The Lifted Heston Stochastic Volatility Model

Ryan Broodryk

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

October 3, 2020

MPhil in Mathematical Finance,
University of Cape Town.

Contents

Abstract

Can we capture the explosive nature of volatility skew observed in the market, without resorting to non-Markovian models? We show that, in terms of skew, the Heston model cannot match the market at both long and short maturities simultaneously. We introduce Abi Jaber (2019)'s Lifted Heston model and explain how to price options with it using both the cosine method and standard Monte-Carlo techniques. This allows us to back out implied volatilities and compute skew for both models, confirming that the Lifted Heston nests the standard Heston model. We then produce and analyze the skew for Lifted Heston models with a varying number N of mean reverting terms, and give an empirical study into the time complexity of increasing N. We observe a weak increase in convergence speed in the cosine method for increased N, and comment on the number of factors to implement for practical use.

Keywords: Stochastic volatility; Implied volatility; Volatility Skew; Monte-Carlo; Cosine method; Riccati equations; Complexity analysis

Acknowledgments

I would like to thank Alex Backwell and Andrew Soane for their attentive reading, insightful comments and frequent discussions. I would also like to thank Andrew specifically for his help on simplifying the integrals in the Lifted Heston characteristic function. Lastly, much thanks must be given to the Banking Sector Education and Training Authority (BANKSETA), for their generous funding towards my participation in the MPhil. Mathematical Finance degree program.

Chapter 1

Introduction

When we create mathematical models, we often need to make a trade-off between precision and tractability. A more complicated model may better capture the underlying phenomena, but find that it becomes too general for many analytical techniques and numerical methods to be applied to it.

One of the simpler models of stock price dynamics is that given by Black-Scholes (BS). Notably, it yields closed-form solutions for call and put option prices. Of all the inputs it takes, $(S_0, r, \sigma, K, t, T)$, only the volatility σ is neither specified by the contract nor known at the time of inception. Holding the other inputs constant, we have a map between volatilities and vanilla option prices. It is then quite natural to consider for each price what volatility under BS would have led to it. We call this the *implied volatility*, and ranging over possible strike prices K and maturities T gives us the *volatility surface*.

A single slice of a volatility surface, for a fixed maturity T is often called a volatility smile. If BS was a perfect model for market prices, we would expect this to be a constant line for the appropriate volatility. Instead, it turns out that a constant volatility is too simple an assumption, and BS implied higher volatilities from market prices further from the money. For a graphical representation of this, see figure 1.1.

The Heston model, Heston (1993), replaces Black-Schole's constant volatility with a stochastic variance process. The risk-neutral dynamics are given as

$$\begin{aligned} dS_t &= rS_t dt + S_t\sqrt{V_t}dB_t, \qquad S_0 > 0, \\ dV_t &= \lambda(\theta - V_t)dt + \nu\sqrt{V_t}dW_t, \qquad V_0 \geq 0, \end{aligned} \tag{1.1}$$

where S_t and V_t are the stock and variance processes respectively, θ is a mean reversion term, λ controls the rate of reversion and ν the impact of the stochastic part of the process. Here B and W are Brownian motions with correlation $\rho \in [-1, 1]$. The volatilities implied by the Heston model match those of the market quite well for longer maturities as seen in El Euch *et al.* (2018).

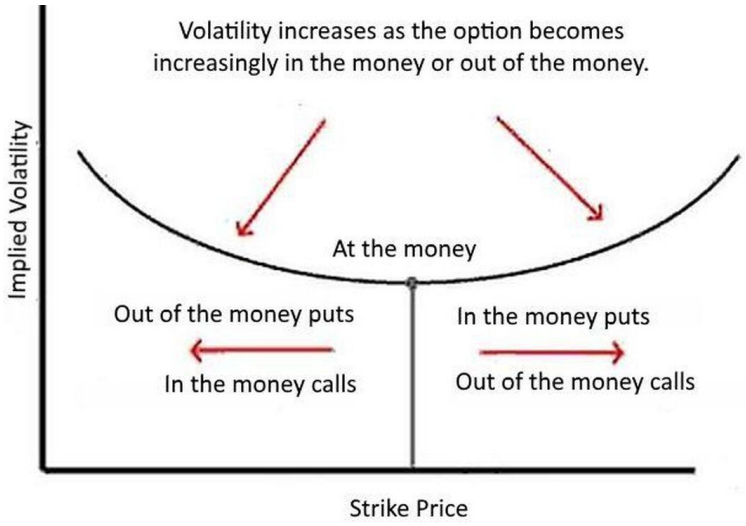

Fig. 1.1: A 'volatility smile' showing higher volatilities for far from the money options at a fixed maturity T. Source: Investopedia.com

At shorter maturities, the degree to which Black-Scholes fits for values of S_0 away from the money gets worse and worse. The Heston model, while a better fit than BS, also struggles to capture the same volatilities as the market here. The *implied volatility skew* (henceforth called skew) is defined as the absolute value

$$|\frac{\partial}{\partial k}\sigma_{\text{implicit}}(k,T)\,|_{k=0}\,| \quad \text{for} \quad k := \ln(K/S_0), \tag{1.2}$$

where $\sigma_{\text{implicit}}(k,T)$ is the volatility surface parameterised by the scaled K. In figure 1.2, we can see that the Heston model has a poor fit for low maturities. Abi Jaber (2019) interprets this as the model struggling to capture the risk of large price movements at short time scales.

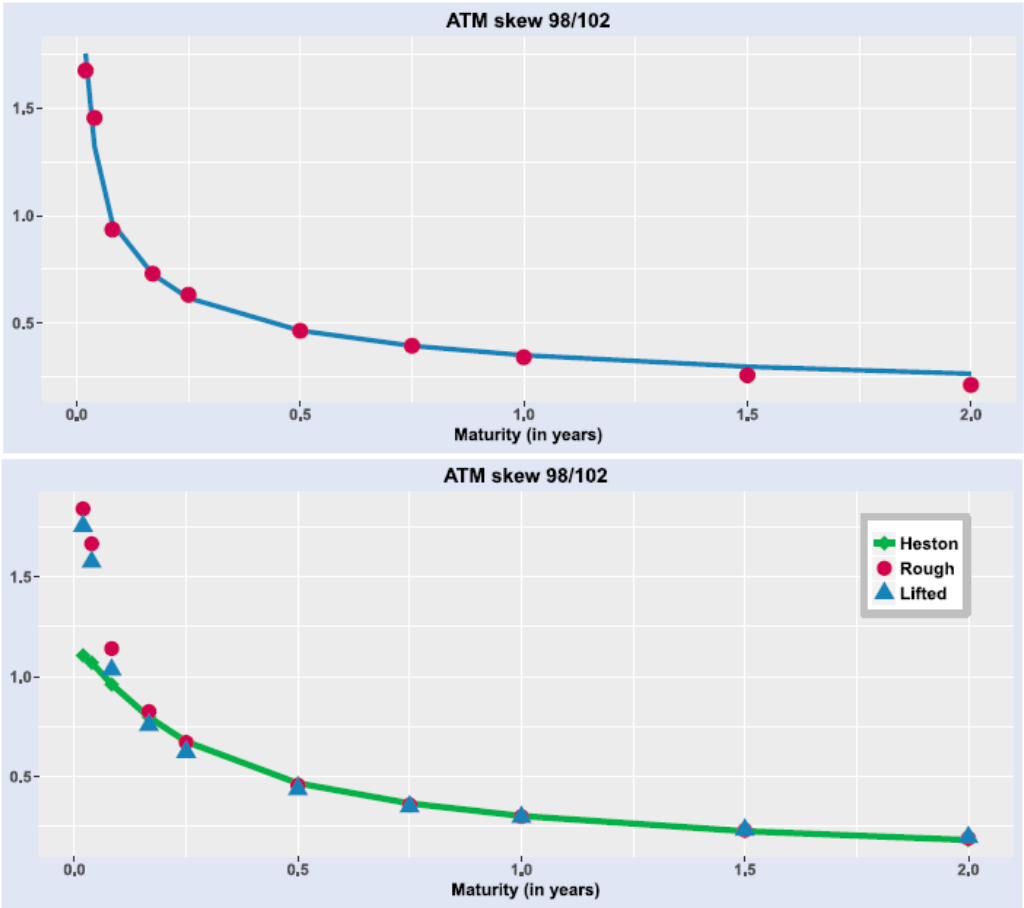

Fig. 1.2: Top: Market skew (red dots) and line of fit. Bottom: Skew of calibrated Rough, Lifted and ordinary Heston models. Both skews are likely computed using a central difference approximation with K at 98% and 102% of Strike. Source: Abi Jaber (2019). For a definition of skew, see (1.2). For an analogous graph on our results, see figure 4.8.

We now turn to Chapter 2 to look at approaches taken in the literature to improve this fit, while bearing in mind the likely trade-off with tractability. Chapter 3 will introduce the numerical methods used to price our options and recover the implied volatilities, with special attention paid to the characteristic functions of our two models. We then present the results of several experiments in Chapter 4, showing how the Lifted Heston reproduces the standard Heston, and how the skew for both models is affected by changes in their parameters. A rudimentary analysis of the time complexity of the Lifted Heston with regards to the number of its factors is contained here as well. Finally, in Chapter 5 we draw conclusions about when the Lifted Heston model should be used, and give our thoughts on a practical choice of N.

Chapter 2

Literature Review

There have been several attempts to improve on the Heston model's ability to fit the market skew at low maturities. In order to jointly model short and long timescales, Cont and Tankov (2003), Gatheral (2011) add jumps to the dynamics. Bergomi (2005), Fouque and Lorig (2011) try to achieve the same by stacking additional random factors. Both of these approaches suffer from the **curse of dimensionality**[1], slowing down calibration as more parameters are introduced. While these improve on the precision of the Heston model, their tractability leaves room for improvement. The Variance-Gamma model, Madan *et al.* (1998), avoids the curse of dimensionality, using a three parameter generalisation of Brownian motion to drive its variance process. The VG process is a Brownian motion that is evaluated at a random time change given by a gamma process. In contrast to Brownian motion, it has finite variation and has no continuous martingale component.

More recently, rough volatility models have been used to fit the implied volatility surface with great precision, as in Bayer *et al.* (2016), El Euch *et al.* (2018) and Gatheral *et al.* (2018). The variance process in these models only uses a single one-dimensional Brownian motion. Few parameters are needed, and therefore the curse of dimensionality is avoided. Unfortunately, rough volatility models are neither semimartingale nor Markovian. Specifically, the variance process is not Markovian and the stock price is not a semimartingale. For $H \neq 0.5$, the variance process has infinite quadratic variation, which is why it is not a semimartingale (El Euch and Rosenbaum (2019)). This is a problem, as the fundamental theorem of asset pricing requires stock prices to be semimartingales in order to have an arbitrage-free system. Additionaly, the non-Markovianity re-introduces the curse of dimensionality as we need to remember the entire history of the process, which only grows as time

[1] **The curse of dimensionality** refers to the phenomena that arise when working with high dimensional spaces, due to the volume of the space scaling exponentialy with each added dimension. In our context each parameter that we have to calibrate for a model is an additional dimension in the space of possible solutions. Adding too many parameters increases the volume of this space to the point where finding an optimal solution quickly becomes intractable.

goes on. This essentially results in an infinite dimensional problem, making models of this type cumbersome to work with.

