Lund University

Faculty of Science, LU

Centre for Mathematical Sciences

Consistent pricing of VIX options

Author

Wilhelm Ålander

wilhelm.alander@gmail.com

Supervisor
Magnus Wiktorsson

February 21, 2020

Contents

Acknowledgements **2**

Abstract **4**

Introduction **5**

Problem description . . . 6
Computing VIX in accordance to the CBOE method . . . 7
Characteristic function . . . 8
The Fourier transform . . . 9
Fourier Gauss-Laguerre . . . 10
Gauss-Laguerre . . . 12
Gold Search Algorithm . . . 12
The 3-dimensional Itô Lemma . . . 13

The models **14**

Cox-Ingersoll-Ross model . . . 14
Multi-dimensional Cox-Ingersoll-Ross model . . . 14
The Heston model . . . 15
Triple Heston model . . . 16
Quintuple Heston model . . . 17
Theoretically calculated VIX for the Heston Model . . . 18

Calibration of the multi-factor Heston model **19**

Calculating the VIX options . . . 20
VIX term structure . . . 20

Results **21**

Discussion **22**

Parameter estimation . . . 22
Model efficiency . . . 23
Stochastic vol-of-vol . . . 23
Fourier optimality . . . 24

Experimental models **25**

General solution using Riccatti ODEs in multivariate affine stochastic volatility models . . . 25
Heston stochastic vol-of-vol . . . 25

Conclusion **26**

References **27**

Appendix **28**

Derivation of Heston characteristic function . . . 28
Approximation of the complex error function . . . 34
Figures . . . 35

Acknowledgements

First, and most of all, I would like to thank my supervisor Associate Professor **Magnus Wiktorsson**, whom with his expertise, assistance, guidance, patience and fruitful discussions throughout has made the creation of this thesis a reality. I would also like to extend my sincere gratitude towards a couple of people,

Gupta, to Ville, Adham and Per-Niklas whom with their greatest intentions of making me procrastinate have made these last couple of months extra fulfilling with fruitful discussions and shared laughs! Especially, thank you Adham for introducing me to your snacking habits.

Andreas Landin, first and foremost the premier reason why I decided to study in Lund. Thanks for showing me around and introducing me to new people, events and especially Göteborgs Nation, Lund would have never been so great without your presence. I will be forever sorry for almost breaking you at bumper-balls a couple of years ago, although I still have the video in case you do not remember!

Ludvig Levay, a disreputable member of the group "Sweden's worst moving-company", but also, thanks for helping me move three consecutive times, I am yet to buy you pizza for these occasions. Apart from these experiences you are one of the premier encouragements of my candidature for marshal and foreman at Göteborgs Nation.

Max Mjörnell, neither have I felt smarter studying with you, nor have I done more stupid things when socialising with you. Thanks for being my partner in crime through hard courses and also for being an exceptionally good friend always having my back. My lifelong achievement will be fulfilled once I pay my debts for all the free food and beverages I have had in your household, cheers!

Fritiof Carling, I am really sorry to inform you that I voted for Jarl during your first candidature at Göteborgs Nation, much offence. I still remember during our first house-party when you where dressed in a skirt as "Candy", memories! I am still forever grateful to call you my friend after being apart for three years.

Hampus Rosvall, one of the very few that share my nerdier habits of occasionally playing computer games and finding mathematics interesting.

Gustav Kullberg, when life gives you lemons, make lemonade! Never have I ever met someone that appreciates life more than you do, feels good to have you around!

Nils Johansson, the notorious inventor of my image "The Mountain" causing a big stir, but which also made me somewhat famous during my first semester in Lund. Still remember when you brought the whole logistics crew premium beer after working 30 hours straight, legend!

Emil Rydén, neither loquacious nor tacit but an infamous workaholic. I can never express my greatest gratitude in you having me, uninvited, at your place for two consecutive weeks last summer, I guess you will not have any trouble with burglars from now on!

Truls Grytli, it feels like yesterday we had our initiation act at Göteborgs Nation, did we do that!? Thanks for being the designated driver of the crazier parts of my student life, I cannot thank you enough!

Amanda Enander, I still remember when you got the bouncer in Durham to buy us drinks and VIP tickets. Thank you for being a true friend during my time in Chapel Hill and throughout!

Jacob Benzow, the true definition of a male! Thank you and your apartment for being the superior getaway from school and responsibilities, football pundits stands no chance against you!

Gustaf Delemark, my best friend as well as one of the biggest inspirations in my life, I cannot believe it has been four and a half long years I broke the news of beginning studying in Lund. Even though you are younger than me, you still substitute as the older brother I never had, thanks for never letting me give up during arduous occasions.

Sofi Gummeson, whether or not you want to buy the fact of us being love at first sight remains unspoken of, but thanks for pretending at least. Thank you for always believing in my extensive projects and weird travel destinations, and the fact that you solely made me stay in the U.S. during my exchange after scoring a hat-trick in failed midterms, you are the best!

Astrid, even though you are not the oldest sibling, I still see you as the most grown up! Thanks for always being such a good listener, a true bro!

Ebba, when I first met you a couple of years ago I never realized your spot on banter that I and the rest of the family have to live with throughout. A real pleasure to have you around, but wow you are getting old bro!

Mamma, arguably the best mother there is. You have always had faith in my actions, whether it was becoming an artist, bartender or mathematician, I have always had your support to the fullest.

Pappa, when we first met you assured me of getting a sports-car when turning eighteen, although I am still waiting for that offer. You are probably the person that know me to the fullest extent which makes you my go to guy when I need to talk about troubles. Thank you for always being a true inspiration and my number one role model in life.

Abstract

This thesis is an extension from the thesis "To what degree is the VIX benchmark computed by CBOE representative of its definition?" presented on June 16 in 2018.
The primary purpose of this thesis is to investigate a consistent way of Fourier pricing with the Heston model and whether or not the estimates can be improved by extending the amount of CIR processes in order to catch the non-linear behavior of VIX options.
Beginning with a brief introduction, explaining the VIX and its theoretical calculations when following the Heston dynamics. The introduction paves way explaining the basic definitions used when defining the choice of the respective model. Later, an in depth analysis of the Fourier Gauss-Laguerre algorithm when pricing European type put and call options will be thoroughly explained.
Finally, discussions regarding the parameter estimates as well as extensions of vol-of-vol terms will be further scrutinized.

Introduction

The VIX measures a 30-day expected volatility on the U.S. stock market, derived from real time, mid-quote prices of S&P500 Index (SPX) call and put options. Often mentioned as the world's most recognized measures of volatility its considered as a "good" daily indicator. Real time VIX values and calculations are further investigated in Ålander et al 2018 [22].
Derivatives on volatility indices such as the VIX index is widely used a tool for hedging against volatility risk. Variance swap contracts, used by operators on the market, are now enabled to a pure exposure to volatility or hedge against exposures of volatility in portfolios.
Liquid markets consisting of VIX options, futures and other over the counter market options have led to the need of continuously improving the consistency of pricing the derivatives on such products. Derivative prices, including their hedging prices, should be based on realistic assumptions of the joint dynamics of underlying assets and variance swaps, including the ability to match observed prices of derivatives, having reasonable liquidity. Any continuous time model with stochastic volatility should imply joint dynamics of underlying assets. The Heston model proposed by Steven Heston, is widely used due to its assumption of stochastic volatility. An attempt in catching the non-linear behaviour of VIX options the thesis experiments with adding up to four extra CIR processes.
Driven by multidimensional Brownian motion, the models are widely used to model consistent prices on volatility derivatives. Although, pricing history underlines the extreme importance of modeling discontinuities in present and future prices. Evidence of jump discontinuities shows in the fluctuations in recent crises, not exclusively purely economy related.

Figure 1 reveals simultaneous large drops in the S&P500 index, with clear volatility spikes, with strong negative correlation between the two levels. The behaviour between the S&P500 and the VIX index is further investigated in Ålander et al 2018 [22].
Realistic models, modeling the variance swaps, needs to be consistent with respect to the dynamics of the empirical evidence in figure 1 as its evident for pricing the derivatives. The jumps are also important when producing positive skew of implied volatilities of VIX options.

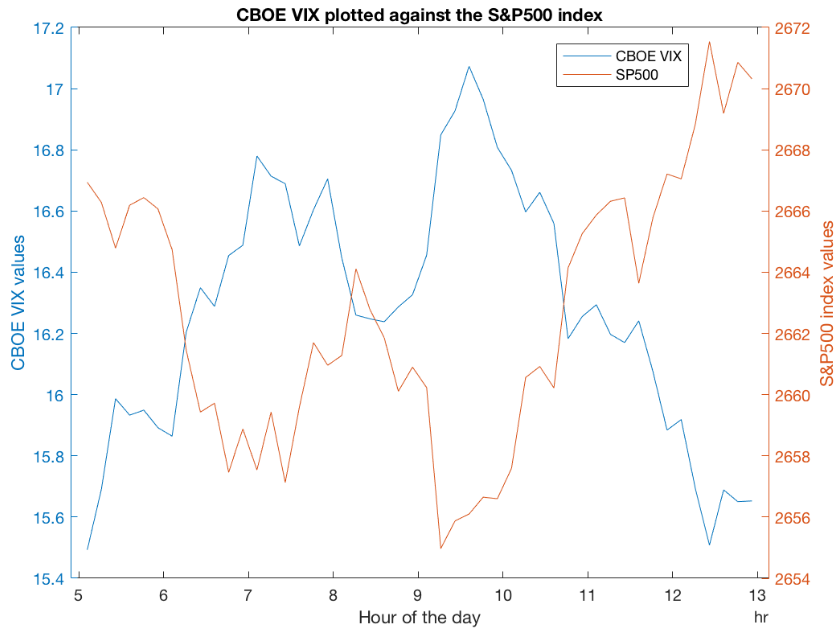

Figure 1: The CBOE computed VIX plotted against the S&P500 index.

Problem description

Writing this thesis, the main objective was to capture the non-linear behavior of volatility options. Throughout, the volatility options will be referred to as the VIX options. As the reader is about to witness, the VIX is derived from out of the money $S\&P500$ options, making it a reasonable assumption to conclude that the VIX options will be dependent on the $S\&P500$ options, i.e. derivative squared. Starting off with the simplest of models, the Black-Scholes model under the risk-neutral $\mathbb{Q}$ measure.

$$dS(t) = rS(t)dt + \sigma S(t)dW(t), \quad r \in \mathbb{R}, \quad \sigma \in \mathbb{R}^+, \quad \sigma > 0$$

The Black-Scholes model defines the dynamics of an underlying asset $S(t)$, with r being the risk-free interest rate and having a constant volatility term σ. Assuming that the interest is known, the model is to be enacted as a one parameter model. The model is very simple and will have trouble in modelling the non-linear behavior of the VIX options. Although, extending the model with a stochastic volatility term would result in the Heston Model.

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)dZ^{(1)}(t), \quad r \in \mathbb{R}$$
$$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW^{(1)}(t), \quad \kappa, \theta, V(t), \sigma \in \mathbb{R}^+$$
$$\langle dZ^{(1)}(t), dW^{(1)}(t)\rangle = \rho dt, \quad -1 \leq \rho \leq 1$$

The Heston model defines the dynamics of an underlying asset $S(t)$, with r being the risk-free interest rate, although having stochastic volatility instead of a constant volatility as in the Black-Scholes model. The Heston model will also have a mean reversion term κ, a long term volatility θ and a volatility of volatility term σ. By extending the Black-Scholes model into a Heston model, the model now consists of five parameters. The increased amount of parameters in the Heston model will have a higher accuracy in modelling the non-linear behavior of the VIX options, because of the stochastic volatility (CIR) process. Throughout the thesis, the Heston model is assumed to be insufficient in order to model the non-linear behavior of the VIX options. In order to to capture the non-linearity of the VIX options, the thesis tries to construct a stochastic volatility of volatility behaviour, i.e. enacting σ as a stochastic process. Modelling the volatility of volatility term as a stochastic process has been tried before in Foque and Saporito (2018), although obtaining an analytical expression without using excessive assumptions and approximations is very hard. This thesis tries to construct a stochastic volatility of volatility factor by extending the amount of independent volatility (CIR) processes in order to derive parameters being able to capture the non-linearity of the VIX options.

$$dV^{(i)}(t) = \kappa_i(\theta_i - V^{(i)}(t))dt + \sigma_i\sqrt{V^{(i)}(t)}dW^{(i)}(t), \quad i \in \mathbb{N} \quad \kappa_i, \theta_i, V^{(i)}(t), \sigma_i \in \mathbb{R}$$

$$\langle dZ^{(i)}(t), dW^{(i)}(t)\rangle = \rho_i dt, \quad \langle dW^{(i)}(t), dW^{(j)}(t)\rangle = 0, \quad i, j \in \mathbb{N} \quad -1 \leq \rho \leq 1$$

Extending the amount of independent volatility processes is assumed to improve the accuracy when pricing VIX options, although increasing the complexity of the model. The thesis will scrutinize three models denoted as the Heston model, the Triple Heston model and the Quintuple Heston model, having one, three and five mutually independent volatility processes. Determining which one of the models having the highest accuracy is done comparing the ask-bid spread of each of the models. The ask-bid spread refers to taking the mean of all calculated model prices lying between the ask and bid price of each option, resulting in a quote value between zero and one, referring to a zero or one-hundred percent fit to real market prices.

Computing VIX in accordance to the CBOE method

Calculating the VIX according to the Chicago Bureau of Exchange (CBOE) follows a simple yet complex three step method.

- Step 1: Out of the money S&P500 puts and calls centered at an at the money strike K_0 are selected as data when calculating the CBOE VIX. Here, only the non-zero bid prices are used.
- Step 2: Calculating the volatilities for near and next term options is done using the formula,

$$\sigma_n^2 = \frac{2}{T_n} \sum_i \frac{\Delta K_i}{K_i^2} e^{r_n T_n} Q(K_i) - \frac{1}{T_n} \left[\frac{F_n}{K_0} - 1 \right]^2, \quad n = 1, 2$$

  The VIX directly reflects the prices of the selected options, whereas ΔK_i is the sole contribution of an option to the VIX value, inversely the contribution is proportional to the square of the strike price corresponding to the option.

- Step 3: Computing the actual VIX level is done using the formula,

$$VIX = 100 \times \sqrt{\left(T_1 \sigma_1^2 \left[\frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}} \right] + T_2 \sigma_2^2 \left[\frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}} \times \right] \frac{N_{365}}{N_{30}} \right)}$$

<u>Parameters</u>

T_n: being time to expiration standardized in minutes per year $(\tau + t)$;

$T_n = \dfrac{M_{currentday} + M_{settelmentday} + M_{otherdays}}{525600}$

N_{T_1}: being the number of minutes to settlement of the near-term options

N_{T_2}: being the number of minutes to settlement of the next-term options

N_{365}: being the number of minutes in a year assuming 365 days

N_{30}: being the number of minutes in 30 days

F: being the forward index level desired from index option prices

r_n: being the interest rate

K_1: being the first strike below the forward index level F for the near-term options

K_2: being the first strike below the forward index level F for the next-term options

$Q(K_i)$: being the mid quote price or the midpoint of the bid-ask spread for the respective option with strike K_i

ΔK_i: being the interval between strike prices which is half the difference between the strike on either side of K_i i.e;

$\Delta K_i = \dfrac{K_{i+1} - K_{i-1}}{2}$

Characteristic function

Throughout this thesis, the main objective pursues an analytical expression for the stock's characteristic equation. When having obtained the characteristic equation, the probability density function can be obtained. This subsection gives the reader a brief definition regarding the characteristic equation.

Definition 1. The characteristic function of a variable X, having real values, defines its probability distribution. If the random variable associates with a specific probability distribution, then the Fourier transform of the probability density function is the characteristic function. Hence, the Fourier inverse of the characteristic function gives the probability density function.

$$F_X(x) = \mathbb{E}\left[\mathbb{I}_{\{X \leq x\}}\right]$$

$$\begin{aligned}\varphi_X(t) &= \mathbb{E}\left[\exp(itX)\right], \quad \varphi_X : \mathbb{R} \to \mathbb{C} \\ &= \int_{\mathbb{R}} \exp(itX) dF_X(x) \\ &= \int_{\mathbb{R}} \exp(itX) f_X(x) dx \\ &= \int_0^1 \exp(itQ_X(p)) dp \end{aligned} \tag{1}$$

Where F_X is the cumulative distribution function of X, with the integral being of Riemann-Stieltjes (integration by parts admitted). Given that the r.v. X has a pdf $f_X(x)$, then the corresponding Fourier transform with sign reversal in the complex exponential, then the formula in equation (1) is valid, due to the Hermitian property. Thus, $Q_X(p)$ is the quantile function of X, i.e. the inverse cumulative distribution function of X.

- $\varphi_X(0) = \mathbb{E}\left[\exp(itX)|t=0\right] = \mathbb{E}\left[\exp(0)\right] = 1$
- $|\varphi_X(t)| \leq 1$
- $\varphi(-t) = \overline{\varphi(t)}$
- $\mathbb{E}[X^k] = i^{-k}\varphi_X^k(0)$
- $\varphi_{a_1X_1+a_2X_2+\ldots+a_nX_n}(t) = \varphi_{a_1X_1}(a_1t)\varphi_{X_1}(a_2t)\ldots\varphi_{X_n}(a_nt)$ (independence)

Definition 2. Let $X(t)$ be a stochastic process on $(\Omega, \mathscr{F}_t, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{Q})$, the conditional characteristic function given the filtration is denoted by,

$$\varphi^{\mathbb{Q}}_{X(T)|\mathscr{F}_t}(z) = \mathbb{E}^{\mathbb{Q}}[\exp(izX(T))|\mathscr{F}_t].$$

The Fourier transform

Having obtained the characteristic function as defined previously, the second objective pursues an Fourier inversion of the characteristic function, obtaining the probability density function. This subsection gives the reader a brief overview of the basic theory behind Fourier transformations. For the observant reader, a characteristic function is a Fourier transform.

