Fourier Methods for Multidimensional Problems and Backward SDEs in Finance and Economics

FOURIER METHODS FOR MULTIDIMENSIONAL PROBLEMS AND BACKWARD SDES IN FINANCE AND ECONOMICS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 6 februari 2015 om 12:30 uur

door

Maria Johanna RUIJTER

wiskundig ingenieur
geboren te Venhuizen, Nederland.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. C. W. Oosterlee

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. C. W. Oosterlee,	Technische Universiteit Delft, promotor
Prof. dr. P. A. Forsyth	University of Waterloo, Canada
Prof. dr. ir. D. Huybrechs	Katholieke Universiteit Leuven, België
Prof. dr. M. R. H. Mandjes	Universiteit van Amsterdam
Prof. dr. R. J. A. Laeven	Universiteit van Amsterdam
Prof. dr. ir. A. W. Heemink	Technische Universiteit Delft
Dr. R. F. T. Aalbers	Centraal Planbureau
Prof. dr. ir. C. Vuik	Technische Universiteit Delft, reservelid

Methods for Multidimensional Problems and Backward SDEs in Finance and Economics
Dissertation at Delft University of Technology

ISBN 978-94-6259-526-2

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`

Cover design by Marie-José Schouten (MarieJ Vormgeving) en Jeroen Timmer

Contents

Summary

In this thesis we deal with processes with uncertainties, such as financial asset prices and the global temperature. We model their evolutions by so-called stochastic processes. Many of these stochastic processes are based on the Wiener process, whose increments are normally distributed. Other models may contain jump components, to model, for example, economic disasters or degradation failures. An important class of models is the Lévy class, where successive increments are independent and statistically identical over different time intervals of the same length. This may give computational advantages.

A well-known application of stochastic processes is in financial mathematics, where the goal is to price financial derivatives or to estimate risk measures. The underlying asset prices may be modeled by, e.g., geometric Brownian motions. More involved models, like the Variance Gamma process, are defined by jumps. In other, for instance economic, personal, or societal, contexts one may face options in the sense of real 'choices'. For example, should one build a new factory now or in the future? Or should one heighten a dike today, and by how much, or in the future? These decisions are called real options and can often be related to financial options. Similar methods can be used to value them.

The numerical problems we consider deal with conditional expectations. Often these problems can be connected to a partial differential equation (PDE) by a Feynman-Kac theorem. Then we can apply PDE methods, such as finite difference schemes and finite volume methods, to approximate the solutions. From the perspective of the probabilistic representation, the class of Monte Carlo methods can be beneficial for high-dimensional problems. They are based on simulated paths of the stochastic process. Besides, the expected value can often be represented by an integral, which offers opportunities to use numerical integration techniques, such as Newton-Cotes formulas.

We focus on a subclass of numerical integration methods, i.e., Fourier-based methods. These 'transform methods' combine a transformation to the Fourier domain with numerical integration. The probability density function of the random variables of interest is usually unknown. Its Fourier transform, i.e., the characteristic function, is however often known and can be used to approximate the corresponding density and distribution function. Our method of choice is the COS method, which is based on Fourier cosine series expansions and the characteristic function. The matrix-vector product appearing may be computed in an efficient way by using a Fast Fourier Transform (FFT) algorithm, especially when dealing with Lévy processes. Besides, the use of the discrete Fourier cosine transform helps us with the approximation of the Fourier coefficients.

After a general introduction of this thesis in Chapter 1, in Chapter 2 we explain the COS formula to compute conditional expectations and provide an error analysis. Then the COS method is applied to a specific class of problems: stochastic control problems, in which an

agent has the possibility to affect the trend or variation of a stochastic process in such a way that his target function is maximized. For example, he can determine his consumption or savings rate.

With the COS method we approximate functions by using Fourier cosine series. Similar to Fourier series they may suffer from the Gibbs phenomenon: trying to recover a function with a jump discontinuity results in undesired oscillations, even if the number of terms in the series goes to infinity. Smooth density functions give rise to a fast exponentially converging error of the COS method, but a density function with a discontinuity in one of its derivatives results in slower algebraic convergence. This is related to the Gibbs phenomenon. A remedy to improve this is by using spectral filters, which smoothen the approximations, see Chapter 3.

Vanilla call and put options are based on one underlying asset. On the other hand, rainbow options are written on multiple assets and the holder may possess a 'basket' of assets. The payoff of, for example, a call-on-max option, depends on the maximum of several assets. In this way, an option holder can manage his risks. Financial options on two assets are discussed in Chapter 4. Also single-asset options under the Heston model, in which an additional stochastic process for the volatility is used, are priced by the so-called 2D-COS method, developed in this chapter.

Chapter 5 deals with a problem from the field of climate change economics. The future global temperature is highly uncertain and there are different damage estimates. In our model an agent can choose the consumption level of his wealth while he is subject to these uncertain climate damages. There is a trade-off between consuming now and saving for later. Economic equilibrium conditions result in a mathematical expression for the appropriate social discount rate. Here the future temperature process and the economic wealth are the uncertain processes and we combine the methods from Chapter 2 and Chapter 4 to solve the problem.

Forward stochastic processes are rather well-known. Their initial value is prescribed and a prescription for the process forwards in time is given. During the last decades the backward stochastic differential equations (BSDEs) have become popular. A BSDE is stochastic differential equation for which a terminal condition, instead of an initial condition, has been specified. Its solution consists of a pair of processes, where one of the processes 'steers' the other towards the terminal condition. The value of a call option can be modeled in this way as the holder of a replicating portfolio aims to end up with a certain payoff at the terminal time. Market imperfections can also be incorporated, such as different lending and borrowing rates for money, the presence of transaction costs, or short selling constraints. In Chapter 6 we extend the COS method to solve such problems and name it the BCOS method.

A forward stochastic process can be approximated by different simulation schemes. The stochastic Euler scheme is a generalization of the Euler scheme for ordinary differential equations. Higher order Taylor schemes include more stochastic terms to obtain a better convergence rate. In Chapter 7 we extend our pricing and valuation methodology by using the characteristic function of these discrete schemes within the BCOS method framework. With the second-order weak Taylor scheme and a θ-time discretization we obtain second order convergence in the number of timesteps for several problems. The techniques in Chapter 7 enable us to generalize the applicability of the BCOS method to forward SDEs for which the 'continuous' characteristic function is not available.

Samenvatting

In dit proefschrift gebruiken we processen met onzekerheden, zoals financiële aandelenprijzen en de toekomstige temperatuurstijging. We modelleren deze bewegingen door middel van zogeheten stochastische processen. Veel stochastische processen zijn gebaseerd op het Wiener proces, waarbij de incrementen normaal verdeeld zijn. Andere modellen kunnen sprongtermen bevatten, om bijvoorbeeld economische rampen of falen door degradatie te modelleren. Een belangrijke groep is de Lévy-klasse, waar de incrementen over verschillende tijdsintervallen, van dezelfde lengte, onafhankelijk en gelijk verdeeld zijn. Dit kan rekenvoordelen opleveren.

Een bekende toepassing van stochastische processen is te vinden in de financiële wiskunde, waarbij het doel is om financiële derivaten of risico's te waarderen. De onderliggende aandelenprijzen kunnen worden gemodelleerd met bijvoorbeeld geometrische Brownse bewegingen. Complexere modellen, zoals het Variantie-Gamma proces, bestaan uit sprongen. Ook in bijvoorbeeld economische, persoonlijke of maatschappelijke context komen opties voor in de zin van echte 'keuzes'. Zou men een nieuwe fabriek nu moeten bouwen of in de toekomst? Zou men de dijken vandaag moeten ophogen, en met hoeveel, of in de toekomst? Deze beslissingsproblemen worden reële opties genoemd en kunnen vaak gerelateerd worden aan financiële opties. Vergelijkbare methoden kunnen worden gebruikt om ze te waarderen.

Wij onderzoeken numerieke problemen met conditionele verwachtingswaarden. Vaak kunnen deze problemen met een Feynman-Kac stelling gelinkt worden aan een partiële differentiaalvergelijking (PDV). Daarop kunnen we dan PDV-methoden toepassen, zoals eindige-differentieschema's en eindige-volumemethoden, om oplossingen te benaderen. Vanuit de probabilistische representatie bekeken kan de klasse van Monte Carlo methoden gunstig zijn voor hoogdimensionale problemen. Hierbij worden paden van het stochastisch proces gesimuleerd. Daarnaast kan de verwachtingswaarde vaak geformuleerd worden als een integraal, wat mogelijkheden biedt voor numerieke integratietechnieken, zoals de Newton-Cotes formules.

Wij richten ons op een subklasse van numerieke integratiemethoden, te weten Fourier-gebaseerde methoden. Deze 'transformatiemethoden' combineren een transformatie naar het Fourierdomein met numerieke integratie. De kansdichtheidsfunctie van de random variabelen waar we in geïnteresseerd zijn is meestal onbekend. Echter de Fouriergetransformeerde, dat is de karakteristieke functie, is vaak wel bekend en kan worden gebruikt om de bijbehorende kansdichtheids- en verdelingsfunctie te benaderen. Wij werken met de COS methode, die is gebaseerd op Fourier-cosinusreeks-expansies en de karakteristieke functie. De bijbehorende matrix-vector-vermenigvuldigingen kunnen op efficiënte wijze worden berekend door middel van een Fast Fourier Transform-algoritme (FFT), met name

bij Lévy processen. Verder gebruiken we discrete Fourier-cosinustransformaties voor het schatten van de Fourier-coëfficiënten.

Na een algemene inleiding van dit proefschrift in Hoofdstuk 1 leggen we in Hoofdstuk 2 de COS formule uit om voorwaardelijke verwachtingswaarden te berekenen en geven we een foutanalyse. Daarna wordt de COS methode toegepast op een specifieke categorie problemen: stochastische regeltechniekproblemen. Hierbij heeft een agent de mogelijkheid om de trend of onzekerheid van een stochastisch proces te beïnvloeden, zodanig dat zijn doelfunctie wordt gemaximaliseerd. Zo kan hij bijvoorbeeld zijn consumptie of spaarquote bepalen.

Met de COS methode benaderen we functies door gebruik te maken van Fourier-cosinusreeksen. Vergelijkbaar met Fourierreeksen lijden zij aan het Gibbs fenomeen: het reconstrueren van een functie met een sprongdiscontinuïteit leidt tot ongewenste oscillaties, zelfs als het aantal termen in de reeks naar oneindig gaat. Gladde kansdichtheidsfuncties geven een snelle exponentieel convergerende fout van de COS methode, maar een dichtheidsfunctie met een discontinuïteit in één van de afgeleiden resulteert in tragere algebraïsche convergentie. Dit houdt verband met het Gibbs fenomeen. Een mogelijkheid om dit te verbeteren is door gebruik te maken van spectrale filters, die de benaderingen gladder maken, zie Hoofdstuk 3.

Vanilla call- en putopties zijn geschreven op één onderliggend aandeel. Daarentegen zijn zogeheten regenboogopties geschreven op meerdere aandelen en de houder kan een 'aandelenmandje' bezitten. De uitbetaling van bijvoorbeeld een call-op-max-optie hangt af van het maximum van meerdere aandelenprijzen. Op deze manier kan een optiehouder zijn risico's managen. Financiële opties op twee aandelen worden besproken in Hoofdstuk 4. Ook opties onder het één-aandeel-Heston model, waarin een extra stochastisch proces voor de volatiliteit wordt gebruikt, kunnen worden geprijsd met de zogeheten 2D-COS methode die ontwikkeld is in dit hoofdstuk.

Hoofdstuk 5 behandelt een probleem uit het onderzoeksgebied klimaateconomie. De toekomstige wereldwijde temperatuur is erg onzeker en er zijn verschillende schadeschattingen. In ons model kan een agent het consumptieniveau van zijn rijkdom kiezen, terwijl hij onderworpen is aan de onzekere klimaatschade. Er is een trade-off tussen het consumeren nu en sparen voor later. Economische evenwichtsvoorwaarden resulteren in een wiskundige uitdrukking voor de bijbehorende sociale discontovoet. Het temperatuursproces en de economische welvaart zijn hier de onzekere processen en we combineren de methoden uit Hoofdstuk 2 en Hoofdstuk 4 om het probleem op te lossen.

Voorwaartse stochastische processen zijn redelijk bekend. Hun startwaarde wordt voorgeschreven en er is een beschrijving voor het proces voorwaarts in de tijd. De laatste decennia zijn terugwaartse stochastische differentiaalvergelijkingen (BSDEs) populair geworden. Een BSDE is een stochastische differentiaalvergelijking waarvoor een eindconditie, in plaats van een startvoorwaarde, is opgegeven. De oplossing bestaat uit twee processen, waarbij één van de processen de andere naar de eindtoestand 'stuurt'. De waarde van een calloptie kan op deze manier worden gemodelleerd: de houder van een replicerende portefeuille heeft als doel te eindigen met een bepaalde uitbetaling op de eindtijd. Ook marktimperfecties kunnen worden opgenomen, zoals verschillende tarieven voor het uitlenen en lenen van geld, transactiekosten, of short selling beperkingen. In Hoofdstuk 6 breiden we de COS methode uit voor het oplossen van dergelijke problemen en noemen het de BCOS methode.