El Euch *et al.* (2018) combine the tractability of the Heston model with the precision of rough volatility models to create what they call the Rough Heston model. This is described by the risk-neutral dynamics

$$\begin{aligned} dS_t &= S_t\sqrt{V_t}dB_t, \qquad S_0 > 0, \\ V_t &= V_0 + \frac{1}{\Gamma(H+1/2)} \times \int_0^t (t-s)^{H-1/2}(\lambda(\theta - V_s)ds + \nu\sqrt{V_s}dW_s), \end{aligned}$$

where Γ is the Gamma function, and H is a measure of the regularity of the paths of V. Note that in the case of $H = 1/2$, the model recovers the standard Heston model. We set rates to zero for convenience here, which will eliminate the drift term in the Heston (1.1). The Rough Heston may be more tractable than other rough volatility models, but it is still non-Markovian and non-semimartingale. This makes it slow to calibrate and unsuitable for simulation using standard techniques.

Abi Jaber (2019) describes the Lifted Heston model, which aims to keep as much of the precision of the Rough Heston as possible, while improving tractability by keeping the model semimartingale and Markovian. The variance process is a weighted sum of N factors, driven by the same one-dimensional Brownian motion, but mean reverting at different speeds. This is thought to capture the effects of different time-scales, allowing better fit to the market skew without sacrificing too much tractability. For a fixed N, the risk-neutral dynamics of an N factor Lifted Heston model are given by

$$\begin{aligned} dS_t &= S_t\sqrt{V_t}dB_t, \qquad S_0 > 0, \\ V_t &= g_0(t) + \sum_{i=1}^N c_i U_t^i, \\ dU_t^i &= (-x_i U_t^i - \lambda V_t)dt + \nu\sqrt{V_t}dW_t, \\ U_0^i &= 0, \quad i = 1, \dots, N, \end{aligned} \tag{2.1}$$

with parameters the function g_0, constants $\lambda, \nu \in \mathbb{R}_+$, $c_i, x_i \geq 0$ and $B = \rho W + \sqrt{1-\rho^2}W^\perp$ a Brownian motion with correlation p to W. Here $(W, W^\perp)$ can be thought of as a two dimensional Brownian motion on a fixed filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ with $\rho \in [-1, 1]$. While more general forms of g_0 can be used, we will stick to the form

$$g_0(t) = V_0 + \lambda\theta \sum_{i=1}^N c_i \int_0^t e^{-x_i(t-s)}ds. \tag{2.2}$$

We see that in the case $N = 1, c = 1, x = 0$ then the model boils down to the classical Heston. Abi Jaber and El Euch (2019) show that in the limit $N \to \infty$, the Lifted Heston converges to the Rough Heston. This allows an interpretation of the Rough Heston as a superposition of mean reverting terms at infinitely many time scales, which could explain why it fits so well even in the short side of the volatility surface. While the Lifted Heston adds $2N$ parameters for the x_i and c_i, it is possible to parameterise these in such a way as to only require two additional parameters rather than $2N$. These are the Hurst index H, which controls the roughness of the variance process, and a parameter r_N which can be picked in such a way as to provide convergence to the Rough Heston in the limit $N \to \infty$. The parameterisation is as follows:

$$\begin{aligned} \alpha &= H + 1/2, \\ c_i &= \frac{(r_N^{1-\alpha} - 1) r_N^{(\alpha-1)(1+N/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_N^{(1-\alpha)i}, \\ x_i &= \frac{(1-\alpha)(r_N^{2-\alpha} - 1)}{(2-\alpha)(r_N^{1-\alpha} - 1)} r_N^{i-1-N/2}. \end{aligned} \tag{2.3}$$

The Lifted Heston calibrates around twenty times faster than the Rough Heston (Abi Jaber (2019)), and its Markovian, semimartingale nature allow it to be used straightforwardly for simulation. This gives a qualitative advantage that the Rough Heston does not enjoy, the ability to use Monte Carlo methods to price exotic options.

Chapter 3

Numerical Methods

We are comparing the Lifted Heston to the standard Heston along two dimensions. We want to know how accurate each model is, and how much computation time this accuracy requires. To measure time, it is simple enough to time how long each algorithm takes to run. With regards to accuracy, recall the definitions of *implied volatility* and *implied volatility skew* (1.2). We will compare qualitatively how well each model captures the explosion of this skew as maturities become short.

To calculate the skew, we need to be able to compute implied volatilities. In order to do this, we need to be able to price a vanilla option under our models. We choose to use call options in all that follows, but the same analysis could be done using put options instead.

3.1 Cosine Method

We primarily use the cosine method of Fang and Oosterlee (2008) to price our options. It has the advantage of being able to price for sets of strikes K very efficiently, as it splits the price into a model specific component and an option specific component, with only the latter varying with K. In the case of a call, this component can be calculated analytically. This makes it almost free to price additional calls, as long as only the strike K is varying. It is worth noting that while the cos method in Fang and Oosterlee (2008) works with distributions in terms of $\log(S_T/K)$, we will work with $\log(S_T)$ for convenience. This amounts to scaling by a constant and does not complicate the method in any way. We present our extrapolation on a summary of the method from McWalter (2019b) below.

Assume that we have a vanilla option with maturity T and strike K and some constant risk-free interest rate r. If we then let $s_T = \log(S_T)$, we can consider the payoff distribution $v(s_T)$ and the density of the log stock price q_{s_T}. It is now

possible to write the price at $t = 0$ as the discounted expectation

$$\begin{aligned} V &= e^{-rT} \int_{-\infty}^{\infty} v(s) q_{s_T}(s) ds \\ &\approx e^{-rT} \int_{a}^{b} v(s) q_{s_T}(s) ds, \end{aligned} \tag{3.1}$$

for a suitable choice of range $[a, b]$.

We often do not know the distribution q_{s_T}, but do know the characteristic function. We will use cosine expansions to move between the two. On the interval $[a, b]$, we can expand as

$$q_{s_T}(s) = \sum_{k=0}^{\infty}{}' A_k \cos\left(k\pi \frac{s-a}{b-a}\right),$$

where the primed summation denotes that the first term is scaled by a half. The Fourier-cosine coefficients A_k for $k \in \mathbb{N}$ are then given by

$$A_k = \frac{2}{b-a} \int_a^b q_{s_T}(s) \cos\left(k\pi \frac{s-a}{b-a}\right) ds, \tag{3.2}$$

we can then express the cosine as the real part of a complex exponent

$$= \frac{2}{b-a} \int_a^b q_{s_T}(s) \text{Re}\{e^{ik\pi \frac{s-a}{b-a}}\} ds,$$

and seperating s allows us to simplify the integral as

$$\begin{aligned} &= \frac{2}{b-a} \int_a^b q_{s_T}(s) \text{Re}\{e^{i\frac{k\pi}{b-a}s} e^{-ik\pi \frac{a}{b-a}}\} ds \\ &= \frac{2}{b-a} \text{Re}\left\{ \left(\int_a^b q_{s_T}(s) e^{i\frac{k\pi}{b-a}s} ds \right) e^{-ik\pi \frac{a}{b-a}} \right\}, \end{aligned}$$

which for a good choice of a and b will approximate the characteristic function

$$\begin{aligned} &\approx \frac{2}{b-a} \text{Re}\left\{ \left(\int_{-\infty}^{\infty} q_{s_T}(s) e^{i\frac{k\pi}{b-a}s} ds \right) e^{-ik\pi \frac{a}{b-a}} \right\} \\ &= \frac{2}{b-a} \text{Re}\left\{ \phi_{s_T}\left(\frac{k\pi}{b-a}\right) e^{-ik\pi \frac{a}{b-a}} \right\}. \end{aligned} \tag{3.3}$$

This lets us move from working with the distribution q_{s_T} which we may not know, to the characteristic ϕ_{s_T} which we often do know. We can now write the price of the option (3.1) as

$$\begin{aligned} V &\approx e^{-rT} \int_a^b v(s) \sum_{k=0}^{\infty}{}' A_k \cos\left(k\pi \frac{s-a}{b-a}\right) ds \\ &= e^{-rT} \sum_{k=0}^{\infty}{}' A_k \int_a^b v(s) \cos\left(k\pi \frac{s-a}{b-a}\right) ds. \end{aligned}$$

It is now useful to define

$$v_k = \frac{2}{b-a}\int_a^b v(s)\cos\left(k\pi\frac{s-a}{b-a}\right)ds, \tag{3.4}$$

for $k \in \mathbb{N}$. Expanding A_k using (3.3), we have the cosine formula

$$V \approx e^{-rT}\sum_{k=0}^{N_{cos}-1}{}'\mathrm{Re}\left\{\phi_{s_T}(\frac{k\pi}{b-a})e^{-ik\pi\frac{a}{b-a}}\right\}v_n, \tag{3.5}$$

where we have concatenated the sum after N_{cos} terms. It is critical to ensure that sufficiently many terms are used when applying this method. In our own experiments, we keep track of the furthest term that changes the price by more than one one-hundredth. If this number is close to N_{cos}, it is prudent to increase the number of terms used in the expansion.

We can now see how the cosine method conveniently splits the problem of pricing into two separate parts. There is the option specific component v_k given by (3.4) and the process specific component inside the brackets in (3.5), the latter of which is mainly a characteristic function evaluation. We now go into detail for the option component of a call, and the process components for the Heston and Lifted Heston.

3.1.1 Call Payoff Component

Let $s_T = \log(S_T)$ and the payoff of our call option be $v(s_T)$. Then we have

$$\begin{aligned} v(s_T) &= (S_T - K)^+ \\ &= (e^{s_T} - K)^+, \end{aligned}$$

and so (3.4) becomes

$$v_k = \frac{2}{b-a}\int_a^b (e^{s_T}-K)^+\cos\left(k\pi\frac{s-a}{b-a}\right)ds,$$

which since the payoff is zero if $s_T < \log K$

$$= \frac{2}{b-a}\int_{\log K}^b (e^{s_T}-K)\cos\left(k\pi\frac{s-a}{b-a}\right)ds. \tag{3.6}$$

Note that a good choice of a should be far less than $\log K$ as we want to approximate the range $[-\infty, \infty]$. Originally we chose the bounds following the definitions in terms of cumulants suggested in Fang and Oosterlee (2008). However, this resulted in the cos method diverging, even when greatly increasing the number of terms in the cosine series (N=10000 vs N=300 used in later results). Through trial and error we arrived at a heuristic that seemed stable for this model. Our rule of thumb is to

use a range $[\log(S_0/h), \log(S_0 h)]$ for some constant $h \in \mathbb{R}$ and starting stock price S_0. In our experiments we use $h = 100$, which corresponds to the stock being able to increase or decrease a hundredfold. Larger values of h will allow a more accurate price at a higher time-cost[1], however, we did not notice much change in the price even for h as large as 100000.