Definition 3. For a continuous $f : \mathbb{R} \to \mathbb{R}$, the corresponding Fourier transform is denoted as,

$$\tilde{f}(z) = \mathcal{F}(f(k))(z) = \int_{-\infty}^{\infty} f(k) \exp(zk) dk.$$

The inverse Fourier transform of $\tilde{f}$, for a complex valued $z = \alpha + i\omega$, us given by,

$$\mathcal{F}^{-1}(\tilde{f}(z))(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\alpha + i\omega) \exp(-k(\alpha + i\omega)) d\omega,$$

such that,

$$\alpha \in \{| \int_{-\infty}^{\infty} \exp(-(\alpha + i\omega)k) \tilde{f}(\alpha + i\omega) d\omega| < \infty\} \cup \{|\tilde{f}(\alpha + i\omega)| < \infty\}.$$

The Fourier transforms exists along the imaginary axis in the complex plane, hence for a complex $z = \alpha + i\omega$,

$$\int_{-\infty}^{\infty} |f(k) \exp(k(\alpha + i\omega))| dk = \int_{-\infty}^{\infty} |f(k) \exp(k\alpha)| dk.$$

In essence, the finiteness of $\mathcal{F}(f(k))(\alpha + i\omega)$ is non dependent on $\mathbf{Im}(z)$.

The transforms have the following convolutional properties for continuous functions $f, g : \mathbb{R} \to \mathbb{R}$,

- $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$
- $\mathcal{F}(fg) = \mathcal{F}(f) * \mathcal{F}(g)$
- $\mathcal{F}^{-1}(\mathcal{F}(f)\mathcal{F}(g)) = f * g$
- $\mathcal{F}^{-1}(\mathcal{F}(f) * \mathcal{F}(g)) = fg$

Fourier Gauss-Laguerre

According to the fundamental theorem of asset pricing, the value of a European-style contract at time t, having maturity at T,

$$\Pi(t, T, S(t), K) = e^{-r(T-t)}\,\mathbb{E}\left[\mathcal{P}(S(T), K)|\mathscr{F}_t\right]$$

Above, $\mathcal{P}$ is denoted as the payoff-function of a European-style contract, generally specified as,

$$\mathcal{P}(e^{\ln(S(T))}, e^{\ln(K)}, p, c) = \max(c(e^{\ln(S(T))} - e^{\ln(K)}, 0))^p$$

Where the parameters p denotes the type of the contract, i.e. one for standard, zero for binary and strictly exceeding one for power options. For simplicity and relevance, p is set to one, since the thesis only concerns standard European options. The Fourier transform with respect to the strike K reads,

$$\mathcal{F}\left(\mathcal{P}(S(T), K, c)\right) = \int_{\mathbb{R}} \max\left(c\left(e^{\ln(S(T))} - e^{\ln(K)}\right)^+\right) e^{zK} dK, \quad z \in \mathbb{C}$$

Considering $c = 1$, i.e. a European call option:

$$\int_{\mathbb{R}} \max\left(\left(e^{\ln(S(T))} - e^{\ln(K)}\right)^+\right) e^{zK} dK = \int_{\infty}^{S(t)} (e^{\ln(S(T))} - e^{\ln(K)}) e^{zK} dK$$

$$= \left[\frac{e^{S(t)+zK}}{z} - \frac{e^{K(z+1)}}{z+1}\right]_{\infty}^{S(t)}$$

$$= \frac{e^{S(t)(z+1)}}{z(z+1)}$$

By the Fourier inversion theorem, the payoff function for the stock at maturity T can be defined as,

$$\mathcal{P}(S(T), K) = \frac{1}{2\pi}\int_{\infty}^{\infty} e^{-(\alpha+i\omega)\ln(K)} \mathcal{F}\left(\mathcal{P}(S(T), \alpha + i\omega, 1)\right) d\omega$$

Taking the conditional expectation with respect to the $\mathbb{Q}$-measure,

$$\Pi(t, T, S(T), K) = \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} e^{-(\alpha+i\omega)\ln(K))}\,\mathbb{E}^{\mathbb{Q}}[\mathcal{F}\left(\mathcal{P}(S(T), \alpha + i\omega, 1)\,|\mathscr{F}_t\right] d\omega$$

Using the definition of the characteristic equation by breaking out $1 = -i^2$,

$$= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-(\alpha+i\omega)\ln(K)}}{(\alpha+i\omega)(\alpha+i\omega)}\,\mathbb{E}^{\mathbb{Q}}[e^{i\ln(S(T))(\omega - i\alpha - i)}|\mathscr{F}_t] d\omega$$

$$= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i) d\omega \tag{2}$$

Recall that the characteristic equation under $\mathbb{Q}$ of the logarithm of $S(T)$ conditional on the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ can be calculated. Furthermore, the parameter α needs to be positive,

$$\Pi(t, T, S(T), K) = \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i) d\omega$$

Assuming the option price is real,

$$= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} \mathbf{Re}\left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i)\right) d\omega$$

First of all, the integral has to exist, making α irrelevant. Although, α has to be chosen such that the integrand must not exhibit overpronounced peaks, or other elements exceeding numerical thresholds such as extreme values and oscillations. Optimizing α,

$$\alpha^* = \underset{\alpha\in A^+_{S(T)}}{\mathbf{argmin}}\left(\max_{\omega\in\mathbb{R}}\left|\frac{\partial}{\partial\omega}\mathbf{Re}\left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i)\right)\right|\right)_{|\omega=0}$$

$$= \underset{\alpha \in A^{+}_{S(T)}}{\mathbf{argmin}} \left(\int_{\omega \in \mathbb{R}} \left| \frac{\partial}{\partial \omega} \mathbf{Re} \left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i) \right) \right| \right)$$

$$A^{+}_{S(T)} := \{\alpha > 0 | \left|\varphi_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i)\right| < \infty\}$$

Assuming the conditional characteristic function exists, conditional on a filtration, the integral in equation (2) has to be truncated to an interval of the form $B := \{B = [a,b] | B \in \mathbb{R}\}$.

Another approach, consider the Gauss-Laguerre Quadrature error term, denoted as E and the weights in the Gauss-Laguerre quadrature denoted as n, (M. Abramovitz, I. Stegun 1972, page 890).

$$E(\epsilon) := \frac{(n!)^2}{(2n)!} \frac{\partial^{2n}}{\partial \omega^{2n}} \left(\omega \to \mathbf{Re} \left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i) \right) \right)_{|\omega=\epsilon}, \quad \epsilon \in (0,\infty)$$

The least computationally expensive way of computing the integrand of equation (2) is given by the inequality,

$$\left| \mathbf{Re} \left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha) \right) \right| \leq \left| \left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha) \right) \right|$$

$$\leq \frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(-i\alpha)$$

Hence, the optimal α^{+} can be chosen st,

$$\alpha^{+} = \underset{\alpha \in A^{+}_{S(T)}}{\mathbf{argmin}} \left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)} \varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(-i\alpha) \middle| \omega = 0 \right)$$

In essence, the range of evident α's is given by Jensen's inequality,

$$\left|\varphi_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i)\right| \leq \mathbb{E}^{\mathbb{Q}} \left[|e^{i \ln(S(T))(\omega - i\alpha - i)}| \middle| \mathscr{F}_t \right]$$

Breaking out a moment generating function,

$$\mathbb{E}^{\mathbb{Q}}[|e^{\ln(S(T))(\alpha+1)}||e^{\ln(S(T))i\omega}||\mathscr{F}_t]$$

Using the boundedness of the characteristic function $\varphi_{\ln(S(T))|\mathscr{F}_t}(\omega) \leq 1$,

$$\implies \left|\varphi_{\ln(S(T))|\mathscr{F}_t}(\omega - i\alpha - i)\right| \leq \mathbb{E}^{\mathbb{Q}}[S(T)^{1+\alpha}|\mathscr{F}_t]$$

Meaning, if $S(T)^{1+\alpha}$ is integrable, i.e.,

$$A^{+}_{S(T)} \supseteq \{\alpha > 0 | \, \mathbb{E}^{\mathbb{Q}}[S^{1+\alpha}(T)|\mathscr{F}_t] < \infty\}$$

Knowing for a fact that every $(1+\hat{\alpha})$-integrable variable is also $(1+\alpha)$-integrable conditional that $\alpha < \hat{\alpha}$. Using a result from Bao Rollin et al. (2010) [13], finding the largest negative and smallest positive zero of the polynomial, conditional on the models setup following Heston, Triple Heston or Quintuple Heston dynamics, is an approximation of the interval A^{+}_{j},

$$a(z) = (-\rho^3\sigma^3 z^3 + \rho\sigma^3 z^3)(3\kappa\rho^2\sigma^2 z^3 - \rho\sigma^3 z^3 - \kappa\sigma^2 z^3 + 6\rho^2\sigma^2 z^2 - 6\sigma^2 z^2) ...$$

$$... + (-3\kappa^2\rho\sigma z^3 + \kappa\sigma^2 z^3 - 12\kappa\rho\sigma z^2 + 6\sigma^2 z^2 - 24\rho\sigma z)\kappa^3 z^3 + 6\kappa^2 z^2 + 24\kappa z + 48$$

In order to prevent eventual numeric overflow, the interval can be further shrunk, using the roots in the above equation to construct the intervals $A^{+} = [a,b]$

$$\tilde{A}^{+} := [0, \min(0.999(b-1), 50000)]$$

Gauss-Laguerre

4 A quadrature rule approximates the definite integral or function as a weighted sum of function values at specified points within the domain of integration, e.g.,

$$\int_{-1}^{1} f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i), \quad \text{polynomial of degree } 2n-1 \text{ or less}$$

Gauss-Laguerre quadrature is an extension of Gaussian quadrature for approximating integrals of the following kind,

$$\int_{0}^{\infty} \exp(-x)f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

Where the Laguerre polynomial $L_n(x)$ defined as:

$$w_i = \frac{x_i}{(n+1)^2[L_{n+1}(x_i)]^2}$$

For a more general function f,

$$\int_{0}^{\infty} f(x)dx = \int_{0}^{\infty} \exp(x)\exp(-x)f(x)dx = \int_{0}^{\infty} \exp(-x)g(x)dx, \quad where \quad g(x) := \exp(x)f(x)$$

Continuing with the integrand from equation (2), by symmetry,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{Re}\left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega-i\alpha-i)\right)d\omega$$
$$= \frac{1}{\pi}\int_{0}^{\infty} \mathbf{Re}\left(\frac{e^{-(\alpha+i\omega)}}{(\alpha+i\omega)(\alpha+i\omega+1)}\varphi^{\mathbb{Q}}_{\ln(S(T))|\mathscr{F}_t}(\omega-i\alpha-i)\right)d\omega$$

Gold Search Algorithm

Definition 5. f is unimodal if for $[a,b] \in I$, $\exists p \in [a,b]$:

$$f \text{ decreasing on } [a,p], \quad f \text{ increasing on } [p,b]$$

Definition 6. Suppose f unimodal on $[a,b]$, choose $c,d \in [a,b]$, st. $a < c < d < b$. If $f(c) \leq f(d)$, the minimum must occur inside $[a,d]$, and b is replaced by c, the search continues on $[a,c]$. If $f(d) \leq f(c)$ the minimum occurs in $[c,b]$, and a is replaced by c, the search continues on $[c,b]$.

$$c = a + (1-r)(b-a) = ra + (1-r)b, \quad d = b - (1-r)(b-a) = (1-r)a + rb, \quad \text{for} \quad \frac{1}{2} < r < 1$$

Wanting r to be constant over each subinterval, meaning that r must be chosen carefully. If r chosen judicially, a new point e. Either an old interior point c or d will be used as intertior point in next subinterval, while the other point becomes endpoint in the next iteration. For every iteration, a point e has to be chosen to calculate $f(e)$, hence the carefully chosen r.

$$c - a = (1-r)(b-a) \quad \text{and} \quad b - c = r(b-a)$$

$$b - d = (1-r)(b-a) \quad \text{and} \quad d - a = r(b-a)$$

Given $f(c) \leq f(d)$,

$$\frac{d-a}{b-a} = \frac{c-a}{d-a}$$

Using equation the above,

$$\frac{r(b-a)}{b-a} = \frac{(1-r)(b-a)}{r(b-a)}, \quad \Longrightarrow \quad r = \frac{1-r}{r} \quad \Longrightarrow \quad r = \frac{-1}{2} \pm \frac{\sqrt{1+4}}{2}$$

Hence is Given by the GR:

$$r = \frac{\sqrt{5}-1}{2},$$

Which corresponds to the lowest root of the Fibonacci sequence:

$$F_n = F_{n-1} + F_{n-2}, \quad \text{where} \quad F_1 = F_2 = 1$$

The 3-dimensional Itô Lemma

Heston's stochastic volatility model treats $t, X(t)$ and $V(t)$ as variables, the Itô Lemma can easily be extended to three variables. Assuming the system of a standard Black-Scholes SDE and an Cox-Ingersoll-Ross process (CIR),

$$dX(t) = \mu_X dt + \sigma_X dW^{(1)}(t)$$
$$dV(t) = \mu_V dt + \sigma_V dW^{(2)}(t)$$

The Wiener processes $W^{(1,2)}(t)$ are dependent, having correlation,

$$\frac{\mathbb{E}\left[dW^{(1)}(t)dW^{(2)}(t)\right]}{t} = \rho dt, \quad -1 \leq \rho \leq 1, \quad t \geq 0$$

Using Taylor's expansion theorem, for $f(t, X, V)$ being twice differentiable, the dynamics can be determined by,

$$df(t,X,V) = \left[f_t + \mu_X f_X + \mu_V f_V + \frac{1}{2}\left(f_{XX}\sigma_X^2 + 2f_{X,V}\sigma_X\sigma_V + f_V\sigma_V^2\right)\right]dt + \left[\sigma_X f_X\right]dW^{(1)}(t) + \left[\sigma_V f_V\right]dW^{(2)}(t)$$

Deriving the three variable Itô Lemma starts with the difference,

$$df(t, X(t), V(t)) = f(t + dt, X(t + dt), V(t + dt)) - f(t, X(t), V(t))$$

Using Taylor's expansion to expand $f(t + dt, X(t + dt), V(t + dt))$,

$$f(t + dt, X(t + dt), V(t + dt)) = f(t, X(t), V(t)) + f_t(dt) + f_x(dX(t)) + f_V(dV(t) + \dots \tag{3}$$

$$\dots + \frac{1}{2}\Big[f_{tt}dt^2 + f_{tX}dtdX(t) + f_{tV}dtdV(t) + f_{Xt}dX(t)dt + f_{XX}dX^2(t) + f_{XV}dX(t)dV(t)...$$

$$+ f_{Vt}dV(t)dt + f_{VX}dV(t)dX(t) + f_{VV}dV^2(t)\Big].$$

The reader should be familiar with,

$$\langle dt, dt\rangle = 0, \quad \langle dt, dX(t)\rangle = \langle X(t), dt\rangle, \quad \langle dt, V(t)\rangle = \langle V(t), dt\rangle = 0,$$

Thus,

$$\langle dW^{(1)}(t), dW^{(2)}(t)\rangle = \langle dW^{(2)}(t), dW^{(1)}(t)\rangle = \rho dt.$$

Accordingly, substituting into equation (3), the resulting expansion becomes,

$$df(t, X(t), V(t)) = \mu_X f_X dt + \sigma_X f_X dW^{(1)}(t) + \mu_V f_V dt + \sigma_V f_V dW^{(2)}(t) + f_t dt...$$

$$... + \frac{1}{2}\Big[f_{XX}\sigma_X^2 dt + f_{VV}\sigma_V^2 dt] + f_{XV}(\sigma_X\sigma_V\rho dt)\Big]$$

$$= \left[\mu_X f_X + \mu_V f_V + f_t + f_{XV}\sigma_{XV}\sigma_X\sigma_V\rho + \frac{1}{2}(f_{XX}\sigma_X^2 + f_{VV}\sigma_V^2)\right]dt + \sigma_X f_X dW^{(1)}dt + \sigma_V f_V dW^{(2)}(t).$$

The models

Cox-Ingersoll-Ross model

Introduced in 1985 as an extension of the Vasicek model, the Cox-Ingersoll-Ross model, i.e. the CIR model, describes the evolution of interest rates. Corresponding to a one factor model, the interest is only driven by market risk, having applications within derivative pricing.

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dW(t)$$

Whereas, $\kappa(\theta - r(t))$ ensures the mean reversion to the long time value of the interest rate θ. The standard deviation factor $\sigma\sqrt{r(t)}$ prohibits the possibility of negative interest rates in the model, through the Feller condition,

$$2\kappa\theta \geq \sigma^2, \quad \sigma > 0$$

The closed form distribution can be interpreted as a r.v. $r_{t+\tau} = Y/\lambda$, where λ is denoted as,

$$\lambda = \frac{4\kappa\exp(-\kappa dt)}{\sigma^2(1-\exp(-\kappa dt))}$$

Whereas, Y follows a non-central chi square distribution with $d = 4\theta\kappa/\sigma^2$ degrees of freedom.

