

The School of Mathematics



THE UNIVERSITY
of EDINBURGH

Equity Modelling Using the Variance Gamma Process

by

Spyridon Dallas

Dissertation Presented for the Degree of
MSc in Computational Mathematical Finance

August 2024

Supervised by
Prof. I. Gyöngy
Dr J. Rodrigues
Dr D. Redfern

Abstract

The Variance Gamma (VG) process has been present in financial mathematics' research since the 1990s, due to its unique characteristics which make it advantageous in modeling asset returns. It is a stochastic process characterized by jumps rather than continuous paths and can have an infinite number of small jumps in any finite time interval, which leads to a fine-grained, jagged trajectory. However, despite the infinite number of small jumps, the overall magnitude of these jumps remains bounded, ensuring that the process does not accumulate infinite variation over time. As a result, it can capture both the frequent small fluctuations and occasional larger jumps in asset prices, while maintaining a controlled total variability. Its general form has an asymmetric Lévy distribution, meaning that the Lévy distribution's tails are able to decay in different rates as we approach positive and negative infinity. Thus it can model asset returns which have different sizes of positive vs negative jumps. In addition, the four parameters of its asset pricing model (r , σ , θ , ν) allow a level of control on the mean, variance, skewness and kurtosis of the log returns. While the VG process has a more leptokurtic distribution than the Normal in short maturities, it converges to the Normal for longer maturities, thus capturing another well known attribute of asset returns. It can capture volatility smiles, the phenomenon of implied volatility increasing as an asset's spot price diverges from its option's strike price. Last, it offers the advantage of being able to offer a closed form expression for arbitrage free option pricing of plain EU options. The purpose of present dissertation is the comprehension and application of these advantageous characteristics of the VG process. Firstly, the existing literature is explored to understand research challenges and recent improvements. The mathematical background is then presented, including the VG process, its distribution, its properties and its option pricing model. Subsequently, the algorithm for producing Monte Carlo (MC) paths of the VG process is presented, and paths are plotted for varying parameters. To better understand the effect of the parameters to the shape of the distribution of the VG log returns, a map for skewness and kurtosis is developed and upper/lower bounds are found for the achievable skewness-kurtosis combinations. A parametric analysis follows of the VG implied volatility (IV) surface. Finally, the VG option pricing model is calibrated based on an IV surface target of a more complex model with more degrees of freedom, as a benchmark of its capability to capture market complexity. The contribution of this thesis to existing research is that (a) it corrects the analytical VG option pricing formula for plain EU options, making its results consistent to those of the VG MC and inverse Fourier transform option pricing formulas, (b) it offers a clear visual map between the parameters of the VG process and the skewness and kurtosis of the VG log returns. It also provides a (c) charting between the parameters of the VG option pricing model and the IV surface, and (d) it benchmarks the VG option pricing model against the more complex Bates option pricing model with regards to their ability to capture the implied volatility of the S&P500 index, and discusses its limitations.

Acknowledgments

I would like to thank my advisors, Prof. Gyöngy from the School of Mathematics of the University of Edinburgh, and Dr Rodrigues and Dr Redfern from Moody's Analytics, for their invaluable guidance, support, and encouragement throughout the course of this research. Their expertise and insights were crucial in shaping the direction and content of this work.

On a personal note, I would like to thank my family for their unwavering support and encouragement. Their belief in me has always been a constant source of strength.

Own Work Declaration

I declare that this dissertation was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Literature review	2
1.3	Thesis contribution	10
1.4	Thesis outline	11
2	Mathematical models	12
2.1	Lévy processes	12
2.2	The Variance Gamma Process	13
2.3	The VG EU option pricing formula	16
2.4	Correcting the analytical VG option pricing formula	20
2.4.1	The MC option pricing formulas	20
2.4.2	The IFT option pricing formula	21
2.4.3	Comparative analysis and observations leading to recognition of the typo	21
2.4.4	Validation through simulations	23
2.5	Implied volatility	24
2.6	Target model	24
3	Results	26
3.1	Simulating VG process paths	26
3.2	Understanding the VG process skewness and kurtosis	26
3.3	Effect of parameters to the VG implied volatility surface	32
3.4	Calibration of the VG option pricing model	36
4	Conclusions	42
	Appendices	47
	A Code	47

List of Tables

1	Optimal parameters $q^* = [r, \theta, \sigma, \nu]^*$ sorted by increasing RMSE and the corresponding initial guesses. The parameter boundaries were [0, 0.21], [-2, 2], [0.01, 0.4], and [0.06, 4], with the exception of #1, and #6, where the boundaries for r were [-0.1, 0.1] and [0.01, 0.05] respectively.	38
2	Optimal parameters $q^* = [\theta, \sigma, \nu]^*$ by increasing RMSE and the corresponding initial guesses for $r = r(T)$ according to Figure 23. The parameter boundaries were [-2, 2], [0.01, 0.4], and [0.06, 4] respectively for θ, σ, ν	39
3	Optimal parameters $q^* = [r, \theta, \sigma, \nu]^*$ or $q^* = [\theta, \sigma, \nu]^*$ for $r = r(T)$ according to Figure 23, sorted by increasing RMSE. The parameter boundaries for θ, σ and ν were [-2, 2], [0.01, 0.4], and [0.06, 4].	40

List of Figures

1	Plain MC, antithetic variates and control variates variance reduction methods for the VG option pricing for (a) $N_p = 10^2$, (b) $N_p = 10^4$ and (c) $N_p = 10^6$	21
2	The solutions of MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	22
3	The error between MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	22
4	The solutions of analytical, CV MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	23
5	The error between CV MC vs analytical and IFT vs analytical VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	23
6	The solutions of analytical, CV MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	24
7	The error between CV MC vs analytical and IFT vs analytical VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.	24
8	(a) One and (b) multiple VG paths for $\nu = 0.17$, $\theta = 0.15$, $\sigma = 0.12$ and a weekly time step for a ten year time range.	27
9	Variance Gamma process paths with varying (a) ν , (b) θ , and (c) σ parameters respectively, from the reference selection $\nu = 0.17$, $\theta = 0.15$, $\sigma = 0.12$	27
10	Monte Carlo estimations and analytical solutions of (a) mean (with 95% confidence interval), (b) variance, (c) skewness and (d) kurtosis, vs time.	28
11	Skewness for varying VG parameters ν, θ, σ and $t = 1$	29
12	Kurtosis for varying VG parameters ν, θ, σ and $t = 1$	29
13	Achievable skewness and kurtosis combinations for $\theta \in [-2, 2]$, $\nu \in [0.001, 1]$, $\sigma \in [0.1, 4]$ and $t \in [\frac{1}{12 \times 4}, 10]$	30
14	Achievable skewness vs kurtosis area for the VG process and the upper and lower boundaries of equations (3.1), (3.5) plotted in blue dash-dot line for (a) $t = \frac{1}{12 \times 4}$ years (1 week) and (b) $t = 10$ years. The colorbar was used to denote parabolic slices of increasing ν of the area from down upwards.	31
15	Reference Variance Gamma implied volatility surface.	32
16	VG implied volatility surface for varying θ	33
17	VG implied volatility surface for varying σ	33
18	VG implied volatility surface for varying ν	34
19	VG implied volatility surface for varying r	34
20	Crude boundaries of parameters (a) r , (b) σ and (c) θ with increasing ν , while keeping the rest of the parameters constant.	35

21	Implied volatility surface of S&P500, as captured by the SVJD model, based on data up to 28/03/2024.	36
22	(a) Best overall calibrated VG IV surface and (b) absolute relative error in comparison to target.	38
23	Risk free rate vs maturity in 28/03/2024.	39
24	(a) Best calibrated VG IV surface for $r = r(T)$ according to Figure 23 and corresponding (b) absolute relative error in comparison to target.	40

Abbreviations

AD	Anderson-Darling test
ATM	At the money
AV	Antithetic variates
APE	Absolute pricing error criterion
ARM	Average relative measure
ARMS	Adaptive rejection Metropolis sampling algorithm
BGDS	Brownian Gamma bridge sampling
BNS	Barndorff-Nielsen and Shephard stochastic volatility model
B-S	Black-Scholes
CDO	Collateralised debt obligations
CGMY	Carr Geman Madan Yor model
CGMYSA	CGMY model prohibiting static arbitrage
CIR	Cox-Ingersoll-Ross model
COS	Option pricing model based on the Fourier-cosine series expansion
CPU	Central processing unit
CV	Control variates
DE	Differential evolution optimization algorithm
DJ	Dow Jones index
DGDS	Double Gamma bridge sampling
EMR	Entropic Monetary Risk
ELVG	Expanded local Variance Gamma model
EU	European
ES	Expected shortfall
ETF	Exchange traded fund
FCO	Foreign currency options
FD	Forward differences
FFT	Fast Fourier transform
FrFT	Fractional Fourier transform
GH	Generalised hyperbolic
HMR	Higher moment rank
IFT	Inverse Fourier transform
IG	Inverse Gaussian
ITM	In the money
IV	Implied volatility
IPA	Infinitesimal perturbation analysis
JD	Jump diffusion
KS	Kolmogorov-Smirnov test
LW	Lax-Wendroff discretisation algorithm
LR	Likelihood ratio
LRVG	Long range dependence Variance Gamma model
LRD	Long range dependence
LR-GL	Likelihood ratio method adjusted by Glasserman and Liu
MAD	Mean Absolute Difference

MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
MLE	Maximum likelihood estimate
MVG	Multivariate Variance Gamma
MUG	Monetary utility gain
NIG	Normal inverse Gaussian
NIGSA	NIG model prohibiting static arbitrage
NLLS	Non-linear least squares
OTM	Out of the money
O-U	Ornstein - Uhlenbeck
PDE	Partial differential equation
PDDE	Partial divided difference equation
PIDE	Partial integro differential equations
PDF	Probability density function
PSO	Particle swarm optimization algorithm
QMC	Quasi Monte Carlo
RMSE	Root mean square error
RV	Random variable
SA	Static arbitrage
SAM	Static arbitrage with martingale
SDE	Stochastic differential equation
s.t.	such that
SR	Sharpe ratio
SV	Stochastic volatility
SVG	Symmetric Variance Gamma
SVVG	Stochastic volatility Variance Gamma
SVJD	Stochastic volatility jump diffusion
TBRC	Three barrier reverse convertible
UNIVG	Univariate Variance Gamma
VaR	Value at Risk
VG	Variance Gamma
VGM	Variance Gamma with Markov regime change of volatility
VG	Variance Gamma
VGG	Variance Generalised Gamma
VGD	Variance Gamma model with additional diffusion term
VGSA	Variance Gamma model prohibiting static arbitrage

1 Introduction

1.1 Motivation

Stock price returns are known to have the following characteristics, [38], [5], [34], [20]:

- In the short term time horizon, their distribution is more leptokurtic than the Normal distribution. This means that in the short term, extreme price movements are more probable to happen than the tails of the Normal distribution predicts.
- In the long term, the returns' distribution is known to better fit the Normal distribution.
- The square returns show long range autocorrelation.
- The size of negative extreme price movements is not the same with the size of positive extreme movements.
- The volatility of returns is non-stable and it interchanges between periods (or clusters) of higher and lower volatility levels.
- Stock prices are affected by the speed of information arrival and transactions' volume in the market, usually referred to as 'activity'.
- Their options tend to demonstrate volatility smiles. As the ratio of the spot price $S(0)$ over strike price K (moneyness: $M = \frac{S(0)}{K}$) moves away from 1, the implied volatility tends to increase. The smile results from the existence of a skewed and heavy-tailed risk-neutral return distribution. We say that $S(0)$ related to K of a call option is at-the-money (ATM) if $M = 1$, in-the-money (ITM) if $M > 1$, and out-of-the-money (OTM) if $M < 1$.

While the Black-Scholes (B-S) model had been a powerful tool in the past for option pricing and still remains a benchmark model for educational purposes, its inability to capture the above important characteristics, especially reflecting to its inability to accurately capture risk exposure in periods of economic instability [50], has paved the way for new models' development. A comprehensive review of a variety of pricing models, their capabilities and limitations can be found in [21]. One of these models, was the VG asset pricing model, which emerged in the early 1990s and got established by the end of the first decade of 2000s. The VG process is a drifted Brownian motion, where the time follows a Gamma distribution:

$$X(t) = \theta G(t) + \sigma W(G(t)), \quad (1.1)$$

where W is the Wiener process, $\theta \in \mathbb{R}$ and $\sigma > 0$ are constants, and $G(t) \stackrel{\text{iid}}{\sim} \gamma(g|t, \nu)$, with

$$\gamma(g|t, \nu) = \frac{g^{\frac{t}{\nu}-1} \exp(-\frac{g}{\nu})}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} \quad (1.2)$$

being the Gamma distribution with shape $\frac{t}{\nu}$ and scale ν , with $t > 0, \nu > 0$. Thus in this thesis, $\gamma(g|t, \nu)$ will be used interchangeably with $\gamma(\frac{t}{\nu}, \nu)$. Here, $\Gamma(z)$ is the gamma function:

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u} du, \quad z > 0. \quad (1.3)$$

Under the VG process, stock prices can be formulated as:

$$S(t) = S(0) \exp [(m + \omega) t + X(t)], \quad (1.4)$$

where $\omega = \frac{1}{\nu} \ln(1 - \theta \nu - \frac{\sigma^2 \nu}{2})$, with $1 - \theta \nu - \frac{\sigma^2 \nu}{2} > 0$. The above formulations have been introduced for the reader's reference in the following introductory sections, and they will be analyzed in more detail in Chapter 2.

The attributes of the VG process that make it an attractive alternative for stock price modelling and option pricing are [41], [38], [11], [34]:

- The plain stock price model has four parameters (m or r , σ , θ, ν), enabling a level of control to the mean, variance, skewness and kurtosis of the log returns.
- Thicker tails than the normal for short maturities for the log returns, with the distribution of returns over longer periods approaching normality.
- Finite moments for at least the lower powers of returns.
- Different jump sizes for the gains vs the losses.
- At each interval of time, the price can move by infinite small jumps. At each time step, real time is translated to "business" or "market" time through the use of a Gamma distribution, capturing market activity.
- Independent stationary increments, with the distribution of any increment belonging to the same simple family of distributions irrespective of the length of time to which the increment corresponds.
- The Black-Scholes (B-S) model is a parametric case.
- It has a closed-form characteristic function allowing for analytical computation of its distribution and its four first moments, and enabling tractable solutions for the arbitrage free plain EU put/call option prices.
- It can capture volatility smiles.
- It can be extended to multivariate models, capturing correlation between underlying assets and long range correlation across time for the second order of returns.
- It can be extended to clustered volatility and stochastic volatility models, capturing volatility smiles better across maturities.

In view of the above remarkable characteristics, matching well with the properties of stock prices' returns, the model's simplicity, but also due to the promising empirical results found in existing literature [38], [35], [30], [1] there was strong motivation to further study existing research conducted until present for different VG models, their various computational implementations and the financial instruments that can be modeled with this process, and by locating the gaps in existing literature, contribute to present research with this thesis.

1.2 Literature review

The VG process model for share market returns was first proposed by D. Madan and E. Seneta in [41]. In this first approach, the VG process proposed was symmetric (SVG), that is the kurtosis parameter was $\theta = 0$, see equation (1.2). The characteristic function of the SVG was located and the method of producing the first four moments described. It was also concluded that the unit period distributions possessed finite moments of all orders. Furthermore, it was shown that kurtosis increased, with increasing ν , and it converged to Normal distribution kurtosis as $t \rightarrow \infty$. Also, that the parameters σ and ν could be calculated either with the moment matching or the transformed maximum likelihood estimation (MLE) method. The probability density function (pdf) of the SVG process was calculated, with the use of a Bessel function of the second kind and the SVG was extended in the multivariate case. Empirical tests showed that the SVG process was a better fit for daily stock market returns in comparison to the normal, the stable, and the Press compound events models in 12 out of 19 stocks studied, based on the χ^2 goodness of fit test.

As the research of the VG process matured in the next decade, Madan, Carr and Chang proved in [38] closed form solutions for the EU call and put options using the asymmetric ($\theta \neq 0$) VG process, and compared its option pricing capability with the B-S model. Their methodology started by locating the characteristic function of the VG process and its first four

central moments. The risk neutral stock price model was then introduced and the distribution of the log price was calculated. The EU option pricing formula was devised in terms of the modified Bessel function of the second kind and the degenerate hypergeometric function of two variables. The performance of the VG, SVG and B-S models were compared, by calibrating them using the MLE method in S&P500 futures options, based on the χ^2 goodness of fit test. The performance tests focused on similarity of calibrated distributions with empirical data and on quality of option prices the models deliver. After calibration, it was found that the drift and volatility of the B-S and VG models were very close in value, but the kurtosis and skewness parameters ν , θ were significantly different than 0 for the VG models. In the first test type, the VG model outperformed the SVG and B-S models in 93% of the cases, based on the log-likelihood ratio test, at 1% significance level. In the second test type, a regression analysis with orthogonality tests was performed on the pricing errors obtained from each model, with predictors moneyness, its square, maturity and interest rate. Again the VG model outperformed the rest based on R^2 and F-test for the regression and t-statistics for the parameters, with non-discriminable from 0 biases for moneyness and reduced maturity biases, at 1% and 5% significance levels.

A year later, Carr and Madan proposed the Fast Fourier Transform (FFT) as a method to value options efficiently by expressing the risk neutral option price as a function of the characteristic function of the log-price and numerically determining the option values, [13]. This paper came as a solution to the computational effort which was previously required to evaluate the analytical expression of the VG option pricing formula, involving elements as the modified Bessel function of the second kind and the degenerate hypergeometric function, terms of which had singularities within the parametric space. In the FFT method, the challenge that needed to be overcome was the singularity of the expression at the origin and the highly oscillatory behavior of the integral for OTM options. These were mitigated by the introduction of an exponential with parameter α , controlled by truncation error, and a hyperbolic sinusoidal transformation. It was shown that the FFT method introduced, required less than half the time of evaluating the analytical solutions of VG option pricing.

Later in [39], the authors proposed an improved VG process based metric in comparison to Sharpe ratio (SR) for portfolio optimization, called the Higher Moment Rank (HMR) statistic. In contrast to SR, which ranks investment options based on their σ^2 -standardized excess mean return, the HMR statistic took into account higher moment parameters. As a result, the paper proposed an improvement in comparison to SR ranking, particularly for investment options which have large positive or negative skews.

Another empirical evaluation of the VG model was published by Lam et al. in 2002, [35]. The VG, SVG and B-S models were calibrated to the prices of the Hang Seng index EU call options. The authors followed both an implied approach (recent 10-20-30 trades data) and a historical approach (recent 30-60-90 days data) for calibration. They compared the accuracy of models based on a parametric and a non-parametric approach, at 1% significance level. The parametric approach took into consideration absolute, squared and relative option pricing errors. The non-parametric, awarded or subtracted points based on model accuracy, and then derived winning probabilities, depending on which the models were tested. Furthermore, the models were compared with regards to systematic biases and hedging capability. With reference to the accuracy criterion, the VG model performed the best regardless of calibration strategy. In terms of systematic biases, the VG model outperformed all the other models, but still exhibited high biases in the implied calibration approach. When used in hedging, the VG model performed only marginally better, especially in the implied approach case. The authors reported though the gain was not economically significant. They concluded that while the VG model outperforms B-S, “..under the implied approach, the VG model continues to exhibit predictable biases and its overall performance in pricing and hedging is still far less than desirable”.

In the same year (2002), a model generalisation and a model improvement were proposed in [10] and [34]. In the first paper, the Lévy process of the VG process was extended by a parameter Y to allow for different combinations of finite/infinite activity and finite/infinite variation. The new 4 parameter process was named CGMY after the authors of the paper, and the stock prices

model was extended to in addition account for diffusion (5 parameters in total). Analytical calculation of its characteristic was followed by FFT MLE calibration, for 13 stock prices and 7 indexes. Interestingly, the diffusion coefficient level of quadratic variation contribution was zero in 13 cases and very small in 5 out of the rest 7 cases. Hence, the authors concluded diffusion components are not priced in the market for risks. In all cases but 3, the Y parameter had values implying infinite activity and finite variation (which is the case of the plain VG process), which thus the authors suggested was representative for modelling with jump diffusion components.

In the second paper, an important observation was brought into the foreground. VG and CGMY models are typically accurate when calibrated for a specific maturity, however, calibrations derived from short-term options do not perform well for long-term options and vice versa. This poor performance across different maturities can be attributed to the strict conditions on the term structure of risk-neutral variance, skewness, and kurtosis. Specifically, these models assume variance proportional to time, skewness that is inversely proportional to the square root of time, and kurtosis that is inversely proportional to time. In contrast, real data often show that variance, skewness, and kurtosis increase with longer times to maturity. An improvement to this drawback was hypothesized to be achieved by clustered or stochastic volatility models. The model introduced in the paper was the two state VG Markov (VGM) model, which proposed a two state volatility regime change. This new option pricing model had 10 parameters: a drift parameter, 2×3 parameters for each VG state process, two transition parameters, and an initial state probability parameter. The authors first located a closed form solution for the characteristic function of VGM, then used FFT to calculate the likelihood of the data and calibrated the model by MLE. Finally, the closed form of plain EU call and put options were found by Fourier inversion of the time value. In terms of testing, state transition parameters were found to be significant, and the VG model was rejected in favor of the VGM model based on log likelihood tests.

An alternative approach to clustering was proposed in [11] to better capture volatility, by introducing stochastic volatility (suffix -SV) to Lévy processes, as the Normal Inverse Gaussian (NIG), the VG and the CGMY processes. The introduction of stochastic volatility was achieved by subordinating the Lévy processes to stochastic time following either the time integral of the Cox-Ingersol-Ross (CIR) model or one of these three one-sided pure jump processes: a variant of the inverse Gaussian (IG) developed in the paper, and the stationary Gamma and Inverse Gaussian processes, being stationary solutions to the Ornstein-Uhlenbeck (O-U) SDE. A first class of risk-neutral stock price models was built by exponentiating these processes and normalizing to reflect initial forward prices to prohibit static arbitrage opportunities (suffix -SA). A second class of stock price models was obtained by compensating the jump terms in the Lévy processes to form martingales, which were then stochastically exponentiated (suffix -SAM). Initial empirical tests on S&P 500 options, showed that the exponential models were significantly better than their stochastic exponential counterparts for all three NIG, VG and CGMY endowed with CIR stochastic volatility. Further tests in this class of models, showed that the best performing model among the CGMYS, NIGSA and VGSA, was CGMYS, with consistently smaller percentage errors across all strikes and maturities.

