Essays on Numerical Evaluation of Derivatives

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Florian Niederstätter
aus Stuttgart

Tübingen
2016

Tag der mündlichen Prfung:	16.03.2017
Dekan:	Professor Dr. rer. soc. Josef Schmid
1. Gutachter:	Professor Dr. rer. pol. Werner Neus
2. Gutachter:	Professor Dr.-Ing. Rainer Schöbel

Contents

List of Figures

List of Tables

List of Symbols

General Symbols

$\mathbb{C}$	Complex numbers
$\mathbb{N}$	Natural numbers
$\mathbb{R}$	Real numbers
$\mathbb{Z}$	Integers

$\mathcal{F}_t$	Filtration
$\mathcal{H}$	Hilbert space
$\mathcal{I}$	Imaginary part of a complex number
$\mathcal{O}(\cdot)$	Asymptotic growth or decay rate
$\mathcal{P}$	Power set; set of all subsets
$\mathcal{Q}$	Risk neutral measure
$\mathcal{R}$	Real part of a complex number

$C^p(\mathbb{R})$	Space of p times differentiable functions with a continuous p-th derivative
$L^p(\mathbb{R})$	Space of Lebesgue measurable functions $\int_{\mathbb{R}} \mid f(x) \mid^p dx < \infty$
$L^2(\mathbb{R})$	Hilbert space among the spaces $L^p(\mathbb{R})$
$S^p(\mathbb{R})$	Schwartz space of rapidly decreasing functions on $\mathbb{R}$

K	Strike value
r	Risk free rate
S	Underlyings value
T, t	Time indices
q	Cost of carry

$\langle f, g \rangle$	Inner product: $\int f(x)\,\overline{g(x)}\,dx$
$\|f\|$	Norm of a function f: $\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int \mid f(x) \mid^2 \; dx}$

$\odot$	Element wise multiplication
$\otimes$	Kronecker products
$\mathbb{1}_A$	Indicator function for a set A, where $\mathbb{1}_A(x) = 1$ if $x \in A$ and zero otherwise
$\overline{z}$	Complex conjugate of $z = a + b \cdot i$ given by $\overline{z} = a - b \cdot i$

Chapter Specific Symbols

Chapter 3

$\mathcal{G}(g, \alpha, \beta)$	Gabor system with generator g as well as parameters α and β
$\mathcal{G}(\gamma, \alpha, \beta)$	Gabor system with dual function γ as well as parameters α and β
$\mathcal{I}_p$	Subinterval of a put option on the truncation interval $[a, b]$
ϵ	Machine precision given floating point arithmetic $\epsilon \approx 2.22 \cdot 10^{-16}$
γ	Dual function
Θ	Parameter set of a specific stochastic process
a, b	Truncation limits
E_n	Error terms
g	Generater function
M, m	Modulation parameter
N, n	Shift parameter
$\hat{N}$	Maximal number of shifts
p	Control variable affecting the decay rate of the dual function
$\lceil x \rceil$	Ceiling function: smallest integer being not less than x
M_ω	Modulation operator $M_\omega g = g(x)\, e^{2\pi i \omega x}$
T_z	Translation operator $T_z g = g(x - z)$

Chapter 4

$\mathcal{T}$	Set of evaluation points of a discretely monitored path dependent option
Ω	Integration domain of a discretely monitored path dependent option
Σ	Covariance matrix
Θ	Contract specific subdomain of Ω
τ	Trigger event
c	Coupon of a swap
$C_{\mathbf{n}}(\cdot)$	Set of cosine functions $C_{\mathbf{n}}(\cdot) = \Pi_{j=1}^{d} \cos(n_j ...)$
H	Hankel matrix
h_j	Standardized log-Barrier
$M_{\mathbf{n}}(\cdot)$	Set of modified sine functions $M_{\mathbf{n}}(\cdot) = \Pi_{j=1}^{d} \sin\left((n_j - \frac{1}{2})...\right)$
$S_{\mathbf{n}}(\cdot)$	Set of sine functions $S_{\mathbf{n}}(\mathbf{x}) = \Pi_{j=1}^{d} \sin(n_j ...)$
T	Toeplitz matrix
$\nu(\cdot)$	Option's value at maturity
$\nu_{\mathbf{n}}$	Fourier payoff coefficients (Shortened version: $\hat{\nu}_{\mathbf{n}}$)

Chapter 5

$\mathcal{P}$	Underlying portfolio of credit default swaps
ϵ	Idiosyncratic factor within a one-factor model
ξ	Continuous portfolio loss distribution
a	Correlation variable within a one-factor model
$c(K_1, K_2)$	Spread payment of the tranche $[K_1, K_2]$
K_1, K_2	Attachment point and detachment point
$L_t^{K_1,K_2}$	Loss of the tranche $[K_1, K_2]$
N	Notional
Up	Upfront payment
Y	System-wide factor within a one-factor model

List of Abbreviations

CGMY	Stochastic process named after the authors Carr, Geman, Madan and Yor
CDO	Collateralized debt obligations
CDS	Credit default swap
DL	Default leg
DTCC	Depository Trust and Clearing Corporation
EDS	Equity default swap
Eq	Equity tranche (losses 0-3 %)
FFT	Fast Fourier transform
FrFFT	Fractional Fast Fourier Transform
GBM	Geometric Brownian motion
GH	Generalized hyperbolic
GTS	Generalized tempered stable
LHP	Large homogeneous portfolio
MAEDS	Multi-asset equity default swap
Me1	First mezzanine tranche (losses 3-6 %)
Me2	Second mezzanine tranche (losses 6-9 %)
Me3	Third mezzanine tranche (losses 9-12 %)
NIG	Normal Inverse Gaussian
PIDE	Partial integro-differential equation
PL	Premium leg
Ref	Reference value
RFL	Random factor loading
Se1	First senior tranche (losses 12-22 %)
Se2	Second senior tranche (losses 22-100 %)
SP	Stochastic process
TD	Tranche deviation
VG	Variance Gamma

Chapter 1

Introduction

1.1 Motivation

Contingent claims are an important ingredient in a huge range of financial contracts. Especially with markets moving ever faster due to electronic trading platforms, automated trading systems and extended trading hours, being able to map the products into a robust and yet intuitive model becomes increasingly difficult. Furthermore, the computational speed of the numerical implementation of a given model is crucial if it is meant to be used in a real time environment. Thus, inherited in the decision which model to choose and how to implement it is a trade-off between simplicity and speed on the one hand as well as the capability of representing market movements accurately on the other hand. This thesis is therefore concerned with the numerical implementation of derivative pricing models.

Market prices move in an apparently coincidental upward and downward movement. Within the seemingly random behavior, stylized facts on the distribution of asset prices and returns can be found within time series. Fat tails and volatility clustering are just two examples that are typically present when dealing, e.g., with share prices. The very basis of modeling market movements as random behavior was introduced by Bachelier (1900). In his work he describes stock price movements by means of a Brownian motion. Even though his work was not appreciated for a long time, it marks the entering of advanced probability theory into finance. Starting from this pioneer of financial mathematics, the theory of stochastic processes made its way into the models.

A Brownian motion assumes movements as being normally distributed which yields to a framework where negative stock prices are possible. To overcome this problem Samuelson (1965) formulates an exponential version, known as geometric Brownian motion (GBM). In contrast to a Brownian motion, log returns under a GBM are normal distributed and prices are now log-normally distributed which prevents negative stock prices. However, within the environment of financial contracts, the assumption of a Normal distribution as the underlying source of randomness is rarely supported by empirical studies. Therefore, more complex stochastic processes with richer behavior were introduced.

With that in mind, we consider the numerical implementation of models that describe markets by means of exponential semimartingale processes. This rich class of stochastic processes includes (jump) diffusion processes as well as pure jump Lévy processes and several stochastic volatility models. For our purposes, we mainly consider pure jump Lévy processes (Lévy processes for short) as well as stochastic volatility models to some extend. The class of Lévy processes contains a wide range of processes with different characteristics. The most basic representatives are Brownian motion and the Poisson process. In fact, it can be shown that all Lévy processes are assembled by a non stochastic drift, a Brownian part and Poisson jumps. Thus, besides other properties, Lévy processes are equipped with the feature of being able to produce discontinuous paths. In financial applications this translates into the capability of prices to jump between two values S_1 and S_2 instead of moving continuously from S_1 to S_2. While, compared to a GBM, more flexible stochastic processes are able to cover a wide range of stylized facts, their incorporation within the pricing models also increases the overall complexity of the models. As a result, closed-form solutions are rare.

This directly leads us to numerical methods in option pricing which can be classified into the following three concepts

i) stochastic differential equation methods,

ii) Monte-Carlo simulations and

iii) numerical integration methods.

Our research is located in the area of numerical integration. However, the methods within this enumeration overlap on multiple occasions and a broad overview is helpful to gain a better understanding of the projects that are described within this thesis. Thus, a short orientation within each of the fields is provided in the following starting with partial differential equation methods.

Fueled by Itô's lemma[1], stochastic differential equations are a common method for describing the value of an option. Hereby, the assumption of the stochastic process driving the underlying asset affects the type of stochastic differential equation that is used to represent the option. If a diffusion process, such as a geometric Brownian motion, is implemented a parabolic partial differential equation occurs. In case of more general jump-diffusion and exponential Lévy market models, partial integro-differential equations (PIDE) are used to represent the option's behavior. Hereby, compared to the former, the latter type of partial differential equation has an additional integral term due to the presence of jumps.

Up to this point, solutions are stated in closed-form. However, only a few stochastic processes allow a closed-form solution of the option price. Thus, from the point on where the partial (integro-) differential equation is defined, numerical methods are used to approximate it. Hereby, tree and Markov chain methods are the most basic methods of choice. In case

[1] Stochastic processes are discontinuous functions which, in turn, implicates that classical calculus cannot be applied to define e.g. a differential. Itô's lemma provides a way to calculate the differential of a time and state dependent function such as an option that depends on time to maturity as well as the state of an underlying asset.

of a geometric Brownian motion, the binomial tree method by Cox et al. (1979) is an approximation of the continuous time process by means of a discrete time Markov chain (Cont and Tankov, 2004, p. 408). Amin (1993) studies multinomial tree methods and thereby generalizes the assumption on the underlying stochastic process by including jump-diffusion models.

The link between tree methods and the approximation of a partial (integro-) differential equation is somehow indirect. Tree methods work on a lattice in (t, S)-space and can be interpreted as an explicit finite difference scheme for the associated partial differential equation (Cont and Tankov, 2004, p. 410). In general, finite difference methods approximate a partial differential equation by replacing derivatives by finite differences. In this context, Cont and Voltchkova (2005) propose a scheme in which European and barrier options under jump-diffusion and exponential Lévy models are priced.

As a final numerical method connected to PIDE appoaches, we would like to mention Galerkin methods. Galerkin methods represent the solution to a PIDE in terms of a series representation based on basis functions. Hereby, the choice of the basis function yields to different numerical schemes. In case of a hat-function on a regular grid, Galerkin methods are equivalent to finite difference methods. However, irregular grids can also be introduces dedicating a higher attention to areas where a higher accuracy is necessary (Cont and Tankov, 2004, p. 425). Besides hat-functions, other basis such as complex exponentials and wavelet basis are common choices. In recent years, especially wavelet basis were studied to a greater extend since they show numerical advantages in terms of the matrix design within the linear system that needs to be solved using Galerkin methods. Hereby, wavelets decrease the density of the matrices in a sense that the number of nonzero elements is decreased which yields to faster numerical calculations. Both, Matache et al. (2004) and Eberlein and Glau (2014) use wavelet methods to price options in an exponential Lévy market model environment.

PIDE methods are a suitable method of choice in case of single underlying contracts. However, it becomes less advantageous if the number of underlyings increases. While the computational complexity of PIDE methods growths exponentially, the complexity of Monte Carlo methods grows linearly given a pre-defined level of accuracy. However, contrary to this advantage that is inherit to simulation methods, the simulation of random variables and therefore the simulation of paths can be quite complex if working with Lévy processes. This is due to the fact that the distribution function of the increments is not known explicitly for most Lévy processes other than (geometric) Brownian motion and Poisson processes. In case of other Lévy processes, simulation can be achieved by exploiting the fact that a general Lévy processes can be written as a subordinated Brownian motion. A subordinator is defined as a non-decreasing Lévy processes that is used to time-change a Brownian motion. Subordination is often mentioned in conjunction with the expression stochastic time-change and business time versus calendar time. Since efficient methods are available for simulating some of the subordinators, the resulting Lévy processes are conveniently being implemented in a pricing routine. Carr and Wu (2004) offers a broad overview to the concept of time-changing

stochastic processes. Approaches to numerically simulate these processes are discussed e.g. in chapter 8 of Schoutens (2003) as well as Madan and Yor (2008).

Both, PIDE methods as well as simulation based techniques, have in common that changing the stochastic process has a rather huge impact on an existing implementation in terms of the adjustments that have to be re-coded. Numerical integration methods, or quadrature methods as they are often called, complete the above trio of numerical option pricing methods and, in part, minimize the need for adaptions within an existing pricing framework.

In its most basic form, quadrature methods discretize the integral that is present in the risk neutral pricing formula. Common discretization rules are (composite) Newton-Cotes, (composite) Gaussian quadrature and adaptive quadrature methods. Newton-Cotes formulas approximate a given integral by a sum composed of the integrand evaluated at finite many points in combination with a weighting function. Depending on the weighting function different rules, such as trapezoid rule and Simpson rule, occur. In this context, the term composite refers to the proceed of not approximating the whole integral at once but subdividing it into smaller areas where the respective rule is applied to each subinterval. Instead of fixed weight functions, Gaussian quadrature methods work with weight functions based on the roots of polynomials that are computed each time to provide an improved order of accuracy. Similar to Newton-Cotes formulas, the choice of a specific polynomial gives rise to different quadrature rules such as Gauss-Legendre, Gauss-Chebyshev, Gauss-Laguerre and Gauss-Hermite. Besides different polynomial functions, these methods differ in the assumption of the integration limits. While the first two define the integration domain to be in between negative one and one, Gauss-Laguerre allows for a domain on the positive reals and Gauss-Hermite includes the whole real line. Thus, depending on which Gaussian quadrature comes into operation, a change of variables may be necessary. Another prominent numerical integration method is given by adaptive quadrature. This method also splits the integration domain. However, instead of using a number of evenly spaced subintervals, the integration range is broken down into ever finer pieces until a given level of accuracy is reached.

A major drawback of all of these quadrature methods when used to directly integrate the risk neutral expectation is that the distribution function of the underlying asset must be known explicitly. In case of a Black-Scholes environment where the asset's behavior is described by a log-Normal distribution, this is not a critical claim. However, as soon as other stochastic processes than a geometric Brownian motion are applied, the density functions are either not known or are based on special functions such as modified Bessel functions which are slow to evaluate. As a different approach Bakshi and Madan (2000) show in a very general setting that the price of an option is a function that depends on, among other things, two ingredients: the probability of finishing in the money and the option's delta. Hereby, both terms are defined by an integral with an integrand that includes characteristic functions of the assumed stochastic process. The advantage of an approach that is based on characteristic functions is that they are known explicitly in most cases and, furthermore, interchanging the assumption on the underlying stochastic process is easily being achieved by interchanging between characteristic functions. In a last step, the approximation of the two integrals can be

done by means of the before-mentioned quadrature methods. However, due to the fact that the integrands are based on complex exponentials, they show an oscillatory nature which can be troublesome for some of the quadrature methods.

Instead of computing the probability of ending up in the money and the option's delta, Carr and Madan (1999) propose a method where the Fourier transform of an adjusted[2] call option has to be inverted to come up with option prices. The inversion step is hereby done by numerically evaluate a single integral. Using a Newton-Cotes formula, to be more precise the trapezoid rule, this step can be done very efficiently by means of the Fast Fourier Transform (FFT) algorithm described by Cooley and Tukey (1965). Carr and Madan's FFT method achieves a high level of computational speed, however, as a disadvantage, prices are related to a grid of strike prices which extends far out of the money. As a result, the method prices a high number of contracts that differ in the respective strike price. However, most of the strike values are not meaningful for most applications and option prices belonging to strike values that are not on the grid have to be interpolated. To overcome this downside, Chourdakis (2005) proposes a fractional FFT method that allows the user to define strike prices as an input variable. Particularly well known models that rely on inversion techniques and can be evaluated efficiently via FFT or fractional FFT methods are Stein and Stein (1991), the extension by Schöbel and Zhu (1999) as well as Heston (1993) and Duffie et al. (2000).

Another way to calculate the risk neutral expected value is to not discretize the integral but to rewrite it in terms of an inner product of series coefficients as in Fang and Oosterlee (2008). Hereby, the integral as a whole can be represented in terms of a sum of two series coefficients. Our research is rooted within this spirit in a way that we also rewrite the risk neutral expectation in terms of an inner product and implement Fourier series and Gabor series methods to numerically solve the pricing problem associated with several derivative type contracts such as plain vanilla European options, multi-asset barrier options as well as swaps and credit derivatives. Fourier and Gabor series are especially suited for this task due to the fact that they also allow the usage of characteristic functions within the computations. As a result, the overall algorithm does not change if we interchange between different stochastic processes described by a particular characteristic function. However, before we elaborate in more detail on how we use these methods, we give an intuition for the nature of pure frequency analysis, to which Fourier series belong, and time-frequency analysis, to which Gabor series belong, by means of two examples: function approximation and musical sound.

At its very core, Fourier series decompose an arbitrarily complex periodic[3] function into basic building blocks of sine and cosine terms with different amplitudes and frequencies. Figure 1.1 depicts the nature of this kind of decomposition in greater detail. Within the figure, the function, which is sometimes referred to as signal, colored in red, has to be approximated.

[2] A damping factor has to be introduced to ensure L^2-convergency of the call price.

[3] Fourier series operate on a closed interval $\mathcal{I} \subset \mathbb{R}$ and assume a periodic extension on $\mathbb{R} \setminus \mathcal{I}$ meaning that the trajectory within the interval is repeated at the outside.

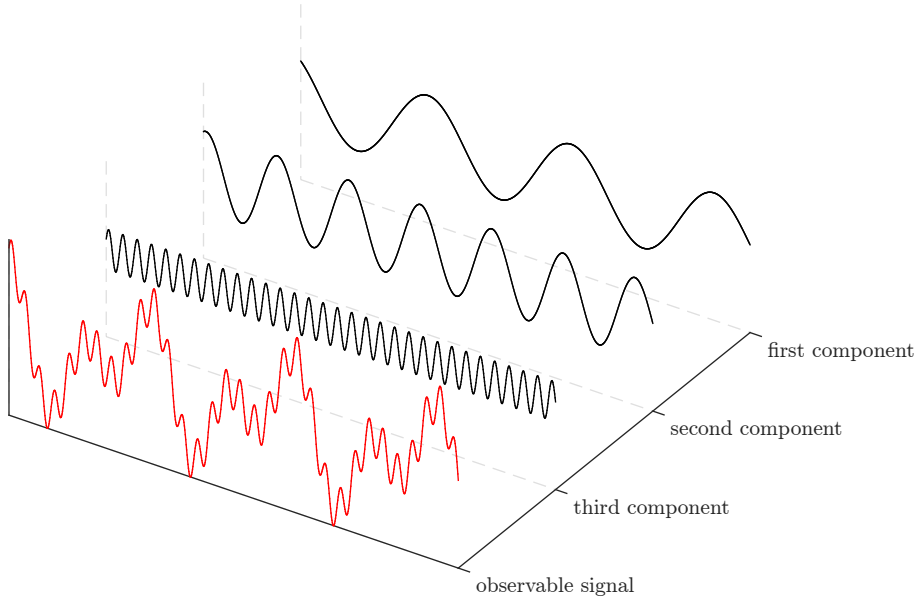

Figure 1.1: Function decomposition of a given observable function into sine waves of different amplitude and frequencies.

Fourier series analysis enables us to realize that the function under consideration is built up by three different components, each of which is a trigonometric function with individual frequency. To come up with a representation of the observed signal in terms of trigonometric functions, we do not even have to know the exact functional relation. Even in case we do know the exact function that creates the signal within the figure, a whole range of mathematical operations are done more conveniently on the simple building blocks of the decomposition than on the function itself.

The second example results quite naturally from function approximation. A musical tone in its pure form is represented by a vibration. Each of the components in Figure 1.1 could therefore also be interpreted as pure tones and the signal in red as a combination of these pure tones played, e.g., on an instrument. The knowledge on which tones are to be played in order to replicate a given sound or signal is handy if we consider a constant input sound or signal. If we concentrate on melodies as a timely order of many tones, Fourier analysis would still come up with the answer which basic frequencies are in use. But it cannot answer the question at which point in time which tone has to be played. Fourier analysis, therefore, only offers frequency information but no information on localization. Such a situation describes a limiting case of Heisenberg's uncertainty principle[4] in a sense that we have all information regarding frequencies but no control on localization.

To illustrate this statement we use Mozart's symphony No. 40 in Figure 1.2. To be more precise, we use a sample containing the first six seconds of the symphony and transform it from time domain (upper part of Figure 1.2) to frequency domain (lower part of Figure 1.2).

[4] Originally Heisenberg's uncertainty principle is rooted in quantum mechanics where it states that the position and the momentum of a particle can only be determined simultaneously with a limited precision (Heisenberg, 1927).

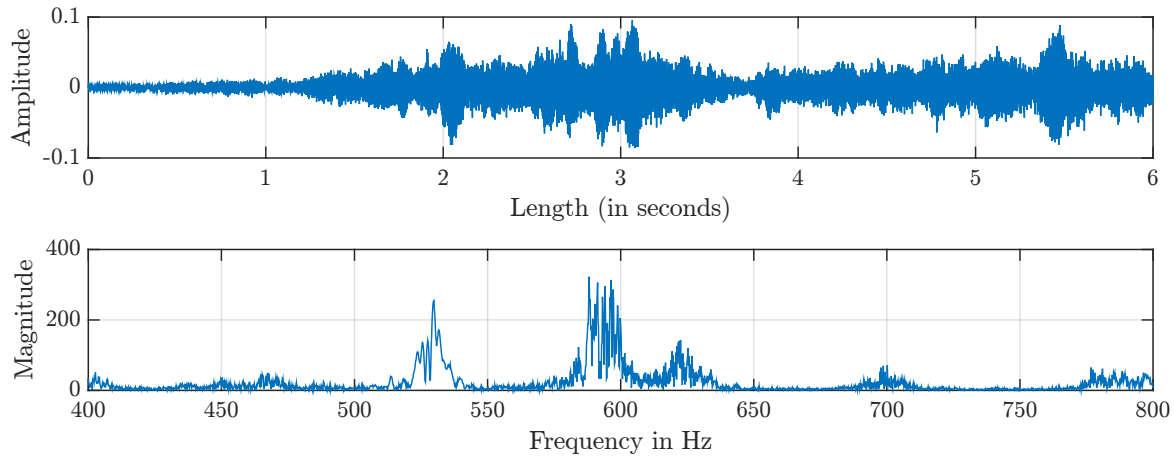

Figure 1.2: Sample file containing the first six seconds of Mozart's symphony No. 40.

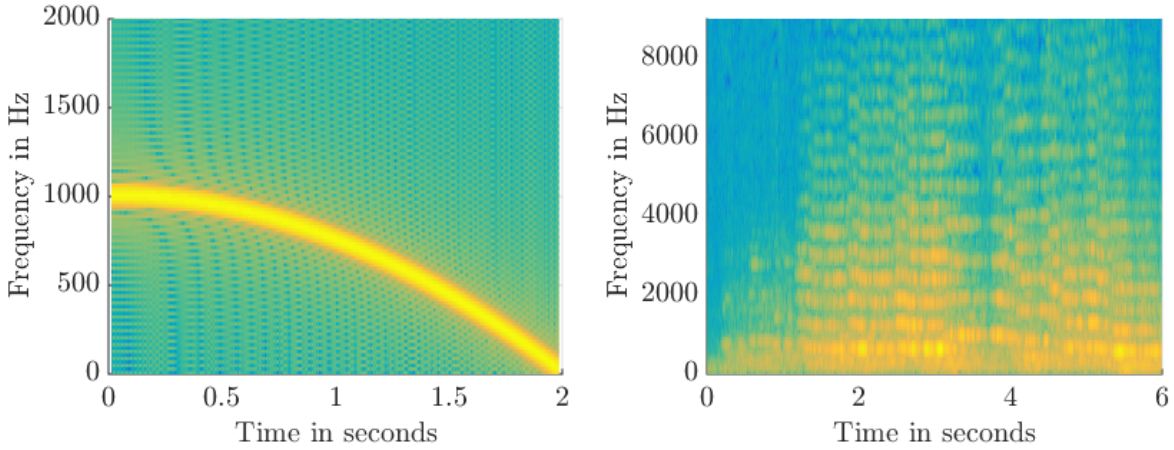

Figure 1.3: Time-frequency analysis of a chirp signal (left-hand side) and Mozart's symphony No. 40 (right-hand side).

While in the upper part the musical movement is observable with respect to time evolving, the lower part shows a summary of which frequencies are played the most during the first six seconds within an interval of 400 Hz to 800 Hz. The plot in the lower part of the figure reveals that certain frequencies are used heavily while others are not. It is even possible that frequencies are used that are out of the spectrum a human ear is able to process. Deleting these frequencies from the sample results in a compressed version of the original signal. A particularly well known standard that is based on this principle is the MP3 format.

In contrast to Fourier analysis, a basic time-frequency analysis separates a signal in smaller sub-signals and takes a look at the frequencies incorporated on the interval of the sub-signal afterwards. As a result, not an overview of the magnitude of all frequencies becomes visible but a time evolving picture of the frequencies used at particular time intervals. An intuitive example is to take a closer look at so-called chirp signals: A chirp refers to a signal which frequencies increase or decrease with time in a monotone way. Such a chirp signal is analyzed on the left-hand side of the spectrogram in Figure 1.3. Within the graph, brighter areas indicate frequencies that are heavily used. By design, the chirp signal of this example

starts at a frequency of 1000 Hz and decreases down to zero Hz with time evolving. However, signals are rarely as clear in structure as chirp signals are. Therefore, on the right-hand side of Figure 1.3, we turn our attention to Mozart's symphony No. 40 again. Compared to a pure Fourier analysis, the figure allows for an allocation of the frequencies to a time scale. Before, a Fourier analysis did show us that frequencies somehow below 600 Hz are played with a high magnitude. The same information is visible in Figure 1.3. But in addition to Figure 1.2, Figure 1.3 indicates that these frequencies are played the most from second two to roughly second three.

This particular analysis is based on an elementary routine called short-time Fourier transform. Gabor analysis is somehow different to a pure short-term Fourier transform in the way it interprets the connection between time and frequency. In a nutshell, the energy of a continuous signal[5] can be thought of as being spread out over the time-frequency plane (Dörfler, 2002, p. 7). Gabor analysis not only cuts the signal into sub-signals but uses a function, called generator function or window function, which has it's own particular time-frequency localization to concentrate the analysis on certain pieces of the original signal within the time-frequency plane. Hereby, the localization of the generator functions plays a crucial role. The theory which tells us how to rebuild the signal based on the individual pieces is known as frame theory. More specific, using Gabor analysis, the theory is known as Gabor frame decomposition or Weyl-Heisenberg frame decomposition. Within our area of interest, this particular frame decomposition is realized by means of a series representation called Gabor series.

Even though time-frequency analysis is able to provide additional information it is still subject to the before-mentioned uncertainty principle which states, in this context, that a function cannot be localized exactly in time and frequency domain simultaneously. However, time-frequency analysis allows us to move in between the limiting cases of either knowing the exact frequency and nothing on timing or of knowing the exact timing but nothing on the frequency spectrum. Besides numerous applications in technical science, these transform methods also became important in financial applications and especially in option pricing.

1.2 Structure

This thesis considers valuation problems related to different types of financial contracts. Within the course of three main chapters, as depicted in Figure 1.4, we move from contracts with a single underlying asset to contracts with multiple underlying assets. To provide the reader with the necessary tools, an additional chapter is included that states the mathematical preliminaries needed to follow the implementations within the main chapters. In case the reader is familiar with the mathematical concepts, Chapter 2 can easily be skipped.

In Chapter 3, we consider European-style options based on a single underlying. European options were the first contracts to be modeled by means of characteristic functions. In this

[5] Given a continuous signal s, the energy is defined as the squared norm of this signal $\int_{-\infty}^{\infty} \mid s(t) \mid^2 dt$.

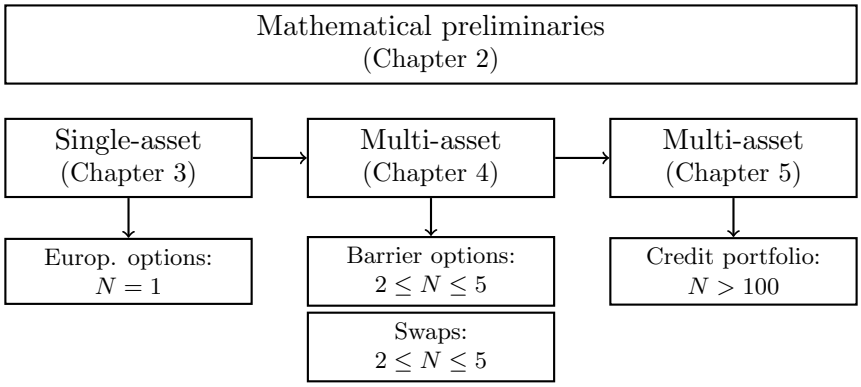

Figure 1.4: Schematic illustration of the structure of the thesis.

spirit, we introduce an option pricing algorithm based on non-orthogonal series expansion methods. More precisely, Gabor frame decomposition is used to split the risk neutral option pricing formula into the sum of two inner products that can be evaluated efficiently by means of Parseval's theorem on complex Fourier series. The first inner product is hereby based on the stochastic process that is assumed to drive the underlying asset and the second one depends on the option contract to be priced. To the best of our knowledge, Gabor series have not been considered yet in literature to calculate option prices. We consider European style plain vanilla call and put options as well as binary options. Compared to Fourier series methods such as Fang and Oosterlee (2008), we find an improved accuracy in terms of options with a short time to maturity as well as a lower sensitivity of the model regarding critical input values.

In Chapter 4, we concentrate on pricing rainbow options. These contracts are defined as derivatives which are exposed to at least two sources of uncertainty. Since closed-form solutions are rare and mostly limited to a Black-Scholes environment, the primary method to evaluate such contracts is to implement Monte Carlo routines. Ruijter and Oosterlee (2012) consider two-dimensional Fourier cosine series to price European and Bermudan options. As an extension, we focus on pricing multivariate discrete barrier options using various Fourier series methods: Besides cosine series, we also consider sine series and modified sine series approximation in a d-dimensional setting to calculate option prices. Especially, the incorporation of different Fourier series methods proofs to be insightful due to the fact that we pinpoint modified sine series to be a better fit for the pricing problem than cosine series and sine series are. The field of application is diverse and ranges from plain vanilla barrier options to multi-asset equity default swaps and structured products such as multi-barrier reverse convertibles. When modeling multi-asset options, the so-called curse of dimensionality, i.e. the issue of computational times increasing fast when incorporating an additional asset, is important. This is also true for the Fourier method considered here. However, the methods turn out to be extremely exact if faced with a basket size of low dimensionality.

In Chapter 5, we put credit derivatives in form of synthetic collateralized debt obligations in focus, and, therefore, analyzes contracts built upon a large portfolio. Factor models based on Gaussian distributions have been widely used to price such synthetic collateralized debt

obligation contracts before the market meltdown starting in 2007. With this chapter, we add value to the credit risk discussion in a twofold way. First, the body of literature researching the impact of replacing the Gaussian by more flexible distribution functions is developed further. Hereby, a special focus is placed on generalized tempered stable and generalized hyperbolic distributions. Moreover, we broaden the perspective by deviating from the usual approach of using identical distributions within the factor model setup. Instead, subclasses of the above-mentioned distributions are mixed in order to combine different characteristics. Second, an extensive calibration study based on standardized iTraxx Europe tranches is conducted that sheds light on the question regarding where the limits of the factor models presented in this chapter are. We hereby find that particularly mixed models assuming extended Variance Gamma distributions are able to reproduce market prices.

Especially in Chapter 3 and Chapter 4 of the thesis, we emphasize on the numerical implementation of the models in their respective field of application. Theoretical parameter sets are used to come up with test scenarios the models can be evaluated on. Thus, in Chapters 3 and 4, the speed of convergency to a pre-defined error tolerance level is central. Within Chapter 5, the term implementation becomes a somehow different meaning. In contrast to before, market prices rather than theoretical values are used to conduct an extensive calibration study.

Chapter 2

Mathematical Preliminaries

The aim of this complementary chapter is to discuss several concepts of the field of probability theory and linear algebra. Each of the following concepts are also introduced in the respective chapters, however, to a lesser extend. The resulting redundancy is therefore made by intention and aims to allow us to focus on the implementation of the concepts within the main chapters which begin with Chapter 3.

We consider a continuous time framework to evaluate derivatives. Hereby, such concepts as probability spaces, filtrations and stochastic processes are crucial to fully understand the resulting pricing formulas and are introduced in the following Section 2.1. Whereas probability theory is needed to specify this pricing relation, numerical methods are needed to evaluate the relation given by an expected value. Thus, in addition to probability theory, Section 2.2 is dedicated to topics which are typically covered in linear algebra such as vector spaces and infinite series approximation.

2.1 Financial modeling in continuous time

In order to come up with a proper environment to price various kinds of financial contracts in, the concepts of probability spaces and filtrations have to be introduced before moving on to stochastic processes and risk neutral pricing methods with the help of an equivalent martingale measure. Figure 2.1 gives an outlook on which topics are covered within this section. For a deeper insight into the theory, Øksendal (2003), Cont and Tankov (2004) and Shreve (2004) are excellent sources from which parts of this section are built upon.

2.1.1 Probability space

A probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is built upon three components: the outcome set Ω, the σ-algebra $\mathcal{A}$, also called event set, and the probability measure $\mathbb{P}$. Hereby, the outcome set Ω includes all possible outcomes. Elements within this set are given by $\omega \in \Omega$. A σ-algebra is

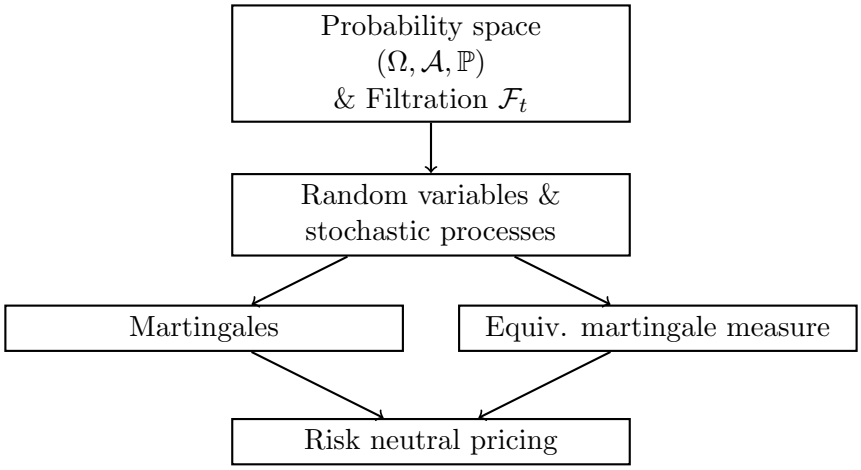

Figure 2.1: Preliminaries in financial modeling in continuous time.

defined as a system of subsets that obey the following conditions:

$$\Omega \in \mathcal{A}$$
$$A \in \mathcal{A} \rightarrow A^C \in \mathcal{A}$$
$$A_1, A_2, ... \in \mathcal{A} \rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$$

While the first condition states that the outcome set itself is part of the σ-algebra, the second condition demands the complement A^C of a given element A to also be part of the σ-algebra. Accordingly, the third statement induces that if an arbitrary number of subsets is included in the σ-algebra, the union of these subsets has to be also part of the σ-algebra.

We use the toss of a coin to elaborate on the concept of probability spaces. To do so, we define head by H and tail by T. Now, the event set is given by $\Omega = \{H, T\}$ and the set of possible subsets[6] is given by $\mathcal{A} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$. To indicate that $\mathcal{A}$ is indeed a σ-algebra, we note that, first, it is true that the event set is included in the set of subsets. Second, the complements are also included if we interpret the complement of head to be tail and the null set – interpreted as not tossing at all – to be the complement of throwing either head or tail. Finally, since it is true that $\emptyset \cup \{H\} \cup \{T\} \cup \{H, T\} = \{H, T\}$, we conclude that the set $\mathcal{A}$ is indeed a σ-algebra.

Up to this point, we only introduced the pair $(\Omega, \mathcal{A})$ which is also known as a measurable space. To come up with a probability space, a probability measure $\mathbb{P}$ has to be added to the pair. The probability measure $\mathbb{P}$ hereby assigns probabilities to the subsets within the σ-algebra. Thus, a probability measure $\mathbb{P}$ can be seen as a function that maps a measurable

[6] Assuming the coin never lands on its edge.

space $(\Omega, \mathcal{A})$ into a closed interval $\mathbb{P}\colon \mathcal{A} \to [0,1]$ and has the following properties:

$$\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1$$
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Hereby, the first and the second condition only state that the empty set has zero probability and the probability that the event has to be within the predefined event set is one. The third statement shows that, given the sets within $\mathcal{A}$ do not overlap, the probability of the union of all subsets is given by the sum of individual probabilities.

Filtration $\mathcal{F}_t$

Using the concept of filtration, the information collected within the σ-algebra can be considered as being time dependent and, for this reason, is labeled with index t, where $t \in [0,T]$. Thus, a filtration can be seen as a sequence of σ-algebras $\{\mathcal{F}_t\}_{t\in[0,T]}$ with the distinct characteristic that each σ-algebra within the sequence contains all the sets of the previous σ-algebra:

$$\mathcal{F}_0 \subseteq \mathcal{F}_1, ..., \mathcal{F}_T \subseteq \mathcal{A}$$

In a way, each σ-algebra within the filtration contains the information that are available up to this point and, therefore, unveils some information by itself.

To be more explicit, we come back to the coin tossing example: When tossing a coin two times, the event set is given by $\Omega = \{H,T\} \times \{H,T\} = \{HH, HT, TH, TT\}$. By incorporating the concept of time, we state three different points in time. At $t = 0$ the coin has not been tossed yet. At $t = 1$ the coin is tossed once, and, at $t = 2$, the coin is tossed twice. Thus, the sequence of σ-algebras is given by

$\mathcal{F}_0 = \{\emptyset, \Omega\}$	contains no information
$\mathcal{F}_1 = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$	contains some information
$\mathcal{F}_2 = \mathcal{P}(\Omega) = \mathcal{A}$	contains all information

Before the coin is tossed for the first time, no information other than the empty set and the event set itself is given. However, the σ-algebra grows at time $t = 1$ and includes two additional sets, $\{HH, HT\}$ and $\{TH, TT\}$. Knowing in which set the outcome of the first toss is located in effectively unveils the nature of the first toss, i.e. if the first toss resulted in head or tail. Therefore, $\mathcal{F}_1$ is said to contain the information of the first toss. σ-algebra $\mathcal{F}_2$ contains all information as described by the power set $\mathcal{P}(\Omega)$ and is identical to $\mathcal{A}$.

The coin tossing example shows that additional information is unveiled with time evolving. In a financial interpretation, this could be seen as stock prices that become observable at some time $t > 0$ which have been viewed as random in $t = 0$. Thus a filtration is nothing but a time-sensitive addition to a given probability space. Such a probability space is then called filtered probability space and is sometimes noted by $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbb{P})$.

2.1.2 Random variables and stochastic processes

Random variables

The importance of a probability space becomes obvious when we try to model the outcome of an experiment or of a stock market movement as a random event. Hereby, we are especially interested in the probability measure which assigns probabilities to the different events. Each outcome is defined as a random variable which lives on the probability space connected to this random variable. Thus, if $(\Omega, \mathcal{A}, \mathbb{P})$ represents a complete[7] probability space, a random variable X is defined as an $\mathcal{F}_t$-measurable function $X\colon \Omega \to \mathbb{R}$ which indicates that the value of the random variable will be known at time t.

Even though the value x of random variable X is not known before time t, the behavior of X can be described by its distribution F_X. In general, a cumulative distribution function $F_X\colon \mathbb{R} \to \mathbb{R}$ is defined by the probability that a given random variable X assumes values less than or equal to x:

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) \\ &= \int_{-\infty}^{x} f_X(s)\, ds \end{aligned}$$

Hereby, the integral based definition of a cumulative distribution function involves the probability density function f_X. With these two concepts at hand, the expectation operator can be defined by

$$E[X] = \int_{\Omega} X(\omega)\, d\mathbb{P}(\omega) = \int_{\mathbb{R}} x\, dF_X(x) = \int_{\mathbb{R}} x\, f_X(x)\, dx,$$

given the condition that $\int_\Omega \mid X(\omega) \mid d\mathbb{P}(\omega) < \infty$. The transition from the calculation of the expected value based on the cumulative distribution function to a density based representation is done by the fact that $dF_X(x) = f_X(x)\, dx$.

A useful feature of an expected value is that cumulative distributions, and therefore probabilities, can be written as such:

$$\begin{aligned} F_X(x) &= E\left[\mathbb{1}_{X \leq x}\right] \\ &= \int_{\mathbb{R}} \mathbb{1}_{s \leq x}\, f_X(s)\, ds \\ &= \int_{-\infty}^{x} f_X(s)\, ds = \mathbb{P}(X \leq x), \end{aligned}$$

where $\mathbb{1}_A$ represents the indicator function. In general the indicator function assumes either of two values: value one if $x \in A$ and zero otherwise.

Another important concept when dealing with random variables is the Fourier transform of the distribution function called characteristic function. In general, Fourier transforms

[7] A probability space is complete if its σ-algebra contains all subsets of the outcome set Ω.

are defined according to their field of application. In physical sciences and engineering the Fourier transform $\hat{f}\colon \mathbb{R} \to \mathbb{C}$ and the inverse Fourier transform $f\colon \mathbb{R} \to \mathbb{C}$ are typically defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x)\ e^{-2\pi i \xi x}\, dx \tag{2.1}$$

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi \tag{2.2}$$

The transform pair in (2.1) and (2.2) are especially handy due to the symmetry inherited within the equations. Every time we use the hat-notation ($\hat{f}$, $\hat{g}$,...) in the following, we refer to this specific transform pair. In a probabilistic environment, standard notation evolved in a way that the angular frequency $u = 2\pi\xi$ is used and, moreover, the signs of the complex exponentials are interchanged. Thus, the characteristic function $\phi_X(u)\colon \mathbb{R} \to \mathbb{C}$ of a random variable X with density function $f_X\colon \mathbb{R} \to \mathbb{R}$ is given by

$$\phi_X(u) = \int_{\mathbb{R}} f_X(x)\, e^{iux}\, dx \tag{2.3}$$

and its inverse by

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} f_X(x)\, e^{-iux}\, dx.$$

Equation (2.3) indicates that, similar to a cumulative distribution function, a characteristic function is defined as an expected value $\phi_X(u) = E\left[e^{iuX}\right]$.

An especially heavily used feature of a characteristic function is its direkt link to the cumulants of a random variable and, therefore, also to the moments of a random variable via

$$c_n = \frac{1}{i^n} \frac{\partial^n \ln(\phi_X)}{\partial u}\Big|_{u=0}$$

The function c_n is called cumulant generating function. Its importance stems from the fact that c_n enables us to calculate the moments of an arbitrary probability distribution with known characteristic function. We use this feature, e.g., in answering the question where to truncate the risk neutral expectation integral. For our purposes, the first four moments

$$\begin{aligned} E[X] &= c_1 \\ Var[X] &= c_2 \\ s[X] &= \frac{c_3}{c_2^{\frac{3}{2}}} \\ \kappa(X) &= \frac{c_4}{c_2^2} \end{aligned}$$

are of special interest. Hereby, besides the expected value $E[X]$ and the variance $Var[X]$, $s[X]$ and $\kappa(X)$ define skewness and kurtosis of a random variable's distribution.

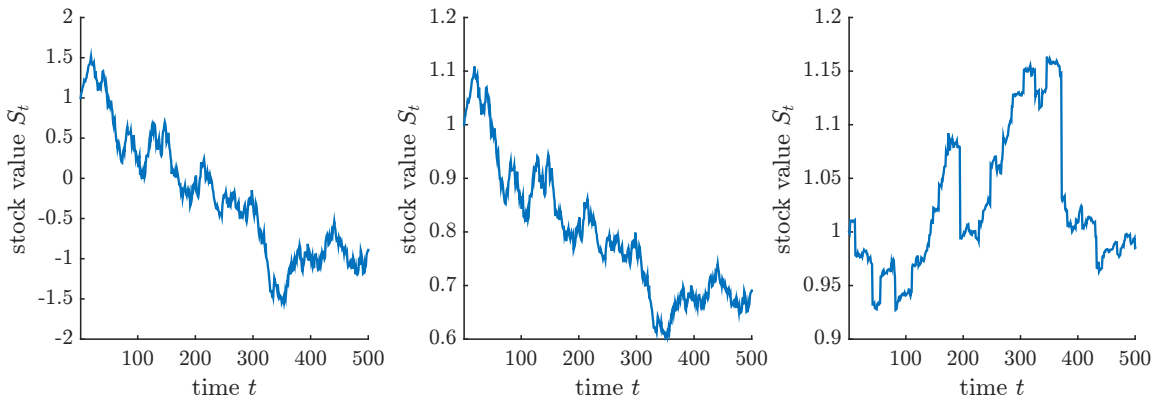

Figure 2.2: Evolution of stochastic processes in finance: Bachelier's model $S_t = S_0 + \mu t + \sigma W_t$ on the left, Black-Scholes $S_t = S_0 e^{\mu t + \sigma W_t}$ in the middle and a Lévy model $S_t = S_0 e^{\mu t + X_t}$ on the right.

Stochastic processes

A stochastic process is a timely ordered collection of random variables $\{X_t\}_{t \in T}$. As it is the case for random variables, stochastic processes are defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and do assume values in $\mathbb{R}^n$. For our application, this probability space is always equipped with a filtration $\mathcal{F}_t$. Similar to an $\mathcal{F}_t$-measurable random variable, a stochastic process whose values are revealed by the information flow within $\mathcal{F}_t$ is considered as being non-anticipating. Calling a stochastic process non-anticipating is equivalent to stating that the process $\{X_t\}_{t \in [0,T]}$ is $\{\mathcal{F}_t\}_{t \in [0,T]}$ adapted which is the most conventional nomenclature.

The use of stochastic processes within financial modeling has underwent different stages. Figure 2.2 indicates this evolution incorporating three distinct examples: Bachelier's model, Black-Scholes and exponential Lévy market models.[8] In its very beginnings, Bachelier (1900) models stock price movements by means of a Brownian motion $S_t = S_0 + \mu t + \sigma W_t$. However, stock prices are able to assume negative values as indicated in the figure. The famous Black and Scholes (1973) framework keeps the assumption of a Brownian motion but puts it in an exponential setting. In order to overcome the assumption of Normal distributed stock price movements, Lévy market models, in a next step, do allow jumps in the stock price path and, thus, introduce excess kurtosis and skewness among other things.

As implied by Figure 2.2, there are different ways to describe the market in which the financial instruments are traded in. To outlay the fundamental concepts, we concentrate on exponential Lévy market models of the form

$$S_t = S_0 \, e^{\mu t + X_t},$$

where μ is a drift factor and the underlying source of randomness is given by the stochastic process X_t. To describe the distribution of the stochastic process at each given step in time,

[8] It should be mentioned that this list is very limited and does not display a full picture of the use of stochastic processes in finance. It is purely indicated as visualization of different concepts in the field of stochastic processes.

characteristic functions are again the concept of choice. As long as we are working with infinitely divisible distributions, the most general starting point to define the characteristic function is through the Lévy-Khintchine formula

$$\begin{aligned}\phi(u) &= e^{t\psi(u)} \\ \psi(u) &= iu\mu - \frac{1}{2}\sigma^2u^2 + \int_{\mathbb{R}}(e^{iux} - 1 - iux\mathbb{1}_{|x|<1})\ \nu(dx),\end{aligned} \tag{2.4}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and ν being a Lévy measure following the usual conditions $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}}(1 \wedge |x|^2)\nu(dx) < \infty$. It can be seen from equation (2.4) that every Lévy process is composed of three elements: a deterministic drift part μ, a diffusion part (given $\sigma^2 > 0$) and a pure jump part (integral part in (2.4)).

Ultimately, we are interested in using stochastic processes to price financial contracts. One way to accomplish this task is to formulate risk neutral pricing relations. However, to understand the dynamics behind the expression risk neutral pricing, martingale processes and the concept of measure transformations have to be introduced.

Risk neutral pricing and equivalent martingale measures

In martingale theory, three different classes are distinguished: martingale processes, submartingales and supermartingales. If a time series shows no trend or periodicity it is a martingale process. Processes whose trajectories do increase on average are called submartingales, processes with a declining trend are labeled as supermartingales.

More technically, a martingale can be defined as a stochastic process $\{X_t\}_{t\in[0,T]}$ that lives on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ which is equipped with a filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ and for which it is true that $E[|\ X_t\ |] < \infty$ as well as

$$E[X_s|\mathcal{F}_t] = X_t, \quad \forall\, s > t.$$

Besides the fact of assuming the expected value to be finite, the above statement says that the best prediction of the future value of the stochastic process at time s is its value at time t. Thus, if we try to price a contingent claim whose value is derived from an underlying stochastic process that is a martingale, we can use the discounted expected value as the contracts value. Unfortunately, most of the observable time series are not martingales. To use the concept of martingales nevertheless, probability measures have to be defined that are equivalent to the real world measures connected to submartingales and supermartingales but induce the resulting process to be a martingale. Such a concept is called an equivalent martingale measure.

The link between the value of a derivative and a risk neutral expected value can be established by the Feynman-Kac theorem. Assuming that the underlying asset follows an Itô process[9], this theorem gives a stochastic representation to solutions of a partial differential equation

[9] An Itô process is a stochastic process following a stochastic differential equation in form of $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ (Øksendal, 2003, p. 110).

(Bingham and Kiesel, 2004, p. 202). In a financial setting this translates to the fact that the solution to the partial differential equation that is connected to an option's value is given by a conditional expectation. Thus, following Cont and Tankov (2004), the value of an arbitrary contingent claim at time t with terminal payoff V that is traded in an arbitrage free market described by the probability measure $\mathbb{P}$ can be represented as

$$v_t(V) = e^{-r(T-t)} E^{\mathbb{Q}}[V|\mathcal{F}_t],$$

where $\mathbb{Q}$ represents an equivalent martingale measure. From an intuitive point of view, an equivalent martingale measure ensures that the value of a financial claim can be computed by means of an discounted expected value. If the process under the real world measure $\mathbb{P}$ is not a martingale, measure $\mathbb{Q}$ must somehow reweight the probabilities within the probability measure $\mathbb{P}$. This reweighting of probabilities is exactly what a change of measure does: It builds a new stochastic process by assigning new probabilities to the events within $\mathcal{A}$. The only reason this approach is called risk neutral is due to the use of the expected value. It does not imply anything about investors risk attitude.

Two questions remain to be answered: What does it mean to be an equivalent martingale measure and how to find them? To answer the first question we state that the probability measure $\mathbb{Q}$ is said to be equivalent to another probability measure $\mathbb{P}$ ($\mathbb{Q} \sim \mathbb{P}$) if

i) they share the same null set and

ii) the discounted stock price process is a martingale under $\mathbb{Q}$.

The former statement herby only clarifies that events that are impossible under the probability measure $\mathbb{P}$ are impossible under $\mathbb{Q}$ as well. In other words, only such events that are feasible under $\mathbb{P}$ are ought to be feasible under $\mathbb{Q}$.

The second question is somehow harder to answer. Kreps (1981) states that in a continuous time framework it can be shown that the existence of an equivalent martingale measure implies an arbitrage free market. The reverse is however not always true. A somehow stronger argument than no-arbitrage needs to be considered. Delbaen and Schachermayer (1994) prove that a equivalent martingale measure exists under the condition that there is no free lunch with vanishing risk. Besides existence of an equivalent martingale measure uniqueness is another important issue. It can be shown that uniqueness implies market completeness, i.e. a market in which a contingent claim can be perfectly hedged. The Black-Scholes framework, e.g., describes a complete market with a unique equivalent martingale measure. In incomplete markets, however, the equivalent martingale measure is not unique and we have to choose it in some way. The exponential Lévy market model is one example of an incomplete market. Within this thesis, we choose to work with mean-correcting martingale measures. Hereby, the original drift term of the stochastic process in (2.4) is modified as follows (Schoutens, 2003, p. 79):

$$m = \mu + r - q - \ln\left[\phi(-i)\right],$$

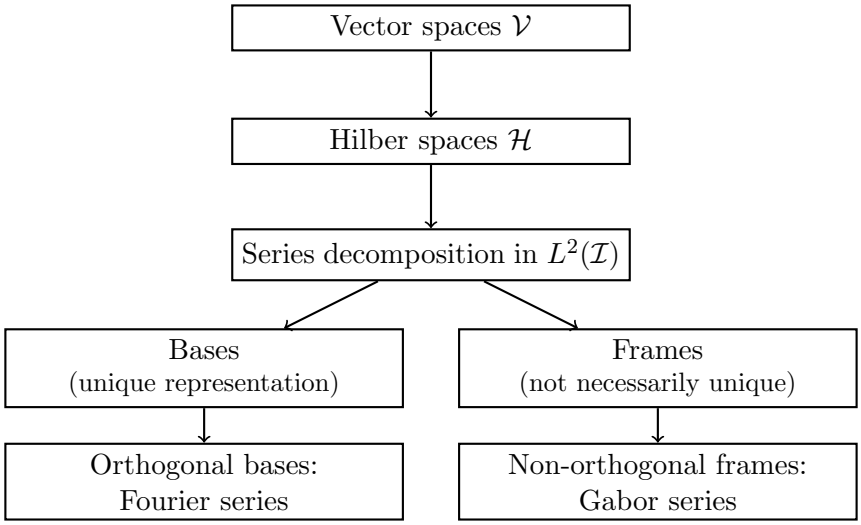

Figure 2.3: Preliminaries in linear algebra.

where r represents the risk free interest rate and q covers costs of carry. Both, explicit forms of the characteristic function and mean correction terms can be found in Schoutens (2003) for a wide range of Lévy processes, Ornstein-Uhlenbeck processes driven by Lévy processes as well as Lévy models with stochastic volatility.