Een voorwaarts stochastisch proces kan worden benaderd door verschillende simulatieschema's. Het stochastische Euler schema is een generalisatie van de Euler methode voor gewone differentiaalvergelijkingen. Bij Taylorschema's van hogere orde zijn meer stochastische termen toegevoegd om een betere convergentie-orde te verkrijgen. In Hoofdstuk 7 breiden wij onze prijzings- en waarderingsmethode uit met behulp van de karakteristieke functie van deze discrete schema's. Met het tweede-orde zwakke Taylorschema en een θ-tijd-discretisatie bereiken we de tweede orde convergentie in het aantal tijdstappen voor verschillende problemen. De technieken in Hoofdstuk 7 stellen ons in staat om de toepasbaarheid van de BCOS methode uit te breiden met voorwaartse SDV'en waarvoor de 'continue' karakteristieke functie is niet beschikbaar is.

CHAPTER 1

Introduction and Outline of this Thesis

The well-known Feynman-Kac theorem links the solution of partial differential equations (PDEs) to stochastic processes. In this chapter we briefly discuss this theorem and some extended versions. The conditional expectations in the probabilistic representations can be approximated by numerical methods, such as Fourier techniques like the COS method. Based on this introduction we lay out the remainder of this thesis.

Suppose we consider a one-dimensional *linear* parabolic PDE of the form

$$\frac{\partial v}{\partial t}(t,x) + \mu(t,x) D_x v(t,x) + \frac{1}{2}\sigma^2(t,x) D_x^2 v(t,x) - \rho v(t,x) = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}, \tag{1.1a}$$

$$v(T,x) = g(x), \quad \forall x \in \mathbb{R}. \tag{1.1b}$$

Function v is called the value function and D_x and D_x^2 denote the first and second derivative with respect to the x-variable, respectively. The functions $\mu : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ satisfy Lipschitz conditions on (t,x). Further, $T > 0$ denotes the finite terminal time, g is a terminal reward function, and ρ is a discount rate, which is common in economic and financial problems.

The Feynman-Kac theorem for linear parabolic PDEs gives as the PDE solution [Pha09]

$$v(t,x) = \mathbb{E}\left[e^{-\rho(T-t)} g(X_T)\right], \tag{1.2}$$

where X_s is governed by the forward stochastic differential equation (FSDE)

$$dX_s = \mu(s,X_s)ds + \sigma(s,X_s)d\omega_s, \quad X_t = x. \tag{1.3}$$

Here ω is a one-dimensional standard Brownian motion on the filtered probability space[1] $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{0\le s\le T}, \mathbb{P})$. Term μ is called the drift of the diffusion process and σ the volatility.

So, we can solve the linear PDE by considering a probabilistic formulation with an FSDE and an expected value. The price of European-style financial options can be formulated in this way, see Section 2.3.4. The most commonly known options are the call and put option, where the holder has the right, but not the obligation, to buy (sell) an asset at a prescribed date for a specified strike price [Hul09].

An extended PDE is the 1D *Hamilton-Jacobi-Bellman (HJB) equation*:

$$\frac{\partial v}{\partial t}(t,x) + \sup_{\alpha \in A}[\mu(t,x,\alpha) D_x v(t,x) + \frac{1}{2}\sigma^2(t,x,\alpha) D_x^2 v(t,x) + f(t,x,\alpha)] - \rho v(t,x) = 0. \tag{1.4}$$

[1] In this thesis we will use the symbol $\mathscr{F}$ to denote Fourier cosine coefficients and $\mathcal{F}$ denotes a σ-algebra.

This dynamic programming equation is associated to a stochastic control problem with value function

$$v(t,x) = \sup_{\boldsymbol{\alpha}\in\mathscr{A}} \mathbb{E}\left[\int_t^T e^{-\rho(s-t)} f(s,X_s,\boldsymbol{\alpha}_s)\,ds + e^{-\rho(T-t)} g(X_T)\right]. \tag{1.5}$$

where X_s follows the controlled FSDE

$$dX_s = \mu(s,X_s,\boldsymbol{\alpha}_s)\,ds + \sigma(s,X_s,\boldsymbol{\alpha}_s)\,d\omega_s, \quad X_t = x. \tag{1.6}$$

The control process $(\boldsymbol{\alpha}_s)_{0\le s\le T}$ is valued in the set $A\subset\mathbb{R}$. Now the drift and volatility terms depend also on this control value. The relation between the HJB equation and the stochastic control problem is an extension of the Feynman-Kac theorem. In Chapter 2, we focus on this type of problems, in which we search for an optimal control law $\boldsymbol{\alpha}$ for specific problems.

The broad class of stochastic problems can be subdivided into several other types, such as optimal stopping problems or impulse control problems [Pha09]. In the first type the controller can influence his expected reward by choosing a time to undertake a particular action. An example is an American-type option, where the holder can exercise his option at any time before the terminal date. Another example is the Bermudan option, which can be exercised at a prescribed set of early-exercise dates.

We apply a Fourier-based method to approximate the expectations appearing by using Fourier cosine series expansions and the characteristic function of the underlying stochastic process. The characteristic function $\varphi(.)$ of a random variable X is defined as

$$\varphi(u) = \mathbb{E}\left[e^{iuX}\right]. \tag{1.7}$$

The probability density function $p(.)$ of a continuous random variable and the characteristic function form a Fourier pair. The characteristic function is the Fourier transform of the density and the density is the inverse Fourier transform of the characteristic function:

$$\varphi(u) = \int_{\mathbb{R}} e^{iuy} p(y)\,dy \quad \text{and} \quad p(y) = \tfrac{1}{2\pi}\int_{\mathbb{R}} e^{-iuy}\varphi(u)\,du. \tag{1.8}$$

In Chapter 2 we explain the COS formula and provide an error analysis. Subsequently the COS method is applied to stochastic control problems.

A special class of underlying stochastic process is formed by Lévy processes. A Lévy process has stationary, independent increments and is right continuous with left limits. For an extensive overview of Lévy processes we refer to [CT03]. This class contains the constant coefficient jump-diffusion processes as well as Variance Gamma and CGMY processes. The Lévy-Khintchine formula gives a representation for the corresponding characteristic function. The specific properties of Lévy processes enable highly efficient computation of matrix-vector products.

Smooth density functions give rise to fast converging errors of the COS method. Density functions with a discontinuity in one of their derivatives result in slower algebraic convergence. This is related to the Gibbs phenomenon and a remedy to improve this is by using spectral filters, see Chapter 3. Also the distribution function of discrete random variables can be recovered by a filtered COS formula.

Rainbow options are a class of financial options where the payoff depends on multiple assets. The COS method for a two-dimensional asset price process is developed in

Chapter 4. Besides, we extend the method to pricing Bermudan options under the Heston stochastic volatility model, where the variance process follows a mean-reverting square root process.

In Chapter 5 we deal with a model from climate change economics. Here the future temperature process and the economic wealth value are correlated stochastic processes. We combine the COS method for stochastic control problems (Chapter 2) and the 2D-COS method (Chapter 4) to solve this optimization problem.

Now suppose we consider a one-dimensional *semilinear* parabolic PDE of the form

$$\frac{\partial v}{\partial t}(t,x)+\mu(t,x)D_x v(t,x)+\tfrac{1}{2}\sigma^2(t,x)D_x^2 v(t,x)+f(t,x,v(t,x),\sigma(t,x)D_x v(t,x))=0,$$
$$(t,x)\in[0,T)\times\mathbb{R}, \tag{1.9a}$$
$$v(T,x)=g(x),\quad x\in\mathbb{R}, \tag{1.9b}$$

where function f may depend on the value function v and its first derivative. This PDE also has a probabilistic representation, by means of the following FSDE and backward stochastic differential equation (BSDE)

$$dX_s=\mu(s,X_s)ds+\sigma(s,X_s)d\omega_s,\quad X_t=x, \tag{1.10a}$$
$$-dY_s=f(s,X_s,Y_s,Z_s)ds-Z_s d\omega_s,\quad Y_T=g(X_T). \tag{1.10b}$$

Contrary to an FSDE, a BSDE is a stochastic differential equation for which a terminal condition, instead of an initial condition, has been specified and its solution consists of a pair of processes, (Y,Z). We refer to Chapter 6 and Chapter 7 for additional assumptions. The generalized Feynman-Kac theorem for semilinear parabolic PDEs gives

$$Y_s=v(s,X_s),\quad Z_s=\sigma(s,X_s)D_x v(s,X_s). \tag{1.11}$$

So, solving the semilinear PDE and the corresponding *decoupled FBSDE* result in the same solution. A PDE can then be solved by either applying numerical discretization techniques or, for FBSDEs, probabilistic numerical methods, like Monte Carlo simulation techniques. We extend the Fourier-based COS method to solve BSDE problems (Chapter 6). For the numerical approximation of the FSDE in Chapter 7 we apply different Taylor schemes, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes, or apply exact simulation. The computation of the conditional expectations appearing relies on the availability of the characteristic function for these discrete schemes. In this way we generalize the applicability of the BCOS method to FSDEs for which the 'continuous' characteristic function is not available.

In Chapter 8 we summarize our main conclusions and give an outlook for future research. Chapters 5, 6, and 7 end with specific appendices and general appendices are placed in Section 9.

CHAPTER 2

On the Fourier Cosine Series Expansion Method for Stochastic Control Problems

2.1. INTRODUCTION

Stochastic optimization can be defined as the optimization of a certain objective function, where an underlying state process is subject to random perturbations. The class of optimization problems can be subdivided into several different types, such as optimal stopping problems or impulse control problems [Pha09]. In this chapter, we focus on stochastic control problems, in which the controller may influence the drift and diffusion terms of the underlying stochastic process. One can derive a partial differential equation (PDE), to be precise the Hamilton-Jacobi-Bellman equation, corresponding to the problem. However, we will use a method based on the dynamic programming principle and solve the problem backwards in time on a fixed time grid.

In financial mathematics, the price of an option can often be formulated as a stochastic optimization problem. In the last decades, financial mathematics has contributed significantly to the theory and the improvement of numerical methods to solve these problems. The techniques employed are closely related to those in the field of real option problems [DP94], encountered in economics, for example. These can also be represented by stochastic optimization problems.

We will generalize the COS technique ([FO08] and [FO09]) to solving stochastic control problems, in which the underlying Lévy process can be controlled. Our method relies on the dynamic programming principle and the COS formula, which is based on Fourier cosine series expansions. A recursive algorithm is developed based on the recursive recovery of the series coefficients. In the backward recursion stage of the algorithm we deal with special matrix structures, namely Hankel and Toeplitz matrices. Matrix-vector products can then be computed efficiently by applying a Fast Fourier Transform (FFT) algorithm, as the resulting matrices are then embedded in a circulant matrix form. Our work builds on both articles [FO08] and [FO09], but it differs on certain points for the accurate treatment of stochastic control problems.

In the resulting method, we need to determine the optimal control law for all possible state values. For that purpose, the value function must be accurately represented in the entire computational domain. It is known, however, that Fourier cosine expansions may

This chapter is based on the articles 'On the Fourier cosine series expansion method for stochastic control problems', published in *Numerical Linear Algebra with Applications*, Vol. 20(4), 598-625, 2013 [ROA13] and 'The COS method for pricing options under uncertain volatility', published in Topics in Numerical Methods for Finance, Vol. 19, 95-113, 2012 [RO12a].

be inaccurate near spatial boundaries, particularly outside the expansion interval. These errors may propagate backwards in time. We give a detailed insight into the source of these local errors and their evolution. Based on this, we propose an extrapolation technique near the domain boundaries as an accurate solution technique in this context.

We test the method by solving two stochastic control problems of practical interest. The first one is the valuation of financial options under uncertain volatility, which was solved in [vdPO12, PFV03, ZW09] by partial differential equation methods for the corresponding nonlinear PDE. The second problem we discuss is a consumption-portfolio problem from economics. The model used is a simplified version of the well-known portfolio-selection problem, which is originally formulated and studied by Merton [Mer69]. Here, an agent allocates his wealth to investments in risky or risk-free assets, and to consumption. The objective is to maximize the expected lifetime utility.

The outline of this chapter is as follows. We start in Section 2.2 with the concepts and notation of stochastic control problems under multidimensional state processes. Then, in Section 2.3, a method based on the dynamic programming principle and the COS formula is derived for solving stochastic control problems with a one-dimensional underlying Lévy process. Section 2.4 provides an extensive error analysis and a way to solve possible propagating errors in the backward recursion. With this, the error converges exponentially in the number of terms in the series expansions for smooth density functions. The two practical examples come up in Section 2.5. Finally, a conclusion for this chapter and our main contributions are presented in Section 2.6.

2.2. STOCHASTIC CONTROL PROBLEMS

We consider the problem class of finite horizon stochastic control problems, where the objective function is optimized over a given finite domain. We start with the notation of control problems and some definitions, based on [Pha09]. The numerical method that we will develop relies on the dynamic programming principle, which is explained in Section 2.2.2.

2.2.1. PROBLEM DESCRIPTION

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $T > 0$ a finite terminal time, $\mathbb{F} = (\mathcal{F}_s)_{0 \le s \le T}$ a filtration satisfying the usual conditions and ω a d-dimensional standard Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. The controlled state process X_t is valued in $\mathbb{R}^n$ and satisfies the stochastic differential equation

$$dX_s = \mu(s, X_s, \boldsymbol{\alpha}_s)ds + \sigma(s, X_s, \boldsymbol{\alpha}_s)d\omega_s, \quad X_0 \text{ given.} \tag{2.2.1}$$

The process here is a controlled diffusion process, which we will use in this section for ease of notation and to stay in line with [Pha09]. Later on, we will work with the class of Lévy processes.

