27-1"></span>
$$
\psi(\tau) := \left| \frac{\partial}{\partial k} \sigma_{\text{BS}}(k, \tau) \right|_{k=0}.
$$
\n(2.14)

[\[15\]](#page-113-2) showed in his paper that for small-time maturities the ATM volatility skew is of the form

$$
\psi(\tau) \sim \tau^{H-1/2},\tag{2.15}
$$

when the volatility process follows a fractional Brownian motion with Hurst parameter H. Hence, [2.14](#page-27-1) is called Fukusawa's approach.

In Gatheral's study [\[1\]](#page-112-0), it is noted that the market's observed term structure of volatility can be effectively approximated by a power-law function of the time to expiry.

$$
\psi(\tau) = A\tau^{H-1/2}.\tag{2.16}
$$

However, stochastic volatility models like the classical Heston model generate an ATM term structure that is constant for small *τ* and somewhat behaves like a sum of decaying exponentials for larger *τ*. As we can see in [2.5](#page-28-0) and [2.6,](#page-28-1) the Heston model and in general any stochastic volatility models generate volatility surfaces that are not consistent with observed market data.

#### (b) **Power-law scaling of the historical volatility process**

[\[1\]](#page-112-0) proved the roughness of the volatility by estimating the smoothness of the variance process from empirical DAX and Bund future data. The authors started by

<span id="page-28-0"></span>

<span id="page-28-1"></span>Figure 2.5: Term structure of the at-the-money implied volatility skew of SPX options from 05.10.2023



Figure 2.6: Term structure of the at-the-money implied volatility skew created using the classical Heston dynamics vs. Fukusawa's approach:  $\psi(\tau) = A \tau^{H-1/2}$  with  $H = 0.12$ .

assuming that there is access to discrete observations  $v_0$ ,  $v_\Delta$ , ...,  $v_{k\Delta}$ ..., with  $k \in$ {0, ⌊*T* /∆⌋} of the volatility process {*v<sup>t</sup>* }*t*∈[0,*<sup>T</sup>* ] on a time grid [0,*T* ] with mesh ∆. They further define for  $q \ge 0$ 

$$
m(q,\Delta) = \frac{1}{N} \sum_{k=1}^{N} \left| \log(v_{k\Delta}) - \log(v_{(k-1)\Delta}) \right|^{q}.
$$
 (2.17)

They assume that for some  $s_q > 0$  and  $b_q > 0$ 

<span id="page-28-2"></span>
$$
N^{qs_q} m(q, \Delta) \to b_q \tag{2.18}
$$

as  $\Delta \rightarrow 0$ .

Under a set of additional technical prerequisites, Equation [2.18](#page-28-2) proposes a compelling interpretation of the volatility process. It suggests that this process is a **2**

member of the Besov smoothness space, specifically denoted as  $\mathscr{B}q,\infty^{s_q}$ . However, it does not extend its membership to the Besov space  $\mathscr{B}q, \infty^{s'q}$  where  $s'_{q} > s_{q}$ .

**Proposition 2.3.3.** *Besov Space Besov spaces are a generalization of the Sobolev function spaces, which are used as a measure for the regularity (smoothness) of a function. Hence, Besov spaces can be used as a smoothness indicator in the rough volatility environment too. For an in depth definition, we refer to [\[16\]](#page-113-3) and [\[17\]](#page-113-4).* 

From this perspective, [\[1\]](#page-112-0) perceive  $s_q$  as a measure of the regularity of the volatility when assessed in *l<sub>q</sub>* norm. They elaborate, functions within the bounds of  $\mathscr{B}_{q,\infty}^s$ for all  $q > 0$  have the privilege of possessing the Hölder property. This property has a parameter *h* that abides by the rule *h* < *s*.

The authors in [\[1\]](#page-112-0) explore a concrete example to further illuminate this concept. Suppose  $log(v_t)$  as a fractional Brownian motion (fBM) with the Hurst parameter *H*. In this scenario, for any  $q \ge 0$ , Equation [2.18](#page-28-2) is verifiable in probability where  $s_q = H$ . What's intriguing here is that the sample paths of this process virtually always belong to  $\mathscr{B}_{q,\infty}^H$ .

Developing further, [\[1\]](#page-112-0) claim that if the increments of the log-volatility process maintain stationarity and if a law of large numbers can be invoked, one could perceive  $m(q, \Delta)$  as an empirical reflection of

$$
\mathbb{E}\left[\left|\log(v_{\Delta}) - \log(v_0)\right|^q\right] = K_q \Delta^{Hq}.
$$

From the Data they conclude that the Hurst parameter is  $H \approx 0.1$ , yielding the conclusion that volatility is indeed rough

We note that in  $[18]$  and  $[19]$  similar results were found in order to prove that the volatility process is indeed a rough one.

#### <span id="page-29-0"></span>**2.3.3.** ROUGH FRACTIONAL STOCHASTIC VOLATILITY MODEL

After showing that the volatility is rough, [\[1\]](#page-112-0) introduced a model called Rough Fractional Stochastic Volatility (RFSV) model where the log-volatility is driven by an Ornstein-Uhlenbeck process. They first suggested that

$$
\log(v_{t+h}) - \log(v_t) = v\left(W_{t+h}^H - W_t^H\right),\,
$$

where  $W^H$  is a fBM and  $\nu$  is a positive constant, since it should be a volatility model with constant smoothness and with a distribution similar to a Gaussian process. However, this model is not stationary and stationarity is desired for mathematical tractability and also to ensure stability of the model at very large times. Adopting a fractional Ornstein-Uhlenbeck process effectively addressed this issue, resulting in a refined model formulation as follows:

$$
v_t = \exp(X_t),
$$
  

$$
dX_t = vdW_t^H - \alpha(X_t - x),
$$

where  $x \in \mathbb{R}$  and  $\alpha$  is positive constant. For any  $\varepsilon \geq 0$ , the volatility process has the Hölder property of order  $H - \varepsilon$ , and since we know that *H* is accepted as approximately 0.1, this process is rough.

# **3**

# <span id="page-30-0"></span>**MATHEMATICAL FRAMEWORK**

In this chapter, all the relevant mathematical tools are presented to formulate, model, and solve the rough volatility option pricing problems.

## <span id="page-30-1"></span>**3.1.** STOCHASTIC CALCULUS

**Definition 3.1.1.** *(Filtration). Let*  $(\Omega, \mathcal{F}, \mathbb{P})$  *be a probability space. A filtration on*  $(\Omega, \mathcal{F}, \mathbb{P})$  $i$ *s* a family of sub- $\sigma$ -fields  $\{\mathscr{F}_t, t \ge 0\}$  of  $\mathscr{F}$  indexed by  $t \in [0,\infty)$ , such that  $\mathscr{F}_s \subset \mathscr{F}_t$  for ev*ery*  $s \le t$  ≤ ∞*.* 

**Definition 3.1.2.** *(Adapted process). A process*  $X = \{X_t, t \ge 0\}$  *is said to be adapted to a*  $filt$  *filtration*  $\{\mathscr{F}_t, t \geq 0\}$  *if for all*  $t \geq 0, X_t$  *is*  $\mathscr{F}_t$  *measurable.* 

**Definition 3.1.3.** *(Martingale). Let*  $M = \{M_t, t \ge 0\}$  *be a process defined on the probability space*  $(\Omega, \mathcal{F}, \mathbb{P})$  *equipped with a filtration* { $\mathcal{F}$ *t*, *t*  $\geq$  0}*. Then M is said to be a martingale if 1. M* is an adapted process. 2. For all  $t \geq 0$ ,  $M_t$  is integrable. 3. M satisfies the martingale *property, which reads*

$$
\mathbb{E}[M_t \mid \mathcal{F}_s] = M_s, \quad \forall 0 \le s < t.
$$

**Definition 3.1.4.** *(Semi-martingale). A stochastic process*  $X = \{X_t, t \ge 0\}$  *is called a semimartingale if it can be decomposed as follows:*

$$
X = X_0 + M + A,
$$

*where the random variable*  $X_0$  *is finite and*  $\mathcal{F}_0$ *-measurable, the stochastic process* M *is a local martingale and the stochastic process A has finite variation.*

**Definition 3.1.5.** *(Brownian Motion). A real-valued process*  $\{W(t), t \ge 0\}$  *is called a Brownian motion if 1. Starting at* 0 : *W* (0) = 0*. 2. Normally distributed increments: For all*  $0 \leq s \leq t$ ,  $W(t) - W(s) \approx N(0, t - s)$ . 3. Independent increments: For  $0 \leq t_0 < t_1 < \cdots < t_n$ , *the random variables*  $Y_i := W(t_i) - W(t_{i-1}), i = 1, \ldots, n$  are independent. 4. Continuous *trajectories: The map*  $t \rightarrow W(t)$  *is continuous.* 

**Definition 3.1.6.** *(Itô Integral). For any square-integrable adapted process*  $g(t)$  *with continuous sample paths, the Itô integral is given by:*

$$
I(T)=\int_0^T g(t)dW(t):=\lim_{m\to\infty}I_m(T),\quad\hbox{ in }L^2.
$$

Here,  $I_m(T) = \int_0^T g_m(t) dW(t)$  for some elementary process  $g_m(t) = \sum_{j=0}^{n-1} \eta_j 1_{[t_j,t_{j+1}]},$  sat*isfying:*

$$
\lim_{m \to \infty} \mathbb{E} \left[ \int_0^T \left( g_m(t) - g(t) \right)^2 dt \right] = 0,
$$

*where*  $\eta_j$  *is*  $\mathcal{F}_{t_j}$  *measurable* for all  $j = 0, 1, ..., n-1$  and square-integrable.

**Theorem 3.1.1.** (Itô's Formula). Let  $f \in C^2(\mathbb{R})$  and consider a continuous semi-martingale *X* with decomposition  $X = M + A$ . Then, the stochastic process  $(f(X_t))_{t \geq 0}$  is also a semi*martingale and holds*

$$
f(X_t) = f(X_0) + \int_0^t \frac{\partial f}{\partial x}(X_u) \, dX_u + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(X_u) \, d[X]_u,
$$

*with* [X] *denotes the quadratic variation of the process*  $(X_t)_{t\geq0}$ . Itô's formula is often ex*pressed in differential form:*

$$
df(X_t) = \frac{\partial f}{\partial x}(X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X_t) d[X]_t
$$

**Definition 3.1.7.** *(Convolutions) For a measurable function K on* R+ *and a measure L on* R<sup>+</sup> *of locally bounded variation, the convolutions K* ∗*L and L* ∗*K are defined by*

$$
(K * L)(t) = \int_{[0,t]} K(t - s)L(ds), \quad (L * K)(t) = \int_{[0,t]} L(ds)K(t - s)
$$

*for t*  $> 0$  *whenever these expressions are well-defined, and extended to t* = 0 *by rightcontinuity when possible [\[20\]](#page-113-7).*

*If F* is a function on  $\mathbb{R}_+$ *, we denote*  $K * F = K * (F dt)$ *, that is,* 

$$
(K * F)(t) = \int_0^t K(t - s)F(s)ds
$$

## <span id="page-31-0"></span>**3.2.** LEVY-PROCESSES

Throughout this thesis, we deal with the classical or the rough Heston model, which both belong to the class of Levy processes.

**Definition 3.2.1.** *(Lévy process). A Lévy process is a stochastic process*  $X = \{X_t : t \geq 0\}$  *that satisfies the following properties:*

- (a) *Initial condition:*  $X_0 = 0$  *almost surely.*
- *(b) Independence of increments: For any*  $0 \le t_1 < t_2 < ... < t_n < \infty$ *, the random variables*  $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \ldots, X_{t_n} - X_{t_{n-1}}$  are mutually independent.
- *(c) Stationary increments: For any*  $0 \le s < t$ , the increment  $X_t X_s$  has the same distri*bution as Xt*−*<sup>s</sup> .*
- *(d) Continuity in probability: For any*  $\epsilon > 0$  *and*  $t \geq 0$ *, it holds that*  $\lim_{h\to 0}$  $\mathbb{P}(|X_{t+h} X_t$   $|> \epsilon$ <sup> $= 0$ </sup>.

*If* X is a Lévy process, it is possible to construct a version of X such that the mapping  $t \mapsto X_t$ *is almost surely right-continuous with left limits.*

**Definition 3.2.2.** *(Characteristic function). Given a stochastic process X and a time t, the characteristic function of the stochastic process at time t is given by:*

$$
\varphi_X(u,t):=\mathbb{E}\left[e^{i u X(t)}\right].
$$

*Moreover, when conditional on the initial value X*(0) = *x it will also be written with the regular φ as:*

$$
\phi_X(u,t;x) := \mathbb{E}\left[e^{i u X(t)} \mid x\right]
$$

**Theorem 3.2.1.** *(Lévy relation ch. f.). Given a Lévy process*  $X = \{X_t, t \ge 0\}$ *, with*  $X_0 = x$ *. The ch. f . of X given initial value follows the relation:*

$$
(u, t; x) = \phi(u, t; 0)e^{iux} = \phi_{levy}(u)e^{iux},
$$

 $where \phi(u, t; x) := \mathbb{E}\left[e^{iuX(t)} | X_0 = x\right].$ 

## <span id="page-32-0"></span>**3.3.** SINGULAR VALUE DECOMPOSITION (SVD)

**Theorem 3.3.1.** *(SVD)* Let  $A \in \mathbb{R}^{m \times n}$  *with*  $m \ge n$ *. Then*  $A = U\Sigma V^T$ *, where* 

- *(a)*  $U ∈ ℝ<sup>m×m</sup>$  *is an orthogonal matrix,*
- *(b)*  $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n}$  *with*  $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$  *and*  $\sigma_i \in \mathbb{R}$  *for all i,*
- *(c)*  $V ∈ ℝ^{n×n}$  *is an orthogonal matrix.*

*The rank of the matrix has value r* = min $\{l | \sigma_l > 0\}$ *, which is the singular value with the smallest index that exceeds 0. In SVD these matrices are referred to as:*

- *(i) The column vectors of*  $U = [u_1 \dots u_m]$  *are called the left singular vectors of A,*
- *(ii) The column vectors of*  $V = [v_1 \dots v_n]$  *are called the right singular vectors of A,*
- *(iii) he values*  $\{\sigma_1, \ldots, \sigma_n\}$  *are called the non-singular values of A.*

#### <span id="page-32-1"></span>**3.3.1.** SOLVING LINEAR SYSTEM USING SVD

In this paper, SVD is employed to achieve a 'best fit' solution for a system of linear equations. This approach is particularly powerful for addressing linear systems, which are over-determined, i.e., where the number of equations (*m*) surpasses the number of unknowns (*n*). Let's consider a linear system represented by the matrix equation:

$$
Ax = b
$$

Here,  $A \in \mathbb{R}^{m \times n}$ ,  $x$  is an  $n$ -dimensional vector of unknowns, and  $\bm{b} \in \mathbb{R}^m$ . The LS solution  $x_L$ S aims to minimize the Euclidean distance between the estimated solution  $Ax_L$ S and the vector of constants *b*. This leads to an unconstrained optimization problem:

<span id="page-33-1"></span>
$$
\boldsymbol{x}_{LS} = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\operatorname{argmin}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2. \tag{3.1}
$$

To derive the LS solution using SVD, we approach as follows

(a) Compute the SVD of matrix *A* :

$$
\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T
$$

(b) Calculate the pseudoinverse of **Σ**, denoted as Σ † , by taking the reciprocal of each non-zero singular value and transposing the resultant matrix. The pseudoinverse is denoted as

$$
\Sigma^{\dagger} = \left[ \begin{array}{ccc} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{array} \right]
$$

where  $\sigma_1, \sigma_2, \ldots, \sigma_n$  are the singular values of A.

(c) Compute the LS solution  $x_{\text{LS}}$  using the formula:

$$
\boldsymbol{x}_{LS} = \boldsymbol{V}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{U}^T\boldsymbol{b}
$$

Here,  $x_{LS}$  is the vector of least square solutions, representing the solution to the minimization problem described by equation [3.1.](#page-33-1)

## <span id="page-33-0"></span>**3.4.** TENSOR CALCULUS

Tensor calculus is an advanced mathematical framework that extends the concepts of matrices and vectors to higher-dimensional objects called tensors. In this thesis, it is used to describe higher-dimensional problems, which is needed for our approach to pricing European options under the rough Heston model. The fundamentals of tensor calculus are explained and illustrated in this section. Tensors introduce additional dimensions and provide a more flexible way to represent and manipulate data. A tensor can be thought of as a multi-dimensional array of numbers, with each element characterized by its position within the array. While tensors might seem initially daunting due to their higher dimensionality, they are indispensable when modeling and analyzing systems with intricate interactions between multiple variables [\[21\]](#page-113-0).

In tensor calculus, one of the fundamental operations is tensor unfolding. This process rearranges the elements of a tensor into a matrix format, unveiling underlying patterns and structures within the tensor. Unfolding tensors enables the application of traditional matrix-based techniques and algorithms. The number of dimensions in a tensor is called the order of a tensor. Here every dimension is called a mode of the tensor. To indicate a specific element of a *d*-th order tensor  $\mathscr{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$  an index set  $(i_n)_{n=1}^d$  is used. Here  $1 \le i_n \le I_n$  for  $n = 1, ..., d$ .

Before delving into tensor operations, it's important to grasp the concept of fibers. Just as matrices are composed of rows and columns, this idea can be extended to higherdimensional tensors through what are known as fibers. A fiber is essentially a line of elements within a tensor that you get by keeping all but one of the tensor's indices fixed. In the context of a three-dimensional tensor, these fibers are categorized as rows, columns, and tubes, which can be visualized in Figure [3.1.](#page-34-0) This concept is crucial for understanding the structure of tensors and how they can be manipulated in mathematical operations.

<span id="page-34-0"></span>

Figure 3.1: Fibers of a third-order tensor [\[21\]](#page-113-0)

Performing operations on higher-order tensors can be quite abstract and involve some manipulations to make it more intuitive. A transformation process called unfolding is applied to transform higher-order tensors into matrices. As an example, we look at unfolding an order-3 tensor with dimension  $2 \times 3 \times 4$ . This tensor can be rearranged into a  $4 \times 6$ ,  $3 \times 8$  or  $2 \times 12$  matrix. There are a lot of different ways one can unfold a tensor into a matrix but for our research, only the mode-  $n$  unfolding is relevant as explained in [\[21\]](#page-113-0). The mode- *n* unfolding of a tensor  $\mathscr{X}$  is denoted as  $\mathscr{X}_{(n)}$ . It is constructed by setting the mode- *n* fibers as columns of the matrix. For the example above for  $\mathbb{R}^{2\times3\times4}$ the three possible mode- *n* unfoldings are:

(i) mode-1 unfolding in  $\mathbb{R}^{4 \times 6}$  : using the columns as fibers of the tensor:

$$
\mathcal{X}_{(1)} = \left[ \begin{array}{cccccc} x_{111} & x_{121} & x_{131} & x_{112} & x_{122} & x_{132} \\ x_{211} & x_{221} & x_{231} & x_{212} & x_{222} & x_{232} \\ x_{311} & x_{321} & x_{331} & x_{312} & x_{322} & x_{332} \\ x_{411} & x_{421} & x_{431} & x_{412} & x_{422} & x_{432} \end{array} \right]
$$

(ii) mode- 2 unfolding in  $\mathbb{R}^{3\times 8}$  : using the rows as fibers of the tensor:

$$
\mathcal{X}_{(2)} = \left[\begin{array}{cccccc} x_{111} & x_{211} & x_{311} & x_{411} & x_{112} & x_{212} & x_{312} & x_{412} \\ x_{121} & x_{221} & x_{321} & x_{421} & x_{122} & x_{222} & x_{322} & x_{422} \\ x_{131} & x_{231} & x_{331} & x_{431} & x_{132} & x_{232} & x_{332} & x_{432} \end{array}\right],
$$

(iii) mode-3 unfolding in  $\mathbb{R}^{2\times12}$  : using the tubes as fibers of the tensor:

$$
\mathcal{X}_{(3)} = \left[ \begin{array}{cccccc} x_{111} & x_{211} & x_{311} & x_{411} & x_{121} & x_{221} & x_{321} & x_{421} & x_{131} & x_{231} & x_{331} & x_{431} \\ x_{112} & x_{212} & x_{312} & x_{412} & x_{122} & x_{222} & x_{322} & x_{422} & x_{132} & x_{232} & x_{332} & x_{432} \end{array} \right]
$$

These three unfolding operations are respectively illustrated in Figure. Once the tensors have been unfolded, the standard matrix operations can be applied. The operations used in this paper will be defined below.

**Definition 3.4.1.** *(Outer Product). Let*  $\mathbf{u} \in \mathbb{R}^m$ ,  $\mathbf{v} \in \mathbb{R}^n$  *be two vectors. Their outer product is denoted with* **u**◦**v** ∈ R *m*×*n , and the resulting matrix can be obtained by multiplying each element of* **u** *by each element of* **v***,*

$$
\mathbf{u} \circ \mathbf{v} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_m v_1 & u_m v_2 & \cdots & u_m v_n \end{bmatrix}
$$
(3.2)  
= 
$$
\begin{bmatrix} \mathbf{u} \cdot v_1 & \mathbf{u} \cdot v_2 & \cdots & \mathbf{u} \cdot v_n \end{bmatrix}
$$

**Definition 3.4.2.** *(Kronecker Product). The Kronecker product of matrices A* ∈ R *<sup>I</sup>*×*<sup>J</sup> and B* ∈  $\mathbb{R}^{K \times L}$  *is denoted with*  $A \otimes B \in \mathbb{R}^{IK \times JL}$ , and *is defined by* 

$$
A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1J}B \\ a_{21}B & a_{22}B & \cdots & a_{2J}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}B & a_{I2}B & \cdots & a_{IJ}B \end{bmatrix}
$$
(3.3)

**Definition 3.4.3.** *(Khatri-Rao Product). The Khatri-Rao product can be viewed as the*  $column \text{-}wise Kronecker product. Given the matrices  $A \in \mathbb{R}^{I \times K}$  and  $B \in \mathbb{R}^{J \times K}$ , the Khatri-$ *Rao product, denoted with*  $A \odot B \in \mathbb{R}^{I J \times K}$ *, is defined by* 

$$
A \odot B = \left[ a_1 \otimes b_1 \quad a_2 \otimes b_2 \quad \cdots \quad a_K \otimes b_K \right] \tag{3.4}
$$

**Definition 3.4.4.** *(Hadamard Product). The Hadamard product is an element-wise matrix product. Therefore, given the matrices*  $A, B \in \mathbb{R}^{I \times J}$ , the Hadamard product  $A \otimes B \in$ R *<sup>I</sup>*×*<sup>J</sup> produces the matrix*

$$
A \circledast B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2J}b_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \cdots & a_{IJ}b_{IJ} \end{bmatrix}
$$
(3.5)

To quantify the similarities between two tensors one can use the matrix norm called the Forbenius norm. This norm can be used to find the distance between two tensors.

<span id="page-35-0"></span>**Definition 3.4.5.** *(Frobenius Norm). Given a tensor* X ∈ R *<sup>I</sup>*1×*I*2×···×*I<sup>d</sup> , its Frobenius norm, often abbreviated with F-norm, is defined as the square root of the sum of the squares of all its elements:*

$$
\|\mathcal{X}\|_F^2 = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_d=1}^{I_d} x_{i_1 i_2 \dots i_d}^2}.
$$
#### **3.5.** FRACTIONAL CALCULUS

We define the fractional integral and the fractional derivative according to [\[22\]](#page-113-0).

