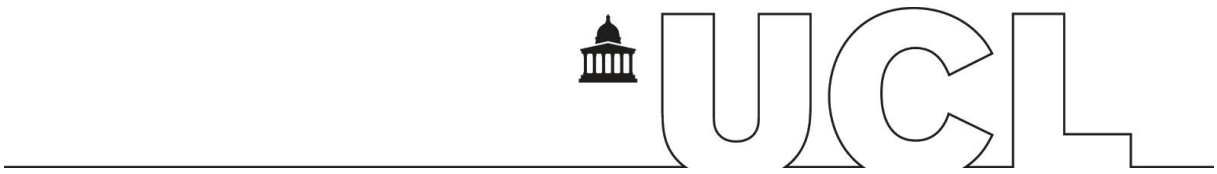


# Analytically Differentiable Expression of the Characteristic Function of the Heston Model with Displacement and its Use in Calibration

**Lin Zhao**

MSc Computational Finance

Department of Computer Science

University College London

Supervised by Dr. Guido Germano

2018

# Disclaimer

This dissertation is submitted as partial requirement for the MSc Computational Finance degree at UCL. It is substantially the result of my own work except where explicitly indicated in the text. The dissertation may be freely copied and distributed provided the source is explicitly acknowledged.

# Acknowledgements

Thanks to Dr. Guido Germano, my supervisor, for seeding the topic of this dissertation and providing me with sage guidance throughout the creation of this work.

Thanks to Dr. Gabriele Pompa for his help expounding on possible future directions and applications of the work. His previous work helped provide many of the fundamental bases for work within this dissertation.

Thanks to Charles Exely for his willingness to collaborate on our related dissertations, which allowed us both to make better progress than we would have alone.

Thanks to Bob M. U. Min, who has been good company to me throughout the last year.

Many thanks to my husband, Chris Down, for providing me with support throughout my dissertation, whether it be moral, financial, or technical, and making sure that I took the time to take a break when I was stressing myself out. Also thanks for brewing me a lot of tea to overcome my academically induced dehydration.

Thank you to my parents for starting me on this journey, despite a heavy financial and emotional burden, over 5 years ago.

# Abstract

The increasing demand of trading and hedging volatility has led to the creation and thriving of a liquid derivative market with the volatility index (VIX) as underlying. This index is defined as the modified volatility of the S&P500 equity index. To give consistent prices of derivatives written on the equity index and of derivatives written on the VIX, stochastic volatility models are used to describe the risk-neutral dynamics of the equity index and its variance. This dissertation introduces analytical gradients of option pricing formulas. Using these analytical gradients, we calibrate a stochastic volatility model in a consistent fashion. We apply this method to the Heston model and to its extension with a displacement, which belongs to the generalized Heston++ class of models. This method has the potential to be extended to calibrate more complex stochastic volatility models. To implement this method, we extend the numerically continuous and easily differentiable expression of the price characteristic function together with its analytical gradient introduced by Cui et al. (2017) for the classic Heston model, and along the approach of the latter paper, we introduce a numerically continuous and easily differentiable form of the volatility characteristic function and its analytical gradient. With these results, we give the analytical gradients of the pricing formulas for both equity options and VIX options. A numerical experiment is conducted to test our theoretical result on the Heston model. Using simulated prices with multiple parameter sets and many initial guesses, 88.8% of the calibrations retrieved the input parameters. For a typical calibration, the average number of iterations is about 30 and the CPU time running on a standard laptop is about half

second. The average times that the objective function is evaluated per iteration is less than 1.2, which is lower than 1.6 when using a numerical gradient. The calibration greatly improved the pricing of both S&P500 index options and VIX options. Due to the time constraint, we did not perform calibration of the Heston model with displacement.

# Contents

# Chapter 1

# Introduction

Due to the increasing market demand of volatility trading, hedging, and risk managing, the Chicago Board Option Exchange (CBOE) introduced the volatility index (VIX) in 1993. After the index was redefined in a more robust way in 2003, CBOE introduced futures on VIX in 2004 and options on it in 2006 (CBOE, 2018). One well-known stylized fact is the negative correlation between return and volatility (Badshah et al., 2016). This makes the out-of-money call option on VIX a valuable asset to hedge against the risk that the market goes down. Futures and options on VIX have been heavily traded ever after their introduction, and pricing them has great interest both for practitioners and academic researchers.

At the centre of VIX derivative pricing lies the dynamics of VIX index. There are broadly two categories in VIX modelling: stand-alone approach and consistent approach (Pompa, 2015). In the stand-alone approach, volatility is modelled directly while the stock price is not considered. The calibration of this kind of model is performed on VIX futures and/or options without considering options written on S&P500 index. These models have the advantage of tractability but are not ideal because of the consistency issue. By definition, the VIX is the volatility that underlies the equity index. This means that a realistic model of VIX should also be able to give proper prices of options written on the equity

index. From a calibration point of view, this means there is information about volatility process contained in the price series of the S&P500 index options. This is a mature and liquid market, if a model can adapt to this market it can extract this information, and provide consistent pricing on both equity index options and VIX options. This property is lost in stand-alone approach because they only model the volatility process. Models that account for this property are categorized as consistent approach. In a consistent approach, the dynamic of stock price is modelled and the volatility is extracted in a model-free way. This type of model can be calibrated using prices from both the equity index option market and the VIX derivative market. The resulting model would contain more complete information and have the ability to give consistent pricing (Pompa, 2015).

According to Pompa (2015), stochastic volatility affine models are predominant within literature studying consistent modelling. A fundamental and arguably most popular model in this category is the Heston model (Heston, 1993; Rouah, 2013) defined as

$$dS(t) = \mu S(t)\,dt + \sqrt{\upsilon(t)}S(t)\,dW_1(t), \tag{1.1a}$$

$$d\upsilon(t) = k(\overline{\upsilon} - \upsilon(t))\,dt + \sigma\sqrt{\upsilon(t)}\,dW_2(t), \tag{1.1b}$$

$$dW_1(t)\,dW_2(t) = \rho\,dt, \tag{1.1c}$$

where $S(t)$ is the stock price and $\upsilon(t)$ is its variance. The variance $\upsilon(t)$ follows a mean-reverting process. The parameter $k$ is the mean-reversion rate, controlling the speed of mean reversion; $\overline{\upsilon}$ is the long-term mean of variance; and $\sigma$ is the volatility of volatility. The parameter $\rho$ is the correlation between the Brownian motions $W_1(t)$ and $W_2(t)$ that drive these two processes. The initial value of the variance $\upsilon_0$ is the fifth parameter of the model.

Pacati et al. (2018) discussed different consistent generalizations of the Heston model. It has been shown that a displaced diffusion term greatly helps the model in fitting the term structure observed from the VIX derivative market. The marginal effect of this deterministic displacement is still significant even after the Heston model was extended

by an extra stochastic volatility process (Rouah, 2013) and jumps (Lin, 2007). Inspired by this fact, we aim to develop a method of fast and consistently calibrating Heston model together with this displacement. This could be a good starting point for developing a general method applicable to a big class of consistent models, thus benefit both further researches and practitioners. After adding this displacement, the new model has the form

$$dS(t) = \mu S(t)\,dt + \sqrt{\upsilon(t) + \phi(t)}S(t)\,dW_1(t), \tag{1.2a}$$

$$d\upsilon(t) = k(\overline{\upsilon} - \upsilon(t))\,dt + \sigma\sqrt{\upsilon(t)}\,dW_2(t), \tag{1.2b}$$

$$dW_1(t)\,dW_2(t) = \rho\,dt, \tag{1.2c}$$

where $\phi(t) \geq 0$ and $\phi(0) = 0$. Through this dissertation, we will refer to this model as the Heston++ model, since it belongs to the class of stochastic volatility model endowed with a displacement. However, it is not exactly the same as the Heston++ in Pacati et al. (2018) or Pompa (2015). The Heston++ model referred in those two papers is actually a two-factor stochastic volatility model with correlated jumps in price and the first variance process together with idiosyncratic jumps in the first variance process. They labelled this model also as 2-SVCVJ++ model. By this convention, our Heston++ model is a SV++ model.

In practice, CBOE calculates the VIX quotation as a model-free replication of the risk-neutral integrated variance in the next 30 days. The used portfolio contains a series of out-of-money options on S&P500. In this paper, since we will adopt the pricing formula from Pacati et al. (2018) and Pompa (2015), we will also use the same definition of VIX quotation as in their papers. This definition is a standard one for variance index that is expressed as the risk-neutral expectation of a log-contract (Jiang et al., 2005, Lin, 2007, Duan et al., 2010, Zhang et al., 2010),

$$\left(\frac{\mathrm{VIX}(t,\overline{\tau})}{S_0}\right)^2 := -\frac{2}{\overline{\tau}}\mathbb{E}\left[\log\frac{S(t+\overline{\tau})}{F(t,t+\overline{\tau})}\right] = \frac{1}{\overline{\tau}}\left(a(\overline{\tau};k)\upsilon(t) + b(\overline{\tau};\overline{\upsilon},k)\right), \tag{1.3}$$

where $F(t, t+\overline{\tau}) = e^{(r-q)\overline{\tau}} S(t)$ is the forward price; $a(\overline{\tau}; k)$ and $b(\overline{\tau}; \overline{v}, k)$ will be defined in later chapters. We also set $S_0 = S(0) = 100$ since $S$ corresponds to the value of an index whose initial value is usually set to 100.

The convention of CBOE is $\overline{\tau} = 30$. Under the Heston++ model, the VIX is defined as

$$\left(\frac{\mathrm{VIX}_{\mathrm{Heston++}}(t, \overline{\tau})}{S_0}\right)^2 = \left(\frac{\mathrm{VIX}(t, \overline{\tau})}{S_0}\right)^2 + \frac{1}{\overline{\tau}} I_\phi(t, t+\overline{\tau}), \tag{1.4}$$

where $I_\phi(t, t+\overline{\tau}) = \int_t^{t+\overline{\tau}} \phi(s)\, ds$.

In this dissertation, following mainly in the approach of Cui et al. (2017), we propose a method to efficiently and consistently calibrate Heston model and displacement across multiple assets. In Section 2, we briefly review the most relevant literature. In section 3, we formalize the problem, derive the analytical gradient, and propose the calibration algorithms for Heston model and Heston++ model. In Section 4, we demonstrate the calibration algorithms by complete sample calibrations performed on the simulated data. We will also analyse the extra cost of calibrating the displacement. In Section 5, we present the main conclusions, difficulties and possible future directions of this topic.

Through this dissertation, bold letters denote matrices, column vectors, or sets; subscript $X$ corresponds with variables related to log stock price and $v$ are attached to variables related to the volatility process.

# Chapter 2

# Literature review

The main idea of consistent calibration is fitting the model to observations from multiple markets to obtain the optimal set of parameters. An efficient method to quickly and fully calibrate Heston model using a single volatility surface from equity option market was proposed by Cui et al. (2017). In this paper, they defined the task as a non-linear least-square problem. Their goal was to minimize the objective function using analytical gradient, thus a characteristic function of stock price in a expression that is both numerically continuous and easily differentiable was required. To find this characteristic function, they explored the literatures, examined a few different representations of characteristic functions, and eventually, on top of the expression proposed by Del Baño Rollin et al. (2010), they reconstructed a new representation that satisfied both features mentioned above. From this characteristic function, they managed to find a numerically efficient expression of the gradient for the minimization problem, then, using Levenberg-Marquardt method, their algorithm succeeded in finding the global optimal parameters while enjoying performance advantage at the same time.

From their work, we can summarize a few key steps in developing a calibration algorithm using analytical gradient: define the objective function; find a expression of characteristic function that is numerically continuous and easily differentiable; take derivative of pricing

formula; and apply Levenberg-Marquardt method to search for optimal parameters. In the rest of this section, we will briefly review a few papers that contain the most relevant results for the topics mentioned above.

## 2.1 Objective function

Schoutens et al. (2004) provided four reasonable error measures: the root mean square error $\varepsilon_{\text{RMS}}$, the average absolute error $\varepsilon_{\text{AAP}}$ as a percentage of the mean price , the average absolute error $\varepsilon_{\text{AA}}$ and the average relative percentage error $\varepsilon_{\text{ARP}}$. Their definitions are

$$\varepsilon_{\text{RMS}} = \sqrt{\frac{1}{N_{\text{option}}} \sum_{i=1}^{N_{\text{option}}} (P_i^{\text{MKT}} - P_i^{\text{MDL}})^2}, \tag{2.1a}$$

$$\varepsilon_{\text{AAP}} = \frac{1}{N_{\text{option}} \overline{P}^{\text{MKT}}} \sum_{i=1}^{N_{\text{option}}} |P_i^{\text{MKT}} - P_i^{\text{MDL}}|, \tag{2.1b}$$

$$\varepsilon_{\text{AA}} = \frac{1}{N_{\text{option}}} \sum_{i=1}^{N_{\text{option}}} |P_i^{\text{MKT}} - P_i^{\text{MDL}}|, \tag{2.1c}$$

$$\varepsilon_{\text{ARP}} = \frac{1}{N_{\text{option}}} \sum_{i=1}^{N_{\text{option}}} \frac{|P_i^{\text{MKT}} - P_i^{\text{MDL}}|}{P_i^{\text{MKT}}}, \tag{2.1d}$$

where we denote $P^{\text{MKT}}$ as market price of a option, $P^{\text{MDL}}$ as model price of a option, $N_{\text{option}}$ as number of options, and $\overline{P}^{\text{MKT}}$ as the mean market option price.