The control process $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_s)_{0 \le s \le T}$ is progressively measurable with respect to $\mathbb{F}$ and is valued in the control set A, a subset of $\mathbb{R}^\ell$. In the class of stochastic optimization problems that we discuss here, the state process is influenced by the control process, $\boldsymbol{\alpha}$, whose value is determined at any time t based on the available information.

The measurable functions $\mu : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^{n \times d}$ satisfy uniform Lipschitz conditions on (t,x) in A. Let $\mathscr{A}$ denote the set of control processes $\boldsymbol{\alpha}$ that satisfy a square integrability condition. With this notation, an element $\boldsymbol{\alpha} \in \mathscr{A}$ is a process over time, with values in set A. With the aforementioned conditions on μ and σ, for all

$\boldsymbol{\alpha} \in \mathscr{A}$ and for any starting condition $(t,x) \in [0,T] \times \mathbb{R}^n$ a unique strong solution to equation (2.2.1) starting from x at $s = t$ exists, which is denoted by $\{X_s^{t,x}, t \le s \le T\}$ [Pha09]. The dependence of $X_s^{t,x}$ on the control process $\boldsymbol{\alpha}$ is omitted for notational convenience. We assume that $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are two measurable functions that satisfy a lower boundedness or a linear growth condition and $\mathbb{E}\left[\int_t^T |f(s, X_s^{t,x}, \boldsymbol{\alpha}_s)| ds\right] < \infty$, $\forall \boldsymbol{\alpha} \in \mathscr{A}$.

The *gain function* on the finite horizon is defined as

$$J(t,x,\boldsymbol{\alpha}) := \mathbb{E}\left[\int_t^T e^{-\rho(s-t)} f(s, X_s^{t,x}, \boldsymbol{\alpha}_s) ds + e^{-\rho(T-t)} g(X_T^{t,x})\right], \tag{2.2.2}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and $\boldsymbol{\alpha} \in \mathscr{A}$. The function f is a so-called *running profit function*, g is a *terminal reward function*, and ρ is a *discount rate*. The objective of the finite horizon problem is to maximize the gain function over all admissible controls in $\mathscr{A}$. We also introduce the so-called *value function*

$$v(t,x) := \sup_{\boldsymbol{\alpha} \in \mathscr{A}} J(t,x,\boldsymbol{\alpha}). \tag{2.2.3}$$

For an initial state $(t,x) \in [0,T) \times \mathbb{R}^n$, we say that $\boldsymbol{\alpha}^* \in \mathscr{A}$ is an optimal control if $v(t,x) = J(t,x,\boldsymbol{\alpha}^*)$. A control process $\boldsymbol{\alpha}$ is called a Markovian control if it has the form $\boldsymbol{\alpha}_s = a(s, X_s^{t,x})$ for some measurable function $a : [0,T] \times \mathbb{R}^n \to A$ [Pha09].

The notion of stochastic control problems can easily be extended with the concepts of optimal stopping or impulse control [ØS07]. Then the controller does not (only) have the disposal of a control process $\boldsymbol{\alpha}$ to optimize his objective, but he can determine the terminal time or can add extra impulses to the state process.

2.2.2. DYNAMIC PROGRAMMING PRINCIPLE

An important principle in the theory of stochastic control is Bellman's optimality principle, also called the dynamic programming principle [Pha09]. It means that if one has taken an optimal control path until some arbitrary observation time θ, then, given this information, it remains optimal to use it after that observation time. The dynamic programming principle is stated as follows:

Result 2.2.1. *(Dynamic programming principle (DPP) (Finite horizon))[Pha09]*
Let $(t,x) \in [0,T] \times \mathbb{R}^n$. Then we have

$$v(t,x) = \sup_{\boldsymbol{\alpha} \in \mathscr{A}} \mathbb{E}\left[\int_t^\theta e^{-\rho(s-t)} f(s, X_s^{t,x}, \boldsymbol{\alpha}_s) ds + e^{-\rho(\theta-t)} v(\theta, X_\theta^{t,x})\right], \tag{2.2.4}$$

for any stopping time $\theta \in [t,T]$. (θ is a stopping time if $\{\theta < t\} \in \mathcal{F}_t$, $\forall t \in [0,T]$, in other words, it should be possible to decide whether or not $\{\theta < t\}$ has occurred on the basis of the knowledge of $\mathcal{F}_t$.)

By the dynamic programming principle, one can split the optimization problem into two parts. An optimal control may be obtained by first searching for an optimal control from a time θ given the state value $X_\theta^{t,x}$, in other words, compute $v(\theta, X_\theta^{t,x})$. Then, the expected value in equation (2.2.4) is maximized over all controls on $[t,\theta]$. We will use this principle to set up a numerical approach for stochastic control problems in Section 2.3.

Remark 2.2.1. *By the dynamic programming principle, one can derive the well-known Hamilton-Jacobi-Bellman (HJB) equation corresponding to problem (2.2.3), see [Pha09]. This*

second-order nonlinear partial differential equation is the infinitesimal version of the dynamic programming principle and reads

$$\frac{\partial v}{\partial t}(t,x) + \sup_{\alpha \in A}[\mathscr{L}^{\alpha} v(t,x) + f(t,x,\alpha)] - \rho v(t,x) = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}^n, \tag{2.2.5}$$

with differential operator of second order

$$\mathscr{L}^{\alpha} v(t,x) = \mu(t,x,\alpha) \cdot D_x v(t,x) + \tfrac{1}{2}\mathrm{Tr}[\sigma\sigma'(t,x,\alpha) D_x^2 v(t,x)], \tag{2.2.6}$$

where $\sigma\sigma'(t,x,\alpha)$ is an $n \times n$ matrix with components $(\sigma\sigma')_{ij} = \sum_{k=1}^{d} \sigma_{ik}\sigma_{jk}$. $\mathscr{L}^{\alpha}$ is called the infinitesimal generator associated to the diffusion X_t with constant control α. The vector D_x denotes the gradient of a function and matrix D_x^2 consists of its second derivatives. The terminal condition is $v(T,x) = g(x), \forall x \in \mathbb{R}^n$, resulting from the definition of the value function.

Stochastic control problems may be solved employing numerical PDE techniques to the corresponding HJB equation. We refer to [FL07] and [PFV03] for numerical discretization methods. Then, issues about convergence to the correct, viscosity solution arise. The viscosity solution concept was introduced by P. L. Lions [Lio83]. We refer to [CIL92] for a general introduction to viscosity solutions and some general uniqueness and existence results. As we will use the dynamic programming approach, we will not go into details about this.

2.3. COS METHOD FOR STOCHASTIC CONTROL PROBLEMS

In this section, we set up a general method to solve stochastic control problems under a one-dimensional *Lévy process*, X_t, for which the characteristic function is known. A Lévy process has stationary, independent increments and is right continuous with left limits. For an extensive overview of Lévy processes we refer to [CT03]. This class contains the constant coefficient jump-diffusion processes. The method is based on the dynamic programming principle and uses the so-called COS formula, which was developed in [FO08] for pricing European options. It results in a recursive algorithm based on the Fast Fourier Transform algorithm. We will explain the COS formula in Section 2.3.1. We start here with the discrete-time framework of the solution method.

Initial time is denoted by t_0 and T is the terminal time. We take a fixed equidistant grid of control times $t_0 < t_1 < \ldots t_m < \ldots < t_M = T$, with $\Delta t := t_{m+1} - t_m$, and a bounded set of possible control values $A \subset \mathbb{R}^{\ell}$. As a discrete approximation, we assume that the control processes are constant during the time intervals $[t_m, t_{m+1}]$. At each control time t_m, with $m < M$, one can choose a control value from the set A, which influences the stochastic process during the time interval $[t_m, t_{m+1}]$. This value is denoted by α_m, where the subscript refers to the control time. The choice may depend on the current value of the state process. With this notation, bold-faced $\boldsymbol{\alpha}$ denotes a control process and α denotes a single control value.

Remark 2.3.1. *For diffusion processes, as in (2.2.1), the stochastic process evolves according to the following dynamics:*

$$dX_s = \mu(s, X_s, \alpha_m)ds + \sigma(s, X_s, \alpha_m)d\omega_s, \quad \text{for } s \in [t_m, t_{m+1}]. \tag{2.3.1}$$

In the examples in Section 2.5 we will use processes of this form, to be precise, geometric Brownian motions. Their log-transformed dynamics belong to the class of Lévy processes, which we consider from now on.

The value function of the discrete-time problem reads

$$v(t,x) = \max_{\boldsymbol{\alpha}\in\hat{\mathscr{A}}} \mathbb{E}\left[\int_t^T e^{-\rho(s-t)} f(s, X_s^{t,x}, \boldsymbol{\alpha}_s) ds + e^{-\rho(T-t)} g(X_T^{t,x})\right]. \tag{2.3.2}$$

$\hat{\mathscr{A}} \subset \mathscr{A}$ denotes the set of all possible control paths $\{\alpha_m\}_{m=0}^{M-1}$, where α_m is valued in the control set A. The terminal condition is $v(T,y) = g(y)$. We deal with a discrete-time stochastic control problem, with M control times. Convergence of the numerical solution to the solution of the original problem (2.2.3) is achieved by increasing the number of time steps (value of M). General convergence results for discrete time problems to their continuous versions can be found in [KD01, DW05, Han05].

The dynamic programming principle now gives:

$$\begin{aligned} v(t_{m-1},x) &= \max_{\alpha_{m-1}\in A} \mathbb{E}^{t_{m-1},x}\left[\int_{t_{m-1}}^{t_m} e^{-\rho(s-t)} f(s, X_s, \alpha_{m-1}) ds + e^{-\rho\Delta t} v(t_m, X_{t_m})\right] \\ &= \max_{\alpha_{m-1}\in A}\left[\int_{t_{m-1}}^{t_m} e^{-\rho(s-t)} \mathbb{E}^{t_{m-1},x}[f(s, X_s, \alpha_{m-1})] ds + e^{-\rho\Delta t}\mathbb{E}^{t_{m-1},x}\left[v(t_m, X_{t_m})\right]\right]. \end{aligned} \tag{2.3.3}$$

For ease of notation, we use the form $\mathbb{E}^{t,x}[X_s]$ instead of $\mathbb{E}[X_s^{t,x}]$. The second equality in (2.3.3) holds by Fubini's theorem. We denote the first term in the maximization operator, i.e., the time integral, by $F(t_{m-1}, x, \alpha_{m-1})$ and call it the *profit function*. We presume that this function is known analytically, or can be approximated using, for example,

$$F(t_{m-1}, x, \alpha_{m-1}) \approx \Delta t f(t_{m-1}, x, \alpha_{m-1}), \tag{2.3.4}$$

or a trapezoidal rule for the time integral, combined with a COS formula or quadrature rule to estimate the expectation. The expectation in the second term, which we call the *continuation value* under control α_{m-1}, is denoted by $c(t_{m-1}, x, \alpha_{m-1})$. So, we use the notation

$$v(t_{m-1},x) = \max_{\alpha_{m-1}\in A}[F(t_{m-1}, x, \alpha_{m-1}) + c(t_{m-1}, x, \alpha_{m-1})]. \tag{2.3.5}$$

2.3.1. FOURIER COSINE EXPANSION FORMULA (COS FORMULA)

Next, we explain our method of choice to approximate the continuation value, in the backward recursion,

$$c(t_{m-1}, x, \alpha) = e^{-\rho\Delta t}\mathbb{E}\left[v(t_m, X_{t_m}) | X_{t_{m-1}} = x\right], \quad \text{with control value } \alpha_{m-1} = \alpha. \tag{2.3.6}$$

We assume a continuous transitional density, which is denoted by $p(y|x,\alpha)$. In other words, $\int_A p(y|x,\alpha)dy = \mathbb{P}(X_{t_m} \in A | X_{t_{m-1}} = x,$ control value $\alpha_{m-1} = \alpha)$, $\forall$ Borel subsets $A \in \mathbb{R}$. We omit the dependence on Δt for notational convenience. We rewrite

$$c(t_{m-1}, x, \alpha) = e^{-\rho\Delta t}\int_{\mathbb{R}} v(t_m, y) p(y|x,\alpha) dy. \tag{2.3.7}$$

The numerical method is based on series expansions of the value function at the next time level and the density function, as we will show below. The resulting equation is called the COS formula, due to the use of Fourier *cos*ine series expansions. Fourier series expansions and their convergence properties have been discussed in [Boy01]. In the derivation of the COS formula, we distinguish three different approximation steps.

Step 1: We assume that the integrand decays to zero as $y \to \pm\infty$, which is usual in the control problems we work on. Because of that, we can truncate the infinite integration range of the expectation to some interval $[a,b] \subset \mathbb{R}$ without losing significant accuracy. This gives the approximation

$$c_1(t_{m-1},x,\alpha;[a,b]) = e^{-\rho\Delta t}\int_a^b v(t_m,y)\,p(y|x,\alpha)\,dy. \tag{2.3.8}$$

The notation c_i is used for the different approximations of c and keeps track of the numerical errors that set in from each step, which are discussed in Section 2.4.1. For final approximations we also use the 'hat'-notation, like $\hat{c}$, $\hat{v}$, etc.