On the simulations' side, Avramidis et al [4] proposed in 2003 efficient methods of simulating Gamma and VG processes. In more detail, they introduced gamma bridge sampling, a method selecting the points of Monte Carlo integration as to concentrate the variance of the integrand to a few coordinates, which cooperates well in efficiency and variance reduction with Quasi Monte Carlo (QMC) methods. In particular, for VG process sampling, two methods were proposed: the Brownian Gamma Bridge Sampling (BGDS) and Double Gamma Bridge Sampling (DGBS). By combining QMC methods (Korobov lattice rules and the Sobol' point sets) with the bridge sampling algorithms, the new method outperformed the plain QMC approach in efficiency for the same accuracy, which in turn performed better than MC. In general, Sobol' point sets performed better than Korobov rules with the bridge methods, and DGBS offered greater variance reduction, while BGDS showed better efficiency.

A systematic approach explaining the relation between the stock returns' skewness, kurtosis

and their effect on volatility smiles was conducted in, [5]. As a first step, the risk-neutral skewness and kurtosis of returns were defined over a period of time and expressed as functions of moneyness. Then, a formula was derived relating risk neutral market skewness, with physical market skewness and kurtosis. The risk neutral skewness of an individual stock was subsequently related with the skewness of the market, and it was concluded that it is less negative than the skew of the market. Following, the risk neutral stock skewness was decomposed into a systematic and idiosyncratic part. An expression was also derived relating implied volatility with risk neutral stock skewness and kurtosis. Based on these models, empirical tests were utilized to prove that the slope of individual stock smiles is flatter than that of the market index, and a more negative risk-neutral skew is related to a steeper negative slope of the implied volatility curve. Finally, it was shown that a higher risk-neutral kurtosis produces a flatter volatility smile.

In 2004, Hirsia and Madan [29] developed a numerical method of pricing American options for VG modeled stock prices, based on solving partial integro-differential equations (PIDE). First, the PIDE was derived by applying the infinitesimal generator to the value function of the American option. The space domain of the integral part was then discretised by finite differences (FD), and it was evaluated explicitly in all regions apart from the area of singularity close to 0, where it was treated implicitly. The time and space domain were treated with first order FDs. By applying a numerical manipulation, the PIDE turned into a difference equation, which was treated as solving a linear system of space equations at each time step. The VG process was calibrated by minimization of the root mean square error (RMSE) of the log market price errors, relating to OTM S&P500 options. Comparing American option prices between the B-S and VG model, B-S presented downwards biased premia in comparison to VG. The VG exercise boundary dominated the B-S boundary, suggesting perhaps thus a more extended continuation area. Fiorani in his PhD thesis [21], extended the method to EU and American barrier options and conducted sensitivity analysis for better understanding of the effect of the VG process parameters to the prices of the options.

In [14], the authors studied the application of the VG process with a diffusion term to foreign currency options (FCO). Four different models were used for the derivation of FCO prices, the B-S, the Merton Jump Diffusion (JD) and the VG and the VG with diffusion (VGD). The German mark option prices between 1996 and 1999 were used for MLE calibration and testing. To compare the models, the following tests were implemented: nested in-sample tests between models based on the likelihood ratio test, non-nested encompassing tests, out of sample tests using absolute and percentage pricing errors and (Cross-entropy and Hellinger) density distance tests. It was concluded that the contribution of the diffusion term in VGD was insignificant, and evidence were present of positive skewness and excess kurtosis. The VG option pricing model outperformed the rest, and while the VG model was able to encompass the JD model, the opposite was not evident.

A year later (2006), a historical anthology for the VG model's development and simulation was included in [24]. Three simulation methods of sequential sampling were presented, based on (a) Gamma time-changed Brownian motions, (b) difference of Gammas and, (c) as approximate compound Poisson processes. Furthermore, two methods of bridge sampling were described. Bridge sampling is a method of filling the MC path backwards, which more efficient with QMC and path dependent options. These bridge sampling methods were described as the (a) the Brownian Gamma time-changed bridge and (b) the Gamma difference bridge.

The same year, a paper was published on pricing collateralised debt obligations (CDO) with correlated VG processes, with the ambition to capture correlation smiles. Three structural VG model extensions were developed, one with correlated subordinators and independent identically distributed Brownian motions, another with independent subordinators but dependent Brownian motions, and the third being a sum of a systemic and an idiosyncratic VG process. Initially, the models were calibrated to the spread curve of DJ iTraxx 5 year (6-12/2005), based on the absolute pricing error (APE) criterion. Calibration took place in two steps. First, the VG parameters were fitted to the index spread curve given by quotes of the indices with different maturities. Then, the additional correlation parameters were calibrated such that the spread

of the equity tranche was matched. Only the third extension (with the addition of symmetry: $\theta = 0$) could simultaneously explain the index spreads for different maturities and the quotes of liquid index tranches of the index, performing better than a benchmark double $t(4)$ copula model in APE. Later in the paper, the dependence structure of the third model was extracted into a factor copula model and was calibrated to tranche prices. This approach led to the best fit overall to the correlation smile, compared to the above structural models and the benchmark double $t(4)$ copula model.

In 2007, Almendral and Oosterlee [2], complemented the research of American option pricing by proposing an improved model to mitigate the non-smooth fitness of American options at the exercise boundary under the VGD process, accompanied by a stability analysis. The system of PIDE was solved by finite differences discretisation, and implicit/ explicit methods for the integration. The Lax-Wendroff (LW) scheme, the Brennan-Schwartz algorithm and fast convolution by FFT were used to mitigate challenges in discretising different terms of the PIDEs and for computational stability and efficiency. Time was discretised by the BDF2 scheme for second order of error. Using von Neumann stability analysis, the size of time steps was found for the numerical solution to be stable. Numerical comparisons took place between the Euler scheme vs the BDF2 scheme, and central vs LW discretisation. It was concluded that while the Euler scheme had lower order accuracy than the BDF2 scheme, it provided more stable solutions for denser discretisations. The existence of a small diffusion term was found to smoothen the option delta. Furthermore, it was suggested that the LW scheme could provide more stable solutions for combination of parameters otherwise leading to non-smoothness.

Two further expansions were proposed subsequently for the VG process pricing model. In the first case, Semeraro [47] introduced dependence by proposing the multivariate VG model, having a multivariate subordinator with Gamma increments. The author first constructed the α -Gamma subordinator, a vector of n rows, each element of which is a linear combination of two random processes: one Gamma process for the whole vector and one Gamma process specific to each row of the vector. She first proved that the n -dimensional subordinator is a Lévy process with Gamma increments, and subsequently proved that a vector of n drifted Brownian motions, subordinated by the α -Gamma subordinator is a Lévy process, with Gamma increments. In the second case [20], Finlay suggested a VG process with long range dependence (LRD), by relaxing a subordinator constraint and independence between time increments. He initially noticed from empirical data that absolute and square returns were LRD dependent, even after 50-100 lags. Then showed that both the VG process model and the t -model are appropriate selections to introduce LRD dependence to squared returns. Among various fitting methods, it was found that the product-density maximum likelihood estimation performed the best, based on the Mean Absolute Difference (MAD) and Average Relative Measure (ARM) criteria between fitted and simulated series. When fitted in financial data (S&P500, Microsoft stock, AUS/USD, WTI), the LRVG model in some cases (Mic, AUS/USD) offered a superior fit to LR- t , where in others the fit between the models was almost the same. In pricing S&P500 options, when the risk neutral model was calculated using a mean correction construction, excluding static arbitrage instead of using skew-correcting martingales, the fit based on the RMSE criterion was better. Reportedly, the proposed model outperformed the 5 parameter Barndorff-Nielsen and Shephard (BNS) SV model of Schoutens but was outperformed by the VGSA and CGMYSA SV models, earlier discussed.

A new efficient approach extending the FFT named COS, was proposed in [16], based on Fourier cosine expansions. This method is applicable to approaches using directly the characteristic function to numerically calculate the pdf and the option price. Firstly, the whole integral of the characteristic function was reconstructed in terms of its Fourier-cosine series expansion. Then the reconstruction was applied to plain EU, digital and gap options. The expression was further simplified for Heston and Lévy processes, to be able to handle multiple strikes simultaneously. Numerical results offered for the B-S, Heston, VG and CGMY models, comparing with the Carr-Madan FFT approach showed that the algorithm converged to the same level of precision with much less computational resources required. For CGMY in particular, 80 vs 512

evaluations were required to achieve a 10^{-6} precision, translated to approx. 3 times less CPU time.

In 2010 [9], Cao and Fu discussed methods of calculating the Greeks for mountain range options, when the underlying assets are assumed to follow the VG process. Indirect methods (such as forward differences (FD)) and three direct methods (infinitesimal perturbation analysis (IPA), the likelihood ratio (LR) method and the variant of Glasserman and Liu (LR-GL)) were compared. The LR-GL method used the characteristic function of the model. In general, IPA was the closest match to FD, but was not applicable to discontinuous payoff cases. While LR-GL performed the same in general with LR, this came at excessive computational cost due to the additional Fourier transformation of the characteristic.

Becker proposed the following year fast and unbiased methods for Monte Carlo valuation of lookback, swing, and barrier options under VG models, called the adaptive DGBS method, [7]. The contribution of the paper was on adjusting the difference of Gammas bridge sampling (truncated DGBS) on the particular structure of these options (requiring only the final, minimal and maximal value of the underlying to be evaluated) and introducing a priori bounds on the bias. In the simulations followed for lookback and barrier option valuation, the performance measurements showed considerable reductions in computational effort and memory requirements in comparison to truncated DGBS, in particular for small precision bounds ϵ .

Wallmeier and Diether in [50] did a comprehensive model comparison between two multivariate VG models, the UNIVG (univariate Gamma subordinator, multivariate Brownian motion) and α -VG model. The empirical analysis was based on all three barrier reverse convertibles (TBRC) outstanding at the end of April 2007. Calibration of UNIVG used two types of input data: (1) the estimated marginal risk-neutral distributions of the underlying stock returns as represented in the strike price profiles of implied volatilities, and (2) the estimated pairwise correlations of stock returns. During fitting (1) was used for locating the optimal VG parameters based on RMSE and (2) was used to fit the correlations of the Brownian motions such that the asset correlations were equal to historical correlations. To calibrate the α -VG model, the marginal parameters were obtained in the first step from replicating the smile in option prices. In the second step, the single common parameter α is fixed such that the average pairwise correlation implied by the α -VG model was equal to the average empirical correlation. The α -VG model allowed better smile replication because the variance rate of the subordinator was stock-specific. Additionally, it was compatible with low correlations, within the upper bound of the α -parameter. In the UNI-VG model, low correlations could not be achieved if the smile profiles were pronounced and the parameters were similar for the underlying assets. However, if correlations were high, the UNI-VG model was better positioned. Thus, a tradeoff was noted between the two VG models: α -VG performed better in capturing the smile, UNIVG was better in capturing the correlation structure. However, it was mentioned that the performance of the α -VG model in capturing the correlation structure could be improved by accepting a less than perfect smile mapping.

A comparison between multivariate VG (MVG) models took place also in [30], this time regarding portfolio allocation. The MVG models compared were these of Madan [41], Semeraro α -VG [47] and Wang [51]. To introduce higher moments, the Taylor expansion of the utility function was considered. An empirical analysis took place based on a portfolio composed of 18 assets taken from the S&P500 index (04/2000-09/2011). A rolling-window optimization of six months in-sample and two months out-of-sample was implemented for the determination of the weights. The monetary utility gain (MUG) was used to analyze the out-of-sample performance. Four moment portfolios performed better than two moment portfolios, with the exception of the Semeraro α -VG, where two moment portfolios performed better. It was found that for all levels of risk aversion, the highest MUG was obtained using the Wang model, observing also the limitations of α -VG to capture more complicated dependence structures.

In 2015, Mozumder et al. [43] proposed an alternative approach to pricing EU options modeled under VG processes, by use of the Fractional Fourier Transformation (FrFT) instead of the FFT, in methods utilizing the characteristic function for option pricing. The VG model

was fitted on weekly S&P500 traded options data over the sample period 01-11/2007. It was found that the FrFT reduced calculation time by 97–98% relative to the FFT, with small impact ($\sim 2\%$) on the RMSE.

The same year, Linders and Stassen [36] proposed a method of efficiently calculating basket option prices, where the underlying assets were modelled with an MVG process. To express the basket options price, the additivity property of comonotonic stop-loss premiums together with Gauss-Laguerre polynomials were used. Two steps were carried out in the calibration procedure. Firstly, the marginal parameters and the distribution of the common time change were calibrated to market quotes of the vanilla options by using the Carr-Madan formula. Secondly, an implied estimate for the average correlation was determined by minimizing the relative error between the market and model index option quotes. As an illustration, it was shown that the multivariate VG model was able to closely match the observed Dow Jones index options.

In [46], the various methods of fitting the plain VG process and its pdf were compared, along with the various tests for parameter estimation accuracy that were available in literature until 2016. Daily Dow Jones (DJ) index and stocks data were used for calibration between 01/1991 and 12/2011. For back-testing the results of the model with the safe haven theory, the yields of daily US Treasuries with a maturity of 6 m, 12 m, 5 y, 10 y and the closing prices of gold were used. A regime switching model was utilised for identifying normal and turbulent phases of the economic time series. For calibrating the regime switching model, a database of weekly DJ closing prices from 01/1991 to 12/2011. The authors found that fitting rates were much better when the Hamilton regime switching model was applied, based on the Kolmogorov-Smirnov (KS) test. Using the Anderson-Darling (AD) test, it was found that for heavily tailed stocks, the χ^2 fitting method worked better than the MLE.

A comparison between the VG, NIG and generalised hyperbolic (GH) model in a Markov regime switching framework for high frequency trades took place in [28]. The calibration of the models was done based on Shanghai Composite Index dataset from 01/2000-09/2013. The models were calibrated based on the MLE method. The authors employed the log-likelihood ratio, the χ^2 and the AD test statistics for the goodness-of-fit analysis. It was shown that while for daily returns the VG model outperformed the rest, as the time scale decreased (from daily to 5 minute intervals), the NIG model outperformed the VG model consistently and with increasing goodness-of-fit statistics.

A further generalisation of the MVG model, the Variance Generalised Gamma convolution (VGG), having a generalised Gamma convolutions subordinator was proposed in [8]. The VMT^d -subclass, a subclass including VG and α -VG, was isolated for concept proving in financial applications. A limitation mentioned for this subclass, is that there is no separation between parameters describing marginal distributions, and parameters describing the dependence, which is a consequence of the fact that the family of Gamma distributions is not stable under convolution. After it was proved the VMT^d -class is invariant under an Esscher transformation, risk neutral pricing was achieved using it as an equivalent martingale measure with respect to the risk free rate numeraire. The method was applied to best-of/worst-of EU and American put options and the authors found the simulated outcomes for varying correlation and strike prices reasonable.

In the topic of portfolio risk exposure, Ivanov in [33] offered explicit risk metric expressions for portfolios with assets modeled under the VG process with dependence, as the Value at Risk (VaR), the Expected Shortfall (ES) and the Entropic Monetary Risk (EMR).

In 2017, Warty et al. proposed a Bayesian estimation method for CIR stochastic volatility inference in VG models, using sequential importance sampling/resampling, [52]. The method combined a sequential learning auxiliary particle filter with a particle learning filter and it was developed to address the issue of intractability of the likelihood for Stochastic Volatility VG (SVVG) models. Proper conjugate priors were selected for the parameters, mainly plain/truncated Gaussians, with the exception of the initial volatility which was taken as Gamma. A Gibbs sampler was used for all parameters except volatility which was sampled by an adaptive rejection Metropolis sampling (ARMS) algorithm. A first test, with benchmark the Markov Chain Monte

Carlo (MCMC) method of Li et al, based on synthetic data, showed that sequential filtering tends to underestimate latent variance when the true latent state attains very large values for short periods. With regards to the parameters estimation, while the sequential method converged very fast, the accuracy of some of the parameters was very poor, in comparison to the MCMC method. The authors commented also on their method's inadequacy to capture the filtering of jumps and Gamma time, for small drift and variance values. The second test was implemented on S&P500 index data from 01/1980-12/2000, showing similar limitations for the same processes and parameters.

Another interesting application of the VG process appeared in 2020, where it was used for modeling Rugby Union matches, [23]. In contrast to soccer, where traditionally Poisson models were considered instructive for modelling, in Rugby due to the complex scoring system the same models are non-descriptive. Hence, the authors modeled the points scored for each team in rugby by Gamma random variables, having the same scale parameter. Then the difference in points between the two competing teams could be modelled as a VG process. Based on this, the probabilities for win, draw and defeat were calculated, the distribution of total points scored was found, and the probability of losing bonus point or try bonus point were also elicited. It was shown that the model could be calibrated either with the use of historical results or using bookmakers' data. The calibrated model, was shown to offer realistic results and insights for a preliminary model with regards to out-of-sample tests.

An information based approach for pricing financial instruments modeled under the VG model was developed in [32]. The information process was constructed based on a normalised SVG bridge, the associated Gamma bridge and a market factor X_T , where the bridges are jointly Markovian. The market filtration was generated by the information process together with the gamma bridge. Hence, given the distribution of the market factor, the payoff function and the calibrated information process, an integral form of the option pricing formula was derived. Using this formula, explicit pricing formulas were devised for a unit-principal credit-risky bond without recovery, a defaultable bond with random recovery, and assets with log-normally and exponentially distributed payoff at maturity.

To address the oscillating behavior of existing VG based option pricing models in short maturities, the author of [1] proposed the use of Mellin transformation for path independent options. More precisely, in this representation, it was possible to express the price of a path-independent option as the product of the Mellin transforms of the VG density and of the option's payoff, where the Mellin transforms were expressed in fast converging series. The OTM, ATM and ITM option prices of various financial instruments (digital asset/cash or nothing call, EU call, EU call-ATM forward, gap and power call options) modelled with the SVG were derived as infinite series. For the asymmetric VG case, the OTM, ATM and ITM option prices were derived for digital cash or nothing call options. Numerical results for the SVG, showed that the method could offer excellent agreement with Fourier transform, Gauss–Laguerre or MC methods, with the evaluation of only 15 terms of the infinite series.

The expanded local VG model (ELVG), an efficient SV based method, being able to capture volatility smiles across different maturities, was proposed in [12]. In this model, instead of subordinating Gamma time to a stochastic process, it was subordinated by a deterministic function. First the authors located the Dupire forward PDE, and then with proper manipulation calculated the forward partial divided-difference equation (PDDE). The local variance was assumed to be a piecewise linear function of strike and piecewise constant function of time. The calibration of the model to market smiles did not require solving any optimization problem or numerical PDEs. In contrast, it could be solved term-by-term as a system of non-linear algebraic equations for each maturity. For comparison, data were taken from XLF and S&P500 indexes, utilized in previous papers. It was concluded that the provided surfaces were improved in comparison to previous research, as they provided arbitrage free prices by construction.

In [44], the author showcased once more the convenience of using FrFT for the calculation of the log-likelihood and its first two derivatives by applying it to VG models and the B-S model, and by calibrating them to SPY ETF historical prices. Based on the KS test, it was found

that the MLE method provided the best fit for the SVG model, with moment matching based methods and the B-S based model being rejected.

In the topic of sensitivities of VG based models, Udoye et al [49], calculated the VG parameter option sensitivities for a Vacicek zero coupon bond model driven by the VG process. The sensitivities of drift parameter θ , volatility parameter σ and kurtosis parameter ν , were found in the simple form of integrals of finite sum factors, based on the application of Malliavin calculus.

In 2022, a VG based model with finite activity was introduced (VG++), for markets with low liquidity (e.g. the energy market). Instead of subordinating the Brownian motion to a Gamma stochastic process, after self decomposition of the Gamma stochastic process, the authors chose to subordinate the Brownian motion to the a-remainder of the Gamma distribution. It was proven that this process is a Lévy process with finite variation and is a subordinator. An algorithm was also constructed for its forward simulation. The characteristic function of the VG++ process and its Lévy measure were calculated. The VG++ process was shown to have finite activity and variation. Then its first four cumulants and its pdf was found. Moreover, a market price model was developed based on this process and an expression for EU call options was derived. A backward simulation scheme based on Lévy bridges was constructed. Based on this method, three approaches were used for the calculation of option prices: the closed form solution, an MC method and the FFT. The FFT errors with regards to the closed form formula were smaller than the MC errors for the test's computational budget. The VG++ and VG model were calibrated on historical data of power future market quotations using the MLE method. Finally, the models were fitted on quoted vanilla contracts using the standard non linear least squares (NLLS) technique and non standard derivatives with backward simulations were priced. VG++ returned higher prices for derivatives in these illiquid markets than the VG model, which the authors found reasonable, since in illiquid markets the hedging strategies are difficult to implement and therefore, option sellers tend to increase the option premia.

VG stochastic volatility models came once more in the foreground in [45]. The model had five parameters in total and had O-U stochastic instantaneous variance, where the Brownian term in O-U SDE was replaced with a background driving Lévy process $z(t)$. It was proved that the stationary marginal distribution of the stochastic instantaneous variance is Gamma, if $z(t)$ is a compound Poisson process. Its characteristic function and pdf was calculated, and the same was done for the equivalent risk neutral measure under the Esscher transformation. A closed form solution was found for the EU-call option, in terms of the risk neutral pdf and also in term of the characteristic. The model was calibrated with SPY ETF daily data from 01/2010-12/2020, based on the MLE method, using either FrFT or the twelve-point composite Newton–Cotes quadrature formulas. It was reported that for small timeframes, FrFT underestimated the peak of the density function, and also that FrFT offered poor results for ITM options. The model option prices across moneyness and maturity were compared with the B-S model, and it was found that they were less than B-S for moneyness 0.8-1 and greater than B-S for moneyness 1-1.2, with values converging for deep ITM or OTM areas.

Finally, two papers were published in spring 2024, focusing on VG's distributional characteristics, [22], [26]. The first one offered a comprehensive review of the VG distribution. It covered probability and cumulative distribution functions, generating functions, moments and cumulants, mode and median, Stein characterisations, representations in terms of other random variables, related distributions and methods for parameter estimation. The contribution of the second paper, was that it offered formulas for all absolute raw and central moments, expressed in terms of infinite series involving the modified Bessel function of the second kind and the modified Lommel function of the first kind.

1.3 Thesis contribution

In view of the previous remarkable research efforts, the following areas were observed that could be complemented with the research in this thesis.

First of all, researchers have recorded biases to the simple VG analytical option pricing formula of [38] in comparison to VG option pricing formulas using the IFT, the MC and other

approaches, [15] and [31], but have not explored them further. In addition, researchers normally avoid the VG analytical option pricing model in their benchmarks with new VG variants' models, [1]. In present thesis, two discrepancies of the analytical option pricing formula of [38] are located, discussed and corrected. The analytical results are validated with simulations. This is a very important finding, as the typos found in [38] not only have been carried over to recent literature, (see [22] equation (4.59)), but also as we will see in Section 2.4 lead to arbitrage opportunities.