From these definitions, we can extract some fundamental ideas about designing a objective function. The central part is summing up the difference, squared or as absolute values, between market price and model price. These differences can be normalised by either market price or the mean market price, thus turned into a percentage. The whole sum may also be turned into an average value by dividing the number of options.

To perform a consistent calibration, the objective function needs to contain multiple terms representing different markets/assets. This type of calibration can be found in Pacati et al. (2018) and Pompa (2015). Here we quote Equation (12) of Pacati et al. (2018) as an

example to present a few element that should be considered when constructing this kind of objective function

$$L = \frac{1}{N_{\text{SPX}}} \sum_{i=1}^{N_{\text{SPX}}} \left( \frac{V_{\text{SPX},i}^{\text{MKT}} - V_{\text{SPX},i}^{\text{MDL}}}{V_{\text{SPX},i}^{\text{MKT}}} \right)^2 + \frac{1}{N_{\text{Fut}}} \sum_{i=1}^{N_{\text{Fut}}} \left( \frac{P_{\text{Fut},i}^{\text{MKT}} - P_{\text{Fut},i}^{\text{MDL}}}{P_{\text{Fut},i}^{\text{MKT}}} \right)^2 + \frac{1}{N_{\text{VIX}}} \sum_{i=1}^{N_{\text{VIX}}} \left( \frac{V_{\text{VIX},i}^{\text{MKT}} - V_{\text{VIX},i}^{\text{MDL}}}{V_{\text{VIX},i}^{\text{MKT}}} \right)^2, \quad (2.2)$$

where $V$ is implied volatility and $P$ is the value of a contract. In this function, three markets/assets were involved: S&P500 index option, VIX future, VIX option, which were represented by the subscripts SPX, Fut, and VIX. The total number of each assets were represented by $N$ together with the respective subscript. We noticed that this objective function contained three weighted sums of squares, and inside the squares were the gaps between a price/implied volatility observed from the markets and its counterparty calculated from the model. Those gaps were in percentages normalized by the market prices or implied volatilities, thus made the gaps from different markets comparable. On the other hand, the weights made sure that the market providing fewer observations do not get dominated by the market providing more. We can also see this objective function as a sum of average percentage price/volatility gaps over three markets .

## 2.2 Different expressions of price characteristic function

Another critical step for the calibration of Heston model in Cui et al. (2017) was obtaining the analytical gradient of the objective function. And the key to this gradient was the analytical gradient of characteristic function of price. As one major block of our consistent calibration, the gradient of characteristic function of stock price will be presented and discussed in Section 3.

As mentioned earlier, in order to obtain a numerically efficient gradient of a characteristic function, one needs to first discuss the numerical continuity of the representation. In Albrecher et al. (2007), the continuity of two representations of price characteristic function of Heston model was examined. The first one was the original characteristic function given in the paper Heston (1993)

$$\begin{aligned}\phi_S(\boldsymbol{\theta}; u, \tau) &= \mathbb{E}\left[\exp\left(iu\log\frac{S(t)}{S_0}\right)\right] \\ &= \exp\left\{iu\log\frac{F}{S_0} + \frac{k\overline{v}}{\sigma^2}\left[(\xi+d)\tau - 2\log\frac{1-g_1e^{d\tau}}{1-g_1}\right] + \frac{v_0}{\sigma^2}(\xi+d)\frac{1-e^{d\tau}}{1-g_1e^{d\tau}}\right\}, \quad (2.3)\end{aligned}$$

where

$$F \coloneqq S(t)e^{(r-q)\tau}, \quad (2.4a)$$

$$\xi \coloneqq k - \sigma\rho iu, \quad (2.4b)$$

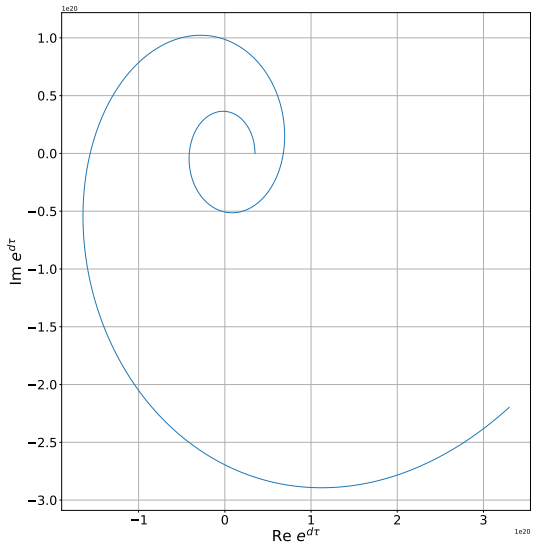
$$d \coloneqq \sqrt{\xi^2 + \sigma^2(u^2+iu)}, \quad (2.4c)$$

$$g_1 \coloneqq \frac{\xi+d}{\xi-d}. \quad (2.4d)$$

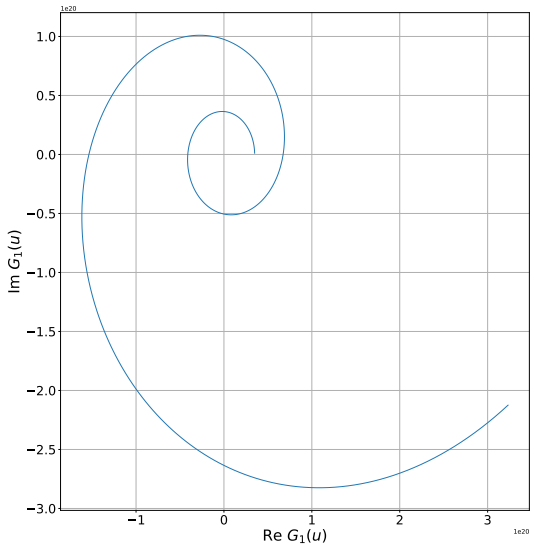
$r$ is the risk-free interest rate and $q$ is the dividend rate of the stock. This form, as first pointed out by Kahl et al. (2005), is not continuous when evaluated as a function of $u$. The reason is that complex functions such as the complex logarithm and the complex square root assume more than one values. When dealing with a multivalued complex function, most programming languages will chose the principal value as the output. Although a natural choice, this is not always the correct one and causes a discontinuity. The following graphs are produced using the parameter set in Table 2.1.

| Model parameters | | | | | Market parameters | | | |
|---|---|---|---|---|---|---|---|---|
| $k$ | $\overline{v}$ | $\sigma$ | $\rho$ | $v_0$ | $S_0$ | $K$ | $r$ | $q$ |
| 3.00 | 0.10 | 0.25 | -0.80 | 0.08 | 1.00 | 1.10 | 0.02 | 0 |

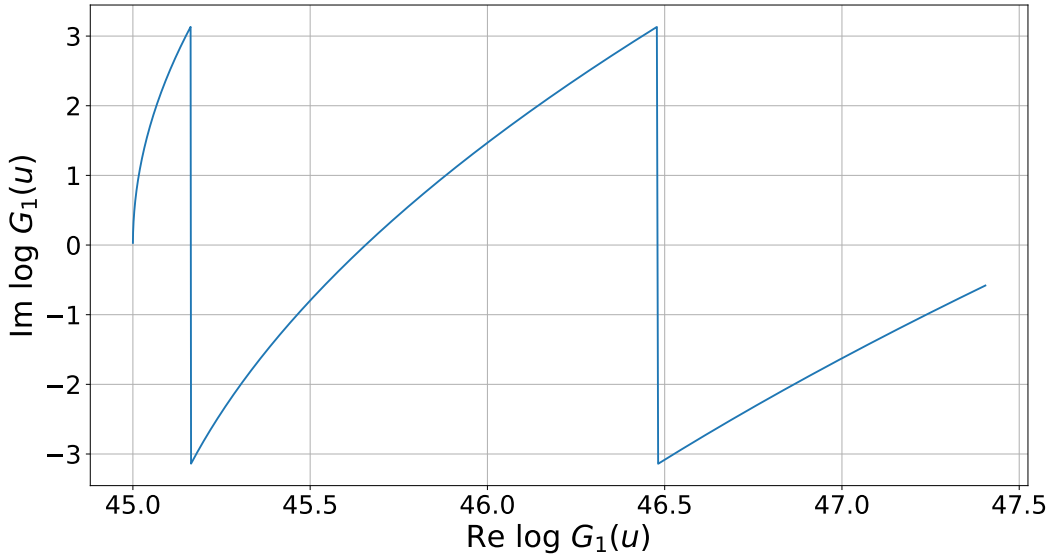
Table 2.1: A set of parameters used to generate Figure 2.1. Five parameters of the Heston model within their reasonable ranges and four market parameters.




(a) Spiral structure of $e^{d\tau}$.




(b) Spiral structure of $G_1(u)$.




(c) Jumps of $\log G_1(u)$.

Figure 2.1: The cause of jumps for characteristic function (Equation (2.3)). $u \in (0, 4)$ and $\tau = 15$. (a) the graph of $e^d\tau$, which has a obvious spiral structure. (b) the graph of $G_1(u)$, where the spiral structure is preserved. (c) the graph of $\log G_1(u)$, we can see two jumps in this graph which correspond with the two time that $G_1(u)$ cross the negative real line.

As shown in Figure 2.1a, the function $e^{d\tau}$ has the structure of a spiral that repeatedly cross the negative half of the real line. This structure is preserved by the form

$$G_1(u) = \frac{1 - g_1 e^{d\tau}}{1 - g_1}, \tag{2.5}$$

as shown in Figure 2.1b. Each time the function cross the negative part of the real line, the principle value of logarithm of $G_1$ jumps, and this is plotted in Figure 2.1c. This jump happened twice when $u$ is within the range of $(0, 4)$, and $\tau = 15$.

The second form of the characteristic function of price originally proposed in Schoutens et al. (2004)

$$\begin{aligned}\phi_S(\boldsymbol{\theta}; u, \tau) &= \mathbb{E}\left[\exp\left(iu \log \frac{S(t)}{S_0}\right)\right] \\ &= \exp\left\{iu \log \frac{F}{S_0} + \frac{k\overline{v}}{\sigma^2}\left[(\xi - d)\tau - 2\log\frac{1 - g_2 e^{-d\tau}}{1 - g_2}\right] + \frac{v_0}{\sigma^2}(\xi - d)\frac{1 - e^{-d\tau}}{1 - g_2 e^{-d\tau}}\right\},\end{aligned} \tag{2.6}$$

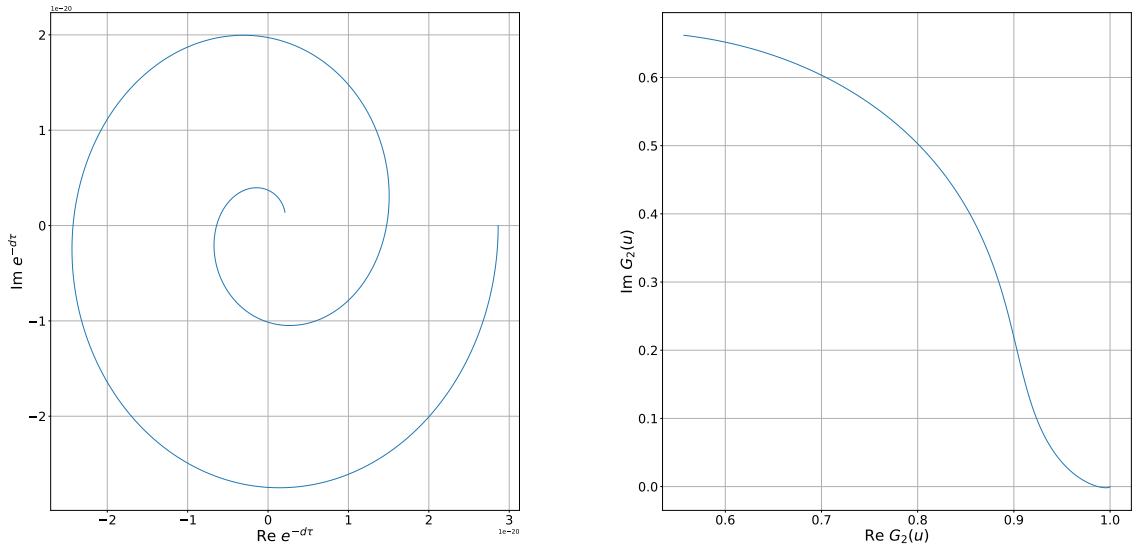
where

$$g_2 \coloneqq \frac{\xi - d}{\xi + d} = \frac{1}{g_1}. \tag{2.7}$$

This form contains $e^{-d\tau}$ which also has a similar spiral structure. However, this spiral structure is not present in

$$G_2(u) = \frac{1 - g_2 e^{-d\tau}}{1 - g_2}, \tag{2.8}$$

thus, does not cause a jump in the logarithm function. In Figure 2.2a and (2.2b), we have plotted these two functions. Since it is obvious that $G_2(u)$ does not cross the negative real line, we omit the plot of $\log G_2(u)$.