Step 2: Next, we consider the Fourier cosine series expansions of the density function and the value function on $[a,b]$:

$$p(y|x,\alpha) = \sum_{k=0}^{\infty}{}' \mathscr{P}_k(x,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right), \tag{2.3.9a}$$

$$\text{and}\quad v(t_m,y) = \sum_{k=0}^{\infty}{}' \mathscr{V}_k(t_m)\cos\left(k\pi\tfrac{y-a}{b-a}\right), \tag{2.3.9b}$$

with series coefficients $\{\mathscr{P}_k\}_{k=0}^{\infty}$ and $\{\mathscr{V}_k\}_{k=0}^{\infty}$ given by

$$\mathscr{P}_k(x,\alpha) = \tfrac{2}{b-a}\int_a^b p(y|x,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \tag{2.3.10a}$$

$$\text{and}\quad \mathscr{V}_k(t_m) = \tfrac{2}{b-a}\int_a^b v(t_m,y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy, \tag{2.3.10b}$$

respectively. $\sum'$ in (2.3.9) indicates that the first term in the summation is weighted by one-half. Replacing the density function by its Fourier cosine series, interchanging summation and integration, using the definition of coefficients $\mathscr{V}_k$, and truncating the series summation, we obtain the next approximation

$$c_2(t_{m-1},x,\alpha;[a,b],N) = \tfrac{b-a}{2}e^{-\rho\Delta t}\sum_{k=0}^{N-1}{}' \mathscr{P}_k(x,\alpha)\mathscr{V}_k(t_m). \tag{2.3.11}$$

Step 3: The coefficients $\mathscr{P}_k(x,\alpha)$ can now be approximated as follows

$$\begin{aligned}\mathscr{P}_k(x,\alpha) &\approx \tfrac{2}{b-a}\int_{\mathbb{R}} p(y|x,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy\\ &= \tfrac{2}{b-a}\Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\middle|x,\alpha\right)e^{-ik\pi\frac{a}{b-a}}\right\} := \Phi_k(x,\alpha).\end{aligned} \tag{2.3.12}$$

$\Re\{.\}$ denotes taking the real part of the input argument. $\varphi(.|x,\alpha)$ is the conditional *characteristic function* of X_{t_m}, given $X_{t_{m-1}} = x$ and $\alpha_{m-1} = \alpha$. The density function of a stochastic process is usually not known, but often its characteristic function is known (see [DPS00, FO08]). For Lévy processes the characteristic function can be represented by the Lévy-Khintchine formula ([CT03]) and there holds

$$\varphi(u|x,\alpha) = \varphi(u|0,\alpha)e^{iux} := \phi_{levy}(u|\alpha)e^{iux}. \tag{2.3.13}$$

Inserting the above equations into (2.3.11) gives us the *COS formula* for approximation of $c(t_{m-1}, x, \alpha)$:

$$\begin{aligned}\hat{c}(t_{m-1}, x, \alpha; [a,b], N) &:= c_3(t_{m-1}, x, \alpha; [a,b], N) \\ &= \tfrac{b-a}{2} e^{-\rho \Delta t} \sum_{k=0}^{N-1}{}' \Phi_k(x, \alpha) \mathcal{V}_k(t_m) \\ &= e^{-\rho \Delta t} \sum_{k=0}^{N-1}{}' \Re\left\{ \phi_{levy}\left(\tfrac{k\pi}{b-a} \middle| \alpha\right) e^{ik\pi \frac{x-a}{b-a}} \right\} \mathcal{V}_k(t_m). \end{aligned} \tag{2.3.14}$$

Since the terms $\mathcal{V}_k(t_m)$ are independent of x, we can calculate the continuation value for many values of x simultaneously.

The value function is now approximated by

$$\hat{v}(t_{m-1}, x) := \max_{\alpha_{m-1} \in A} [F(t_{m-1}, x, \alpha_{m-1}) + \hat{c}(t_{m-1}, x, \alpha_{m-1})]. \tag{2.3.15}$$

Remark 2.3.2. *(Density function) Conditions for the existence of a continuous density are given in [CT03]. We have the following normalizing property*

$$\int_a^b p_N(y|x, \alpha) dy = \int_a^b p(y|x, \alpha) dy, \tag{2.3.16}$$

where p_N denotes the approximation of the density function by the Fourier cosine series with N terms, equation (2.3.9a). The integral equals one for sufficiently wide interval $[a,b]$. There holds $\lim_{\Delta t \to 0} p(y|x,\alpha) = \delta(x)$. The absolute value of $p_N(y|x,\alpha)$ is bounded by $\sum_{k=0}^{N-1}{}' |\Phi_k(x,\alpha)| \leq \frac{2N}{b-a}$, but converges to the true value if N goes to infinity. So, if Δt decreases, we need to increase N in order to achieve the same accuracy.

2.3.2. RECURSION FORMULA FOR COEFFICIENTS $\mathcal{V}_k(t_m)$

The algorithm for solving stochastic control problem (2.3.2) is based on the recursive recovery of the coefficients $\mathcal{V}_k(t_m)$, starting with the coefficients at the terminal time:

$$\mathcal{V}_k(t_M) = \tfrac{2}{b-a} \int_a^b v(T, y) \cos\left(k\pi \tfrac{y-a}{b-a}\right) dy, \tag{2.3.17}$$

for which often an analytic solution is available. This is the case for, among others, exponential and polynomial terminal reward functions. If these coefficients are not known analytically, they can be approximated by numerical integration rules or discrete Fourier cosine transforms (see Appendix C). These coefficients are used for the approximation of the continuation value at time t_{M-1}.

Next we consider the coefficients that are used to approximate the continuation value at time t_{m-1}, i.e., $\mathcal{V}_k(t_m)$, for $m \leq M-1$. The value function, equation (2.3.5), at time t_m appears in the terms $\mathcal{V}_k(t_m)$ and we need to find an optimal control law for all state values $y \in [a,b]$. We propose two techniques for this:

- Firstly, suppose that the set A of possible control values is finite, or an infinite set is approximated by a finite one, $A = \{\alpha^1, \ldots, \alpha^q, \ldots, \alpha^Q\}$, where Q is a finite number and $\alpha^i \in \mathbb{R}^\ell$. Then it may be possible to determine subdomains $\mathcal{D}_m^q \subset [a,b]$, so that for each $y \in \mathcal{D}_m^q$ it is optimal to choose control α_m^q at control time t_m. The subscript of

α_m^q indicates the time level and the superscript the control value. To be precise, we find the value(s) x^* for which an optimal policy changes to another one, i.e.,

$$F(t_m, x^*, \alpha_m^p) + \hat{c}(t_m, x^*, \alpha_m^p) = F(t_m, x^*, \alpha_m^q) + \hat{c}(t_m, x^*, \alpha_m^q). \tag{2.3.18}$$

Newton's method can be applied to accurately determine these values x^* and specify the subdomains $\mathcal{D}_m^q$. This approach will be applied to the problem in Section 2.5.1.

- If this procedure cannot be applied, or in case the control set A is a continuous range, then the interval $[a,b]$ is divided into Q subintervals, which span the interval $[a,b]$. They are denoted by $\mathcal{D}_m^q$, $q = 1,2,\ldots,Q$, where Q is a finite number. On each subinterval, $\mathcal{D}_m^q$, we determine the optimal control, α_m^q, for the time interval $[t_m, t_{m+1}]$. One can use an optimization algorithm or, for a finite control set, compute the value function for all possible control values on a grid to determine the maximum. We here assume that the control value is constant over the spatial subinterval. With many subintervals this approximation is sufficiently accurate. This piecewise constant control timestepping approach is also described in [Kry99, Kry00] and will be used in the application in Section 2.5.2.

In both approaches, we split the integral for the definition of $\mathcal{V}_k$ into different parts:

$$\begin{aligned}\mathcal{V}_k(t_m) &= \tfrac{2}{b-a}\sum_{q=1}^{Q}\int_{\mathcal{D}_m^q} F(t_m, y, \alpha_m^q)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &\quad + \tfrac{2}{b-a}\sum_{q=1}^{Q}\int_{\mathcal{D}_m^q} c(t_m, y, \alpha_m^q)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &:= \sum_{q=1}^{Q}\mathcal{F}_k(t_m, \mathcal{D}_m^q, \alpha_m^q) + \sum_{q=1}^{Q}\mathcal{C}_k(t_m, \mathcal{D}_m^q, \alpha_m^q), \qquad (m \neq M).\end{aligned} \tag{2.3.19}$$

Here assume that the terms $\mathcal{F}_k$ are known analytically, or can be approximated sufficiently accurate. In practical applications these terms may be independent of time. The coefficients $\mathcal{C}_k$ at time t_m can be approximated by using the coefficients $\mathcal{V}_k(t_{m+1})$ from the next time level, as we will explain shortly. This results in a backward recursion of the coefficients $\mathcal{V}_k(t_m)$.

For approximation of the value function at time t_{M-2} we need the coefficients $\mathcal{V}_k(t_{M-1})$. We will use the approximated values, $\hat{c}(t_{M-1}, y, \alpha)$, to approximate the terms $\mathcal{C}_k(t_{M-1}, z_1, z_2, \alpha)$. This approximation is then denoted by $\hat{\mathcal{C}}_k(t_{M-1}, z_1, z_2, \alpha)$. On the integrands of terms $\hat{\mathcal{C}}_k$ we again apply the Fourier cosine series expansion by inserting equation (2.3.14):

$$\begin{aligned}\hat{\mathcal{C}}_k(t_{M-1}, z_1, z_2, \alpha) &:= \tfrac{2}{b-a}\int_{z_1}^{z_2}\hat{c}(t_{M-1}, y, \alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &= e^{-\rho\Delta t}\Re\left\{\sum_{j=0}^{N-1}{}'\phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)\mathcal{V}_j(t_M)\mathcal{M}_{k,j}(z_1, z_2)\right\},\end{aligned} \tag{2.3.20}$$

where the elements of matrix $\mathcal{M}(z_1, z_2)$ are given by:

$$\mathcal{M}_{k,j}(z_1, z_2) := \tfrac{2}{b-a}\int_{z_1}^{z_2} e^{ij\pi\frac{y-a}{b-a}}\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{2.3.21}$$

The parameters of the matrices $\mathcal{M}$ are the boundary values of their respective integration ranges. Finally, we end up with the vector form

$$\hat{\mathcal{V}}(t_{M-1}) = \sum_{q=1}^{Q} \mathcal{F}(t_{M-1}, \mathcal{D}^q_{M-1}, \alpha^q_{M-1}) + \sum_{q=1}^{Q} e^{-\rho\Delta t} \Re\left\{\mathcal{M}(\mathcal{D}^q_{M-1})\mathbf{w}^q\right\}, \tag{2.3.22}$$

where $\mathbf{w}$ is a vector with entries

$$\mathbf{w}^q = \{w^q_j\}_{j=0}^{N-1}, \quad \text{with} \quad w^q_j = \phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha^q_{M-1}\right)\mathcal{V}_j(t_M), \quad w^q_0 = \tfrac{1}{2}\phi_{levy}(0|\alpha^q_{M-1})\mathcal{V}_0(t_M). \tag{2.3.23}$$

For the other coefficients, $\mathcal{V}_k(t_m)$, $1 \le m \le M-2$, the approximations $\hat{c}(t_m, y, \alpha)$ *and* $\hat{\mathcal{V}}_k(t_{m+1})$ will be used to approximate the terms $\mathcal{C}_k(t_m, z_1, z_2, \alpha)$. The same arguments give the following numerical approximation of the Fourier cosine coefficients at time t_m:

$$\hat{\mathcal{V}}(t_m) = \sum_{q=1}^{Q} \mathcal{F}(t_m, \mathcal{D}^q_m, \alpha^q_m) + \sum_{q=1}^{Q} e^{-\rho\Delta t} \Re\left\{\mathcal{M}(\mathcal{D}^q_m)\hat{\mathbf{w}}^q\right\}, \qquad (m = 1, \ldots, M-2), \tag{2.3.24}$$

where $\hat{\mathbf{w}}$ is a vector with entries

$$\hat{\mathbf{w}}^q = \{\hat{w}^q_j\}_{j=0}^{N-1}, \quad \text{with} \quad \hat{w}^q_j = \phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha^q_m\right)\hat{\mathcal{V}}_j(t_{m+1}), \quad \hat{w}^q_0 = \tfrac{1}{2}\phi_{levy}(0|\alpha^q_m)\hat{\mathcal{V}}_0(t_{m+1}). \tag{2.3.25}$$

An additional error is introduced because the coefficients are approximated using the approximated elements $\hat{\mathcal{V}}_j(t_{m+1})$. We will examine this evolving error in Sections 2.4.2 and 2.4.3 and propose a more accurate approximation for the Fourier coefficients $\mathcal{V}_k(t_m)$ in Section 2.4.2.

2.3.3. ALGORITHM

The matrix-vector products $\mathcal{M}\mathbf{w}$ in the terms $\hat{\mathcal{C}}$ can be computed by a Fourier-based algorithm, as stated in the next result:

Result 2.3.1. *(Efficient computation of $\hat{\mathcal{C}}(t_m, z_1, z_2, \alpha)$)*
The matrix-vector product $\mathcal{M}(z_1, z_2)\mathbf{w}$ can be computed in $\mathcal{O}(N\log_2 N)$ operations, with the help of the Fast Fourier Transform (FFT) algorithm.