2.2 Series approximation

This upcoming section introduces the necessary tools from linear algebra needed to install different types of series approximations[10]. Figure 2.3 contains an overview of the topics that are covered. The starting point is a definition of vectors and vector spaces $\mathcal{V}$ as well as function spaces as a part of vector spaces. The concept of bases are introduced twice: on an exemplary level when defining vector spaces and in a more detailed level when considering series decompositions. However, before series decomposition techniques on a bounded interval $\mathcal{I}$ are discussed, Hilbert spaces $\mathcal{H}$ are introduced.

2.2.1 Vector spaces, function spaces and bases

Nearly all of the calculations that are done in the following chapters are carried out in vector spaces. Thus the question arises: What is a vector space? The answer to this question is supposed to serve as a starting point from which concepts such as bases and frames and their role in function approximation are introduced.

Most importantly, a vector space is not always a space with traditional vectors in it. Following the definition of a vector, functions can be seen as vectors as well. In its most general definition, a vector is a mathematical construct that obeys the following axioms:

[10] Within this thesis, we use the terms series approximation, series expansion and series decomposition as synonyms.

i) Given the vectors a, b from some set $\mathcal{V}$, there is a vector $a + b$ which also comes from $\mathcal{V}$.

ii) Commutativity: $a + b = b + a$.

iii) Associativity: $a + (b + c) = (a + b) + c$.

iv) There is a zero vector included in $\mathcal{V}$ such that $a + 0 = a$.

v) Each $a \in \mathcal{V}$ comes with an inverse element $-a$ such that $a + (-a) = 0$.

vi) Given a scalar λ and a vector a it is true that $\lambda \cdot a \in \mathcal{V}$.

Since a function also follows these axioms, functions can be considered as being vectors. Thus, when defining a vector space as a construct whose elements are vectors, a function space is an algebraic concept whose elements are considered as functions. While each function space is a vector space, not all vector spaces are function spaces.

Within a vector space, bases are an important concept. Bases form a set of elements which can be used in a linear combination to uniquely construct all vectors within the associated space. As an example, we consider the vector space $\mathbb{R}^3$. Associated to $\mathbb{R}^3$ is the set

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

which forms the basis for this vector space. Using a linear combination of these elements, all vectors in $\mathbb{R}^3$ can be described. Thus, bases can be used to decompose a given vector into basic elements of the vector space itself.

In most cases, however, we are not only interested in the vector space itself but rather in a vector space that is equipped with an inner product which yields directly to Hilbert spaces $\mathcal{H}$. Given y, z are two vectors and f, g are two functions, the inner product is defined by

$$\langle y_k, z_k \rangle = \sum_{k \in \mathbb{N}} y_k \, \overline{z_k} \quad \text{and} \tag{2.5}$$

$$\langle f, g \rangle_{\mathcal{I}} = \int_{\mathcal{I}} f(x) \, \overline{g(x)} \, dx = \langle f, g \rangle \,, \tag{2.6}$$

respectively.

$$\langle f, f \rangle = \int_{\mathcal{I}} f(x) \, \overline{g(x)} \, dx$$

Attached to the inner product in (2.6), and sometimes written as an index, is the integration domain $\mathcal{I}$ on which the functions are defined on. In most cases, however, the index is dismissed. In addition, if $y, z \in \mathbb{C}^n$ are complex valued vectors or if $f, g \colon \mathbb{R} \to \mathbb{C}$ are complex valued functions, $\overline{z_k}$ and $\overline{g(\cdot)}$ indicate complex conjugates[11]. Given inner products as in

[11] The complex conjugate of a complex number shows identical real part but an imaginary part with opposite sign: $z = a + b \cdot i \to \overline{z} = a - b \cdot i$.

(2.5)-(2.6) exist on a vector space, the combination of both is called Hilbert space or inner product space. Within this thesis, the relevant Hilbert space is given by the space of Lebesgue measurable functions $L^2([a,b])$ on a bounded interval $\mathcal{I} = [a,b] \subset \mathbb{R}$. Hereby, the Hilbert space is defined by the vector space and its associated inner product:

$$L^2([a,b]) = \left\{ f \colon \mathbb{R} \to \mathbb{C} \;\middle|\; \int_{[a,b]} \mid f(x) \mid^2 dx < \infty \right\}$$
$$\langle f, g \rangle = \int_{[a,b]} f(x)\, \overline{g(x)}\, dx.$$

Hilbert spaces are of major importance due to the fact that a series decomposition is built upon inner products. Thus, working within a Hilbert space ensures the existence of a series representation. These series representations rely on basic building functions that are weighted by a set of coefficients.

Ultimately, to reconstruct a function $f \in \mathcal{H}$ as a superposition of a sequence of coefficients $\{f_k\}$, we are looking for conditions on this sequence $\{f_k\}$. In most cases, an exact decomposition cannot be done with a finite number of elements. Thus, the theory needed to accomplish a given series approximation is carried out in infinite-dimensional vector spaces rather than finite-dimensional vector spaces. However, in real-world applications, a truncation of the series is necessary which introduces some noise to the system. To control the level of accuracy, we are interested the most in finding finite sequences in infinite dimensional vector spaces. In other words, we are looking for sequences $\{f_k\}_{k=1}^{\infty}$ where a finite number of elements is non-zero. Both, bases and frames are instruments that utilize that thought.

2.2.2 Orthogonal bases and Fourier series

In the example of an Euclidean vector space $\mathbb{R}^3$, bases are introduced as a set which elements can be used in a unique way to describe all other vectors in the space. A similar linear combination can be found when using functions $e_k \colon \mathbb{R} \to \mathbb{C}$ as bases. Hereby, a sequence of functions $\{e_k\}_{k=1}^{\infty}$ is a basis with respect to a Hilbert space $\mathcal{H}$ if each function $f \in \mathcal{H}$ is uniquely described by a set of coefficients $\{f_k\}_{k=1}^{\infty}$ in a linear way:

$$f = \sum_{k=1}^{\infty} f_k \; e_k$$

Although other specifications of e_k are possible, orthonormal bases are a common choice and convenient to implement. For a basis to be orthonormal on a given interval $\mathcal{I} = [a,b]$ it must be true that

$$\langle e_k, e_j \rangle_{[a,b]} = \langle e_k, e_j \rangle = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases} . \tag{2.7}$$

Condition (2.7) separates an orthonormal basis from an orthogonal basis. For a system to be orthogonal, it must be true that $\langle e_k, e_j \rangle = 0$, $k \neq j$. For a system to be orthonormal, it must also be true that $\langle e_k, e_k \rangle = 1$.

When using an orthonormal basis to decompose an arbitrary function in terms of a linear combination, two facts are important for the progress of this thesis:

$$f = \sum_{k=1}^{\infty} \langle f, e_k \rangle \, e_k = \sum_{k=1}^{\infty} f_k \, e_k \qquad \forall f \in L^2([a,b]) \tag{2.8}$$

$$\|f\|^2 = \sum_{k=1}^{\infty} | \langle f, e_k \rangle |^2 \qquad \forall f \in L^2([a,b]) \tag{2.9}$$

$$\text{where} \quad \|f\|^2 = \int_a^b | f(x) |^2 \, dx$$

Equation (2.8) states that every function that lies within a Hilbert space $\mathcal{H}$, here we assume $L^2([a,b])$, can be expressed by an infinite sum that is built upon an orthonormal bases $\{e_k\}_{k=1}^{\infty}$ and a set of coefficients. The set of coefficients $\{f_k = \langle f, e_k \rangle\}_{k=1}^{\infty}$ is given by the inner product of the orthonormal basis itself and the function that is to be decomposed. Another important insight is stated in equation (2.9) and is often named as Parseval's equation. If Parseval's equation holds, the system $\{e_k\}$ is said to be complete. On a more intuitive basic, equation (2.9) shows that the sum of the squared coefficients converge for an arbitrary square integrable function due to the fact that the left-hand side of the equation is finite. As a result, the coefficients also converge to zero as k grows large. The practical importance of equation (2.9) lies in its immediate consequences on the calculation of integrals that consist of two functions by means of their series coefficients.

Up to this point, we dealt with orthonormal basis in a general setting. Now, we introduce Fourier series on a bounded interval $\mathcal{I} = [0, \frac{1}{b}]$ (Christensen, 2008, p. 69). Fourier series come in different flavors. One way is to use an exponential function as the building block of the orthonormal system $\left\{ e_k(x) = b^{\frac{1}{2}} e^{2\pi i k b x} \right\}_{k \in \mathbb{Z}}$ on the interval $[0, \frac{1}{b}]$. A series representation based on this particular choice of orthonormal system is called complex Fourier series due to the complex valued basic function. We use this type of Fourier series heavily in Chapter 3. Other types of Fourier series are given by building functions being sine functions, cosine functions or a combination of both. Chapter 4 relies on a multivariate version of sine and cosine series and Chapter 5 incorporates univariate cosine series.

Although the system $\{e_k\}_{k \in \mathbb{Z}}$ is an orthonormal system on the interval $[0, \frac{1}{b}]$, standard notation of Fourier series write the expansion $f = \sum_{k \in \mathbb{Z}} \langle f, e_k \rangle \, e_k$ only in terms of the exponential function

$$f(x) = \sum_{k \in \mathbb{Z}} f_k \, e^{2\pi i k b x} \tag{2.10}$$

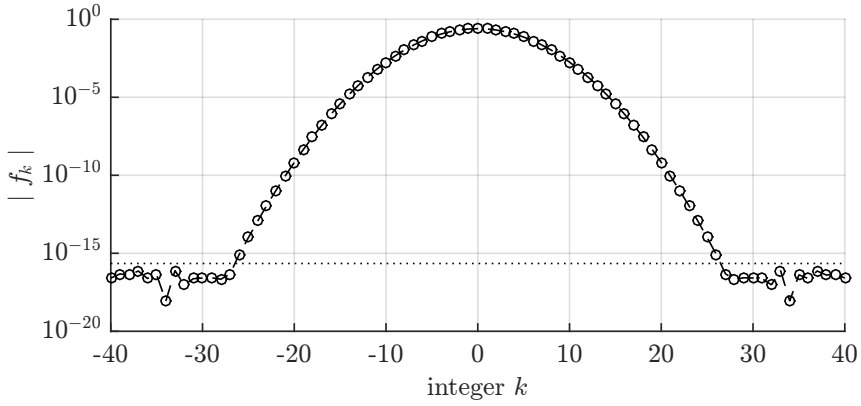

Figure 2.4: Fourier series coefficients of a Normal probability density function $N(2, 0.2)$.

and pull the factor $b^{\frac{1}{2}}$ into the set of coefficients

$$f_k = b^{\frac{1}{2}} \langle f, e_k \rangle = b \int_0^{\frac{1}{b}} f(x)\, e^{-2\pi i k b x}\, dx. \tag{2.11}$$

Using orthonormal basis such as the sequence $\{e_k\}_{k\in\mathbb{Z}}$ within the series representation (2.10) yields to a unique definition of the according dual basis $\{\overline{e_k}\}_{k\in\mathbb{Z}}$ and, therefore, to a unique set of coefficients f_k given by (2.11).

To visualize the concept of a sequence with a finite number of non-zero elements, we consider a Normal probability density function with mean μ and standard deviation σ given by $N(\mu, \sigma) = N(2, 0.2)$ and calculate the coefficient vector f_k which is displayed in Figure 2.4. Within the figure, two components are included: the coefficients itself and, as a reference value, a horizontal line at $\epsilon = 2.2 \cdot 10^{-16}$ which indicates the floating-point relative accuracy. With a slight abuse of its meaning, we consider this level to be the threshold below which coefficients can be considered as being zero. Thus, when defining $|\, f_k \,| < \epsilon$ to be zero, the example in Figure 2.4 implies that the set of coefficients $\{f_k\}_{k\in\mathbb{Z}}$ contains 52 non-zero entries. An accurate approximation of the above Normal density function on the given interval can be accomplished using only these non-zero elements in a linear combination with the belonging basis function.[12]

After outlying a specific example of a complex Fourier series, we would like to pick up Parseval's equation once again to elaborate on the calculation of an integral by means of Fourier coefficients: Assuming two functions $f, g \in L^2(\mathcal{I})$ with Fourier coefficients f_k and g_k, it can be shown that the sum

$$f + g = \sum_{k\in\mathbb{Z}} f_k e_k + \sum_{k\in\mathbb{Z}} g_k e_k = \sum_{k\in\mathbb{Z}} (f_k + g_k) e_k$$

[12] The above example is solely chosen for illustrative purposes and does not necessarily represent the coefficients' convergency behavior of any arbitrary square integrable function. When assuming a Normal distribution with a different standard deviation, the coefficient vector assumes different values. However, the overall convergency behavior is preserved in this case.

has Fourier coefficients f_k+g_k and, accordingly, the difference has Fourier coefficients f_k-g_k. Applying Parseval's equation to the sum and the difference individually, it is true that

$$\int_{\mathcal{I}} [f(x)+g(x)]^2 \, dx = \sum_{k\in\mathbb{Z}} (f_k+g_k)^2 \quad \text{and}$$
$$\int_{\mathcal{I}} [f(x)-g(x)]^2 \, dx = \sum_{k\in\mathbb{Z}} (f_k-g_k)^2 .$$

When we subtract both equations from each other, square-terms cancel out leaving only cross-terms left. Now, after multiplying by $\frac{1}{4}$, the following result appears:

$$\int_{\mathcal{I}} f(x)\, g(x)\, dx = \sum_{k\in\mathbb{Z}} f_k\, g_k \tag{2.12}$$

Thus, equation (2.12) uncloses a technique to calculate an integral by means of the inner product of the functions' Fourier coefficients.

2.2.3 Non-orthogonal frames and Gabor series

Orthogonality is an inherit characteristic of bases which is at the same time a strong condition on the family of vectors that are used to decompose a function $f \in \mathcal{H}$. Thus, the question arises whether orthogonality is necessary in any circumstances. On the one hand, orthogonality prevents redundancy but, on the other hand, comes with the disadvantage that it includes a low fault tolerance. A concept that is far more general than bases are is the concept of frames. Frames are also used to decompose a function by means of a linear combination of vectors. These vectors are, in contrast to the vectors used in the theory of bases, linear dependent.

Within this subsection, we intend to provide the reader with an intuition for the differences between function approximation using bases and using frames. Moreover, we introduce Gabor frames as a special class of frames that are used in Chapter 3 to price European options.

Frames, similar to bases, rely on basic building functions $F_k \colon \mathbb{R} \to \mathbb{C}$. In its most general version, following Christensen (2008), a frame is defined as a sequence of these basic elements $\{F_k\}_{k=1}^{\infty}$ in $\mathcal{H}$ for which it is true that, given two constants $A, B > 0$,

$$A\|f\|^2 \leq \sum_{k=1}^{\infty} |\, \langle f, F_k \rangle \,|^2 \leq B\|f\|^2, \quad \forall f \in \mathcal{H} \tag{2.13}$$

Comparing equations (2.13) and (2.9), shows that the function norm is no longer identically preserved in the sum of the coefficient norms but rather lies within an interval that is defined by the frame bounds A and B. Thus, equation (2.13) can be considered as a generalized version of Parsevals equation given by (2.9). If $A = B$ equality is restored and the resulting

frame is called a tight frame. Tight frames share many properties with bases. Since one of these properties is being quite restrictive, we do not consider tight frames in the following.

Within the theory of frames, we are interested the most in procedures to reconstruct an arbitrary function. Using bases to approximate a given function in $\mathcal{H}$ yields to a unique definition between a basis and its dual basis. As for orthonormal bases, a simple link is present: The dual basis to the basis e_k is uniquely given by its complex conjugate $\overline{e_k}$. This simple connection is, however, no longer given when considering frames to decompose a function. Now, in addition to a frame itself, a frame operator $S\colon \mathcal{H} \to \mathcal{H}$ is necessary to construct a dual frame in form of the sequence $\{S^{-1}F_k\}_{k=1}^{\infty}$. This dual frame is the counter piece in frame theory to a dual basis in the theory of bases.

Given a frame $\{F_k\}_{k=1}^{\infty}$ and a dual frame $\{S^{-1}F_k\}_{k=1}^{\infty}$, we can write down an arbitrary function $f \in \mathcal{H}$ in terms of

$$f = \sum_{k=1}^{\infty} \left\langle f, S^{-1}F_k \right\rangle F_k \tag{2.14}$$

$$= \sum_{k=1}^{\infty} \langle f, F_k \rangle \, S^{-1}F_k \tag{2.15}$$

A major difference between the series expansions in (2.14) and (2.15) compared to (2.8) is given by the fact that in most cases the coefficients of a frame decomposition are not unique, and, finding and calculating the inverse of the frame operator S may be highly cumbersome. Nevertheless, due to the fact that the elements within a frame do not necessarily be orthogonal, allows frames to be a more general solution to the problem of function approximation than orthonormal expansion methods are.

We now turn to a special class of frames called Gabor frames. Gabor frames are built upon Gabor systems, also known as Weyl-Heisenberg systems, which, in turn, are built upon the operators of translation T_z and modulation M_ω:

$$\begin{aligned}(T_z g)(x) = T_z g \;\; &= g(x-z) \quad \text{and} \\ (M_\omega g)(x) = M_\omega g &= g(x)\, e^{2\pi i \omega x}\end{aligned}$$

While the modulation operator is also included in the complex Fourier series (2.10), the translation operator adds a new element. Hereby, the distinctive feature of T_z is that it allows to analyze local behavior within a function more closely.

As mentioned before, frames rely on basic building blocks. In the environment of Gabor frames, these building blocks are called generator functions. Given a generator function $g \in L^2(\mathbb{R})$ and two positive variables $\alpha, \beta > 0$, the set of functions

$$\mathcal{G}(g, \alpha, \beta) = \left\{ g(x - n\alpha)\, e^{2\pi i m \beta x} = T_{n\alpha} M_{m\beta}\, g \right\}_{m,n \in \mathbb{Z}}$$

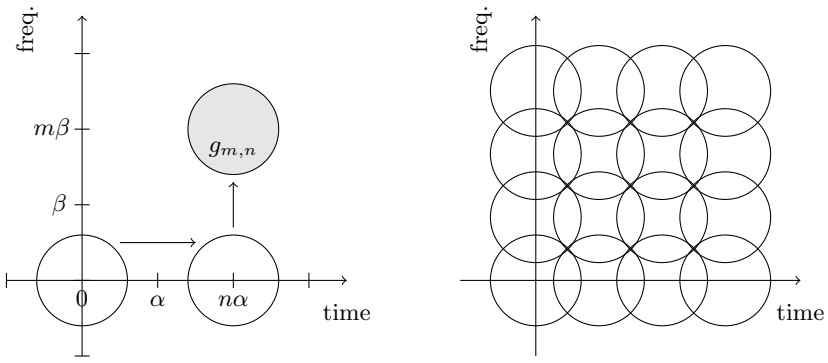

Figure 2.5: Schematic representation of a time-frequency lattice with Gabor particles included.

forms a Gabor system.[13] An element out of this system can be interpreted as a particle occupying one node on the time-frequency lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$. Hereby the localization of the generator g as well as the product $\alpha{\cdot}\beta$ are crucial ingredients. Figure 2.5 displays the grid in greater detail. Moving from one integer value to the other shifts the generator function along the time-frequency axes.

How to choose α and β is one of the central questions and is related to the uncertainty principle. In general, any time-frequency application is subject to Heisenberg's uncertainty principle

$$\int_{\mathbb{R}} |\ x\ g(x)\ |^2\ dx\ \int_{\mathbb{R}} |\ \omega\ \hat{g}(\omega)\ |^2\ d\omega \geq \frac{\|g\|^4}{16\pi^2} \tag{2.16}$$

which states that a better localization in time domain x goes along with a worse localization in frequency domain ω (Stein and Shakarchi, 2011, p. 158). In a way, equation (2.16) marks the lower limit of time-frequency localization.

Within the theory of Gabor systems, localization heavily depends on the product $\alpha{\cdot}\beta$. Figure 2.5 shows a redundant system with $\alpha{\cdot}\beta < 1$. The redundancy is characterized by the overlapping areas on the right hand side of the figure. Such a redundant system yields to a good localization in time and frequency. In addition to a redundant system, a parameter set with $\alpha{\cdot}\beta > 1$ results in an incomplete scheme where no function approximation is feasible. Finally, we consider a system with $\alpha{\cdot}\beta = 1$, called critical density (Gröchenig, 2001, p. 163). Such a system is redundancy-free and is even able to form an orthonormal basis. However, regardless the choice of function g in $\mathcal{G}(g, \alpha, \beta)$, there will be no system with good time-frequency localization at the critical density. In a mathematical sense, this implies that the inequality in (2.16) equals infinity

$$\int_{\mathbb{R}} |\ x\ g(x)\ |^2\ dx\ \int_{\mathbb{R}} |\ \omega\ \hat{g}(\omega)\ |^2\ d\omega = \infty \tag{2.17}$$

[13] By referring to Gabor systems, we concentrate on regular Gabor systems where the combinations $(n\alpha, m\beta)_{m,n\in\mathbb{Z}}$ form a lattice in $\mathbb{R}^2$.

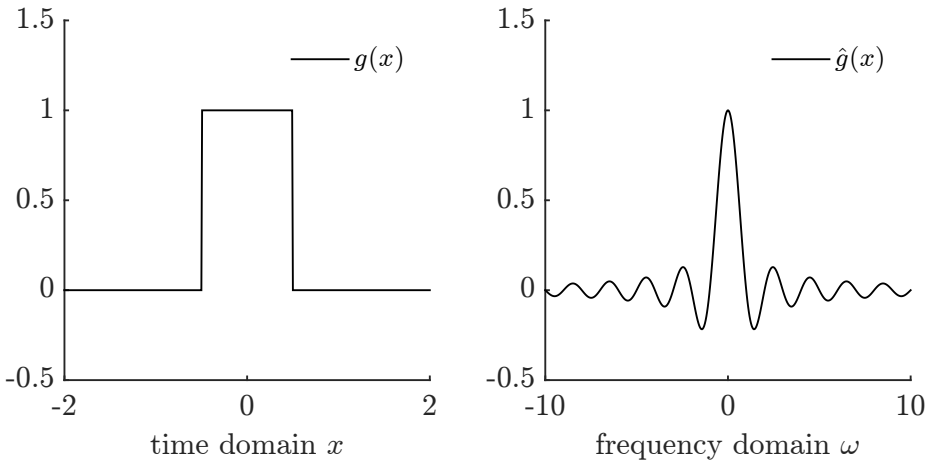

Figure 2.6: Rectangular function in time domain and frequency domain.

In a time-frequency environment, equation (2.17) is known as Balian-Low theorem. It states that, given an an orthonormal basis, either g or its Fourier transform $\hat{g}$ is not decaying and is therefore not localized in either time or frequency.

To visualize this thought, we use the standard example of a rectangular function that assumes non-zero values within the interval $\mathcal{A} = -\frac{1}{2} \leq x \leq \frac{1}{2}$:

$$g(x) = \mathbb{1}_{\mathcal{A}}(x)$$
$$\hat{g}(\omega) = \int_{-\infty}^{\infty} g(x)\, e^{-2\pi i \omega x}\, dx = \frac{\sin(\pi\omega)}{\pi\omega}$$

Figure 2.6 shows both the evolution of the rectangular function g in time domain and in frequency domain. The system $\mathcal{G}(\mathbb{1}_{\mathcal{A}}(x), 1, 1)$ does indeed form an orthonormal basis. However, whereas an extremely fast convergency is present in time domain, the Fourier transform $\hat{g}$ is not converging at all in frequency domain. As a result of $\hat{g}$ not converging, the integral $\int_{-\infty}^{\infty} \mid \omega\, \hat{g}(\omega) \mid^2 \, d\omega$ tends to infinity and, thus, the Balian-Low theorem in (2.17) holds. An expansion based on such a Gabor system will fail to reconstruct the original function.

Instead of an orthonormal bases, a less demanding concept is introduced with frames. The difference between bases and frames is that the elements of a basis are linearly independent while the elements of frames are not. That is, using frames instead of bases brings some redundancy into the approximation. In other words, the price for keeping the property of an unconditionally convergent expansion is to give up uniqueness of the coefficients (Christensen (2008, p. 93)) and therefore move to a system with $\alpha \cdot \beta < 1$.

To ensure a proper Gabor frame approximation, some preliminaries have to be put in place that cope with the existence of Gabor frames; moreover, the concept of dual frames has to be discussed. Similar to equation (2.13), the system $\mathcal{G}(g, \alpha, \beta)$ forms a Gabor frame if there

exist two constants $A, B > 0$ such that

$$A\|f\|^2 \leq \sum_{m,n\in\mathbb{Z}} |\langle f, T_{n\alpha}M_{m\beta}g\rangle|^2 \leq B\|f\|^2. \tag{2.18}$$

Thus, in case of Gabor frames the norm of function f is also no longer identically preserved in the coefficients but is rather defined within an interval between the lower frame bound A and the upper frame bound B. For our purposes, however, it is not necessary to determine these frame bounds. Instead, once the existence of a Gabor frame is ensured, a second frame is needed to facilitate any kind of function approximation. This frame is based on the Gabor frame operator S_g,

$$S_g f = \sum_{m,n\in\mathbb{Z}} \langle f, T_{n\alpha}M_{m\beta}g\rangle\, T_{n\alpha}M_{m\beta}g.$$

Since the operator is invertible, function f can be decomposed into

$$f = \sum_{m,n\in\mathbb{Z}} \langle f, T_{n\alpha}M_{m\beta}g\rangle\, T_{n\alpha}M_{m\beta}\gamma \tag{2.19}$$

$$= \sum_{m,n\in\mathbb{Z}} \langle f, T_{n\alpha}M_{m\beta}\gamma\rangle\, T_{n\alpha}M_{m\beta}g, \tag{2.20}$$

where $\gamma = S_g^{-1}g$. The expression γ is called dual function due to the fact that it is part of a (canonical) dual frame $\mathcal{G}(\gamma, \alpha, \beta)$ to the original frame $\mathcal{G}(g, \alpha, \beta)$ (Gröchenig (2001, p. 94)). As can be seen from (2.19) and (2.20), the frame and its dual can be interchanged within the function approximation.

Chapter 3

Non-Orthogonal Option Pricing

3.1 Introduction

In recent years, option pricing based on more sophisticated assumptions about the associated stochastic process for the underlying asset has received a high level of attention. While the centerpiece of the famous Black-Scholes formula is the assumption of a geometric Brownian motion as a driving factor, current option pricing methods often rest upon continuous-time asset models based on exponential semimartingale processes. Besides pure jump Lévy processes, this class of stochastic processes also encompasses jump diffusion processes and affine processes in general, which allows them to represent stylized facts of asset returns more closely. The higher level of complexity in the assumptions translates into a higher level of complexity in the numerical implementation of the evaluation methods. Special attention is therefore set on efficient methods that lower the computational load.

In many cases, the implementation is based on the knowledge of the characteristic functions. As the characteristic function is a mere Fourier transformation of the probability density function, Fourier transform methods became of interest. Since Carr and Madan (1999) published their method to price European options based on characteristic functions using the Fast Fourier Transform (FFT) algorithm, this area of research picked up pace and developed further. Although the FFT algorithm is highly efficient, the implementation is subject to the Nyquist relation[14] yielding to a situation where more computations than necessary have to be executed. A refinement to the FFT method of Carr and Madan was developed by Chourdakis (2005). He introduced the fractional fast Fourier transform (FrFFT) into option pricing and thereby disconnected the strike price grid from the integration variable grid.

Alongside other developments, Fourier series methods were introduced by Fang and Oosterlee (2008). By means of their procedure, called Cos method, they present a way to use orthogonal expansion methods, more precisely Fourier cosine series, to calculate the risk neutral pricing

[14] The Nyquist relation states in terms of option pricing that the grid size of the integration domain s and the grid size of the strike domain K are connected through $\Delta s\,\Delta K = \frac{2\pi}{N}$, where N is the number of terms within the FFT algorithm.

formula. Hurn et al. (2013) recently added complete Fourier series and Fourier sine series to the picture. As a next step, Ortiz-Gracia and Oosterlee (2013) proposed an algorithm based on Haar wavelets and B-spline wavelets to price European-type options. Due to the higher dimensionality of the algorithm, computational time increases when compared to, e.g., the Cos method. This disadvantage, however, is compensated by a better performance in pricing long term call options.

The focus of the present chapter[15] is related to Ortiz-Gracia and Oosterlee (2013) in the sense that our work uses tools from the field of time-frequency analysis. Instead of considering wavelet theory, we propose Gabor series expansion techniques for the given task. Both approaches are built on coherent systems that work with operators acting on given generator functions. However, the operators used in wavelet theory (translation and dilatation) differ from the ones used in time-frequency and Gabor theory (translation and modulation). To the best of our knowledge, Gabor series have not been considered yet in literature to calculate option prices.

In contrast to Fourier series theory, where expansions are supported by orthogonal bases, the kernel of a Gabor series is rather a frame than a basis, and, in most cases, this frame is non-orthogonal. Based on these frames, our procedure (Gabor method) consists of the decomposition of the risk neutral expected value of the payoff at maturity into two inner products in Hilbert space that are evaluated by means of Parseval's theorem and complex Fourier series. Using this framework, we are interested in the question whether time-frequency methods are capable of providing a stable and efficient pricing tool for European options. In more detail, we research whether a refinement of the expansion method, compared to e.g. cosine series as in the work of Fang and Oosterlee (2008), yields to a higher pricing accuracy? We find that time-frequency expansions can yield to an improved pricing accuracy. Especially in terms of contracts with a short time to maturity, this effect is most pronounced. In addition, compared to other pricing methods, our method shows a higher robustness with respect to the input parameters truncating the pricing domain.

The chapter is organized as follows: In Section 2, we discuss non-orthogonal expansion methods and the conditions needed to ensure a good approximation behavior. In Section 3, the pricing formula is introduced and Gabor frames are used to solve this option pricing problem. Section 4 presents an error analysis and Section 5 contains a numerical implementation of the Gabor method and tests its robustness. A conclusion can be found in Section 6 followed by an appendix.

3.2 Non-orthogonal expansion methods

This section presents a short summary of the theory. For a more detailed overview of frame theory, we refer to Gröchenig (2001) and Christensen (2008) as well as to the information stated in Section 2.2. As for the rest of this chapter, except stated otherwise, we operate in a

[15] A version of this chapter is available on SSRN as Niederstätter (2014).

Hilbert space $\mathcal{H}$ of Lebesgue measurable functions in $L^2(\mathbb{R})$ equipped with the norm $\|f\|^2 = \int_{\mathbb{R}} \mid f(x) \mid^2 dx$ and the inner product $\langle f, g \rangle = \int_{\mathbb{R}} f(x)\, \overline{g(x)}\, dx$ or any subset $L^2([a,b]) \subset L^2(\mathbb{R})$ with corresponding norm and inner product.

In 1946, Dennis Gabor proposed a series expansion of an arbitrary function $f \in L^2(\mathbb{R})$ by means of

$$f = \sum_{m,n \in \mathbb{Z}} c_{mn}\, T_{n\alpha} M_{m\beta}\, g \tag{3.1}$$

and thereby established the field of modern time-frequency analysis among others. In order to be able to use this relation for any kind of function approximation, we take a closer look at each of the components of equation (3.1). The expansion belongs to the class of atomic decomposition (Feichtinger and Gröchenig, 1992, p. 360) and is based on three different components: the fundamental operator of translation T_z in combination with the operator of modulation M_ω, Gabor series coefficients c_{mn} and a generator function $g \in L^2(\mathbb{R})$ also called Gabor atom. To start with the operators, we define for $z, \omega \in \mathbb{R}$ translation and modulation as

$$\begin{aligned} (T_z g)(x) = T_z g \quad &= g(x - z) \quad \text{and} \\ (M_\omega g)(x) = M_\omega g &= g(x)\, e^{2\pi i \omega x}, \end{aligned}$$

where i is the imaginary number. While operator T_z shifts a function g on the abscissa, in signal processing often referred to as time axis, operator M_ω influences the frequency of function g in complex space. As a consequence, combining both operations is called time-frequency shifting.[16]

In signal processing the compression or transmission of speech is an important topic. Instead of decomposing the audio signal as a whole into frequencies without knowledge about which frequencies are used at a given point in time, time-shifts allow to map a frequency decomposition such as Fourier series to specific time windows. If the relevant domain with respect to a audio signal is time, the relevant domain with respect to option pricing changes to the domain of feasible payoffs. This is also the domain on which the 'time-shifts' are made in context of this chapter.

As part of Chapter 2, differences between basis and frames are discussed. Using orthonormal basis has the advantage of coming up with a unique set of coefficients. However, uniqueness is not always needed. Other features such as an improved handling of local characteristics of the function that is to be approximated are more important in some circumstances. Frame theory, and Gabor theory in particular, are suitable methods to overcome some of the disadvantages of bases.

[16] Within the environment of this chapter the term 'time-shifting' might be misleading somehow since it does not relate to the option's time to maturity. We use the term nevertheless since it has been established as standard vocabulary for this kind of expansion methods.

Given a generator function $g \in L^2(\mathbb{R})$ and two positive variables $\alpha, \beta > 0$, the set of functions

$$\mathcal{G}(g, \alpha, \beta) = \left\{ g(x - n\alpha)\, e^{2\pi i m \beta x} = T_{n\alpha} M_{m\beta}\, g \right\}_{m,n \in \mathbb{Z}}$$

forms a Gabor system. As addressed in Section 2.2.3, it has to be ensured that the product $\alpha \cdot \beta < 1$ which yields to a redundant system. With that in mind, two additional elements of equation (3.1), besides the operators of modulation and translation, have to be discussed: the Gabor coefficients c_{mn} as well as the choice of the generator function g within the frame $\mathcal{G}(g, \alpha, \beta)$ and how wo calculate a dual frame connected to the original frame.

Similar to the concept of a basis and its dual basis, every frame has at least one dual frame. In contrast to a basis, there are a number of dual frames that can be used in combination with a single frame. However, before we move on to define the dual frame we use in this chapter, we elaborate on the coefficients first. Within Gabor frame theory, the existence and invertibility of a Gabor frame operator S_g is crucial. Assuming $\mathcal{G}(g, \alpha, \beta)$ is a proper Gabor frame, it is true that

$$S_g f = \sum_{m,n \in \mathbb{Z}} \langle f, T_{n\alpha} M_{m\beta} g \rangle\, T_{n\alpha} M_{m\beta} g.$$

Thus, since S_g^{-1} exists, a function f can be decomposed into

$$f = \sum_{m,n \in \mathbb{Z}} \langle f, T_{n\alpha} M_{m\beta} g \rangle\, T_{n\alpha} M_{m\beta} \gamma \tag{3.2}$$

$$= \sum_{m,n \in \mathbb{Z}} \langle f, T_{n\alpha} M_{m\beta} \gamma \rangle\, T_{n\alpha} M_{m\beta} g, \tag{3.3}$$

using a generator function g and its dual $\gamma = S_g^{-1} g$. With that in mind, we are able to define the Gabor series coefficients from the beginning of this section via equation (3.3) as

$$c_{mn} = \langle f, T_{n\alpha} M_{m\beta} \gamma \rangle .$$

Up to this point, we only claim that there exists a Gabor frame operator S_g which enables us to build the dual frame based on a generator function g. Since constructing and inverting the operator S_g can be highly cumbersome, we choose a special case in which formulating Gabor frames and its duals becomes much more convinient. This concept is known as 'painless non-orthogonal expansion' and goes back to the work of Daubechies et al. (1986). The following remarks, however, refer to Gröchenig (2001), Chapter 6.4. We define the generator function to be part of the Schwartz class $g \in \mathcal{S}(\mathbb{R})$ with support on an interval of length $L \leq \frac{1}{\beta}$ and the additional condition $\alpha \leq L$. Under these conditions, the dual window is defined as

$$\gamma(x) = \frac{\beta\ g(x)}{\sum\limits_{q \in \mathbb{Z}} g(x - q\alpha) \overline{g(x - q\alpha)}}, \tag{3.4}$$

where the bar indicates complex conjugates. With the relation in (3.4) we are able to explicitly compute a dual frame to any given frame that obeys the assumption of a limited

support on an interval of length L. Other methods of determining the dual frame involve the numerical calulation of the pseudo-inverse of a matrix based on the generator g. These procedures, however, are not feasible for the option pricing task at hand due to the fact that both the frame and its dual have to be evaluated within an integral (see Chapter 3.3). Therefore, an analytic expression for the dual frame as in equation (3.4) is necessary.

As already indicated by the Balian-Low theorem, a high approximation quality is only feasible under the condition of a redundant system. This, in turn, implies that $\alpha{\cdot}\beta < 1$ has to be guaranteed. A Gabor system $\mathcal{G}(g, \alpha, \beta)$ will never form a frame if $\alpha{\cdot}\beta > 1$. At the critical density $\alpha{\cdot}\beta = 1$, the Balian-Low theorem clarifies the lack of good time-frequency localization (see Section 2.2.3), which also yields to poor results.

3.3 Option pricing

Given an equivalent martingale measure $\mathbb{Q}$ and a filtration $\mathcal{F}_t$, the value of a European-style option with payoff $v(x, T)$ at maturity T is described by the expected value

$$\begin{aligned} v_t &= e^{-r(T-t)}\, \mathbb{E}^{\mathcal{Q}}\left[v(x,T)|\mathcal{F}_t\right] \\ &= e^{-r(T-t)} \int_{\mathbb{R}} v(x,T) f_X(x)\, dx. \end{aligned} \tag{3.5}$$

By truncating the moneyness $x = \log \frac{S_T}{K}$, with S_T being the value of the underlying at maturity and K being the strike price, on a given interval $x \in [a, b]$ and by inserting the Gabor expansion (3.2) of the risk neutral density function, the above pricing formula changes to

$$v_t = e^{-r(T-t)} \int_a^b v(x,T) \sum_{m,n\in\mathbb{Z}} \langle f_X, T_{n\alpha} M_{m\beta} g\rangle \, (T_{n\alpha} M_{m\beta} \gamma)(x)\, dx.$$

Interchanging the integral and the sums shows that option prices are defined by two inner products that will be discussed independently in the following.

$$\begin{aligned} v_t &= e^{-r(T-t)} \sum_{m,n\in\mathbb{Z}} \langle f_X, T_{n\alpha} M_{m\beta} g\rangle \int_a^b v(x,T)(T_{n\alpha} M_{m\beta} \gamma)(x)\, dx \\ &= e^{-r(T-t)} \sum_{m,n\in\mathbb{Z}} \langle f_X, T_{n\alpha} M_{m\beta} g\rangle \, \overline{\langle v, T_{n\alpha} M_{m\beta} \gamma\rangle}, \end{aligned} \tag{3.6}$$

where the bar above the second inner product indicates complex conjugates. The main part of this section is denoted to the calculation of (3.6). For this purpose, besides the choice of a Gabor frame $\mathcal{G}(g, \alpha, \beta)$ and its dual frame $\mathcal{G}(\gamma, \alpha, \beta)$, complex Fourier series as well as Parseval's theorem are crucial instruments that will be applied. Thus, we define a complex

Fourier series expansion of a function $f \in L^2([a,b])$ as

$$f(x) = \sum_{k\in\mathbb{Z}} f_k \, e^{-ik\pi\frac{x-a}{b-a}} \quad \text{, with}$$
$$f_k = \frac{1}{2(b-a)} \int_a^b f(x) e^{ik\pi\frac{x-a}{b-a}} \, dx.$$

Based on the completeness of this trigonometric system (see Section 2.2.2), Parseval's theorem states that if f and g are square integrable functions on $[a,b]$ the following holds:

$$\int_a^b f(x) \, \overline{g(x)} \, dx = 2(b-a) \sum_{k\in\mathbb{Z}} f_k \, \overline{g_k}.$$

Equipped with these relations and the information about Gabor systems from Section 2.2.3 as well as Section 3.2, we concentrate on the efficient computation of the inner products in equation (3.6) in the following two subsections. To do so, we first have to choose a generator function g that builds the foundation of the Gabor frame $\mathcal{G}(g,\alpha,\beta)$. The importance of fast decay has already been emphasized in the previous section. Thus, from the Schwarz space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing functions, we select the Gaussian function as (original) generator:

$$g(x) = e^{-p\pi(x-c)^2}$$

The Gaussian function by itself is defined on the whole real line and therefore harms the assumption of compact support on an interval of length $L = 1/\beta$. To overcome this problem, we multiply the indicator function $\mathbb{1}_A(\cdot)$ by the Gaussian. As a result, an enhanced generator function is given by

$$g(x) = e^{-p\pi(x-c)^2} \, \mathbb{1}_A(x-c),$$

where $A = \left\{ -\frac{1}{2\beta} \leq x \leq \frac{1}{2\beta} \right\}$ and $c \in \mathbb{R}$ represents a constant shift parameter. The corresponding dual function based on (3.4) assumes the form

$$\gamma(x) = \frac{\beta \, e^{-p\pi(x-c)^2} \, \mathbb{1}_A(x-c)}{\sum\limits_{q\in\mathbb{Z}} e^{-2p\pi(x-c)^2} \, \mathbb{1}_A(x-c-q\alpha)}.$$

The sharp edges that are introduced by the indicator function do not harm when controlling variable p of the generator function appropriately. In general, cutting the generator function at the lower and upper end by means of the indicator function implies slow convergency in the frequency domain.[17] However, in every numerical implementation, calculations have to be done with a finite number of digits. This means that a number is defined through an interval rather than a single point on the axis. Within Matlab's floating point arithmetic, the machine precision − or machine epsilon[18] − is around $\epsilon \approx 2.22 \cdot 10^{-16}$. Thus, to ensure that the Gaussian generator approaches zero before the limits of the indicator function $\mathbb{1}_A(\cdot)$

[17] Within the theory of Fourier series, this effect is known as Gibbs phenomenon.
[18] See Matlab's eps(1) command.

are reached and discontinuities are introduced, we formulate

$$p = -\frac{\log(1 \cdot 10^{-17})}{\pi(\frac{1}{2\beta})^2}$$

to ensure convergence towards zero. On the one hand, the enhanced generator now obeys the assumption of a compact support. On the other hand, all numerical calculations can be executed up to a machine precision level by only considering the p-controlled Gaussian function.

3.3.1 Calculation of $\langle f_X, T_{n\alpha} M_{m\beta} g \rangle$

The first inner product within formula (3.6) is based on the risk neutral density function as well as a shifted and modulated version of the Gabor frame. By the definition of inner products in Hilbert spaces, these Gabor coefficients read as

$$\langle f_X, T_{n\alpha} M_{m\beta}\, g \rangle = \int_a^b f_X(x)\, e^{-p\pi(x-c-n\alpha)^2} \mathbb{1}_A(x-c-n\alpha)\, e^{-2\pi i m\beta x}\, dx. \tag{3.7}$$

When the model is supposed to operate under more complex stochastic processes such as Lévy or stochastic volatility processes, the above integral cannot be solved in closed form. However, due to Parseval's theorem, it is possible to approximate it to a discretionary level of accuracy. Equation (3.7) therefore becomes the sum of two Fourier series coefficients:

$$\begin{aligned} \langle f_X, T_{n\alpha} M_{m\beta}\, g \rangle &= \int_a^b f_X(x)\, e^{-p\pi(x-c-n\alpha)^2} \mathbb{1}_A(x-c-n\alpha)\, e^{-2\pi i m\beta x}\, dx \\ &= 2\,(b-a) \sum_{k\in\mathbb{Z}} f_k\, \overline{g_k}, \end{aligned} \tag{3.8}$$

where the coefficients are given by

$$f_k = \frac{1}{2(b-a)} \int_a^b f_X(x)\, e^{-2\pi i m\beta x}\, e^{ik\pi \frac{x-a}{b-a}}\, dx \tag{3.9}$$

$$g_k = \frac{1}{2(b-a)} \int_a^b e^{-p\pi(x-c-n\alpha)^2} \mathbb{1}_A(x-c-n\alpha)\, e^{ik\pi \frac{x-a}{b-a}}\, dx. \tag{3.10}$$

In the case of (3.9), f_X is defined by the stochastic process (SP) chosen with according parameter vector Θ_{SP}. For a large class of processes, the characteristic function

$$\phi(u; \Theta_{SP}) = \int_{\mathbb{R}} f_X(x; \Theta_{SP})\, e^{iux}\, dx \quad , u \in \mathbb{R}$$

is known and most often given in closed form[19]. When the truncation interval $[a, b]$ in (3.9) and (3.10) is chosen sufficiently large, the expressions can be approximated highly accurately

[19] Within the environment of diffusion models and pure jump models as well as stochastic volatility models and (Lévy driven) Ornstein-Uhlenbeck processes, expressions for the characteristic function are known in closed form. Where a closed form is not given, it is mostly due to the presence of special functions such as, e.g., Bessesl functions.

by characteristic functions which are defined on the whole real line $\mathbb{R}$. Thus, coefficients f_k can be rewritten to

$$f_k = \frac{1}{2(b-a)}\,\phi_{SP}\left(\frac{k\pi}{b-a} - 2\pi m\beta; \Theta_{SP}\right) e^{ix_0\left(\frac{k\pi}{b-a}-2\pi m\beta\right)-i\frac{k\pi}{b-a}a}.$$

In the case of equation (3.10), a similar procedure yields a solution: Rewriting g_k in a way that we incorporate the indicator function into the integration limits and extract the part of the exponential function that does not rely on the integration variable

$$g_k = \frac{1}{2(b-a)} \int_{c+n\alpha-\frac{1}{2\beta}}^{c+n\alpha+\frac{1}{2\beta}} e^{-p\pi(x-c-n\alpha)^2}\, e^{ik\pi\frac{x}{b-a}}\, dx\, e^{-ik\pi\frac{a}{b-a}} \tag{3.11}$$

enables us to use the Fourier transform (see equation (2.1)) of the original generator function

$$\begin{aligned}
\hat{g}(\xi) &= \int_{\mathbb{R}} g(x-n\alpha)\, e^{-2\pi i\xi x}\, dx \\
&= \int_{\mathbb{R}} e^{-p\pi(x-c-n\alpha)^2}\, e^{-2\pi i\xi x}\, dx && (3.12)\\
&= \frac{1}{\sqrt{p}}\, e^{-2\pi i\xi(c+n\alpha)-\frac{\pi\xi^2}{p}} && (3.13)
\end{aligned}$$

to represent the coefficients g_k. To do so we use equation (3.12) with $\xi = -\frac{1}{2}\frac{k}{b-a}$ to calculate the integral in (3.11). Finally, due to the fact that (3.12) is given in closed form by (3.13), we can write g_k as

$$g_k = \frac{1}{2(b-a)}\,\hat{g}\left(-\frac{1}{2}\frac{k}{b-a}\right) e^{-ik\pi\frac{a}{b-a}}. \tag{3.14}$$

In order for equation (3.14) to be accurate, variable p has to be chosen as described earlier to ensure convergence of the Gaussian function on the interval A. Only under this circumstance, the integral over the given finite interval which is defined by the indicator function is accurately calculated by an infinite interval.

3.3.2 Calculation of $\langle v, T_{n\alpha} M_{m\beta}\, \gamma\rangle$

The second part of the pricing formula (3.6) includes the contract's payoff function $v(S_T, T)$ as well as the dual frame $\mathcal{G}(\gamma, \alpha, \beta)$. Once again, Parseval's theorem can be used to compute the inner product based on the individual Fourier series coefficients. To outline the general procedure, we implement an European put option. There are, however, several other European type derivatives for which closed form solutions in terms of Fourier series coefficients exist. Next to the coefficients of a call option, Appendix 3.A lists several coefficients for binary options.

In line with the procedure in the previous subsection, the inner product based on a put option is defined by

$$\begin{aligned}\langle v, T_{n\alpha} M_{m\beta}\, \gamma\rangle &= \int_a^b K\,[1-e^x]^+ \; \frac{\beta\, e^{-p\pi(x-c-n\alpha)^2}\mathbb{1}_A(x-c-n\alpha)}{\sum\limits_{q\in\mathbb{Z}} e^{-2p\pi(x-c-n\alpha-q\alpha)^2}\mathbb{1}_A(x-c-n\alpha-q\alpha)}\, e^{2\pi i m\beta x}\, dx \\ &= 2(b-a)\sum_{k\in\mathbb{Z}} v_k\, \overline{e_k}, \end{aligned} \tag{3.15}$$

and the coefficients are given by

$$v_k = \frac{1}{2(b-a)}\int_a^b K\,[1-e^x]^+ \; e^{2\pi i m\beta x}\, e^{ik\pi\frac{x-a}{b-a}}\, dx \tag{3.16}$$

$$e_k = \frac{1}{2(b-a)}\int_a^b \frac{\beta\, e^{-p\pi(x-c-n\alpha)^2}\mathbb{1}_A(x-c-n\alpha)}{\sum\limits_{q\in\mathbb{Z}} e^{-2p\pi(x-c-n\alpha-q\alpha)^2}\mathbb{1}_A(x-c-n\alpha-q\alpha)}\, e^{ik\pi\frac{x-a}{b-a}}\, dx. \tag{3.17}$$

In terms of v_k, integration yields to a closed form solution:

$$v_k = \frac{1}{2(b-a)}\int_a^0 K\,(1-e^x)\; e^{2\pi i m\beta x}\, e^{ik\pi\frac{x-a}{b-a}}\, dx \tag{3.18}$$

$$= \begin{cases} -\frac{K}{2}\left(\frac{i\left(e^{ik\pi\frac{a}{a-b}} - e^{2ia\beta m\pi}\right)}{\zeta} + \frac{e^{ik\pi\frac{a}{a-b}}\left(e^{a-\frac{ia\zeta}{a-b}}-1\right)}{a-b-i\zeta}\right) & \text{for } \zeta \neq 0 \\ , \frac{K}{2(b-a)}\,(e^a - 1 - a) & \text{for } \zeta = 0 \end{cases}, \tag{3.19}$$

where $\zeta = (k + 2(b-a)\beta m)\,\pi$.

In contrast to these payoff coefficients, the coefficients e_k containing the dual frame are not given in closed form. However, using a fractional Fast Fourier Transform algorithm allows us to compute e_k efficiently. Before the FrFFT can be unleashed, equation (3.17) first has to be discretized and reorganized. Considering the range in which the indicator function of the numerator assumes nonzero values, equation (3.17) becomes

$$e_k = \frac{\beta}{2(b-a)}\int_{c+n\alpha-\frac{1}{2\beta}}^{c+n\alpha+\frac{1}{2\beta}} \frac{e^{-p\pi(x-c-n\alpha)^2}}{\sum_{q\in\mathbb{Z}} e^{-2p\pi(x-c-n\alpha-q\alpha)^2}\mathbb{1}_A(x-c-n\alpha-q\alpha)}\, e^{ik\pi\frac{x-a}{b-a}}\, dx. \tag{3.20}$$

In a next, step we define a new function $h(x)$ which contains the fraction resulting in:

$$e_k = \frac{\beta}{2(b-a)}\int_{c+n\alpha-\frac{1}{2\beta}}^{c+n\alpha+\frac{1}{2\beta}} h(x)\, e^{ik\pi\frac{x}{b-a}}\, dx\, e^{-ik\pi\frac{a}{b-a}}$$

Discretization involves transformation of the continuous variable x into a discrete set $\{x_j\}_{j=0}^{J-1}$. We choose a unitary step size $\Delta x = \frac{1/\beta}{J-1}$ which leads to a grid in the x-domain according to

$x_j = c + n\alpha - \frac{1}{2\beta} + \Delta x \cdot j$ and the following approximation:

$$e_k = \frac{\beta}{2(b-a)} \left\{ \sum_{j=0}^{J-1} h(x_j)\, e^{ik\pi \frac{1/\beta}{(J-1)(b-a)} j} \right\} e^{ik\pi \frac{c+n\alpha-\frac{1}{2\beta}-a}{b-a}} \tag{3.21}$$

Equation (3.21) represents a discretized version of equation (3.20) and inherits a structure that allows us to use the FrFFT to calculate the expression in curly brackets: Based on Bailey and Swarztrauber (1994), a sum

$$G_k(x,\xi) = \sum_{j=0}^{J-1} h(x_j)\, e^{-2\pi i j k \xi}$$

can be efficiently calculated using a generalization of the discrete Fourier transform. The main advantage over a non-fractional FFT algorithm is that the step size of variables k as well as the step size of variable ξ can be chosen independently. This is a major benefit compared to calculating the sum by means of a non-fractional FFT algorithm in which case the step size of the two variables are connected. This fact is known in literature as Nyquist relation and describes the effect that a finer grid in one variable comes with a coarser grid in the other variable. A drawback of the FrFFT algorithm is that instead of only one execution of the FFT algorithm, three FFTs are needed. However, the freedom of choosing k and ξ independently from each other outweighs this disadvantage in terms of overall accuracy as well as in terms of computational time needed to reach a given accuracy.

