Martin Glanzer, BSc

SPREAD OPTION VALUATION IN ORNSTEIN-UHLENBECK TYPE STOCHASTIC VOLATILITY MODELS

MASTERARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

Masterstudium Finanz- und Versicherungsmathematik

eingereicht an der

Technischen Universität Graz

Betreuer:
O. Univ.-Prof. Dr. phil. Robert TICHY
Assoc. Prof. Carlo SGARRA, Ph.D.

Institut für Analysis und Computational Number Theory (Math A)
Department of Mathematics - Politecnico di Milano

Graz, Mai 2014

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Datum/Date	Unterschrift/Signature

Abstract

This thesis examines the valuation problem for Spread Options in market models, where the volatility process is of Ornstein-Uhlenbeck type. An appropriate formulation of these models in a multidimensional setting requires a matrix subordination approach, i.e., the Ornstein-Uhlenbeck process is driven by a matrix-valued Lévy process, whose increments only take values in the cone of positive semidefinite symmetric matrices. Such models have gained some popularity in the modelling of equity markets and, more recently, in commodity and energy markets. The literature about pricing methods for Spread Options in non-Gaussian setups is sparse; however, in models where the joint characteristic function of the log-return process is known in closed form, some techniques based on the fast Fourier transform (FFT) are available and allow for efficient valuation.

The contribution of this thesis is the following: We compute explicit prices for Spread Options in the two-dimensional OU-Wishart model. From a computational point of view, we compare the results obtained by a Monte-Carlo simulation with those obtained by an FFT method. We investigate the FFT method in detail and implement it in two different ways in order to deal with any given contract-characteristics. Realizing the drawbacks of the particular FFT method, which are revealed by our analyses, we discuss alternative possibilities in order to (possibly) avoid them. In particular, we comment on the Integration-Along-Cut method and take first steps for possible applicability (future work required).

From a modelling perspective, we particularly address the issue of specifying the stationary distribution of the volatility process. Under some simplifying assumptions, we formulate a particular specification of the general multivariate IG-OU type stochastic volatiliy model and derive the joint characteristic function of the two-dimensional log-return process within our model. Moreover, we examine a first testing of parameter sensitivities for this special case that we consider.

Acknowledgements

This thesis marks the end of my student days. Therefore, I would like to take the opportunity to express my gratitude to a number of people, who supported me regarding this thesis, or accompanied me on my way during the past years.

I want to thank Dr. Markus Hofer, who initially suggested to me the interesting subject of Ornstein-Uhlenbeck type Stochastic Volatility models, and who facilitated my visit to the Politecnico di Milano. I want to thank my supervisor, Professor Robert Tichy, for his support and cooperativeness. I particularly want to express my deepest gratitude to Professor Carlo Sgarra from the Politecnico di Milano, who guided me through this thesis, always took his time for all my questions, and made my stay in Milan such an interesting and educational time; I was incredibly lucky to have the privilege to profit from his expertise and passion for mathematics and research. I am thankful for the support of Assistant Professor Daniele Marazzina, who helped me regarding the numerical part of this thesis.

I am indebted to the group of students at TU Graz, who accompanied me all the years of my studies: for all the exercise sessions we prepared together and all the pre-exam discussions. Without this support, it would have never ever been possible for me to proceed in the same way.

On a personal level, I want to say that I am very glad that I have always had friends and companions, who also shared the pleasures of life with me outside the university. There are so many nice experiences and memories that will always remain in my mind and make me smile, whenever I will think back to my student days. Special thanks to everybody, who will be part of one or more of these memories.

Finally, but most of all, I want to thank my parents; not only for their support and for enabling me to live an exciting and pleasant student life without any sorrows, but rather for giving me a strong background and equipping me with motivation and confidence, making it a pleasure to meet the challenges of life and persistently work on my goals.

List of Abbreviations

a.s.	..... almost surely
BM	..... Brownian motion
BDLP	..... Background Driving Lévy process
BNS	..... Barndorff-Nielsen and Shephard (model)
CI	..... confidence interval
CPP	..... Compound Poisson process
CV	..... control variate
DFT	..... Discrete Fourier transform
EMM	..... equivalent martingale measure
FFT	..... Fast Fourier transform
GBM	..... Geometric Brownian motion
GH	..... Generalized Hyperbolic (distribution)
GIG	..... Generalized inverse Gaussian (distribution)
IAC	..... Integration-Along-Cut method
IG	..... inverse Gaussian (distribution/process)
i.i.d	..... independent and identically distributed
MC	..... Monte-Carlo (simulation)
mgf	..... moment generating function
NIG	..... normal inverse Gaussian (distribution/process)
OTC	..... over-the-counter
OU	..... Ornstein-Uhlenbeck
SDE	..... stochastic differential equation
SV	..... Stochastic Volatility (models)

Contents

Chapter 1

Introduction

Spread Options are derivative contracts, where the payoff is made up of the spread between two underlying assets and a predetermined strike price. These kind of derivatives represent a popular instrument for traders in various markets. The vast majority of these contracts is traded over-the-counter. Therefore, efficient and accurate pricing methods as well as appropriate market modelling are of great interest.

The valuation of exotic options is in many cases limited to Monte-Carlo methods. Due to the Fundamental Theorem of Asset Pricing, the arbitrage-free price of any contingent claim can be obtained by computing the discounted expected value of the payoff at maturity, where the expectation is taken with respect to a risk-neutral pricing measure. Therefore, by applying Monte-Carlo methods in order to compute this expectation numerically, it is possible to determine option prices via simulation. However, Monte-Carlo methods require a lot of computation time. In many situations, particularly, when prices of numerous contracts are required at the same time, their computational costs are prohibitive and they can hence not be applied.

The famous work of Carr and Madan [27] has introduced the use of the Fast Fourier transform (FFT) for pricing plain vanilla options by Fourier inversion, in cases where a closed-form expression of the characteristic function of the underlying asset price model is available. Their suggestion represents a fast and accurate pricing technique, thus enjoying great popularity. Hurd and Zhou [51] have extended the work of Carr and Madan by presenting an FFT method for Spread Option pricing. The heart of their work is that they derive a representation of the Fourier transform of the payoff function of a Spread Option in terms of the Gamma function. Option values can then be computed by a numerical bivariate Fourier inversion. They apply their method to three kinds of asset models: the two-asset Black-Scholes model (i.e., bivariate GBM), a three factor Stochastic Volatility model, as well as the Variance-

Gamma model as an example for exponential Lévy models. In principle, however, their method is applicable to any asset model, where the two-dimensional log-asset price process shows an analytic joint characteristic function, which, in addition, satisfies a certain factorization assumption. This approach of Hurd and Zhou [51] represents the best Spread Option pricing technique we are aware of. Therefore, we want to examine its performance regarding Ornstein-Uhlenbeck type stochastic volatility models in this thesis.

It should not be left unmentioned that the very recently suggested lower bound approximation by Caldana and Fusai [24] is presented to be competitive or even superior to Hurd and Zhou's method in some aspects. However, we do not consider these aspects helpful for our purposes.

The development of appropriate models for financial markets has been the stimulus for a great amount of literature. The fundamental work of Black and Scholes [20] and Merton [64] was a milestone in 1973 (see, e.g., the books of Hull [49] or Wilmott [86] for details). In practice, their model based on geometric Brownian motion (GBM) is still widespread and commonly used as a reference model, particularly due to the fact that several explicit results are easily obtainable. However, the shortcomings of this model are very well known: it is not capable of coping with many features that are empirically observed in asset return data. Among these so-called *stylized facts* are aggregational Gaussianity, fat tails, volatility clustering, leverage effects and volatility smiles. Therefore, approaches have been advanced and the resulting models became more complex. As contrary to the Black-Scholes model, where volatility is a constant parameter, in so-called Stochastic Volatility (SV) models, the volatility is assumed to evolve according to a stochastic diffusion process; the asset price process and the volatility process are driven by two (correlated) Brownian motions. Among the most popular specifications belonging to this class are the models proposed by Hull and White [50], Stein and Stein [81], and, foremost, Heston [45]. An alternative approach to describe the features of the market behaviour in a more realistic way is the introduction of jumps in the dynamics of the asset price. Famous specifications in this context are the Jump-Diffusion models of Merton [65] and Kou [58]. The model suggested by Bates [13] combines stochastic modelling of volatility with jumps in the asset price; it is an extension of Heston's model, in the sense that a compound Poisson process is added in the dynamics of the asset price. Diffusion-based models can cope with many stylized facts, in particular, if parameters are fine-tuned in a proper way. Though this flexibility may be appreciated, research has shown that various of the *obtainable* properties of diffusion models

are *generic* in models based on jump processes. Financial modelling with jump processes generally gained a lot of popularity during the 1990's; in particular, the class of Lévy processes, i.e., stochastic processes with independent and stationary increments, became the focus of attention. Hence, numerous Lévy-based models have been proposed and studied in the literature. For a comprehensive survey of Lévy processes in mathematical finance, see the book of Cont and Tankov [31].

Barndorff-Nielsen and Shephard [6, 7] have suggested a model, which combines stochastic volatility and jumps in a non-trivial way. In this model (henceforth termed BNS model), the instantaneous variance is an Ornstein-Uhlenbeck (OU) type process, which is driven by a subordinator, i.e., a (pure) jump Lévy process taking only positive values. OU processes are mean-reverting, stationary processes. There is a "one-to-one correspondence" between the stationary distribution of the process and its so-called Background Driving Lévy Process (BDLP). Hence, these processes offer both analytic tractability as well as modelling-flexibility. The BDLP of the volatility process is also incorporated in the dynamics of the log-asset price. Consequently, volatility and asset prices jump at the same time, which makes the model account for the leverage-effect. Naturally, the BNS model is more complex than, for instance, the model of Bates, where the stochastic component and the jump component are independent from one another. Thus, sometimes computations may get quite involved. However, the BNS model is yet an affine stochastic volatility model in the sense of Keller-Ressel [56], referring to the notion of affine processes introduced by Duffie, Filipović, and Schachermayer [40]. Hence, it is analytically tractable and allows very explicit results. The characteristic function of the log-asset price process is given in closed form in the work of Nicolato and Venardos [68].

It is argued for the preferability of the model of Bates in the book of Cont and Tankov [31], particularly due to its simplicity and greater flexibility, yet offering both stochastic volatility and jumps. Furthermore, the BNS model was strongly criticized for example by Mandelbrot, as part of the discussion of the original paper by Barndorff-Nielsen and Shephard [7]: He saw "little purpose or merit to it", arguing that a successful model should be parsimonious, while, according to him, the BNS model would propose a family of building blocks "of staggering and unmotivated complication". However, considering the fact that the presentation of the BNS model stimulated a considerable amount of literature (e.g., [43,47,48,68]), its attractivity from a mathematical point of view can definitely not be denied. Moreover, OU type SV models have gained some popularity in the modelling of commodity and energy markets (e.g., [14–16]), since it turned out that they are capable of describing particularly well the typical features of these markets' behaviour.

Spread Options are bivariate contingent claims. Though it might be the simplest case of only two dimensions, multivariate modelling is required and, in principle, already brings along all the issues connected to modelling in multiple dimensions.

The BNS model has been generalized to multiple dimensions in the work of Pigorsch and Stelzer [70], and Muhle-Karbe, Pfaffel, and Stelzer [66]. In the multivariate case, the volatility process is represented by a matrix-valued process of Ornstein-Uhlenbeck type, as defined by Barndorff-Nielsen and Stelzer [8]. The essential point is the use of matrix subordinators (see Barndorff-Nielsen and Pérez-Abreu [10]) as driving processes. The resulting multivariate Ornstein-Uhlenbeck type stochastic volatility model incorporates a (non-trivial) stochastic dependence structure among the underlying assets.

Spread Options can be seen as a "bet" on the correlation between two underlyings. Therefore, appropriate modelling of the dependence structure is crucial. The multivariate OU type SV model appears especially interesting from this perspective, as well as, in addition, due to the interest in this model with respect to the commodity and energy markets, where Spread Options are particularly popular.

This thesis is organized as follows. In Chapter 2, we introduce Spread Options; we start with the properties of these derivative contracts and explain their relevance in the markets. Then we discuss available pricing methods for Spread Options. Chapter 3 is dedicated to Lévy processes: We explain the properties of this class of stochastic processes and discuss the theory, where the chapter about OU type SV models will be based on. Moreover, we review the most famous market models and particularly point out their shortcomings, which have led to the development of those models, which are the focus of this thesis. Then the way is paved for Chapter 4, where we enter the matter of OU type SV modelling. We start with a thorough discussion of the one-dimensional BNS model and proceed to its generalization to multiple dimensions, which requires an introduction of the theory of matrix-subordinators. We examine a concrete specification in the two-dimensional case, which has been presented in the literature. Furthermore, we address the issue of prespecifying the stationary distribution of the volatility process in the multi-dimensional case and, under some simplifying assumptions, we define an own model-specification in this context. For this model, we derive the joint characteristic function of the (two-dimensional) log-return process. Chapter 5 contains all our numerical studies and investigations from a computational perspective. Finally, in Chapter 6, we conclude and discuss open issues for future work.

Chapter 2

Spread Options

A Spread Option (of European type) is a derivative contract, where for two underlying processes $S^{(1)}$ and $S^{(2)}$, as well as an exercise price (strike) $K \neq 0$, which is agreed in the contract, the payoff at maturity T is of the form

$$(S_T^{(1)} - S_T^{(2)} - K)^+,$$

where $x^+ := \max\{x, 0\}$. In other words, at the time of maturity, the difference - the *spread* - between two underlyings is compared to a predetermined strike price. One might interpret this as the right to change one asset against the other, with costs K. Figure 2.1 illustrates the exercise region (blue-coloured) of a Spread Option, wheras Figure 2.2 shows the corresponding payoff diagram.

If the strike is set to be $K = 0$ then the contract is referred to as Exchange Option. In this case, the famous Margrabe's formula offers an easy way to evaluate the option. This method, which has been suggested in Margrabe [63], is based on the assumption that the underlying asset price processes follow a bivariate geometric Brownian motion.

Spread Options are traded across many different markets; they are popular for both hedging and speculative intentions. Famous examples are options on

- *Crack Spreads* in the commodity markets:
  The term Crack Spread refers to the price difference between crude oil and refined products (such as heating oil, gasoline, or diesel), which are obtained by "cracking" the crude oil during the refining process. A refiner faces the risk of stagnation or even decline of prices of finished products on the market, while prices of the raw material increase. In order to hedge against this risk, Spread Options are popular instruments. A detailed description of Crack Spreads and the possibilities of hedging the corresponding risk, including the use of Crack

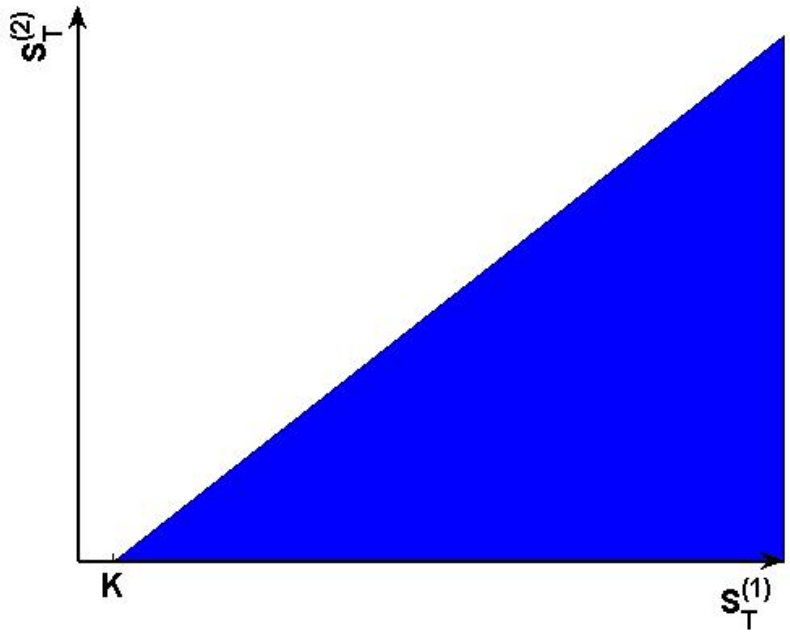


Figure 2.1: Exercise region for a Spread option with strike K

Spread Options, can be found in [29].

- *Spark Spreads* in the energy markets:
  The difference between the price that a power producer receives for selling electricity and the cost of the fuels he needs in order to produce it, is called Spark Spread. The general relevance of Spark Spread Options in the energy derivatives market is discussed in Deng and Oren [39], some stochastic modelling approaches for energy-commodity prices, particularly applied on Spark Spread Options, are presented in Deng [37].

- *Credit Spreads* in the fixed income markets:
  A Credit Spread is the difference in yield between two bonds, which is due to different credit qualities. Consider a defaultable bond with yield Y^* until maturity T and an identical riskless bond with yield Y during the same period. A contract having a payoff which depends on the difference $Y^* - Y$ is called a Credit Spread Option. In [41], Credit Spread Options are discussed and a tree algorithm to price these kind of derivatives is suggested.

- *Index Spreads* in the equity markets:
  The difference between the values of two different stock market indices. Taking a position in a contract having a payoff that depends on the connection between

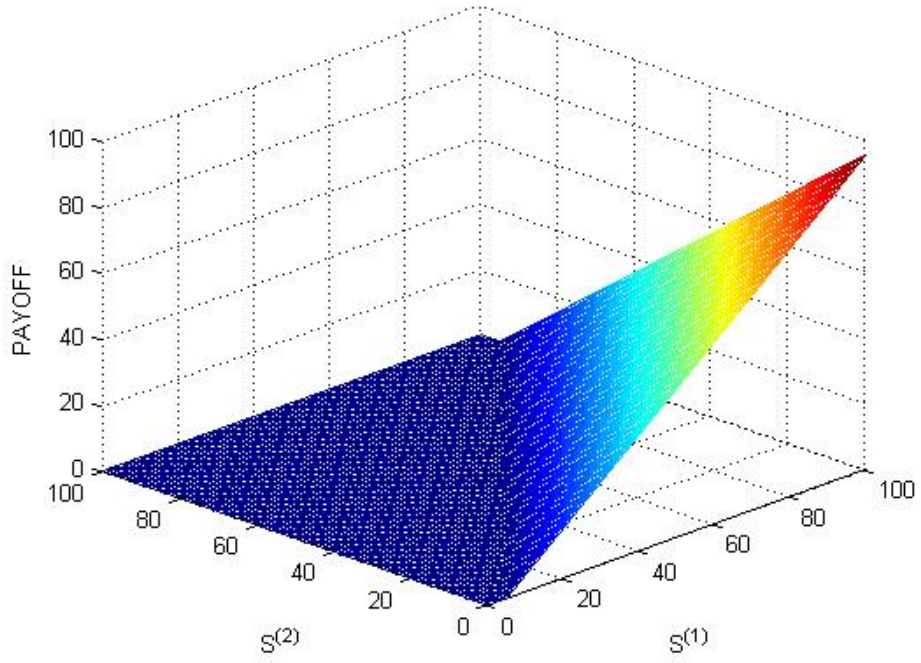


Figure 2.2: Payoff diagram for spot prices $S^{(1)}$ and $S^{(2)}$ of the underlyings and strike $K = 5$.

whole indices rather than on single shares, corresponds to betting on the large movements in the markets.

Sometimes, Spread Options are traded on exchanges (e.g., some types of Crack Spread Options are traded on the New York Mercantile Exchange (NYMEX)). The vast majority is traded over-the-counter (OTC), though. Due to their wide applicability indicated above, there is a remarkable demand for this class of options. However, appropriate pricing is difficult: on the one hand, there is the issue of postulating an adequate model for the spread between two stochastic processes with a non-trivial dependence structure, and on the other hand, there is the need for accurate and efficient techniques in order to determine explicit prices.

2.1 Pricing methods for Spread Options

Carmona and Durrleman [25] provide a comprehensive survey of the available literature and discuss the theoretical and computational problems associated with the different approaches. Based on a bivariate log-normal process, as a model for the two underlying asset prices, several approaches have been presented: For instance, the approximation formula suggested by Kirk [57] turned out to be very popular in

practice. A further analytic approximation formula is due to Carmona and Durrleman [26]; they derive a family of lower as well as an upper price bound, which can yield a very tight interval. Venkatramanan and Alexander [85] suggest a method where they obtain the price of a Spread Option as the sum of the prices of two compound options, one of which is to exchange vanilla Call Options on the two underlying assets and the other one is to exchange the corresponding Put Options. Other methods are, e.g., due to Pearson [69], Deng et al. [38], Borovka et al. [21], or Bjerksund and Stensland [19].

The available literature on Spread Option valuation for general models, i.e., without the assumption of a bivariate log-normal process, is rather sparse. The ansatz proposed by Dempster and Hong [36] was the first efficient method applicable to non-Gaussian models. They extended the Fast Fourier transform technique of Carr and Madan [27] (see sect. 2.1.2) to a multi-factor setting, providing applicability for instance to many models of the affine jump-diffusion type. Analogously to the idea of integrating by Riemann sums, they calculate the Spread Option price by forming tight upper and lower bounds for the integral over a non-polygonal region (the exercise region in logarithmic variables is non-linear).

Hurd and Zhou [51] have proposed a more elegant method as an extension of the logic of Carr and Madan [27]: the core of their procedure is the representation they derive of the Fourier transform of the Spread Option payoff function in terms of the complex Gamma function.

Most recently, Caldana and Fusai [24] have generalized the work of Bjerksund and Stensland [19] by deriving a lower bound approximation for the Spread Option price, which can be applied for any model, where the joint characteristic function of the log-returns of the two underlying assets is available in closed form. Basically, the idea of this approach is the following: Define the event

$$A := \left\{ \omega : \frac{S_T^{(1)}}{\left(S_T^{(2)}\right)^{\alpha}} > \frac{e^k}{\mathbb{E}\left[\left(S_T^{(2)}\right)^{\alpha}\right]} \right\},$$

and consider the lower bound of the Spread Option payoff

$$\left(S_T^{(1)} - S_T^{(2)} - K\right)^+ \geq \left(S_T^{(1)} - S_T^{(2)} - K\right) \mathbb{1}_A.$$

Extending the work of Bjerksund and Stensland [19], who worked in the bivariate GBM setting, Caldana and Fusai [24] suggest how to approximate the exact Spread Option price by $C_K^{k,\alpha}(0) := e^{-rT}\mathbb{E}[(S_T^{(1)} - S_T^{(2)} - K)\mathbb{1}_A]$ (for a suitable choice

of the parameters k and α) for any stock price model, where the joint characteristic function of $(\log S_T^{(1)}, \log S_T^{(2)})$ is available in closed form. In particular, they give a representation of the approximate Spread Option price $C_K^{k,\alpha}(0)$ in terms of a Fourier inversion formula. Their bound has turned out to be very accurate and easily computable. The authors argue that their method improves upon the one due to Hurd and Zhou [51] on some points: First, unlike the Hurd and Zhou method, which is not applicable to Exchange Options (i.e., in the case $K = 0$), their procedure also copes with this case. Second, Hurd and Zhou's technique requires the assumption that the characteristic function of the log-asset price process X_t factorizes as $\mathbb{E}[e^{\mathrm{i}uX_T}|X_0] = e^{\mathrm{i}uX_0}\mathbb{E}[e^{\mathrm{i}u(X_T-X_0)}]$. This assumption rules out mean-reverting asset models, while there is no issue for Caldana and Fusai's method regarding such models. Third, Caldana and Fusai's technique involves only a univariate Fourier inversion, rather than a bivariate one, as is the case for Hurd and Zhou's approach. This improves the computational speed considerably, considering a single contract.

However, the advantages provided by the Caldana and Fusai method are not really relevant for our purposes: The asset models considered in this thesis are not of mean-reverting nature (only the volatility processes are, but this is not an issue). Moreover, we want to study "real" Spread Options rather than Exchange Options. The improvement in terms of the computational speed may be true for a single contract; however, as is also admitted by the authors, considering the valuation of many contracts, Hurd and Zhou's method will still be superior. They implement an interpolation procedure, where the valuation of every further contract only requires very little time, while the computational cost for Caldana and Fusai's lower bound increases linearly in the number of evaluated contracts. Furthermore, even if it has turned out to be very accurate for a number of asset models, Caldana and Fusai still propose a procedure to determine price bounds, while Hurd and Zhou aim for the exact price. For these reasons, we are going to follow the technique suggested by Hurd and Zhou [51]. Now we want to explain this method in detail and previously also give a general introduction to transform-based option pricing.

2.1.1 Fourier transform methods for option pricing

Let $S_t = S_0 e^{X_t}$ be the price process of a risky asset. The arbitrage-free pricing paradigm based on the Fundamental Theorem of Asset Pricing means that the price at time zero of a claim with terminal payoff $f(S_T)$ is the discounted expected value of this payoff, where the expectation is taken with respect to an equivalent

martingale measure, i.e.,

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}_T}[f(S_T)] = e^{-rT}\int_{-\infty}^{\infty} f(x)\rho_T(x)dx,$$

where ρ_T denotes the density of the log-returns under the risk-neutral pricing measure $\mathbb{Q}_T$. However, in many models, this density ρ_T is not available; the probability density of a Lévy processes, for example, is typically not known in closed form. On the other hand, the characteristic function of the process is in most cases available in terms of elementary functions. Therefore, one wants to work with representations, where the distribution of the log-returns appears in terms of its characteristic function. The concrete idea is the following: If the payoff $f(S_T)$ can be represented as an integral of the form $\int g(z)(S_T)^z dz$, and the moment generating function[1] Φ_{X_T} of the log-returns is available in closed form, then we can write

$$\begin{aligned} V_0 &= e^{-rT}\int g(z)\mathbb{E}^{\mathbb{Q}_T}[(S_T)^z]dz \\ &= e^{-rT}\int g(z)\mathbb{E}^{\mathbb{Q}_T}[e^{zX_T}](S_0)^z dz \\ &= e^{-rT}\int g(z)\Phi_{X_T}(z)(S_0)^z dz. \end{aligned}$$

Hence, the price can be determined by a one-dimensional (numerical) integration. The desired representation of the payoff can, e.g., be obtained by using the Fourier transform.

The d-dimensional Fourier transform of a function f is, for $v \in \mathbb{R}^d$, defined by

$$\mathbf{F}\{f(x)\}(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{\mathrm{i}v^\top x} f(x)\; dx_1\cdots dx_d.$$

The function f can be "recovered" by the corresponding Fourier inversion, which is given by

$$f(x) = \frac{1}{(2\pi)^d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-\mathrm{i}x^\top v}\; \mathbf{F}\{f(x)\}(v)\; dv_1\cdots dv_d.$$

2.1.2 Using the Fast Fourier transform for option pricing: The method of Carr and Madan

In order to explain the idea, we first consider a Call Option on the underlying asset price S_t, with maturity T and strike K. Let $k := \log K, s_T := \log S_T$, and denote

[1] We use the moment generating function here for ease of notation; due to the relationship $\Phi(z) = \phi(-\mathrm{i}z)$ to the characteristic function, the same lines can of course analogously be written in terms of ϕ.

the risk-neutral density of the log-asset price by q_T. The following strategy is along the lines of Carr and Madan [27]:

The value of the Call Option at time zero is given by

$$C_T(k) = e^{-rT} \int_{-\infty}^{\infty} (e^x - e^k)^+ q_T(x) dx = e^{-rT} \int_k^{\infty} (e^x - e^k) q_T(x) dx.$$

However, the Call price as a function in k is not (square-)integrable, since

$$\begin{aligned}
\lim_{k \to -\infty} C_T(k) &= \lim_{k \to -\infty} e^{-rT} \int_k^{\infty} (e^x - e^k) q_T(x) dx \\
&= e^{-rT} \int_{-\infty}^{\infty} e^x q_T(x) dx \\
&= \mathbb{E}^{\mathbb{Q}_T}[e^{-rT} S_T] \\
&= S_0.
\end{aligned}$$

L^1-integrability of a function is however a sufficient condition for its Fourier transform to exist. The idea in order to avoid this problem is to modify the Call-pricing function in terms of choosing some $\alpha > 0$ and working with the "damped" function

$$c_T(k) := e^{\alpha k} C_T(k)$$

instead. If we now consider the Fourier transform of $c_T(k)$, which can be written as

$$\begin{aligned}
\psi_T(v) &= \int_{-\infty}^{\infty} e^{\mathrm{i}vk} c_T(k) dk \\
&= \int_{-\infty}^{\infty} e^{\mathrm{i}vk} e^{\alpha k} \left(e^{-rT} \int_k^{\infty} (e^x - e^k) q_T(x) dx \right) dk \\
&= \int_{-\infty}^{\infty} e^{-rT} q_T(x) \int_{-\infty}^{x} (e^{x+\alpha k} - e^{(1+\alpha)k}) e^{\mathrm{i}vk} dk \; dx \\
&= \int_{-\infty}^{\infty} e^{-rT} q_T(x) \left(\frac{e^{(\alpha+1+\mathrm{i}v)x}}{\alpha + \mathrm{i}v} - \frac{e^{(\alpha+1+\mathrm{i}v)x}}{\alpha + 1 + \mathrm{i}v} \right) dx \\
&= \frac{e^{-rT}}{(\alpha + \mathrm{i}v)(\alpha + 1 + \mathrm{i}v)} \int_{-\infty}^{\infty} q_T(x) e^{(\alpha+1+\mathrm{i}v)x} dx \\
&= \frac{e^{-rT}}{\alpha^2 + \alpha - v^2 + \mathrm{i}v(2\alpha + 1)} \int_{-\infty}^{\infty} e^{\mathrm{i}(-\mathrm{i}(\alpha+1)+v)x} q_T(x) dx \\
&= \frac{e^{-rT} \phi_T(v - \mathrm{i}(\alpha + 1))}{\alpha^2 + \alpha - v^2 + \mathrm{i}v(2\alpha + 1)},
\end{aligned} \tag{2.1}$$

then we can represent the option price by Fourier inversion of $\psi_T(v)$ (and undamp-

ing) as

$$C_T(k) = e^{-\alpha k} c_T(k) = e^{-\alpha k} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv$$
$$= \frac{\exp(-\alpha k)}{\pi} \int_0^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv, \tag{2.2}$$

where we have a closed-form expression of $\psi_T(v)$ at hand by (2.1). The last equality holds due to the fact that option prices are real, which means that the imaginary part must vanish, which in turn is the case if the integrand is odd in its imaginary part (for any integration limits symmetrical to the origin). On the other hand, its real part is even, which implies that the integrals over the two half-axes coincide.

In order to obtain an explicit price for the Call Option, we are only left with the evaluation of the integral in (2.2). An easy way for approximating this integral is to apply the rectangle rule: Fix N, choose a step size η, set $v_j = \eta(j-1)$ for $j = 1, \ldots, N$, and use the approximation

$$C_T(k) \approx \frac{\exp(-\alpha k)}{\pi} \sum_{j=1}^{N} e^{-iv_j k} \psi_T(v_j) \eta. \tag{2.3}$$

The Fast Fourier transform (FFT) is an efficient algorithm for computing a discrete Fourier transform (DFT), i.e., a sum of the form

$$X(k) = \sum_{j=1}^{N} e^{-\mathrm{i}\frac{2\pi}{N}(j-1)(k-1)} x(j) \qquad \text{for } k = 1, \ldots, N, \tag{2.4}$$

for an input vector $x = (x_1, \ldots, x_N)$, and where N is (typically) a power of 2. The FFT algorithm reduces the number of operations necessary to compute (2.4) from $\mathcal{O}(N^2)$ (corresponding to computing all the sums directly) to $\mathcal{O}(N \log N)$. Particularly due to the fact that, with respect to (2.3), this would correspond to obtaining prices for a whole range of N different strike values in a very fast way (i.e., with just one evaluation of the very efficient FFT), it is appealing to use the Fast Fourier transform.