Secondly, while current research has touched on the relationship between returns' skewness and kurtosis for stocks and markets [5], in this thesis a clear visual linkage between each parameter of the VG process vs the skewness and kurtosis is offered. In addition, it is shown that achievable skewness and kurtosis pairs at each time step are bounded by two parabolas of second order, with analytical expressions.

Thirdly, in [50], a tradeoff between correlation and smile steepness of the volatility surface was observed. Here, by creating a visual map between the VG IV surface and the VG option pricing models' parameters, it was confirmed that for the single dimension model the parameter σ controls smile steepness, and that less shallow smiles can be achieved for higher σ values. In addition, the role of ν in the steepness of the smiles was emphasized. It was also concluded that the strength of the smile short term for the VG IV is coupled with the strength of the slope across maturity in the long term. This means that the VG IV surface can either have steep smiles short term and steep slopes long term across maturity or shallow smiles short term and flat slopes long term across maturity. The above observation sheds additional light to the observation of [34] of the inability of the VG option pricing model to capture the implied volatility surface across maturity, due to the skewness and kurtosis of the log returns being inverse proportional to time.

Finally, this thesis demonstrates a straightforward calibration process, utilizing different optimization methods, and benchmarks the VG option pricing model with the Bates option pricing model, based on the IV surface. Thus, it provides insights on the VG model's capabilities and shortcomings, in comparison to a model with more than twice degrees of freedom. To the author's knowledge this comparison has not been presented before in literature.

1.4 Thesis outline

The rest of the paper is structured as follows. Chapter 2 offers the definitions and mathematical formulation of Lévy processes, the Lévy measure, the Gamma distribution, the VG process and its application to risk neutral option pricing. In addition, the analytical VG plain EU option pricing formula found in literature is compared to other methods (MC, IFT), its inconsistencies are found and corrected. In Chapter 3, the VG process is simulated, and the skewness and kurtosis of the VG log returns are mapped for varying parameters (θ, σ, ν). Moreover, the achievable (skewness, kurtosis) pairs at each time step are shown to be bounded by two parabolas. The corrected analytical VG option pricing formula is used to produce the VG IV surface. The VG IV surface is in turn calibrated to the S&P500 IV surface produced by the Stochastic Volatility Jump Diffusion (SVJD) model of Moody's Analytics (also known as Bates model). The two IV surfaces are compared to understand the VG option pricing model's capabilities and limitations. Chapter 4 concludes, offering motivation for future work.

2 Mathematical models

This chapter introduces the reader to Lévy processes, subordinators and their properties. It then presents the VG process, its characteristic function, distribution and moments. It continues by demonstrating the corrected VG option pricing model, and the method which lead to its correction. While some proofs were considered beyond the scope of this thesis, interested readers can refer to [3] for the first part and [38] for the second and third parts. Finally, the chapter concludes by explaining the concept of implied volatility, volatility smiles and providing details regarding the target model.

2.1 Lévy processes

Definition 2.1. Let $X = (X(t))_{t \geq 0}$ be a stochastic process defined on a probability space (Ω, \mathcal{F}, P) . We say that it has independent increments if for each $n \in \mathbb{N}$ and each $0 \leq t_1 < t_2 \leq \dots < t_{n+1} < \infty$ the random variables $(X(t_{j+1}) - X(t_j), 1 \leq j \leq n)$ are independent, and that it has stationary increments if each

$$X(t_{j+1}) - X(t_j) \stackrel{d}{=} X(t_{j+1} - t_j) - X(0).$$

We say that X is a Lévy process if:

- (L1) $X(0) = 0$ (a.s.);
- (L2) X has independent and stationary increments;
- (L3) X is stochastically continuous, i.e., for all $a > 0$ and for all $s \geq 0$,

$$\lim_{t \rightarrow s} P(|X(t) - X(s)| > a) = 0.$$

Note that in the presence of (L1) and (L2), (L3) is equivalent to the condition

$$\lim_{t \downarrow 0} P(|X(t)| > a) = 0$$

for all $a > 0$.

Definition 2.2. A probability measure μ on the Borel σ -algebra of \mathbb{R}^d is infinitely divisible, if for every integer $n \geq 2$, there is a probability measure μ_n , such that $\mu = \underbrace{\mu_n * \mu_n * \dots * \mu_n}_{n \text{ times}}$, with “ $*$ ” denoting convolution.

Theorem 2.1 (Lévy-Khintchine). A Borel probability measure μ on \mathbb{R}^d is infinitely divisible if there exists a vector $b \in \mathbb{R}^d$, a positive semi-definite symmetric $d \times d$ matrix A , and a Lévy measure ν on $\mathbb{R}^d \setminus \{0\}$ such that, for all $u \in \mathbb{R}^d$,

$$\phi_\mu(u) = \exp \left(i(b, u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d \setminus \{0\}} (e^{i(u, y)} - 1 - i(u, y)\chi_{\hat{B}}(y)) \nu(dy) \right),$$

where ϕ_μ denotes the characteristic function of μ , $\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u, x)} \mu(dx)$ and $\hat{B} = B_1(0)$.

Conversely, any mapping of the form above is the characteristic function of an infinitely divisible probability measure on \mathbb{R}^d .

Remark 2.1. The triple (b, A, ν) are called the characteristics of the infinitely divisible random variable X , when X is a random vector in \mathbb{R}^d with distribution μ .

Remark 2.2. The Lévy measure is denoted here by bold ν to distinguish it from the kurtosis parameter of the VG process, ν .

Remark 2.3. Note that the characteristic function can be written in the form $\phi_\mu(u) = e^{\eta(u)}$. We will refer to the map $\eta : \mathbb{R}^d \rightarrow \mathbb{C}$ as a Lévy symbol, since it serves as the symbol for a pseudo-differential operator. Other authors call η the characteristic exponent or Lévy exponent. Since, for all $u \in \mathbb{R}^d$, $|\phi_\mu(u)| \leq 1$ for any probability measure μ and $\phi_\mu(u) = e^{\eta(u)}$, when μ is infinitely divisible we deduce that $\operatorname{Re}\eta(u) \leq 0$.

Remark 2.4. Suppose that an infinitely divisible probability measure is such that its Lévy measure ν is absolutely continuous with respect to the Lebesgue measure. We write $g_\nu = \frac{d\nu}{dx}$ and call it the Lévy density.

Theorem 2.2. If $X = (X(t))_{t \geq 0}$ is a Lévy process, then the characteristic function $\phi_{X(t)}(u)$ is given by

$$\phi_{X(t)}(u) = e^{t\eta(u)}$$

for each $u \in \mathbb{R}^d$ and $t \geq 0$, where η is the Lévy symbol of $X(1)$.

Definition 2.3. The Lévy–Khinchine formula for a Lévy process $X = (X(t))_{t \geq 0}$ is,

$$E(e^{i(u, X(t))}) = \exp \left(t \left[i(b, u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d \setminus \{0\}} (e^{i(u,y)} - 1 - i(u,y)\hat{\chi}_B(y)) \nu(dy) \right] \right)$$

for each $t \geq 0$, $u \in \mathbb{R}^d$, where (b, A, ν) are the characteristics of $X(1)$.

Definition 2.4 (Subordinator). A subordinator is a one-dimensional Lévy process that is non-decreasing (a.s.). Such processes can be thought of as a random model of time evolution, since if $T = (T(t))_{t \geq 0}$ is a subordinator we have, $T(0) = 0$ and $T(t) \geq 0$ a.s. for each $t > 0$, and $T(t_1) \leq T(t_2)$ a.s. whenever $t_1 \leq t_2$.

Theorem 2.3 (Lévy Symbol of a Subordinator). If T is a subordinator, then its Lévy symbol takes the form

$$\eta(u) = ibu + \int_0^\infty (e^{iuy} - 1)\lambda(dy),$$

where $b \geq 0$ and the Lévy measure λ satisfies the additional requirements:

$$\lambda((-\infty, 0)) = 0 \quad \text{and} \quad \int_0^\infty (y \wedge 1)\lambda(dy) < \infty.$$

Conversely, any mapping from $\mathbb{R} \rightarrow \mathbb{C}$ of the form above is the Lévy symbol of a subordinator. We call the pair (b, λ) the characteristics of the subordinator T .

Theorem 2.4. Let X be an arbitrary Lévy process and let T be a subordinator defined on the same probability space as X such that X and T are independent. Define a new stochastic process $Z = (Z(t))_{t \geq 0}$ by the prescription $Z(t) = X(T(t))$, for each $t \geq 0$, so that for each $\omega \in \Omega$, $Z(t)(\omega) = X(T(t)(\omega))(\omega)$. Then Z is a Lévy process.

2.2 The Variance Gamma Process

The VG process is a drifted Brownian process subordinated to a Gamma process, as seen in equation (1.1). As both the drifted Brownian process and the Gamma subordinator are Lévy processes, then by Theorem 2.4 the VG process is a Lévy process.

Proposition 2.1. The drifted Brownian motion defined as $B(t) = \theta t + \sigma W(t)$, is a Lévy process, where $t > 0$, $W(t)$ is a Wiener process and $\sigma > 0$, $\theta \in \mathbb{R}$.

Proof. To show that $B(t) = \theta t + \sigma W(t)$ is a Lévy process, we need to verify that it satisfies the defining properties of a Lévy process:

1. $B(0) = 0$: By definition, $W(0) = 0$, hence $B(0) = \theta \cdot 0 + \sigma \cdot 0 = 0$.

2. Independent increments: Since $W(t)$ has independent increments, the increments of $B(t)$, expressed as $\theta(t-s) + \sigma(W(t) - W(s))$, are also independent $\forall s, t$ s.t. $0 < s < t$. The term $\theta(t-s)$ is deterministic and does not affect the independence imparted by the Wiener process part.
3. Stationary increments: The increments of $B(t)$ are of the form $\theta(t-s) + \sigma(W(t) - W(s))$. Since $W(t) - W(s) \sim N(0, t-s)$, it follows that $\sigma(W(t) - W(s)) \sim N(0, \sigma^2(t-s))$ and thus $B(t) - B(s) \sim N(\theta(t-s), \sigma^2(t-s))$, indicating stationary increments.
4. Continuity of paths: The paths of $B(t)$ are continuous, given that both θt and $\sigma W(t)$ are continuous functions of t . We recall here that continuous processes are also continuous in probability, but not the other way around.

Therefore, $X(t)$ meets all the criteria for being a Lévy process. \square

Proposition 2.2. *The Gamma process G defined as $G = (G(t))_{t>0}$, where $G(t) \stackrel{iid}{\sim} \gamma(g | t, \nu)$ with the γ distribution being defined by equation (1.2), is a Lévy process.*

Proof. We prove that G is a Lévy process by verifying the defining properties:

1. Starts at 0: For $t = 0$, $G(0) \sim \gamma(0, \nu)$, with $E(G(0)) = 0 \cdot \nu = 0$ and $Var(G(0)) = 0 \cdot \nu^2 = 0$, so the pdf is concentrated around the origin and hence $G(0) = 0$ (a.s.).
2. Independent increments: The Gamma process is constructed such that the increments $G(t) - G(s)$ for $s < t$ are independent. This is inherent in the construction of the Gamma process where each increment is defined by an independent Gamma distribution.
3. Stationary increments: For any $t > s$, the increment $G(t) - G(s)$ follows a Gamma distribution with parameters $\frac{t-s}{\nu}$ and ν due to the additive property of independent Gamma-distributed increments. Hence, the distribution of the increments depends only on the difference $t - s$, satisfying the stationary increments condition.
4. Continuity in probability: Given that the Gamma process is constructed from Gamma distributions, the probability of making a large jump in any finite interval is 0 (no jumps). Thus, $G(t+s) \rightarrow G(t)$ in probability as $s \rightarrow 0$, showing continuity in probability.

These properties confirm that G satisfies all the criteria for being a Lévy process. \square

Proposition 2.3. *The characteristic function of $G(t)$ with pdf defined by equation (1.2) is:*

$$\phi_G(u) = (1 - iu\nu)^{-\frac{t}{\nu}}$$

Proof.

$$\begin{aligned} \phi_G(u) &= \mathbb{E}[e^{iuG}] = \int_0^\infty e^{iug} \gamma(g|t, \nu) dg = \int_0^\infty \frac{g^{\frac{t}{\nu}-1} \exp[-\frac{g}{(\frac{1}{\nu}-iu)^{-1}}]}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg = \\ &= \frac{(\frac{1}{\nu} - iu)^{-\frac{t}{\nu}}}{\nu^{\frac{t}{\nu}}} \int_0^\infty \frac{g^{\frac{t}{\nu}-1} \exp[-\frac{g}{(\frac{1}{\nu}-iu)^{-1}}]}{(\frac{1}{\nu} - iu)^{-\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg = \frac{1}{(1 - iu\nu)^{\frac{t}{\nu}}}. \end{aligned}$$

\square

Proposition 2.4. *The characteristic function of the drifted Brownian motion defined in Proposition 2.1 is:*

$$\phi_B(u) = e^{(iu\theta - \frac{\sigma^2 u^2}{2})t}$$

Proof.

$$\begin{aligned}\phi_B(u) &= \mathbb{E}[e^{iuB}] = \mathbb{E}[e^{iu\theta t + iu\sigma W(t)}] = e^{iu\theta t} \mathbb{E}[e^{iu\sigma W(t)}] = \\ &= e^{iu\theta t} \int_{-\infty}^{\infty} e^{iu\sigma x} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx = e^{iu\theta t - \frac{\sigma^2 u^2}{2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-iu\sigma t)^2}{2t}} dx = \\ &= e^{iu\theta t - \frac{\sigma^2 u^2 t}{2}}.\end{aligned}$$

□

Proposition 2.5. *The characteristic function of the VG process defined in equation (1.1) is:*

$$\phi_X(u) = (1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2)^{-\frac{t}{\nu}}$$

Proof. The characteristic function of the VG process is derived by applying the tower rule, as follows:

$$\begin{aligned}\phi_X(u) &= \mathbb{E}[e^{iuX}] = \mathbb{E}[\mathbb{E}[e^{iuB(G)} \mid G = g]] = \int_0^{\infty} \phi_B(u|g)\gamma(g|t,\nu) dg = \\ &= \int_0^{\infty} e^{(iu\theta - \frac{\sigma^2 u^2}{2})g} \frac{g^{\frac{t}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg = [\nu(1/\nu - iu\theta + \frac{\sigma^2 u^2}{2})]^{-\frac{t}{\nu}} = (1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2)^{-\frac{t}{\nu}}.\end{aligned}$$

□

Remark 2.5. *The distribution of the VG process can be located by applying inverse Fourier transformation:*

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} (1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2)^{-\frac{t}{\nu}} du.$$

Alternatively, it can be expressed as a conditional Gaussian, times a marginal Gamma, where the g variable is integrated out:

$$f_X(x) = \int_0^{\infty} \frac{1}{\sqrt{2\pi g}} e^{-\frac{(x-\theta g)^2}{2g}} \frac{g^{\frac{t}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg,$$

noting that $\mathbb{E}[X|G=g] = \mathbb{E}[\theta g + \sigma W(g)] = \theta g$ and $Var(X|G=g) = Var(\theta g + \sigma W(g)) = \sigma^2 g$.

Proposition 2.6. *The expectation, variance, skewness and kurtosis of the VG process are:*

- $\mathbb{E}[X] = \theta t$
- $Var(X) = (\theta^2\nu + \sigma^2)t$
- $Skewness(X) = \left[\frac{2\theta^2\nu^2 + 3\sigma^2\nu}{(\theta^2\nu + \sigma^2)^{3/2}} \right] \frac{\theta}{t^{1/2}}, \quad \text{where } \left[\frac{2\theta^2\nu^2 + 3\sigma^2\nu}{(\theta^2\nu + \sigma^2)^{3/2}} \right] \geq 0$
- $Kurtosis(X) = 3 + 3 \left[\frac{\sigma^4 + 4\sigma^2\theta^2\nu + 2\theta^4\nu^2}{(\theta^2\nu + \sigma^2)^2} \right] \frac{\nu}{t}, \quad \text{where } \left[\frac{\sigma^4 + 4\sigma^2\theta^2\nu + 2\theta^4\nu^2}{(\theta^2\nu + \sigma^2)^2} \right] \geq 0.$

Proof. The first four moments are calculated from the derivatives of the characteristic as:

$$\mathbb{E}[X^k] = i^{-k} \left. \frac{d^k \phi}{du^k} \right|_{u=0}$$

In detail:

$$\mathbb{E}[X] = i^{-1} \left(-\frac{t}{\nu} \right) \left(1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2 \right)^{-\frac{t}{\nu}-1} (-i\theta\nu + \nu\sigma^2 u) \Big|_{u=0} = -\frac{ti}{\nu i} (-\theta\nu) = \theta t$$

$$\begin{aligned}\mathbb{E}[X^2] &= i^{-2} \left[-\frac{t}{\nu} \left(-\frac{t}{\nu} - 1 \right) \left(1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2 \right)^{-\frac{t}{\nu}-2} (-i\theta\nu + \nu\sigma^2 u)^2 \right. \\ &\quad \left. - \frac{t}{\nu} \left(1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2 \right)^{-\frac{t}{\nu}-1} \nu\sigma^2 \right] \Big|_{u=0} = \theta^2 t^2 + (\theta^2\nu + \sigma^2)t\end{aligned}$$

$$\mathbb{E}[X^3] = \theta^3 t^3 + 3(\theta^3\nu + \theta\sigma^2)t^2 + (2\theta^3\nu^2 + 3\theta\sigma^2\nu)t$$

$$\mathbb{E}[X^4] = \theta^4 t^4 + (6\theta^4\nu + 6\sigma^2\theta^2)t^3 + (11\theta^4\nu^2 + 12\sigma^2\theta^2\nu + 3\sigma^4\nu)t^2 + (6\theta^4\nu^3 + 6\sigma^2\theta^2\nu^2 + 3\sigma^4\nu^2)t$$

Subsequently, the central moments are derived as:

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\theta^2\nu + \sigma^2)t$$

$$\mathbb{E}[(X - \mathbb{E}[X])^3] = \mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mathbb{E}[X] + 3\mathbb{E}[X](\mathbb{E}[X])^2 - (\mathbb{E}[X])^3 = (2\theta^3\nu^2 + 3\theta\sigma^2\nu)t$$

$$\begin{aligned}\mathbb{E}[(X - \mathbb{E}[X])^4] &= \mathbb{E}[X^4] - 4\mathbb{E}[X^3]\mathbb{E}[X] + 6\mathbb{E}[X^2](\mathbb{E}[X])^2 - 4\mathbb{E}[X](\mathbb{E}[X])^3 + (\mathbb{E}[X])^4 \\ &= (3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2)t^2 + (3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3)t\end{aligned}$$

Finally, by applying the formulas of skewness and kurtosis:

$$\text{Skewness}(X) = \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{[\text{Var}(X)]^{3/2}} = \frac{2\theta^3\nu^2 + 3\theta\sigma^2\nu}{(\theta^2\nu + \sigma^2)^{3/2} t^{1/2}}$$

$$\begin{aligned}\text{Kurtosis}(X) &= \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{[\text{Var}(X)]^2} = \frac{(3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2)t + (3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3)}{((\theta^2\nu + \sigma^2)^2 t)} \\ &= \frac{3(\theta^2\nu + \sigma^2)^2 t + (3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3)}{(\theta^2\nu + \sigma^2)^2 t} = 3 + \frac{3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3}{((\theta^2\nu + \sigma^2)^2 t)}\end{aligned}$$

□

Remark 2.6. It is clear from these expressions that the skewness and kurtosis $\rightarrow \infty$ as $t \rightarrow 0$. Furthermore, skewness $\rightarrow 0$ and kurtosis $\rightarrow 3$ as as $t \rightarrow \infty$, which are the skewness and kurtosis of the Normal. In addition, the skewness has the same sign as θ and skewness $\propto \theta$, and excess kurtosis $\propto \nu$, where the symbol “ \propto ” denotes proportionality.

2.3 The VG EU option pricing formula

It is recalled that the B-S asset price formula under the risk-neutral measure \mathbb{Q} is:

$$S_{BS}(t) = S_{BS}(0)e^{(r-\frac{\sigma^2}{2})t+\sigma W^Q(t)},$$

where $r \in \mathbb{R}$ the risk free rate, $\sigma > 0$ the volatility and W^Q is a Wiener process under the risk neutral measure \mathbb{Q} . Since the discounted stock price is a martingale under \mathbb{Q} with regards to its filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, the following property should hold:

$$\mathbb{E}^{\mathbb{Q}} [S_{BS}(t)e^{-rt} | \mathcal{F}_s] = S_{BS}(s)e^{-rs}, \quad 0 \leq s \leq t \tag{2.1}$$

$$\begin{aligned} &\iff \mathbb{E}^{\mathbb{Q}} \left[\frac{S_{BS}(t)}{S_{BS}(s)} e^{-r(t-s)} \middle| \mathcal{F}_s \right] = 1 \iff \mathbb{E}^{\mathbb{Q}} \left[\frac{S_{BS}(0) e^{(r-\frac{\sigma^2}{2})t + \sigma W^Q(t)}}{S_{BS}(0) e^{(r-\frac{\sigma^2}{2})s + \sigma W^Q(s)}} e^{-r(t-s)} \middle| \mathcal{F}_s \right] = 1 \\ &\iff e^{-\frac{\sigma^2}{2}(t-s)} \mathbb{E}^{\mathbb{Q}} \left[e^{\sigma(W^Q(t) - W^Q(s))} \middle| \mathcal{F}_s \right] = 1, \end{aligned}$$

and since $W^Q(t) - W^Q(s)$ is independent of \mathcal{F}_s :

$$\iff e^{-\frac{\sigma^2}{2}(t-s)} \mathbb{E}^{\mathbb{Q}} \left[e^{\sigma(W^Q(t) - W^Q(s))} \right] = 1 \iff \mathbb{E}^{\mathbb{Q}} \left[e^{\sigma(W^Q(t) - W^Q(s))} \right] = e^{\frac{\sigma^2}{2}(t-s)},$$

which clearly holds because $W^Q(t) - W^Q(s) \sim N(0, t-s)$ and hence holds equation (2.1).