The key insight of this efficient computation is the equality

$$\mathcal{M}_{k,j}(z_1, z_2) = -\tfrac{i}{\pi}\left(\mathcal{M}^c_{k,j}(z_1, z_2) + \mathcal{M}^s_{k,j}(z_1, z_2)\right), \tag{2.3.26}$$

where matrix $\mathcal{M}^c$ is a Hankel matrix and $\mathcal{M}^s$ a Toeplitz matrix ($\mathcal{M}^c_{i,j} = \mathcal{M}^c_{i-1,j+1}$ and $\mathcal{M}^s_{i,j} = \mathcal{M}^s_{i+1,j+1}$). The matrices $\mathcal{M}^c_{k,j}$ and $\mathcal{M}^s_{k,j}$ can be found in Appendix B. The special matrix structure enables us to use the FFT algorithm for the matrix-vector products. If a process does not possess the property in equation (2.3.13), the FFT algorithm cannot be employed in a straightforward way (see [ZO13a]).

We can recover the terms $\hat{\mathcal{V}}_k(t_m)$ recursively, starting with $\mathcal{V}_k(t_M)$. The algorithm to solve the discrete-time stochastic control problem (2.3.2) backwards in time then reads:

Algorithm 1. *(COS method for stochastic control problems)*
Initialization: *Calculate coefficients* $\mathcal{V}_k(t_M)$ *for* $k = 0, 1, \ldots, N-1$.
Main loop to recover $\hat{V}(t_m)$***:***
For $m = M-1$ *to* 1*:*

- *Determine the subdomains* $\mathcal{D}_m^q$ *for which the optimal control value is* α_m^q*, or determine the optimal control values* α_m^q *for given subdomains* $\mathcal{D}_m^q$.
- *Compute* $\hat{V}(t_m)$*, equation (2.3.22) or (2.3.24), with the help of the FFT algorithm.*

Final step: *Compute* $\hat{v}(t_0, x_0)$ *by inserting* $\hat{V}_k(t_1)$ *into equation (2.3.14).*

The computational complexity of the algorithm is $\mathcal{O}\left(MQN\log_2 N\right)$, as we need to compute M time steps, and Q subintervals. The computation time also depends on the efficiency of the optimization method to find the optimal control α_m^q.

Remark 2.3.3. *We elaborate on the differences between using the COS method for pricing Bermudan and barrier options and for solving stochastic control problems. In Algorithm 1, we search for an optimal control law for all state values* y *in the computational domain* $[a, b]$*. For this, the numerical continuation values need to be accurate over the entire interval. Errors may arise, however, in the vicinity of the boundaries and propagate backwards in time, as we will show in Sections 2.4.1 and 2.5.2. In Section 2.4.2, a remedy for this problem will be proposed. When pricing barrier and Bermudan call options, as in [FO09], one only searches for the early-exercise points, where the continuation value equals the payoff. For this task, interval* $[a, b]$ *can be chosen large so that the early-exercise points are not close to the interval boundaries and the boundary errors do not affect the resulting price.*

Another difference in applying the COS method to stochastic control problems is the dependency of the characteristic function on the control values. As this function is evaluated often during the optimization procedure, it may be time-consuming.

2.3.4. INTERMEZZO: EUROPEAN OPTIONS

In this section we discuss an important application from financial mathematics, the pricing of a European option. In this financial setting, the asset price at time t is denoted by S_t. The risk-neutral option pricing formula ([Shr08]) for a European option with payoff function $g(.)$ reads

$$v(t_0, x) = e^{-r\Delta t}\mathbb{E}^{t_0,x}\left[v(T, X_T)\right] = e^{-r\Delta t}\int_{\mathbb{R}} g(y)p(y|x)dy, \tag{2.3.27}$$

where X_t is the state process, which can be any monotone function of the underlying asset price S_t. We take the log-asset price process $X_t = \log S_t$ here. r is the risk-neutral interest rate and $\Delta t := T - t_0$. The problem considered is a simplified, reduced version [1] of stochastic control problem (2.2.3). $p(y|x)$ is the conditional probability density of X_T, given $X_{t_0} = x$. In other words, the option price is equal to the expected value of its discounted future payoff, under a certain probability measure.

[1] Note that there is no control process $\boldsymbol{\alpha}$ and no running profit function.

Explicit expressions for probability density functions $p(y|x)$ encountered in finance are often not known, or involve some mathematical special functions, which make them impractical to calculate. Instead the characteristic function $\varphi(u|x)$ corresponding to $p(y|x)$ is often known [DPS00] and for Lévy processes we can use the equality

$$\varphi(u|x) = \varphi(u|0)e^{iux} := \phi_{levy}(u)e^{iux}. \tag{2.3.28}$$

At first [Hes93] found a closed-form solution for European options with stochastic volatility by means of the Fourier transform. In [CM99], the Fourier transform of the damped payoff function, together with a Fast Fourier Transform (FFT), was used to evaluate European options under a broad class of models. Fourier methods for Bermudan options were then developed in, among others, [O'S05, LFBO08]. For recent developments in pricing exotic options, like Asian and multi-asset options, we refer to [FMM11, Sur09, ZO13b, RO12b].

In this thesis we consider a specific Fourier technique, the COS method. Fourier cosine series expansion of the payoff function gives the approximation:

$$g_N(y) = \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_M)\cos\left(k\pi\tfrac{y-a}{b-a}\right), \tag{2.3.29}$$

in which

$$\mathcal{V}_k(t_M) = \tfrac{2}{b-a}\int_a^b g(y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{2.3.30}$$

An analytic solution for the coefficients at the terminal time is available for the options we discuss in this chapter. The COS formula yields

$$\hat{v}(t_0,x) := e^{-r\Delta t}\sum_{k=0}^{N-1}{}' \Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}\mathcal{V}_k(t_M). \tag{2.3.31}$$

The authors of [FO08] provide the following rule-of-thumb for the computational domain for European options

$$[a,b] = \left[\kappa_1 - L\sqrt{\kappa_2+\sqrt{\kappa_4}}, \kappa_1 + L\sqrt{\kappa_2+\sqrt{\kappa_4}}\right], \qquad L\in[8,10], \tag{2.3.32}$$

where $\kappa_1,\kappa_2,\ldots$ are the cumulants of the underlying stochastic process. For the cumulants of the log-jump-diffusion and Variance Gamma process, we refer to [FO08].

If N, the number of terms in the cosine expansion, is chosen sufficiently large, then a larger computational domain should not affect the option price. For numerical experiments on European options we refer to Sections 2.4.1.1 and 3.4.2.2.

2.4. ERROR ANALYSIS AND EXTRAPOLATION TECHNIQUE

In this section, we analyze the error of the COS method for stochastic control problems and base our analysis on [FO08], [FO09], and [Rui10]. Errors are introduced by the COS formula and by evolution through time via the coefficients $\hat{\mathcal{V}}_k(t_m)$ and a possibly incorrect control process α. We start with the local error of the COS formula where backward recursion of the approximated coefficients $\hat{\mathcal{V}}_k(t_m)$ and control are not taken into account, in Section

2.4.1. In the financial context, this corresponds to a European option with an uncontrolled asset price process. The local error may be significant in the vicinity of the boundaries. We give an example in which the COS formula is inaccurate in the vicinity of a domain boundary (Section 2.4.1.1). This may give difficulties during the recursive recovery of the Fourier cosine coefficients $\mathcal{V}_k(t_m)$. In Section 2.4.2, we propose an improved approximation for $\mathcal{V}_k(t_m)$, which is in that case more accurate than $\hat{\mathcal{V}}_k(t_m)$ from equation (2.3.24). Finally, the propagating error in the backward recursion is studied and bounded (Section 2.4.3).

2.4.1. LOCAL ERROR COS FORMULA

We define the local error of the COS formula for the continuation value by

$$\epsilon_{COS}(t_{m-1},x,\alpha;[a,b],N) := c(t_{m-1},x,\alpha) - \hat{c}(t_{m-1},x,\alpha;[a,b],N). \tag{2.4.1}$$

The above notation includes the parameters used for the approximations, namely $[a,b]$ and N. The error

$$\max_{\alpha_{m-1}\in A} |\epsilon_{COS}(t_{m-1},x,\alpha_{m-1};[a,b],N)| \tag{2.4.2}$$

bounds the absolute error of the approximated value function $\hat{v}(t_{m-1},x)$, assuming that the correct optimal control law has been chosen and that the function $F(t_{m-1},x,\alpha)$ is known analytically, or can be approximated sufficiently accurate.

We first assume that the terms $\mathcal{V}_k(t_m)$ are exact. Errors are introduced in three steps. An upper bound for the error of the European option pricing COS formula with respect to the truncation range and the convergence rate, in dependence of N, has been derived in [FO08]. We discuss the errors one after the other.

Step 1: The integration range truncation error:

$$\begin{aligned}\epsilon_1(t_{m-1},x,\alpha;[a,b]) &:= c(t_{m-1},x,\alpha) - c_1(t_{m-1},x,\alpha;[a,b])\\ &= e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]} v(t_m,y)p(y|x,\alpha)dy.\end{aligned} \tag{2.4.3}$$

If $v(t_m,y)p(y|x,\alpha)$ is sufficiently small outside the interval $[a,b]$, then the error ϵ_1 can be ignored.

Step 2: The series truncation error on $[a,b]$:

$$\begin{aligned}\epsilon_2(t_{m-1},x,\alpha;[a,b],N) &:= c_1(t_{m-1},x,\alpha;[a,b]) - c_2(t_{m-1},x,\alpha;[a,b],N)\\ &= \tfrac{b-a}{2}e^{-\rho\Delta t}\sum_{k=N}^{\infty}\mathscr{P}_k(x,\alpha)\mathcal{V}_k(t_m)\\ &= e^{-\rho\Delta t}\int_a^b v(t_m,y)\left[p(y|x,\alpha) - p_N(y|x,\alpha)\right]dy.\end{aligned} \tag{2.4.4}$$

The functions $v_N(t_m,y)$ and $p_N(y|x,\alpha)$ denote the Fourier cosine series expansions of the value function and the density function, using N terms in the series summations.

The convergence rate of Fourier cosine series depends on the properties of the approximated functions in the expansion interval. Information about different convergence types can be found in [Boy01]. Based on that theory, we find that the series truncation error converges exponentially for density functions in the class $C^\infty([a,b])$. A density function with discontinuity in one of its derivatives results in an algebraic convergence.

Step 3: The error related to approximating $\mathscr{P}_k(x,\alpha)$ by $\Phi_k(x,\alpha)$ (see equation (2.3.12)):

$$\begin{aligned}\epsilon_3(t_{m-1},x,\alpha;[a,b],N) &:= c_2(t_{m-1},x,\alpha;[a,b],N) - c_3(t_{m-1},x,\alpha;[a,b],N)\\ &= \tfrac{b-a}{2}e^{-\rho\Delta t}\sum_{k=0}^{N-1}{}'(\mathscr{P}_k(x,\alpha)-\Phi_k(x,\alpha))\mathscr{V}_k(t_m)\\ &= -e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]}\left[\sum_{k=0}^{N-1}{}'\mathscr{V}_k(t_m)\cos\left(k\pi\tfrac{y-a}{b-a}\right)\right]p(y|x,\alpha)\,dy\\ &= -e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]}v_N(t_m,y)\,p(y|x,\alpha)\,dy. \end{aligned} \tag{2.4.5}$$

Remark 2.4.1. *Note that the Fourier cosine series expansions used in Section 2.3.1 are defined for $y\in[a,b]$, whereas here the function $v_N(t_m,y)$ is evaluated on $\mathbb{R}\setminus[a,b]$. Here we denote by function $v_N(t_m,y)$, for $y\in\mathbb{R}\setminus[a,b]$, the infinitely periodic, even extension of the Fourier cosine series expansion outside the expansion interval. This value will usually be different from $v(t_m,y)$, even if N tends to infinity.*

The integration range truncation error, ϵ_1, enters by truncation of the infinite domain to the finite domain $[a,b]$. Conversely, error ϵ_3 is due to replacing the finite domain by an infinite domain in equation (2.3.12). The third error 'compensates', completely or partly, for the first error. Addition of both errors gives

$$\begin{aligned}&\epsilon_1(t_{m-1},x,\alpha;[a,b]) + \epsilon_3(t_{m-1},x,\alpha;[a,b],N) =\\ &\qquad e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]}\left[v(t_m,y)-v_N(t_m,y)\right]p(y|x,\alpha)\,dy.\end{aligned} \tag{2.4.6}$$

So, error $\epsilon_1+\epsilon_3$ results from using $e^{-\rho\Delta t}v_N(t_m,y)$ instead of the true discounted value function $e^{-\rho\Delta t}v(t_m,y)$. We can write

$$\begin{aligned}&\hat{c}(t_{m-1},x,\alpha;[a,b],N) =\\ &\qquad e^{-\rho\Delta t}\int_a^b v(t_m,y)\,p_N(y|x,\alpha)\,dy + e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]}v_N(t_m,y)\,p(y|x,\alpha)\,dy.\end{aligned} \tag{2.4.7}$$

We can write the local error of the COS formula as

$$\begin{aligned}&\epsilon_{COS}(t_{m-1},x,\alpha;[a,b],N) =\\ &\qquad \epsilon_1(t_{m-1},x,\alpha;[a,b]) + \epsilon_2(t_{m-1},x,\alpha;[a,b],N) + \epsilon_3(t_{m-1},x,\alpha;[a,b],N).\end{aligned} \tag{2.4.8}$$

If, for given x, the integration interval $[a,b]$ is chosen sufficiently wide, then the series truncation error ϵ_2 dominates the overall local error. This implies that for smooth density functions the local error converges exponentially to zero, otherwise it converges algebraically.