It is now useful to introduce cosine series for $g(s) = e^s$ and $g(s) = 1$ on a subrange[2] $[c, d]$ of $[a, b]$ in order to simplify (3.6). First, letting $g(s) = e^s$, we have cosine coefficients of

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

and for $g(s) = 1$ we have cosine coeffiecients of

$$\psi_k(c,d) = \int_c^d \cos\left(k\pi \frac{s-a}{b-a}\right) ds.$$

These have analytical solutions given by

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

and

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

These expressions can be implemented straightforwardly in code and allow us to simplify (3.6) to the analytic expression

$$v_k = \frac{2}{b-a}\left[\chi_k(\log(K), b) - K\psi_k(\log(K), b)\right],$$

for $k \in \mathbb{N}$.

[1] Often this is because more terms are needed for the cosine series to converge for a larger range, which in turn requires more activations of the characteristic function.

[2] To keep in line with (3.2), you can think of these functions as being $g(s) = \frac{b-a}{2} e^s \mathbb{I}_{c \le s \le d}$ and $g(s) = \frac{b-a}{2}\mathbb{I}_{c \le s \le d}$ in order to have χ and ψ as their cosine coefficients over $[a, b]$, but this is a technicality that needlessly complicates the analysis — the integrals in question will solve in the same way.

3.1.2 Heston Characteristic

As part of his seminal paper, Heston (1993) provides a characteristic function for the Heston model. In our computations, we instead use the numerically robust specification given by Albrecher *et al.* (2007). For consistency with our implementation, we use the notation given by McWalter (2019a) when presenting this specification.

Again, let $s_T = \log(S_T)$ be the log stock price under the risk-neutral measure. Recall from (1.1) that a Heston model is parameterised by $(V_0, r, \nu, \lambda, \theta, \rho)$ under the risk-neutral measure. Then the characteristic function $\phi_{s_T}(u) = \mathbb{E}\left[e^{ius_T}|\mathcal{F}_t\right]$ can be represented as

$$\phi_{s_T}(u) = \exp\left(C + DV_t + iu\log(S_t)\right),$$

where

$$\begin{aligned} C &= rTiu + \theta\lambda\left(Tx_- - \frac{1}{a}\log\left(\frac{1-ge^{-Td}}{1-g}\right)\right), \\ D &= \frac{1-e^{-Td}}{1-ge^{-Td}}x_-, \end{aligned}$$

with

$$a = \frac{\nu^2}{2}, \quad b = \lambda - \rho\nu i\mu, \quad c = -\frac{\mu^2 + i\mu}{2}, \quad d = \sqrt{b^2 - 4ac},$$

$$x_\pm = \frac{b \pm d}{2a} \quad \text{and} \quad g = \frac{x_-}{x_+}.$$

Here we are still assuming that $r = 0$. This simplifies our equations by removing the first term of C.

One of the appeals of the Heston model is that its characteristic function can be calculated analytically. This makes pricing routines like the cosine method very fast, as their computation time is usually dominated by the time taken to evaluate the characteristic function in each term.

3.1.3 Lifted Heston Characteristic

For the Lifted Heston, we turn again to Abi Jaber (2019). He provides the Fourier-Laplace transform $L(u) = \mathbb{E}\left[e^{us_T}|\mathcal{F}_t\right]$, which can be used to compute the characteristic function simply by making the substitution $\phi_{s_T}(u) = L(iu)$. We now unwrap this specification. Recall from (2.1) that a Lifted Heston model with N mean-reverting factors for some fixed $N \in \mathbb{N}$ is parameterised by $(V_0, \nu, \lambda, \theta, \rho)$ and either

(r_N, H)[3] or $(x_i, c_i), i = 1, \ldots, N$. The transform L is given by

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

where ψ_i from $i = 1, \ldots, N$ is the solution to an N-dimensional system of Riccati equations with dynamics:

$$\begin{aligned}(\psi^i(s))' &= -x_i\psi^i(s) + F(u, \sum_{j=1}^{N} c_j \psi^j(s)), \qquad (3.7)\\ \psi^i(0) &= 0, \qquad i = 1, \ldots, N.\end{aligned}$$

The solutions to these equations are then used to compute $\phi(t,T)$. We will explain how we compute these at the end of the section. Given our Riccati solution, we then need to introduce two non-linear functions F and g_0 given by

$$F(u,z) = \frac{1}{2}(u^2 - u) + (\rho\nu u - \lambda)z + \frac{\nu^2}{2}z^2, \qquad (3.8)$$

$$\begin{aligned}g_0(t) &= V_0 + \lambda\theta \sum_{i=1}^{N} c_i \int_0^t e^{-x_i(t-s)} ds\\ &= V_0 + \lambda\theta \sum_{i=1}^{N} \frac{c_i}{x_i}(1 - e^{-x_i t}), \qquad (3.9)\end{aligned}$$

for $x_i \neq 0$. Note that when $x_i = 0$, such as for the case $N = 1, c = 1, x = 0$ that reproduces the Heston, the integral reduces to $\int_0^t 1 ds = t$. In this special case we have

$$g_0(t) = V_0 + \lambda\theta t.$$

The first input into (3.8) is always the point we are evaluating the characteristic function at, and the second is a weighted sum of the Riccati components. Note that this uses the same weights c_i present in the dynamics (2.1) of the Lifted Heston. Abi Jaber (2019) shows that the function $g_0(t)$ (3.9) can be used to fit an initial volatility structure when calibrating, and while more general formulations are possible, this version is sufficient for our purposes. This is the same g_0 that appears in the dynamics (2.1). We can now define $\phi(t,T)$ as

$$\phi(t,T) = \int_0^{T-t} F(u, \sum_{j=1}^{N} c_j \psi^j) g_0(T-s) ds.$$

Notably, for $t = 0$ we have that each mean reverting factor $U^i = 0$ and so

$$L(u) = \mathbb{E}[\exp(u\log(S_T)|\mathcal{F}_0) = \exp\left(\phi(0,T) + u\log(S_0)\right)$$

[3] see (2.3)

Simplifying $\phi(t,T)$

It is possible to simplify $\phi(t,T)$ further in general, such that it does not involve F. We start by breaking $g_0(t)$ into a constant and a function of t.

$$\begin{aligned} g_0(t) &= V_0 + \lambda\theta \sum_{i=1}^{N} \frac{c_i(1-e^{-x_i t})}{x_i} \\ &= (V_0 + \lambda\theta \sum_{i=1}^{N} \frac{c_i}{x_i}) - \lambda\theta \sum_{i=1}^{N} \frac{c_i e^{-x_i t}}{x_i} \\ &= A + B \sum_{i=1}^{N} \frac{c_i e^{-x_i t}}{x_i} \end{aligned}$$

Now we manipulate the Ricatti equations:

$$(\psi^i)' = -x_i\psi^i + F(u, \sum_{i=1}^{N} c_i\psi^i),$$

can be arranged with an integrating factor as

$$\frac{d}{dt}\left(\psi^i(t)e^{-x_i(T-t)}\right) = e^{-x_i(T-t)}F(u, \sum_{i=1}^{N} c_i\psi^i),$$

or

$$F(u, \sum_{i=1}^{N} c_i\psi^i) = e^{x_i(T-t)}\frac{d}{dt}\left(\psi^i(t)e^{-x_i(T-t)}\right), \tag{3.10}$$

for every $i \in \{1,\dots,N\}$. So now

$$\begin{aligned} \phi(t,T) &= \int_0^{T-t} F(u, \sum_{i=1}^{N} c_i\psi^i(s))g_0(T-s)ds \\ &= A\int_0^{T-t} F(u, \sum_{i=1}^{N} c_i\psi^i(s))ds + B\sum_{i=1}^{N} \frac{c_i \int_0^{T-t} F(u, \sum_{i=1}^{N} c_i\psi^i(s))e^{-x_i(T-s)}}{x_i}ds. \end{aligned} \tag{3.11}$$

Using (3.10) and the fundamental theorem of calculus we get the second term of (3.11) to simplify to

$$B\sum_{i=1}^{N} \frac{c_i\psi^i(T-t)e^{-x_i t}}{x_i}.$$

Using (3.10) and integration by parts, the first term of (3.11) becomes

$$\begin{aligned}
A\int_0^{T-t} F(u,\sum_{i=1}^N c_i\psi^i(s))ds &= \left(V_0+\lambda\theta\sum_{i=1}^N\frac{c_i}{x_i}\right)\int_0^{T-t}F(u,\sum_{i=1}^N c_i\psi^i(s))ds \\
&= \left(\sum_{i=1}^N\frac{V_0}{N}+\lambda\theta\frac{c_i}{x_i}\right)\int_0^{T-t}F(u,\sum_{i=1}^N c_i\psi^i(s))ds \\
&= \sum_{i=1}^N\left(\frac{V_0}{N}+\lambda\theta\frac{c_i}{x_i}\right)\int_0^{T-t}e^{x_i(T-t)}\frac{d}{dt}\left(\psi^i(t)e^{-x_i(T-t)}\right)ds \\
&= \sum_{i=1}^N\left(\frac{V_0}{N}+\lambda\theta\frac{c_i}{x_i}\right)\left(\psi^i(T-t)+\int_0^{T-t}x_i\psi^i(s)ds\right).
\end{aligned}$$

Putting these results back into (3.11), we get

$$\begin{aligned}
\phi(t,T) &= \sum_{i=1}^N\left(\frac{V_0}{N}+\lambda\theta\frac{c_i}{x_i}\right)\left(\psi^i(T-t)+\int_0^{T-t}x_i\psi^i(s)ds\right)-\lambda\theta\frac{c_i\psi^i(T-t)e^{-x_it}}{x_i} \\
&= \lambda\theta\sum_{i=1}^N\frac{c_i}{x_i}\left[\left(1+\frac{x_iV_0}{c_i\lambda\theta N}\right)\left(\psi^i(T-t)+\int_0^{T-t}x_i\psi^i(s)ds\right)-\psi^i(T-t)e^{-x_it}\right].
\end{aligned}$$

Numerical Riccati

We make use of Abi Jaber (2019) Appendix A.3 for an explicit-implicit discretisation scheme of (3.7). This works by initializing all N components to zero, and then evolving over a discrete number of evenly spaced time-steps according to

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

We use 1000 time-steps in all of our experiments instead of the 300 used in Abi Jaber (2019). This was necessary in order to have enough accuracy for the skew of the Heston embedded in the Lifted Heston ($N=1, x=0, c=1$) to agree with that of a standard Heston (See figure (4.2)). Unfortunately, we have to solve this set of N Riccati equations every time we evaluate the characteristic function, as their dynamics are dependent on the point u being evaluated.