**Definition 3.5.1.** *(Fractional Integral) We define the fractional integral of order r*  $\in$   $(0,1]$ *of a function f as*

$$
I^{r} f(t) = \frac{1}{\Gamma(r)} \int_{0}^{t} (t - s)^{r-1} f(s) ds
$$
 (3.6)

**Definition 3.5.2.** *(Fractional Derivative) We define the fractional derivative of order r* ∈ (0,1] *of a function f as*

$$
D^{r} f(t) = \frac{1}{\Gamma(1-r)} \frac{d}{dt} \int_{0}^{t} (t-s)^{-r} f(s) ds.
$$
 (3.7)

#### **3.6.** HAWKES PROCESSES

In order to later derive the rough Heston model and its characteristic function, we need to introduce Hawkes processes [\[23\]](#page-113-1). Hawkes processes (HP) recently became more popular in financial models as a non-markovian extension of the well-known Poisson process. HP is therefore a counting process, that models, similar to the Poisson process, a sequence of arrivals over time. In the application of finance, HP for example can model the arrival time of trade orders.

We start by giving some definitions introduced by [\[24\]](#page-113-2).

**Definition 3.6.1.** *(Counting process) A counting process is a stochastic process*  $(N(t): t \geq 1)$ 0) *taking values in*  $\mathbb{N}_0$  *that satisfies*  $N(0) = 0$ *, is almost surely (a.s.) finite, and is a rightcontinuous step function with increments of size +1. Further, denote by*  $(\mathcal{H}(u): u \ge 0)$  the *history of the arrivals up to time u. (Thus,*  $\mathcal{H}(\cdot)$  *can be viewed as a filtration.)* 

A counting process can be viewed as a cumulative count of the number of 'arrivals' into a system up to the current time. Another way to characterize such a process is to consider the sequence of random arrival times  $T = \{T_1, T_2, ...\}$  at which the counting process  $N(\cdot)$  has jumped. The process defined as these arrival times is called a point process.

**Definition 3.6.2.** *(Point process) If a sequence of random variables*  $T = \{T_1, T_2, ...\}$ *, taking values in*  $[0, \infty)$ *, has*  $\mathbb{P}$   $(0 \le T_1 \le T_2 \le ...)$  = 1*, and the number of points in a bounded region is a.s. finite, then T is a (simple) point process.*

It is important to note that the meaning of a point process or a counting process is often interchangeable. The point process, however, may be defined via a distribution function of the next arriving point conditioned on the past. Nevertheless, it is difficult to work with conditional arrival distributions because of the dependency on the distribution of the whole history up until the last arrival. Hence, the conditional intensity function is used to uniquely characterize the finite-dimensional distributions of the point process.

**Definition 3.6.3.** *(Conditional intensity function) Consider a counting process N*(·) *with associated histories*  $H(·)$ *. If a (non-negative) function*  $λ<sup>*</sup>(t)$  *exists such that* 

$$
\lambda^*(t) = \lim_{h \downarrow 0} \frac{\mathbb{E}[N(t+h) - N(t) \, | \, \mathcal{H}(t)]}{h}
$$

 $which only relies on the information of  $N(\cdot)$  in the past (that is,  $\lambda^*(t)$  is  $\mathcal{H}(t)$ -measurable),$ *then it is called the conditional intensity function of N*(·)*.*

As a result of this characteristic, the point process can be described as 'self-exciting', i.e. if an arrival causes the conditional intensity to increase, or 'self-regulating', i.e. if the conditional intensity drops after an arrival.

<span id="page-37-0"></span>**Definition 3.6.4.** *(Hawkes process) Consider*  $(N(t): t \ge 0)$  *a counting process, with associated history* ( $\mathcal{H}(t)$  :  $t \ge 0$ ), that satisfies

$$
\mathbb{P}(N(t+h) - N(t) = m | \mathcal{H}(t)) = \begin{cases} \lambda^*(t)h + o(h), & m = 1 \\ o(h), & m > 1 \\ 1 - \lambda^*(t)h + o(h), & m = 0 \end{cases}
$$

*Suppose the process' conditional intensity function is of the form*

$$
\lambda^*(t) = \lambda + \int_0^t \mu(t - u) dN(u)
$$

*for some*  $\lambda > 0$  *and*  $\mu$ :  $(0, \infty) \rightarrow [0, \infty)$  *which are called the background intensity and excitation function respectively. Assume that*  $\mu(\cdot) \neq 0$  *to avoid the trivial case, that is, a homogeneous Poisson process. Such a process N*(·) *is a Hawkes process.*

# **4**

### **THE ROUGH HESTON MODEL**

In this chapter, we derive the rough Heston model as well as the normal Heston model as the limit of nearly unstable Hawkes processes. The following approach is proposed in [\[25\]](#page-113-3).

#### **4.1.** TICK-BY-TICK PRICE MODEL

The tick-by-tick price model is defined by [\[25\]](#page-113-3) as a bi-dimensional Hawkes process. Taking into account [3.6.4,](#page-37-0) we find that the bi-dimensional Hawkes process is given by a bivariate point process  $(N_t^+)$  $t^+, N_t^ \left(t_{t}^{-}\right)_{t\geq0}$  taking values in  $(\mathbb{R}_{+})^2$  and with intensity  $\left(\lambda_t^{*+}\right)$ *t* ,*λ* ∗− *t* ¢ , i.e.

$$
\begin{pmatrix} \lambda_t^{*+} \\ \lambda_t^{*-} \end{pmatrix} = \begin{pmatrix} \lambda^+ \\ \lambda^- \end{pmatrix} + \int_0^t \begin{pmatrix} \mu_1(t-s) & \mu_3(t-s) \\ \mu_2(t-s) & \mu_4(t-s) \end{pmatrix} \cdot \begin{pmatrix} dN_s^+ \\ dN_s^- \end{pmatrix}
$$
(4.1)

where,  $\lambda^+$  and  $\lambda^-$  are positive constants and the functions  $(\mu_i)_{i=1,...4}$  are non-negative with associated matrix called kernel matrix. It is stated in  $[25]$  that for a large tick asset, which is an asset whose bid-ask spread is almost always equal to one tick and therefore essentially moves by one tick jumps, the price is denoted by

$$
P_t = N_t^+ - N_t^-.
$$
\n(4.2)

In this model,  $N_t^+$  $_t^\texttt{+}$  denotes the number of upward jumps of the asset in the time interval [0, t] whereas  $N_t^{\sim}$  $\tilde{t}_t^-$  corresponds to the number of downward jumps. As a result, the instantaneous probability of getting an upward or downward jump depends on the arrival times of the past upward and downward jumps.

We further simplify the model by setting

$$
\phi = \left(\begin{array}{cc} \mu_1 & \mu_3 \\ \mu_2 & \mu_4 \end{array}\right)
$$

which yields

$$
\begin{pmatrix} \lambda_t^{*+} \\ \lambda_t^{*-} \end{pmatrix} = \lambda_T \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t a_T \phi(t-s) \cdot \begin{pmatrix} dN_s^{T,+} \\ dN_s^{T,-} \end{pmatrix}
$$
 (4.3)

The tick-by-tick price model enables us to incorporate the following 4 characteristics of modern electronic trading.

- (1) Most of the orders hitting the market have no real economic reason but are rather put by algorithms as a reaction to other orders.
- (2) Statistical arbitrage in high-frequency markets is hard to create, meaning creating trading strategies that are on average possible is hardly possible.
- (3) The liquidity in the order book is not symmetrically distributed between the bid and ask sides, indicating that buying and selling are not equivalent actions. For instance, a market maker usually maintains a predominantly positive inventory. When receiving a buy order, a market maker is inclined to increase the price by a smaller margin compared to decreasing the price following a sell order of equivalent size. This is because a buy order reduces the market maker's inventory, which is favorable for them, while a sell order increases it.
- (4) A substantial number of transactions are attributed to large orders, known as metaorders. These are not executed immediately in their entirety but are instead broken down over time by trading algorithms.

In order to preserve the characteristics (2) and (3) in the assumed Hawkes framework, [\[25\]](#page-113-3) state that

$$
\mathbb{E}\left[N_t^+\right] = \int_0^t \mathbb{E}\left[\lambda_s^{*+}\right]ds, \quad \mathbb{E}\left[N_t^-\right] = \int_0^t \mathbb{E}\left[\lambda_s^{*-}\right]ds \tag{4.4}
$$

and

$$
\mathbb{E}\left[\lambda_t^{*+}\right] = \lambda^+ + \int_0^t \mu_1(t-s)\mathbb{E}\left[\lambda_s^{*+}\right]ds + \int_0^t \mu_3(t-s)\mathbb{E}\left[\lambda_s^{*-}\right]ds,
$$
\n
$$
\mathbb{E}\left[\lambda_t^{*-}\right] = \lambda^- + \int_0^t \mu_2(t-s)\mathbb{E}\left[\lambda_s^{*+}\right]ds + \int_0^t \mu_4(t-s)\mathbb{E}\left[\lambda_s^{*-}\right]ds.
$$
\n(4.5)

By setting

$$
\mu^+ = \mu^- \quad \text{and} \quad \mu_1 + \mu_3 = \mu_2 + \mu_4 \tag{4.6}
$$

we find a natural no-arbitrage argument, namely  $\mathbb{E}\big[\lambda_t^{**} \big]$  $\left[ \lambda_t^{*+} \right] = \mathbb{E} \left[ \lambda_t^{*-} \right]$ *t* ¤ . Characteristic (iii), which states that the ask side is more liquid than the bis side is translated as the conditional probability of observing a downward jump after an upward jump is smaller than the conditional probability of observing a downward jump right after a downward jump. This yields in the Hawkes framework to

 $\mu_1(t) < \mu_4(t)$  or similarly  $\mu_3(t) > \mu_2(t)$  when *t* is close to zero.

[\[25\]](#page-113-3) therefore makes the assumption that there exists

 $\beta$  > 1 such that

$$
\mu_3=\beta\mu_2.
$$

Thus, we assume the following structure for the intensity process:

$$
\left(\begin{array}{c} \lambda_t^{*+} \\ \lambda_t^{*-} \end{array}\right) = \lambda \left(\begin{array}{c} 1 \\ 1 \end{array}\right) + \int_0^t \phi(t-s) \cdot \left(\begin{array}{c} dN_s^{+} \\ dN_s^{-} \end{array}\right),
$$

where

$$
\phi = \begin{pmatrix} \mu_1 & \beta \mu_2 \\ \mu_2 & \mu_1 + (\beta - 1)\mu_2 \end{pmatrix},
$$

with  $\lambda > 0$  and  $\beta \ge 1$ .

In order to deal with characteristic 1, [\[25\]](#page-113-3) consider a bi-dimensional Hawkes process on an interval [0, *T*]. They assume that  $\lambda_T > 0$ 

$$
\phi^T = a_T \phi, \quad \phi = \begin{pmatrix} \mu_1 & \beta \mu_2 \\ \mu_2 & \mu_1 + (\beta - 1)\mu_2 \end{pmatrix}
$$

where  $\beta \geq 1$ ,  $\mu_1$  and  $\mu_2$  are two positive measurable functions such that

$$
\mathscr{S}\left(\int_0^\infty \phi(s) \, ds\right) = \|\mu_1\|_1 + \beta \, \|\mu_2\|_1 = 1
$$

and  $a<sub>T</sub>$  is an increasing sequence of positive numbers converging to one.

$$
\mathcal{S}\left(\int_0^\infty \phi(s)ds\right) = \left\|\varphi_1\right\|_1 + \beta \left\|\varphi_2\right\|_1,
$$

is defined as the spectral radius of the kernel matrix integral, where  $\mathscr S$  denotes the spectral radius operator.

#### **4.1.1.** THE CLASSICAL HESTON MODEL AS A LIMIT OF NEARLY UNSTABLE HAWKES PROCESSES

For additional details and a more comprehensive understanding, we direct the reader to the works of  $[25]$  and  $[3]$ . These references provide deeper insights into the underlying assumptions. Now, as T tends to infinity the re-scaled microscopic price

$$
\frac{1}{T}P_{tT}^T = \frac{N_{tT}^{T,+} - N_{tT}^{T,-}}{T}
$$

converges in law for the Skorokhod topology to the following Heston model:

$$
P_t = \frac{1}{1 - (\|\phi_1\|)_1 - (\|\phi_2\|)_1} \sqrt{\frac{2}{1 + \beta}} \int_0^t \sqrt{X_s} dW_s
$$

with

$$
dX_t = \frac{\lambda}{m} \left( \frac{(\beta + 1)\bar{\lambda}}{\lambda} - X_t \right) dt + \frac{1}{m} \sqrt{\frac{1 + \beta^2}{1 + \beta}} \sqrt{X_t} dB_t, X_0 = 0
$$

where (*W*,*B*) is a correlated bi-dimensional Brownian motion with

$$
d\langle W, B\rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}} dt
$$

and positive parameters  $\lambda$ ,  $\bar{\lambda}$ , and *m* such that

<span id="page-41-0"></span>
$$
T(1 - a_T) \to_{T \to \infty} \lambda, \lambda_T = \bar{\lambda}
$$
\n(4.7)

and

<span id="page-41-1"></span>
$$
\mathcal{S}\left(\int_0^{+\infty} x\phi(x)dx\right) = m < \infty.
$$
 (4.8)

#### **4.1.2.** THE ROUGH HESTON MODEL AS A LIMIT OF NEARLY UNSTABLE HAWKES **PROCESSES**

The model described previously does not account for the impact of metaorders in the market. To incorporate this effect within the Hawkes process framework, it is necessary to ensure that the kernel matrix reflects the heavy-tailed distribution that is typically observed in market data. This requires a modification of the existing conditions—specifically, assumptions [4.7](#page-41-0) and [4.8](#page-41-1) must be revised to allow for a kernel matrix that exhibits a slow decay, aligning with empirical observations. This implies a modification in the asymptotic setting in order to retrieve a non-degenerate scaling limit. More precisely, [4.7](#page-41-0) and [4.8](#page-41-1) can be exchanged by the following argument.

There exist  $\alpha \in (1/2, 1)$  and  $C > 0$  such that

$$
\alpha x^{\alpha} \int_{x}^{\infty} \mu_1(s) + \beta \mu_2(s) ds \rightarrow_{x \rightarrow \infty} C
$$

Moreover, for some  $\lambda^* > 0$  and  $\bar{\lambda} > 0$ ,

$$
T^{\alpha}(1-a_T) \to_{T \to \infty} \lambda^* > 0, T^{1-\alpha}\lambda_T \to_{T \to \infty} \bar{\lambda}
$$

Note that in practice, estimated values for  $\alpha$  are actually close to 1/2.

**Theorem 4.1.1.** *Under the previously mentioned assumptions, as T tends to infinity, the re-scaled microscopic price*

$$
\sqrt{\frac{1-a_T}{\mu T^{\alpha}}}P_{tT}^T, t\in[0,1]
$$

*converges in the sense of finite dimensional laws to the following macroscopic price process:*

$$
P_t = \frac{1}{1 - (\|\phi_1\|)_1 - (\|\phi_2\|)_1} \sqrt{\frac{2}{1 + \beta}} \int_0^t \sqrt{Y_s} dW_s,
$$

*with Y the unique solution of*

$$
Y_t = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda \left( (1+\beta) - Y_s \right) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda v \sqrt{Y_s} dB_s
$$

*where* (*W*,*B*) *is a correlated bi-dimensional Brownian motion with correlation*

$$
\rho = \frac{1 - \beta}{\sqrt{2(1 + \beta^2)}}
$$

*and*

$$
v = \sqrt{\frac{2(1+\beta^2)}{\lambda^* \bar{\lambda}(1+\beta)^2}}, \lambda = \frac{\alpha \lambda^*}{C\Gamma(1-\alpha)}
$$

*Furthermore, the process*  $Y_t$  *has Holder regularity*  $\alpha - 1/2 - \epsilon$  *for any*  $\epsilon > 0$ *.* 

Hence, this result shows that the main stylized facts of modern electronic markets naturally give rise to a very rough behavior of volatility.

Building upon additional assumptions and technical details—for instance, finding the appropriate Poisson rate and a specific choice of the Kernel matrix—for which we refer to [\[3\]](#page-112-0) and [\[25\]](#page-113-3), we can now provide a statement about the limiting behavior of the specific sequence of bi-dimensional nearly unstable Hawkes processes with heavy tails. For this purpose, we define for  $t \in [0,1]$ 

$$
X_t^T = \frac{1 - a_T}{T^{\alpha} \mu} N_{tT}^T, \quad \Lambda_t^T = \frac{1 - a_T}{T^{\alpha} \mu} \int_0^{tT} \lambda_s^T ds, \quad Z_t^T = \sqrt{\frac{T^{\alpha} \mu}{1 - a_T}} \left( X_t^T - \Lambda_t^T \right).
$$

<span id="page-42-0"></span>**Theorem 4.1.2.** As  $T \to \infty$ , the process  $(\Lambda_t^T, X_t^T, Z_t^T)_{t \in [0,1]}$  converges in law for the Sko*rokhod topology to* (Λ,*X*,*Z*)*, where*

$$
\Lambda_t = X_t = \int_0^t Y_s ds \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad Z_t = \int_0^t \sqrt{Y_s} \begin{pmatrix} dB_s^1 \\ dB_s^2 \end{pmatrix}
$$

*and Y is the unique solution of the rough stochastic differential equation*

$$
Y_t = \xi + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda (1-Y_s) ds + \lambda \sqrt{\frac{1+\beta^2}{\lambda \bar{\lambda} (1+\beta^2)}} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{Y_s} dB_s,
$$

*where*

$$
B = \frac{B^1 + \beta B^2}{\sqrt{1 + \beta^2}}
$$

*and*  $(B^1, B^2)$  *is a bi-dimensional Brownian motion. Furthermore, for any*  $\varepsilon > 0$ , *Y* has *Hölder regularity α*−1/2−*ε.*

Taking into account [4.1.2](#page-42-0) yields a microscopic process converging to the log price. More precisely, for  $\theta > 0$ , [\[3\]](#page-112-0) define

$$
P^{T} = \sqrt{\frac{\theta}{2}} \sqrt{\frac{1 - a_{T}}{T^{\alpha} \mu}} \left( N_{.T}^{T,+} - N_{.T}^{T,-} \right) - \frac{\theta}{2} \frac{1 - a_{T}}{T^{\alpha} \mu} N_{.T}^{T,+} = \sqrt{\frac{\theta}{2}} \left( Z^{T,+} - Z^{T,-} \right) - \frac{\theta}{2} X^{T,+}.
$$

 ${\bf Corollary 4.1.2.1.} \ \ As \ T \to \infty,$  the sequence of processes  $\left(P_t^T\right)_{t \in [0,1]}$  converges in law for the *Skorokhod topology to*

$$
P_t = \int_0^t \sqrt{V_s} dW_s - \frac{1}{2} \int_0^t v_s ds
$$

*where V is the unique solution of the rough stochastic differential equation*

$$
v_t = \theta \xi + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda (\theta - v_s) ds + \lambda \sqrt{\frac{\theta (1+\beta^2)}{\lambda \bar{\lambda}(1+\beta)^2} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \sqrt{V_s} dB_s}
$$

*with* (*W*,*B*) *a correlated bi-dimensional Brownian motion whose bracket satisfies*

$$
d\langle W,B\rangle_t = \frac{1-\beta}{\sqrt{2(1+\beta^2)}} dt.
$$

By setting

$$
v_0 = \xi \theta, \rho = \frac{1 - \beta}{\sqrt{2(1 + \beta^2)}}, v = \sqrt{\frac{\theta(1 + \beta^2)}{\lambda \mu(1 + \beta^2)}}, \theta = \bar{v}, \lambda = \kappa
$$

we define our rough Heston model as used in this thesis as

**Definition 4.1.1.** *(Rough Heston model)*

$$
dS_t = S_t \sqrt{\nu_t} dW_t
$$
  

$$
\nu_t = \nu_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \kappa (\bar{\nu} - \nu_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \nu \sqrt{\nu_s} dB_s,
$$
 (4.9)

*with*  $\alpha = H + 0.5$ *. v* denotes the long-term mean of the variance process,  $\kappa$  is the speed *of mean reversion, ν is the volatility of the volatility, ρ is the correlation between the two Brownian motion and v*<sup>0</sup> *defines the initial variance.*

Note that applying Ito's Lemma to the asset price process *dS<sup>t</sup>* yields the form of the Price process  $P_t$  in the Corollary above.

#### **4.2.** LIFTED HESTON

Another model which has significant importance throughout this thesis is the Lifted Heston model. Introduced by [\[26\]](#page-113-4), we construct a multi-factor stochastic volatility model which, in contrast to the standard rough Heston model, enjoys the Markov property. A similar model has been developed by [\[27\]](#page-113-5).

Recalling that the rough Heston model is defined as

$$
dS_t = S_t \sqrt{\nu_t} dW_t
$$
  

$$
\nu_t = \nu_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \kappa (\bar{\nu} - \nu_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \nu \sqrt{\nu_s} dB_s,
$$
(4.10)

with  $\alpha = H + 0.5$ . The idea of [\[2\]](#page-112-1) is to write the fractional Kernel

$$
K(t)=\frac{t^{H-0.5}}{\Gamma(H+0.5)}
$$

as a Laplace transform of a positive measure *µ*, i.e.

$$
K(t) = \int_0^\infty e^{-xt} \mu(dx); \quad \mu(dx) = \frac{x^{-H-\frac{1}{2}}}{\Gamma(H+1/2)\Gamma(1/2-H)} dx. \tag{4.11}
$$

Afterwards, the measure  $\mu$  is approximated by a finite sum of Dirac measures  $\mu^n = \sum_{i=1}^n c_i^n \delta_{x_i^n}$ with positive weights  $\left(c_i^n\right)_{1\leq i\leq n}$  and mean reversions  $\left(x_i^n\right)_{1\leq i\leq n}$ , for  $n\geq 1$ . This yields an approximation of the fractional Kernel by a sum of smoothed Kernels as follows

$$
K^{n}(t) = \sum_{i=1}^{n} c_{i}^{n} e^{-x_{i}^{n}t}, \quad n \ge 1.
$$
 (4.12)

As a result, we can define the Lifted Heston model as follows. We fix an *n* ∈ N,

<span id="page-44-0"></span>
$$
dS_t^n = S_t^n \sqrt{\nu_t^n} dW_t, \quad S_0^n > 0
$$
  

$$
\nu_t^n = g_0^n(t) + \sum_{i=1}^n c_i^n U_t^{n,i}
$$
  

$$
dU_t^{n,i} = \left(-x_i^n U_t^{n,i} - \kappa \nu_t^n\right) dt + \nu \sqrt{\nu_t^n} dB_t, \quad U_0^{n,i} = 0, \quad i = 1, ..., n,
$$
\n(4.13)

with parameters the function  $g_0^n$ ,  $\kappa$ ,  $v \in \mathbb{R}_+$ ,  $c_i^n$ ,  $x_i^n \ge 0$ , for  $i = 1,...,n$ , and  $W = \rho B + \rho B$  $\sqrt{1-\rho^2}B^{\perp}$ , with  $(B,B^{\perp})$  a two-dimensional Brownian motion on a fixed filtered probability space  $(\Omega, \mathscr{F}, \mathbb{F} := (\mathscr{F}_t)_{t \geq 0}, \mathbb{Q})$ , with  $\rho \in [-1, 1]$ .