For a given interval $[a,b]$, the local error may however be large if x is in the vicinity of the domain boundaries, resulting from error $\epsilon_1+\epsilon_3$. We will show this by an example in the next section. A local error may propagate via the backward recursion for a stochastic control problem.

2.4.1.1. ERROR IN THE VICINITY OF THE BOUNDARIES

Here, an example of a large error close to the spatial boundaries is presented, when we use the COS formula to price a European call option (see Section 2.3.4). We model the asset

price by a geometric Brownian motion,

$$dS_s = rS_s ds + \bar{\sigma} S_s d\omega_s, \quad S_0 \text{ given.} \tag{2.4.9}$$

The payoff of a call option at terminal time, T, with log-asset price $y = \log(S_T)$, is given by the function

$$g(y) = (e^y - K)^+, \tag{2.4.10}$$

where $(z)^+ := \max(z, 0)$ and K denotes the strike price. The Fourier cosine coefficients of the payoff function are given by

$$\begin{aligned} \mathcal{V}_k(t_M) &= \tfrac{2}{b-a} \int_a^b g(y) \cos\left(k\pi \tfrac{y-a}{b-a}\right) dy \\ &= \tfrac{2}{b-a} \left(\chi_k(\log K, b, a, b) - K\psi_k(\log K, b, a, b)\right), \qquad (a \le \log K \le b). \end{aligned} \tag{2.4.11}$$

The analytic solution of the functions χ_k and ψ_k can be found in Appendix A. The COS formula yields

$$\begin{aligned} \hat{v}(t_0, x; [a,b], N) &= e^{-r\Delta t} {\sum_{k=0}^{N-1}}' \Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\right) e^{ik\pi \frac{x-a}{b-a}}\right\} \mathcal{V}_k(t_M) \\ &= e^{-r\Delta t} \int_{[a,b]} g(y) p(y|x) dy + e^{-r\Delta t} \int_{\mathbb{R}\setminus[a,b]} g_N(y) p(y|x) dy - \epsilon_2(t_0, x; [a,b], N). \end{aligned} \tag{2.4.12}$$

We take N sufficiently large, so that error ϵ_2 can be neglected. The equation above shows that a significant error is introduced if $g_N(y)p(y|x)$ is not close to $g(y)p(y|x)$ outside the expansion interval $[a,b]$, which is the case in the example to follow.

For the call option under geometric Brownian motion an analytic solution is available, i.e., the Black-Scholes price [BS73], so that the numerical option value can be compared with the exact solution. The following parameters are used for the tests in this section:

$$K = 100,\ S_0 = 100\,(x_0 \approx 4.6),\ r = 0.1,\ T = 0.1,\ \bar{\sigma} = 0.25,\ L = 10,\ N = 2^{10}. \tag{2.4.13}$$

Log-asset price, x, is varied and the results are shown in Figure 2.4.1. In the left-side plot, function $g(y)$ and the series expansion $g_N(y)$ and its infinitely periodic, even extension outside $[a,b]$, which is symmetric in a and b, are presented. The COS formula 'sees' the latter function, see equation (2.4.12). Function $g_N(y)$ resembles the true payoff function well at the left-hand side of a, as the function is constant there. Hence, no error is introduced in the vicinity of that boundary. However, at the right-side of b a difference between the two functions is observed, which gives a significant error $\epsilon_1 + \epsilon_3$ *in* the computational domain at that boundary. This is shown in the right-side plot, where the exact Black-Scholes price and the COS approximation are presented.

The numerical option value for initial log-asset price, x_0, is highly accurate, with an error less than 10^{-14}. However, inaccuracies at some place in the domain $[a,b]$ may seriously affect the backward recursion of the Fourier coefficients, $\mathcal{V}_k(t_m)$, when a stochastic control problem is solved recursively.

2.4.2. IMPROVEMENT BY EXTRAPOLATION

The coefficients $\mathcal{V}_k(t_m)$ are recovered recursively, backwards in time. In that case, the local error, ϵ_{COS}, described in the previous section, may propagate through time. Here, we

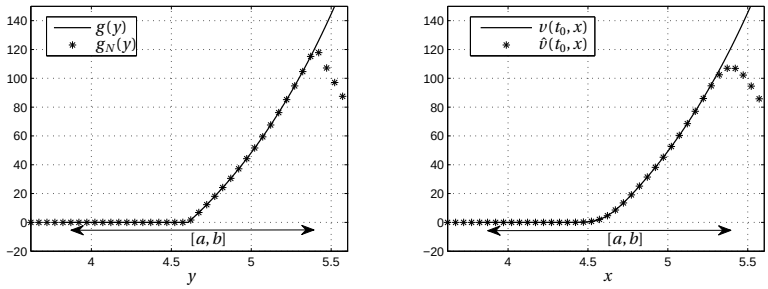


Figure 2.4.1: Example of an error, $\epsilon_1 + \epsilon_3$, close to one of the boundaries.

propose a technique to deal with this issue. In Section 2.4.3, we prove that the error of the approximated Fourier coefficients converges exponentially in N, for probability density functions in the class $C^\infty([a,b])$.

Recall that the approximated Fourier cosine coefficients at time t_{m-1} are given by

$$\hat{\mathcal{V}_k}(t_{m-1}) = \frac{2}{b-a}\int_a^b \left[\max_{\alpha_{m-1}\in A}[F(t_{m-1},x,\alpha_{m-1}) + \hat{c}(t_{m-1},x,\alpha_{m-1})]\right]\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx. \quad (2.4.14)$$

From this definition it follows that inaccurate numerical continuation values at time t_{m-1} may affect the choice of the optimal control value and the coefficients, which in turn affect the continuation value at time t_{m-2}, and so on.

The idea to deal with the propagating error is to determine the area in which inaccurate approximate values from the COS method occur. In this area, we employ an *extrapolation technique* to compute a value with improved accuracy, using accurate numerical continuation values from the neighboring region.

In practical applications, it may be possible to estimate the area in which $\hat{c}(t_{m-1},x,\alpha)$ is inaccurate, assuming that coefficients $\mathcal{V}_k(t_m)$ are exact. The density function, together with the value function, give the desired information. For instance, suppose we can calculate a value x_L, so that the continuation value is well approximated for $x \in [x_L, b]$ and is inaccurate for $x \in [a, x_L]$ [2]. The continuation function $c(t_{m-1},x,\alpha)$, on $[a,x_L]$, can then be approximated by an extrapolation technique. For this, we employ a second-order Taylor expansion in x_L:

$$\hat{c}^{ex}(t_{m-1},x,\alpha) := \hat{c}(t_{m-1},x_L,\alpha) + \hat{c}_x(t_{m-1},x_L,\alpha)(x-x_L) + \tfrac{1}{2}\hat{c}_{xx}(t_{m-1},x_L,\alpha)(x-x_L)^2. \quad (2.4.15)$$

The derivatives can easily be computed in this setting, as:

$$\hat{c}_x(t_{m-1},x_L,\alpha) = e^{-\rho\Delta t}\sum_{k=0}^{N-1}{}'\,\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\middle|\alpha\right)e^{ik\pi\frac{x_L-a}{b-a}}\,\tfrac{ik\pi}{b-a}\right\}\mathcal{V}_k(t_m), \quad (2.4.16a)$$

$$\hat{c}_{xx}(t_{m-1},x_L,\alpha) = e^{-\rho\Delta t}\sum_{k=0}^{N-1}{}'\,\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\middle|\alpha\right)e^{ik\pi\frac{x_L-a}{b-a}}\left(\tfrac{ik\pi}{b-a}\right)^2\right\}\mathcal{V}_k(t_m). \quad (2.4.16b)$$

[2] The methodology above can also be applied if the approximated continuation value is inaccurate in a certain area $[x_R, b]$.

We denote the extrapolated continuation value by

$$\tilde{c}(t_{m-1},x,\alpha) := \begin{cases} \hat{c}^{ex}(t_{m-1},x,\alpha), & \text{for } x\in[a,x_L], \\ \hat{c}(t_{m-1},x,\alpha), & \text{for } x\in[x_L,b]. \end{cases} \tag{2.4.17}$$

The local error of the COS formula *with* extrapolation technique is denoted by

$$\tilde{\epsilon}_{COS}(t_{m-1},x,\alpha;[a,b],N) := c(t_{m-1},x,\alpha) - \tilde{c}(t_{m-1},x,\alpha;[a,b],N), \tag{2.4.18}$$

and we have:

$$\tilde{\epsilon}_{COS}(t_{m-1},x,\alpha;[a,b],N) = \mathcal{O}((x-x_L)^3), \quad \text{for } x\in[a,x_L]. \tag{2.4.19}$$

We use continuation value $\tilde{c}$ to determine the optimal control law and to approximate the terms $\mathcal{C}_k$ by

$$\tilde{\mathcal{C}}_k(t_{m-1},z_1,z_2,\alpha) := \tfrac{2}{b-a}\int_{z_1}^{z_2} \tilde{c}(t_{m-1},y,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{2.4.20}$$

The corresponding Fourier coefficients are denoted by $\tilde{\mathcal{V}}_k(t_{m-1})$. Suppose the interval $[z_1,z_2]\subset[a,b]$ can be divided into $[z_1,x_L]$ and $[x_L,z_2]$. Then the Fourier cosine coefficients read

$$\begin{aligned}\tilde{\mathcal{C}}_k(t_{m-1},z_1,z_2,\alpha) = {} & \tfrac{2}{b-a}\int_{z_1}^{x_L} \hat{c}^{ex}(t_{m-1},x_L,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ & + e^{-\rho\Delta t}\Re\left\{\sum_{j=0}^{N-1}{}' \phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)\mathcal{V}_j(t_m)\mathcal{M}_{k,j}(x_L,z_2)\right\},\end{aligned} \tag{2.4.21}$$

where the analytic solution to the first part can be found with the help of functions ψ_k, ξ_k, and ξ_k^2 (see Appendix A). The extrapolation technique can be improved by using a higher-order Taylor expansion. In that case, a similar approach can be applied, as the derivatives of approximated continuation values can be computed easily.

Remark 2.4.2. *If we know that the continuation value should be of exponential form, which is sometimes the case for the type of problems we are interested in, it may be more accurate to use an exponential extrapolation. We will use this form in an example in Section 2.5.2.*

2.4.3. ERROR PROPAGATION IN THE BACKWARD RECURSION

In this section, we study the error of the Fourier cosine coefficients in the backward recursion. We start with the algorithm *without the extrapolation technique* and define

$$\varepsilon_k(t_m,z_1,z_2,\alpha) := \mathcal{C}_k(t_m,z_1,z_2,\alpha) - \hat{\mathcal{C}}_k(t_m,z_1,z_2,\alpha). \tag{2.4.22}$$

The terms $\mathcal{F}_k$ are assumed to be exact, or are approximated sufficiently accurate, so that the error in the Fourier coefficients is given by

$$\varepsilon_k(t_m) := \mathcal{V}_k(t_m) - \hat{\mathcal{V}}_k(t_m) = \sum_{q=1}^{Q} \varepsilon_k(t_m,\mathcal{D}_m^q,\alpha_m^q). \tag{2.4.23}$$

We assume that an accurate, true control law [3] is found, which is denoted by α_m^*. Here the dependency on the state value is omitted. The error may be larger in case of incorrect control values.

[3] which is an optimistic assumption if the continuation value is inaccurate.

At time lattice $M-1$ we have:

$$\begin{aligned}\varepsilon_k(t_{M-1},z_1,z_2,\alpha) &= \tfrac{2}{b-a}\int_{z_1}^{z_2}(c(t_{M-1},y,\alpha)-\hat{c}(t_{M-1},y,\alpha))\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &= \tfrac{2}{b-a}\int_{z_1}^{z_2}\epsilon_{COS}(t_{M-1},y,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \end{aligned} \tag{2.4.24}$$

Here we omit the dependency of the error on the interval $[a,b]$ and N. Coefficients $\mathcal{V}_k(t_M)$ are assumed to be known analytically, so that the only error introduced by the COS formula is the local error, ϵ_{COS}. This error may be significant close to the domain boundaries a and b. We end up with

$$\varepsilon_k(t_{M-1}) = \tfrac{2}{b-a}\int_a^b \epsilon_{COS}(t_{M-1},y,\alpha^*_{M-1})\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{2.4.25}$$

Note that the terms $\varepsilon_k(t_{M-1})$ are the Fourier cosine coefficients of the local error at time t_{M-1} on the expansion interval $[a,b]$.