Following the algorithm by Bailey and Swarztrauber, two $2J$-long sequences have to be defined as

$$y = \begin{bmatrix} \left\{ h(x_j)\, e^{-\pi i j^2 \xi} \right\}_{j=0}^{J-1} \\ \\ \left\{ 0 \right\}_{j=J}^{2J-1} \end{bmatrix} \quad \text{and} \quad z = \begin{bmatrix} \left\{ e^{\pi i j^2 \xi} \right\}_{j=0}^{J-1} \\ \\ \left\{ e^{\pi i (j-2J)^2 \xi} \right\}_{j=J}^{2J-1} \end{bmatrix}.$$

To calculate $G_k(x,\xi)$, three $2J$-point discrete Fourier transforms D_j are implemented using the FFT algorithm

$$G_k(x,\xi) = \left\{ e^{-\pi i k^2 \xi}\, D_k^{-1} \left[D_j(y) \odot D_j(z) \right] \right\}_{k=0}^{J-1}, \tag{3.22}$$

where the symbol $\odot$ describes element-wise multiplication. Since parameter k is restricted to the set of integers $\mathbb{Z}$, we define

$$\xi = -\frac{1}{2\beta(J-1)(b-a)}.$$

Thus, equation (3.21) changes to

$$e_k = \frac{\beta}{2(b-a)}\, G_k\left(x, -\frac{1}{2\beta(J-1)(b-a)} \right) e^{ik\pi \frac{c+n\alpha-\frac{1}{2\beta}-a}{b-a}}.$$

3.4 Error analysis

Similar to Fang and Oosterlee (2008) and Ortiz-Gracia and Oosterlee (2013), we also include an error analysis before we concentrate on numerical implementations. We indicate four major sources of errors due to

(i) truncation of the integral in (3.5):

$$E_1 = \int_{\mathbb{R}\setminus[a,b]} v(x,T) f_X(x)\, dx$$

(ii) usage of Parseval's theorem in (3.8) and (3.15):

$$E_2 = 2(b-a) \sum_{k\in\mathbb{Z}\setminus[-K,K]} f_k\, \overline{g_k} \quad \text{and}$$
$$E_3 = 2(b-a) \sum_{k\in\mathbb{Z}\setminus[-K,K]} v_k\, \overline{e_k}$$

(iii) discretization of equation (3.17):

$$E_4 = \int_a^b \gamma(x-n\alpha) e^{-ik\pi\frac{x-a}{b-a}}\, dx - \sum_{j=0}^{J} \gamma(x_j-n\alpha) e^{-ik\pi\frac{x_j-a}{b-a}}\, w_j \Delta x_j$$

(iv) truncation of series (3.6):

$$E_5 = e^{-r(T-t)} \sum_{m\in\mathbb{Z}\setminus[-M,M]} \sum_{n\in\mathbb{Z}\setminus[-N,N]} \langle f, T_{n\alpha} M_{m\beta} g\rangle\, \overline{\langle v, T_{n\alpha} M_{m\beta}\gamma\rangle}.$$

Term E_1 is discussed in both Fang and Oosterlee (2008) and Ortiz-Gracia and Oosterlee (2013). If the interval $[a,b]$ is chosen appropriately large, E_1 is negligibly small. In Section 3.5.1, we outline three different truncation schemes to handle E_1.

Parseval's theorem in (3.8) and (3.15) is based on Bessel's inequality

$$\sum_{k=-\infty}^{\infty} |f_k|^2 \le \frac{1}{2(b-a)} \int_a^b |f(x)|^2\, dx. \tag{3.23}$$

The fact that (3.23) holds for any square integrable function $f \in L^2([a,b])$ ensures that the coefficients on the left-hand side converge to zero as $k \to \pm\infty$. Since the error terms E_2 and E_3 depend on the products $f_k\,\overline{g_k}$ as well as $v_k\,\overline{e_k}$, only the faster converging coefficients are relevant for the overall error. As soon as the faster converging coefficients within the respective products reach zero, the summation as a whole converges. Thus, convergency depends on g_k and e_k for the most part. In contrast to the coefficients f_k of a highly skewed distribution and the coefficients v_k of the options contract payoff function, the Fourier coefficients of the p-controlled Gaussian function g_k are fast decaying. The same is true for

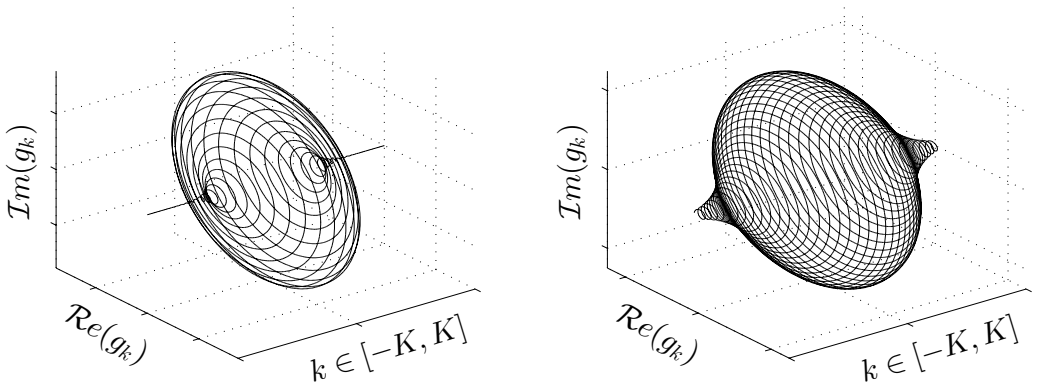

Figure 3.1: Fourier coefficients of the Gaussian generator g with different control variables p_1 and p_2, where $p_1 > p_2$, on a fixed interval $[a, b]$.

the coefficients e_k, since the dual function conserves the smoothness of the original generator (Gröchenig, 2001, p. 118).

To bind the speed of convergency, we take a look at the Fourier coefficients of the Gaussian function g_k. Figure 3.1 displays two exemplary curves of the coefficients g_k in the complex plane $\mathbb{C}$.[20] The speed of convergency only depends on the length of interval $[a, b]$ and the control variable p. A change in the shift parameter changes the values of the coefficients that are assumed on the above function, but not the convergency itself. The mean squared error δ_K is an appropriate tool to analyze this convergency. We therefore define

$$\begin{aligned}\delta_K &= \frac{1}{b-a}\|g(x) - S_K(x)\|^2 \\ &= \frac{1}{b-a}\|g(x)\|^2 - \|S_K(x)\|^2,\end{aligned}$$

where $S_K(x)$ is the partial sum $S_K(x) = \sum_{k=-K}^{K} g_k \; e^{-ik\pi\frac{c-a}{b-a}}$. Using the p-contolled version of g allows us to calculate the norm $\|g(x)\|^2$ over the entire real line. Since, in addition, corresponding coefficients $g(k)$ and $g(-k)$ in the partial sum are complex conjugates, the mean squared error of approximating function g by means of a series of coefficients g_k is defined by

$$\begin{aligned}\delta_K &= \frac{1}{b-a}\int_{\mathbb{R}} g(x)^2 dx - g_0^2 - 2\sum_{k=1}^{K}|g_k|^2 \\ &= \frac{1}{\sqrt{2p}(b-a)} - \frac{1}{4p^2(b-a)^2} - \frac{1}{2}\frac{1}{p^2(b-a)^2}\sum_{k=1}^{K} e^{-k^2\frac{1}{p^2(b-a)^2}}\end{aligned}$$

Thus, the mean squared error converges exponentially due to $\delta_K = \mathcal{O}(e^{-k^2})$ as $k \to \infty$.

[20] It has to be mentioned that within the figure, variable k not only assumes integer values but is also able to assume any real number. Therefore, the plots are rather shells in which the coefficients themselves only occupy a finite number of points depending on the shape of function g.

Discretization error E_4 depends on the numerical integration scheme chosen. Due to the complex exponential in E_4, the integrand $I(\psi) = \gamma(\psi - n\alpha)e^{-ik\pi\frac{x-a}{b-a}}$ is oscillatory which makes a trapezoidal rule a good choice. In this case, E_4 is defined by

$$E_4 = -\frac{1}{12J^2\beta^3}\ I''(\psi), \quad \psi \in \left[c + n\alpha - \frac{1}{2\beta}, c + n\alpha + \frac{1}{2\beta}\right]$$

and depends heavily on the curvature of $I(\psi)$ which can be controlled by variable α. How to choose α in accordance with a minimal error term E_4 is covered by Section 3.5.1.

Error term E_5 originates from truncating variables M and N to finite values. Akin to the argumentation in case (ii), the magnitude of the error depends on the convergency behavior of the coefficients that are defined by the inner products. In this case, however, convergency depends on two dimensions, M and N. Thus, we can split E_5 into $E_5 = E_5^N + E_5^M$ and examine both parts individually. Error E_5^N interacts with E_1 since the N-dimension is directly connected to the interval $[a, b]$. Thus, by setting $[a, b]$ sufficiently large, E_5^N converges to zero and E_5 is dominated by E_5^M. The M-dimension, however, is driven by the inner products themselves as well as by their components. Once again, only the faster converging coefficients are crucial for the magnitude of E_5^M. Since many payoff functions v show slowly converging coefficients, we pay attention to the coefficients $c_{mn} = \langle f, T_{n\alpha}M_{m\beta}g\rangle$ instead. Equation (3.9) is responsible for the M-dimension in the calculation of c_{mn}. For the purpose in this section, we rename it from f_k to $f(m; k)$ and state that, as long as the density function is well behaved $f \in C^\infty([a, b])$, the function $f(m; k)$ shows an exponential convergency in k for any given integer value m. Equation (3.30) in Section (3.5.2) provides a rule of thumb how to determine M subject to a given target level of pricing accuracy.

3.5 Numerical implementation

This section gives an insight into the convergency behavior of the model. However, before the model from Section 3.3 can be used to price options, we have to prepare the model parameters first. To do so we address error terms E_1, E_5 and E_4 individually. In contrast to E_2 and E_3, these terms depend on the input parameters of the model itself and, therefore, have to be determined with some care. Subsequently we conduct several numerical tests. The test scenarios are chosen from standard Brownian motions to more complex jump processes and stochastic volatility models. The belonging input parameters are either defined with the task of analyzing specific characteristics[21] in mind or are derived from parameter sets that are commonly used in related literature.

[21] E.g. model behavior in the presence of skewness and kurtosis or model behavior given different times to maturity.

3.5.1 Specification of the model

Up to this point, the pricing formula is given by equation (3.6). To be able to implement the model in a programing environment such as Matlab, these sums have to be truncated at a given level

$$v_0 = e^{-r(T-t)} \sum_{m=-M}^{M} \sum_{n=-N}^{N} \langle f, T_{n\alpha} M_{m\beta} g \rangle \, \overline{\langle v, T_{n\alpha} M_{m\beta} \gamma \rangle}. \tag{3.24}$$

This implies that the frame coefficients can only be calculated with a finite sequence of numbers and, thus, a finite precision. As indicated by the error analysis in the former section, truncation introduces some noise in the system. Therefore, some care has to be exercised to minimize this source of error. We will address the error sources E_1, E_5 and E_4 consecutively.

Specification of error term E_1

Due to the fact that determination of the maximal number of shifts N depends on the lower bound a, we start with the truncation of the integral in (3.5) and the first error term. E_1 approaches zero if the values of the integrand evaluated at the integration limits is below machine precision

$$\left| v\left(\begin{bmatrix} a \\ b \end{bmatrix}, T \right) f_X \left(\begin{bmatrix} a \\ b \end{bmatrix} \right) \right| < \epsilon.$$

To define appropriate values for a and b, we concentrate on the decay of the density function. To be more precise, we are interested in the question at which point the density function assumes values below machine precision. One option would be to fall back on inverse distribution functions. However, since in most cases these inverse distributions are not readily available, other schemes need to be defined. Fang and Oosterlee (2008) determine the lower and upper boundary values in terms of cumulants of the stochastic process

$$[a, b] = \left[c_1 - b_1 \sqrt{c_2 + \sqrt{c_4}}, \; c_1 + b_1 \sqrt{c_2 + \sqrt{c_4}} \right] \quad \text{with } b_1 = 10, \tag{3.25}$$

where $b_1 \in \mathbb{R}$ determines the spread of the interval and the cumulants $\{c_i\}_{i \in \mathbb{N}}$ are calculated based on the cumulant generating function which is readily available by means of the logarithm of the characteristic function of the stochastic process (see Section 2.1.2). Thus, using the cumulant generating function, the sequence $\{c_n\}_{n=1}^{4}$ is known as soon as input parameters Θ_{SP} are chosen.

Since the above approach does not consider skewness other than its impact on kurtosis, we enhance (3.25) in two ways: First, we add cumulant c_3 and, second, we allow b_i to be determined individually for a and b. As a result, the following three truncation schemes

influence the convergency of the model and therefore the error term E_1:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c_1 - b_1^a \sqrt{c_2 + \sqrt{c_4}} \\ c_1 + b_1^b \sqrt{c_2 + \sqrt{c_4}} \end{bmatrix} \tag{3.26}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c_1 + b_1^a c_3 - b_2^a \sqrt{c_2 + \sqrt{c_4}} \\ c_1 + b_1^b c_3 + b_2^b \sqrt{c_2 + \sqrt{c_4}} \end{bmatrix} \tag{3.27}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b_1^a + b_2^a \Phi_N^{-1}\left(q_a, c_1, \sqrt{c_2}\right) + b_3^a c_3 + b_4^a c_4^{b_5^a} + b_6^a c_3 c_2 + b_7^a c_3 c_4 \\ b_1^b + b_2^b \Phi_N^{-1}\left(q_b, c_1, \sqrt{c_2}\right) + b_3^b c_3 + b_4^b c_4^{b_5^b} + b_6^b c_3 c_2 + b_7^b c_3 c_4 \end{bmatrix}. \tag{3.28}$$

The methods from (3.25) and (3.26) only differ in the fact that in (3.26) the coefficient b_1 is calculated individually for a and b. In (3.27), skewness c_3 shifts the interval of interest depending on the magnitude of asymmetry of the stochastic process' marginal distribution. Truncation method (3.28) is based on an inverse Normal distribution at a given level $q_a = \epsilon$ and $q_b = 1 - \epsilon$, as well as on a number of interaction terms. Including the inverse Normal distribution is due to two arguments: first, it offers an initial guess for the location of the true values, and, second, it is readily available and fast to evaluate. Adjustments to the initial guess are made by adding additional interaction terms. This method, however, should only be considered in a non-symmetric setup since it heavily depends on cumulant c_3. In case of a symmetric probability distribution, the method in (3.28) only adds a markup term to the inverse Normal distribution.

To pre-calculate the coefficients b_i, we compose a dataset consisting of the dependent variables a and b, which are defined as inverse distributions F_X^{-1} of various density functions evaluated again at q_a and q_b:

$$\begin{aligned} a &= F_X^{-1}(q_a) = F_X^{-1}(\epsilon) \\ b &= F_X^{-1}(q_b) = F_X^{-1}(1-\epsilon) \end{aligned}$$

We use data points stemming from NIG distributions, CGMY distributions and VG distributions to provide dependent variables. In case of NIG, we use the Normal Inverse Gaussian distribution toolbox[22] which includes a function to calculate inverse distributions. In case of CGMY and VG, simulations are used to create distribution functions and their respective inverse distributions. As independent variables, the first four cumulants $X = [c_1\ c_2\ c_3\ c_4]$ are considered. To come up with suitable coefficients, we use (nonlinear) regressions to estimate $b_i^\bullet$ according to the schemes in (3.26)-(3.28) and store the parameters for later usage. Results are shown in Table 3.1.

The table shows that equations (3.26) and (3.27) coincide given a symmetric distribution. Overall, it should be mentioned that these parameters are calculated based on a time horizon of one year. They deliver quite robust boundary values even if the time to maturity differs. Since the parameters are pre-stored, they do not create additional computational load during the price calculation. Thus, similar to Table 3.1, additional coefficients can be calculated and pre-stored with a special focus on extremely short or very long time horizons.

[22] Downloadable at Matlab Central File Exchange http://de.mathworks.com/matlabcentral/fileexchange.

	(3.26)		(3.27)		(3.28)	
	a	b	a	b	a	b
$b_1^\bullet$	8.3169	6.6669	6.7775	5.3835	3.5790	-2.1726
$b_2^\bullet$			8.3169	6.6669	-2.0528	1.4768
$b_3^\bullet$					3.8471	2.8479
$b_4^\bullet$					-4.5670	3.3609
$b_5^\bullet$					0.5639	0.5895
$b_6^\bullet$					7.3642	6.7739
$b_7^\bullet$					-0.0667	-0.0325
$\overline{R^2}$	0.36	0.30	0.72	0.69	0.97	0.97

Table 3.1: Determination of the lower and upper bound a and b by means of (nonlinear) regression models (3.26)-(3.28).

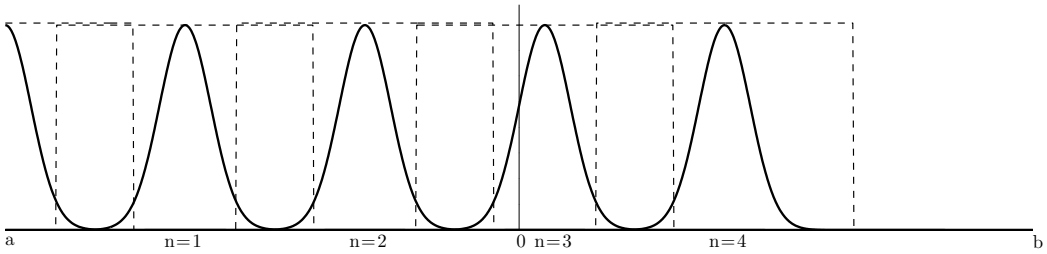

Figure 3.2: Shifted generator function.

Specification of error term E_5

Error term E_5 deals with truncation of the sums in (3.24). In this case, the modulation parameter M stays an exogenous variable. The shift variable N, however, can be internalized by bearing in mind that the truncation area $[a, b]$ has contract specific relevant subintervals. When pricing a put option, only the subinterval $\mathcal{I}_p = [a, 0]$ is of interest.[23] Therefore, the maximal number of shifts is defined by a given number $\hat{N}$ that relocates the generator function right to the point where the lower limit of the shifted frame is within the interval $[a, 0]$. That is, for $n = \hat{N}$ it is no longer true that $\text{supp}\{g(x - n\alpha)\} \subset \mathcal{I}_p$; instead, both sets form an intersection $\text{supp}\{g(x - n\alpha)\} \cap \mathcal{I}_p$. Figure 3.2 exemplarily plots a scenario where the shifted versions of g moves to the right by a factor $n\alpha = n \cdot 0.7$. As indicated in the figure, the maximal amount of shifts needed in this particular case is $n = 3 = \hat{N}$. Given that number of shifts, a part of the final frame lies within $\mathcal{I}_p$. Therefore, this area has to be incorporated. That is, the interval of interest given a put option is actually larger than the contract specific domain $\mathcal{I}_p = [a, 0]$ due to the final shift $\hat{N}$ which can be defined by

$$\hat{N} = \left\lceil \frac{2|a|\beta + 1}{2\alpha\beta} \right\rceil - 1 + \frac{1}{\beta}.$$

[23] Similar arguments can be formulated for e.g. a call option.

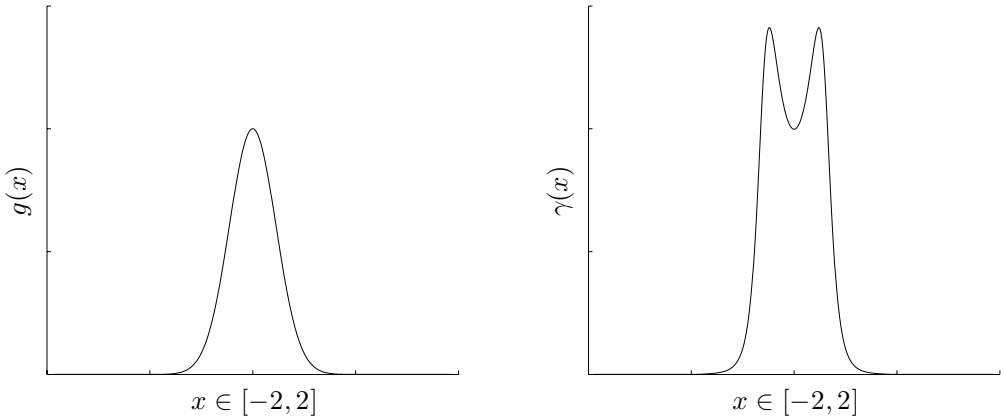

Figure 3.3: Gaussian generator and its dual function. Parameters are $\alpha = 0.6$ and $\beta = 0.25$.

Given that, formula (3.24) changes to

$$v_0 = e^{-r(T-t)} \sum_{m=-M}^{M} \sum_{n=0}^{\hat{N}} \langle f, T_{n\alpha} M_{m\beta} g \rangle \, \overline{\langle v, T_{n\alpha} M_{m\beta} \gamma \rangle}. \tag{3.29}$$

The inner summation runs from $n = 0, ..., \hat{N}$. The reason why it starts at $n = 0$ (in equation (3.24), the starting value is defined as $n = -N$) is that we choose the shift parameter c within the generator function to be equal to the lower bound of the interval, $c = a$. Therefore, the generator starts at the lower bound and moves to the right.

Now that the number of terms in the sums of equation (3.29) has been taken care of, we have to focus on variables α and β. The support of the generator function is directly linked to variable β via $\text{supp}(g) = [l_0, l_0 + L]$, with $l_0 \in \mathbb{R}$ being some constant and $L = \frac{1}{\beta}$ (see Section 3.2). For the numerical implementation, we choose $l_0 = -\frac{1}{2\beta}$. Thus, variable β not only influences the frequency of the modulation operator, but also the length of the Gabor frame and its dual frame. Since the generator and its dual can also be interpreted as magnifiers or spotlights of a function segment, β defines the size of this spotlight. If not stated otherwise, we use the length of the relevant subinterval of the contract that is to be priced, yielding $\beta = \frac{1}{|a|}$ in the case of a put option.

Specification of error term E_4

Parameter α primarily controls the magnitude of the shifts on the 'time'-axis. It also influences the shape of the dual frame γ defined in (3.4) and therefore error term E_4. In contrast to orthogonal basis, the shape of the dual frame function differs considerably from the generator function it is based on. Due to the denominator in (3.4) and using a Gaussian as generator function, α defines whether the dual function γ shows a single hump or a double dip shoulder as shown in Figure 3.3.

Depending on α, the double dip that is present in the dual function can be more distinct or may even vanish. Multiple maxima, as on the right-hand side of Figure 3.3, increase the curvature of the function which, in turn, yields to a higher error term E_4. Given multiple

maxima, E_4 can only be lowered at the expense of a higher computational effort. To prevent this effect, we choose we choose α to ensure a smooth dual frame with a plateau in the center. An analytic way to ensure this claim would be to restrict $\left[\frac{\partial^2}{\partial x^2}\gamma(x)\right]_{x=l_0+L/2} = 0$ and solve for α. Since this expression becomes quite cumbersome and is not given in closed form, we use a more heuristic approach. As a first step, we define a function

$$\nu(\alpha) = \frac{1}{\sum\limits_q e^{-2p\pi q^2\alpha^2}} - \frac{e^{-p\pi(x_1-q\alpha)^2}}{\sum\limits_q e^{-2p\pi(x_1-q\alpha)^2}} \quad \text{with } x_1 = x_0 + \tau,$$

that, given τ being small, assumes positive values in case the center point of γ is a maximum and negative values in case of a local minimum. Given a plateau, the function shows a result of zero. Thus, as a second step, the optimal α is given by

$$\alpha = \sup\left\{\alpha \in \mathbb{R}^+ : \nu(\alpha) \geq 0\right\}$$

and can be determined by a root finding algorithm with a very low number of iteration steps.

In the remaining part of this section, we test the level of precision of equation (3.29). As reference models, we implement the Cos method by Fang and Oosterlee (2008) as well as the FFT method by Carr and Madan (1999).

3.5.2 Numerical tests

To be able to not only test symmetric marginal distributions, we use a Geometric Brownian Motion (GBM) and, in addition, a Normal Inverse Gaussian (NIG) process as the underlying source of uncertainty. The four-parameter NIG marginal distribution was introduced by Barndorff-Nielsen (1977) and is capable of producing a purely discontinuous Lévy process with heavy-tailed as well as skewed marginal probability distribution. The characteristic function of a GBM as well as the characteristic function of a NIG process is given by

$$\begin{aligned}\phi_{GBM}(u) &= e^{i\mu ut - \frac{1}{2}\sigma^2u^2t} \\ \phi_{NIG}(u) &= e^{i\mu ut + \delta\left(\sqrt{\alpha_{NIG}^2-\beta_{NIG}^2} - \sqrt{\alpha_{NIG}^2-(\beta_{NIG}+iu)^2}\right)}.\end{aligned}$$

A wide range of test scenarios is conducted in this NIG environment. To also include further stochastic processes we use reference values from Fang and Oosterlee (2008). Therefore we are able to additionally test the numerical method assuming a Variance Gamma (VG) and a CGMY Lévy process introduced by Madan and Seneta (1990) and Carr et al. (2002), respectively. Characteristic function are given by:

$$\begin{aligned}\phi_{VG}(u) &= e^{i\mu ut}\left[1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2\right]^{-\frac{t}{\nu}} \\ \phi_{CGMY}(u) &= e^{i\mu ut}e^{tC\Gamma(-Y)\left[(M-iu)^Y - M^Y + (G+iu)^Y - G^Y\right]}\end{aligned}$$

Next to pure jump Lévy processes, we include Heston's stochastic volatility model published in Heston (1993). However, due to a restriction of the complex logarithm to its principal branch, the characteristic function of the Heston model is defined in a slightly different way than it has been stated in the original paper. We implement the version that is also used in Fang and Oosterlee (2008) to preserve comparability in notation.

$$\phi_{Heston}(u) = e^{i\mu u t} e^{\frac{u_0}{\eta^2}\left(\frac{1-e^{-D\cdot t}}{1-Ge^{-D\cdot t}}\right)(\lambda - i\rho\eta u - D)} e^{\frac{\lambda\bar{u}}{\eta^2}\left[t(\lambda - i\rho\eta u - D) - 2\log\left(\frac{1-Ge^{-D\cdot t}}{1-G}\right)\right]},$$

with $D = \sqrt{(\lambda - i\rho\eta u)^2 + (u^2 + iu)\eta^2}$ and $G = \frac{\lambda - i\rho\eta u - D}{\lambda - i\rho\eta u + D}$.

Numerical tests assuming a geometric Brownian motion

The first series of tests is done within a GBM environment. In this case, the precision can be tested against the closed form solution by Black and Scholes (1973). In a first step, we define the parameter set Θ^{GBM} as

$$\Theta^{\text{GBM}} = [\mu, \sigma] = \left[r - q - \frac{1}{2}\sigma^2, \sigma\right].$$

In general, a GBM has two parameters. However, parameter μ is pre-defined due to the drift correction motivated by the equivalent martingale measure. The remaining test parameters[24] are given by a current price of the underlying $S = 10$, strike price $K = 10$, a risk free interest rate $r = 0.02$, a standard deviation of the GBM process $\sigma = 0.35$ and time to maturity between 10 days and one year[25]. Within this test setup, we do not yet use equations (3.26)-(3.28) in combination with the coefficients in Table 3.1 but rather test the overall robustness by assuming different truncation intervals and varying maturities.

Figure 3.4 is based on pricing errors defined by the difference between the price calculated by means of the closed form solution and the price approximation of the numerical methods. While the Gabor method is considered in the upper part of the figure, Cos method is examined in the lower part. Besides a varying time to maturity the spread variable b_1 from equation (3.25) assumes integer values within $8 \leq b_1 \leq 30$. From the figure two observations are important: First, in case of the Gabor method, besides the outliers that are marked by a plus-symbol in the figure, most of the deviations from the analytical price are on a level of 10^{-15}. These outliers, however, can be avoided by using the afore-mentioned truncation schemes. Second, compared to the Cos method, the Gabor method is far less sensitiv to a particular choice of b_1. The Cos method looses its accuracy quite fast given a value for b_1 that is chosen either too high or too low. This effect is more severe the larger time to maturity is assumed to be.

Numerical tests assuming jump processes

While a single spread variable b_1 performs quite well using a GBM as the underlying source of randomness, the effect of including cumulants c_3 and c_4 in combination with the drift term

[24] Test parameters within this section are based on hypothetical parameter sets that are either commonly used in literature or show specific features of the models.

[25] We define a year to consist of 250 trading days.

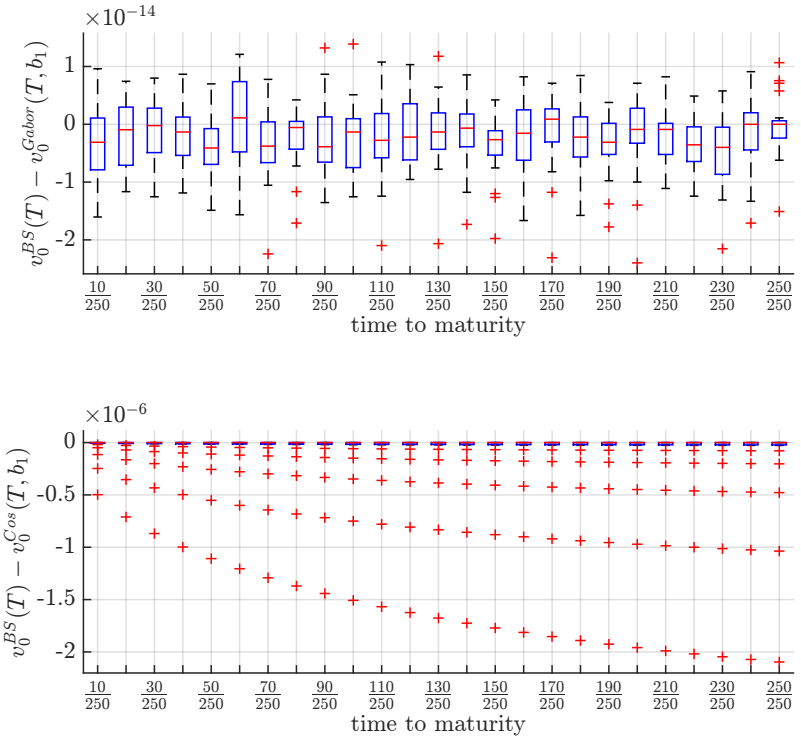

Figure 3.4: Box plot of pricing errors of the Gabor method (upper part) and the Cos method (lower part). Variable b_1 in (3.25) assumes values $b_1 = \{8, ..., 30\}$ in case of both methods. In addition, $S = K = 10$, $r - q = 0.02$ and $\sigma = 0.35$. Frequency variable is set to M = 40 for the Gabor method and an equivalent number of terms is used for the Cos method.

and volatility to the truncation is more distinct given fat tailed and asymmetric distributions. Therefore, the second series of tests employs fat-tailed NIG processes. We use different parameter sets with excess kurtosis and non-zero skewness

$$\begin{aligned}\Theta^{\text{NIG}} &= [\mu_{NIG}, \alpha_{NIG}, \beta_{NIG}, \delta] \\ &= [\mu_{NIG}, 5, \beta_{NIG}, 1]\,.\end{aligned}$$

Hereby, variable $\mu_{NIG} = r - q - \delta\left(\sqrt{\alpha_{NIG}^2 - \beta_{NIG}^2} - \sqrt{\alpha_{NIG}^2 - (\beta_{NIG}+1)^2}\right)$ is chosen in accordance to a drift correction term which assures the process to be a martingale (see Chapter 2.1.2). We assume $\alpha_{NIG} = 5$ and $\delta = 1$ which results in fat tails ($\alpha_{NIG} \to \infty$ generates a GBM in the limit). Skewness parameter β_{NIG} is chosen out of three different values $\beta_{NIG} = \{-2, 0, 2\}$. Picking β_{NIG} out of these three different values results in a negatively skewed distribution if $\beta_{NIG} = -2$, in a symmetric distribution if $\beta_{NIG} = 0$ and a positively skewed distribution if $\beta_{NIG} = 2$.

Before examining the parameter sets individually, we first compare the convergency behavior with respect to the frequency variable M of all three NIG setups combined. Figure 3.5 shows that the speed of convergency is dependent on the truncation scheme used as well as on the characteristic of the stochastic process. Reference values are calculated by means of Carr

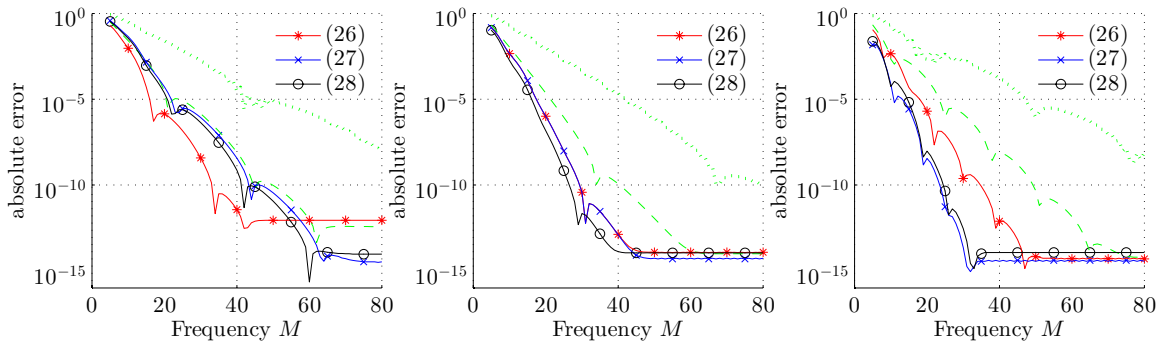

Figure 3.5: Convergency of the Gabor method with respect to three different truncation schemes. On the left, a negatively skewed distribution ($\beta_{NIG} = -2$) is implemented, in the middle part a symmetric ($\beta_{NIG} = 0$) and a positively skewed is used on the right ($\beta_{NIG} = 2$). Cos method convergency is displayed as a comparison in dotted lines and also in dashed lines with twice the number of terms $M_{Cos} = 2\,M$.

and Madan's FFT method and the Cos method[26] is included as a comparison. By the term 'absolute error' we refer to the absolut deviation of prices calculated with the help of the Gabor method compared to prices calculated by the reference models.

When pricing under the assumption of a fat-tailed symmetric distribution (middle part of the figure), the three truncation schemes perform similarly. The approximation accuracy is higher for left skewed distributions (left part of the figure) given the cumulant c_3 is considered. This higher accuracy, however, requires more terms in the approximation. In the presence of a stochastic process with a right skewed marginal distribution (right part of the figure), the truncation schemes (3.27) and (3.28) are again advantageous. In general, the skewness enhanced schemes (3.27) and (3.28) show robust truncation goodness even when departing from a one-year time horizon. Even when pricing under the assumption of a fairly skewed marginal distribution of the stochastic process, the Gabor method only needs about 40-50 terms to come up with an accurate price approximation.

Taking a closer look at the convergency under the aspect of the computational time needed to reach a given level of precision, Figure 3.6 displays the model's behavior assuming a time to maturity of half a year and a positively skew. Within the figure, for each of the truncation schemes (3.26) - (3.28) a plot of CPU time[27] in milliseconds against frequency parameter M is included in the upper part. Independent from the truncation scheme used, CPU time is nearly linear in parameter M.

The lower part of the figure is concerned with the question how absolute error values, i.e. absolute deviations from reference values, are connected to CPU time. Hereby, the figure implies that truncation schemes (3.27) and (3.28) converge faster compared to (3.26).

[26] The inclusion of the Cos method in Figure 3.5 is supposed to serve as a classification rather than a comparison. This is due to the fact that the Cos method approximates the whole function at once, while the Gabor method works on a shift-wise basis. The number of time shifts, however, is not incorporated here.

[27] Calculations are implemented within the programming environment Matlab 2014b using a machine with a 2.3 GHz Quad-Core CPU and 16 GB RAM.

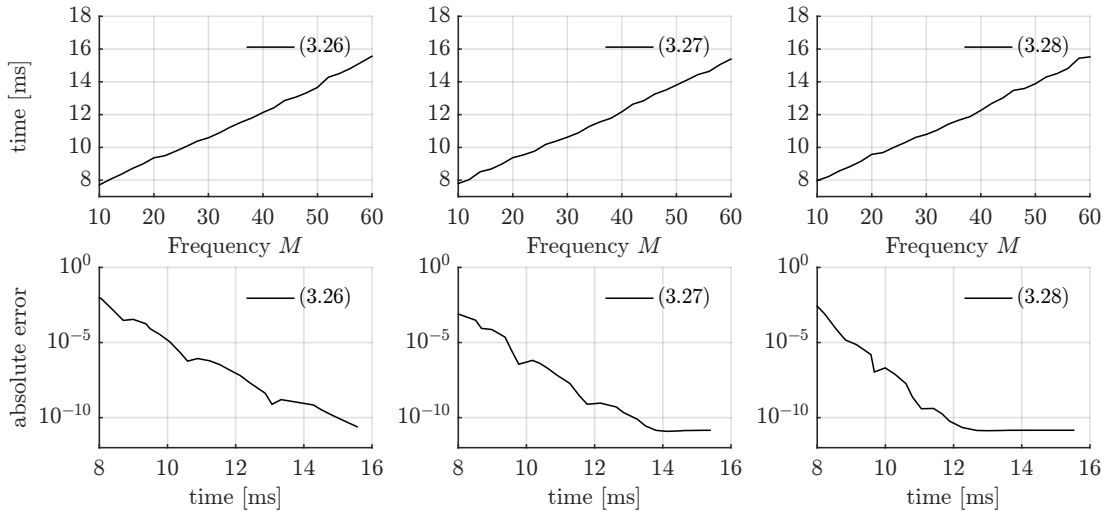

Figure 3.6: CPU time based on different truncation schemes. Time to maturity is half a year and skewness is positive with $\beta_{NIG} = 2$. Dashed red lines indicate an error level of 10^{-3}, 10^{-5} and 10^{-10} respectively.

This observation is even more pronounced for truncation scheme (3.28) than for truncation scheme (3.27). Similar findings can be described when analyzing left skewed or symmetric distributions.

In the next part of this section, we concentrate on shortening the time to maturity. Besides the choice of short-dated contracts, we take a closer look at the parameter environment described in the middle as well as on the right side of Figure 3.5. That is, we use the general setup Θ^{NIG} and define two individual setups Θ_1^{NIG} and Θ_2^{NIG} that differ in the value parameter β_{NIG} is attached to. To begin with a symmetric setup ($\beta_{NIG} = 0$), we specify the parameter vector

$$\begin{aligned}\Theta_1^{\text{NIG}} &= \Big[\mu, \alpha_{NIG}, \beta_{NIG}, \delta\Big] \\ &= \Big[\mu_{NIG}, 5, 0, 1\Big].\end{aligned}$$

Figure 3.7 plots the resulting risk neutral probability density function given the parameter set described by Θ_1^{NIG} and assuming a time to maturity of one year. Both, the NIG distribution in black solid lines as well as the Normal distribution in red dashed lines are symmetric. The NIG distribution, however, shows a much richer tail behavior as shown in the log-density plot on the right-hand side. Based on Θ_1^{NIG}, the NIG distribution assumes a level of excess kurtosis of $\kappa = 3.6$.

Given the above parameter set Θ_1^{NIG}, Table 3.2 summarizes deviations from reference values. These reference values are again calculated by the FFT method down to a time to maturity of 10 days. Since this particular method looses precision given very short time horizons, the Cos method with $M_{Cos} = 2^{17}$ is consulted for maturities of only 5 days. We implement at-the-money put options with $S = K = 10$, a risk free interest rate of $r = 0.02$ and costs of carry given by $q = 0$. The Gabor method again shows small errors even with regard to low

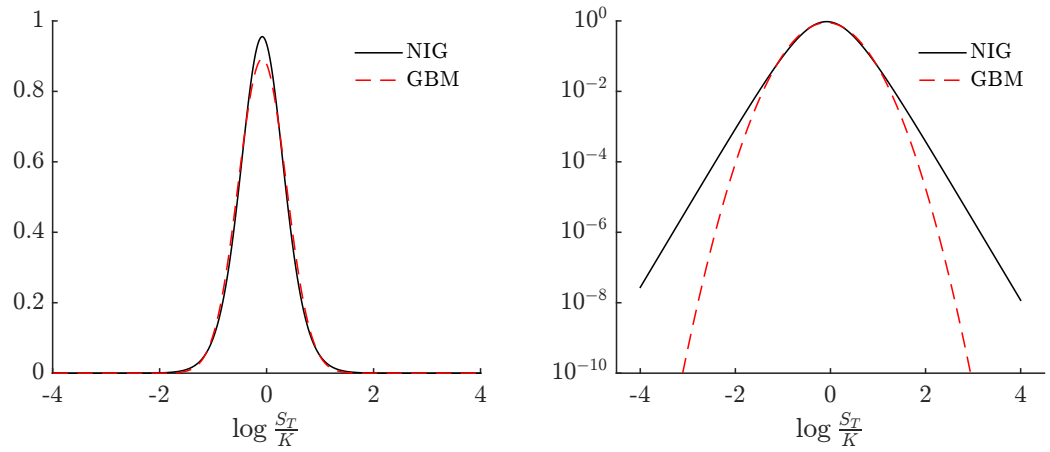

Figure 3.7: Density (left) and log-density (right) plot given the above parameter set and a time horizon of $T = 1$. As comparison, a Normal distribution with matching standard deviation is plotted in dashed lines.

		Frequency Parameter M						
T		10	50	100	200	400	800	1000
1 year	Gabor	-4.5e-3	4.9e-15	4.9e-15	4.9e-15	4.9e-15	4.9e-15	4.9e-15
	Cos	-0.22	-8.9e-07	1.12e-12	9.8e-15	9.8e-15	9.8e-15	9.8e-15
1/2 year	Gabor	-2.6e-2	4.1e-12	-8.9e-16	-8.9e-16	-8.9e-16	-8.9e-16	-8.9e-16
	Cos	-0.54	-8.7e-4	-6.2e-7	1.4e-12	1.1e-15	1.1e-15	1.1e-15
1/4 year	Gabor	-7.8e-2	-5.2e-07	1.1e-12	9.1e-15	9.1e-15	9.1e-15	9.1e-15
	Cos	-0.58	-4.2e-3	-3.7e-5	-5.1e-9	-1.1e-15	-1.1e-15	-1.1e-15
10 days	Gabor	-0.3	-5.1e-3	-2.8e-4	-2.4e-6	-3.4e-10	5.44e-15	5.44e-15
	Cos	-0.8	-5.4e-2	-9.7e-3	-7.7e-4	-1.3e-5	-9.7e-9	-2.6e-10
5 days	Gabor	-0.3	-1.5e-2	-2.2e-3	-1.3e-4	-1.0e-6	-1.4e-10	-8.9e-13
	Cos	-0.9	-9.8e-2	-2.8e-2	-5.2e-3	-4.6e-4	9.5e-6	-1.6e-6

Table 3.2: Pricing errors for at-the-money put options. Parameters are chosen as $S = K = 10$, risk free interest rate $r = 0.02$ and $q = 0$. Truncation in the case of the Gabor method is done by scheme (3.27). In the case of the Cos method, $M = 2^{17}$ cosine series terms are used.

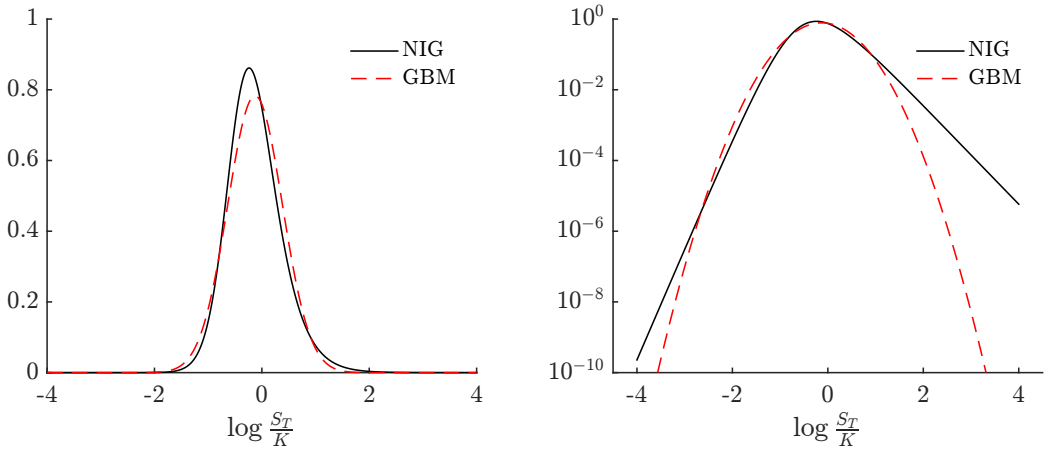

Figure 3.8: Density (left) and log-density (right) plot given the above parameter set and a time horizon of $T = 1$. For comparison purposes, a Normal distribution with matching standard deviation is plotted in dashed lines.

values of M. In particular, when confronted with a combination of a short time to maturity and a highly non-normal marginal distribution, this feature becomes important.

In addition to heavy-tailed symmetric distributions, we now test against skewed distributions. Thus, we define a second parameter setup

$$\begin{aligned}\Theta_2^{\text{NIG}} &= \Big[\mu, \alpha_{NIG}, \beta_{NIG}, \delta\Big] \\ &= \Big[\mu_{NIG}, 5, 2, 1\Big].\end{aligned}$$

Again, as a comparison, we plot the behavior of the resulting density functions on the basis of a one-year time horizon in Figure 3.8. In contrast to Θ_1^{NIG}, there is skewness $s = 0.56$ and kurtosis $\kappa = 4.07$ present based on Θ_2^{NIG}. The effect of skewness in comparison to a symmetric Normal distribution is especially visible in a log-density plot that is included in the right part of the figure.

The convergency of the Gabor method and the Cos method is shown in Figure 3.9. Both methods show fast convergency to the reference value. Especially with shrinking maturities, the Gabor method needs less terms M for the task. Overall, The Gabor method seldom needs more than $M = 400$ terms to achieve a useful price approximation, and, in many cases, a high level of accuracy is already reached with $M \leq 100$.

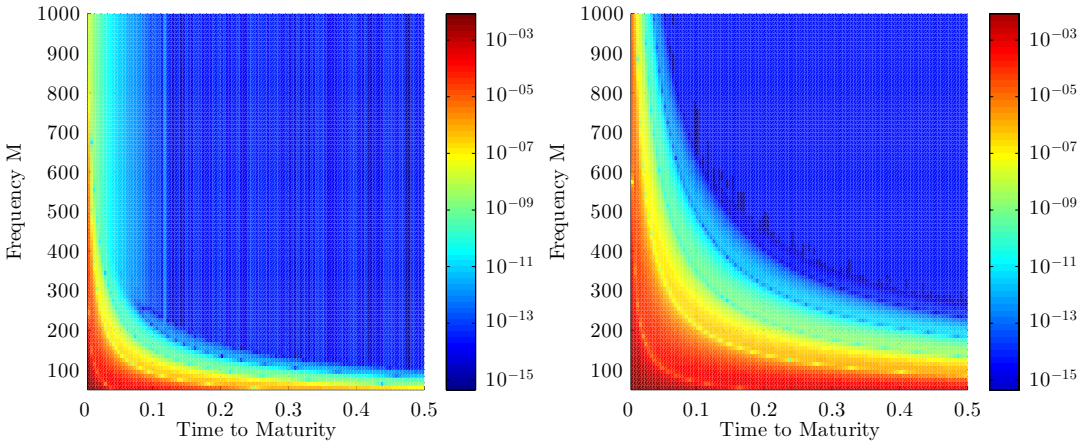

Figure 3.9: Pricing errors of the Gabor method (left) and the Cos method (right) as a function of M and time to maturity $T = \{1/250, 5/250, ..., 125/250\}$. Parameters as in Θ_2^{NIG} mentioned above.

To complete the numerical test scenarios we turn to VG, CGMY and Heston processes. Parameter sets are given by

$$
\begin{aligned}
VG: S = 100, \quad K = 90, \quad r = 0.1, \quad q = 0 \\
\mu_{VG} &= r - q + \frac{1}{\nu}\log\left(1 - \theta\nu - \frac{\sigma\nu}{2}\right) \\
\Theta^{\text{VG}} &= [\mu_{VG}, \sigma, \nu, \theta] \\
\Theta^{\text{VG}} &= [\mu_{VG}, 0.12, 0.2, -0.14] \\
CGMY: S = 100, \quad K = 100, \quad r = 0.1, \quad q = 0 \\
\mu_{CGMY} &= r - q - C\Gamma(-Y)\left[(M-1)^Y - M^Y + (G+1)^Y - G^Y\right] \\
\Theta^{\text{CGMY}} &= [\mu_{CGMY}, C, G, M, Y] \\
\Theta^{\text{CGMY}} &= [\mu_{CGMY}, 1, 5, 5, Y_j] \\
Heston: S = 100, \quad K = 100, \quad r = 0, \quad q = 0 \\
\mu_{Heston} &= r - q \\
\Theta^{\text{Heston}} &= [\mu_{Heston}, \lambda, \eta, \bar{u}, u_0, \rho] \\
\Theta^{\text{Heston}} &= [\mu_{Heston}, 1.5768, 0.5751, 0.0398, 0.0175, -0.5711]
\end{aligned}
$$

and are chosen identical to the parameters in Fang and Oosterlee (2008). Besides the parameter setup, reference values are also obtained from this source.

Table 3.3 contains an overview of the level of accuracy depending on variable M. In case of VG and Heston, the time horizon is defined within the table. As for the CGMY process, time horizon is fixed at $T = 1$. Here, instead of the time horizon, factor Y assumes three different states. In the upper part of Table 3.3, pricing errors of the Gabor method are summarized up to a level of M where the price converges. In the lower part of the table, the maximal precision which is reachable by means of the Cos method is also included. Even though parameter constellations are chosen to produce extremely peaked density functions,

	VG		CGMY			Heston	
M	$T = 1$	$T = 0.1$	$Y_1 = 0.5$	$Y_2 = 1.5$	$Y_3 = 1.98$	$T = 1$	$T = 10$
10	3.2e-03	4.8e-05	6.5e-03	2.9e-03	3.7e-03	0.2	0.4
20	4.9e-06	3.7e-04	9.1e-07	4.7e-10	8.2e-10	5.6e-03	2.1e-03
30	1.2e-07	2.0e-04	5.3e-09	4.8e-10	6.0e-11	1.6e-03	3.8e-06
40	2.2e-09	8.6e-06	6.6e-09			7.9e-05	3.4e-07
50	5.5e-10	4.3e-05				1.5e-05	1.0e-08
75	2.0e-10	3.3e-06				1.0e-07	5.4e-11
100		5.8e-06				2.9e-08	
150		1.6e-06					
200		4.8e-07					
500		4.6e-08					
1000		9.4e-09					
48				3.6e-11	1.2e-11		
128			3.1e-09				
160	1.9e-11						1.9e-10
192						3.2e-07	
1024		2.5e-08					

Table 3.3: Convergency behavior with respect to various stochastic processes. Reference values are, in order of the processes' appearance in the table, given by: 19.099354724, 10.993703187, 19.812948843, 49.790905469, 99.999905510, 5.785155450, 22.318945791.

as in the case of VG, the method is highly accurate. Both, the pricing errors as well as convergency is in line with the reference model chosen.

Based on the various test scenarios, the accuracy of the model is controllable by means of parameter M. It is worth mentioning that a sufficient parametrization of input value M is dependent on time to maturity T as well as kurtosis c_4 to a large extend. Therefore, as a rule of thumb, we suggest

$$M = \left\lceil 20 + \frac{3}{100} \cdot S + 2 \cdot \frac{c_4}{T} \right\rceil . \tag{3.30}$$

Skewness is hereby not a direct part of the rule of thumb. However, since skewness interacts with kurtosis, it is accounted for to some degree by using cumulant c_4. To come up with (3.30), we use an identical dataset of skewed and fat tailed marginal distributions as before and implement the optimal value of M as a dependent variable within a median regression.[28] When using a linear regression, outliers due to short maturities (given time to maturity is extremely low, the fraction $\frac{c_4}{T}$ becomes high) result in an overshooting of the estimated coefficient. Thus, we choose to implement a quantile regression estimating the conditional

[28] Several robustness checks have shown that the effect of including further variables besides a constant term, c_4 and T is marginal.

median instead of an estimation of the arithmetic mean. Judging on the behavior of this rule of thumb on several random datasets[29], the specification in (3.30) yields to a level of accuracy in the price approximation of at least $1 \cdot 10^{-5}$ in more than 95% of cases. This level of accuracy is sufficient for most practical applications. If a higher level of accuracy is needed, equation (3.30) provides a starting point from which adjustments according to the specific needs can be made.

3.6 Conclusion

In this paper, we present a new method for evaluating European-type options via time-frequency theory. More specifically, we use Gabor frames with truncated Gaussian generator functions to split the risk neutral pricing integral into two independent inner products. The first inner product is affected by the stochastic process of the underlying asset and the second one is based on the type of contract to be evaluated. Besides European put and call options, we also outline procedures for pricing various binary options in Appendix 3.A.

In Gabor theory, as in most parts of non-orthogonal frame theory, a major issue lies in computation of the dual frame. A function approximation is only feasible by means of a generator in combination with a dual function belonging to this particular generator. Instead of time consuming numerical computation and inversion of the frame operator, we use a method known in the literature as 'painless non-orthogonal expansion' to achieve a considerable simplification. Even though the technique eases the formulation of dual frames, the inner product that is built upon this dual frame cannot be solved in closed form. Thus, while the inner product containing the generator function can be solved in closed form, a fractional FFT procedure is used to numerically evaluate the inner product containing the dual frame. The shape of the dual frame can differ considerably from its generating frame. To assure a fast and accurate calculation of the dual frame inner product, we propose a method to control parameter α and thereby generate a smooth dual frame.

We implement the FFT method by Carr and Madan (1999) as well as the Cos method by Fang and Oosterlee (2008) as reference models. Numerical tests show that the Gabor method achieves very high levels of accuracy combined with short computational times. Compared to the Cos method—which relies on a cosine series expansion—the Gabor method becomes more involved in terms of computations needed. This heavier computational load, however, has the advantage that even extremely short dated contracts as well as heavily-skewed and fat-tailed stochastic processes can be priced reliably. In addition, the Gabor method is less sensitive to truncation limits which are needed to implement series expansion methods. Compared to the FFT method, the Gabor method does not rely on a damping parameter which is needed to ensure L^2-convergency of the pricing integral. Therefore, the only truly exogene model parameter is the maximal number of terms in the frequency domain. The method at hand shares this feature with the Cos method. In case of the Gabor method,

[29] Besides the value of the underlying S, strike K and time to maturity T, input parameters of the stochastic process are randomized as well yielding to random samples.

we formulate a rule of thumb to come up with sufficiently high parameter values assuring a price accuracy on a level of at least 10^{-5} which is sufficient for most practical applications. Higher levels of accuracy can be obtained by simply increasing the parameter value that is suggested by the rule of thumb.

Appendix: Non-Orthogonal Option Pricing

3.A European-type contracts

To simplify notation, we introduce the variable

$$\zeta = (k + 2(b-a)\beta m)\,\pi$$

with $k, m \in \mathbb{Z}$ and $\alpha, \beta > 0$.