The function  $g_0^n$  is often referred to as the forward variance curve. Taking expectation on both sides of the variance process in  $4.13$  and taking into account that  $g_0^n$  is deterministic, we find that

<span id="page-44-1"></span>
$$
\mathbb{E}\left[v_t^n\right] + \kappa \sum_{i=1}^n c_i^n \int_0^t e^{-x_i^n(t-s)} \mathbb{E}\left[v_s^n\right] ds = g_0^n(t), \quad t \ge 0. \tag{4.14}
$$

 $g_0^n$  allows one to give initial term structure curves. In practice, these forward variance curves can be obtained from variance swaps and are plugged in instead of  $\mathbb{E}\left[ \nu_t^n \right]$  in [4.14.](#page-44-1) As long as the forward variance curve satisfies certain conditions for which we refer to [\[26\]](#page-113-4), we can define a whole set of admissible input curves. Going further, we define the forward variance curve as

$$
g_0^n : t \to v_0 + \kappa \bar{\nu} \sum_{i=1}^n c_i^n \int_0^t e^{-x_i^n(t-s)} ds, \quad \text{with } v_0, \bar{\nu} \ge 0.
$$
 (4.15)

#### **4.2.1.** PARAMETRIZATION OF THE WEIGHTS AND THE MEAN REVERSIONS

As we saw above the weights  $(c_i^n)_{1 \le i \le n}$  and mean reversions  $(x_i^n)_{1 \le i \le n}$ , for  $n \ge 1$  play a major role in defining the Lifted Heston model. We want to choose these parameters in a way such that the Lifted Heston model actually converges to the rough Heston model, whenever  $n \to \infty$  [\[26\]](#page-113-4). For this purpose we first state the theorem mentioned in [26].

**Theorem 4.2.1.** *(Representation of the limiting rough process). The fractional kernel appearing in the variance process of the rough Heston model has the following Laplace representation*

$$
\frac{t^{H-1/2}}{\Gamma(H+1/2)} = \int_0^\infty e^{-xt} \mu(dx), \quad \text{with } \mu(dx) = \frac{x^{-H-1/2}}{\Gamma(1/2-H)\Gamma(H+1/2)},
$$

*so that the stochastic Fubini theorem, after setting*  $V_0 \equiv 0$  *in* [5.2,](#page-49-0) *leads to* 

$$
v_t = \int_0^\infty U_t(x)\mu(dx), \quad x > 0
$$

*where, for all*  $x > 0$ *,* 

<span id="page-45-0"></span>
$$
U_t(x) := \int_0^t e^{-x(t-s)} \left( \kappa (\bar{v} - v_s) \, ds + \nu \sqrt{v_s} \, dW_s \right). \tag{4.16}
$$

*This can be interpreted as the mild formulation of the subsequent stochastic differential equation (SDE):*

$$
dU_t(x) = \left(-xU_t(x) + \kappa \left(\bar{v} - \int_0^\infty U_t(y)\mu(dy)\right)\right)dt + \nu \sqrt{\int_0^\infty U_t(y)\mu(dy)}dW_t
$$
  

$$
U_0(x) = 0, \quad x > 0
$$

By fixing  $n \ge 1$ , we can interpret [4.13](#page-44-0) as a discretization of [4.2.1](#page-45-0) once the measure  $\mu$ is approximated by a sum of Dirac measures  $\mu^n = \sum_{i=1}^n c_i^n \delta_{x_i^n}$ .

Taking into account this discretization, we can approximate

$$
v_t = \int_0^\infty U_t(x) \mu(dx), \quad x > 0
$$

by means of [\[28\]](#page-113-6) and [\[2\]](#page-112-1) by

$$
\hat{v}_t = \sum_i c_i U(x_i, t).
$$

Thus, we can choose according to [\[26\]](#page-113-4)

$$
c_i^n = \int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx), \quad x_i^n = \frac{1}{c_i^n} \int_{\eta_{i-1}^n}^{\eta_i^n} x \mu(dx), \quad i \in \{1, \dots, n\}.
$$
 (4.17)

As a result, by fixing  $r_n > 1$  and using the geometric partition  $\eta_i^n = r_n^{i-n/2}$  for  $i = 0, ..., n$ as proposed in [\[28\]](#page-113-6), we can set

<span id="page-45-1"></span>
$$
c_i^n = \frac{(r_n^{1-\alpha} - 1) r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)i}, \quad i = 1, ..., n
$$
 (4.18)

and

<span id="page-45-2"></span>
$$
x_i^n = \frac{1 - \alpha}{2 - \alpha} \frac{r_n^{2 - \alpha} - 1}{r_n^{1 - \alpha} - 1} r_n^{i - 1 - n/2}, \quad i = 1, ..., n
$$
 (4.19)

where  $\alpha := H + 1/2$  for some  $H \in (0, 1/2)$ . Last but not least [\[26\]](#page-113-4) show that once the sequence  $(r_n)_{n\geq 1}$  satisfies  $r_n \downarrow 1$  and  $n \ln(r_n) \rightarrow \infty$ , as  $n \rightarrow \infty$ , the Lifted Heston model converges to the rough Heston model for  $n \rightarrow \infty$ . According to [\[26\]](#page-113-4), we define  $r_n$ as the following sequence

$$
r_n = 1 + 10n^{-0.9}, \quad n \ge 1.
$$

## **5**

### **CHARACTERISTIC FUNCTION OF THE ROUGH HESTON MODEL**

The characteristic function plays an important role in option pricing theory as it enables to use Fourier-based methods to efficiently price options. In this section, we provide a detailed analysis of the characteristic function of the rough Heston and lifted Heston model introduced by [\[3\]](#page-112-0) and [\[26\]](#page-113-4), respectively.

#### **5.1.** CHARACTERISTIC FUNCTION OF THE CLASSICAL HESTON MODEL

In the classical Heston model, the characteristic function can be derived as follows. For

$$
\mathbf{X}(t) = [X(t) = log(S(t)), v(t)]^{\mathrm{T}},
$$

the solution  $\phi$ **X** :=  $\phi$ **X**(*u*; *t*, *T*) satisfies the following pricing PDE :

$$
0 = -\frac{\partial \phi_{X}}{\partial \tau} - \frac{1}{2} \nu \frac{\partial \phi_{X}}{\partial X} + \kappa (\bar{\nu} - \nu(t)) \frac{\partial \phi_{X}}{\partial \nu} + \frac{1}{2} \nu^{2} \nu \frac{\partial^{2} \phi_{X}}{\partial \nu^{2}} + \frac{1}{2} \nu \frac{\partial^{2} \phi_{X}}{\partial X^{2}} + \rho \nu \nu \frac{\partial^{2} \phi_{X}}{\partial X \partial \nu}
$$
(5.1)

subject to the initial condition  $\phi_{\mathbf{X}}(u; T, T) = \exp(i u X(0))$ , for  $\tau = T - t = 0$ . We can easily prove that the process  $X(t)$  has an affine structure and according to  $[13]$  the solution of the PDE is of the following form

$$
\phi_{\mathbf{X}}(u; t, T) = \exp(\bar{A}(\mathbf{u}, \tau) + \bar{B}(\mathbf{u}, \tau)X(t) + \bar{C}(\mathbf{u}, \tau)v(t)).
$$
\n(5.2)

By substituting the proposed solution into the pricing PDEs, we find the following set of Riccati ODEs

$$
\frac{d\bar{B}}{dt} = 0, \quad \bar{B}(\mathbf{u}, 0) = i u,\n\frac{d\bar{C}}{dt} = \bar{B}(\bar{B} - 1)/2 - (\kappa - v\rho \bar{B})\bar{C} + v^2 \bar{C}^2/2, \quad \bar{C}(\mathbf{u}, 0) = 0,\n\frac{d\bar{A}}{dt} = \kappa \bar{v}\bar{C}, \quad \bar{A}(\mathbf{u}, 0) = 0.
$$
\n(5.3)

According to [\[8\]](#page-112-3) the solution of the set of ODEs is given by

$$
\bar{B}(\mathbf{u}, \tau) = iu
$$
\n
$$
\bar{C}(\mathbf{u}, \tau) = \frac{1 - e^{-D_1 \tau}}{v^2 (1 - g e^{-D_1 \tau})} \left( \kappa - v \rho i u - D_1 \right),
$$
\n
$$
\bar{A}(\mathbf{u}, \tau) = \frac{\kappa \bar{v} \tau}{v^2} \left( \kappa - v \rho i u - D_1 \right) - \frac{2\kappa \bar{v}}{v^2} \log \left( \frac{1 - g e^{-D_1 \tau}}{1 - g} \right),
$$
\n(5.4)

with  $D_1 = \sqrt{(\kappa - v\rho i u)^2 + (u^2 + i u) v^2}$  and  $g = \frac{\kappa - v\rho i u - D_1}{\kappa - v\rho i u + D_1}$ *κ*−*νρi u*+*D*<sup>1</sup> .

Hence, the characteristic function of the classical Heston model has indeed a closedform solution.

#### <span id="page-49-0"></span>**5.2.** CHARACTERISTIC FUNCTION OF THE ROUGH HESTON MODEL

Again, for further technicalities and an in-depth treatment of the following, we refer to [\[3\]](#page-112-0). Here, we only state their main result. To do so, we first set

$$
\nu_0 = \xi \theta, \rho = \frac{1 - \beta}{\sqrt{2(1 + \beta^2)}}, \nu = \sqrt{\frac{\theta(1 + \beta^2)}{\lambda \mu(1 + \beta^2)}}, \theta = \bar{\nu}, \lambda = \kappa
$$

This implies that  $\rho \in (-1/\sqrt{2},1/\sqrt{2}]$ . We also write  $P_t = \log(S_t/S_0)$  and hence consider the following rough Heston model:

$$
dS_t = S_t \sqrt{\nu_t} dW_t
$$
  

$$
\nu_t = \nu_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \kappa (\bar{\nu} - \nu_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \nu \sqrt{\nu_s} dB_s
$$

The parameters  $\kappa$ ,  $\bar{v}$ ,  $V_0$  and  $\nu$  are positive and play the same role as in the Heston model. *W* and *B* are two Brownian motions with correlation *ρ*.

**Theorem 5.2.1.** *Consider the rough Heston model [5.2](#page-49-0) with a correlation between the two* **Brownian 5.2.1.** Consider the rough Heston model 5.2 with a correlation *Brownian motions ρ satisfying ρ* ∈ (−1/√2,1/√2]*. For all t* ≥ 0*, we have* 

$$
\phi_{\mathbf{X}}(a;t) = \exp\left(\bar{\nu}\kappa I^{1}h(a,t) + \nu_{0}I^{1-\alpha}h(a,t)\right)
$$

*where h is the solution of the fractional Riccati equation*

$$
D^{\alpha}h(a,t) = \frac{1}{2}(-a^2 - ia) + \kappa(ia\rho v - 1)h(a,s) + \frac{v^2}{2}h^2(a,s), \quad I^{1-\alpha}h(a,0) = 0,
$$

#### *which admits a unique continuous solution.*

However, in contrast to the classical model, the fractional Riccati differential equation does not allow for a closed-form solution. Thus, we have to numerically approximate the solution, which will be explained in more detail in the next chapter.

#### **5.3.** CHARACTERISTIC FUNCTION OF THE LIFTED HESTON MODEL

Before defining the characteristic function, we emphasize that the lifted Heston model is a n-dimensional multi-factor approximation of the rough Heston model. Hence, we will see in order to get the characteristic function a n-dimensional Ricatti equation has to be solved. Thus, we will continue working with  $\log S_t^n$  as our n-dimensional log-stock.

The Fourier-Laplace transform of the Lifted Heston model with respect to the logstock price is introduced by [\[26\]](#page-113-4) and denoted as

$$
\mathbb{E}\left[\exp\left(u\log S_t^n\right)|\mathcal{F}_t\right] = \exp\left(\phi^n(t,T) + u\log S_t^n + \sum_{i=1}^n c_i^n \psi^{n,i}(T-t)U_t^{n,i}\right),\tag{5.5}
$$

where  $(\psi^{n,i})_{1\leq i\leq n}$  solves the following *n*-dimensional system of Riccati ODE's

<span id="page-50-0"></span>
$$
\left(\psi^{n,i}\right)' = -x_i^n \psi^{n,i} + F\left(u, \sum_{j=1}^n c_j^n \psi^{n,j}\right), \quad \psi^{n,i}(0) = 0, \quad i = 1, \dots, n,\tag{5.6}
$$

with

$$
F(u, v) = \frac{1}{2} (u^2 - u) + (\rho v u - \kappa) v + \frac{v^2}{2} v^2,
$$
\n(5.7)

and

$$
\phi^{n}(t,T) = \int_{0}^{T-t} F\left(u, \sum_{i=1}^{n} c_{i}^{n} \psi^{n,i}(s)\right) g_{0}^{n}(T-s) ds, \quad t \leq T.
$$

 $x_i^n$  and  $c_i^n$  in [5.6](#page-50-0) are defined in [4.18](#page-45-1) and [4.19.](#page-45-2)

Note that for  $t = 0$ , we find that  $U_0^{n,i} = 0$  and hence the unconditioned Fourier-Laplace transform is given by

$$
\phi_{X^n}(u,T) = \mathbb{E}\left[\exp\left(u\log S_t^n\right)\right] = \exp\left(u\log S_0^n + \int_0^T F\left(u, \sum_{i=1}^n c_i^n \psi^{n,i}(s)\right) g_0^n(T-s) ds\right). \tag{5.8}
$$

Even though the solution of the n-dimensional Riccati equation cannot be obtained in closed form, it is of great computational advantage to numerically solve the n-dimensional system of Riccati equations instead of the fractional Riccati equation in the rough Heston model.

# **6**

## **PRICING EUROPEAN OPTIONS UNDER THE ROUGH HESTON MODEL**

#### **6.1.** 1D COS METHOD FOR EUROPEAN OPTIONS

In this section, we introduce the one-dimensional COS method developed by [\[4\]](#page-112-4). The Feynman-Kac theorem is a cornerstone in the pricing of European options, serving as a fundamental starting point whether one employs Monte Carlo simulations or Fourierbased numerical integration schemes such as the Carr-Madan method. Accordingly, we begin with the risk-neutral pricing formula:

$$
\nu(x, t_0) = e^{-r\Delta t} \mathbb{E}^{\mathbb{Q}}[\nu(y, T) | x] = e^{-r\Delta t} \int_{\mathbb{R}} \nu(y, T) f(y | x) dy \tag{6.1}
$$

Here, *v* represents the option value,  $\Delta t$  is the time interval between the maturity date  $T$  and the initial date  $t_0$ , and  $\mathbb{E}^{\mathbb{Q}}[\cdot]$  denotes the expectation under the risk-neutral measure  $\mathbb{Q}$ . The variables *x* and *y* correspond to state variables at times  $t_0$  and *T*, respectively. Furthermore,  $f(y \mid x)$  is the probability density of *y* given *x*, and *r* is the riskneutral interest rate. However, in practical scenarios, the probability density of more complex price processes is often unknown, while their characteristic function is typically available. To leverage the availability of the characteristic function, [\[4\]](#page-112-4) developed the COS method. This innovative pricing method is particularly effective for Levy and Heston processes, allowing for the efficient pricing of options with various strikes in a single computation.

We commence by establishing the Fourier pair, which delineates the relationship between the density and the characteristic function, as follows:

<span id="page-52-0"></span>
$$
\phi(\omega) = \int_{\mathbb{R}} e^{ix\omega} f(x) dx, \tag{6.2}
$$

<span id="page-53-0"></span>
$$
f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x} \phi(\omega) d\omega.
$$
 (6.3)

The idea of [\[4\]](#page-112-4) is to reconstruct the whole integral [6.3](#page-53-0) from Fourier-cosine expansion.

For an arbitrary function f supported on [0,*T* ] the cosine expansion is denoted by

$$
f(\theta) = \sum_{k=0}^{\infty} A_k \cdot \cos(k\theta) \quad \text{with} \quad A_k = \frac{2}{\pi} \int_0^{\pi} f(\theta) \cos(k\theta) d\theta,\tag{6.4}
$$

where  $\sum'$  reads that the first term in the summation is multiplied by  $\frac{1}{2}$ . In case the function f is supported on an arbitrary interval  $[a, b] \in \mathbb{R}$ , the cosine expansion can be derived by the following change of variables

$$
\theta := \frac{x-a}{b-a}\pi, \quad x = \frac{b-a}{\pi}\theta + a
$$

As a result, we find that

$$
f(x) = \sum_{k=0}^{\infty} A_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right),\tag{6.5}
$$

with

<span id="page-53-2"></span>
$$
A_k = \frac{2}{b-a} \int_a^b f(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx.
$$
 (6.6)

A pivotal aspect of the COS method is its reliance on the relationship between the Fourier cosine coefficients, denoted as  $A_k$ , and the characteristic function. Given that the density function is confined to the interval [*a*,*b*], the corresponding characteristic function is evaluated within this finite range. The accuracy of the density representation over the interval  $[a, b]$  is only marginally compromised, allowing the truncated integral to serve as a highly precise approximation of its infinite counterpart, as referenced in [6.2.](#page-52-0) This approximation is expressed as follows:

<span id="page-53-1"></span>
$$
\hat{\phi}(\omega) := \int_{a}^{b} e^{i\omega x} f(x) dx \approx \int_{\mathbb{R}} e^{i\omega x} f(x) dx = \phi(\omega). \tag{6.7}
$$

In this formulation,  $\hat{\phi}(\omega)$  represents the approximation of the characteristic function over the finite interval [ $a$ , $b$ ], closely mirroring the true characteristic function  $\phi(\omega)$ , which is integrated over the entire real line. This approach effectively harnesses the power of the COS method in capturing the essential characteristics of the option pricing model within a constrained domain.

Comparing  $6.7$  with  $A_k$  in  $6.6$  yields to

$$
A_k = \frac{2}{b-a} \operatorname{Re} \left\{ \hat{\phi} \left( \frac{k\pi}{b-a} \right) \cdot \exp \left( -i \frac{k a \pi}{b-a} \right) \right\},\tag{6.8}
$$

where Re{·} indicates taking the real part of the argument.

Taking into account [6.7](#page-53-1) with  $\hat{\phi}(\omega) \approx \phi(\omega)$  yielding to  $A_k \approx F_k$ , we find that

$$
F_k = \frac{2}{b-a} \operatorname{Re} \left\{ \phi \left( \frac{k\pi}{b-a} \right) \cdot \exp \left( -i \frac{k a \pi}{b-a} \right) \right\}.
$$
 (6.9)

By replacing  $A_k$  with  $F_k$ , we find for the approximation of the density function  $f(x)$ on [*a*,*b*]

$$
\hat{f}(x) = \sum_{k=0}^{\infty} F_k \cos\left(k\pi \frac{x-a}{b-a}\right)
$$
\n(6.10)

and truncate the series summation such that

$$
\hat{f}(x) = \sum_{k=0}^{N-1} F_k \cos\left(k\pi \frac{x-a}{b-a}\right).
$$
\n(6.11)

For pricing European options, we have the following formula

$$
v_1(x, t_0) = e^{-r\Delta t} \int_a^b v(y, T) f(y | x) dy
$$
 (6.12)

The conditional density  $f(y | x)$  is often not known but the characteristic function is. As a result, we replace the density by its cosine expansion in  $\gamma$ , i.e.

$$
f(y \mid x) = \sum_{k=0}^{+\infty} A_k(x) \cos\left(k\pi \frac{y-a}{b-a}\right)
$$
 (6.13)

with

$$
A_k(x) := \frac{2}{b-a} \int_a^b f(y \mid x) \cos\left(k\pi \frac{y-a}{b-a}\right) dy.
$$
 (6.14)

This yields the following expression for the price of a European option

<span id="page-54-0"></span>
$$
v(x, t_0) = e^{-r\Delta t} \int_a^b v(y, T) \sum_{k=0}^{+\infty} A_k(x) \cos\left(k\pi \frac{y-a}{b-a}\right) dy.
$$
 (6.15)

By interchanging summation and integration in  $6.15$ , we find that

$$
\nu(x, t_0) = \frac{1}{2}(b - a) \cdot e^{-r\Delta t} \cdot \sum_{k=0}^{+\infty} A_k(x) \cdot V_k,
$$
\n(6.16)

with

$$
V_k := \frac{2}{b-a} \int_a^b v(y, T) \cos\left(k\pi \frac{y-a}{b-a}\right) dy \tag{6.17}
$$

Lastly, by replacing  $A_k(x)$  by  $F_k(x)$  and setting as seen above and defining  $x := \ln(S_0/K)$ and  $y := \ln(S_T/K)$  we have,

$$
\nu(x, t_0) \approx e^{-r\Delta t} \sum_{k=0}^{N-1'} \text{Re}\left\{\phi\left(\frac{k\pi}{b-a}; x\right) e^{-ik\pi \frac{a}{b-a}}\right\} V_k.
$$
 (6.18)

As described in [\[4\]](#page-112-4), we apply the following property when pricing options under the (rough) Heston model.

$$
\nu(x, t_0) \approx e^{-r\Delta t} \sum_{k=0}^{N-1} \text{Re}\left\{\phi\left(\frac{k\pi}{b-a}\right) e^{-ik\pi \frac{x-a}{b-a}}\right\} V_k.
$$
 (6.19)

#### TRUNCATION RANGE

The accuracy of option price predictions using the COS method significantly depends on the precise determination of the integration bounds. Setting these bounds too wide requires employing a larger number of basis functions, which consequently increases computational complexity. Conversely, overly narrow bounds can lead to increased truncation errors. To achieve an optimal balance, it is recommended to set the bounds as follows:

$$
[a, b] = [c_1 + x_0 - L\sqrt{c_2}, c_1 + x_0 + L\sqrt{c_2}]
$$

with  $L = 8$ ,

$$
c_1 = T + \left(1 - e^{-\kappa T}\right) \frac{\bar{v} - v_0}{2\kappa} - \frac{\bar{v}T}{2},
$$

and

$$
c_2=\bar{v}(1+v)\,T.
$$

This approach ensures that the bounds are neither excessively wide, which would complicate the computation, nor too narrow, which might compromise the accuracy due to truncation. By carefully selecting these parameters, the COS method can be utilized effectively for efficient and accurate option pricing.