3.2 Monte Carlo

We are unable to use standard Monte-Carlo techniques to simulate the variance process of a Rough Heston model. One of the main reasons for this is that the variance process is non-Markovian (El Euch and Rosenbaum (2019)).

In contrast to this, the Lifted Heston is a Markovian and semimartingale model. Specifically, that is to say that the variance process is itself Markovian, and the stock price is a semimartingale. Using this fact, Abi Jaber (2019) has adapted standard Euler-Maruyama schemes to simulate the variance process at each time step, and use this updated variance to simulate the stock price for that time point.

Let $(V_0, \nu, \lambda, \theta, \rho, r_N, H)$ be the parameters of the Lifted Heston model in question and fix $N \in \mathbb{N}$ for the number of mean-reverting factors. Also let

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

We then use following algorithm:

- Choose some suitable large $n \in \mathbb{N}$ and number of time-steps m.
- Initialize $S = S_0$, $V = V_0$ and $U^i = 0$ for $i = 1, \dots, N$.
- For each time-step t_j for $j = 1, \dots, m$:
  - Let Z be a sample from a Normal distribution with covariance Σ, with elements Z_1 and Z_2.
  - Update S from the previous variance as $S_{t_j} = S_{t_{j-1}} e^{(r - \frac{1}{2} V^+_{t_{j-1}})\Delta t + \sqrt{V^+_{t_{j-1}} \Delta t} Z_1}$.
  - Update the vector of mean-reverting factors U as $U_{t_j} = \frac{1}{x_i} . (U_{t_{j-1}} - \lambda V_{t_{j-1}})\Delta t + \nu \sqrt{V^+_{t_{j-1}} \Delta t} Z_2$ where . denotes element-wise operation.
  - Update the variance process as $V_{t_j} = g_0(t_j) + \sum_{i=1}^N c_i U^i_{t_j}$
- Calculate and store the final discounted payoff of the option from S_T.
- Average the above over the n simulations.

The main drawback of using Monte-Carlo methods is that we cannot simulate a range of strikes K cheaply like with the cosine method. This makes it a lot slower to compute an implied volatility surface. It is useful, however, when we want to price more exotic options for which we cannot so easily compute the v_k from (3.4).

Chapter 4

Results

We now move on to some numerical experiments. First we test the assertion that the Lifted Heston accurately reproduces the standard Heston under the parameterisation $N = 1$, $x_1 = 0$, $c_1 = 1$. Then we explore the effects of changing the various parameters of the standard Heston on its skew and show that it cannot satisfactorily recover the explosive nature of the Rough Heston at low maturity. We then show how the skew of the Lifted Heston changes with increased number of mean-reverting factors, and analyze a cross section of implied volatilities at maturities of one week and one year. We proceed to show the effects of the Lifted Heston's parameters for a particular N. We then finish our results on the cosine method with a simple empirical attempt to understand the time complexity as N increases and better understand the trade-off with accuracy. Finally, a Monte-Carlo simulation of price is shown for completion.

4.1 Embedded Heston

Here we compare a Lifted Heston with a single mean-reverting term, with $x_1 = 0$ and $c_1 = 1$, to the standard Heston that it theoretically agrees with. Here our parameter set is given by

S_0	V_0	θ	λ	ν	ρ	N_{cos}
100	0.02	0.02	0.3	0.3	-0.7	300

and additionally our log-money strikes $k = \log(K/S_0)$ come from the Matlab command \linspace(-1.2,0.2,80).

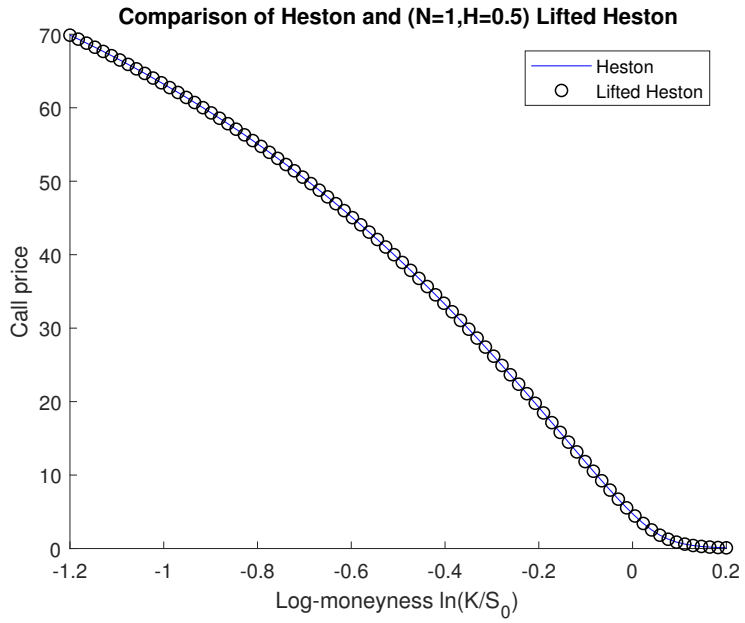

Fig. 4.1: Call prices for Heston and Heston-reproducing Lifted

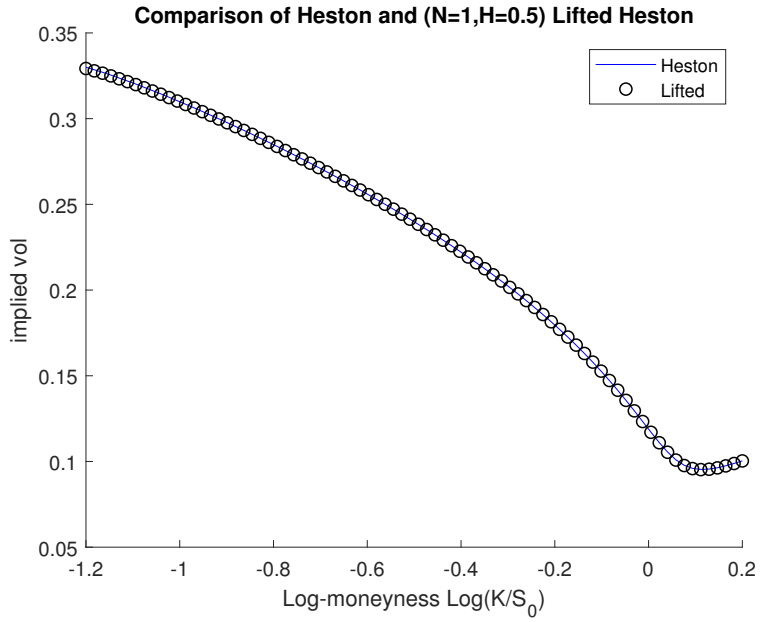

Fig. 4.2: Implied volatilities for Heston and Heston-reproducing Lifted

From figures 4.1 and 4.2 we can see that both prices and implied volatilities are essentially the same. The maximum error in 4.2 is 0.0057, and it seems that the Lifted Heston is accurately reproducing the standard Heston.

It is worth taking a moment here to reflect on the nature of this reproduction. It is only in the case that both $N = 1$ and the parameters x_1 and c_1 are chosen to be 0 and 1 that we attain the Heston. In this case, from (2.2) we have

$$g_0(t) = V_0 + \lambda\theta t,$$

which makes the variance process

$$V_t = g_0(t) + U_1.$$

Written in dynamics notation this is

$$\begin{aligned} dV_t &= dg_0(t) + dU_1 \\ &= \lambda\theta dt + (-0U_1 - \lambda V_t)dt + \nu\sqrt{V_t}dW_t \\ &= \lambda(\theta - V_t)dt + \nu\sqrt{V_t}dW_t, \end{aligned}$$

which matches the dynamics of the standard Heston (1.1). Now in our experiments we follows Abi Jaber (2019) in using the (H, r_N) parameterisation of (x_i, c_i) with $r_N = 1 + 10N^{-0.9}$ in order to have one extra parameter instead of $2N$ extra parameters. The main consequence of this here, is that $x_1 \neq 0$ and c_1 will likely not equal one. This changes the above analysis. We now have

$$g_0(t) = V_0 + \lambda\theta\frac{c_1}{x_1}(1 - e^{-x_1 t}),$$

which makes the variance process

$$V_t = g_0(t) + c_1 U_1.$$

This time the dynamics are

$$\begin{aligned} dV_t &= dg_0(t) + c_1 dU_1 \\ &= \lambda\theta c_1 e^{-x_1 t}dt + c_1(-x_1 U_1 - \lambda V_t)dt + \nu\sqrt{V_t}dW_t \\ &= c_1\lambda(\theta e^{-x_1 t} - V_t)dt - c_1 x_1 U_1 dt + \nu\sqrt{V_t}dW_t, \end{aligned}$$

It is not clear that this should correspond to any particular Heston model. While we could define a new mean reversion rate $\lambda' = c_1\lambda$, the new long term level $\theta' = \theta e^{-x_1 t}$ would have to be time-dependent. Even if we do make this extension, it is not obvious how to reconcile the presence of the extra drift term $-c_1 x_1 U_1 dt$. It is for this reason that we require $x_1 = 0$ when reproducing the Heston. Since this is incompatible with our (H, r_N) parameterisation, we do not include a Heston base case when plotting the skew of Lifted Heston models in our experiments, as it is not clear what would be a fair and natural comparison.

4.2 Skew - Heston

We have claimed that the Heston model cannot reproduce the explosion of skew seen in the short maturity end of the market. Here we will plot the skew for a Heston as figure 4.3 and show how changing parameters affects that plot. Our parameters are as before, but with $N_{cos} = 1000$. Our maturities are the range $T_i = i/52, i = 1, \ldots, 52$, ranging in weeks from one week to a year. In general we notice an increase in the number of cos method terms needed, the shorter the maturity is. Since the Heston is so fast, we have stuck to 1000 throughout. It would be possible to accurately price at most of the maturities with 300 terms.