In the same manner, we want the discounted VG stock price to be a martingale under \mathbb{Q} . The asset price of the VG model under \mathbb{Q} is:

$$S_{VG}(t) = S_{VG}(0) e^{(r+\omega)t + X^Q(t)}, \quad (2.2)$$

where $X^Q(t) = \theta G(t) + \sigma W^Q(G(t))$ is \mathcal{F}_t adapted and $\omega \in \mathbb{R}$ a normalizing constant, necessary for the martingale property to hold. For the discounted VG stock price to be a martingale under \mathbb{Q} with regards to its filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, we demand that:

$$\mathbb{E}^{\mathbb{Q}}[e^{X^Q(t)}] < \infty,$$

which holds as

$$\mathbb{E}^{\mathbb{Q}}[e^{X^Q(t)}] = \mathbb{E}^{\mathbb{Q}}[e^{iuX^Q(t)}] \Big|_{u=-i} = \phi_{X^Q(t)}(u) \Big|_{u=-i} = (1 - \theta\nu - \frac{\nu\sigma^2}{2})^{-\frac{t}{\nu}} < \infty,$$

and in addition we demand:

$$\mathbb{E}^{\mathbb{Q}} \left[S_{VG}(t) e^{-rt} \middle| \mathcal{F}_s \right] = S_{VG}(s) e^{-rs}, \quad 0 < s \leq t. \quad (2.3)$$

Following the same route as before and by recalling that the VG process is a Lévy process with independent increments, such that $X^Q(t) - X^Q(s)$ is independent of \mathcal{F}_s :

$$\begin{aligned} &\iff e^{\omega(t-s)} \mathbb{E}^{\mathbb{Q}} \left[e^{X^Q(t) - X^Q(s)} \right] = 1 \iff \mathbb{E}^{\mathbb{Q}} \left[e^{X^Q(t) - X^Q(s)} \right] = e^{-\omega(t-s)} \\ &\iff (1 - \theta\nu - \frac{\nu\sigma^2}{2})^{-\frac{(t-s)}{\nu}} = e^{-\omega(t-s)} \end{aligned}$$

Solving for ω :

$$\iff \omega = \frac{1}{\nu} \ln(1 - \theta\nu - \frac{\nu\sigma^2}{2}), \quad (2.4)$$

with $1 - \theta\nu - \frac{\nu\sigma^2}{2} > 0$.

Remark 2.7. By demanding that $1 - \theta\nu - \frac{\nu\sigma^2}{2} > 0$, we get $\nu(\theta + \frac{\sigma^2}{2}) < 1$ or alternatively $\theta < \frac{1}{\nu} - \frac{\sigma^2}{2}$. Clearly, this is a necessary condition for the discounted VG stock price process to be a martingale.

Remark 2.8. The log returns $L = \ln \frac{S_{VG}(t)}{S_{VG}(0)} = (r + \omega)t + X^Q(t)$ have mean $\mathbb{E}[L] = (r + \omega)t + \mathbb{E}[X^Q] = (\theta + r + \omega)t$ and the same variance, skewness and kurtosis with the VG process, see Proposition 2.6.

Theorem 2.5. The density for the log price $Z = \ln \left(\frac{S_{VG}(t)}{S_{VG}(0)} \right)$ when prices follow the VG process dynamics of equation (2.2) is given by:

$$f_Z(z) = \frac{2 \exp(\frac{\theta}{\sigma^2}x)}{\nu^{\frac{t}{\nu}} \sqrt{2\pi\sigma^2} \Gamma(\frac{t}{\nu})} \left(\frac{x^2}{2\sigma^2/\nu + \theta^2} \right)^{\frac{t}{2\nu}-\frac{1}{2}} K_{\frac{t}{\nu}-\frac{1}{4}} \left(\frac{\sqrt{2\sigma^2/\nu + \theta^2}}{\sigma^2} |x| \right),$$

where K is the modified Bessel function of the second kind, and $x = z - rt - \frac{t}{\nu} \ln(1 - \theta\nu - \frac{\nu\sigma^2}{2})$.

Proof. Conditional on the Gamma variate G , Z is normally distributed with $\mathbb{E}[Z|G = g] = (r + \omega)t + \theta g$ and $Var(Z|G = g) = \sigma^2 g$. Hence the distribution of Z can be found by multiplying the conditional with the marginal and integrating out g :

$$f_Z(z) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 g}} e^{-\frac{(z-(r+\omega)t+\theta g))^2}{2\sigma^2 g}} \frac{g^{\frac{t}{\nu}-1} e^{-\frac{g}{\nu}}}{\nu^{\frac{t}{\nu}} \Gamma(\frac{t}{\nu})} dg,$$

which with use of Lemma [27], 3.471(9.):

$$\int_0^\infty x^{\alpha-1} e^{-\frac{\beta}{x} - \gamma x} dx = 2 \left(\frac{\beta}{\gamma} \right)^{\frac{\alpha}{2}} K_\alpha \left(2\sqrt{\beta\gamma} \right),$$

takes the form of the theorem, noting $\alpha = \frac{t}{\nu} - \frac{1}{2}$, $\beta = \frac{x^2}{2\sigma^2}$ and $\gamma = \frac{\theta^2}{2\sigma^2} + \frac{1}{\nu}$. \square

Remark 2.9. It is recalled the modified Bessel function of the second kind K_α can be expressed as:

$$K_\alpha(u) = \int_0^\infty e^{-u \sinh t} e^{-\alpha t} dt$$

Theorem 2.6. The European call option price of an asset, when the risk-neutral dynamics of the asset price is given by the VG process of equation (2.2) is,

$$\begin{aligned} p_{VG}(S(0), K, t) &= S(0) \Psi \left(d \sqrt{\frac{1-c_1}{\nu}}, (a+s) \sqrt{\frac{\nu}{1-c_1}}, \frac{t}{\nu} \right) \\ &\quad - K e^{-rt} \Psi \left(d \sqrt{\frac{1-c_2}{\nu}}, a \sqrt{\frac{\nu}{1-c_2}}, \frac{t}{\nu} \right), \end{aligned}$$

where $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$, $s = \frac{\sigma^2}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$, $c_1 = \frac{\nu(a+s)^2}{2}$, $c_2 = \frac{\nu a^2}{2}$, and

$$d = \frac{1}{s} \left(\ln \left(\frac{S(0)}{K} \right) + rt + \frac{t}{\nu} \ln \left(\frac{1-c_1}{1-c_2} \right) \right),$$

with $\frac{1-c_1}{1-c_2} > 0$ and Ψ is expressed explicitly in terms of the modified Bessel function of the second

kind and the degenerate hypergeometric function Φ of two variables.

$$\begin{aligned} \Psi(x_1, x_2, x_3) = & \frac{c^{x_3+\frac{1}{2}} \exp(\text{sign}(x_1)c)(1+u)^{x_3}}{\sqrt{2\pi}\Gamma(x_3)x_3} \\ & K_{x_3+\frac{1}{2}}(c)\Phi\left(x_3, 1-x_3, 1+x_3; \frac{1+u}{2}, -\text{sign}(x_1)c(1+u)\right) \\ & - \text{sign}(x_1) \frac{c^{x_3+\frac{1}{2}} \exp(\text{sign}(x_1)c)(1+u)^{1+x_3}}{\sqrt{2\pi}\Gamma(x_3)(1+x_3)} \\ & K_{x_3-\frac{1}{2}}(c)\Phi\left(1+x_3, 1-x_3, 2+x_3; \frac{1+u}{2}, -\text{sign}(x_1)c(1+u)\right) \\ & + \text{sign}(x_1) \frac{c^{x_3+\frac{1}{2}} \exp(\text{sign}(x_1)c)(1+u)^{x_3}}{\sqrt{2\pi}\Gamma(x_3)x_3} \\ & K_{x_3-\frac{1}{2}}(c)\Phi\left(x_3, 1-x_3, 1+x_3; \frac{1+u}{2}, -\text{sign}(x_1)c(1+u)\right), \end{aligned}$$

where $c = |x_1|\sqrt{2+x_2^2}$, $u = \frac{x_2}{\sqrt{2+x_2^2}}$ and,

$$\Phi(y_1, y_2, y_3; p_1, p_2) = \frac{\Gamma(y_3)}{\Gamma(y_1)\Gamma(y_3-y_1)} \int_0^1 x^{y_1-1} (1-x)^{y_3-y_1-1} (1-p_1x)^{-y_2} e^{p_2 x} dx,$$

where $S(0)$ is the spot price of the underlying asset, K is the strike price, r is the risk free rate, and T the option maturity.

Proof. The proof is omitted, because it is presented in detail in the Appendix of [38]. \square

Remark 2.10. The careful reader will notice two deliberate differences to this corrected option pricing formula, in comparison to the formula presented in Theorem 2 of p.88 of [38].

1. In p.88 of [38], the second input of Ψ of the strike term is incorrectly presented as “as $\sqrt{\frac{\nu}{1-c_2}}$ ”. The correct expression, found in the Appendix of [38] (p. 99), is “a $\sqrt{\frac{\nu}{1-c_2}}$ ”.
2. In p.88 of [38], the “a” parameter is mentioned to be $a = \zeta s$, where $\zeta = -\frac{\theta}{\sigma^2}$ and $s = \frac{\sigma}{\sqrt{1+(\frac{\theta}{\sigma})^2 \frac{\nu}{2}}}$. This results to $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$. This is incorrect as proved in the next section. The correct expression is $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$.

This finding is very important, because with $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$, the necessary martingale property (2.3) does not hold, and the option pricing model therefore admits arbitrage.

Remark 2.11. By demanding in the corrected formula $\frac{1-c_1}{1-c_2} > 0$, we get $\frac{2-\nu(a+s)^2}{2-\nu a^2} > 0$ and by replacing $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$ and $s = \frac{\sigma^2}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$ we get:

$$\begin{aligned} \frac{2\sigma^2 + \nu\theta^2 - \nu(\sigma^2 + \theta)^2}{2\sigma^2 + \nu\theta^2 - \nu\theta^2} > 0 \iff & \frac{2\sigma^2 + \nu\theta^2 - \nu\sigma^4 - \nu\theta^2 - 2\nu\theta\sigma^2}{2\sigma^2} > 0 \\ & 1 - \frac{\nu\sigma^2}{2} - \nu\theta > 0 \end{aligned}$$

which validates the constraint of Remark 2.7.

2.4 Correcting the analytical VG option pricing formula

Looking into past research, it was observed that the results of the analytical VG option pricing formula (see p.88 of [38]), did not exactly match those obtained using the Monte Carlo (MC) method, even with a sufficiently large number of paths, nor did they align with the inverse Fourier transformation (IFT) method, as referenced in [15], [31]. This section offers the path of reason that led to the tracking of the second typo of Remark 2.10 and the correction of the analytical formula. Firstly, the MC and the IFT option pricing formulas are introduced, which were the benchmarks against which the erroneous analytical option pricing formula was compared.

2.4.1 The MC option pricing formulas

The VG risk neutral price for the i -th MC path is:

$$S_i(t) = S(0) e^{(\omega+r)t + X_i(t)}, \quad (2.5)$$

where $S(0)$ is the spot price, r is the risk free rate, $\omega = \frac{1}{\nu} \ln(1 - \theta\nu - \frac{\nu\sigma^2}{2})$ and

$$X_i(t) = \theta \Delta G_i(t) + \sigma \sqrt{\Delta G_i(t)} Z_i(t), \quad (2.6)$$

with $\Delta G_i(t) \stackrel{\text{iid}}{\sim} \gamma(\frac{t}{\nu}, \nu)$ and $Z_i(t) \stackrel{\text{iid}}{\sim} N(0, 1)$. Then the fair option price of the vanilla EU call option with contingent claim $h = (S(T) - K)_+$ is:

Plain MC (P):

$$p_P(T) = e^{-rT} \frac{1}{N_p} \sum_{i=1}^{N_p} \max(S_i(T) - K, 0), \quad (2.7)$$

where N_p is the number of paths, T is the maturity, and K the strike price. Since the plain vanilla EU call option is not path dependent, from the Central Limit Theorem, as $N_p \rightarrow \infty$, then $p_{MC}(T) \rightarrow p_{VG}(S(0), K, T)$. For the sake of comparison, the following variance reduction methods were also implemented:

Antithetic Variates (AV):

$$p_{AV}(T) = e^{-rT} \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{\left[\max(S_i(T) - K, 0) + \max(\tilde{S}_i(T) - K, 0) \right]}{2}, \quad (2.8)$$

with $\tilde{S}_i(T) = S(0) e^{(\omega+r)t - X_i(t)}$.

Control Variates (CV):

$$p_{CV}(T) = \bar{Y}_2 - b (\bar{Y}_1 - S(0)) \quad (2.9)$$

where

$$\begin{aligned} Y_{1,i} &= e^{-rT} S_i(T) \\ Y_{2,i} &= e^{-rT} \max(S_i(T) - K, 0) \\ \bar{Y}_1 &= \frac{1}{N_p} \sum_{i=1}^{N_p} Y_{1,i} \\ \bar{Y}_2 &= \frac{1}{N_p} \sum_{i=1}^{N_p} Y_{2,i} \end{aligned}$$

$$b = \frac{\sum_{i=1}^{N_p} [(Y_{1,i} - \bar{Y}_1)(Y_{2,i} - \bar{Y}_2)]}{\sum_{i=1}^{N_p} (Y_{1,i} - \bar{Y}_1)^2}$$

For all three MC methods, the fair call option price was plotted against maturity for $N_p = \{10^2, 10^4, 10^6\}$ and $r = 0.042, \nu = 0.1686, \theta = -0.1436, \sigma = 0.1213, S(0) = K = 100$, see Figure 1. It was evident that all the methods converged to the same line with $N_p \rightarrow \infty$, and also that $\text{var}(P) > \text{var}(AV) > \text{var}(CV)$.

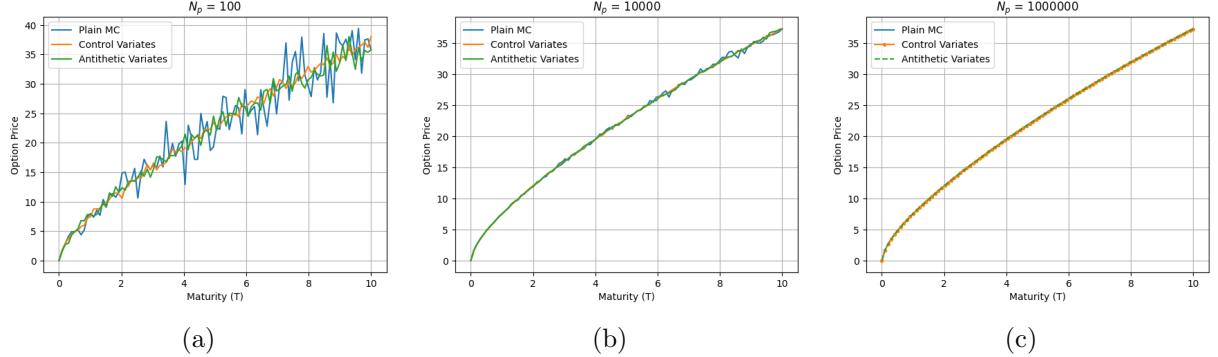


Figure 1: Plain MC, antithetic variates and control variates variance reduction methods for the VG option pricing for (a) $N_p = 10^2$, (b) $N_p = 10^4$ and (c) $N_p = 10^6$.

2.4.2 The IFT option pricing formula

The formula in this case originates from [13], and is as follows for the plain EU call option:

$$p_F(T) = S(0)P_1 - Ke^{-rT}P_2, \quad (2.10)$$

where

$$\begin{aligned} P_1 &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[\frac{e^{-iu \ln(K)} \phi_T(u-i)}{iu \phi_T(-i)} \right] du \\ P_2 &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[\frac{e^{-iu \ln(K)} \phi_T(u)}{iu} \right] du \\ s_T &= \ln[S(T)] = \ln[S(0)] + (\omega + r)T + X(T) \end{aligned}$$

$$\begin{aligned} \phi_T(u) &= E[e^{iust}] = \exp\{iu(\ln[S(0)] + (r + \omega)T)\} \exp\{iuX(T)\} \stackrel{\text{Pr. 2.5}}{=} \\ &= \exp\{iu(\ln[S(0)] + (r + \omega)T)\} \left(1 - i\theta\nu u + \frac{\nu\sigma^2}{2}u^2\right)^{-\frac{T}{\nu}}, \end{aligned}$$

with $\Re(\cdot)$ denoting the real part of a complex function.

2.4.3 Comparative analysis and observations leading to recognition of the typo

Theoretically, as $N_p \rightarrow \infty$, then $p_{MC}(T) \rightarrow p_{VG}(T)$ and $p_{MC}(T) \rightarrow p_F(T)$. Firstly, the latter was tested by plotting the results of the MC VG option pricing formula for $N_p = 10^5$ against those of the IFT formula. The VG MC and IFT plots coincided across moneyness and maturity, see Figure 2. The MC errors were found to oscillate around 0, as is evident in Figure 3.

The same comparative analysis was repeated for the incorrect analytical option pricing formula found in [38], (where $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$). To avoid congestion, this time the plots included the

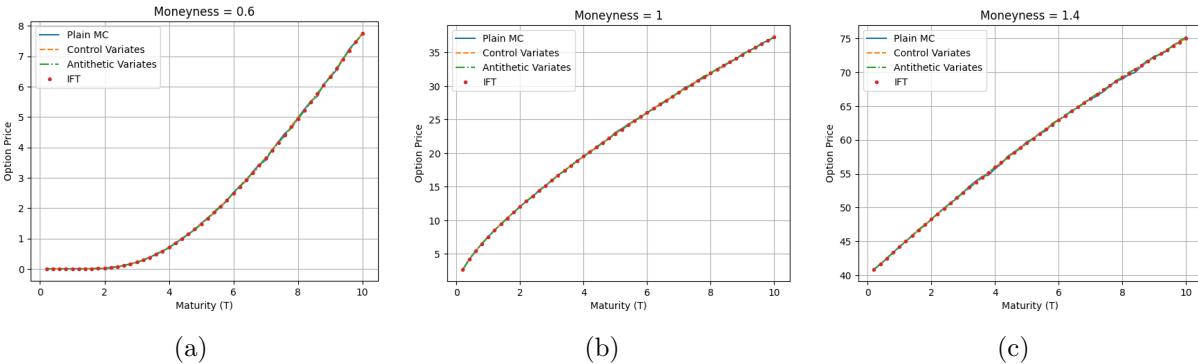


Figure 2: The solutions of MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

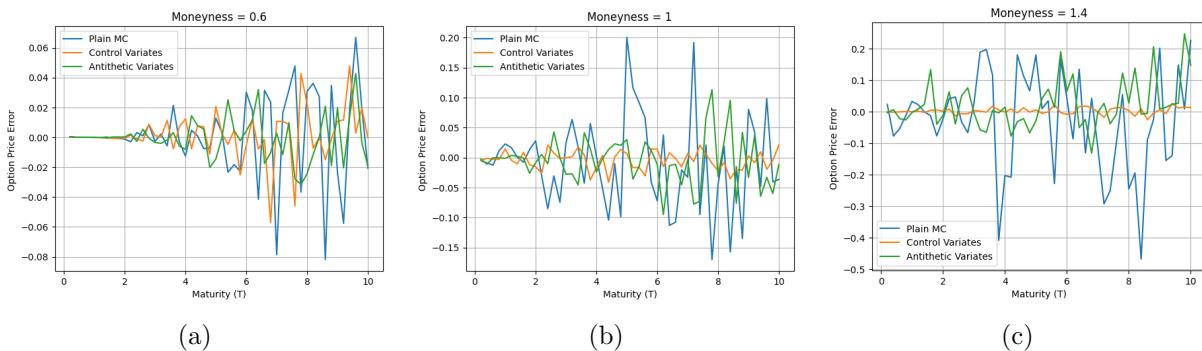


Figure 3: The error between MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

analytical, IFT and CV MC solutions and the error was plotted between CV MC vs analytical and IFT vs analytical. It was observed that the analytical option prices seeminly coincided with the CV MC and IFT option prices for ATM and ITM spot prices (as shown in Figures 4b and 4c), but not for OTM spots, (Figure 4a). The expectation for the error plots would be for the CV MC vs analytical to oscillate symmetrically around 0 and for the IFT vs analytical to be a constant horizontal line at 0. Looking more carefully however at the error plots, it was noted that the errors had certain trends for all moneyness values (see Figure 5). This led to Observation 1.

Observation 1. *Comparing the MC with the IFT VG option prices we note their trends coincide and the error between them is random and around 0. This implies that both formulas' results are consistent. However, comparing the MC and IFT prices with the analytical option prices of [38], it is evident that there is non-random error between the analytical and the rest of the formulas' results, showing a large trend with maturity, and being the largest for OTM spot prices.*

Subsequently, it was noted that for the incorrect option pricing expression of [38], (where $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$), the constraint of Remark 2.11, would have a different sign in front of $\nu\theta$ it comparison to Remark 2.7. In more detail, by demanding $\frac{1-c_1}{1-c_2} > 0$, the calculations would result to $1 + \nu\theta - \frac{\nu\sigma^2}{2} > 0$. This implied that ω in this formula might have been incorrectly taken as $\omega = \frac{1}{\nu} \ln(1 + \nu\theta - \frac{\nu\sigma^2}{2})$. This is an important observation, as we've seen that $\omega = \frac{1}{\nu} \ln(1 - \nu\theta - \frac{\nu\sigma^2}{2})$ is a necessary condition for the discounted VG stock price process to be a martingale, see equations (2.3) and (2.4). And unless the discounted stock price is a martingale, it is well known that there can be found arbitrage opportunities.