The possibility of evaluating discrete Fourier transforms in an efficient way has been crucial for many applications; FFT algorithms are among the most important algorithms in various applied fields. There are many different versions of FFT algorithms available, however, the original[2] formulation due to Cooley and Tukey

[2]In fact, it has been discovered that a similar idea had already been used by Gauss at the beginning of the nineteenth century, but was not really noticed for a long time (cf. Heidemann et al. [44]).

[32] is the most common one, which can certainly be seen as a cornerstone in numerical analysis. The idea of this algorithm is as follows: It is a classical "divide-and-conquer" procedure. Recall that we want to compute the DFT

$$X(k) = \sum_{n=0}^{N-1} x(j) e^{-\mathrm{i}\frac{2\pi}{N}jk} \qquad \text{for } k = 0, \ldots, N-1, \tag{2.5}$$

which corresponds to N^2 operations. The algorithm is designed for values of N that are a power of 2. We introduce the notation

$$w := e^{-\mathrm{i}\frac{2\pi}{N}}.$$

Let us now, as an illustrative example, consider the case $N = 4$. Then we can write the system of equations given in (2.5) in matrix form as

$$\begin{pmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{pmatrix} = \begin{pmatrix} w^0 & w^0 & w^0 & w^0 \\ w^0 & w^1 & w^2 & w^3 \\ w^0 & w^2 & w^4 & w^6 \\ w^0 & w^3 & w^6 & w^9 \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{pmatrix}, \tag{2.6}$$

where the matrix-vector multiplication corresponds to $N^2 = 16$ operations. By noting that $w^0 = 1, w^2 = -1$, as well the general periodicity of the exponent of w modulo N, we can write the Fourier matrix in (2.6) in a simplified form as

$$\mathbf{F}_4 := \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & w^1 & w^2 & w^3 \\ 1 & w^2 & 1 & w^2 \\ 1 & w^3 & w^2 & w^1 \end{pmatrix}.$$

The crucial point is now that $\mathbf{F}_4$ can be factorized as

$$\mathbf{F}_4 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & w^1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -w^1 \end{pmatrix} \begin{pmatrix} 1 & w^0 & 0 & 0 \\ 1 & w^2 & 0 & 0 \\ 0 & 0 & 1 & w^0 \\ 0 & 0 & w^0 & w^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.7}$$

where the last matrix is a permutation matrix, which "resorts" a vector by first

listing all even components, then followed by all the odd ones, i.e.,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(2) \\ x(1) \\ x(3) \end{pmatrix}.$$

We notice from equation (2.7) that $\mathbf{F}_4$ is of the form

$$\mathbf{F}_4 = \begin{pmatrix} I_2 & W_2 \\ I_2 & -W_2 \end{pmatrix} \begin{pmatrix} \mathbf{F_2} & 0 \\ 0 & \mathbf{F_2} \end{pmatrix} \begin{pmatrix} \text{permutation} \\ \text{matrix} \end{pmatrix},$$

where I_2 is the (2-dim.) identity matrix, $W_2 = \begin{pmatrix} 1 & 0 \\ 0 & w^1 \end{pmatrix}$, and $\mathbf{F_2} = \begin{pmatrix} 1 & w^0 \\ w^0 & w^2 \end{pmatrix}$. Therefore, we have reduced the number of operations from $N^2 = 16$ to $2\left(\frac{N}{2}\right)^2 = \frac{N^2}{2} = 8$.

The factorization we just performed on $\mathbf{F_4}$, where we split the problem into two smaller problems of half size, resulting in a reduction of the number of computational operations, is generally applicable; it yields the recursion

$$\mathbf{F}_n = \begin{pmatrix} I_{n/2} & W_{n/2} \\ I_{n/2} & -W_{n/2} \end{pmatrix} \begin{pmatrix} \mathbf{F_{n/2}} & 0 \\ 0 & \mathbf{F_{n/2}} \end{pmatrix} \begin{pmatrix} \text{permutation} \\ \text{matrix} \end{pmatrix},$$

where $W_{n/2} = \text{diag}(w^0, w^1, \ldots, w^{n/2-1})$. Hence, the problem of a DFT of size N can be fully reduced to two DFT's, each of size $N/2$. Since we assumed that $N = 2^m$, after $m = \log_2 N$ steps, a DFT of size N can be reduced to N Fourier transforms, each of size 1. Due to the fact that the Fourier transform of a single number is the number itself (see equation (2.5) for $N = 1$), the algorithm terminates trivially. This means that there is a total of $\log_2 N$ "stages" of computation. Each of them requires $\mathcal{O}(N)$ complex operations. Therefore, the FFT algorithm reduces the number of operations required to evaluate a DFT from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$.

Let us now get back to the aim of evaluating the option by applying the FFT in order to compute (2.2): We choose a step size λ in the "strike-world" and consider the values k_u for k along a regularly-spaced grid, which are given by

$$k_u = -b + \lambda(u-1), \qquad \text{for } u = 1, \ldots, N,$$

where $b = \frac{N\lambda}{2}$. Consequently, for $u = 1, \ldots, N$, we get

$$\begin{aligned} C_T(k_u) &\approx \frac{\exp(-\alpha k_u)}{\pi} \sum_{j=1}^{N} e^{-\mathrm{i}v_j(-b+\lambda(u-1))} \psi_T(v_j)\eta, \\ &= \frac{\exp(-\alpha k_u)}{\pi} \sum_{j=1}^{N} e^{-\mathrm{i}\lambda\eta(j-1)(u-1)} e^{\mathrm{i}bv_j} \psi_T(v_j)\eta. \end{aligned} \tag{2.8}$$

Thus, if we choose the "Fourier grid" and the "strike grid" such that

$$\lambda\eta = \frac{2\pi}{N},$$

then we can apply the FFT (cf. (2.4)) on (2.8) and we obtain as a result all the values $C_T(k_u)$, for $u = 1, \ldots, N$.

An alternative approach regarding the integrability issue

The approach explained above makes use of a damping factor $e^{\alpha k}$ and works with the damped Call pricing function $c_T(k) = e^{\alpha k} C_T(k)$, since the function $C_T(k)$ itself is not in L^1. Another way of circumventing the lacking integrability is to work with time-values instead of damped option prices. This idea is also due to Carr and Madan [27]. They suggested it, since they noticed that for short maturities and strike values far from the at-the-money level, the integrand in (2.2) becomes highly oscillatory, and hence difficult to integrate numerically. We want to explain this technique here, since we will come back to this idea at a later stage. The particular strategy is the following:

The time-value of an option is defined as the option price subtracted by the intrinsic value. It can be interpreted as the premium an investor pays over the current exersice value of the option. Naturally, the time-value decays exponentially to zero when approaching maturity. Let now

$$z_T(k) := e^{-rT} \mathbb{E}^{\mathbb{Q}_T} \left[\left(e^{X_T} - e^k \right)^+ \right] - \left(1 - e^k \right)^+$$

be the time-value of a Call Option as a function in the log-strike k, and denote its Fourier transform by

$$\zeta_T(v) := \mathbf{F}\{z_T(k)\}(v) = \int_{-\infty}^{\infty} e^{\mathrm{i}vk} z_T(k) dk. \tag{2.9}$$

Exploiting the fact that the discounted price process is a $\mathbb{Q}_T$-martingale, we get

$$\begin{aligned} z_T(k) &= e^{-rT} \int_{-\infty}^{\infty} \rho_T(x)(e^x - e^k)\mathbb{1}_{\{k \leq x\}}\, dx - (1 - e^k)\mathbb{1}_{\{k \leq 0\}} \\ &= e^{-rT} \int_{-\infty}^{\infty} \rho_T(x)(e^x - e^k)(\mathbb{1}_{\{k \leq x\}} - \mathbb{1}_{\{k \leq 0\}})\, dx. \end{aligned}$$

If we plug this representation of $z_T(k)$ into (2.9) then we can write[3]

$$\begin{aligned} \zeta_T(v) &= e^{-rT} \int_{-\infty}^{\infty} \rho_T(x) \left(\int_0^x e^{\mathrm{i}vk}(e^x - e^k)dk \right) dx \\ &= e^{-rT} \int_{-\infty}^{\infty} \rho_T(x) \left(\frac{(1 - e^x)}{1 + \mathrm{i}v} + \frac{\mathrm{i}e^x}{(1 + \mathrm{i}v)v} - \frac{\mathrm{i}e^{x + \mathrm{i}vx}}{(1 + \mathrm{i}v)v} \right) dx \\ &= e^{-rT} \left(\frac{1 - \mathbb{E}^{\mathbb{Q}_T}[S_T]}{1 + \mathrm{i}v} + \frac{\mathrm{i}}{(1 + \mathrm{i}v)v} \left(\mathbb{E}^{\mathbb{Q}_T}[S_T]) - \int_{-\infty}^{\infty} \rho_T(x) e^{\mathrm{i}(v - \mathrm{i})x} dx \right) \right) \\ &= \frac{e^{-rT}}{1 + \mathrm{i}v} - \frac{1}{1 + \mathrm{i}v} + \frac{\mathrm{i}}{(1 + \mathrm{i}v)v} - \frac{\mathrm{i}e^{-rT}\phi_T(v - \mathrm{i})}{(1 + \mathrm{i}v)v} \\ &= \frac{e^{-rT}}{1 + \mathrm{i}v} + \frac{-v + \mathrm{i}}{(\mathrm{i} - v)(-\mathrm{i}v)} - \frac{e^{-rT}\phi_T(v - \mathrm{i})}{v(v - \mathrm{i})} \\ &= e^{-rT} \left(\frac{1}{1 + \mathrm{i}v} - \frac{e^{rT}}{\mathrm{i}v} - \frac{\phi_T(v - \mathrm{i})}{v(v - \mathrm{i})} \right). \end{aligned}$$

With this closed-form expression of the Fourier transform at hand, we can again proceed as before; we obtain option prices by numerically inverting the Fourier transform (with possible use of the FFT analogously to above).

2.1.3 The method of Hurd and Zhou for Spread Option pricing

Hurd and Zhou [51] have extended the idea of Carr and Madan, i.e, the application of the FFT for option-pricing, to two dimensions: they propose an efficient strategy for the pricing of Spread Options. The core of their work is the following theorem, which gives a representation of the Fourier transform of the payoff of a Spread Option in terms of the complex Gamma function. Since we will follow their strategy for our numerical studies in chapter 5, we also want to explain the proof of this central result in detail.

Theorem 2.1: *Let $P(x_1, x_2) = (e^{x_1} - e^{x_2} - 1)^+$ be the payoff function of a Spread Option with strike 1. Then, for any vector of real numbers $\epsilon = (\epsilon_1, \epsilon_2)$ with $\epsilon_2 > 0$*

[3] In order to ensure the feasibility of all the rearrangements, particularly, of the interchanging of the integration order, the following technical condition must be assumed: $\exists\ \alpha > 0 : \int_{-\infty}^{\infty} \rho_T(x) e^{(1+\alpha)x} dx < \infty$. However, as contrary to the strategy with the damping factor, we do not need to choose a specific value of α in order to obtain the option price.

and $\epsilon_1 + \epsilon_2 < -1$ *and* $x = (x_1, x_2)$*, the payoff function has the representation*

$$P(x) = \frac{1}{(2\pi)^2} \int \int_{\mathbb{R}^2 + \mathrm{i}\epsilon} e^{\mathrm{i}ux^t} \hat{P}(u) d^2u, \tag{2.10}$$

where

$$\hat{P}(u) = \frac{\Gamma(\mathrm{i}(u_1 + u_2) - 1)\Gamma(-\mathrm{i}u_2)}{\Gamma(\mathrm{i}u_1 + 1)}.$$

Here, $\Gamma(z)$ *denotes the complex Gamma function defined for* $\mathrm{Re}(z) > 0$ *by the integral* $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$*.*

Proof. Let $\epsilon_2 > 0$ and $\epsilon_1 + \epsilon_2 < -1$. Application of the Fourier inversion theorem on $e^{\epsilon x} P(x)$ (the factor $e^{\epsilon x}$ ensures that this is in $L^2(\mathbb{R}^2)$) yields

$$\begin{aligned}
e^{\epsilon x} P(x) &= \mathbf{F}^{-1}\{\mathbf{F}\{e^{\epsilon x} P(x)\}(u)\}(x) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{\mathrm{i}ux} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-\mathrm{i}ux} e^{\epsilon x} P(x) du_1 du_2 dx_1 dx_2 \\
&= \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{\epsilon x} e^{\mathrm{i}(u + \mathrm{i}\epsilon)x} \int_{\mathbb{R}^2} e^{-\mathrm{i}(u + \mathrm{i}\epsilon)x} P(x) d^2x d^2u \\
&= \frac{1}{2\pi} e^{\epsilon x} \int_{\mathbb{R}^2 + \mathrm{i}\epsilon} e^{\mathrm{i}ux} \int_{\mathbb{R}^2} e^{-\mathrm{i}ux} P(x) d^2x d^2u \\
&= e^{\epsilon x} \left(\frac{1}{2\pi} \int_{\mathbb{R}^2 + \mathrm{i}\epsilon} e^{\mathrm{i}ux} g(u) d^2u \right).
\end{aligned}$$

Therefore,

$$P(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2 + \mathrm{i}\epsilon} e^{\mathrm{i}ux} g(u) d^2u,$$

where

$$g(u) = \int_{\mathbb{R}^2} e^{-\mathrm{i}ux} P(x) d^2x.$$

The payoff function $P(x) = (e^{x_1} - e^{x_2} - 1)^+$ means that the domain of the integral can be restricted to $\{x = (x_1, x_2) : x_1 > 0, e^{x_2} < e^{x_1} - 1\}$, and the function g can be hence be written as

$$\begin{aligned}
g(u) &= \int_0^\infty e^{-\mathrm{i}u_1 x_1} \left(\int_{-\infty}^{\log(e^{x_1} - 1)} e^{-\mathrm{i}u_2 x_2} (e^{x_1} - e^{x_2} - 1) dx_2 \right) dx_1 \\
&= \int_0^\infty e^{-\mathrm{i}u_1 x_1} (e^{x_1} - 1)^{1 - \mathrm{i}u_2} \left(\frac{1}{-\mathrm{i}u_2} - \frac{1}{1 - \mathrm{i}u_2} \right) dx_1.
\end{aligned}$$

Substituting e^{-x_1} by z leads to

$$g(u) = \frac{1}{(1-\mathrm{i}u_2)(-\mathrm{i}u_2)} \int_0^1 z^{\mathrm{i}u_1} \left(\frac{1-z}{z}\right)^{1-\mathrm{i}u_2} \frac{dz}{z}.$$

Using the Beta function

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

which is defined on $\{a, b \in \mathbb{C} : \mathrm{Re}(a), \mathrm{Re}(b) > 0\}$ as

$$B(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz,$$

and the property $\Gamma(z) = (z-1)\Gamma(z-1)$ of the Gamma function, the function g can further be written as

$$\begin{aligned} g(u) &= \frac{\Gamma(\mathrm{i}(u_1+u_2)-1)\Gamma(-\mathrm{i}u_2+2)}{(1-\mathrm{i}u_2)(-\mathrm{i}u_2)\Gamma(\mathrm{i}u_1+1)} \\ &= \frac{\Gamma(\mathrm{i}(u_1+u_2)-1)(-\mathrm{i}u_2+1)(-\mathrm{i}u_2)\Gamma(-\mathrm{i}u_2)}{(1-\mathrm{i}u_2)(-\mathrm{i}u_2)\Gamma(\mathrm{i}u_1+1)} \\ &= \frac{\Gamma(\mathrm{i}(u_1+u_2)-1)\Gamma(-\mathrm{i}u_2)}{\Gamma(\mathrm{i}u_1+1)}. \end{aligned}$$

□

If we assume that the process X has independent and stationary increments (this asssumption will hold true for all models studied in chapter 4), we can use (2.10) to obtain the following representation of the price of a Spread Option:

$$\begin{aligned} \mathrm{Spr}(X_0;T) &= \mathbb{E}[e^{-rT}P(X_T)|\mathcal{F}_0] \\ &= e^{-rT}\mathbb{E}\left[(2\pi)^{-2}\int\int_{\mathbb{R}^2+\mathrm{i}\epsilon} e^{\mathrm{i}uX_T^t}\hat{P}(u)d^2u \,\middle|\, \mathcal{F}_0\right] \\ &= (2\pi)^{-2}e^{-rT}\int\int_{\mathbb{R}^2+\mathrm{i}\epsilon} e^{\mathrm{i}uX_0^t}\mathbb{E}\left[e^{\mathrm{i}u(X_T-X_0)^t}\,\middle|\,\mathcal{F}_0\right]\hat{P}(u)d^2u \\ &= (2\pi)^{-2}e^{-rT}\int\int_{\mathbb{R}^2+\mathrm{i}\epsilon} e^{\mathrm{i}uX_0^t}\mathbb{E}\left[e^{\mathrm{i}uX_T^t}\right]\hat{P}(u)d^2u \\ &= \frac{1}{(2\pi)^2}e^{-rT}\int\int_{\mathbb{R}^2+\mathrm{i}\epsilon} e^{\mathrm{i}uX_0^t}\phi_{X_T}(u)\hat{P}(u)d^2u. \end{aligned} \tag{2.11}$$

For a numerical evaluation of the double integral in 2.11, the authors of [51] suggest to follow the logic of Carr and Madan (see section 2.1.2) in the following way: In

the two-dimensional setting it means an approximation of the double integral by choosing a truncation interval $[-\bar{u}, \bar{u}]$, as well as a number of steps N (we will always use powers of 2, which is a convenient choice and ensures that N is even, what will be assumed implicitly for some rearrangements in the sequel) within this interval, and evaluating the double sum over the corresponding lattice

$$\Gamma = \{u(k) = (u_1(k_1), u_2(k_2)) \mid k = (k_1, k_2) \in \{0, \dots, N-1\} \times \{0, \dots, N-1\}\},$$

with $u_i(k_i) = -\bar{u} + k_i\eta$, and where η is the step size, resulting from the choice of $\bar{u}$ and N, by $\eta = 2\bar{u}/N$. The reciprocal lattice is given by

$$\Gamma^* = \{x(l) = (x_1(l_1), x_2(l_2)) \mid l = (l_1, l_2) \in \{0, \dots, N-1\} \times \{0, \dots, N-1\}\},$$

with $x_i(k_i) = -\bar{x} + l_i\eta^*$, and where $\eta^* = 2\pi/(N\eta) = \pi/\bar{u}$ and $\bar{x} = N\eta^*/2$.

In particular, under the assumption that the vector of the logarithmic initial values of the underlying asset is lying on the lattice Γ^*, i.e., that there exists a vector $l^* = (l_1^*, l_2^*) \in \{0, \dots, N-1\}^2$ such that $X_0 = x(l^*) \in \Gamma^*$, for $S_0 = e^{X_0}$, a simple application of the rectangle method on the double integral in (2.11) yields

$$\mathrm{Spr}(X_0; T) \approx \frac{\eta^2 e^{-rT}}{(2\pi)^2} \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} e^{\mathrm{i}(u(k)+\mathrm{i}\epsilon)x(l^*)^t} \phi_{X_T}(u(k) + \mathrm{i}\epsilon)\hat{P}(u(k) + \mathrm{i}\epsilon). \tag{2.12}$$

The exponential function in (2.12) can be written as

$$\begin{aligned}
e^{\mathrm{i}(u(k)+\mathrm{i}\epsilon)x(l^*)^t} &= e^{-\epsilon x(l^*)^t} e^{\mathrm{i}\left(u_1(k_1)x_1(l_1^*) + u_2(k_2)x_2(l_2^*)\right)} \\
&= e^{-\epsilon x(l^*)^t} \cdot \exp\left(\mathrm{i}\left((-\bar{u} + k_1\eta)(-\bar{x} + l_1^*\eta^*) + (-\bar{u} + k_2\eta)(-\bar{x} + l_2^*\eta^*)\right)\right) \\
&= e^{-\epsilon x(l^*)^t} \cdot \exp(\mathrm{i}((-\tfrac{N\eta}{2} + k_1\eta)(-\tfrac{\pi}{\eta} + l_1^*\eta^*) + (-\tfrac{N\eta}{2} + k_2\eta)(-\tfrac{\pi}{\eta} + l_2^*\eta^*))) \\
&= e^{-\epsilon x(l^*)^t} \cdot e^{\mathrm{i}\pi N} \cdot e^{-\mathrm{i}\pi(k_1+k_2+l_1^*+l_2^*)} \cdot e^{\frac{2\pi\mathrm{i}}{N}(k_1 l_1^* + k_2 l_2^*)} \\
&= e^{-\epsilon x(l^*)^t} \cdot (-1)^{\mathrm{i}\pi(k_1+k_2+l_1^*+l_2^*)} \cdot e^{2\pi\mathrm{i}kl^{*t}/N},
\end{aligned}$$

which leads to the representation

$$\mathrm{Spr}(X_0; T) \approx (-1)^{l_1^*+l_2^*} e^{-rT} \left(\frac{\eta N}{2\pi}\right)^2 e^{-\epsilon x(l^*)^t} \left[\frac{1}{N^2} \sum_{k_1,k_2=0}^{N-1} e^{\frac{2\pi\mathrm{i}kl^{*t}}{N}} h(k)\right], \tag{2.13}$$

where

$$h(k) = (-1)^{k_1+k_2} \phi_{X_T}(u(k) + \mathrm{i}\epsilon)\hat{P}(u(k) + \mathrm{i}\epsilon).$$

The term in square brackets on the right hand side of equation (2.13) is now of a form such that it corresponds exactly to that value of a double inverse discrete Fourier transform of the function h, which corresponds to l^*. The double inverse discrete Fourier transform can be efficiently computed by using the FFT algorithm, and is a standard-routine in all well-known software packages. In MATLAB, for instance, the function[4] `ifft2()` applied on the $N \times N$-dimensional input array H, where $H[k_1, k_2] := h(k)$ for $k = (k_1, k_2) \in \{0, \ldots, N-1\}^2$, yields the output array Y, where, for $l = (l_1, l_2) \in \{0, \ldots, N-1\}^2$,

$$Y[l] = \frac{1}{N^2} \sum_{k_1,k_2=0}^{N-1} e^{\frac{2\pi i k l^t}{N}} H[k].$$

In this way, one gets prices for Spread Options with strike 1 and log-initial values $X_0^{(1)}, X_0^{(2)}$ lying on Γ^*. The assumption regarding the strike is, in fact, not a curtailment of generality: For a general strike $K \in \mathbb{R}$, we can write

$$\begin{aligned} \mathrm{Spr}(X_0; T) &= \mathbb{E}[e^{-rT}(S_T^{(1)} - S_T^{(2)} - K)^+] \\ &= K \cdot e^{-rT}\mathbb{E}[(S_T^{(1)}/K - S_T^{(2)}/K - 1)^+] \\ &= K \cdot e^{-rT}\mathbb{E}[(e^{X_T^{(1)}} - e^{X_T^{(2)}} - 1)^+], \end{aligned}$$

which means that by setting $X^{(1)} := \log(S^{(1)}/K)$ and $X^{(2)} := \log(S^{(2)}/K)$ instead of $\log(S^{(1)})$ resp. $\log(S^{(2)})$, i.e., by using "moneyness" instead of the absolute values of the underlying assets, we can use the pricing procedure for a Spread Option with strike 1 as described above and simply have to multiply the resulting option price by K.

The assumption that both of the log-initial values $X_0^{(1)}$ and $X_0^{(2)}$ exactly lie on the lattice in the Fourier-space, is of theoretical nature, though. In practice, if one wants to use this method for the pricing of arbitrary Spread Option contracts (as we will do in chapter 5), one has to think about how to deal with any given set $\{S_0^{(1)}, S_0^{(2)}, K\}$. We want to use two different approaches: The first one is to choose the lattice Γ, i.e., the lattice in the Fourier-space, not just by specifying any truncation interval and the number of steps, but rather exactly in such a way that the log-initial values fall on grid points of $\Gamma^{(*)}$. In particular, we do this in the following way: We first fix the number of steps N and the truncation-interval "roughly" as $[-\bar{u}_0, \bar{u}_0]$. Then, for each of the two assets, we keep stretching the

[4] One needs to be careful when using these built-in functions: in different software packages, Fourier transforms are implemented with different parameters. In MATHEMATICA, for example, the function `InverseFourier` corresponds to MATLAB's `ifft`; however, exact correspondence is only given if one uses the option `InverseFourier[...,FourierParameters->{1,-1}]`.

interval until it is exactly of such a length that one of the N points in this interval matches the log-initial value of the corresponding asset. It means that, in general, we will get different truncation intervals along the two axes, i.e. $[-\bar{u}^{(i)}, \bar{u}^{(i)}]$, for $i = 1, 2$. The following algorithm describes this strategy:

Algorithm 2.1:

FOR $i = 1 : 2$
  $\bar{u}^{(i)} = \bar{u}_0$;
  FOR $j = 1 : N$
    $\bar{u}_{\texttt{temp}} := \pi \frac{j - N/2}{X_0^{(i)}}$;
    IF $\bar{u}_{\texttt{temp}} > \bar{u}_0$
      $\bar{u}^{(i)} = \bar{u}_{\texttt{temp}}$;
      BREAK;
    END
  END
  IF $\bar{u}^{(i)} == \bar{u}_0$
    PRINT "Error: Choose a smaller value for $\bar{u}_0$.";
  END
END

Proof. Having run this algorithm, for $i = 1, 2$, there is a $j \in 1, \ldots, N$ such that

$$\eta^{*(i)} = \frac{\pi}{\bar{u}^{(i)}} = \frac{\pi}{\pi \left(\frac{j - N/2}{X_0^{(i)}} \right)} = \frac{X_0^{(i)}}{j - N/2}.$$

Therefore,

$$\begin{aligned} x_i(j) &= -\bar{x} + j\eta^{*(i)} \\ &= \eta^{*(i)}(-N/2 + j) \\ &= \frac{X_0^{(i)}}{(j - N/2)}(j - N/2) \\ &= X_0^{(i)}. \end{aligned}$$

□

For a single Spread Option contract with arbitrary characteristics $\{S_0^{(1)}, S_0^{(2)}, K\}$, we can hence evaluate the price. However, if we want to price a set of different contracts

(think of a calibration of model parameters, for instance, where one usually needs to calculate prices for a lot of different options at the same time), we will need to run the whole selection procedure of the lattice as well as the Fourier inversion for every single contract. This coincides with an increase in computation time.
Therefore, as an alternative approach, we want to exploit that, in fact, with every Fourier inversion we get a whole $N \times N$ matrix of prices, corresponding to contracts with characteristics $\{e^{x_1(l_1)}, e^{x_2(l_2)}, K\}$, for $l = (l_1, l_2) \in \{0, \dots, N-1\}^2$. Given all those prices, we can approximate any contract with $S_0^{(i)} \in [\min\{e^{x_i(l_i)} : l_i \in \{0, \dots, N-1\}\}, \max\{e^{x_i(l_i)} : l_i \in \{0, \dots, N-1\}\}]$, for $i = 1, 2$.

Figure 2.3 is intended to emphasize that, by interpolating between grid points, we get a whole surface of prices (by only one (double) Fourier inversion); values between lattice points are approximated linearly here. For our numerical studies in chapter 5, in order to get a better fit, we will approximate the values of the contracts in question by fitting a third-degree polynomial to the values in the neighbourhood, which is made up of the 3×3 matrix of the closest lattice points.

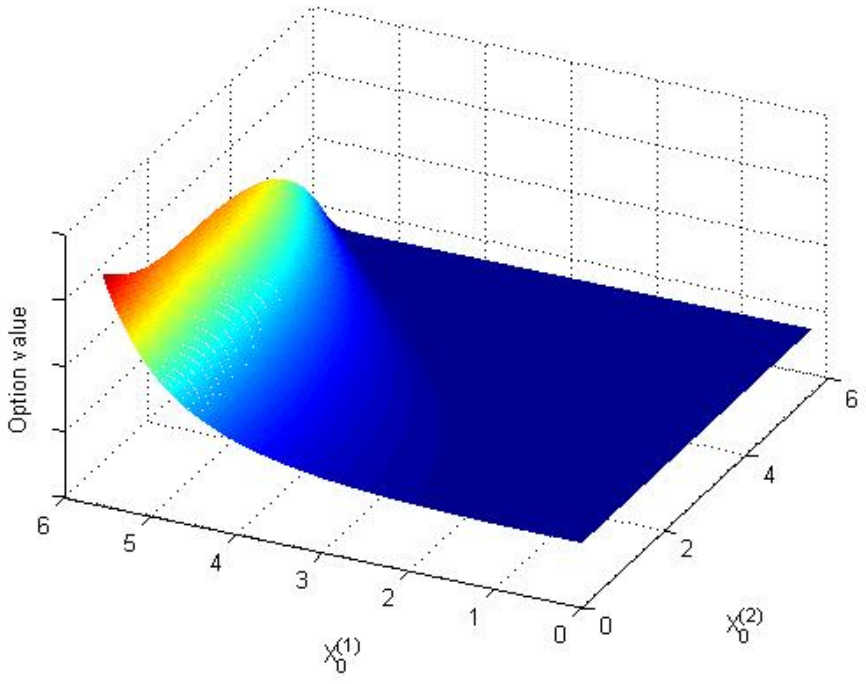


Figure 2.3: Surface of Spread Option prices for corresponding log-initial values $X_0^{(1)}$ and $X_0^{(2)}$.

Chapter 3

Theory of Lévy Processes and Market Modelling

3.1 Lévy Processes

3.1.1 A review of general theory

Definition 3.1 (Lévy Process)**:** *A stochastic process $X = (X_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a* Lévy process*, if it satisfies the following conditions:*

1. $X_0 = 0$, $\mathbb{P}$*-a.s.*
2. *The increments are independent:*
   For any increasing series of times $t_0, t_1, \ldots, t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.
3. *The increments are stationary:*
   For any $t_1, t_2, h \geq 0$, it holds that $X_{t_1+h} - X_{t_1} \stackrel{d}{=} X_{t_2+h} - X_{t_2}$. In particular, the distribution of the increment $X_{t+h} - X_t$ does not depend on t.
4. *X is stochastically continuous:*
   $\forall \epsilon > 0 : \lim_{h\to 0} \mathbb{P}\left[|X_{t+h} - X_t| \geq \epsilon\right] = 0.$

The requirement of stochastic continuity means that the probability of a discontinuity in the trajectory, at a given concrete time t, equals zero, i.e., jumps only occur randomly. Brownian motion is the only Lévy process having indeed continuous sample paths.

The distribution of the increments of a Lévy process cannot be chosen arbitrarily; there are constraints on feasible choices, in particular, the distribution of X_t has to be infinitely divisible. The formal statement is given below in Proposition 3.1.

Definition 3.2 (Infinite divisibility)**:** *A probability distribution μ is called* infinitely divisible *if for any $n \in \mathbb{N}, n \geq 2$, there exist independent and identically distributed random variables $Y_1, \ldots, Y_n$ such that the distribution of the sum $Y_1 + \ldots + Y_n$ is given by μ.*

Proposition 3.1 (Infinite divisibility and Lévy processes (cf. [31], Prop. 3.1))**:** *Let $X = (X_t)_{t \geq 0}$ be a Lévy process. Then for any given time t, the distribution of X_t is infinitely divisible. Conversely, for any given infinitely divisible distribution μ there exists a Lévy process X such that the distribution of X_1 is given by μ.*

Two cornerstones of the theory of Lévy processes are the so-called Lévy-Khintchine formula (see Theorem 3.2), which describes the distributional properties of a Lévy process by giving a representation of its characteristic function, and the Lévy-Itō decomposition (see Theorem 3.1), which gives indication of the structure of the trajectories of the process.