#### **6.2.** NUMERICAL APPROXIMATION OF THE ROUGH HESTON CHAR-ACTERISTIC FUNCTION

#### **6.2.1.** ADAMS SCHEME

As seen before, the unconditioned characteristic function is entirely defined through the fractional Riccati equation

$$
D^{\alpha}h(a,t) = \frac{1}{2} \left( -a^2 - ia \right) + \kappa(ia\rho v - 1)h(a,s) + \frac{v^2}{2} h^2(a,s), \quad I^{1-\alpha}h(a,0) = 0.
$$

Using the fractional integral operator on both sides yields

<span id="page-55-0"></span>
$$
h(a,t) = \frac{1}{\Gamma(a)} \int_0^t (t-s)^{\alpha-1} F(a, h(a,s)) \, ds. \tag{6.20}
$$

For [6.20](#page-55-0) we use numerical methods and start with the well-known Adams method as proposed in [\[3\]](#page-112-0). The method is described as follows:

First, we define  $g(a, t) = F(a, h(a, t))$  and define the discrete time grid  $(t_k)_{k \in \mathbb{N}}$  with mesh  $\Delta(t_k = k\Delta)$ . Now, we estimate

$$
h(a, t_{k+1}) = \frac{1}{\Gamma(\alpha)} \int_0^{t_{k+1}} (t_{k+1} - s)^{\alpha - 1} g(a, s) \, ds
$$

by

$$
\frac{1}{\Gamma(\alpha)}\int_0^{t_{k+1}}(t_{k+1}-s)^{\alpha-1}\hat{g}(a,s)\ ds,
$$

where

$$
\hat{g}(a,t)=\frac{t_{j+1}-t}{t_{j+1}-t_j}\hat{g}(a,t_j)+\frac{t-t_{j+1}}{t_{j+1}-t_j}\hat{g}(a,t_{j+1}),\;\;t\in[t_j,t_{j+1}),\;\;0\leq j\leq k.
$$

This leads to the following expression:

$$
\hat{h}(a,t_{k+1}) = \sum_{0 \le j \le k} a_{j,k+1} F(a,\hat{h}(a,t_j)) + a_{k+1,k+1} F(a,\hat{h}(a,t_{k+1}))
$$

with

<span id="page-56-0"></span>
$$
a_{0,k+1} = \frac{\Delta^{\alpha}}{\Gamma(\alpha+2)} (k^{\alpha+1} - (k-\alpha)(k+1)^{\alpha})
$$
 (6.21)

$$
a_{j,k+1} = \frac{\Delta^{\alpha}}{\Gamma(\alpha+2)} \big( (k-j+2)^{\alpha+1} + (k-j)^{\alpha+1} - 2(k-j+1)^{\alpha+1} \big), \ \ 1 \le j \le k \tag{6.22}
$$

and

$$
a_{k+1,k+1} = \frac{\Delta^{\alpha}}{\Gamma(\alpha+2)}.
$$

We realize that the  $6.21$  is implicit. However, according to  $[3]$  we can use a so-called predictor, denoted by  $\hat{h}^P(a,t_{k+1})$  and defined as

$$
\hat{h}^{P}(a, t_{k+1}) = \frac{1}{\Gamma(\alpha+1)} \int_{0}^{t} (t-s)^{\alpha-1} \tilde{g}(a, s) ds
$$
\n(6.23)

with

$$
\tilde{g}(a,t) = \hat{g}(a,t_k), \ \ t \in [t_k,t_{k+1}).
$$

Hence, we have

$$
\hat{h}^{P}(a, t_{k+1}) = \sum_{0 \le j \le k} b_{j,k+1} F(a, \hat{h}(a, t_j))
$$
\n(6.24)

with

$$
b_{j,k+1} = \frac{\Delta^{\alpha}}{\Gamma(\alpha+1)}((k-j+1)^{\alpha} - (k-j)^{\alpha})
$$

As a result, we have the final implicit scheme

$$
\hat{h}(a, t_{k+1}) = \sum_{0 \le j \le k} a_{j,k+1} F(a, \hat{h}(a, t_j)) + a_{k+1,k+1} F(a, \hat{h}^P(a, t_j)), \ \hat{h}(a, 0) = 0. \tag{6.25}
$$

However, we observe that in terms of computation time, the Adams Scheme is rather inefficient. The computational complexity is of  $O\big(N_{cos}N^2\big)$ , where  $N_{cos}$  is the number of space steps taken for the COS method and *N* for the use of the fractional Adams method [\[3\]](#page-112-0). We realize that especially for small *H* the Adams method needs a lot of steps to sufficiently converge to the solution of the fractional Riccati equation.

<span id="page-57-0"></span>

Figure 6.1: Convergence in the number of cosine basis functions with model parameter  $\rho = -0.6$ ,  $\bar{v} = 0.05$ ,  $v = 0.4$ ,  $\kappa = 2$  and  $v_0 = 0.0625$ . Moreover, we have  $S_0 = 100$ ,  $K = 100$  and  $T = 1.0$ 

#### **6.2.2.** NUMERICAL RESULTS

We first assume  $H = \frac{1}{2}$ . As a result, the rough Heston model coincides with the classical Heston model. In order to show the convergence and validity of the Adams method, we benchmark our prices using the 1D COS method and the classical Heston model.

It can be seen in  $6.1$  that with around 60 basis functions there is a good convergence towards the benchmark price. However, choosing the right number of cosine basis functions plays a major role in the overall speed of the pricing process. We note that it also heavily depends on the set of parameters chosen and the time to maturity. We can see that the amount of time it takes to calculate one pricing grows more or less linearly with the number of *Ncos*.

<span id="page-57-1"></span>

Figure 6.2: Convergence in the number of time steps taken in the Adams method with model parameter  $\rho = -0.8$ ,  $\bar{v} = 0.05$ ,  $v = 0.4$ ,  $\kappa = 2$  and  $v_0 = 0.0625$ . Moreover, we have  $S_0 = 100$ ,  $K = 100$  and  $T = 1.0$ 

In [6.2,](#page-57-1) the convergence with respect to the number of time steps in the Adams method is considered. Again, the benchmark is the price of the option under the classical Heston model using the 1D COS method. Remarkable is the amount of time to archive a very good approximation of the option price under the classical Heston model. However, choosing 20,000-time steps to archive an accuracy of 10−<sup>5</sup> and better, yields a computational time of around 9 minutes, which is not really practical. In the literature, the Adams scheme is often seen as the benchmark method to get the characteristic function which

is used for pricing options under the rough Heston model. For the following calculations, we use  $N = 1,500$ .

We calculate the implied volatility surface and especially the ATM skew under the rough Heston model and compare it to the classical Heston model.



Figure 6.3: Implied volatility surface created using the rough Heston dynamics with parameters *ρ* = −0.4,  $\bar{v}$  = 0.0678,  $v$  = 0.3,  $\kappa$  = 2.6,  $v_0$  = 0.1 and *H* = 0.12

In [6.5,](#page-59-0) we can observe that for larger maturities the term structure tends to be the same for both models. However, we notice that for short-term maturities the ATM skew of the classical Heston model tends to be constant, while the rough Heston dynamics generate ATM skews in line with  $\psi(\tau) = A \tau^{H-\frac{1}{2}}$  as it can be seen in [6.4](#page-59-1)

Implied Volatility Surface

<span id="page-59-1"></span>

<span id="page-59-0"></span>Figure 6.4: Term structure of the at-the-money implied volatility skew created using the rough Heston dynamics vs. Fukusawa's approach: *ψ*(*τ*) = *Aτ <sup>H</sup>*−1/2 with *<sup>H</sup>* <sup>=</sup> 0.12.



Figure 6.5: Term structure of the at-the-money implied volatility skew created using the rough Heston dynamics vs. classical Heston dynamics.

#### **6.3.** NUMERICAL APPROXIMATION OF THE LIFTED HESTON CHAR-ACTERISTIC FUNCTION

#### **6.3.1.** IMPLICIT-EXPLICIT EULER SCHEME

We first recall that in order to get the characteristic function for the lifted Heston model, we have to solve the following n-dimensional Riccati equation

<span id="page-60-0"></span>
$$
\left(\psi^{n,i}\right)' = -x_i^n \psi^{n,i} + F\left(u, \sum_{j=1}^n c_j^n \psi^{n,j}\right), \quad \psi^{n,i}(0) = 0, \quad i = 1, \dots, n,\tag{6.26}
$$

with

$$
F(u, v) = \frac{1}{2} (u^2 - u) + (\rho v u - \kappa) v + \frac{v^2}{2} v^2.
$$
 (6.27)

In order to gain a first idea of the method introduced by [\[26\]](#page-113-4), consider first the case where  $F \equiv 0$  so that [6.26](#page-60-0) coincides with

<span id="page-60-1"></span>
$$
\left(\psi^{n,i}\right)' = -x_i^n \psi^{n,i}, \quad i = 1, ..., n,
$$
\n(6.28)

where its solution is denoted by

$$
\psi^{n,i}(t) = \psi^{n,i}(0)e^{-x_i^n t}, \quad i = 1, \dots, n.
$$

We could first consider an explicit Euler scheme for [6.28,](#page-60-1) that is

$$
\hat{\psi}_{t_{k+1}}^{n,i} = \hat{\psi}_{t_k}^{n,i} - x_i^n \Delta t \hat{\psi}_{t_k}^{n,i} = (1 - x_i^n \Delta t) \hat{\psi}_{t_k}^{n,i}, \quad i = 1, ..., n,
$$

for a regular time grid  $t_k = (kT)/N$  for all  $k = 1, ..., N$ , where *T* is the terminal time, *N* the number of time steps and  $\Delta t = T/N$ . A sufficient condition for the stability of the scheme reads

$$
\Delta t \le \min_{1 \le i \le n} \frac{1}{x_i^n}.
$$

The biggest pain point is that  $x_n^n$  grows very large as *n* increases. For instance, for  $n = 120$ ,  $r_{120} = 1.13$  and  $H = 0.1$ ,  $x_n^n = 6449.88$ . Hence, if we need to secure the stability of the explicit scheme, we would need a very large number of time steps *N*. In contrast, the implicit Euler scheme

$$
\hat{\psi}_{t_{k+1}}^{n,i} = \hat{\psi}_{t_k}^{n,i} - x_i^n \Delta t \hat{\psi}_{t_{k+1}}^{n,i}, \quad i = 1, \dots, n,
$$

is stable for any number of time steps *N*. The implicit scheme yields

$$
\hat{\psi}_{t_{k+1}}^{n,i} = \frac{1}{1 + x_i^n \Delta t} \hat{\psi}_{t_k}^{n,i}, \quad i = 1, ..., n.
$$
\n(6.29)

As a result, the authors of [\[26\]](#page-113-4) consider the following explicit-implicit discretization scheme of the *n* dimensional Riccati system of equations denoted by

$$
\hat{\psi}_0^{n,i} = 0, \quad \hat{\psi}_{t_{k+1}}^{n,i} = \frac{1}{1 + x_i^n \Delta t} \left( \hat{\psi}_{t_k}^{n,i} + \Delta t F\left(u, \sum_{j=1}^n c_j^n \hat{\psi}_{t_k}^{n,j}\right) \right), \quad i = 1, ..., n
$$

for a time grid  $t_k = k\Delta t$  for all  $k = 1, ..., N$ , with step size  $\Delta t = T/N$ , terminal time *T* and number of time steps *N*.

#### **6.3.2.** NUMERICAL RESULTS

We first do similar convergence tests for validating the use of the COS method with the Lifted Heston model. Thus, we choose  $H = 0.12$  and test convergence against the rough Heston price generated using the 1D COS method. For the first experiment, we chose a relatively high number of factors, i.e. *n* = 5000.



Figure 6.6: Convergence in the number of cosine basis functions with model parameter  $\rho = -0.6$ ,  $\bar{v} = 0.05$ ,  $v = 0.4$ ,  $\kappa = 2$  and  $v_0 = 0.0625$ . Moreover, we have  $S_0 = 100$ ,  $K = 100$  and  $T = 1.0$ 

Next, we test the convergence in the number of factors used in the lifted Heston model.

<span id="page-61-0"></span>

Figure 6.7: Convergence in the number factors *n* used in the Lifted Heston method with model parameter  $\rho = -0.6$ ,  $\bar{v} = 0.05$ ,  $v = 0.4$ ,  $\kappa = 2$  and  $v_0 = 0.0625$ . Moreover, we have  $S_0 = 10$ ,  $K = 10$  and  $T = 1.0$ 

We can observe in [6.7](#page-61-0) that a very low number of factors still results in a very good approximation. Going forward we use  $n = 120$ , which marks a good tradeoff between accuracy and computation time.

In the next step, we benchmark the results to option prices generated by the 1D COS method under the rough Heston model using the Adams scheme with  $N = 2500$ .

In the following table, we take a closer look at the error at different strike levels. We consider OTM as *x* = −0.2, ATM as *x* = 0.0, and ITM as *x* = 0.2.

As seen in [6.8,](#page-62-0) for options far out of the money the relative error seems quite high. However, these prices are anyway close to zero in a range of  $10^{-12} - 10^{-14}$ . Hence, for

<span id="page-62-0"></span>

Figure 6.8: Comparison option price using 1D COS method under the Lifted Heston vs. rough Heston with the Adams scheme.



Table 6.1: Relative error analysis over various strike prices.

these prices, we rather take into account the absolute error which seems more than accurate for out-of-the-money options.

Important to note is that going further the Lifted Heston model becomes our benchmark model due to the massive improvement in terms of computational time. Hence, we benchmark further results against the option prices obtained via the Lifted Heston model with parameter  $n = 120$ .

In the next numerical Experiment, we plot the implied volatility surface created using the Lifted Heston model and plot the error against the rough Heston implied volatility. The volatility surface consists of 50 different strike prices and 50 different maturities between 0.1 and 1.0.

Nevertheless, we note that the computation time for the whole volatility surface for the lifted Heston is 12.21 seconds, while for the rough Heston using only  $N = 2,500$  steps the computation time is already at 253.46 seconds, which can be seen in the following table



Table 6.2: CPU Time analysis

Implied Volatility Surface



Figure 6.9: Implied volatility surface created using the lifted Heston dynamics with parameters *ρ* = −0.4,  $\bar{v}$  = 0.0678,  $v$  = 0.3,  $\kappa$  = 2.6,  $v_0$  = 0.1 and *H* = 0.12

#### AbsError in Rough Heston vs Lifted



rough Heston dynamics with parameters  $ρ = -0.4$ ,  $ν = 0.0678$ ,  $ν = 0.3$ ,  $κ = 2.6$ ,  $ν_0 = 0.1$  and  $H = 0.12$ 

Last but not least, we see in  $6.11$ , that the lifted Heston model is creating the same typical rough volatility behavior of the ATM skew.

However, both models the lifted Heston as well as the rough Heston model find their bottleneck whenever the density of the underlying log-stock is heavily peaked as shown in [6.13.](#page-64-1) In order for the density to be sufficiently smooth, a high number of cosine basis functions is needed to overcome the Gibbs phenomena, which significantly slows down the pricing process of both models. The effect of the number of cosine basis functions is

<span id="page-64-0"></span>

Figure 6.11: Term structure of the at-the-money implied volatility skew created using the lifted Heston dynamics vs. rough Heston dynamics.

displayed in [6.13](#page-64-1) as well. Moreover, for the implicit-explicit Euler scheme for the lifted Heston a very small number of steps *T* /*N* are needed in order to compensate for the high number of cosine basis functions and to stay numerically stable. In the following table, we price an option with parameters given in [6.13](#page-64-1) and *Ncos* = 2000, in order to visualize the effect.

<span id="page-64-1"></span>

In comparison to the extreme, in  $6.15$ , we see a density that is less peaked and has way-fetter tails. We observe that already a small number of basis functions is sufficient to have a smooth density.

We note that the above-given examples are displayed for a maturity of 0.8 years.

<span id="page-65-0"></span>

Figure 6.14: Example of a wider density with fatter tails (Zoomed out)



Figure 6.15: Example of a wider density with fatter tails (Zoomed in

|                             | CPU Time (s) | Price |
|-----------------------------|--------------|-------|
| Lifted Heston( $N = 600$ )  | 1.109        | NaN   |
| Lifted Heston( $N = 9000$ ) | 5.63         | 0.873 |
| Rough Heston( $N = 600$ )   | 1.153        | NaN   |
| Rough Heston( $N = 2500$ )  | 135.6        | 0.880 |

Table 6.3: Computational time for the peaked density example [6.13](#page-64-1)

**7**

## **OUR APPROACH OF PRICING EUROPEAN OPTIONS UNDER THE ROUGH HESTON MODEL**

Efficient and fast option pricing is pivotal in the finance industry as it enables traders to make timely, informed decisions in the fast-paced derivatives market, where even fractional delays can significantly impact profitability and risk management. In this chapter, we introduce a novel method of pricing European options in the rough Heston setting. Our aim in this chapter is to effectively solve the bottleneck in computational time in pricing options or more explicitly the approximation of the characteristic function in the pricing process under the rough Heston model.

Introducing the COS-CPD method, which aims to approximate the characteristic function in a supervised machine learning process by reducing the computational complexity that arises from calculating the multi-dimensional Fourier coefficient tensor resulting from the global decomposition of the characteristic function. Hence, we can efficiently approximate the characteristic function to have a fast option pricing method for the rough Heston model.

In the first section of this chapter, we describe the initial problem and give an introduction to our global decomposition using a cosine expansion. Moreover, we give general model assumptions and explain how we make use of our COS-CPD method in the rough Heston option pricing environment. Hence, the next sections are devoted to the implementation of the COS-CPD method in the calculation of European options. It is demonstrated that the COS-CPD method can greatly reduce the computational time in the option pricing process.

#### **7.1.** GENERAL PROBLEM AND ASSUMPTIONS

$$
v(x, t_0) = \frac{1}{2}(b - a) \cdot e^{-r\Delta t} \cdot \sum_{k=0}^{N-1} F_k(x) \cdot V_k,
$$

As we have already seen the rough Heston model does not take into account any interest rate. This yields

$$
v(x, t_0, v_0) = \sum_{k=0}^{N-1} F_k(x, v_0) \cdot V_k.
$$

However, this model assumes fixed model parameters, i.e. in the rough Heston model the parameter  $(\rho, \nu, H, \bar{\nu}, \kappa)$ . So whenever we change the model parameters, we obviously change the characteristic function, which is captured in the *Fk*− coefficients. As a result, we have

$$
\nu(x, t_0, v_0, \theta) = \sum_{k=0}^{N-1} {}'F_k(x, v_0, \theta) \cdot V_k.
$$

We further assume that the initial  $v_0$ -value is now part of our parameter set  $\theta$ . Hence, from now on  $\theta$  consists of all model parameters as well as initial market conditions.

Before we start with the actual method, we recall that the  $F_k$ -coefficients are given by

$$
F_k(x,\theta) = \text{Re}\left\{\phi\left(\frac{k\pi}{b-a}; x,\theta\right) e^{-ik\pi \frac{a}{b-a}}\right\}.
$$
 (7.1)

Furthermore, by taking into account the Heston model- property for factoring out the *x*-values as described in  $[4]$ , we find that

$$
F_k(\mathbf{x}, \theta) = \text{Re}\left\{\phi\left(\frac{k\pi}{b-a}; \theta\right) e^{ik\pi \frac{\mathbf{x}-a}{b-a}}\right\}.
$$
 (7.2)

For solving our optimization problem, it is important that the  $F_k$ 's are drawn from the same function (characteristic function). As a result, we use cosine-series expansion for the characteristic function, which results in

$$
\varphi(u|\theta) = f_R(u) + i \cdot f_I(u) = \sum_{j=0}^{N_{\varphi}-1} A_j(\theta) \cos\left(j\pi \frac{u - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right) + i \cdot B_j(\theta) \sin\left(j\pi \frac{u - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right), \tag{7.3}
$$

where *I* and *R* represent the real part and imaginary part respectively. Recalling that

$$
\nu(\mathbf{x}; t, \theta) \approx \sum_{k=0}^{N-1} \text{Re}\left\{\varphi\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{\mathbf{x}-a}{b-a}\right)\right\} V_k \tag{7.4}
$$

and using Euler identity, i.e.

$$
e^{ix} = \cos(x) + i\sin(x),
$$

we find that

$$
v(\mathbf{x}; t, \theta) \approx \sum_{k=0}^{N-1} \text{Re}\left\{\varphi\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{\mathbf{x}-a}{b-a}\right)\right\} V_k
$$
  
= 
$$
\sum_{k=0}^{N-1} \text{Re}\left\{\sum_{j=0}^{N_{\varphi}-1} \left[A_j \cos\left(j\pi\frac{u_k-a_{\varphi}}{b_{\varphi}-a_{\varphi}}\right)+i \cdot B_j \sin\left(j\pi\frac{u_k-a_{\varphi}}{b_{\varphi}-a_{\varphi}}\right)\right]\right\}
$$
(7.5)  

$$
\cdot \left[\cos\left(k\pi\frac{\mathbf{x}-a}{b-a}\right)+i \cdot \sin\left(k\pi\frac{\mathbf{x}-a}{b-a}\right)\right] V_k.
$$

Switching sums and taking a real part leads to

$$
v(\mathbf{x}; t, \theta) \approx \sum_{j=0}^{N_{\varphi}-1} A_j \left[ \sum_{k=0}^{N-1} \cos\left(j\pi \frac{u_k - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right) \cos\left(k\pi \frac{\mathbf{x} - a}{b - a}\right) V_k \right] - \sum_{j=0}^{N_{\varphi}-1} B_j \left[ \sum_{k=0}^{N-1'} \sin\left(j\pi \frac{u_k - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right) \sin\left(k\pi \frac{\mathbf{x} - a}{b - a}\right) V_k \right].
$$
 (7.6)

Now for the sake of simplicity, we denote

$$
C_j(\mathbf{x}) = \left[ \sum_{k=0}^{N-1} \cos \left( j \pi \frac{u_k - a_\varphi}{b_\varphi - a_\varphi} \right) \cos \left( k \pi \frac{\mathbf{x} - a}{b - a} \right) V_k \right]
$$
(7.7)

$$
S_j(\mathbf{x}) = \left[ \sum_{k=0}^{N-1} \sin\left(j\pi \frac{u_k - a_\varphi}{b_\varphi - a_\varphi}\right) \sin\left(k\pi \frac{\mathbf{x} - a}{b - a}\right) V_k \right].
$$
 (7.8)

This yields

$$
\nu(\mathbf{x}; t, \theta) = \sum_{j=0}^{N_{\varphi}-1} \left( A_j(\theta) C_j(\mathbf{x}) - B_j(\theta) S_j(\mathbf{x}) \right). \tag{7.9}
$$

This linear system can be written as follows

<span id="page-68-0"></span>
$$
\boldsymbol{M} \left[ \begin{array}{c} \boldsymbol{A} \\ \boldsymbol{B} \end{array} \right] = \boldsymbol{\nu}_{\text{bench}} \left( \boldsymbol{x}; t, \theta \right), \tag{7.10}
$$

The elements of the respective matrices are defined as follows

$$
\begin{aligned} \left(\mathbf{M}_{A}\right)_{ij} &= C_{j}\left(x_{i}\right) = \sum_{k=0}^{N-1} \cos\left(j\pi \frac{u_{k} - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right) \cos\left(k\pi \frac{x_{i} - a}{b - a}\right) V_{k},\\ \left(\mathbf{M}_{B}\right)_{ij} &= -S_{j}\left(x_{i}\right) = -\sum_{k=0}^{N-1} \sin\left(j\pi \frac{u_{k} - a_{\varphi}}{b_{\varphi} - a_{\varphi}}\right) \sin\left(k\pi \frac{x_{i} - a}{b - a}\right) V_{k}. \end{aligned} \tag{7.11}
$$

#### **7.2.** SVD FOR SOLVING THE LINEAR SYSTEM

As we're trying to estimate the vectors  $\bm{A}$  and  $\bm{B}$  the matrix  $\bm{M}\in\mathbb{R}^{n\times 2N_\varphi}$  is split up in two parts, i.e.  $\boldsymbol{M} = [M_A \quad M_B]$ , where  $M_A \in \mathbb{R}^{n \times N_\varphi}$  and  $M_B \in \mathbb{R}^{n \times N_\varphi}$  correspond respectively to the vectors *A* and *B*. This procedure has greatly been researched at FFQuant [\[29\]](#page-113-7). We solve the linear equation denoted in [7.10](#page-68-0) by SVD representation, i.e. we try to find the least square solution of the following set of equations

$$
\left[\begin{array}{c} A_{LS} \\ B_{LS} \end{array}\right] = \underset{A, B \in \mathbb{R}^{N\varphi}}{\operatorname{argmin}} \left\| M \left[\begin{array}{c} A \\ B \end{array}\right] - \nu_{\text{bench}}\left(x; t, \theta\right) \right\|_2. \tag{7.12}
$$

Using the LS solution and plugging this into the Fourier expansion of the chf. gives back the full chf.

We recall that the price of the option is denoted by

<span id="page-69-1"></span>
$$
\nu(\mathbf{x}; t, \theta) = \sum_{j=0}^{N_{\phi}-1} \left( A_j(\theta) C_j(\mathbf{x}) - B_j(\theta) S_j(\mathbf{x}) \right). \tag{7.13}
$$

Subsequently, we can also check the fit of the LS solution directly on the option rice.