Tracking the derivation of the analytical option pricing formula back to equation (6.1) of p. 49 of [40] for the symmetrical VG case ($\theta = 0$), it became evident that the initial assumption for the option pricing model for the asymmetric VG model ($\theta \neq 0$) was:

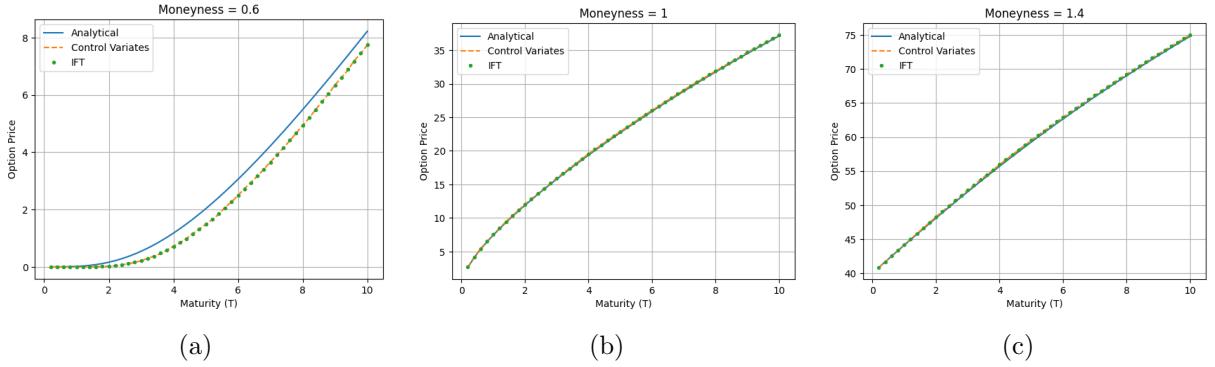


Figure 4: The solutions of analytical, CV MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

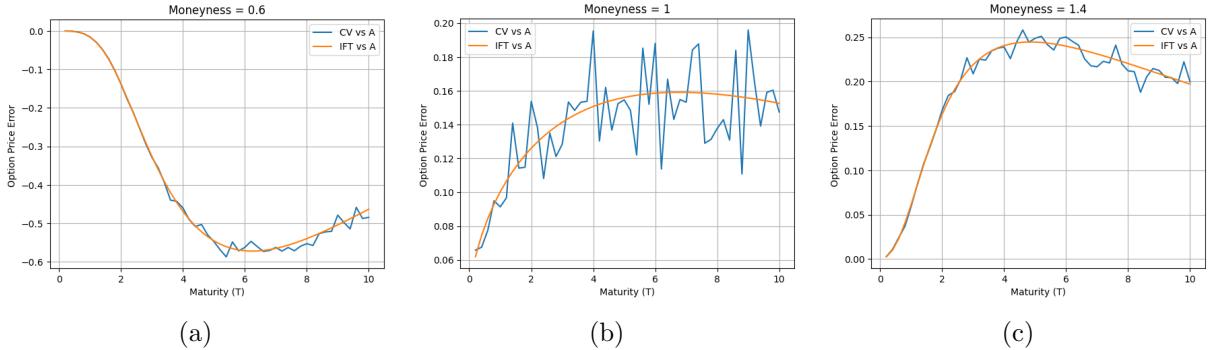


Figure 5: The error between CV MC vs analytical and IFT vs analytical VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

$$S(t) = S(0) e^{[r + \frac{1}{\nu} \ln(\frac{(1-\nu(a+s)^2/2)}{1-\nu a^2/2})]t + X(t)}$$

By demanding that the necessary martingale condition of equation (2.3) holds, it was concluded that the following condition must hold:

$$\frac{1 - \nu(a + s)^2/2}{1 - \nu a^2/2} = 1 - \theta\nu - \frac{\nu\sigma^2}{2},$$

which for $s = \frac{\sigma^2}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}} > 0$ holds only if $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$. The above analysis led to Observation 2.

Observation 2. *The necessary martingale condition of equation (2.3) holds only if $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$. Hence, if $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$ the VG option pricing model admits arbitrage.*

The above observations led to the conclusion that the correct analytical VG option pricing formula should have $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$.

2.4.4 Validation through simulations

To validate this conclusion, the sign of “ a ” was changed to $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$ resulting to the formula of presented in Theorem 2.6 of this thesis and the CV MC and IFT option prices were plotted once more against the analytical for varying moneyness and ($N_p = 10^5$). In this case, it was observed in Figure 6 that the CV MC and IFT solutions matched the analytical ones across maturity and moneyness. Furthermore, the CV MC error vs the analytical was random around

0 and the IFT vs analytical error was very close to 0, (see Figure 7). As with the correction of “ a ” to $a = +\frac{\theta}{\sqrt{\sigma^2 + \nu\theta^2}}$ all three methods’ solutions matched and the error demonstrates the expected behavior, the analytical results of the previous section were computationally validated.

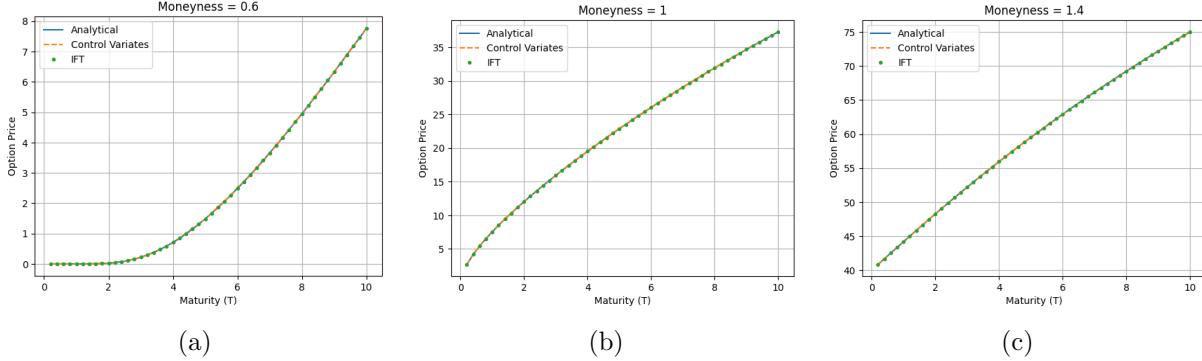


Figure 6: The solutions of analytical, CV MC and IFT VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

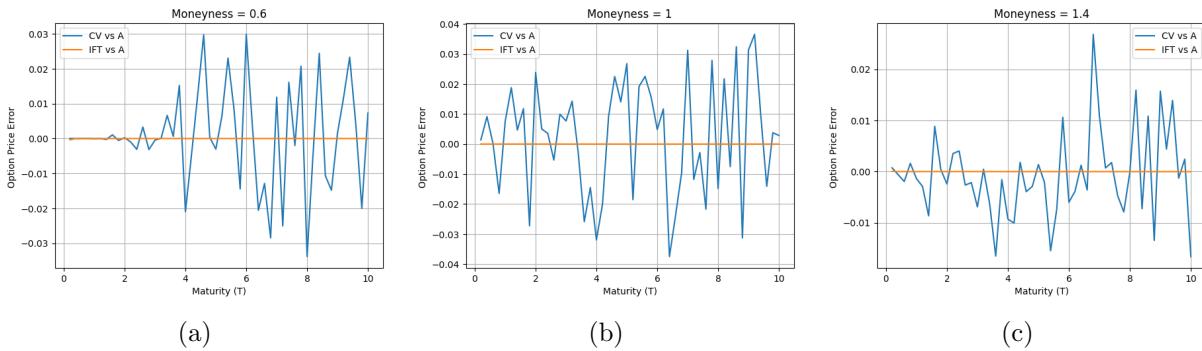


Figure 7: The error between CV MC vs analytical and IFT vs analytical VG option pricing methods across maturity with moneyness equal to (a) 0.6 , (b) 1 and (c) 1.4 respectively.

2.5 Implied volatility

Given a target option price $p^*(0)$, with origin either from market data directly or from a model that can capture volatility curvature, the implied volatility (IV) is found by equating the B-S model price with the target price $p_{BS}(0) = p^*(0)$ and solving with regards to the volatility term σ of the B-S model. The IV surface is produced by plotting the IV values for different moneyness values $M = \frac{S(0)}{K}$ and maturities T . We say that a call option price is at-the-money (ATM) if $S(0) = K$, in-the-money (ITM) if $S(0) > K$ and out-of-the-money (OTM) if $S(0) < K$. Clearly, the inequality signs have the opposite directions to the above for ITM and OTM put options. For a given maturity, if the IV curves are convex, we say they display a “volatility smile”, with lower IV values close to ATM regions and higher IV values for deep OTM and ITM values. Intuitively, this means that the volatility of an underlying asset grows, as its spot price moves from ATM to ITM and OTM regions.

2.6 Target model

One of the penultimate purposes of this thesis is to understand the capabilities of the VG process option pricing model, having four calibration parameters (r , θ , ν and σ), in capturing the IV surface of a more complex model used by Moody’s Analytics, the nine parameter Stochastic Volatility Jump Diffusion (SVJD) model (known also as the Bates model [6]), which is essentially

a Merton jump model with Heston stochastic volatility. The underlying asset is hence modeled as:

$$dS(t) = S(t_-) \left[(\mu - \lambda \bar{\mu})dt + \sqrt{v(t)}dW^{(1)}(t) + (J-1)dN(t) \right]$$

$$dv(t) = a(\theta - v(t))dt + \xi \sqrt{v(t)}dW^{(2)}(t),$$

where μ is the cost of carry, $W^{(1)}(t)$ and $W^{(2)}(t)$ are Wiener processes with correlation ρ , $N(t)$ is a Poisson counter with risk-neutral intensity λ , $N(t) \sim Poisson(\lambda t)$ and $(J-1)$ is the random relative jump size with $J = LogNormal(\mu_J, \sigma_J^2)$.

3 Results

This chapter provides an analysis of how to simulate Monte Carlo paths for the VG process, alongside an examination of its distributional properties, including skewness and kurtosis, across varying maturities. A comprehensive map is developed charting the skewness and kurtosis of the VG log returns for varying parameters, and highlighting the achievable combinations of (skewness, kurtosis) with time and their upper/lower boundaries. Finally, the VG option pricing model is calibrated to replicate the IV surface of the SVJD model and the results are discussed.

3.1 Simulating VG process paths

This section aims to illustrate the appearance of VG MC paths and explore the impact of parameter changes on these paths. For this reason, the plain MC method was implemented. The interested reader however can further adjust the algorithm presented with variance reduction techniques, discussed in Section 2.4.1. The algorithm used to simulate paths of the VG process was adapted from [24], the original shown below:

Input: VG parameters θ, σ, ν ; time spacing $\Delta t_1, \dots, \Delta t_N$ s.t. $\sum_{i=1}^N \Delta t_i = T$.

Initialization: Set $X_0 = 0$.

Loop from $i = 1$ to N :

1. Generate $\Delta G_i \sim \gamma\left(\frac{\Delta t_i}{\nu}, \nu\right)$, $Z_i \sim N(0, 1)$ independently and independent of past r.v.s.
2. Return $X_{t_i} = X_{t_{i-1}} + \theta \Delta G_i + \sigma \sqrt{\Delta G_i} Z_i$.

In the version adapted for this thesis, the time spacing was taken to be constant and instead of a for-loop, matrix operations were used for more efficient simulation. As such, the adapted algorithm is the following:

Input: VG parameters θ, σ, ν ; constant time spacing Δt s.t. $\Delta t = \frac{T}{N}$.

Initialization: Set $X_0 = 0$.

Build the VG path vector: with N elements for which the i -th element is:

1. Generate $\Delta G_i \sim \gamma\left(\frac{\Delta t}{\nu}, \nu\right)$, $Z_i \sim N(0, 1)$ independently and independent of past r.v.s.
2. Return $X_i = X_0 + \theta \sum_{j=1}^i \Delta G_j + \sigma \sum_{j=1}^i \left[\sqrt{\Delta G_j} Z_j \right]$.

Now to simulate N_p paths, the same process was followed with the exception that $\mathbf{X}_0, \mathbf{\Delta G}_i$ and \mathbf{Z}_i were column vectors of length N_p . An example of a single and multiple VG process paths can be found in Figure 8.

The next aspect to naturally explore was the effect of the three parameters to the shape of a path, see Figure 9. It was found that for smaller ν values there were more frequent, smaller jumps and for greater ν values less frequent larger jumps. Moreover, that a positive θ resulted to more frequent positive jumps and a negative θ to more frequent negative jumps, with the magnitude of jumps increasing with the magnitude of θ . Also θ had a direct effect on the slope of the path. Finally, it was evident that with increasing σ , increased the variance of the path, with the magnitude of σ having a respective effect to the magnitude of the jumps.

3.2 Understanding the VG process skewness and kurtosis

For the reference selection of parameters ($\nu = 0.17$, $\theta = 0.15$, $\sigma = 0.12$), the analytical mean, variance, skewness and kurtosis expressions were compared to MC ones for $N_p = 5000$, as a means of visually cross-referencing the results of Proposition 2.6, see Figure 10. In all cases, the MC moment metrics closely followed the analytical ones. The mean and variance are clearly

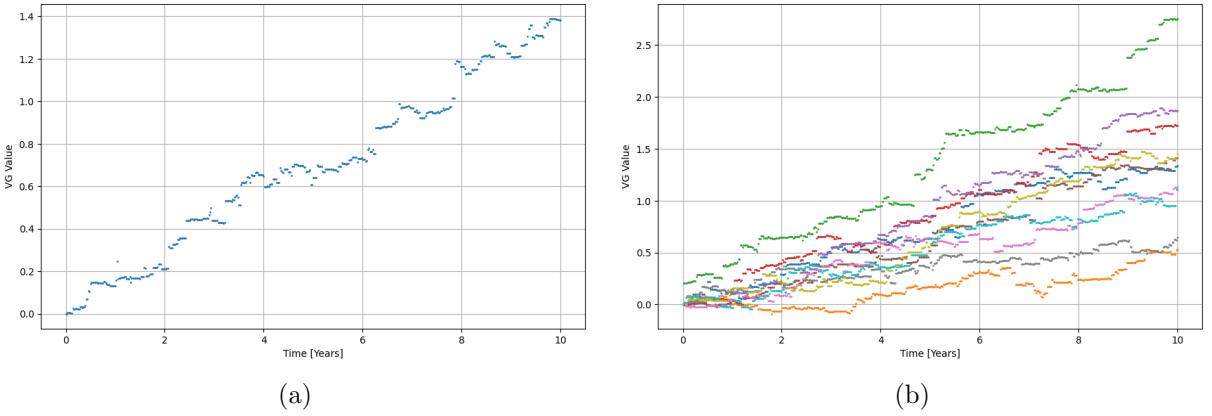


Figure 8: (a) One and (b) multiple VG paths for $\nu = 0.17$, $\theta = 0.15$, $\sigma = 0.12$ and a weekly time step for a ten year time range.

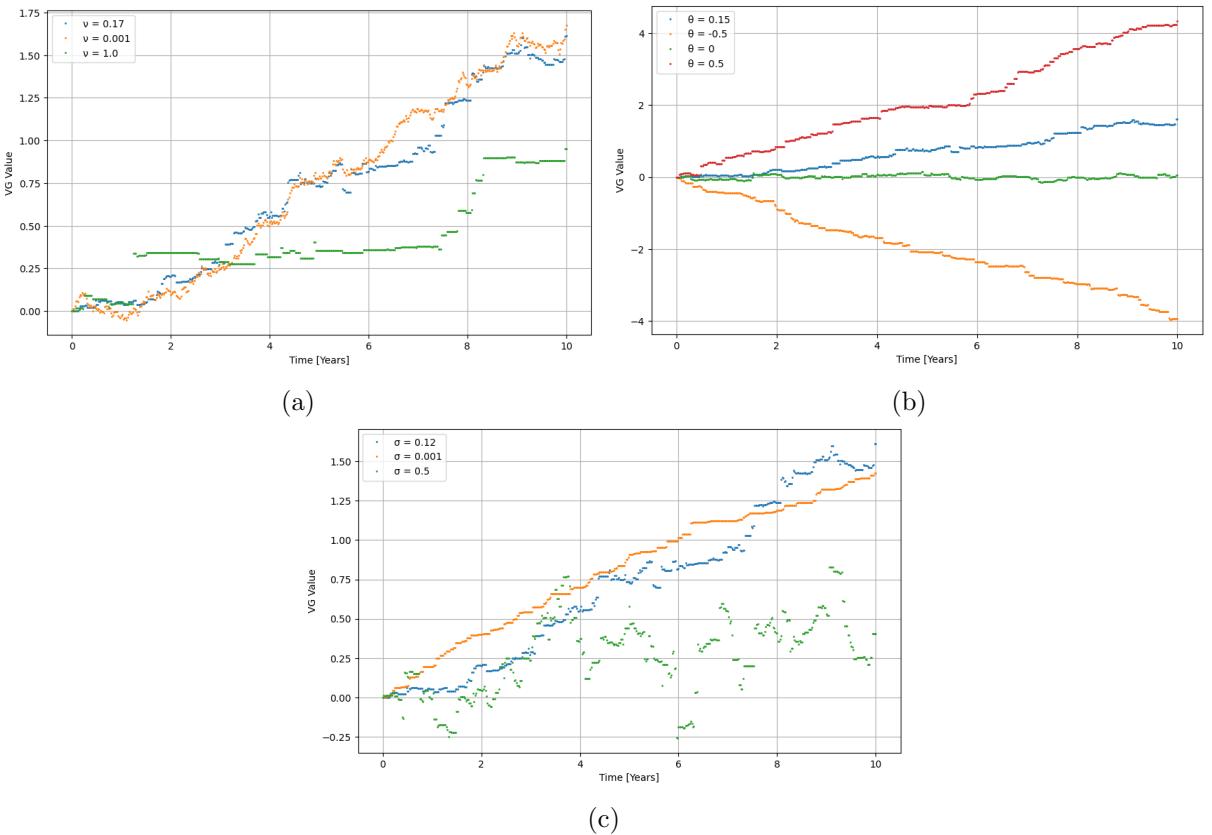


Figure 9: Variance Gamma process paths with varying (a) ν , (b) θ , and (c) σ parameters respectively, from the reference selection $\nu = 0.17$, $\theta = 0.15$, $\sigma = 0.12$.

increasing with time (the mean would be decreasing in time if $\theta < 0$) and the skewness and kurtosis are decreasing with time. It is noteworthy that while both skewness and kurtosis start from infinite values at $t=0$, they converge to steady state values, which for this selection of parameters were 0.17 for skewness and 3.07 for kurtosis. The convergence of skewness and kurtosis to these of the normal (0 and 3 respectively) with increasing time, is consistent to Remark 2.6.

Having studied the behavior of moment metrics vs time for a single set of parameters, skewness and kurtosis were subsequently mapped for a range of parameters, using the analytical skewness and kurtosis expressions. As time $t \in [0, 10]$, it only affects the scale of skewness and kurtosis, see Proposition 2.6. Time was therefore taken as $t = 1$ for the following parametric mapping, but the interested reader can scale the results by the desired time, by multiplying by

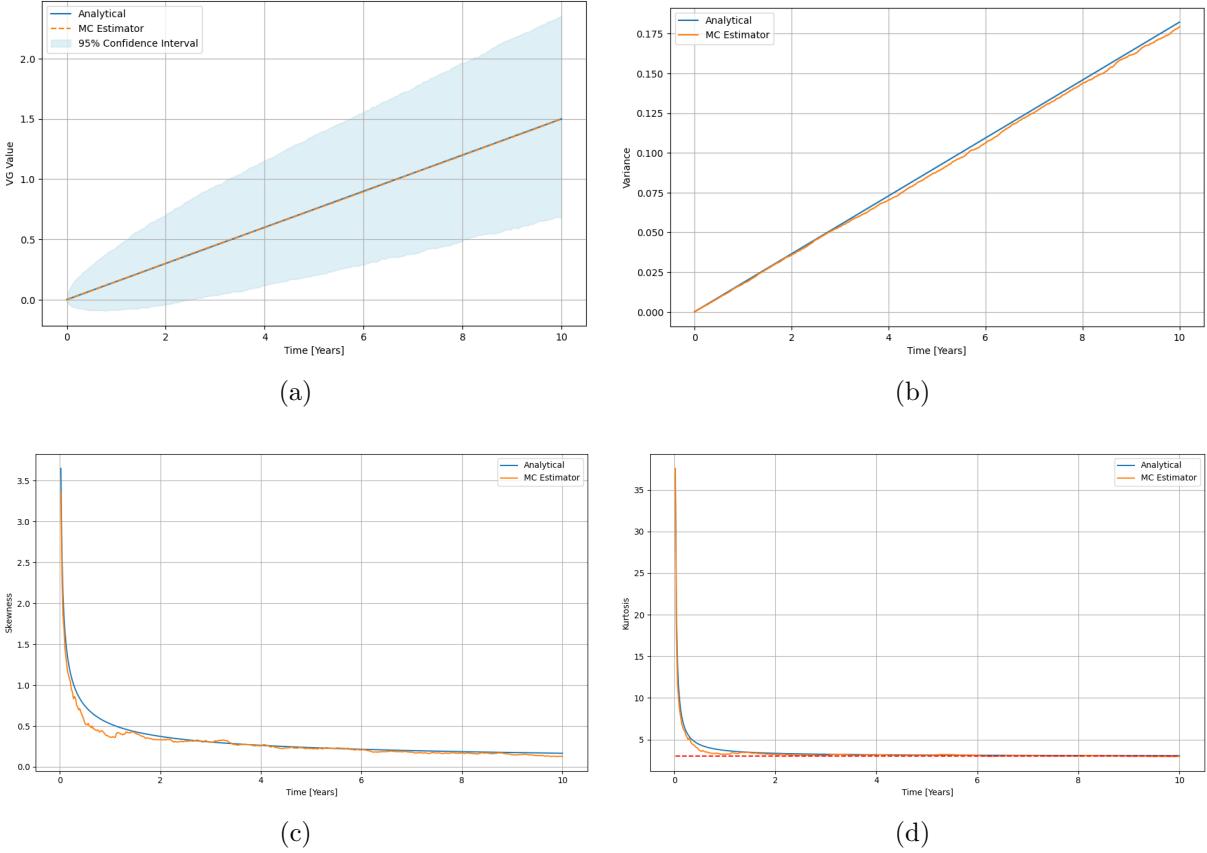


Figure 10: Monte Carlo estimations and analytical solutions of (a) mean (with 95% confidence interval), (b) variance, (c) skewness and (d) kurtosis, vs time.

$t^{-1/2}$ for skewness and t^{-1} for kurtosis. For $\theta \in [-2, 2]$, $\nu \in [0.001, 1]$ and $\sigma \in [0.1, 4]$, the plots of Figures 11 and 12 were produced, accompanied by the following insights.

Both skewness and kurtosis:

The volatility parameter σ affects the magnitude and smoothness in θ of skewness and kurtosis. For greater values of σ , skewness and kurtosis are smooth in theta, but as $\sigma \rightarrow 0$, skewness and kurtosis resemble the step and upside-down impact functions. Smaller values of σ lead to greater magnitudes of infimum and supremum for skewness and supremum for kurtosis.

Skewness:

- For $\theta < 0$, skewness is decreasing and convex in ν . For $\theta > 0$, skewness is increasing and concave in ν .
- For $\nu = 0$, skewness = 0, regardless θ .
- For greater volatilities σ , skewness is increasing in θ . For σ in the vicinity of 0, skewness becomes a step function in θ symmetric around 0, with the sign of θ determining the sign of skewness.