Definition 3.3 (Lévy measure)**:** *For a Lévy process $X = (X_t)_{t \geq 0}$ with jumps $\Delta X_t = X_t - X_{t-}$, its* Lévy measure *$\nu$ is defined as*

$$\nu(A) := \mathbb{E}[\#\{t \in (0,1] : 0 \neq \Delta X_t \in A\}], \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

Definition 3.4 (Poisson random measure)**:** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a propability space and define $\mathbb{M}_{\mathbb{N}}(\mathbb{R} \times [0,\infty)) := \{\mu : \mu \text{ is a measure on } \mathbb{R} \times [0,\infty);\ \mu(A \times I) \in \mathbb{N},\ \forall A \in \mathcal{B}(\mathbb{R}),\ \forall I \in \mathcal{B}([0,\infty))\}$. A mapping $N : \Omega \to \mathbb{M}_{\mathbb{N}}(\mathbb{R} \times [0,\infty))$ is called a* Poisson random measure *if*

1. *$\forall A \in \mathcal{B}(\mathbb{R}), \forall I \in \mathcal{B}([0,\infty))$:*
   $N(A \times I)$ is a Poisson random variable with parameter $\nu(A) \cdot \lambda(I)$
   (λ denoting the Lebesgue measure, $\nu(A) < \infty$).
2. *$\forall A_1 \times I_1, A_2 \times I_2 \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}([0,\infty))$ with $A_1 \times I_1 \cap A_2 \times I_2 = \emptyset$:*
   - *$N(A_1 \times I_1)$ and $N(A_2 \times I_2)$ are independent,*
   - *$N(A_1 \times I_1 \cup A_2 \times I_2) = N(A_1 \times I_1) + N(A_2 \times I_2)$.*

Definition 3.5 (Compensated Poisson random measure)**:** *Let N be a Poisson random measure. Then its* compensated *measure $\tilde{N}$ is defined as*

$$\tilde{N}(A \times I) := N(A \times I) - \nu(A) \cdot \lambda(I).$$

Theorem 3.1 (Lévy-Itō decomposition; cf. [3], Prop. 3.7)**:** *Let $X = (X_t)_{t \geq 0}$ be a Lévy process on $\mathbb{R}^d$. Then there exists a vector $\gamma \in \mathbb{R}^d$, a d-dimensional Brownian*

motion B^A with covariance matrix A, and an independent Poisson random measure N on $\mathbb{R}^d\backslash\{0\} \times \mathcal{B}(\mathbb{R}^+)$ such that, for each $t \geq 0$,

$$X_t = \gamma t + B_t^A + \int_{|x|<1} x\tilde{N}(dx, [0,t]) + \int_{|x|\geq 1} xN(dx, [0,t]). \tag{3.1}$$

Theorem 3.2 (Lévy-Khintchine representation; cf. [77], Thm 8.1)**:** *Let μ be an infinitely divisible distribution on $\mathbb{R}^d$. Then, for $z \in \mathbb{R}^d$, the characteristic function of μ is given by $\phi(z) = e^{\psi(z)}$, where*

$$\psi(z) = -\frac{1}{2}\langle z, Az\rangle + \mathrm{i}\langle \gamma, z\rangle + \int_{\mathbb{R}^d} \left(e^{\mathrm{i}\langle z,x\rangle} - 1 - \mathrm{i}\langle z, x\rangle \mathbb{1}_{\{|x|\leq 1\}}\right) \nu(dx), \tag{3.2}$$

for a vector $\gamma \in \mathbb{R}^d$, a symmetric positive-semidefinite $d\times d$ matrix A, and a measure ν on $\mathbb{R}^d$ satisfying

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty. \tag{3.3}$$

The function $\psi(z)$ is called the characteristic exponent *of μ and its representation in* (3.2) *by γ, A, and ν is unique.*

Conversely, for any vector $\gamma \in \mathbb{R}^d$, symmetric positive-semidefinite $d \times d$ matrix A, and measure ν satisfying (3.3)*, there exists an infinitely divisible distribution μ whose characteristic exponent is given by* (3.2)*.*

Corollary 3.1: *For a real-valued Lévy process $X = (X_t)_{t\geq 0}$, the characteristic function is given by*

$$\phi_{X_t}(z) = \mathbb{E}[e^{\mathrm{i}zX_t}] = e^{t\psi(z)}, \;\; z \in \mathbb{R},$$

where

$$\psi(z) = -\frac{1}{2}az^2 + \mathrm{i}\gamma z + \int_{-\infty}^{\infty} \left(e^{\mathrm{i}zx} - 1 - \mathrm{i}zx\mathbb{1}_{\{|x|\leq 1\}}\right) \nu(dx),$$

for $a \geq 0, \gamma \in \mathbb{R}$, and a Lévy measure ν.

From the Lévy-Itō decomposition (3.1) we see that any Lévy process X is, in fact, a combination of a Brownian motion with drift and a sum of (possibly infinitely many) independent compound Poisson processes. The drift vector γ and the covariance matrix A of the Brownian motion describe the continuous part of the process. The jump component of X is described by the two other terms in (3.1), which are characterized by the Lévy measure ν of X: the first integral, describing the small jumps, can be interpreted as a superposition of independent compensated Poisson

processes, while the second integral, describing the large jumps of X, is a compound Poisson process. The number of jumps with an absolute value greater than one is finite, which we particularly know from (3.3). The triplet $[\gamma, A, \nu]$ consisting of the linear drift, the Gaussian covariance matrix and the Lévy measure, respectively, which uniquely determines the distribution of a Lévy process, is called *characteristic triplet* or *Lévy triplet* of X. If the Brownian component is zero, we call X a *Lévy jump process*, if the drift term is zero as well, X is called a *Lévy pure jump process*.

In addition to infinite divisibility, another property classifying probability distributions and, in further consequence, Lévy processes, is represented by the concept of *selfdecomposability*. Selfdecomposable distributions will play a crucial role when we investigate OU type SV models, due to the fact that the class of selfdecomposable distributions and the class of stationary distributions of OU processes (which are driven by general Lévy processes) coincide. However, let us proceed step by step.

Definition 3.6 (Selfdecomposability)**:** *Let μ be a probability measure on $\mathbb{R}$ and denote its characteristic function by ϕ_μ. Then μ is said to be* selfdecomposable *or to* belong to Lévy's class L*, if for all $t \in \mathbb{R}$ and all $c \in (0,1)$ there exists a probability measure μ_c on $\mathbb{R}$, with corresponding characteristic function ϕ_{μ_c}, such that*

$$\phi_\mu(t) = \phi_\mu(ct)\phi_{\mu_c}(t). \tag{3.4}$$

Remark: *Selfdecomposability in terms of random variables means that given a random variable X, for any $c \in (0,1)$ there exists a random variable X_c, independent of X, such that*

$$X \stackrel{d}{=} cX + X_c.$$

Remark: *A Lévy process corresponding to a selfdecomposable distribution is called a selfdecomposable process.*

Selfdecomposability is a stronger concept than infinite divisibility. More specifically, one important relation can be formulated as follows:

Proposition 3.2 (cf. [77], Prop. 15.5)**:** *All probability measures $\mu \in L$ are infinitely divisible, i.e., for any $n \geq 1$ there exists a characteristic function ϕ_n such that*

$$\phi(t) = (\phi_n(t))^n \ \forall t \in \mathbb{R}.$$

Furthermore, for any $c \in (0,1)$, μ_c in (3.4) is uniquely determined and infinitely divisible.

The converse statement, that is a necessary and sufficient condition for a probability measure to be selfdecomposable, is given in the next proposition.

Proposition 3.3: *A probability measure μ on $\mathbb{R}$ is selfdecomposable, if and only if it is infinitely divisible with Lévy triplet $[\gamma, A, \nu]$, where $\gamma \in \mathbb{R}$, $A \geq 0$, and the Lévy measure ν is of the form*

$$\nu(dx) = \frac{k(x)}{|x|}dx,$$

for a nonnegative function $k(x)$, which is increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$.

For a proof of Proposition 3.2 and Proposition 3.3, and a comprehensive discussion of the concept of selfdecomposability of probability measures in general, see the book of Sato [77]. Further interesting remarks on the class L and its relevance in various contexts are explained by Jurek [53]. One critical characteristic of the class L is given in the subsequent theorem, which is due to Jurek and Vervaat [54]: the relation between selfdecomposability and Lévy processes.

Theorem 3.3: *A random variable X has law in L if and only if X has a representation of the form*

$$X = \int_0^\infty e^{-t} dZ_t,$$

where $Z = (Z_t)_{t \geq 0}$ is a Lévy process.
In this case, the resp. Lévy measures ν and ρ of X and Z_1 are related by

$$\nu(dx) = \int_0^\infty \rho(e^t dx)dt.$$

We call Z the background driving Lévy process or, in short, the BDLP corresponding to X.

We now also formulate the above relation in terms of the corresponding Lévy densities (cf. [6]), since we will use it in this form at a later stage, that is in the context of constructing Ornstein-Uhlenbeck processes.

Proposition 3.4: *Suppose that the Lévy density f corresponding to ν is differentiable. Then the Lévy measure ρ has a density g, and f and g are related by*

$$g(x) = -f(x) - xf'(x). \tag{3.5}$$

Using the notation

$$\nu^+(x) := \nu([x,\infty)) = \int_x^\infty f(y)dy \tag{3.6}$$

for the upper tail integral of the Lévy density, in Barndorff-Nielsen [5] it is derived from the above formulae that

$$\nu^+(x) = xg(x).$$

The inverse function of ν^+, denoted by ν^{-1}, is given by

$$\nu^{-1}(x) = \inf\{y > 0 : \nu^+(y) \leq x\}. \tag{3.7}$$

3.1.2 Increasing Lévy processes: Subordinators

A Lévy process taking values only in the positive half-plane is called a *subordinator*. The non-negativeness implies that sample paths of such a process are increasing, as is formally stated in Proposition 3.5. The terminology refers to the use of this class of Lévy processes as random models for time evolution in order to time-change other (independent) Lévy processes; the theoretical basis of this technique is given in Theorem 3.4. A process modified in such a way is called *subordinate* to the original one.

Proposition 3.5 (cf. [31], Prop. 3.10)**:** *Let $X = (X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$. The following conditions are equivalent:*

(i) $X_t \geq 0$ *a.s. for some* $t > 0$.

(ii) $X_t \geq 0$ *a.s. for all* $t > 0$.

(iii) Sample paths of X are a.s. non-decreasing:

$$s \leq t \Rightarrow X_s \leq X_t \ \ a.s.$$

(iv) The process X has a non-negative drift, no diffusion component, and only positive jumps of finite variation. In other words, the characteristic triplet $[\gamma, A, \nu]$ of X satisfies $\gamma \geq 0$, $A = 0$, and $\nu((-\infty, 0]) = 0$, $\int_0^\infty (x \wedge 1)\nu(dx) < \infty$.

Theorem 3.4 (Subordination of a Lévy process; cf. [31], Thm. 4.2)**:** *Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $X = (X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic exponent $\psi(u)$, characterized by the triplet $[\gamma, A, \nu]$. Moreover, let $S = (S_t)_{t\geq 0}$ be a subordinator with Laplace exponent $l(u)$ and triplet $[b, 0, \rho]$. Then the process*

$Y = (Y_t)_{t\geq 0}$ defined by $Y_t(\omega) := X_{S_t(\omega)}(\omega)$, for all $\omega \in \Omega$, is again a Lévy process. Its characteristic function is given by

$$\phi_{Y_t}(u) = \mathbb{E}[e^{\mathrm{i}uY_t}] = e^{tl(\psi(u))}.$$

This means that the characteristic exponent of Y is obtained by composition of the Laplace exponent of S with the characteristic exponent of X. The characteristic triplet $[\gamma^Y, A^Y, \nu^Y]$ of Y is given by

$$\begin{aligned}
\gamma^Y &= b\gamma + \int_0^\infty \rho(ds) \int_{|x|\leq 1} x p_s^X(dx), \\
A^Y &= bA, \\
\nu^Y(B) &= b\nu(B) + \int_0^\infty p_s^X(B)\rho(ds), \quad \forall B \in \mathcal{B}(\mathbb{R}^d),
\end{aligned}$$

where p_t^X denotes the probability distribution of X.

In the book of Bertoin [18], a comprehensive chapter is dedicated to the theory of subordinators. We refer to this work for a thorough discussion of this class of Lévy processes from a theoretical viewpoint. Our special interest in subordinators is due to their importance as building blocks for Lévy-based models in mathematical finance. In particular, we will use them as driving processes for the Stochastic Volatility models considered in chapter 4.

3.1.3 Selected examples of Lévy processes

The two fundamental members of the class of Lévy processes are Brownian motion and the Poisson process. We have seen that any Lévy process can be represented as a superposition of a Brownian motion and (possibly infinitely many) Poisson processes. We now want to briefly present two examples of special Lévy processes, namely, the inverse Gaussian process and the normal inverse Gaussian process. We will consider these processes later on as driving processes in different models. In order to be able to study them, we first devote our attention to some theory about the occurring distributions in this context.

The inverse Gaussian distribution

The inverse Gaussian (IG) distribution is a two-parametric distribution family. The density function of the $IG(\delta, \gamma)$ distribution, for $\delta > 0, \gamma \geq 0$, is given by[1]

$$f_{\mathrm{IG}}(x; \delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta\gamma) x^{-3/2} \exp\left(-\frac{1}{2}(\delta^2 x^{-1} + b^2 x)\right), \quad x > 0.$$

It is a special case of the generalized inverse Gaussian distribution family $GIG(\lambda, \delta, \gamma)$, corresponding to $\lambda = -\frac{1}{2}$. Regarding the moments, the mean is given by δ/γ, the variance by δ/γ^3. The characteristic function of the $IG(\delta, \gamma)$ law is of the form

$$\phi_{\mathrm{IG}}(u) = \exp\left(-\delta\left(\sqrt{-2\mathrm{i}u + \gamma^2} - \gamma\right)\right). \tag{3.8}$$

A scaling property satisfied by the inverse Gaussian distribution is as follows: Let $X \sim IG(\delta, \gamma)$. Then, for $c > 0$, $cX \sim IG(\sqrt{c}\delta, \gamma/\sqrt{c})$.

The inverse Gaussian law is well known as the distribution of first passage times of Brownian motions. The particular relation (in our notation) can be formulated as follows: Let $W = (B_t + \gamma t)_{t \geq 0}$ be a standard Brownian motion with drift. Then the time $\tau^{(\delta,\gamma)} := \inf\{t > 0 : W_t = \delta\}$, i.e. the random time when W reaches the positive level δ for the first time, has an $IG(\delta, \gamma)$ distribution.

Literature on the inverse Gaussian distribution can, for instance, be found in terms of the book of Chhikara and Leroy Folks [28], or the work of Seshadri [80].

The normal inverse Gaussian distribution

The normal inverse Gaussian (NIG) distribution is defined as a normal variance-mean mixture taking the IG law as mixing distribution. In particular, let $\sigma^2 \sim IG(\delta, \sqrt{\alpha^2 - \beta^2})$ and $\epsilon \sim \mathrm{N}(0, 1)$, independent of σ^2. Let X be a random variable such that the conditional distribution of X given σ^2 is the normal distribution, with $\mathbb{E}[X|\sigma^2] = \mu + \beta\sigma^2$ and $\mathrm{Var}[X|\sigma^2] = \sigma^2$. We take $X = \mu + \beta\sigma^2 + \epsilon\sigma$. Then X is an $NIG(\alpha, \beta, \delta, \mu)$ distributed random variable and its density function is given by

$$f_{\mathrm{NIG}}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} \frac{K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\sqrt{\delta^2 + (x-\mu)^2}} \exp\{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\},$$

where K_λ denotes the modified Bessel function[2] of the third kind and order λ. The parameter assumptions are $0 \leq |\beta| \leq \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$. Apart from the

[1] An alternative parameterization (often) found in the literature, is given by $IG(\mu, \lambda)$ for $\mu = \delta/\gamma, \lambda = \delta^2$.

[2] For a comprehensive survey of Bessel functions and its properties, see Abramowitz and Stegun [1].

location-scale parameters μ and δ, the parameters α and β are responsible for the concrete rate of the decay of the tails ("tail-heaviness") and the degree of asymmetry, respectively. In general, the NIG distribution is skewed. Only the choice $\beta = 0$ corresponds to a symmetric density function around μ. The larger the value of $|\beta|$, the more distinct is the skewness. Regarding the asymptotic behaviour, we call the NIG distribution semi-heavy tailed; for $x \to \pm\infty$ and a constant c, the specific relation is given by

$$f_{\mathrm{NIG}}(x; \alpha, \beta, \delta, \mu) \sim c|x|^{-3/2} \exp(-\alpha|x| + \beta x).$$

In particular, the tails tend to zero much slower than it is the case for the normal distribution.

In fact, the NIG distribution is a special case of the class of generalized hyperbolic (GH) distributions. The density function of the GH law is defined by

$$\begin{aligned} f_{\mathrm{GH}}(x; \lambda, \alpha, \beta, \delta, \mu) = a(\lambda, \alpha, \beta, \delta) \left(\delta^2 + (x-\mu)^2\right)^{(\lambda - \frac{1}{2})/2} \cdot \\ \cdot K_{\lambda - 1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left(\beta(x-\mu)\right), \\ a(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda - 1/2}\delta^{\lambda} K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})}. \end{aligned}$$

Specifically, for $\lambda = -1/2$, we have

$$f_{\mathrm{NIG}}(x; \alpha, \beta, \delta, \mu) = f_{\mathrm{GH}}(x; -1/2, \alpha, \beta, \delta, \mu).$$

The dissertation of Prause [72] is an extensive work on generalized hyperbolic distributions. It comprises the analysis of the probability-theoretical properties, as well as approaches for parameter estimation and models based on the GH law in the context of financial derivative pricing and risk measures. We refer to the derivation of the moment generating function of the GH distribution carried out therein (Lemma 1.13). Using the well-known relations $K_{1/2}(z) = \sqrt{\pi/2} z^{-1/2} \exp(-z)$ and $K_{\lambda}(z) = K_{-\lambda}(z)$ for the modified Bessel function (listed, e.g., in Schoutens [78]), we then easily get the corresponding result for the NIG distribution: The moment generating function (mgf) of the NIG distribution is given by

$$\Phi(u; \alpha, \beta, \delta, \mu) = \exp\left(\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}\right) + \mu u\right).$$

Let $X^{(1)} \sim NIG(\alpha, \beta, \delta_1, \mu_1)$ and $X^{(2)} \sim NIG(\alpha, \beta, \delta_2, \mu_2)$ be two independent random variables, and define $X := X^{(1)} + X^{(2)}$. Then

$$\begin{aligned}\Phi_X(u) &= \mathbb{E}[e^{uX}] = \left(\mathbb{E}[e^{uX^{(1)}}]\right)\left(\mathbb{E}[e^{uX^{(2)}}]\right) = \left(\Phi_{X^{(1)}}(u)\right)\left(\Phi_{X^{(2)}}(u)\right) \\ &= e^{\left(\delta_1(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})+\mu_1 u\right)} e^{\left(\delta_2(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})+\mu_2 u\right)} \\ &= \exp\left((\delta_1+\delta_2)\left(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2}\right)+(\mu_1+\mu_2)u\right),\end{aligned}$$

which corresponds to the mgf of an NIG$(\alpha, \beta, \delta_1+\delta_2, \mu_1+\mu_2)$ distribution. The mgf characterizes a distribution uniquely. Therefore, we may conclude that the NIG distribution family is closed under convolution; in particular,

$$NIG(\alpha, \beta, \delta_1, \mu_1) * NIG(\alpha, \beta, \delta_2, \mu_2) = NIG(\alpha, \beta, \delta_1+\delta_2, \mu_1+\mu_2). \tag{3.9}$$

The inverse Gaussian process

The inverse Gaussian process with parameters δ and γ is defined as the process $X = (X_t)_{t\geq 0}$ such that the increments are independent and inverse Gaussian distributed; in particular, $X_{t+h} - X_t \sim IG(\delta h, \gamma)$. Equivalently, we can define the process "directly" as a "first-hitting-time process" by

$$X_t := \inf\{s > 0 : B_s + \gamma s \geq \delta t\},$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion. The $IG(\delta, \gamma)$ process is a subordinator; its Lévy triplet is given by $[b, 0, \nu_{\mathrm{IG}}]$, with drift component

$$b = \frac{\delta}{\gamma}(2N(\gamma) - 1),$$

where $N(\cdot)$ denotes the Normal distribution function, and the Lévy measure

$$\nu_{\mathrm{IG}}(dx) = (2\pi)^{-1/2}\delta x^{-3/2}\exp\left(-\frac{1}{2}\gamma^2 x\right)\mathbb{1}_{\{x>0\}}dx.$$

The normal inverse Gaussian process

A normal inverse Gaussian process $X = (X_t)_{t\geq 0}$ is a Lévy process with NIG distributed increments. Property (3.9) of the NIG distribution implies that for an NIG process X it holds

$$X_{t+s} - X_t \sim NIG(\alpha, \beta, \delta s, \mu s).$$

Hence, in particular, X_t has an $NIG(\alpha, \beta, \delta t, \mu t)$ distribution.

An NIG process can be obtained by using an IG subordinator for the time evolution of a Brownian motion. Therefore, it is quite easy to simulate such processes (cf. [31], Algorithm 6.12).

The NIG process is a pure jump process. In Barndorff-Nielsen [4], the Lévy-Khintchine formula is derived; the Lévy triplet is given by $[\gamma, 0, \nu]$, where

$$\gamma = \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) dx,$$
$$\nu(dx) = \frac{\delta\alpha}{\pi} \frac{\exp(\beta x) K_1(\alpha|x|)}{|x|} dx.$$

From a theoretical point of view, a detailed discussion of the NIG distribution and NIG processes can be found in the work of Barndorff-Nielsen [4, 5], where these processes were also introduced originally. Numerous authors have discovered the potential of the NIG law to fit financial data and suggest its application for various modelling approaches; amongst others, Lillestøl [59] has used the NIG distribution in the context of risk analysis, Albrecher and Predota [2] have proposed NIG processes to price Asian Options and Asmussen et al. [74] have studied the pricing of further exotic options on the basis of NIG models.

3.2 Processes of Ornstein-Uhlenbeck type

3.2.1 Gaussian Ornstein-Uhlenbeck processes

The Gaussian Ornstein-Uhlenbeck (OU) process is the unique (up to indistinguishability) solution of the SDE

$$\begin{cases} dX_t &= \kappa(\theta - X_t)dt + \sigma dB_t, \\ X_0 &= x_0. \end{cases}$$

In order to calculate the explicit solution, we first set $Y_t = X_t - \theta$. This does not have an impact on the differential, since $dY_t = dX_t$, but gives us a simplified expression for the SDE:

$$\begin{cases} dY_t &= -\kappa Y_t dt + \sigma dB_t, \\ Y_0 &= y_0. \end{cases}$$

As a next step, we set $Z_t := e^{\kappa t} Y_t$. Using the product rule then gives

$$\begin{aligned} dZ_t &= \kappa e^{\kappa t} Y_t dt + e^{\kappa t} dY_t + d(e^{\kappa t}) dY_t \\ &= e^{\kappa t}(\kappa Y_t dt - \kappa Y_t dt + \sigma dB_t) \\ &= \sigma e^{\kappa t} dB_t, \end{aligned}$$

and integration yields

$$Z_t - Z_0 = \sigma \int_0^t e^{\kappa s} dB_s.$$

Changing back the substituted variables gives

$$Y_t = e^{-\kappa t} Y_0 + \sigma e^{-\kappa t} \int_0^t e^{\kappa s} dB_s,$$

and finally

$$X_t = Y_t + \theta = \theta + e^{-\kappa t}(X_0 - \theta) + \sigma e^{-\kappa t} \int_0^t e^{\kappa s} dB_s.$$

This representation of the process as an Itō-integral with respect to a Brownian motion reveals that $X = (X_t)_{t \geq 0}$ has continuous trajectories. Moreover, X is a Gaussian process, i.e., for any finite set of points in time $t_1, ..., t_n$, the vector $(X_{t_1}, ..., X_{t_n})$ has a multivariate normal distribution. Because of these features and its mean-reversion property, the Gaussian OU process has proved attractive for various applications. One famous example is the approach chosen by Vasicek [84] in order to model the term structure of interest rates.

However, there are many fields - including volatility modelling - where the behaviour of the Gaussian OU process does not make it a convenient modelling-instrument. BM is the only Lévy process having continuous sample paths (a.s.). Therefore, we now want to generalize the "classical" version of an OU process studied above, by substituting the BM as driving noise by a general Lévy process. This approach will, in further consequence, enable us to prescribe certain characteristics of the process, such as positivity or the marginal distribution.

3.2.2 General Ornstein-Uhlenbeck processes

Definition 3.7 (Ornstein-Uhlenbeck process)**:** *Let $Y = (Y_t)_{t\geq 0}$ be the solution of the SDE*

$$\begin{cases} dY_t &= -\lambda Y_t dt + dZ_t, \\ Y_0 &= y_0, \end{cases} \tag{3.10}$$

where $Z = (Z_t)_{t\geq 0}$ is a Lévy process. Then Y is called an OU process driven by Z; in turn, Z is termed the Background Driving Lévy process (BDLP) corresponding to the process Y.

Proceeding with the same strategy as above in the Gaussian case, we obtain the explicit solution

$$Y_t = e^{-\lambda t} Y_0 + e^{-\lambda t} \int_0^t e^{\lambda s} dZ_s. \tag{3.11}$$

Having the intention to model volatility, one's particular interest is in stochastic processes having sample paths with values in the positive half-plane. From (3.11) it is clear that Y_t is (a.s.) a strictly positive process, given that we choose $y_0 > 0$ and take a subordinator for Z. That is exactly the approach of the Barndorff-Nielsen and Shephard model, which will be investigated in detail in chapter 4.

The following two propositions deal with important distributional properties of OU processes.

Proposition 3.6 (cf. [31], Prop. 15.1)**:** *Let $L = (L_t)_{t\geq 0}$ be a Lévy process with characteristic triplet $[\gamma, A, \nu]$. The distribution of Y_t, defined by equation* (3.10)*, is infinitely divisible for every t and has characteristic triplet $[\gamma_t^Y, A_t^Y, \nu_t^Y]$ with*

$$\begin{aligned} A_t^Y &= \frac{A}{2\lambda}\left(1 - e^{-2\lambda t}\right), \\ \gamma_t^Y &= \frac{\gamma}{\lambda}\left(1 - e^{-\lambda t}\right) + Y_0 e^{-\lambda t}, \\ \nu_t^Y(B) &= \int_1^{e^{\lambda t}} \nu(\xi B)\frac{d\xi}{\lambda \xi} \quad \forall B \in \mathcal{B}(\mathbb{R}), \end{aligned}$$

where ξB is a shorthand notation for $\{\xi x : x \in B\}$.

Proposition 3.7 (cf. [31], Prop. 15.4)**:** *Let $L = (L_t)_{t\geq 0}$ be a Lévy process with*

characteristic triplet $[\gamma, A, \nu]$. *If*

$$\mathbb{E}\left[\log(1+|L_1|)\right] = \int_{|x|\geq 1} \log|x|\nu(dx) < \infty \tag{3.12}$$

then the OU process Y *defined in* (3.10) *has a stationary distribution* μ *which is selfdecomposable, has characteristic exponent*

$$\psi^Y(u) = \lim_{t\to+\infty} \psi_t^Y(u) = \lim_{t\to+\infty} \int_0^t \psi(ue^{\lambda(s-t)})ds,$$

and Lévy triplet $[\gamma^Y, A^Y, \nu^Y]$ *with* $\gamma^Y = \frac{\gamma}{\lambda}$, $A^Y = \frac{A}{2\lambda}$, *and*

$$\nu^Y(B) = \int_1^\infty \nu\left(\xi B\right) \frac{d\xi}{\lambda\xi} \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Conversely, for every selfdecomposable distribution μ *there exists a Lévy process* L *such that* μ *is the stationary distribution of the OU process driven by* L.

In other words, Proposition 3.7 contains two important informations: First, (3.12) is a necessary and sufficient condition for the defining SDE of the OU process to have a stationary solution. Second, Lévy's class L of all selfdecomposable probability measures can, in this context, also be seen as the class of all stationary distributions of OU processes. If we specify an arbitrary selfdecomposable law μ, we can hence always construct an OU process having stationary distribution μ.

3.3 Market modelling

3.3.1 The Black-Scholes model

By far the most famous market model, is the model suggested by Black and Scholes [20] and Merton [64]. For the formulation of this model and, in particular, the derivation of the associated Black-Scholes formulae for the pricing of European Call and Put Options, Merton and Scholes also received the Nobel Prize for Economic Sciences in 1997 (Black had passed away in 1995) [82]. Their model consists of a risk-free bond with price process $S^0 = (S_t^0)_{0\leq t\leq T}$, where $S_t^0 = e^{rt}$ for a riskless interest rate r, and a risky asset whose price dynamics are modelled as

$$dS_t = S_t(\mu dt + \sigma dB_t), \tag{3.13}$$

with a deterministic drift $\mu \in \mathbb{R}$, volatility of the asset price $\sigma > 0$, a given initial asset price $S_0 > 0$, and where $B = (B_t)_{0\leq t\leq T}$ is a standard Brownian motion. The

explicit solution of the stochastic differential equation (3.13) is given by

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t\right), \qquad 0 \leq t \leq T.$$

There is most probably no textbook on mathematical finance, which does not contain a detailed chapter about the Black-Scholes model. We are hence not going into any further detail about the properties of this model. We only want to sketch its very well-known shortcomings, which have motivated the development of the market models discussed in the sequel. Logarithmic returns in the Black-Scholes model are normally-distributed; the distribution of real return data, however, shows a certain skewness, as well as tails that are much heavier than those of a Gaussian distribution. Moreover, the model cannot describe the so-called *leverage-effect*; this term refers to the empirically observed phenomenon that large downward movements in asset prices coincide with upward moves in volatility. Considering the implied volatility curve as a function in strike (that value σ, which is the unique solution resulting from equating the Black-Scholes price of a liquidly-traded vanilla option with its market price, is called "implied volatility"), one observes *volatility smiles* and *smirks*, that is implied volatility is not constant as a function in all other parameters, which should be the case, however, if the model described the market correctly. The observation that there are periods of high volatiliy alternating with periods of low volatility, or, as Mandelbrot [61] put it, "large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes", is referred to as *volatility clustering*. The Black-Scholes model, having a constant volatility parameter, naturally does not offer this feature either.