Cash-or-nothing call

A Cash-or-Nothing call option pays the holder of the contract a certain amount P if $S_T \geq K$ at maturity. The contract's coefficients v_k needed to calculate equation (3.15) are given by

$$\begin{aligned} v_k &= \frac{1}{2(b-a)} \int_a^b (P\ \mathbb{1}_{x\geq 0}(x))\ e^{2\pi i m\beta x}\ e^{ik\pi\frac{x-a}{b-a}}\ dx \\ &= \begin{cases} \frac{i\left(e^{ik\pi\frac{a}{a-b}} - e^{i\pi(2b\beta m+k)}\right)P}{\zeta} & \text{for } \zeta \neq 0 \\ \frac{bP}{2(b-a)} & \text{for } \zeta = 0 \end{cases} . \end{aligned}$$

Cash-or-nothing put

A Cash-or-Nothing put option pays the holder of the contract a certain amount P if $S_T < K$ at maturity. The contract's coefficients v_k needed to calculate equation (3.15) are given by

$$\begin{aligned} v_k &= \frac{1}{2(b-a)} \int_a^b (P\ \mathbb{1}_{x<0}(x))\ e^{2\pi i m\beta x}\ e^{ik\pi\frac{x-a}{b-a}}\ dx \\ &= \begin{cases} \frac{-i\left(e^{ik\pi\frac{a}{a-b}} - e^{2\pi i m\beta a}\right)P}{\zeta} & \text{for } \zeta \neq 0 \\ -\frac{aP}{2(b-a)} & \text{for } \zeta = 0 \end{cases} . \end{aligned}$$

Asset-or-nothing put

An Asset-or-Nothing put option hands out the underlying asset if $S_T < K$ at maturity. The contract's coefficients v_k needed to calculate equation (3.15) are given by

$$\begin{aligned} v_k &= \frac{1}{2(b-a)} \int_a^0 (Ke^x)\ e^{2\pi i m \beta x}\ e^{ik\pi \frac{x-a}{b-a}}\ dx \\ &= \begin{cases} \frac{e^{ik\pi \frac{a}{a-b}} \left(e^{a - \frac{ia\zeta}{a-b}} - 1 \right) K}{2(a-b-i\zeta)} & \text{for } \zeta \neq 0 \\ \frac{K(1-e^a)}{2(b-a)} & \text{for } \zeta = 0 \end{cases} . \end{aligned}$$

Asset-or-nothing call

An Asset-or-Nothing call option hands out the underlying asset if $S_T \geq K$ at maturity. The contract's coefficients v_k needed to calculate equation (3.15) are given by

$$\begin{aligned} v_k &= \frac{1}{2(b-a)} \int_0^b (Ke^x)\ e^{2\pi i m \beta x}\ e^{ik\pi \frac{x-a}{b-a}}\ dx \\ &= \begin{cases} \frac{e^{ik\pi \frac{a}{a-b}} \left(e^{b - \frac{ib\zeta}{a-b}} - 1 \right) K}{2(a-b-i\zeta)} & \text{for } \zeta \neq 0 \\ \frac{K(e^b-1)}{2(b-a)} & \text{for } \zeta = 0 \end{cases} . \end{aligned}$$

European call options

A call option pays the holder of the contract an amount $S_T - K$ if $S_T > K$ at maturity. The contract's coefficients v_k needed to calculate equation (3.15) are given by

$$\begin{aligned} v_k &= \frac{1}{2(b-a)} \int_a^b K\left(e^x - 1\right)^+\ e^{2\pi i m \beta x}\ e^{ik\pi \frac{x-a}{b-a}}\ dx \\ &= \begin{cases} -\frac{K}{2} \left(\frac{i\left(e^{ik\pi \frac{a}{a-b}} - e^{i\pi(k+2b\beta m)} \right)}{\zeta} + \frac{e^{ik\pi \frac{a}{a-b}} \left(e^{b - \frac{ib\zeta}{b-a}} - 1 \right)}{a-b-i\zeta} \right) & \text{for } \zeta \neq 0 \\ \frac{K}{2(b-a)} \left(e^b - b - 1 \right) & \text{for } \zeta = 0 \end{cases} . \end{aligned}$$

An alternative approach is to compute the value of a put option and use the Put-Call-Parity to draw a conclusion on the value of the call option.

Chapter 4

A Trigger to Rule Them All: Valuation of Multi-Asset Barrier Options

4.1 Introduction

From Black and Scholes (1973), option pricing methods have been developed in many directions. Today, a vast literature is present dealing with both the coverage of stylized facts that are regularly observed in markets as well as the numerical implementation of the models. Most of the research is thereby based on one-dimensional contracts rather than multi-asset contracts. However, within the class of exotic options, multi-asset derivatives are an important subgroup traded on over-the-counter markets to a large extend. Although representing an important branch of option contracts, models to value multi-asset options are researched to a lesser extent compared to their univariate counterparts.

Especially within the class of barrier options, numerical implementation is often not straight forward. Monte Carlo methods are a natural choice since the computational complexity increases only linearly with the number of underlying assets. However, Monte Carlo simulations converge very slowly when barrier levels are involved (see e.g. Boyle and Lau, 1994). Next to Monte Carlo methods, lattice approaches and grid methods are candidates for a numerical implementation as well. In case of lattice approaches, Boyle (1988) discusses an implementation given two underlying assets. The assumption of a two dimensional basket is broadened in Boyle et al. (1989) to a d-dimensional[30] basket size. A drawback of tree methods is an oscillatory convergency behavior which prevents an implementation of extrapolation methods. Korn and Müller (2009) present a method which is based on the before mentioned tree methods but assures a smooth convergency. However, it is still the case that the model approximates a d-dimensional geometric Brownian motion as the underlying

[30] Throughout this chapter we refer to the number of assets within the basket as the option's dimensionality. Thus, a two-dimensional contract refers to a contingent claim with two underlying assets.

source of randomness. Related to lattice approaches are finite difference methods also known as grid methods. A well known example of using finite difference methods to value derivative securities is Hull and White (1990). Both approaches, lattice and grid, have in common that incorporating path dependency can be cumbersome, especially when time varying barriers are in place.

Within this chapter, we contribute to this area of research by analyze multivariate Fourier series expansions and their ability to price financial derivatives. Fourier series expansion methods are originally used to solve the heat equation, i.e. a partial differential equation. Over the years the field of applications of Fourier series vastly increased and the methodology became an integral part of signal processing, quantum mechanics and electrical engineering, among others. In finance, and especially in option pricing, Fourier series expansions are introduced in Fang and Oosterlee (2008) within a univariate setting. Based on this contribution, this chapter is an enhancement to Ruijter and Oosterlee (2012) in which Fourier cosine series expansion methods are used to price rainbow options[31] with Bermudan features[32] as well as to Meng and Ding (2013) which use modified sine series to price European style options based on two underlyings.

Since multivariate Fourier series have been implemented successfully to price European and Bermudan options our main interest lies within the following question: Are there structural differences within the class of multivariate Fourier series when using them to price derivatives? More precisely, we are interested in answering the question whether a trader who is interested in implementing Fourier series methods to price various basket options should prefer one type of Fourier series over another. Our main findings are that pure sine series are dominated by both modified sine series as well as cosine series. Furthermore, when also considering computational complexity in addition to pure convergency arguments, modified sine series also dominate cosine series and should therefore be the method of choice.

Compared to Ruijter and Oosterlee (2012), we consider discrete multi-asset barrier options of various kinds instead of Bermudan options. Moreover, we include not only cosine series but also sine series and modified sine series. As mentioned before, modified sine series are also studied in Meng and Ding (2013). The authors assert a similar convergency behavior compared to cosine series. However, our analysis shows that modified sine series unveil their strength to a full extend only when pricing higher basket sizes or when introducing path dependency which has not been done in Meng and Ding (2013).

Fourier series methods offer a flexible framework which enables the pricing of derivatives assuming more complex stochastic processes such as pure jump Lévy processes as well as the implementation of various kinds of contracts. Besides outlying the procedure to incorporate path dependency in a similar fashion as in Ruijter and Oosterlee (2012), we concentrate on

[31] The term rainbow options is equivalent to the term multi-asset option. On occasion the various sources of randomness within a multi-asset option are referred to as the colors of a rainbow.

[32] In its ability to exercise early, a Bermudan option ranges in between European options and American options. While early exercise is not possible in case of European options and always possible in case of American options, Bermudan options offer a finite number of predefined points in time on which exercise is possible.

different areas of application and thereby stroll from pure equity basket options (put and call options as well as binary options) to hybrid equity and credit derivatives (multi-asset equity swaps) along to structured products (multi-barrier reverse convertibles).

Within the core of the numerical algorithm used to price these products is a multivariate version of Parseval's theorem that describes how to calculate an integral of the product of two functions as the multidimensional sum of their Fourier coefficients. For this sum to be finite it is important that the coefficients converge given a finite number of terms. For the numerical implementation by a finite number of terms within the sums, the rate of convergency becomes crucial. In context of the option pricing problems chosen in this chapter, we consider two broader types of functions and their coefficients: First, density functions which, as part of a stochastic process, control the individual behavior of the underlying assets and the dependency between each other. Density functions $f \in C^{\infty}(\mathbb{R}^d)$ show a high degree of smoothness, which yields to a fast decay of their respective Fourier coefficients. Second, payoff functions as defined by the option contract under consideration. Payoff functions $\nu \in L^2([a,b]^d)$ are often non-analytic functions[33] which yields to a slower degree of convergency. Especially in relation to payoff functions, the choice between cosine series coefficients, sine series coefficients and modified sine series coefficients affects the overall level of accuracy. Primary, we consider step function-like payoffs since they can be used to calculate probabilities of hitting a barrier which becomes an important ingredient of more exotic options or credit derivatives.

The course of the chapter is two-folded: Within the first part (Sections 2 and 3) the convergency behavior of the Fourier series expansions considered are central in Section 2. Based on these insights, approximation quality and computational speed of European type options are subject to Section 3. The second part of the chapter (Section 4) introduces the algorithm to price discrete barrier options of various types. Section 5 contains a conclusion and is followed by an appendix.

4.2 Multidimensional Fourier series

We approach the problem of evaluating multi-asset option contracts with the help of d-dimensional series expansion methods. Common techniques within the field of series expansion are either full Fourier series, Fourier cosine series, Fourier sine series, or mixed Fourier series. Rather recently in Iserles and Nørsett (2008) as well as Adcock (2010), modified Fourier series joint as a new alternative to the traditional series expansion methods. Common to all of the before mentioned methods is that they rely on basis functions that share the feature of being orthogonal with respect to an inner product on a given interval. The basis functions considered here are trigonometric functions. Thus, each of the resulting series expansions approximate a given function by means of a superposition of trigonometric functions.

[33] Analytic functions are such functions that can be described locally by a convergent power series.

With the goal of option pricing in mind, we focus on the approximation of two distinct kind of functions: on the one hand multivariate density functions and on the other hand payout profiles of derivative contracts. In case of the last-mentioned, next to call and put options, especially digital[34] payout profiles are considered. Digital options are hereby defined to pay a pre-defined amount if stock prices at maturity stay above their respective strike prices. If at least one of the stock prices is below the strike level, the contract expires worthless. The fact that every stock price has to stay above its strike value is an assumption to simplify argumentations in the following. It would also be possible to reduce the final payoff if some underlyings stay below their strike level instead of expiring worthless.

The notation within this paper is chosen as follows: variables printed in bold indicate vectors, matrices, or block matrices such as $\mathbf{n} = (\mathbf{n}_1, ..., \mathbf{n}_d)$, where, depending on the expansion method, $\mathbf{n}_j$ is e.g. defined as the vector $\mathbf{n}_j = (0, ..., N_j - 1)$. The symbol $\mathcal{R}(\cdot)$ indicates taking the real part of a complex number, $i = \sqrt{-1}$ defines the imaginary unit itself and $\mathbf{n}^\top$ is associated with taking the transpose of vector $\mathbf{n}$. In addition, we consider a d-dimensional space $\Omega = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$ as well as several subsets $\Theta \subset \Omega$ that are defined later on. Lower case variables ω_j and θ_j are used to describe the length of an interval at a given dimension, e.g. $\omega_j = b_j - a_j$.

We consider three different types of orthogonal sets to approximate a given function $f(\mathbf{x}) \in \Omega$, $\mathbf{x} \in \mathbb{R}^d$. Hereby, we use the variables

$$u_j = \frac{n_j \pi}{b_j - a_j}, \quad p_j = \frac{\left(n_j - \frac{1}{2}\right)\pi}{b_j - a_j}$$

heavily in the following. As a first orthogonal set we consider the set of cosine functions

$$\begin{aligned} C_{\mathbf{n}}(\mathbf{x}) &= \cos\left(\mathbf{n}_1 \pi \frac{x_1 - a_1}{b_1 - a_1}\right) \cdot ... \cdot \cos\left(\mathbf{n}_d \pi \frac{x_d - a_d}{b_d - a_d}\right) \\ &= \cos\left(\mathbf{u}_1 (x_1 - a_1)\right) \cdot ... \cdot \cos\left(\mathbf{u}_d (x_d - a_d)\right), \end{aligned}$$

second, the set of sine functions

$$\begin{aligned} S_{\mathbf{n}}(\mathbf{x}) &= \sin\left(\mathbf{n}_1 \pi \frac{x_1 - a_1}{b_1 - a_1}\right) \cdot ... \cdot \sin\left(\mathbf{n}_d \pi \frac{x_d - a_d}{b_d - a_d}\right) \\ &= \sin\left(\mathbf{u}_1 (x_1 - a_1)\right) \cdot ... \cdot \sin\left(\mathbf{u}_d (x_d - a_d)\right), \end{aligned}$$

and third, the set of modified sine functions which are abbreviated by the term 'modsine' occasionally in the following.

$$\begin{aligned} M_{\mathbf{n}}(\mathbf{x}) &= \sin\left(\left(\mathbf{n}_1 - \frac{1}{2}\right) \pi \frac{x_1 - a_1}{b_1 - a_1}\right) \cdot ... \cdot \sin\left(\left(\mathbf{n}_d - \frac{1}{2}\right) \pi \frac{x_d - a_d}{b_d - a_d}\right) \\ &= \sin\left(\mathbf{p}_1 (x_1 - a_1)\right) \cdot ... \cdot \sin\left(\mathbf{p}_d (x_d - a_d)\right). \end{aligned}$$

[34] The terms 'binary option' and 'digital option' are used as synonyms within this chapter.

Given the individual base functions, we can define the respective series approximation expressions as follows: Fourier cosine series are given by

$$f(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{N}^d} c_{\mathbf{n}}^C \; C_{\mathbf{n}}(\mathbf{x}) \tag{4.1}$$
$$c_{\mathbf{n}}^C = \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \int_\Omega f(\mathbf{x}) \; C_{\mathbf{n}}(\mathbf{x}) \, d\mathbf{x},$$

where d_n defines the number of non-zero terms in the vector $\mathbf{n} = (\mathbf{n}_1, \mathbf{n}_2, ..., \mathbf{n}_d)$ (Pivato, 2010). Fourier sine series are given by

$$f(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{N}_+^d} c_{\mathbf{n}}^S \; S_{\mathbf{n}}(\mathbf{x}) \tag{4.2}$$
$$c_{\mathbf{n}}^S = \frac{2^{d}}{\prod_{k=1}^{d} \omega_k} \int_\Omega f(\mathbf{x}) \; S_{\mathbf{n}}(\mathbf{x}) \, d\mathbf{x},$$

and modified sine series by

$$f(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{N}_+^d} c_{\mathbf{n}}^M \; M_{\mathbf{n}}(\mathbf{x}) \tag{4.3}$$
$$c_{\mathbf{n}}^M = \frac{2^{d}}{\prod_{k=1}^{d} \omega_k} \int_\Omega f(\mathbf{x}) \; M_{\mathbf{n}}(\mathbf{x}) \, d\mathbf{x}.$$

It is worth noticing that the sums in (4.1) - (4.3) are d-dimensional sums starting at different integers. While integer zero is included in (4.1), the sums in (4.2) and (4.3) start at integer one as indicated by $\mathbf{n} \in \mathbb{N}_+^d$. Within the above series approximations, the dimension of the block matrix $\mathbf{n}$ defines the number of actual sums. That is, in case of three underlying assets, we define $\mathbf{n} = [\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3]$. Each individual sum's starting and ending integer is then defined by the respective vectors $\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{n}_3$. Keeping these relations in mind, we start with elaborating on the approximation of multivariate distribution functions

Approximation of multivariate distribution functions

We hereby consider the marginal density function $f_{\mathbf{X}_t}(\mathbf{x}_t|\mathbf{x}_s) \in \mathbb{R}^d$, $s < t$. Where writing the whole expression it is not crucial for the understanding, we write $f_{\mathbf{X}}(\mathbf{x})$ for the marginal distribution. When trying to define series coefficients of a multivariate density function $f_{\mathbf{X}}(\mathbf{x}) \in \mathbb{R}$, the concept of characteristic functions is of special use.

$$\phi(\omega) = \int_{\mathbb{R}^d} f_{\mathbf{X}}(\mathbf{x}) \; e^{i\omega^\top \mathbf{x}} \; d\mathbf{x}. \tag{4.4}$$

As can be seen in equation (4.4), The characteristic function $\phi\colon \mathbb{R}^d \to \mathbb{C}^d$ transforms the information of a real valued density function into complex space. The major advantage

hereby is that these functions are known in closed form for a wide range of stochastic processes and underlying marginal distributions.

To be able to approximate a multivariate distribution function, we need to prepare the series coefficients $f_{\mathbf{n}}$ of a density function before returning to the concept of characteristic functions. We state the procedure in detail for cosine coefficients. Similar steps based on sine and modified sine series can be used to reproduce the formulas stated in the following. In terms of cosine series approximation, series coefficients are calculated by

$$\begin{aligned} f_{\mathbf{n}}^{C} &= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \int_{\Omega} f_{\mathbf{X}}(\mathbf{x})\ C_{\mathbf{n}}(\mathbf{x})\, d\mathbf{x} \\ &= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \int_{\Omega} f_{X_1,\ldots,X_d}(x_1,\ldots,x_d) \prod_{j=1}^{d} \cos\left(\mathbf{u}_j(x_j - a_j)\right)\ dx_1 \cdots dx_d. \end{aligned}$$

Compared to the form of an arbitrary characteristic function in (4.4), we have to rewrite the product of cosine terms into complex exponentials to approach the aim of approximating density functions by means of multi-dimensional Fourier series. Hereby, the trigonometric identities

$$\begin{aligned} \cos(z) &= \frac{1}{2}\left(e^{iz} + e^{-iz}\right), \quad \text{and} \\ \sin(z) &= \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) \end{aligned}$$

become important. Incorporating theses identities, the coefficients can be written as

$$f_{\mathbf{n}}^{C} = \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \frac{1}{2^d} \int_{\Omega} f_{\mathbf{X}}(x_1,\ldots,x_d) \prod_{j=1}^{d} \left(e^{i\mathbf{u}_j(x_j-a_j)} + e^{-i\mathbf{u}_j(x_j-a_j)}\right)\ dx_1 \cdots dx_d \tag{4.5}$$

$$= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \frac{1}{2^d} \int_{\Omega} f_{\mathbf{X}}(\mathbf{x}) \sum_{j=1}^{2^d} e^{i(u_{\mathbf{n}}\mathrm{diag}(\xi_j))^{\top}(\mathbf{x}-\mathbf{a})} d\mathbf{x}, \tag{4.6}$$

where $\xi \in \{-1, 1\}^{2^d \times d}$ in (4.6) is a matrix of binary combinations. The entries of each row vector within ξ, hereby, represent the signs of the exponentials when expanding the product $\prod_{j=1}^{d} \left(e^{i\mathbf{u}_j(x_j-a_j)} + e^{-\mathbf{u}_j(x_j-a_j)}\right)$ in (4.5). In case of $d = 2$, the matrix shows the following structure:

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$$

However, since the trigonometric identities above imply $2\,\mathcal{R}(e^{iz}) = 2\,\mathcal{R}(e^{-iz}) = e^{iz} + e^{-iz} \in \mathbb{R}$, it is sufficient to only use the first d row vectors in ξ and drop the second half with

complementary signs. Thus, the summation in (4.6) can be shortened by half and, given the interval Ω is chosen sufficiently large, characteristic functions can be introduced. In addition, to shorten notations, we define $\bar{\xi} = \text{diag}(\xi)$.

$$\begin{aligned} f_{\mathbf{n}}^{C} &= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \frac{1}{2^d} \int_{\Omega} f_{\mathbf{X}}(\mathbf{x}) \sum_{j=1}^{2^{d-1}} 2 \; \mathcal{R}\left\{ e^{i(u_{\mathbf{n}}\bar{\xi}_j)^{\top}(\mathbf{x}-\mathbf{a})} \right\} d\mathbf{x} \\ &= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} 2^{1-d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \int_{\Omega} f(\mathbf{x}) \; e^{i(u_{\mathbf{n}}\bar{\xi}_j)^{\top}(\mathbf{x}-\mathbf{a})} d\mathbf{x} \right\} \end{aligned} \tag{4.7}$$

Equation (4.7) almost shows the structure of a characteristic function within the brackets. To complete the picture, we use the fact that the characteristic function of a stable distribution can be written independently of the state s in which the individual values of the d-dimensional random variable $\mathbf{X}_t$ are:

$$\phi(\omega, \mathbf{x}_s) = \int_{\mathbb{R}^d} f_{\mathbf{X}_t}(\mathbf{x}_t | \mathbf{x}_s = 0) \; e^{i\omega^{\top}\mathbf{x}} \, d\mathbf{x} \; e^{i\omega^{\top}\mathbf{x}_s} = \phi(\omega) \; e^{i\omega^{\top}\mathbf{x}_s},$$

where the characteristic function to the very right indicates initial values of zero. In a final step, multiplying the characteristic function by an exponential term of the form

$$\phi(\omega, \mathbf{x}_s) \; e^{i\omega^{\top}\mathbf{a}}$$

and assigning $\omega = u_{\mathbf{n}}\bar{\xi}_j$ yields to the following cosine series coefficients:

$$f_{\mathbf{n}}^{C} = \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} 2^{1-d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \phi\left(u_{\mathbf{n}}\bar{\xi}_j\right) e^{i(u_{\mathbf{n}}\bar{\xi}_j)^{\top}(\mathbf{x}_s-\mathbf{a})} \right\} \tag{4.8}$$

In a similar fashion, the coefficients for Fourier sine series and modified Fourier sine series can be obtained as:

$$f_{\mathbf{n}}^{S} = \frac{2}{\prod_{k=1}^{d} \omega_k} \frac{1}{(2i)^d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \left(\prod_{k=1}^{d} \xi_j^k \right) \phi\left(u_{\mathbf{n}}\bar{\xi}_j\right) e^{i(u_{\mathbf{n}}\bar{\xi}_j)^{\top}(x_{\mathbf{n}}-a_{\mathbf{n}})} \right\} \tag{4.9}$$

$$f_{\mathbf{n}}^{M} = \frac{2}{\prod_{k=1}^{d} \omega_k} \frac{1}{(2i)^d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \left(\prod_{k=1}^{d} \xi_j^k \right) \phi\left(p_{\mathbf{n}}\bar{\xi}_j\right) e^{i(p_{\mathbf{n}}\bar{\xi}_j)^{\top}(x_{\mathbf{n}}-a_{\mathbf{n}})} \right\} \tag{4.10}$$

In theory, using equations (4.8)-(4.10) in combination with their respective series representation (4.1) - (4.3), a given multivariate density function in $L^2(\Omega)$ can be approximated by incorporating an infinite number of coefficients. Since every numerical implementation has to be done with a finite number of coefficients, the convergence of the coefficients towards zero is of crucial importance. The faster the coefficients converge towards zero, the fewer terms are needed within the series representation, the faster computations are done.

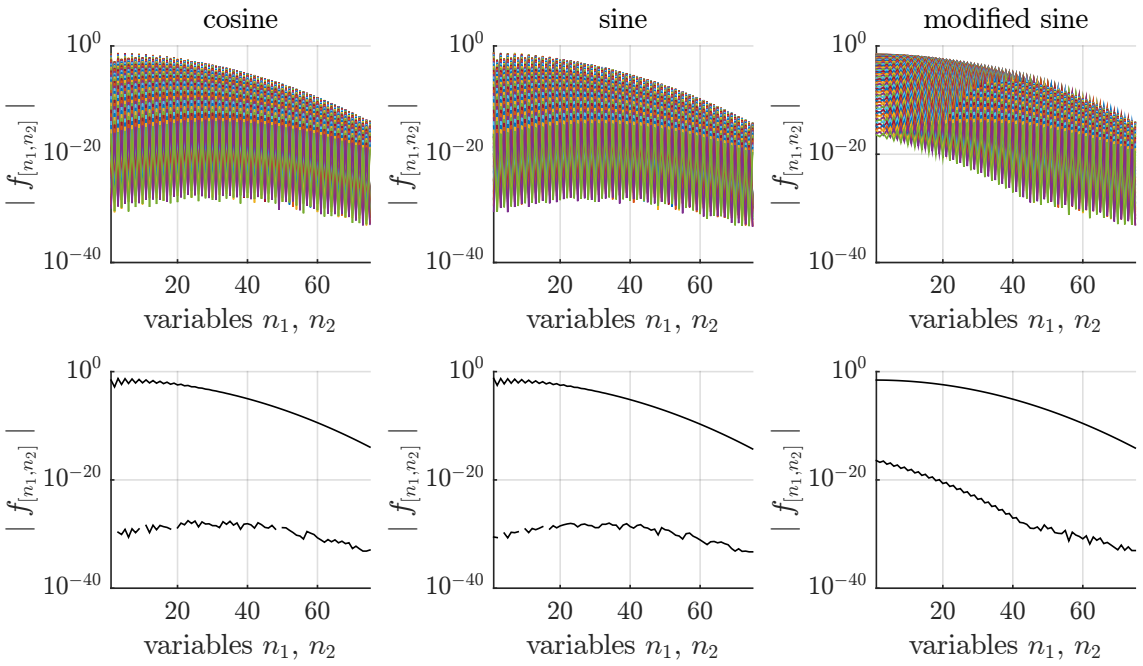

Figure 4.1: Decay behavior of Fourier coefficients approximating a bivariate normal density with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 0.3$ and $\rho = 0.4$. The respective characteristic function is given by $\phi(\mathbf{u}) = e^{-\frac{1}{2}\mathbf{u}^\top \Sigma \mathbf{u}}$, with Σ being the covariance matrix.

To show the convergency behavior in more detail we approximate a bivariate normal distribution with parameters chosen as $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 0.3$ and $\rho = 0.4$. Since a bivariate Normal distribution $f_{XY} \in C^\infty(\mathbb{R})$ is a particularly smooth function, convergency will be fast. The example is supposed to provide the reader with an intuition for how the convergency of the coefficients in Figure 4.1 interacts with the overall quality of approximation in Figure 4.2 as well as elaborate on differences between the three types of Fourier series.

Both, Figure 4.1 and Figure 4.2 indicate that all three approximation schemes behave similar. Figure 4.1 depicts the series coefficients in a log-plot. The respective Fourier coefficients of a bivariate Normal distribution are stored in a matrix of size $\mathbb{R}^{N_1 \times N_2}$, in this case $\mathbb{R}^{75\times 75}$. Plotting the matrix results in a three-dimensional graph. However, to emphasize on the decay rate of the coefficients, we display the results in a two dimensional plot. The entire matrices are shown in the upper part of Figure 4.1 where the oscillatory nature of the coefficients is visible. The same information is plotted in the lower part of the figure. However, instead of all coefficients, only the largest and smallest values are included. Cosine and sine series coefficients are almost identical in its behavior to decay. In case of the modified sine series, the overall picture varies to a certain amount but without any pronounced differences in terms of the speed of convergency. Concentrating on the lower thee subgraphs, the slowest coefficients to converge, indicated by the respective upper line, reach a level below 10^{-15} at about $N_1 = N_2 = 75$ terms within the series. Given a machine precision of $\epsilon = 2.22 \cdot 10^{-16}$, this, in turn, implies that incorporating coefficients beyond $N_1 = N_2 \approx 75$ does not improve the numerical approximation quality.

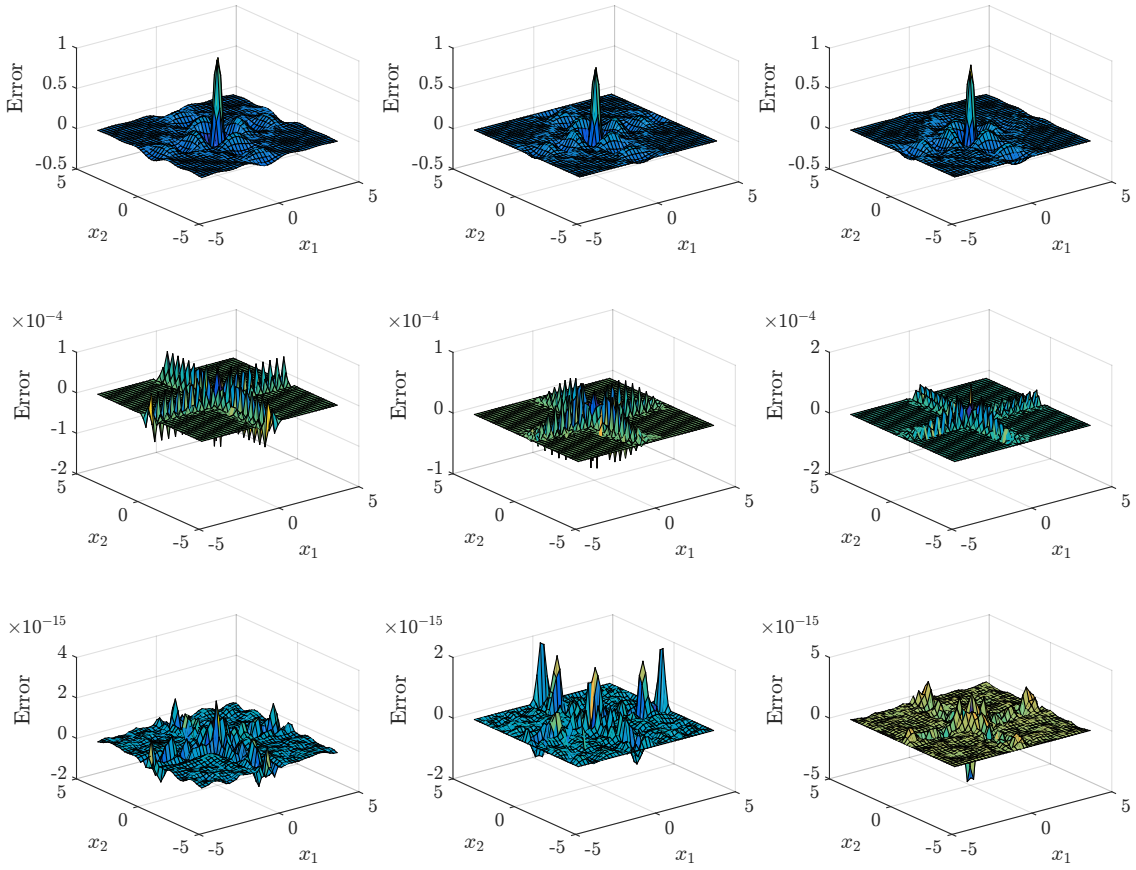

Figure 4.2: Approximation quality of a bivariate normal density with $\sigma_1 = \sigma_2 = 0.3$ and $\rho = 0.4$. Each series approximation is implemented with $N_1 = N_2 = (10, 40, 80)$ terms. Graphs are vertically connected, meaning that the graphs to the left belong to a cosine approximation, the graphs in the middle belong to a sine approximation and the graphs to the right belong to a modified sine series.

To complete the picture, Figure 4.2 shows the error terms, i.e. the deviation of the series representation from the true function, based on the coefficients displayed in Figure 4.1. The parameters chosen suggests that the orthogonal set used within the series expansion does not yield to dissimilar results in terms of approximation quality. All of the three methods show a fast convergency of the coefficients and based on that an accurate approximation given a quite low number of terms within the summation. Figure 4.2 also displays a common pattern of Fourier series: the approximation is based on wave functions. This characteristic is most pronounced when considering the deviation pattern of Fourier series with a low number of terms as in the upper part of the figure. Hereby, the error terms move in waves comparable to movements of water when throwing a stone into a pond.

Approximation of a cube

This type of exponential convergency, however, does not translate over to functions that do not show the same degree of smoothness. As long as the function under consideration is still square integrable $\nu \in L^2(\Omega)$, it can be approximated using Fourier series. However, convergency slows down given a lower degree of smoothness. As a contrast to the smooth

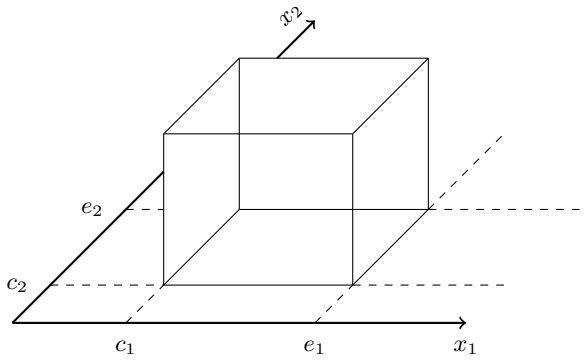

Figure 4.3: Cube with support on the domain $\Theta = [c_1, e_1] \times [c_2, e_2]$.

and symmetric bivariate density function before, we therefore consider the approximation of a d-dimensional cube that lives on the subinterval $\Theta = [c_1, e_1] \times ... \times [c_d, e_d] \subset \Omega$ given by:

$$\nu(\mathbf{x}) = \mathbb{1}_{\Theta}(\mathbf{x}) = \mathbb{1}_{\{[c_1,e_1]\}}(x_1) \cdot ... \cdot \mathbb{1}_{\{[c_d,e_d]\}}(x_d),$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function. Figure 4.3 displays the scenario in a setting where variable $\mathbf{x}$ assumes two domains $\mathbf{x} = (x_1, x_2)$.

Again, starting with cosine series coefficients, the following holds:

$$\begin{aligned}
\nu_{\mathbf{n}}^{C} &= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \int_{\Omega} \mathbb{1}_{\{[c_1,e_1]\}}(x_1) \cdot ... \cdot \mathbb{1}_{\{[c_d,e_d]\}}(x_d) \\
&\qquad \cdot \cos\left(\mathbf{u}_1(x_1 - a_1)\right) \cdot ... \cdot \cos\left(\mathbf{u}_d(x_d - a_d)\right)\ dx_1 ... dx_d \\
&= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \prod_{j=1}^{d} \int_{\Omega_j} \mathbb{1}_{\{[c_j,e_j]\}}(x_j)\ \cos\left(\mathbf{u}_j(x_j - a_j)\right)\ dx_j \\
&= \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \prod_{j=1}^{d} \int_{\Theta_j} \cos\left(\mathbf{u}_j(x_j - a_j)\right)\ dx_j
\end{aligned} \tag{4.11}$$

Equation (4.11) shows that the d-dimensional coefficients of the function ν are Kronecker products of the univariate coefficients. Thus it is sufficient to solve the integral only in a univariate case $\Theta = [c, e]$

$$\begin{aligned}
\nu_n^C(\omega, \theta) &= \int_{\Theta} \cos\left(u(x - a)\right)\ dx \\
&= \begin{cases} e - c & n = 0 \\ -\frac{(b-a)\left(\sin\left(n\pi\frac{a-c}{a-b}\right) - \sin\left(n\pi\frac{a-e}{a-b}\right)\right)}{n\pi} & n \neq 0 \end{cases} \\
&= \begin{cases} \theta & n = 0 \\ -\frac{\omega\left(\sin\left(n\pi\frac{c-a}{\omega}\right) - \sin\left(n\pi\frac{e-a}{\omega}\right)\right)}{n\pi} & n \neq 0 \end{cases}
\end{aligned} \tag{4.12}$$

and use (4.12) to calculate the multivariate cosine coefficients as

$$\nu_{\mathbf{n}}^{C} = \frac{2^{d_n}}{\prod_{k=1}^{d} \omega_k} \, \nu_{n_1}^{C}(\omega_1, \theta_1) \otimes \nu_{n_2}^{C}(\omega_2, \theta_2) \otimes ... \otimes \nu_{n_d}^{C}(\omega_d, \theta_d).$$

The symbol $\otimes$ hereby represents taking a Konecker product. Following similar steps, sine coefficients are given by

$$\nu_{\mathbf{n}}^{S} = \frac{2^{d}}{\prod_{k=1}^{d} \omega_k} \, \nu_{n_1}^{S}(\omega_1, \theta_1) \otimes \nu_{n_2}^{S}(\omega_2, \theta_2) \otimes ... \otimes \nu_{n_d}^{S}(\omega_d, \theta_d),$$

where

$$\nu_{n}^{S}(\omega, \theta) = \frac{\omega \left(\cos\left(n\pi \frac{c-a}{\omega} \right) - \cos\left(n\pi \frac{e-a}{\omega} \right) \right)}{n\pi}, \tag{4.13}$$

and modified sine coefficients are given by

$$\nu_{\mathbf{n}}^{M} = \frac{2^{d}}{\prod_{k=1}^{d} \omega_k} \, \nu_{n_1}^{M}(\omega_1, \theta_1) \otimes \nu_{n_2}^{M}(\omega_2, \theta_2) \otimes ... \otimes \nu_{n_d}^{M}(\omega_d, \theta_d),$$

where

$$\nu_{n}^{M}(\omega, \theta) = \frac{2\,\omega \left(\cos\left(\frac{(2n-1)\pi}{2} \frac{c-a}{\omega} \right) - \cos\left(\frac{(2n-1)\pi}{2} \frac{e-a}{\omega} \right) \right)}{n\pi}. \tag{4.14}$$

From equations (4.12), (4.13) and (4.14) we infer that each of the series coefficients decay like $\mathcal{O}\left(\frac{1}{n_1 \cdot ... \cdot n_d}\right)$, independent of the orthogonal set used. This type of asymptotic behavior indicates that convergency is slowest given one of the variables, say n_1, is fixed at a low value, e.g. $n_1 = 1$. Increasing all other variables results in a decay but not in a decay as rapid as when all variables are increased together.

In a two-dimensional setting, this means that the largest magnitudes in the absolute values of the coefficients are given when fixing one of the parameters at its lowest possible value and increasing the other one. If we take, as an example, the function

$$\nu(x_1, x_2) = \mathbb{1}_{x_1 > 1} \cdot \mathbb{1}_{x_2 > 1}$$

on the cube $\Omega = [-\pi, \pi]^2$, Figure 4.4 displays the convergency behavior of all three orthogonal sets considered here. As indicated above, convergency is rather slow which is due to the fact that step functions are non-analytic with at least one jump in the function profile. In case of choosing one orthogonal set over another, there is no clear-cut decision available. Thus, calculating the price of a binary-type option contract, the three series approximation methods ought to perform in a similar fashion (see Section 4.3). This, however changes when considering smoother payoff functions compared to step functions.

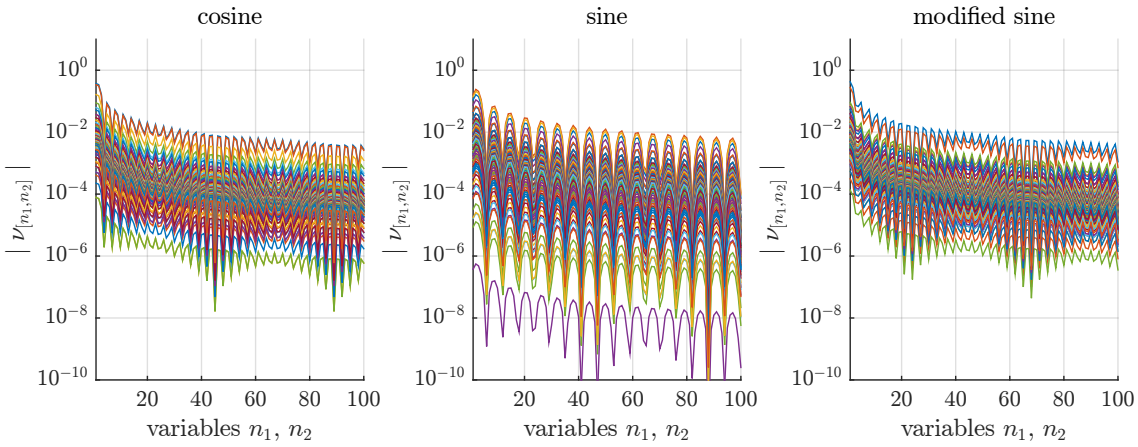

Figure 4.4: Decay behavior of Fourier coefficients (cosine coefficients are shown in the left part, sine coefficients in the middle and modified sine coefficients in the right part) approximating a step function $\nu(\mathbf{x}) = \mathbb{1}_{\mathbf{x}>1} \in [-\pi, \pi]^2$.

Approximation of exponential functions

Up to this point, we considered functions that produce coefficients which can be localized on different endings of a convergency scale: on the one hand a normal density with exponential convergency and, on the other hand, step functions with a slower $\mathcal{O}\big(\frac{1}{n_1 \cdot \ldots \cdot n_d}\big)$ convergency. As a final example within this section, we include exponential functions which can be ranged in between the two extrem cases of before.

Next to step functions in form of binary options, payoff contracts that incorporate the exponential function are important. In contrast to step functions, the exponential function is a smooth analytic function. The higher degree of smoothness hereby translates into a faster decay of the series coefficients. This is, however, only true for cosine series as well as modified sine series. To elaborate on this result, we introduce the series coefficients in a univariate setting. In a multivariate setting, findings are identical.

We define $\nu(x) = e^x$ on a domain $\nu \in [a, b]$ such that the coefficients are given by

$$\nu_n(x) = \frac{2}{b-a} \int_a^b e^x \, \varphi_n(x) \, dx.$$

Solving the above equation for each orthogonal set used in our analysis, the series coefficients using an orthogonal cosine set is given by equation (4.15), coefficients of a sine series by

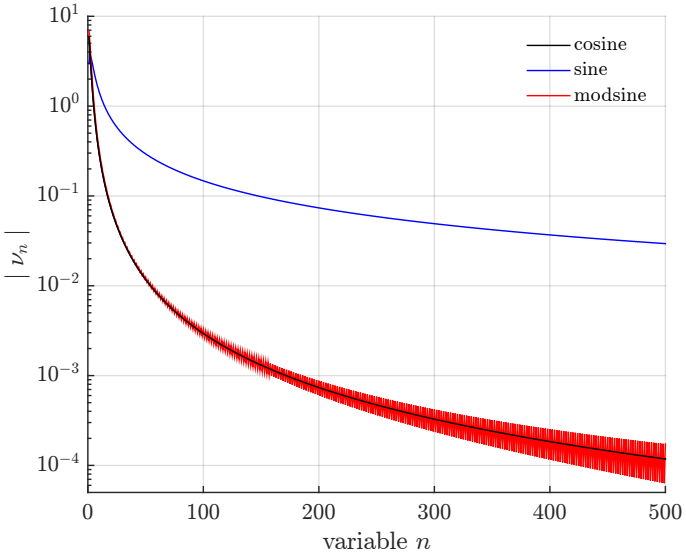

Figure 4.5: Decay behavior of Fourier coefficients approximating an exponential function $\nu(\mathbf{x}) = e^x \in [-\pi, \pi]^2$.

equation (4.16), and equation (4.17) defines coefficients of a modified sine series:

$$\frac{2}{b-a}\int_a^b e^x \cos\left(u(x-a)\right)\, dx = \frac{2}{b-a}\cdot\frac{e^b(-1)^n - e^a}{n^2\left(\frac{\pi}{b-a}\right)^2 + 1} \tag{4.15}$$

$$\frac{2}{b-a}\int_a^b e^x \sin\left(u(x-a)\right)\, dx = \frac{2}{b-a}\cdot\frac{n\pi(b-a)\left(e^a - e^b(-1)^n\right)}{n^2\pi^2 - (b-a)^2} \tag{4.16}$$

$$\frac{2}{b-a}\int_a^b e^x \sin\left(p(x-a)\right)\, dx = \frac{2}{b-a}\cdot\frac{e^a\frac{(n-\frac{1}{2})\pi}{b-a} - e^b(-1)^n}{\left(\frac{(n-\frac{1}{2})\pi}{b-a}\right)^2 + 1} \tag{4.17}$$

From the equations in (4.15)-(4.17), we conclude that cosine series coefficients as well as modified sine series coefficients decay at a rate of $\mathcal{O}\left(\frac{1}{n^2}\right)$. Sine series coefficients, however, converge only at a rate of $\mathcal{O}\left(\frac{1}{n}\right)$ as is also illustrated in Figure 4.5. Coefficients based on a sine series converge rather slow. Cosine and modified sine series show the same speed of convergency. However, while cosine coefficients converge in a smooth way, modified sine series oscillate in their decay.[35]

This rather specific example of an exponential function, can be generalized: modified sine series as well as cosine series show a superior convergency in case of analytic functions that are non-periodic (Iserles and Nørsett, 2008). How distinct these differences in convergency have a influence on the actual task of pricing multi-asset options is part of the following section.

[35] When looking at Figure 4.5 a technical note is in order: Modified sine series do converge in an oscillatory nature. However the point of break at around $n = 150$ is due to a graphical issue and not due to a pattern in convergency.

4.3 Path independent option pricing

4.3.1 Option pricing framework

In a risk neutral setting, the price of a European basket option is given by an discounted expected value under a risk neutral measure $\mathbb{Q}$ of the option's payoff function ν at maturity

$$\begin{aligned}\nu_t &= e^{-r(T-t)}\,\mathbb{E}^{\mathbb{Q}}[\nu(\mathbf{S}_T) \mid \mathcal{F}_t] \\ &= e^{-r(T-t)} \int_{\mathbb{R}^d_+} \nu(\mathbf{S}_T)\, f(\mathbf{S}_T \mid \mathbf{S}_t)\; d\mathbf{S}_T. \end{aligned} \tag{4.18}$$

To effectively implement Fourier series techniques, some changes in variables are helpful. First, we use log returns instead of the underlying asset price S_t. Second, variables such as asset specific strike levels K_j are also normalized by their respective underlying at a given point in time:

$$x_j^t = \log \frac{S_j(t)}{S_j(0)}, \quad k_j = \log \frac{K_j}{S_j(0)}, \quad j = 1, ..., d$$

For reasons of simplicity, we use $t = 0$ as initial value. To keep notation clearly laid out, we try to avoid multiple indices and, therefore, write time indications in the variables lower index when using matrices such as $\mathbf{x}_t = (x_1^t, ..., x_d^t)$. By means of these changes, equation (4.18) becomes

$$\nu_t = e^{-r(T-t)} \int_{\mathbb{R}^d} \nu(\mathbf{x}_T)\, f(\mathbf{x}_T \mid \mathbf{x}_t)\; d\mathbf{x}_T. \tag{4.19}$$

Since neither of functions ν and f are periodic, the above d-dimensional integral has to be truncated at a suitable point. This means that the domain Ω has to be chosen at points in which the integral has already converged to zero. To accomplish this task, we follow Ruijter and Oosterlee (2012) and use the statistical moments of the stochastic process to determine appropriate bounds for the above integral.

For now, we assume that the space Ω is a suitable truncation of $\mathbb{R}^d$ with respect to the above integral. As a result, equation (4.19) can be approximated by

$$\nu_t = e^{-r(T-t)} \int_{\Omega} \nu(\mathbf{x}_T)\, f(\mathbf{x}_T \mid \mathbf{x}_t)\; d\mathbf{x}_T. \tag{4.20}$$

Based on Parseval's identity, we can use Fourier series coefficients to calculate a given integral if the integral under consideration can be written as the product of two square integrable functions. For two arbitrary functions $g, k \in L^2(\mathbb{R}^d)$, this general statement can be formalized

by the following expression:

$$\int_{\Omega} g(\mathbf{x})\, k(\mathbf{x})\; d\mathbf{x} = \begin{cases} \frac{\prod_{j=1}^{d} \omega_j}{2^{d_n}} \sum\limits_{\mathbf{n}\in\mathbb{Z}^d} g_{\mathbf{n}}\, k_{\mathbf{n}} & \text{if a cosine series} \\ \frac{\prod_{j=1}^{d} \omega_j}{2^{d}} \sum\limits_{\mathbf{n}\in\mathbb{Z}_+^d} g_{\mathbf{n}}\, k_{\mathbf{n}} & \text{if a (modified) sine series.} \end{cases} \tag{4.21}$$

The d-dimensional version of Parseval's identity in (4.21) can be used to solve the integral in (4.20). In the following, due to the fact that sine series will show to be less effective, we limit the formulas to versions based on cosine and modified sine series. However, sine series are included within the numerical examples of this section.

To get rid of the factor in front of the sums in (4.21) we use the full version of the density coefficients $f_{\mathbf{n}}$ as shown in (4.8) for cosine series and (4.10) for modified sine series but only a shortened version of the payoff coefficients $\hat{\nu}_{\mathbf{n}}$, defined by

$$\hat{\nu}_{\mathbf{n}} = \begin{cases} \int\limits_{\Omega} \nu(\mathbf{x})\, C_{\mathbf{n}}(\mathbf{x})\; d\mathbf{x} & \text{if a cosine series} \\ \int\limits_{\Omega} \nu(\mathbf{x})\, M_{\mathbf{n}}(\mathbf{x})\; d\mathbf{x} & \text{if a modified sine series.} \end{cases} \tag{4.22}$$

Strictly speaking the values stored in $\hat{\nu}_{\mathbf{n}}$ are no longer Fourier coefficients of the payoff function since they miss their weights. Even being incorrect in a strict sense, due to reasons of convenience, we continue to refer to $\hat{\nu}_{\mathbf{n}}$ as series coefficients or Fourier coefficients. Thus, using equations (4.8)-(4.10) in combination with equation (4.20), the value of an European basket option evaluated at time $0 \le t < T$ can be calculated by means of

$$v_t = \frac{e^{-r(T-t)} 2^{d_n}}{\prod_{j=1}^{d} \omega_j} 2^{1-d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \phi\left(u_{\mathbf{n}} \bar{\xi}_j\right)\; e^{i(u_{\mathbf{n}}\bar{\xi}_j)^{\top}(\mathbf{x}_t - \mathbf{a})}\, \hat{\nu}_{\mathbf{n}} \right\} \tag{4.23}$$

in case of cosine series coefficients. Using modified sine coefficients, the expression changes slightly:

$$v_t = \frac{e^{-r(T-t)} 2^{d}}{\prod_{j=1}^{d} \omega_j} \left(\frac{1}{i}\right)^{d} 2^{1-d} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \left(\prod_{k=1}^{d} \xi_j^k \right) \phi\left(p_{\mathbf{n}} \bar{\xi}_j\right)\; e^{i(p_{\mathbf{n}}\bar{\xi}_j)^{\top}(\mathbf{x}_t - \mathbf{a})}\, \hat{\nu}_{\mathbf{n}} \right\} \tag{4.24}$$

For the rest of the chapter we use a unified version of equations (4.23) and (4.24). For this purpose, we define two variables ψ_1, ψ_2 which depend on the choice of basis functions used in the series expansion

$$\psi_1 = \begin{cases} \frac{2^{d_n}}{\prod_{j=1}^{d} \omega_j} 2^{1-d} & \text{if cosine} \\ \frac{2^{d}}{\prod_{j=1}^{d} \omega_j} 2^{1-d} \left(\frac{1}{i}\right)^{d} & \text{if modsine} \end{cases}, \quad \psi_2 = \begin{cases} 1 & \text{if cosine} \\ \left(\prod_{k=1}^{d} \xi_j^k \right) & \text{if modsine} \end{cases},$$

$$q_j = \begin{cases} \frac{n_j \pi}{b_j - a_j} & \text{if cosine} \\ \frac{(n_j - 0.5)\pi}{b_j - a_j} & \text{if modsine} \end{cases}.$$

The resulting unified pricing formula is now given by

$$v_t = e^{-r(T-t)}\,\psi_1 \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{\psi_2\,\phi\left(q_{\mathbf{n}}\bar{\xi}_j\right)\; e^{i(q_{\mathbf{n}}\bar{\xi}_j)^\top(\mathbf{x}_t-\mathbf{a})}\,\hat{\nu}_{\mathbf{n}}\right\}. \tag{4.25}$$

4.3.2 European digital option

With the help of equations (4.25), the first numerical test can be implemented. As a first test scenario, European digital options are introduced. The payoff function is assumed to pay an amount of $\nu = 1$ if each of the underlying assets stays above their individual strike level at maturity:

$$\nu(\mathbf{x}) = \mathbb{1}_{\{\mathbf{x}>\mathbf{k}\}} = \mathbb{1}_{\{x_1>k_1\}} \cdot \ldots \cdot \mathbb{1}_{\{x_d>k_d\}}$$

The evolution of the basket is described by a 3-dimensional geometric Brownian motion, $j = 1, 2, 3$, with the following covariance matrix:

$$\Sigma = \begin{bmatrix} 0.43^2 & 0.1204 & 0.10105 \\ 0.1204 & 0.56^2 & 0.1316 \\ 0.10105 & 0.1316 & 0.47^2 \end{bmatrix}$$

This parameter constellation is labeled Scenario I. Within the covariance matrix Σ, correlation is chosen to be at a level of $\rho = 0.5$. However, this choice of a flat correlation structure is not crucial. Non-identical correlation structures do not change the overall behavior of the model. We always use Scenario I when testing a model setup under the assumption of a multivariate geometric Brownian motion. The characteristic function of a multivariate geometric Brownian motion is hereby given by

$$\phi(\mathbf{u}) = e^{i\,\mathbf{m}^\top\mathbf{u}\,t - \frac{1}{2}\mathbf{u}^\top\Sigma\,\mathbf{u}\,t},$$

where $\mathbf{m}$ represents a vector of drift correction terms according to $m_j = (r - \frac{1}{2}\sigma_j)t$. When assuming two instead of three underlying assets we also use the parametrization of Scenario I but drop values which are assigned to asset three. Expressions for series coefficients of digital options can be found in Appendix 4.A.