As the reader is about to witness, the Heston model has a stochastic volatility term following the dynamics of a CIR process, meaning that substituting the interest for volatility gives,

$$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW(t)$$

The characteristic function can be written on affine form,

$$\varphi_{V(t)}(z) = \mathbb{E}[e^{izV(t)}] = \exp(A(T-t) + B(T-t)V(t))$$

Whereas the terms $A(T-t)$, $B(T-t)$ can be defined as follows.

$$A(T-t) = \frac{z\exp(-\kappa(T-t))}{(1 - \frac{\sigma^2}{2(\kappa z(1-\exp(-\kappa(T-t)))})}$$

$$B(T-t) = -\frac{2\kappa\theta}{\sigma^2}\log\left(1 - \frac{\sigma^2}{2(\kappa z(1-\exp(-\kappa(T-t))))}\right)$$

Whereas the derivation of the CIR characteristic function is very similar to the Heston model shown in appendix.

Multi-dimensional Cox-Ingersoll-Ross model

Having the same assumptions as above, the multi-dimensional CIR model can be written, for simplicity, for five dimensions,

$$dV^{(i)}(t) = \kappa_i(\theta_i - V^{(i)}(t))dt + \sigma_i\sqrt{V^{(i)}(t)}dW^{(i)}(t),$$

$$\langle dW^i(t), dW^j(t)\rangle = 0, \quad i,j = 1,2,3,4,5$$

Since the characteristic function of the volatility can be written on affine form, the function can be written as,

$$\varphi_{V(t)}(z) = \exp\Big(A_1(T-t) + B_1(T_t)V^{(1)}(t) + A_2(T-t) + B_2(T-t)V^{(2)}(t)...$$

$$... + A_3(T-t) + B_3(T-t)V^{(3)}(t) + A_4(T-t) + B_4(T-t)V^{(4)}(t) + A_4(T-t) + B_5(T-t)V^{(5)}(t)\Big)$$

Where the the terms,

$$A_i(T-t) = \frac{z\exp(-\kappa_i(T-t))}{(1 - \frac{\sigma_i^2}{2(\kappa_i z(1-\exp(-\kappa_i(T-t)))})}$$

$$B_i(T-t) = -\frac{2\kappa_i\theta_i}{\sigma_i^2}\log\left(1 - \frac{\sigma_i^2}{2(\kappa_i z(1-\exp(-\kappa_i(T-t))))}\right), \quad i = 1,2,3,4,5$$

The Heston model

Describing the evolution of the volatility of a specific underlying asset, whereas the volatility lacks deterministic properties, can be done using Heston's model. Throughout, the volatility is considered as stochastic and follows a random process.

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)dZ^{(1)}(t)$$

$$dV(t) = \kappa\left(\theta - V(t)\right)dt + \sigma\sqrt{V(t)}dW^{(1)}(t)$$

Where the dynamics of the volatility process follows a CIR model. Scrutinizing the CIR, $V(t)$ can hit zero an infinite amount of times, even though assuming $V(t) > 0$. Solving the problem is done introducing the Feller condition,

$$2\kappa\theta \geq \sigma^2, \quad \sigma > 0$$

In essence, the Feller condition implies that the upward drift is sufficiently large in order to make the origin absolute inaccessible. Determining the dynamics of the log stock price,

$$X(t) = \ln(S(t))$$

The Itô dynamics are given by,

$$dX(t) = \frac{dS(t)}{S(t)} - \frac{(dS(t))^2}{2S^2(t)}$$

$$= rdt + \sqrt{V(t)}dZ^{(1)}(t) - \frac{V(t)}{2}dt$$

$$= \left(r - \frac{V(t)}{2}\right)dt + \sqrt{V(t)}dZ^{(1)}(t)$$

The characteristic function for $X(t) = \ln(S(t))$ can be written on affine form,

$$\varphi_{X(t)}(z) = \exp(C_1(T-t) + B_1(T-t)X(t) + D_1(T-t)V(t) + izX(t))$$

Although, keeping in mind that the CIR processes and the logarithm of the stock price does have a correlation term,

$$\langle dZ^{(1)}(t), dW^{(1)}(t)\rangle = \rho_1 dt$$

The analytical expression for th Heston model is stated as follows, given that we can the logarithm of the stock can be written on affine form and using the independence property of the characteristic function,

$$\varphi_{\ln(S(t))}(z) = \exp\left(iz(X(t) + r(T-t)) + C_1(T-t) + D_1(T-t)V^{(1)}(t)\right)$$

In essence, the parameters are distinguished in the following way, derived further in appendix, p: 27-33.

$$d_1 = \sqrt{(iz\rho_1\sigma_1 - \kappa_1)^2 + \sigma_1^2(iz + z^2)},$$

$$C_1(T-t) = \frac{\kappa_1\theta_1}{\sigma_1^2}\Bigg(\left(iz\rho_1\sigma_1 - \kappa_1 - d_1\right)(T-t)...$$

$$... - 2\ln\left(\frac{(\kappa_1 - iz\rho_1\sigma_1)(1 - \exp(-d_1(T-t))) + d_1(\exp(-d_1(T-t)) + 1))}{2d_1}\right)\Bigg)$$

$$D_1(T-t) = (1 - \exp(-d_1(T-t))\left(\frac{(iz)^2 - iz}{(\kappa_1 - iz\rho_1\sigma_1)(1 - \exp(-d_1(T-t))) + d_1(\exp(-d_1(T-t)) + 1)}\right)$$

Triple Heston model

Obtaining dynamics of a model replicating a stochastic volatility of stochastic volatility is assumed to be able to capture the non-linear behaviour of the term structure of the VIX options. This subsection provides the reader with sufficient information, based on the definition in the previous section. The model studied here is an extension of the Heston model.

$$dS(t) = rS(t)dt + \sqrt{V^{(1)}(t)}S(t)dZ^{(1)}(t) + \sqrt{V^{(2)}(t)}S(t)dZ^{(2)}(t)\dots$$

$$\dots + \sqrt{V^{(3)}(t)}S(t)dZ^{(3)}(t)$$

$$dV^{(1)}(t) = \kappa_1\left(\theta_1 - V^{(1)}(t)\right)dt + \sigma_1\sqrt{V^{(1)}(t)}dW^{(1)}(t)$$

$$dV^{(2)}(t) = \kappa_2\left(\theta_2 - V^{(2)}(t)\right)dt + \sigma_2\sqrt{V^{(2)}(t)}dW^{(2)}(t)$$

$$dV^{(3)}(t) = \kappa_3\left(\theta_3 - V^{(3)}(t)\right)dt + \sigma_3\sqrt{V^{(3)}(t)}dW^{(3)}(t)$$

Deriving the characteristic equation from the above equations is surprisingly easy, since we allow the models Wiener processes corresponding to each CIR term to independent of each other, in essence,

$$\langle dW^{(i)}(t), dW^{(j)}(t)\rangle = 0, \quad i, j = 1, 2, 3$$

Although, keeping in mind that the CIR processes and the logarithm of the stock price does have a correlation term,

$$\langle dZ^{(i)}(t), dW^{(i)}(t)\rangle = \rho_i dt, \quad i = 1, 2, 3$$

The analytical expression for the Triple Heston model is stated as follows, given that we can the logarithm of the stock can be written on affine form and using the independence property of the characteristic function ,

$$\varphi_{\ln(S(t))}(z) = \exp\Bigg(iz(X(t) + r(T-t)) + C_1(T-t) + C_2(T-t) + C_3(T-t)\dots$$

$$\dots + D_1(T-t)V^{(1)}(t) + D_2(T-t)V^{(2)}(t) + D_3(T-t)V^{(3)}(t)\Bigg)$$

In essence, the parameters are distinguished in the following way, further derived in the appendix,

$$d_j = \sqrt{(iz\rho_j\sigma_j - \kappa_j)^2 + \sigma_j^2(iz + z^2)},$$

$$C_j(T-t) = \frac{\kappa_j\theta_j}{\sigma_j^2}\left(\left(iz\rho_j\sigma_j - \kappa_j - d_j\right)(T-t)...\right.$$

$$\left.... - 2\ln\left(\frac{(\kappa_j - iz\rho_j\sigma_j)(1 - \exp(-d_j(T-t))) + d_j(\exp(-d_j(T-t)) + 1))}{2d_j}\right)\right)$$

$$D_j(T-t) = (1 - \exp(-d_j(T-t)))\left(\frac{(iz)^2 - iz}{(\kappa_j - iz\rho_j\sigma_j)(1 - \exp(-d_j(T-t))) + d_j(\exp(-d_j(T-t)) + 1)}\right)$$

$$j = 1, 2, 3$$

Approximating A_j^+ is done taking the largest negative and smallest positive zero of the polynomial,

$$a_j(z) = (-\rho_j^3\sigma_j^3z^3 + \rho_j\sigma_j^3z^3)(3\kappa_j\rho_j^2\sigma_j^2z^3 - \rho_j\sigma_j^3z^3 - \kappa_j\sigma_j^2z^3 + 6\rho_j^2\sigma_j^2z^2 - 6\sigma_j^2z^2)...$$

$$... + (-3\kappa_j^2\rho_j\sigma_jz^3 + \kappa_j\sigma_j^2z^3 - 12\kappa_j\rho_j\sigma_jz^2 + 6\sigma_j^2z^2 - 24\rho_j\sigma_jz)\kappa_j^3z^3 + 6\kappa_j^2z^2 + 24\kappa_jz + 48$$

$$j = 1, 2, 3$$

Quintuple Heston model

Obtaining dynamics of a model replicating a stochastic volatility of stochastic volatility is assumed to be able to capture the non-linear behaviour of the term structure of the VIX options. This subsection provides the reader with sufficient information, based on the definition in the previous section. The model studied here is an extension of the Heston model.

$$dS(t) = rS(t)dt + \sqrt{V^{(1)}(t)}S(t)dZ^{(1)}(t) + \sqrt{V^{(2)}(t)}S(t)dZ^{(2)}(t)\dots$$

$$\dots + \sqrt{V^{(3)}(t)}S(t)dZ^{(3)}(t) + \sqrt{V^{(4)}(t)}S(t)dZ^{(4)}(t) + \sqrt{V^{(5)}(t)}S(t)dZ^{(5)}(t)$$

$$dV^{(1)}(t) = \kappa_1\left(\theta_1 - V^{(1)}(t)\right)dt + \sigma_1\sqrt{V^{(1)}(t)}dW^{(1)}(t)$$

$$dV^{(2)}(t) = \kappa_2\left(\theta_2 - V^{(2)}(t)\right)dt + \sigma_2\sqrt{V^{(2)}(t)}dW^{(2)}(t)$$

$$dV^{(3)}(t) = \kappa_3\left(\theta_3 - V^{(3)}(t)\right)dt + \sigma_3\sqrt{V^{(3)}(t)}dW^{(3)}(t)$$

$$dV^{(4)}(t) = \kappa_4\left(\theta_4 - V^{(4)}(t)\right)dt + \sigma_4\sqrt{V^{(4)}(t)}dW^{(4)}(t)$$

$$dV^{(5)}(t) = \kappa_5\left(\theta_5 - V^{(5)}(t)\right)dt + \sigma_5\sqrt{V^{(5)}(t)}dW^{(5)}(t)$$

Deriving the characteristic equation from the above equations is surprisingly easy, since we allow the model to Wiener processes corresponding to each CIR term to independent of each other, in essence,

$$\langle dW^i(t), dW^j(t)\rangle = 0, \quad i,j = 1,2,3,4,5$$

Although, keeping in mind that the CIR processes and the logarithm of the stock price does have a correlation term,

$$\langle dZ^i(t), dW^i(t)\rangle = \rho_i dt, \quad i = 1,2,3,4,5$$

The analytical expression for the quintuple Heston model is stated as follows, given that we can the logarithm of the stock can be written on affine form and using the independence property of the characteristic function,

$$\varphi_{\ln(S(t))}(z) = \exp\Bigg(iz(X(t) + r(T-t)) + C_1(T-t) + C_2(T-t) + C_3(T-t) + C_4(T-t) + C_5(T-t)\dots$$

$$\dots + D_1(T-t)V^{(1)}(t) + D_2(T-t)V^{(2)}(t) + D_3(T-t)V^{(3)}(t) + D_4(T-t)V^{(4)}(t) + D_5(T-t)V^{(5)}(t)\Bigg)$$

In essence, the parameters are distinguished in the following way, further derived in the appendix,

$$d_j = \sqrt{(iz\rho_j\sigma_j - \kappa_j)^2 + \sigma_j^2(iz + z^2)},$$

$$C_j(T-t) = \frac{\kappa_j\theta_j}{\sigma_j^2}\left(\left(iz\rho_j\sigma_j - \kappa_j - d_j\right)(T-t)...\right.$$

$$\left. ... - 2\ln\left(\frac{(\kappa_j - iz\rho_j\sigma_j)(1 - \exp(-d_j(T-t))) + d_j(\exp(-d_j(T-t)) + 1))}{2d_j}\right)\right)$$

$$D_j(T-t) = (1 - \exp(-d_j(T-t))\left(\frac{(iz)^2 - iz}{(\kappa_j - iz\rho_j\sigma_j)(1 - \exp(-d_j(T-t))) + d_j(\exp(-d_j(T-t)) + 1)}\right)$$

$$j = 1,2,3,4,5$$

Approximating A_j^+ is done taking the largest negative and smallest positive zero of the polynomial,

$$a_j(z) = (-\rho_j^3\sigma_j^3z^3 + \rho_j\sigma_j^3z^3)(3\kappa_j\rho_j^2\sigma_j^2z^3 - \rho_j\sigma_j^3z^3 - \kappa_j\sigma_j^2z^3 + 6\rho_j^2\sigma_j^2z^2 - 6\sigma_j^2z^2)...$$

$$... + (-3\kappa_j^2\rho_j\sigma_jz^3 + \kappa_j\sigma_j^2z^3 - 12\kappa_j\rho_j\sigma_jz^2 + 6\sigma_j^2z^2 - 24\rho_j\sigma_jz)\kappa_j^3z^3 + 6\kappa_j^2z^2 + 24\kappa_jz + 48$$

$$j = 1,2,3,4,5$$

Theoretically calculated VIX for the Heston Model

The theoretical formula for the VIX can be described as the future, (τ units), expected squared volatility, with respect to the filtration $\mathscr{F}_t$, under the risk neutral measure $\mathbb{Q}$.

$$\frac{1}{\tau}\,\mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+\tau} V(u)du\big|\mathscr{F}_t\right]$$

Beginning from the CIR model extracted directly from the Heston model.

$$dV(t) = \kappa\left(\theta - V(t)\right)dt + \sigma\sqrt{V(t)}dW^{(1)}(t)$$

Directly integrating for $t < s \leq u$, similar to the slution of an ODE,

$$V(u) - V(t) = \int_t^u \kappa(\theta - V(s))ds + \int_t^u \sigma\sqrt{V(s)}dW^{(1)}(s)$$

$$\implies V(u) = V(t) + \int_t^u \kappa(\theta - V(s))ds + \int_t^u \sigma\sqrt{V(s)}dW^{(1)}(s)$$

By taking the expectation, conditionally on the martingale measure $\mathbb{Q}$, one obtains,

$$\mathbb{E}^{\mathbb{Q}}\left[V(u)\big|\mathscr{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[V(t) + \int_t^u \kappa(\theta - V(s))ds\big|\mathscr{F}_t\right] \tag{4}$$

$$= V(t) + \kappa\theta(u-t) - \kappa\,\mathbb{E}^{\mathbb{Q}}\left[\int_t^u V(s)ds\big|\mathscr{F}_s\right]$$

Making the substitution,

$$m(u) = \mathbb{E}^{\mathbb{Q}}\left[V(u)\big|\mathscr{F}_t\right]$$

yields,

$$m(u) = V(t) + \int_t^u \kappa(\theta - m(s))ds \quad .$$

Which results in the following Ordinary differential equation,

$$\dot{m}(u) = \kappa(\theta - m(u)), \quad m(t) = V(t).$$

Changing the variables, one obtains the standard ODE form,

$$\tilde{m}(u) = (\theta - m(u)) \implies \dot{\tilde{m}}(u) = -\kappa\tilde{m}(u)$$

The ODE can obviously be solved by finding an integrating factor, $P_u = e^{-\int_t^u \kappa dx}$,

$$\tilde{m}(u) = (\theta - V(t))e^{-\kappa(u-t)} \implies m(u) = \theta - (\theta - V(t))e^{-\kappa(u-t)}.$$

Substituting the analytical solution for $m(u)$ into equation (4) gives,

$$\frac{1}{\tau}\,\mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+\tau} V(u)du\big|\mathscr{F}_t\right] = \frac{1}{\tau}\int_t^{t+\tau} m(u)du$$

$$= \frac{1}{\tau}\int_t^{t+\tau}(\theta - (\theta - V(t))e^{-\kappa(u-t)})du$$

$$= \left(\theta - \frac{\theta - V(t)}{\kappa\tau}(1 - e^{-\kappa\tau})\right)$$

$$= V(t)\frac{(1 - e^{-\kappa\tau})}{\kappa\tau} + \frac{(e^{-\kappa\tau} - 1 + \kappa\tau)}{\kappa\tau}\theta$$

The above definition will be referred to as the theoretical VIX under the assumption that the volatility follows the Heston model under the $\mathbb{Q}$ measure. Calculating the realized volatility for the multidimensional Heston model is gathered as follows,

$$\sum_{j=1}^{N} Y_j(t) = \sum_{j=1}^{N}\left(V_j(t)\frac{(1 - e^{-\kappa_j\tau_j})}{\kappa_j\tau_j} + \frac{(e^{-\kappa_j\tau_j} - 1 + \kappa_j\tau_j)}{\kappa_j\tau_j}\theta_j\right) = \sum_{j=1}^{N}\left((C_j + D_jV_j(t)\right), \quad N \in \mathbb{N}$$