Considering the just explained examples of so-called *stylized facts* (i.e., empirical properties of asset returns), which the Black-Scholes model cannot properly cope with, the model has gradually been altered in various ways. One approach has been to add jumps to the asset price process. The resulting *Jump-Diffusion models*, however, are still not capable of dealing with, for instance, the leverage effect or volatility clustering. The point is that in the Black-Scholes model, as well as in these Jump-Diffusion models of the type[3]

$$dS_t = S_{t-}(\mu dt + \sigma dB_t + dJ_t), \tag{3.14}$$

with a Brownian motion B and a Compound Poisson process J (independent of B), such as, e.g., the models of Merton (see [65]) or Kou (see [58]), the volatility σ of the asset price process $S = (S_t)_{t\geq 0}$ is assumed to be constant. Now we move on

[3]The notation S_{t-} denotes the left-sided limit at time t.

to models where a stochastic process $\sigma = (\sigma_t)_{t\geq 0}$ is introduced in order to model the diffusion coefficient itself as a (positive) stochastic process. The process σ can then be interpreted as the instantaneous volatility of the underlying asset, i.e., σ_t represents the volatility of the returns over the infinitesimal interval $(t, t + dt)$.

3.3.2 Stochastic Volatility models

A general (solely diffusion-based) Stochastic Volatility model is of the form

$$dS_t = S_t(\mu dt + \sigma_t dB_t), \tag{3.15}$$

and is determined by the specific choice of the process σ. Note, however, that σ is latent; only the price process S can be observed on the market, volatility itself is not empirically observable.

The Heston model

The most common diffusion-based Stochastic Volatility model is the model introduced by Heston [45], where σ is chosen as a Cox-Ingersoll-Ross (CIR) process. This choice guarantees for positiveness and entails the feature of mean-reversion, which particularly ensures that the model-volatility does not tend to infinity in the long run. The CIR process (also called square root process) was introduced in Cox et al. [33] in the context of the study of term structures of interest rates. They actually modelled the short rate via these dynamics. The CIR process $Y = (Y_t)_{t\geq 0}$ is defined as satisfying the SDE

$$\begin{cases} dY_t &= \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dB_t, \\ Y_0 &= y_0, \end{cases} \tag{3.16}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, $\sigma > 0$, $\kappa > 0$, $\theta > 0$, and $y_0 > 0$. Under these positivity-assumptions for all parameters and a positive starting value, the process stays nonnegative. It can be shown that Y almost surely never touches zero, if the parameters are chosen such that $2\kappa\theta \geq \sigma^2$. A main characteristic of the CIR process is the mean-reverting drift term. If $Y_t > \theta$ it pulls the process downwards, whereas if $Y_t < \theta$ we have a positive drift. Consequently, the solution will oscillate around the value of θ. We therefore interpret θ as the long-term mean and κ as the rate of the mean-reversion.

The model of Heston defines the dynamics of the price process as

$$dS_t = S_t(\mu dt + \sqrt{V_t} dB_t), \tag{3.17}$$

where the volatility process satisfies the SDE

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dB_t^V. \tag{3.18}$$

As already mentioned before, we interpret V as the instantaneous variance process. In general, the Brownian motion B driving the asset price process and the BM B^V driving the volatility process, are correlated in the sense of the relation

$$B_t^V = \rho B_t + \sqrt{1-\rho^2} W_t,$$

where W is a BM independent of B. We can see $\rho \in [-1, 1]$ as instantaneous correlation coefficient, since the differential of the quadratic covariation is given by

$$\begin{aligned} d[B, B^V]_t &= d[B, \rho B + \sqrt{1-\rho^2} W]_t = d([B, \rho B]_t + [B, \sqrt{1-\rho^2} W]_t) \\ &= d(\rho [B, B]_t + \sqrt{1-\rho^2}[B, W]_t) = \rho dt. \end{aligned}$$

Choosing a negative value for ρ incorporates a leverage-effect in the model. Moreover, the model accounts for volatility clustering and long-term volatility smiles and skews. However, the short-term behaviour of implied volatiliy has turned out not to be modelled in a realistic way. This issue has been tackled by additionally introducing jumps in stochastic volatiliy models.

The Bates model

Bates [13] has proposed to add an independendent jump component to the price-dynamics of the Heston model. In particular, the Bates model in its original formulation is defined via

$$\begin{aligned} dS_t &= S_t(\mu dt + \sqrt{V_t} dB_t + dJ_t), \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dB_t^V, \end{aligned}$$

where $J = (J_t)_{0 \le t \le T}$ is a compound Poisson process with log-normally distributed jump sizes (anything else defined as in the model of Heston). The jump-size distribution can also be varied. In any case, the Bates model does have the capability to generate appropriate implied-volatility patterns even for short maturities.

An alternative approach to combine stochastic volatility modelling and jumps, is to introduce jumps in the evolution of volatility. Positive Ornstein-Uhlenbeck processes, i.e. OU processes driven by subordinators, have been considered to model volatility in this context. In this way, volatility and jumps are not independent in the model, but rather combined in a non-trivial way. Hence, the model gets more sophisticated offering a more realistic structure, especially from a qualitative point of view. In particular, it provides the possibility to make an a priori choice of the stationary distribution of the volatility process. Moreover, such models turn out to be of very tractable nature and allow many explicit calculations. We are now going to examine models of this type in detail in the subsequent chapter.

Chapter 4

Ornstein-Uhlenbeck Type Stochastic Volatility Models

4.1 The Barndorff-Nielsen and Shephard Model

The Barndorff-Nielsen and Shephard model (henceforth BNS model) is a continuous time market model consisting of a risk-free asset S^0 with price process $S_t^0 = e^{rt}$, where r denotes the constant risk-free interest rate, and a stock $S = (S_t)_{0\leq t\leq T}$, with $S_t = e^{X_t}$ and where the logarithmic asset price process $X = (X_t)_{0\leq t\leq T}$ satisfies

$$dX_t = (\mu + \beta\sigma_t^2)dt + \sigma_t dW_t + \rho dZ_{\lambda t}. \tag{4.1}$$

The volatility process σ^2 is defined as a process of Ornstein-Uhlenbeck type, i.e., σ_t^2 satisfies

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}, \quad \sigma_0^2 > 0. \tag{4.2}$$

For the parameters we assume $\mu, \beta \in \mathbb{R}, \lambda > 0$ and $\rho \leq 0$. $W = (W_t)_{0\leq t\leq T}$ is a standard Brownian motion and $Z = (Z_{\lambda t})_{0\leq t\leq T}$ is a Lévy subordinator, independent from W. We assume further that Z has no deterministic drift term. We denote the density of its Lévy measure by $\mathrm{w}(x)$. The cumulant transform $\kappa(\theta) = \log \mathbb{E}[e^{\theta Z_1}]$ is then of the form

$$\kappa(\theta) = \int_{\mathbb{R}^+} (e^{\theta x} - 1)\mathrm{w}(x)dx. \tag{4.3}$$

The process S lives on a probabibility space $(\Omega, \mathcal{A}, \mathbb{P})$, with a filtration $\mathcal{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ satisfying the usual hypotheses. In particular, we take $\mathcal{F}$ to be the usual augmentation of the filtration generated by the pair (W, Z).

We obtain the dynamics of the asset price process simply by application of Itō's Lemma. Setting $f(x) = e^x$ yields

$$dS_t = S_{t-}\left(dX_t + \frac{1}{2}d[X]_t\right).$$

The differential of the quadratic variation process of X is given by

$$\begin{aligned} d[X]_t =&(\mu + \beta\sigma_t^2)^2 dtdt + \sigma_t^2 d[W,W]_t + \rho^2 d[Z,Z]_{\lambda t} \\ &+ 2\sigma_t(\mu + \beta\sigma_t^2)dtdW_t + 2\rho(\mu + \beta\sigma_t^2)dtdZ_{\lambda t} + 2\rho\sigma_t dW_t dZ_{\lambda t} \\ =&\sigma_t^2 dt + \rho^2 d[Z,Z]_{\lambda t}. \end{aligned}$$

Hence, we have

$$\begin{aligned} dX_t + \frac{1}{2}d[X]_t &= \left(\mu + \left(\beta + \tfrac{1}{2}\right)\sigma_t^2\right)dt + \sigma_t dW_t + \rho dZ_{\lambda t} + \tfrac{1}{2}\rho^2 d[Z,Z]_{\lambda t} \\ &= \left(\mu + \left(\beta + \tfrac{1}{2}\right)\sigma_t^2\right)dt + \sigma_t dW_t + (e^{\rho dZ_{\lambda t}} - 1), \end{aligned}$$

where the second step follows from looking at the terms $\rho dZ_{\lambda t} + \frac{1}{2}\rho^2(dZ_{\lambda t})^2$ as the first two elements of the power series representation of $e^{\rho dZ_{\lambda t}}$, subtracted by the element $e^0 = 1$. This works, since the product of three or more differential forms of a Lévy process, i.e. $(dZ_{\lambda t})^n$ for $n \geq 3$, equals zero due to the stochastic continuity of the process. Expanding the last expression by $\pm\lambda\kappa(\rho)$ yields the representation

$$dX_t + \frac{1}{2}d[X]_t = \underbrace{\left(\mu + \lambda\kappa(\rho) + \left(\beta + \tfrac{1}{2}\right)\sigma_t^2\right)}_{=:b_t}dt + \underbrace{(e^{\rho dZ_{\lambda t}} - 1 - \lambda\kappa(\rho)dt)}_{=:dM_t} + \sigma_t dW_t.$$

Now we can write the dynamics of the asset price in the BNS model in the form

$$dS_t = S_{t-}(b_t dt + \sigma_t dW_t + dM_t).$$

Since Z is a subordinator, i.e., increasing only in terms of jumps, for $M = (M_t)_{0\leq t\leq T}$ we get the representation

$$M_t = \sum_{0\leq s\leq t}(e^{\rho\Delta Z_{\lambda s}} - 1) - \lambda\kappa(\rho)t. \tag{4.4}$$

Moreover, $M = (M_t)_{0\leq t\leq T}$ is a martingale: By application of the Lévy-Khintchin formula (given in (3.2)) and the power series representation of the exponential func-

tion, we get

$$\begin{aligned}\mathbb{E}[e^{\rho dZ_{\lambda t}}] &= e^{d(\lambda t)\int_{\mathbb{R}^+}(e^{\rho x}-1)\mathrm{w}(x)dx} \\ &= \sum_{n=0}^{\infty} \frac{(\lambda dt \cdot \kappa(\rho))^n}{n!} \\ &= 1 + \lambda\kappa(\rho)dt,\end{aligned}$$

since $(dt)^n = 0$ for $n \geq 2$. Hence, $\mathbb{E}[dM_t] = 0$.

The term $\beta\sigma_t^2$ in the drift of the log-price, i.e., the introduction of dependence between the drift of the asset price and the volatility, is due to the fact that in practice investors usually require a premium for the risk they take in comparison to alternatively invest in riskless assets. Therefore, we want the drift to depend on the volatility level.

Most authors choose the modified timing λt for the BDLP of the volatility process. The reason for this choice is the intention to decouple the marginal distribution and the autocorrelation structure: the marginal distribution of σ_t^2 is not affected by the particular value of λ.

By means of the term $\rho Z_{\lambda t}$ in the log-price process, a leverage-effect is incorporated in the model. Without this term, i.e., in the case where $\rho = 0$, there are jumps in the volatility, the sample paths of the asset price are continuous, though. Since Z is a subordinator, i.e., only exhibits positive jumps, for $\rho \neq 0$ (recall our general assumption $\rho \leq 0$), each jump in the volatility coincides with a downwards jump of the log-asset price of proportional size. The capability of a market model to account for this stylized-fact is a highly appreciated feature. Extensive investigations of the financial markets' behaviour have led to a broad consensus in the literature that the arrival of some kind of negative information on the market triggers both, a fall in stock prices as well as an increase in volatility (see, e.g., Nelson [67], Cont [30], or Bouchaud et al. [22]). Therefore, in the sequel, we will always assume $\rho < 0$. However, it should not be left unmentioned that in the work of Benth [14] the BNS model in the no-leverage case is used as a part of their modelling approach for commodity spot prices. They propose this model and outline the feasibility of their choice in this context with an empirical study in terms of UK gas spot prices.

A possible extension of the model described above is to use a superposition of OU processes for the volatility. In particular, for n independent OU processes σ_j^2

and weights w_j such that $\sum_{j=1}^{n} w_j = 1$, we define the volatility process as

$$\sigma^2(t) = \sum_{j=1}^{n} w_j \sigma_j^2(t).$$

The BDLPs of the different processes σ_j^2 are not necessarily identically distributed. Theoretical aspects of superpositions of OU processes are extensively studied in [9], whereas the idea of adopting those results in the BNS model is considered in [6] and [7]. The intention behind this idea is to gain more flexibility in the modelling of the volatility, since in this way one could use some components σ_j^2 to represent the variation over short periods, whereas other components could account for long-term movements.

4.1.1 Moment generating function

Theorem 4.1: *The moment generating function of the log-asset price process under the BNS model as defined by* (4.1) *and* (4.2) *is given by*

$$\Phi_{X_t}(z) = \exp\left(z(X_0 + \mu t) + \frac{(z^2 + 2\beta z)(1 - e^{-\lambda t})\sigma_t^2}{2\lambda} + \lambda \int_0^t \kappa(f(s,z))ds \right),$$

where f is defined as the deterministic function

$$f(s,z) = \rho z + \frac{(z^2 + 2\beta z)(1 - e^{-\lambda(t-s)})}{2\lambda}.$$

For a proof of this result see [68, pg. 450], where the more general conditional Laplace transform at time T given the information at time $0 \leq t \leq T$, i.e., $\mathbb{E}[e^{zX_T}|\mathcal{F}_t]$ for $z \in \mathbb{C}$, together with the corresponding strip of regularity, is derived. The salient point of this derivation is the application of the so-called *key formula*, which represents the result that for a subordinator Z with cumulant transform κ and a left-continuos function $f : \mathbb{R}^+ \to \mathbb{C}$ such that $\mathrm{Re}(f) \leq 0$, it holds that

$$\mathbb{E}\left[\exp\left(\int_0^t f(s) dZ_{\lambda s} \right) \right] = \exp\left(\lambda \int_0^t \kappa(f(s))ds \right).$$

4.1.2 Equivalent martingale measures

In the original work of Barndorff-Nielsen and Shephard [7], it is shown that the market model described above does not allow arbitrage. Using the Esscher transformation technique, the existence of an equivalent martingale measure (EMM) under which the asset price process is a local martingale is verified. The model

specification, in particular, the combination of jumps and stochastic components, however brings along the matter of an incomplete market model. In other words, there exist infinitely many EMMs, under which the BNS model is free of arbitrage. Therefore, we want to present the result of [7] in terms of the more general results, which are due to Nicolato and Venardos [68], who derive the entire set of EMMs. Additionally, they describe subclasses of this set, whose members preserve the structure of the model; that is under such an EMM the log-returns are again described by a BNS model, even if the parameters are possibly different and stationary distributions might change. For the proofs we refer to their work. The notation in Theorem 4.2 is adopted from Jacod and Shiryaev [52] and is to be understood in the sense that on the basis of expression (4.4) we write

$$\begin{aligned} M_t &= \int_{[0,t]\times\mathbb{R}} (e^{\rho x} - 1)(\mu_Z - \nu_Z)(ds, dx) \\ &= (e^{\rho x} - 1) \star (\mu_Z - \nu_Z)_t, \end{aligned}$$

where μ_Z is the random measure associated with the jumps of Z and

$$\nu_Z(\omega, dt, dx) = \lambda \mathrm{w}(x) dx dt$$

is its compensator. Further, for a semimartingale X and a predictable X-integrable process H, the stochastic integral is denoted by $H \bullet X$.

The first theorem characterizes the set of equivalent martingale measures for the BNS model, which we denote by $\mathcal{M}$.

Theorem 4.2 ([68], pg. 452)**:** *Let $\mathbb{Q} \in \mathcal{M}$. Then the density process $L_t = \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t}$ is given by the Doléans-Dade exponential process*

$$\begin{aligned} L_t &= \mathcal{E}\left(\psi \bullet W + (Y-1) \star (\mu_Z - \nu_Z)\right)_t \\ &= \mathcal{E}\left(\int_0^t \psi_s dW_s + \int_0^t \int_{\mathbb{R}^+} (Y(s,x) - 1)(\mu_Z - \nu_Z)(dx, ds)\right), \end{aligned} \tag{4.5}$$

where $\psi = (\psi_t)_{0 \le t \le T}$ is a predictable process and $Y = Y(\omega, t, x)$ is a strictly positive predictable function such that

$$\int_0^t ds \int_{\mathbb{R}^+} \left(\sqrt{Y(s,x)} - 1\right)^2 \mathrm{w}(x) dx < +\infty, \quad \mathbb{P} - a.s.$$

The function Y and the process ψ are linked by

$$\mu + \left(\beta + \tfrac{1}{2}\right)\sigma_t^2 + \sigma_t \psi_t + \lambda \int_{\mathbb{R}^+} Y(t,x)(e^{\rho x} - 1)\mathrm{w}(x) dx - r = 0,$$

$d\mathbb{P} \otimes dt$ almost surely.

An interpretation of (4.5) could be given in the sense that ψ and $Y-1$ can be seen as the risk-premia associated with the diffusion component and the jump component of the model, respectively. For any $\mathbb{Q} \in \mathcal{M}$, the discounted price process is a $\mathbb{Q}$-martingale and the model under $\mathbb{Q}$ is hence free of arbitrage. However, under $\mathbb{Q}$, the BDLP Z is neither necessarily again a Lévy process, nor does it have to be independent of $W^{\mathbb{Q}}$. Therefore, the next theorem characterizes the subset of $\mathcal{M}$, which contains those EMMs, under which the structure of the BNS model is preserved. We denote this set by $\mathcal{M}'$ and use the following definitions:

$$\mathcal{Y}' := \left\{ y : \mathbb{R}^+ \to \mathbb{R}^+ \,\middle|\, \int_{\mathbb{R}^+} (\sqrt{y(x)} - 1)^2 \mathrm{w}(x) dx < +\infty \right\},$$

and for $y \in \mathcal{Y}'$ we set

$$\mathrm{w}^y(x) = y(x)\mathrm{w}(x). \tag{4.6}$$

We further define

$$\kappa^y(\theta) = \int_{\mathbb{R}^+} (e^{\theta x} - 1)\mathrm{w}^y(x) dx \quad \text{for } \mathrm{Re}(\theta) < 0, \tag{4.7}$$

which exists, since from the definitions above it follows that $\int_{\mathbb{R}^+} (1 \wedge x)\mathrm{w}^y(x) dx < \infty$.

Theorem 4.3 ([68], pg. 453)**:** *Let $y \in \mathcal{Y}'$. Then the process*

$$\psi_t = \sigma_t^{-1} \left(r - \mu - \left(\beta + \frac{1}{2} \right) \sigma_t^2 - \lambda \kappa^y(\rho) \right),$$

where κ^y is given in (4.7)*, is a.s. square integrable, i.e.,*

$$\mathbb{P}\left(\int_0^T \psi_s^2 ds < \infty \right) = 1$$

and $L^y = (L_t^y)_{0 \le t \le T}$, where

$$L_t^y = \mathcal{E}(\psi \bullet W + (y-1) \star (\mu_Z - \nu_Z))_t,$$

is a density process. The probability measure defined by

$$d\mathbb{Q}^y = L_T^y d\mathbb{P}$$

is an EMM and the dynamics of X under $\mathbb{Q}^y$ are given by

$$dX_t = \left(r - \lambda\kappa^y(\rho) - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t^y + \rho dZ_{\lambda t} \tag{4.8}$$

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}, \tag{4.9}$$

where $W_t^y = W_t - \int_0^t \psi_s ds$ is a $\mathbb{Q}^y$-Brownian motion and $Z_{\lambda t}$ is a $\mathbb{Q}^y$-Lévy process. Z_1 has Lévy density $w^y(x)$ and cumulant transform $\kappa^y(\theta)$ given by (4.6) *and* (4.7) *respectively, and the processes W^y and Z are independent under $\mathbb{Q}^y$. Hence, $\mathbb{Q}^y \in \mathcal{M}'$.*
Conversely, for any $\mathbb{Q} \in \mathcal{M}'$, there exists a function $y \in \mathcal{Y}'$ such that $\mathbb{Q}$ coincides with $\mathbb{Q}^y$.

An investigation of the class of EMMs for the BNS model, in the sense of a discussion of several optimal martingale measures, is carried out by Hubalek and Sgarra [47]. Specifically, they apply the Esscher martingale transform on some popular concrete specifications, such as, e.g., the Γ-OU or the IG-OU BNS model. We want to mention a particular result of their work, in order to exemplarily emphasize the simplifications resulting from the model where the BDLP does not appear in the dynamics of the asset-price: They show that in the no-leverage case, existence of the resp. measures assumed, the minimal martingale measure[1] coincides with both the Esscher martingale transform for the exponential process e^X as well as the Esscher martingale transform for the linear process $\tilde{X}$, where $S = S_0 e^X = S_0\mathcal{E}(\tilde{X})$. In the general case, all these measures are different.

4.1.3 Popular specifications

Strictly speaking, *the* BNS model corresponds to a class of models, where the actual model is characterized by the particular choice of the BDLP. We now want to present some concrete examples, which are prevalently used in the literature.

In terms of constructing the model, Proposition 3.7 shows that there are two different ways we can pursue: First, we can specify the model by postulating the characteristics of the BDLP. The second possibility is to choose a parametric form for the one-dimensional marginal distribution of the OU process and work out the characteristics of the BDLP implicitly specified thereby. This approach is, e.g, described in [6] for the generalized inverse Gaussian distribution family, where the inverse Gaussian, the positive hyperbolic, the reciprocal gamma, or the gamma

[1] The minimal martingale measure $\hat{\mathbb{P}}$ is defined as the unique equivalent local martingale measure, for a continuous adapted process X, with the property that local $\mathbb{P}$-martingales strongly orthogonal to the $\mathbb{P}$-martingale part of X are also local $\hat{\mathbb{P}}$-martingales (cf., e.g., Schweizer [79]).

distribution are prominent special cases thereof.

In case that the $IG(\delta, \gamma)$-law is chosen for $\sigma^2(t)$, the upper tail integral $\nu^+(x)$ (as defined in (3.6)) is given by

$$\nu^+(x) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{1}{2}\gamma^2 x}.$$

The BDLP of an IG-OU process can be represented as the sum of two independent Lévy processes. In particular, we can write $Z = Z^{(1)} + Z^{(2)}$, where $Z^{(1)}$ is an $IG(\frac{\delta}{2}, \gamma)$-process, and $Z^{(2)}$ is the compound Poisson process

$$Z_t^{(2)} = \frac{1}{\gamma} \sum_{n=1}^{N_t} x_n^2, \tag{4.10}$$

with a Poisson process $N = (N_s)_{s \geq 0}$ having intensity $\delta\gamma/2$ and a sequence $(x_n)_{n \geq 0}$ of independent standard normally distributed random variables, which are independent of N_t. Because of $Z^{(1)}$ being an IG-process, also Z is not a process of finite activity. The Lévy measure of the positive real axis ($\nu^+(0)$ in our notation) equals $+\infty$ and the IG-OU process hence jumps infinitely often in every finite time interval. The IG-OU BNS model without leverage, i.e., the BNS model where the stationary distribution of the variance process is the inverse Gaussian distribution and $\rho = 0$, is particularly interesting in that the log-returns in this model are approximately normal inverse Gaussian distributed (cf. [68]). We mentioned the ability of this distribution to fit financial return data, which has been observed in the literature, in section 3.1.3.

For the $\Gamma(\beta, \alpha)$-law as marginal distribution of the OU process, the corresponding tail integral is given by

$$\nu^+(x) = \beta e^{-\alpha x},$$

where the inverse function can be computed anlytically and is given by

$$\nu^{-1}(x) = \max\left\{0, -\frac{1}{\alpha} \log\left(\frac{x}{\beta}\right)\right\}. \tag{4.11}$$

The BDLP Z of a $\Gamma(\beta, \alpha)$-OU process is a compound Poisson process, i.e.,

$$Z_t = \sum_{n=1}^{N_t} x_n,$$

where $N = (N_s)_{s \geq 0}$ is a Poisson process with intensity β and $(x_n)_{n \geq 0}$ is a sequence

of indepent $\Gamma(1,\alpha)$-distributed random variables, independent of N_t. Consequently, the number of jumps of the Γ-OU process is finite in finite time intervals. If the Γ-OU process is used as volatility process in the BNS model, then log returns are approximately distributed according to a variance-gamma law.

For the $NIG(\alpha,\beta,0,\delta)$-OU process, the BDLP is derived in [5] to be representable as $Z = Z^{(1)} + Z^{(2)} + Z^{(3)}$. For $\rho = \beta/\alpha$, the part $Z^{(1)}$ is a $NIG(\alpha,\beta,0,(1-\rho)\delta)$-process. The second part, $Z^{(2)}$, is of the form

$$Z_t^{(2)} = \frac{1}{2\alpha(1-\rho^2)^{\frac{1}{2}}} \sum_{n=1}^{N_t} (x_n^2 - \tilde{x}_n^2),$$

where N_t is a Poisson process with intensity $[((1-\rho)/(1+\rho))^{\frac{1}{2}}\delta\alpha]^{-1}$ and $(x_n)_{n\geq 0}$ and $(\tilde{x}_n)_{n\geq 0}$ are sequences of independent standard normally distributed random variables, independent of N_t. The process $Z^{(3)}$ is characterized in terms of its Laplace transform, given by

$$\mathbb{E}[e^{\theta Z_t^{(3)}}] = \exp\left(t\rho\delta\left(\beta\left(\frac{\alpha-\beta}{\alpha+\beta}\right)^{\frac{1}{2}} - (\theta+\beta)\left(\frac{\alpha-\theta-\beta}{\alpha+\theta+\beta}\right)^{\frac{1}{2}}\right)\right).$$

4.1.4 Simulation

In order to simulate from the variance process in the BNS model, i.e., from the process σ^2 defined by

$$\sigma_t^2 = e^{-\lambda t}\sigma_0^2 + e^{-\lambda t}\int_0^{\lambda t} e^s dZ_s, \tag{4.12}$$

the most natural way would be to simulate from the BDLP and then approximate the corresponding integral. However, since Z is a jump process, this might not be the most accurate way. We therefore want to investigate alternative strategies. One other approach is the use of infinite series representations. The particular result for the integral in (4.12) we take from [6], where for a deterministic function f it is given by

$$\int_0^t f(s)dZ(s) \stackrel{d}{=} \sum_{i=1}^{\infty} \nu^{-1}\left(\frac{a_i}{t}\right) f(r_i t), \tag{4.13}$$

with $(r_i)_{i\geq 0}$ being a sequence of indepent random variables, uniformly distributed on $[0,1]$, and $(a_i)_{i\geq 0}$ being a (increasing) sequence of random variables, corresponding to the arrival times of a Poisson process with intensity 1. The two sequences are further assumed to be independent of one another. The representation in (4.13) is

obtained as an immediate consequence of their previously proven result that, given the sequences $(r_i)_{i\geq 0}$ and $(a_i)_{i\geq 0}$ as above, a Lévy process Z is representable in law on the unit interval as

$$\{Z_s : 0 \leq s \leq 1\} \stackrel{d}{=} \{\tilde{Z}_s : 0 \leq s \leq 1\}, \tag{4.14}$$

where

$$\tilde{Z}_s = \sum_{i=1}^{\infty} \nu^{-1}(a_i) \mathbb{1}_{[0,s]}(r_i).$$

A general survey of various methods of generating series representations of Lévy processes is presented by Rosiński [76]. The results above are based on the work of Marcus [62] and Rosiński [75], where these techniques have generally been introduced.

In case of the Γ-OU process ($\Gamma(\beta, \alpha)$ marginal law), we have seen that it is possible to analytically invert the upper tail mass function. By combining (4.11) with (4.13) and defining a Poisson process $N = (N_s)_{s\geq 0}$ with intensity $\beta\lambda t$ ($\mathbb{E}[N_s] = \beta\lambda t s$) and arrival times $b_1 < b_2 < \ldots$, we thus obtain

$$\begin{aligned}
e^{-\lambda t} \int_0^{\lambda t} e^s dZ_s &\stackrel{d}{=} e^{-\lambda t} \sum_{i=1}^{\infty} \nu^{-1}\left(\frac{a_i}{\lambda t}\right) e^{r_i \lambda t} \\
&= e^{-\lambda t} \sum_{i=1}^{\infty} \max\left\{0, -\frac{1}{\alpha} \log\left(\frac{a_i}{\beta\lambda t}\right)\right\} e^{r_i \lambda t} \\
&= -\frac{1}{\alpha} e^{-\lambda t} \sum_{i=1}^{\infty} \mathbb{1}_{(0,1)}\left(\frac{a_i}{\beta\lambda t}\right) \log\left(\frac{a_i}{\beta\lambda t}\right) e^{r_i \lambda t} \\
&= +\frac{1}{\alpha} e^{-\lambda t} \sum_{i=1}^{\infty} \mathbb{1}_{(0,1)}(b_i) \log(b_i^{-1}) e^{r_i \lambda t} \\
&= \frac{1}{\alpha} e^{-\lambda t} \sum_{i=1}^{N(1)} \log(b_i^{-1}) e^{r_i \lambda t}.
\end{aligned}$$

A sample path of a Γ-OU process with parameter values $\beta = 10, \alpha = 100, \lambda = 10$ and an initial volatility level of $\sigma_0^2 = 0.08$, where we use 500 discrete time points for simulation in the interval $[0, 1]$, is given in Figure 4.1.

Since we adopted the BDLP in the dynamics of the asset price, we also need to simulate the BDLP itself. The BDLP of the Γ-OU process is a compound Poisson process (cf., e.g., Cont and Tankov [31]). Using (4.14) we proceed as above to

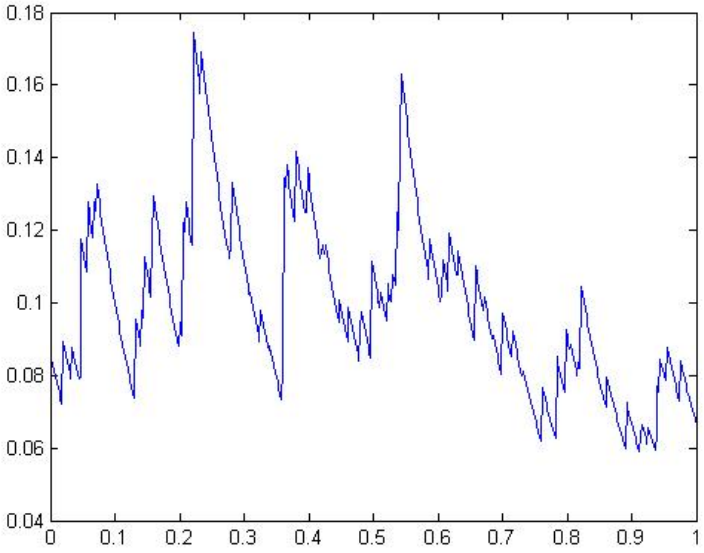


Figure 4.1: Γ-OU process

simplify the infinite series representation of the process:

$$\begin{aligned} Z_s &\stackrel{d}{=} \sum_{i=1}^{\infty} \nu^{-1}(a_i)\mathbb{1}_{[0,s]}(r_i) \\ &= \sum_{i=1}^{\infty} \mathbb{1}_{(0,1)}\left(\frac{a_i}{\beta}\right)\left(-\frac{1}{\alpha}\log\left(\frac{a_i}{\beta}\right)\right)\mathbb{1}_{[0,s]}(r_i) \\ &= \frac{1}{\alpha}\sum_{i=1}^{N(1)} \log(c_i^{-1})\mathbb{1}_{[0,s]}(r_i), \end{aligned}$$

where N now is a Poisson process with intensity β and arrival times $c_1 < c_2 < \dots$.