Besides the convergency behavior of the model using cosine series, sine series and modified sine series, Figure 4.6 shows CPU[36] time needed to calculate the prices for a given number of terms within the summation. The reference value is calculated as follows: First, a Monte Carlo simulation is used to calculate the price on a level of accuracy of 10^{-5} to assure the model price is not off target in a fundamental way. Second, a reference value on a higher level of accuracy is calculated by means of cosine series using $N_{\max} = (1000, 1000, 1000)$ terms. This second reference value is mainly used to show convergency behavior.

[36] Calculations are implemented within the programming environment Matlab 2014b using a machine with a 2.3 GHz Quad-Core CPU and 16 GB RAM.

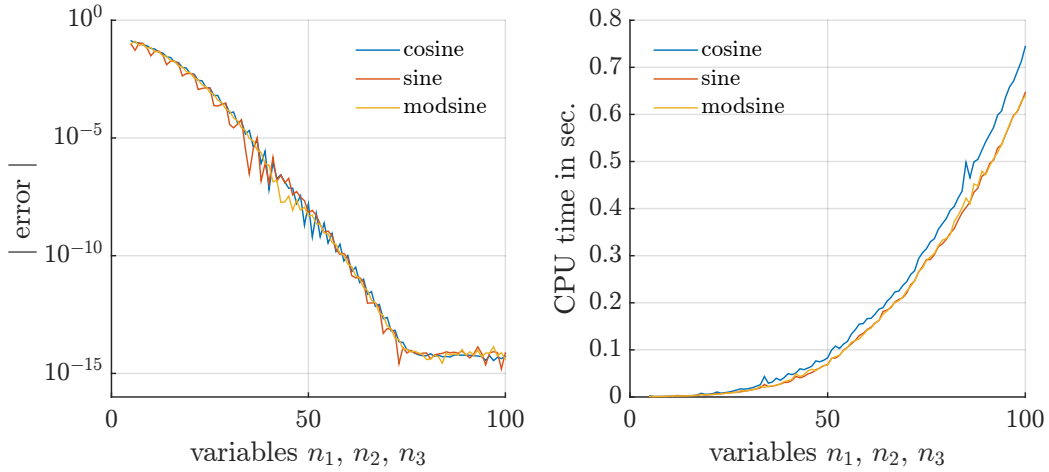

Figure 4.6: Approximation of a 3-dimensional European digital option with parametrization as in Scenario I. Additional parameters are given by $K_j = 0.85$ and $r = 0.01$.

	computational time in seconds		
Accuracy	cosine	sine	modsine
10^{-2}	0.0037	0.0034	0.0043
10^{-5}	0.0229	0.0223	0.0202
10^{-10}	0.1341	0.1155	0.1170

Table 4.1: Computational time needed to reach a predefined level of accuracy approximating a 3-dimensional European digital option with parametrization as in Szenario I. Additional parameters are given by $K_j = 0.85$ and $r = 0.01$.

The overall convergency is plotted on the left part of Figure 4.6. Each expansion method has its individual pattern of convergency. However, in terms of the actual error at a given set (n_1, n_2, n_3) none of the three versions systematically outperforms the other two.

The reason CPU why time is included is observable on the right-hand side of Figure 4.6. With respect to calculation time needed, cosine series are slower. This is in part due to the factor 2^{d_n} which is a crucial ingredient of a cosine series. While the value 2^d used in (modified) sine series is a scalar, the term 2^{d_n} represents a d-dimensional matrix. Calculation time and storage costs of this variable increases with the number of underlyings as well as the number of terms in the expansion.

To show critical levels of accuracy in more detail, Table 4.1 summarizes the computational time needed to converge below a given level of accuracy. The before-mentioned slightly slower convergency of cosine series is also observable within the table. Sine and modified sine series do not differ in a fundamental way.

A geometric Brownian motion is based on a Normal distribution as the source of randomness. As displayed in Section 4.2, coefficients of a Normal distribution converge fast which is due to the smoothness and symmetry of the function itself. To also reveal the methods abilities

under more extreme and more realistic market situations, we also define a second test scenario under the assumption of a d-dimensional variance gamma process.

There are several ways to define a multivariate variance gamma process. In this chapter we implement the version introduced in Leoni and Schoutens (2008) which defines the variance gamma process as a gamma-time changed brownian motion with a single background parameter ν_{VG} as well as individual skewness parameters θ_j, volatilities σ_j and volatility adjustment parameters ω_j.[37]

The characteristic function of this particular version of the variance gamma process is defined by

$$\phi(\mathbf{u}) = e^{i\,\mathbf{u}^\top \mathbf{m}\, t} \cdot \left(1 - i\,\nu_{VG}\,\mathbf{u}^\top\theta + \frac{1}{2}\nu_{VG}\,\mathbf{u}^\top(\omega^\top\omega \circ \Sigma_{\mathrm{BM}})\mathbf{u}\right)^{-\frac{t}{\nu_{VG}}},$$

where $\circ$ represents element-wise multiplication. In addition, the drift correction term needed to define an equivalent martingale measure is given by

$$m_j = r + (\nu_{VG}\log(1 - \theta_j\nu_{VG} - 0.5\nu_{VG}(\omega_j\sigma_j).^2))^{-1}.$$

For numerical implementations we define a Scenario II as following:

$$\Sigma_{\mathrm{BM}} = \begin{bmatrix} 0.18390 & 0.11960 & 0.09965 \\ 0.11960 & 0.31296 & 0.13048 \\ 0.09965 & 0.13048 & 0.21894 \end{bmatrix}$$

$$\theta = \begin{bmatrix} -1 \\ -0.95 \\ -0.9 \end{bmatrix}, \quad \omega = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \nu = 0.05$$

Hereby, the covariance matrix of the subordinated Brownian motion is chosen in a way that the overall asset correlation fits the correlation structure of Scenario I. However, instead of a symmetric marginal distributions, the distributions are skewed and fat-tailed.

The convergency behavior of the model considering cosine series, sine series and modified sine series are shown on the left hand side of Figure 4.7. Due to the non-symetric nature of the marginal density functions, convergency of the according density coefficients slows down compared to normal distributed marginal densities coefficients. While the series converge at about 75 terms in each summation in Scenario I, the environment in Scenario II demands at least 100 terms in all three dimensions.

Similar conclusions can be drawn from CPU times shown on the right part of Figure 4.7 as well as based on the entries of Table 4.2. In Scenario II, a very high level of accuracy (10^{-10}) is reached in a time period of well below a second. However, due to a higher number of terms, calculation time increased by a factor of 2.7 on average compared to a multivariate geometric

[37] This adjustment parameters became important within a calibration routine. Since we do not consider such a task within this chapter, we refer to Leoni and Schoutens (2008) for further details.

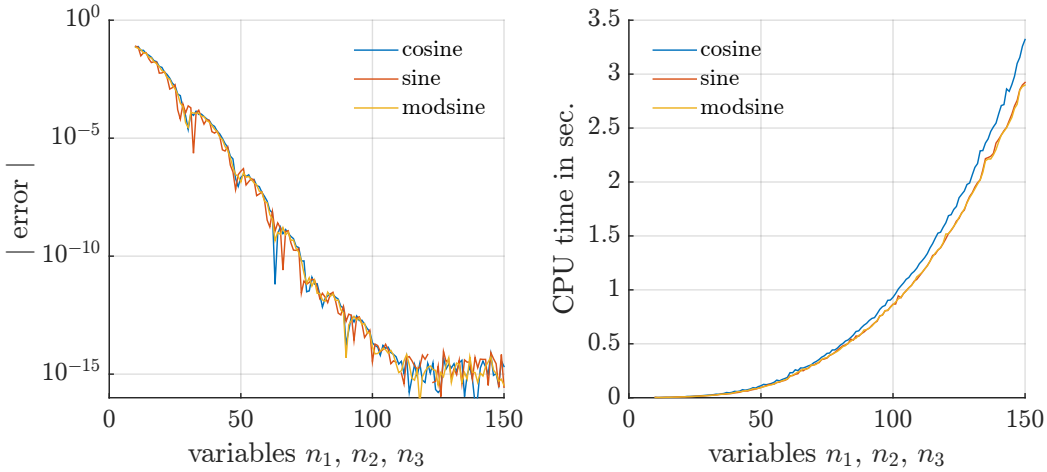

Figure 4.7: Approximation of a 3-dimensional European digital option with parametrization as in Scenario II. Additional parameters are given by $K_j = 0.85$ and $r = 0.01$.

	computational time in seconds		
Accuracy	cosine	sine	modsine
10^{-2}	0.0087	0.0072	0.0084
10^{-5}	0.0592	0.0564	0.0568
10^{-10}	0.3391	0.3225	0.3307

Table 4.2: Computational time needed to reach a predefined level of accuracy approximating a 3-dimensional European digital option with parametrization as in Scenario II. Additional parameters are given by $K_j = 0.85$ and $r = 0.01$.

Brownian motion in Scenario I. When considering an accuracy of 10^{-5}, computational times differ by a factor of 2.6. Differences become smaller the lower the level of accuracy is set. Considering an accuracy of 10^{-2}, time differences are defined by a factor of 2.1 on average. From the point of view of a practitioner, a level of accuracy in between 10^{-2} and 10^{-5} should be sufficient for most cases of application.

4.3.3 European exchange option

In a second test, exchange options based on Margrabe (1978) are implemented. This two dimensional option contract has a payoff function according to

$$\nu(S_1, S_2) = \max\left(S_1(T) - S_2(T), 0\right)$$

which translates to

$$\nu(x_1, x_2) = \max\left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}, 0\right).$$

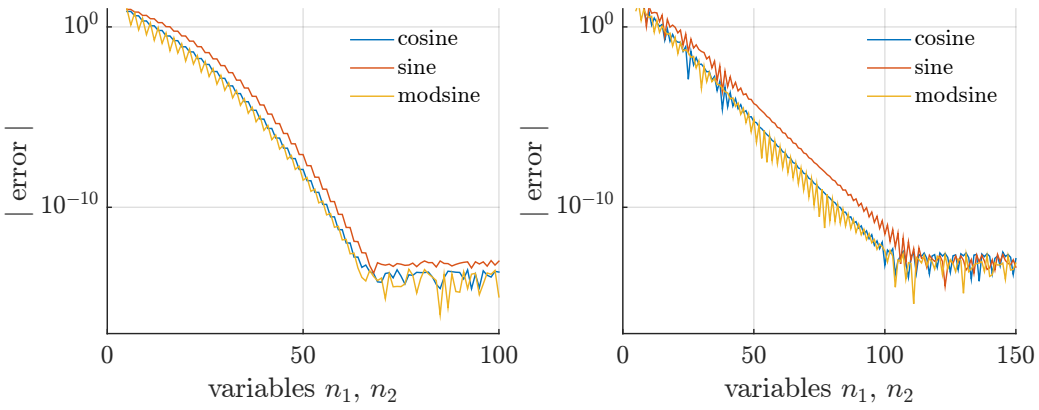

Figure 4.8: Approximation of a 2-dimensional European Margrabe option with parametrization as in Scenario I on the left hand side and parametrization as in Scenario II on the right hand side.

Thus, the series coefficients on $\Omega = [a_1, b_1] \times [a_2, b_2]$ as defined by (4.22) read as follows:

$$\begin{aligned}\hat{\nu}_{[n_1,n_2]} &= \int_\Omega \max\left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}, 0\right)\, \varphi_{[n_1,n_2]}(x_1, x_2)\, dx_1 dx_2 \\ &= \int_{a_2}^{b_2} \int_{z+x_2}^{b_1} \left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}\right)\, \varphi_{[n_1,n_2]}(x_1, x_2)\, dx_1 dx_2, \end{aligned} \tag{4.26}$$

where $z = \log \frac{S_2(T)}{S_1(T)}$. Since individual solutions assuming a given basis function $\varphi_{[n_1,n_2]}(x_1, x_2)$ are quite lengthy, we refer to Appendix 4.B in which closed form solution to the integral in (4.26) can be found.

As can be seen in (4.26), this type of option incorporates an analytic non-periodic function. Thus, from the findings in Section 4.2, we expect a pure sine series expansion to show a slower convergency. Reference values in case of Scenario I are given by Margrabe (1978). Reference values in case of Scenario II are not as readily available. We use simulations to retrieve reference values on a level of accuracy of about 10^{-4} to assure the overall correctness of the methods and use the individual Fourier series with a very high number of terms to analyze their convergency behavior.

Figure 4.8 combines the convergency behavior of all three series types within a log-plot. From the figure, the slower convergency of a sine series is observable. This effect is even more pronounced when pricing Margrabe options under the assumption of non symmetric marginal density functions as in Scenario II. Cosine series and modified sine series, however, behave similar. Once again, selected error levels are displayed in Table 4.3.

At least in a European-type environment, we conclude that pure sine series are dominated by both cosine series and modified sine series in terms of convergency behavior of the payout coefficients. When considering higher dimensional contracts, the storage costs of cosine series are higher. This is due to the fact that the summation in a cosine series expansion starts at integer $n = 0$, which yields to case differentiation in the cosine series coefficients.

		computational time in seconds		
	Accuracy	cosine	sine	modsine
Scen. I	10^{-2}	0.0022	0.0025	0.0017
	10^{-5}	0.0030	0.0041	0.0031
	10^{-10}	0.0054	0.0057	0.0051
Scen. II	10^{-2}	0.0025	0.0034	0.0025
	10^{-5}	0.0049	0.0067	0.0055
	10^{-10}	0.0074	0.0090	0.0071

Table 4.3: Computational time needed to reach a predefined level of accuracy approximating a 2-dimensional Margrabe option.

In the next section, we will add path dependency as in Fang and Oosterlee (2009) to the algorithm. Instead of considering all three expansion types, we completely drop pure sine series in the following and only use cosine series and modified sine series.

4.4 Discrete barrier options

As shown in Fang and Oosterlee (2009) and Ruijter and Oosterlee (2012), the incorporation of path dependency at a discrete time grid involves the recursive calculation of vector-matrix products and matrix-matrix products respectively. Two particular types of matrices are within the center of an efficient calculation as they occur within the derivation of these products: Hankel and Toeplitz matrices. Since they can be embedded within a circular matrix, vector-matrix products can be calculated in $\mathcal{O}(N \log_2 N)$ complex calculations instead of $\mathcal{O}(N^2)$ calculations by means of the FFT algorithm. We discuss the algorithm in a general manner before moving forward to enhancing the work of Ruijter and Oosterlee (2012) by deriving formulas to price d-dimensional binary barrier options. Besides this classical type of barrier option, we also consider more complex types of derivatives and structured products that are regularly priced via slow converging Monte Carlo simulations.

4.4.1 Introducing path dependency

Discrete barrier options distinct themselves from their continuous counterparts by the fact that the contract's underlying assets are evaluated to whether a corresponding barrier event has occurred only at pre-defined dates in time. Therefore, matching the notation of Fang and Oosterlee (2009), we define a finite set of evaluation points $\mathcal{T} = \{t_0, ..., t_E = T\}$.

In the context of defining monitoring dates also the domain Ω on which the pricing integral is evaluated on needs to be adjusted. In case of European options the number of subdomains included in Ω is defined by the number of underlying assets that are incorporated within the option contract. In case of barrier options this structure is enhanced by the time dimen-

sion which is defined by the number of monitoring events. Thus, when pricing discretely monitored path dependent options the relevant integration domain is described by:

$$\Omega = \begin{bmatrix} \Omega_1(t_1) \times \Omega_2(t_1) \times ... \times \Omega_d(t_1) \\ \Omega_1(t_2) \times \Omega_2(t_2) \times ... \times \Omega_d(t_2) \\ \vdots \\ \Omega_1(T) \times \Omega_2(T) \times ... \times \Omega_d(T) \end{bmatrix} = \begin{bmatrix} \Omega(t_1) \\ \Omega(t_2) \\ \vdots \\ \Omega(T) \end{bmatrix} \tag{4.27}$$

For each time step an individual d-dimensional integration domain is given. The same is true for the subset $\Theta \subset \Omega$ which is not defined by the convergency of the probability density function itself but rather includes the specific characteristics of the barrier levels given in the contract. Hereby, asset specific barrier levels H_j, $j = 1, ..., d$, are normalized in the same manner as strike levels k_j and the underlyings x_j itself have been:

$$h_j^t = \log \frac{H_j(t)}{S_j(0)}$$

Taking a three-asset binary barrier option as an example, the following payout structure

$$\begin{aligned} \nu(\mathbf{x}_T) &= \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_1^s > h_1^s \right\}} \cdot \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_2^s > h_2^s \right\}} \cdot \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_3^s > h_3^s \right\}} \\ &= \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} \mathbf{x}_s > \mathbf{h}_s \right\}} \end{aligned} \tag{4.28}$$

could be defined. Assuming this particular structure describes a knock-out product that ceases to exist as soon as any one of the underlying assets hits their barrier level. Other d-dimensional payout structures such as

$$\nu(\mathbf{x}_T) = d - \frac{1}{d} \left(\mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_1^s > h_1^s \right\}} + ... + \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_d^s > h_d^s \right\}} \right) \tag{4.29}$$

are, however, also convenient to implement. In contrast to (4.28), equation (4.29) describes a scenario where the final payout is lowered by the fraction of underlying assets any time one of the assets hits a barrier. Based on the payout structure chosen, the payout coefficients are specified identical to a European option contract. Taking equation (4.28) as an example once again, we write the coefficients as

$$\begin{aligned} \hat{\nu}_{\mathbf{n}}^{t_E} &= \int_{a_3}^{b_3} \int_{a_2}^{b_2} \int_{a_1}^{b_1} \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_1^s > h_1^s \right\}} \cdot \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_2^s > h_2^s \right\}} \cdot \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} x_3^s > h_3^s \right\}} \\ &\qquad \varphi_{[n_1,n_2,n_3]}(x_1^{t_E}, x_2^{t_E}, x_3^{t_E})\, dx_1^{t_E}\, dx_2^{t_E}\, dx_3^{t_E} \\ &= \int_{\Omega(t_E)} \mathbb{1}_{\left\{ \min\limits_{0 \le s \le T} \mathbf{x}_s > \mathbf{h}_s \right\}} \varphi_{\mathbf{n}}(\mathbf{x}_{t_E})\, d\mathbf{x}_{t_E}. \end{aligned} \tag{4.30}$$

Incorporating the barrier levels into the integration limits results in the definition of the subdomain $\Theta \subset \Omega$. At time step t_E, the corresponding version to equation (4.30) is given

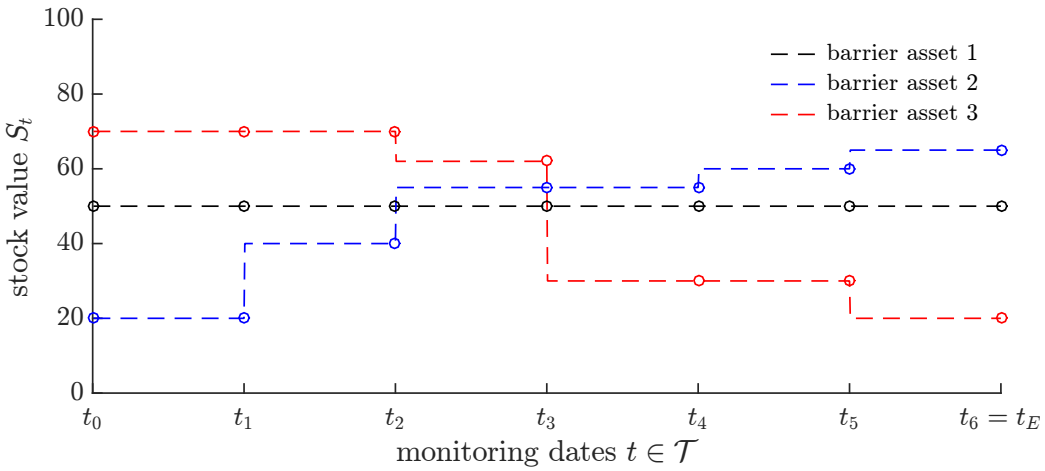

Figure 4.9: Exemplary barrier structure of a 3-dimensional contract. Each of the barrier follows an individual pattern: A constant barrier in case of the first asset, a decreasing barrier in case of the second asset and an increasing barrier in case of the third asset.

by the term

$$\hat{\nu}_{\mathbf{n}}^{t_E} = \int_{\Theta(t_E)} 1 \, \varphi_{\mathbf{n}}(\mathbf{x}_{t_E}) \, d\mathbf{x}_{t_E},$$

and, thus, the subdomain is defined by $\Theta(T) = [h_1, b_1] \times [h_2, b_2] \times [h_3, b_3]$. Equivalently to equation (4.27), the contract specific multi-dimensional domain Θ is, therefore, formulated in general as follows:

$$\Theta = \begin{bmatrix} \Theta_1(t_1) \times \Theta_2(t_1) \times ... \times \Theta_d(t_1) \\ \Theta_1(t_2) \times \Theta_2(t_2) \times ... \times \Theta_d(t_2) \\ \vdots \\ \Theta_1(T) \times \Theta_2(T) \times ... \times \Theta_d(T) \end{bmatrix} = \begin{bmatrix} \Theta(t_1) \\ \Theta(t_2) \\ \vdots \\ \Theta(T) \end{bmatrix} \tag{4.31}$$

By means of (4.31), implementing a three-asset contract with flat barrier levels is just as convenient as combining a flat barrier structure for the first asset with an increasing barrier level for the second underlying and a decreasing structure for a third asset within a single option contract. The only parameters to change are the individual subdomains $\Theta_j(t)$, $j = 1, ..., d$, at given monitoring events $t \in \mathcal{T}$. Figure 4.9 displays the barrier structure of such a contract.

Depending on the contract, monitoring a barrier level can be defined daily, weekly, monthly, quarterly or whatever needs are to be served for. In between each of these points the contract behaves like an European option. This is also the reason for why we can use the formulas introduced for European options in Section 4.3 to evaluate path dependent options of discrete type. The foundation of the algorithm is still a matrix-matrix product that is composed by density coefficients and payout coefficients. However, as can be seen in (4.32), now the payoff

coefficients at time step t_1 instead of at maturity T are crucial.

$$\nu_{t_0} = e^{-r\Delta t} \sum_{\mathbf{n}} f_{\mathbf{n}}^{\Delta t}\, \hat{\nu}_{\mathbf{n}}^{t_1}, \quad \Delta t = t_1 - t_0 \tag{4.32}$$

As shown in Fang and Oosterlee (2009), these coefficients can be retrieved recursively starting at maturity. At $T = t_E$, under consideration of (4.27), a general d-dimensional payout coefficient matrix is given by:

$$\hat{\nu}_{\mathbf{n}}^{t_E} = \int_{\Omega(t_E)} \nu(\mathbf{x}_{t_E})\, \varphi_{\mathbf{n}}(\mathbf{x}_{t_E})\, d\mathbf{x}_{t_E} \tag{4.33}$$

To keep arguments general, we do not yet assume a particular option contract but rather work with the unspecified[38] version from (4.33). It has to be kept in mind, however, that the coefficients in (4.33) can be expressed in closed form for a wide range of contracts. On the contrary, coefficients before maturity are not known from the beginning but can be retrieved by a recursion. Similar to the calculation of the coefficients at maturity, the coefficients one time step ahead $\hat{\nu}_{\mathbf{n}}^{t_{E-1}}$ are given by

$$\hat{\nu}_{\mathbf{n}}^{t_{E-1}} = \int_{\Omega(t_{E-1})} \nu(\mathbf{x}_{t_{E-1}})\, \varphi_{\mathbf{n}}(\mathbf{x}_{t_{E-1}})\, d\mathbf{x}_{t_{E-1}}.$$

The term $\nu(\mathbf{x}_{t_{E-1}})$, however, is defined by the pricing formula (4.25) of an European option with time horizon $\Delta t_{E-1} = t_E - t_{E-1}$:

$$\begin{aligned}
\hat{\nu}_{\mathbf{n}}^{t_{E-1}} &= \int_{\Omega(t_{E-1})} e^{-r\Delta t_{E-1}}\, \psi_1 \sum_{\mathbf{m}} \sum_{j=1}^{2^{d-1}} \mathcal{R} \left\{ \psi_2\, \phi\left(q_{\mathbf{m}} \bar{\xi}_j\right)\, e^{i(q_{\mathbf{m}}\bar{\xi}_j)^\top (\mathbf{x}_{t_{E-1}} - \mathbf{a})} \hat{\nu}_{\mathbf{m}}^{t_E} \right\} \varphi_{\mathbf{n}}(\mathbf{x}_{t_{E-1}})\, d\mathbf{x}_{t_{E-1}} \\
&= e^{-r\Delta t_{E-1}}\, \psi_1 \sum_{\mathbf{m}} \sum_{j=1}^{2^{d-1}} \mathcal{R} \left\{ \psi_2\, \phi\left(q_{\mathbf{m}} \bar{\xi}_j\right) \hat{\nu}_{\mathbf{m}}^{t_E} \int_{\Omega(t_{E-1})} e^{i(q_{\mathbf{m}}\bar{\xi}_j)^\top (\mathbf{x}_{t_{E-1}} - \mathbf{a})} \varphi_{\mathbf{n}}(\mathbf{x}_{t_{E-1}}) d\mathbf{x}_{t_{E-1}} \right\}
\end{aligned} \tag{4.34}$$

Equation (4.34) shows that the payoff coefficients one step backwards in time $\hat{\nu}_{\mathbf{n}}^{t_{E-1}}$ are dependent on the coefficients one time step ahead $\hat{\nu}_{\mathbf{m}}^{t_E}$. Since this is true for any given time step within the set $\mathcal{T}$, the payoff coefficients at $t = t_1$ are given by:

$$\hat{\nu}_{\mathbf{n}}^{t_1} = e^{-r\Delta t_1}\, \psi_1 \sum_{\mathbf{m}} \sum_{j=1}^{2^{d-1}} \mathcal{R} \left\{ \psi_2\, \phi\left(q_{\mathbf{m}} \bar{\xi}_j\right) \hat{\nu}_{\mathbf{m}}^{t_2} \int_{\Omega(t_1)} e^{i(q_{\mathbf{m}}\bar{\xi}_j)^\top (\mathbf{x}_{t_1} - \mathbf{a})} \varphi_{\mathbf{n}}(\mathbf{x}_{t_1}) d\mathbf{x}_{t_1} \right\} \tag{4.35}$$

Now that the recursive nature of the algorithm is introduced, the d-dimensional integral included in (4.34) and (4.35)

$$\int_{\Omega(t)} \bullet\, d\mathbf{x}_{t_1} = \int_{\Omega_1(t)} \int_{\Omega_2(t)} \cdots \int_{\Omega_d(t)} \bullet\, dx_1^{t_1} dx_2^{t_1} \cdots dx_d^{t_1}$$

[38] Using a specified version implies to fix a specific payout structure $\nu(\mathbf{x}_{t_E})$, as e.g. in (4.28) or (4.29), and switch from the integration domain Ω to Θ in the calculation of the payoff coefficients.

needs to be solved at each time step $t \in \mathcal{T}$ before we are able to use formula in (4.35) in combination with equation (4.32) for pricing barrier options. By disentangling the d-dimensional integral, the task comes down to solving an integral of the type

$$\int_{\Omega_j(t)} e^{iv(x-a)}\, \varphi_n(x)\, dx$$

multiple times, where $\Omega_j(t)$ defines the j^{th} subinterval of the d-dimensional domain at time step t. It can be shown, that such integrals are defined by the sum of a Hankel and a Toeplitz matrix as has been done in Fang and Oosterlee (2009) in terms of cosine series. Thus, in Appendix 4.C, we show a derivation of this claim based on modified sine series. Within this section, however, we only state that the above integral is given by the sum of a Hankel matrix H and a Toeplitz matrix T multiplied by some pre-factors

$$\int_{\Omega_j(t)} e^{iv(x-a)}\, \varphi_n(x)\, dx = \begin{cases} \frac{\omega_j}{2}\frac{i}{\pi}\left(H^C + T^C\right) & \text{if cosine} \\ \frac{\omega_j}{2}\frac{1}{\pi}\left(H^M + T^M\right) & \text{if modsine,} \end{cases} \tag{4.36}$$

where superscript C and M stand for matrices that are build on cosine series and modified sine series respectively. Incorporating this relation into equation (4.35), the series coefficients $\hat{\nu}^t_{\mathbf{n}}$ can be retrieved for any given discrete time step $t \in \mathcal{T}$.

To outline the procedure in an explicit structure, we restart at maturity $t_E = T$ and incorporate the knowledge about the recursive structure as well as the knowledge about the calculation of the multivariate integral by means of Hankel and Toeplitz matrices:

$$\begin{aligned}
\hat{\nu}^{t_E}_{\mathbf{n}} &= \int_{\Omega(t_E)} \nu(\mathbf{x}_{t_E})\, \varphi_{\mathbf{n}}(\mathbf{x}_{t_E})\, d\mathbf{x}_{t_E} \\
\hat{\nu}^{t_{E-1}}_{\mathbf{n}} &= \int_{\Omega(t_{E-1})} \nu(\mathbf{x}_{t_{E-1}})\, \varphi_{\mathbf{n}}(\mathbf{x}_{t_{E-1}})\, d\mathbf{x}_{t_{E-1}} \\
&= e^{-r\Delta t}\, \psi_1 \sum_{\mathbf{m}} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \psi_2\, \phi\left(u_{\mathbf{m}} \bar{\xi}_j\right) \hat{\nu}^{t_E}_{\mathbf{m}} \right. \\
&\qquad \left. \prod_{k=1}^{d} \int_{\Omega_k(t_{E-1})} e^{i(u_{m_k}\bar{\xi}_j)^\top (x^k_{t_{E-1}} - a_k)} \varphi_{n_k}(x^k_{t_{E-1}}) dx^k_{t_{E-1}} \right\} \\
&= e^{-r\Delta t}\, \psi_1 \sum_{\mathbf{m}} \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{ \psi_2\, \phi\left(u_{\mathbf{m}} \bar{\xi}_j\right) \hat{\nu}^{t_E}_{\mathbf{m}} \prod_{k=1}^{d} \psi_3^k (H_k + T_k) \right\}
\end{aligned} \tag{4.37}$$

with

$$\psi_3^j = \begin{cases} \frac{\omega_j}{2}\frac{i}{\pi} & \text{if cosine} \\ \frac{\omega_j}{2}\frac{1}{\pi} & \text{if modsine.} \end{cases}$$

Equation (4.37) formulates an explicit way to retrieve the series coefficients of the payoff function a single time step away from maturity based on the coefficients at maturity itself. Hereby, the payoff coefficients are a function g of several variables. Most importantly they

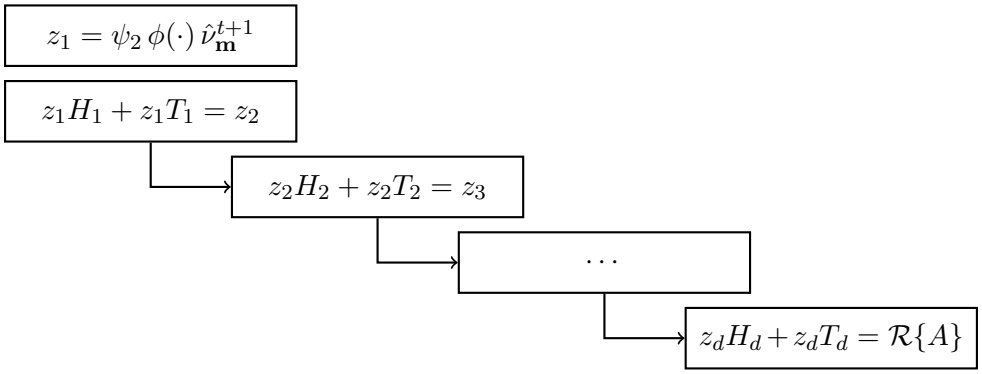

Figure 4.10: Schematic illustration of the backward reduction of the term $\mathcal{R}\{A\} = \mathcal{R}\left\{\psi_2\,\phi\left(u_{\mathbf{m}}\bar{\xi}_j\right)\hat{\nu}_{\mathbf{m}}^{t_2}\prod_{k=1}^{d}\psi_3^k(H_k+T_k)\right\}$ in equation (4.38).

depend on the payoff coefficients one time step ahead. As a consequence the overall recursive structure can be formulated in a stylized way as follows:

$$\begin{aligned}
\hat{\nu}_{\mathbf{n}}^{t_{E-2}} &= g(..., \hat{\nu}_{\mathbf{m}}^{t_{E-1}}, ...) \\
&\vdots \\
\hat{\nu}_{\mathbf{n}}^{t_z} &= g(..., \hat{\nu}_{\mathbf{m}}^{t_{z+1}}, ...) \\
&= e^{-r\Delta t_z}\,\psi_1 \sum_{\mathbf{m}}\sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{\psi_2\,\phi\left(u_{\mathbf{m}}\bar{\xi}_j\right)\hat{\nu}_{\mathbf{m}}^{t_{z+1}}\prod_{k=1}^{d}\psi_3^k(H_k+T_k)\right\} \qquad (4.38) \\
&\vdots \\
\hat{\nu}_{\mathbf{n}}^{t_1} &= g(..., \hat{\nu}_{\mathbf{m}}^{t_2}, ...)
\end{aligned}$$

In principle, equation (4.38) can be used $(E-1)$-times throughout the recursion to obtain $\hat{\nu}_{\mathbf{n}}^{t_1}$ which, plugged into

$$\nu_{t_0} = e^{-r\Delta t_1}\,\psi_1 \sum_{j=1}^{2^{d-1}} \mathcal{R}\left\{\psi_2\,\phi_{\Delta t}\big(q_{\mathbf{n}}\bar{\xi}_j\big)\,e^{i(q_{\mathbf{n}}\bar{\xi}_j)^\top(\mathbf{x}_t-\mathbf{a})}\,\hat{\nu}_{\mathbf{n}}^{t_1}\right\},$$

enables the calculation of the price of a discretely monitored barrier option at time $t = t_0$. However, the computational efficiency of (4.38) can be increased: basically, the calculations done in equation (4.38) come down to different matrix multiplications of the form $y = z_j \cdot C$, where z_j and C are d-dimensional complex matrices. Here, matrix z_j is defined by the term $z_1 = \psi_2\,\phi\left(q_{\mathbf{m}}\bar{\xi}_j\right)\hat{\nu}_{\mathbf{m}}^{t_2}$ at the beginning. In terms of equation (4.38), the d-dimensional matrix z_1 is multiplied with a product series of special matrices $z_1 \cdot \prod_{k=1}^{d}\psi_3^k(H_k+T_k)$ which yields to an updated version of z_j. Figure 4.10 depicts a scheme showing the evolution of these products more clearly.

To increase efficiency of theses calculations, we do not use straight multiplication but a fast Fourier routine that allows to compute these products in a lower number of complex computations. In Appendix 5.A, a standard procedure to calculate vector-matrix products

is reviewed that involves a so-called Toeplitz-plus-Hankel matrix. With some adjustments, this standard procedure also works on multidimensional matrix-matrix multiplications.

4.4.2 Digital barrier options

Based on the method outlined to evaluate discrete Barrier options so far, digital options once again form the basic contracts on which more complex structures are build upon in the following. Due to the fact that digital barrier options can be used to calculate probabilities of some underlying share price path hitting a barrier, they are especially suitable as a starting point in terms of the subsections to come.

Independent of the payoff at maturity, the first step to evaluating any type of barrier contract is to incorporate the barrier structure into the time varying integration domain Θ. Using a percentage value of today's asset value and keeping this value constant over time is the most common type of barrier structure. We will refer to this case as a flat barrier structure. Within the numerical part of this section we focus on two issues: First, the barrier structure and, second, the price convergency from a discrete to a continuous setting.

As mentioned before, a digital payout structure

$$\begin{aligned}\nu(\mathbf{S}_{t_E}) &= \mathbb{1}_{\{S_1(t_E)>K_1\}} \cdot \ldots \cdot \mathbb{1}_{\{S_d(t_E)>K_d\}} \\ \nu(\mathbf{x}_{t_E}) &= \mathbb{1}_{\{x_1^{t_E}>k_1\}} \cdot \ldots \cdot \mathbb{1}_{\{x_d^{t_E}>k_d\}} = \mathbb{1}_{\{\mathbf{x}^{t_E}>\mathbf{k}\}}\end{aligned}$$

is assumed. In addition to the payout structure itself, series coefficients $\hat{\nu}$ of the payout structure are crucial. Since we focus on cosine series and modified sine series, the initial matrix $\hat{\nu}_{\mathbf{n}}^{t_E}$ is defined by

$$\begin{aligned}\hat{\nu}_{\mathbf{n}}^{t_E} &= \int_{\Omega} \mathbb{1}_{\{\mathbf{x}^{t_E}>\mathbf{k}\}} \varphi_{\mathbf{n}}(\mathbf{x})\, d\mathbf{x} \\ \hat{\nu}_{\mathbf{n}}^{t_E} &= \begin{cases} \mathcal{D}^C(h_1^{t_E}, b_1^{t_E}, u_1) \otimes \mathcal{D}^C(h_2^{t_E}, b_2^{t_E}, u_2) \otimes \ldots \otimes \mathcal{D}^C(h_d^{t_E}, b_d^{t_E}, u_d) \\ \mathcal{D}^M(h_1^{t_E}, b_1^{t_E}, u_1) \otimes \mathcal{D}^M(h_2^{t_E}, b_2^{t_E}, u_2) \otimes \ldots \otimes \mathcal{D}^M(h_d^{t_E}, b_d^{t_E}, u_d) \end{cases},\end{aligned}$$

where the symbol $\otimes$ indicates Kronecker products. Moreover, the terms $\mathcal{D}^C$ and $\mathcal{D}^M$ are defined in Appendix 4.A.

With regard to the barrier structure, Figure 4.11 illustrates the base scenario for a barrier contract with two underlying assets. The left part of the figure indicates the overall integration domain Ω as well as the contract specific subset Θ at two consecutive monitoring events t_1 and t_2. The barrier level at every step can be chosen individually. Thus, implementing a flat barrier structure is just as convenient as choosing a barrier structure that allows for higher downward moves with time evolving (step-down-barrier) as well as upward sloping barriers (step-up-barrier). It is worth mentioning that, while the domain Ω is kept constant over time, Θ changes if at least one barrier level is chosen to be non-flat.

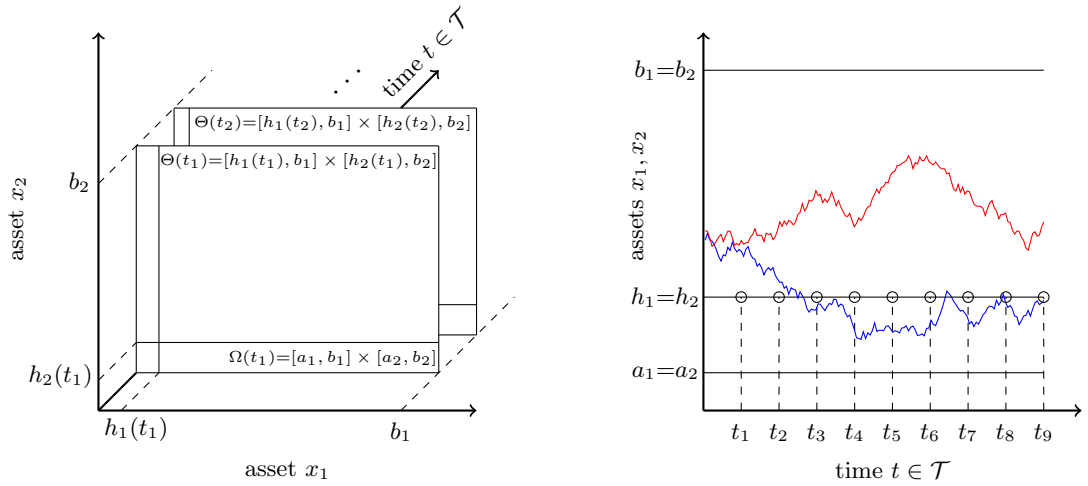

Figure 4.11: 2-dimensional barrier contract with flat barrier structure and $E = 9$ monitoring dates.

On the right-hand side of Figure 4.11, a scenario with not only a flat barrier structure but also identical barrier levels for both underlyings is displayed. Given the stock price movements in the figure (right part), a trigger event occurs at monitoring date t_3. Depending on the contract, touching the barrier either knocks-out the product as a whole or decreases payments at maturity.[39] Producing step-down or step-up types of barrier levels is equivalent to adjusting the individual layers on the left-hand side of the figure.

Having fixed a suitable barrier structure, we are especially interested in an analysis of the price convergency from European options to discretely monitored barrier options and, in the limit, to options with continuously monitored trigger levels. Option contracts with a barrier structure that is monitored not on a continuous basis but only on discretely distributed dates in time differ from their continuous counterparts. In case of a down-and-out option, the price of a discretely monitored contract shows a monotone decreasing behavior with respect to an increasing number of monitoring dates.

Based on Scenario I and Scenario II, Figure 4.12 shows this convergency. As can be seen in the figure, the differences from a discretely monitored contract with only a low number of monitoring dates to a contract which is frequently monitored is substantial. However, with monitoring dates increasing, the percentage price changes converge to zero which allows us to approximate an continuously monitored contract with this algorithm as well. It should be mentioned, however, that in case of a very high number of monitoring dates combined with a skewed underlying stochastic process, the computational time increases very fast. However, given these circumstances, a Richardson extrapolation[40] is well suited to calculate an approximation of a continuous version based on a low number of discretely monitored contracts.

[39] The method at hand is especially suitable for knock-out products. Contracts with a knock-in feature, however, can also be modeled by making use of the in-out parity: combining a down-and-out option with a down-and-in option describes a plain vanilla option.

[40] The Richardson extrapolation is a method that uses a weighted sum of a finite number of Bermudan options of different monitoring steps to approximate a continuous version. It has first been used in terms of calculating American options based on Bermudan options in Geske and Johnson (1984).

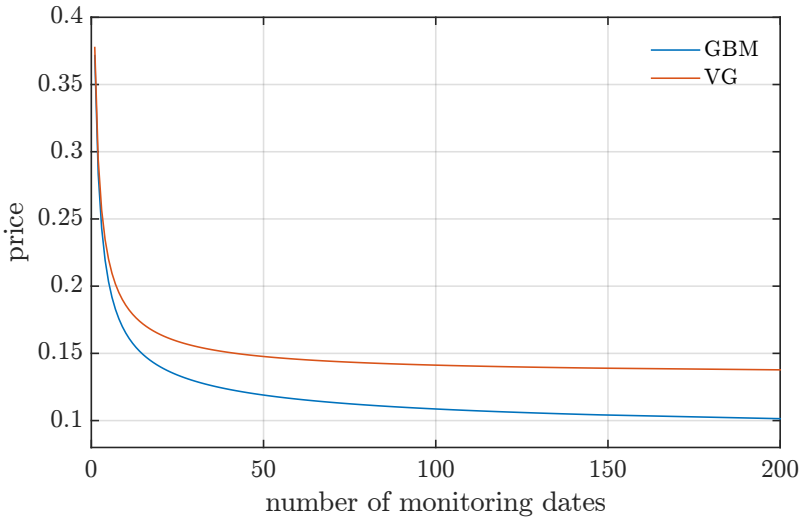

Figure 4.12: Price convergency of a digital down-and-out option based on Scenario I and Scenario II with barrier levels at 85 percent of the stock prices.

4.4.3 Multi-asset equity default swap

An equity default swap (EDS) is a financial instrument that pays a certain amount at the trigger event of the underlying stock price hitting a lower barrier. The structure is therefore similar to a credit default swap (CDS). However, while a CDS pays on the event of a default, an EDS pays on an arbitrary, pre-defined, stock price decline, e.g. 50 percent within the next year. One advantage of an EDS over a CDS can be seen in the fact that the trigger event or, respectively, the distance to this trigger event is observable through time.

In this subsection, we introduce a way to price equity default swaps that are written on a basket of underlying stock prices, called multi-asset equity default swaps (MAEDS), rather than only one underlying. For each underlying asset, a lower barrier is defined. If a stock price hits its respective barrier a predefined amount A is paid and the contract becomes worthless. The distinct difference to an ordinary barrier option lies within the swap-type structure of the contract. While option contracts demand for an option premium at the beginning of the time horizon, swap contracts usually include periodic coupon payments c which equalize the present values of a premium leg and a protection leg.

In terms of the premium leg, a percentage value c of a notional amount N is paid as long as none of the underlying stock prices S_j, $j = 1, ..., J$, hit a lower barrier H_j. The time of the first triggering event is defined by τ. Thus, the present value of a premium leg is defined by

$$\text{PV(Premium)} = N\, c \sum_{i \in \mathcal{I}} \Delta t_i \, e^{-r t_i} \, \mathbb{P}[\tau > t_i].$$

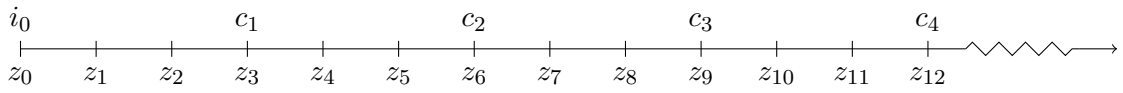

Figure 4.13: Timeline of coupon payments and monitoring days. In this example coupon payments are quarterly and monitoring dates are monthly. Given the contract matures after one year, $\mathcal{I} = \{1, ..., I = 4\}$ and $\mathcal{Z} = \{1, ..., Z = 12\}$.

Similar to Fang et al. (2010), the survival probability $\mathbb{P}[\tau > t_i]$ can be expressed as the price of a binary down-and-out option without discounting:

$$\text{PV(Premium)} = N\, c \sum_{i \in \mathcal{I}} \Delta t_i\, e^{-r t_i}\, \mathbb{E}\left[\prod_{j=1}^{J} \mathbb{1}_{\left\{ \min\limits_{0 \le s \le t_i} x_j(s) > h_j \right\}} \right]$$

Coupon payments c are made periodically, e.g. quarterly. The set $\mathcal{I}$ therefore incorporates the dates of coupon payments. A trigger event, however, can also happen in between. Thus, the set of monitoring days $\mathcal{Z}$ is at least as large as the set of payment days. In most cases, however, the set $\mathcal{Z}$ is larger as shown in Figure 4.13. Coupon payments are assumed to be paid quarterly and monitoring dates are timed on a monthly bases in the figure.

We also include an accrual on trigger A_{Tr} which is defined in the set $\mathcal{Z}$ and allows for trigger events in between two successive coupon payments:

$$A_{Tr} = N\, c\, \frac{I}{Z} \sum_{z \in \mathcal{Z}} e^{-r t_z} \left\{ \mathbb{P}[\tau > t_{z-1}] - \mathbb{P}[\tau > t_z] \right\}.$$

The resulting premium leg, including an accrual on trigger, is therefore defined by the term

$$\text{PV(Premium)} = N\, c \left(\sum_{i \in \mathcal{I}} \Delta t_i\, e^{-r t_i}\, \mathbb{P}[\tau > t_i] + \frac{I}{Z} \sum_{z \in \mathcal{Z}} e^{-r t_z} \left\{ \mathbb{P}[\tau > t_{z-1}] - \mathbb{P}[\tau > t_z] \right\} \right). \tag{4.39}$$

Next to the premium leg, the protection leg needs to be defined:

$$\text{PV(Protection)} = A \sum_{z \in \mathcal{Z}} e^{-r t_z}\, \mathbb{P}[t_{z-1} \le \tau \le t_z]$$

Once again, the probability $\mathbb{P}[t_{z-1} \le \tau \le t_z]$ can be calculated as the difference between two down-and-out digital barrier options without discounting:

$$\text{PV(Protection)} = A \sum_{z \in \mathcal{Z}} e^{-r t_z} \left\{ 1 - \mathbb{E}\left[\prod_{j=1}^{J} \mathbb{1}_{\left\{ \min\limits_{t_{z-1} \le s \le t_z} x_j(s) > h_j \right\}} \right] \right\}. \tag{4.40}$$

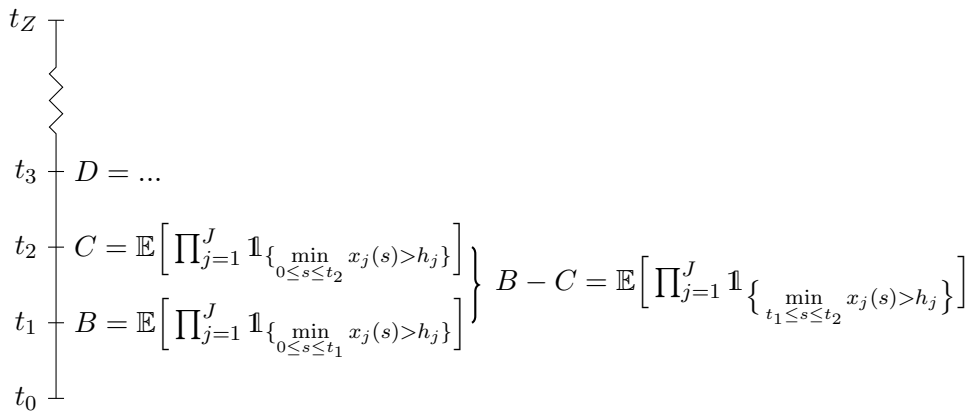

Figure 4.14: Timeline of monitoring dates with according survival probabilities.

In a last step, the coupon payments c are chosen to equalize the premium leg in (4.39) and the protection leg in (4.40):

$$c = \frac{A}{N} \frac{\sum_{z\in\mathcal{Z}} e^{-rt_z}\, \mathbb{P}[t_{z-1} \le \tau \le t_z]}{\sum_{i\in\mathcal{I}} \Delta t_i\, e^{-rt_i}\, \mathbb{P}[\tau > t_i] + \frac{I}{Z} \sum_{z\in\mathcal{Z}} e^{-rt_z} \left\{\mathbb{P}[\tau > t_{z-1}] - \mathbb{P}[\tau > t_z]\right\}}$$

Thus, given a MAEDS, an efficient way to calculate the expected values in (4.39) and (4.40) are central. Figure 4.14 elaborates on the relation of these expected values: within the figure, the terms labeled as B, C and D are d-dimensional binary down-and-out options with barrier levels at the pre-defined trigger event. The contracts only differ in their time to maturity. Thereby, contract B shows the shortest time span $t_1 - t_0$, contract C is based on a maturity $t_2 - t_0$, and so forth. With a time to maturity of the swap contract of $t_Z - t_0$, the prices of these shorter-dated binary options can be calculated on the run as a byproduct of the recursive algorithm. Thus, no noteworthy extra amount of computational time has to be spent.

In terms of the fair spread payments c, the probability term structure of the trigger events are most important. These term structures, in turn, depend on the number of monitoring dates that are incorporated within the contract. In Figure 4.15 the evolution of the probabilities of not hitting one of the triggers is shown. The less monitoring dates are present, the higher the overall curve is located in the plane. The differences between the specific paths, however, declines with an increasing number of monitoring dates. Within the figure, this becomes most obvious when comparing a contract with daily monitoring and a contract with monitoring twice a day. This convergency of the probabilities is also present in the swap rates that are calculated based on these curves as can be seen in Figure 4.16. Resulting fair swap spreads are shown on the left-hand side of the figure. As indicated by the probabilities in Figure 4.15, swap payments increase with an increasing number of monitoring dates. The curve, however converges to a limit. In addition to Scenario I, the asymmetric constellation of Scenario II is also included. Due to the presence of skewness and excess kurtosis, the overall level of the curve based on Scenario II is higher.

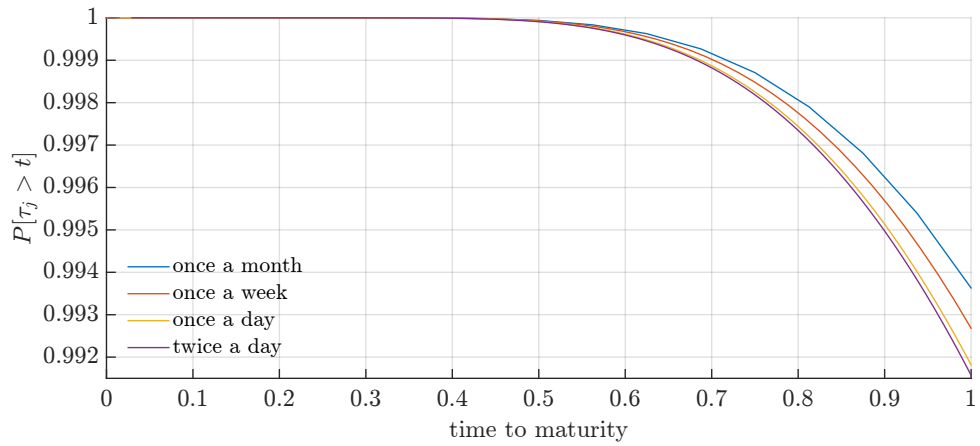

Figure 4.15: Probabilities of a decline of one of the underlying assets of less than 80 percent. Parameters as in Scenario I.

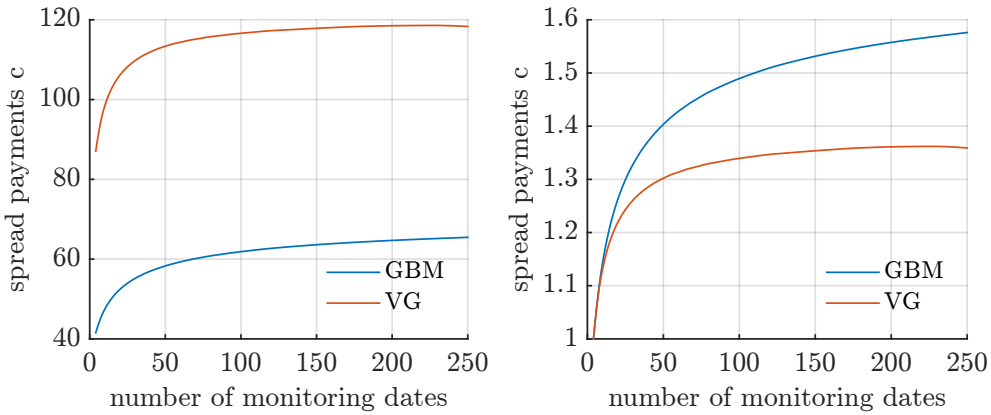

Figure 4.16: Fair spread payments of MAEDS based on varying number of monitoring dates. Parameter constellation as in Scenario I (GBM) and Scenario II (VG). Pure swap rates on the left-hand side and swap rates divided by a quarterly monitored contract on the right-hand side.