Calibration of the multi-factor Heston model

Determining the multi-factor Heston models respective parameters to match market prices of real VIX options, referred to as the model calibration problem, i.e. the inverse pricing problem.
When computing the option prices, following the Heston multi-factor dynamics, the computation needs non observable parameters from the market data, hence the problem of using empirical estimates. Pragmatically, it is neither valid nor conceivable to directly match market prices. In essence, the calibration problem becomes an optimization problem, aiming to minimize the pricing error between model and market prices for put and call options on the S&P500. The squared differences between the model and market prices, results in a measure error, defined as the nonlinear least square method, applied on the Heston model.
Finding the set of parameters Θ, minimizing the objective function G is calculated using the built in Matlab function **lsqnonlin**.

$$\min_{\theta \in \Theta} ||G(\theta)||_2^2 = \min_{\theta \in \Theta} (G(\theta_1)^2 + G(\theta_2)^2 + \ldots + G(\theta_n)^2)$$

If we let $G(\Theta)$ be the deviance, i.e. error, between the market data price and the Fourier Gauss Laguerre priced set of options,

$$G(\Theta) = \sum_{i=1}^{N} \left(\Pi^*(T_i, K_i) - \Pi^{\Theta}(t, \log(S(t)), T_i, K_i)\right)$$

Resulting in,

$$\min_{\theta \in \Theta} ||G(\theta)||_2^2 = \min_{\theta \in \Theta} \left(\sum_{i=1}^{N} \left(\Pi^*(T_i, K_i) - \Pi^{\Theta}(t, \log(S(t)), T_i, K_i)\right)^2 \right), \quad \Theta \subseteq \mathbb{R}^m, \quad m \in \mathbb{N}$$

The function allows options such as setting a lower and an upper bound for the parameters, whereas the bound is set to $\Theta \in [-\infty, \infty]$, allowing the parameters to be set to any value. The initial parameters are set to the old, non-calibrated parameters. Establishing a penalty function, set to the inverse value of the matrix consisting of the variance of each parameter. Introducing a weight, usually having a value around ATM prices. The aim is to replicate the model prices, as close to the mid quote prices as possible, therefore the weights are put on the very smallest spreads. The weight is therefore defined as,

$$\left(\max \left(\frac{\textbf{spread}}{2 * 1.96}, 0.01 \right) \right)^{-1}$$

The market data prices consists of bid and ask prices of European put and call options on the S&P500 index, information that the **lsqnonlin** can take as inputs.

When adding extra CIR processes to the Heston model, acquiring multi-factor Heston models, the processes are assumed to be mutually independent, although dependent on the underlying asset $S(t)$. Assuming independence, the optimization holds for the extended parameter space,

$$\Theta = \{V^{(1)}(0), \kappa_1, \theta_1, \sigma_1, \rho_1, V^{(2)}(0), \kappa_2, \theta_2, \sigma_2, \rho_2, V^{(3)}(0), \kappa_3, \theta_3, \sigma_3, \rho_3, V^{(4)}(0), \kappa_4, \theta_4, \sigma_4, \rho_4, V^{(5)}(0), \kappa_5, \theta_5, \sigma_5, \rho_5\}$$

Throughout, when using the function **lsqnonlin** the parameters are allowed to obtain any value. Nevertheless, since $\kappa, \theta, V(0)$ are assumed to be positive, an exponential transform is executed in order to secure those values. Similarly, the ρ parameters are transformed by using a hyperbolic tangent function, since the asymptotic values are -1 and 1 respectively.
Having acquired new, calibrated, parameters using the S&P500 option data, the Fourier Gauss Laguerre method is used once again to calculate the new price of each option, using the new parameters.

Calculating the VIX options

To begin with, the calibrated parameters achieved are used and substituted into the one, three and five dimensional log characteristic function of the CIR process, although the set of ρ parameters can be neglected throughout, since the CIR processes are mutually independent. The instant value of VIX, is given by,

$$Y(t) = C + DV(t)$$

where the constants C, D are defined as of the Theoretical VIX calculations,

$$C = \frac{1 - \exp(-\kappa(T_M))}{\kappa T_M}$$

$$D = \theta\left(\frac{\kappa T_M + \exp(-\kappa T_M) - 1}{\kappa T_M}\right), \quad T_M = \frac{30}{365}$$

Starting off with the call case of the square root of the VIX, with its Fourier transform evaluated as, evaluated for $\mathbf{Re}(z) < 0$

$$\mathcal{F}((\sqrt{Y(t)} - K)^+)(z) = \int_{K^2}^{\infty} \exp(zy)(\sqrt{y} - K)dy$$

$$= \int_{K^2}^{\infty} \exp(zy) \int_{K}^{\sqrt{y}} dvdy$$

Rearranging the order of integration,

$$K < v \leq \sqrt{y}, \quad K^2 \leq y < \infty \iff K \leq v < \infty, \quad v^2 \leq y < \infty$$

and breaking out constants gives,

$$\mathcal{F}((\sqrt{Y(t)} - K)^+)(z) = \int_{K}^{\infty} \int_{v^2}^{\infty} \exp(zy)dydv$$

$$= \int_{K}^{\infty} \frac{\exp(-(-z)v^2)}{-z} dv = \int_{K}^{\infty} \sqrt{\frac{2\pi}{-2z}} \frac{1}{\sqrt{\frac{2\pi}{-2z}}} \exp\left(-\frac{v^2}{\frac{2}{-2z}}\right) dv.$$

The observant reader immediately realises the standard normal cdf occurring,

$$\mathcal{F}((\sqrt{Y(t)} - K)^+)(z) = \sqrt{\frac{\pi}{-z}}\left(\frac{1 - \mathcal{N}(K\sqrt{-2z})}{-z}\right) = \sqrt{\frac{2\pi}{-2z}}\left(\frac{\mathcal{N}(-K\sqrt{-2z})}{-z}\right)$$

The inverse Fourier transform can then be calculated as,

$$\mathcal{F}^{-1}\left(\left(\sqrt{Y(t)} - K\right)^+\right)(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-zy)\sqrt{\pi}\left(\frac{\mathcal{N}(-K\sqrt{-2z})}{(-z)^{-3/2}}\right)\Big|_{z=\bar{z}+i\omega} d\omega$$

Taking the conditional expected value w.r.t. the risk neutral measure $\mathbb{Q}$,

$$\mathbb{E}^{\mathbb{Q}}\left[\mathcal{F}^{-1}((\sqrt{Y(t)} - K)^+|\mathscr{F}_t)\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbb{E}[\exp(-zY(0))|\mathscr{F}_t]\sqrt{\pi}\left(\frac{\mathcal{N}(-K\sqrt{-2z})}{(-z)^{-3/2}}\right)\Big|_{z=\bar{z}+i\omega} d\omega$$

Since $\mathbb{E}[\exp(-zY(0))|\mathscr{F}_t]$ is equal to the moment generating function $M_{Y(0)|\mathscr{F}_t}(-z)$. Changing variable $\hat{z} = z$, st $\bar{z} > 0$,

$$= \frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty} \exp(-r(T-t))M_{Y(0)|\mathscr{F}_t}(\hat{z})\left(\frac{\mathcal{N}(-K\sqrt{-2\hat{z}})}{(-\hat{z})^{-3/2}}\right)\Big|_{\hat{z}=\bar{z}+i\omega} d\omega$$

Secondly, the analytical strip of the moment generating function is calculated. The normal function is calculated with the help of the complex valued Faddeeva function found in appendix on p:33. The put case is evaluated similarly.

VIX term structure

The VIX term structure for a set of calculated expected volatilities on the S&P500 index having different time to maturity. These are the **VIX9D**, **VIX**, **VIX3M**, **VIX6M**, **VIX1Y** corresponding to 9-day, 30-day, 3-month, 6-month and yearly expectations of future volatility.

Results

The S&P500 option prices are captured at 04:36:00 Chicago time and the VIX option prices are captured at 04:35:00 Chicago time, in essence GMT minus six, seen in figure(3-8). Calibration of the option prices for Heston, Triple Heston and Quintuple Heston are shown in figure(9,11,13). Figure(10,12,14) shows the inside spread with respect to S&P500 options of the respective model. Figure (15-23) shows the theoretical VIX plotted against the real VIX prices. The below tables shows the calibrated parameters on S&P500 options captured at 04:36 and shows the corresponding inside spreads when the parameters are used into calculating the theoretical VIX against options captured at the three respective times.

Parameter estimates			
Params	Heston	Triple Heston	Quintuple Heston
$V^{(1)}(0)$	0.009885	0.005308	0.006627
κ_1	11.84	0.005657	6.507
θ_1	0.02593	0.01728	0.03542
σ_1	0.001696	7.352	1.166
ρ_1	-0.9999	-0.9054	-0.9536
$V^{(2)}(0)$		0.01140	0.00002463
κ_2		8.537	1.597
θ_2		0.03250	0.001121
σ_2		1.191	0.8797
ρ_2		-0.9862	-0.5742
$V^{(3)}(0)$		0.003773	0.01788
κ_3		0.000005970	32.72
θ_3		1.011	0.0001799
σ_3		0.4831	5.066
ρ_3		-0.2708	-0.9899
$V^{(4)}(0)$			0.000009984
κ_4			0.1534
σ_4			1.990
ρ_4			0.5500
$V^{(5)}(0)$			0.005245
κ_5			0.0003012
θ_5			0.6344
σ_5			0.5618
ρ_5			-0.4332

Inside spreads with respect to VIX options			
Time	Heston	Triple Heston	Quintuple Heston
04:35	0.2961	0.3007	0.1617
09:58	01177	0.1246	0.0666
13:38	0.1395	0.1122	0.0680

Discussion

Throughout the thesis, experimenting with the original Heston model, extending with extra Cox-Ingersoll-Ross terms, in order to capture the non linear behaviour of VIX options. The non-linearity explained, since the S&P500 index is constructed by the stock performance of 500 big companies listed on stock exchanges in the US. Thereafter, sloppy explained, acquiring traded options on the S&P500 index constructs the VIX index. Finally, acquiring options in the VIX index grants, obviously, VIX options. Since there are two underlying platforms, in essence the stock market and the options market, having different dynamics and liquidity, the main idea was to model the VIX options with a stochastic-vol-of-vol term.
Stating the obvious and as mentioned before, the Heston model consists of a stochastic volatility term following a CIR process. With the CIR processes having a deterministic volatility term, defined as the vol-of-vol term. With the vol-of-vol being deterministic, it could face troubles trying to model the non-linearity of the VIX options. Therefore, by extending the original Heston model with adding independent CIR processes, the model could be able to capture the looks of the VIX options.

Parameter estimation

Due to choice of the Golden Search algorithm, an initial guess is always done when trying to optimize the characteristic function. Dealing with the initial guess, the worse the initial guess, the higher amounts of steps is needed to find extreme values. With a higher amount of steps comes longer computational time for the algorithm to price the options. Stating the obvious, the higher dimension of the parameter space, the longer the calculation time, e.g. it takes longer time to price options having five parameters than twenty-five parameters, as for the Heston and Quintuple Heston models. Finally, if the algorithm is way off, the algorithm could face problems calculating the moments of the characteristic functions.
Throughout the thesis, the S&P500 data from 04:36 was used when calibrating the parameters with the non-linear least squares algorithm. The dynamics of the S&P500 index was assumed to follow a single, triple or a quintuple Heston model. The same parameters was then used when calculating the VIX prices, whereas the dynamics of the VIX was assumed to follow a single, triple or quintuple Cox-Ingersoll-Ross process. Concluding, the main difference between the two approaches is the stock dynamics found in the Heston model, as the VIX explains the volatility found in the S&P500 index. Hence, the parameter ρ is not used when calculating the theoretical VIX under the Heston model.
The initial values in the Gold-search algorithm used points following a standard normal distribution. Although, having estimated a set of parameters, the same set was used as new initial values when re-running the algorithm. The tactic was a great success, resulting in higher ask-bid spreads. The tactic was continuously used until fully optimized.

Scrutinizing the parameters found under the results, especially the Triple Heston model. The model violates the Feller condition as the product of two times κ_3 and θ_3 is lower than the vol-of-vol term σ_3. This means that the volatility is allowed to hit zero. Whether the violation inflicts damage on the parameter could be interpreted in the low estimates of σ_1 of the Triple Heston model.

Interestingly, the Quintuple Heston model does acquire very low parameter estimates for $V^{(1)}(0)$ and $V^{(4)}(0)$, these parameter estimates could be a result of the over-fitting of the Quintuple Heston's 25 parameters.

Observing the last two pairs of parameters in the Quintuple Heston model, there is a significant change in the parameters. Whilst ρ_4 indicates a positive correlation, ρ_5 indicates a correlation closer to zero than minus one. The problem might occur in the failure of the optimizer to converge to a good point. The Quintuple Heston model surely is over fitted, although gaining a better inside spread. The correlation terms are truly interesting, indicating that higher stock prices lowers volatility. In essence, through vast downfalls in the stock market, the overall volatility would interestingly increase.

Regarding the results posted in figure 10,12 and 14, the inside spread, i.e. the best bid and ask prices offered among put and call options, significantly increases with the amount of CIR processes, e.g. the Quintuple Heston achieves better results than the Triple Heston model and the Triple Heston outperforms the Heston model, by reading the results table.

Model efficiency

Handling the characteristic functions of each model was surprisingly easy since each of the CIR processes are assumed to be mutually independent, although correlated with the actual stock. Nevertheless, due to the fact that the characterstic functions are one dimensional, the Fourier transform can also be carried out in one dimension as well. A next step in adding extra complexity could be adding a non-zero correlation between the CIR processes, i.e. a square root diffusion process. Although, adding correlation to the CIR processes would obviously make the analytical expression for the characteristic as well as calibrating the model a proper nightmare.
The advantages of using the Heston model over the more complex ones is clearly the computation time. To acquire proper optimized parameters, the optimizer has to be run several times, making the Heston model super efficient due to the fact that it only takes about a minute to calculate the parameters. Although, extending the set of parameters significantly increases the computational time for optimizing the parameters with the Quintuple Heston model usually stretching over five hours, making it very time consuming to rerun the optimizer several times. Nevertheless, the Quintuple Heston gets the best ask-bid spread when plotted against the S&P500 options, being slightly over 0.5, whilst the Heston model achieves significantly less than 0.5. However, the Triple Heston model is neither the least computationally efficient nor the most efficient due to it's ask-bid spread being slightly less than the Quintuple Heston model when plotted against the S&P500 options.
Finally, the real results, the calibrated parameters from each of the three models are used respectively to calculate the theoretical VIX. The results are found in figure 15 to 23. The Heston model does a great job in replicating the VIX options, even though the CIR process is one dimensional and should have trouble with non-linearity of the VIX options, with an ask-bid spread of 0.2961. The Triple Heston gets the best results, with an ask-bid spread staggering 0.3007 and the Quintuple Heston model achieves the worst results with an ask-bid spread of 0.1617. The Quintuple Heston model seems to have problem in converging to good points to optimize its parameters. Although, the Triple Heston model might suffice in understanding the non-linearity of the VIX options, it might actually be able to capture the behavior, i.e. a Heston model might not be sufficient. Thus, calculating the theoretical VIX options with the same parameters against new data acquired at 09:58 and 13:38 did have a negative impact on the ask-bid spreads. Stating the obvious, the ask-bid spreads decreases over time. Although, interesting enough, the Heston model seems to outperform the Triple Heston model regarding their ask-bid spreads at the options captured at 13:38 with an ask-bid spread of 0.1395 and 0.1122 respectively. The Heston and Quintuple Heston models gets a better fit at 13:38 than 09:58.
Overall, the theoretical VIX calculations seems to be more efficient for the first three maturities and it certainly does have issues modelling the higher maturities. Scrutinizing, there seems to be a systematic error since the theoretical prices are almost always higher than the ask/bid prices, thus never below the ask/bid prices.

Stochastic vol-of-vol

The main idea behind this thesis was to capture the non linear behavior of the VIX options. As seen in the end of this thesis, alternative ways of modeling the dynamics had a common factor, in essence, trying to model the underlying stock with a stochastic vol-of-vol term. Since the Heston model has a stochastic volatility term, but solely a deterministic vol-of-vol constant σ, the idea of expanding the amount of CIR process was basically to replicate a stochastic σ term. Scrutinizing the extensions from the Heston model, adding extra CIR processes certainly adds extra complexity. Indicating, what distinguishes the Heston model from the Quintuple model is that there are five instead of twenty-five parameters pricing the VIX option, although theoretically, it's still not a stochastic vol-of-vol term. In order to construct a stochastic vol-of-vol term could be to extend the independent Wiener processes $W(t)$ to be dependent of each other, i.e. adding an extra correlation parameter to every set of five parameters. Another way of modeling the stochastic vol-of-vol could be by adding a jump diffusion process, following an arbitrary distribution. Although, calibrating the models would be very complex and time consuming but could be a great topic for future work. Nevertheless, adding extra CIR processes seems to work pretty well. The Heston model might have troubles capturing the more extreme non-linear behavior and the problem seems to be fixed by adding extra complexity in the form of extra CIR processes.