#### **7.3.** COS-CPD FOR PRICING OPTIONS

Assuming we fix  $j = 0, ..., N_\omega - 1$  we make use of a multidimensional cosine expansion per *j*. This approach has already been research at FFQuant by [\[30\]](#page-113-8) and [\[31\]](#page-113-9). This results in the following:

$$
A_j(\theta) = \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} F_{j,k_1,\dots,k_d} \prod_{i=1}^d \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right).
$$
 (7.14)

<span id="page-69-0"></span>
$$
B_j(\theta) = \sum_{k_1=0}^{K-1'} \cdots \sum_{k_d=0}^{K-1'} G_{j,k_1,\dots,k_d} \prod_{i=1}^d \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right).
$$
 (7.15)

As a result, the price of a European option under the rough Heston model can be written as

$$
\nu(\mathbf{x}; t, \theta) \approx \sum_{j=0}^{N_{\varphi}-1} \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} \left( F_{j,k_1,\dots,k_d}(\theta) \cdot C_j(\mathbf{x}) - G_{j,k_1,\dots,k_d}(\theta) \cdot S_j(\mathbf{x}) \right) \prod_{i=1}^d \cos \left( k_i \pi \frac{\theta_i - a_i}{b_i - a_i} \right)
$$
(7.16)

We further note that  $\sum_{k=0}^{N-1}$ ' means that the first element is taken by  $\frac{1}{2}$ . However, going further we try to estimate the matrices *F* and *G* and hence for the sake of simplicity, we assume the  $\frac{1}{2}$  is already taken into account. To avoid this exponential growth in the computational complexity for increasing d, the CPD technique can be used to approximate the Fourier coefficient tensor with a sum of rank-one components, i.e.

$$
F_j \approx [\mathbf{A}_{1,j}, \mathbf{A}_{2,j}, \dots, \mathbf{A}_{d,j}]_R = \sum_{r=1}^R a_r^{1,j} \circ a_r^{2,j} \circ \dots \circ a_r^{d,j}, \tag{7.17}
$$

$$
G_j \approx [\mathbf{B}_{1,j}, \mathbf{B}_{2,j}, \dots, \mathbf{B}_{d,j}]_R = \sum_{r=1}^R b_r^{1,j} \circ b_r^{2,j} \circ \dots \circ b_r^{d,j},
$$
(7.18)

where ∘ represents the outer product.

Consequently, the CPD approximation is obtained by computing the 2*d* factor matrices  $\{A_{i,j}\}_{i=1}^d$  and  $\{B_{i,j}\}_{i=1}^d$  of sizes  $K \times R$  each, as mentioned by [\[31\]](#page-113-9). Therefore, the computational complexity of computing the Fourier coefficients drops from  $\mathscr{O}\big(K^{2d}\big)$  to  $\mathcal{O}(2dKR)$ , which makes this approach more suitable.

Once the factor matrices are computed, the Fourier coefficient tensor can be written as

$$
F_j\left[k_1,\ldots,k_d\right] = \sum_{r=1}^R \prod_{i=1}^d a_r^{i,j}\left[k_i\right],
$$

and

$$
G_j\left[k_1,\ldots,k_d\right] = \sum_{r=1}^R \prod_{i=1}^d b_r^{i,j}\left[k_i\right].
$$

This ultimately leads to the following representation of the option price

$$
\nu(\mathbf{x};t,\theta) = \sum_{j=0}^{N_{\phi}-1} \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} \sum_{r=1}^{R} \prod_{i=1}^{d} \left( a_r^{i,j} \left[ k_i \right] \cdot C_j - b_r^{i,j} \left[ k_i \right] \cdot S_j \right) \cos \left( k_i \pi \frac{\theta_i - a_i}{b_i - a_i} \right). \tag{7.19}
$$

For the sake of brevity, we denote  $\mathbf{v}_i$  [ $k_i$ ] := cos $\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right)$  $\left(\frac{\theta_i - a_i}{b_i - a_i}\right)$ . Hence, we find that

$$
\nu(\mathbf{x}; t, \theta) = \sum_{j=0}^{N_{\varphi}-1} \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} \sum_{r=1}^{R} \prod_{i=1}^{d} \left( a_r^{i,j} \left[ k_i \right] \cdot C_j(\mathbf{x}) - b_r^{i,j} \left[ k_i \right] \cdot S_j(\mathbf{x}) \right) \mathbf{v}_i \left[ k_i \right]. \tag{7.20}
$$

By defining

<span id="page-70-0"></span>
$$
f_j(\theta) \approx \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} \sum_{r=1}^R \prod_{i=1}^d a_r^{i,j} [k_i] \mathbf{v}_i [k_i]
$$
 (7.21)

and

<span id="page-70-1"></span>
$$
g_j(\theta) \approx \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} \sum_{r=1}^R \prod_{i=1}^d b_r^{i,j} [k_i] \mathbf{v}_i [k_i]
$$
(7.22)

we find that the price of the option can be written as

$$
\nu(\mathbf{x};t,\theta) = \sum_{j=0}^{N_{\varphi}-1} f_j(\theta) \cdot C_j(\mathbf{x}) - g_j(\theta) \cdot S_j(\mathbf{x})
$$

By setting

$$
f_{n,r}^j = \sum_{k_n=0}^{K-1} a_r^n [k_n] \mathbf{v}_n [k_n] = \mathbf{v}_n^T \mathbf{A}_n[:, r]
$$

and

$$
g_{n,r}^j = \sum_{k_n=0}^{K-1} a_r^n [k_n] \mathbf{v}_n [k_n] = \mathbf{v}_n^T \mathbf{B}_n[:, r],
$$

expression [7.21](#page-70-0) and [7.22](#page-70-1) can greatly be simplified to

$$
f_j(\theta) \approx \sum_{r=1}^{R} f_{1,r}^j(\theta_1) \cdots f_{d,r}^j(\theta_d)
$$
 (7.23)

and

$$
g_j(\theta) \approx \sum_{r=1}^R g_{1,r}^j(\theta_1) \cdots g_{d,r}^j(\theta_d).
$$
 (7.24)

Each component  $f_{n,r}^j$  represents a univariate Fourier-cosine series. Consequently, the Fourier-cosine series derived through Canonical Polyadic Decomposition (CPD) can be articulated as a product of these univariate series.

By stating that for each *j* the output  $y_i$  given an input vector  $(\theta_1, \ldots, \theta_d)$  yields

$$
y_j^A = (\mathbf{v}_1^T \mathbf{A}_{1,j} \otimes \cdots \otimes \mathbf{v}_d^T \mathbf{A}_{d,j}) \mathbf{1} = \left( \otimes_{i=1}^d \mathbf{v}_i^T \mathbf{A}_{i,j} \right) \mathbf{1}.
$$
 (7.25)

$$
\mathbf{y}_{j}^{B} = (\mathbf{v}_{1}^{T} \mathbf{B}_{1,j} \otimes \cdots \otimes \mathbf{v}_{d}^{T} \mathbf{B}_{d,j}) \mathbf{1} = \left( \otimes_{i=1}^{d} \mathbf{v}_{i}^{T} \mathbf{B}_{i,j} \right) \mathbf{1},
$$
\n(7.26)

where ⊛ denotes the Hadamad product as defined in [3.4.4](#page-35-0) and **1** the unit vector of size R. Thus, we find that the price of the option is denoted by

$$
\nu(\mathbf{x};t,\theta) = \sum_{j=0}^{N_{\varphi}-1} \left( \circledast_{i=1}^d \mathbf{v}_i^T \mathbf{A}_{i,j} \right) \mathbf{1} \cdot C_j - \left( \circledast_{i=1}^d \mathbf{v}_i^T \mathbf{B}_{i,j} \right) \mathbf{1} \cdot S_j. \tag{7.27}
$$

#### **7.4.** OUR COSINE NETWORK

To enhance the comprehension of the mathematical concepts presented in the preceding section, we aim to provide further clarity through graphical illustrations. These visualizations not only aid in understanding but also highlight the advantages of this approach compared to traditional machine-learning techniques.

We begin by examining the multidimensional cosine expansion for each *j*, as outlined in [7.15,](#page-69-0) which originates from [7.13.](#page-69-1) It's important to note that while our discussion focuses on *Aj*(*θ*), the process for *Bj*(*θ*) is entirely analogous.

$$
A_j(\theta) = \sum_{k_1=0}^{K-1'} \cdots \sum_{k_d=0}^{K-1'} F_{j,k_1,\dots,k_d} \prod_{i=1}^d \cos \left( k_i \pi \frac{\theta_i - a_i}{b_i - a_i} \right).
$$

As mentioned above, the goal is to approximate the Fourier coefficient tensor with a sum of rank-one tensors, i.e,

$$
F_j \approx [\mathbf{A}_{1,j}, \mathbf{A}_{2,j}, \dots, \mathbf{A}_{d,j}]_R = \sum_{r=1}^R a_r^{1,j} \circ a_r^{2,j} \circ \dots \circ a_r^{d,j}.
$$

Each rank one component consists of a product of three univariate Fourier-cosine series. As a result, our supervised machine learning method boils down to understanding
the above-given figure. Even though the mathematics behind it might not be straightforward to understand, the intuition however is quite clear. We approximate the multidimensional Fourier-cosine series by a sum of rank-one components with each component given by a product of a univariate Fourier-cosine series, which can be displayed using the Hadamard product as seen in [7.25](#page-71-0) and [7.26.](#page-71-1) We recap that we thus can write the price of an option given input parameter vector as

<span id="page-72-0"></span>
$$
\nu(\mathbf{x};t,\theta) = \sum_{j=0}^{N_{\phi}-1} \left( \circledast_{i=1}^{d} \mathbf{v}_{i}^{T} \mathbf{A}_{i,j} \right) \mathbf{1} \cdot C_{j} - \left( \circledast_{i=1}^{d} \mathbf{v}_{i}^{T} \mathbf{B}_{i,j} \right) \mathbf{1} \cdot S_{j}.
$$
 (7.28)

However, we can write of the European option in terms of our cosine network as follows. Before we can actually introduce our cosine network, we have to introduce some special notation according to [\[32\]](#page-114-0).



Figure 7.1: Notation for our cosine network

In order to understand how our notation exactly works, we break it down step by step.

- (1) vector-vector multiplication. This is just the inner product between two vectors of the same size. The result is a scalar
- (2) vector-matrix multiplication. Having a matrix of  $K \times R$  and a vector of size  $R \times 1$ results in a vector of  $K \times 1$
- (3) vector-tensor multiplication. Having a tensor of  $N \times K \times R$  and a vector of size  $R \times 1$ results in a matrix of size of *N* ×*K*
- (4) Tensor copy. Copying the result of the previous computation for the next step. I.e. if the previous calculation was vector-matrix multiplication, then the resulting vector is copied.
- (5) Element-wise multiplication. Refers to the Hadamard product explained in [3.4.4.](#page-35-0)

In the following, we will methodically break down and explain each step to provide a clear understanding of the process with respect to [7.28.](#page-72-0) We emphasize that our cosine network is displayed for  $d = 4$ . However, the methodology translated to an arbitrary number of *d*.



Figure 7.2: Our novel cosine network

- (1) In the first step, we calculate  $\mathbf{v}_1^T \mathbf{A}_1$ , which corresponds to the first vector-tensor multiplication in the upper row. We point out that  $A_1$  is now of form  $K \times N_\varphi \times R$ , while  $\mathbf{v}_1^T$  is still of form  $1 \times K$ . The resulting matrix is of the form  $N_\varphi \times R$ . This matrix will be element-wise-multiplied using the Hadamard product with the resulting matrix of the  $\mathbf{v}_2^T \mathbf{A}_2$  vector-tensor multiplication. We continue with this process until we reach  $\overline{d} = 4$ . We stress that the same procedure applies to  $\mathbf{v}_i^T \mathbf{B}_i$  in the lower row.
- (2) In the next step we copy our resulting  $N_{\phi} \times R$ -matrix from the Hadamard multiplication (black dot).
- (3) Next, we multiply this matrix with the unit vector of size  $R \times 1$ . This corresponds to the sum of rank-1 components in [7.23](#page-71-2) and [7.24.](#page-71-3) Again, we note that we do the same computation for both the upper and lower rows in the figure. The resulting vector of each computation is  $N_{\varphi} \times 1$ .
- (4) Going further, we multiply our resulting vectors with the matrix *C* and *S* defined in [7.7](#page-68-0) and [7.8,](#page-68-1) respectively.
- (5) Lastly, We subtract both results, in order to get our vector of option prices with respect to various strike prices.

In general, we emphasize that by using our cosine network we can loosen up all parameters in the respective stochastic volatility model, i.e. the rough Heston model up to 6 as well as the strike dimension which is captured in the matrices *C* and *S* respectively. In summary, our novel cosine network presents a tractable and supervised machine learning approach that efficiently approximates the necessary cosine coefficients, enabling us to derive the desired characteristic function with superior speed in comparison to existing benchmark methods. This innovative technique streamlines the process, ensuring both accuracy and efficiency in obtaining the characteristic function for complex models.

Our cosine network distinguishes itself from conventional artificial neural networks (ANNs), such as those discussed in [\[33\]](#page-114-1), by maintaining tractability throughout its operational process. Unlike standard ANNs, which often operate as "black boxes" leaving practitioners and researchers puzzled about the rationale behind specific decisions, our cosine network offers a transparent and comprehensible decision-making pathway. This transparency is crucial, as the opacity of ANNs can erode trust and hinder their practical application in real-world scenarios, a concern highlighted by [\[34\]](#page-114-2). Our cosine network's clear, tractable nature thus represents a significant advancement, enhancing the interpretability and reliability of neural network decisions in complex tasks.

# **7.5.** TRAINING DATA

<span id="page-74-0"></span>In order to find the factor matrices training data is needed. First, we fix the index *j* and only focus on the *A* part, since the other part will be analogous. Secondly, let us consider the grid of *M* combinations of *d* model parameters *θ*. In order to be able to loosen up all 6 parameters in the rough Heston model, which in the end will result in the calculation *M* = 1.000.000 options pricing per maturity, we want to avoid densities that are too peaked, as shown in figure [6.13.](#page-64-0) As a result, our parameters are restricted to the following intervals. The resulting combinations will be our training input.

| <b>Parameters</b>                  | Range         |
|------------------------------------|---------------|
| Correlation, $\rho$                | $(-0.9, 0.9)$ |
| Reversion speed, $\kappa$          | (1.5,5)       |
| Long average variance, $\bar{\nu}$ | (0.04, 0.2)   |
| Initial variance, $v_0$            | (0.04, 0.15)  |
| Hurst parameter, H                 | (0.05, 0.15)  |
| Volatility of volatility, $\nu$    | (0.1, 0.4)    |

Table 7.1: Parameter Ranges for the Model

Before we further elaborate, we note that the upcoming explanation is done for  $A_i(\theta_i)$ . However,  $B_i(\theta_i)$  follows analogously.

For  $i = 1,..., M$  the training output will be  $A_j(\theta_i)$ . These values are derived using SVD as explained in the previous sections. With this a set of log-asset prices **x** is needed to calculate a set of benchmark values per  $\theta_i$ . This means that first the whole vector  $A(\theta_i) = \left[A_0(\theta_i),...,A_{N_\varphi-1}(\theta_i)\right]^T \in \mathbb{R}^{N_\varphi}$  is computed using minimization problem [7.12.](#page-69-0) To get the correct output, we split them apart afterward such that for index *j* we get the output vector  $A_j(\bm{\theta}) = \left[A_j(\theta_1),...,A_j(\theta_M)\right]^T \in \mathbb{R}^M$ . We denote the estimated outputs of  $A_j(\boldsymbol{\theta})$  and  $B_j(\boldsymbol{\theta})$  by  $\mathbf{y}_j^j$  $\frac{j}{A}$  and  $\mathbf{y}_I^j$  $B$ <sup>,</sup> In total both

# **7.6.** ALS ALGORITHM

Using the CPD technique yields to solving the following minimization problem.

$$
\min_{\left[{\bf A}_{i,j}\right]_{i=1}^d} \frac{1}{M} \sum_{m=1}^M \left( y_{A,m}^j - \left(\circledast_{i=1}^d {\bf v}_i^T {\bf A}_{i,j}\right) {\bf 1} \right)^2
$$

and

$$
\min_{\left|\mathbf{B}_{i,j}\right|_{i=1}^d} \frac{1}{M} \sum_{m=1}^M \left(y_{B,m}^j - \left(\circledast_{i=1}^d \mathbf{v}_i^T \mathbf{B}_{i,j}\right) \mathbf{1}\right)^2,
$$

where we use the squared error as the loss function

where  $\mathbf{V}_i \in \mathbb{R}^{K \times M}$  with  $\mathbf{V}_i[k,m] = \cos\left(k\pi \frac{\theta_i^m - a_i}{b_i - a_i}\right)$  $\frac{\partial_i^{m} - a_i}{\partial_i - a_i}$ . The value  $\theta_i^m$  denotes the value of the *i*−th coordinate in sampling vector *θm*. The Alternate Least Square (ALS) algorithm can now be exploited to solve the factor matrices iteratively as derived in [\[21\]](#page-113-0) and [\[31\]](#page-113-1). Fixing all factor matrices except for **A***i*,*<sup>j</sup>* (the case **B***i*,*<sup>j</sup>* follows analogously) we have,

$$
\min_{\mathbf{A}_{i,j}} \frac{1}{M} \sum_{m=1}^M \left( y_{A,m}^j - \mathbf{V}_i[:,m]^T \mathbf{A}_{i,j} \mathbf{Q}_{i,j}^A[:,m] \right)^2,
$$

where  $\mathbf{Q}_{i,j}^{A}$  with

$$
\mathbf{Q}_{i,j}^A = \left(\circledast_{n \neq i} \left(\mathbf{A}_{n,j}^T \mathbf{V}_n\right)\right)
$$

By setting the gradient equal to zero, we find that

$$
\frac{2}{M} \sum_{m=1}^{M} \left( y_{A,m}^{j} - \mathbf{V}_{i}[:,m]^{T} \mathbf{A}_{i,j} \mathbf{Q}_{i,j}^{A}[:,m] \right) = 0.
$$
 (7.29)

Rearranging terms leads to the following

<span id="page-75-0"></span>
$$
\frac{1}{M} \sum_{m=1}^{M} \left( \mathbf{V}_{i}[:,m]^{T} \mathbf{A}_{i,j} \mathbf{Q}_{i,j}^{A}[:,m] \right) \left( \mathbf{V}_{i}[:,m] (\mathbf{Q}_{i,j}^{A})^{T}[:,m] \right) = \frac{1}{M} \sum_{m=1}^{M} y_{A,m}^{j} \left( \mathbf{V}_{i}[:,m] (\mathbf{Q}_{i,j}^{A})^{T}[:,m] \right).
$$
\n(7.30)

Note that  $7.30$  is not a standard linear system of the form  $Ax = b$ , which has to be solved for the unknown vector *x*. Instead, the system can be seen as a matrix equation, in which the unknown matrix  $A_n$  has to be solved. A previous thesis done at FF Quant showed how to transform  $7.30$  into a equation of the form  $Ax = b$  by vectorizing the factor matrix  $A_n$  [\[31\]](#page-113-1). Furthermore, it was shown that the Conjugate Gradient (CG) method is the fastest method in terms of convergence to solve [7.30.](#page-75-0)

# **7.7.** TRUNCATION RANGE AND FOURIER EXPANSION TERMS

In order to have a good convergence in the CPD method, the choice of the truncation range and the number of cosine basis functions ( $N_\omega$  and  $N_\text{cos}$ ) in the COS method is therefore of high importance.

Even though it is yet to develop a formula for the truncation range, we try to give an idea of how to carefully choose a proper truncation range in order to avoid reverse engineering. However, knowing that testing truncation intervals with different numbers of *N<sup>ϕ</sup>* and *Ncos* in order to check convergence is computationally not feasible for all parameters. Assume we have  $d = 6$  with  $K = 10$ . Hence, for every different combination of  $N_\omega$ and *Ncos* we have to go through the whole COS-CPD calculation and generate a training set of 1 million option pricings. Hence, we fix three parameters and use COS-CPD on 3/6 parameters with  $K = 10$  in the rough Heston model, which then boils down to only have 1000 option pricing processes. Once we were satisfied with the convergence, we tested it on the value extremes of the three fixed parameters and their combinations.

Afterward, we loosen up one more parameter and check its fit and convergence of the CPD method of that specific combination in 5D. In the end, this was done for all 5 different maturities on which we later based on our implied volatility surface. Nonetheless, with the explanations below choosing a suitable truncation range is less of a guessing game but rather a matter of a careful choice.

Before we analyze the different truncation ranges, we recall that in [7.3,](#page-67-0) we have to deal with two different Truncation ranges. First, the one with respect to the series expansion of the characteristic function. We truncate the series of the characteristic function with respect to the interval  $\left[0, \frac{N_{cos} \pi}{(b-a)}\right]$ . Hence, we find that  $\left[a_{\varphi}=0, b_{\varphi}=\frac{N_{cos} \pi}{(b-a)}\right]$ . The other truncation range we deal with is the one of the 1D COS-method as described in [\[4\]](#page-112-0). However, these two truncation ranges behave opposite of each other as choosing [*a*,*b*] large results in a tighter interval of  $[a_\varphi,b_\varphi],$  while choosing [*a, b*] small results in a bigger interval  $[a_\varphi,b_\varphi].$  Hence, it is important to find a good trade-off between both of them.

<span id="page-76-0"></span>

Figure 7.3: Absolute and relative error for an interval of  $[-1,1]$  and  $N<sub>\phi</sub> = 32$  with Rank  $= 15$ 



Figure 7.4: Absolute and relative error for an interval of  $[-1,1]$  and  $N<sub>\phi</sub> = 64$  with Rank  $= 15$ 

As we can see in figure [7.4,](#page-76-0) for a very tight interval [*a*,*b*] we need a larger number of  $N_{phi}$  in order to have a less oscillating error, which is accordance to the intuition that a wider interval needs more cosine basis functions in order to have a good approximation. However, even though the error seems to improve we observe that when plotting the density there is not much of an improvement of the overall density function even though we further increased  $N_{\text{phi}}$ . Even though the initial error in [7.4](#page-76-0) seems quite promising using this density in the option pricing method would indeed result in significant errors. One could argue that choosing a sufficiently wide intervals [ $a$ ,  $b$ ] and  $\left[ a_{\varphi},b_{\varphi}\right]$  and a high number of cosine basis functions in both expansions would solve the problem. However, first and foremost a very high number of  $N_\varphi$  would increase the computational time in



our COS-CPD method by a multiple. Remember we have to find  $2N_{\varphi}d$  Factor matrices.

Figure 7.5:  $[-1, 1]$  and  $N_{\varphi} = 64$ . Bad density due to truncation error

<span id="page-77-0"></span> $10$ 

 $10^{\circ}$ 

 $10<sup>6</sup>$ 

 $10^{-}$ 

 $-1.00$   $-0.75$   $-0.50$   $-0.25$ 

Error

RelError and AbsError CPD  $r=1$ , a = 3, N\_phi = 64

AbsFrron RelErro

 $0.75$ 

1.00





Density CPD

wwwwww

 $0.50$  $0.75$ 

Figure 7.7: Absolute and relative error for an interval of  $[-3,3]$  and  $N<sub>ω</sub> = 32$  with Rank = 1

 $0.00$ 

 $0.25$  $0.50$ 

Figure 7.8: Absolute and relative error for an interval of  $[-3,3]$  and  $N_{\varphi} = 32$  with Rank  $= 15$ 

As we can see in [7.8,](#page-77-0) as mentioned before very wide interval [*a*,*b*] results in a very tight interval  $[a_\varphi,b_\varphi]$ . As a result, by looking at the non-improvement of the rank that there is some kind of dominant truncation error most probably resulting from the very tight interval  $[a_{\varphi}, b_{\varphi}]$ .