	12		52		250	
Accuracy	(N_1, N_2)	sec.	(N_1, N_2)	sec.	(N_1, N_2)	sec.
10^{-0}	27	0.006	35	0.038	50	0.327
10^{-2}	45	0.015	68	0.159	113	1.929
10^{-5}	60	0.022	87	0.190	169	2.691
Ref	48.78207704		58.5112337		65.47550658	
10^{-0}	49	0.017	130	0.320	292	12.755
10^{-2}	162	0.125	285	1.54	*	
10^{-5}	437	1.272	*		*	
Ref	100.63212282		113.45451342	47.59	111.5675509	

Table 4.4: Accuracy in basis points based on the number of terms (N_1, N_2), where $N_1 = N_2$ within the double summation. If the cells are marked by a * sign the level of precision could not be reached. This, however, is not a characteristic of the model but a lack of RAM on the computer that calculated the values.

To better compare not only the height of the curve but also their evolution, the right-hand side of the figure displays percentage increases in a way that each curve is standardized by a contract with four monitoring dates.

In Tabel 4.4 we summarize information about how many terms are needed to reach a given level of accuracy and pair it with computational times. Hereby, the upper part of the table is based on Scenario I and the lower part is based on Scenario II. Reference values (Ref) are obtained by a Monte Carlo simulation are included at the lower part of the respective subtable. Within the table we consider three different contract types: While the time to maturity is kept fix at $T = 1$ year, the number of monitoring dates varies from monthly (12) monitoring to weekly (52) monitoring and, finally, to daily (250) monitoring.

It is worth mentioning that spread payments are in basis points. Thus, in most cases an integer value precision is sufficient. In rare cases more than two digits are necessary. According to Table 4.4, an integer precision is reached very fast. Higher levels of accuracy come with more terms in the summation which increases the need for random access memory (RAM) plugged in into the working station.

As a supplement to Table 4.4, we visualize the price differences $\Delta_c^{N_1,N_2}$ between two consecutive prices $\Delta_c^{N_1,N_2} = c(N_1 + 1, N_2 + 1) - c(N_1, N_2)$ in Figure 4.17. Hereby, the horizontal dashed line on the upper part of the figure indicates a difference lower than $\pm 10^{-2}$. In contrast, in the lower part, a price difference an a level of $\pm 10^{-5}$ is marked. The oscillatory nature of the model with respect to the number of terms within the summation is observable in both parts of the figure.

To counteract the oscillatory behavior and therefore to avoid misspricing, a sufficiently high value for the tuple (N_1, N_2) is crucial. One possibility is to check the series coefficients for

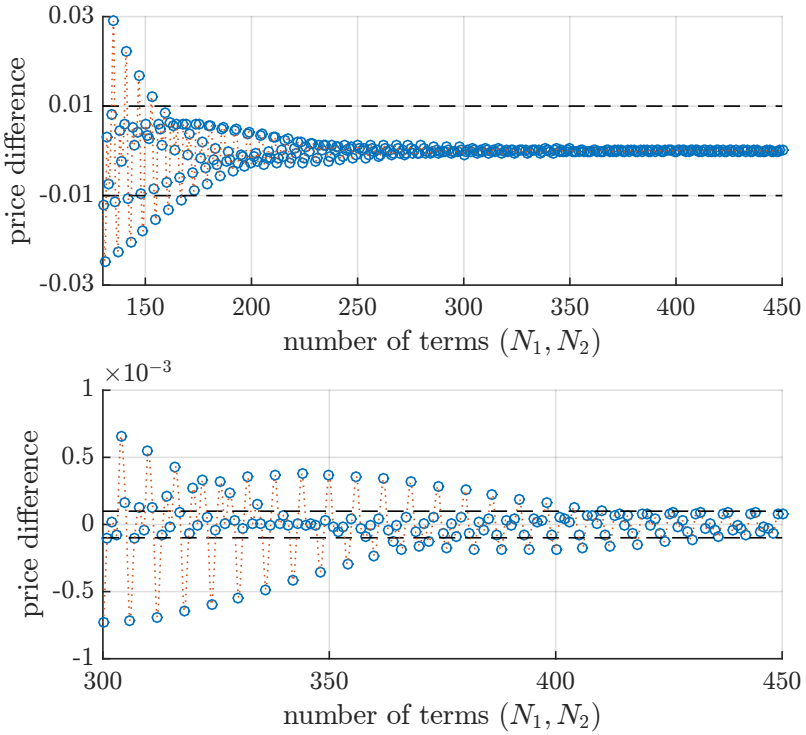

Figure 4.17: Evolution of the price differences between two consecutive prices when increasing the number of terms (N_1, N_2) by integer one $\Delta_c^{N_1,N_2} = c(N_1+1, N_2+1) - c(N_1, N_2)$. Parameter constellation as in Scenario II (VG).

convergency. We use the term

$$\sum_{n_1=N_1-10}^{N_1} \sum_{n_2=N_2-10}^{N_2} \mid \nu_{[n_1,n_2]} \mid < \epsilon,$$

with ϵ being a user defined tolerance level close to zero, to determine (N_1, N_2). If the sum of the absolute values of the coefficients is above some ϵ we increase (N_1, N_2) until the condition above is fulfilled.

4.4.4 Multi-barrier reverse convertibles

Up to this point, we only considered derivative types of contracts. Within this subsection, however, we focus on a structured product which is defined by combining some type of derivative product with another financial contract, a bond in this case. Multi-barrier reverse convertibles are products that became popular recently, foremost in the Swiss marketplace for derivatives. We choose this contract due to the fact that it's usual basket size is either two or three assets which makes it especially suitable for our algorithm at hand.

On the technical side, the contracts combines a state independent coupon payment with a state dependent repayment of the denomination: if neither of the underlyings hits a lower barrier level, the face value is paid back in full. However, if either of the assets within the

basket falls below their barrier levels, the repayment of the face value depends on the state of the worst performing underlying in terms of the quotient of the assets value at maturity over the respective initial values. If a barrier is triggered during the lifetime of the contract and either one of the assets' stock price at maturity is below the initial value, investors are paid in shares instead of cash. The overall structure of the derivative implemented within the structured product can be split in three parts: First, the fact that only the worst performing underlying determines the payoff refers to a minimum option[41]. Second, depending on a trigger event defines a barrier option. In this case the option is activated if a barrier is reached which calls for a down-and-in structure. Third, since investors are paid in shares below a certain strike value, they essentially write a put option. Thus, a multi-barrier reverse convertible includes a straight bond as well as a (shorted) down-and-in put-on-minimum option.

In other words, the payment at maturity ν_T of a multi-barrier reverse convertible is defined by a final coupon payment c and a redemption of the face value[42] N which, however, is lowered when a barrier event has occurred during the lifetime of the contract.

$$\nu_T = (1+c)\,N - \left\{ N \left[1 - \min\left(\frac{S_1(T)}{S_1(0)}, ..., \frac{S_d(T)}{S_d(0)} \right) \right]^+ \right\} \mathbb{1}_{\left\{ \min\limits_{0\le s\le T} x_j(s) \le h_j(s) \right\}} \quad \forall\, j \in \mathcal{T}$$

Since our numerical method is specialized on down-and-out barrier options, we make use of the fact that the price of a down-and-out barrier put-on-minimum option can be decomposed into a plain vanilla put-on-minimum option less a down-and-in put-on-minimum contract. This subsection, therefore combines elements from the section on path-independent options with the present section on path-dependent options.

$$\begin{aligned} \nu_T = (1+c)\,N - &\left\{ N \left[1 - \min\left(\frac{S_1(T)}{S_1(0)}, ..., \frac{S_d(T)}{S_d(0)} \right) \right]^+ \right\} + \\ &\left\{ N \left[1 - \min\left(\frac{S_1(T)}{S_1(0)}, ..., \frac{S_d(T)}{S_d(0)} \right) \right]^+ \right\} \mathbb{1}_{\left\{ \min\limits_{0\le s\le T} x_j(s) > h_j(s) \right\}} \quad \forall\, j \in \mathcal{T} \end{aligned} \tag{4.41}$$

Equation (4.41) reflects this combination by including a final payment linked to the bond contract, a European (plain vanilla) put-on-minimum option part within the first set of braces and a path dependent (down-and-out) option within the second set of braces. Discounting the payments in (4.41), yields to the following structure

$$\nu_0 = \sum_{t=1}^{T} D(0,t)\, c_t\, N + D(0,T)\, N - \nu_0^{\text{Euro}} + \nu_0^{\text{down-and-out}}, \tag{4.42}$$

[41] Options on the minimum (maximum) of at least two risky underlying assets are basket options that define their payoff by the difference between a pre-defined strike value and the asset's value performing worst (best).

[42] Within this subsection, we refer to variable N as the face value of a contract. In contrast, the upper limit of a counting variable is labeled by an additional index, e.g N_1.

where ν_0^{Euro} describes todays price of a path independent put-on-minimum option and $\nu_0^{\text{down-and-out}}$ the price of a down-and-out put-on-minimum option at $t = 0$.

As a central part of equation (4.42), the valuation of a down-and-out put-on-minimum is crucial for the pricing of the overall structure. Within the calculation of these kinds of derivatives, an efficient coding of the Fourier coefficients of the payoff function is central. Especially when trying to evaluate a structure as complicated as a put-on-minimum option, modified sine series have the clear cut advantage of needing less distinctions of cases within the closed form solution compared to a cosine series. Thus, within this subsection, we only consider a solution based on modified sine series.

In terms of the payout coefficients that are to be calculated, the following integral has to be defined:

$$\nu_T^{\text{Euro}} = \int_\Omega N \left[1 - \min\left(e^{x_1^T}, ..., e^{x_d^T}\right)\right]^+ \varphi_{\mathbf{n}}(\mathbf{x})\, d\mathbf{x}_T \tag{4.43}$$

For demonstrative purposes, we confine ourself further to a two dimensional basket. In this case equation (4.43) can be written as

$$\begin{aligned}
\hat{\nu}_T^{\text{Euro}} &= N \int_{a_2}^{b_2} \int_{a_1}^{b_1} \max\left[1 - \min\left(e^{x_1^T}, ..., e^{x_d^T}\right), 0\right] \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\
&= N \int_{a_2}^{0} \int_{a_1}^{0} \left[1 - \min\left(e^{x_1^T}, e^{x_2^T}\right)\right] \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\
&\quad + N \int_{0}^{b_2} \int_{a_1}^{0} \left(1 - e^{x_1^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T + N \int_{a_2}^{0} \int_{0}^{b_1} \left(1 - e^{x_2^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T
\end{aligned} \tag{4.44}$$

Within (4.44), the second and third double integral are straight forward. However, the first double integral containing the minimum function can be rewritten even further resulting in the following equation:

$$\begin{aligned}
\hat{\nu}_T^{\text{Euro}} &= N \int_{a_2}^{0} \int_{a_1}^{0} \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\
&\quad - N \left[\int_{a_2}^{0} \left\{ \int_{a_1}^{x_2^T} e^{x_1^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T + \int_{y_2^T}^{0} e^{x_2^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T \right\} dx_2^T \right] \\
&\quad + N \int_{0}^{b_2} \int_{a_1}^{0} \left(1 - e^{x_1^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T + N \int_{a_2}^{0} \int_{0}^{b_1} \left(1 - e^{x_2^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T
\end{aligned} \tag{4.45}$$

The solution to equation (4.45) can be expressed in closed form. Within this section, however, we only state that equation (4.45) is built on four double integrals E_1 to E_4 which can be solved individually. For individual solutions, we refer to Appendix 4.D.

Formula (4.45) only considers European type option contracts. As discussed at the beginning of Section 4.4.1, path dependent options differ from their European counterparts most obviously in the integration domain Θ. When pricing e.g. down-and-out barrier options

$[a_1, b_1] \times [a_2, b_2] \to [h_1, b_1] \times [h_2, b_2]$, where h_1 and h_2 can variate over time. Thus, to be able to also use the above statements when pricing barrier options, integration limits are changed to variables in case they are different to zero. Expressing the solution in terms of general integration limits enables us to calculate the payoff coefficients for both, European and barrier options. In general, the coefficients $\hat{\nu}_T$ can therefore be written as

$$\begin{aligned}\hat{\nu}_T &= N \int_g^0 \int_e^0 \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\ &\quad - N \left[\int_g^0 \left\{ \int_e^{x_2^T} e^{x_1^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T + \int_{y_2^T}^0 e^{x_2^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T \right\} dx_2^T \right] \\ &\quad + N \int_0^j \int_e^0 \left(1 - e^{x_1^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T + N \int_g^0 \int_0^f \left(1 - e^{x_2^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\ &= E_1(e,g) - E_2(e,g) + E_3(e,j) + E_4(f,g).\end{aligned}$$

By means of this notation, coefficients of a European type contract are given by

$$\hat{\nu}_T^{\text{Euro}} = N\Big[E_1(a_1,a_2) - E_2(a_1,a_2) + E_3(a_1,b_2) + E_4(b_1,a_2)\Big] \tag{4.46}$$

and down-and-out barrier option coefficients are described by

$$\hat{\nu}_T^{\text{down-and-out}} = N\Big[E_1(h_1,h_2) - E_2(h_1,h_2) + E_3(h_1,b_2) + E_4(b_1,h_2)\Big]. \tag{4.47}$$

Detailed information on E_1 to E_4 are provided in Appendix 4.D. Equations (4.46) and (4.47) can now be used in combination with (4.42) to evaluate a multi-barrier reverse convertible contract.

The focus of our interest, however, is not rooted in the final value of the overall contract but in the numerically demanding step of the evaluation process: The time consuming part of equation (4.42) lies within the two embedded options which are monitored on a discrete time scale. Table 4.5 displays the convergency of the model assuming weekly monitoring, parameter set as defined in Scenario II and a barrier close to the actual stock prices. Reference values (Ref) are obtained by a Monte Carlo simulation which is used as an indicator to whether the calculated value is correct on a level of 10^{-5}. However, for a level of accuracy higher than 10^{-5}, Monte Carlo methods are not a suitable method of choice. Final reference values are, thus, calculated by the Fourier series method itself with a large number of terms $(N_1, N_2) = (1000, 1000)$ within the summation.

Also the parameter environment chosen is rather challenging, the convergency behavior of the model is fast. If using a working precision of at least 10^{-4}, a number of 50 terms within the summation is sufficient.

(N_1, N_2)	European	CPU time	Down-&-in	CPU time
(50,50)	$2.8 \cdot 10^{-5}$	0.006	$6.9 \cdot 10^{-4}$	0.074
(100,100)	$5.0 \cdot 10^{-12}$	0.009	$1.5 \cdot 10^{-5}$	0.171
(150,150)	$1.0 \cdot 10^{-12}$	0.019	$3.3 \cdot 10^{-6}$	0.499
Ref	273.66772168		274.01851288	

Table 4.5: Pricing errors of put-on-minimum options based on Scenario II, barrier levels at 90% of the initial stock prices, weekly monitoring and a nominal amount of $N = 1000$.

4.5 Conclusion

Within this chapter, the evaluation of exotic options by means of series expansion methods is analyzed. Hereby we are especially interested in a method that is suitable for low dimensional basket sizes with up to five assets. Given higher basket sizes and many monitoring dates the so-called curse of dimensionality yields to a more efficient evaluation via Monte Carlo methods. However, given a low dimensionality, pricing multi-asset contracts with series expansion methods is a very efficient way. Due to the usage of characteristic functions, it also allows for a high level of flexibility in the choice of the stochastic process assumed for modeling assets behavior.

As an enhancement to the work of Ruijter and Oosterlee (2012) and Meng and Ding (2013) which are concerned with Bermuda options and European options, we lay out a method to price discrete Barrier options. To indicate the broad field of application, we implement various equity option contracts. Furthermore, with multi-asset equity default swaps, a contract which is located in between equity and credit derivatives is shown and analyzed. Pure credit derivatives such as first-to-default swaps are also realizable following similar arguments. Finally, by means of multi-barrier reverse convertibles, a structured product is priced which contains a short put-on-minimum option.

In contrast to Ruijter and Oosterlee (2012) and Meng and Ding (2013), we include three different types of basis functions within the expansion: cosine, sine and modified sine. Each of the resulting orthogonal series expansions has different characteristics in terms of the convergency behavior and approximation quality. Cosine and modified sine series share the feature of faster convergency with respect to the number of terms needed in the series expansion. In this matter, pure sine series are equally efficient at best. In case of analytic functions that are non-periodic such as call or put options, sine series show an inferior behavior compared to the other two basis functions. We illustrated this finding using exchange options based on Margrabe (1978).

A disadvantage of cosine series is the fact that the indices within the summation of the coefficients starts at $n = 0$ instead at $n = 1$ as it is the case for (modified) sine series. Due to this differences in starting values, the closed form solutions of payoff coefficients are more involved for cosine series since case differentiations are necessary. This distinction of

	Digital options	Exchange options	Barrier options
speed of convergency	$S = C = M$	$S < C = M$	$S < C = M$
computational complexity	$C < S < M$	$C < S = M$	$C < S = M$
overall	$C \approx S = M$	$S < C < M$	$S < C < M$

Table 4.6: Classification of the pricing abilities of different Fourier series: sine series (S), modified sine series (M) and cosine series (C).

cases not only exacerbates the solution itself but also demands for additional computational time. Modified sine series showed to be a convincing alternative in this context. Table 4.6 summarizes the results using the contracts analyzed within this chapter.

Based on the analysis throughout the different sections we conclude that, given the contracts considered in this chapter, pure sine series are dominated by both cosine series as well as modified sine series. This statement is true given the speed of convergency with respect to the number of terms in the summation. When also including computational time and the overall complexity of the coefficients, modified sine series are in advantage over cosine series.

Appendix: A Trigger to Rule Them All: Valuation of Multi-Asset Barrier Options

4.A Digital coefficients

d-dimensional digital options are defined by the following payoff structure at maturity:

$$\nu(\mathbf{S}_T) = \mathbb{1}_{\{S_1(T)>K_1\}} \cdot ... \cdot \mathbb{1}_{\{S_d(T)>K_d\}}$$
$$\nu(\mathbf{x}_T) = \mathbb{1}_{\{x_1^T>k_1\}} \cdot ... \cdot \mathbb{1}_{\{x_d^T>k_d\}}$$

Coefficients are therefore given by

$$\begin{aligned}\hat{\nu}_{\mathbf{n}} &= \int_{\Omega} \mathbb{1}_{\{\mathbf{x}^T>\mathbf{k}\}} \varphi_{\mathbf{n}}(\mathbf{x})\, d\mathbf{x} \\ &= \int_{\Theta} \varphi_{\mathbf{n}}(\mathbf{x})\, d\mathbf{x} \\ &= \int_{k_1}^{b_1} \varphi_{n_1}(x_1)\, dx_1 \cdot ... \cdot \int_{k_d}^{b_d} \varphi_{n_d}(x_d)\, dx_d \\ &= \prod_{j=1}^{d} \int_{k_j}^{b_j} \varphi_{n_j}(x_j)\, dx_j. \end{aligned} \tag{4.48}$$

According to equation (4.48), the coefficients are Kronecker products of the function $\mathcal{D}(k_j, b_j) = \int_{k_j}^{b_j} \varphi_{n_j}(x_j)\, dx_j$. To define the individual series coefficients, function $\mathcal{D}(k_j, b_j)$ has to be solved for each series representation $\mathcal{D} = \{\mathcal{D}^C, \mathcal{D}^S, \mathcal{D}^M\}$. In terms of cosine series, the following is true

$$\begin{aligned}\mathcal{D}^C(k_j, b_j, u_j) &= \int_{k_j}^{b_j} \cos[u_j(x_j - a_j)]\, dx_j \\ &= \begin{cases} b_j - k_j & u_j = 0 \\ \frac{\sin[u_j(a_j - k_j)]}{u_j} & u_j > 0 \end{cases}. \end{aligned}$$

When using sine series, we state that

$$\begin{aligned}\mathcal{D}^S(k_j, b_j, u_j) &= \int_{k_j}^{b_j} \sin[u_j(x_j - a_j)]\, dx_j \\ &= \frac{\cos[u_j(a_j - k_j)]) - (-1)^{n_j}}{u_j},\end{aligned}$$

while in case of modified sine series the function is given by

$$\begin{aligned}\mathcal{D}^M(k_j, b_j, p_j) &= \int_{k_j}^{b_j} \sin[p_j(x_j - a_j)]\, dx_j \\ &= \frac{\cos[p_j(a_j - k_j)]}{p_j}.\end{aligned}$$

Coefficients $\hat{\nu}_{\mathbf{n}}$ are, therefore, given by the expression

$$\hat{\nu}_{\mathbf{n}} = \mathcal{D}(k_1, b_1, \cdot) \otimes \mathcal{D}(k_2, b_2, \cdot) \otimes ... \otimes \mathcal{D}(k_k, b_k, \cdot),$$

where $\mathcal{D}$ is chosen from $\mathcal{D} = \{\mathcal{D}^C, \mathcal{D}^S, \mathcal{D}^M\}$.

4.B Margrabe coefficients

The terminal payoff of a Margrabe option is defined by

$$\begin{aligned}\hat{\nu}_{[n_1,n_2]} &= \int_\Omega \max\left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}, 0\right)\, \varphi_{[n_1,n_2]}(x_1, x_2)\, dx_1 dx_2 \\ &= \int_{a_2}^{b_2} \int_{z+x_2}^{b_1} \left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}\right)\, \varphi_{[n_1,n_2]}(x_1, x_2)\, dx_1 dx_2.\end{aligned}$$

To state closed form solutions the general term $\varphi_{[n_1,n_2]}$ needs to be filled in by either cosine, sine, or modified sine terms. Thus, we start with cosine coefficients and define sine and modified sine coefficients afterwards.

$$\hat{\nu}^C_{[n_1,n_2]} = \int_{a_2}^{b_2} \int_{z+x_2}^{b_1} \left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}\right)\, \cos(q_1(a_1 - x_2 - z)) \cos(q_2(x_2 - a_2))\, dx_1 dx_2$$

Given $n_1, n_2 > 0$, the closed form solution reads as

$$\begin{aligned}\hat{\nu}^C_{[n_1,n_2]} =& \frac{e^{b_1} S_1 \sin(A_4)\left(q_1 \sin\left(q_1\left(b_1-a_1\right)\right)+\cos\left(q_1\left(b_1-a_1\right)\right)\right)}{\left(q_1^2+1\right) q_2} \\ &+ \frac{S_2 \sin\left(q_1\left(a_1-b_1\right)\right)\left(e^{b_2}\left(q_2 \sin(A_4)+\cos(A_4)\right)-e^{a_2}\right)}{q_1\left(q_2^2+1\right)} \\ &- \frac{S_1\left(\frac{1}{2} e^{b_2+z}\left(B_4+C_4\right)-\frac{1}{2} e^{a_2+z}\left(B_2+C_2\right)\right)}{q_1^2+1} \\ &- \frac{q_1 S_1\left(\frac{1}{2} e^{b_2+z}\left(B_3+C_3\right)-\frac{1}{2} e^{a_2+z}\left(B_1+C_1\right)\right)}{q_1^2+1} \\ &+ \frac{S_2\left(\frac{1}{2} e^{b_2}\left(B_3+C_3\right)-\frac{1}{2} e^{a_2}\left(B_1+C_1\right)\right)}{q_1}\end{aligned}$$

Since a cosine expansion includes $n_1 = 0$ and $n_2 = 0$, a distinction of case is necessary. If $n_1 = 0$ it is true that

$$\begin{aligned}\hat{\nu}^C_{[0,n_2]} =& \frac{S_1\left(e^{a_2+z} q_2+\left(q_2^2 \sin(A_3)\left(e^{b_1}-e^{b_2+z}\right)-q_2 \cos(A_3) e^{b_2+z}+e^{b_1} \sin(A_3)\right)\right)}{q_2\left(q_2^2+1\right)} \\ &- \frac{S_2 e^{a_2} q_2\left(q_2^2\left(a_2+z+1\right)+a_2-b_1\left(q_2^2+1\right)+z-1\right)}{q_2\left(q_2^2+1\right)^2} \\ &+ \frac{S_2 e^{b_2} q_2\left(\cos(A_3)\left(z+b_2-1+q_2^2(z+b_2+1)\right)+\sin(A_3)\left(z+b_2-2+q_2^3(z+b_2)\right)\right)}{q_2\left(q_2^2+1\right)^2}\end{aligned}$$

Equivalently, assuming $n_2 = 0$, it is true that

$$\begin{aligned}\hat{\nu}^C_{[n_1,0]} =& \frac{q_1 S_1 e^z\left(e^{a_2}\left(\sin(A_1)-q_1 \cos(A_1)\right)+e^{b_2}\left(q_1 \cos(A_2)-\sin(A_2)\right)\right)}{\left(q_1^2+1\right)^2} \\ &- \frac{S_1 e^z\left(e^{b_2}\left(q_1 \sin(A_2)+\cos(A_2)\right)-e^{a_2}\left(q_1 \sin(A_1)+\cos(A_1)\right)\right)}{\left(q_1^2+1\right)^2} \\ &+ \frac{S_2\left(e^{a_2}\left(q_1 \cos(A_1)-\sin(A_1)\right)+e^{b_2}\left(\sin(A_2)-q_1 \cos(A_2)\right)\right)}{q_1^3+q_1} \\ &+ \frac{e^{b_1} q_1 S_1\left(b_2-a_2\right) \sin\left(q_1\left(b_1-a_1\right)\right)}{q_1^2+1}+\frac{S_2\left(e^{b_2}-e^{a_2}\right) \sin\left(q_1\left(a_1-b_1\right)\right)}{q_1} \\ &+ \frac{e^{b_1} S_1\left(b_2-a_2\right) \cos\left(q_1\left(b_1-a_1\right)\right)}{q_1^2+1}\end{aligned}$$

and, given $n_1 = n_2 = 0$, coefficients read as

$$\begin{aligned}\hat{\nu}^C_{[0,0]} =& S_1\left(e^z\left(e^{a_2}-e^{b_2}\right)+e^{b_1}\left(b_2-a_2\right)\right) \\ &+ S_2\left(e^{b_2}\left(b_2-b_1+z-1\right)-e^{a_2}\left(a_2-b_1+z-1\right)\right)\end{aligned}$$

Finally, using the known abbreviation $M_{[n_1,n_2]}(x_1,x_2) = \sin(q_1(x_1-a_1))\ \sin(q_2(x_2-a_2))$ and let $[q_1,q_2] = \left[\frac{(n_1-\frac{1}{2})\pi}{b_1-a_1}, \frac{(n_2-\frac{1}{2})\pi}{b_2-a_2}\right]$, modified sine series coefficients are given by:

$$\hat{\nu}_{[n_1,n_2]} = \int_{a_2}^{b_2}\int_{z+x_2}^{b_1} \left(S_1(0)\, e^{x_1} - S_2(0)\, e^{x_2}\right)\, M_{[n_1,n_2]}(x_1,x_2)\, dx_1 dx_2$$

$$\begin{aligned}
\hat{\nu}_{[n_1,n_2]} =& \frac{e^{b_1} S_1 \left(\cos\left(A_4\right)-1\right)\left(q_1\cos\left(q_1\left(b_1-a_1\right)\right)-\sin\left(q_1\left(b_1-a_1\right)\right)\right)}{\left(q_1^2+1\right)q_2} \\
&+ \frac{S_2\cos\left(q_1\left(b_1-a_1\right)\right)\left(q_2\left(e^{a_2}-e^{b_2}\cos\left(A_4\right)\right)+e^{b_2}\sin\left(A_4\right)\right)}{q_1\left(q_2^2+1\right)} \\
&- \frac{S_1\left(\frac{1}{2}e^{b_2+z}\left(C_4-B_4\right)-\frac{1}{2}e^{a_2+z}\left(C_2-B_2\right)\right)}{q_1^2+1} \\
&+ \frac{q_1 S_1\left(\frac{1}{2}e^{b_2+z}\left(B_3-C_3\right)-\frac{1}{2}e^{a_2+z}\left(B_1-C_1\right)\right)}{q_1^2+1} \\
&- \frac{S_2\left(\frac{1}{2}e^{b_2}\left(B_3-C_3\right)-\frac{1}{2}e^{a_2}\left(B_1-C_1\right)\right)}{q_1}
\end{aligned}$$

with

$$\begin{aligned}
A_1 &= q_1\left(-a_1+a_2+z\right) & B_1 &= \frac{-\cos\left(A_1\right)\left(q_1+q_2\right)+\sin\left(A_1\right)}{\left(q_1+q_2\right)^2+1} \\
A_2 &= q_1\left(-a_1+b_2+z\right)+q_2\left(a_2-b_2\right) & B_2 &= \frac{\sin\left(A_1\right)\left(q_1+q_2\right)+\cos\left(A_1\right)}{\left(q_1+q_2\right)^2+1} \\
A_3 &= q_1\left(-a_1+b_2+z\right)+q_2\left(b_2-a_2\right) & B_3 &= \frac{-\cos\left(A_3\right)\left(q_1+q_2\right)-\sin\left(A_3\right)}{\left(q_1+q_2\right)^2+1} \\
A_4 &= q_2\left(b_2-a_2\right) & B_4 &= \frac{\sin\left(A_3\right)\left(q_1+q_2\right)+\cos\left(A_3\right)}{\left(q_1+q_2\right)^2+1}
\end{aligned}$$

$$\begin{aligned}
C_1 &= \frac{-\cos\left(A_1\right)\left(q_1-q_2\right)+\sin\left(A_1\right)}{\left(q_1-q_2\right)^2+1} \\
C_2 &= \frac{\sin\left(A_1\right)\left(q_1-q_2\right)+\cos\left(A_1\right)}{\left(q_1-q_2\right)^2+1} \\
C_3 &= \frac{-\cos\left(A_2\right)\left(q_1-q_2\right)+\sin\left(A_2\right)}{\left(q_1-q_2\right)^2+1} \\
C_4 &= \frac{\sin\left(A_2\right)\left(q_1-q_2\right)+\cos\left(A_2\right)}{\left(q_1-q_2\right)^2+1}
\end{aligned}$$

4.C Vector matrix multiplication

A key ingredient to the computation of discrete path dependent options in general is the efficient calculation of matrix products. The usage of the FFT algorithm is a standard procedure that can be implemented in a wide range of applications. In Fang and Oosterlee (2009) the FFT algorithm is used to calculate one dimensional options based on cosine series expansions. In Ruijter and Oosterlee (2012), two dimensional options are priced with the

help of FFT. Here we describe a d-dimensional case with modified sine series as underlying expansion method.

As already shown in Section 4.4, the starting point is given by the following integral:

$$\int_{\Omega_j} e^{iv(x-a)}\,\varphi_n(x)\,dx = \int_c^d e^{iv(x-a)}\,\sin\left(u(x-a)\right)\,dx,$$

where u and v are given by $u = \frac{(m-0.5)\pi}{b-a}$ and $v = \frac{(n-0.5)\pi}{b-a}$. Solving the above integral yields the expression

$$\begin{cases} \frac{-e^{i(c-a)}[u\cos(u(a-c))+iv\sin(u(a-c))]+e^{i(d-a)}[u\cos(u(a-d))+iv\sin(u(a-d))]}{(v-u)(u+v)} & u \neq v \\ \frac{1}{2}(d-c)i & u = v \end{cases}$$

Considering the first case $u \neq v$, rearranging terms and replacing trigonometric with complex exponential functions by means of $\cos(z) = \frac{1}{2}e^{iz} + \frac{1}{2}e^{-iz}$ and $\sin(z) = \frac{1}{2}e^{-iz} - \frac{1}{2}e^{iz}$, changes the above expression to

$$\frac{1}{2}\left\{\frac{e^{i(u+v)(c-a)} - e^{i(u+v)(d-a)}}{u+v} + \frac{e^{i(v-u)(d-a)} - e^{i(v-u)(c-a)}}{v-u}\right\} \tag{4.49}$$

which already shows two structured matrices: a Hankel and a Toeplitz matrix. A Hankel matrix H is a square matrix with constant positive sloping diagonals. A Toeplitz matrix T shows a similar structure. But instead of positive sloping diagonals being constant, negative sloping diagonals are constant. Inserting u and v unveils the sum of a Hankel and a Toeplitz matrix in more detail.

$$\frac{b-a}{2\pi}\left\{\frac{e^{\frac{i\pi}{b-a}(m+n-1)(c-a)} - e^{\frac{i\pi}{b-a}(m+n-1)(d-a)}}{m+n-1} + \frac{e^{\frac{i\pi}{b-a}(n-m)(d-a)} - e^{\frac{i\pi}{b-a}(n-m)(c-a)}}{n-m}\right\}$$

Here, each entry within matrix H and T only depend on an integer value $j = m+n-1$ and $k = n-m$. Thus the above integral can be represented by

$$\frac{b-a}{2\pi}\left\{\frac{e^{i\pi j\frac{c-a}{b-a}} - e^{i\pi j\frac{d-a}{b-a}}}{j} + \frac{e^{i\pi k\frac{d-a}{b-a}} - e^{i\pi k\frac{c-a}{b-a}}}{k}\right\} = \frac{b-a}{2\pi}\Big\{h_j + t_k\Big\}, \tag{4.50}$$

where

$$h_j = \frac{e^{i\pi j\frac{c-a}{b-a}} - e^{i\pi j\frac{d-a}{b-a}}}{j} \quad \forall j$$

$$t_k = \begin{cases} \frac{e^{i\pi k\frac{d-a}{b-a}} - e^{i\pi k\frac{c-a}{b-a}}}{k} & k > 0 \\ \frac{d-c}{b-a}\,i\pi & k = 0 \end{cases}.$$

In terms of equation (4.50), the Hankel and Toeplitz matrix can be displayed as:

$$H = \begin{bmatrix} h_1 & h_2 & \cdots & h_{N-1} & h_N \\ h_2 & h_3 & \cdots & h_N & h_{N+1} \\ h_3 & h_4 & \cdots & h_{N+1} & h_{N+2} \\ \vdots & \iddots & & & \\ h_N & h_{N+1} & \cdots & & h_{2N-1} \end{bmatrix} \quad T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{2-N} & t_{1-N} \\ t_1 & t_0 & \cdots & t_{3-N} & t_{2-N} \\ t_2 & t_1 & \cdots & t_{4-N} & t_{3-N} \\ \vdots & \ddots & & & \\ h_{N-1} & h_{N+1} & \cdots & t_1 & t_0 \end{bmatrix}$$

When embedded into a circulant matrix, only a fraction of the calculations are needed compared to straight matrix multiplication. A product of the form

$$y = z \cdot H \quad z \in \mathbb{C}^N,\, H \in \mathbb{C}^{N \times N}$$

involving a Hankel matrix can be calculated by means of

$$\tilde{y} = \mathbb{F}^{-1}\left(\mathbb{F}(z) \cdot \mathbb{F}(c_h)\right),$$

where the vectors $z \in \mathbb{C}^{2N}$ and $c_h \in \mathbb{C}^{2N}$ are defined by

$$z = [z_1\ z_2\ ...\ z_N\ 0\ 0\ ...\ 0]$$
$$c_h = [h_N\ h_{N-1}\ ...\ h_1\ 0\ h_{2N-1}\ ...\ h_{N+1}].$$

and the resulting vector y is given by the final N entries of $\tilde{y}$ in reverse order

$$y = [\tilde{y}_N\ \tilde{y}_{N-1}\ ...\ \tilde{y}_1].$$

Calculations involving a Toeplitz matrix are very similar. A product

$$y = z \cdot T \quad z \in \mathbb{C}^N,\, T \in \mathbb{C}^{N \times N}$$

is efficiently solved by

$$y = \mathbb{F}^{-1}\left(\mathbb{F}(z) \cdot \mathbb{F}(c_t)\right),$$

where the vector z is defined as above and $c_t \in \mathbb{C}^{2N}$ are defined by

$$c_t = [t_0\ t_{-1}\ ...\ t_{1-N}\ 0\ t_{N-1}\ ...\ t_1].$$

The resulting vector y is given by the first N entries of $\tilde{y}$

$$y = [\tilde{y}_1\ \tilde{y}_2\ ...\ \tilde{y}_N].$$

The procedure shown is a standard algorithm for the calculation of vector matrix products. However, the calculations at hand demand for multiplications of multidimensional matrices. Using a programing environment where the implemented FFT-algorithm also allows for multidimensional inputs, the steps shown above can be equivalently used to calculate

matrix matrix products. The programming environment Matlab enables multidimensional inputs into the FFT routine. However, to save computational time, some remarks are made. We assume a three-dimensional case for the purpose of illustration. Variable z is now a multidimensional matrix of dimension $z \in \mathbb{C}^{M_1 \times M_2 \times M_3}$.

4.D Put-on-minimum coefficients

To be able to also use the following statements when pricing barrier options, integration limits are changed to variables in case they are different to zero. This is due to the fact that using a European option as a starting point, the relevant integration domain changes when pricing e.g. down-and-out barrier options $[a_1, b_1] \times [a_2, b_2] \to [h_1, b_1] \times [h_2, b_2]$.

Formula (4.45) can be subdivided into four individual expressions E_1 to E_4:

$$\begin{aligned} E_1(g,e) &= \int_g^0 \int_e^0 \varphi^M_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\ &= \frac{\cos(a_1 q_1) - \cos[q_1(a_1 - e)]}{q_1} - \frac{\cos(a_2 q_2) - \cos[q_2(a_2 - g)]}{q_2} \end{aligned}$$

$$\begin{aligned} E_2(g,e) =& \int_g^0 \left\{ \int_e^{x_2^T} e^{x_1^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T + \int_{y_2^T}^0 e^{x_2^T} \varphi_{[n_1,n_2]}(\mathbf{x})\, dx_1^T \right\} dx_2^T \\ =& \frac{e^e \left(q_1 \cos[q_1(a_1 - e)] + \sin[q_1(a_1 - e)]\right)}{1 + q_1^2} \cdot \frac{\cos[q_2(a_2 - g)] - \cos[q_2(a_2 - j)]}{q_2} \\ & - \frac{\cos(a_1 q_1)}{q_1} \left(\frac{e^g \left(q_2 \cos[q_2(a_2 - g)] + \sin[q_2(a_2 - g)]\right)}{1 + q_2^2} - \frac{q_2 \cos(a_2 q_2) + \sin(a_2 q_2)}{1 + q_2^2} \right) \\ & + \left(\frac{1}{q_1} - \frac{q_1}{1 + q_1^2} \right) \left(-\frac{1}{2} e^g (B_1 - B_2) + \frac{1}{2}(B_3 - B_4) \right) \\ & - \frac{1}{1 + q_1^2} \left(\frac{1}{2} e^g (C_1 - C_2) - \frac{1}{2}(C_3 - C_4) \right) \end{aligned}$$

with

$$\begin{aligned} A_1 &= q_1(a_1 - g) - q_2(a_2 - g), & A_2 &= q_1(a_1 - g) + q_2(a_2 - g) \\ A_3 &= q_1 a_1 - q_2 a_2, & A_4 &= q_1 a_1 + q_2 a_2 \end{aligned}$$

$$B_1 = \frac{(q_1 - q_2)\cos(A_1) + \sin(A_1)}{1 + (q_1 - q_2)^2}, \qquad C_1 = \frac{\cos(A_1) - (q_1 - q_2)\sin(A_1)}{1 + (q_1 - q_2)^2}$$
$$B_2 = \frac{(q_1 + q_2)\cos(A_2) + \sin(A_2)}{1 + (q_1 + q_2)^2}, \qquad C_2 = \frac{\cos(A_2) - (q_1 + q_2)\sin(A_2)}{1 + (q_1 + q_2)^2}$$
$$B_3 = \frac{(q_1 - q_2)\cos(A_3) + \sin(A_3)}{1 + (q_1 - q_2)^2}, \qquad C_3 = \frac{\cos(A_3) - (q_1 - q_2)\sin(A_3)}{1 + (q_1 - q_2)^2}$$
$$B_4 = \frac{(q_1 + q_2)\cos(A_4) + \sin(A_4)}{1 + (q_1 + q_2)^2}, \qquad C_4 = \frac{\cos(A_4) - (q_1 + q_2)\sin(A_4)}{1 + (q_1 + q_2)^2}$$

$$\begin{aligned} E_3(e,j) &= \int_0^j \int_e^0 \left(1 - e^{x_1^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\ &= \frac{\cos(a_2 q_2) - \cos[q_2(a_2 - j)]}{q_2} \\ &\quad \cdot \frac{-\cos(a_1 q_1) + (1 + q_1^2 - e^e q_1^2)\cos[q_1(a_1 - e)] + q_1\Big(\sin(a_1 q_1) - e^e \sin[q_1(a_1 - e)]\Big)}{q_1(1 + q_1^2)} \end{aligned}$$

$$\begin{aligned} E_4(f,g) &= \int_g^0 \int_0^f \left(1 - e^{x_2^T}\right) \varphi_{[n_1,n_2]}(\mathbf{x})\, d\mathbf{x}_T \\ &= -\frac{\cos(a_1 q_1) - \cos[q_1(a_1 - f)]}{q_1} \\ &\quad \cdot \frac{\cos(a_2 q_2) - (1 + q_2^2 - e^g q_2^2)\cos[q_2(a_2 - g)] - q_2\Big(\sin(a_2 q_2) + e^g \sin[q_2(a_2 - g)]\Big)}{q_2(1 + q_2^2)} \end{aligned}$$

Coefficients are, thus, given by

$$\hat{\nu}_{[n_1,n_2]} = N\Big[E_1(a_1,a_2) - E_2(a_1,a_2) + E_3(a_1,b_2) + E_4(b_1,a_2)\Big]$$

in terms of European options and

$$\hat{\nu}_{[n_1,n_2]} = N\Big[E_1(h_1,h_2) - E_2(h_1,h_2) + E_3(h_1,b_2) + E_4(b_1,h_2)\Big]$$

in terms of down-and-out barrier options.

Chapter 5

Yet Another Factor Model

5.1 Introduction

The market for credit risk is still in the middle of a substantial period of transition. Instruments such as synthetic collateralized debt obligations (CDO) and the models used to price and monitor these financial contracts faced a lot of criticism during the subprime crisis and its aftermath. Particularly modeling via Gaussian copula was soon spotted as the source of the crisis and branded as 'the formula that killed Wall Street' (Salmon, 2009). However, when taking a closer look and moving beyond such colorful headlines, it becomes apparent that the weaknesses of the models were known and alternatives did exist. Alongside numerous other developments, one such alternative is to change the linear dependence structure implied by a pure Gaussian approach. Detaching a factor model from a pure Gaussian environment is within the core of this chapter. To concentrate on this effect, we adopt the market standard for pricing synthetic CDO, i.e., a one-factor model in combination with Vasicek's large homogeneous portfolio approach (LHP). Being a static concept, factor models are not able to cope with dynamics of the aggregated loss process in a way that more complex products such as constant proportion debt obligations demand (Burtschell et al., 2009). However, keeping the focus on CDO pricing, factor models are an adequate way of mapping the overall structure into a model while still remaining intuitive.

Since it is well-known that the Normal distribution is not able to represent the nonlinear dependence structure that can be observed within the underlying pool of credit derivatives, we study the behavior of more flexible distributions. By changing the distributional assumptions, the dependence structure changes implicitly. Whereas in case of Gaussian random variables correlation is a complete measure of dependence, it no longer is when turning to more flexible distributions. Our analysis unveils that, predominantly, the degree of skewness s and kurtosis κ are of major importance for the model's behavior when testing it at real world data. Andersen and Sidenius (2004) assert that market loss distributions have fatter upper tails and assign only a relatively small probability to small losses compared to a

loss distribution generated by a Gaussian factor model[43]. Within our model setup, we have higher moments of the factors' distribution at hand to realize this empirical evidence.

Two broader distribution families known as generalized tempered stable (GTS) and generalized hyperbolic (GH) distributions as well as their subclasses form the basis of our analysis. Within this extension of the Gaussian framework, we build upon the work of previous publications: In a first step, Hull and White (2004) used a factor model with Student's t-distributed parameters which – due to its fat-tails – is able to capture market spreads[44] more precisely. However, similar to the normal distribution, the Student's t-distribution is a symmetric function. In order to allow for both kurtosis and skewness, researchers moved on to distributions allowing for non-zero skewness. Within the class of infinitely divisible distributions, Kalemanova et al. (2007) utilize Normal Inverse Gaussian distributions (NIG) whereas Moosbrucker (2006b) installs Variance Gamma distributions (VG). Both articles linked the parameters by means of the distribution's stability under convolution property which indicates that the sum of two identically distributed random variables follows the same distribution function as the individual variables. While it has some implementation advantages, this simplification effectively implicates that one factor is entirely determined by the other. Eberlein et al. (2008) also use NIG and VG, but do not connect the parameters by means of some stability criteria. Nimmanunta et al. (2008) use Meixner distributions in a copula-based model as well as a structural credit model. Since none of these studies explicitly analyze the fitting properties of a mixed factor model approach, we apply several infinitely divisible semi-heavy-tailed distributions and a heavy-tailed distribution both in a mixed and in a straight setting.

Our study is related to that of Albrecher et al. (2007) in the sense that we are also interested in unifying the above-mentioned approaches. However, besides the fact that we use different distribution functions and allow for mixed factor assumptions, a period of time is analyzed rather than one point in time within the calibration study.

Our research question can thus be summarized as follows: What is the actual effect on theoretical tranche prices when introducing probability distributions that allow for skewness and kurtosis? Moreover, we are interested in the question whether factor models based on a mix of infinitely divisible distributions are able to represent real market conditions? Thus, in the proceeding we analyze new types of mixed factor models and thereby fill in a gap of scarce empirical evidence about the fitting properties of credit risk models. In particular, we consider 16 different setups of mixed factor models, 12 of which are pure mixed models in a sense that they assume different probability distributions for the individual components of the factor model. To our knowledge, only the factor model based on Student's t-distributed parameters has been described in literature before. All other model setups are new specifications.

Within the empirical section of this chapter, we do not only state deviations of the mixed factor models from market prices but also use two reference models to classify the perfor-

[43] We use the term 'Gaussian factor model' and 'Normal factor model' as synonyms in the following.
[44] Within this paper, we follow market conventions and use the terms 'spread', 'price' and 'quote' as synonyms.

mance: While all of the above-mentioned literature has in common that it stays within the classical factor model setup with its fixed correlation structure. Andersen and Sidenius (2004) introduced a different approach by allowing correlation to switch between regimes. For comparison purposes, we also implement this model to produce benchmark values from another class of models. Afterwards, we use the data to conduct two-tailed Wilcoxon-signed rank tests in order to analyze whether or not a model setup produces significantly different results compared to the reference model.

The remainder of this chapter proceeds as follows: Section 2 briefly introduces the theoretical basics needed to price synthetic CDO. While the algorithms used to numerically evaluate (cumulative) density functions and their inverse are presented in Section 3, the distribution functions we use in our framework and the impact of changing the underlying distributions are discussed in Section 4. A calibration study based on the tranched iTraxx Europe from 2006 to 2011 is performed in Section 5 and Section 6 concludes. An appendix is provided afterwards.

5.2 Modeling synthetic collateralized debt obligations

Synthetic collateralized debt obligations are complex instruments which convert an underlying pool of credit default swaps (CDS) through a special purpose vehicle into a tranched security. Each tranche has an attachment point K_1 and a detachment point K_2. Hereby, losses below the attachment point and above the detachment point are not to be borne by an investor holding a particular tranche.

An investor selling protection on a tranche (protection seller) receives payments from the protection buyer periodically. As soon as cumulative losses of the reference pool exceed the attachment point K_1, the protection buyer is compensated by the protection seller up to the point where cumulative losses exceed K_2. Obviously the contract is two-sided: From the protection seller's point of view, there is an incoming cash flow consisting of spread payments $c(K_1, K_2)$ (premium leg, PL) and an outgoing cash flow depending on losses (default leg, DL). Starting with the incoming cash flow, the premium leg reads as

$$\text{PL} = \sum_{i=1}^{n} c(K_1, K_2) \cdot N \cdot \Delta t_i \ (1 - \mathbb{E}[L_{t_i}^{K_1,K_2}])D(0, t_i) + N \cdot Up,$$

where $L_{t_i}^{K_1,K_2}$ denotes the loss of tranche $[K_1, K_2]$, N is the notional, Up indicates an upfront payment which is intended to reduce counterparty credit risk the protection seller is exposed to and $D(0, t_i)$ is a discounting factor.

Within the chapter's framework, the default leg is defined as

$$\begin{aligned} \text{DL} &= \int_0^T N \cdot D(0,t)\ dL_t^{K_1,K_2} \\ &\approx \sum_{i=1}^{n} (\mathbb{E}[L_{t_i}^{K_1,K_2}] - \mathbb{E}[L_{t_{i-1}}^{K_1,K_2}]) N \cdot D(0,t_i). \end{aligned} \tag{5.1}$$

With this kind of default leg, it is assumed that default payments are made solely at times of spread payments. This simplification can be relaxed by including an accrual on default. However, given the chapter's purpose of comparing different model setups, we stay within the version of formula (5.1).

In case of zero upfront, the fair price of a tranche is defined as the spread payment

$$c(K_1,K_2) = \frac{\sum_{i=1}^{n} \left[\mathbb{E}[L_{t_i}^{K_1,K_2}] - \mathbb{E}[L_{t_{i-1}}^{K_1,K_2}]\right]\ D(0,t_i)}{\sum_{i=1}^{n} \left[1 - \mathbb{E}[L_{t_i}^{K_1,K_2}]\right]\ D(0,t_i)\ \Delta t_i}, \tag{5.2}$$

equalizing the premium leg and the default leg. In case of a nonzero upfront, spread payments are fixed at a certain level and the variable Up is determined by

$$Up = \sum_{i=1}^{n} (\mathbb{E}[L_{t_i}^{K_1,K_2}] - \mathbb{E}[L_{t_{i-1}}^{K_1,K_2}]) D(0,t_i) - \sum_{i=1}^{n} c(K_1,K_2)\ \Delta t_i\ (1 - \mathbb{E}[L_{t_i}^{K_1,K_2}]) D(0,t_i). \tag{5.3}$$

The critical unknown in equations (5.2) and (5.3) are the expected values $\mathbb{E}[L_{t_i}^{K_1,K_2}]$ over time. An intuitive way to think of the expected loss is to treat it as a call spread option written on the underlying portfolio losses. Therefore, it can be expressed by

$$\begin{aligned} \mathbb{E}[L_{t_i}^{K_1,K_2}] &= \frac{1}{K_2 - K_1} \left\{ \int_{K_1}^{1} (u - K_1) d\xi(t,u) - \int_{K_2}^{1} (u - K_2) d\xi(t,u) \right\} \\ &= \frac{1}{K_2 - K_1} \int_{K_2}^{K_1} (u - K_1) d\xi(t,u) + 1 - \xi(t,K_2) \end{aligned} \tag{5.4}$$

where $\xi(t,x)$ describes the continuous portfolio loss distribution. Equation (5.4) outlines that the challenge in modeling synthetic CDO contracts is to describe $\xi(t,x)$ in accordance with market data.

To be able to deduce a portfolio loss distribution, we first have to cope with dependence. Therefore, it has to be clarified how the portfolio's single CDS behave on an individual basis, and, in addition, how they interact with other entities. More precisely, we are interested in the interaction of default times τ_i of the items within the credit pool. First, marginal distributions $q_i(t) = \mathbb{P}[\tau_i \leq t]$ deal with the behavior of a single entity and can be bootstrapped either from an associated CDS quote or from a portfolio index level. Second, supposing an underlying portfolio of $\mathcal{P} = \{1, \ldots, N\}$ different CDS, a homogeneous one-factor model

$$X_i = aY + \sqrt{1 - a^2}\epsilon_i \qquad \forall i \in \mathcal{P} \tag{5.5}$$

can be used to connect a system-wide factor Y with an idiosyncratic factor ϵ_i by means of variable a which is associated to linear dependence in this context.[45] Thus, a homogeneous one-factor model assumes a flat correlation structure with identical entries within the correlation matrix.[46] Within this setting, a reference entity X_i defaults when crossing or touching a lower barrier B_i. Therefore, the individual probability of default is given by

$$q_i(t) = \mathbb{P}\left[X_i \leq B_i(t)\right] = F_X[B_i(t)]. \tag{5.6}$$

For equation (5.6) to hold, it must be true that $B_i(t) = F_X^{-1}\left[q_i(t)\right]$. Since we assume a homogeneous case, the threshold $B_i(t)$ is the same for all titles within the portfolio, i.e., $B_i(t) = B(t) = F_X^{-1}\left[q(t)\right]$.

While the standard model is defined by standard normally distributed factors Y and ϵ_i, we analyze more flexible distributions with zero mean and unit variance but varying higher moments. A portfolio loss formula, however, can be achieved even without assuming specific distribution functions in equation (5.5): Conditioning on factor Y, the components of vector X_i are independent from each other. Therefore, it is true that

$$\begin{aligned}
\mathbb{P}[X_i \leq B|Y = y] &= \mathbb{P}[ay + \sqrt{1-a^2}\epsilon_i \leq B|Y = y] \\
&= \mathbb{P}\big[\epsilon_i \leq \frac{B - ay}{\sqrt{1-a^2}}|Y = y\big] \\
&= F_\epsilon\left[\frac{B(t) - ay}{\sqrt{1-a^2}}\right].
\end{aligned}$$

Due to conditional independence, the conditional portfolio loss distribution follows a binomial distribution:

$$\mathbb{P}[\mathcal{L} = k|Y = y] = \binom{N}{k} F_\epsilon\left[\frac{B(t) - ay}{\sqrt{1-a^2}}\right]^k \left(1 - F_\epsilon\left[\frac{B(t) - ay}{\sqrt{1-a^2}}\right]\right)^{N-k}. \tag{5.7}$$

When integrating equation (5.7), we obtain the unconditional loss distribution (5.8) and cumulative loss distribution (5.9), respectively, for an underlying pool consisting of exactly N different CDS:

$$\mathbb{P}[\mathcal{L} = k] = \int_{\mathbb{R}} \binom{N}{k} F_\epsilon\left[\frac{B(t) - au}{\sqrt{1-a^2}}\right]^k \left(1 - F_\epsilon\left[\frac{B(t) - au}{\sqrt{1-a^2}}\right]\right)^{N-k} dF_Y(u) \tag{5.8}$$

$$\mathbb{P}[\mathcal{L} \leq k] = \int_{\mathbb{R}} \sum_{k=0}^{\lfloor Nx \rfloor} \binom{N}{k} F_\epsilon\left[\frac{B(t) - au}{\sqrt{1-a^2}}\right]^k \left(1 - F_\epsilon\left[\frac{B(t) - au}{\sqrt{1-a^2}}\right]\right)^{N-k} dF_Y(u) \tag{5.9}$$

[45] The correlation coefficient describing linear dependence is defined by $Corr(X_i, X_j) = a^2$.