The cumulant function of the BDLP Z, which we need for the risk free dynamics of the log-asset price process X, given in (4.8) as

$$dX_t = \left(r - \lambda\kappa(\rho) - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t + \rho dZ_{\lambda t},$$

is given by

$$\kappa(\theta) = \log \mathbb{E}[e^{\theta Z_1}] = \frac{\beta\theta}{\alpha - \theta}, \quad \mathrm{Re}(\theta) < \alpha.$$

If we then simulate the increments of a standard Brownian motion and use, for instance, an Euler-Maruyama scheme for X, we can finally simulate trajectories of

the asset price.

A sample path simulated with this strategy, together with the volatility process and its BDLP, is given in Figure 4.2. We use the parameter values from [68], where the Γ-OU BNS model is calibrated to a data set of Call Options on the S&P 500 index. Their obtained values are $\beta = 1.0071, \alpha = 116.01, \lambda = 1.6787, \rho = -4.4617, \sigma_0^2 = 0.065883^2$. Further, we take a risk free interest rate of $r = 0.5538\%$, an initial asset price of $S_0 = 100$, and use 500 time steps of size 0.002 for simulation in the interval $[0, 1]$.

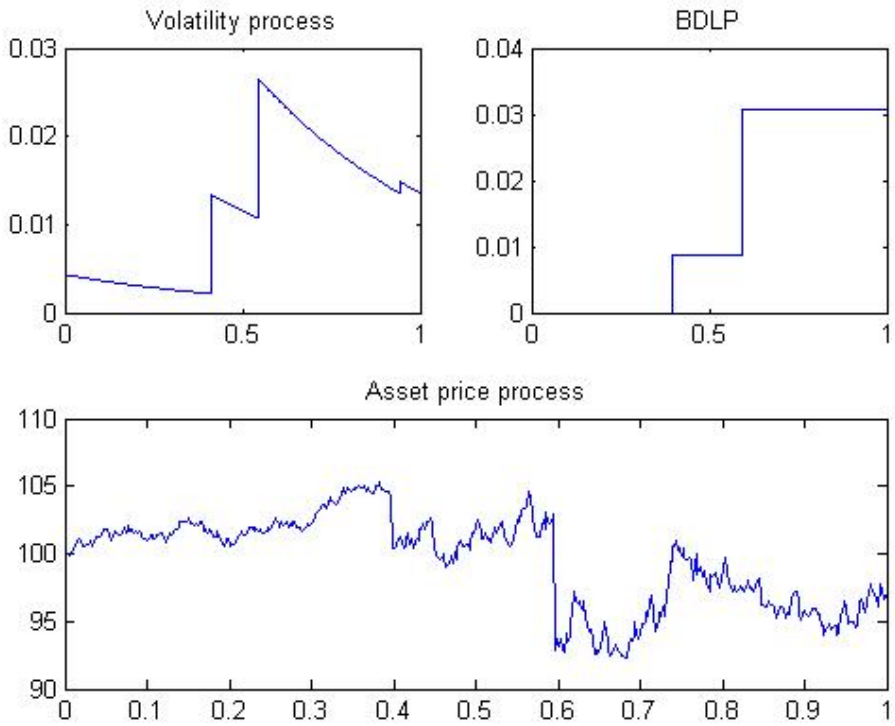


Figure 4.2: Γ-OU BNS model

For the IG-OU process, the series in (4.13) converges slowly, since the decay of $\nu^{-1}(x)$ for $x \to \infty$ is only of quadratic order in this case (cf. [6]). A more efficient way to simulate from the IG-OU process has thus been suggested by Valdivieso, Schoutens, and Tuerlinckx [83], who make use of a special path rejection technique. Based on the decomposition of the BDLP given above, that is into an IG-process and a compound Poisson process as in (4.10), and using an Euler discretization scheme, they derive the following recursive scheme for simulation of an $IG(\delta, \gamma)$-

OU process:

$$\sigma^2(k\Delta) = e^{-\lambda\Delta}\left(\sigma^2((k-1)\Delta) + J_M(\Delta) + \sum_{j=1}^{N(\lambda\Delta)} e^{T_j}Y_j\right).$$

The process J_M is defined by

$$J_M(t) = 2\sum_{j=1}^{M} \min\left\{\left(\frac{\delta\lambda T}{2b_j\Gamma\left(\frac{1}{2}\right)}\right)^2, \frac{e_j v_j^2}{\gamma^2}\right\} e^{\lambda T u_j}\mathbb{1}_{\{Tu_j\leq t\}},$$

which converges a.s. and uniformly to the process $J = (J_t)_{0\leq t\leq T}$ with

$$J_t = \int_0^{\lambda t} e^s dZ_s^{(1)},$$

for $M \to \infty$. In practice, values of M around 25000 are reasonable choices for the truncation of the series (cf. Schoutens [78], pg. 111 & 116). The sequence $(e_j)_{j\geq 0}$ consists of i.i.d $Exp(1)$ random variables, $(u_j)_{j\geq 0}$ and $(v_j)_{j\geq 0}$ are sequences of i.i.d $uniform([0,1])$ random variables, and $b_1 < b_2 < \ldots$ are arrival times of a Poisson process with intensity 1. Mutual independence of the sequences is further assumed. Figure 4.3 shows a sample path of an IG-OU process, simulated with this method. The parameter values are chosen to be $\delta = 10, \gamma = 20, \lambda = 0.3, \sigma_0^2 = 0.5$, and $M = 25000$.

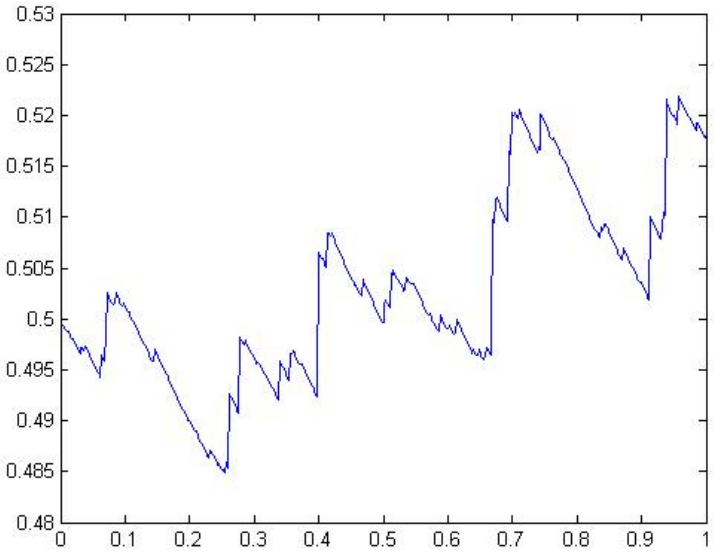


Figure 4.3: IG-OU process

Proceeding as above for the Γ-OU BNS model, we construct trajectories of the asset price also in the IG-OU BNS model. The cumulant function of the BDLP in case of the $IG(\delta, \gamma)$-OU process is given by

$$\kappa(\theta) = \frac{\delta\theta}{\sqrt{\gamma^2 - 2\theta}}, \qquad \mathrm{Re}(\theta) < \frac{\gamma^2}{2}. \tag{4.15}$$

Again we use calibration results obtained in [68], where the parameters are given as $\delta = 0.0872, \gamma = 11.98, \lambda = 2.4958, \sigma_0^2 = 0.064262^2, \rho = -4.7039$. The initial asset price is chosen to be $S_0 = 100$, the risk free interest rate is 0.5538%. We simulate 50 sample paths over the interval $[0, 1]$, with a step size of 0.002, corresponding to 500 discrete time points. The resulting picture is given in Figure 4.4.

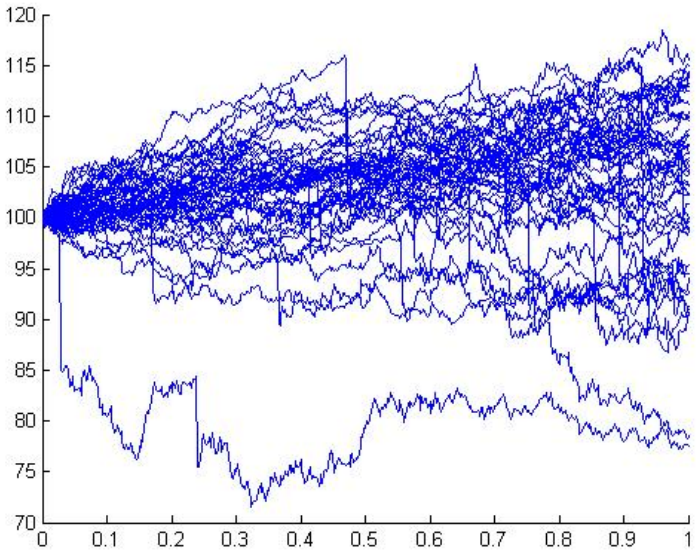


Figure 4.4: IG-OU BNS model

4.2 Multi-dimensional SV models of OU type

In the fundamental paper of Barndorff-Nielsen and Shephard [7], section 6.5 is dedicated to a first indication of the possible structure of a multivariate extension of their model. It includes their suggestion of a factor approach for the integrated covariance in terms of their model structure. This approach has subsequently been followed by Hubalek and Nicolato [46]. However, a very elegant structure for a multivariate generalization of the BNS model has been proposed by Pigorsch and Stelzer [70]. We are now going to follow their approach and investigate their model,

particularly in the context of (Spread) Option pricing, in the remainder of this thesis.

An even more general set-up, namely the supOU Stochastic Volatility model, has been introduced in Barndorff-Nielsen and Stelzer [12]. This model is based on describing the volatility by a positive semidefinite supOU process, as defined and studied in Barndorff-Nielsen and Stelzer [11] as a multivariate extension of the theory of superpositions of univariate OU processes, which have been introduced in Barndorff-Nielsen [9]. The property, which makes the supOU model appealing for financial modelling, is its feature of producing long-range dependence effects in the volatility and in logarithmic returns. This behaviour, which is usually observed in financial return data, cannot be reproduced by the model, which uses only one positive semidefinite OU type process. Furthermore, in [12] a discussion of the possibility of factor modelling within their framework is included and the model of Hubalek and Nicolato [46] is finally explained to be a special case thereof.

4.2.1 Notation

We mainly adopt the notation used in [8]. By $\mathbb{M}_{m,n}(\mathbb{K})$ we denote the set of all $\mathbb{K}$-valued $(m \times n)$-matrices. We abbreviatory write $\mathbb{M}_n$ in case $m = n$ and $\mathbb{K} = \mathbb{R}$. We denote the subspace of $\mathbb{M}_n$ of all symmetric matrices by $\mathbb{S}_n$, the cone of all posititve semidefinite symmetric matrices by $\mathbb{S}_n^+$, and the open cone of all positive definite symmetric matrices by $\mathbb{S}_n^{++}$. For the identity matrix in $\mathbb{R}^{n\times n}$ we write I_n. $\sigma(\cdot)$ denotes the spectrum of an operator, being the set of eigenvalues for a matrix $A \in \mathbb{M}_n(\mathbb{R})$.

The addition of two sets $A, B \subseteq \mathbb{R}$ is defined by $A + B := \{a + b : a \in A, b \in B\}$. The Kronecker product of two matrices $A \in \mathbb{M}_{m,n}, B \in \mathbb{M}_{p,r}$ is denoted by $A \otimes B$. It is defined as the matrix C, where $C_{ij} = a_{ij} \cdot B$, i.e., every element of A is multiplied by the matrix B. This means that $C \in \mathbb{M}_{mp,nr}$. Further, $\langle x, y\rangle := x^\top y$, being the Euclidean scalar product on $\mathbb{R}^n$. On $\mathbb{M}_n(\mathbb{R})$ we use the scalar product $\langle A, B\rangle := \mathrm{tr}(A^\top B)$ and on $\mathbb{R}^n \times \mathbb{M}_n$ we define $\langle (x_1, Y_1), (x_2, Y_2)\rangle := x_1^\top x_2 + \mathrm{tr}(Y_1^\top Y_2)$. We use the superscript asterisk to indicate adjoints of operators or matrices. For a linear operator $\mathbf{A} : V \to W$, where V, W are vector spaces, the adjoint operator $\mathbf{A}^*$ is the unique linear operator such that $\langle \mathbf{A}x, y\rangle_W = \langle x, \mathbf{A}^* y\rangle_V$ for all $x \in V, y \in W$. For a matrix $A \in \mathbb{M}_{m,n}(\mathbb{C})$, one gets the adjoint matrix A^* by taking the complex conjugate of each entry and transposing the matrix.

Let X, Y be two stochastic processes with values in $\mathbb{M}_{d,m}$ and $\mathbb{M}_{n,p}$, respectively. Let L be an $\mathbb{M}_{m,n}$-valued semimartingale. Then $\int_0^t X_t dL_t Y_t$ stands for the $(d \times p)$

random matrix, where the element in the i-th row and j-th column is given by $\sum_{k=1}^{m}\sum_{l=1}^{n}\int_0^t X_s^{ik}Y_s^{lj}dL_s^{kl}$.

4.2.2 The general multi-dimensional OU type SV model

As already mentioned above, the model we deal with has in principle been introduced by Pigorsch and Stelzer [70]. However, they define the model with jumps only occurring in the volatility process. We want to formulate the model in the sense of Muhle-Karbe, Pfaffel, and Stelzer [66], who also include a jump component in the asset price process itself, in order to make allowance for a leverage effect in the model.

Aiming for the model formulation, we first address some of the theory, where the model is based on. The crucial point of the model is the use of matrix subordinators for the construction of the volatiliy process. The concept of matrix subordinators has first been considered in Barndorff-Nielsen and Pérez-Abreu [10].

Definition 4.1: *An $\mathbb{S}_d$-valued Lévy process $L = (L_t)_{t\in\mathbb{R}_+}$ is called* matrix subordinator*, if for its increments it holds that $L_t - L_s \in \mathbb{S}_d^+$ a.s. for all $t \geq s$.*

For a matrix subordinator L, we consider the characteristic function $\phi_{L_1}(Z) = \mathbb{E}[e^{\mathrm{i}\,\mathrm{tr}(ZL_1)}]$, where $Z \in \mathbb{M}_d(\mathbb{R})$. In terms of the Lévy-Khintchine representation, the characteristic exponent is of the form

$$\psi_L(Z) = \mathrm{i}\,\mathrm{tr}(\gamma_L Z) + \int_{\mathbb{S}_d^+}(e^{\mathrm{i}\,\mathrm{tr}(XZ)} - 1)\kappa_L(dX),$$

with a drift matrix $\gamma_L \in \mathbb{S}_d^+$ and a Lévy measure κ_L on $\mathbb{S}_d$, which satisfies the conditions $\kappa_L(\mathbb{S}_d\backslash\mathbb{S}_d^+) = 0$ and $\int_{\{\|X\|\leq 1\}}\|X\|\kappa_L(dX) < \infty$.

We also consider the cumulant transform Θ_L, which is, for $Z \in \mathbb{M}_d(\mathbb{C})$, defined by

$$\Theta_L(Z) = \psi_L(-\mathrm{i}Z) = \mathrm{tr}(\gamma_L Z) + \int_{\mathbb{S}_d^+}(e^{\mathrm{tr}(XZ)} - 1)\kappa_L(dX).$$

Analogously to the univariate BNS model, the volatility process will be defined as a process of Ornstein-Uhlenbeck type. However, in the multivariate case we are dealing with matrix-valued OU type processes. Such processes have been introduced by Barndorff-Nielsen and Stelzer [8], in the light of the general study of processes of finite variation, which take values in the space of positive semidefinite matrices.

The crucial idea of generalising OU processes to matrix valued processes, is to identify $\mathbb{M}_d(\mathbb{R})$ with $\mathbb{R}^{d^2}$. In this context, we will use a vectorisation operator

denoted by vec: $\mathbb{M}_d(\mathbb{R}) \to \mathbb{R}^{d^2}$, which is the bijective linear operator that stacks the columns of a matrix below one another.

Definition 4.2: *Let $L = (L_t)_{t\in\mathbb{R}}$ be a Lévy process with values in $\mathbb{M}_d(\mathbb{R})$ and let $\mathbf{A} : \mathbb{M}_d(\mathbb{R}) \to \mathbb{M}_d(\mathbb{R})$ be a linear operator. A solution of the SDE*

$$dX_t = \mathbf{A}X_t dt + dL_t \tag{4.16}$$

is called a (matrix-valued) process of Ornstein-Uhlenbeck type.

For some given value X_0, it can be shown that (4.16) has a unique solution, which is - analoguesly to the univariate case - given by

$$X_t = e^{\mathbf{A}t}X_0 + \int_0^t e^{\mathbf{A}(t-s)}dL_s.$$

Moreover, the two conditions $\mathbb{E}[\log^+ \|L_1\|] < \infty$ and $\sigma(\mathbf{A}) \in (-\infty, 0)+i\mathbb{R}$ guarantee that the unique stationary solution of (4.16) is given by

$$X_t = \int_{-\infty}^t e^{\mathbf{A}(t-s)}dLs.$$

With the intention of constructing a process, which is suitable for modelling the stochastic evolution of a covariance matrix, we are interested in positive semidefinite Ornstein-Uhlenbeck processes. Therefore, we need to make use of matrix subordinators as driving Lévy processes. Then, for an initial matrix $X_0 \in \mathbb{S}_d^+$ and for a linear operator $\mathbf{A}$, which satisifies the condition $e^{\mathbf{A}t}(\mathbb{S}_d^+) \subseteq \mathbb{S}_d^+$ for all $t \in \mathbb{R}^+$, $(X_t)_{t\in\mathbb{R}^+}$ only takes values in $\mathbb{S}_d^+$. In other words, the condition on $\mathbf{A}$ means that we are now searching for linear operators $\mathbf{A}$ such that the exponential operator $e^{\mathbf{A}t}$ preserves positive semidefiniteness for all $t \in \mathbb{R}^+$. A sufficient condition for ensuring this is given in [8] (Proposition 4.4), as follows:

Proposition 4.1: *Assume the operator $\mathbf{A} : \mathbb{M}_d(\mathbb{R}) \to \mathbb{M}_d(\mathbb{R})$ is representable as $X \mapsto AX + XA^*$ for some $A \in \mathbb{M}_d(\mathbb{R})$. Then $e^{\mathbf{A}t}$ has the representation $X \mapsto e^{At}Xe^{A^*t}$ and $e^{\mathbf{A}t}(\mathbb{S}_d^+) = \mathbb{S}_d^+$ for all $t \in \mathbb{R}$.*

Remark: *The operator $\mathbf{A}$ associated with the matrix A can be represented as*

$$\mathrm{vec}^{-1} \circ ((A \otimes I) + (I \otimes A)) \circ \mathrm{vec}\,.$$

The above results are combined in the subsequent theorem, which will in further consequence enable us to formulate the model.

Theorem 4.4: *Let $L = (L_t)_{t\in\mathbb{R}}$ be a matrix subordinator satisfying $\mathbb{E}[\log^+ \|L_1\|] < \infty$ and $A \in \mathbb{M}_d$ such that $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$. Then the stochastic differential equation of Ornstein-Uhlenbeck type*

$$dX_t = (AX_t + X_tA^*)dt + dL_t \tag{4.17}$$

has a unique stationary solution, given by

$$X_t = \int_{-\infty}^{t} e^{A(t-s)} dL_s e^{A^*(t-s)}.$$

Moreover, $X_t \in \mathbb{S}_d^+$ for all $t \in \mathbb{R}$. For a given initial value X_0, the unique strong solution of (4.17) satisfies

$$X_t = e^{At}X_0e^{A^*t} + \int_0^t e^{A(t-s)} dL_s e^{A^*(t-s)}.$$

Remark: *Applying the vectorization operator on X_t, we get the representation*

$$\operatorname{vec}(X_t) = \int_{-\infty}^{t} e^{(I_d\otimes A + A\otimes I_d)(t-s)} d\operatorname{vec}(L_s).$$

Equipped with the theory introduced above, we are now able to formulate the model along the lines of Muhle-Karbe, Pfaffel, and Stelzer [66]. Let L be a matrix subordinator and let W be a $\mathbb{R}^d$-valued Brownian motion, independent of L. The *multivariate stochastic volatility model of OU type* is defined via the following dynamics of the logarithmic asset returns and the volatility, respectively:

$$dY_t = (\mu + \beta(\Sigma_t))dt + \Sigma_t^{\frac{1}{2}} dW_t + \rho(dL_t), \tag{4.18}$$
$$d\Sigma_t = (A\Sigma_t + \Sigma_t A^\top)dt + dL_t, \tag{4.19}$$

with initial values $Y_0 \in \mathbb{R}^d$, $\Sigma_0 \in \mathbb{S}_d^+$, and parameters $\mu \in \mathbb{R}^d$ and $A \in M_d(\mathbb{R})$ such that $0 \notin \sigma(A) + \sigma(A)$. The risk premium and the leverage term are modeled using the linear operators $\beta, \rho : M_d(\mathbb{R}) \to \mathbb{R}^d$. A common way of specifying these operators is to choose coefficients $\beta_1, \ldots, \beta_d \in \mathbb{R}$ and $\rho_1, \ldots, \rho_d \in \mathbb{R}$ for the diagonal elements of the argument, i.e.,

$$\beta(X) = \begin{pmatrix} \beta_1 X_{11} \\ \vdots \\ \beta_d X_{dd} \end{pmatrix}, \quad \rho(X) = \begin{pmatrix} \rho_1 X_{11} \\ \vdots \\ \rho_d X_{dd} \end{pmatrix}, \quad \forall X \in \mathbb{M}_d(\mathbb{R}).$$

In this case, we call β and ρ *diagonal*. If β and ρ are taken to be diagonal and

also A is chosen as a diagonal matrix, then the model for each asset is equivalent in distribution to a univariate BNS model.

Since, in the formulation of the model, the square root of a positive semidefinite OU process is used, we want to mention a general result for such square root processes, in order to explicitly address its unique existence and finite variation property. This result is due to Barndorff-Nielsen and Stelzer [8] (Proposition 6.7).

Proposition 4.2: *Let $(X_t)_{t\in\mathbb{R}^+}$ be an $\mathbb{S}_d^+$-valued process of Ornstein-Uhlenbeck type driven by a matrix subordinator L with drift $\gamma \in \mathbb{S}_d^+$ and associated Poisson random measure μ. Then the unique positive square root $Y_t = \sqrt{X_t}$ is of finite variation and has the following representation:*

$$\begin{aligned} dY_t &= \mathbf{X}_{t-}^{-1}(AX_{t-} + X_{t-}A^* + \gamma)dt + \int_{\mathbb{S}_d^+\setminus\{0\}} \left(\sqrt{X_{t-}+x} - \sqrt{X_{t-}}\right)\mu(dt,dx) \\ &= \mathbf{Y}_{t-}^{-1}(AY_{t-}^2 + Y_{t-}^2A^* + \gamma)dt + \int_{\mathbb{S}_d^+\setminus\{0\}} \left(\sqrt{Y_{t-}^2+x} - Y_{t-}\right)\mu(dt,dx), \end{aligned}$$

provided that the process X_t is locally bounded within $\mathbb{S}_d^{++}$ or the integrals

$$\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)ds \text{ and } \int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} \left(\sqrt{X_{s-}+x} - \sqrt{X_{s-}}\right)\mu(ds,dx)$$

exist a.s. for all $t \in \mathbb{R}$. Here, $\mathbf{X}_{t-}$ is the linear operator $Z \mapsto \sqrt{X_{t-}}Z + Z\sqrt{X_{t-}}$ and $\mathbf{Y}_{t-}$ is the mapping $Z \mapsto Y_{t-}Z + ZY_{t-}$.

Particularly the application of transform-based methods in the context of option pricing, as we will do in chapter 5, requires the characteristic function of the asset price. For the multivariate OU type SV model, we may use the following result, computed in [66] (Theorem 2.5).

Theorem 4.5: *For every $(y,z) \in \mathbb{R}^d \times \mathbb{M}_d(\mathbb{R})$ and $t \in \mathbb{R}_+$, the joint characteristic function of (Y_t, Σ_t) is given by*

$$\begin{aligned} &\mathbb{E}[\exp(\mathrm{i}\langle(y,z),(Y_t,\Sigma_t)\rangle)] \\ &= \exp\Bigg\{\mathrm{i}y^\top(Y_0+\mu t) + \mathrm{i}\,\mathrm{tr}(\Sigma_0 e^{A^\top t}ze^{At}) \\ &\quad + \mathrm{i}\,\mathrm{tr}\left(\Sigma_0\left(e^{A^\top t}\mathbf{A}^{-*}\left(\beta^*(y)+\frac{\mathrm{i}}{2}yy^\top\right)e^{At} - \mathbf{A}^{-*}\left(\beta^*(y)+\frac{\mathrm{i}}{2}yy^\top\right)\right)\right) \\ &\quad + \int_0^t \psi_L\left(e^{A^\top s}ze^{As} + \rho^*(y) + e^{A^\top s}\mathbf{A}^{-*}\left(\beta^*(y)+\frac{\mathrm{i}}{2}yy^\top\right)e^{As} - \mathbf{A}^{-*}\left(\beta^*(y)+\frac{\mathrm{i}}{2}yy^\top\right)\right)ds\Bigg\}, \end{aligned}$$

where $\mathbf{A}^{-*} := (\mathbf{A}^*)^{-1}$ denotes the inverse of the adjoint of $\mathbf{A} : X \mapsto AX + XA^\top$,

that is, the inverse of $\mathbf{A}^* : X \mapsto A^\top X + XA$.

In the remainder of this subsection, we want to study some theoretical properties of the model. All of the three subsequent results are due to [66]; the proofs can be found therein.

The first theorem deals with the regularity of the moment generating function. We will use this result in section 4.2.3, where we derive the moment generating function for the concrete specification of the two-dimensional OU Wishart model along the lines of [66]. This will particularly enable us to apply the method introduced in section 2.1.3 in order to price Spread Options, what will be done within the context of our numerical investigations in chapter 5.

Theorem 4.6: *Suppose the matrix subordinator* L *satisfies*

$$\int_{\{\|X\|\geq 1\}} e^{\mathrm{tr}(RX)} \kappa_L(dX) < \infty \text{ for all } R \in \mathbb{M}_d(\mathbb{R}) \text{ with } \|R\| < \epsilon,$$

for some $\epsilon > 0$. *Then the moment generating function* Φ_{Y_t} *of* Y_t *is analytic on the open convex set*

$$S_\theta := \{y \in \mathbb{C}^d : \| \operatorname{Re}(y)\| < \theta\},$$

where

$$\theta := -\frac{\|\rho\|}{(e^{2\|A\|t}+1)\|\mathbf{A}^{-1}\|} - \|\beta\| + \sqrt{\Delta} > 0$$

with

$$\Delta := \left(\frac{\|\rho\|}{(e^{2\|A\|t}+1)\|\mathbf{A}^{-1}\|} + \|\beta\|\right)^2 + \frac{2\epsilon}{(e^{2\|A\|t}+1)\|\mathbf{A}^{-1}\|}.$$

Moreover,

$$\Phi_{Y_t}(y) = \exp\left(y^\top(Y_0 + \mu t) + \mathrm{tr}(\Sigma_0 H_y(t)) + \int_0^t \Theta_L(H_y(s) + \rho^*(y))ds\right)$$

for all $y \in S_\theta$, *where*

$$H_y(s) := e^{A^\top s}\mathbf{A}^{-*}\left(\beta^*(y) + \frac{1}{2}yy^\top\right)e^{As} - \mathbf{A}^{-*}\left(\beta^*(y) + \frac{1}{2}yy^\top\right). \tag{4.20}$$

In accordance with the Fundamental Theorem of Asset Pricing, we need to address the question of equivalent martingale measures. First, we study the conditions under which the discounted price process in the multivariate OU type SV model

is a martingale. The second theorem gives a characterization of the class of those equivalent martingale measures, which preserve the structure of the model, i.e., under which L remains a Lévy process independent of W.

Theorem 4.7: *The discounted price process $(e^{-rt}S_t)_{t\in[0,T]}$ is a martingale if and only if, for $i = 1, \dots, d$,*

$$\int_{\{\|X\|>1\}} e^{\rho^i(X)}\kappa_L(dX) < \infty,$$

and

$$\beta^i(X) = -\frac{1}{2}X_{ii}, \quad X \in \mathbb{S}_d^+,$$
$$\mu_i = r - \int_{\mathbb{S}_d^+}(e^{\rho^i(X)} - 1)\kappa_L(dX).$$

Theorem 4.8: *Let $\Lambda(t) : X \mapsto e^{At}Xe^{A^\top t} - X$. Let $y : \mathbb{S}_d^+ \to (0,\infty)$ such that*

(i) $\int_{\mathbb{S}_d^+}(\sqrt{y(X)} - 1)^2\kappa_L(dX) < \infty$,

(ii) $\int_{\{\|X\|>1\}} e^{\rho^i(X)}\kappa_L^y(dX) < \infty, \quad i = 1, \dots, d,$

where $\kappa_L^y(B) := \int_B y(X)\kappa_L(dX)$ for $B \in \Lambda(\mathbb{S}_d^+)$. Define the $\mathbb{R}^d$-valued process $(\psi_t)_{t\in[0,T]}$ as

$$\psi_t = -\Sigma_t^{-\frac{1}{2}}\left(\mu + \beta(\Sigma_t) + \frac{1}{2}\begin{pmatrix}\Sigma_t^{11}\\ \vdots \\ \Sigma_t^{dd}\end{pmatrix} + \begin{pmatrix}\int_{\mathbb{S}_d^+}(e^{\rho^1(X)} - 1)\kappa_L^y(dX)\\ \vdots \\ \int_{\mathbb{S}_d^+}(e^{\rho^d(X)} - 1)\kappa_L^y(dX)\end{pmatrix} - r\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}\right).$$

Then $Z = \mathcal{E}(\int \psi dW + (y-1) * (\mu^L - \nu^L))$ is a density process, and the probability measure $\mathbb{Q}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T$ is an equivalent martingale measure. Moreover, $W^{\mathbb{Q}}$ with $W_t^{\mathbb{Q}} := W_t - \int_0^t \psi_s ds$ is a $\mathbb{Q}$-standard Brownian motion, and L is an independent driftless $\mathbb{Q}$-matrix subordinator with Lévy measure κ_L^y. The $\mathbb{Q}$-dynamics of (Y, Σ) are given by

$$dY_t^i = \left(r - \int_{\mathbb{S}_d^+}(e^{\rho^i(X)} - 1)\kappa_L^y(dX) - \frac{1}{2}\Sigma_t^{ii}\right)dt + \left(\Sigma_t^{\frac{1}{2}}dW_t^{\mathbb{Q}}\right)^i + \rho^i(dL_t), i = 1, \dots, d,$$
$$d\Sigma_t = (\gamma_L + A\Sigma_t + \Sigma_t A^\top)dt + dL_t.$$

4.2.3 The two-dimensional OU-Wishart model

Now we want to discuss a concrete specification of the general model introduced above, which has been considered in [66]. In chapter 5, we will use this two-dimensional specification to model the logarithmic returns of the underlying assets of Spread Options and determine the corresponding option value. However, before we define the model, we want to comment on the most important aspects about the Wishart distribution.