(a) Spiral structure of $e^{-d\tau}$, $u \in (0, 4)$.

(b) Graph of $G_2(u)$, $u \in (0, 500)$.

Figure 2.2: Continuity analysis of characteristic function (Equation (2.6)), $\tau = 15$. (a) graph of $e^{-d\tau}$, due to the rapid outward movement of the spiral, we only show the graph of $u \in (0, 4)$ here. (b) a graph of $G_2(u)$. We plotted its graph on a relatively big domain ($u \in (0, 500)$) and showed that it does not preserve the spiral structure.

To summarize, as done in Cui et al. (2017), we plotted the real part of each representation as a function of $u$ to show the different continuity property in Figure 2.3. We can see an obvious jump when $u = 1$ on the plot of expression in Equation (2.3), while expression in Equation (2.6) looks continuous. The continuity of the expression in Equation (2.6) is formally proved by Lord et al. (2010).

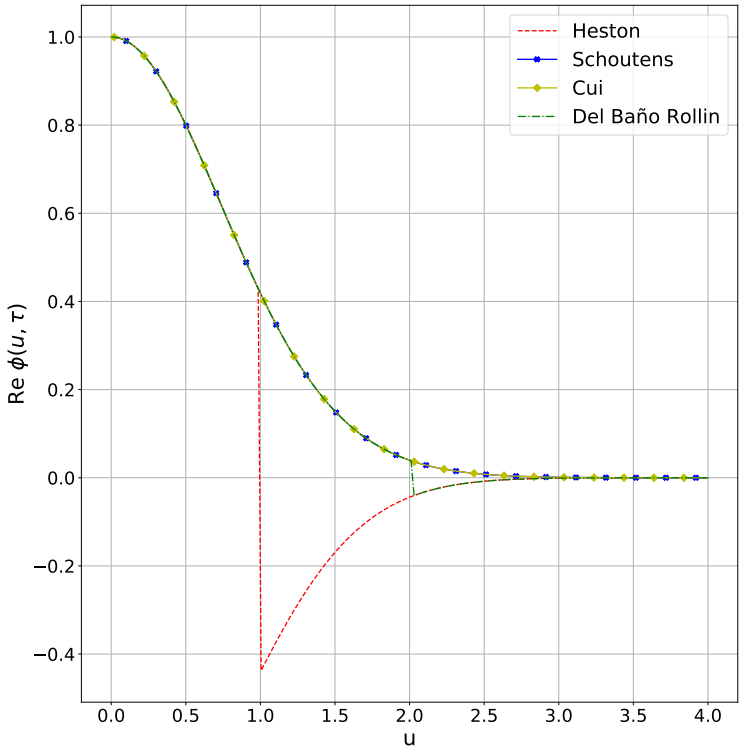



Figure 2.3: Real parts of characteristic functions (Equation (2.3), (2.6), (2.9), and (3.8)) as functions of $u$. Jumps are observed at $u = 1$ and $u = 2$. In all functions, $u \in (0, 4)$ and $\tau = 15$.

To obtain a form of the price characteristic function that is easier to handle analytically, Del Baño Rollin et al. (2010) re-derived this characteristic function through the complex moment-generating function. Starting from the Heston model (Heston, 1993), using Ito formula, they constructed a PDE of $h(u, v, t) = \mathbb{E}\left[e^{uX(t)+vV(t)}\right]$ where $X(t) = \log(S(t)/S_0)$. Solving this PDE gives an expression of $\mathbb{E}\left[e^{uX(t)+vV(t)}\right]$. Then they turned off $V(t)$ by setting $v = 0$, thus got the expression of the moment generating function $M(u) = \mathbb{E}\left[e^{\{uX(t)\}}\right]$, which is Equation (3) in Del Baño Rollin et al. (2010). Substituting

$u$ by $iu$ in this expression gives $\mathbb{E}\left[\exp\left\{iuX(t)\right\}\right]$, which is the characteristic function. The closed-form expression is Equation (6) in Del Baño Rollin et al. (2010). Here we quote a equivalent representation from Cui et al. (2017), Equation (14)

$$\phi_S(\boldsymbol{\theta}; u, \tau) = \mathbb{E}\left[\exp\left(iu \log \frac{S(t)}{S_0}\right)\right] = \exp\left(iu \log \frac{F}{S_0} - \frac{k\overline{v}\rho\tau iu}{\sigma} - A\right) B^{2k\overline{v}/\sigma^2}, \quad (2.9)$$

where

$$A \coloneqq \frac{A_1}{A_2}, \quad (2.10a)$$

$$A_1 \coloneqq (u^2 + iu) \sinh \frac{d\tau}{2}, \quad (2.10b)$$

$$A_2 \coloneqq \frac{d}{v_0} \cosh \frac{d\tau}{2} + \frac{\xi}{v_0} \sinh \frac{d\tau}{2}, \quad (2.10c)$$

$$B \coloneqq \frac{de^{k\tau/2}}{v_0 A_2}. \quad (2.10d)$$

This form is not necessarily more compact than the one in Equation (2.3) or in Equation (2.6), but it is in general easier to handle. Multiple exponential functions were replaced by hyperbolic functions making it easier to take derivatives. This was the prototype of the final representation used in Cui et al. (2017), which can also been seen in Figure 2.3. Numerically this expression is not a continuous function of $u$ (Cui et al., 2017) as can be seen in Figure 2.3.

## 2.3 A displacement of the diffusion term

A non-negative displacement acting as a lower bound of volatility was discussed in Pompa (2015) and Pacati et al. (2018). When we model volatility processes, by definition, they should always be non-negative, which means there should be a lower bound at zero for these processes. However, this lower bound may not be enough. In reality, the volatility almost always hovers in a range that some distance away from zero, and this range may vary with time. This means that while zero might be a reasonable bound for the whole

process, it is not a good one for every time point. In another word, to capture more detail of the volatility process, the lower bound for the whole process should not be a single constant. The displacement studied in Pompa (2015) and Pacati et al. (2018), $\phi(t)$, served as a deterministic and time-related lower bound to the volatility process at time $t$. It was defined as a function with range greater or equal to zero, which intuitively matches with its role as a lower bound of volatility. This is also the only assumption about this function since it is allowed to take any form. These two papers also provided close-form pricing for VIX futures and options assuming the stock price process has a displaced diffusion term. In these pricing functions, this displacement was used as "an analytically tractable correction for the corresponding pricing formulas for the non-displaced models" (Pompa, 2015). A numerical method to calibrate this displacement was discussed in the appendixes of these papers. In this dissertation, we will develop a calibration scheme to find $\phi(t)$ together with other Heston model parameters through an analytical gradient.

# Chapter 3

# Methodology

In this chapter, we will first define our objective function, and then find the gradients of the pricing formulas with respect to the Heston model parameter set. In the last section, we will extend this gradient to include the derivatives of the integrations of the displacement. In this way, we can include the calibration of the displacement into our general scheme for the Heston parameters.

## 3.1 Objective function for consistent calibration

As shown by Equation (2.2), the objective function of a consistent calibration may contain multiple terms that represent multiple assets. Here we define an objective function for calibrating parameters across S&P500 option market and VIX option market

$$L = \frac{1}{2N_{\text{SPX}}} \sum_{i=1}^{N_{\text{SPX}}} \left( \frac{P^{\text{MKT}}_{\text{SPX},i} - P^{\text{MDL}}_{\text{SPX},i}}{P^{\text{MDL}}_{\text{SPX},i}} \right)^2 + \frac{1}{2N_{\text{VIX}}} \sum_{i=1}^{N_{\text{VIX}}} \left( \frac{P^{\text{MKT}}_{\text{VIX},i} - P^{\text{MDL}}_{\text{VIX},i}}{P^{\text{MDL}}_{\text{VIX},i}} \right)^2, \tag{3.1}$$

where $N$ denotes the number of options of the type specified by its subscript. In this stage, we assume that the stock price follows the stochastic process defined in Equation (1.1), then $P_{\text{SPX},i} = P_{\text{SPX}}(\boldsymbol{\theta}; K_i, \tau_i)$ where $\boldsymbol{\theta}$ is the parameter vector $[\upsilon_0, \overline{\upsilon}, \rho, k, \sigma]^{\mathsf{T}}$, $K_i$ is

the strike price of option $i$ and $\tau_i$ is the time to maturity of option $i$. Each sum of the squared difference is then weighted according to the number of options observed in that market. This can also be seen as taking the average of squared relative differences in each market.

To convert this objective function into vector form, we define

$$r_{\mathrm{SPX},i}(\boldsymbol{\theta}) \coloneqq \frac{P^{\mathrm{MKT}}_{\mathrm{SPX},i} - P^{\mathrm{MDL}}_{\mathrm{SPX},i}}{P^{\mathrm{MDL}}_{\mathrm{SPX},i}}, \tag{3.2a}$$

$$r_{\mathrm{VIX},i}(\boldsymbol{\theta}) \coloneqq \frac{P^{\mathrm{MKT}}_{\mathrm{VIX},i} - P^{\mathrm{MDL}}_{\mathrm{VIX},i}}{p_{\mathrm{VIX},i}}, \tag{3.2b}$$

$$\boldsymbol{r}_{\mathrm{SPX}}(\boldsymbol{\theta}) \coloneqq [r_{\mathrm{SPX},1}(\boldsymbol{\theta}), r_{\mathrm{SPX},2}(\boldsymbol{\theta}), r_{\mathrm{SPX},3}(\boldsymbol{\theta}), \ldots, r_{\mathrm{SPX},N_{\mathrm{SPX}}}(\boldsymbol{\theta})]^{\mathsf{T}}, \tag{3.2c}$$

$$\boldsymbol{r}_{\mathrm{VIX}}(\boldsymbol{\theta}) \coloneqq [r_{\mathrm{VIX},1}(\boldsymbol{\theta}), r_{\mathrm{VIX},2}(\boldsymbol{\theta}), r_{\mathrm{VIX},3}(\boldsymbol{\theta}), \ldots, r_{\mathrm{VIX},N_{\mathrm{VIX}}}(\boldsymbol{\theta})]^{\mathsf{T}}. \tag{3.2d}$$

Now Equation (3.1) can be expressed more compactly as

$$L = \frac{1}{2N_{\mathrm{SPX}}} r^2_{\mathrm{SPX}} + \frac{1}{2N_{\mathrm{VIX}}} r^2_{\mathrm{VIX}}, \tag{3.3}$$

where $r^2_{\mathrm{SPX}} \coloneqq \boldsymbol{r}^{\mathsf{T}}_{\mathrm{SPX}}(\boldsymbol{\theta})\boldsymbol{r}_{\mathrm{SPX}}(\boldsymbol{\theta})$ and $r^2_{\mathrm{VIX}} \coloneqq \boldsymbol{r}^{\mathsf{T}}_{\mathrm{VIX}}(\boldsymbol{\theta})\boldsymbol{r}_{\mathrm{VIX}}(\boldsymbol{\theta})$.

There are 5 parameters in the Heston model, and we should expect that both numbers of S&P500 options and VIX options to be higher than 5. Thus the calibration problem is overdetermined.

The major difference between our objective function and Equation (2.2) is that we use the difference between option prices instead of implied volatilities. As discussed in Cui et al. (2017), the squared difference of implied volatilities are used in the industry, however, for our purpose of calibrating the model through an analytical gradient, the availability of closed-form formula is important. Compared with the implied volatility, the pricing functions of options are in general widely studied and easy to get.

Besides the specific form of the objective function, the strike prices and the maturities used in the calibration are also important choices. We will discuss our choice in Chapter 4.

To find the gradient of our objective function $L$, we need the Jacobian matrices of the vectors $\boldsymbol{r}_{\text{SPX}}$ and $\boldsymbol{r}_{\text{VIX}}$. And we will solve this part of the problem in the next three sections.

## 3.2 Gradient of equity index option price

In this section, we will find the Jacobian matrix $\boldsymbol{J}_{\text{SPX}} \coloneqq \nabla \boldsymbol{r}_{\text{SPX}}^{\mathsf{T}} \in \mathbb{R}^{m \times N_{\text{SPX}}}$ where $m = \dim \boldsymbol{\theta}$. The elements of this matrix are

$$J_{\text{SPX},ij} = \frac{\partial r_{\text{SPX},j}}{\partial \theta_i} = \frac{1}{p_{\text{SPX},j}} \frac{\partial P_{\text{SPX}}(\boldsymbol{\theta}; K_j, \tau_j)}{\partial \theta_i}, \tag{3.4}$$

which means that each column in this matrix is a gradient vector evaluated at strike $K_j$ and maturity $\tau_j$, and normalized by the market price observed with the same strike price and maturity, $\frac{1}{p_{\text{SPX},j}} \nabla P_{\text{SPX}}(\boldsymbol{\theta}; K_j, \tau_j)$. Next, we need to find gradient of $P_{\text{SPX}}$.