[46] For a different approach see e.g. Hager and Schöbel (2006). Within this publication, the authors use an evolutionary algorithms to fit the individual entries of the correlation matrix to a given implied correlation structure.

Following Vasicek (1987), as $N \to \infty$ the limit of the portfolio loss distribution comes down to

$$\begin{aligned}\xi(t,x) &= \int_0^1 \mathbb{1}_{\{0\le s\le x\}}\, dF_Y\left(\frac{\sqrt{1-a^2}F_\epsilon^{-1}(s) - B(t)}{a}\right) \\ &= -\int_0^x dF_Y\left(\frac{B(t) - \sqrt{1-a^2}F_\epsilon^{-1}(s)}{a}\right) \\ &= 1 - F_Y\left(\frac{B(t) - \sqrt{1-a^2}F_\epsilon^{-1}(x)}{a}\right),\end{aligned} \tag{5.10}$$

where $s = F_\epsilon\left(\frac{B(t)-ax}{\sqrt{1-a^2}}\right)$. At first sight, approximating a finite portfolio by means of an infinite portfolio may seem like a radical simplification. However, it is quite remarkable how close the approximation gets for a portfolio size starting at 100 titles (Schönbucher, 2003). In principle, equation (5.10) could be used in combination with equations (5.4) and (5.2) or (5.3) to price specific tranches. But due to the fact that evaluating the inverse distribution F_ϵ^{-1} from (5.10) in the integral of (5.4) is numerically slow, we follow Kalemanova et al. (2007) and rewrite the integral $\int_{K_2}^{K_1}(u-K_1)d\xi(t,u)$ in (5.4) to

$$\int_{K_2}^{K_1}(u-K_1)d\xi(t,u) = \frac{\sqrt{1-a^2}}{a}\int_{F_\epsilon^{-1}(K_1)}^{F_\epsilon^{-1}(K_2)}(F_\epsilon(u)-K_1)\, f_Y\left(\frac{B(t)-\sqrt{1-a^2}\,u}{a}\right)du. \tag{5.11}$$

Now, inverse distributions only have to be evaluated in the integration limits. Nonetheless, there are at least two facts that can exacerbate the calculation of (5.11). First, it may be the case that there is no known version of F_Y and F_ϵ, respectively, or second, the inverse distribution F_X^{-1} – i.e., the inverse distribution of the sum of two random variables – needed to calculate $B(t)$ is not known. However, both problems can be solved by using the concepts of characteristic functions rather than distribution functions.

5.3 Approximation algorithm

Characteristic functions are defined as Fourier transforms of density functions. Since they represent a mere transformation from real to imaginary space, characteristic functions contain exactly the same information compared to their real valued counterpart. The back and forth transformation constitute a transform pair:

$$\varphi(u) = \int_{\mathbb{R}} e^{iux} f(x)dx \tag{5.12}$$

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iux}\varphi(u)du. \tag{5.13}$$

Besides this relation between characteristic functions and their according density functions, there is also a direct connection between characteristic functions and cumulative density

functions given by Gil-Pelaez's version of the inversion theorem (Gil-Pelaez, 1951):

$$F(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{e^{iux}\varphi(-u) - e^{-iux}\varphi(u)}{2iu}.$$

Due to the fact that φ is hermitian[47], the above inversion formula can also be written as:

$$F(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{I}\left\{\varphi(u)\frac{e^{-iux}}{u}\right\} du, \tag{5.14}$$

where $i = \sqrt{-1}$ and $\mathcal{I}\{\cdot\}$ describes the imaginary part of a complex number. Equations (5.13) and (5.14) could be used to solve the first problem mentioned when there is no known closed form version of F_Y or F_ϵ. With (5.13) and (5.14), one possibility is to use numerical integration methods to obtain either the density function f or the cumulative density function F. In both cases, some implementations do exist based on the fast Fourier transform (FFT) algorithm (see e.g. Scherer et al., 2009). However, we use a Fourier cosine series approximation (COS) developed by Fang and Oosterlee (2008) for evaluation of equation (5.13). One advantage of this approximation is that, compared to direct integration of (5.13) or (5.14), COS is able to calculate distribution values for a whole vector x simultaneously. It shares this feature with an FFT implementation. On a given level of accuracy, however, COS outperforms FFT density approximation methods in the calculation of small to moderate-sized vectors since it is not subject to the Nyquist relation. In addition, we use a series representation for cumulative density functions which simplifies the calculation of inverse distributions that are otherwise rather cumbersome.

Generally speaking, the COS method utilizes the fact that a Fourier series can approximate a periodic function up to an arbitrary level of accuracy. By truncating the density function at an appropriate level, it can be considered as being quasi-periodic and the full set of Fourier series analysis can be applied. A short derivation of the cosine series used can be found in the following. For a broader overview, we refer to Tolstov (1976) as well as to Section 2.2.2.

A partial Fourier cosine series on $[a, b]$ is defined by a complete orthogonal set $e_n(x)$ and its Fourier coefficients a_n by the series expansion

$$f(x) = \sum_{n=0}^{N-1} \delta_n a_n e_n(x), \tag{5.15}$$

where δ_n is 0.5 for $n = 0$ and 1 otherwise and the orthogonal set is specified by $e_n(x) = \left\{\cos\left(n\pi\frac{x-a}{b-a}\right)\right\}_{n\in\mathbb{N}}$. The Fourier coefficients a_n are to be calculated as the inner product[48] of the function f and e_n divided by the squared norm[49] of e_n (Tolstov, 1976, p. 43):

$$a_n = \frac{\langle f, e_n\rangle}{\|e_n\|^2} = \frac{2}{b-a}\langle f, e_n\rangle. \tag{5.16}$$

[47] A function f is considered being hermitian if its complex conjugate is equal to the function itself evaluated with opposite sign: $\overline{f(x)} = f(-x)$

[48] The inner product of two continuous functions f and g is defined as $\langle f, g\rangle = \int_a^b f(x)g(x)dx$.

[49] The norm of a function f is $\|f\| = \sqrt{\langle f, f\rangle}$.

Inserting (5.16) in (5.15) yields

$$f(x) = \frac{2}{b-a} \sum_{n=0}^{N-1} \delta_n \langle f, e_n \rangle \; e_n(x).$$

Since the characteristic function can be written as $\varphi(u) = \left\langle e^{iux}, f \right\rangle_{\mathbb{R}}$ it follows that

$$\mathcal{R}\left\{ e^{-in\frac{\pi}{b-a}} \varphi\left(\frac{n\pi}{b-a}\right) \right\} = \mathcal{R}\left\{ \left\langle e^{-in\frac{\pi}{b-a}(a-x)}, f \right\rangle_{\mathbb{R}} \right\} \approx \langle f, e_n \rangle ,$$

which, in turn, implies that the Fourier cosine series representation of a density function f reads as

$$\begin{aligned} f(x) &= \frac{2}{b-a} \sum_{n=0}^{N-1} \delta_n \mathcal{R}\left\{ e^{-in\frac{\pi}{b-a}} \varphi\left(\frac{n\pi}{b-a}\right) \right\} \; e_n(x) \\ &= \frac{2}{b-a} \mathcal{R}\left\{ \sum_{n=0}^{N-1} \delta_n e^{-in\frac{\pi}{b-a}} \varphi\left(\frac{n\pi}{b-a}\right) \cos\left(n\pi \frac{x-a}{b-a}\right) \right\}, \end{aligned} \tag{5.17}$$

where $\mathcal{R}\{\cdot\}$ describes the real part of a complex number. Cumulative distribution functions can be described by means of integration:

$$F(x) = \frac{2}{b-a} \int_{\mathbb{R}} \mathcal{R}\left\{ \sum_{n=0}^{N-1} \delta_n e^{-in\frac{\pi}{b-a}} \varphi\left(\frac{n\pi}{b-a}\right) \cos\left(n\pi \frac{x-a}{b-a}\right) \right\} dx \tag{5.18}$$

$$= \frac{2}{b-a}(x-a) + \mathcal{R}\left\{ \sum_{n=1}^{N-1} \delta_n \frac{2}{\pi n} \; \varphi\left(\frac{n\pi}{b-a}\right) \; e^{in\pi\frac{-a}{b-a}} \sin\left(n\pi \frac{x-a}{b-a}\right) \right\}. \tag{5.19}$$

With these relations, semi-analytic approximation equations for density functions as well as cumulative density functions are available.

As for any approximation method based on series representation, convergence is a matter of interest. However, since we work with distributions within the class of infinitely continuous functions ($C^\infty[a,b] \in \mathbb{R}$) we note that the approximation shows geometric convergence, while algebraic convergence is reached for densities with a discontinuity in the function itself or in one of its derivatives (Fang and Oosterlee, 2008).

After having outlined our approximation scheme for both f in (5.17) and F in (5.19), evaluation of the inverse function needs to be clarified. More precisely, based on equation (5.11), the inverse distribution function of the sum of two random variables needs to be implemented. Some distribution functions are stable under convolution, meaning that the sum of two identically distributed random variables follows the same distribution function as the individual variables. This, however, is not always the case. Moreover, when the sum does not consist of identically distributed random variables, we most often do not know which probability law the resulting random variable obeys. For this issue, it is useful to notice that the characteristic function of the sum of two independent random variables is defined by the product of the respective characteristic functions of each random variable. Taking the factor

loadings into account as well, we get the following relation:

$$\varphi_X(u) = \varphi_Y(au)\ \varphi_\epsilon(\sqrt{1-a^2}u). \tag{5.20}$$

With that in mind, we are still not able to calculate inverse distributions, but we know the respective characteristic function and, in consequence, also the (cumulative) density function by means of (5.17) and (5.19). Since inverse distribution values $F^{-1}(q)$ can be calculated by solving $F(x) - q = 0$ for x, equation (5.19) and its derivative (5.17) could be used in a Newton-Raphson root searching algorithm. Even though this method is fast, the disadvantage of depending heavily on initial values made us use a more robust root finding procedure. Matlab's built-in routine *fzero* combines a robust bisection with a secant method yielding sufficient accuracy and speed.

5.4 Distribution families and factor model behavior

Since we are working with infinitely divisible distributions F on $\mathbb{R}$, the most general starting point to describe the characteristic function is through the Lévy-Khintchine formula

$$\begin{aligned}\varphi(u) &= e^{\psi(u)} \\ \psi(u) &= iu\mu - \frac{1}{2}\sigma^2u^2 + \int_{\mathbb{R}}(e^{iux} - 1 - iux\mathbb{1}_{|x|<1})\ \nu(dx),\end{aligned} \tag{5.21}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and ν being a Lévy measure following the usual conditions $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}}(1 \wedge |x|^2)\nu(dx) < \infty$. To more directly connect this rather technical statement about Lévy measures to the present problem, we note that the existence of higher moments, which will become an essential ingredient in our analysis, depends on the integrability of $\nu(dx)$. Sato (2005) shows that the n^{th} moment of a random variable only exists if $\int_{|x|\geq 1}|x|^n\,\nu(dx) < \infty$. Formula (5.21) can be understood as a decomposition of the probability law into three individual parts. First, a deterministic drift part μ, second, a volatility parameter σ and, third, a Poisson part that introduces non-Normal behavior.

The distributions used in the following calibration analysis are representatives of one of two broader distribution families: the family of generalized tempered stable distributions (GTS) (see e.g., Koponen 1995; Rosinski 2007) and the family of generalized hyperbolic distributions (GH) (see e.g. Barndorff-Nielsen, 1977). At least in the former case, taking a closer look at the Lévy density $\nu(x)$, which defines the Lévy measure $\nu(dx)$ by means of $\nu(dx) = \nu(x)dx$, leads to an intuition for the behavior of the distribution:

$$\begin{aligned}&GTS(c^-, c^+, \alpha^-, \alpha^+, \lambda^-, \lambda^+) \\ &\nu_{GTS}(dx) = \left(\frac{c^-}{|x|^{1+\alpha^-}}e^{-\lambda^-|x|}\mathbb{1}_{x<0} + \frac{c^+}{|x|^{1+\alpha^+}}e^{-\lambda^+|x|}\mathbb{1}_{x>0}\right)dx,\end{aligned} \tag{5.22}$$

with $c^- > 0$, $c^+ > 0$, $\alpha^- < 2$, $\alpha^+ < 2$, $\lambda^- > 0$, and $\lambda^+ > 0$. Equation (5.22) shows the classical structure of a stable distribution. However, to yield finite moments, the Lévy measure is tempered, or, to be more precise, the tails of the Lévy measure are exponentially damped by the factor $e^{-\lambda^{\mp}|x|}$ (Cont and Tankov, 2004, p. 110). With that in mind, λ^+ and λ^- have to be interpreted as parameters regulating the decay rate of the positive and negative side of the Lévy density $\nu(x)$ and therefore controlling the tail behavior of the resulting probability distribution. A combination of equations (5.21) and (5.22) defines the generating triplet (μ, σ, ν_{GTS}) of the distribution family and leads to the associated characteristic exponent:

$$\psi_{GTS}(u) = iu\mu - \frac{1}{2}\sigma^2 u^2 + \Gamma(-\alpha^+)\lambda_+^{\alpha^+} c^+ \left\{ \left(1 - \frac{iu}{\lambda^+}\right)^{\alpha^+} - 1 + \frac{iu\alpha^+}{\lambda^+} \right\} + \Gamma(-\alpha^-)\lambda_-^{\alpha^-} c^- \left\{ \left(1 - \frac{iu}{\lambda^-}\right)^{\alpha^-} - 1 + \frac{iu\alpha^-}{\lambda^-} \right\}, \quad \forall\ \alpha_\pm < 2 \setminus \{0, 1\}. \tag{5.23}$$

When $\sigma = 0$ in (5.23), we are dealing with a purely non-Gaussian law (Sato, 2005, p. 38). Within this class of GTS distributions, we are interested in two subclasses. A particularly well-known distribution was introduced into finance by Carr et al. (2002). The characteristic exponent of this so-called CGMY distribution can be derived from (5.23) by setting $c^- = c^+ = C$, $\alpha^+ = \alpha^- = Y$, $\lambda^- = G$, and $\lambda^+ = M$:

$$\psi_{CGMY}(u) = iu\mu + \Gamma(-Y)C\left[(M - iu)^Y - M^Y + (G - iu)^Y - G^Y\right]. \tag{5.24}$$

The CGMY distribution is a generalization of a distribution function previously defined by Madan and Seneta (1990) that is called Variance Gamma (VG). By means of the gamma function's property

$$\Gamma(-Y) = \frac{1}{Y}\Gamma(1 - Y)$$

on the one hand and l'Hospital's rule on the other hand, the characteristic exponent can be retrieved as the limiting case of (5.24) for $Y \to 0$:

$$\psi_{VG}(u) = iu\mu + C \ln\left[\frac{MG}{(M - iu)(G - iu)}\right]. \tag{5.25}$$

Equation (5.25) is only one way of defining a VG distribution. Since a representation as limiting case of a GH distribution makes normalization[50] more convenient, we get back to the VG distribution in the following.

As for generalized hyperbolic distributions, the Lévy measure does not have a form as intuitive as in (5.22) which is mainly based on the fact that it contains Bessel functions of the

[50] We normalize the probability distribution of factors Y and ϵ_i in equation (5.5) to zero mean and unit volatility to preserve the interpretation of parameter a as a measure of correlation.

first kind, $J_\lambda(x)$, and the second kind, $Y_\lambda(x)$, in the integrand (Prause, 1999, see e.g.):

$$GH(\lambda, \alpha, \beta, \delta)$$
$$\nu_{GH}(dx) = \begin{cases} \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \frac{exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y(J_\lambda^2(\delta\sqrt{2y})+Y_\lambda^2(\delta\sqrt{2y}))}\, dy + \lambda e^{-\alpha|x|}\right) dx, & \lambda \geq 0 \\ \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \frac{exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y(J_{-\lambda}^2(\delta\sqrt{2y})+Y_{-\lambda}^2(\delta\sqrt{2y}))}\, dy\right) dx & \lambda < 0 \end{cases},$$

with

$$\begin{aligned} &\delta \geq 0, |\beta| < \alpha \quad \text{if } \lambda > 0 \\ &\delta > 0, |\beta| < \alpha \quad \text{if } \lambda = 0 \\ &\delta > 0, |\beta| \leq \alpha \quad \text{if } \lambda < 0. \end{aligned}$$

In contrast to GTS distributions, GH distributions have a known density representation. However, these representations still include Bessel functions which are slow to evaluate. Contrary to density functions, characteristic functions of subclasses do not rely on these special functions.

Barndorff-Nielsen (1977) includes both, an expression for the probability density function as well as the characteristic function of GH distributions:

$$\varphi_{GH}(u) = e^{iu\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\left(\delta\sqrt{\alpha^2-(\beta+iu)^2}\right)}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}, \tag{5.26}$$

with $K_\lambda(x)$ being the modified Bessel function of the second kind. However, these Bessel functions are only present within the general expression of the characteristic function. As before, we are interested in particular subclasses rather than in the broader probability distribution family itself. Thus, in the following, we extract from the general version in (5.26) the Normal distribution, the Normal Inverse Gaussian distribution as well as the Variance Gamma and the Student's t-distribution as special cases:

Whereas the Normal distribution is obtained as a limiting case for $\delta \to \infty$ and $\frac{\delta}{\alpha} \to \sigma^2$

$$\psi_{Normal}(u) = iu\mu - \frac{1}{2}u^2\sigma^2,$$

the Normal Inverse Gaussian distribution is specified by $\lambda = -\frac{1}{2}$:

$$\psi_{NIG}(u) = iu\mu - \delta\left(\sqrt{\alpha^2-(\beta+iu)^2} - \sqrt{\alpha^2-\beta^2}\right). \tag{5.27}$$

	Tranche [%]					
	Eq	Me1	Me2	Me3	Se1	Se2
Attachment	0	3	6	9	12	22
Detachment	3	6	9	12	22	100

Table 5.1: Attachment and detachment points in percentage of cumulated losses. Attachment and detachment points are hereby chosen in accordance with standard iTraxx Europe tranches.

As mentioned before, the VG distribution can also be expressed in terms of the GH family[51] with $\lambda = \frac{\sigma^2}{\nu}$, $\alpha = \sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^4}}$, $\beta = \frac{\theta}{\sigma^2}$ and $\delta \to 0$:

$$\psi_{VG}(u) = iu\mu - \frac{1}{\nu} \log\left(1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu^2\right). \tag{5.28}$$

Finally, the Student's t-distribution can be obtained from GH distributions by $\lambda = -\frac{\nu}{2}$, $\alpha = \beta = 0$ and $\delta = \sqrt{\nu}$ (Prause, 1999, p. 4). Using this parametrization as well as the asymptotic equality $K_\lambda(\omega) = \Gamma(\lambda)\ 2^{\lambda-1}\omega^{-\lambda}$, Hurst (1995) proved the following analytic form of the characteristic function:

$$\varphi_{St-t}(u) = \frac{K_{\frac{1}{2}\nu}\left(\sqrt{\nu}\,|u|\right)\left(\sqrt{\nu}\,|u|\right)^{\frac{1}{2}\nu}}{\Gamma\left(\frac{1}{2}\nu\right) 2^{\frac{1}{2}\nu-1}}.$$

This completes the portfolio of distribution functions {CGMY, VG, NIG, St.t} that are used in our analysis. Since the CGMY distribution has one degree of freedom in excess compared to VG and NIG, we enhance VG and NIG distributions by means of a Brownian part within the calibration study (see equation (5.32) in Section 5.5). In addition to an equal number of parameters, standardization becomes more convenient thereby.

However, before turning to actual market data, we give an intuition of what it means to change the underlying distributional assumption. In essence, we focus on the effect of distributional changes on individual tranche behavior and, thereby, combine concepts from Section 2 with elements from Sections 3 and 4. To study these tranche effects, attachment and detachment pairs as shown in Table 5.1 have to be defined. The values within the table are, hereby, not chosen by random choice but according to contracts written on the iTraxx Europe index.

Effect of skewness and kurtosis on correlation smiles

To point to the core of the theoretical problem, we start the analysis with a phenomenon known as implied correlation smile: A standard factor model assumes a flat correlation structure, meaning that every tranche is priced under the assumption of a fixed correlation

[51] Differences in the final representation of the VG characteristic function are rooted in the starting points: In case we interpret a VG process or law as a time changed Brownian motion, a representation as in (5.28) is the result. Starting with the definition of a VG process as the difference of two independent gamma processes, equation (5.25) occurs (see e.g. Madan et al., 1998, p. 83). However, both representations contain the same information and can be transferred into one another.

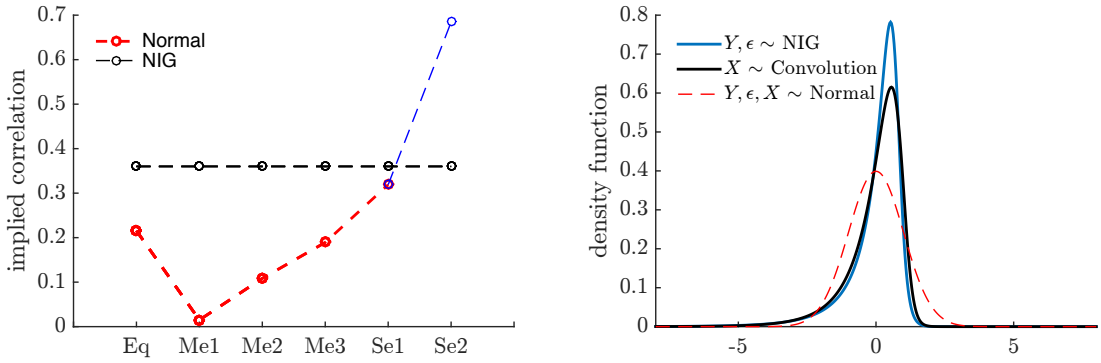

Figure 5.1: Embedded implied correlation smile of mixed models and density functions used in the model setup.

value. However, the correlation parameter varies across tranches when calibrating each tranche individually to market quotes counteracting the assumption made by a factor model. Since, in general, optimization yields higher parameter values for equity and senior tranches than for mezzanine tranches, the typical smile structure also known from implied volatilities in option markets occurs. We will come back to this point using actual market data, but for now, theoretical prices produced by a double NIG[52] model are used to study the nature of default dependence in more detail.

By construction, Figure 5.1 presents a flat correlation structure within the double NIG model at a correlation level of $a^2 = 0.36$. Using optimization to fit a Normal factor model, the smile-shape described above occurs.[53] The fact that the double NIG model produces reference prices with an embedded smile can be referred to the density functions used within the factor model setup. Within the double NIG model, factors Y and ϵ_i follow a non-Gaussian behavior and, as a result, the convolution X_i of the two random variables is non-Gaussian (and non NIG) as well. To illustrate this, densities are also plotted in Figure 5.1. Judging from the figure, it seems that mixed factor models are able to cope with smile structures but require higher levels of linear dependence compared to a Normal factor model. This observation of lower normal implied correlations can also be found in Eberlein et al. (2008) and Moosbrucker (2006a) amongst others.

Effect of skewness and kurtosis on probabilities of (joint) default

Senior tranches are protected by subordinated equity and mezzanine tranches. Therefore, they only bear losses when many credit events occur which, in turn, only happens given a high individual probability of default or a strong dependence on default. Because the ability to create strong dependence structures is crucial especially for the pricing of senior tranches, we concentrate on these tranches first before coming back to the whole tranche structure. As mentioned above, the probability mass attached to senior tranches using the Gaussian is too low. Figure 5.1 illustrates this for the second (super) senior tranche Se2. To

[52] We use the term 'double' to indicate that both factors follow a distribution stated afterwards (NIG in this case).

[53] Since the second (super) senior tranche Se2 is not traded actively, it is marked by different coloring.

reproduce the theoretical price, a correlation of almost 0.7 is needed. Therefore, in order to fit high senior tranche prices, the model must be capable of allocating a high probability to conjoint defaults, which, according to the existence of the correlation smile, cannot be done with linear dependence alone. However, the fact that market prices for senior tranches can reach very high levels became obvious in 2008. Within the market turmoil, prices for the first senior tranche of the iTraxx Europe widened from 2 bps (01/31/2007) to over 127 bps (11/25/2008). Thus, market participants evaluated the risk of conjoint defaults as being underestimated and allocated a higher probability mass to the before-mentioned tail events.

In this context, the measure of tail dependence is often used to analyze the probability of extreme (tail) events. Especially the lower tail dependence (see e.g. Albrecher et al., 2007, p. 272)

$$\lim_{x \to -\infty} \mathbb{P}\left[X_i \leq x | X_j \leq x\right] \tag{5.29}$$

is of interest for our research. However, we do not use the limit in (5.29) but rather the following conditional probability:

$$\lambda(q) = \mathbb{P}\left[X_i \leq F_X^{-1}(q) | X_j \leq F_X^{-1}(q)\right]. \tag{5.30}$$

Since in a one-factor model framework the inverse distribution function $F_X^{-1}(q)$ describes the default threshold of each entity, equation (5.30) measures the probability of item i defaulting given the fact that item j has already defaulted. The higher the above conditional probability of default gets for a given model, the wider the spread for the most senior tranche.

Keeping in mind the structure of the random variable X_i from (5.5) and using conditional independence results in a formula for the conditional probability that can be solved numerically:

$$\lambda(q) = \frac{1}{q} \int_{\mathbb{R}} \left\{ F_\epsilon \left(\frac{F_X^{-1}(q) - au}{\sqrt{1-a^2}} \right) \right\}^2 dF_Y(u). \tag{5.31}$$

In a first step we focus on the interaction of higher moments and the conditional probability of default given in (5.31). Afterwards, we focus on tranche prices rather than probabilities. We use double NIG factor models as well as a Normal factor model to conduct our analysis. While the first NIG factor model is based on symmetric but semi-heavy-tailed distributions ($Y \sim \text{NIG}(2,0)$), the second model builds upon negatively skewed distributions ($Y \sim \text{NIG}(\sqrt{3+2\sqrt{6}}, -1)$). Hereby, parameter values are chosen in a way that kurtosis ($\kappa = 3.75$) stays the same while skewness assumes negative values ($s = -0.43$). In a third model, we do not fix kurtosis but add skewness to the kurtosis level of the symmetric double NIG model ($Y \sim \text{NIG}(2,-1)$). Here, kurtosis and skewness assume values of $\kappa = 5.67$ and $s = -1$. Up to this point, factor ϵ_i is kept symmetric and follows a NIG(2,0) distribution.

Based on distributional assumptions from above, we compare conditional probabilities of default λ connected to different levels of individual probabilities of default q. The left part

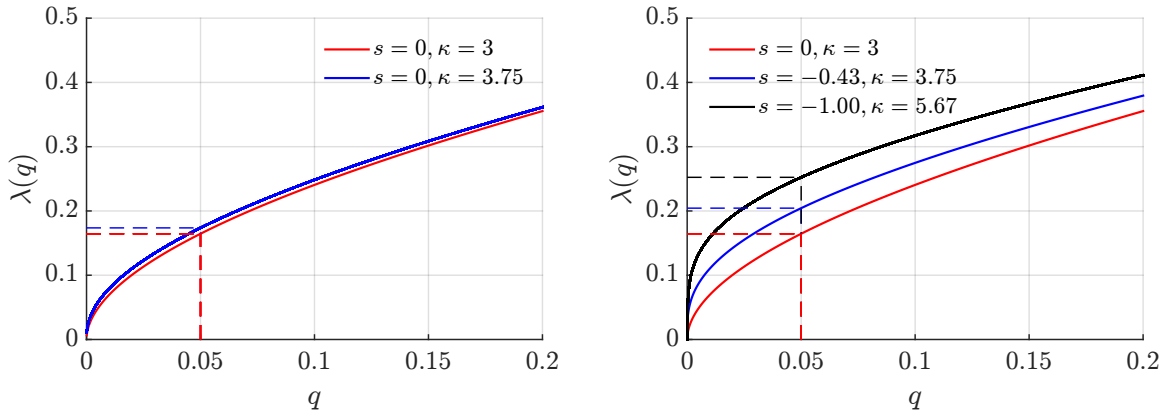

Figure 5.2: Conditional probability of default for Gaussian and double NIG factor models. Correlation factor $a^2 = 0.35$, individual probability of default $q \in (0, 0.2]$.

of Figure 5.2 illustrates a scenario where even a mild level of excess kurtosis has a positive impact on the level of λ compared to probabilities generated by a Gaussian model. Taking $q = 0.05$ as an example, the conditional probability of default is higher by a value of 92 bps. Differences in probabilities created by asymmetric models are even higher and reach values of 393 bps and 878 bps, as shown in the right part of Figure 5.2. Even though kurtosis values are the same for the two blue curves, conditional probability of default values are not. This indicates the importance of non-symmetric distribution functions.

A more direct connection to model-intern pricing behavior is the interaction between linear and non-linear parts of the dependence structure. Thus, we need to focus on correlation, skewness and kurtosis combined. To approach this issue, we compare conditional probability of default parameters $\lambda(q)$ for a given individual default probability $\bar{q} = 0.05$ and varying distribution and correlation parameters. Thus, we cut the functions from Figure 5.2 in a single point and expand that point on the correlation line. The effect of skewness and kurtosis are shown in more detail in Figure 5.3. To visualize different aspects, one of the two factor distributions is kept symmetric while the other one assumes skewness. In the upper part of the figure, the firm specific factor ϵ_i follows a symmetric NIG(2,0) distribution. In the lower part, the market wide factor Y is treated equivalently.

Introducing zero skewness but excess kurtosis in the distribution of factor Y yields higher conditional probabilities of default. The increased kurtosis, however, is subdivided into the left and right tail of the distribution. A more direct effect in terms of conditional default probabilities can be achieved through the left tail alone. When increasing negative skewness, probabilities λ increase faster compared to the scenario before. The opposite procedure, increasing positive skewness, increases the probability of survival. Factor ϵ_i behaves differently. Increasing negative skewness of this factor directly lowers conditional probabilities of default and inverse statements can be formulated for positive skewness. The reason for this can be visualized best in a bivariate distribution plot. In contrast to Figure 5.3, we do not directly aim at default scenarios in Figure 5.4. As a result, the whole spectrum becomes observable: areas where both random variables X_i and X_j survive (above and right to the

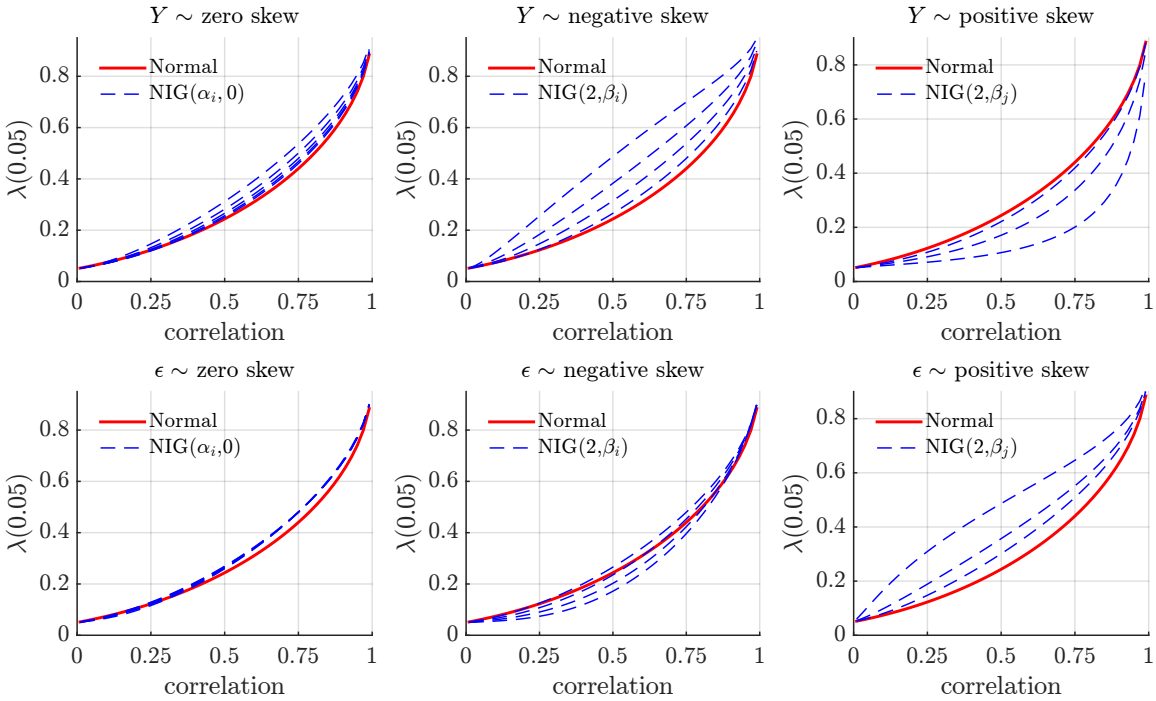

Figure 5.3: Tail dependence for Gaussian and double NIG models with varying correlation parameters. Parameters: $\alpha_i \in [0, 3]$, $\beta_i \in [-1.5, 0]$, and $\beta_j \in [0.5, 1.5]$.

dashed red lines), areas where only one title survives (below or left to the dashed red lines), as well as regions where both default (below and left to the dashed red lines).

Increasing left skewness of factor Y leads to a higher probability of conjoint default. The bivariate distribution is therefore stretched to the lower left. Factor ϵ_i, in contrast to the economic wide factor Y, deals with firm specific behavior. Increasing the negative skewness of ϵ_i yields a higher probability mass in the area of individual default but not in the region of conjoint default.

The probability of individual default is an exogenous variable that moves the default barriers (dashed red lines) in Figure 5.4. Likewise, positive skewness of f_{ϵ_i} increases individual probability of survival. However, the probability of conjoint default and dependent default respectively also increase. As a result, a combination of a negatively skewed factor Y and a positively skewed factor ϵ_i allocates the most probability mass into the area of conjoint defaults.

Increasing the probability of conjoint defaults through skewness or kurtosis is a valid instrument for increasing super senior tranche prices compared to a Gaussian scenario. The effect on the other tranches, however, is not that clear. Hence, we expand our analysis to tranche spreads themselves.

Effect of skewness and kurtosis on tranche prices

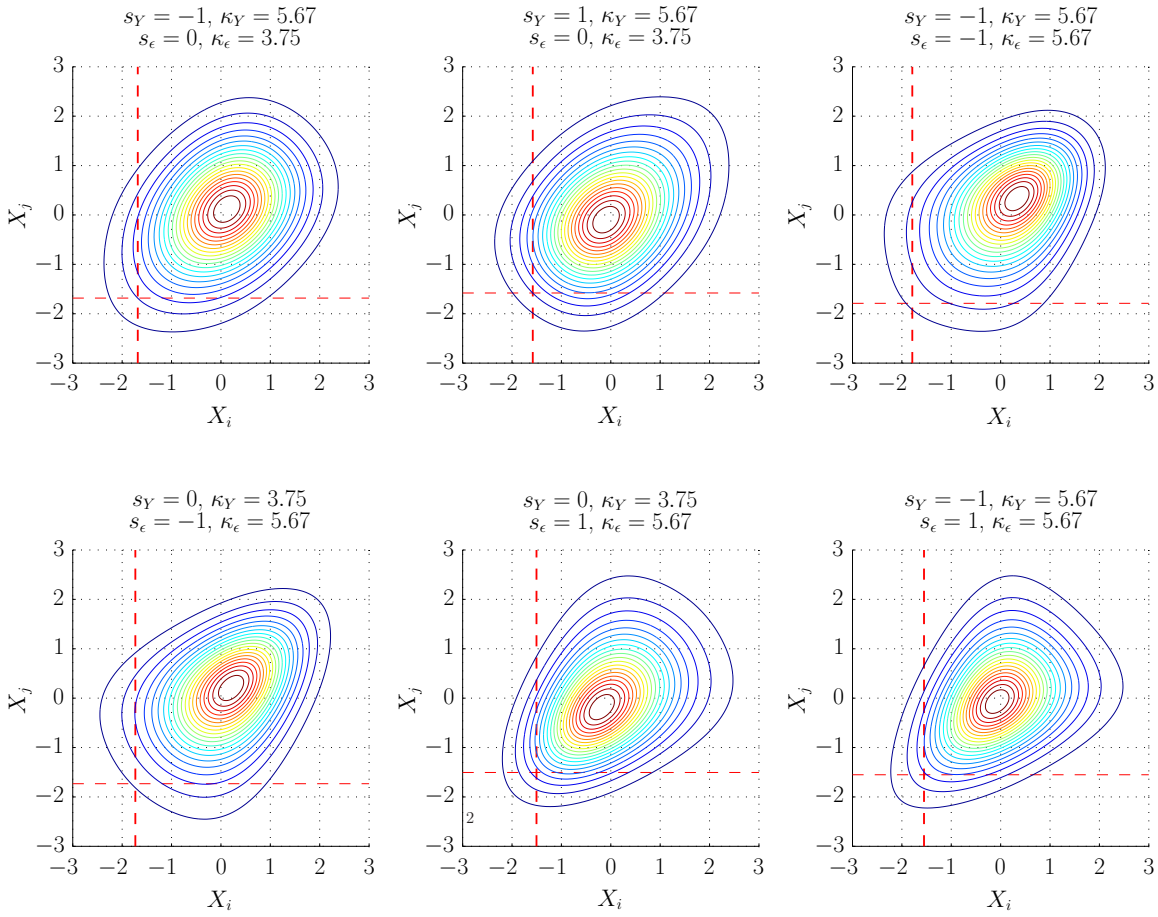

Figure 5.4: Bivariate distribution functions created by different NIG factor models based on a correlation of $a^2 = 0.35$.

In a pure Gaussian approach, the only possibility for producing a similar behavior of higher senior tranche prices is to raise correlation which, in turn, directly lowers equity payments.[54] This must not be the case for asymmetric or fat-tailed models. Tranche prices c_{NIG} produced by a double NIG model[55] are used in Table 5.2 to calculate markups of the form $M = \frac{c_{NIG}}{c_{Gaussian}}$. The respective Gaussian tranche prices[56] are reported in the table indicated by $s = 0$ and $\kappa = 3$. Since we use a Gaussian model for standardization, or more precisely a model built only on linear dependence, the following effects are based solely on non-linear dependence structures.

	s	κ	Eq	Me1	Me2	Me3	Se1	Se2	$\lambda(q)$
corr. $a^2 = 0.15$	0	3	51.01	467.54	138.54	44.53	6.40	0.01	8.8
	0	4	1.09	0.83	0.76	0.90	1.57	15.65	8.7
	0	5	1.13	0.73	0.65	0.85	1.89	34.97	8.8
	0	6	1.16	0.66	0.58	0.82	2.11	55.32	9.1
	0	7	1.18	0.61	0.54	0.80	2.28	75.48	9.4
	0	8	1.19	0.56	0.51	0.79	2.42	94.99	9.7
	0	9	1.20	0.53	0.48	0.78	2.54	113.68	10.1
corr. $a^2 = 0.35$	0	3	31.77	468.66	238.45	135.53	50.38	1.17	12.3
	0	4	1.14	0.84	0.79	0.82	0.99	2.79	17.7
	0	5	1.21	0.74	0.68	0.73	0.97	3.89	18.9
	0	6	1.26	0.67	0.61	0.67	0.96	4.75	20.0
	0	7	1.29	0.61	0.56	0.63	0.96	5.45	20.9
	0	8	1.31	0.57	0.52	0.60	0.95	6.03	21.8
	0	9	1.33	0.54	0.49	0.58	0.94	6.54	22.5
corr. $a^2 = 0.55$	0	3	16.67	404.13	255.42	177.11	94.15	5.34	27.6
	0	4	1.18	0.85	0.82	0.84	0.92	1.87	30.9
	0	5	1.28	0.76	0.72	0.75	0.88	2.23	33.0
	0	6	1.35	0.69	0.66	0.69	0.84	2.48	34.6
	0	7	1.41	0.64	0.61	0.65	0.82	2.67	35.8
	0	8	1.45	0.60	0.57	0.62	0.80	2.82	36.8
	0	9	1.48	0.57	0.55	0.59	0.79	2.94	37.6

Table 5.2: Markups of heavy-tailed models as ratio of model price and Gaussian reference price. Dependent probabilities of default $\lambda(q = 0.05)$ are given in percentages.

In Table 5.2, kurtosis is increased while still keeping distributions symmetric. As indicated above, we observe a decreasing equity upfront payment when increasing correlation. However, for a given correlation value, higher values of kurtosis yield higher markups for the equity tranche. That is, introducing kurtosis counteracts declining equity prices when raising correlation. The most pronounced markups in terms of kurtosis alone can be found for the most senior tranche (Se2). Especially for low correlation values, markups are substantial. The same applies for the first senior tranche (Se1), but only on lower levels of markups and

[54] Correlation measures joint behavior. Thus, higher correlation indicates a higher probability of joint defaults but also a higher probability of joint survival.

[55] We consult the NIG distribution as one example of a distribution function that is able to produce skewness and kurtosis. We could have used other distributions as well, the results are quite robust in the way conclusions can be drawn from it.

[56] Prices for equity tranches are quoted in percentage upfront with fixed 500 bps running spread. Prices for higher tranches are in basis points without upfront payments. As an immediate reaction to the market meltdown, higher tranches begun to trade with an upfront as well. For the analysis at this point of the paper, we, however, abstract from this convention.

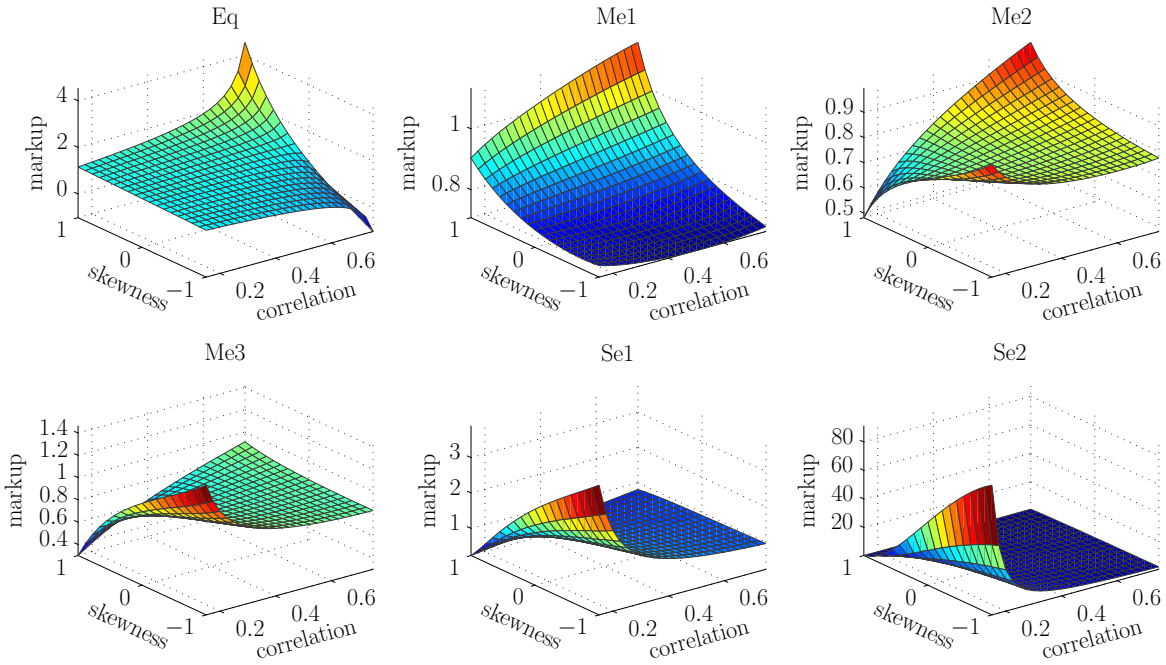

Figure 5.5: Markups of heavy-tailed and asymmetric models as ratio of model price and Gaussian reference price. Parameters: $s_Y = -1, ..., 1$, $\kappa_Y = 5.67$, $s_\epsilon = 0$, $\kappa_\epsilon = 3.75$, $a^2 = 0.15, ..., 0.75$.

low correlation parameters. This observation can be related to the question whether implied normal correlation tends to be lower. The Normal factor model prices equity tranches too high and senior tranches too low. Using more flexible models to change skewness and kurtosis yields markups in the equity tranche. In order to counteract these markups, correlation needs to be increased, which, in turn, also increases the absolute level of senior tranches. In other words, keeping equity prices on a certain level and simultaneously increasing senior tranche prices is not possible with correlation alone. However, the combination of increasing correlation and varying higher moments creates markups in the equity tranche that neutralize the effect of correlation on this tranche and raises prices of the first senior tranche (mainly due to higher correlation) and the second senior tranche (mainly due to a higher dependent probability of default). When concentrating on mezzanine tranches, an actual markdown is given.

To also incorporate non symmetric model setups, we fix kurtosis at $\kappa = 5.67$ and vary skewness from negative to positive $s \in [-1; 1]$. Even though kurtosis levels are constant, the model behaves quite differently to positive and negative skewness related to its distributions. Since interpretations are more involved for the skewed models, Figure 5.5 and 5.6 visualize and expand Table 5.2.

In Figure 5.5, the distribution function of the firm specific factor ϵ_i is kept symmetric while the distribution function of factor Y varies. Given a fixed kurtosis level, the equity tranche undergoes markups. Markdowns are only present for high correlation and negative skewness. Both senior tranches are priced higher compared to Gaussian model prices especially for low to medium correlation and negatively skewed distributions. In terms of equity and senior

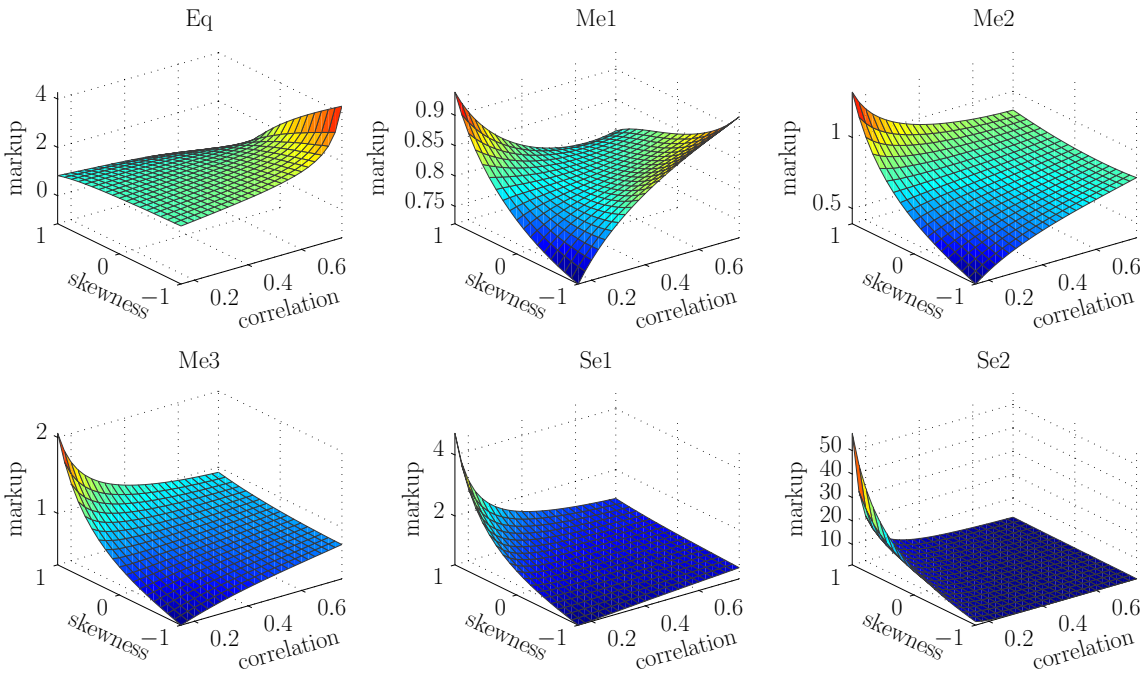

Figure 5.6: Markups of heavy-tailed and asymmetric models as ratio of model price and Gaussian reference price. Parameters: $s_Y = 0$, $\kappa_Y = 3.75$, $s_\epsilon = -1, ..., 1$, $\kappa_\epsilon = 5.67$, $a^2 = 0.15, ..., 0.75$.

tranches, effects from non-linear dependence are most pronounced on different ends of the correlation spectrum. Mezzanine tranches are generally marked down. There are, however, (s, κ, a^2) combinations where an actual markup is present which is not the case for symmetric models.

A change of perspective from factor Y to factor ϵ_i also changes the influence on tranche prices (see Figure 5.6). The model's sensitivity to changes in skewness can be described as inversely related to changes described for factor Y. As in the case of varying shapes of the distribution function f_Y, the highest markups for senior tranches are present in a low correlation environment. While at higher states of correlation linear dependence is more important for these tranches, non-linear dependence is crucial given low to medium states of correlation. The equity tranche is marked up the most. Again, the inverse structure plays a central role. Where a positively skewed system-wide factor distribution has yielded the highest markups before, negative skewness produces this outcome for the firm specific factor ϵ_i.

Robustness checks showed that changing kurtosis to different fixed levels does not change the overall pattern but only the height of the surface. Of course, Figures 5.5 and 5.6 are only special cases. An infinity of combinations can be used to create different markup and markdown patterns making mixed models much more flexible than their Gaussian counterparts. Since our analysis so far has used theoretical model parameters, in the next section, a calibration study brings the models to real world data.

	Gross Notional (billion USD)	Net Notional (billion USD)	Contracts
Index Tranched Total	2,543	453	39,833
iTraxx Tranched Total	1,040	204	14,571
iTraxx Europe S9	489	42	7,559
iTraxx Europe S15	14	3	312

Table 5.3: Outstanding positions of index tranches as of May 27, 2011. Source: Depository Trust and Clearing Corporation's (DTCC) Trade Information Warehouse.

5.5 Calibration

In order to gain some insight into the fitting performance of our model setups, we turn towards iTraxx Europe tranches. This standardized portfolio contains the 125 most liquid investment grade CDS and is rolled into a new series every March 20 and September 20. Every new series is labeled with sequential numbers. As an example, starting in September 2006, the 125 most liquid CDS are composed into the iTraxx Europe Series 6. This index was on-the-run until March 2007 where it was rolled over into Series 7. The old series, however, does not disappear. As long as a market exists, trading continues.

Maturities range from 3 to 5, 7 and 10-year contracts; in addition, recovery rates (R) are assumed to be 40 percent. It can be argued that the assumption of $R = 0.4$ has been adjusted within the financial crisis (Ascheberg et al., 2013). However, since a new market standard has not developed yet, we keep the assumption unchanged during the calibration study. Since the focus here is on comparing different model setups, this seems noncritical as long as all models are specified in the same way regarding R. Tranche attachment and detachment points are also standardized as shown in Table 5.1.

The dataset we use for our calibration study is provided by Creditex Group Inc. and Markit Group Limited.[57] It is composed of tranche quotes for 5-year contracts. To provide an overview, parts of the dataset are stated in Appendix 5.B. Starting in September 2006, monthly quotes are available through June 2007 and bimonthly spreads are available from that date. In terms of iTraxx series counts, that translates to the following: We start with Series 6 and move on to Series 7, Series 8 and Series 9. By the time the iTraxx index was rolled over to Series 10, the market for tranched iTraxx products kept the Series 9 portfolio as the most liquid one, as can be seen in Table 5.3.

Even though Series 15 was on-the-run at the end of the dataset's time period, there were more outstanding contracts based on Series 9 than on all other iTraxx Series combined. Even until the end of 2012, Series 9 stayed the top tranched iTraxx portfolio in terms of gross[58] and net[59] notional. These numbers can be interpreted in at least two ways. On the one hand, it could be argued that Series 9 tranches are the most liquid products within this market. But, at the same time, the overall liquidity must be questioned. It seems reasonable that

[57] The dataset can be viewed on the webpage http://www.creditfixings.com/CreditEventAuctions/itraxx.jsp
[58] Sum of protection bought (and sold) on a per-trade basis.
[59] Sum of protection bought (and sold) aggregated on counterparty entities.

Series 9 contracts were outstanding since netting them would have resulted in high losses at that point in time. Nevertheless, in terms of marking-to-market necessity, Series 9 tranche quotes were the most important ones. Therefore, we also use Series 9 prices from March 2008 onwards until May 2011 at which point our dataset ends.

Since we anticipate the model to fit market quotes nicely in pre-crisis times and lose performance during the more recent market situation in times of crisis, we split the dataset in two parts. While the first part contains Series 6 to Series 9, the second part includes data from the Series 9 portfolio, which was not rolled. That is, we interpret the event of not rolling-over to Series 10 as an indicator of major market distortion. Further, we use Euribor rates to calculate discounting factors $D(0,t)$. However, several robustness tests showed that results are rather insensitive concerning the choice of risk-free rates. This kind of insensitivity has been stated in literature before (see e.g. Longstaff and Rajan, 2008, p. 537). With respect to the determination of the individual probabilities of default, we use a reduced form model with an intensity parameter that follows an Inverse Gaussian Ornstein-Uhlenbeck process as recommended in Schoutens and Cariboni (2009).

Reference models

Before analyzing mixed factor models, the reference models are introduced: Figure 5.7 exemplifies portfolio loss distributions generated by three different models. Next to the double NIG factor model, which again serves as a representative model for the class of mixed models, Figure 5.7 also introduces two reference models that are used in the following. As a natural choice of benchmark, the Gaussian factor model is used. However, due to the fact that every distribution function presented in the beginning of the last section includes the normal distribution as a limiting case, they can be expected to fit market data more closely. Therefore, the Normal factor model can be seen as a lower bound of precision. The second reference model stems from a slightly different line of research. Andersen and Sidenius (2004) introduced an approach called 'random factor loadings' (RFL). Within their setup, parameters Y and ϵ_i are still distributed normally but correlation is allowed to switch between regimes. In the version we use here, there are two states the model is allowed to switch between, a low correlation environment a_L and a high correlation environment a_H. In addition, a threshold θ separates the regimes.