Fourier optimality

Deciding whether or not pricing with Fourier methods is optimal or not can be described in figure 2. The Fourier method is briefly compared with using Monte Carlo pricing for put and call options. Concluding, the two methods are pricing the same thing, i.e. European puts and calls. The most significant thing happens when scrutinizing the execution time for each method, where the Monte Carlo method executes the calculation at 5.85 and 5.81 seconds for puts and calls respectively, meanwhile the Fourier method prices at 0.251 and 0.285 respectively.
Speaking about inefficiency, the Monte Carlo is in this case, not generally, the more inefficient method of pricing European put and call options, since the Fourier method gives good results in order to its low execution time. The comparison was based underlying assets following the Heston dynamics.

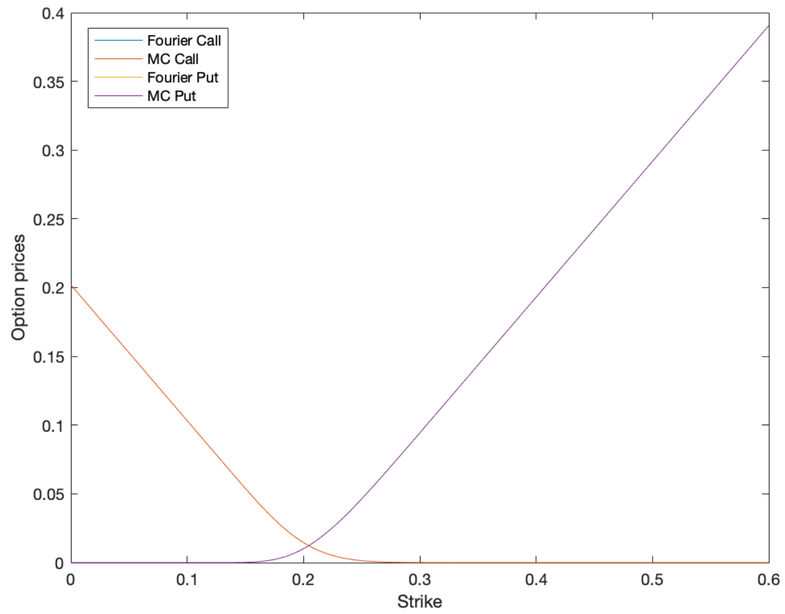

Figure 2: At the money prices for put and call options with Fourier/MC methods

Experimental models

Throughout, as in the main topic of this thesis, the main purpose was to investigate new models in order to get a consistent pricing of volatility options. Even though the thesis successfully used multi-factor Heston models, several other approaches was also scrutinized.

General solution using Riccatti ODEs in multivariate affine stochastic volatility models

Backward Stochastic Differential Equations (BSDEs) has been granted extra attention during the past few years as a new way of modelling multivariate affine stochastic models. Although, there are very few explicit solutions for these equations. By generalising results from Kallsen and Muhle-Karbe (2010), Anja Richter from Baruch College, US, proposed a method of solving generalized Riccatti ordinary differential equations. The method was suppose to introduce a flexible and comprehensive class of quadratic BSDEs, having the solution of and explicit solution of an ODE, i.e. analytically tractable.

$$Y(t) = F(X(T), O(T)) - \int_t^T \mathrm{Tr}(Z(s)^* dW(s)) - \int_t^T \mathrm{Tr}(\hat{Z}^*(s) d\hat{Q}(s))...$$

$$... - \int_0^t \int_{S_d^+ \setminus \{0\}} K_s(\xi)(\mu^X(ds, d\xi) - (ds, d\xi)) + \int_t^T (f, s, X(s), Y(s), Z(s), \hat{Z}(s), K(s))ds$$

For $t \in [0, T]$, and F depends on X and O. The Brownian motion W of the underlying (affine) process X with the generator being a Borel measureable function $f : [0, T] \times S_d^+ \times \mathbb{R} \times M_d \times M_d \times \mathbb{R} \rightarrow \mathbb{R}$
The biggest problem, acquiring a complex valued matrix case for the underlying stock dynamics, as we could only find a real valued case as in Marie-France Bru (1990) [24]. Even though more effort could have been put on this model, it would become a nightmare to calibrate the model to real data.

Heston stochastic vol-of-vol

Another approach that was proposed was modelling a Heston Stochastic vol-of-vol model, worked out by Foque and Saporito (2018) [10]. Scrutinizing a parsimonious generalization of the Heston model, where the term modelled as σ_i throughout this thesis, in essence the vol of vol was considered stochastic. Clarifying, a stochastic vol-of-vol Heston model. The model had its fundamentals in the course litterature proposed by Foque in Multiscale Stochastic Volatility of equity, Interest Rate and Credit Derivatives (2011) [17]. The objective, finding a first-order approximation, given by the Heston's quasi closed formula with corresponding parameters. Foque and Saporito tricked us into this topic by indicating that the model could be efficiently calculated since only Fourier integrals and simple ODEs was to be solved.

$$dS(t) = (r - q)S(t)dt + \sqrt{V(t)}S(t)dW^S(t)$$

$$dV(t) = \kappa(m - V(t))dt + \eta_t \sqrt{V(t)}$$

$$dW^S(t)dW^V(t) = \rho_{SV}$$

$$\eta_t = \eta(Y^\epsilon(t), Z^\delta(t))$$

$$dY^\epsilon(t) = \frac{V(t)}{\epsilon}\alpha(Y(t)^\epsilon)dt + \sqrt{\frac{V(t)}{\epsilon}}\beta(Y^\epsilon(t))dW^Y(t)$$

$$dZ^\delta(t) = V(t)\delta c(Z^\delta)dt + \sqrt{\delta V(t)}g(Z^\delta(t)g(Z^\delta(t))dW^Z(t)$$

$$dW^i(t)dW^j(t) = \rho_{ij}dt, \quad i, j = S, V, Y, Z$$

$$2\kappa m \geq \sup_{t\in[0,T]} \eta(Y^\epsilon(t), Z^\delta(t)), \quad \inf_{t\in[0,T]} \eta(Y\epsilon(t), Z^\delta(t)) > 0$$

Stating the obvious, assuming that the system of differential equations has a unique strong solution for fixed (ϵ, δ) in the fast and slow reverting stochastic vol-of-vol components. The stochastic vol-of-vol should also be measurable, bounded, bounded away from zero (Feller condition), smooth and integrable for complex valued z. Finally, α and β should follow a unique invariant distribution whereas they are also mean-reverting.

To begin with, the model was unable to find an analytical expression for the characteristic function without the use of overwhelming assumptions and approximation. Setting up the calibration, realizing the complex nightmare of trying to calibrate with the stochastic vol-of-vol term η_t. Similar to Anja Richter's generalization, this topic would be great in a future dissertation.

Conclusion

This thesis provides encouraging results regarding the application of the Heston model, extended with multi factor CIR processes to price volatility derivatives. The results show that extensions of the Heston model is favourable for further exploration, e.g. by adding extra components. Further, the method of calibrating the model to real S&P500 option prices and plotting them against real VIX option prices permits measurement of the favourable success of the technique, used when estimating the option prices. Summarizing, the Triple Heston model seems to be the best one to use, since it is better in capturing the non-linear behavior of the VIX options as well as being very computationally efficient and inexpensive.

Nevertheless, the biggest setback writing this thesis was due to the complexity of other models, when modeling stochastic vol-of-vol. Authors around the world have a tendency to favor approximations crafted by themselves and since there are a lot of authors researching in this specific topic, there are a huge number of approximate solutions when it comes to consistent modeling of VIX derivatives.

Finally, the thesis would have benefited by comparing the Heston model and its extensions to other models measuring stochastic vol-of stochastic vol.

References

[1] S. Heston (1993): *A Closed Form Solutions for Options with Stochastic Volatility with Applications to Bond and Currency Options*, Review of Financial Studies

[2] R. Cont, T. Kokholm (2009): *A Consistent Pricing Model for Index Options and Volatility Derivatives* , Finance Research Group

[3] T. Björk (2009): *Arbitrage Theory in Continuous Time, 3rd ed*, Oxford University Press S. Åberg (2018): *Lecture Notes on Derivative Pricing*. Lund University

[4] Milan Mrázek and Jan Pospíšil (2017): *Calibration and simulation of Heston model*, Open Math, vol 15, 679-704

[5] J. Jeon, G. Kim, J. Huh (2019): *Consistent and Efficient Pricing of SPX and VIX Options under Multiscale Stochastic Volatility*

[6] J. Gatheral (2008): *Consistent Modeling of SPX and VIX options*, The Fifth World Congress of the Bachelier Finance Society, London

[7] A. Richter (2014): *Explicit solutions to quadratic BSDEs and applications to utility maximization in multivariate affine stochastic volatility models*, Baruch College

[8] A. Papanicolaou (2010): *Extreme Strike Comparisons and Structural Bounds for SPX and VIX Options*

[9] M. Abramovitz, I.A Stegun (1972): *Handbook of mathematical functions with formulas, graphs and mathematical tables* National Burea of Standards, Applied Mathematics Series 55

[10] J.P. Foque, Y.F. Saporito (2018): *Heston Stochastic vol-of-vol model for joint calibration of VIX and S&P500 options* , Quantitative Finance, 18:6

[11] P. Carr, D.B. Madan (2014): *Joint modeling of VIX and SPX options at a single and common maturity with risk management applications* , IIE Transactions, 46:11

[12] M. Wiktorsson (2015): *Notes on the Benchop implementations for the fourier Gauss Laguerre FGL method*

[13] B. Rollin, S. Ferreiro-Castilla, F. Utzet (2010): *On the density of log-spot in the Heston volatility model* Stochastic Processes and their Applications, 2037-2063

[14] R. Lord, C. Kahl (2006): *Optimal Fourier Inversion in Semi -Analytical Option Pricing*, Tinbergen Institute Discussion Paper

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

[16] O.E Barndorff-Nielsen (2011): *Stochastic volatility of volatility and variance risk premia*, Aarhus University

[17] J.P Foque (2010)*Multiscale Stochastic Volatility of equity, Interest Rate and Credit Derivatives*

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

[19] P. Carr, H. Geman, D. Madan, M. Yor (2002): *The Fine Structure of Asset Returns: An Empirical Investigation* , The Journal of Business

[20] D.B. Madan, M. Yor (2010): *The S&P500 Index as a Sato Process travelling at the speed of the VIX* University of Maryland - Robert H. Smith School of Business

[21] D.B. Madan, E. Seneta (1990): *The variance Gamma (VG) Model for Share Market returns*, Journal of Business

[22] P. Liedbeck, W. Ålander (2018): *To what degree is the VIX benchmark computed by CBOE representative of its definition?*

[23] White paper VIX CBOE Volatility Index, *https://www.cboe.com/micro/vix/vixwhite.pdf*

[24] M.F. Bru (1991): *Wishart Processes* , Journal of Theoretical Probability, vol. 4, no. 4

Appendix

Derivation of Heston characteristic function

By Itô's lemma

$$df(t, X(t), V(t)) = \varphi_{X(t)}\Bigg(\left(C'(T-t) + B'(T-t)X(t) + D'(T-t)V(t)\right) dt + \dots$$

$$\dots + \left(B(T-t) + iz\right) dX(t) + D(T-t)dV(t) + \dots$$

$$\dots + \frac{1}{2}\left(B(T-t) + iz\right)^2 (dX(t))^2 + B(T-t)D(T-t)dX(t)dV(t) + \frac{1}{2}C^2(T-t)(dV(t))^2\Bigg)$$

$$= \varphi_{X(t)}\Bigg(\left(C'(T-t) + B'(T-t)X(t) + D'(T-t)V(t) + iz\right) dt + \dots$$

$$\dots + \left(B(T-t) + iz\right)\left(r - \frac{1}{2}V(t)\right) dt + \left(B(T-t) + iz\right)\sqrt{V(t)}dZ(t) + \dots$$

$$\dots + \frac{1}{2}\left(B(T-t) + iz\right)\sigma V(t)dt + B(T-t)D(T-t)\sigma V(t)\rho dt + \frac{1}{2}\sigma^2 V(t)dt\Bigg)$$

According to the Tower Property of conditional expectation, the drift equals zero. The initial conditions that $t = T$ and $f(iz) = \exp(izX(T))$, implies, $C(0) = B(0) = D(0) = 0$. Setting $X(t) = V(t) = 0$ to solve the system of equations.

$$-B'(T-t) = 0, \quad B(0) = 0 \quad \Longrightarrow \quad B(T-t) = F \in \mathbb{N} \quad \Longrightarrow \quad B(T-t) = 0$$

Meaning that the characteristic function looks like,

$$\mathbb{E}[\exp(izX(t))] = \exp(C(T-t) + D(T-t)V(t) + izX(t))$$

$$D'(T-t) = \frac{-1}{2}iz - \kappa D(T-t) - \frac{1}{2}z^2 + \frac{1}{2}\sigma^2(D(T-t)^2 + \rho\sigma iz D(T-t)$$

$$= -\frac{1}{2}\left(-\frac{iz}{\sigma^2} - \frac{z^2}{\sigma^2} + \left(\frac{2\rho iz}{\sigma} - \frac{2\kappa}{\sigma^2}\right) D(T-t) + C^2(T-t)\right)$$

$$\Longrightarrow \frac{D'(T-t)}{D^2(T-t) + \left(\frac{2\rho iz}{\sigma} - \frac{2\kappa}{\sigma^2} D(T-t) - \left(\frac{iz+z^2}{\sigma^2}\right)\right)} = -\frac{\sigma^2}{2} \tag{5}$$

Solving the roots for $D(T-t)$ using the quadratic formula,

$$r_{1,2}(T-t) = -\frac{iz\rho\sigma - \kappa}{\sigma^2} \pm \sqrt{-\frac{\rho^2 z^2}{\sigma^2} - \frac{2i\rho z\kappa}{\sigma^3} + \frac{\kappa^2}{\sigma^4} + \frac{iz + z^2}{\sigma^2}}$$

Substitute $\beta = \sqrt{-\frac{\rho^2 z^2}{\sigma^2} - \frac{2i\rho z\kappa}{\sigma^3} + \frac{\kappa^2}{\sigma^4} + \frac{iz+z^2}{\sigma^2}}$

$$r_{1,2}(T-t) = \alpha \pm \sqrt{B}$$

Rearranging the denominator in equation (5),

$$\frac{D'(T-t)}{\left(D(T-t) - (\alpha - \beta)\right)\left(D(T-t) - (\alpha + \beta)\right)} = -\frac{1}{2}\sigma^2$$

Partial fraction decomposition gives,

$$\frac{1}{\left(D(T-t) - (\alpha - \beta)\right)\left(D(T-t) - (\alpha + \beta)\right)} = \frac{G}{\left(D(T-t) - (\alpha - \beta)\right)} + \frac{H}{\left(D(T-t) - (\alpha + \beta)\right)} \tag{6}$$

$$\implies \quad 1 = G\left(D(T-t) - (\alpha - \beta)\right) + H\left(D(T-t) - (\alpha + \beta)\right)$$

Dividing into two cases,

$$\text{Let } D(T-t) = \alpha + \beta \implies 2\beta H = 1 \implies H = \frac{1}{2\beta}$$

$$\text{Let } D(T-t) = \alpha - \beta \implies -2\beta G = 1 \implies G = -\frac{1}{2\beta}$$

Substituting the coefficients into equation (6),

$$\frac{D'(T-t)}{\left(D(T-t) - (\alpha - \beta)\right)\left(D(T-t) - (\alpha + \beta)\right)} = \left(\frac{-\frac{1}{2\beta}}{(D(T-t) - (\alpha - \beta))} + \frac{\frac{1}{2\beta}}{(D(T-t) - (\alpha + \beta))}\right) D'(T-t)$$

The differential equation can now be solved for $D(T-t)$,

$$-\frac{\sigma^2}{2} = \left(\frac{-\frac{1}{2\beta}}{(D(T-t) - (\alpha - \beta))} + \frac{\frac{1}{2\beta}}{(D(T-t) - (\alpha + \beta))}\right) D'(T-t)$$

$$\implies -\int_{s=t}^{s=T} \frac{\sigma^2}{2} ds = \int_{s=t}^{s=T} \left(\frac{-\frac{1}{2\beta}}{(D(T-t) - (\alpha - \beta))} + \frac{\frac{1}{2\beta}}{(D(T-t) - (\alpha + \beta))}\right) dC(T-s)$$

$$\implies -\left[\frac{\sigma^2}{2} s\right]_{s=t}^{s=T} = \left[-\frac{1}{2\beta}\ln(C(T-s) - (\alpha - \beta)) + \frac{1}{2\beta}\ln(C(T-s) - (\alpha + \beta))\right]_{s=t}^{s=T}$$

Assuming that $D(0) = 0$,

$$\implies \frac{\sigma^2}{2}(T-t) = -\frac{1}{2\beta}\Big(\ln(\beta - \alpha) - \ln(D(T-t) - (\alpha - \beta))\Big) + \frac{1}{2\beta}\Big(\ln(-(\alpha + \beta)) - \ln(D(T-t) - (\alpha + \beta))\Big)$$

Rearranging,

$$= \frac{1}{2\beta}\Big(\ln(D(T-t) - (\alpha - \beta)) + \ln(-(\alpha + \beta))\Big) - \frac{1}{2\beta}\Big(\ln(D(T-t) - (\alpha + \beta)) + \ln(\beta - \alpha)\Big)$$

Knowing that the logarithmic law indicates $\ln(x) + \ln(y) = \ln(xy)$

$$= \frac{1}{2\beta}\Big(\ln((D(T-t) - (\alpha - \beta))(-\alpha - \beta))\Big) - \frac{1}{2\beta}\Big(\ln((D(T-t) - (\alpha + \beta))(\beta - \alpha))\Big)$$