In Pacati et al. (2018), the arbitrage-free price at time $t$ of a European call option on underlying $S(t)$, with strike price $K$ and maturity $\tau = T - t$ is given in closed-form by Equation (6). When we set to zero the displacement term in this equation, we get the representation

$$\begin{aligned} &P_{\text{SPX}}(\boldsymbol{\theta}; K, \tau) = S(t) e^{-q\tau} \\ &- \frac{1}{\pi} \sqrt{S(t)K} e^{-\frac{1}{2}(r+q)\tau} \int_0^\infty \text{Re}\left[ e^{iu\kappa - i(u-\frac{i}{2})[X(t)-Y(t)+(r-q)\tau]} \phi_X\left(\boldsymbol{\theta}; u - \frac{i}{2}\right)\right] \frac{1}{u^2 + 1/4}\, du, \end{aligned} \tag{3.5}$$

where $\kappa = \log \frac{S(t)}{K} + (r-q)\tau$ and $Y(t)$ is defined by the equation $S(t) = S_0 e^{(r-q)t + Y(t)}$. The gradient of this pricing formula can be expressed as

$$\begin{aligned} &\nabla P_{\text{SPX}}(\boldsymbol{\theta}; K, \tau) = \\ &- \frac{1}{\pi} \sqrt{S(t)K} e^{-\frac{1}{2}(r+q)\tau} \int_0^\infty \text{Re}\left[ e^{iu\kappa - i(u-\frac{i}{2})[X(t)-Y(t)+(r-q)\tau]} \nabla \phi_X\left(\boldsymbol{\theta}; u - \frac{i}{2}\right)\right] \frac{1}{u^2 + 1/4}\, du, \end{aligned} \tag{3.6}$$

where $\nabla\phi_X(\boldsymbol{\theta}; u, \tau)$ is available in Cui et al. (2017), Section 3.3.1.

As we discussed in the previous section, before taking derivative of a characteristic function, one need to examine the property of numerical continuity. Del Baño Rollin et al. (2010) proposed a representation of the Heston price characteristic function that is easy to differentiate. However, it is not numerically continuous as can be seen from the spiral structure of $A_2$ (Equation 2.10). Cui et al. (2017) modified the expression of $\log A_2$ and re-defined a part of the characteristic function

$$D := \log B = \log \frac{d}{v_0} + \frac{k\tau}{2} - \log A_2 = \log \frac{d}{v_0} + \frac{(k-d)\tau}{2} - \log\left(\frac{d+\xi}{2v_0} + \frac{d-\xi}{2v_0}e^{-d\tau}\right), \tag{3.7}$$

leading to a new form of characteristic function of Heston stock price, which is both numerically continuous and easily differentiable. We modified it to become a characteristic function of log price

$$\phi_X(\boldsymbol{\theta}; u, \tau) \;=\; \mathbb{E}\left[\exp\left(iuX(t)\right)\right] \;=\; \exp\left[iu(X(t) + (r-q)\tau) - \frac{k\overline{v}\rho\tau iu}{\sigma} - A + \frac{2k\overline{v}}{\sigma^2}D\right]. \tag{3.8}$$

Theorem 1 in Cui et al. (2017) shows that the gradient of this characteristic function is $\nabla\phi_X(\boldsymbol{\theta}; u, \tau) = \phi_X(\boldsymbol{\theta}; u, \tau)\boldsymbol{h}_X(u)$, where

$$\boldsymbol{h}_X(u) := \begin{bmatrix} h_{X,v_0}(u) \\ h_{X,\overline{v}}(u) \\ h_{X,\rho}(u) \\ h_{X,k}(u) \\ h_{X,\sigma}(u) \end{bmatrix}. \tag{3.9}$$

The elements of $\boldsymbol{h}_X$ are

$$h_{X,\upsilon_0}(u) = -\frac{A}{\upsilon_0}, \tag{3.10a}$$

$$h_{X,\overline{v}}(u) = \frac{2k}{\sigma^2}D - \frac{k\rho\tau iu}{\sigma}, \tag{3.10b}$$

$$h_{X,\rho}(u) = -\frac{\partial A}{\partial \rho} + \frac{2k\overline{v}}{\sigma^2 d}\left(\frac{\partial d}{\partial \rho} - \frac{d}{A_2}\frac{\partial A_2}{\partial \rho}\right) - \frac{k\overline{v}\tau iu}{\sigma}, \tag{3.10c}$$

$$h_{X,k}(u) = \frac{1}{\sigma iu}\frac{\partial A}{\partial \rho} + \frac{2\overline{v}}{\sigma^2}D + \frac{2k\overline{v}}{\sigma^2 B}\frac{\partial B}{\partial k} - \frac{\overline{v}\rho\tau iu}{\sigma}, \tag{3.10d}$$

$$h_{X,\sigma}(u) = -\frac{\partial A}{\partial \sigma} - \frac{4k\overline{v}}{\sigma^3}D + \frac{2k\overline{v}}{\sigma^2 d}\left(\frac{\partial d}{\partial \sigma} - \frac{d}{A_2}\frac{\partial A_2}{\partial \sigma}\right) + \frac{k\overline{v}\rho\tau iu}{\sigma^2}, \tag{3.10e}$$

where $\xi, d$ are defined in Equation (2.4), $A, A_1, A_2, B$ are defined in Equation (2.10), and $D$ is defined in Equation (3.7). The intermediate terms in the derivative of $\rho$ are

$$\frac{\partial d}{\partial \rho} = -\frac{\xi\sigma iu}{d}, \tag{3.11a}$$

$$\frac{\partial A_2}{\partial \rho} = -\frac{\sigma iu(2+\xi\tau)}{2d\upsilon_0}\left(\xi\cosh\frac{d\tau}{2} + d\sinh\frac{d\tau}{2}\right), \tag{3.11b}$$

$$\frac{\partial B}{\partial \rho} = \frac{e^{k\tau/2}}{\upsilon_0}\left(\frac{1}{A_2}\frac{\partial d}{\partial \rho} - \frac{d}{A_2^2}\frac{\partial A_2}{\partial \rho}\right), \tag{3.11c}$$

$$\frac{\partial A_1}{\partial \rho} = -\frac{iu(u^2+iu)\tau\xi\sigma}{2d}\cosh\frac{d\tau}{2}, \tag{3.11d}$$

$$\frac{\partial A}{\partial \rho} = \frac{1}{A_2}\frac{\partial A_1}{\partial \rho} - \frac{A}{A_2}\frac{\partial A_2}{\partial \rho}, \tag{3.11e}$$

while the derivatives of $k$ and $\sigma$ can be expressed in term of derivatives of $\rho$

$$\frac{\partial A}{\partial k} = \frac{i}{\sigma u}\frac{\partial A}{\partial \rho}, \tag{3.12a}$$

$$\frac{\partial B}{\partial k} = \frac{i}{\sigma u}\frac{\partial B}{\partial \rho} + \frac{B\tau}{2}, \tag{3.12b}$$

$$\frac{\partial d}{\partial \sigma} = \left(\frac{\rho}{\sigma} - \frac{1}{\xi}\right)\frac{\partial d}{\partial \rho} + \frac{\sigma u^2}{d}, \tag{3.12c}$$

$$\frac{\partial A_1}{\partial \sigma} = \frac{(u^2+iu)\tau}{2}\frac{\partial d}{\partial \sigma}\cosh\frac{d\tau}{2}, \tag{3.12d}$$

$$\frac{\partial A_2}{\partial \sigma} = \frac{\rho}{\sigma}\frac{\partial A_2}{\partial \rho} - \frac{2+\tau\xi}{\upsilon_0\tau\xi iu}\frac{\partial A_1}{\partial \rho} + \frac{\sigma\tau A_1}{2\upsilon_0}, \tag{3.12e}$$

$$\frac{\partial A}{\partial \sigma} = \frac{1}{A_2}\frac{\partial A_1}{\partial \sigma} - \frac{A}{A_2}\frac{\partial A_2}{\partial \sigma}. \tag{3.12f}$$

The fact that Cui et al. (2017) managed to merge and rearrange all the derivatives and eventually boil them down to a group of equations with many common terms makes the analytical gradient even more convenient: when evaluating the analytical gradient, one only needs to evaluate once the common parts and this gives an advantage in performance.

To summarize this subsection, we found the Jacobian matrix of $\boldsymbol{r}_{\text{SPX}}$ by combining results from Pacati et al. (2018) and Cui et al. (2017). Multiplying the vector in Equation (3.9) with the characteristic function in Equation (3.8) gives the gradient $\nabla\phi_X(\boldsymbol{\theta}; u, \tau)$; inserting this gradient into Equation (3.6) gives the analytical gradient of the stock option price.

## 3.3 Gradient of VIX option price

Next step in constructing the gradient of our objective function is to find the Jacobian matrix $\boldsymbol{J}_{\text{VIX}} \coloneqq \nabla \boldsymbol{r}_{\text{VIX}}^{\mathsf{T}} \in \mathbb{R}^{m \times N_{\text{VIX}}}$. Similarly to Equation (3.4), the element of this matrix is defined as

$$J_{\text{VIX},ij} = \frac{\partial r_{\text{VIX},j}}{\partial \theta_i} = \frac{1}{p_{\text{VIX},j}} \frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K_j, \tau_j)}{\partial \theta_i}, \tag{3.13}$$

Each column of the matrix is a gradient vector evaluated at strike price $K_j$ and time to maturity $\tau_j$, and normalized by the option price observed from the market with the same strike and term. Thus, each column of $\boldsymbol{J}_{\text{VIX}}$ can be expressed as $\frac{1}{p_{\text{VIX},j}} \nabla P_{\text{VIX}}(\boldsymbol{\theta}; K_j, \tau_j)$.

The closed-form pricing formula of VIX option is available in Pacati et al. (2018), Equation (11). This formula is for a model that has two stochastic volatility processes and a displacement with one of the diffusion term. To obtain a suitable pricing formula, we turn off the second volatility process, jumps in price and volatility, and set the displacement equal to 0. The resulting equation gives price at time $t = 0$ of a call option with strike $K$

and maturity $\tau = T - t$

$$P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau) = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re}\left[\phi_v(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau};\overline{v},k)}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}}\right] d\text{Re}(u), \quad (3.14)$$

where

$$U(u, k, \overline{\tau}) \coloneqq -u\frac{a(\overline{\tau}; k)}{\overline{\tau}}, \quad (3.15a)$$

$$a(\overline{\tau}; k) \coloneqq \frac{1 - e^{-\overline{\tau}k}}{k}, \quad (3.15b)$$

$$b(\overline{\tau}; \overline{v}, k) \coloneqq \overline{v}(\overline{\tau} - a(\overline{\tau}; k)), \quad (3.15c)$$

and $\text{erf}(\cdot)$ is the error function with complex argument. Here $u \in \mathbb{C}$ and the complex part of $u$ is not zero. The integral should be performed along a straight line parallel to the real axis and a range of the complex part of $u$ is given in Pacati et al. (2018) and Lian et al. (2013). $\overline{\tau}$ denotes the length of time into the future that is used when extracting the volatility index from option prices. Following the convention of CBOE, we set $\overline{\tau} = 30$ days. However, the actual number used in the function need to coincide with the unit of other time period involved in the function, such as time to maturity. For example, if $\tau = 1$ means time to maturity is one year then $\overline{\tau}$ should be set to equal 30/365 or 30/252. From the definition Equation (3.15) we can see that both $U$ and $b$ are function of $k$, and $b$ is also a function of $\overline{v}$. This makes the derivatives of $k$ and $\overline{v}$ more complicated than their counterparty in Section 3.2. Here we derive the partial derivatives of $a(\overline{\tau}; k)$ and $b(\overline{\tau}; \overline{v}, k)$

$$\frac{\partial a(\overline{\tau}; k)}{\partial k} = \frac{\overline{\tau} - a(\overline{\tau}; k)(k\overline{\tau} + 1)}{k} \quad (3.16a)$$

$$\frac{\partial b(\overline{\tau}; \overline{v}, k)}{\partial \overline{v}} = \overline{\tau} - a(\overline{\tau}; k) \quad (3.16b)$$

$$\frac{\partial b(\overline{\tau}; \overline{v}, k)}{\partial k} = -\overline{v}\frac{\partial a(\overline{\tau}; k)}{\partial k} = -\overline{v}\frac{\overline{\tau} - a(\overline{\tau}; k)(k\overline{\tau} + 1)}{k} \quad (3.16c)$$

For now we will denote the gradient of volatility characteristic function $\nabla\phi_v(\boldsymbol{\theta}; u, \tau) = \phi_v(\boldsymbol{\theta}; u, \tau)\boldsymbol{h}_v(u)$, where

$$\boldsymbol{h}_v(u) \coloneqq \begin{bmatrix} h_{v,v_0}(u) \\ h_{v,\overline{v}}(u) \\ h_{v,\sigma}(u) \\ h_{v,k}(u) \end{bmatrix}. \tag{3.17}$$

The expressions of elements of $\boldsymbol{h}_v(u)$ will be defined later in this chapter.