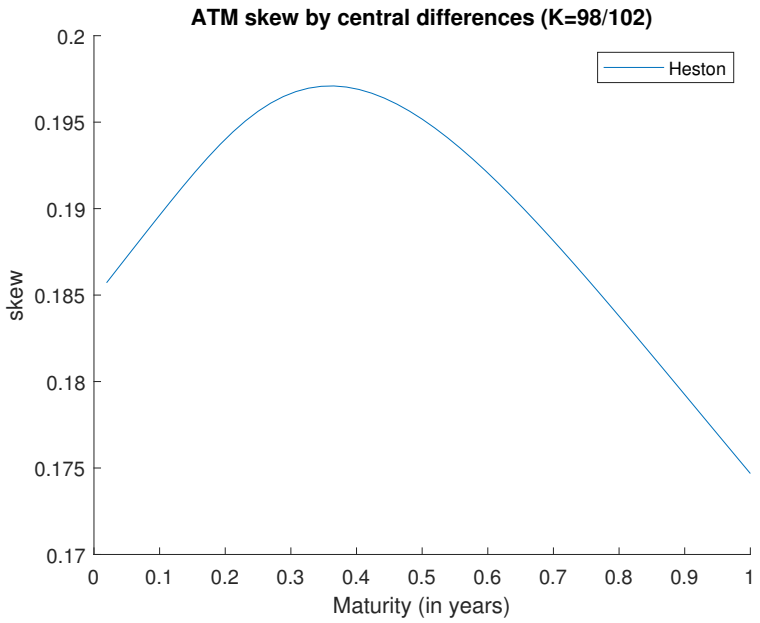

Fig. 4.3: Implied skew for a standard Heston model

Notice how the skew peeks at around $T = 0.4$, and overall only differs by at most $.025$ over the first year. It is this kind of complete lack of short maturity explosion that leads us to investigate rougher models. Let us now see if varying any of the parameters can give us the explosive behavior we want.

4.2.1 Effects of θ

Here we vary θ. We plot five skew curves as figure 4.4, using the set

$$\theta \in \{0.002, 0.0067, 0.02, 0.06, 0.2\},$$

and plotting the skew for each value.

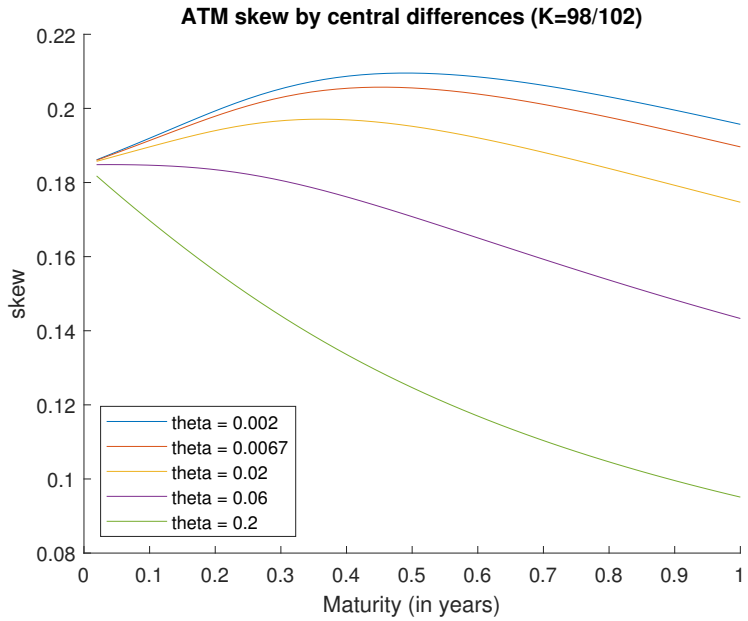

Fig. 4.4: Implied skew for Heston varying θ

We can see θ having an effect on the convexity of the skew curve. For the largest value the skew curve appears convex, becoming more and more concave as θ decreases. While a high θ could achieve the increasing skew we want as maturity decreases, it still seems to gentle compared with the market skew curve found in Abi Jaber (2019).

4.2.2 Effects of λ

In similar fashion, we plot five skew curves as figure 4.5 using

$$\lambda \in \{0.03, 0.1, 0.3, 0.9, 3\}.$$

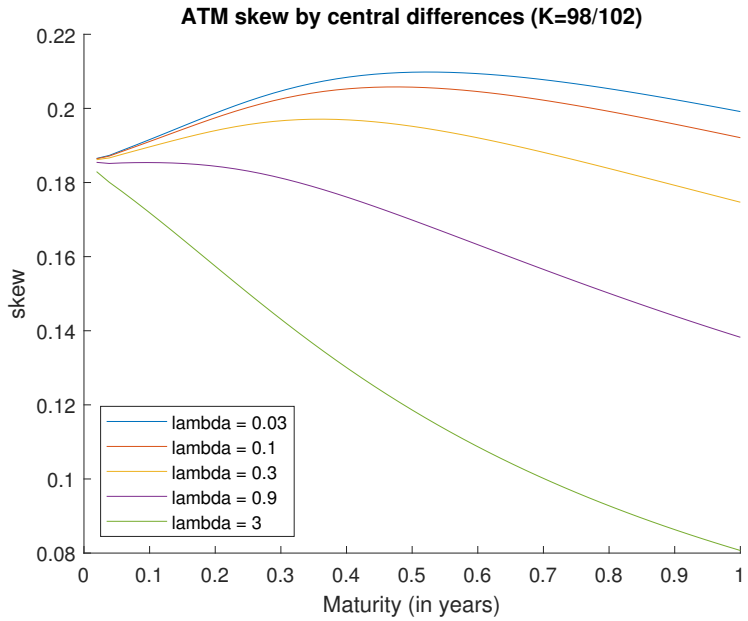

Fig. 4.5: Implied skew for Heston varying λ

The effect here seems similar, if slightly more pronounced. We can see that $\lambda = 3$ has a smaller skew at $T = 1$ than the highest θ plot did, but qualitatively the plots seem very similar.

4.2.3 Effects of ν

We repeat the treatment for

$$\nu \in \{0.03, 0.1, 0.3, 0.9, 3\},$$

as figure 4.6 This plot is the reason we chose a value of $N_{cos} = 1000$. This is necessary for the prices to converge with high ν for the 1 week maturity.

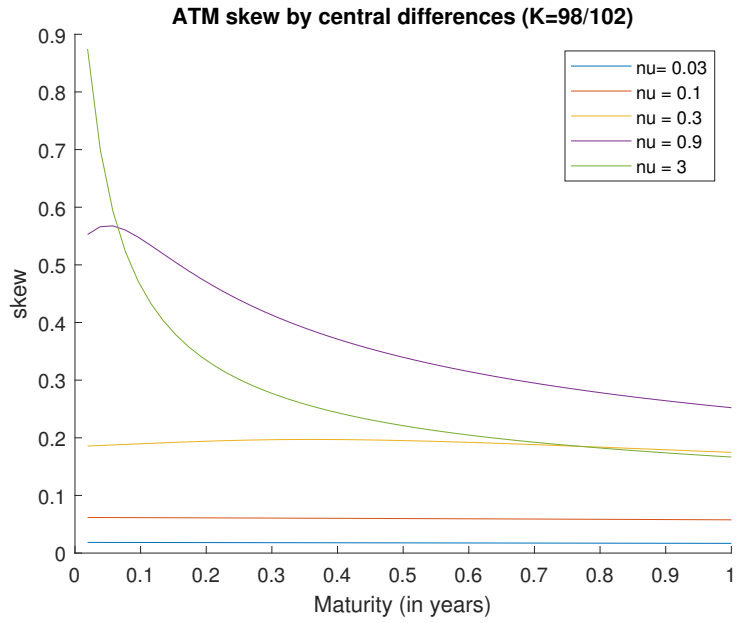

Fig. 4.6: Implied skew for Heston varying ν

Here we actually recover quite explosive skew in the lower maturities. This is somewhat expected, as we are heavily increasing the volatility of variance to 3, which is an unreasonably high amount. Unfortunately this higher volatility has an adverse effect on our fit for later maturities — they are still affected by the increased volatility and so the skew is pulled higher than it ought to be for later maturities. This is perhaps the weakness of the standard Heston, in adapting to one part of the maturity curve it must make sacrifices elsewhere.

4.2.4 Effects of ρ

Finally we try varying rho in figure 4.7. We pick the values

$$\rho \in \{0, -0.4, -0.7, -0.8, -1, \}$$

as our candidates. We are constrained to the range $[-1, 1]$ as these are correlations. Furthermore, since stock prices and volatilities have been observed to be negatively correlated, Bouchaud *et al.* (2001), we restrict ourselves to $\rho \leq 0$.

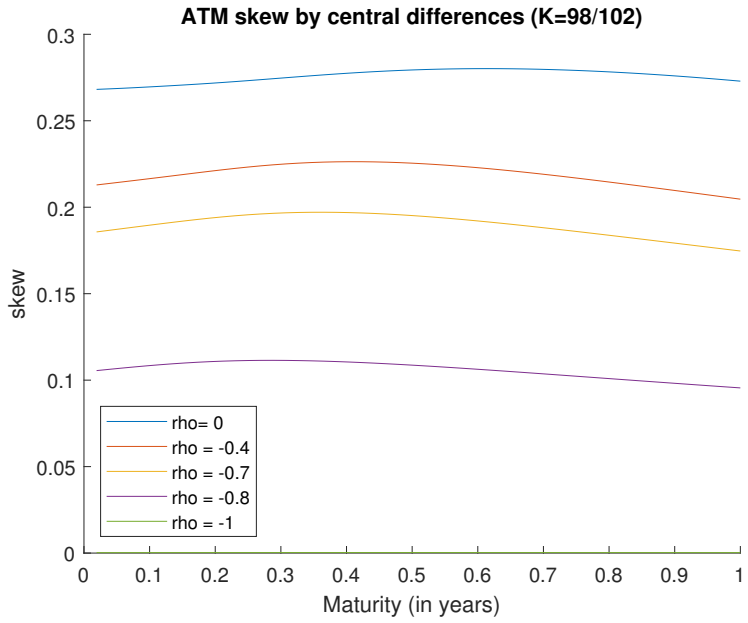

Fig. 4.7: Implied skew for Heston varying ρ

The overall level of skew seems to be controlled by ρ. A zero correlation results in the highest skew, with increasingly negative correlation bringing the skew further and further down in a level manner.

4.3 Skew - Varying N

None of our Heston parameters gave us a satisfactory short-maturity explosive behavior, without being intrinsically tied drawbacks elsewhere along the curve. We now turn to the Lifted Heston model, and plot its skew for varying values of N. Throughout this we use a Hurst index of 0.1, and our parameter r_N is calculated as $r_N = 1 + 10N^{-0.9}$, following the example of Abi Jaber (2019). Since the Lifted Heston takes between 10 to 1000 times longer than the standard Heston, depending on N, we limit ourselves to the smaller time set given by

$$T \in \{3/365, 4/365, 5/365, 6/365, 1/52, 2/52, 3/52, 1/12, 2/12, 3/12, 6/12, 9/12, 1\}.$$

We let N_{cos} range from 1000 to 300 as appropriate. The exact numbers used are

$$\{1000, 1000, 1000, 1000, 1000, 1000, 1000, 600, 600, 300, 300, 300, 300\},$$

corresponding to the maturities above.