Knowing that the logarithmic law indicates $\ln(x) - \ln(y) = \ln(\frac{x}{y})$

$$= \frac{1}{2\beta}\ln\left(\frac{(D(T-t) - (\alpha - \beta))(-\alpha - \beta)}{(D(T-t) - (\alpha + \beta))(\beta - \alpha)}\right)$$

Multiplying $\frac{1}{2\beta}$ on each side,

$$\implies -\beta\sigma^2(T-t) = \ln\left(\frac{(D(T-t) - (\alpha - \beta))(-\alpha - \beta)}{(D(T-t) - (\alpha + \beta))(\beta - \alpha)}\right)$$

Executing an exponential transform on each side,

$$\implies \exp(-\beta\sigma^2(T-t)) = \frac{(D(T-t) - (\alpha - \beta))(-\alpha - \beta)}{(D(T-t) - (\alpha + \beta))(\beta - \alpha)}$$

$$\implies -\left(\frac{\beta - \alpha}{\alpha + \beta}\right)\exp(-\beta\sigma^2(T-t)) = \frac{D(T-t) - (\alpha - \beta)}{D(T-t) - (\alpha + \beta)}$$

$$\implies -\left(\frac{\beta-\alpha}{\alpha+\beta}\right)\exp(-\beta\sigma^2(T-t))\left(D(T-t)-(\alpha+\beta)\right)=D(T-t)-(\alpha-\beta)$$

$$\implies D(T-t)+\frac{\beta-\alpha}{\beta+\alpha}\exp(-\beta\sigma^2(T-t))D(T-t)=(\alpha-\beta)+(\beta+\alpha)\exp(-\beta\sigma^2(T_t))$$

$$\implies D(T-t)\left(1+\frac{\beta-\alpha}{\beta+\alpha}\exp(-\beta\sigma^2(T-t))\right)=(\alpha-\beta)+(\beta+\alpha)\exp(-\beta\sigma^2(T-t))$$

$$\implies D(T-t)=\frac{(\alpha-\beta)+(\beta-\alpha)\exp(-\beta\sigma^2(T-t))}{1+\left(\frac{\beta+\alpha}{bb+\alpha}\right)\exp(-\beta\sigma^2(T-t))}$$

Which is the closed expression for $D(T-t)$, solving for β,

$$\beta=\sqrt{-\frac{\rho^2z^2}{\sigma^2}-\frac{2i\rho z\kappa}{\sigma^3}+\frac{\kappa^2}{\sigma^4}+\frac{iz+z^2}{\sigma^2}}$$

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

Setting $d=\sqrt{(\rho-\kappa)+\sigma^2(iz+z^2)}$ giving,

$$\beta=\frac{1}{\sigma^2}d$$

Simplifying $D(T-t)$ even further,

$$D(T-t)=\frac{(\alpha-\frac{M^2}{\sigma^4})+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))}{1+\left(\frac{\frac{M^2}{\sigma^4}-\alpha}{\frac{M^2}{\sigma^4}+\alpha}\right)\exp(-d(T-t))}$$

$$=\frac{(\alpha-\frac{M^2}{\sigma^4})+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))}{\frac{\frac{M^2}{\sigma^4}+\alpha}{\frac{M^2}{\sigma^4}+\alpha}+\left(\frac{\frac{M^2}{\sigma^4}-\alpha}{\frac{M^2}{\sigma^4}+\alpha}\right)\exp(-d(T-t))}$$

$$=\frac{(\alpha-\frac{M^2}{\sigma^4})+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))}{\frac{\frac{M^2}{\sigma^4}+\alpha+\left(\frac{M^2}{\sigma^4}-\alpha\right)\exp(-d(T-t))}{\frac{M^2}{\sigma^4}+\alpha}}$$

Multiply each side of numerator and denominator by $\frac{M^2}{\sigma^4}+\alpha$,

$$=\frac{(\frac{M^2}{\sigma^4}+\alpha)\left((\alpha-\frac{M^2}{\sigma^4})+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))\right)}{\frac{M^2}{\sigma^4}+\alpha+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))}$$

Multiplying by $\frac{\exp(d(T-t))}{\exp(d(T-t))}$,

$$=\frac{(\frac{M^2}{\sigma^4}+\alpha)\left((\alpha-\frac{M^2}{\sigma^4})+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))\right)}{\frac{M^2}{\sigma^4}+\alpha+(\frac{M^2}{\sigma^4}-\alpha)\exp(-d(T-t))}\left(\frac{\exp(d(T-t))}{\exp(d(T-t))}\right)$$

$$=\frac{(\alpha^2-\frac{M^2}{\sigma^4})\exp(d(T-t))+(\frac{M^2}{\sigma^4}-\alpha^2)}{(\frac{M^2}{\sigma^4}+\alpha)\exp(d(T-t))+(\frac{M^2}{\sigma^4}-\alpha)}$$

$$=\frac{(\exp(d(T-t))-1)(\alpha^2-\frac{M^2}{\sigma^4})}{(\frac{M^2}{\sigma^4}?\alpha)\exp(d(T-t))+(\frac{M^2}{\sigma^4}-\alpha)}$$

$$=\frac{(\exp(d(T-t))-1)\left(-\frac{z^2\rho^2}{\sigma^2}-\frac{2iz\rho\kappa}{\sigma^3}+\frac{\kappa^2}{\sigma^4}-\frac{1}{\sigma^4}\left(-z^2\rho^2\sigma^2-2iz\rho\sigma\kappa\kappa^2\sigma^2iz+\sigma^2z^2\right)\right)}{(\frac{M^2}{\sigma^4}+\alpha)\exp(d(T-t))+(\frac{M^2}{\sigma^4}-\alpha)}$$

$$= \frac{(\exp(d(T-t))-1)\left(-\frac{z^2\rho^2}{\sigma^2}-\frac{2iz\rho\kappa}{\sigma^3}+\frac{\kappa^2}{\sigma^4}+\frac{z^2\rho^2}{\sigma^2}+\frac{2iz\rho\kappa}{\sigma^3}-\frac{\kappa^2}{\sigma^4}-\frac{iz}{\sigma^2}-\frac{z^2}{\sigma^2}\right)}{(\frac{M^2}{\sigma^4}+\alpha)\exp(d(T-t))+(\frac{M^2}{\sigma^4}-\alpha)}$$

$$= \left(\frac{\exp(d(T-t))-1}{\sigma^2}\right)\left(\frac{-iz-z^2}{(\frac{M^2}{\sigma^4}-\frac{iz\rho\sigma-\kappa}{\sigma^2})\exp(d(T-t))+(\frac{M^2}{\sigma^4}+\frac{iz\rho\sigma-\kappa}{\sigma^2}}\right)$$

$$= \frac{\exp(d(T-t))-1}{\sigma^2}\left(\frac{iz+z^2}{\frac{1}{\sigma^2}(iz\rho\sigma-\kappa-d))\exp(d(T-t))-(iz\rho\sigma-\kappa+d)}\right)$$

$$= \frac{(\exp(d(T-t))-1)(iz+z^2)}{(iz\rho\sigma-\kappa-d)\exp(d(T-t))-(iz\rho\sigma)-\kappa+d)}\left(\frac{\frac{1}{-(izps-k+d)}}{\frac{1}{-(izps-k+d}}\right)$$

$$= \frac{(\exp(d(T-t))-1)\left(\frac{-(iz+z^2)}{izps-k+d}\right)}{1-\left(\frac{iz\rho\sigma-\kappa-d}{izps-k+d}\right)\exp(d(T-t))}$$

$$N = \frac{iz\rho\sigma-\kappa-d}{iz\rho\sigma-\kappa+d}$$

$$\implies D(T-t) = \frac{(\exp(d(T-t))-1)\left(-\frac{iz+z^2}{iz\rho\sigma-\kappa+d}\right)}{1-N\exp(d(T-t))}$$

$$= \frac{(\exp(d(T-t))-1)\left(-\sigma^2\frac{iz+z^2}{iz\rho\sigma-\kappa+d}\right)}{\sigma^2(1-N\exp(d(T-t)))}$$

$$= \frac{(\exp(d(T-t))-1)\left(-\sigma^2\frac{(iz+z^2)(iz\rho\sigma-\kappa-d)}{(iz\rho\sigma-\kappa+d)(iz\rho\sigma-\kappa-d)}\right)}{\sigma^2(1-N\exp(d(T-t)))}$$

$$= \frac{(\exp(d(T-t))-1)\left(\frac{(-\sigma^2z^2-iz\sigma^2)(iz\rho\sigma-\kappa-d)}{-(\sigma^2+iz\sigma^2)}\right)}{\sigma^2(1-N\exp(d(T-t)))}$$

$$\implies D(T-t) = \frac{(\exp(d(T-t))-1)(iz\rho\sigma-\kappa-d)}{\sigma^2(1-N\exp(d(T-t)))}$$

Where,

$$d = \sqrt{(iz\rho\sigma-\kappa)^2+\sigma^2(iz+z^2)}, \quad N = \frac{iz\rho\sigma-\kappa-d}{iz\rho\sigma-\kappa+d}$$

Now, to $C(T-t)$,

$$C'(T-t) = irz+\kappa\theta D(T-t)$$

Integrating each side,

$$\int_{s=t}^{t=T} A'(T-s)ds = \int_{s=t}^{t=T}(irz+\kappa\theta C(T-s))\,ds$$

$$\implies C(T-t) = irz(T-t)+\kappa\theta\int_{s=t}^{t=T} C(T-s)ds$$

Substituting for $D(T-t)$,

$$\implies C(T-t) = irz(T-t)+\kappa\theta\int_{s=t}^{t=T}\frac{(\exp(d(T-s))-1)(iz\rho\sigma-\kappa-d)}{\sigma^2(1-N\exp(d(T-s)))}ds$$

$$= irz(T-t)+\frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma-\kappa-d)\int_{s=t}^{t=T}\frac{\exp(d(T-s))-1}{1-N\exp(d(T-s))}ds$$

$$= irz(T-t)+\frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma-\kappa-d)\left[-\frac{\ln(N\exp(d(T-t))-1)}{M}+\frac{\ln(N\exp(d(T-t))-1)}{MN}+\frac{\ln(\exp(d(T-t)))}{M}\right]_{s=t}^{s=T}$$

$$= irz(T-t) + \frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma - \kappa - d)\Bigg(\Bigg(-\frac{\ln(N-1)}{M} + \frac{\ln(N-1)}{MN}\Bigg)\dots$$

$$\dots - \Bigg(-\frac{\ln(N\exp(d(T-t))-1)}{M} + \frac{\ln(N\exp(d(T-t))-1)}{MN} + \frac{\ln(\exp(d(T-t)))}{M}\Bigg)\Bigg)$$

$$= irz(T-t) + \frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma - \kappa - d)\Bigg(\frac{1}{d}\left(\ln(N\exp(d(T-t))-1) - \ln(N_1)\right) + \frac{1}{dN}(\ln(N-1)...$$

$$... - \ln(N\exp(d(T-t))-1)) - \frac{d(T-t)}{d}\Bigg)$$

$$= irz(T-t) + \frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma - \kappa - d)\left(\frac{1}{d}\ln(\frac{N\exp(d(T-t))-1}{N-1}) + \frac{1}{dN}\ln(\frac{N-1}{N\exp(d(T-t))-1}) - (T-t)\right)$$

$$= irz(T-t) + \frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma - \kappa - d)\left(\frac{1}{d}\ln(\frac{1-N\exp(d(T-t))}{1-N}) + \frac{1}{dN}\ln(\frac{1-N}{1-N\exp(d(T-t))}) - (T-t)\right)$$

$$= irz(T-t) + \frac{1}{\sigma^2}\kappa\theta(iz\rho\sigma - \kappa - d)\left(\frac{N\ln(\frac{1-N\exp(d(T-t))}{1-N}) + \ln(\frac{1-N}{1-N\exp(d(T-t))})}{dN} - (T-t)\right)$$

$$= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) + (\rho\sigma iz - \kappa + d)\left(\frac{N\ln(\frac{(1-N\exp(d(T-t))}{1-N} + \ln(\frac{1-N}{1-N\exp(d(T-t))}}{d}\right)\right)$$

$$= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) + (\rho\sigma iz - \kappa + d)\left(\frac{N\ln(\frac{(1-N\exp(d(T-t))}{1-N} + \ln(\frac{1-N\exp(d(T-t))}{1-N}}{d}\right)\right)$$

$$= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) + (\rho\sigma iz - \kappa + d)\left(\frac{N-1}{d}\ln(\frac{1-N\exp(d(T-t))}{1-N}\right)\right)$$

$$= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) + (\rho\sigma iz - \kappa + d)\left(\frac{\frac{\rho\sigma iz-\kappa-d}{\rho\sigma iz-\kappa+d} - 1}{d}\ln(\frac{1-N\exp(d(T-t))}{1-N}\right)\right)$$

$$= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) + \left(\frac{\rho\sigma iz - \kappa - d - \rho\sigma iz + \kappa - d}{d}\right)\left(\ln(\frac{1-N\exp(d(T-t))}{1-N}\right)\right)$$

$$\implies C(T-t) = riz(T-t) + \frac{\kappa\theta}{\sigma^2}\left(-(\rho\sigma iz - \kappa - d)(T-t) - 2\ln(\frac{1-N\exp(d(T-t))}{1-N})\right)$$

Simplifying even further, with the objective to reduce the amount of deterministic constants, the right hand side expression is targeted,

$$-2\ln\left(\frac{1-N\exp(d(T-t))}{1-N}\right) = -2\ln\left(\frac{1-\left(\frac{iz\rho\sigma-\kappa-d}{iz\rho\sigma-\kappa+d}\right)\exp(d(T-t))}{1-\left(\frac{iz\rho\sigma-\kappa-d}{iz\rho\sigma-\kappa+d}\right)}\right)$$

The obvious thing is to multiply by the denominator of N,

$$= -2\ln\left(\frac{iz\rho\sigma - \kappa + d - (iz\rho\sigma - \kappa + d)\exp(d(T-t))}{2d}\right)$$

$$= -2\ln\left(\frac{(\kappa - iz\rho\sigma)\left(\exp(d(T-t))-1\right) + d\left(\exp(d(T-t))+1\right)}{2d}\right)$$

Knowing that using the logarithmic properties,

$$-(iz\rho\sigma - \kappa - d) = \ln(\exp(-1))(iz\rho\sigma - \kappa - d)$$

For notational issues, the interest rate term will be separated from the expression, giving the new expression for $C(T-t)$,

$$\implies C(T-t) = \frac{\kappa\theta}{\sigma^2}\left(\left(iz\rho\sigma - \kappa - d\right)(T-t) - 2\ln\left(\frac{(\kappa - iz\rho\sigma)(1-\exp(-d(T-t))) + d(\exp(-d(T-t))+1))}{2d}\right)\right)$$

Doing the same for the expression $D(T-t)$, focusing on removing the N term,

$$D(T-t) = \frac{(1-\exp(-d(T-t)))(iz\rho\sigma - \kappa - d)}{\sigma^2(\exp(-d(T-t)) - N)}$$

Multiplying with the denominator of N on each side of the fraction,

$$= \frac{(1-\exp(-d(T-t)))(d^2+\kappa^2-2i\kappa\rho\sigma z - \rho^2\sigma^2 z^2)}{\sigma^2(\exp(-d(T-t))(iz\rho\sigma - \kappa + d) - iz\rho\sigma + \kappa + d)}$$

$$= \frac{(1-\exp(-d(T-t)))(d^2+\kappa^2-2i\kappa\rho\sigma z - \rho^2\sigma^2 z^2)}{\sigma^2(\kappa - iz\rho\sigma)(1-\exp(-d(T-t)) + d(\exp(-d(T-t))+1)}$$

Using the expression for d^2,

$$d^2 = (iz\rho\sigma - \kappa)^2 + \sigma^2(iz+z^2)$$

$$\implies D(T-t) = (1-\exp(-d(T-t))\frac{(iz)^2 - iz}{(\kappa - iz\rho\sigma)(1-\exp(-d(T-t))) + d(\exp(-d(T-t))+1)}$$

In essence,

$$d = \sqrt{(iz\rho\sigma-\kappa)^2 + \sigma^2(iz+z^2)}$$

$$C(T-t) = \frac{\kappa\theta}{\sigma^2}\left(\left(iz\rho\sigma - \kappa - d\right)(T-t) - 2\ln\left(\frac{(\kappa - iz\rho\sigma)(1-\exp(-d(T-t))) + d(\exp(-d(T-t))+1))}{2d}\right)\right)$$

$$D(T-t) = (1-\exp(-d(T-t))\left(\frac{(iz)^2 - iz}{(\kappa - iz\rho\sigma)(1-\exp(-d(T-t))) + d(\exp(-d(T-t))+1)}\right)$$

Approximation of the complex error function

Defining an error function for complex inputs is done using Gauss error function, i.e. the complex error function, for $z \in \mathbb{C}$,

$$\mathbf{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty \exp(-t^2)dt = 1 - \mathbf{erf}(z)$$

= Which obviously is symmetric and the error function for $z \in \mathbb{R}$,

$$\mathbf{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-t^2)dt$$

The complex error function is entire and has no singularities, except at infinity. Assuming that the function's Taylor series always converges, the asymptotic expansion by 7.1.23 in [9],

$$\sqrt{\pi} z \exp(z^2)\mathbf{erfc}(z) \approx 1 + \sum_{m=1}^{\infty} (-1)^m \frac{1 * 3 ...(2m-1)}{(2z^2)^m}, \quad z \to \infty, \quad \left| \arg(z) < \frac{3\pi}{4} \right|$$