In comparison to the above two approaches, in [7.10,](#page-78-0) we see that when having a very good trade-off between both truncation ranges the COS-CPD method works really well. We can also see that with an increasing number of *Nphi* the approximation of the density is becoming really smooth as seen in [7.12.](#page-78-1)

As a result, we go with the following set of parameters for the COS-CPD method.



We emphasize that the table mentioned is suitable only for the parameter set defined in [7.1.](#page-74-0) Generally speaking, choosing a suitable [*a*,*b*] might be a challenging task, but it is

<span id="page-78-0"></span>

Figure 7.9: Absolute and relative error for an interval of  $[-2.5, 2.5]$  and  $N_\varphi = 32$  with  $Rank = 1$ 

<span id="page-78-1"></span>

Figure 7.11: [−2.25,2.25] and *N<sup>ϕ</sup>* = 16. Bad density due to very little *Nϕ*



Figure 7.10: Absolute and relative error for an interval of  $[-2.5, 2.5]$  and  $N_\varphi = 32$  with  $Rank = 15$ 



Figure 7.12: [−2.25,2.25] and *N<sup>ϕ</sup>* = 32. Nice density recovered from COS-CPD chf.

important to note that a lot depends on the parameter combinations and the maturities the COS-CPD method is given as input.



Figure 7.13: Absolute and relative error for a different number of cosine basis functions in each parameter dimension. The rank is set to 15. Here, we consider an example for  $\rho = 0.4$ ,  $\bar{\nu} = 0.060$  and  $H = 0.08$ 



Figure 7.14: Absolute and relative error for a different number of cosine basis functions in each parameter dimension. The rank is set to 15. Here, we consider an example for  $\rho = -0.7$ ,  $\bar{\nu} = 0.078$  and  $H = 0.12$ 

# **7.8.** NUMBER OF COSINE BASIS FUNCTIONS IN THE PARAME-TER DIMENSIONS

For choosing a suitable number of cosine basis functions (*K*) for the different parameter dimensions, we first fix the rank at  $r = 15$ . Throughout this analysis we consider ATM options, i.e. keep  $x = 0$  fixed, and plot the different errors against the number of cosine basis functions. We consider ATM options because the error seems to be the most sensitive for ATM options. As a result, we see if the error improves with a finer grid of training points and to what degree.

We tested the outcome for various different sets of parameters in 3D, which means we keep  $v_0 = 0.0625$ ,  $v = 0.2$  and  $\kappa = 2.5$  fixed and "loosen up" the rest of the parameters.

It is important to note that a higher number of N does not automatically mean a better error as seen above. However, the generation of the training set grows exponentially with *K*. *K* is the number of cosine basis function in in each dimension. As we aim to lose up all of the 6 parameters in the rough Heston model,  $K = 10$  already yields a training set of 1,000,000 option prices, while  $K = 14$  would yield a training set of over 7.5 million option prices.

# **7.9.** OPTIMAL CHOICE OF RANK

Throughout this section, we fix  $K = 10$  as the number of cosine basis functions in each parameter dimension and try to find an optimal rank. As mentioned in [\[35\]](#page-114-3) and [\[31\]](#page-113-1), finding the rank is a NP-hard problem. However, we find an "optimal" rank as follows. Again, we fix 3 parameters as we think this is sufficient to prove the point we make. Having fixed three parameters, we apply COS-CPD on the remaining 3D tensor and calculate option prices with various different strikes. We then take the mean absolute error of all the options against our benchmark method. We note that it is suitable to choose strikes that are not too far from each other as for deep OTM options the absolute error is around 10−<sup>6</sup> , which might lead to misleading mean errors. We recommend choosing strikes around ATM. Next, we repeat this procedure for all sorts of ranks and plot the resulting error against the rank to see which rank suits best. The results can be seen in

<span id="page-80-0"></span>

Figure 7.15: Absolute and relative error for a different number for the Rank. Here, we consider an example for  $\rho = -0.2$ ,  $\bar{v} = 0.060$ and  $H = 0.08$ 



Figure 7.16: Absolute and relative error for a different number for the Rank. Here, we consider an example for  $\rho = -0.7$ ,  $\bar{v} = 0.078$ and  $H = 0.12$ 

# [7.15](#page-80-0)

Even though it might look like a very small number of rank might be sufficient, we experienced that for higher dimensions, we have to choose a higher number for the rank in order for the COS-CPD method to sufficiently converge. As a result, we choose rank = 10 going forward.

# **7.10.** NUMERICAL RESULTS

# **7.10.1.** OPTION PRICING

For the first numerical experiment, we chose two different sets of parameters from [7.1,](#page-74-0) i.e.

(1)  $\rho = -0.6$ ,  $\kappa = 2.0$ ,  $\bar{\nu} = 0.055$ ,  $\nu_0 = 0.0625$ ,  $H = 0.08$ ,  $\nu = 0.15$ 

(2) 
$$
\rho = -0.26
$$
,  $\kappa = 1.8$ ,  $\bar{\nu} = 0.95$ ,  $\nu_0 = 0.0756$ ,  $H = 0.12$ ,  $\nu = 0.36$ 

for *T* = 0.75. We see in [7.17](#page-81-0) that for both sets of parameters, the price of the COS-CPD is indistinguishable from the Lifted Heston as well as from the Rough Heston price.

For the first set of parameters the COS-CPD error against the rough Heston and the lifted Heston can be seen in [7.19.](#page-81-1) We can see that the error is better against the lifted Heston, which is expected as that method was used for benchmarking. However, the COS-CPD pricing performs sufficiently accurately against the rough Heston as well.

Furthermore, we have a closer look at the performance of one single option pricing process and the relative error analysis through different strike levels. Again, we consider OTM as  $x = -0.2$ , ATM as  $x = 0.0$ , and ITM as  $x = 0.2$ . Important to note that in the following analysis, we include the rough Heston price using  $N = 2,500$  time steps just for the sake of performance comparison. As previously mentioned the lifted Heston was used as a benchmark.

We conclude that our method is roughly 16 times faster than the lifted Heston and about 180 times faster than the benchmark most often being used in the literature while having very promising accuracy through various strike levels.

Lastly, we price 729 options with different parameter values and maturities and see the accuracy is consistent with the observation we made in  $7.17$  as well as  $7.19$ . The max-

<span id="page-81-0"></span>



<span id="page-81-1"></span>





Figure 7.19: Absolute and relative error COS-CPS against lifted Heston



Table 7.2: Relative error analysis over various strike prices for Testset (1). The 'BM' denotes the benchmark being used.

imum absolute error throughout the 729 option pricings is 0.00945. In [7.20,](#page-82-0) we observe that the absolute error is consistently in the range of  $10^{-3}$  throughout the whole set of strikes (displayed as log-moneyness x).

<span id="page-82-0"></span>

Figure 7.20: Mean Average and Mean Relative Error

# **7.10.2.** IMPLIED VOLATILITY SURFACE

In order to get a reasonable implied volatility surface the training step has been done for 5 different maturities, i.e. 0.1,0.25,0.5,0.75 and 1.0. Afterward, the whole surface was interpolated using cubic interpolation in order to get a sufficiently smooth surface with respect to log-moneyness and maturity [\[36\]](#page-114-4).











AbsError in rough Heston CPD vs lifted Heston

Figure 7.23: Absolute Error whole implied volatility surface Interpolated COS-CPD surface vs. lifted Heston implied volatility.

We clearly see that even though we interpolate between the above-mentioned maturities the absolute error is indeed small. We can see the spike in the short-term deepin-the-money options. This is due to some Gibbs phenomena when approximating the density using COS-CPD. Especially for short-term maturities, we find this problem as we only chose 10 cosine basis functions in each of the parameter dimensions.

Next, we compare the methods with respect to performance.





Last but not least, we check the ATM skew of the COS-CPD method in comparison to our benchmark.



Figure 7.24: Term structure of the at-the-money implied volatility skew created using our COS-CPD approach vs. lifted Heston dynamics.

We highlighted the points that are actually computed by the COS-CPD method, namely maturities 0.1,0.25,0.50,0.75 and 1.0. As a result the error of the ATM skew mainly results from the interpolation procedure. We further highlight that the whole implied volatility surface can be calculated in about half a second around 8 times faster than the interpolated Lifted Heston surface.

# **8**

# **OUR CALIBRATION APPROACH**

One of the most important tasks in quantitative finance is the calibration of the model to actual market data, which is equal to finding the optimal parameters such that the model output is as close as possible to the actual market data. In this chapter, we introduce a novel method of calibrating the rough Heston model to market data using the COS-CPD method using the first derivative of the option price with respect to each of the considered parameters in the rough Heston model. Our cosine network enables us to derive the derivative with respect to each parameter analytically which in combination with the very efficient pricing method from [7,](#page-66-0) results in a calibration within a minute while current benchmark methods take hours in calibration time.

# **8.1.** OUR COS-CPD CALIBRATION APPROACH

# **8.1.1.** ROUGH HESTON MODEL DERIVATIVES

The COS-CPD environment enables us to compute the derivatives of the option price with respect to each of the parameters in the rough Heston model. We recall that the European option price in the COS-CPD environment can be denoted by

$$
\nu(\mathbf{x}; t, \theta) = \sum_{j=0}^{N_{\varphi}-1} \left( A_j(\theta) C_j(\mathbf{x}) - B_j(\theta) S_j(\mathbf{x}) \right)
$$
(8.1)

with

$$
A_j(\theta) = \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} F_{j,k_1,\dots,k_d} \prod_{i=1}^d \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right).
$$
 (8.2)

$$
B_j(\theta) = \sum_{k_1=0}^{K-1} \cdots \sum_{k_d=0}^{K-1} G_{j,k_1,\dots,k_d} \prod_{i=1}^d \cos \left( k_i \pi \frac{\theta_i - a_i}{b_i - a_i} \right).
$$
 (8.3)

Let us first denote *ζ* as a placeholder for any of the rough Heston parameters

$$
\theta = \rho, \bar{\nu}, \nu, \nu_0, H, \kappa.
$$

Then, we find that

$$
\frac{d\,\nu(\mathbf{x};t,\theta)}{d\zeta} = \sum_{j=0}^{N_{\varphi}-1} \frac{d}{d\zeta} \left( A_j(\theta) \right) C_j(\mathbf{x}) - \frac{d}{d\zeta} \left( B_j(\theta) \right) S_j(\mathbf{x}). \tag{8.4}
$$

We note that  $C_i(x)$  and  $S_i(x)$  are both independent of any model parameters. Hence, we only have to take care of  $\frac{d}{d\zeta}A_j(\theta)$  and  $\frac{d}{d\zeta}B_j(\theta)$ . The following calculations are done for  $\frac{d}{d\zeta}A_j(\theta)$ . However,  $\frac{d}{d\zeta}B_j(\theta)$  follows analogously.

$$
\frac{d}{d\zeta}A_j(\theta) = \frac{d}{d\zeta}\sum_{k_1=0}^{K-1'}\cdots\sum_{k_d=0}^{K-1'}F_{j,k_1,\dots,k_d}\prod_{i=1}^d\cos\left(k_i\pi\frac{\theta_i-a_i}{b_i-a_i}\right)
$$
\n
$$
= \sum_{k_1=0}^{K-1'}\cdots\sum_{k_d=0}^{K-1'}\frac{d}{d\zeta}\left(F_{j,k_1,\dots,k_d}\prod_{i=1}^d\cos\left(k_i\pi\frac{\theta_i-a_i}{b_i-a_i}\right)\right)
$$
\n(8.5)

Realizing that  $F_{j,k_1,\dots,k_d}$  is the coefficient matrix, which is independent of any parameter leaves us with the derivative of  $\prod_{i=1}^d \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right)$  $\frac{\theta_i - a_i}{b_i - a_i}$  with respect to *ζ*. We recall that *ζ* is a placeholder for any of the rough Heston parameters in *θ*. As a result, taking the derivative with respect to *ζ* yields

$$
\frac{d}{d\zeta}A_j(\theta) = -\sum_{k_1=0}^{K-1'} \cdots \sum_{k_d=0}^{K-1'} F_{j,k_1,\dots,k_d} \sin\left(k_\zeta \pi \frac{\zeta - a_\zeta}{b_\zeta - a_\zeta}\right) \prod_{i=1}^{d-1} \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right) \tag{8.6}
$$

To make things clearer, we set  $\zeta = H$ . Then the derivative of the option price with respect to *H* is given by

$$
\frac{d\,v(\mathbf{x};t,\theta)}{dH} = \sum_{j=0}^{N_{\varphi}-1} \frac{d}{dH} \left( A_j(\theta) \right) C_j(\mathbf{x}) - \frac{d}{dH} \left( B_j(\theta) \right) S_j(\mathbf{x}),\tag{8.7}
$$

with

$$
\frac{d}{dH}A_j(\theta) = -\sum_{k_1=0}^{K-1'} \cdots \sum_{k_d=0}^{K-1'} F_{j,k_1,\dots,k_d} \sin\left(k_H \pi \frac{H-a_H}{b_H-a_H}\right) \prod_{i=1}^{d-1} \cos\left(k_i \pi \frac{\theta_i - a_i}{b_i - a_i}\right) \tag{8.8}
$$

and consisting of all other parameters but *H*.

In order to check if the derived gradients are correct, we make use of Taylor expansion, with respect to each of the parameters and check if the Taylor expansion converges to the actual price. For this purpose, we take derivatives of the option price with respect to  $ρ$ ,  $H$ ,  $ν$  respectively. This is enough to validate our derivatives. Let's consider the following parameter set

$$
\theta = (H = 0.12, \rho = -0.6, \nu = 0.27),
$$

and

$$
\tilde{\theta} = (\tilde{H} = 0.13, \tilde{\rho} = -0.55, \tilde{v} = 0.25).
$$

Next, we consider the following Taylor expansion

<span id="page-88-0"></span>
$$
v(\mathbf{x}; t, \theta) = v(\mathbf{x}; t, \tilde{\theta}) + \frac{dv(\mathbf{x}; t, \theta)}{dH} \cdot (H - \tilde{H}) + \frac{dv(\mathbf{x}; t, \theta)}{d\rho} \cdot (\rho - \tilde{\rho}) + \frac{dv(\mathbf{x}; t, \theta)}{dv} \cdot (v - \tilde{v}). \tag{8.9}
$$



Figure 8.1: 1st order Taylor expansion with respect to Parameters *H*,*ρ*,*ν*.

We can easily observe in [8.1](#page-88-0) that the Taylor expansion significantly improves the error against the initial option price with parameter set *θ*. "AbsError no Taylor" is the absolute error of  $v(x; t, \theta)$  and  $v(x; t, \theta)$ . and As a result, we can be confident that the derivatives are correctly derived.

In order to use the gradients, we tried different Python built-in optimizers. Whenever a good initial guess is at hand for the parameters then *L* − *BFGS* − *B*, a quasi-Newton method suits best as it converges really fast [\[37\]](#page-114-5). As the algorithm together with our pricing method converges quite fast, one can allow a couple of calibration procedures with different initial guesses. Nevertheless, another method of choice can be the conjugate gradient (CG) method. Nonetheless, *L* −*BFGS* −*B* seems to converge a bit faster once a good initial guess is at hand.

# L-BFGS-B - ALGORITHM

**Objective Function**: Accoring to [\[37\]](#page-114-5), we define an objective function  $f(\theta)$  that measures the difference between model-generated option prices and observed market prices for a set of parameters *θ*.

**Steps**:

- 1. **Initialization**: Choose an initial guess for the parameters  $\theta_0$  and set  $k = 0$ .
- 2. **Evaluate**: Calculate the gradient of the objective function with respect to the parameters,  $\nabla f(\theta_k)$ .

3. **Quasi-Newton Direction**: Determine the search direction  $p_k$  using the inverse Hessian approximation:

$$
p_k = -H_k \nabla f(\theta_k)
$$

- 4. **Line Search**: Find a step size  $\alpha_k$  that reduces  $f(\theta)$  along  $p_k$ .
- 5. **Update**: Adjust the parameters:

$$
\theta_{k+1} = \theta_k + \alpha_k p_k
$$

- 6. **Limited Memory Update**: Update the inverse Hessian approximation.
- 7. **Check for Convergence**: If criteria are satisfied (e.g., gradient close to zero), stop. Otherwise, return to Step 2.

**Box Constraints**: Enforce constraints on the parameter values, if needed.

# CG-METHOD (CONJUGATE GRADIENT METHOD)

**Objective Function:** As before, define an objective function  $f(\theta)$  [\[37\]](#page-114-5).

**Steps**:

- 1. **Initialization**: Choose an initial guess for the parameters  $\theta_0$  and compute the initial gradient  $g_0 = \nabla f(\theta_0)$ . Set the search direction  $p_0 = -g_0$  and  $k = 0$ .
- 2. **Compute step size**:

$$
\alpha_k = \frac{g_k^T g_k}{p_k^T \nabla^2 f(\theta_k) p_k}
$$

3. **Update parameters**:

$$
\theta_{k+1} = \theta_k + \alpha_k p_k
$$

4. **Update gradient**:

$$
g_{k+1} = g_k + \alpha_k \nabla^2 f(\theta_k) p_k
$$

- 5. **Check for Convergence**: If ∥*gk*+1∥ is sufficiently small, stop.
- 6. **Compute beta**:

$$
\beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}
$$

7. **Update direction**:

$$
p_{k+1} = -g_{k+1} + \beta_k p_k
$$

8. Return to Step 2.

# **8.1.2.** CALIBRATION TO OPTION PRICE

In order to properly calibrate with respect to the option price, we use as an objective function

$$
\min_{\theta} \frac{1}{M} \sum_{m=1}^{M} \frac{(v_{market} - v_{model}(\theta))^{2}}{v_{market}^{2}}.
$$
\n(8.10)

We start by calibrating our model to synthetic market prices generated by the COS method. For this purpose we generate around 216 option prices with maturities varying between (0.1,0.25,0.5,0.75,1.0), different strike levels and a random combination of parameters within the range defined in [7.1.](#page-74-0) In the following table, we show the mean absolute and maximal absolute error throughout the whole calibration process. We note that the calibrated option values are very close to the the synthetic market prices with maximum absolute error over all calibrations being 0.00463.



In order to calibrate against actual market data, we fix two parameters  $v_0$  and  $\kappa$ . The choice of these two are quite practical.  $v_0$  can very well be estimated by ATM implied volatility of some short term maturity. For this purpose, consider the following lemma by  $[11]$ 

**Lemma 8.1.1.** *(Term structure of the Black-Scholes implied volatility in the Heston Model).*

$$
\sigma_{ATM}^2(T) \approx \frac{1}{T} \int_0^T \left[ \left( v - \theta' \right) e^{-\kappa' t} + \theta' \right] dt
$$

$$
= \left( v - \theta' \right) \frac{1 - e^{-\kappa' T}}{\kappa' T} + \theta'
$$

*where*  $\kappa' = \kappa - \frac{\rho \eta}{2}$  $\frac{\partial \eta}{\partial z}$ ,  $\theta' = \frac{\theta \kappa}{\kappa'}$ . The ATM BS implied variance  $\lim_{T \to 0} \sigma_{BS}^2(T) \big|_{ATM} \to v_0$ .

As a result we conclude that the relationship between the ATM black-scholes variance and the initial variance  $\nu$ 0 can be assumed as linear. Hence, we estimate the initial variance with an ATM options with maturity up to 1 months.

In order to justify the fixing of *κ*, we make a small sensitivity analysis of the parameters  $\kappa$ ,  $\bar{\nu}$  and  $\nu$ . In the figures [8.2](#page-91-0)[-8.4](#page-92-0) below, we find that different values of  $\kappa$  have a bigger influence for options that a further out-of-the-money, while *ν* drives the skewness of the implied volatility skew. Lastly, we observe that higher  $v_0$ - values shift the whole smile/skew.

These characteristics enable us to fix *κ* and by slightly changing the skewness as well as the up or downward shift, to overcome the defiance of fixing *κ*. In Figure [8.5,](#page-92-1) we can

<span id="page-91-0"></span>





Figure 8.3: Implied volatility with different *κ*-values.

see by carefully alternating  $\bar{v}$  and *ν*, we can even overcome large differences in *κ*. Again, we note the values for *v*̄ and *ν* were chosen based on careful consideration and without any optimization.

Now, we try to calibrate SPX options from 06-Oct-2023 for different maturities. We estimated the initial variance as mentioned above with a value of  $v_0 = 0.025$ , while we fixed  $\kappa = 0.5$ .

Important to know, is that the calibration procedure for both examples took less than 1 minute while a the calibration procedure using the lifted Heston model, took in between 3-6 hours. We can see in [8.7](#page-93-0) and [8.6](#page-92-2) that the COS-CPD calibration is nicely able to handle the the skewness of the price. Even though our model prices seems deviating quite a lot for deep in the money option prices, we are still in a relative error range of  $10^{-1}$ .

<span id="page-92-0"></span>

Figure 8.4: Implied volatility with different *ν*-values.

<span id="page-92-1"></span>

Figure 8.5: Implied volatility fit for different *κ*-, *ν*- and *v*-values

<span id="page-92-2"></span>

Figure 8.6: Calibrated SPX option prices on 06-Oct-2023 for maturity 0.8 years

<span id="page-93-0"></span>

Figure 8.7: Calibrated SPX option prices on 06-Oct-2023 for maturity 0.2 years

# **9**

# **NOVEL METHOD OF PRICING BARRIER AND BERMUDAN OPTIONS UNDER THE ROUGH HESTON MODEL**

We want to price Barrier and Bermudan options with the COS method under the rough Heston model. Pricing such options are essentially pricing path-dependent options. As [\[9\]](#page-112-2) states, we have to make use of the a priori not known joint density function between the log-stock price and the variance process given by a CIR process. For the classical Heston model, however, it can be recovered from the characteristic function in closed form. As we have already seen in the previous sections there is only a semi-closed-form solution available when dealing with the rough Heston model.