To obtain partial derivative of $P_{\text{VIX}}$ with respect to $\overline{v}$, we apply the product rule and get

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial \overline{v}} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re}\left[ \frac{\partial \phi_v(\boldsymbol{\theta}; U)}{\partial \overline{v}} \frac{e^{-iu\frac{b(\overline{\tau},\overline{v},k)}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} \right.$$
$$\left. -\phi_v(\boldsymbol{\theta}; U)\frac{1 - \text{erf}(K\sqrt{-iu}/S_0)}{(-iu)^{3/2}} e^{-iu\frac{b(\overline{\tau},\overline{v},k)}{\overline{\tau}}} \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial \overline{v}} \right] d\text{Re}(u), \tag{3.18}$$

It is a rather messy expression, and to enjoy the performance advantage of analytical gradient, we need to clean it up a bit. We can first substitute $\frac{\partial \phi_v(\boldsymbol{\theta};U)}{\partial \overline{v}}$ with $\phi_v(\boldsymbol{\theta}; U, \tau)h_{v,\overline{v}}(U)$, and then with some rearrangement, we got

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial \overline{v}} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}}$$
$$\int_0^\infty \text{Re}\left\{ \phi_v(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau},\overline{v},k)}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} \left[ h_{v,\overline{v}}(U) - \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial \overline{v}} \right] \right\} d\text{Re}(u), \tag{3.19}$$

Next, we apply the product rule and derive the partial derivative of $P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)$ with respect to $k$

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial k} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re}\left[ \frac{\partial \phi_v(\boldsymbol{\theta}; U)}{\partial k} \frac{e^{-iu\frac{b(\overline{\tau},\overline{v},k)}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} \right.$$
$$\left. -\phi_v(\boldsymbol{\theta}; U)\frac{1 - \text{erf}(K\sqrt{-iu}/S_0)}{(-iu)^{3/2}} e^{-iu\frac{b(\overline{\tau},\overline{v},k)}{\overline{\tau}}} \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial k} \right] d\text{Re}(u), \tag{3.20}$$

This expression looks a lot like Equation (3.18), however, since $U$ is also a function of $k$, we need to obtain $\frac{\partial \phi_v(\boldsymbol{\theta};U(k))}{\partial k}$ first. Using chain rule, we derived

$$\frac{\partial \phi_v(\boldsymbol{\theta}; U(k))}{\partial k} = \frac{\partial \phi_v(\boldsymbol{\theta}; U)}{\partial k} + \frac{\partial \phi_v(\boldsymbol{\theta}; U(k))}{\partial U(k)} \frac{\partial U}{\partial k}. \tag{3.21}$$

For each term involved above, we have one expression

$$\frac{\partial \phi_\upsilon(\boldsymbol{\theta}; U)}{\partial k} = \phi_\upsilon(\boldsymbol{\theta}; U, \tau) h_{\upsilon,k}(U), \tag{3.22a}$$

$$\frac{\partial \phi_\upsilon(\boldsymbol{\theta}; U(k))}{\partial U(k)} = \phi_\upsilon(\boldsymbol{\theta}; U, \tau) \left[ \frac{F}{U} - \frac{1}{G} \left( \frac{2k\overline{\upsilon}}{\sigma^2} + F \right) \frac{\partial G}{\partial U} \right], \tag{3.22b}$$

$$\frac{\partial U}{\partial k} = -\frac{u}{\overline{\tau}} \frac{\partial a(\overline{\tau}; k)}{\partial k}, \tag{3.22c}$$

where $G$ and $F$ are defined as in Equation (3.33) and

$$\frac{\partial G}{\partial U} = \frac{G - e^{\frac{k\tau}{2}}}{U}. \tag{3.23}$$

We define here

$$h'_{\upsilon,k}(u) \coloneqq h_{\upsilon,k}(U) - \frac{u}{\overline{\tau}} \left[ \frac{F}{U} - \frac{1}{G} \left( \frac{2k\overline{\upsilon}}{\sigma^2} + F \right) \frac{\partial G}{\partial U} \right] \frac{\partial a(\overline{\tau}; k)}{\partial k}, \tag{3.24}$$

then $\frac{\partial \phi_\upsilon(\boldsymbol{\theta}; U(k))}{\partial k} = \phi_\upsilon(\boldsymbol{\theta}; U, \tau) h'_{\upsilon,k}(u)$. And substitute this into Equation (3.20), and with some rearrangement, we got a expression of derivative of $k$ which parallel with Equation (3.19)

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial k} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re} \left\{ \phi_\upsilon(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau};\overline{\upsilon},k)}{\overline{\tau}}} (1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} \left[ h'_{\upsilon,k}(u) - \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial k} \right] \right\} d\text{Re}(u), \tag{3.25}$$

For the derivative with respect to $\upsilon_0$ and $\sigma$, we get the similar expressions below

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial \upsilon_0} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re} \left\{ \phi_\upsilon(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau};\overline{\upsilon},k)}{\overline{\tau}}} (1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} h_{\upsilon,\upsilon_0}(U) \right\} d\text{Re}(u), \tag{3.26}$$

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial \sigma} = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re} \left\{ \phi_\upsilon(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau};\overline{\upsilon},k)}{\overline{\tau}}} (1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} h_{\upsilon,\sigma}(U) \right\} d\text{Re}(u), \tag{3.27}$$

Combining Equation (3.19), (3.25), (3.26), and (3.27), we can define a vector $\boldsymbol{H}_{\upsilon}(u)$ such that

$$\nabla\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau) = S_0 \frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^{\infty} \text{Re}\left\{ \phi_{\upsilon}(\boldsymbol{\theta}; U) \frac{e^{-iu\frac{b(\overline{\tau};\overline{\upsilon},k)}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}} \boldsymbol{H}_{\upsilon}(u) \right\} d\text{Re}(u), \quad (3.28)$$

where the elements of $\boldsymbol{H}_{\upsilon}(u)$ are defined as

$$\boldsymbol{H}_{\upsilon}(u) \coloneqq \begin{bmatrix} h_{\upsilon,\upsilon_0}(U) \\ h_{\upsilon,\overline{\upsilon}}(U) - \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial \overline{\upsilon}} \\ h_{\upsilon,\sigma}(U) \\ h'_{\upsilon,k}(u) - \frac{iu}{\overline{\tau}} \frac{\partial b}{\partial k} \end{bmatrix}. \quad (3.29)$$

Thus we obtain the closed-form expression of the Jacobian matrix of $\boldsymbol{r}_{\text{VIX}}^{\mathsf{T}}$ in terms of partial derivatives of the volatility characteristic function. In the next part, we will derive a closed-form expression for each element in vector $\boldsymbol{h}_{\upsilon}(u)$ (3.17).

## 3.4 Representations and gradient of volatility characteristic function

Following the footsteps of Cui et al. (2017), we try to find a expression of characteristic function of volatility that is both numerically continuous and easily differentiable.

To obtain a characteristic function of volatility in Heston model, we found the characteristic function of volatility in a double Heston model (Heston model extended with a second stochastic volatility process) in Pacati et al. (2018) (equation (A.4)). Turning off the second volatility process in this characteristic function gives the characteristic function of

the Heston volatility

$$\begin{aligned}\phi_v(\boldsymbol{\theta}; u, \tau) &= \mathbb{E}\left[e^{iuv(t)}\right] \\ &= \exp\left\{\frac{-2k\overline{v}}{\sigma^2}\log\left[1 - iu\frac{\sigma^2}{2k}(1 - e^{-k\tau})\right] + v_0\frac{iue^{-k\tau}}{1 - iu\frac{\sigma^2}{2k}(1 - e^{-k\tau})}\right\}. \quad (3.30)\end{aligned}$$

Since the volatility process itself is a Cox Ingersoll Ross (CIR) process (Cox et al., 1985), we also verified that this expression matches the characteristic function of a CIR process. This expression is numerically continuous (see Figure (3.1)), however, the parameters are tangled and not easy to take derivatives.

We mentioned in Section 2 that Del Baño Rollin et al. (2010) has derived the price characteristic function through the complex moment generating function and this moment generating function is derived by turning off $v(t)$ in the expression of $\mathbb{E}\left[e^{uX(t)+vv(t)}\right]$. By the same logic, if we turn off instead $X(t)$ by setting $u = 0$, we will get a complex moment generating function for the volatility process

$$\begin{aligned}M(v) &= \mathbb{E}\left[e^{vv(t)}\right] \\ &= \left(\frac{e^{\frac{k\tau}{2}}}{\cosh\frac{k\tau}{2} + \frac{k-\sigma^2 v}{k}\sinh\frac{k\tau}{2}}\right)^{\frac{2k\overline{v}}{\sigma^2}} \exp\left(v_0 v\frac{\cosh\frac{k\tau}{2} - \sinh\frac{k\tau}{2}}{\cosh\frac{k\tau}{2} + \frac{k-\sigma^2 v}{k}\sinh\frac{k\tau}{2}}\right). \quad (3.31)\end{aligned}$$

This equation is parallel with Equation (3) in Del Baño Rollin et al. (2010). Substituting $v$ in this equation with $iu$ gives us the expression of $\mathbb{E}\left[e^{iuv(t)}\right]$, which is, by definition, the characteristic function of $v(t)$. After rearranging, here is the final expression

$$\phi_v(\boldsymbol{\theta}; u, \tau) = M(iu) = \mathbb{E}\left[e^{iuv(t)}\right] = \left(\frac{e^{\frac{k\tau}{2}}}{G}\right)^{\frac{2k\overline{v}}{\sigma^2}} e^F, \quad (3.32)$$

where

$$F \coloneqq \frac{v_0 iu}{G}e^{-\frac{k\tau}{2}} \quad (3.33a)$$

$$G \coloneqq \cosh\frac{k\tau}{2} + \frac{k - \sigma^2 iu}{k}\sinh\frac{k\tau}{2}. \quad (3.33b)$$

The two intermediate parameters defined above makes this representation more compact and makes it easier to break the expression of gradient into separate blocks.The hyperbolic functions makes the parameters in this expression less tangled and thus it is easier to take derivatives of this expression.

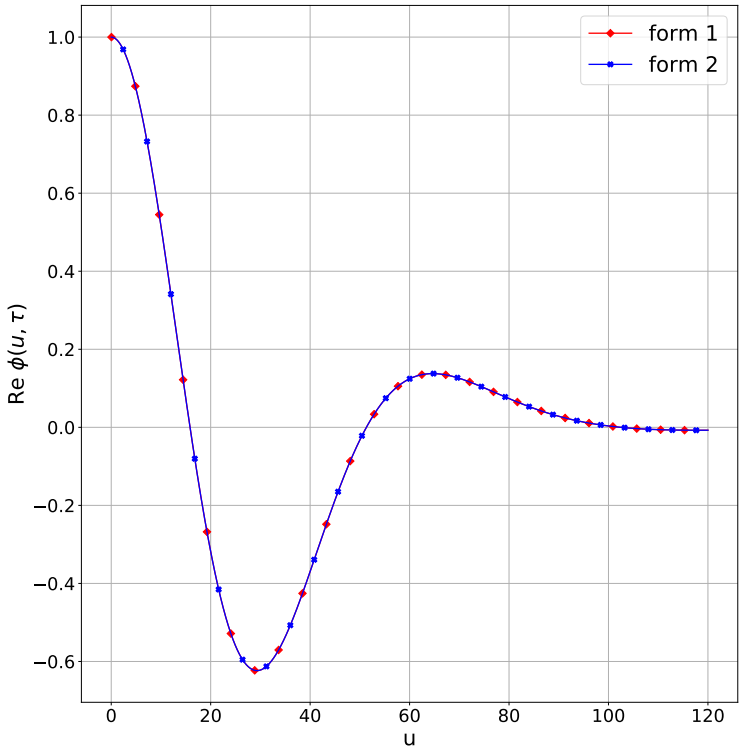



Figure 3.1: Real parts of characteristic functions (Equation (3.30) and (3.32)) as functions of $u$. Form 1 is the characteristic function obtained by turning off second volatility process in a double Heston model; Form 2 is the expression obtain from the complex moment generating function. $u \in (0, 120)$ and $\tau = 15$. Other parameters are the same as in Table 2.1.

Before proceed, we need to analyse the numerical continuity of this expression. To present

it in a straight forward way, we plotted the real part of both representations as functions of $u$ in Figure 3.1. From this figure we can see that both representations are numerically continuous.

For the rest of this dissertation, we will only refer Equation (3.32) as $\phi_\upsilon(\boldsymbol{\theta}; u, \tau)$ if without specification, since it is a expression that is both numerically continuous and easily differentiable.

We took derivative of each parameter in $\phi_\upsilon(\boldsymbol{\theta}; u, \tau)$ and combined the duplicated part. The purpose is to reduce the amount of calculation need to be done by the algorithm and increase the efficiency of this method.