The infinite series for the Complex error function 7.1.29 in [9], for $z := x + iy$,

$$\mathbf{erf}(x+iy) = \mathbf{erf}(x) + \frac{\exp(-x^2)}{2\pi x}\left((1-\cos(2xy)) + i\sin(2xy)\right) + \frac{2}{\pi}\exp(-x^2)\sum_{n=1}^{\infty}\frac{\exp(-n^2/2)}{n^2+4x^2}\left(f_n(x,y) + ig_n(x,y)\right) + \epsilon(x,y)$$

Defining the polynomials,

$$f_n(x, y) = 2x - 2x\cosh(ny)\cos(2xy) + n\sinh(ny)\sin(2xy)$$

$$g_n(x, y) = 2x\cosh(ny)\sin(2xy) + n\sinh(ny)\cos(2xy)$$
$$|\epsilon(x, y)| \approx 10^{-16}|\mathbf{erf}(x + iy)|$$

Now, defining the scaled complex complimentary error function, i.e. the Faddeeva function, for $z \in \mathbb{C}$,

$$w(z) := \exp(-z^2)\mathbf{erfc}(-iz) \tag{7}$$
$$= \exp(-z^2)\left(1 + \frac{2i}{\sqrt{\pi}}\int_0^z \exp(t^2)dt\right)$$