The Wishart distribution, denoted by $\mathcal{W}_d(n, \Sigma)$, is a probability distribution of a positive-semidefinite $(d \times d)$ random matrix. The parameter $n \in \mathbb{N}$ represents the degrees of freedom, while $\Sigma \in \mathbb{S}_d^+$ is the scale matrix of the distribution. Let X be a $(d \times n)$ matrix, where the columns are independently drawn from multivariate standard normal distributions, i.e., $X_{(i)} \sim N(0, I_d)$. Then $(XX^\top) \sim \mathcal{W}_d(n, I_d)$. Since, in general, any vector $Y \sim N(\mu, \Sigma)$ can be written in the form $Y = \mu + CZ$, where $Z \sim N(0, I_d)$ and $CC^\top = \Sigma$ (using, e.g., Cholesky decomposition), we can get a $\mathcal{W}_d(n, \Sigma)$ distribution by considering the matrix $M = (CXX^\top C^\top)$. Assuming $n > d - 1$ and Σ is invertible, the probability density function of a matrix $M \sim \mathcal{W}_d(n, \Sigma)$ is given by

$$f(Z, n, \Sigma) = \frac{(\det(Z))^{\frac{n-d-1}{2}} \, e^{-\frac{1}{2}\operatorname{tr}\left(Z\Sigma^{-1}\right)}}{2^{\frac{nd}{2}} \pi^{\frac{d(d-1)}{4}} \, (\det(\Sigma))^{\frac{n}{2}} \prod_{i=1}^{d} \Gamma\left(\frac{n+1-i}{2}\right)}.$$

The moment generating function is the mapping

$$Z \mapsto \det\left(I_d - 2Z\Sigma\right)^{-\frac{n}{2}}, \tag{4.21}$$

where Z is any symmetric matrix of real numbers (cf., e.g., [42]).

The *two-dimensional OU-Wishart model* is (already in terms of a risk-neutral pricing measure $\mathbb{Q}$) defined via the dynamics

$$\begin{pmatrix} dY_t^1 \\ dY_t^2 \end{pmatrix} = \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \Sigma_t^{11} \\ \Sigma_t^{22} \end{pmatrix} \right) dt + \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} + \begin{pmatrix} \rho_1 dL_t^{11} + \rho_{12} dL_t^{12} \\ \rho_2 dL_t^{22} + \rho_{21} dL_t^{12} \end{pmatrix} \tag{4.22}$$

$$\begin{pmatrix} d\Sigma_t^{11} & d\Sigma_t^{12} \\ d\Sigma_t^{12} & d\Sigma_t^{22} \end{pmatrix} = \left(\begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix} + \begin{pmatrix} 2a_1\Sigma_t^{11} & (a_1+a_2)\Sigma_t^{12} \\ (a_1+a_2)\Sigma_t^{12} & 2a_2\Sigma_t^{22} \end{pmatrix} \right) dt + \begin{pmatrix} dL_t^{11} & dL_t^{12} \\ dL_t^{12} & dL_t^{22} \end{pmatrix} \tag{4.23}$$

with initial values

$$Y_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \Sigma_0^{11} & \Sigma_0^{12} \\ \Sigma_0^{12} & \Sigma_0^{22} \end{pmatrix} \in \mathbb{S}_2^{++}.$$

For the parameters, we assume $\gamma_1, \gamma_2 \geq 0, a_1, a_2 < 0$, and $\rho_1, \rho_2, \rho_{12}, \rho_{21} \in \mathbb{R}$. The BDLP L is taken to be a compound Poisson process with intensity λ and where for the jump-size distribution we choose a $\mathcal{W}_2(2, \Theta)$-law with dispersion matrix

$$\Theta = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{12} & \Theta_{22} \end{pmatrix} \in \mathbb{S}_2^+.$$

The drift vector $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, which ensures risk neutrality, results from Theorem 4.7. It is given by

$$\begin{aligned} \mu_1 &= r - \int_{\mathbb{S}_d^+} (e^{\rho^1 X^{11} + \rho^{12} X^{12}} - 1) \kappa_L(dX) \\ &= r - \lambda \left(\int_{\mathbb{S}_d^+} e^{\rho^1 X^{11} + \rho^{12} X^{12}} f(X) dX - \int_{\mathbb{S}_d^+} f(X) dX \right), \\ \mu_2 &= r - \int_{\mathbb{S}_d^+} (e^{\rho^2 X^{22} + \rho^{21} X^{12}} - 1) \kappa_L(dX) \\ &= r - \lambda \left(\int_{\mathbb{S}_d^+} e^{\rho^2 X^{22} + \rho^{21} X^{12}} f(X) dX - \int_{\mathbb{S}_d^+} f(X) dX \right), \end{aligned}$$

where the Lévy measure of the BDLP, κ_L, in our case is given by λ times the density f of the Wishart distribution. Regarding μ_1, the second integral equals 1 (being the integral of the density over the whole domain), and for the first integral we use the expression for the moment generating function of the Wishart distibution given

in (4.21). Hence, we can write

$$\begin{aligned}\mu_1 &= r - \lambda\left(\det\left(I_2 - 2\left(\begin{smallmatrix}\rho_1 & \rho_{12}\\ 0 & 0\end{smallmatrix}\right)\Theta\right)^{-1} - 1\right)\\ &= r - \lambda\frac{2\rho_1\Theta_{11} + 2\rho_{12}\Theta_{12}}{1 - 2\rho_1\Theta_{11} - 2\rho_{12}\Theta_{12}}.\end{aligned}$$

Analogously, we obtain

$$\mu_2 = r - \lambda\frac{2\rho_2\Theta_{22} + 2\rho_{21}\Theta_{12}}{1 - 2\rho_2\Theta_{22} - 2\rho_{21}\Theta_{12}}.$$

An interesting observation is the following link between the two-dimensional OU-Wishart model and the Γ-OU BNS model. Let ρ be diagonal, i.e., $\rho_{12} = \rho_{21} = 0$ in (4.22). Recall from the definition above that a matrix $M = \Theta^{\frac{1}{2}}XX^\top\Theta^{\frac{1}{2}}$ is $\mathcal{W}_d(n,\Theta)$-distributed, if the entries of the $(d \times n)$ matrix X are i.i.d standard normal. In our case, where $d = n = 2$, this means that $(XX^\top)_{11}$ as well as $(XX^\top)_{22}$ are sums of the squares of two standard normally distributed random variables, i.e., they both have a chi-squared distribution with two degrees of freedom. The probability density function of the χ^2_n-distribution is given by

$$f_n(x) = \begin{cases}\frac{x^{\frac{n}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}, & x > 0\\ 0 & x \le 0\end{cases} \quad \text{which for } n = 2 \text{ means} \quad f_2(x) = \begin{cases}\frac{e^{-\frac{x}{2}}}{2}, & x > 0\\ 0 & x \le 0\end{cases}.$$

Since this corresponds to the probability density of an exponential distribution with parameter 2, it means that a χ^2_2-distrution is equal to an $Exp(2)$ distribution. In other words, the jump size distribution of L^{ii} is an exponential distribution. Consequently, L^{ii} is a compound Poisson process with exponentially distributed jumps. The sum of n independent $Exp(\lambda)$ distributed random variables is $Gamma(n,\lambda)$-distributed. Hence, the model for each asset is equivalent in distribution to a Γ-OU BNS model.

The joint moment generating function of (Y^1, Y^2) is derived in [66] and given by

$$\begin{aligned}&\Phi_{Y_t}(y) = \mathbb{E}[e^{y^\top Y_t}]\\ &= \exp\left(y^\top\mu t + \operatorname{tr}(\Sigma_0 H_y(t)) + \int_0^t \operatorname{tr}(\gamma_L H_y(s))ds + \lambda\int_0^t \frac{1}{\det(I_2 - 2(H_y(s) + \rho^*(y))\Theta)}ds - \lambda t\right),\end{aligned} \tag{4.24}$$

where $A = \left(\begin{smallmatrix}a_1 & 0\\ 0 & a_2\end{smallmatrix}\right), \gamma_L = \left(\begin{smallmatrix}\gamma_1 & 0\\ 0 & \gamma_2\end{smallmatrix}\right), \rho^*(y) = \left(\begin{smallmatrix}\rho_1 y_1 & \rho_{12} y_1\\ \rho_{21} y_2 & \rho_2 y_2\end{smallmatrix}\right)$ and the function H_y is defined as in (4.20).

Assuming $a_1 = a_2 =: a$, the integrations in the exponent in formula (4.24) can be

performed analytically after some simplifications of the occurring integrands. In particular, applying the rule $\det(A+B) = \det(A) + \det(B) + \operatorname{tr}(A)\operatorname{tr}(B) - \operatorname{tr}(AB)$, which is valid for (2×2) matrices A, B, the determinant in the denominator of the second integral can be simplified sufficiently. However, this does not seem to be possible unless the mean reversion speeds a_1 and a_2 are assumed to be equal. Just to give a very rough indication of the occurring technical issue in this case, if $a_1 \neq a_2$ one faces difficulties in getting rid of the matrix exponentials e^{As} in $H_y(s)$, which otherwise can be written as $e^{as} I_2$, which in turn renders it possible to find a "sufficiently nice" expression for the respective denominator. The closed-form expression of the joint moment generating function of (Y^1, Y^2), in case that the two mean reversion speeds in the volatility process are assumed to coincide, is given by

$$\begin{aligned}\Phi_{Y_t}(y) = \exp\Bigg\{ & y_1\mu_1 t + y_2\mu_2 t + \frac{e^{2at}-1}{4a} \operatorname{tr}\left(\Sigma_0 \begin{pmatrix} y_1^2 - y_1 & y_1 y_2 \\ y_1 y_2 & y_2^2 - y_2 \end{pmatrix}\right) \\ & + \frac{1}{4a}\left(\gamma_1(y_1^2 - y_1) + \gamma_2(y_2^2 - y_2)\right)\left(\frac{1}{2a}(e^{2at}-1) - t\right) \\ & + \frac{\lambda}{2ab_0}\left[\frac{b1}{\Delta}\left(\arctan\left(\frac{2b_2 + b_1}{\Delta}\right) - \arctan\left(\frac{2b_2 e^{2at} + b_1}{\Delta}\right)\right)\right. \\ & \left. + \frac{1}{2}\ln\left(\frac{b_0 + b_1 + b_2}{b_2 e^{4at} + b_1 e^{2at} + b_0}\right)\right] + \frac{\lambda}{b_0} t - \lambda t \Bigg\},\end{aligned}$$

where

$$\begin{aligned} b_0 &:= 1 + 4\det(B-C) + 2\operatorname{tr}(B-C), \\ b_1 &:= -8\det(B) + 4\operatorname{tr}(B)\operatorname{tr}(C) - 4\operatorname{tr}(BC) - 2\operatorname{tr}(B), \\ b_2 &:= 4\det(B), \\ \Delta &:= \sqrt{4b_0 b_2 - b_1^2}, \\ B &:= \frac{1}{4a}\begin{pmatrix} y_1^2 - y_1 & y_1 y_2 \\ y_1 y_2 & y_2^2 - y_2 \end{pmatrix}\Theta, \\ C &:= \begin{pmatrix} \rho_1 y_1 & \rho_{12} y_1 \\ \rho_{21} y_2 & \rho_2 y_2 \end{pmatrix}\Theta. \end{aligned}$$

In order to evaluate the moment generating function in a model, where Y_1 and Y_2 are not assumed to have equal speeds of mean reversion, one has to apply numerical integration methods. However, in the light of our context of transform-based option pricing, we want to emphasize that this will bring a further approximated dimension

in the pricing model and hence lead to adding up approximation errors.

4.2.4 A two-dimensional IG-OU model

In section 4.2.3, we investigated a special case of the general multivariate OU type SV model by specifying the BDLP as a CPP with Wishart-distributed jumps. This case turned out to be rather amenable; a closed-form expression for the moment generating function had already been presented in the literature.

Analogously to the one-dimensional BNS model (cf. sect. 4.1.3), as a next step, instead of specifying the details of the BDLP in advance, we would now like to prespecify the invariant distribution of the volatility process. In particular, we are interested in the inverse Gaussian distribution. The inverse Gaussian law on a symmetric cone is defined and investigated in the work of Bernadac [17]. In order to define this probability distribution in a way easier to handle, a generalization of the Bessel function is previously introduced as follows:

Definition 4.3: *For $A, B \in \mathbb{S}_n^{++}(\mathbb{R})$ and $\lambda \in \mathbb{R}$, we define the function $K(\cdot)$ by*

$$\begin{aligned} K(\lambda, A, B) :=& \frac{1}{2^n} (\det A)^{\frac{\lambda}{2}} (\det B)^{-\frac{\lambda}{2}} \\ & \times \int_{\mathbb{S}_n^{++}(\mathbb{R})} \exp\left\{ \operatorname{tr}\left(-\frac{1}{2}\left(AX + BX^{-1} \right) \right) \right\} (\det(X))^{\frac{1}{2}\lambda(n+1)}\, dX. \end{aligned}$$

Definition 4.4: *The* inverse Gaussian probability distribution on $\mathbb{S}_n^{++}(\mathbb{R})$ *is defined by*

$$\begin{aligned} \mu_{\lambda,A,B}(dX) =& \frac{(\det A)^{\frac{\lambda}{2}} (\det B)^{-\frac{\lambda}{2}}}{2^n K(\lambda, A, B)} (\det X)^{\lambda - \frac{n+1}{2}} \\ & \times \exp\left\{ \operatorname{tr}\left(-\frac{1}{2}(AX + BX^{-1}) \right) \right\} \mathbb{1}_{\mathbb{S}_n^{++}(\mathbb{R})}(X) dX, \end{aligned}$$

where $A, B \in \mathbb{S}_n^{++}(\mathbb{R})$ and $\lambda \in \mathbb{R}$.

Proposition 4.3 ([17], pg. 234)**:** *The Laplace transform of the inverse Gaussian probability distribution on $\mathbb{S}_n^{++}(\mathbb{R})$ is given by*

$$\begin{aligned} L_{\mu_{\lambda,A,B}}(Z) :=& \int e^{\operatorname{tr}(-ZX)} \mu_{\lambda,A,B}(dX) \\ =& \frac{K(\lambda, A + 2Z, B)}{K(\lambda, A, B)} \left(\det(I_n + 2A^{-1}Z) \right)^{-\frac{\lambda}{2}}, \end{aligned} \tag{4.25}$$

where $Z \in \mathbb{S}_n^{++}(\mathbb{R})$.

In the very general setting, now two questions arise: First, if we indeed prespecify the invariant distribution of the volatility process, what does the particular BDLP look like? And second, how can we use (4.25) in order to perform transform-based option pricing?

The adaptation of the concept of operator selfdecomposability to the matrix-valued setting[2], which will be needed in order to address the questions raised above, as well as the general result regarding the link between the stationary distribution of an OU type process and the specification of its BDLP, we take from the work of Pigorsch and Stelzer [71, Thm. 4.9].

Definition 4.5: *Let $\mathcal{Q} : \mathbb{S}_d \to \mathbb{S}_d$ be a linear operator. A probability distribution μ on $\mathbb{S}_d$ is called* operator selfdecomposable with respect to the operator $\mathcal{Q}$ *if there exists a probability distribution ν_t on $\mathbb{S}_d$ such that $\mu = (e^{\mathcal{Q}t}\mu) * \nu_t$ for all $t \in \mathbb{R}^+$.*

Theorem 4.9: *Let $\mathbf{A} : \mathbb{S}_d \to S_d$ be a linear operator such that there is an $A \in \mathbb{M}_d(\mathbb{R})$ with $\sigma(A) \subset (-\infty, 0) + \mathrm{i}\mathbb{R}$ satisfying $\mathbf{A}X = AX + XA^\top$ for all $X \in \mathbb{S}_d^+$. Furthermore, let μ be an operator selfdecomposable (with respect to $\mathbf{A}$) distribution on $\mathbb{S}_d^+$ such that its characteristic function is of the form*

$$\hat{\mu}(Z) = \exp\left(\mathrm{i}\,\mathrm{tr}(\gamma_\mu Z) + \int_{\mathbb{S}_d} \left(e^{\mathrm{i}\,\mathrm{tr}(XZ)} - 1\right) \nu_\mu(dX)\right), \quad Z \in \mathbb{S}_d,$$

where $\gamma_\mu \in \mathbb{S}_d^+$ and ν_μ is a Lévy measure on $\mathbb{S}_d$ satisfying

$$\nu_\mu\left(\mathbb{S}_d \backslash \mathbb{S}_d^+\right) = 0 \quad \textit{and} \quad \int_{\|X\| \le 1} \|X\| \nu_\mu(dX) < \infty.$$

Let $\psi(Z) = \log\left(\int_{\mathbb{S}_d^+} e^{\mathrm{i}\,\mathrm{tr}(XZ)} \mu(dX)\right), Z \in \mathbb{S}_d$, be its cumulant transform (logarithm of the characteristic function $\hat{\mu}$). Assume that $-A\gamma_\mu - \gamma_\mu A^\top \in \mathbb{S}_d^+$, that $\psi(Z)$ is differentiable for all $Z \neq 0$ with derivative $D\psi(Z)$ and that

$$\psi_L : Z \mapsto \begin{cases} -D\psi(Z)(A^\top Z + ZA), & \textit{for } Z \in \mathbb{S}_d \backslash \{0\}, \\ 0, & \textit{for } Z = 0, \end{cases}$$

is continuous at zero. Then $\hat{\mu}_L : Z \mapsto \exp(\psi_L(Z))$ is the characteristic funtion of an infinitely divisible distribution μ_L on $\mathbb{S}_d^+$.

Let L be a matrix subordinator with characteristic function $\hat{\mu}_L$ at time one. Then the positive semidefnite Ornstein-Uhlenbeck type process $d\Sigma_t = \mathbf{A}\Sigma_t dt + dL_t$ driven by L has stationary distribution μ.

[2]Operator selfdecomposable distributions, a subclass of the class of infinitely divisible distributions, are extensively studied in the book of Jurek and Mason [55].

In case of the IG distribution on $\mathbb{S}_n^{++}(\mathbb{R})$, however, the characteristic function is not available in closed form; combining (4.3) with the result from Theorem 4.9, does not straightforwardly yield a suitable representation for applying tranform-based option pricing methods. Future work will be needed to surmount the technical hurdles in order to find such a representation in this case.

Therefore, at this stage, we want to put aside the model, where the invariant distribution of the variance process is specified as an IG law on $\mathbb{S}_n^{++}(\mathbb{R})$; in turn, we define a different multivariate model of IG-OU type in the following way:

The logarithmic return process $Y = (Y^{(1)}, Y^{(2)})$ of the underlying assets and the resp. variance process $\Sigma = (\Sigma^{(1)}, \Sigma^{(2)})$ satisfy

$$\begin{pmatrix} dY_t^{(1)} \\ dY_t^{(2)} \end{pmatrix} = \begin{pmatrix} r - \lambda_1\kappa(\rho_1) - \frac{1}{2}\Sigma_t^{(1)} \\ r - \lambda_2\kappa(\rho_2) - \frac{1}{2}\Sigma_t^{(2)} \end{pmatrix} dt + \Sigma_t^{\frac{1}{2}} \begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \end{pmatrix} + \begin{pmatrix} \rho_1 dZ_{\lambda_1 t} \\ \rho_2 dZ_{\lambda_2 t} \end{pmatrix}, \tag{4.26}$$

$$\begin{pmatrix} d\Sigma_t^{(1)} & 0 \\ 0 & d\Sigma_t^{(2)} \end{pmatrix} = \begin{pmatrix} -\lambda_1\Sigma_t^{(1)} & 0 \\ 0 & -\lambda_2\Sigma_t^{(2)} \end{pmatrix} dt + \begin{pmatrix} dZ_{\lambda_1 t} & 0 \\ 0 & dZ_{\lambda_2 t} \end{pmatrix}. \tag{4.27}$$

$W^{(1)}$ and $W^{(2)}$ are two independent Brownian motions and the (one-dimensional) variance processes $\Sigma^{(1)}$ and $\Sigma^{(2)}$ are taken to be $IG(\delta, \gamma)$-OU processes, both driven by one and the same BDLP Z. However, λ_1 and λ_2 may be chosen differently, which means that the timing of the BDLPs as well as the mean reversion speeds of the two processes differ. Recall from section 4.1 that the value of λ_i does not affect the marginal distribution of $\Sigma_t^{(i)}$.

As we have seen in section 4.1.3, the BDLP of an $IG(\delta, \gamma)$-OU process can be represented as the sum of an $IG(\frac{\delta}{2}, \gamma)$ process and an independent compound Poisson process of the form (4.10). We denote this decomposition by

$$Z = Z^{(\mathrm{IG})} + Z^{(CPP)}. \tag{4.28}$$

We further want to write the BDLP Z in (4.27) in the form

$$dZ_t = d\ dt + d\tilde{Z}_t,$$

where $\tilde{Z}$ is a martingale; we will need this kind of representation in the sequel, when we derive the characteristic function of the log-returns in our model.

A Lévy process L has the decomposition

$$L_t = \int_0^t \int_0^\infty z\nu(dz)dt + \int_0^t \int_0^\infty z\left(N(dz,dt) - \nu(dz)dt\right), \tag{4.29}$$

where the second integral on the right-hand side is a martingale. In this representation, N denotes a Poisson random measure and ν is the Lévy measure of L. Loosely speaking, in (4.29) we split the process by writing its compensated version as one part and its expectation as a separate second part.

Let us first have a look at the part $Z^{(\mathrm{IG})}$ in (4.28): The Lévy measure of an $IG(a,b)$ process is given by

$$\nu(dx) = \frac{a}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2}b^2 x} \mathbb{1}_{\{x>0\}} dx,$$

and the corresponding integral

$$\int_0^\infty z\nu(dz) = \frac{a}{b}.$$

Consequently, we get from (4.29) that, for $i = 1, 2$, we can write

$$dZ^{(\mathrm{IG})}_{\lambda_i t} = \frac{\lambda_i \delta}{2\gamma} dt + d\tilde{Z}^{(\mathrm{IG})}_{\lambda_i t}, \tag{4.30}$$

where $\tilde{Z}^{(\mathrm{IG})}$ is a martingale.

For a compound Poisson process, the Lévy measure is given by

$$\nu(dz) = \theta f(z)dz,$$

where θ denotes the intensity of the corresponding Poisson process and f denotes the density of the jump-size distribution. In our case (see section 4.1.3), the intensity of $Z^{(\mathrm{CPP})}$ is given by $\frac{\delta\gamma}{2}$, and the jumps, being squares of independent standard normally distributed random variables, each follow a chi-squared distribution with one degree of freedom. The expected value of a chi-squared distributed random variable is its degree of freedom. Hence,

$$\int_0^\infty z\nu(dz) = \frac{1}{\gamma}\frac{\delta\gamma}{2} = \frac{\delta}{2},$$

and, for $i = 1, 2$, we get the representation

$$dZ^{(\mathrm{CPP})}_{\lambda_i t} = \frac{\lambda_i \delta}{2} dt + \tilde{Z}^{(\mathrm{CPP})}_{\lambda_i t}, \tag{4.31}$$

with a martingale $\tilde{Z}^{(\mathrm{CPP})}$.

Combining (4.30) and (4.31), we can now write

$$\begin{aligned} dZ_{\lambda_1 t} &= \frac{\lambda_1 \delta}{2\gamma} (1+\gamma)\, dt + d\tilde{Z}_{\lambda_1 t}, \\ dZ_{\lambda_2 t} &= \frac{\lambda_1 \delta}{2\gamma} (1+\gamma)\, dt + d\tilde{Z}_{\lambda_2 t}, \end{aligned}$$

where $\tilde{Z} = \tilde{Z}^{(\mathrm{IG})} + \tilde{Z}^{(\mathrm{CPP})}$ is a martingale. We adopt this notation in the definition of the model and reformulate it as follows:

$$\begin{aligned} dY^{(1)}_t &= \left(r - \lambda_1 \kappa(\rho_1) - \tfrac{1}{2}\Sigma^{(1)}_t\right) dt + \sqrt{\Sigma^{(1)}_t} dW^{(1)}_t + \rho_1 dZ_{\lambda_1 t}, \\ dY^{(2)}_t &= \left(r - \lambda_2 \kappa(\rho_2) - \tfrac{1}{2}\Sigma^{(2)}_t\right) dt + \sqrt{\Sigma^{(2)}_t} dW^{(2)}_t + \rho_2 dZ_{\lambda_2 t}, \\ d\Sigma^{(1)}_t &= \left(-\lambda_1 \Sigma^{(1)}_t + \frac{\lambda_1 \delta}{2\gamma}(1+\gamma)\right) dt + d\tilde{Z}_{\lambda_1 t}, \\ d\Sigma^{(2)}_t &= \left(-\lambda_2 \Sigma^{(2)}_t + \frac{\lambda_2 \delta}{2\gamma}(1+\gamma)\right) dt + d\tilde{Z}_{\lambda_2 t}. \end{aligned} \tag{4.32}$$

Remark: *From an economic point of view, this model can be advocated in the sense that it should be interpreted as offering two assets whose usual fluctuations behave differently, but which are both affected by the same shocks in the market.*

Remark: *In* (4.32) *we have given the dynamics of all the individual processes of the model. However, when we initially formulate this model in* (4.26) *and* (4.27)*, we have deliberately written also the volatility process in matrix notation (even if the non-diagonal elements are zero), in order to emphasize that this model should still be seen as a special case of the general multivariate OU type SV model* (4.18) - (4.19)*, and thus inherits all the properties derived in the general setting.*

Now that we have formulated the model and we set about deriving the joint characteristic function of $(Y^{(1)}, Y^{(2)})$, we want to exploit that the model is an affine stochastic volatility model in the sense of Keller-Ressel [56]. In order to add clarity and due to the general importance of affine models in mathematical finance, at this stage, we want to undertake at least a very short excursus in the theory of affine models.

We start with the class of affine processes, which has originally been defined by Duffie, Filipović, and Schachermayer [40, Def. 2.1] in terms of a rather general formulation, using, for a state space D, the semigroup $(P_t)_{t\geq 0}$ of operators applied to functions f in the set of bounded Borel-measurable functions $b\mathcal{B}(D)$, which can be associated to a time-homogeneous Markov process $X = (X_t)_{t\geq 0}$, by setting

$$P_t f(x) = \mathbb{E}[f(X_t)|X_t = x], \quad \forall x \in D.$$

We want to use a slightly modified formulation of this definition according to [56], since it fits better in the present context. Compared with the original definition in [40], here stochastic continuity is already part of the definition.

Definition 4.6: *Let $X = (X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ be a stochastically continuous, time-homogeneous Markov process with state space $D = \mathbb{R}_+^m \times \mathbb{R}^n$ and where $\mathbb{P}^x$ denotes the distribution of X with initial value x, i.e., $X_0 = x$, $\mathbb{P}^x$-almost surely. Then X is called an* affine process*, if its characteristic function is an exponentially-affine function of the state vector. This means that on $\mathrm{i}\mathbb{R}^d$ there exist functions $\phi : \mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d \to \mathbb{C}$ and $\psi : \mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d \to \mathbb{C}^d$ such that*

$$\mathbb{E}^x \left[e^{\langle X_t, u\rangle}\right] = e^{\phi(t,u)+\langle x,\psi(t,u)\rangle},$$

for all $x \in D$, and for all $(t, u) \in \mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d$.

A Stochastic Volatility model with asset price process $X = (X_t)_{0\leq t\leq T}$ and volatility process $V = (V_t)_{0\leq t\leq T}$ is then called an *affine Stochastic Volatility model*, if the joint process (X, V) satisfies the following two assumptions:

(i) (X, V) is a stochastically continuous, time-homogeneous Markov process.

(ii) The moment generating function $\Phi_t(u, w)$ of (X_t, V_t) is of particular affine form: There exist functions $C(t, u, w)$ and $D(t, u, w)$ such that

$$\Phi_t(u, w) := \mathbb{E}[\exp(uX_t + wV_t)|X_0, V_0] = e^{uX_0+C(t,u,w)+V_0 D(t,u,w)}.$$

The functions $C(\cdot)$ and $D(\cdot)$ can be obtained as solutions of a system of ordinary differential equations of the (generalized) Riccati type.

It is shown in [56] that the one-dimensional BNS model belongs to this class. The multivariate generalization of the BNS model is explained to be an affine stochastic volatility model in the work of Cuchiero, Filipović, Mayerhofer, and Teichmann [34], which is dedicated to affine processes on positive semidefinite matrices. Moreover, almost all other relevant stochastic volatility models in the literature, as, e.g., the

models of Heston and Bates (see sect. 3.3.2), fall in the class of affine stochastic volatility models (cf. [56]), as well as popular term structure models are usually of the affine type (cf. Cuchiero, Filipović and Teichmann [35]). In general, it can be said that an affine structure of a model offers plenty of analytic tractability and statistical flexibility; due to this fact, these kind of models are predominantly used in the finance literature.