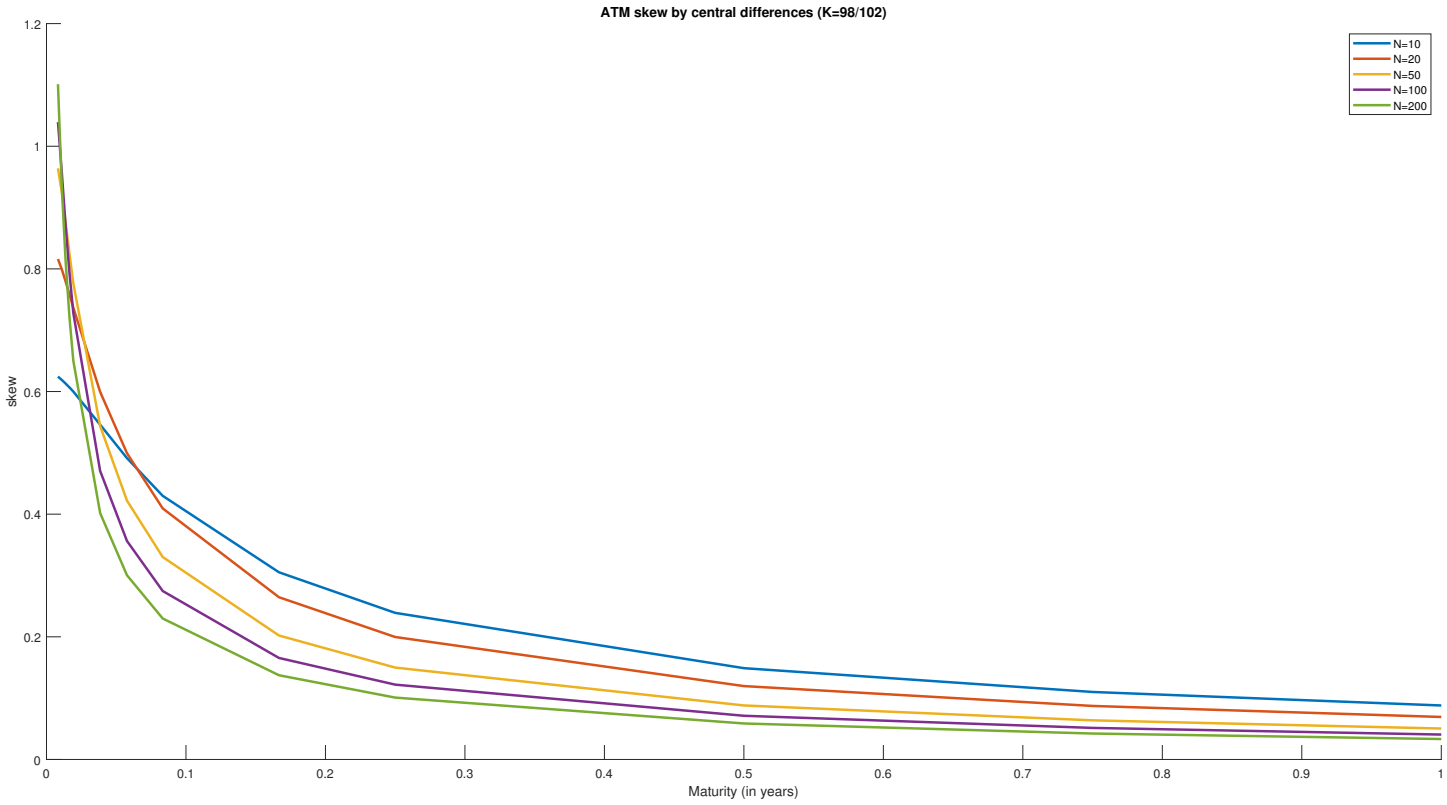

Fig. 4.8: Implied skew for Lifted Heston with varying N

We see in figure 4.8 that as N increases, the slope of the skew at low maturities rapidly increases, while for longer maturities it remains relatively flat. This matches the qualitative behavior we were hoping to see. For all but the early maturities, skew is decreasing with N. It appears to be decreasing less and less despite effectively doubling N. This may be a sign that it is converging to a Rough Heston model. We would suggest, however, that it is prudent to first pick a value of N and then to calibrate to the market. If the explosive qualities are insufficient and a larger N is required, we would advise recalibrating to the market to avoid this level decrease. We will now look at two specific cross-sections of the implied volatility surface.

4.3.1 Cross Sections

Maturity 1 Week

First we look at the 1 week cross section of figure 4.9. We can see that further out of the money, the implied volatility increases with N. Slightly in the money, it is difficult to see, but the order is roughly reversed. This results in the slope at the money becoming steeper as N increases, which agrees with what we see in figure 4.8 at the short end, with the skew increasing at the 1 week point with N.

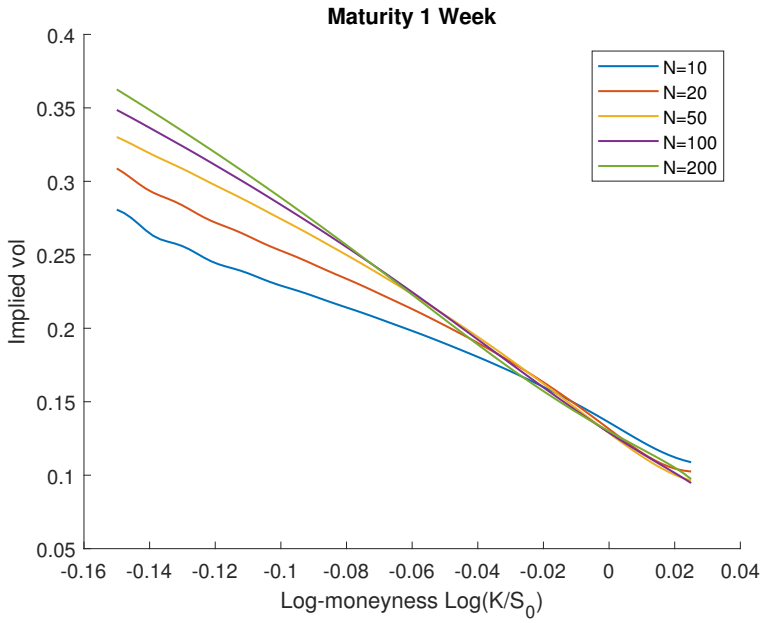

Fig. 4.9: Maturity 1 week implied volatility surface

Maturity 1 Year

At a maturity of 1 year on the other hand, see figure 4.10, the inverse relationship holds. Further out of the money, the implied volatility decreases with N, and slightly in the money it increases. This results in a slope that becomes more gentle as N increases which in turn leads to the decreasing skew we see at the 1 year mark in figure 4.8.

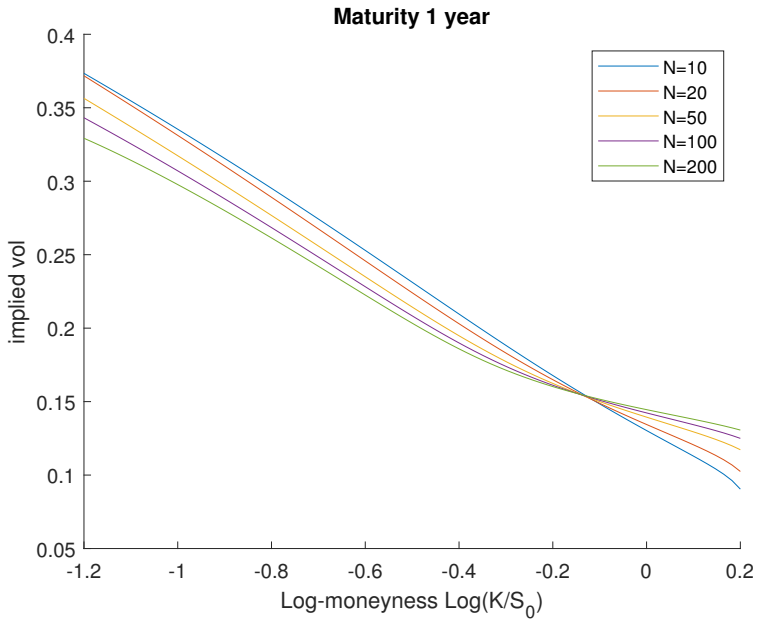

Fig. 4.10: Maturity 1 year implied volatility surface

4.4 N=20 Parameter Effects

We now fix $N = 20$ and repeat the procedure used in section 4.2 to see how changing model parameters affects the skew for the Lifted Heston. The maturities used here are the same as before. We use 1000 cosine terms for the variation of ν and ρ, and 600 for the rest.

4.4.1 Effects of θ

First we vary θ. We plot five skew curves as figure 4.11, using the set

$$\theta \in \{0.002, 0.0067, 0.02, 0.06, 0.2\},$$

and plotting the skew for each value.

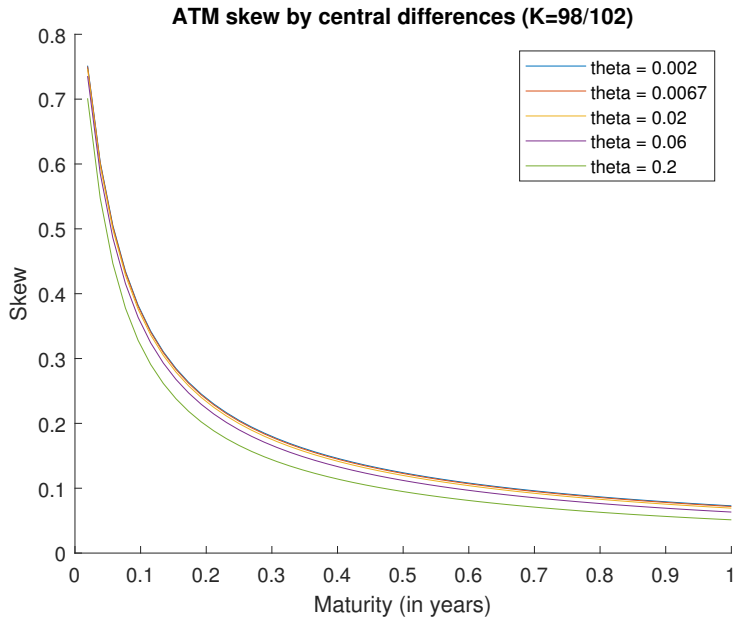

Fig. 4.11: Implied skew for Lifted Heston varying θ

The effect here is quite different to the Heston. Unlike figure 4.4, we see that increasing θ decreases skew across the board, without really affecting convexity.