While the shape of a Normal factor model's portfolio loss distribution is determined solely by its correlation coefficient, the RFL and the double NIG models offer additional degrees of freedom. In case of RFL models, the regime switching nature produces loss distributions with at least one point of discontinuity or kink[60]. As for the double NIG model, a smooth portfolio loss distribution results. In contrast to an RFL model, the flexibility of this approach comes from higher moments[61].

[60] Depending on the number of different states the model can switch in between.

[61] In our example, the skewness tuple is defined by $\{s_Y, s_\epsilon\} = \{-1.33, 0.05\}$ and the kurtosis tuple by $\{\kappa_Y, \kappa_\epsilon\} = \{6.3, 3.12\}$.

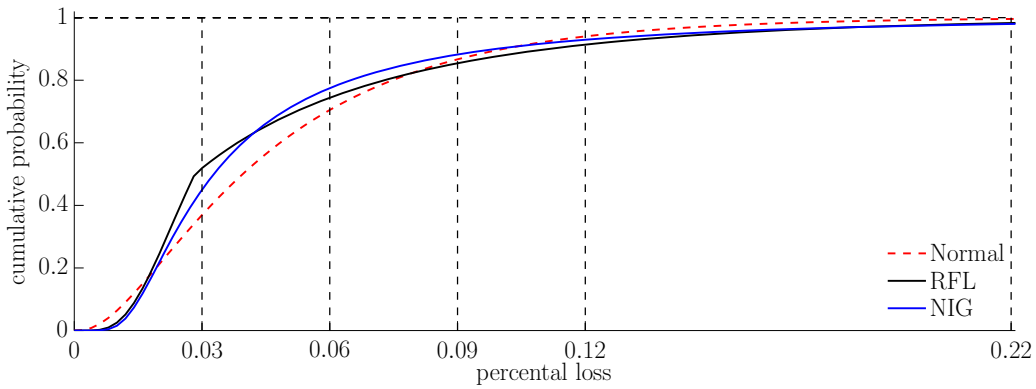

Figure 5.7: Portfolio loss distributions for Gaussian ($a = 0.35$), RFL ($a_H = 0.5$, $a_L = 0.2$, $\theta = 0$) and double NIG ($a = 0.35$, $\alpha_{1/2} = 5$, $\beta_1 = -4$, $\beta_2 = 0.4$). Individual probability of default $q = 0.05$.

To derive implicit distribution parameters $\hat{\theta}$ from market quotes, we minimize the sum of squared tranche deviations (TD) using non-linear least square methods as well as a simplex search algorithm. While for the first part Matlab's *lsqnonlin* routine is used, the second part is implemented via Matlab's *fminsearch* in combination with *fminsearchbnd*[62]. Thus the minimization problem reads as:

$$\hat{\theta} = \arg\min_{\theta} \sum_{n=1}^{5} \left(c_n(\theta) - c_n^{\text{market}} \right)^2 = \arg\min_{\theta} \sum_{n=1}^{5} \text{TD}(n)^2.$$

Minimizing the sum of squared deviations implicitly stresses the importance of the first loss tranches. An economic reasoning for this particular choice can be found in the fact that the equity tranche carries the highest risk and should therefore be priced with a maximum level of precision (for similar arguments see Eberlein et al., 2008). A different approach of defining an objective function would be to use percentage deviations.

Figure 5.8 shows cumulative absolute tranche errors $\sum_{n=1}^{5} |\text{TD}(n)|$ produced by the two reference models through time. The before-mentioned split in the dataset is marked by a dashed vertical line. With the beginning of the global financial crisis in mid-2007, absolute deviations produced by the Normal factor model start to increase sharply. As economic conditions became even more serious in 2008, this trend continues until pricing errors reach a maximum deviation of about 31 percentage points absolute deviation from market to model prices cumulated over all five tranches in May 2010. Even when considering only the first part of the dataset (left to the dashed vertical line), median cumulative absolute tranche deviations reach 59 bps.[63] While the equity tranche is priced with a relatively high level of precision – at least in the first part of the dataset – the first mezzanine tranche is usually the one with the highest deviation from market quotes when using the Normal factor model.

[62] Downloadable at Matlab Central File Exchange http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd.

[63] Judging from Figure 5.8 and keeping in mind the small sample size, it should be mentioned that the median is a more adequate measure of the average than the arithmetic mean. In the following, median expectations are abbreviated by $\mathbb{ME}(\cdot)$.

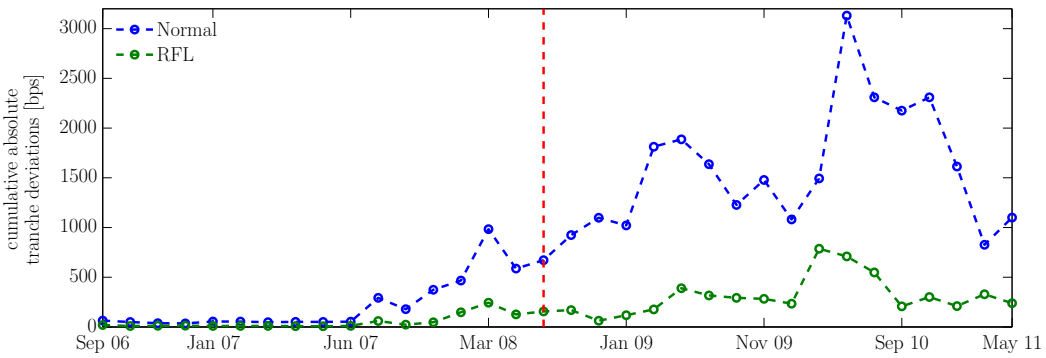

Figure 5.8: Cumulative absolute tranche mispricing of reference models.

Due to the fact that we do not calibrate to single tranches but to all of them simultaneously, the correlation parameter backed out from minimization typically differs most from implied mezzanine correlation values based on tranche-specific optimization. Therefore, the Normal factor model is not able to adequately reproduce the dependence structure with only one degree of freedom. This is one reason why market participants do not use a Normal factor model the way we implement it here. Typically, traders use the Gaussian factor model to quote implied correlations of individual tranches. Here, another analogy to option markets and the use of the Black-Scholes model appears.

The second reference model, RFL, shows a much better performance. Not only median cumulative absolute tranche deviations are lower but, moreover, the RFL model is also able to price the first mezzanine tranche quite accurately. It should be mentioned, however, that in some cases the better fit of the RFL model is based on suspicious implied parameter sets where the upper correlation parameter $\bar{a}$ hits the limit of 100 percent.

Mixed factor models: Cumulative deviations

The situation presented above is supposed to serve as a starting point from which the performance of alternative distributional assumptions is tested. To ensure that any differences in calibration performance between GH and GTS distributions are not due to the difference in free parameters, we enhance the NIG and VG characteristic function with a Brownian term. Now, as indicated in Section 5.4, the new characteristic exponents read as

$$\psi_i^*(u) = -\frac{1}{2}u^2\sigma^2 + \psi_i(u) \quad i = \text{NIG}, \text{VG}. \tag{5.32}$$

To distinguish (5.32) from (5.27) and (5.28), we label the new enhanced versions NIGJD and VGJD in dependence on the fact that a stochastic process built out of these distributions is no longer a pure jump process but a jump diffusion process. Procedures to standardize the distribution functions to zero mean and unit volatility are outlined in Appendix 5.A.

We consider 16 different mixed model setups, 12 of which are pure mixed models in the sense that they assume different distributions for factor Y and ϵ_i, respectively. To our knowledge, only the double t factor model has been described in literature before. All other models are

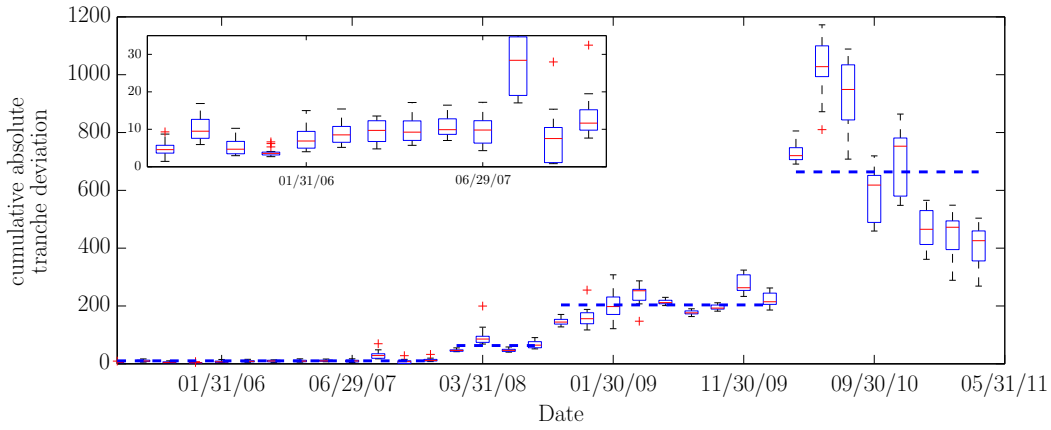

Figure 5.9: Cumulative absolute tranche mispricing of mixed models.

	TD1 [bps]				TD2 [bps]			
$F_Y \setminus F_\epsilon$	CGMY	VGJD	NIGJD	St.t	CGMY	VGJD	NIGJD	St.t
CGMY	−9.5***	−9.1***	−9.1***	−9.0***	68.0	68.4	−0.7	72.0*
VGJD	−9.4***	−10.0***	−9.2**	−6.1*	31.8	26.3	43.0	50.9*
NIGJD	−9.5***	−9.9***	−9.1**	3.4	88.5*	85.4*	101.7*	108.9**
St.t	−9.7***	−9.9***	−8.6**	−7.4**	108.4*	111.8**	124.3**	185.8**

Table 5.4: Median of cumulative absolute deviations produced by mixed models less deviations based on the RFL model. ***/**/* indicate significance at the 1%-/5%-/10%-level.

new specifications. The double NIG model has been used by Kalemanova et al. (2007) and Eberlein et al. (2008). However, our model differs from the first publication in not linking the parameters of the two factors. In addition, by enhancing the characteristic function with a Brownian part, we enlarge the set of feasible parameter combinations compared to both publications mentioned. The same arguments hold for the double VG model introduced by Moosbrucker (2006a) and Eberlein et al. (2008).

Figure 5.9 shows tranche deviations of all 16 models aggregated by time steps. There are four different states observable, two in each of the samples' halves. The horizontal dashed lines mark average values of tranche deviations to visualize the different states. Within the graph, state one is magnified since deviations are quite small. The period from 2006 until the end of 2007 can be labeled as calm market conditions with the beginning of the financial crisis at the end. In the second state tranche deviations jumped for the first time. Another jump occurred due to such events as the Lehman bankruptcy, the AIG bailout and the rejection of the bank bailout plan which afterwards led to the so-called Troubled Asset Relief Plan. These events are encompassed in state three. On top, the European sovereign crisis took place mainly in state four.

Figure 5.9 does not address model combinations but only displays ranges of tranche deviations. Thus, Table 5.4 outlines the combinations of distribution functions that are used and shows differences between the respective mixed models and the RFL model in terms of median values $v_{i,j} = \mathbb{MIE}\left(\sum_i^5 |\mathrm{TD}(i,t)|_{\mathrm{mix}} - \sum_i^5 |\mathrm{TD}(i,t)|_{\mathrm{RFL}}\right)$.

Since a comparison to the RFL model based on eyeballing average values is inconclusive, we use two-tailed Wilcoxon-signed rank tests to answer the question whether or not a model setup produces significant different results than RFL does. In particular, the alternative hypothesis states that the data consisting of tranche deviations produced by mixed models less tranche deviations produced by the RFL model stems from a distribution with a median unequal to zero. In period TD1, most of the mixed models outperform the RFL model on a level of significance of one percent. In period TD2, the results are different: Besides in cases where factor Y follows a CGMY or VGJD distribution, the reference model is able to cope significantly better with the new market conditions. In addition to a comparison to the RFL model, we also compare mixed models with one another. Table 5.5 summarizes median values of cumulative deviations produced by mixed models $\mathbb{MIE}\left(\sum_i^5 |\mathrm{TD}(i,t)|_{\mathrm{mix}} - \sum_i^5 |\mathrm{TD}(i,t)|_{\mathrm{mix}}\right)$. Based on the Wilcoxon-signed rank test, models built up by the factor distribution's combination {CGMY, CGMY}, {CGMY, VGJD}, and {VGJD, VGJD} are especially suitable in period TD1. Considering period TD2, a combination of {VGJD, VGJD} distributions for factor Y and ϵ_i, respectively, outperform almost all other mixed models.

		CGMY CGMY	CGMY VGJD	CGMY NIGJD	CGMY St.t	VGJD CGMY	VGJD VGJD	VGJD NIGJD	VGJD St.t	NIGJD CGMY	NIGJD VGJD	NIGJD NIGJD	NIGJD St.t	St.t CGMY	St.t VGJD	St.t NIGJD	St.t St.t
CGMY	CGMY	0															
CGMY	VGJD	−0.1	0														
CGMY	NIGJD	0.8**	0.9**	0													
CGMY	St.t	1.8***	2.2**	0.9**	0												
VGJD	CGMY	1.5	1.7	−0.2	−0.4	0											
VGJD	VGJD	0.1	0.2	−0.9**	−2.1***	−1.3**	0										
VGJD	NIGJD	5.3***	6.0***	4.1***	3.5**	2.8***	5.1***	0									
VGJD	St.t	7.5***	7.6***	6.1***	4.1***	5.9***	6.9***	2.1*	0								
NIGJD	CGMY	2.8***	3.3***	1.5	−0.1	0.8**	2.8***	−2.2***	−4.5***	0							
NIGJD	VGJD	1.4*	1.1*	−0.3	−0.8	0.1	1.0***	−3.1**	−5.2***	−1.1**	0						
NIGJD	NIGJD	5.1***	5.0***	3.8**	2.3	2.6***	4.7***	−0.2	−2.4**	2.0	2.4*	0					
NIGJD	St.t	10.1***	10.2***	9.8***	7.4***	8.2***	10.1***	4.9**	2.2***	6.9***	7.5***	5.1***	0				
St.t	CGMY	2.1**	3.2**	2.0	0.7	1.1***	2.8***	−2.2***	−4.1***	0.4	1.2*	−1.4**	−7***	0			
St.t	VGJD	2.2*	2.6	0.8	−0.2	−0.1	1.4**	−3.1***	−4.9***	−0.6	0.2	−2.4***	−6.9***	−0.5	0		
St.t	NIGJD	6.1***	6.2***	4.8***	3.9**	3.5***	5.6***	0.7**	−1.8	2.6***	4.0**	0.6**	−4.4	2.6***	3.6***	0	
St.t	St.t	8.2***	8.2***	7.5***	5.7***	5.9***	7.8***	2.6***	0.8**	5.0***	5.8***	2.9***	−2.3	4.7***	6.0***	2.1	0
CGMY	CGMY	0															
CGMY	VGJD	1.5	0														
CGMY	NIGJD	7.0*	6.3*	0													
CGMY	St.t	13.7**	7.4***	8*	0												
VGJD	CGMY	−10.4**	−16.1***	−17.7**	−31.9***	0											
VGJD	VGJD	−26.7***	−35.5***	−25.1***	−50.5***	−6.9	0										
VGJD	NIGJD	−25.1*	−23.7*	−30.0*	−29.1***	0.9	6.4	0									
VGJD	St.t	−0.5	3.1	−2.2	−10.9**	5.6	14.3***	6	0								
NIGJD	CGMY	7.6	10.4*	7.9	2.9	44.4***	60.6***	47.4**	27.2**	0							
NIGJD	VGJD	10.2	11.4	4.5	−8.8	15.9***	32.1***	27.0**	13.4*	−2.5	0						
NIGJD	NIGJD	10.4	14.7	0.4	2.7	25.1***	24.8***	27.2**	15.4**	−2.4	−0.9	0					
NIGJD	St.t	43.0***	49.3***	47.4***	20.1***	61.4***	76.1***	58.0***	33.0***	10.3**	29.2***	9.7**	0				
St.t	CGMY	30.4***	41.7***	22.1***	19.0**	43.9***	55.1***	50.3***	39.9***	18.2**	23.0***	16.3***	−7.8	0			
St.t	VGJD	25.2***	23.8***	21.5***	12.1**	45.9***	58.5***	53.7***	43.7***	17.6***	32.0***	15.6**	−4.4	−0.7	0		
St.t	NIGJD	68.1***	73.1***	58.0***	28.2***	63.1***	89.9***	91.9***	60.3***	28.4***	46.3***	41.9***	10.0**	16.1***	12.7**	0	
St.t	St.t	99.5***	101.2***	80.2***	40.9***	121.7***	137.9***	133.1***	77.2***	52.6***	62.0***	57.9***	18.6***	40.7***	30.0***	9.2*	0

Table 5.5: Median differences between mixed models. Values above the main diagonal correspond to their counterparts below the main diagonal multiplied by negative one. Parameters on the upper half (lower half) belong to period TD1 (TD2). ***/**/* indicate significance at the 1%-/5%-/10%-level.

		TD1 [bps]					TD2 [bps]				
F_Y	F_ϵ	Eq	Me1	Me2	Me3	Se	Eq	Me1	Me2	Me3	Se
Normal	Normal	4.9	41.6	7.4	2.8	2.2	144.8	433.8	300.5	273.3	119.6
RFL		0.0	1.2	5.8	0.8	3.7	27.6	81.5	17.0	107.9	39.8
CGMY	CGMY	0.1	0.7	1.5	1.7	2.2	3.4	8.9	28.9	101.1	83.4
CGMY	VGJD	0.1	0.2	1.6	1.6	2.3	3.9	7.4	35.6	101.2	81.3
CGMY	NIGJD	0.1	0.8	1.8	2.1	2.4	2.0	12.6	44.6	102.5	81.9
CGMY	St.t	0.1	0.6	2.2	2.7	2.7	2.9	20.1	31.1	101.9	80.6
VGJD	CGMY	0.0	1.0	2.8	2.5	2.4	3.3	14.5	23.8	96.7	75.6
VGJD	VGJD	0.0	1.2	2.1	1.3	2.0	1.9	7.5	23.3	101.6	72.7
VGJD	NIGJD	0.2	1.2	4.1	3.4	2.8	1.5	11.6	34.1	98.5	72.2
VGJD	St.t	0.2	0.6	4.9	4.4	3.2	3.7	18.1	40.4	97.2	77.0
NIGJD	CGMY	0.0	1.2	3.2	2.8	2.4	1.7	20.4	35.7	103.4	71.5
NIGJD	VGJD	0.0	1.4	2.6	1.9	2.1	3.9	19.5	33.8	107.0	77.3
NIGJD	NIGJD	0.1	1.1	3.9	3.4	2.8	2.9	11.9	39.8	100.9	79.4
NIGJD	St.t	0.1	0.9	7.0	5.9	3.6	3.0	24.9	70.3	104.8	82.8
St.t	CGMY	0.1	1.1	3.6	2.9	2.4	2.9	48.1	32.6	99.2	74.8
St.t	VGJD	0.0	1.2	2.8	2.5	2.1	3.4	48.3	27.1	100.8	73.9
St.t	NIGJD	0.0	1.3	4.2	3.7	2.9	1.4	49.3	67.3	94.3	78.7
St.t	St.t	0.2	0.7	5.2	4.7	3.1	4.1	15.2	93.9	106.5	83.0

Table 5.6: Median absolute deviation $\mathbb{ME}\left(|\mathrm{TD}(i,t)|\right)$ in periods TD1 and TD2.

Mixed factor models: Tranche deviations

To analyze in more detail whether or not there are differences in terms of the models' ability to price individual tranches, we change the perspective from cumulative absolute values to tranche based absolute values in Table 5.6. In terms of the reference models, as mentioned before, the Normal factor model misprices the first mezzanine tranche most. In contrast, the RFL model is able to reproduce the market structure more accurately.

In terms of results for mixed models, Table 5.6 should be read in combination with Table 5.7. While the former gives an impression of the absolute level of tranche based deviations from market prices, the latter table analyzes which tranches are priced significantly better than by means of the RFL model.[64]

Considering period TD1 first, tranches Me1, Me2 and Se exhibit significant lower deviations using mixed models. In the case of Me3, the RFL model shows better performance for some models. In most cases, however, a positive but non-significant median is observable for Me3. Both RFL and mixed models are able to fit the equity tranche with less than one basis point pricing error precision. In period TD2, performances of RFL and mixed models converge. None of the mixed models is able to outperform RFL significantly on a basis of cumulated average deviations (see Table 5.4). When focusing on individual tranche levels, differences in behavior become observable. During times of severe crisis, mixed models seem to be able to reproduce tranches Eq as well as Me1 more adequately. High levels of significance are reached for the equity tranche. For the first mezzanine tranche, this is only true for some of the mixed models. Given the overall poor fit in period TD2, mixed models based on {CGMY,

[64] Differences in values between Tables 5.4 and 5.7 are due to the fact that, in contrast to an arithmetic mean, the difference of two median values does not equal the median of the respective differences.

		TD1 [bps]					TD2 [bps]				
F_Y	F_ϵ	Eq	Me1	Me2	Me3	Se	Eq	Me1	Me2	Me3	Se
CGMY	CGMY	0.0	−0.8***	−4.6***	0.9	−2.8***	−21.8***	−10.2	25.7	48.4	45.8***
CGMY	VGJD	0.1	−0.8***	−4.8***	0.7	−2.4***	−23.3***	−18.9*	24.5	45.8	46.4***
CGMY	NIGJD	0.1	−0.8***	−4.6***	1.1	−2.3***	−24.4***	−7.5	39.7	48.2	34.5***
CGMY	St.t	0.1	−0.6***	−4.4***	2.1	−2.6***	−25.9***	−8.6	28.6	48.8	46.7***
VGJD	CGMY	0.0	−0.6***	−4.6***	1.5	−2.6***	−17.1***	−13.7**	0.3	37.9	41.1***
VGJD	VGJD	0.0	−0.3**	−5.1***	0.7	−3.0***	−18.0***	−19.7***	20.8	35.9	32.9***
VGJD	NIGJD	0.1	−0.1*	−4.5***	2.6	−2.3***	−11.2***	−12.8**	21.3*	42.0	32.6***
VGJD	St.t	0.1	−0.9***	−2.7***	3.8**	−2.2***	−20.8***	−8.4	27.9	45.1	35.6***
NIGJD	CGMY	0.0	−0.2**	−3.5***	2.1	−2.6***	−13.9***	−1.0	25.4	46.2*	32.1***
NIGJD	VGJD	0.0	−0.3	−4.2***	1.1	−2.9***	−8.4**	−9.6	14.2	47.0	32.4***
NIGJD	NIGJD	0.0	−0.1	−4.7***	2.6	−2.3***	−19.7***	−3.2	30.6	51.2*	38.6***
NIGJD	St.t	0.1	−0.7**	−2.2*	5.2**	−2.1**	−20.5***	−4.5	57.5**	51.3*	44.7***
St.t	CGMY	0.0	−0.2***	−4.5***	2.1	−2.7***	−14.2***	3.6	22.5	50.6*	41.5***
St.t	VGJD	0.0	−0.3*	−4.2***	1.4	−2.8***	−7.0***	−1.0	4.2	49.4*	31.9***
St.t	NIGJD	0.0	−0.1	−3.6***	3.1*	−2.3***	−23.0***	15.1	65.1	52.0*	43.1***
St.t	St.t	0.1**	−0.8**	−2.2**	4.0*	−2.3***	−22.8***	−5.1	65.1**	79.9**	49.7***

Table 5.7: Median absolute model differences $\mathbb{ME}\left(|\mathrm{TD}(i,t)|_{\mathrm{mix}} - |\mathrm{TD}(i,t)|_{\mathrm{RFL}}\right)$ in periods TD1 and TD2. ***/**/* indicate significance at the 1%-/5%-/10%-level.

VGJD} as well as on {VGJD, CGMY}, {VGJD, VGJD} and {VGJD, NIGJD} distributions have to be highlighted in this context. However, due to its ability to produce a discontinuous portfolio loss distribution, RFL is able to price higher tranches more accurately within this period of time.

Mixed factor models: Case study

After having analyzed the models on an aggregate level, we conclude this section with a treatment of selected points in time. We consider four events in this case study: The first date, September 29, 2006, marks the beginning of our dataset and acts as representative of calm market conditions. A second point in time is September 28, 2007. Former rumors of an upcoming crisis came true at least as soon as the interbank lending market became illiquid. Therefore, this second date marks the beginning of the financial crisis. Right before the third date, September 30, 2008, the investment bank Lehman Brothers filed for Chapter 11 bankruptcy on September 15 (Reuters 2008). Earlier that month, the Federal Housing Finance Agency announced the federal takeover of Freddie Mac and Fannie Mae. These events had a global impact and therefore also moved the iTraxx Europe and its tranches. The fourth date, May 27, 2010, brings the event itself right into Europe since on April 23, 2010, the Greek government requested an EU/IMF bailout (Reuters 2010).

We choose to consider only these mixed model setups that showed superior performance compared to RFL and compared to other mixed models. In addition to tranche deviations, skewness and kurtosis of the convoluted distribution F_X as well as conditional probability of default according to formula (5.31) are presented in Table 5.8. At least given the first two dates, the mixed models chosen are able to fit market quotes almost perfectly. This also reflects a general observation within our dataset: a mixed factor model is able to fit market quotes extremely accurate when market conditions are regular as well as at the

	F_Y	F_ϵ	Eq	Me1	Me2	Me3	Se	$\sum \lvert \cdot \rvert$	s	κ	λ
09/29/06	Normal	Normal	5.5	47.8	2.5	−4.5	−3.5	63.8	0	3	5.4
	RFL		0.3	1.7	−4.7	6.4	6.8	19.9			7.0
	VGJD	CGMY	0.0	0.7	−0.2	0.6	1.1	2.6	−1.2	10.0	9.6
	VGJD	VGJD	0.0	−1.0	−2.1	−0.8	0.6	4.6	−1.1	8.6	10.3
	VGJD	NIGJD	0.0	−1.7	0.2	1.4	1.8	5.2	−1.1	11.9	9.5
09/28/07	Normal	Normal	9.0	119.2	36.1	5.1	−8.9	178.2	0	3	4.0
	RFL		0.0	0.7	−10.9	1.5	9.6	22.7			15.4
	VGJD	CGMY	0.0	0.1	−0.2	0.0	0.7	1.1	−1.3	7.7	21.0
	VGJD	VGJD	0.0	0.1	−0.3	0.0	0.6	1.1	−1.1	7.4	20.9
	VGJD	NIGJD	−0.1	−1.0	2.6	2.4	2.6	8.7	−1.6	16.1	21.3
09/30/08	Normal	Normal	−46.2	319.5	230.0	212.6	115.8	924.2	0	3	16.7
	RFL		−4.5	60.8	−33.3	−14.4	58.2	171.3			42.4
	VGJD	CGMY	2.5	0.3	−23.8	39.5	62.5	128.6	−2.2	15.9	44.7
	VGJD	VGJD	1.1	7.5	−20.4	40.7	62.2	131.9	−2.1	15.1	44.6
	VGJD	NIGJD	0.0	−3.4	−34.1	31.2	58.8	127.6	−0.2	15.9	46.3
05/27/10	Normal	Normal	−175.1	559.5	−155.8	1131.7	482.5	2504.7	0	3	19.1
	RFL		81.9	−81.1	−118.9	296.1	143.0	721.0			44.5
	VGJD	CGMY	7.0	−94.1	−76.4	483.0	211.5	872.0	3.4	34.0	45.7
	VGJD	VGJD	18.1	−88.0	14.7	560.2	234.6	915.6	−1.3	11.6	43.6
	VGJD	NIGJD	0.9	−58.9	−126.1	434.6	189.4	809.9	5.8	56.1	46.7

Table 5.8: Deviations from market prices ($c_{\text{Model}} - c_{\text{Market}}$) on different dates and belonging third and fourth moments of convoluted distributions F_X.

beginning of the financial crisis. Even in a scenario of an upcoming crisis, the model is able to reproduce market prices. However, in a crisis as severe as the financial crisis, market conditions (especially market illiquidity) are not to be captured within a mixed factor model framework at a satisfying level of accuracy. The same is true for the reference models though.

In terms of implied correlations, Figure 5.10 illustrates correlation structures with respect to each tranche. The good fit of the first two dates hereby translates into an almost flat correlation pattern. The Normal factor model, in comparison, creates the smile structure mentioned before with a correlation level well below the one of mixed factor models. We discussed this effect in Section 5.4 where we state that, given a higher correlation, mixed models are able to counteract decreasing equity prices while increasing senior tranche prices. As a result, mixed factor models provide markups and markdowns in each tranche at a given correlation level to fit market prices.

Performances of all models considered in this case study worsen when we switch over to dates three and four. Implied correlations of the mixed models become unstable. Especially higher tranches such as Me3 and Se1 face large deviations. In contrast to the flat correlation structure from before, implied correlations form a downward sloping function first and even lose structure on date four. One explanation is that, on the technical side, we use the sum of squared deviations in our optimization procedure and, therefore, put more weight on lower and thus more risky tranches compared to, e.g., implementing the sum of relative deviations. Within the course of the financial crisis, a reevaluation of risk occurred and especially high level tranches were to be found underpriced within the market conditions given in that particular period. As a result, prices of these tranches increased disproportionately high.

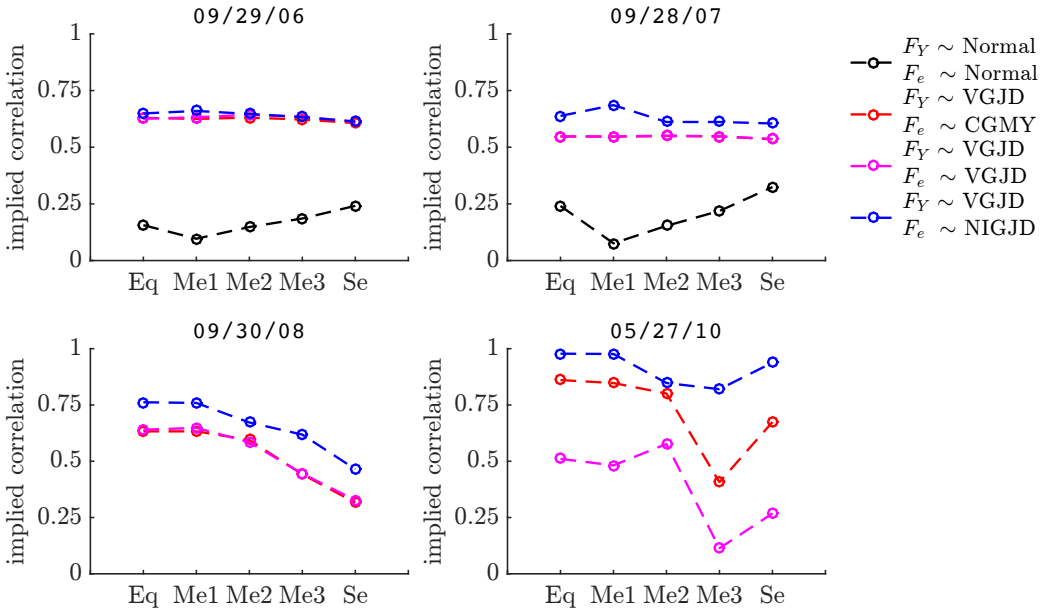

Figure 5.10: Implied correlation on different dates.

The focus of the minimization routine, however, is still on the first tranches. Table 5.8 shows that tranches Eq and Me1 are priced accurately by mixed models on date three. Given the high level of skewness and kurtosis to match these tranches, implied correlation is too high for more senior tranches to be priced close to market quotes.

5.6 Conclusion

In this paper, we present new extensions to the one-factor Gaussian copula. A focus is placed on mixing distributions that are part of either one of two broader distribution families: generalized tempered stable and generalized hyperbolic distributions. Since we deal with probability distributions that are not stable under convolution, we use an algorithm based on their respective characteristic functions to deal with convolution issues. In essence, we analyze the impact of skewness and kurtosis of the factors' distribution on tranche prices based on theoretical values as well as real market conditions. For the latter part, we use iTraxx Europe tranche quotes from 2006 to 2011 within a calibration study to elaborate fitting properties of mixed models compared to different reference models. Mixed factor models are able to provide markups on equity and senior tranches while simultaneously marking down mezzanine tranches compared to a Gaussian factor model. This feature enables mixed models to fit market data very well, at least when the market itself is not in a state of severe crisis. The calibration study shows that before the financial crisis until mid-2007, mixed models are able to almost flatten out a phenomenon known as correlation smile. Especially factor models based on Variance Gamma distributions that are extended by a Brownian part show a good fitting performance. However, when market conditions grow more severe, mixed models also lose fitting performance. Nevertheless, mixed factor

models can be used to accurately reproduce a wide range of stylized facts from empirical portfolio loss distributions while still remaining intuitive.

5.7 Acknowledgement

We gratefully thank the bwGRiD project[65] for the computational resources.

[65] bwGRiD (http://www.bw-grid.de), member of the German D-Grid initiative, funded by the Ministry for Education and Research (Bundesministerium für Bildung und Forschung) and the Ministry for Science, Research and the Arts Baden-Württemberg (Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg).

Appendix: Yet Another Factor Model

5.A Standardizing distribution functions

As indicated in Section 5.2, the distributions used for the two factors in equation (5.5) have to be standardized to zero mean and unit variance. The cumulant generating function

$$c_n = \frac{1}{i^n} \frac{\partial^n \psi(u)}{\partial u^n}\Big|_{u=0} = \begin{cases} 0 & n = 1 \\ 1 & n = 2 \end{cases} \tag{5.33}$$

can be used to assure the above criteria for the expected value ($c_1 = 0$) and variance ($c_2 = 1$). Therefore, the knowledge of the characteristic exponent $\psi(u)$ is used to standardize the distribution functions with the help of its derivatives. In the case of this paper's infinitely divisible distributions, extended versions of NIG and VG distributions as well as the CGMY distribution are of interest. Considering the extended NIG distribution first yields to the following conditions:

$$\mu = \frac{-\delta\beta}{\alpha^2 - \beta^2}, \tag{5.34}$$

$$\bar{\delta} = \min\left(\delta, \frac{(\alpha^2 - \beta^2)^{\frac{3}{2}}}{\alpha^2}\right), \tag{5.35}$$

$$\sigma^2 = 1 - \frac{\bar{\delta}\alpha^2}{(\alpha^2 - \beta^2)^{\frac{3}{2}}}. \tag{5.36}$$

The extended VG distribution standardization can be achieved by means of:

$$\mu = -\theta, \tag{5.37}$$

$$\bar{\sigma}_{VG} = \sqrt{\min\left(\sigma_{VG}^2, 1 - \nu\theta^2\right)}, \tag{5.38}$$

$$\sigma^2 = 1 - \sigma_{VG}^2 - \nu\theta^2. \tag{5.39}$$

The $\min(\cdot)$ function in both the NIG and VG standardization guarantees that δ and σ_{VG}, respectively, are free parameters in the optimization routine as long as the non-enhanced distributions exhibit $c_2 < 1$. In that case, the Brownian part yields unit variance. In the

case of $c_2 > 1$, the Brownian part becomes zero, and δ or σ_{VG} are restricted to assure $c_2 = 1$. In addition, CGMY distributions are standardized given the following relations:

$$\mu = -CY\Gamma(-Y)\left(G^{Y-1} - M^{Y-1}\right) \tag{5.40}$$

$$C = \left[Y\Gamma(-Y)(G^{Y-2} + M^{Y-2})(Y-1)\right]^{-1}. \tag{5.41}$$

Standardizing the Student's t-distribution to zero mean and unit variance implies taking the limit $\nu \to \infty$, which, in turn, implicates convergence to a standard Normal distribution. Thus, for a double t model, a scaled version of the one-factor model in (5.5)

$$X_i = z_Y a Y + z_\epsilon \sqrt{1 - a^2}\epsilon_i \tag{5.42}$$

is introduced by Hull and White (2004) instead. To allow for mixes, the scaling factors z_1 and z_2 for each factor $k = Y, \epsilon$ have to be defined as

$$z_k = \begin{cases} \sqrt{\frac{\nu_k - 2}{\nu_k}} & \text{if Student's t-distributed} \\ 1 & \text{otherwise.} \end{cases} \tag{5.43}$$

As we do not restrict the parameter ν to be integer valued, our implementation is related to the fractional degrees of freedom copula model of Wang et al. (2009). Apart from the fact that the tail-heaviness is allowed to vary continuously, another advantage is that, while still demanding a finite variance[66], ν is no longer restricted to a minimum value of three but is allowed to reach down to $\nu > 2$, which increases the heaviest tail decay rate possible.

[66] The variance of a Student's t-distributed random variable X is given by $\mathbb{V}[X] = \frac{\nu}{\nu-2}$.

5.B Dataset

Date	Eq (500)	Me1 (0)	Me2 (0)	Me3 (0)	Se (0)
January 31, 2007	10.34 %	41.59	11.95	5.60	2.00
March 01, 2007	9.98 %	46.17	12.88	5.77	2.16
March 30, 2007	11.23 %	57.75	14.28	6.24	2.58
April 30, 2007	9.94 %	49.82	12.45	5.53	2.54
May 31, 2007	6.33 %	39.90	10.33	4.39	1.93
June 29, 2007	11.75 %	62.05	16.29	7.48	3.10
July 31, 2007	30.23 %	227.08	90.50	49.60	26.70
September 28, 2007	19.13 %	91.96	37.17	24.04	15.04
November 30, 2007	25.75 %	161.42	85.50	62.00	36.00
January 31, 2008	30.98 %	316.90	212.40	140.00	73.60
March 31, 2008	40.28 %	483.50	310.00	216.40	109.50
May 30, 2008	34.03 %	300.92	188.92	126.96	61.63
July 31, 2008	31.48 %	355.70	220.00	140.70	69.80
September 30, 2008	46.92 %	672.06	387.93	208.02	96.83
November 25, 2008	64.03 %	1175.83	600.56	325.00	127.33
January 30, 2009	64.28 %	1185.63	606.69	315.63	97.13
	Eq (500)	Me1 (500)	Me2 (500)	Me3 (0)	Se (0)
March 31, 2009	66.83 %	31.23 %	11.53 %	418.80	155.00
May 28, 2009	53.13 %	13.75 %	−0.11 %	242.13	99.38
July 31, 2009	38.63 %	2.75 %	−7.52 %	143.33	65.17
September 30, 2009	36.81 %	2.83 %	−6.95 %	147.75	58.75
November 30, 2009	36.27 %	−0.73 %	−7.80 %	134.81	54.31
January 29, 2010	28.81 %	−4.00 %	−10.48 %	104.83	41.33
	Eq (500)	Me1 (300)	Me2 (100)	Me3 (100)	Se (100)
March 31, 2010	27.03 %	−4.18	−3.99	94.01	37.13
	Eq (500)	Me1 (500)	Me2 (500)	Me3 (100)	Se (100)
May 27, 2010	41.39 %	7.43 %	3.45 %	189.75	77.30
July 30, 2010	35.14 %	4.70 %	−0.38 %	143.85	59.75
September 30, 2010	27.71 %	−1.56 %	−2.90 %	111.00	50.50
	Eq (500)	Me1 (500)	Me2 (300)	Me3 (100)	Se (100)
November 30, 2010	29.80 %	0.08 %	−2.29 %	106.50	49.67
January 31, 2011	22.17 %	−3.77 %	−3.66 %	69.13	30.06
March 31, 2011	20.18 %	−3.75 %	−4.02 %	53.13	22.88
May 31, 2011	22.27 %	−3.56 %	−3.69 %	57.00	22.38

Table 5.9: iTraxx Europe tranche prices on which the calibration study is partially based on. Up-front payments are indicated in brackets below the label of the respective tranche. The table head is repeated every time a change in how the individual tranches are quoted occurred.

Chapter 6

Conclusion

Over the last decades, the financial world has evolved to a higher level of technical integration and automatization. A large portion of daily trading is executed by autonomously acting computer systems. In such an environment, the importance of thoroughly implemented models and knowledge of their capabilities in terms of accuracy and speed is crucial. Simultaneously, the complexity of financial markets as well as the complexity of the financial contracts that are traded in these markets stand in contrast to many valuation methods that have been used in the past. The combination of market complexity on the one hand and a demand for real time evaluation on the other hand, defines the need for robust and flexible models which meet the requirements that are connected to the task of pricing and risk monitoring.

In general, this thesis contemplates the potential of series expansion methods in various ways. It is written with the scope of delivering insight into new numerical methods as well as extending present methods that can be used within a widespread area of applications. Thus, besides interested readers from academia, especially quantitative analysts and model developers of financial institutions are a potential audience.

Recently, instead of relying on closed-form solutions, the usage of numerical methods come to the fore and become more popular. Within this development, our focus of attention is dedicated to derivatives. The contribution of the thesis at hand is three-folded: First, we presented and analyzed a new method to price European options that are written on a single underlying asset by introducing Gabor series methods into option pricing in Chapter 3. The resulting procedure shows to be a very robust pricing tool with a special strength in calculating short-term contracts. Since European options are heavily considered when determining input parameters of stochastic processes implicitly from traded contracts, the Gabor method is predestinated to be used within a calibration routine handling short term contracts.

Second, we dedicate Chapter 4 to multi-asset derivatives. Multi-asset contracts are notoriously hard to deal with in case the user demands both, advanced stochastic processes and fast evaluation. Compared to single underlying derivatives, these multi-asset exotic options

are studied to a much lower degree. We focused on European multi-asset options as well as on discrete barrier multi-asset options. To the best of our knowledge, the valuation of multi-asset barrier options in terms of multi-dimensional Fourier series methods has not been addressed in literature before. The method provides a fast and flexible alternative to Monte Carlo simulations which are used predominantly. This fact and the knowledge drawn from comparing different types of Fourier series is important for quantitative modelers providing pricing frameworks for market maker as well as traders of such exotic options.

Third, as recent events in the market for credit risk have shown, there is a need for further research in methods to evaluate credit derivatives. Therefore, in Chapter 5, we focused on extending the standard Gaussian factor model to price synthetic collateralized debt obligations. The new models are able to cope with a wide range of market conditions. However, given a crisis as severe as the financial crisis of 2008, questions, such as market liquidity, that are outside the scope of pure modeling overlay the approximation quality. Nevertheless, the models presented are flexible instruments to price synthetic collateralized debt obligations while still staying in the intuitive framework of a factor model.

Bibliography

Adcock, B. (2010). Multivariate Modified Fourier Series and Application to Boundary Value Problems. *Numerische Mathematik 115*(4), 511–552.

Albrecher, H., S. A. Ladoucette, and W. Schoutens (2007). A Generic One-Factor Lévy Model for Pricing Synthetic CDOs. In M. Fu, R. Jarrow, J.-Y. Yen, and R. Elliot (Eds.), *Advances in Mathematical Finance*, pp. 259–277. Boston: Birkhäuser.

Amin, K. (1993). Jump Diffusion Option Valuation in Discrete Time. *Journal of Finance 48*(5), 1833–1863.

Andersen, L. and J. Sidenius (2004). Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings. *Journal of Credit Risk 1*(1), 29–70.

Ascheberg, M., B. Bick, and H. Kraft (2013). Hedging Structured Credit Products During the Credit Crisis: A Horse Race of 10 Models. *Journal of Banking and Finance 37*, 1687–1705.

Bachelier, L. (1900). Théorie de la Spéculation. *Annales Scientifiques de l'École Normale Supérieure 3*(17), 21–86.

Bailey, D. and P. Swarztrauber (1994). A Fast Method for the Numerical Evaluation of Continuous Fourier and Laplace Transforms. *SIAM Journal on Scientific Computing 15*(5), 1105–1110.

Bakshi, G. and D. Madan (2000). Spanning and Derivative-Security Valuation. *Journal of Financial Economics 55*(2), 205–238.

Barndorff-Nielsen, O. (1977). Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Royal Society of London Proceedings Series A 353*, 401–419.

Bingham, N. and R. Kiesel (2004). *Risk-neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer Science & Business Media.

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*(3), 637–654.

Boyle, P. (1988). A Lattice Framework for Option Pricing with Two State Variables. *Journal of Financial and Quantitative Analysis 23*(01), 1–12.

Boyle, P., J. Evnine, and S. Gibbs (1989). Numerical Evaluation of Multivariate Contingent Claims. *Review of Financial Studies 2*(2), 241–250.

Boyle, P. and S. H. Lau (1994). Bumping Up Against the Barrier with the Binomial Method. *The Journal of Derivatives 1*(4), 6–14.

Burtschell, X., J. Gregory, and J.-P. Laurent (2009). A Comparative Analysis of CDO Pricing Models under the Factor Copula Framework. *Journal of Derivatives 16*(4), 9–37.

Carr, P., H. Geman, D. Madan, and M. Yor (2002). The Fine Structure of Asset Returns: An Empirical Investigation. *Journal of Business 75*(2), 305–332.

Carr, P. and D. Madan (1999). Option Valuation Using the Fast Fourier Transform. *Journal of Computational Finance 2*(4), 61–73.

Carr, P. and L. Wu (2004). Time-changed Lévy Processes and Option Pricing. *Journal of Financial Economics 71*(1), 113–141.

Chourdakis, K. (2005). Option Pricing Using the Fractional FFT. *Journal of Computational Finance 8*, 1–18.

Christensen, O. (2008). *Frames and Bases*. Birkhäuser.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Boca Raton: Chapman & Hall/CRC.

Cont, R. and E. Voltchkova (2005). A Finite Difference Scheme for Option Pricing in Jump-Diffusion and Exponential Lévy Models. *SIAM Journal on Numerical Analysis 43*(4), 1596–1626.

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

Cox, J., S. Ross, and M. Rubinstein (1979). Option pricing: A Simplified Approach. *Journal of Financial Economics 7*(3), 229–263.

Daubechies, I., A. Grossmann, and Y. Meyer (1986). Painless Nonorthogonal Expansions. *Journal of Mathematical Physics 27*(5), 1271–1283.

Delbaen, F. and W. Schachermayer (1994). A General Version of the Fundamental Theorem of Asset Pricing. *Mathematische Annalen 300*(1), 463–520.

Dörfler, M. (2002). *Gabor Analysis for a Class of Signals called Music*. Ph. D. thesis, University of Vienna.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*(6), 1343–1376.

Eberlein, E., R. Frey, and E. A. von Hammerstein (2008). Advanced Credit Portfolio Modeling and CDO Pricing. In W. Jäger and H.-J. Krebs (Eds.), *Mathematics – Key Technology for the Future*, pp. 253–279. Springer.

Eberlein, E. and K. Glau (2014). Variational Solutions of the Pricing PIDEs for European Options in Lévy Models. *Applied Mathematical Finance 21*(5), 417–450.

Fang, F., H. Jönsson, C. Oosterlee, and W. Schoutens (2010). Fast Valuation and Calibration of Credit Default Swaps Under Levy Dynamics. *Journal of Computational Finance 14*(2), 57–86.

Fang, F. and C. Oosterlee (2008). A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing 31*(2), 826–848.

Fang, F. and C. Oosterlee (2009). Pricing Early-Exercise and Discrete Barrier Options by Fourier-Cosine Series Expansions. *Numerische Mathematik 114*, 27–62.

Feichtinger, H. and K. Gröchenig (1992). Gabor Wavelets and the Heisenberg Group: Gabor Expansions and Short Time Fourier Transform from the Group Theoretical Point of View. In *Wavelets: A Tutorial in Theory and Applications*, pp. 359–398.

Gabor, D. (1946). Theory of Communication. *Journal of the Institution of Electrical Engineers 93*(26), 429–441.

Geske, R. and H. Johnson (1984). The American Put Option Valued Analytically. *Journal of Finance 39*(5), 1511–1524.

Gil-Pelaez, J. (1951). Note on the Inversion Theorem. *Biometrika 38*, 481–482.

Gröchenig, K. (2001). *Foundations of Time-Frequency Analysis*. Birkhäuser.

Hager, S. and R. Schöbel (2006). *Computational Science – ICCS 2006*, Volume 3994 of *Lecture Notes in Computer Science*, Chapter Deriving the Dependence Structure of Portfolio Credit Derivatives Using Evolutionary Algorithms, pp. 340–347. Springer.

Heisenberg, W. (1927). Über den Anschaulichen Inhalt der Quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik 43*(3), 172–198.

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

Hull, J. and A. White (1990). Valuing Derivative Securities using the Explicit Finite Difference Method. *Journal of Financial and Quantitative Analysis 25*(01), 87–100.

Hull, J. and A. White (2004). Valuation of a CDO and nth to Default CDS without Monte Carlo Simulation. *Journal of Derivatives 12*(2), 8–23.

Hurn, S., K. Lindsay, and A. McClelland (2013). On the Efficacy of Fourier Series Approximations for Pricing European and Digital Options. Technical report, National Centre for Econometric Research.

Hurst, S. (1995). The Characteristic Function of the Student t Distribution. Financial Mathematics Research Report 006-95, Australian National University.

Iserles, A. and S. Nørsett (2008). From High Oscillation to Rapid Approximation I: Modified Fourier Expansions. *IMA Journal of Numerical Analysis 28*(4), 862–887.

Kalemanova, A., B. Schmid, and R. Werner (2007). The Normal Inverse Gaussian Distribution for Synthetic CDO Pricing. *Journal of Derivatives 14*(3), 88–94.

Koponen, I. (1995). Analytic Approach to the Problem of Convergence of Truncated Lévy Flights towards the Gaussian Stochastic Process. *Physical Review E 52*(1), 1197–1199.

Korn, R. and S. Müller (2009). The Decoupling Approach to Binomial Pricing of Multi-Asset Options. *Journal of Computational Finance 12*(3), 1–30.

Kreps, D. (1981). Arbitrage and Equilibrium in Economies with Infinitely Many Commodities. *Journal of Mathematical Economics 8*(1), 15–35.

Leoni, P. and W. Schoutens (2008). Multivariate Smiling. *Wilmott magazine*.

Longstaff, F. A. and A. Rajan (2008). An Empirical Analysis of the Pricing of Collateralized Debt Obligations. *Journal of Finance 63*(2), 529–563.

Madan, D., P. Carr, and E. Chang (1998). The Variance Gamma Process and Option Pricing. *European Finance Review 2*(1), 79–105.

Madan, D. B. and E. Seneta (1990). The VG Model for Share Market Returns. *Journal of Business 63*(4), 511–524.

Madan, D. B. and M. Yor (2008). Representing the CGMY and Meixner Lévy Processes as Time Changed Brownian Motions. *Journal of Computational Finance 12*(1).

Margrabe, W. (1978). The Value of an Option to Exchange One Asset for Another. *Journal of Finance 33*(1), 177–186.

Matache, A.-M., T. v. Petersdorff, and C. Schwab (2004). Fast Deterministic Pricing of Options on Lévy Driven Assets. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 38*(1), 37–71.

Meng, Q.-J. and D. Ding (2013). An efficient pricing method for rainbow options based on two-dimensional modified sine-sine series expansion. *International Journal of Computer Mathematics 90*(5), 1096–1113.

Moosbrucker, T. (2006a). Explaining the Correlation Smile Using Variance Gamma Distributions. *Journal of Fixed Income 16*(1), 71–87.

Moosbrucker, T. (2006b). Pricing CDOs with Correlated Variance Gamma Distributions. Technical report, Department of Banking, University of Cologne.

Niederstätter, F. (2014). Non-orthogonal option pricing. *SSRN-id2432870*, 1–27.

Nimmanunta, K., A. Chiarawongse, and S. Tirapat (2008). On Pricing CDOs with Meixner Distributions. *Journal of Fixed Income 18*(1), 86–99.

Øksendal, B. (2003). *Stochastic Differential Equations*. Springer.

Ortiz-Gracia, L. and C. Oosterlee (2013). Robust Pricing of European Options with Wavelets and the Characteristic Function. *SIAM Journal on Scientific Computing 35*(5), B1055–B1084.

Pivato, M. (2010). *Linear Partial Differential Equations and Fourier Theory*. Cambridge University Press.

Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measure*. Ph. D. thesis, Albert-Ludwigs-Universität Freiburg i. Br.

Reuters (2008, September). Lehman Files for Bankruptcy, Plans to Sell Units. [accessed on 01/03/2013].

Reuters (2010, April). Greek PM Announces Activation of EU/IMF Aid Package. [accessed on 01/03/2013].

Rosinski, J. (2007). Tempering Stable Processes. *Stochastic Processes and their Applications 117*(6), 677–707.

Ruijter, M. and C. Oosterlee (2012). Two-Dimensional Fourier Cosine Series Expansion Method For Pricing Financial Options. *SIAM Journal on Scientific Computing 34*(5), B642–B671.

Salmon, F. (2009). Recipe for Disaster: The Formula That Killed Wall Street. *Wired Magazine (March 17th)*.

Samuelson, P. (1965). Rational theory of warrant pricing. *Industrial Management Review 6*(2), 13–32.

Sato, K.-I. (2005). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics 68. New York: Cambridge University Press.

Scherer, M., S. T. Rachev, Y. S. Kim, and F. J. Fabozzi (2009). A FFT-based Approximation of Tempered Stable and Tempered Infinitely Divisible Distributions. Technical report.

Schöbel, R. and J. Zhu (1999). Stochastic Volatility with an Ornstein–Uhlenbeck Process: An Extension. *European Finance Review 3*(1), 23–46.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models*. John Wiley & Sons.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. Chichester: Wiley.

Schoutens, W. and J. Cariboni (2009). *Levy Processes in Credit Risk*. Wiley.

Shreve, S. (2004). *Stochastic Calculus for Finance II: Continuous-time Models*. Springer.

Stein, E. and J. Stein (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies 4*(4), 727–752.

Stein, E. M. and R. Shakarchi (2011). *Fourier analysis: An Introduction*. Princeton University Press.

Tolstov, G. P. (1976). *Fourier Series*. Dover Publications.

Vasicek, O. A. (1987). Probability of Loss on Loan Portfolio. Technical report, KMV Corporation.

Wang, D., S. T. Rachev, and F. J. Fabozzi (2009). Pricing of Credit Default Index Swap Tranches with One-Factor Heavy-Tailed Copula Models. *Journal of Empirical Finance 16*(2), 201–215.