Let us now get back to our model. Equipped with the knowledge of the theory mentioned above, in order to derive the joint characteristic function of $(Y^{(1)}, Y^{(2)})$ from (4.32), we define

$$f(\vec{y}, \vec{v}, t, \vec{u}) := \phi_{Y_T|\mathcal{F}_t}(\vec{u}) = \mathbb{E}^{\mathbb{Q}}\left[e^{\mathrm{i}\vec{u}Y_T} \mid \vec{y} = Y_t, \vec{v} = \Sigma_t\right] = e^{\mathrm{i}\vec{u}\vec{y} + C(T-t) + \vec{v}\vec{D}(T-t)}. \quad (4.33)$$

We use the notation with arrows on top to emphasize vectors here. We want to apply Itō's Lemma on (4.33), so we first have a look at all the quadratic covariations between the occuring processes (recall that quadratic covariation is symmetric):

$$\begin{aligned}
d[y^{(1)}, y^{(1)}] &= \Sigma_t^{(1)} dt + \rho_1^2 d[Z]_{\lambda_1 t}, \\
d[y^{(2)}, y^{(2)}] &= \Sigma_t^{(2)} dt + \rho_2^2 d[Z]_{\lambda_2 t}, \\
d[y^{(1)}, y^{(2)}] &= 0, \\
d[v^{(1)}, v^{(1)}] &= d[\tilde{Z}]_{\lambda_1 t}, \\
d[v^{(2)}, v^{(2)}] &= d[\tilde{Z}]_{\lambda_2 t}, \\
d[v^{(1)}, v^{(2)}] &= 0, \\
d[y^{(1)}, v^{(1)}] &= \rho_1 d[Z, \tilde{Z}]_{\lambda_1 t}, \\
d[y^{(2)}, v^{(2)}] &= \rho_2 d[Z, \tilde{Z}]_{\lambda_2 t}, \\
d[y^{(1)}, v^{(2)}] &= 0, \\
d[y^{(2)}, v^{(1)}] &= 0.
\end{aligned}$$

Itō's Lemma then yields

$$\begin{aligned}
df =&\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial y^{(1)}}\left(\left(r-\lambda_1\kappa(\rho_1)-\tfrac{1}{2}v^{(1)}\right)dt+\sqrt{v^{(1)}}dW_t^{(1)}+\rho_1 dZ_{\lambda_1 t}\right)\\
&+\frac{\partial f}{\partial y^{(2)}}\left(\left(r-\lambda_2\kappa(\rho_2)-\tfrac{1}{2}v^{(2)}\right)dt+\sqrt{v^{(2)}}dW_t^{(2)}+\rho_2 dZ_{\lambda_2 t}\right)\\
&+\frac{\partial f}{\partial v^{(1)}}\left(\left(-\lambda_1 v^{(1)}+\frac{\lambda_1\delta}{2\gamma}(1+\gamma)\right)dt+d\tilde{Z}_{\lambda_1 t}\right)\\
&+\frac{\partial f}{\partial v^{(2)}}\left(\left(-\lambda_2 v^{(2)}+\frac{\lambda_2\delta}{2\gamma}(1+\gamma)\right)dt+d\tilde{Z}_{\lambda_2 t}\right)\\
&+\frac{1}{2}\frac{\partial^2 f}{\partial {y^{(1)}}^2}\left(v^{(1)}dt+\rho_1^2 d[Z]_{\lambda_1 t}\right)+\frac{1}{2}\frac{\partial^2 f}{\partial {y^{(2)}}^2}\left(v^{(2)}dt+\rho_2^2 d[Z]_{\lambda_2 t}\right)\\
&+\frac{1}{2}2\frac{\partial^2 f}{\partial y^{(1)}\partial y^{(2)}}\cdot 0\\
&+\frac{1}{2}\frac{\partial^2 f}{\partial {v^{(1)}}^2}d[\tilde{Z}]_{\lambda_1 t}+\frac{1}{2}\frac{\partial^2 f}{\partial {v^{(2)}}^2}d[\tilde{Z}]_{\lambda_2 t}\\
&+\frac{1}{2}2\frac{\partial^2 f}{\partial y^{(1)}\partial v^{(1)}}\rho_1 d[\tilde{Z},Z]_{\lambda_1 t}+\frac{1}{2}2\frac{\partial^2 f}{\partial y^{(2)}\partial v^{(2)}}\rho_2 d[\tilde{Z},Z]_{\lambda_2 t},
\end{aligned}$$

and, setting the drift term equal to zero, we obtain

$$\begin{aligned}
0 =&\frac{\partial f}{\partial t}+\frac{\partial f}{\partial y^{(1)}}\left(r-\lambda_1\kappa(\rho_1)-\tfrac{1}{2}v^{(1)}\right)+\frac{\partial f}{\partial y^{(2)}}\left(r-\lambda_2\kappa(\rho_2)-\tfrac{1}{2}v^{(2)}\right)\\
&+\frac{\partial f}{\partial v^{(1)}}\left(-\lambda_1 v^{(1)}+\frac{\lambda_1\delta}{2\gamma}(1+\gamma)\right)+\frac{\partial f}{\partial v^{(2)}}\left(-\lambda_2 v^{(2)}+\frac{\lambda_2\delta}{2\gamma}(1+\gamma)\right)\\
&+\frac{1}{2}\frac{\partial^2 f}{\partial {y^{(1)}}^2}v^{(1)}+\frac{1}{2}\frac{\partial^2 f}{\partial {y^{(2)}}^2}v^{(2)}\\
=&-C'-\vec{v}\vec{D}'+\mathrm{i}u^{(1)}\left(r-\lambda_1\kappa(\rho_1)-\tfrac{1}{2}v^{(1)}\right)+\mathrm{i}u^{(2)}\left(r-\lambda_2\kappa(\rho_2)-\tfrac{1}{2}v^{(2)}\right)\\
&+D^{(1)}\left(-\lambda_1 v^{(1)}+\frac{\lambda_1\delta}{2\gamma}(1+\gamma)\right)+D^{(2)}\left(-\lambda_2 v^{(2)}+\frac{\lambda_2\delta}{2\gamma}(1+\gamma)\right)\\
&+\frac{1}{2}\mathrm{i}{u^{(1)}}^2 v^{(1)}+\frac{1}{2}\mathrm{i}{u^{(2)}}^2 v^{(2)}.
\end{aligned}$$

By rearranging, we get

$$\begin{aligned}
C'+\vec{v}\vec{D}' =&\left(-\frac{1}{2}\mathrm{i}u^{(1)}-\lambda_1 D^{(1)}-\frac{1}{2}{u^{(1)}}^2\right)v^{(1)}\\
&+\left(-\frac{1}{2}\mathrm{i}u^{(2)}-\lambda_2 D^{(2)}-\frac{1}{2}{u^{(2)}}^2\right)v^{(2)}\\
&+\mathrm{i}u^{(1)}\left(r-\lambda_1\kappa(\rho_1)\right)+\mathrm{i}u^{(2)}\left(r-\lambda_2\kappa(\rho_2)\right)\\
&+\frac{\lambda_1\delta}{2\gamma}(1+\gamma)D^{(1)}+\frac{\lambda_2\delta}{2\gamma}(1+\gamma)D^{(2)},
\end{aligned}\tag{4.34}$$

and, by comparing coefficients, we finally obtain

$$\begin{aligned}
D^{(1)\prime} &= -\frac{1}{2}\mathrm{i}u^{(1)} - \lambda_1 D^{(1)} - \frac{1}{2}u^{(1)2}, \quad D^{(1)\prime}(0) = 0, \\
D^{(2)\prime} &= -\frac{1}{2}\mathrm{i}u^{(2)} - \lambda_2 D^{(2)} - \frac{1}{2}u^{(2)2}, \quad D^{(2)\prime}(0) = 0, \\
C' &= \mathrm{i}u^{(1)}\left(r - \lambda_1\kappa(\rho_1)\right) + \mathrm{i}u^{(2)}\left(r - \lambda_2\kappa(\rho_2)\right) \\
&\quad + \frac{\lambda_1\delta}{2\gamma}\left(1+\gamma\right)D^{(1)} + \frac{\lambda_2\delta}{2\gamma}\left(1+\gamma\right)D^{(2)}, \quad C(0) = 0.
\end{aligned}$$

This is a system of ordinary linear differential equations, where the solution is given by:

$$D^{(1)}(x) = \frac{e^{-\lambda_1 x}\left(1 - e^{\lambda_1 x}\right)u^{(1)}(\mathrm{i} + u^{(1)})}{2\lambda_1}, \tag{4.35}$$

$$D^{(2)}(x) = \frac{e^{-\lambda_2 x}\left(1 - e^{\lambda_2 x}\right)u^{(2)}(\mathrm{i} + u^{(2)})}{2\lambda_2}, \tag{4.36}$$

$$\begin{aligned}
C(x) &= \mathrm{i}u^{(1)}x\left(r - \lambda_1\kappa(\rho_1)\right) + \mathrm{i}u^{(2)}x\left(r - \lambda_2\kappa(\rho_2)\right) \\
&\quad - \frac{\delta(1+\gamma)u^{(1)}(\mathrm{i} + u^{(1)})(\lambda_1 x + e^{-\lambda_1 x} - 1)}{4\lambda_1\gamma} \\
&\quad - \frac{\delta(1+\gamma)u^{(2)}(\mathrm{i} + u^{(2)})(\lambda_2 x + e^{-\lambda_2 x} - 1)}{4\lambda_2\gamma}.
\end{aligned} \tag{4.37}$$

The cumulant function $\kappa(\theta)$ of the BDLP of $\Sigma^{(1)}$ and $\Sigma^{(2)}$ is (see (4.15)) given by

$$\kappa(\theta) = \frac{\delta\theta}{\sqrt{\gamma^2 - 2\theta}}.$$

Putting together the results obtained above, we can now formulate the following proposition:

Proposition 4.4: *The conditional joint characteristic function of $(Y^{(1)}, Y^{(2)})$, at time T, given the information at time $t \leq T$, i.e. $Y_t = \vec{y}$ and $\Sigma_t = \vec{v}$, in the model* (4.26) - (4.27)*, is given by*

$$\phi_{Y_T|\mathcal{F}_t}(\vec{u}) = e^{\mathrm{i}\vec{u}\vec{y} + C(T-t) + \vec{v}\vec{D}(T-t)},$$

where the functions $C(\cdot), D^{(1)}(\cdot)$, and $D^{(2)}(\cdot)$ are given by (4.35) - (4.37)*.*

Remark: *In principle, another interesting approach to model the dependence between the two underlying assets would be to use a correlated Brownian motion, i.e., a two-dimensional BM $W = (W^{(1)}, W^{(2)})$ such that $dW_t^{(1)}dW_t^{(2)} = \varrho dt$, for $\varrho \in (-1, 1)$. However, when following the strategy as above, it turns out that this*

would correspond to abandoning the affine structure of the BNS model. The point is that in this case the quadratic covariation of the log-returns of the two underlying assets is $d[Y^{(1)}, Y^{(2)}]_t = \varrho\sqrt{\Sigma_t^{(1)}}\sqrt{\Sigma_t^{(2)}}dt$. *If we, for instance, added this kind of correlation in the model* (4.32) *above, then in equation* (4.34) *we would have an additional term* $\varrho\sqrt{v^{(1)}}\sqrt{v^{(2)}}$, *and we would hence not obtain a system of Riccati-type linear differential equations. It does not seem to be possible to circumvent this issue without using a somehow artificial way of specifying the model.*

Chapter 5

Computational Studies

Let $S_0^{(1)} = 100$ and $S_0^{(2)} = 95$ be the resp. values of two assets at time zero. We investigate a Spread Option on these two assets, where the option has a maturity of one year, i.e. $T = 1$.

5.1 The two-dimensional OU-Wishart model from sect. 4.2.3

For our numerical studies we take the parameter values from [66]. For their calibration of the two-dimensional OU-Wishart model they consider the foreign exchange market: They calibrate the model on the prices of Call Options on the mutual exchange rates between EUR, USD, and GBP on April 29, 2010. Due to the fact that, on the one hand, multi-asset options are mainly traded over-the-counter, as well as, on the other hand, a Call Option on some exchange rate can be seen as a Spread Option on two other ones, this represents a convenient and reasonable approach. The particular values that we use are:

$r = 0.676, \lambda = 0.901, a = -3.008, \gamma_1 = 0.034, \gamma_2 = 0.00,$

$\rho_1 = -5.364, \rho_{12} = 0.679, \rho_{21} = 0.896, \rho_2 = -0.661$

$$\Theta = \begin{pmatrix} 0.011 & 0.023 \\ 0.023 & 0.067 \end{pmatrix}, \Sigma_0 = \begin{pmatrix} 0.019 & 0.013 \\ 0.013 & 0.018 \end{pmatrix}.$$

5.1.1 Pricing by Simulation

Table 5.1 shows the option prices that we have obtained as a result of a Monte-Carlo simulation for a range of strike prices, that is from $2.0 - 4.0$ and from $6.0 - 8.0$ with steps of size 0.5, whereas for the "most relevant" interval between 4.0 and 6.0 we observe prices for a strike-step-size of 0.2. We have used an Euler-Maruyama discretization scheme for the time-interval with 256 points, which approximately corresponds to "daily monitoring" (there are averagely 251 trading days a year).

The number of replications of the simulation that we have run for each strike price is 1 million. In order to enhance the efficiency of our simulation, we have adopted a control variate for the purpose of variance reduction; in particular, we exploit that our discounted asset price processes are martingales, which means that the spread between the two underlying assets itself serves perfectly well as a control variate.

Table 5.1: MC Simulation results

STRIKE	OPTION VALUE	CONFIDENCE INTERVAL
2.0	6.7029	$[6.6937, 6.7120]$
2.5	6.4014	$[6.3923, 6.4105]$
3.0	6.0976	$[6.0887, 6.1066]$
3.5	5.8099	$[5.8010, 5.8188]$
4.0	5.5372	$[5.5284, 5.5461]$
4.2	5.4148	$[5.4060, 5.4236]$
4.4	5.3089	$[5.3001, 5.3177]$
4.6	5.1996	$[5.1908, 5.2084]$
4.8	5.0974	$[5.0886, 5.1061]$
5.0	4.9883	$[4.9796, 4.9971]$
5.2	4.8865	$[4.8778, 4.8952]$
5.4	4.7938	$[4.7850, 4.8025]$
5.6	4.6846	$[4.6760, 4.6933]$
5.8	4.5870	$[4.5783, 4.5957]$
6.0	4.4846	$[4.4760, 4.4933]$
6.5	4.2487	$[4.2401, 4.2574]$
7.0	4.0167	$[4.0081, 4.0252]$
7.5	3.7989	$[3.7903, 3.8075]$
8.0	3.5832	$[3.5746, 3.5917]$

In Table 5.2, we illustrate the variance reduction in our simulation that we obtain by the usage of the control variate; we compare the resp. confidence intervals with and without using the control variate, for a range of strikes, and observe the narrowing of the intervals due to this variance reduction technique.

Table 5.2: Effect of control variate

STRIKE	CI WITHOUT CV	CI WITH CV
4.0	$[5.5188, 5.5486]$	$[5.5284, 5.5461]$
4.2	$[5.4000, 5.4294]$	$[5.4060, 5.4236]$
4.4	$[5.2893, 5.3185]$	$[5.3001, 5.3177]$
4.6	$[5.1887, 5.2176]$	$[5.1908, 5.2084]$
4.8	$[5.0758, 5.1044]$	$[5.0886, 5.1061]$
5.0	$[4.9717, 5.0001]$	$[4.9796, 4.9971]$

Table 5.3 illustrates the impact of the choice of the number of discrete time points N in the interval $[0, 1]$. We notice that in some instances - see, e.g., the case where the strike price is set to be $K = 4.0$ (which was the most extreme case we observed) - the confidence interval can be shifted by increasing N. This means that the discretization error is rather big in this case. In most cases, however, the results turn out to be rather stable in the sense that an increase of N does not have a remarkable impact.

Table 5.3: Impact of discretization

STRIKE: 4.0		STRIKE: 4.5		STRIKE: 5.0	
N	CI (with CV)	N	CI (with CV)	N	CI (with CV)
256	$[5.5284, 5.5461]$	256	$[5.2433, 5.2608]$	256	$[4.9796, 4.9971]$
512	$[5.5070, 5.5247]$	512	$[5.2440, 5.2615]$	512	$[4.9798, 4.9973]$
		1024	$[5.2423, 5.2599]$		

Now we want to use the results of the Monte-Carlo simulation as benchmark values for a comparison with the results obtained by the FFT method.

5.1.2 Pricing by the FFT Method

Table 5.4 shows the option prices that we obtain by the two different implementations of the FFT method described in chapter 2, i.e., the method where, for each strike, we choose the grid according to Algorithm 2.1 and the method where we run the Fourier inversion once and use third-degree polynomials in order to interpolate between grid points. We choose the truncation interval as $[-\bar{u}, \bar{u}] = [-40, 40]$ and the number of grid points on each axis in the resp. interval as $N = 256$. For the grid-selection algorithm we start with the interval $[-\bar{u}_0, \bar{u}_0] = [-40, 40]$. Column "MC" contains the resp. Monte-Carlo estimators.

Table 5.4: Comparison of FFT method(s) and MC simulation
$[-\bar{u}, \bar{u}] = [-40, 40], N = 256, \bar{u}_0 = 40, \epsilon = [-3, 1]$

STRIKE	MC	FFT (grid selection)	FFT (interpolation)
3.0	6.0976	6.0233	6.1845
3.5	5.8099	5.7339	5.7261
4.0	5.5372	5.4529	5.5234
4.2	5.4148	5.3429	5.3395
4.4	5.3089	5.2342	5.3159
4.6	5.1996	5.1330	5.1560
4.8	5.0974	5.0211	5.0336
5.0	4.9883	4.9159	4.8673
5.2	4.8865	4.8135	4.8820
5.4	4.7938	4.7112	4.8253
5.6	4.6846	4.6114	4.7764
5.8	4.5870	4.5197	4.7489
6.0	4.4846	4.4150	4.5048
6.5	4.2487	4.1774	4.2482
7.0	4.0167	3.9486	3.9700

By looking at Table 5.4, we notice that the values obtained by the two different FFT implementations differ notably. We observe that the results of the interpolation procedure between grid points may match the results of the Monte-Carlo simulation extremely well in some cases (e.g., $K = 6.5$); however, this only happens if we "are lucky". If we take a look at those values where the interpolation is based on, we notice that they represent quite a big range of prices. In this case, we cannot consider the interpolation method as satisfactorily reliable. Table 5.5 shows some results of our investigation of the impact of the choices of the parameters $\bar{u}$ and N, i.e., of the truncation interval and the number of grid points in this interval, respectively. The matrices given in the resp. columns show the prices obtained for the grid points in the "neighbourhood" of that point which corresponds to the actual option characteristics (i.e. the values $\log(S_0^{(1)}/K)$ on the x-axis, and $\log(S_0^{(2)}/K)$ on the y-axis).

Table 5.5: Investigation of approximation method

$\bar{u} = 40$	$\bar{u} = 80$	$\bar{u} = 120$
$N = 256$	$N = 1024$	$N = 1024$
$\begin{pmatrix} 5.5131 & 2.4594 & 0.9293 \\ 11.0005 & 6.2077 & 2.7988 \\ 18.0754 & 12.2256 & 6.9644 \end{pmatrix}$	$\begin{pmatrix} 4.0320 & 2.6247 & 1.6334 \\ 6.2075 & 4.2882 & 2.7983 \\ 8.9644 & 6.5778 & 4.5560 \end{pmatrix}$	$\begin{pmatrix} 4.6968 & 3.5967 & 2.6885 \\ 6.2153 & 4.8904 & 3.7493 \\ 7.9934 & 6.4609 & 5.0896 \end{pmatrix}$

We notice that, for the interpolation method, the truncation $[-40, 40]$ is too small to yield reliable results. Stretching the truncation interval, we shall of course also increase the number of points N, since otherwise the discretization error will get big. For the truncation interval $[-120, 120]$ and $N = 1024$, we observe that the prices of options corresponding to "neighbouring" lattice points are reasonably close to each other. Therefore, an interpolation makes sense. Table 5.6 outlines the reliability of the results by comparing them with the results of the method where we choose the grid for each given strike: the prices obtained by the two different implemenations are rather close.

Table 5.6: Comparison of FFT method(s) and MC simulation
$[-\bar{u}, \bar{u}] = [-120, 120], N = 1024, \bar{u}_0 = 120, \epsilon = [-3, 1]$

STRIKE	MC	FFT (grid selection)	FFT (interpolation)
3.0	6.0976	6.0229	6.0221
3.5	5.8099	5.7326	5.7386
4.0	5.5372	5.4599	5.4610
4.2	5.4148	5.3418	5.3537
4.4	5.3089	5.2342	5.2369
4.6	5.1996	5.1270	5.1280
4.8	5.0974	5.0285	5.0290
5.0	4.9883	4.9165	4.9225
5.2	4.8865	4.8206	4.8207
5.4	4.7938	4.7116	4.7325
5.6	4.6846	4.6113	4.6179
5.8	4.5870	4.5185	4.5240
6.0	4.4846	4.4149	4.4366
6.5	4.2487	4.1773	4.2068
7.0	4.0167	3.9485	3.9481

As opposed to the interpolation method, if we compare the results of Table 5.4 and Table 5.6 regarding the method with the grid selection, we notice that the resp. results are very close to each other. Therefore, we may be confident that for this method the truncation interval $[-40, 40]$, with 256 steps therein, is already sufficient.

Now we want to compare the computational time that is required to obtain option prices by the different methods. We have implemented all of the procedures in *MATLAB* and run them on a standard PC. Table 5.7 shows the time that elapsed on the CPU in the resp. case, while running the FFT method (with gridselection) for a single contract for various numbers N of discrete points within the truncation interval, in comparison with the required time for a Monte-Carlo simulation consisting of $n = 1000000$ replications and N discrete time points.

Table 5.7: Elapsed CPU time for the different methods

Method	Time
FFT: $N = 2^8$	$\sim$ 17 sec.
FFT: $N = 2^9$	$\sim$ 68 sec.
FFT: $N = 2^{10}$	$\sim$ 4.5 min.
MC: $n = 10^6, N = 2^8$	$\sim$ 6 h.
MC: $n = 10^6, N = 2^9$	$\sim$ 11.7 h.
MC: $n = 10^6, N = 2^{10}$	$\sim$ 23 h.

We notice that, regarding computation times, there is by no means a discussion about the outstanding performance of the FFT method. In many situations, such as, e.g., the calibration of a market model, the long time that is required to price options via simulation is not acceptable. The FFT method yields the option price within about 17 seconds. Since, in such situations, one needs the prices for a large number of contracts, the FFT method where we interpolate between grid points is usually favourable. We found out above that this methods requires a larger truncation interval (say $[-120, 120]$), coinciding with a higher number of necessary grid points (say 1024), in order to yield the "same" results as in the case where we choose the grid for each contract. Therefore, in fact, 17 seconds face 4.5 minutes regarding a single contract. On the other hand, we should emphasize once again that the interpolation method means that we only have to run the Fourier inversion once (yielding prices for a whole set of different contracts simultaneously) and then, by interpolation, we obtain prices for any new contract within very little time. This is

the reason why the FFT method with interpolation is usually unchallenged by other methods, when it comes to simultaneously pricing many contracts. However, in our case, we have observed that we need to use many more discrete points N to obtain close results with the interpolation method. Then, after we have run the Fourier inversion once (which takes for around 4.5 minutes, as mentioned above), the interpolation (with a third degree polynomial) for every further contract yet (averagely) takes another 20 seconds. Therefore, we are still in favour of our method with the grid selection, even regarding the pricing of a high number of contracts.

Our hitherto investigations within this chapter have shown the behaviour of the FFT method: we have illustrated for which parameter choices regarding truncation and discretization we may expect reliable results, and we have outlined the great advantage in terms of the computational speed. However, if we compare the resulting option prices with those obtained by the Monte-Carlo method, the variation is not negligible. It might be the case that, due to the jumps in the trajectories, some simulation results change a bit if one uses higher numbers of discrete points. We already indicated this above, where we considered some examples (cf. Table 5.3). However, we have seen that with 1024 points, together with one million replications, it already takes almost one day to evaluate a single contract. Hence, it gets tedious to further increase the number of discretization points. In any case, regarding the FFT method, there is one parameter left that we have not investigated so far: the vector $\epsilon = [\epsilon_1, \epsilon_2]$, which is used to circumvent the issue with the integrability of the payoff function.

On the choice of ϵ

We can see ϵ as the extension to two dimensions of what Carr and Madan [27] call the damping factor α, as described in section 2.1.2. In theory, this is only a technical finesse. In practice, however, recall that the modified function is subject to the numerical two-dimensional Fourier inversion. The only condition that we have required was $\epsilon_1 + \epsilon_2 < -1$, where $\epsilon_1, \epsilon_2 \in \mathbb{R}$ with $\epsilon_2 > 0$. It seems evident that a reasonable choice of ϵ is neither obvious, nor is it a negligible issue. Particularly in the two-dimensional case, the choice of ϵ is, in fact, a crucial part of the pricing model. Table 5.8 shows some selected examples, outlining the strong impact of this parameter.

Table 5.8: Investigation of the sensitivity of the model to changes in ϵ.
Method with grid selection; $\bar{u} = [-40, 40], N = 256, K = 5$

EPSILON $\epsilon = [\epsilon_1, \epsilon_2]$	OPTION PRICE
$[-3, 1]$	4.9159
$[-3.1, 1.79]$	4.9567
$[-3.5, 1]$	4.9166
$[-5, 1]$	4.9164
$[-10, 8]$	4.7750
$[-7, 3]$	3.8877
$[-9, 4]$	8.4779
$[-8, 1]$	17190
$[-30, 1]$	$2618685340536278 \cdot 10^{18}$

We observe that, apparently, there is a region of "indeed feasible" choices; the possible impact of ϵ in general is worrying, though. For all the numerical results shown so far, i.e. before Table 5.8, we have always assumed $\epsilon = [-3, 1]$. Hurd and Zhou [51] have shown that there is an upper bound on $\hat{P}$ if one chooses some $\epsilon_2 > 0$ and sets $\epsilon_1 = -1 - 2\epsilon_2$. However, this represented our only reference point and we cannot present a motivation for our choice apart from empirical convenience. Hurd and Zhou neither tell us about their specific choices for ϵ for their numerical studies. However, if we examplarily consider their results for Spread Option prices within the two-factor GBM model (see [51], pg. 151), for a strike of 2.0, they obtain a relative error of magnitude 10^{-8} using a polynomial with a degree of 8 for interpolation, while the magnitude of our relative error, for the same model and parameter values, interpolating with a third-degree polynomial, is 10^{-4} for $\epsilon = [-3, 1]$. Moreover, the choice $\epsilon = [-3, 1]$ seemed to be reasonable for all of our further checkings as well. That there is no straightforward way which leads to an objectively superior choice of ϵ is also indicated in the work of Hurd and Zhou, where they mention that "the selection of suitable values for ϵ, N, and η" when implementing this kind of FFT approximation "is a somewhat subtle issue whose details depend on the asset model in question".

Even in the one-dimensional case, the particular choice of the damping parameter α is not at all a crystal-clear issue. Most authors use values that have empirically turned out to be convenient, or they apply rules of thumb: Carr and Madan [27] suggest to derive an upper bound on α from the condition $\mathbb{E}\left[S_T^{\alpha+1}\right] < \infty$ and then take one fourth of this upper bound. Other authors suggest ad-hoc choices; Raible [73], for example, considers the value $\alpha = 25$ as the best choice for the

models he studies. A rigorous investigation of the issue of an appropriate choice of α has been presented by Lord and Kahl [60]: they address the problem that, as they formulate it in particular, "by changing α, the integrand can become either strongly peaked when getting close to the poles of the integrand, or highly oscillatory when reaching the maximum allowed α", by formulating an optimality criterion and solving the resulting optimization problem. In particular, they both suggest an optimal choice of α depending on the payoff as well as a payoff-independent alternative of choosing the damping parameter.

In the case of Spread Options, i.e., in the case of two dimensions, the method of Lord and Kahl cannot be applied directly. We face a two-dimensional vector ϵ instead of the one-dimensional α, which renders it more complicated to formulate an appropriate optimization problem. For a satisfactory degree of reliability of the pricing method it would be necessary, though, to pose some optimality criterion on ϵ. This would be a way to detach the model from any empirically motivated ad-hoc decisions and add confidence about the results. However, we do not consider this as a scope of this thesis and hence do not try to follow this path here.

Since, as described above, the impact of the choice of ϵ does not indeed allow holistic satisfaction with the method, we now want to check alternative approaches for applicability in the context of Spread Options, with the intention to possibly avoid the concept of multiplying the payoff in order to ensure integrability.

The method with time-values

In section 2.1.2 we described the idea of Carr and Madan [27] to use the time-value of an option, i.e., to subtract the intrinsic value from the option price, in order to avoid the damping factor but ensure integrability of the modified option price as a function in strike. However, this idea is unfortunately not extendable to Spread Options: What Hurd and Zhou have suggested as an extension of the Carr and Madan method to the case of Spread Options, as we described in detail in section 2.1.3, is that they have found a representation of the modified (multiplied by a factor which ensures integrability) payoff function (not option price as function in strike) by Fourier inversion; the (discounted) expected value then corresponds to a representation of the option price which contains the joint characteristic function of the underlying assets.

The idea with subtracting the intrinsic value only makes sense if one considers the option price as a function in strike and wants to circumvent the problem that this option pricing function is not (square-)integrable. This works perfectly well in the

one-dimensional case: they've worked out the details for a Call Option, where they obtain a closed-form expression for the Fourier transform of the option's time-value, which contains the characteristic function of the underlying asset. Consequently, Fourier inversion yields the desired representation.

However, considering Spread Options, the problem with this strategy is that the strike is still (as for a vanilla option) one-dimensional, whereas now one faces two asset-price processes and the density depends on both assets. Following an analogous strategy as for a Call Option, i.e., to start with the option's time-value as a function in strike and proceed with rearranging its Fourier transform, hence does not render it possible to obtain a closed-form expression of the Fourier transform of the time-value, which contains the joint characteristic function of the two underlying assets.

The Integration-Along-Cut method

As we have seen, the FFT method has its drawbacks; particularly, we are not glad about the fact that there is no general effective procedure available in order to choose the damping parameter (which can also be seen as choosing the line of integration). A different approach, which does not make use of the FFT and where there is no need for a damping parameter, has been suggested in the book of Boyarchenko and Levendorskiĭ [23]. Their method, which they call the Integration-Along-Cut (IAC) method, allows for an effective control of the computational error and, moreover, in many cases they even observe advantages with respect to the FFT method in terms of the computational speed when pricing single options. We therefore want to take a closer look at the IAC method and check its applicability for our purposes. We do not want to go into full detail in terms of technicalities but rather comment on the crucial aspects. For details see [23].

The method is presented for a market model where the log-asset price is assumed to follow a CGMY$(c, -\lambda_-, \lambda_+, \nu)$ (the authors prefer the less common term KoBoL in their original work, which is an acronym for Koponen, Boyarchenko, Levendorskiĭ) process of order $\nu \in (0, 1)$. The skewness parameters are assumed to satisfy $\lambda_- < -1 < 0 < \lambda_+$. Consider a European Put Option with strike price $K = 1$, maturity T, and spot price of the asset given by $S_0 = e^x$. The Fourier transform of the payoff

is given by

$$\begin{aligned}\hat{P}(\xi) &= \int_{-\infty}^{\infty} e^{-\mathrm{i}x\xi}(1-e^x)^+ dx \\ &= \frac{1}{(-\mathrm{i}\xi)(-\mathrm{i}\xi+1)}.\end{aligned}$$

It is well-defined in the half-plane $\mathrm{Im}(\xi) > 0$ and, having two poles in $\xi = 0$ and $\xi = -\mathrm{i}$, it admits a meromorphic extension into the complex plane. We set $\kappa := \lambda_+ - \lambda_-, \chi_+ := \lambda_+/\kappa, \chi_- := -\lambda_-/\kappa, \tau := T - t, c_\nu := c \cdot \Gamma(-\nu), U := \kappa(x + \tau\mu), V := -\kappa^\nu \tau c_\nu, \rho := 1/\kappa$, and $\sigma_1 = \sigma/\kappa$, where $\sigma \in (0, \lambda_+)$. Then they show that the price of a Put Option at time t can be written as

$$f(x,t) = R_1 \int_{-\infty+\mathrm{i}\sigma_1}^{+\infty+\mathrm{i}\sigma_1} \frac{\exp[\mathrm{i}U\xi - V((\chi_- - \mathrm{i}\xi)^\nu + (\chi_+ + \mathrm{i}\xi)^\nu)]}{-\mathrm{i}\xi(-\mathrm{i}\xi+\rho)} d\xi, \tag{5.1}$$

where

$$R_1 = \frac{\exp(-\tau r + V(\chi_-^\nu + \chi_+^\nu))}{2\pi\kappa}.$$

The integral in (5.1) can now be computed by applying the IAC method.

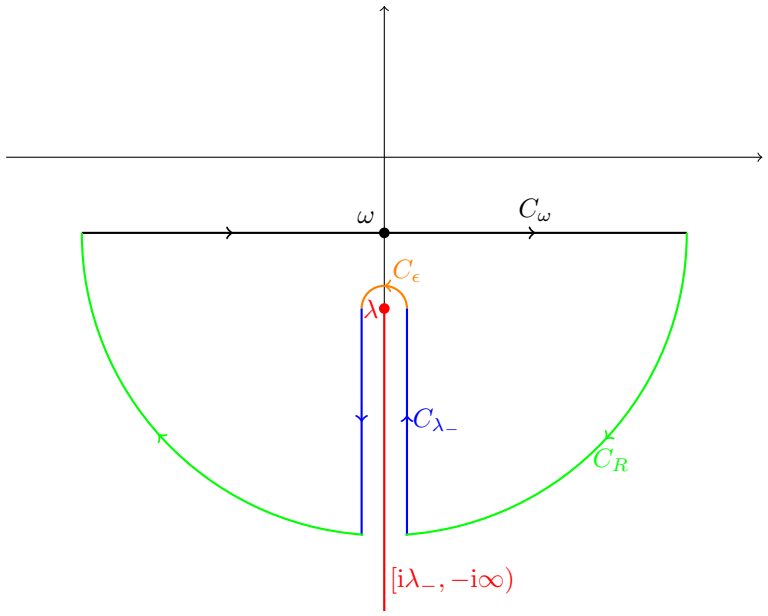
In principle, the idea of the method is the following: Usually, one knows the strip of regularity of the characteristic function of the log-asset price process. In particular, the characteristic function is analytic on the whole complex plane except for the discontinuities along two cuts of the type $(-\mathrm{i}\infty, \mathrm{i}\zeta_-]$ and $[\mathrm{i}\zeta_+, \mathrm{i}\infty)$. What Boyarchenko and Levendorskiĭ now propose is, instead of considering an integral of the type $\int_{-\infty+\mathrm{i}\omega}^{+\infty+\mathrm{i}\omega} f(x)dx$, where f is, in essence, a combination of the characteristic function of the log-asset price process and the Fourier transform of the payoff function, to apply Cauchy's integral theorem and transform the contour of integration to one of the cuts. The shifting of the integration contour is exemplarily sketched in Figure 5.1.