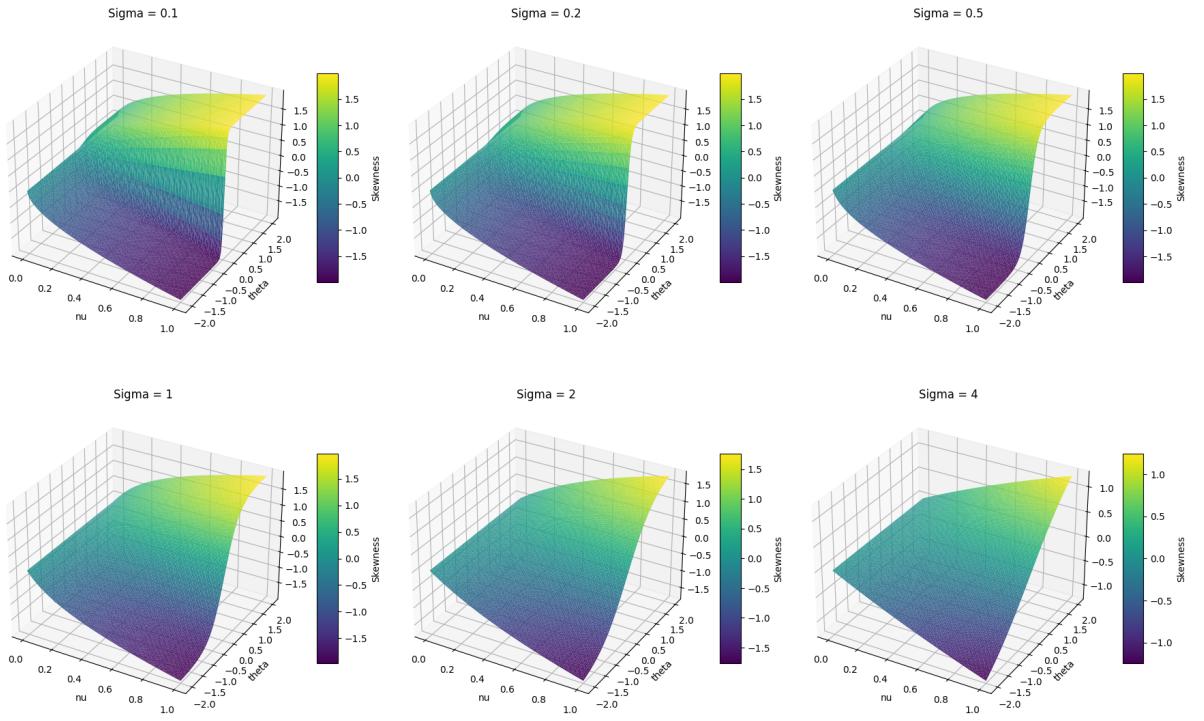


Figure 11: Skewness for varying VG parameters ν , θ , σ and $t = 1$.

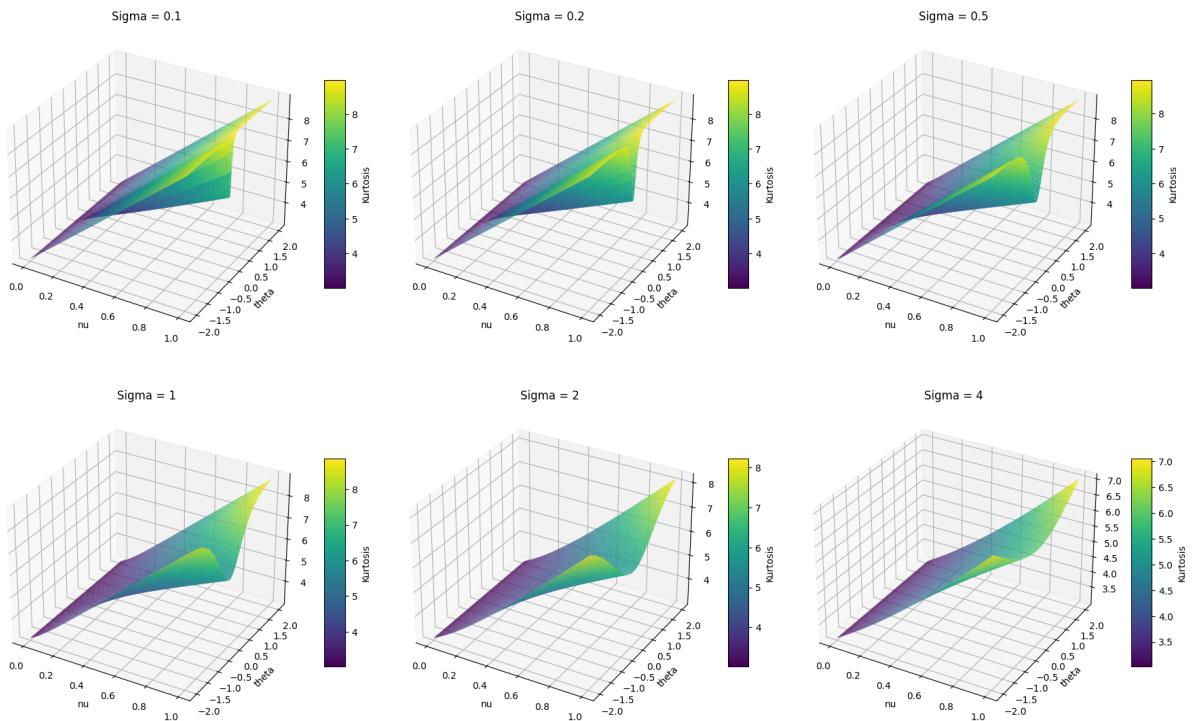


Figure 12: Kurtosis for varying VG parameters ν , θ , σ and $t = 1$.

Kurtosis:

- For $\theta < 0$, kurtosis is decreasing and concave in θ . For $\theta > 0$, kurtosis is increasing and concave in θ .
- Kurtosis is increasing in ν .
- For $\nu = 0$, kurtosis = 3, regardless θ .
- For greater volatilities σ , kurtosis has a parabolic shape in θ . For σ in the vicinity of 0, kurtosis looks like an upside-down impact function symmetric around $\theta = 0$.

Clearly, for the aforementioned range of parameters, it is deducted from the VG skewness and kurtosis ranges of the graphs (see respective colorbars of Figures 11 and 12 for $\sigma = 0.1$), that the VG skewness and kurtosis for varying time lays approximately in $[-2t^{-1/2}, 2t^{-1/2}]$ and $[3, 3 + 6t^{-1}]$ respectively. This observation provided motivation to study in more detail the achievable combinations of skewness and kurtosis for $t \in [\frac{1}{12 \times 4}, 10]$ years ($\frac{1}{12 \times 4}$ years = 1 week) and the parameter ranges above, see Figure 13. It is evident the achievable skewness and kurtosis are bounded by two parabolas characterized for the first week ($t = \frac{1}{12 \times 4} = 0.021$, area in blue) by the (skewness, kurtosis) pairs $(-13.86, 291)$, $(13.86, 291)$, $(0, 147)$ and $(0, 3)$, the surface of which declines with increasing time.

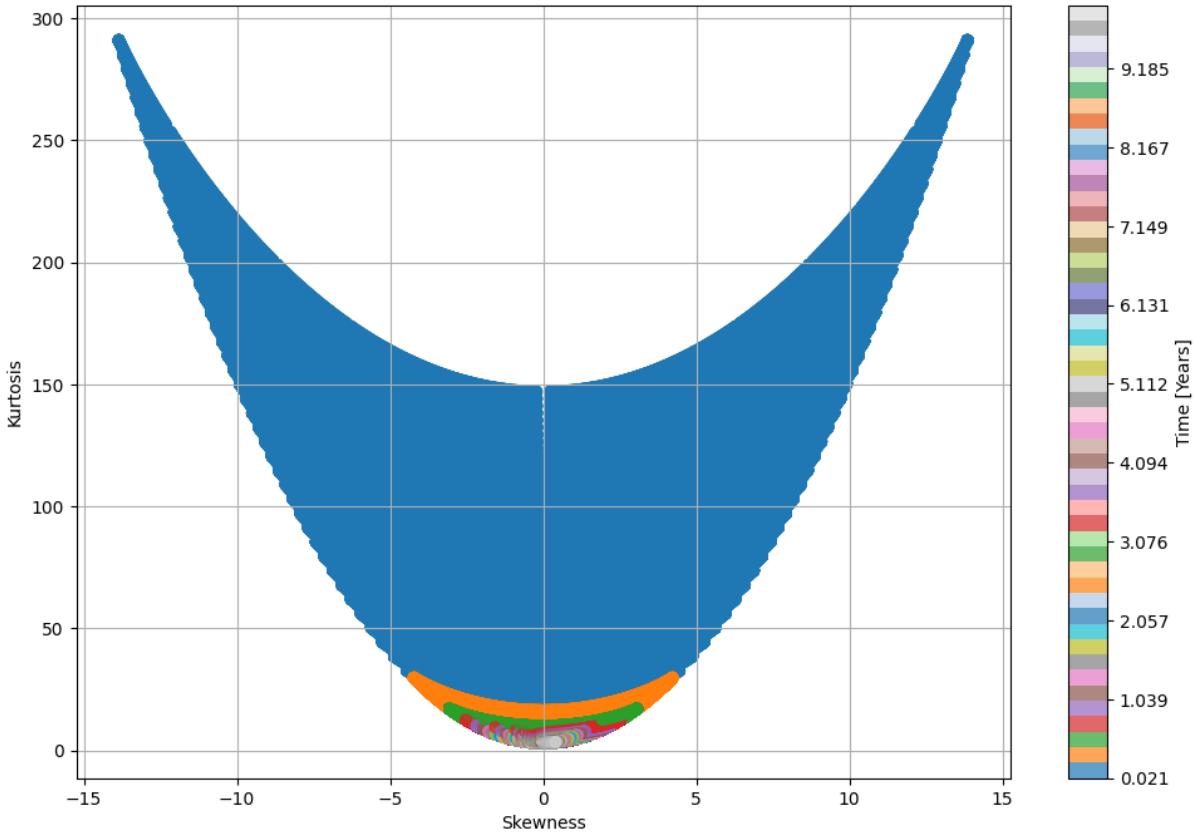


Figure 13: Achievable skewness and kurtosis combinations for $\theta \in [-2, 2]$, $\nu \in [0.001, 1]$, $\sigma \in [0.1, 4]$ and $t \in [\frac{1}{12 \times 4}, 10]$.

The figure provides the extent of the capacity of the VG model to capture different skewness and kurtosis combinations of the target data, across time. Interestingly, comparing Figure 13 with Figures 11 and 12 this leads us to the following observations:

- The extrema of skewness and suprema for kurtosis are present where $\sigma^2 \ll \theta^2\nu$ (notice the yellow areas of the surfaces of Figures 11 and 12 have greater values where σ^2 is small in comparison to θ and ν). By replacing this condition to the expressions of Proposition

2.6, the extrema of skewness and suprema for kurtosis are simplified to be dependent only on ν and t :

$$(\text{Sk})_{ex} = 2 \operatorname{sign}(\theta) \sqrt{\frac{\nu}{t}} \quad (3.1)$$

$$(\text{Ku})_{sup} = 3 + 6 \frac{\nu}{t}, \quad (3.2)$$

- For each ν the extrema of skewness and suprema of kurtosis form the lower boundary of the skewness vs kurtosis surface. Thus, by replacing equation (3.1) to (3.2) we can find the lower bounding parabola:

$$(\text{Ku})_{lb} = \frac{6}{4} (\text{Sk})^2 + 3 \quad (3.3)$$

- The minimum kurtosis is 3 for $(\text{Sk}) = 0$ ($\nu = 0$). The maximum kurtosis value for $(\text{Sk}) = 0$ ($\theta = 0$) is:

$$(\text{Ku}) = 3 + 3 \frac{\nu_{ub}}{t}, \quad (3.4)$$

where ν_{ub} is the upper bound of the parametric search for ν .

- The upper bounding parabola, is the upper bound of kurtosis for each given skewness, which happens for $\nu = \nu_{ub}$.
- Assuming for the upper bounding line a second order parabola which has the intercept of equation (3.4) and passes from the 2 symmetric points of equation (3.1) for $\nu = \nu_{ub}$:

$$(\text{Ku})_{ub} = \frac{3}{4} (\text{Sk})^2 + 3 + 3 \frac{\nu_{ub}}{t} \quad (3.5)$$

The boundaries of equations (3.3) and (3.5) were plotted against the skewness vs kurtosis surface, see Figure 14. It was observed that the lower bounding line of the area was exact and effectively bounded the area from below for all t . While the upper bound of the area also effectively bounded the surface from above for all t , it did not exactly coincide with $(\text{Ku})|_{\nu_{ub}}$, except at the minimum and the suprema of kurtosis.

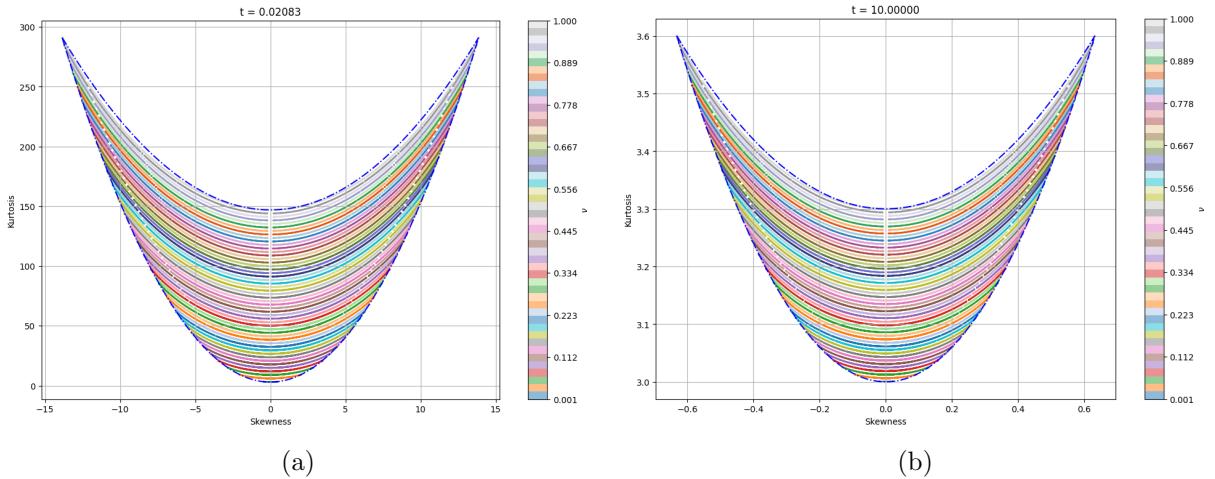


Figure 14: Achievable skewness vs kurtosis area for the VG process and the upper and lower boundaries of equations (3.1), (3.5) plotted in blue dash-dot line for (a) $t = \frac{1}{12 \times 4}$ years (1 week) and (b) $t = 10$ years. The colorbar was used to denote parabolic slices of increasing ν of the area from down upwards.

3.3 Effect of parameters to the VG implied volatility surface

This section focuses on the method used to produce the VG IV surface and the effect of each parameter's value to the shape of the VG IV surface. As discussed in Section 2.5, the IV surface of the VG option pricing model is produced by finding for each pair of moneyness and maturity (M, T) the volatility σ_{BS} for which the option price of the B-S model equates that of the VG model, $p_{BS} = p_{VG}$. The corrected analytical expression of Theorem 2.6 was utilized for the VG option prices. The method used to locate the VG IV surface is summarized as follows:

- **Optimizing parameter:** σ_{BS}
- **Parameter boundaries:** $\sigma_{BS} \in [0.01, 2]$
- **Inputs:** $S_0, \theta, \sigma, \nu, r, T \in [0.25, 9], M = \frac{S_0}{K} \in [0.6, 1.4]$
- **Objective function:** $p_{BS} - p_{VG}$
- **Algorithm:** Brent's root-finding method, which is based on the bisection, secant and inverse quadratic interpolation methods, as developed in Python's `scipy.optimize` library.
- **Output:** The implied volatility value $\sigma_{BS}^*(M, T)$ which solves $p_{BS} = p_{VG}$ for each pair of (M, T) .

Initially, the reference VG IV surface was produced by using the parameter combination found in [38], i.e. $(\nu, \theta, \sigma) = (0.1686, -0.1436, 0.1213)$, r was taken as 4.2% which was the yield on the 10-year U.S. treasury bond on the last trading day of March 2024 (28/03/2024) according to [18] and $S_0 = \$5248.49$ which was the closing price of S&P500 the same day, [53]. The resulting VG IV surface is shown in Figure 15. It demonstrates volatility smiles for short maturities and becomes flatter for longer maturities, with implied volatilities found in reasonable ranges.

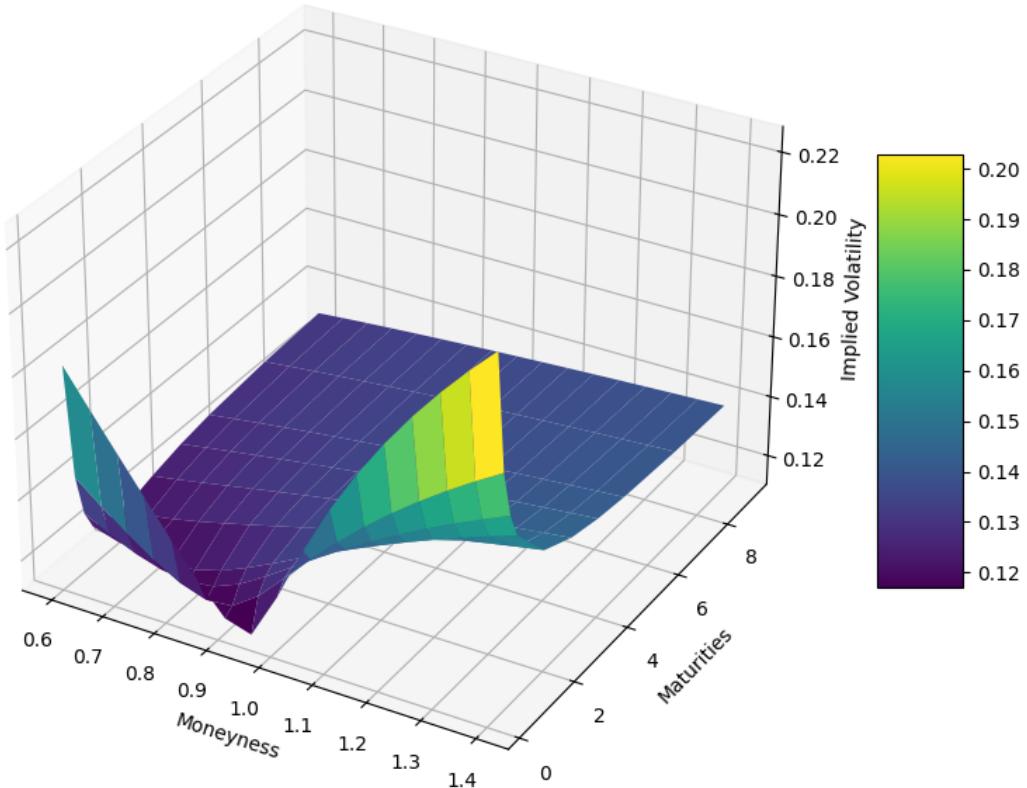


Figure 15: Reference Variance Gamma implied volatility surface.

Subsequently, the parameters were deviated one at a time in the following ranges $S_0 \in \{100, 10000\}$, $\theta \in \{0, 0.2\}$, $\sigma \in \{0.3, 0.5\}$, $\nu \in \{0.07, 1\}$, $r \in \{0, 0.15\}$. The shape of the VG IV did not change at all with deviating S_0 , which was reasonable as when solving the equation $p_{BS} = p_{VG}$, both sides can be divided by S_0 , in which case only the level of moneyness M matters. For the rest of parameters the various IV shapes where plotted in Figures 16 to 19.

By changing θ , for long maturities the slope of the surface changes in moneyness. A positive θ makes the surface decreasing in moneyness, a negative θ makes it increasing in moneyness, and $\theta = 0$ makes the surface flat in moneyness. For short maturities, it affects the location of the minimum of the smile. A positive θ places the location of the minimum in ITM moneyness, $\theta = 0$ places it to ATM areas and $\theta < 0$ places it to OTM areas. Furthermore, it is observed that with increasing absolute θ the overall IV is increased.

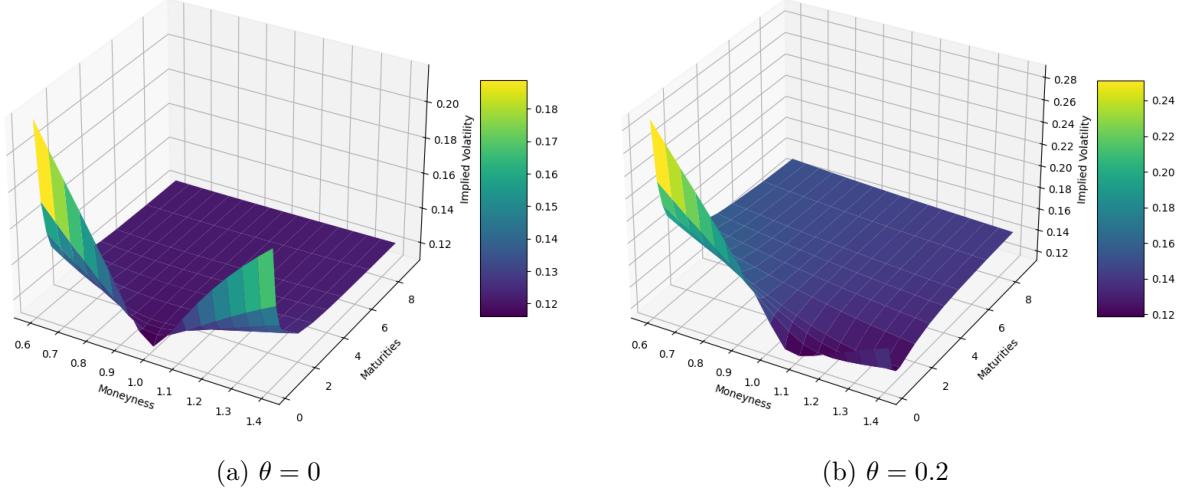


Figure 16: VG implied volatility surface for varying θ .

The volatility parameter σ seems to be responsible for the overall level of IV. By increasing σ to 0.3 and 0.5 it is observed that the IV surface ranges from 0.28 to 0.35 and 0.47 to 0.53 respectively. Interestingly, greater σ levels make the smile smoother/less steep at short maturities, (compare Figures 15 and 17), and the IV surface flat at long maturities.