4.4.2 Effects of λ

Again, we plot five skew curves as figure 4.12 using

$$\lambda \in \{0.03, 0.1, 0.3, 0.9, 3\}.$$

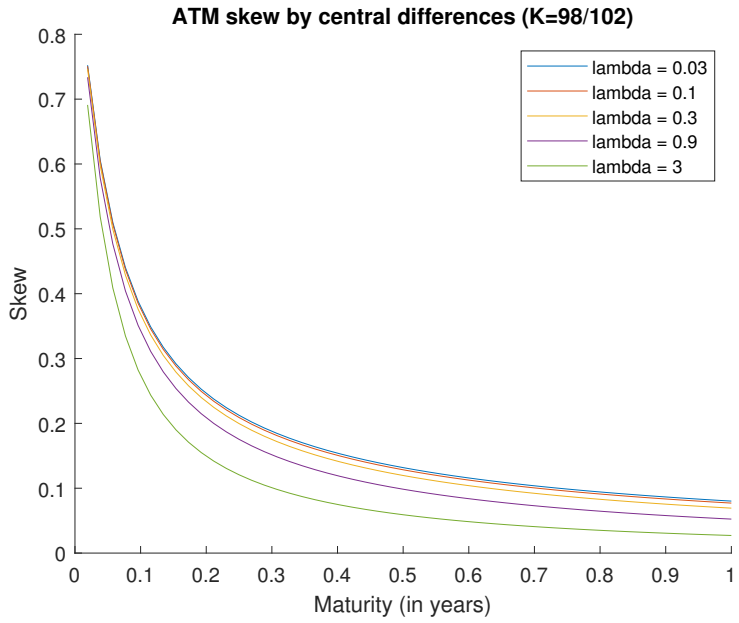

Fig. 4.12: Implied skew for Lifted Heston varying λ

The effect here is pretty much identical, just more pronounced. This matches the relation between θ and λ seen in the standard Heston in figures 4.4 and 4.5. Once more, the difference can mainly be seen at longer maturities.

4.4.3 Effects of ν

As before, we repeat the treatment for

$$\nu \in \{0.03, 0.1, 0.3, 0.9\},$$

as figure 4.13 This time we only plot four curves. The reason for this is that with $\nu = 3$, or even 2, we end up with diverging prices from the cosine method. This is likely due to our choice of range $[a, b] = [S_0/100, 100S_0]$ being to small to capture almost all the possible outcomes for the stock price under such a high volatility parameter.

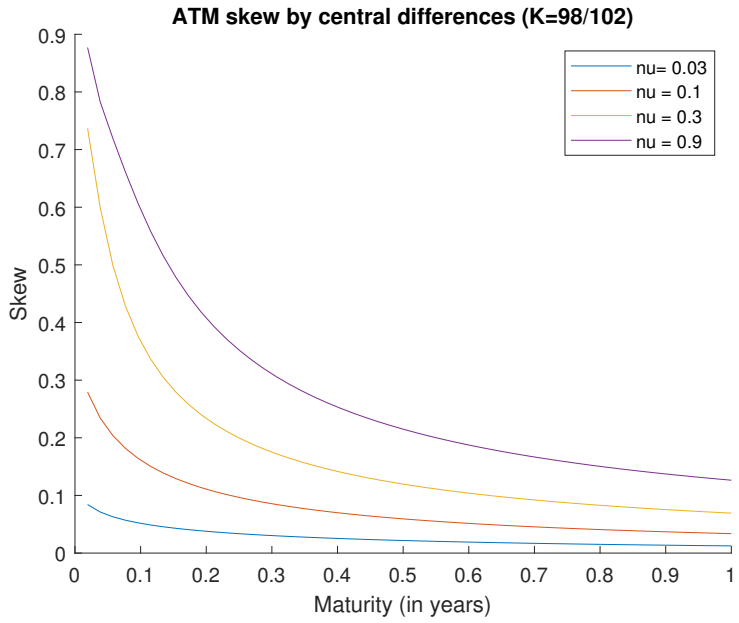

Fig. 4.13: Implied skew for Lifted Heston varying ν

Like the Heston case, see figure 4.6, increasing ν pulls up the slope of the curve at all maturities. Unlike the standard Heston, we can see the that the short-maturity explosion is preserved in the shape of the curve, even at low volatilities of variance.

4.4.4 Effects of ρ

We now go back to varying rho in figure 4.14. Again we use

$$\rho \in \{0, -0.4, -0.7, -0.8, -1, \}$$

as our values.

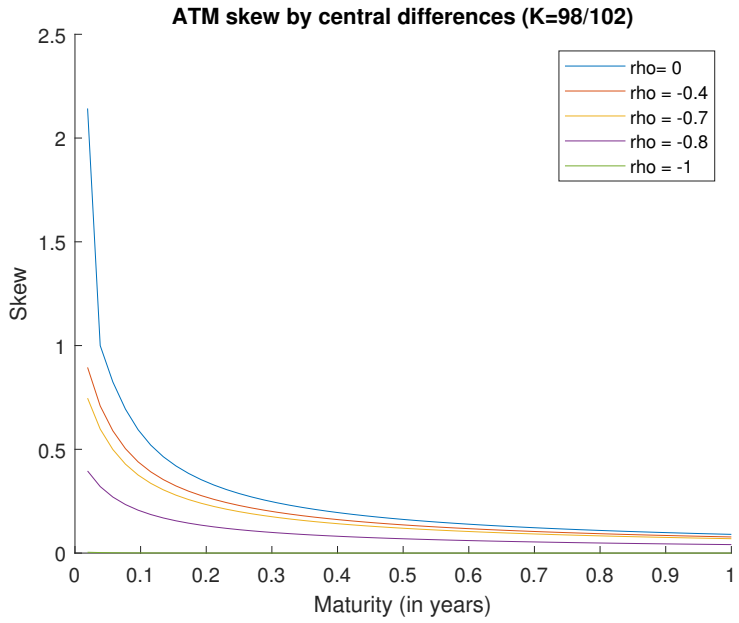

Fig. 4.14: Implied skew for Lifted Heston varying ρ

We see the same decrease in skew as ρ decreases as we do for the Heston (figure 4.7), but this time the effect is greater at short maturities, becoming fairly minimal by the time we reach T of one year. Again, when the Brownian motions are perfectly negatively correlated, we see the skew collapse to zero.

4.4.5 Effects of H

For the Lifted Heston, we have one additional parameter. The Hurst index, $H \in [0, \frac{1}{2}]$, is supposed to control the roughness of the variance process, with and H of zero being very rough and an H of half being as rough as Brownian motion. As such, we expect the explosive effect to be magnified for lower H, while being diminished for H close to $\frac{1}{2}$ (which would be used in the standard Heston case). We use values of

$$H \in \{0, 0.05, 0.1, 0.2, 0.4, \}$$

and plot the resulting curves in figure 4.15

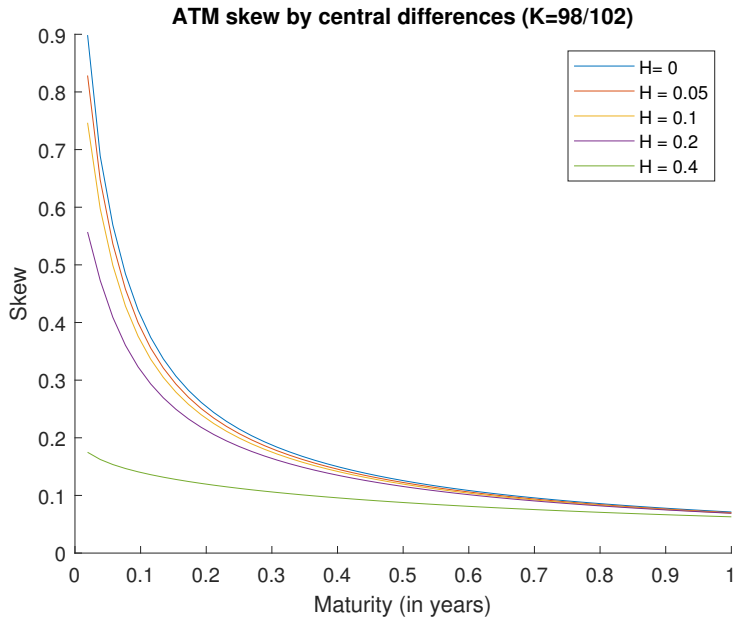

Fig. 4.15: Implied skew for Lifted Heston varying H

We see that for short maturities, the slope of skew increases as H gets closer to zero. Meanwhile, skew at later maturities remains fairly unchanged. This is exactly the behavior that theory leads us to expect, and we can see how we lose the short-maturity explosion as H grows to large.

4.5 Time Complexity

We have seen that the Lifted Heston effectively captures the short-maturity skew explosion commonly seen in the market. We now turn to the question of how much time this takes. To do this, we will fix maturity as $T = 1$ and time how long it takes to price the same set of 80 strikes used in the first experiment. In all cases we use 300 cosine terms. This will be repeated for N ranging from 10 to 200 in intervals of 10. The whole experiment will be repeated for 10 runs, and the times averaged. This is to minimize the effects of background processes on the timing, and to compensate for the bias where earlier runs are slower as the program is loaded into faster levels of memory cache.

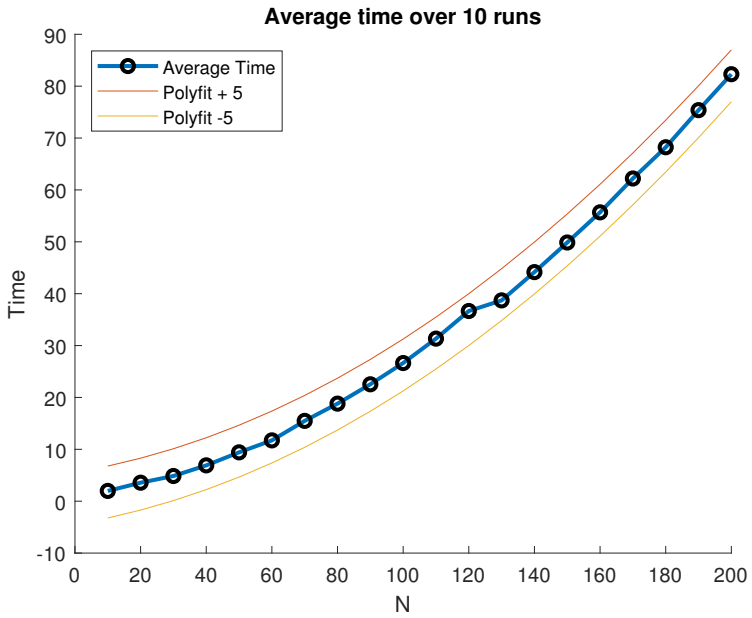

Fig. 4.16: Average time for Lifted Heston of varying N

The average times are plotted in figure 4.16. Additionally, we fit a polynomial to these times using Matlab's \polyfit command. The lowest degree polynomial to offer a reasonable fit is a low-coefficient quadratic, and this is plotted with a plus and minus 5 constant. This seems to bound the times quite well, and we can conclude that at least empirically times appear to scale as $\mathrm{O}(N^2)$.