Here are the closed-form expressions for elements in $\boldsymbol{h}_\upsilon(u)$ (3.17)

$$h_{\upsilon,\upsilon_0}(u) = \frac{F(u)}{\upsilon_0}, \tag{3.34a}$$

$$h_{\upsilon,\overline{\upsilon}}(u) = \frac{2k}{\sigma^2}\log\frac{e^{\frac{k\tau}{2}}}{G(u)}, \tag{3.34b}$$

$$h_{\upsilon,\sigma}(u) = -\frac{2\overline{\upsilon}}{\sigma}h_{\upsilon,\overline{\upsilon}}(u) - \frac{2k\overline{\upsilon}}{\sigma^2 G}\frac{\partial G}{\partial \sigma} - \frac{\upsilon_0 iu}{G^2 e^{\frac{k\tau}{2}}}\frac{\partial G}{\partial \sigma}, \tag{3.34c}$$

$$h_{\upsilon,k}(u) = -\frac{\sigma}{2k}h_{\upsilon,\sigma}(u) + \frac{\overline{\upsilon}\tau iu}{Ge^{\frac{k\tau}{2}}} - \frac{\upsilon_0 u\tau}{2kG^2}(2ki + u\sigma^2), \tag{3.34d}$$

where

$$\frac{\partial G}{\partial \sigma} = -\frac{2\sigma iu}{k}\sinh\frac{k\tau}{2}. \tag{3.35}$$

In this and previous section together, we derived the gradient of VIX option pricing equation. The key steps are summarized by Equation (3.24), (3.28), (3.29), and (3.34) while the details can be found in Equation (3.15), (3.16), (3.23), and (3.35). With the work in the last three sections, we have all the elements needed to construct the analytical gradient of our objective function (Equation (3.1)).

## 3.5 Minimization method

As one routine method for non-linear least-square problem, Levenberg-Marquardt is our choice for minimizing our objective function (Equation (3.1)).

First we are going to introduce the basic theory related to this method. Assume that $f$ is a function from $\mathbb{R}^n$ to $\mathbb{R}$, and $\boldsymbol{a}$ is the vector of variables. Further assume that the Taylor series of $f$ exist, then $f$ can be approximated by

$$f(\boldsymbol{a}) \approx \gamma - \boldsymbol{d}^{\mathsf{T}}\boldsymbol{a} + \frac{1}{2}\boldsymbol{a}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{a}, \tag{3.36}$$

where $\boldsymbol{d} = -\boldsymbol{\nabla} f(\boldsymbol{0})$ and $\boldsymbol{H}$ is Hessian matrix of $f$ evaluated at the origin. To find the optimal $\boldsymbol{a}$ that minimize $f$, we need to discuss two scenarios: first, if this quadratic form is a good approximation, we can use the inverse-Hessian method (Press et al., 2007) to search for the optimal

$$\boldsymbol{a}_{\text{next}} = \boldsymbol{a}_{\text{current}} - \boldsymbol{H}^{-1}\boldsymbol{\nabla} f(\boldsymbol{a}_{\text{current}}); \tag{3.37}$$

otherwise, if the quadratic form is a poor approximation, we can only step along the gradient, which is the steepest descent method (Press et al., 2007)

$$\boldsymbol{a}_{\text{next}} = \boldsymbol{a}_{\text{current}} - C\boldsymbol{\nabla} f(\boldsymbol{a}_{\text{current}}), \tag{3.38}$$

where $C$ is a constant that make sure the size of the step does not exhaust the gradient direction. By combining these two method together, Levenberg-Marquardt method allows us to move between the inverse-Hessian method and steepest descent method smoothly

$$\boldsymbol{a}_{\text{next}} = \boldsymbol{a}_{\text{current}} - (\boldsymbol{H} + c\boldsymbol{I})^{-1}\boldsymbol{\nabla} f(\boldsymbol{a}_{\text{current}}), \tag{3.39}$$

where $c$ is another constant that controls the moving between these two methods. When the iterate is far from the optimal parameters, $c$ is set to be large, term $cI$ dominates the Hessian matrix and the method is closer to a steepest descent method; when the iterate is close to the optimal parameters, $c$ is given a small value so Hessian matrix dominates the term $cI$, and the method is closer to an inverse-Hessian method.

For the Levenberg-Marquardt method to work, we need a reasonable close start point. In general, we cannot assume that our objective function only has one global minimum, and our method here play no magic to change that. Press et al. (2007) suggests to use it as an "endgame" trick applied after one already have a reasonably good guess about the optimal parameters. Practically, initial guess input into the Levenberg-Marquardt method can be a set of optimal parameters obtained form the older data. The very first set of parameters can be obtained using numerical method on a relatively big searching range, and the optimal parameters for one period can be good initial guess for the next period.

Parallel to Equation (21) in Cui et al. (2017), we derived the gradient and Hessian of our objective function, they can be expressed respectively as

$$\boldsymbol{\nabla} L = \frac{1}{N_{\text{SPX}}} \boldsymbol{J}_{\text{SPX}} \boldsymbol{r}_{\text{SPX}} + \frac{1}{N_{\text{VIX}}} \boldsymbol{J}_{\text{VIX}} \boldsymbol{r}_{\text{VIX}}, \tag{3.40}$$

$$\begin{aligned} \boldsymbol{\nabla}\boldsymbol{\nabla}^{\mathsf{T}} L = \frac{1}{N_{\text{SPX}}} \left( \boldsymbol{J}_{\text{SPX}} \boldsymbol{J}_{\text{SPX}}^{\mathsf{T}} + \sum_{i=1}^{N_{\text{SPX}}} r_{\text{SPX},i} \boldsymbol{H}(r_{\text{SPX},i}) \right) \\ + \frac{1}{N_{\text{VIX}}} \left( \boldsymbol{J}_{\text{VIX}} \boldsymbol{J}_{\text{VIX}}^{\mathsf{T}} + \sum_{i=1}^{N_{\text{VIX}}} r_{\text{VIX},i} \boldsymbol{H}(r_{\text{VIX},i}) \right), \end{aligned} \tag{3.41}$$

where $\boldsymbol{H}(r)$ represents Hessian matrix of the residual $r$. The Hessian matrices can be ignored if the term $\sum_{i=1}^{N_{\text{SPX}}} r_{\text{SPX},i} \boldsymbol{H}(r_{\text{SPX},i})$ and $\sum_{i=1}^{N_{\text{VIX}}} r_{\text{VIX},i} \boldsymbol{H}(r_{\text{VIX},i})$ are small. That means either the residuals $r$ are small or our Hessian matrices are small. If our model is successful, the residuals defined in Equation (3.2) should be small and have either sign. Under this assumption, the sum of the products of residual and Hessian matrix should be small and we can approximate the Hessian (Equation (3.41)) of the objective function (Equation (3.1)) by

$$\boldsymbol{\nabla}\boldsymbol{\nabla}^{\mathsf{T}} L = \frac{1}{N_{\text{SPX}}} \boldsymbol{J}_{\text{SPX}} \boldsymbol{J}_{\text{SPX}}^{\mathsf{T}} + \frac{1}{N_{\text{VIX}}} \boldsymbol{J}_{\text{VIX}} \boldsymbol{J}_{\text{VIX}}^{\mathsf{T}}, \tag{3.42}$$

where we ignore all the terms that contains $\boldsymbol{H}$.

We derived the closed-form expressions of both $\boldsymbol{J}_{\text{SPX}}$ and $\boldsymbol{J}_{\text{VIX}}$ in the previous sections, together with Equation (3.40) and Equation (3.42), we can apply Levenberg-Marquardt method on our objective function (Equation (3.1))

$$\boldsymbol{\theta}_{\text{next}} = \boldsymbol{\theta}_{\text{current}} - (\boldsymbol{\nabla}\boldsymbol{\nabla}^{\mathsf{T}} L(\boldsymbol{\theta}_{\text{current}}) + c\boldsymbol{I})^{-1}\boldsymbol{\nabla} L(\boldsymbol{\theta}_{\text{current}}). \tag{3.43}$$

## 3.6 Gradient with displacement

In this section, we will extend the calibration to include not only the parameters of Heston model but also the displacement introduced in Pompa (2015) and Pacati et al. (2018). As we mentioned in Section 2.3, this displacement severs as a handy lower bound of the stochastic volatility. It significantly improved the ability of the model to fit to multiple markets (Pacati et al., 2018). To calibrate this displacement with analytical gradients, we need the closed-form pricing formulas for both equity options and the VIX options. Both of the formulas are provided in Pacati et al. (2018), and the price of the equity index option is

$$\begin{aligned} P_{\text{SPX}}(\boldsymbol{\theta}; K, \tau) &= S(t)e^{-q\tau} - \frac{1}{\pi}\sqrt{S(t)K}e^{-\frac{1}{2}(r+q)\tau} \\ &\times \int_0^\infty \text{Re}\left[e^{iu\kappa - i(u-\frac{i}{2})[X(t)-Y(t)+(r-q)\tau]}\phi_X\left(\boldsymbol{\theta}; u-\frac{i}{2}\right)\right]\frac{e^{-(u^2+1/4)I_\phi(t,T)}}{u^2+1/4}\,du, \end{aligned} \tag{3.44}$$

where $I_\phi(t,T) = \int_t^T \phi(s)\,ds$. And the price of the VIX index option is

$$\begin{aligned} &P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau) = \\ &S_0\frac{e^{-r\tau}}{2\sqrt{\pi}}\int_0^\infty \text{Re}\left[\phi_\upsilon(\boldsymbol{\theta}; U)\frac{e^{-iu\frac{b(\overline{\tau};\overline{v},k)+I_\phi(T,T+\overline{\tau})}{\overline{\tau}}}(1-\text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}}\right] d\text{Re}(u). \end{aligned} \tag{3.45}$$

In both pricing formulas, the displacement $\phi(t)$ only appears in the integrations, thus the goal of the calibration is to obtain not the displacement but its integrations on different intervals. Before starting to take the gradients, we first study which are the intervals that

we need to calibrate. Assume that for our calibration, we have a set of maturities for the equity options, $A = \{a_1, a_2, a_3, \ldots, a_{N_1}\}$, and a set of maturities for the VIX options, $B = \{b_1, b_2, b_3, \ldots, b_{N_2}\}$. If we denote today as $t = 0$, then $I_\phi(t, T) = I_\phi(0, \tau)$, where $\tau = T - t$. So the intervals we need to calibrate for our set of options are $I_\phi(0, a_i)$ for $i = 1, 2, \ldots, N_1$ and $I_\phi(b_j, b_j + \overline{\tau})$ where $j = 1, 2, \ldots, N_2$. Next, we define a new set of maturities $\tau = \{0 \cup A \cup B \cup (B + \overline{\tau})\} = \{\tau_0, \tau_1, \tau_2, \tau_3, \ldots, \tau_{N_3}\}$, where $\tau_0 = 0$, the duplicated maturities are only counted once so $N_3 \leq N_1 + N_2$, and all the maturities are sorted in ascending order. If we have all the $\Delta\Phi_i = I_\phi(\tau_i, \tau_{i+1})$, $i \in (0, N_3 - 1)$, we can obtain all the integrals we need. Since every maturity $a_i \in A$ is in this new set, we can get $I_\phi(0, a_i) = \sum_{k=0}^{n-1} \Delta\Phi_k$ where $\tau_n = a_i$. Any $I_\phi(b_j, b_j + \overline{\tau})$ can be obtained by $\sum_{k=p}^{q-1} \Delta\Phi_k$ where $\tau_p = b_j$ and $\tau_q = b_j + \overline{\tau}$, and since all the unique elements from sets $B$ and $B + \overline{\tau}$ are in set $\tau$, one should have no problem finding these two elements. In this way, we convert our goal into calibrating $\Delta\Phi_i$ where $i = 0, 1, \ldots, N_3 - 1$. So far we have followed the calibrating method in Pacati et al. (2018), appendix B.1. Next, we will derive the analytical derivative for each $\Delta\Phi_i$ where $i = 0, 1, \ldots, N_3 - 1$.

For the equity option pricing formula, we have

$$\frac{\partial P_{\text{SPX}}(\boldsymbol{\theta}; K, \tau)}{\partial \Delta\Phi_k} = -\frac{1}{\pi}\sqrt{S(t)K}e^{-\frac{1}{2}(r+q)\tau}\int_0^\infty \text{Re}\left[e^{iu\kappa - i(u-\frac{i}{2})[X(t)-Y(t)+(r-q)\tau]}\phi_X\left(\boldsymbol{\theta}; u - \frac{i}{2}\right)\right] \times \frac{e^{-(u^2+1/4)I_\phi(0,a_i)}}{u^2 + 1/4}\left(-(u^2 + \frac{1}{4})\right)\mathbf{1}_{k\in[0,n-1]}\, du, \tag{3.46}$$

and for the VIX option pricing formula, we have

$$\frac{\partial P_{\text{VIX}}(\boldsymbol{\theta}; K, \tau)}{\partial \Delta\Phi_k} = S_0\frac{e^{-r\tau}}{2\sqrt{\pi}} \int_0^\infty \text{Re}\left[\phi_v(\boldsymbol{\theta}; U)\frac{e^{-iz\frac{b(\overline{\tau},\overline{v},k)+I_\phi(b_j,b_j+\overline{\tau})}{\overline{\tau}}}(1 - \text{erf}(K\sqrt{-iu}/S_0))}{(-iu)^{3/2}}\frac{-iu}{\overline{\tau}}\mathbf{1}_{k\in[p,q-1]}\right] d\text{Re}(u). \tag{3.47}$$

To obtain the new Jacobian matrix, for each column, we only need to append the derivatives above, normalized with the market price of the corresponding option, under the

original gradient vector. We define the vector $\boldsymbol{\theta}' = [\upsilon_0, \overline{v}, \rho, k, \sigma, \Delta\Phi_0, \Delta\Phi_1, \ldots, \Delta\Phi_{N_3}]^\mathsf{T}$, an entry of the new Jacobian matrix can be expressed as

$$J_{ji} = \frac{\partial r_i}{\partial \theta'_j} = \frac{1}{p_i} \frac{\partial P(\boldsymbol{\theta}; K_i, \tau_i)}{\partial \theta_j}, \tag{3.48}$$

when $j = 1, 2, \ldots, 5$ and as

$$J_{ji} = \frac{\partial r_i}{\partial \theta'_j} = \frac{1}{p_i} \frac{\partial P(\boldsymbol{\theta}; K_i, \tau_i)}{\partial \Delta\Phi_{j-6}}, \tag{3.49}$$

when $j = 6, 7, \ldots, 6 + N_3$. We do not specify the type of option, SPX or VIX, because this definition applies to both.