For the approximation of $c(t_{M-2},x,\alpha)$ the COS formula with coefficients $\hat{\mathcal{V}}_k(t_{M-1})$ is used. The approximated value is here denoted by $\overline{c}(t_{M-2},x,\alpha)$. The use of the approximations $\hat{\mathcal{V}}_k(t_{M-1})$ in equation (2.3.24) gives rise to an additional error in $\mathcal{C}_k(t_{M-2},z_1,z_2,\alpha)$:

$$\varepsilon_k(t_{M-2},z_1,z_2,\alpha) = \tfrac{2}{b-a}\int_{z_1}^{z_2}(c(t_{M-2},y,\alpha)-\overline{c}(t_{M-2},y,\alpha))\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy, \tag{2.4.26}$$

with $\overline{c}$ obtained by inserting $\hat{\mathcal{V}}_k(t_{M-1})$ into the COS formula:

$$\begin{aligned}\overline{c}(t_{M-2},y,\alpha) &= e^{-\rho\Delta t}{\sum_{j=0}^{N-1}}'\Re\left\{\phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)e^{ij\pi\frac{y-a}{b-a}}\right\}\hat{\mathcal{V}}_j(t_{M-1}) \\ &= e^{-\rho\Delta t}{\sum_{j=0}^{N-1}}'\Re\left\{\phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)e^{ij\pi\frac{y-a}{b-a}}\right\}(\mathcal{V}_j(t_{M-1})-\varepsilon_j(t_{M-1})) \\ &= \hat{c}(t_{M-2},y,\alpha) - e^{-\rho\Delta t}{\sum_{j=0}^{N-1}}'\Re\left\{\phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)e^{ij\pi\frac{y-a}{b-a}}\right\}\varepsilon_j(t_{M-1}). \end{aligned} \tag{2.4.27}$$

The error in the coefficients $\mathcal{C}_k$ can now be separated into two parts:

$$\begin{aligned}\varepsilon_k(t_{M-2},z_1,z_2,\alpha) &= \tfrac{2}{b-a}\int_{z_1}^{z_2}(c(t_{M-2},y,\alpha)-\hat{c}(t_{M-2},y,\alpha) \\ &\quad +\hat{c}(t_{M-2},y,\alpha)-\overline{c}(t_{M-2},y,\alpha))\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &= \tfrac{2}{b-a}\int_{z_1}^{z_2}(\epsilon_{COS}(t_{M-2},y,\alpha)+\overline{\epsilon}(t_{M-2},y,\alpha))\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy, \end{aligned} \tag{2.4.28}$$

where

$$\overline{\epsilon}(t_{M-2},y,\alpha) = e^{-\rho\Delta t}{\sum_{j=0}^{N-1}}'\Re\left\{\phi_{levy}\left(\tfrac{j\pi}{b-a}\middle|\alpha\right)e^{ij\pi\frac{y-a}{b-a}}\right\}\varepsilon_j(t_{M-1}). \tag{2.4.29}$$

$\overline{\epsilon}$ resembles the COS formula (2.3.14), now with Fourier coefficients $\varepsilon_j(t_{M-1})$. For suffi-

ciently large values of N we find

$$\begin{aligned}\overline{\epsilon}(t_{M-2},y,\alpha) &= e^{-\rho\Delta t}\int_{[a,b]}\epsilon_{COS}(t_{M-1},z,\alpha^*_{M-1})\,p(z|y,\alpha)\,dz\\ &\quad + e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]}\hat{\epsilon}_{COS}(t_{M-1},z,\alpha^*_{M-1})\,p(z|y,\alpha)\,dz\\ &\approx e^{-\rho\Delta t}\mathbb{E}^{t_{M-2},y}[\epsilon_{COS}(t_{M-1},X_{t_{M-1}},\alpha^*_{M-1})],\end{aligned}\tag{2.4.30}$$

where $\hat{\epsilon}_{COS}$ is the Fourier cosine expansion of the local error with N terms. So,

$$\begin{aligned}c(t_{M-2},y,\alpha)-\overline{c}(t_{m-1},y,\alpha) &\approx \epsilon_{COS}(t_{M-2},y,\alpha)\\ &\quad + e^{-\rho\Delta t}\mathbb{E}^{t_{M-2},y}[\epsilon_{COS}(t_{M-1},X_{t_{M-1}},\alpha^*_{M-1})].\end{aligned}\tag{2.4.31}$$

The recursive algorithm gives

$$\begin{aligned}\varepsilon_k(t_m) \approx \tfrac{2}{b-a}\int_a^b &(\epsilon_{COS}(t_m,y,\alpha^*_m)\\ &+ e^{-\rho\Delta t}\mathbb{E}^{t_m,y}[\epsilon_{COS}(t_{m+1},X_{t_{m+1}},\alpha^*_{m+1})])\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy.\end{aligned}\tag{2.4.32}$$

The first part of the error is due to the use of the COS formula at time t_m. The possible propagation of errors from time level t_{m+1} causes the second part to appear. The value of the expectation depends on the drift and diffusion of the stochastic process[4]. A clarifying example is provided in Section 2.5.2.1. It is not possible to bound the error $\varepsilon_k(t_m)$ by increasing the number of terms in the Fourier series expansions, N, as the error $\epsilon_1+\epsilon_3$ remains. The propagation of the error may give rise to incorrect results of the algorithm.

Next we discuss the error convergence if we employ the extrapolation methodology from Section 2.4.2. The error in the terms $\mathscr{C}_k$ is redefined by

$$\varepsilon_k(t_m,z_1,z_2,\alpha) := \mathscr{C}_k(t_m,z_1,z_2,\alpha)-\tilde{\mathscr{C}}_k(t_m,z_1,z_2,\alpha)\tag{2.4.33}$$

and

$$\varepsilon_k(t_m) := \mathscr{V}_k(t_m)-\tilde{\mathscr{V}}_k(t_m) = \sum_{q=1}^{Q}\varepsilon_k(t_m,\mathscr{D}^q_m,\alpha^q_m).\tag{2.4.34}$$

The error of the approximated Fourier coefficients now converges exponentially in N under certain conditions:

Result 2.4.1. *With a sufficiently accurate extrapolation technique, with $[a,b]\subset\mathbb{R}$ chosen sufficiently wide and a probability density function p in $C^\infty([a,b])$, error $\varepsilon_k(t_m)$ converges exponentially in N for $1\le m\le M-1$.*

The proof of this result is similar to that for pricing Bermudan options, which can be found in [FO09]. The difference is that we explained how errors in the vicinity of the boundaries may propagate. Note that error $\epsilon_1+\epsilon_3$ has now been reduced by using an extrapolation technique. It can also be proved that if the local error converges algebraically, then so does $\varepsilon(t_m)$.

[4] A large error ϵ_{COS} at the left-hand side of the computational domain may 'disappear' by a large positive drift. It may travel through the domain for other drift terms.

2.5. NUMERICAL EXPERIMENTS

In this section, we apply Algorithm 1 to two different stochastic control problems. In the first example in Section 2.5.1, we calculate the price of financial options under uncertain volatility.

The second example, Section 2.5.2, deals with an optimal consumption-portfolio problem, with which we can demonstrate the impact of an error near the spatial boundaries, and its propagation. The continuous-time variant of this stochastic control problem admits an analytic solution. This problem is thus instructive as we can show the propagation of a local error and the improvement by extrapolation of the continuation value.

2.5.1. FINANCIAL OPTIONS UNDER UNCERTAIN VOLATILITY

The model we use for pricing financial options under uncertain volatility is based on the problem described in [PFV03, ZW09]. The setting of this problem is the financial option market. The risk-neutral dynamics of the asset price is assumed to evolve according to either a geometric Brownian motion (GBM),

$$dS_s = rS_s ds + \boldsymbol{\alpha}_s S_s d\omega_s, \quad S_0 \text{ given.} \tag{2.5.1}$$

or Merton's jump-diffusion process,

$$dS_s = (r - \lambda\kappa)S_s ds + \boldsymbol{\alpha}_s S_s d\omega_s + (e^J - 1)S_s dq_s, \quad S_0 \text{ given.} \tag{2.5.2}$$

Here $\kappa := \mathbb{E}[e^J - 1]$, r is the risk-neutral interest rate, and q_s is an $\mathcal{F}_s$-adapted Poisson process with intensity rate λ. The jumps J are normally distributed with mean μ_J and standard deviation σ_J. $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_s)_{0\le s\le T}$ represents an uncertain volatility process, which is valued in the bounded set $A = [\alpha^-, \alpha^+]$.

We consider the worst-case scenario for an investor with a long position in a European-style option. Then, the option value at time t reads:

$$v(t,S) = \inf_{\boldsymbol{\alpha}\in\mathcal{A}} J(t,S,\boldsymbol{\alpha}) = \inf_{\boldsymbol{\alpha}\in\mathcal{A}} \mathbb{E}^{t,S}[e^{-r(T-t)} g(S_T)], \tag{2.5.3}$$

where $g(.)$ is a prescribed payoff function. The pricing problem is now formulated as a stochastic control problem. Note that there is no running profit function included in this problem formulation, but only a terminal reward function. Contrary to the problem in equation (2.2.3), the infimum over gain functions $J(.)$ is taken. However, similar theory and solution algorithms can be developed for minimization problems.

Remark 2.5.1. *The corresponding Hamilton-Jacobi-Bellman (HJB) equations read*

$$\frac{\partial v}{\partial t}(t,S) + \min_{\alpha\in[\alpha^+,\alpha^-]} \left[rS\frac{\partial v}{\partial S}(t,S) + \tfrac{1}{2}\alpha^2 S^2 \frac{\partial^2 v}{\partial S^2}(t,S) \right] - r\,v(t,S) = 0, \qquad \forall (t,S) \in [0,T)\times\mathbb{R}_+ \tag{2.5.4}$$

under GBM, and

$$\frac{\partial v}{\partial t}(t,S) + \min_{\alpha\in[\alpha^+,\alpha^-]} \left[(r-\lambda\kappa)S\frac{\partial v}{\partial S}(t,S) + \tfrac{1}{2}\alpha^2 S^2 \frac{\partial^2 v}{\partial S^2}(t,S) + \lambda\mathbb{E}[v(t,e^J S) - v(t,S)] \right] - r\,v(t,S) = 0, \qquad \forall (t,S) \in [0,T)\times\mathbb{R}_+, \tag{2.5.5}$$

under the jump-diffusion process. This yields

$$\begin{aligned} &\text{if} \quad \frac{\partial^2 v}{\partial S^2} \leq 0 \quad \Rightarrow \text{take } \alpha = \alpha^+, \\ &\text{if} \quad \frac{\partial^2 v}{\partial S^2} > 0 \quad \Rightarrow \text{take } \alpha = \alpha^-, \end{aligned} \tag{2.5.6}$$

This allows us to restrict the set of possible control values to $A = \{\alpha^-, \alpha^+\}$. We see that the control value, that is the volatility, is a function of the Greek $\Gamma = \partial^2 v/\partial S^2$. The same partial differential equation as (2.5.4) is derived for a transaction costs model in [Lel85].

The valuation of financial options under uncertain volatility was solved in [vdPO12] and [PFV03] by fully implicit discretization methods for the corresponding nonlinear partial differential equation. In [PFV03], an implicit discretization method for the governing PDE to solve a similar problem has been applied. Numerical experiments were performed using the butterfly spread and digital options. It was demonstrated that a non-monotone scheme may lead to incorrect, non-viscosity solutions. The authors in [ZW09] combined an exponentially fitted finite volume method for the space direction and an implicit scheme in time for the PDE. The method was tested by using butterfly spread, double barrier call, and digital call options. The option prices in these papers will serve as references to which we will compare our results.

As before, we consider an equidistant time grid, $t_0, t_1, \ldots, t_M = T$, with $\Delta t := t_m - t_{m-1}$. In the numerical approximation a constant volatility $\alpha_m \in \{\alpha^-, \alpha^+\} = A$ is applied within the time intervals $[t_m, t_{m+1}]$. The choice depends on the current asset price.

We switch to log-asset price processes, $X_s := \log S_s$, that belong to the class of Lévy processes. For GBM we then deal with the Brownian motion

$$dX_s = (r - \tfrac{1}{2}\alpha_m^2)ds + \alpha_m d\omega_s, \quad \text{for } s \in [t_m, t_{m+1}], \tag{2.5.7}$$

whereas the log-jump-diffusion process reads

$$dX_s = (r - \lambda\kappa - \tfrac{1}{2}\alpha_m^2)ds + \alpha_m d\omega_s + J dq_s, \quad \text{for } s \in [t_m, t_{m+1}]. \tag{2.5.8}$$

The value function of the discrete-time problem reads

$$v(t,x) = \min_{\boldsymbol{\alpha} \in \hat{\mathscr{A}}} \mathbb{E}^{t,x}\left[e^{-r(T-t)} g(e^{X_T})\right]. \tag{2.5.9}$$

$\hat{\mathscr{A}} \subset \mathscr{A}$ denotes the set of all possible control paths $\{\alpha_m\}_{m=0}^{M-1}$, where α_m is valued in the control set A. The dynamic programming principle gives:

$$\begin{aligned} v(t_{m-1}, x) &= \min_{\alpha_{m-1} \in A} e^{-r\Delta t} \mathbb{E}[v(t_m, X_{t_m}) | X_{t_{m-1}} = x, \alpha_{m-1}] \\ &= \min\left[c(t_{m-1}, x, \alpha^-),\, c(t_{m-1}, x, \alpha^+)\right], \end{aligned} \tag{2.5.10}$$

for which we use the COS formula (2.3.14). For Brownian motion, equation (2.5.7), the characteristic function reads

$$\phi_{levy}(u|\alpha) = \exp(iu(r - \tfrac{1}{2}\alpha^2)\Delta t - \tfrac{1}{2}u^2\alpha^2\Delta t) \tag{2.5.11}$$

and for the log-jump-diffusion process, equation (2.5.8), we have

$$\phi_{levy}(u|\alpha) = \exp(iu(r - \lambda\kappa - \tfrac{1}{2}\alpha^2)\Delta t - \tfrac{1}{2}u^2\alpha^2\Delta t) e^{\lambda\Delta t(\exp(i\mu_J u - \frac{1}{2}u^2\sigma_J^2) - 1)}. \tag{2.5.12}$$

Algorithm 1 is used to solve the pricing problem. The coefficients at the terminal time are known analytically for the financial options we use. For the other time levels the Fourier coefficients are approximated by

$$\hat{V}_k(t_m) = \frac{2}{b-a}\int_a^b \min\left[\hat{c}(t_m,y,\alpha^-),\hat{c}(t_m,y,\alpha^+)\right]\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{2.5.13}$$

We can divide the integration interval $[a,b]$ into subdomains $\mathcal{D}_m^-$ and $\mathcal{D}_m^+$, for which the optimal control values at control time t_m are α_m^- and α_m^+, respectively. For this we use Newton's method to determine the value(s) x^* for which the optimal policy changes to another one, i.e.,

$$\hat{c}(t_m,x^*,\alpha^-) = \hat{c}(t_m,x^*,\alpha^+). \tag{2.5.14}$$

We test the COS method by pricing butterfly, digital, and bull split-strike combo options under uncertain volatility. In the numerical experiments we use $T = 0.25$, $S_0 = 100$, and for the jumps

$$\lambda = 0.1,\ \mu_J = -0.90, \text{ and } \sigma_J = 0.45. \tag{2.5.15}$$

These parameters are approximately the same as those reported in [AA00] using European call options on the S&P 500 stock index in April of 1999 (see [dFV05, KFV09]). For the control values we use here $\alpha^- = 0.15$, $\alpha^+ = 0.25$.