By Cauchy's theorem, we know that the integral over any closed path equals zero, i.e.,

$$\oint_C f(x)dx = \int_{C_\omega} f(x)dx + \int_{C_R} f(x)dx + \int_{C_\epsilon} f(x)dx + \int_{C_{\lambda_-}} f(x)dx = 0.$$

As $R := |x| \to +\infty$, the integral $\int_{C_R} f(x)dx \to 0$, and as $\epsilon \to 0$, the integral

Figure 5.1: Contour shift



$\int_{C_\epsilon} f(x)dx \to 0$. Therefore, we can write

$$\int_{\mathrm{Im}(x)=\omega} f(x)dx = \int_{-\mathrm{i}\infty-0}^{\mathrm{i}\lambda_- -0} f(x)dx + \int_{\mathrm{i}\lambda_-+0}^{-\mathrm{i}\infty+0} f(x)dx.$$

Let us now get back to the particular case of the CGMY process, where we "stopped" at (5.1):

In the case $U \geq 0$, the contour of integration is shifted to the cut in the upper half-plane (the cuts result from the strip of analyticity of the characteristic function of a CGMY process). The choice of the cut is due to the intention that $\mathrm{i}U\xi$ be negative at the cut, i.e., $\exp(\mathrm{i}U\xi)$ shall not grow to infinity. By setting $a_\nu := \cos(\nu\pi), b_\nu := \sin(\nu\pi)$ and performing some rearranging steps, including variable substitutions, the option price can be written as

$$f(x,t) = 2R_1 e^{-U\chi_+} \mathcal{I}(\nu, U, V, \chi_+, \chi_+ + \rho), \tag{5.2}$$

where

$$\mathcal{I}(\nu, u, v, \rho_1, \rho_2) := \int_0^\infty \frac{e^{-zu}\sin(vb_\nu z^\nu)}{(z+\rho_1)(z+\rho_2)} \exp(-v[(1+z)^\nu + a_\nu z^\nu])dz. \tag{5.3}$$

The integral (5.3) can be evaluated using any numerical integration prodecure, yielding the option price as a result of (5.2), for the case $U \geq 0$.

In the case $U < 0$, the contour of integration is shifted to the cut in the lower half-plane. This transformation means that one has to take into account the two poles $\xi = 0$ and $\xi = -\mathrm{i}$. The residue theorem can then be used to express the integral in terms of the new contour, and by rearranging and changing variables one can obtain the representation

$$f(x,t) = \exp(-\tau r) - \exp(x) + 2R_1 e^{U\chi_-}\mathcal{I}(\nu, -U, V, \chi_-, \chi_- - \rho),$$

with the function $\mathcal{I}$ as defined in (5.3).

Now we have seen how the IAC technique works and that it is very well suited for the pricing of vanilla options in the CGMY model. It is also applicable in many other situations, and the performance of the method is particularly good if the option is close to maturity. However, the method is not generally applicable to any model; for example, Lord and Kahl [60] mention some cases where problems arise. Let us now take a look at our situation of Spread Options in the two-dimensional OU type SV model (where the OU Wishart model is a special case thereof). Theorem 4.6 gives information about the regularity of the moment-generating function in the multi-dimensional OU type SV model. The strip of analyticity of the characteristic function is consequently given by

$$S_\theta^\phi := \{y \in \mathbb{C}^d : \| \operatorname{Im}(y)\| < \theta\},$$

with θ as defined in the theorem. As a consequence of Theorem 2.1, equation (2.11) gives the price of a Spread Option as

$$\operatorname{Spr}(X_0; T) = \frac{1}{(2\pi)^2} e^{-rT} \int\int_{\mathbb{R}^2 + \mathrm{i}\epsilon} e^{\mathrm{i}uX_0^t} \phi_{X_T}(u) \hat{P}(u) d^2u,$$

where

$$\hat{P}(u) = \frac{\Gamma(\mathrm{i}(u_1 + u_2) - 1)\Gamma(-\mathrm{i}u_2)}{\Gamma(\mathrm{i}u_1 + 1)}.$$

The Gamma function has poles of first order in all non-positive integers. Therefore, shifting the contour of integration to the cuts means that we cross infinitely many poles. A straightforward application of the residue theorem is thus not possible and the IAC method is in further consequence not easily applicable.

There are generalizations of the residue theorem, which deal with infinite numbers of isolated singularities; roughly speaking, the real axis is "closed at infinity" by a semi-circle. Carefulness is required in this case, though, and some more work will

be necessary in order to succeed in working out the full details for Spread Options in the two-dimensional OU type SV model. However, particularly the potential of the method in terms of error control makes it appealing; hence, this should be an interesting subject for future work.

5.1.3 Comments on the Behaviour of the Market Model

So far in this section, we have studied the pricing of Spread Options in the two-dimensional OU-Wishart model. By investigating the different pricing techniques, we have also seen the resulting option prices within this model. The parameter values we used stem from the literature. The two-dimensional OU-Wishart model includes numerous parameters; recall that a matrix-valued compound Poisson process is used as driving process, involving already four different parameters $(\lambda, \Theta_{11}, \Theta_{12}, \Theta_{22})$, which in turn brings along four leverage parameters $(\rho_1, \rho_{12}, \rho_{21}, \rho_2)$. The plethora of parameters together with their mutual influences, do not allow for a revealing investigation of the particular impact of most of the model parameters; in many cases, the interconnections are too complicated to make it meaningful to study the variation in single parameters. The only two parameters we want to investigate are the mean reversion speed a of the volatility process, as well as the intensity λ of the driving CPP: Figure 5.2 and Figure 5.3 show the results of varying the resp. parameter values:

We observe the parameter a in the interval $[-10, -0.5]$, where we notice an exponential increase in option prices with growing values for a. Varying the intensity λ of the CPP between 0.1 and 2.0, we observe that the option price decreases approx. linearly. The range of obtained prices lies within one unit, though.

Figure 5.2: Option values for varying a

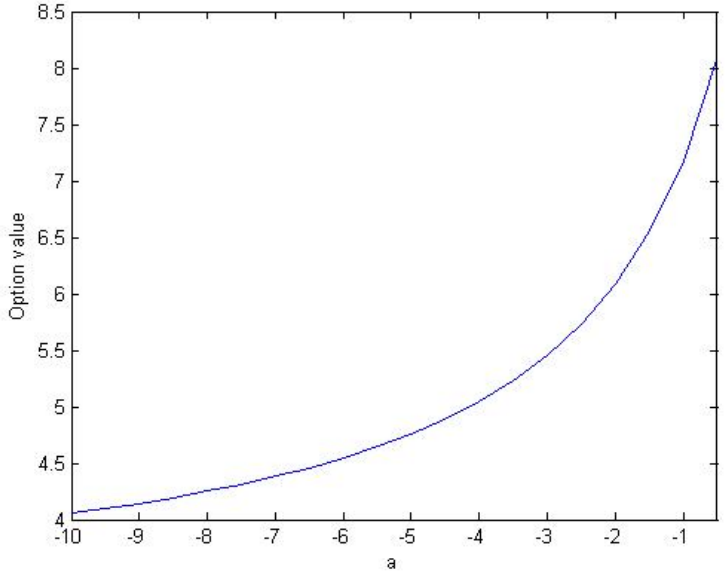
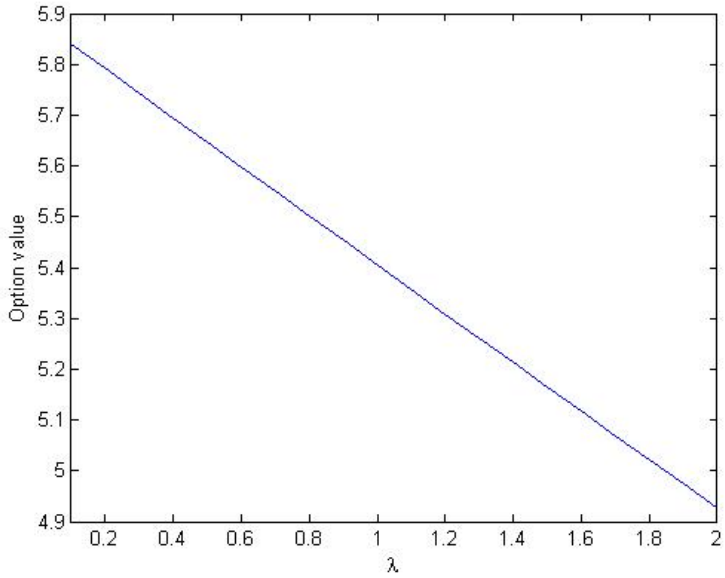


Figure 5.3: Option values for varying λ



5.2 The two-dimensional IG-OU model from sect. 4.2.4

Compared to the OU-Wishart model, the two-dimensional model of IG-OU type, as we formulated it in section 4.2.4, is not so complicated: in this model we consider two one-dimensional volatility processes, which means that there are fewer parameters

involved. The significance of the single parameters is hence clearer. However, for a rigorous survey of the model, it would be necessary to calibrate it to market prices; this is beyond the scope of this thesis, though. We orientate ourselves towards the parameter values obtained by Nicolato and Venardos [68] for their calibration of the one-dimensional BNS model and observe our model in terms of its sensitivity to changes in the parameters. Let us emphasize that we do not investigate the obtained absolute prices; such a discussion would not be feasible due to the lack of an appropriate calibration. The following investigations should rather be seen as a first testing in order to gain some understanding of the model behaviour.

Figure 5.4 illustrates the impact of the initial volatility level; increasing volatility results in increasing option prices.

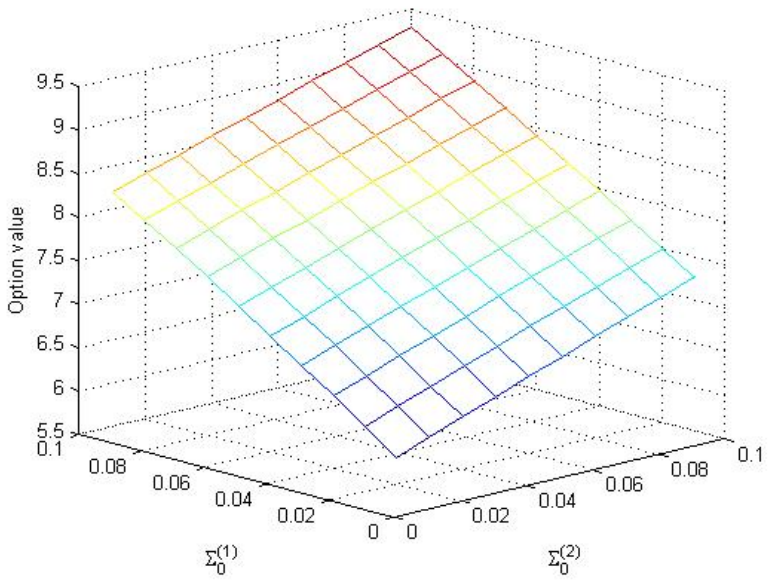


Figure 5.4: Option prices for varying $\Sigma_0^{(1)}$ and $\Sigma_0^{(2)}$.
$S_0^{(1)} = 100; S_0^{(2)} = 95; K = 4; T = 1; r = 0.00676; \rho_1 = -0.7039; \rho_2 = -4.7039; \lambda_1 = 2.4958; \lambda_2 = 4.4958; \gamma = 11.98; \delta = 0.0872;$

Figure 5.5 shows the impact of the mean reversion speeds of the volatility processes, i.e., the way λ_1 and λ_2 influence the resulting option prices. We notice that increasing λ_1 coincides with a slow increase in option prices, while increasing λ_2 goes along with a much faster decrease in prices.

Figure 5.6 gives indication of the model sensitivity to variation in the leverage parameters (recall that leverage parameters are assumed to be negative). We notice that option prices increase for decreasing ρ_1, while prices decrease for decreasing ρ_2.

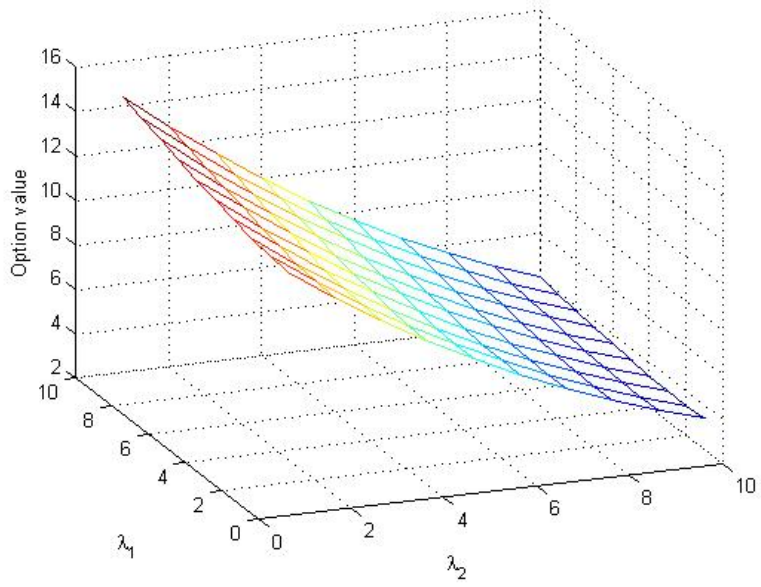


Figure 5.5: Option prices for varying λ_1 and λ_2.
$S_0^{(1)} = 100; S_0^{(2)} = 95; K = 4; T = 1; r = 0.00676; \rho_1 = -0.7039; \rho_2 = -4.7039; \gamma = 11.98; \delta = 0.0872; \Sigma_0^{(1)} = 0.064262; \Sigma_0^{(2)} = 0.054262;$

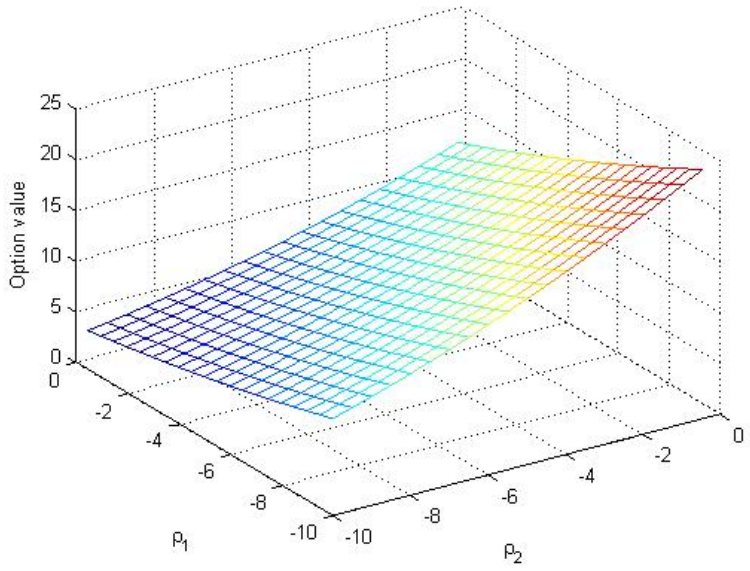


Figure 5.6: Option prices for varying ρ_1 and ρ_2.
$S_0^{(1)} = 100; S_0^{(2)} = 95; K = 4; T = 1; r = 0.00676; \lambda_1 = 2.4958; \lambda_2 = 4.4958; \gamma = 11.98; \delta = 0.0872; \Sigma_0^{(1)} = 0.064262; \Sigma_0^{(2)} = 0.054262;$

Chapter 6

Conclusion

In this thesis, we have discussed the pricing of Spread Options in multivariate Ornstein-Uhlenbeck type stochastic volatility models, which are driven by matrix subordinators. On the one hand, this included an investigation of different pricing techniques: we ran Monte-Carlo simulations and we compared the results with those obtained by an FFT method, which had been developed by Hurd and Zhou [51]. On the other hand, we examined different specifications of the market model, both from a qualitative and a quantitative point of view.

Regarding the pricing techniques, we analyzed the FFT method in detail, which revealed its advantages and weaknesses. We compared two possible ways of dealing with the issue of applying the method to any given set of spot prices of the underlying assets and strike price; we presented an algorithm to choose the lattice for approximation in an appropriate way for any particular contract on the one hand, and we implemented an interpolation procedure on the other hand. The computational speed of the FFT method is outstanding. The short computation time required to calculate option prices makes it appealing, particularly in cases where many contracts are involved. The main shortcoming of the method, which gives us some cause for concern, is the sensitivity to that parameter vector, which is used to modify the payoff function in order to ensure integrability and, in further consequence, existence of its Fourier transform. We observed that the impact of the particular choice of this parameter vector can be very strong; a sensible choice is hence an absolute necessity. However, there are no routines available to find an objectively superior choice. One has to make an ad-hoc decision, solely based on empirical convenience. We consider this as a clear drawback of this method. Therefore, we also studied alternative approaches for applicability; particularly, the IAC method appears to have potential, which represents an interesting starting point for future work.

In terms of the multivariate Ornstein-Uhlenbeck type stochastic volatility model, we first surveyed the literature about this model. We started with the one-dimensional BNS model (Barndorff-Nielsen and Shephard [6,7]) and discussed the necessary theory in order to be able to finally formulate the multivariate extension along the lines of Muhle-Karbe, Pfaffel, and Stelzer [66]. We took a particular specification of this model (which has also been presented in their work), namely, the two-dimensional OU-Wishart model and computed explicit prices for Spread Options. From a modelling point of view, we particularly addressed the issue of prespecifying the invariant distribution of the volatility process in the multi-dimensional case. Considering the inverse Gaussian distribution on a positive semidefinite cone, this is still an open issue in the very general case. However, under some simplifying assumptions, we formulated a model as a special case: we took two one-dimensional volatility processes with stationary distribution of the inverse Gaussian type, both driven by one and the same BDLP. We derived the joint characteristic function of the (two-dimensional) log-return process and finally also examined a first testing of parameter sensitivities. Moreover, as a part of our modelling attempts, we noticed that using two correlated Brownian motions would break the affine structure of the model.

In the course of the development of this thesis, various questions have arisen. As already mentioned above, the details of the correspondence between the BDLP and the stationary distribution of an Ornstein-Uhlenbeck process need to be extended to the multi-dimensional case: It would be an interesting advancement to be able to prespecify a (matrix-valued) IG-law for the volatility process and work out the full details for the BDLP, making it possible to simulate from it. On the other hand, succeeding in deriving the joint characteristic function of the log-return process in this model would enable us to apply transform-based option pricing methods.

As a further step, a calibration to market prices of the multivariate IG-OU model will be necessary. For the special case that we considered, given an appropriate data set, we would already have the tools to do so by applying the FFT method and using the characteristic function, which we successfully derived. It would be especially interesting to calibrate the model on energy derivatives, since in the modelling of energy markets, stochastic volatility models with jumps have attracted particular attention.

Finally, enhancement of the computational methods will be necessary in order to deal with the numerical issues raised above. From a mathematical point of view, further development of the rather elegant IAC method seems particularly attractive.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth Dover printing, tenth GPO printing edition, 1964.

[2] H. Albrecher and M. Predota. On Asian Option Pricing for NIG Lévy Processes. *J. Comput. Appl. Math.*, 172(1):153–168, November 2004.

[3] D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2004.

[4] O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Stat.*, 24(1):1–13, 1997.

[5] O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68, 1997.

[6] O. E. Barndorff-Nielsen and N. Shephard. Modelling by Lévy Processes for Financial Econometrics. Economics Series Working Papers 2000-W03, University of Oxford, Department of Economics, January 2000.

[7] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society Series B*, 63(2):167–241, 2001.

[8] O. E. Barndorff-Nielsen and R. Stelzer. Positive-Definite Matrix Processes of Finite Variation. *Probability and Mathematical Statistics*, 27(1):3–43, 2007.

[9] O.E. Barndorff-Nielsen. Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability and its Applications*, 45:175–194, 2001.

[10] O.E. Barndorff-Nielsen and V. Pérez-Abreu. Matrix subordinators and related upsilon transformations. *Theory of Probability and its Applications*, 52(1):1–23, 2008.

[11] O.E. Barndorff-Nielsen and R. Stelzer. Multivariate supOU processes. *Annals of Applied Probability*, 21:140–182, 2010.

[12] O.E. Barndorff-Nielsen and R. Stelzer. The multivariate supOU stochastic volatility model. *Mathematical Finance*, 23(2):275–296, 2013.

[13] D. S. Bates. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies*, 9(1):69–107, 1996.

[14] F. E. Benth. The Stochastic Volatility Model of Barndorff-Nielsen and Shephard in Commodity Markets. *Mathematical Finance*, 21(4):595–625, 2011.

[15] F. E. Benth and L. Vos. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets. *Advances in Applied Probability*, 45(2):545–571, 06 2013.

[16] F. E. Benth and L. Vos. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets. *Advances in Applied Probability*, 45(2):572–594, 06 2013.

[17] E. Bernadac. Random continued fractions and inverse Gaussian distribution on a symmetric cone. *Journal of Theoretical Probability*, 8(2):221–259, 1995.

[18] J. Bertoin. *Lévy Processes*. Cambridge Tracts in Mathematics. Cambridge University Press, 1998.

[19] P. Bjerksund and G. Stensland. Closed form spread option valuation. *Quantitative Finance*, iFirst:1–10, 2011.

[20] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637, 1973.

[21] S. Borovkova, F. Permana, and H. van der Weide. A Closed Form Approach to the Valuation and Hedging of Basket and Spread Options. *The journal of derivatives : the official publication of the International Association of Financial Engineers*, 14(4):8–24, 2007.

[22] J.-P. Bouchaud, A. Matacz, and M. Potters. The leverage effect in financial markets: retarded volatility and market panic. Science & Finance (CFM) working paper archive 0101120, Science & Finance, Capital Fund Management, 2001.

[23] S.I. Boyarchenko and S.Z. Levendorskiĭ. *Non-Gaussian Merton-Black-Scholes Theory*. Advanced Series on Statistical Science and Applied Probability. World Scientific Publishing Company, Incorporated, 2002.

[24] R. Caldana and G. Fusai. A general closed-form spread option pricing formula. *Journal of Banking & Finance*, 37(12):4893–4906, 2013.

[25] R. Carmona and V. Durrleman. Pricing and Hedging Spread Options. *SIAM Review*, 45:627–685, 2003.

[26] R. Carmona and V. Durrleman. Pricing and Hedging Spread Options in a Log-Normal Model. Working paper, Princeton Univ., 2003.

[27] P. P. Carr and D. B. Madan. Option Valuation Using the Fast Fourier Transform. *Journal of Computational Finance*, 2:61–73, 1999.

[28] R. Chhikara and J. Leroy Folks. *The Inverse Gaussian Distribution: Theory: Methodology, and Applications*. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1988.

[29] CME Group. *Introduction to Crack Spreads*. cmegroup.com, 2013.

[30] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[31] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC financial mathematics series. Chapman & Hall/CRC, Boca Raton, Fla., London, New York, 2004.

[32] J. Cooley and J. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, 19(90):297–301, 1965.

[33] J. Cox, J. Ingersoll, and St. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407, March 1985.

[34] Ch. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *The Annals of Applied Probability*, 21(2):397–463, 04 2011.

[35] Ch. Cuchiero, D. Filipović, and J. Teichmann. Affine Models. In *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Ltd, 2010.

[36] M.A.H. Dempster and S.S.G. Hong. Spread Option Valuation and the Fast Fourier Transform. In H. Geman, D. Madan, St. R. Pliska, and T. Vorst,

editors, *Mathematical Finance – Bachelier Congress 2000*, Springer Finance, pages 203–220. Springer Berlin Heidelberg, 2002.

[37] S.J. Deng. Stochastic Models of Energy Commodity Prices and Their Applications: Mean-reversion with Jumps and Spikes, 2000.

[38] S.J. Deng, M. Li, and J. Zhou. Multi-asset spread option pricing and hedging. *Quant. Finance*, 10(3):305–324, 2010.

[39] S.J. Deng and S.S. Oren. Electricity derivatives and risk management. *Energy*, 31(6):940 – 953, 2006. Electricity Market Reform and Deregulation.

[40] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

[41] J. B. C. Garcia, H. van Ginderen, and R. C. Garcia. On the Pricing of Credit Spread Options: A Two Factor HW-BK Algorithm. *International Journal of Theoretical and Applied Finance*, 06(05):491–505, 2003.

[42] F. Gordon and S. Gordon. Characterizations of the Wishart distribution using regression properties. *Annals of the Institute of Statistical Mathematics*, 28(1):201–220, 1976.

[43] M. Groth. *Topics in Computational Finance: The Barndorff-Nielsen & Shephard Stochastic Volatility Model*. PhD thesis, University of Oslo, 2007.

[44] M. T. Heideman, D. H. Johnson, and C. S. Burrus. Gauss and the history of the fast Fourier transform. *Archive for History of Exact Sciences*, 34(3):265–277, 1985.

[45] S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6:327–343, 1993.

[46] F. Hubalek and E. Nicolato. On multivariate extensions of Lévy driven Ornstein-Uhlenbeck type stochastic volatility models and multi-asset options. Preprint, 2005.

[47] F. Hubalek and C. Sgarra. On the Esscher transforms and other equivalent martingale measures for Barndorff-Nielsen and Shephard stochastic volatility models with jumps. *Stochastic Processes and their Applications*, 119(7):2137 – 2157, 2009.

[48] F. Hubalek and C. Sgarra. On the explicit evaluation of the Geometric Asian options in stochastic volatility models with jumps. *J. Computational Applied Mathematics*, 235(11):3355–3365, 2011.

[49] J. C. Hull. *Options, Futures, and Other Derivatives*. Pearson Prentice Hall, Upper Saddle River, NJ [u.a.], 6. ed., pearson internat. ed edition, 2006.

[50] J. C. Hull and A. D. White. The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*, 42(2):281–300, June 1987.

[51] T. R. Hurd and Z. Zhou. A Fourier Transform Method for Spread Option Pricing. *SIAM J. Financial Math.*, 1(1):142–157, 2010.

[52] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1987.

[53] Z. J. Jurek. Remarks on the selfdecomposability and new examples. *Demonstratio Mathematica*, XXXIV:241–250, 2001.

[54] Z. J. Jurek and W. Vervaat. An integral representation for selfdecomposable Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 62(2):247–262, 1983.

[55] Z.J. Jurek and J.D. Mason. *Operator-limit Distributions in Probability Theory*. Wiley series in probability and mathematical statistics: Probability and mathematical statistics. Wiley, 1993.

[56] M. Keller-Ressel. *Affine processes - theory and applications in finance*. Phd thesis, Vienna University of Technology, 2008.

[57] E. Kirk. Correlation in the Energy Markets. In V. Kaminski, editor, *In: Managing Energy Price Risk*, pages 71–78. London: Risk Publication, first edition, 1995.

[58] S. G. Kou. A Jump-Diffusion Model for Option Pricing. *Manage. Sci.*, 48(8):1086–1101, August 2002.

[59] J. Lillestøl. Risk analysis and the NIG distribution. *Journal of Risk*, 2000.

[60] R. Lord and Ch. Kahl. Optimal Fourier inversion in semi-analytical option pricing. Tinbergen Institute Discussion Papers 06-066/2, Tinbergen Institute, 2006.

[61] B. Mandelbrot. The Variation of Certain Speculative Prices. *The Journal of Business*, 36:394, 1963.

[62] M. B. Marcus. *Xé -radial Processes and Random Fourier Series*. Number 368 in Memoirs Series. American Mathematical Society, 1987.

[63] W. Margrabe. The Value of an Option to Exchange One Asset for Another. *Journal of Finance*, 33(1):177–86, March 1978.

[64] R. C. Merton. Theory of Rational Option Pricing. *Bell Journal of Economics*, 4(1):141–183, Spring 1973.

[65] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[66] J. Muhle-Karbe, O. Pfaffel, and R. Stelzer. Option Pricing in Multivariate Stochastic Volatility Models of OU Type. *SIAM J. Financial Math.*, 3(1):66–94, 2012.

[67] D. B. Nelson. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370, March 1991.

[68] E. Nicolato and E. Venardos. Option Pricing in Stochastic Volatility Models of the Ornstein-Uhlenbeck Type. *Mathematical Finance*, 13(4):445–466, 2003.

[69] N. D. Pearson. An Efficient Approach for Pricing Spread Options. *The journal of derivatives : the official publication of the International Association of Financial Engineers.*, 1995.

[70] Ch. Pigorsch and R. Stelzer. A Multivariate Ornstein-Uhlenbeck Type Stochastic Volatility Model. *Preprint*, 2009.

[71] Ch. Pigorsch and R. Stelzer. On the definition, stationary distribution and second order structure of positive semidefinite Ornstein-Uhlenbeck type processes. *Bernoulli*, 15(3):754–773, 08 2009.

[72] K. Prause. *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*. PhD thesis, University of Freiburg, 1999.

[73] S. Raible. *Lévy Processes in Finance: Theory, Numerics and Empirical Facts*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2000.

[74] S. Rasmus, S. Asmussen, and M. Wiktorsson. Pricing of Some Exotic Options with NIG-Lévy Input. In Marian Bubak, G. Dick van Albada, Peter M. A. Sloot, and Jack Dongarra, editors, *International Conference on Computational Science*, volume 3039 of *Lecture Notes in Computer Science*, pages 795–802. Springer, 2004.

[75] J. Rosiński. On A Class of Infinitely Divisible Processes Represented as Mixtures of Gaussian Processes. In St. Cambanis, G. Samorodnitsky, and M. S. Taqqu, editors, *Stable Processes and Related Topics*, volume 25 of *Progress in Probabilty*, pages 27–41. Birkhäuser Boston, 1991.

[76] J. Rosiński. Series Representations of Lévy Processes from the Perspective of Point Processes. In O. E. Barndorff-Nielsen, S. Resnick, and T. Mikosch, editors, *Lévy Processes*, pages 401–415. Birkhäuser Boston, 2001.

[77] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[78] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley Series in Probability and Statistics. Wiley, 2003.

[79] M. Schweizer. A minimality property of the minimal martingale measure. *Statistics & Probability Letters*, 42(1):27 – 31, 1999.

[80] V. Seshadri. The inverse Gaussian distribution: Some properties and characterizations. *Canadian Journal of Statistics*, 11(2):131–136, 1983.

[81] J. Stein and E. Stein. Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies*, 4:727–752, 1991.

[82] The Royal Swedish Academy of Sciences. Press release, 1997.

[83] L. Valdivieso, W. Schoutens, and F. Tuerlinckx. Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type. *Statistical Inference for Stochastic Processes*, 12(1):1–19, 2009.

[84] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

[85] A. Venkatramanan and C. Alexander. Closed Form Approximations for Spread Options. *Applied Mathematical Finance*, 18(5):447–472, 2011.

[86] P. Wilmott. *Paul Wilmott on Quantitative Finance 3 Volume Set*. Wiley, second edition, March 2006.