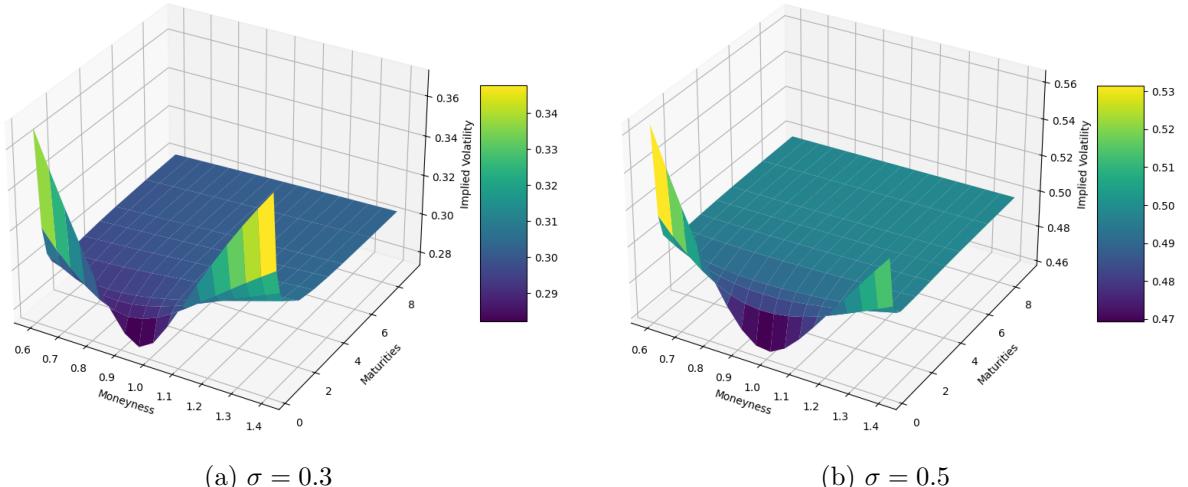


Figure 17: VG implied volatility surface for varying σ .

The kurtosis parameter ν , seems to be affecting the overall level of implied volatility, the steepness of the smile at short maturities and the strength of the slope at long maturities, see Figure 18. Taking ν close to 0, the surface flattens out for long maturities close to the value of σ . For short maturities the volatility smile is smoother. For greater ν values, the overall level of implied volatility increases and particularly at short maturities the smile becomes steeper.

Furthermore, for greater ν , the slope of IV for longer maturities becomes stronger, which becomes evident by comparing Figures 18b and 18a.

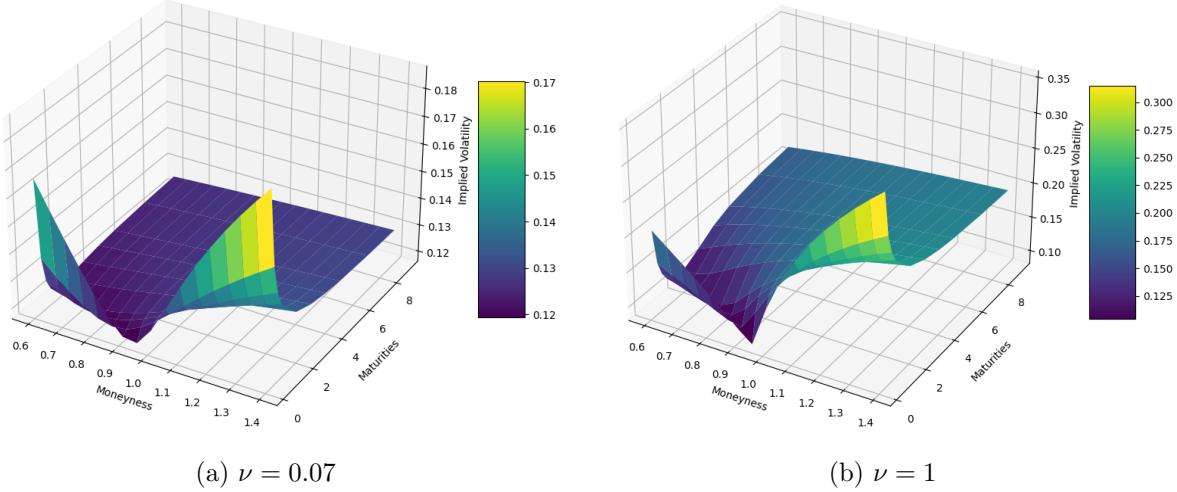


Figure 18: VG implied volatility surface for varying ν .

Finally, the interest rate r affects overall the implied volatility surface, see Figure 19. A large positive r in comparison to $r = 0$, results to an IV surface with higher values both in short and long maturities, with a stronger impact at long maturities.

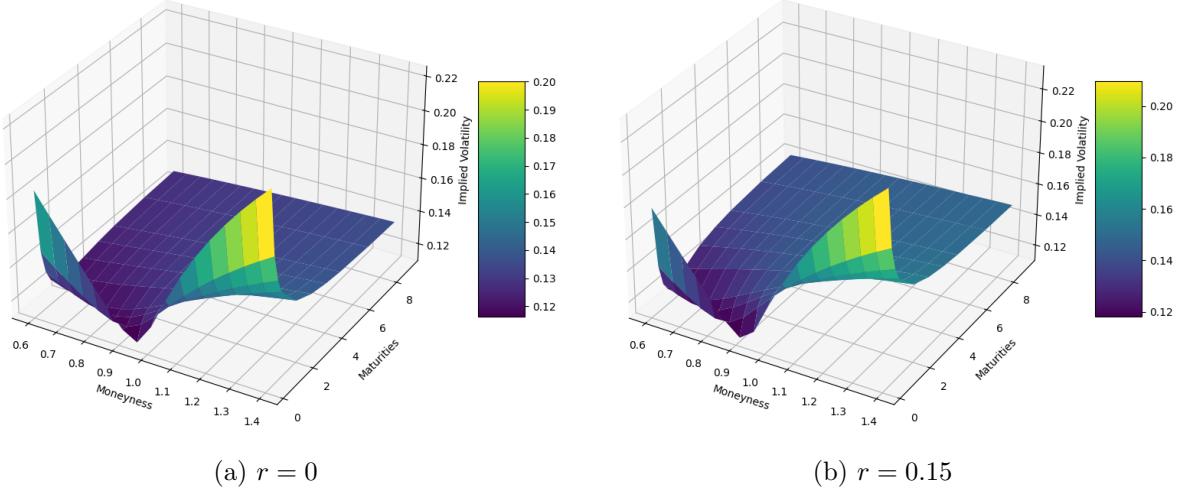


Figure 19: VG implied volatility surface for varying r .

Apart from the parametric evaluation of the VG IV surface, another interesting takeaway of this section is that certain terms of function Ψ of Theorem 2.6 are pushed to infinity for $\nu < 0.0657$ for larger maturities and ITM areas, with the rest of the parameters kept to the reference level discussed in the second paragraph of this section. In more detail, the term $\Gamma(x_3 = T/\nu) \rightarrow \infty$ found in the denominators of Ψ , the term $\Gamma(y_1 = T/\nu) \rightarrow \infty$ found in the denominator of Φ , the power term $c^{x_3+0.5} \rightarrow \infty$ and the exponential term $\exp(\text{sign}(x_1)c) \rightarrow 0$, found in the numerator of Ψ . By changing the sign of θ , the problematic area becomes the one for large maturities and OTM moneyness values.

This observation provided motivation for crude exploration of boundaries for the parameters for increasing ν , by trial and error changes of one parameter at a time. For the boundary value of $\nu = 0.0657$, the option pricing formula moves away from the singularity for $r \in [0, 0.042]$, $\sigma \in [0.1208, 1.9]$, $\theta \in [-0.1445, 0.351]$. By raising the value of ν to 0.13, the parameters ranges which ensure non-singularity are $r \in [0, 0.21]$, $\sigma \in [0.036, 1.8]$, $\theta \in [-0.60, 0.69]$. Further increasing ν to 0.26, the acceptable parameter bounds are $r \in [0, 0.21]$, $\sigma \in [0.026, 1.7]$, $\theta \in [-0.87, 0.69]$. The above crude boundaries were plotted in Figure 20 for easy reference. While

they are not very precise, they provide a general understanding of the achievable parameter ranges for the analytical option pricing formula of Theorem 2.6, particularly that for small ν , there are restrictions in the selection of r and θ , which are relaxed as ν increases.

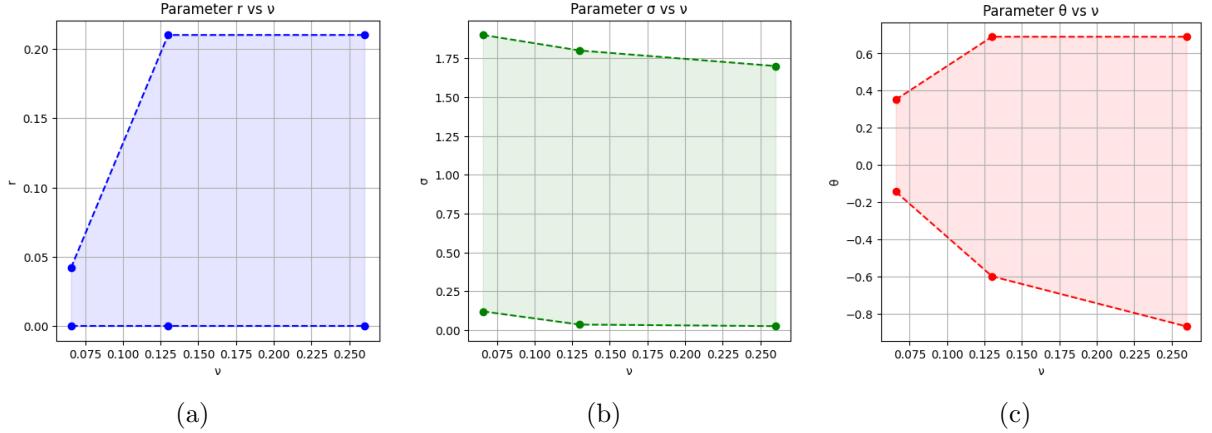


Figure 20: Crude boundaries of parameters (a) r , (b) σ and (c) θ with increasing ν , while keeping the rest of the parameters constant.

Finally, it is interesting to notice that for small values of ν , the acceptable ranges of the parameters for which the analytical option pricing formula stays away from singularities, are tighter than the constraint incurred from the condition of Remark 2.7. For example, for $\nu = 0.0657$ and $\sigma \leq 1.9$, from the condition of Remark 2.7, we need $\theta < 13.41$, which holds as $\theta \in [-0.60, 0.69]$. The same is evident for $\nu = 0.26$, where to stay away from singularities we need $\sigma \leq 1.7$, which through remark leads to $\theta < 2.4$, again satisfied as $\theta \in [-0.87, 0.69]$. For larger values of ν however the constraint of Remark 2.7 becomes important and will be therefore included in the calibration process of the next chapter.

3.4 Calibration of the VG option pricing model

In this section, it is investigated how well the corrected four parameter VG option pricing model of Theorem 2.6 can capture the IV surface produced by the nine parameter SVJD model, calibrated to S&P 500 data up to the end of March 2024 (refer to Figure 21). The surface demonstrates smiles for short maturities and becomes decreasing in moneyness for longer maturities.

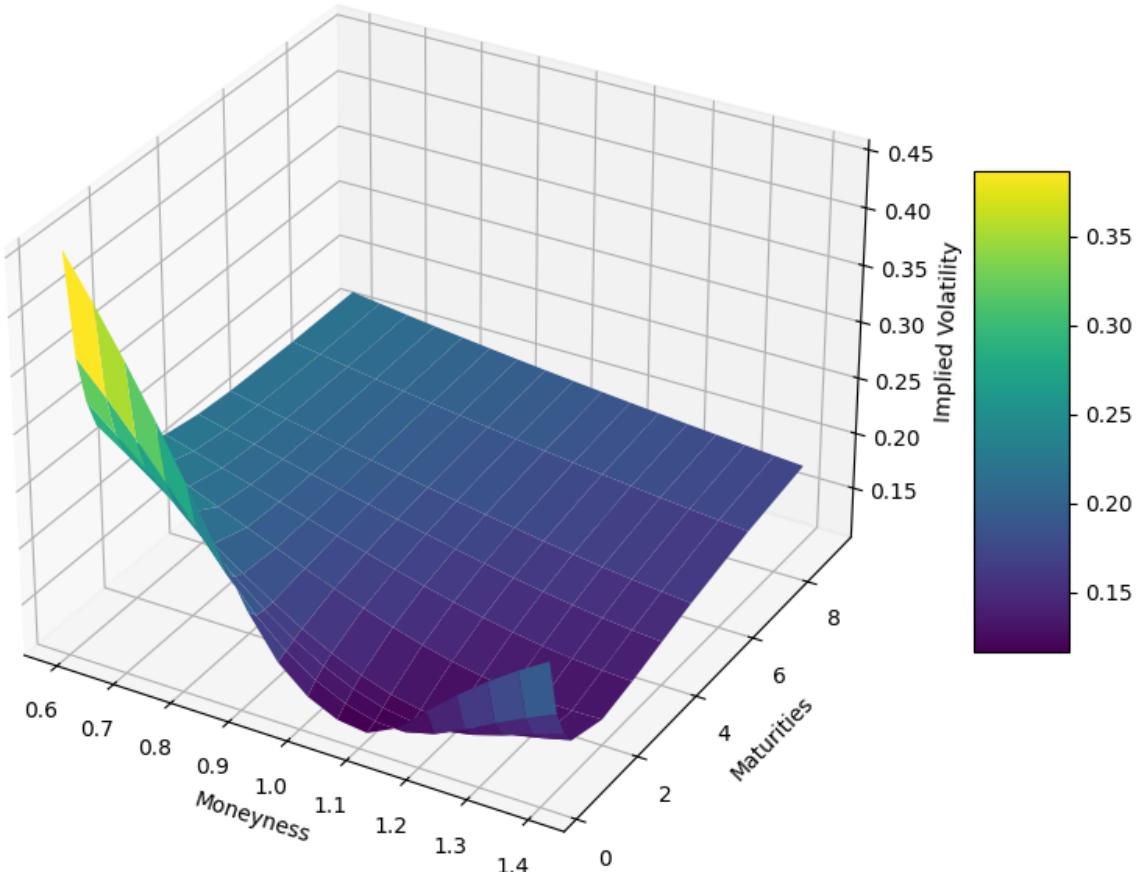


Figure 21: Implied volatility surface of S&P500, as captured by the SVJD model, based on data up to 28/03/2024.

For the calibration of the VG process, two optimization levels were implemented, an external and an internal. The external optimization level sought to find the parameter combination \mathbf{q} to minimize the RMSE of the VG IV surface from the target. The internal optimization level sought, for a given combination of moneyness M , maturity T and parameters \mathbf{q} , to find the implied volatility σ_{BS} for which the B-S option price is equal to the VG option price, $p_{BS} = p_{VG}$. So it essentially locates the VG IV value for the parameter combination given. There were also two different approaches regarding the risk free rate, r . It was taken either as a constant parameter to be optimized (see Table 1) or as a given varying parameter with maturity, (refer to Table 2 and Figure 23). Additional details on the algorithm are provided below:

External optimization:

- **Set of optimizing parameters:** $\mathbf{q} = [r, \theta, \sigma, \nu]$ for r constant across maturity T or $\mathbf{q} = [\theta, \sigma, \nu]$ for $r = r(T)$ in line with Figure 23.
- **Parameter boundaries:** See Tables 1 and 2

- **Inputs:** $S_0 = 5639.17$, $T \in [0.25, 9]$, $M = \frac{S_0}{K} \in [0.6, 1.4]$
- **Objective function:** $RMSE = \frac{1}{N_{IV}} \sqrt{\sum_{T, M} (IV_{target}(M, T) - IV_{VG, q}(M, T))^2}$,
where N_{IV} is the number of distinct moneyness-maturity (M, T) pairs, $IV_{target}(M, T)$ is the given IV surface target value for each pair and $IV_{VG, q}(M, T)$ is the VG IV value for the parameter combination q at each (M, T) pair.
- **Constraint:** $1 - \theta\nu - \frac{\nu\sigma^2}{2} > 0$
- **Initialization:** See Tables 1 and 2
- **Algorithm:** Limited-memory Broyden-Fletcher-Goldfarb-Shanno with Box constraints (L-BFGS-B), using the minimize() function from Python's scipy.optimize library.
- **Output:** The IV_{VG, q^*} surface that most closely matches IV_{target} based on RMSE, and the corresponding q^* .

Internal optimization:

- **Optimizing parameter:** σ_{BS}
- **Parameter boundaries:** $\sigma_{BS} \in [0.01, 2]$
- **Inputs:** $S_0 = 5639.17$, $T \in [0.25, 9]$, $M = \frac{S_0}{K} \in [0.6, 1.4]$, q
- **Objective function:** $p_{BS} - p_{VG}$
- **Algorithm:** Brent's root-finding method, which is based on the bisection, secant and inverse quadratic interpolation methods, as developed in Python's scipy.optimize library.
- **Output:** The implied volatility value σ_{BS}^* which solves $p_{BS} = p_{VG}$.

Since the L-BFGS-B algorithm is not a global optimization algorithm, different initial values and boundaries were taken that led to different optimal results. Taking into account the outcome of the parametric analysis of Section 3.3, the general setup of boundaries of the parameters r , θ , σ , ν was taken as $[0, 0.21]$, $[-2, 2]$, $[0.01, 0.4]$, and $[0.06, 4]$ respectively. Among the only exceptions was row #1 of Table 1, where the r boundaries were taken symmetrically around 0 to figure out where the parameter would naturally be led by the optimization process. Moreover, in row #6 of the same table, the parameter r was localized in $[0.01, 0.055]$, to understand the effect of localizing boundaries on converging to an optimal r value.

In general it was found that the minimum $RMSE = 0.009914$ was achieved for $r = -8.25\%$, a risk free rate which while offered the best fit (see Figure 22a) has never been observed in the US economy. Moreover, that generally smaller risk free rates led to better fits. And in addition that the increasing IV error across maturity (see Figure 22b), might mean that the risk free rate is not constant but increases for longer maturities. This led to the second round of optimizations (see next paragraph) where r was a function of T . The rest of the optimal parameters ranged in reasonable regions with regards to the parametric analysis of Section 3.3 and the smile/curvature characteristics of the target of Figure 21, discussed earlier in this chapter. These were in the range of 0.0394 to 0.1348 for θ , 0.1100 to 0.1357 for σ and 0.32 to 2.45 for ν , see Table 1.

In the subsequent approach r was taken as the yield of the US zero coupon bond (ZCB) for increasing maturities on the last trading day of March 2024 (28/03/2024), the date of the data based on which the SVJD IV surface was developed. The ZCB yields vs maturity (1 to 9 years) were collected from the database of the Federal Reserve Bank of St. Louis, [17]. For the 3 and 6 month rates, the yield of US treasury securities was considered, [19]. The 9 month risk free rate was found by linearly interpolating the neighbouring 6 months and 1 year values. The resulting risk free rate curve vs time $r = r(T)$ is shown in Figure 23.

#	Initial Guess [q_0]	Optimal Parameters [q^*]	RMSE
1	[0.01, 0.1436, 0.1213, 0.1681]	[-0.084587, 0.038248, 0.130129, 2.515679]	0.009912
2	[0.009, 0.01, 0.12, 2]	[0, 0.078411, 0.13436, 1.533644]	0.0112781
3	[0.05, 0.14, 0.14, 0.14]	[0, 0.078627, 0.134479, 1.521556]	0.0112782
4	[0.01, 0.1436, 0.1213, 0.1686]	[0, 0.079529, 0.134271, 1.503393]	0.011280
5	[0.007, 0.1, 0.1, 0.1]	[0, 0.081506, 0.134163, 1.454608]	0.011302
6	[0.035, 0.14, 0.14, 0.14]	[0.01, 0.083695, 0.134805, 1.455853]	0.011535
7	[0.042, 0.15, 0.15, 0.30]	[0.002903, 0.090494, 0.130857, 1.233863]	0.012092
8	[0.042, 0.01, 0.12, 0.80]	[0.013616, 0.100198, 0.135733, 1.019744]	0.012794
9	[0.042, -0.15, 0.13, 0.30]	[0.147523, 0.239875, 0.110069, 0.493113]	0.019772
10	[0.035, 0, 0.14, 0.14]	[0.122776, 0.271993, 0.118241, 0.328246]	0.021954

Table 1: Optimal parameters $q^* = [r, \theta, \sigma, \nu]^*$ sorted by increasing RMSE and the corresponding initial guesses. The parameter boundaries were $[0, 0.21]$, $[-2, 2]$, $[0.01, 0.4]$, and $[0.06, 4]$, with the exception of #1, and #6, where the boundaries for r were $[-0.1, 0.1]$ and $[0.01, 0.055]$ respectively.

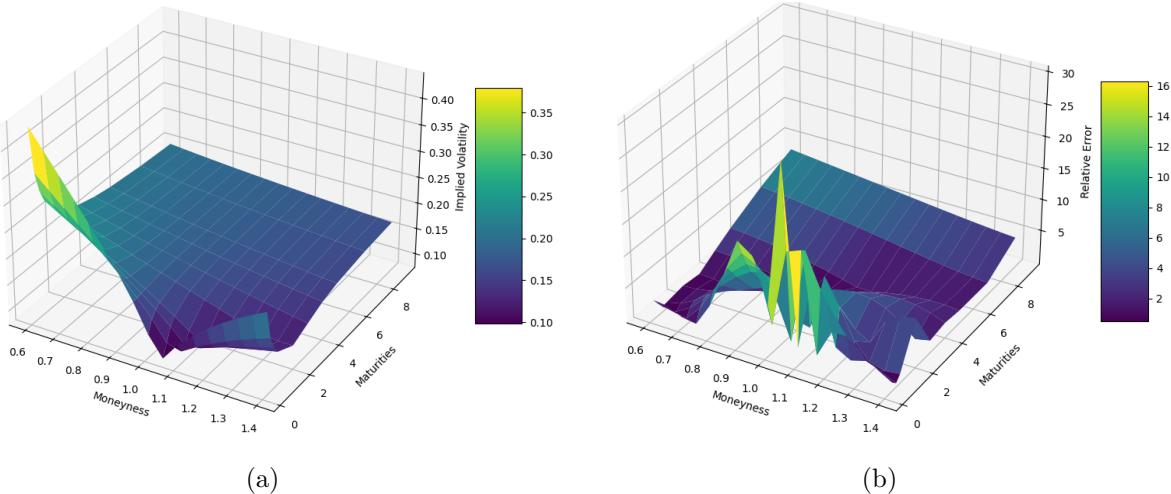


Figure 22: (a) Best overall calibrated VG IV surface and (b) absolute relative error in comparison to target.