Figures

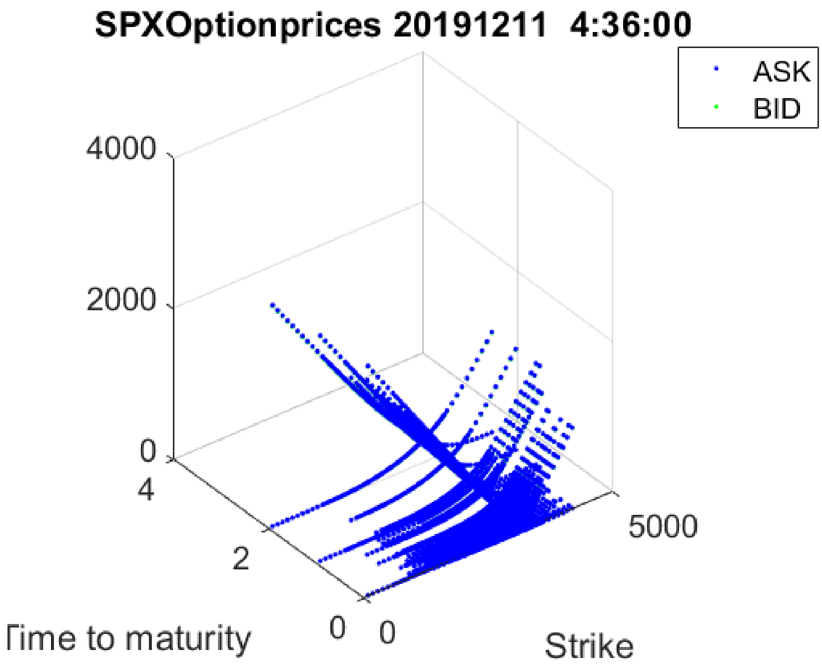

Figure 3: S&P500 option prices

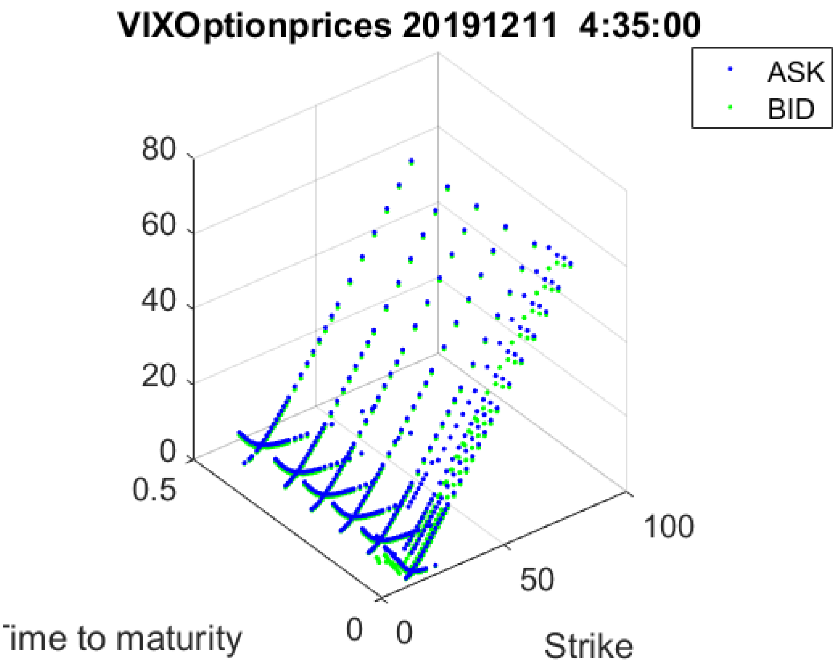

Figure 4: VIX option prices

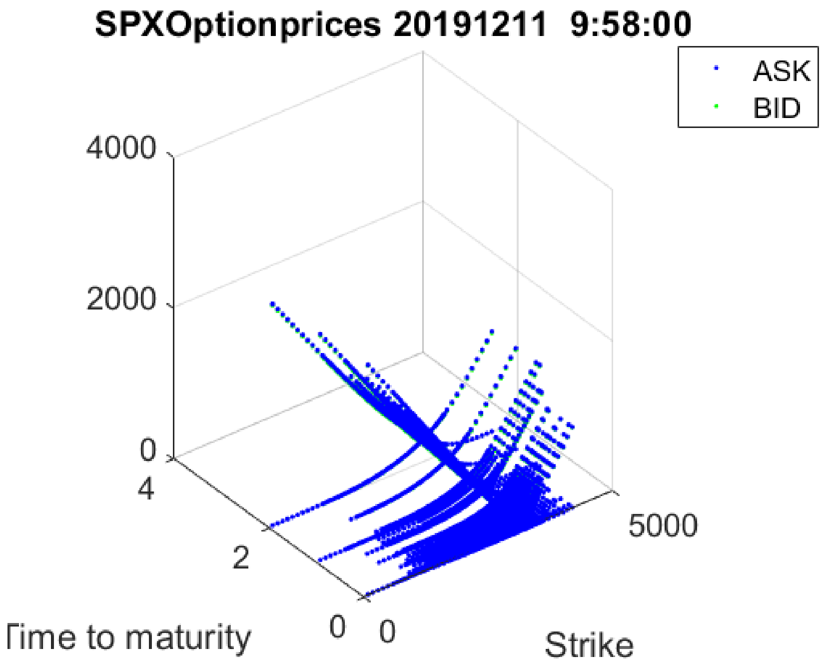

Figure 5: S&P500 option prices

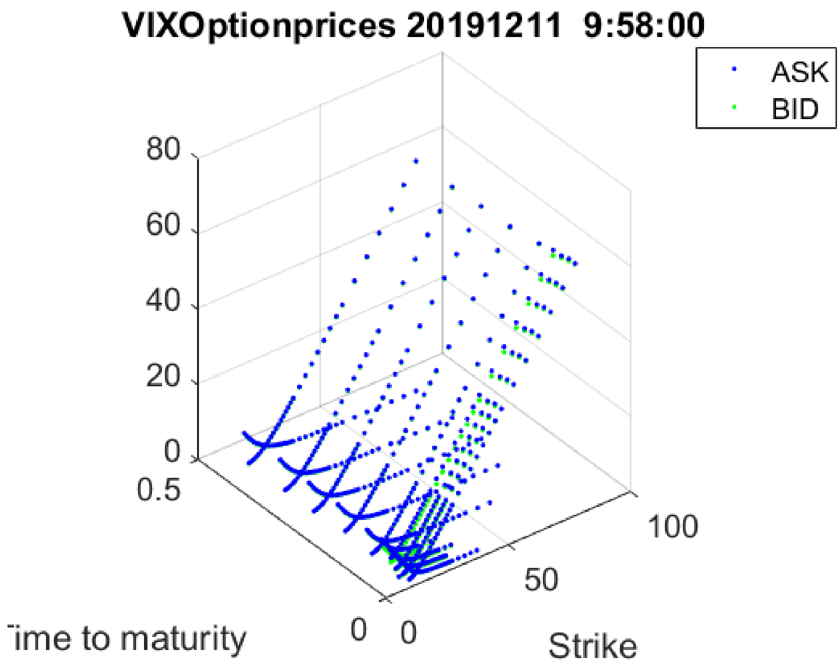

Figure 6: VIX option prices

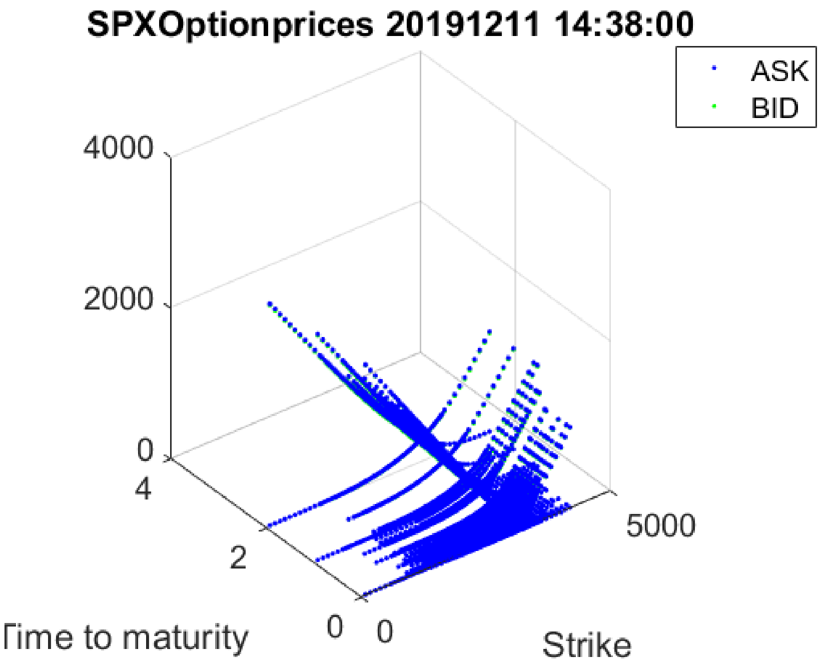

Figure 7: S&P500 option prices

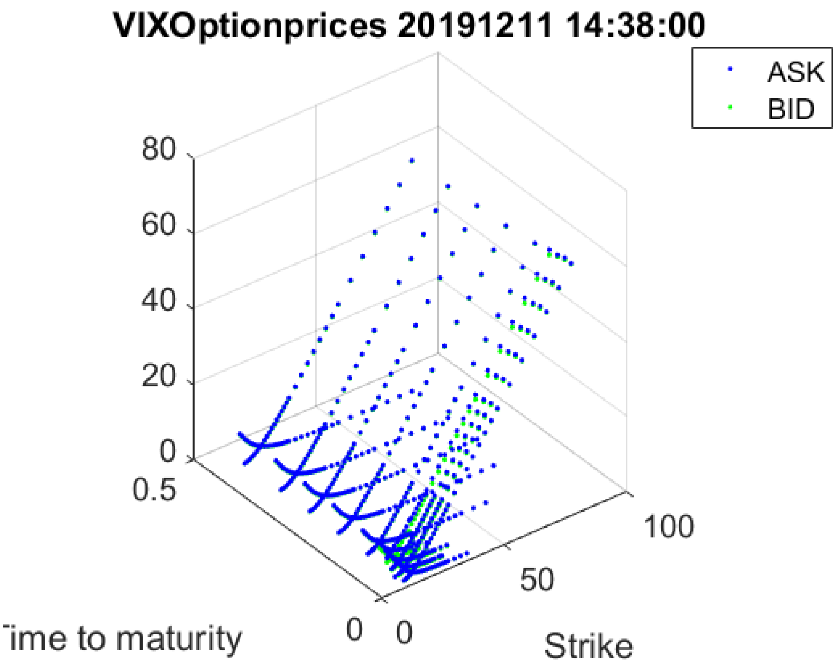

Figure 8: VIX option prices

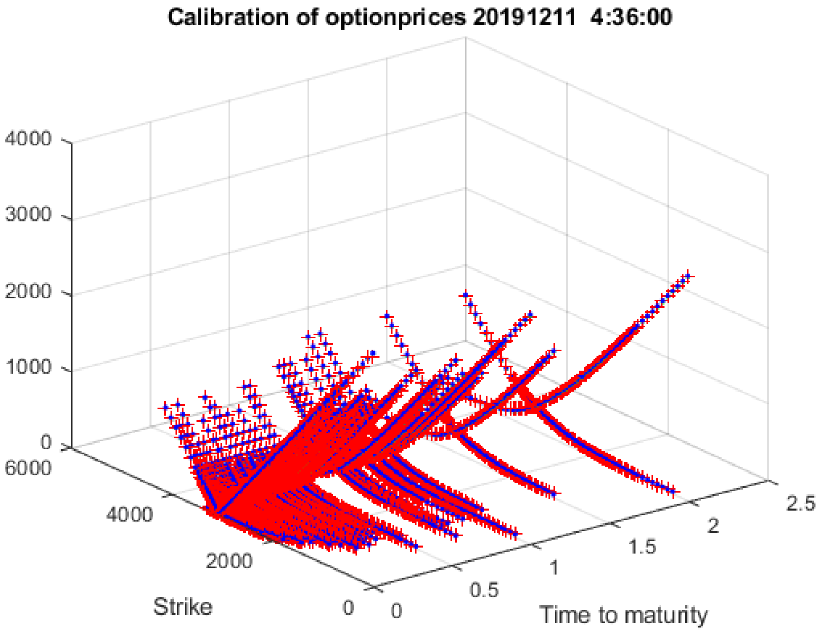

Figure 9: Calibration of option prices using the Heston model

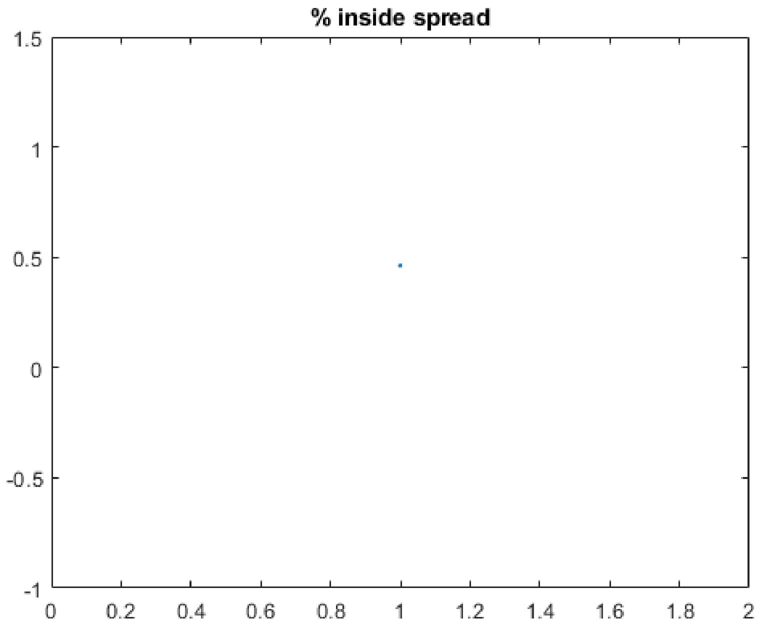

Figure 10: Inside spread of the Heston model

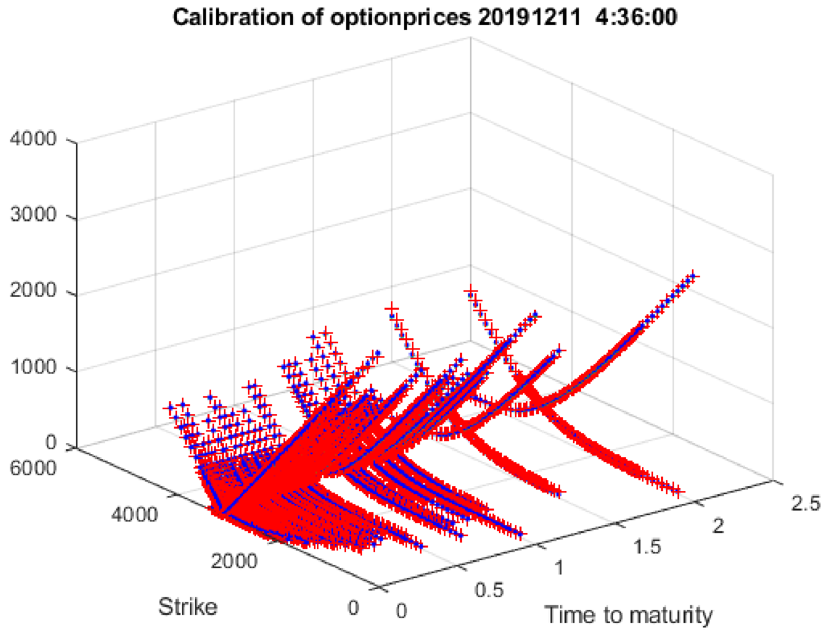

Figure 11: Calibration of option prices using the Triple Heston model

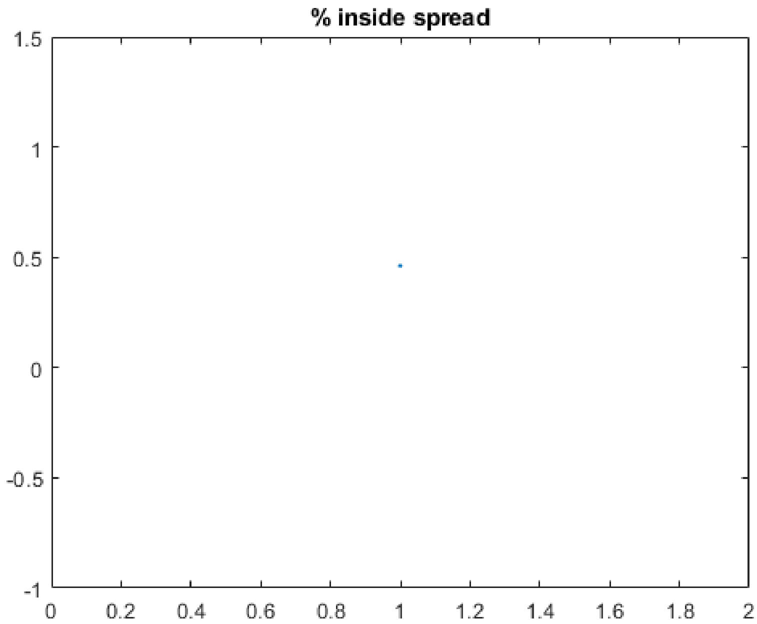

Figure 12: The CBOE computed VIX plotted against the S&P500 index

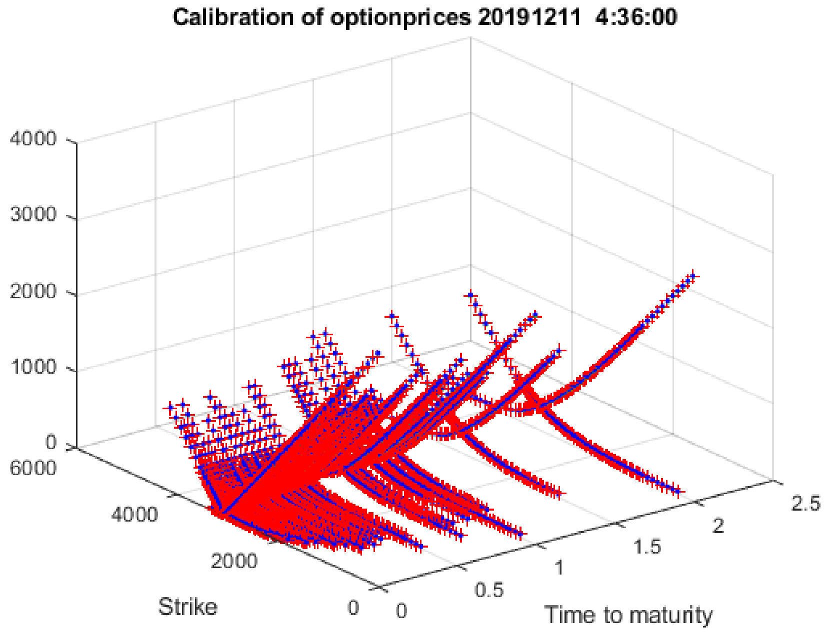

Figure 13: Calibration of option prices using the Quintuple Heston model

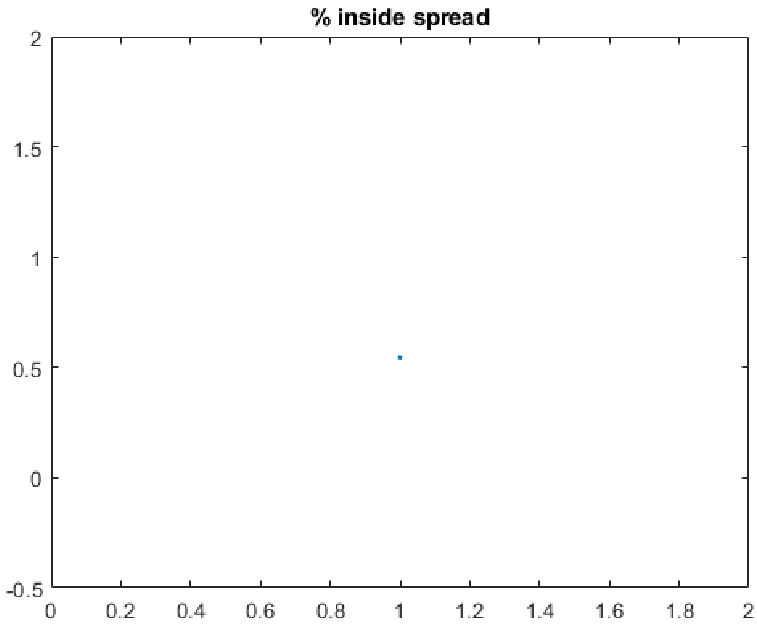

Figure 14: Inside spread of the Quintuple Heston model

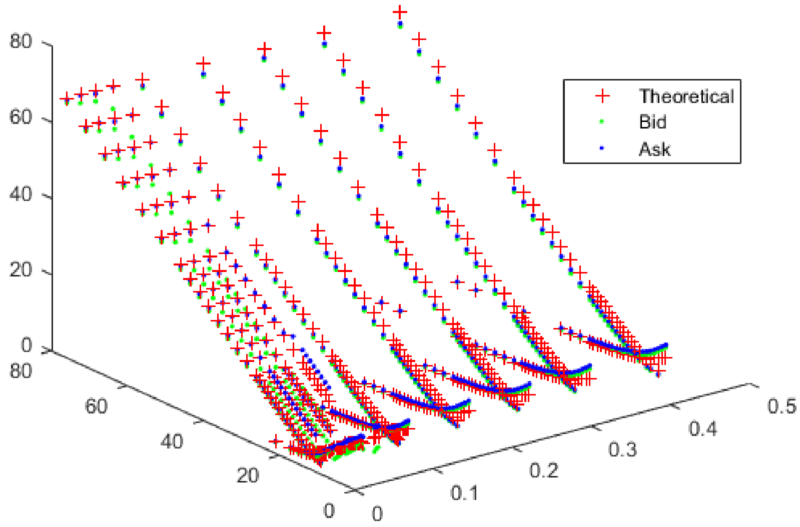

Figure 15: 04.35, VIX options priced with 1 dimensional CIR

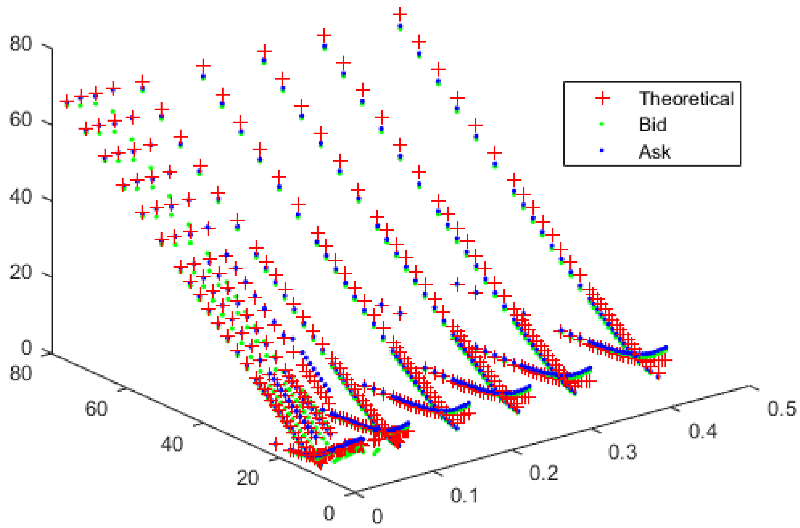

Figure 16: 04:35, VIX options priced with 3 dimensional CIR

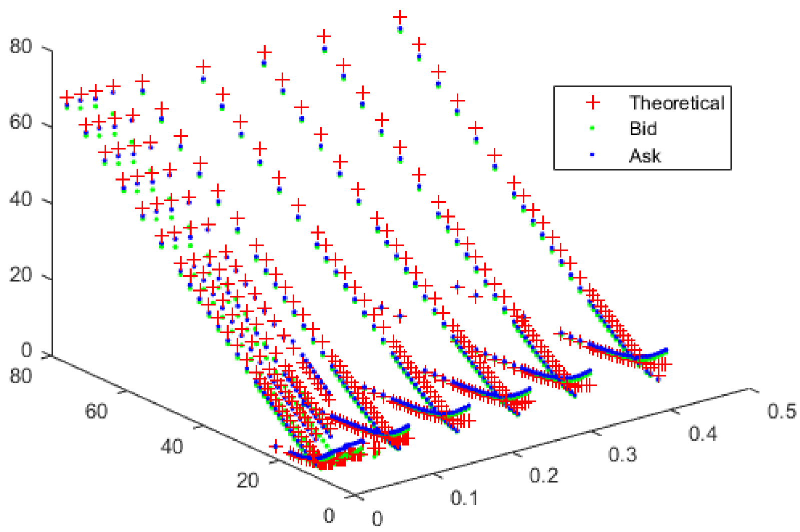

Figure 17: 04:35, VIX options priced with 5 dimensional CIR

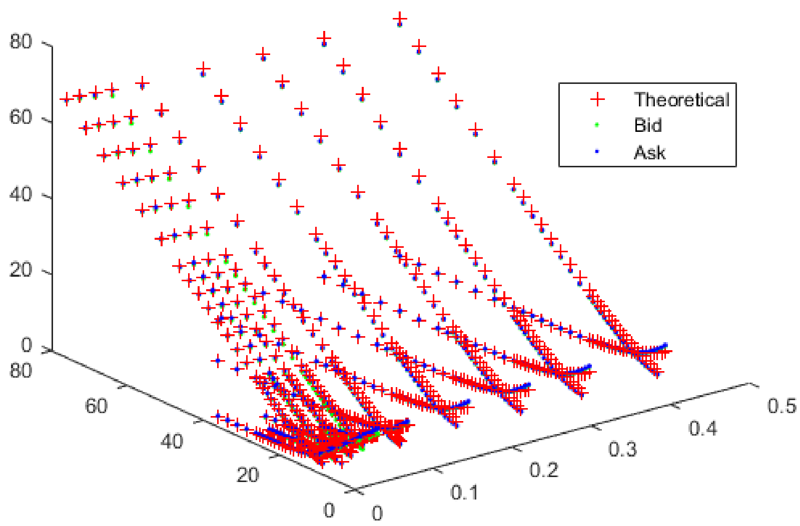

Figure 18: 09:58, VIX options priced with 1 dimensional CIR

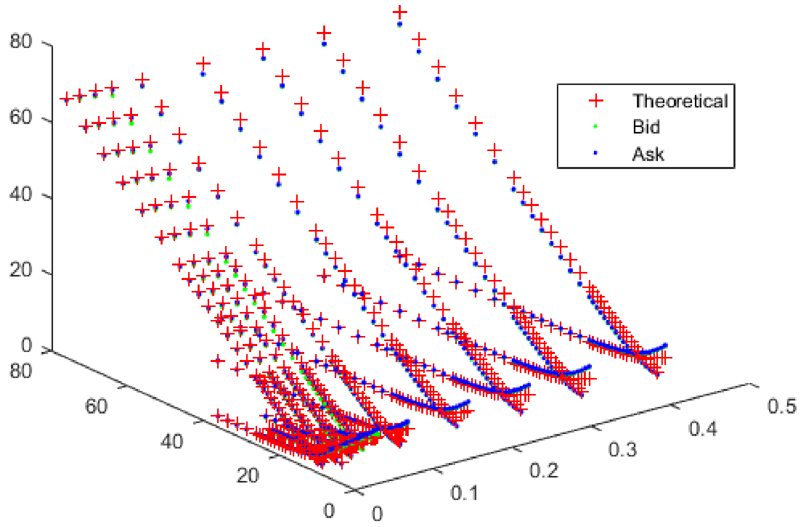

Figure 19: 09:58, VIX options priced with 3 dimensional CIR

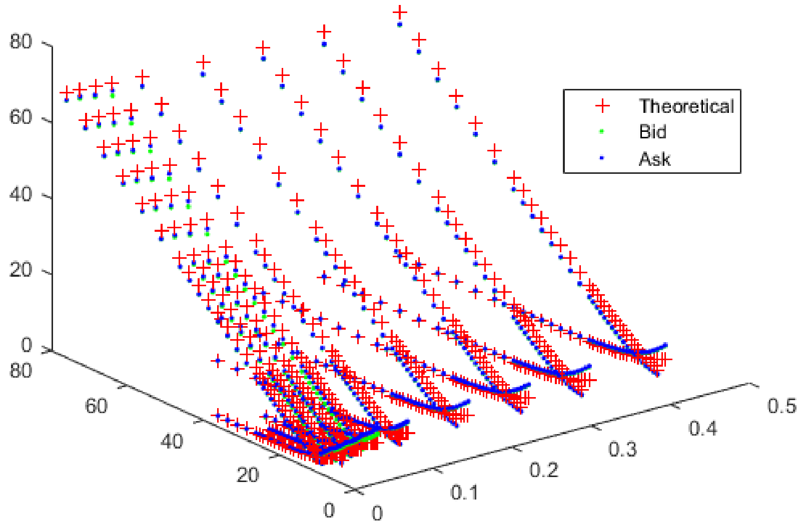

Figure 20: 09:58, VIX options priced with 5 dimensional CIR

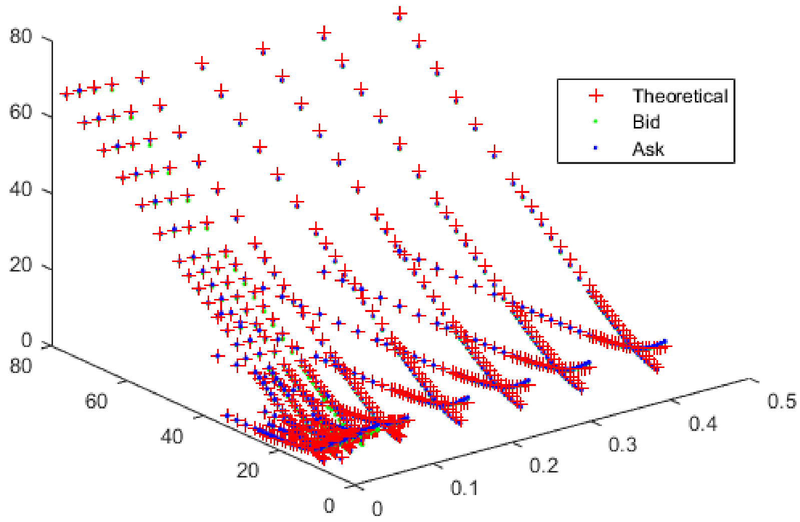

Figure 21: 14:38, VIX options priced with 1 dimensional CIR

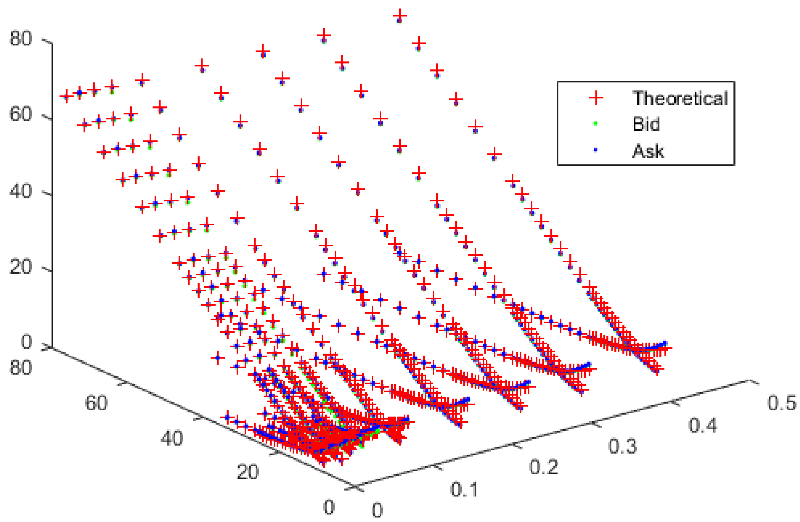

Figure 22: 14:38, VIX options priced with 3 dimensional CIR

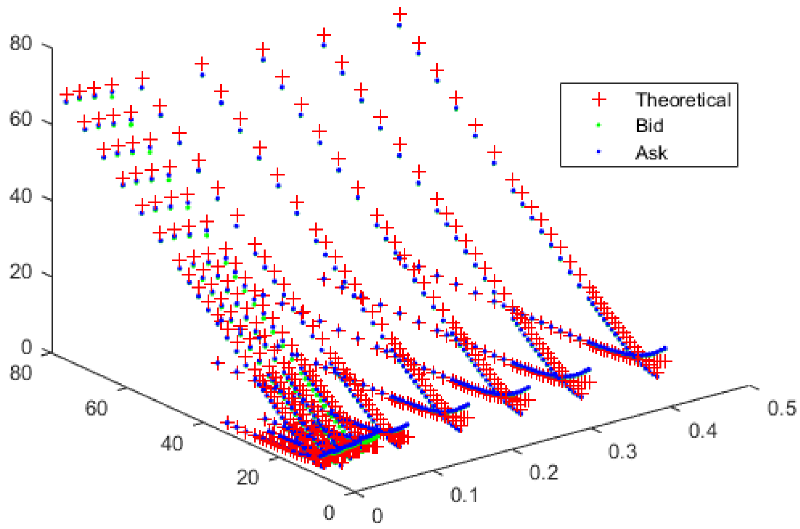

Figure 23: 14:38, VIX options priced with 5 dimensional CIR