# **9.1.** BI-VARIATE CHARACTERISTIC FUNCTION OF LOG-STOCK AND VARIANCE IN THE CLASSICAL HESTON MODEL

As previously mentioned by using the 2D COS method, the computation for the bivariate characteristic function of the log-stock and variance process is needed. Accord-ing to [\[5\]](#page-112-3) the bi-variate characteristic function for the affine process  $\mathbf{X}_t = (log(S_t))$  $X_t$ ,  $v_t$ ) is of the form

$$
\phi(\mathbf{a}, \mathbf{X}_t, t, T) = \mathbb{E}\left[e^{ia_1X_T + ia_2v_T} \mid X_t, v_t\right] = \exp\left(B_1(T - t, \mathbf{a})X_t + B_2(T - t, \mathbf{a})v_t + A(T - t, \mathbf{a})\right),\tag{9.1}
$$

with

$$
\beta := \lambda - i \rho v a_1
$$
  
\n
$$
D := \sqrt{\beta^2 + v^2 a_1 (i + a_1)}
$$
  
\n
$$
h := (\beta - D - i a_2 v^2) / (\beta + D - i a_2 v^2).
$$
\n(9.2)

The functions  $A, B_1$ , and  $B_2$  are solutions to a system of ordinary differential equations (ODEs) of Riccati type:

$$
\frac{\partial B_1(t, \mathbf{a})}{\partial t} = 0,
$$
\n
$$
\frac{\partial B_2(t, \mathbf{a})}{\partial t} = \frac{1}{2} \eta^2 B_2^2(t, \mathbf{a}) - \beta B_2(t, \mathbf{a}) - \frac{1}{2} a_1 (i + a_1),
$$
\n
$$
\frac{\partial A(t, \mathbf{a})}{\partial t} = \lambda \theta B_2(t, \mathbf{a}),
$$
\n(9.3)

with initial conditions  $B_1(0, \mathbf{a}) = i a_1, B_2(0, \mathbf{a}) = i a_2, A(0, \mathbf{a}) = 0$ . Solving the ODEs gives

$$
B_1(t, \mathbf{a}) = ia_1
$$
  
\n
$$
B_2(t, \mathbf{a}) = \frac{1}{v^2} \frac{\beta - D - (\beta + D)he^{-Dt}}{1 - he^{-Dt}}
$$
  
\n
$$
A(t, \mathbf{a}) = \frac{\lambda \theta}{v^2} \left[ (\beta - D)t - 2\ln\left(\frac{he^{-Dt} - 1}{h - 1}\right) \right].
$$
\n(9.4)

# **9.2.** BI-VARIATE CHARACTERISTIC FUNCTION OF LOG-STOCK AND VARIANCE IN THE ROUGH HESTON MODEL

In order to use the 2D COS method for the rough Heston model, we first have to find the bi-variate characteristic function for the rough Heston case. The authors of [\[20\]](#page-113-2) expand the idea of [\[3\]](#page-112-4) define the full Fourier-Laplace functional as

$$
\mathbb{E}\left[e^{u_1\log S_T + u_2v_T + (f_1*\log S)_T + (f_2*v)_T}\right] = \exp\left(\phi(T) + \psi_1(T)\log S_0 + I^{1-\alpha}\psi_2(T)v_0\right),\tag{9.5}
$$

where  $\psi_1 = u_1 + 1 * f_1$  and  $\phi$  and  $\psi_2$  solve the fractional Riccati equations

$$
\phi' = \kappa \bar{\nu} \psi_2, \qquad \phi(0) = 0,
$$
  

$$
D^{\alpha} \psi_2 = f_2 + \frac{1}{2} (\psi_1^2 - \psi_1) + (\rho \sigma \psi_1 - \lambda) \psi_2 + \frac{\nu^2}{2} \psi_2^2, \quad I^{1-\alpha} \psi_2(0) = u_2,
$$
 (9.6)

where ∗ is the convolution between *f*<sup>1</sup> and *log S* defined in [3.1.7](#page-31-0)

2

For our purpose, we set  $f_1 = f_2 = 0$ ,  $u_1 = ia_1$ ,  $u_2 = ia_2$  and obtain

$$
\mathbb{E}\left[e^{ia_1\log S_T + ia_2\nu_T}\right] = \exp\left(\phi(T) + ia\log S_0 + I^{1-\alpha}\psi_2(T)\nu_0\right),\tag{9.7}
$$

where  $\psi_1 = ia_1$  and  $\phi$  and  $\psi_2$  solve the fractional Riccati equations

$$
\phi' = \kappa \bar{\nu} \psi_2, \qquad \phi(0) = 0,
$$
  

$$
D^{\alpha} \psi_2 = -\frac{1}{2} a(a+i) + (\rho \sigma i a - \lambda) \psi_2 + \frac{v^2}{2} \psi_2^2, \quad I^{1-\alpha} \psi_2(0) = i a_2.
$$
 (9.8)

# **9.3.** 2D COS METHOD FOR OPTION PRICING

For pricing path-dependent options like Barrier or Bermudan options under the Heston dynamics, we are dealing with a multidimensional asset price process. For this purpose [\[5\]](#page-112-3) established the two-dimensional Fourier cosine series expansion or short 2D COS method. The idea behind this is similar to the 1D COS method introduced above.

# **9.3.1.** GENERAL IDEA AND FORMULA FOR EUROPEAN OPTIONS

We again start with the risk-neutral option valuation formula

$$
\nu(t_0, \mathbf{x}) = e^{-r\Delta t} \mathbb{E}^{t_0, \mathbf{x}} \left[ g\left( \mathbf{X}_T \right) \right] = e^{-r\Delta t} \iint_{\mathbb{R}^2} g(\mathbf{y}) f(\mathbf{y} \mid \mathbf{x}) d\mathbf{y},\tag{9.9}
$$

where in this case  $\mathbf{x} = (x_1, x_2)$  is the current state,  $f(y_1, y_2 | x_1, x_2)$  is the conditional density function, *r* is the risk-free rate, and time to expiration is denoted by  $\Delta t := T - t_0$ . In our application of the Heston model  $\mathbf{x} = (x_1, x_2)$  is equivalent to  $\mathbf{x} = (x = log(S), v)$ . By truncating **x** to a domain  $[a_1,b_1] \times [a_2,b_2] \subset \mathbb{R}^2$  without losing significant accuracy, we find that using multidimensional Fourier cosine expansion

$$
v(t_0, \mathbf{x}) = e^{-r\Delta t} \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(\mathbf{y}) f(\mathbf{y} | \mathbf{x}) dy_1 dy_2
$$
  
=  $e^{-r\Delta t} \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(\mathbf{y}) \sum_{k_1=0}^{+\infty} \sum_{k_2=0}^{+\infty} A_{k_1, k_2}(\mathbf{x}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) dy_1 dy_2.$  (9.10)

We note that - equivalent to the 1D cae -  $\Sigma'$  means that the first term of the summation is weighted by  $\frac{1}{2}$ . In the second line, we replaced the conditional density with its Fourier cosine expansion on  $[a_1, b_1] \times [a_2, b_2]$  with respect to **y**, with the series coefficients defined by

$$
A_{k_1,k_2}(\mathbf{x}) := \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(\mathbf{y} \mid \mathbf{x}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) dy_1 dy_2. \tag{9.11}
$$

Following the approach of [\[5\]](#page-112-3), we truncate the series and approximate the series coefficients by

$$
F_{k_1,k_2}(\mathbf{x}) := \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \iint_{\mathbb{R}^2} f(\mathbf{y} \mid \mathbf{x}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) dy_1 dy_2. \tag{9.12}
$$

Now using

$$
2\cos(\alpha)\cos(\beta) = \cos(\alpha + \beta) + \cos(\alpha - \beta),
$$

we obtain

$$
2F_{k_1,k_2}(\mathbf{x}) = F_{k_1,k_2}^+(\mathbf{x}) + F_{k_1,k_2}^-(\mathbf{x}),
$$

where

$$
F_{k_1,k_2}^{\pm}(\mathbf{x}) := \frac{2}{b_1-a_1} \frac{2}{b_2-a_2} \iint_{\mathbb{R}^2} f(\mathbf{y} \mid \mathbf{x}) \cos \left( k_1 \pi \frac{y_1-a_1}{b_1-a_1} \pm k_2 \pi \frac{y_2-a_2}{b_2-a_2} \right) dy_1 dy_2.
$$

Analogously to the 1D COS method,  $F_{k_1,k_2}$  are now given by

<span id="page-97-0"></span>
$$
F_{k_1,k_2} = \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \operatorname{Re} \left( \varphi \left( \frac{k_1 \pi}{b_1 - a_1}, \pm \frac{k_2 \pi}{b_2 - a_2} \mid \mathbf{x} \right) \exp \left( -ik_1 \pi \frac{a_1}{b_1 - a_1} \mp ik_2 \pi \frac{a_2}{b_2 - a_2} \right) \right),\tag{9.13}
$$

with *ϕ*(· | **x**) representing the bivariavte conditional characteristic function of **X***<sup>T</sup>* , given  $\mathbf{X}_{t_0} = \mathbf{x}$ .

As a result, we find for the 2D COS method formula for pricing European options that,

$$
\hat{\nu}(t_0, \mathbf{x}) := \frac{b_1 - a_1}{2} \frac{b_2 - a_2}{2} e^{-r \Delta t} \sum_{k_1 = 0}^{N_1 - 1} \sum_{k_2 = 0}^{N_2 - 1} \frac{1}{2} \left[ F_{k_1, k_2}^+(\mathbf{x}) + F_{k_1, k_2}^-(\mathbf{x}) \right] V_{k_1, k_2}(T). \tag{9.14}
$$

Taking into account [\(9.13\)](#page-97-0) finally yields

$$
\hat{v}(t_0, \mathbf{x}) = e^{-r\Delta t} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \frac{1}{2} \left[ \text{Re} \left( \varphi \left( \frac{k_1 \pi}{b_1 - a_1}, + \frac{k_2 \pi}{b_2 - a_2} \, | \, \mathbf{x} \right) \, \exp \left( i k_1 \pi \frac{-a_1}{b_1 - a_1} + i k_2 \pi \frac{-a_2}{b_2 - a_2} \right) \right) \right] \n+ \text{Re} \left( \varphi \left( \frac{k_1 \pi}{b_1 - a_1}, - \frac{k_2 \pi}{b_2 - a_2} \, | \, \mathbf{x} \right) \, \exp \left( i k_1 \pi \frac{-a_1}{b_1 - a_1} - i k_2 \pi \frac{-a_2}{b_2 - a_2} \right) \right) \right] V_{k_1, k_2}(T). \tag{9.15}
$$

# **9.3.2.** 2D COS METHOD FOR PRICING DISCRETE BARRIER OPTIONS

Throughout the following section, we adjust the method presented by [\[5\]](#page-112-3) to price discretely monitored barrier options. Throughout this section, only the knock-out event is considered, or even more specifically the up-and-out barrier option. This means that the option becomes worthless if on one of the *M* monitoring dates  $\mathcal{T} = \{t_m : 1 \le m \le M\}$ , with  $t_M = T$ , the underlying stock price process  $(S_t)_{t \geq 0}$  is above the barrier level *B*. Moreover, the option can only be exercised at maturity *T*. Equivalently this means that the payoff is:

$$
\nu(x,T) = \left(\alpha \left(S_T - K\right)\right)^+ \cdot 1_{\tau > T},\tag{9.16}
$$

where  $\tau := \inf\{t_m \geq t_0 : S_{t_m} \geq H\}$  is the first monitoring date that registers a breach of  $S_t$  on the barrier level *B*. As a result, the indicator  $1_{\tau>T}$  is zero once the barrier level is crossed at one of the monitoring dates by the underlying stock process. Furthermore, *x* is the log-asset price and  $\alpha = 1$  for a call option and -1 for a put option.

Knowing the payoff at time *T* , this problem is solved backward in time using the following relation

$$
\begin{cases}\nv(t_{\mathcal{M}}, \mathbf{x}) = g(\mathbf{x}) \\
c(t_{m-1}, \mathbf{x}) = e^{-r\Delta t} \mathbb{E}\left[v\left(t_m, \mathbf{X}_{t_m}\right) | \mathbf{X}_{t_{m-1}} = \mathbf{x}\right] \\
v(x, t_{m-1}) = \begin{cases}\n0, & \text{if } x \geq h \\
c(x, t_{m-1}), & \text{if } x < h\n\end{cases},\n\end{cases} \tag{9.17}
$$

where  $h = log(B/K)$  and  $c(t_{m-1}, \mathbf{x})$  is called the continuation value and can be approximated by

<span id="page-98-0"></span>
$$
\hat{c}\left(t_{m-1},\mathbf{x}\right) := \frac{b_1 - a_1}{2} \frac{b_2 - a_2}{2} e^{-r\Delta t} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \frac{1}{2} \left[ F_{k_1,k_2}^+(\mathbf{x}) + F_{k_1,k_2}^-(\mathbf{x}) \right] V_{k_1,k_2}\left(t_m\right). \tag{9.18}
$$

The Fourier coefficients of the value function in  $(9.18)$  are given by

$$
V_{k_1,k_2}(t_m) := \frac{2}{b_1-a_1} \frac{2}{b_2-a_2} \int_{a_2}^{b_2} \int_{a_1}^{b_1} v(t_m, \mathbf{y}) \cos\left(k_1 \pi \frac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \frac{y_2-a_2}{b_2-a_2}\right) dy_1 dy_2.
$$

As shown in equation [\(9.18\)](#page-98-0), the continuation value at *tm*−<sup>1</sup> is calculated using the 2D COS method but depends on the coefficients of the value function at *tm*. Hence, we have to find a recursive relation of these coefficients and solve them backward in time. Using

$$
\nu(x, t_{m-1}) = \hat{\nu}(x, t_{m-1}) = \begin{cases} 0, & \text{if } x \ge h \\ \hat{c}(x, t_{m-1}), & \text{if } x < h \end{cases}
$$

and

$$
c(x,t_{m-1})\approx \hat{c}(x,t_{m-1})
$$

from  $9.18$ , we can approximate  $V_{k_1,k_2}(t_m)$  by

$$
\hat{V}_{k_1,k_2}(t_{m-1}) = \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \int_{a_2}^{b_2} \int_{a_1}^{b_1} \hat{v}(t_{m-1}, \mathbf{y}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) dy_1 dy_2
$$
\n
$$
= \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \int_{a_2}^{b_2} \int_{a_1}^{h} \hat{c}(t_{m-1}, \mathbf{y}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) dy_1 dy_2
$$
\n
$$
= \hat{C}_{k_1,k_2}(a_2, b_2, a_1, h, t_{m-1}), \tag{9.19}
$$

with

$$
\hat{C}_{k_1,k_2}(a_2,b_2,x_1,x_2,t_{m-1}) = \frac{2}{b_1-a_1} \frac{2}{b_2-a_2} \int_{a_2}^{b_2} \int_{x_1}^{x_2} \hat{c}(t_{m-1},y) \cos\left(k_1 \pi \frac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \frac{y_2-a_2}{b_2-a_2}\right) dy_1 dy_2.
$$
\n(9.20)

Taking into account [\(9.18\)](#page-98-0) yields

<span id="page-98-1"></span>
$$
\hat{C}_{k_1,k_2}(a_2,b_2,a_1,h,t_{m-1}) = \int_{a_2}^{b_2} \int_{a_1}^{h} \sum_{j_1=0}^{N_1-1/N_2-1} \sum_{j_2=0}^{N_2-1} e^{-r\Delta t} \frac{1}{2} \left[ F_{j_1,j_2}^+(y) + F_{j_1,j_2}^-(y) \right]
$$
\n
$$
\hat{V}_{j_1,j_2}(t_m) \cos \left( k_1 \pi \frac{y_1 - a_1}{b_1 - a_1} \right) \cos \left( k_2 \pi \frac{y_2 - a_2}{b_2 - a_2} \right) dy_1 dy_2.
$$
\n(9.21)

Knowing  $\hat{V}_{k_1,k_2}(t_{\mathscr{M}})$  as the Fourier coefficients of the value function for normal European option this recursion formula is well-defined.

**9**

### 84 9. NOVEL METHOD OF PRICING BARRIER AND BERMUDAN OPTIONS UNDER THE ROUGH HESTON MODEL

# **9.4.** PRICING BARRIER OPTIONS UNDER THE ROUGH HESTON MODEL

For a fixed time step  $\Delta t = t_m - t_{m-1}$  defining  $u_1 = ia_1$  and  $u_2 = ia_2$ , we find that

$$
\varphi(u_1, u_2 \mid x, v) := \phi(\mathbf{u}, \mathbf{x}, t_m, t_{m+1}) := e^{u_1 x} e^{\phi(\Delta t, \mathbf{u}) v} \varphi_A(\mathbf{u}), \text{ where}
$$
\n
$$
\varphi_A(\mathbf{u}) := e^{I^{1-\alpha} \psi_2(\Delta t, \mathbf{u})}.
$$
\n(9.22)

As a result, we find that

$$
\hat{C}_{k_1,k_2}(a_2, b_2, a_1, h, t_{m-1})
$$
\n
$$
= \text{Re}\left(\sum_{j_1=0}^{N_1-1} \sum_{j_2=0}^{N_2-1} \frac{1}{2} \varphi_A\left(\frac{j_1\pi}{b_1-a_1}, +\frac{j_2\pi}{b_2-a_2}\right) V_{j_1,j_2}(t_M)
$$
\n
$$
M^+_{k_1,j_1}(a_1, h, a_1, b_1) H^+_{k_2,j_2}\left(a_2, b_2, a_2, b_2, \frac{j_1\pi}{b_1-a_1}\right)\right)
$$
\n
$$
+ \text{Re}\left(\sum_{j_1=0}^{N_1-1} \sum_{j_2=0}^{N_2-1} \varphi_A\left(\frac{j_1\pi}{b_1-a_1}, -\frac{j_2\pi}{b_2-a_2}\right) V_{j_1,j_2}(t_M)
$$
\n
$$
M^+_{k_1,j_1}(a_1, h, a_1, b_1) H^-_{k_2,j_2}\left(a_2, b_2, a_2, b_2, \frac{j_1\pi}{b_1-a_1}\right)\right)
$$
\n
$$
:= \text{Re}\left(\sum_{j_1=0}^{N_1-1} M^+_{k_1,j_1}(a_1, h, a_1, b_1) \mathscr{A}_{j_1,k_2}\right),
$$
\n(9.23)

where the matrices  $H^{\pm}$  and  $M^{+}$  are given by

$$
H_{k_2,j_2}^{\pm}\left(a_2,b_2,a_2,b_2,\frac{j_1\pi}{b_1-a_1}\right)
$$
  
:= 
$$
\frac{2}{b_2-a_2}\int_{a_2}^{b_2}e^{y_2\phi(\left(\Delta t,\frac{j_1\pi}{b_1-a_1},\pm\frac{j_2\pi}{b_2-a_2}\right)}e^{\pm i j_2\pi\frac{-a_2}{b_2-a_2}}\cos\left(k_2\pi\frac{y_2-a_2}{b_2-a_2}\right)dy_2,
$$
\n(9.24)

and

$$
M_{m,n}^+(x_1, x_2, a, b) := \frac{2}{b-a} \int_{x_1}^{x_2} e^{in\pi \frac{y-a}{b-a}} \cos\left(m\pi \frac{y-a}{b-a}\right) dy.
$$
 (9.25)

The  $(N_1 \times N_2)$ -matrix  $\mathscr{A}^q$  is now calculated in a row-wise fashion and the row-vector  $\mathscr{A}_{j_1,\dots} = {\{\mathscr{A}_{j_1,k_2}\}}_{k_2=0}^{N_2-1}$  is denoted as two matrix-vector multiplications, i.e.,

$$
\mathcal{A}_{j_1,..} = H^+ \left( a_2, b_2, a_2, b_2, \frac{j_1 \pi}{b_1 - a_1} \right) \mathbf{w}_{j_1..}^+ + H^- \left( a_2, b_2, a_2, b_2, \frac{j_1 \pi}{b_1 - a_1} \right) \mathbf{w}_{j_1..}^- \tag{9.26}
$$

with

$$
\mathbf{w}_{j_{1},}^{\pm} := \left\{ w_{j_{1},j_{2}}^{\pm} \right\}_{j_{2}=0}^{N_{2}-1}, \quad \text{with} \quad w_{j_{1},j_{2}}^{\pm} := \frac{1}{2} \varphi_{A} \left( \frac{j_{1} \pi}{b_{1} - a_{1}}, \pm \frac{j_{2} \pi}{b_{2} - a_{2}} \right) V_{j_{1},j_{2}}(t_{m}). \tag{9.27}
$$

# **9.4.1.** NUMERICAL RESULTS

# CHOICE OF BENCHMARK

For validating the pricing method, we choose  $H = \frac{1}{2}$ . As a result, the rough Heston model coincides with the classical Heston model. Next, we use Monte-Carlo simulation for continuously monitored Barrier options under the classical Heston model and check if the Barrier option price using the 2D COS-method lies in between a 5%-confidence interval.

### CONVERGENCE IN THE NUMBER OF BASIS FUNCTIONS

<span id="page-100-0"></span>

Figure 9.1: Convergence in the number of cosine basis functions with model parameter  $\rho = -0.6$ ,  $\bar{v} = 0.16$ , *ν* = 0.5, *κ* = 1.5 and *v*<sup>0</sup> = 0.0925. Moreover, we have *S*<sup>0</sup> = 90, *K* = 90 and Barrier *B* = 120 as well as *T* = 0.5 with number of watchtimes being 30.

As we can see in [9.1,](#page-100-0) with an increasing number of cosine basis functions the price of the up and out-call option becomes constant and lies in between a 5%− confidence interval, which is desired. However, the computational time increases drastically with the number of "cos-" terms.

# COMPUTATIONAL COMPLEXITY ADAMS SCHEME

In the numerical experiment above, we set  $N = 500$ . However, we observe a similar behaviour in the computational time as in the 1D-approximation [\(6.2\)](#page-57-0) of the characteristic function whenever we increase the number of steps taken in the Adams scheme. In the 2D case, an increase in the number of timesteps taken in the Adams scheme leads to a significant increase in the overall pricing process as we can see in [9.2.](#page-101-0) The computational complexity is  $O(N_{cos}^2 N^2)$ , where  $N_{cos}$  is the number of space steps taken in the COS method.

## CONVERGENCE IN THE NUMBER OF BASIS FUNCTIONS VS EUROPEAN OPTIONS

In the next numerical experiment, we validate our model by pricing our Barrier option against a European one by setting the watchtime to 1, which is equivalent to considering a European option observed at time T. We set the Barrier level *B* to a value high enough

<span id="page-101-0"></span>

Figure 9.2: Computational time with different numbers of time steps N in the Adams scheme with model  $p$ arameter *ρ* = −0.6,  $\bar{v}$  = 0.16, *ν* = 0.5, *κ* = 1.5 and *v*<sub>0</sub> = 0.0925. Moreover, we have *S*<sub>0</sub> = 90, *K* = 90, and Barrier  $B = 120$  as well as  $T = 0.5$  with the number of watchtimes being 30.

to consider it "unreachable". We observe that indeed the prices coincide with a sufficient number of cosine basis functions. As we can see in [9.3,](#page-101-1) in this specific example 250 cosine basis functions are needed to get an error of  $10^{-13}$ .

<span id="page-101-1"></span>

Figure 9.3: Convergence in the number of cosine basis functions against 1D Cos-European option price with model parameter  $\rho = 0.4$ ,  $\bar{v} = 0.060$ ,  $v = 0.3$ ,  $\kappa = 2.5$  and  $v_0 = 0.065$ . Moreover, we have  $S_0 = 90$ ,  $K = 90$  and Barrier  $B = 1200$  as well as  $T = 0.5$  with number of watchtimes being 1.

However, throughout this validation, we considered parameter combinations such that the Feller condition, i.e.  $\frac{2\kappa\bar{v}}{v^2} \ge 1$ , is satisfied and the densities are well-behaved in a sense that they do not spike too much. In the following, we consider two parameter sets. one for which the Feller condition is satisfied and one for which it is violated.

We can see in [9.5,](#page-102-0) that for the case where the feller condition is highly violated the density has its peak where  $\nu = 0.0$ , whereas when the Feller condition is satisfied the density has almost no mass whenever  $\nu = 0.0$ . Moreover, we see that the density is more well-behaved in the sense that the values of the density are not peaking too much. In our example, the maximum value is around 20. In contrast, the value of the peaked density

<span id="page-102-0"></span>

Figure 9.4: Density with values satisfying the feller condition.



is around 2000. As a result, similar to the 1D case, we need more cosine basis functions in order to have a good approximation for the density function. Another challenge is pricing discrete barrier options boils down to pricing European options on the very small end. For example, a barrier option with a maturity of 1 year and 24 watchtimes is, in terms of density, similar to pricing European options with a maturity of 2 weeks. The density naturally peaks way more with smaller maturities, as seen in the 1D case as well as in [9.6.](#page-102-1)

<span id="page-102-1"></span>

Figure 9.6: Same parameters as in wider [9.5](#page-102-0) case but with 20 more watchtimes.