The boundaries of the remaining parameters were taken in the range of $[-0.5, 0.5]$, $[0.01, 0.4]$, and $[0.06, 4]$. The optimal parameters' set located for this approach leading to the minimum RMSE of 0.01169609 is shown in row #1 of Table 2. The optimal θ was positive (0.104361) implying positive skewness, σ was as expected a little bit below the minimum IV value of the target for large maturities (0.134901) and the level of ν (1.216148) implied thick tails at short maturities. The optimal IV surface in this case and associated absolute relative error surface are found in Figure 24.

To validate the optima found via local optimization methods with different starting points were global optima, the same algorithm was applied with global optimization methods. Two global optimization methods were hence implemented in the external optimization level, the particle swarm optimization method and the global differential evolution method, see “pso” of the pyswarm library and “differential_evolution” of the scipy.optimize library of the Python suite respectively.

The particle swarm optimization method (PSO) is inspired by the social behavior of birds flocking or fish schooling and is commonly used in various optimization problems. A group of candidate solutions (called particles) is initialized randomly in the search space. Each particle has a position and a velocity. Each particle's velocity is updated based on its own experience, the experience of its neighbors, or the global experience. The particle's position is then updated

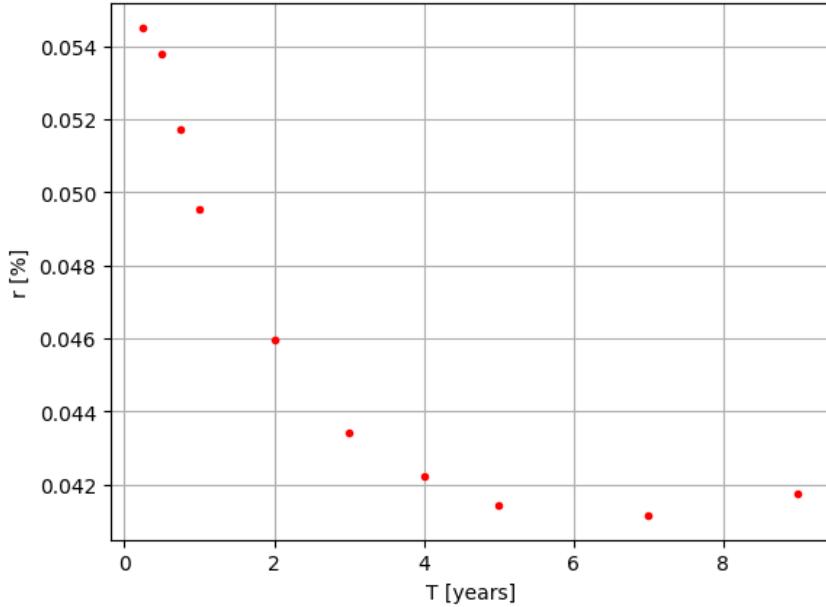


Figure 23: Risk free rate vs maturity in 28/03/2024.

#	Initial Guess $[q_0]$	Optimal Parameters $[q^*]$	RMSE
1	[0.104365, 0.134885, 1.216146]	[0.104361, 0.134901, 1.216148]	0.01169609
2	[-0.12, 0.12, 0.17]	[0.104365, 0.134885, 1.216146]	0.01169611
3	[0.01, 0.12, 0.17]	[0.104679, 0.135004, 1.209195]	0.01169645
4	[0.1045, 0.1349, 1.2031]	[0.104989, 0.1349, 1.203078]	0.01169683
5	[0.24, 0.135, 1]	[0.10462, 0.135471, 1.197062]	0.011701
6	[0.12, 0.12, 0.17]	[0.106005, 0.13488, 1.182293]	0.011703
7	[0.24, 0.135, 1.6]	[0.102715, 0.13504, 1.247452]	0.011705
8	[-0.24, 0.15, 0.06]	[0.106447, 0.134948, 1.175078]	0.011708
9	[-0.01, 0.12, 0.17]	[0.108507, 0.134743, 1.134665]	0.011744
10	[0.15, 0.09, 1.2]	[0.1277630.0872631.295791]	0.025221

Table 2: Optimal parameters $q^* = [\theta, \sigma, \nu]^*$ by increasing RMSE and the corresponding initial guesses for $r = r(T)$ according to Figure 23. The parameter boundaries were [-2, 2], [0.01, 0.4], and [0.06, 4] respectively for θ , σ , ν .

using the new velocity.

The global Differential Evolution (DE) is a population-based stochastic optimization algorithm, inspired by the concept of natural selection and evolution. DE works with a population of candidate solutions (called individuals), which are vectors representing potential solutions in the search space. In each iteration, candidate solutions might be mutated and are subsequently crossovered. During mutation, DE creates a new candidate solution by adding a weighted difference between two population vectors to a third vector. During the crossover phase, candidate solutions in the population are combined to create a trial vector. This crossover process introduces diversity into the population. At the end of each iteration, the trial vector is evaluated against the objective function. If it provides a better solution (lower cost for a minimization problem), it replaces the corresponding vector in the population present in the previous iteration. Otherwise, the original vector is retained.

As it is evident in Table 3, by considering r a constant optimization parameter, the optimal parameters/RMSEs for the DE and the PSO methods were very close to #1 of Table 1, for $r \in [-0.1, 0.1]$ and #2 for $r \in [0, 0.21]$. It was again evident that smaller values of r led to better RMSEs. On the other hand, by taking $r = r(T)$ in line with Figure 23, the optimal solutions for both the PSO and the DE methods were very close to row #1 of Table 2.

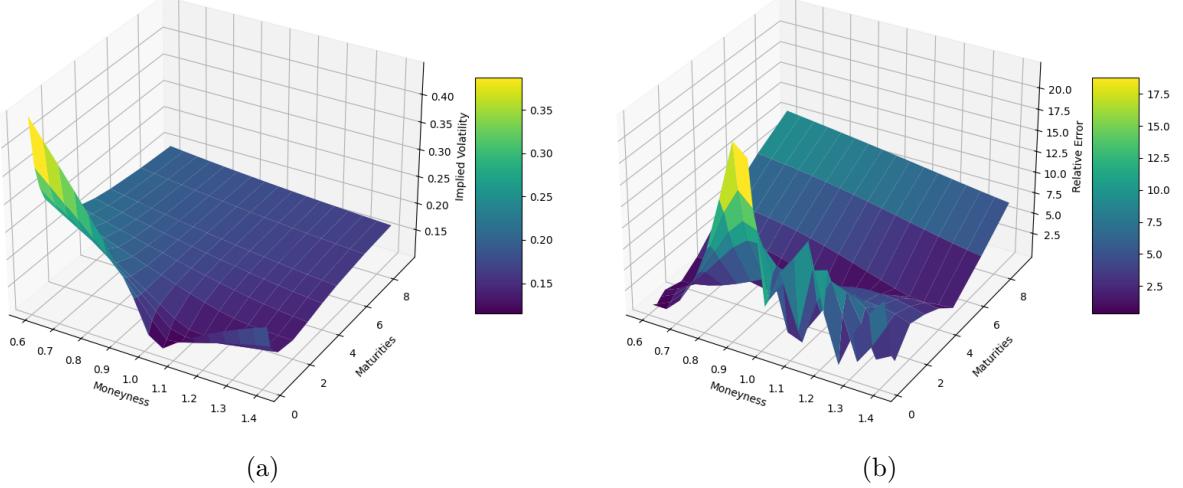


Figure 24: (a) Best calibrated VG IV surface for $r = r(T)$ according to Figure 23 and corresponding (b) absolute relative error in comparison to target.

Method	r	Optimal Parameters $[q^*]$	RMSE
DE	[-0.1, 0.1]	[-0.086426, 0.037261, 0.129981, 2.56453]	0.009914
PSO	[-0.1, 0.1]	[-0.080954, 0.042169, 0.131348, 2.267901]	0.009974
DE	[0, 0.21]	[0, 0.078613, 0.134357, 1.527392]	0.0112779
PSO	[0, 0.21]	[0, 0.079538, 0.134199, 1.502293]	0.011280
PSO	$r(T)$	[0.104397, 0.13494, 1.214359]	0.01169606
DE	$r(T)$	[0.104533, 0.134935, 1.211672]	0.01169607

Table 3: Optimal parameters $q^* = [r, \theta, \sigma, \nu]^*$ or $q^* = [\theta, \sigma, \nu]^*$ for $r = r(T)$ according to Figure 23, sorted by increasing RMSE. The parameter boundaries for θ , σ and ν were $[-2, 2]$, $[0.01, 0.4]$, and $[0.06, 4]$.

Having validated the optimal results, the optimal VG IV surfaces were compared with the target SVJD IV. First of all, optically comparing the target of Figure 21 with the optimal VG IV surfaces of Figures 22 and 24, it was observed that the overall IV ranges match (approx. 0.10 to 0.38). However, for large maturities (e.g. $T = 9$ years) the VG IV surface could not capture very accurately the SVJD IV surface. Furthermore, for very short maturities (e.g. $T = 0.25$ years), the VG IV surface had a more steep smile than the SVJD IV surface. The same was reflected to the absolute relevant error, which peaked for very short maturities ($\simeq 20\%$) and was increased for very long maturities ($\simeq 10\%$), but remained in lower levels for mid range maturities. It was thus concluded that while the VG model is capable of replicating well overall the IV surface, it has limitations to the level of complexity it can capture in very short or very long maturities within the studied ranges $T \in [0.25, 9]$, in comparison to the SVJD model.

These limitations may be present due to the fact that the parameters of the VG model determining the steepness of the smile in short maturities are not disengaged with the overall IV level and the steepness of the slope in long maturities. The SVJD S&P500 IV had a smooth smile in short maturities and a steep slope for long maturities. For the VG model, the parameters that determine the steepness of the smile in short maturities are σ and ν . As seen in Section 3.3, increasing σ , makes the smile less steep. However, increasing σ increases also the overall level of IV. And in order to capture subject SVJD S&P500 IV target, there would be no meaning in increasing σ more than $\simeq 0.15$. On the other hand, a smaller ν makes the smile smoother for short maturities, but it concurrently flattens the IV surfaces slope for long maturities. For this reason, the optimization algorithm preferred to capture the slope of the target in long maturities by converging to a high ν , in the expense of a more steep smile for short maturities. In summary,

the VG process might be better in capturing IV surfaces, with steep smiles short term and strong slopes long term or smooth smiles short term and flatter slopes long term or with higher overall IV.

An additional justification would be due to the fact that Moody's Analytics might be using a more complex risk free rate model than the models used here: (a) r optimized but constant across maturity and (b) $r = r(T)$ from online sources. A risk free rate model closer to the one Moody's Analytics used for the calibration of the SVJD IV surface might have resulted in better calibration results across maturity.

4 Conclusions

The VG process is considered to be a powerful process used in option pricing models, due to its capability to account for the skewness and kurtosis of the underlying asset's returns' distribution, especially in its volatility regime change variant, its stochastic volatility variants and its multivariate variant, where implied volatility surfaces and correlations can be captured better.

In this thesis, the mathematical theory behind the simple form VG process was presented, its skewness and kurtosis were explored and bounded and the capabilities of its analytical option pricing model for plain EU options was tested against a benchmark model. Its main contributions were (1) correcting the analytical VG option pricing formula, (2) offering a clear understanding of the effect of: (a) the VG process' parameters (θ, σ, ν) to its skewness and kurtosis, and (b) the VG option pricing model's parameters (r, θ, σ, ν) to the implied volatility surface and (3) understanding its capabilities and limitations in comparison to a benchmark model, i.e. SVJD. Among the key takeaways were:

- The analytical VG option pricing formula of [38] was corrected, replacing the incorrect term “ $as\sqrt{\frac{\nu}{1-c_2}}$ ” with “ $a\sqrt{\frac{\nu}{1-c_2}}$ ”, and the incorrect parameter $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$ with the correct $a = +\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$. It was shown that if incorrectly $a = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\nu\theta^2}{2}}}$, then the discounted stock price is not a martingale and hence the VG option pricing model admits arbitrage.
- The expectation of the VG log returns under the risk free measure is linear with time, with its slope being the sum of $\theta + r + \omega$, see Remark 2.8.
- The variance of the VG log returns is linear with time with slope equal to the sum of $\theta^2\nu + \sigma^2$. The variance has a linear dependence with regards to ν and square dependence with regards to θ, σ .
- The skewness is proportional to the inverse of the square root of time and has the sign of θ . As $t \rightarrow \infty$ it converges to 0.
- The kurtosis is proportional to the inverse of time and has intercept 3. As $t \rightarrow \infty$ it converges to 3.
- The area of achievable pairs of (skewness, kurtosis) at each time step for the VG log returns can be bounded by two parabolas, see equations (3.3) and (3.5).
- The parameter θ affects the direction of the slope of the VG IV surface across moneyness for long maturities, and the location of the minimum of the smile for short maturities.
- The parameter σ affects the overall level of the VG IV surface, with its value being close to the minimum IV VG value. Larger σ values make the IV surface concave for ITM areas and offer smoother smiles for short maturities.
- The parameter ν affects the overall level of IV, the intensity of the slope of the IV surface across moneyness for longer maturities and the steepness of the smile for shorter maturities.
- The parameter r affects the slope of the IV surface across maturity and thus the final value of IV in time.
- As ν decreases below 0.06, the analytical VG EU option pricing formula approaches a singularity, which means that certain terms of it go to infinity. As ν increases, the range of acceptable r and θ widens.
- The VG IV surface can overall capture the IV ranges of the SVJD S&P500 IV, but has difficulty capturing the complexity of surface simultaneously for short maturities ($T = 0.25$) and longer maturities ($T = 9$). This happens because the VG model is not capable of both having a shallow smile for short maturities and a stronger slope for long maturities.

In view of the above conclusions, it would be interesting to explore in future work how the limitations of the VG process could be improved by:

- Using the variant of the option pricing formula utilizing the characteristic function with the fast or fractional Fourier transformation methods (FFT or FrFT).
- Using a VG model, where σ is a deterministic function of moneyness and maturity, $\sigma = \sigma(M, T)$.
- Using a VG regime change model or VG stochastic volatility model.
- Using a multivariate VG process model to calibrate various assets, and investigating in addition its correlation and dependence structure.

References

- [1] J.-P. Aguilar. Some pricing tools for the Variance Gamma model. *International Journal of Theoretical and Applied Finance*, 23(5):35, 2020.
- [2] A. Almendral and C. W. Oosterlee. On American options under the Variance Gamma process. *Applied Mathematical Finance*, 14(2):131–152, 2007.
- [3] D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2009.
- [4] A. N. Avramidis, P. L’Ecuyer, and P.-A. Tremblay. Efficient simulation of Gamma and Variance Gamma processes. In *Proceedings of the 2003 Winter Simulation Conference, 2003.*, volume 1, pages 319–326, 2003.
- [5] G. Bakshi, N. Kapadia, and D. Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1):101–143, 2003.
- [6] D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutschemark options. *Review of Financial Studies*, 9(1):69–107, 1996.
- [7] M. Becker. Unbiased Monte Carlo valuation of lookback, swing and barrier options with continuous monitoring under Variance Gamma models. *The Journal of Computational Finance*, 13(4):35–61, 2010.
- [8] B. Buchmann, B. Kaehler, R. Maller, and A. Szimayer. Multivariate subordination using generalised Gamma convolutions with applications to Variance Gamma processes and option pricing. *Stochastic Processes and their Applications*, 127(7):2208–2242, 2017.
- [9] L. Cao and M. C. Fu. Estimating Greeks for Variance-Gamma. In *Proceedings of the 2010 Winter Simulation Conference*, pages 2620–2628. IEEE, 2010.
- [10] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *The Journal of Business*, 75(2):305–332, 2002.
- [11] P. Carr, H. Geman, D. B. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13(3):345–382, 2003.
- [12] P. Carr and A. Itkin. An expanded local Variance Gamma model. *Computational Economics*, 57(4):949–987, 2021.
- [13] P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *The Journal of Computational Finance*, 2(4):61–73, 1999.
- [14] E. A. Daal and D. B. Madan. An empirical examination of the Variance-Gamma model for foreign currency options. *The Journal of Business*, 78(6):2121–2152, 2005.
- [15] J. Drahokoupil. Variance Gamma process in the option pricing model. FFA Working Paper 2/2021, FFA, Prague University of Economics and Business, Prague, 2021.
- [16] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-cosine series expansions. *Society for Industrial and Applied Mathematics*, 31(2):826–848, 2008.
- [17] Federal Reserve Bank of St Louis. *Fitted Yield on a XX Year Zero Coupon Bond*, 2024. <https://fred.stlouisfed.org/searchreults?st=zero+coupon+yields>. Accessed: 2024-08-01.
- [18] Federal Reserve Bank of St Louis. *Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis (DGS10)*, 2024. <https://fred.stlouisfed.org/series/DGS10/>. Accessed: 2024-08-01.

- [19] Federal Reserve Bank of St Louis. Market Yield on U.S. Treasury Securities at 3/6-Month Constant Maturity, Quoted on an Investment Basis, 2024. <https://fred.stlouisfed.org/searchresults/?st=Market%20Yield%20on%20U.S.%20Treasury%20Securities>. Accessed: 2024-08-01.
- [20] R. Finlay. *The Variance Gamma (VG) Model with Long Range Dependence*. Ph.d. thesis, University of Sydney, 2009.
- [21] F. Fiorani. *Option Pricing Under the Variance Gamma Process*. Ph.d. thesis, University of Trieste, 2004.
- [22] A. Fischer, R. Gaunt, and A. Sarantsev. The Variance-Gamma distribution: A review. *Statistical Science*, page 31, 2024.
- [23] J. Fry, O. Smart, J.-P. Serbera, and B. Klar. A Variance Gamma model for rugby union matches. *Journal of Quantitative Analysis in Sports*, 17(1):67–75, 2021.
- [24] M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott. *Advances in Mathematical Finance*. Birkhäuser, Boston-Basel-Berlin, 2006.
- [25] M. Gardini, P. Sabino, and E. Sasso. The Variance Gamma++ process and applications to energy markets. *Applied Stochastic Models in Business and Industry*, 38(1):71–90, 2022.
- [26] R. E. Gaunt. Absolute moments of the Variance-Gamma distribution. *Working paper pre-print*, 2024.
- [27] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 1980.
- [28] A. Göncü and H. Yang. Variance-Gamma and Normal-Inverse Gaussian models: Goodness-of-fit to Chinese high-frequency index returns. *North American Journal of Economics and Finance*, 36:279–292, 2016.
- [29] A. Hirsch and D. B. Madan. Pricing American options under Variance Gamma. *Journal of Computational Finance*, 7(2):63–80, 2004.
- [30] A. Hitaj and L. Mercuri. Portfolio allocation using multivariate Variance Gamma models. *Financial Markets Portfolio Management*, 27:65–99, 2013.
- [31] A. Hoyyi, A. Abdurakhman, and D. Rosadi. Variance Gamma process with Monte Carlo simulation and closed form approach for European call option price determination. *Media Statistika*, 14:183–193, 01 2022.
- [32] L. P. Hughston and L. Sánchez-Betancourt. Pricing with Variance Gamma information. *Risks*, 8(2):105, 2020.
- [33] R. V. Ivanov. On risk measuring in the Variance-Gamma model. *Statistics and Risk Modeling*, 35(1-2):23–33, 2017.
- [34] M. Konikov and D. B. Madan. Option pricing using Variance Gamma Markov chains. *Review of Derivatives Research*, 5:81–115, 2002.
- [35] K. Lam, E. Chang, and M. C. Lee. An empirical test of the Variance Gamma option pricing model. *Pacific-Basin Finance Journal*, 10(3):267–285, 2002.
- [36] D. Linders and B. Stassen. The multivariate Variance Gamma model: Basket option pricing and calibration. *Quantitative Finance*, 16(4):555–572, 2016.
- [37] G. Ludvigsson. Numerical methods for option pricing under the CGMY process. Master’s thesis, Uppsala Universitet, 2015.

- [38] D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma process and option pricing. *Review of Finance*, 2(1):79–105, 1998.
- [39] D. B. Madan and G. S. McPhail. Investing in skews. *The Journal of Risk Finance*, 2(1):10–24, 2000.
- [40] D. B. Madan and F. Milne. Option pricing with V.G. martingale components. *Mathematical Finance*, 1(4):39–55, 1991.
- [41] D. B. Madan and E. Seneta. The Variance Gamma (V.G.) model for share market returns. *The Journal of Business*, 63(4):511–24, 1990.
- [42] T. Moosbrucker. Pricing CDOs with correlated Variance Gamma distributions. Research report, Department of Banking, University of Cologne, 2006.
- [43] S. Mozumder, G. Sorwar, and K. Dowd. Revisiting Variance Gamma pricing: An application to S&P500 index options. *International Journal of Financial Engineering*, 2(2):24, 2015.
- [44] A. H. Nzokem. Gamma Variance model: Fractional Fourier transform (FRFT). *Journal of Physics: Conference Series*, 2090:7, 2021.
- [45] A. H. Nzokem. Pricing European options under stochastic volatility models: Case of five-parameter Variance-Gamma process. *Journal of Risk and Financial Management*, 16(55), 2023.
- [46] A. W. Rathgeber, J. Stadler, and S. Stöckl. Modeling share returns: An empirical study on the Variance Gamma model. *Journal of Economics and Finance*, 40:653–682, 2016.
- [47] P. Semeraro. A multivariate Variance Gamma model for financial applications. *International Journal of Theoretical and Applied Finance*, 11(1):1–18, 2008.
- [48] E. Seneta. Fitting the Variance-Gamma model to financial data. *Journal of Applied Probability*, 41:177–87, 2004.
- [49] A. M. Udoye, L. S. Akinola, M. N. Annorzie, and Y. Yakubu. Sensitivity analysis of Variance Gamma parameters for interest rate derivatives. *IAENG International Journal of Applied Mathematics*, 52(2):112–123, 2022.
- [50] M. Wallmeier and M. Diethelm. Multivariate downside risk: Normal versus Variance Gamma. *Journal of Futures Markets*, 32(5):431–458, 2011.
- [51] J. Wang. *The Multivariate Variance Gamma Process and its Applications in Multi-asset Option Pricing*. PhD thesis, University of Maryland, 2009.
- [52] S. P. Warty, H. F. Lopes, and N. G. Polson. Sequential Bayesian learning for stochastic volatility with Variance-Gamma jumps in returns. *Applied Stochastic Models in Business and Industry*, 34(4):460–479, 2017.
- [53] Yahoo-Finance. *SNP - Free Realtime Quote [USD] - GSPC*, 2024. <https://finance.yahoo.com/quote/%5EGSPC/>. Accessed: 2024-08-01.

Appendices

A Code

The Python notebooks related to the computational results found in this thesis can be located in the following Github repository: <https://github.com/SpDallas/VG-Equity-Modelling>