This seems reasonable. When we increase N, we increase the dimension of the Riccati equations that have to be solved in every characteristic function evaluation, and many parts of the characteristic function and dynamics involve terms that sum N components. A proper theoretical analysis of the time complexity, unfortunately, is out of the scope of this minor dissertation.

One more result worth mentioning here, however, is that the number of cosine terms needed to converge seems to decrease as we increase N. We plot this in figure 4.17, where by last significant term we mean the last term that changed the price by more than one one-hundredth. In all cases, all 300 terms where used. This should partially offset the cost of using a higher N. It does appear, however, that even accounting for the more rapid convergence, the cost of high N quickly becomes prohibitive. This is magnified by considering that we are interested in high N when we are more focused on what is happening at low maturities, and it is at these same low maturities that we tend to need more cosine terms in order to converge anyway.

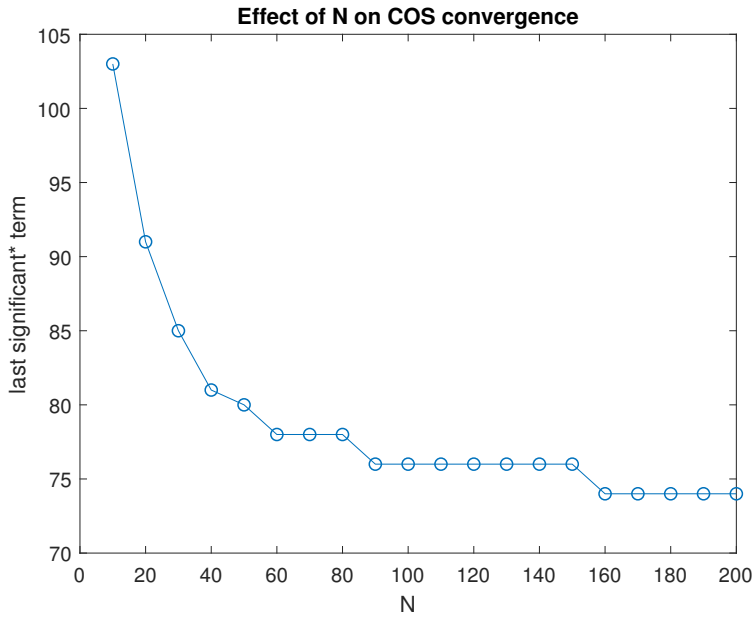

Fig. 4.17: Last term to change price by more than 1 cent.

4.6 Monte Carlo

For completeness, we include a plot of Monte-Carlo prices for a Lifted Heston model. We let $N = 20$ and the other parameters are the same as throughout the dissertation. We use sample sizes of 1000 to 50000 in steps of 1000. We plot the resulting prices in figure 4.18.

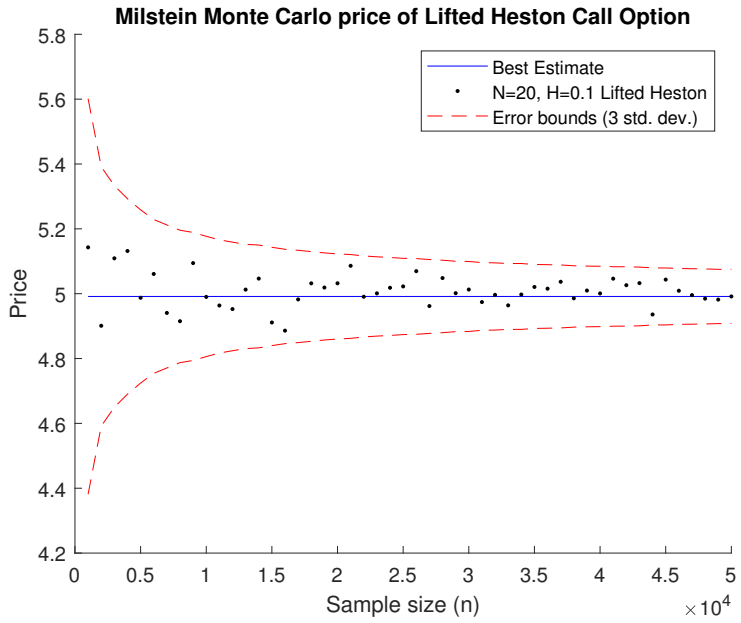

Fig. 4.18: Monte-Carlo pricing with Lifted Heston

We see that the prices stay within the 3 standard deviation error bounds, and converge towards a final price. While Monte-Carlo methods can give accurate prices with a high enough sample size, and come with the benefit of knowing how wide your error bounds are,they are a lot slower than batch methods like the cosine method, even if we are only evaluating a single strike.

Chapter 5

Conclusion

We began by reviewing the flaws of the existing Heston and Rough Heston models. The Heston was numerically fast and mathematically simpler, but unable to replicate the explosive skew [1] observed in the market at short maturities without heavily increasing volatility of variance — a trade-off that greatly mis-priced later maturity options. The Rough Heston was able to capture this explosion without sacrificing later maturities, but was numerically slow and mathematically complex due to its fractional Brownian motion.

We introduced Abi Jaber (2019)'s reduced parameterisation of his Lifted Heston model, and showed that it was able to reconcile the standard Heston and Rough Heston. We first confirmed that it embeds the Heston, by comparing the implied volatilities generated by both models under the equivalent parameters. We then plotted the effects of changing θ, λ, ν and ρ on the skew generated by the Heston, showing that none of these are sufficient to satisfactorily reproduce the explosive skew we desired. Inspired by this, we computed the skew for Lifted Heston models with varying number N of mean-reverting factors. Even at a fairly small N of 20, the explosive skew was achieved for short maturities while still keeping the skew of longer maturities relatively flat. We then explored the effects of changing parameters in the Lifted Heston case, and compared this to the Heston. The effects of each parameter were largely the same, with the exception of θ and λ having an impact on the skew inversely proportional to the maturity. The explosive skew at low maturity scaled inversely with the Hurst index as expected.

Finally we provided an empirical analysis of the time complexity of the Lifted Heston as a function of the number of mean-reverting factors. We observed a relation of order $\mathrm{O}(N^2)$ and so cautioned against using too many factors. A secondary effect is observed in an inverse relationship between number of factors N and the

[1] Here skew is not to be confused with the general shape of a volatility plot. It is defined as the absolute value of the log-moneyness partial derivative of the implied volatility surface, evaluated at the money. $\sigma_{\text{skew}} = |\frac{\partial}{\partial k}\sigma_{\text{implicit}}(k,T)\mid_{k=0}|$ for $k := \ln(K/S_0)$,

number of terms needed by the cosine method to converge to the correct price. This partially offset the numerical cost of using a higher N, but an N of around 20 still seemed to be the sweet spot in the trade-off between explanatory power and numerical efficiency. Even at $N = 20$, the Lifted Heston was around 100 times slower than the standard Heston, and may not have been worth this cost if dealing primarily with options with maturities further out than a year.

While out of the scope of this minor dissertation, there are several possible extensions that could be added. To start, it would be useful to compute the skew for the Rough Heston model, and ensure that it bounds the skew of the Lifted Heston from above. This would also allow for more accurate comments on any convergent behavior of the Lifted Heston. Secondly, all the analysis in this work has been done on preselected parameter sets. It would be instructive to perform the same analysis in terms of models calibrated to actual market data, specifically that from a South African context. This could then be compared to the work of Abi Jaber (2019), calibrated to the S&P index, to see whether the model fairs differently in the local market. Finally, while we have looked at using standard Monte-Carlo methods to price vanilla options under the Lifted Heston model, we have not found evidence in the literature of applying these to more exotic options. While it is quite possible that these methods would behave as they do for simpler models, it would still be worth the exercise in testing this.

Bibliography

Abi Jaber, E. (2019). Lifting the Heston model, *Quantitative Finance* **19**(12): 1995–2013.
URL: *https://doi.org/10.1080/14697688.2019.1615113*

Abi Jaber, E. and El Euch, O. (2019). Multifactor approximation of rough volatility models, *SIAM Journal on Financial Mathematics* **10**(2): 309–349.

Albrecher, H., Mayer, P., Schoutens, W. and Tistaert, J. (2007). The little Heston trap, *Wilmott* **2007**(1): 83–92.

Bayer, C., Friz, P. and Gatheral, J. (2016). Pricing under rough volatility, *Quantitative Finance* **16**(6): 887–904.

Bergomi, L. (2005). Smile dynamics II, *Risk* **18**: 67–73.

Bouchaud, J.-P., Matacz, A. and Potters, M. (2001). Leverage effect in financial markets: The retarded volatility model, *Phys. Rev. Lett.* **87**: 228701.
URL: *https://link.aps.org/doi/10.1103/PhysRevLett.87.228701*

Cont, R. and Tankov, P. (2003). *Financial modelling with jump processes*, Vol. 2, CRC Press, Taylor and Francis Group: Boca Raton, London, New York, Washington.

El Euch, O., Gatheral, J. and Rosenbaum, M. (2018). Roughening Heston, *SSRN* .
URL: *https://ssrn.com/abstract=3116887 or https://dx.doi.org/10.2139/ssrn.3116887.*

El Euch, O. and Rosenbaum, M. (2019). The characteristic function of rough Heston models, *Mathematical Finance* **29**(1): 3–38.

Fang, F. and Oosterlee, C. (2008). A novel pricing method for european options based on Fourier-cosine series expansions., *SIAM J. Sci. Comput.* **31**(2): 826–848.

Fouque, J.-P. and Lorig, M. J. (2011). A fast mean-reverting correction to Heston's stochastic volatility model, *SIAM Journal on Financial Mathematics* **2**(1): 221–254.

Gatheral, J. (2011). *The volatility surface: a practitioner's guide*, Vol. 357 (Wiley Finance Book series), John Wiley & Sons.

Gatheral, J., Jaisson, T. and Rosenbaum, M. (2018). Volatility is rough, *Quantitative Finance* **18**(6): 933–949.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Rev. Financ. Stud* **6**(2): 327–343.

Madan, D. B., Carr, P. P. and Chang, E. C. (1998). The variance gamma process and option pricing, *Review of Finance* **2**(1): 79–105.

McWalter, T. A. (2019a). Numerical methods in finance II — characteristic function pricing & stochastic volatility.

McWalter, T. A. (2019b). Numerical methods in finance II — the cosine method.