As a result, the computational times for pricing barrier options under the rough Heston model vary quite a lot depending on which set of parameters we choose. In the following, we show the option pricing process for 3 different sets of parameters. The first one, whenever the Feller condition is satisfied, i.e.  $\frac{2\kappa \bar{v}}{v^2} >> 1$ , the second set for which the parameters are such that  $\frac{2\kappa \bar{v}}{v^2} \approx 1$ . Lastly, we choose parameters such that the feller condition is violated, i.e.  $\frac{2\kappa \bar{v}}{v^2} < 1$ . Moreover we choose  $T = 0.5$ ,  $S_0 = 90$ ,  $K = 80, 90, 100$ 

and  $B = 120$ . Last but not least, we chose the number of watchtimes to 6, equivalent to monthly monitored Barrier options.

- (a)  $\rho = -0.6$ ,  $\bar{v} = 0.16$ ,  $v = 0.5$ ,  $v_0 = 0.0925$ ,  $\kappa = 1.5$  and  $H = 0.12$ ,
- (b)  $\rho = -0.6$ ,  $\bar{v} = 0.125$ ,  $v = 0.5$ ,  $v_0 = 0.0925$ ,  $\kappa = 1.0$  and  $H = 0.12$ ,
- (c)  $\rho = -0.6$ ,  $\bar{v} = 0.16$ ,  $v = 0.8$ ,  $v_0 = 0.0925$ ,  $\kappa = 0.5$  and  $H = 0.12$ .

We note that 128, 150 and 400 cosines basis functions respectively were chosen in order to be certain that the densities are sufficiently smooth. Moreover, we chose  $N =$ 500, *N* = 500 and *N* = 1000 respectively.

|                   | CPU Time (s) | Price      |       |       |
|-------------------|--------------|------------|-------|-------|
|                   |              | <b>OTM</b> | A'I`M | ITM   |
| Parameter set (a) | 61.3         | 2.362      | 5.172 | 6.603 |
| Parameter set (b) | 96.5         | 2.913      | 5.505 | 6.464 |
| Parameter set (c) | $45.2$ (min) | 2.915      | 5.509 | 6.462 |

Table 9.1: CPU time analysis for different parameters.

# **9.5.** PRICING BERMUDAN OPTIONS UNDER THE ROUGH HES-TON MODEL

Throughout this section, we introduce a method to price Bermudan options under the rough Heston model. As proposed in [\[5\]](#page-112-3), to calculate Bermudan options, we have to divide the domain where the log-stock and variance process live on into early exercise and continuation regions. To do so, we divide the domain  $[a_1, b_1] \times [a_2, b_2]$  into rectangular subdomains  $\mathscr{C}^q$  and  $\mathscr{G}^p$ , so that approximately for all states  $\mathbf{y} \in \mathscr{C}^q$  it is optimal to continue and for all  $y \in \mathcal{G}^p$  it is optimal to exercise the option.

The idea is to split the  $V_{k_1,k_2}$  − coefficients into two parts, namely

$$
V_{k_1,k_2}(t_m) = \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \sum_{p} \iint_{\mathcal{G}^p} g(\mathbf{y}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) d\mathbf{y}
$$
  
+ 
$$
\frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \sum_{q} \iint_{\mathcal{G}^q} c(t_m, \mathbf{y}) \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right) d\mathbf{y}
$$
  
:= 
$$
\sum_{p} G_{k_1,k_2}(\mathcal{G}^p) + \sum_{q} C_{k_1,k_2}(t_m, \mathcal{G}^q) \quad (m \neq 0, \mathcal{M}).
$$
 (9.28)

We straight away see that once we determined the early exercise and continuation regions  $\sum_{p} G_{k_1,k_2}(\mathscr{G}^p)$  has a closed form solution as  $g(y)$  is just the payoff function of a European call option. For  $\sum_{q} C_{k_1,k_2} (t_m, \mathcal{C}^q)$   $(m \neq 0, \mathcal{M})$ , we apply the same recursion formula as for Barrier options given by [9.21,](#page-98-1) i.e.

$$
\hat{C}_{k_1,k_2} (t_{\mathcal{M}-1}, [z_q, z_{q+1}] \times [w_q, w_{q+1}])
$$
\n
$$
= \text{Re}\left(\sum_{j_1=0}^{N_1-1/N_2-1} \sum_{j_2=0}^{N_2-1} \frac{1}{2} e^{-r\Delta t} \varphi_A \left(\frac{j_1 \pi}{b_1-a_1}, +\frac{j_2 \pi}{b_2-a_2}\right) V_{j_1,j_2} (t_{\mathcal{M}})
$$
\n
$$
M^+_{k_1,j_1} (z_q, z_{q+1}, a_1, b_1) H^+_{k_2,j_2} \left(w_q, w_{q+1}, a_2, b_2, \frac{j_1 \pi}{b_1-a_1}\right)\right)
$$
\n
$$
+ \text{Re}\left(\sum_{j_1=0}^{N_1-1/N_2-1/1} \frac{1}{2} e^{-r\Delta t} \varphi_A \left(\frac{j_1 \pi}{b_1-a_1}, -\frac{j_2 \pi}{b_2-a_2}\right) V_{j_1,j_2} (t_{\mathcal{M}})
$$
\n
$$
M^+_{k_1,j_1} (z_q, z_{q+1}, a_1, b_1) H^-_{k_2,j_2} \left(w_q, w_{q+1}, a_2, b_2, \frac{j_1 \pi}{b_1-a_1}\right)\right)
$$
\n
$$
:= \text{Re}\left(\sum_{j_1=0}^{N_1-1} M^+_{k_1,j_1} (z_q, z_{q+1}, a_1, b_1) \mathcal{A}_{j_1,k_2}^q\right)
$$
\n
$$
(9.29)
$$

The procedure for getting the early exercise points is described in [\[5\]](#page-112-3) and [\[9\]](#page-112-2). For this, we divide the domain of the variance dimension,  $[a_2, b_2]$ , into *J* sub-intervals:

$$
[a_2, b_2] = [w_0, w_1] \cup [w_1, w_2] \dots [w_q, w_{q+1}] \dots [w_{J-1}, w_J].
$$

At the center of each sub-interval, we determine the value(s)  $y^*$  for which the optimal exercise policy changes to the optimal continuation, i.e.,

<span id="page-104-0"></span>
$$
g\left(y^*, \frac{1}{2}\left(w_q + w_{q+1}\right)\right) = c\left(t_m, y^*, \frac{1}{2}\left(w_q + w_{q+1}\right)\right). \tag{9.30}
$$

For a call option, we can then define an early-exercise region  $\mathscr{G}^q = [y^*, b_1] \times [w_q, w_{q+1}]$ and a continuation region  $\mathcal{C}^q = [a_1, y^*] \times [w_q, w_{q+1}]$ . Therefore, the computational domain is divided into *J* early-exercise and continuation regions. Note that, in order to solve [9.30,](#page-104-0) we need some root-finding technique. We chose Brent's method. However, the Newton method or any other root-finding method works as well. The values  $w_q$  are computed beforehand on an equidistant grid, i.e.

$$
w_q = a_2 + (b_2 - a_2) q / J.
$$

In order to create a benchmark pricing method for Bermudan options under the rough Heston model, we choose  $H = \frac{1}{2}$  and benchmark our values against Bermudan  $20$  priors priced under the classical Heston model using the 1D COS method as described by [\[9\]](#page-112-2). We consider again the three parameters as above, i.e.

- (a)  $\rho = -0.6$ ,  $\bar{\nu} = 0.16$ ,  $v = 0.5$ ,  $\nu_0 = 0.0925$  and  $\kappa = 1.5$ ,
- (b)  $\rho = -0.6$ ,  $\bar{v} = 0.125$ ,  $v = 0.5$ ,  $v_0 = 0.0925$  and  $\kappa = 1.0$ ,
- (c)  $\rho = -0.6$ ,  $\bar{v} = 0.16$ ,  $v = 0.8$ ,  $v_0 = 0.0925$  and  $\kappa = 0.5$ .

For simplicity, we consider the three parameter sets with  $K = 90$ ,  $S_0 = 90$  and again  $T = 0.5$  with watch times being 6.

|                   | 1D COS method | Price  | Absolute Error |
|-------------------|---------------|--------|----------------|
| Parameter set (a) | 8.1548        | 8.1493 | 0.0055         |
| Parameter set (b) | 7.6422        | 7.6338 | 0.0084         |
| Parameter set (c) | 7 6202        | 7.6105 | 0.0097         |

Table 9.2: Error analysis against 1D COS method for different parameter sets.

We note that, we only used  $N = 800$  for the Adams scheme to approximate the characteristic function. Hence, the error can very well be optimized. However, due to increasing computational time, we believe this is sufficient in order to validate our method.

Next, we again consider the same 3 different parameter sets used above (now extended by  $H = 0.12$ ). We use  $2<sup>4</sup>$  early exercise points. However, we note that especially for parameters for which the feller condition is not satisfied, a lot of cosine basis functions are needed in order to get sufficiently smooth density (Gibbs phenomena). It can be already be seen in [\[5\]](#page-112-3) that for 100 basis functions in each dimension the computational time for the classical Heston model is more than 50 seconds. In the following, we consider  $S_0 = 90$  and  $K = 90$ .

|                   | CPU(s)       | Price  |
|-------------------|--------------|--------|
| Parameter set (a) | 93.4         | 8.2196 |
| Parameter set (b) | 266.4        | 7.7145 |
| Parameter set (c) | $55.6$ (min) | 7.7067 |

Table 9.3: Error analysis against 1D COS method for different parameter sets.
# **10**

### **CONCLUSION AND FUTURE RESEARCH**

The primary objective of this thesis is to provide a comprehensive suite of derivative pricing methodologies within the framework of the rough Heston model. Our exploration begins with European options, highlighting the challenge of efficiently pricing options under the rough Heston model due to significant computational demands. This work initiates with an in-depth introduction to the rough Heston model and its multi-factor approximation, the Lifted Heston model, subsequently comparing the computational efficiency of these two models.

Given the inherent roughness in the volatility process, conventional industry methods for option pricing typically resort to the Monte Carlo method or Fourier-based approaches, such as the COS method developed by [\[4\]](#page-112-0). However, even when employing the relatively efficient COS method for European options, the limitations of both the rough and lifted Heston models become evident, particularly in terms of computational performance during the pricing and calibration processes. This thesis endeavors to address and surmount these challenges, offering novel insights and solutions in the field of derivative pricing under complex models.

Indeed, in Chapter 6 a novel method, the COS-CPD option pricer, was derived in order to efficiently price European options. Throughout a whole set of tests using different parameter sets, strikes and time to maturities a very promising accuracy is maintained while being able to significantly speed up the pricing process in comparison to industry standard methods.

Rough volatility models are complex and harder to work on, but they have some better qualities than other option pricing models. First, the real market data for option prices implies that the volatility behavior is rough as shown in Chapter 2. Moreover, it is shown that the implied volatility behavior of the rough volatility models has a great resemblance to the market-implied volatility behavior. Hence, it can be said that the rough volatility models tend to reflect real market behavior. As a result, it is shown that our novel method of pricing options under the rough Heston model is very well capturing the market-implied behavior.

In quantitative Finance, it is of utmost importance to not only have an efficient pricing method but also be able to calibrate the model to actual market data in an efficient way. Hence, we developed a novel calibration method based on the derivatives of the option price with respect to the model parameters for faster convergence. In Chapter 7, it is shown with actual market data that our calibration approach is indeed able to find suitable optimal parameters to capture the market prices.

Barrier and Bermudan options are exotic derivatives whose payoffs depend on the underlying asset's price path; the former is activated or deactivated by breaching a preset price level, while the latter can be exercised at several predetermined dates, offering a blend of American and European option features. In the last chapter, we derive a novel benchmark method to price Bermudan and Barrier options under the rough Heston model. Even though the computational time is significant, especially when the socalled Feller condition is not satisfied, our method can be used in order to benchmark future improvements in path-dependent option pricing.

#### **10.1.** FUTURE RESEARCH

For further research, we start with our novel COS-CPD option pricing method. Finding the right truncation range is indeed a difficult task and should definitely be the starting point for further research. As of now, one has to have extensive knowledge of the method to have a proper guess for the things truncation range in the first place. Hence, deriving some mathematical proof in order to justify a good choice can be advantageous. Next, as of today we first have to generate the cosine coefficients via SVD in order to have the input data for the COS-CPD method. In the future, one can try to use the COS-CPD method directly on the option price to avoid further numerical errors and gain even better accuracy.

Another approach to optimize the COS-CPS option pricing method is trying to find an optimal number of cosine basis functions. As of now for each parameter, the same number of basis functions is used in the CPD-method. Nevertheless, this might not even be needed as [\[30\]](#page-113-0) showed. This would also help to further reduce the bottleneck in the training set generation.

Our calibration method leads to the idea of hedging under the rough Heston mode. Due to the non-markovianity and the lack of a pricing PDE hedging under the rough Heston model is more difficult. However, due to our COS-CPD approach derivatives can easily be implemented with respect to each of the parameters. Hence, the hedging process can be greatly simplified and might be researched and tested against market data in the future. On pricing path-dependent options like Bermudan or Barrier options, a good starting point for further research is to speed up the whole pricing process. Either trying to find a way to speed up the approximation of the characteristic function or to directly speed up the option price. This might even be doable with our COS-CPD method. However, at some point, the bottleneck in computational time to generate the training set will arise.

### **BIBLIOGRAPHY**

- [1] J. Gatheral, T. Jaisson, and M. Rosenbaum, *Volatility is rough*, 2014. arXiv: [1410.](https://arxiv.org/abs/1410.3394) [3394 \[q-fin.ST\]](https://arxiv.org/abs/1410.3394).
- [2] E. A. Jaber and O. E. Euch, *Multi-factor approximation of rough volatility models*, 2018. arXiv: [1801.10359 \[math.PR\]](https://arxiv.org/abs/1801.10359).
- [3] O. E. Euch and M. Rosenbaum, *The characteristic function of rough heston models*, 2016. DOI: [10.48550/ARXIV.1609.02108](https://doi.org/10.48550/ARXIV.1609.02108). [Online]. Available: [https://arxiv.](https://arxiv.org/abs/1609.02108) [org/abs/1609.02108](https://arxiv.org/abs/1609.02108).
- <span id="page-112-0"></span>[4] F. Fang and C. Oosterlee, "A novel pricing method for european options based on fourier-cosine series expansions," *SIAM J. Scientific Computing*, vol. 31, pp. 826– 848, Jan. 2008. DOI: [10.1137/080718061](https://doi.org/10.1137/080718061).
- [5] M. Ruijter and C. Oosterlee, "Two-dimensional fourier cosine series expansion method for pricing financial options," *SIAM Journal on Scientific Computing*, vol. 34, Jan. 2012. DOI: [10.1137/120862053](https://doi.org/10.1137/120862053).
- [6] P. Wilmott, *Paul Wilmott introduces quantitative finance*. John Wiley Sons, 2007.
- [7] M. Schweizer, "On bermudan options," Jan. 2002. DOI: [10.1007/978- 3- 662-](https://doi.org/10.1007/978-3-662-04790-3_13) [04790-3\\_13](https://doi.org/10.1007/978-3-662-04790-3_13).
- [8] C. W. Oosterlee, *Mathematical Modeling and Computation in Finance: With Exercises and Python and MATLAB Computer Codes*. World Scientific Publishing Co., 2019. DOI: [10.1142/q0236](https://doi.org/10.1142/q0236). [Online]. Available: [https://doi.org/10.1142/](https://doi.org/10.1142/q0236) [q0236](https://doi.org/10.1142/q0236).
- [9] F. Fang and C. Oosterlee, "Pricing early-exercise and discrete barrier options by fourier-cosine series expansions," *Numerische Mathematik*, vol. 114, pp. 27–62, Oct. 2009. DOI: [10.1007/s00211-009-0252-4](https://doi.org/10.1007/s00211-009-0252-4).
- [10] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *The Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.
- [11] J. Gatheral, *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, 2011.
- [12] A. Ben-Israel, "A newton-raphson method for the solution of systems of equations," *Journal of Mathematical Analysis and Applications*, vol. 15, no. 2, pp. 243– 252, 1966, ISSN: 0022-247X. DOI: [https://doi.org/10.1016/0022-247X\(66\)](https://doi.org/https://doi.org/10.1016/0022-247X(66)90115-6) [90115- 6](https://doi.org/https://doi.org/10.1016/0022-247X(66)90115-6). [Online]. Available: [https://www.sciencedirect.com/science/](https://www.sciencedirect.com/science/article/pii/0022247X66901156) [article/pii/0022247X66901156](https://www.sciencedirect.com/science/article/pii/0022247X66901156).
- [13] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, vol. 6, no. 2, pp. 327–343, 1993, ISSN: 08939454, 14657368. [Online]. Available: [http://](http://www.jstor.org/stable/2962057) [www.jstor.org/stable/2962057](http://www.jstor.org/stable/2962057) (visited on 10/25/2023).
- [14] B. B. Mandelbrot and J. W. Van Ness, "Fractional brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.
- [15] M. Fukasawa, "Asymptotic analysis for stochastic volatility: Martingale expansion," *Finance and Stochastics*, vol. 15, pp. 635–654, Dec. 2010. DOI: [10.1007/s00780-](https://doi.org/10.1007/s00780-010-0136-6)  $010-0136-6.$  $010-0136-6.$
- [16] G. Leoni, *A first course in fractional sobolev spaces*, 2023. arXiv: [2303.05940 \[math.AP\]](https://arxiv.org/abs/2303.05940).
- [17] J. Brasseur, *On restrictions of besov functions*, 2018. arXiv: [1706.04462 \[math.FA\]](https://arxiv.org/abs/1706.04462).
- [18] G. Livieri, S. Mouti, A. Pallavicini, and M. Rosenbaum, "Rough volatility: Evidence from option prices," *IISE Transactions*, pp. 1–21, 2018.
- [19] M. Bennedsen, A. Lunde, and M. S. Pakkanen, *Decoupling the short and long term behavior of stochastic volatility*, Available at SSRN: 2846756, 2016.
- [20] E. A. Jaber, M. Larsson, and S. Pulido, *Affine volterra processes*, 2019. arXiv: [1708.](https://arxiv.org/abs/1708.08796) [08796 \[math.PR\]](https://arxiv.org/abs/1708.08796).
- [21] N. Kargas and N. D. Sidiropoulos, "Supervised learning and canonical decomposition of multivariate functions," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1097–1107, 2021. DOI: [10.1109/TSP.2021.3055000](https://doi.org/10.1109/TSP.2021.3055000).
- [22] V. Daftardar-Gejji, *Fractional Calculus and Fractional Differential Equations*. Birkhäuser, 2019.
- [23] A. G. Hawkes, "Point spectra of some mutually exciting point processes," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 33, no. 3, pp. 438–443, 1971.
- [24] P. J. Laub, T. Taimre, and P. K. Pollett, *Hawkes processes*, 2015. arXiv: [1507.02822](https://arxiv.org/abs/1507.02822) [\[math.PR\]](https://arxiv.org/abs/1507.02822).
- [25] E. E. Omar, F. Masaaki, and R. Mathieu, *The microstructural foundations of leverage effect and rough volatility*, 2016. arXiv: [1609.05177 \[q-fin.TR\]](https://arxiv.org/abs/1609.05177).
- [26] E. A. Jaber, *Lifting the heston model*, 2019. arXiv: [1810.04868 \[q-fin.CP\]](https://arxiv.org/abs/1810.04868).
- [27] C. Bayer and S. Breneis, *Markovian approximations of stochastic volterra equations with the fractional kernel*, 2022. arXiv: [2108.05048 \[q-fin.CP\]](https://arxiv.org/abs/2108.05048).
- [28] P. Carmona, L. Coutin, and G. Montseny, "Approximation of some gaussian processes," *Statistical Inference for Stochastic Processes*, vol. 3, pp. 161–171, 2000, Cite this article.
- [29] L. Klomp, "Pricing barrier options under lévy processes using dimension-reduced cosine expansion," Available at http://repository.tudelft.nl/, M.S. thesis, Delft University of Technology, 2023.
- <span id="page-113-0"></span>[30] Z. Cheng, "Dimension reduction techniques for multi-dimensional numerical integrations based on fourier-cosine series expansion," Available at [https://reposi](https://repository.tudelft.nl/islandora/object/uuid:f6ffbcd6-df14-425b-84bb-63e4e105205c)tory. [tudelft.nl/islandora/object/uuid:f6ffbcd6-df14-425b-84bb-63e4e105205c](https://repository.tudelft.nl/islandora/object/uuid:f6ffbcd6-df14-425b-84bb-63e4e105205c), M.S. thesis, Delft University of Technology, 2022.
- [31] Z. Cheng and M. Brands, "Finding dimension-reduced fourier-cosine series expansion via supervised machine-learning and its application in multi-dimensional integration," Manuscript submitted for publication, 2023.
- [32] J. Miller, G. Roeder, and T.-D. Bradley, "Probabilistic graphical models and tensor networks: A hybrid framework," Jun. 2021.
- [33] S. Liu, A. Borovykh, L. A. Grzelak, and C. W. Oosterlee, "A neural network-based framework for financial model calibration," *Journal of Mathematics in Industry*, vol. 9, no. 1, pp. 1–28, 2019.
- [34] J. M. Benítez, J. L. Castro, and I. Requena, "Are artificial neural networks black boxes?" *IEEE transactions on neural networks*, vol. 8 5, pp. 1156–64, 1997. [Online]. Available: <https://api.semanticscholar.org/CorpusID:8127923>.
- [35] J. Håstad, "Tensor rank is np-complete," *Journal of Algorithms*, vol. 11, no. 4, pp. 644– 654, 1990, ISSN: 0196-6774. DOI: [10.1016/0196- 6774\(90\)90014- 6](https://doi.org/10.1016/0196-6774(90)90014-6). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0196677490900146>.
- [36] M. Marsden, "Cubic spline interpolation of continuous functions," *Journal of Approximation Theory*, vol. 10, no. 2, pp. 103–111, 1974, ISSN: 0021-9045. DOI: [https:](https://doi.org/https://doi.org/10.1016/0021-9045(74)90109-9) [//doi.org/10.1016/0021- 9045\(74\)90109- 9](https://doi.org/https://doi.org/10.1016/0021-9045(74)90109-9). [Online]. Available: [https:](https://www.sciencedirect.com/science/article/pii/0021904574901099) [//www.sciencedirect.com/science/article/pii/0021904574901099](https://www.sciencedirect.com/science/article/pii/0021904574901099).
- [37] L. Nazareth, "A relationship between the bfgs and conjugate gradient algorithms and its implications for new algorithms," *SIAM Journal on Numerical Analysis*, vol. 16, no. 5, pp. 794–800, 1979, ISSN: 00361429. [Online]. Available: [http://www.](http://www.jstor.org/stable/2156634) [jstor.org/stable/2156634](http://www.jstor.org/stable/2156634) (visited on 11/05/2023).

## **A**

### **SPX OPTIONS DATA**

This appendix contains the option quotes in order to calibrate our COS-CPD rough volatility model to actual market data [8.](#page-86-0)



Figure A.1: Maturity 1week



Figure A.2: Maturity 2weeks

5,500.00

 $-0.256$ 

 $0.05$ 

100



Figure A.3: Maturity 3weeks

**A**



#### Figure A.4: Maturity 1month



Figure A.5: Maturity 2months





105

## **B**

### **HISTORICAL S&P DATA**

This appendix contains the historical S&P 500 data in order to create the ATM Skew used in [2.](#page-20-0)