One might ask why we need to express each integration as a sum and calibrate the elements in the sum instead of just calibrating the integration as a whole thing. We believe there are two main reasons: Firstly, this method gives us a term structure of the whole time period which we can use to price any option with a maturity that falls in, while the other method will give us many integrations with intervals overlapping each other and may or may not cover the whole time period in the calibration. This expression cannot be easily adjusted to price an option with a new maturity that is not in the set used for calibration. Secondly, from a calibration point of view, this method takes full advantage of the number of data points available by calibrating the term structure at the most fine grid that is possible.

# Chapter 4

# Numerical results

We test the theory for calibrating the Heston model we proposed in Chapter 3. The code is available in the supplementary material. We first simulate a set of parameters from some reasonable ranges specified in Cui et al. (2017, Table 5), and use them to price a group of equity options and a group of VIX options. We quote these ranges in Table 4.1.

| | |
|---|---|
| $k$ | ( 0.50, 5.00) |
| $\overline{v}$ | (0.05, 0.95) |
| $\sigma$ | (0.05, 0.95) |
| $\rho$ | (-0.90, -0.10) |
| $v_0$ | (0.05, 0.95) |

Table 4.1: Reasonable ranges of the Heston model parameter.

Then we use these prices as market prices in Equation (3.1). To calibrate the parameters, we started from initial guesses from the same ranges. The reason for setting the initial guesses in the same ranges is that, as we mentioned in the precious chapter, Levenberg-Marquardt method needs a reasonable close start point. To implement the Levenberg-Marquardt method, we used the LEVMAR (Lourakis, 2004) library in C++. This is a

robust and stable implementation which has been used in many mature products such as Hugin and SciGraphica, and has also been the choice in Cui et al. (2017). The search will stop when one of the following three criteria has been met:

$$\|\boldsymbol{r}(\boldsymbol{\theta}_k)\| \leq \varepsilon_1, \tag{4.1a}$$

$$\|\boldsymbol{J}_k \boldsymbol{e}\|_\infty \leq \varepsilon_2, \tag{4.1b}$$

$$\frac{\|\Delta\boldsymbol{\theta}_k\|}{\|\boldsymbol{\theta}_k\|} \leq \varepsilon_3, \tag{4.1c}$$

where $\varepsilon_1$, $\varepsilon_2$, and $\varepsilon_3$ are the chosen tolerance. Following the practice in Cui et al. (2017), we set all three to $1e-10$. The first criterion indicates that the objective function, Equation (3.1), has reached the desired value; the second criterion is the threshold for stopping the iteration due to a overly small gradient; the last criterion indicates that the update of the parameters are too small. We used Gauss-Legendre as the quadrature with the same truncation as in Cui et al. (2017), i.e. $u = 200$.

## 4.1 Data

In the calibration, we use each set of parameters to price 40 equity options and 30 VIX options. Following Cui et al. (2017), we calculate the strike prices according to the Black-Shores deltas. In this calculation, we set the parameters to the values in Table 2.1. We selected 5 delta values of call option: $90\%, 75\%, 50\%, 25\%, 10\%$. Because of the put-call parity, these are equivalent to the delta set for put option: $-10\%, -25\%, -50\%, -75\%, -90\%$. Our selected maturities for these two types of options are presented in Table 4.2, and 4.3.

| $\tau$ (days) | $\Delta_{\text{put}} = -10\%$ | $\Delta_{\text{put}} = -25\%$ | $\Delta_{\text{call}} = 50\%$ | $\Delta_{\text{call}} = 25\%$ | $\Delta_{\text{call}} = 10\%$ |
|---|---|---|---|---|---|
| 30 | 93.71 | 99.56 | 104.27 | 122.87 | 139.39 |
| 60 | 86.03 | 98.68 | 104.63 | 123.99 | 141.02 |
| 90 | 81.12 | 97.28 | 104.99 | 124.85 | 142.91 |
| 120 | 77.60 | 95.88 | 105.30 | 126.59 | 144.56 |
| 150 | 74.70 | 94.64 | 105.62 | 126.46 | 146.03 |
| 180 | 72.16 | 93.58 | 105.93 | 127.15 | 147.36 |
| 270 | 66.99 | 91.75 | 106.63 | 128.59 | 150.05 |
| 360 | 61.37 | 90.25 | 107.66 | 130.46 | 153.28 |

Table 4.2: Strike prices for equity options.

| $\tau$ (days) | $\Delta_{\text{put}} = -10\%$ | $\Delta_{\text{put}} = -25\%$ | $\Delta_{\text{call}} = 50\%$ | $\Delta_{\text{call}} = 25\%$ | $\Delta_{\text{call}} = 10\%$ |
|---|---|---|---|---|---|
| 7 | 7.66 | 7.82 | 8.01 | 8.20 | 8.37 |
| 14 | 7.53 | 7.75 | 8.02 | 8.28 | 8.53 |
| 21 | 7.43 | 7.71 | 8.02 | 8.35 | 8.66 |
| 30 | 7.33 | 7.65 | 8.03 | 8.43 | 8.81 |
| 60 | 7.08 | 7.53 | 8.07 | 8.64 | 9.19 |
| 90 | 6.91 | 7.45 | 8.10 | 8.81 | 9.50 |

Table 4.3: Strike prices for VIX options.

In our calibration, $S(t) = S_0$ because any start price can be normalized before the actual calibration and the change will be trivial. We also set the risk-free rate $r = 0.02$ and the dividend rate $q = 0$. Through the whole calibration, we normalized all the maturities with one calendar year (365 days).

## 4.2 Performance

We performed 10,000 calibrations of the Heston parameters. When select one initial guess for each calibration, about 70% (6947 out of 10,000) of the calibrations was stopped by the desired low value of the objective function. And a large amount of the calibrations are stopped by a stagnating update of parameters. We then allow the calibrations that are stopped by stagnating updates to restart from another randomly drew point within the reasonable range mentioned in Table 4.1 as long as the number of restarts is below certain fixed number (100 in our code). After allowing the restart, 8,878 out of 10,000 calibrations successfully find the optimal set of parameters and are terminated by the desired low value of objective function, which is set to $1e-10$ in our programme. When looking at the optimal parameters found by the calibration programme, we noticed that although the calibration is capable of finding the very accurate absolute value of parameter $\rho$ and $\sigma$, it seems to be confused about the sign of these two. We found 2,111 out of 8,878 solved cases that has the correct absolute value but the opposite sign to the input parameters. One possible solution for this issue it to add constraint to the search. A summary of all the solved cases can be found in Table 4.4 and Table 4.5.

| | | | |
|---|---|---|---|
| $\lvert k^0 - k^\star\rvert$ | $1.3485e+00$ | $\lvert k^\dagger - k^\star\rvert$ | $1.6924e-03$ |
| $\lvert \overline{\upsilon}^0 - \overline{\upsilon}^\star\rvert$ | $2.9363e-01$ | $\lvert \overline{\upsilon}^\dagger - \overline{\upsilon}^\star\rvert$ | $4.5000e-05$ |
| $\lvert \sigma^0 - \sigma^\star\rvert$ | $3.0195e-01$ | $\lvert\lvert\sigma^\dagger\rvert - \lvert\sigma^\star\rvert\rvert$ | $3.1224e-05$ |
| $\lvert \rho^0 - \rho^\star\rvert$ | $2.6517e-01$ | $\lvert\lvert\rho^\dagger\rvert - \lvert\rho^\star\rvert\rvert$ | $2.1297e-04$ |
| $\lvert \upsilon_0^0 - \upsilon_0^\star\rvert$ | $2.9537e-01$ | $\lvert \upsilon_0^\dagger - \upsilon_0^\star\rvert$ | $8.7753e-06$ |

Table 4.4: Superscript 0 indicate the initial guess, $\star$ indicate the real optimal value, and †indicates the optimal value found by the calibration.

| Error measure | | Computational cost | |
|---|---|---|---|
| $\|\boldsymbol{r}^0\|$ | $7.3081e+08$ | CPU time (seconds) | 0.5077 |
| $\|\boldsymbol{r}^\dagger\|$ | $5.6567e-11$ | LM iterations | 31.2988 |
| $\|\boldsymbol{J}^\dagger \boldsymbol{e}\|_\infty$ | $2.0113e-05$ | Price evaluations | 37.6922 |
| $\|\Delta\boldsymbol{\theta}^\dagger\|$ | $1.4258e-05$ | Gradient evaluations | 31.2988 |

Table 4.5: The left half of the table reports some error measures and the right half of the table reports the computational cost of the calibrations. In the table, superscript 0 indicate the initial guess, and †indicates the optimal value found by the calibration.

We notice that the number of iterations are much higher than that in Cui et al. (2017). This might due to the more complicated model or error in the implement.

In the rest of this section, we will present the result of one representative calibration. The initial guess, optimal values found by calibration, and real values of the parameters are reported in Table 4.6, together with the absolute differences between the these values.

| Par. | Initial guess | Found optimal | Real value | $\|\theta_i^0 - \theta_i^\star\|$ | $\|\theta_i^\dagger - \theta_i^\star\|$ |
|---|---|---|---|---|---|
| $k$ | $1.4910e+00$ | $3.9394e+00$ | $3.9393e+00$ | $2.4483e+00$ | $7.7262e-05$ |
| $\overline{v}$ | $2.6249e-01$ | $6.0711e-01$ | $6.0711e-01$ | $3.4461e-01$ | $7.2592e-07$ |
| $\sigma$ | $7.3247e-01$ | $3.5040e-01$ | $3.5028e-01$ | $3.8219e-01$ | $1.1438e-04$ |
| $\rho$ | $-6.3405e-01$ | $-5.5150e-01$ | $-5.5163e-01$ | $8.2417e-02$ | $1.2869e-04$ |
| $v_0$ | $7.6963e-01$ | $1.6244e-01$ | $1.6244e-01$ | $6.0718e-01$ | $8.0092e-07$ |

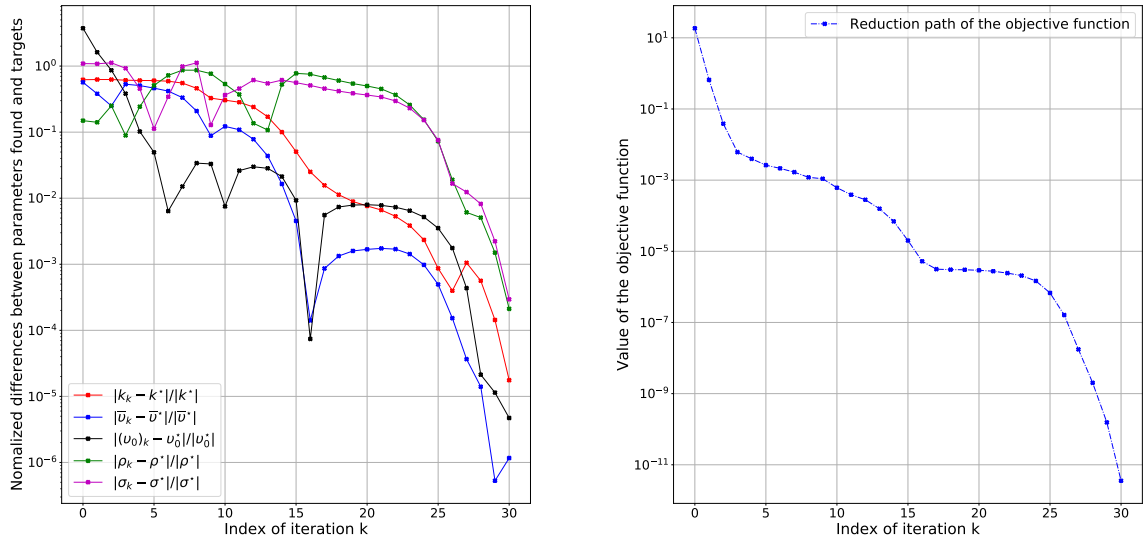
Table 4.6: The initial guess, optimal found, and real parameters of the representative calibration. Last two columns are the absolute differences between the initial guess and the real value, and absolute differences between optimal value found by calibration and real value. In the table, superscript 0 indicate the initial guess, and † indicates the optimal value found by the calibration.

The error measures and computational cost of this particular calibration are reported in Table 4.7.

| Error measure | | Computational cost | |
|---|---|---|---|
| $\|\boldsymbol{r}^0\|$ | $1.8563e+01$ | CPU time (seconds) | 0.5467 |
| $\|\boldsymbol{r}^\dagger\|$ | $4.2622e-12$ | LM iterations | 30.0000 |
| $\|\boldsymbol{J}^\dagger\boldsymbol{e}\|_\infty$ | $1.3356e-05$ | Price evaluations | 37.0000 |
| $\|\Delta\boldsymbol{\theta}^\dagger\|$ | $1.3427e-06$ | Gradient evaluations | 30.0000 |

Table 4.7: The left half of the table reports error measures of the representative calibration and the right half of the table reports the computational cost. In the table, superscript 0 indicate the initial guess, and † indicates the optimal value found by the calibration.

In the graph below, we can see the paths of each parameter through all 30 iterations. We also report the values of objective function after each iteration. This value is monotonically decreasing, which matches the logic of Levenberg-Marquardt method: the update of parameters that cannot reduce the objective function will be rejected and re-calculate with a different $u$ in Equation (3.43).

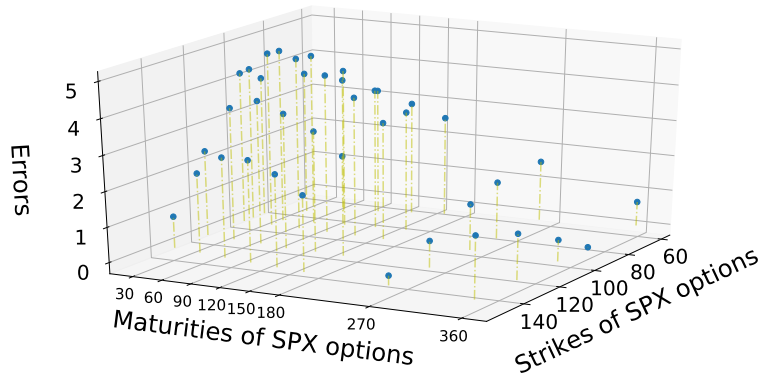



(a) Paths of each parameter in Heston model through iterations.

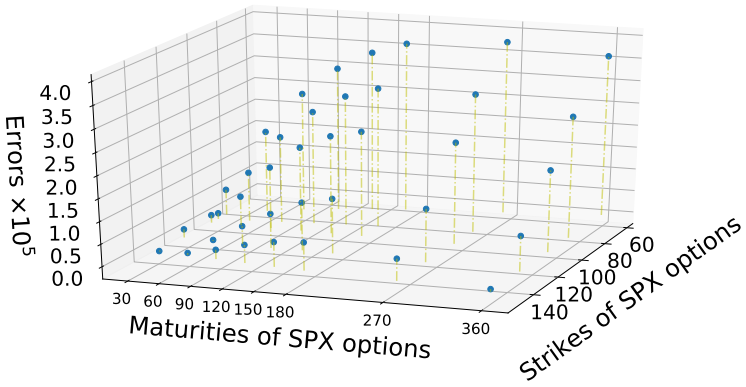
(b) Reduction of objective function (Equation (3.1)).

Although all the parameters converged to their real values, the paths they took are not as smooth as showed in Cui et al. (2017).

Since the ultimate goal for modelling the equity price and volatility is to price the corresponding options, we also report here the absolute pricing error of both equity options and VIX options. We can see that, as expected, the optimal parameters found by our method greatly reduced the pricing error for both.

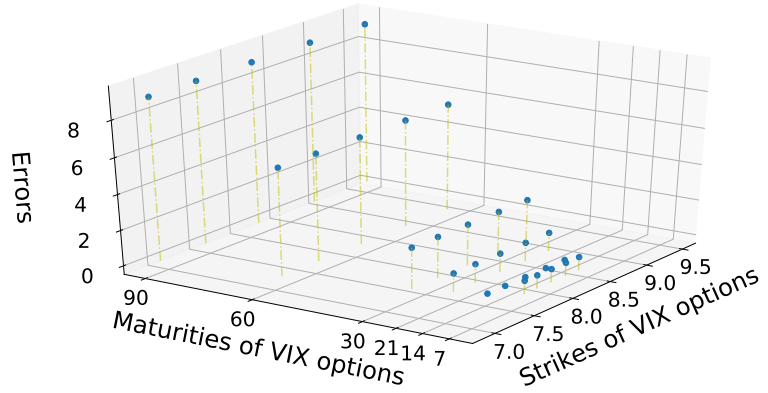



(a) Pricing errors of equity options for initial parameters.

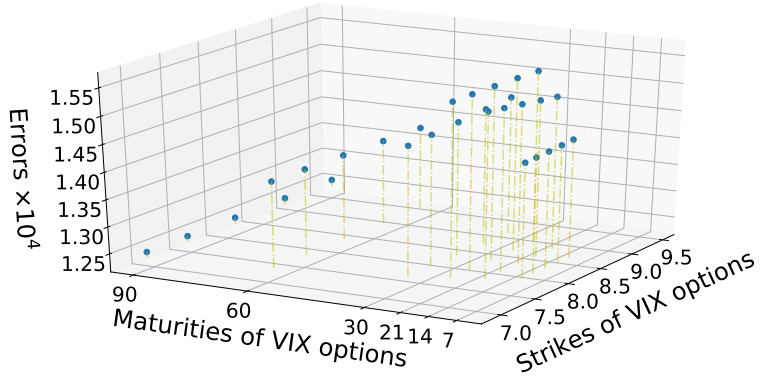



(b) Pricing errors of equity options for optimal parameters found by calibration.

The error for the equity option pricing was reduced to a very low level. However, the errors for prices of options that have short maturities and high strike prices are lower than others. For VIX option pricing, errors for long maturity options are lower.




(a) Pricing errors of VIX options for initial parameters.




(b) Pricing errors of VIX options for optimal parameters found by calibration.

Our numerical test for calibrating the Heston model in a consistent fashion shows reasonably good result. Since not all the initial guesses lead to the same optimal parameters, the objective function we chose might have local minima.

# Chapter 5

# Conclusions and future directions

## 5.1 Conclusions

We proposed a new expression of the Heston volatility characteristic function. This expression is numerically continuous and easily differentiable. We derived the analytical form of the gradient of this characteristic function and checked this form with the result from the symbolic maths toolbox in Matlab. This gradient is efficient to evaluate and can be applied in any gradient-based algorithm. We presented a method to calibrate the Heston model in a consistent fashion using market prices of both S&P500 index options and VIX options. The pricing formula we used to price S&P500 index options contains only one instead of two integrations, and is thus more efficient. The calibration conducted using the Levenberg-Marquardt method managed to find the optimal parameters in 88.8% of the 10,000 cases tested. In all of the test cases, we used randomly generated parameters and initial guesses. In the case of the S&P500 index options, the calibration reduced pricing error from a magnitude of $1e+00$ to $1e-05$; in the case of VIX options, pricing error was reduced from a magnitude of $1e+00$ to $1e-04$. This calibration algorithm has the potential to be applied both by practitioners and academic researchers.

## 5.2 Future directions

There are still many ways to improve this algorithm. The average number of iterations is higher than the one in Cui et al. (2017), and the paths of the parameters are not as smooth as would be ideally desired. It is possible that out implementation of the theory is not totally correct. One possible benchmark would be to perform a calibration of the same model using the same objective function with numerical approximation of gradients, for example forward or central finite differences. If the analytical gradient method is implemented correctly, it should take fewer iterations than the numerical approximation method. Also, the properties of the objective function reflected by the two calibration methods, for example the existence of local minima, should match. One may add constraints to the searching ranges. This may increase the proportion of cases that this method can solve. However, the constraints need to be carefully chosen. The smaller the ranges, the more helpful, but overly small ranges might exclude the real optimal parameters. It is possible to extend this method to calibrate the Heston model with jumps in the volatility process, and thus price VIX options. It is known that the Heston model is non-ideal for pricing VIX options (Gatheral, 2008). To properly price VIX options, modifications to the Heston model, such as adding jumps, or including an extra volatility process, are needed. The characteristic functions for some of these models exist and are known. When the closed-form prices are available, it may also be interesting to extend this method to Lévy-driven stochastic volatility models (Gong et al., 2016) since they are frequently discussed for option pricing.

# References


Albrecher, H., P. Mayer, W. Schoutens, and J. Tistaert (2007). "The little Heston trap". In: *Wilmott Magazine* 2007 (January), pp. 83–92. URL: https://www.wilmott.com/the-little-heston-trap-wilmott-magazine-article-hansjorg-albrecher-philipp-mayer-wim-schoutens-jurgen-tistaert/.

Badshah, I., B. Frijins, J. Knif, and A. Tourani-Rad (2016). "Asymmetries of the intra-day return-volatility relation". In: *International Review of Financial Analysis* 48 (12), pp. 182–192. DOI: 10.1016/j.irfa.2016.09.016.

CBOE (2018). *The CBOE volatility index-VIX*. White Paper. URL: https://www.cboe.com/micro/vix/vixwhite.pdf.

Cox, J., J. Ingersoll, and S. Ross (1985). "A theory of the term structure of interest rates". In: *Econometrica* 53 (2), pp. 385–407. DOI: 10.2307/1911242.

Cui, Y., S. Del Baño Rollin, and G. Germano (2017). "Full and fast calibration of the Heston stochastic volatility model". In: *European Journal of Operational Research* 263 (2), pp. 625–638. DOI: 10.1016/j.ejor.2017.05.018.

Del Baño Rollin, S., A. Ferreiro-Castilla, and F. Utzet (2010). "On the density of log-spot in the Heston volatility model". In: *Stochastic Processes and their Applications* 120 (10), pp. 2037–2063. DOI: 10.1016/j.spa.2010.06.003.

Duan, J.-C. and C.-Y. Yeh (2010). "Jump and volatility risk premiums implied by VIX". In: *Journal of Economic Dynamics and Control* 34 (11), pp. 2232–2244. DOI: 10.1016/j.jedc.2010.05.006.

Gatheral, J. (2008). *Consistent modeling of SPX and VIX options*. Bachelier Congress.

Gong, X.-L. and X.-T. Zhuang (2016). "Option pricing and hedging for optimized Levy driven stochastic volatility models". In: *Chaos, Solitons and Fractals* 91 (10), pp. 118–127. DOI: 10.1016/j.chaos.2016.05.012.

Heston, S. I. (1993). "A closed-form solution for options with stochastic volatility with applications to bond and currency options". In: *Review of Financial Studies* 6 (2), pp. 327–342. DOI: 10.1093/rfs/6.2.327.

Jiang, G. and Y. Tian (2005). "The model-free implied volatility and its information content". In: *Review of Financial Studies* 18 (4), pp. 1305–1342. DOI: 10.1093/rfs/hhi027.

Kahl, C. and P. Jäckel (2005). "Not-so-complex logarithms in the Heston model". In: *Wilmott Magazine* 2005 (September), pp. 94–103. URL: https://www.wilmott.com/not-so-complex-logarithms-in-the-heston-model/.

Lian, G.-H. and S.-P. Zhu (2013). "Pricing VIX options with stochastic volatility and random jumps". In: *Decisions in Economics and Finance* 36 (1), pp. 71–88. DOI: 10.1007/s10203-011-0124-0.

Lin, Y.-N. (2007). "Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures". In: *Journal of Futures Markets* 27 (12), pp. 1175–1217. DOI: 10.1002/fut.20291.

Lord, R. and C. Kahl (2010). "Complex logarithms in Heston-like models". In: *Mathematical Finance* 20 (4), pp. 671–694. DOI: 10.1111/j.1467-9965.2010.00416.x.

Lourakis, M. I. A. (2004). *levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++*. URL: http://www.ics.forth.gr/~lourakis/levmar.

Pacati, C., G. Pompa, and R. Renò (2018). "Smiling twice: The Heston++ model". In: *Journal of Banking and Finance*. SSRN 2697179. DOI: 10.2139/ssrn.2697179.

Pompa, G (2015). *Deterministic Shift Extension of Affine Models for Variance Derivatives*. PhD thesis, IMT Lucca. DOI: 10.6092/imtlucca/e-theses/185.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007). *Numerical Recipes: The Art of Scientific Computing, 3rd Edition*. Cambridge: Cambridge University Press.

Rouah, F. D. (2013). *The Heston model and its extensions in Matlab and C#*. Hoboken, New Jersey: John Wiley & Sons, Inc. DOI: 10.1002/9781118656471.

Schoutens, W., E. Simons, and J. Tistaert (2004). "A perfect calibration! Now what?" In: *Wilmott Magazine* 2004 (March), pp. 66–78.

Zhang, J. E., J. Shu, and M. Brenner (2010). "The new market for volatility trading". In: *Journal of Futures Markets* 30 (9), pp. 809–833. DOI: 10.1002/fut.20448.