The numerical method converges in M, the number of time steps, to the solution of the original problem (2.5.3). We will use a 4-point Richardson-extrapolation scheme on the option values with small M to obtain more accurate values. This method is used in [FO09] to approximate American option values with the help of a few Bermudan option prices. Let $\hat{v}(t_0,x_0;M)$ denote the option value with M time steps. We calculate the extrapolated value, $\hat{v}_R(t_0,x_0;M)$, by

$$\hat{v}_R(t_0,x_0;M) := \tfrac{1}{21}\left[64\hat{v}(t_0,x_0;8M) - 56\hat{v}(t_0,x_0;4M) + 14\hat{v}(t_0,x_0;2M) - \hat{v}(t_0,x_0;M)\right]. \tag{2.5.16}$$

Note that this is another extrapolation method than described in Section 2.4.2.

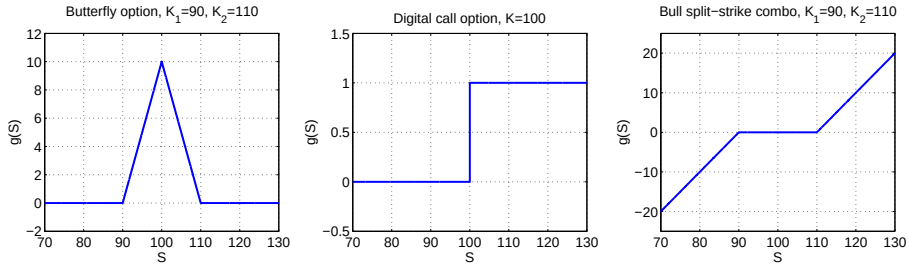


Figure 2.5.1: Payoff functions butterfly, digital call, and bull option.

2.5.1.1. BUTTERFLY OPTION

The payoff function of a butterfly option at terminal time T is given by:

$$g(S) = (S-K_1)^+ - 2\left(S-\tfrac{K_1+K_2}{2}\right)^+ + (S-K_2)^+, \tag{2.5.17}$$

for certain strike prices K_1 and K_2, see Figure 2.5.1. We switch to the log-asset price domain and the coefficients at the terminal time are known analytically

$$\begin{aligned}\mathcal{V}_k(t_M) &= \tfrac{2}{b-a}\int_a^b g(e^y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy\\ &= \tfrac{2}{b-a}\Big[\chi_k(\log K_1,b,a,b) - 2\chi_k\left(\log\tfrac{K_1+K_2}{2},b,a,b\right) + \chi_k(\log K_2,b,a,b)\\ &\quad - K_1\psi_k(\log K_1,b,a,b) + (K_1+K_2)\psi_k\left(\log\tfrac{K_1+K_2}{2},b,a,b\right)\\ &\quad + K_2\psi_k(\log K_2,b,a,b)\Big], \qquad (a\le\log K_1,\log K_2\le b),\end{aligned} \tag{2.5.18}$$

see Appendix A for the analytic solution to the functions χ_k and ψ_k.

In the numerical experiment we use the following additional model parameters:

$$r = 0.1,\ K_1 = 90,\ \text{and}\ K_2 = 110. \tag{2.5.19}$$

2

For the integration interval under GBM we take $[a,b] = [\log 70, \log 130]$, similar as in [vdPO12]. For the problem with jump-diffusion we need a larger computation domain and $[a,b] = [\log 10, \log 200]$ suffices.

It is worth mentioning that we do not need to use the coefficients $\tilde{\mathcal{V}}_k$, with extrapolated continuation values $\tilde{c}$, from Section 2.4.2. The reason for this is that the value function converges to zero if the log-asset price goes to plus or minus infinity. So, for sufficiently large intervals $[a,b]$, the value function on time lattice t_{m+1} is zero outside the expansion interval. Then, by assuming that N is chosen sufficiently large, the function $v_N(t_{m+1},x)$ is also accurate just outside $[a,b]$ and the local error of the COS formula at time t_m is small for all $y\in[a,b]$. Because of this, extrapolation for the continuation value is not necessary.

Table 2.5.1: Results butterfly option under uncertain volatility.

(a) GBM.

M	$\hat{v}(t_0,x_0)$
60	2.3178
120	2.3078
240	2.3027
480	2.3002
960	2.2990
$\hat{v}_R(t_0,x_0;120)$	2.2977

(b) CPU time (s).

M	N				
	400	600	800	1000	1200
60	0.29	0.35	0.43	0.51	0.58
120	0.50	0.61	0.74	0.86	1.05
240	0.97	1.18	1.38	1.63	1.84
480	1.92	2.33	2.73	3.19	3.58
960	3.85	4.67	5.39	6.29	7.13

(c) Jump-diffusion.

M	$\hat{v}(t_0,x_0)$
60	2.2859
120	2.2760
240	2.2710
480	2.2685
960	2.2672
$\hat{v}_R(t_0,x_0;120)$	2.2660

Geometric Brownian Motion In Tables 2.5.1a and 2.5.1b, results for different values of M and N are shown. The option values have converged in N up to nine decimal places, for $N \geq 400$. Increasing the number of control times, M, gives convergence to the true option value. The results are highly satisfactory and match the prices in [PFV03, vdPO12, ZW09]. The computation time is linear in M and $\mathcal{O}(N \log_2 N)$ in the number of terms in the Fourier cosine series expansions.

Jump-Diffusion Process Table 2.5.1c presents the results for the option pricing problem where the underlying asset price is a jump-diffusion. The only difference compared to the GBM is the usage of another characteristic function and a different computational domain size. Again the COS method performs highly satisfactorily. This demonstrates the applicability of the COS method for the broad class of Lévy processes.

2.5.1.2. DIGITAL OPTION

The payoff of a digital call option is given by

$$g(S) = \begin{cases} 1, & \text{for} \quad S > K, \\ 0, & \text{for} \quad S < K, \end{cases} \tag{2.5.20}$$

for certain strike price K. We switch to the log-asset price domain and the terminal coefficients for the digital call option are given by

$$\mathcal{V}_k(t_M) = \tfrac{2}{b-a}\psi_k(\log K, b, a, b), \qquad (a \leq \log K \leq b). \tag{2.5.21}$$

For the functions χ_k and ψ_k we refer to Appendix A. For the tests, we take $K = 100$ and $r = 0.05$. Figure 2.5.1 shows the corresponding payoff function. We take N sufficiently large, so that convergence of the series approximation is reached.

Geometric Brownian Motion We use the same spatial domain as in [ZW09], namely $[a,b] = [\log 50, \log 160]$, and compare our results. Because we are dealing with constant option values at both boundaries, interval $[a,b]$ can be chosen sufficiently large and extrapolation of the continuation value is not required for accuracy.

The results are presented in Table 2.5.2 and they are very similar to the prices in [ZW09, PFV03]. Also the extrapolated values $\hat{v}_R(t_0, x_0; M)$, by means of a 4-point Richardson-extrapolation scheme, are accurate. The option value, the Delta ($\partial v/\partial S$) and Gamma ($\partial^2 v/\partial S^2$) of the option value, and the control domains are shown in Figure 2.5.2.

Table 2.5.2: Digital call under uncertain volatility, GBM.

M	100	200	400	800	1600
$\hat{v}(t_0, x_0)$	0.452690	0.449547	0.447311	0.445723	0.444597
$\hat{v}_R(t_0, x_0; M)$	0.443707	0.443169			

Jump-Diffusion Process We end this section with a digital option pricing problem where the underlying asset prices is a jump-diffusion. Table 2.5.3 presents the results for a digital call option. Again the COS method performs highly satisfactorily.

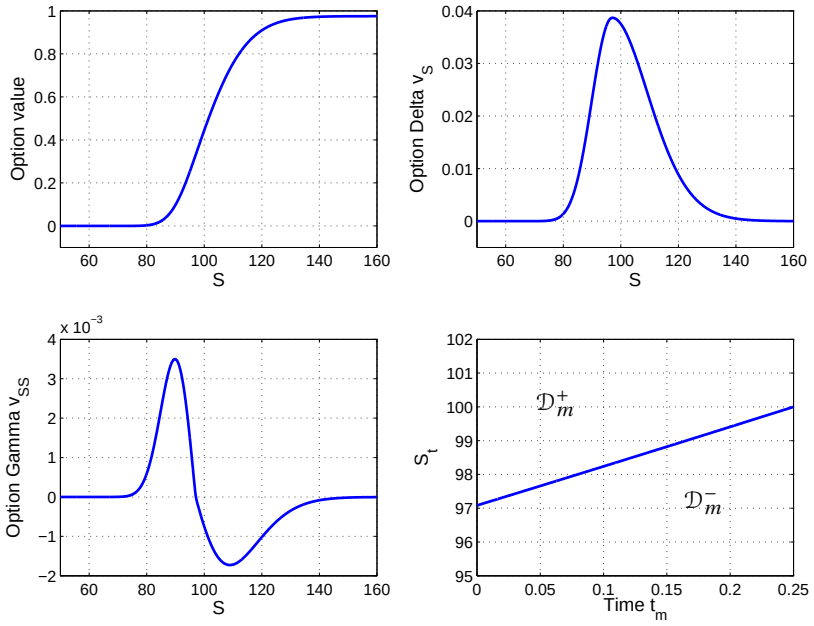


Figure 2.5.2: Digital call under uncertain volatility (GBM, $M = 1600$).

Table 2.5.3: Digital call under uncertain volatility, jump-diffusion.

M	100	200	400	800	1600
$\hat{v}(t_0,x_0)$	0.4523049	0.4492045	0.4469982	0.4454318	0.4443220
$\hat{v}_R(t_0,x_0;M)$	0.4434415	0.4429120			

2.5.1.3. BULL OPTION

The bull split-strike combo (bull) option is a combination of a short position in a put option with strike K_1 and a long position in a call with strike K_2, where $K_1 < K_2$. The payoff function (see Figure 2.5.1) is given by

$$g(S) = \begin{cases} S-K_1, & \text{for} \quad S \le K_1, \\ 0, & \text{for} \quad K_1 \le S \le K_2, \\ S-K_2, & \text{for} \quad K_2 \le S. \end{cases} \tag{2.5.22}$$

Again we switch to the log-asset price domain. The terminal coefficients for the bull split-strike combo option are given by

$$\begin{aligned} \mathcal{V}_k(t_M) = \tfrac{2}{b-a}\Big[&\chi_k(\log K_2, b, a, b) - \psi_k(\log K_2, b, a, b) \\ &- \psi_k(\log K_1, b, a, b) + \chi_k(\log K_1, b, a, b)\Big], \qquad (a \le \log K_1, \log K_2 \le b). \end{aligned} \tag{2.5.23}$$

For the tests, we take $K_1 = 90$, $K_2 = 110$, and $r = 0.05$. The periodic, even extension of the payoff function outside the interval $[a,b]$ (equation (2.4.12)) may cause problems at both

ends of the computational domain. Because of that, we now use extrapolation by a second-order Taylor expansion (see Section 2.4.2) to determine accurate values of the continuation values close to both boundaries a and b. The function $p(y|x,\alpha)$ represents a normal density function with distribution

$$\mathcal{N}\left(x+(r-\tfrac{1}{2}\alpha^2)\Delta t, \alpha^2\Delta t\right). \tag{2.5.24}$$

We can presume that the continuation value is well approximated on $[x_L, x_R]$, with

$$x_L := a-(r-\tfrac{1}{2}(\alpha^+)^2)\Delta t+5\alpha^+\sqrt{\Delta t}, \tag{2.5.25a}$$

$$x_R := b-(r-\tfrac{1}{2}(\alpha^+)^2)\Delta t-5\alpha^+\sqrt{\Delta t}. \tag{2.5.25b}$$

We take a computation domain $[a,b]=[\log 50, \log 160]$. Table 2.5.4 presents the option values. Figure 2.5.3 shows the option value calculated without and with extrapolation technique for the continuation values. The true value is calculated by using a larger computational domain, so that errors in the vicinity of the boundaries do not affect the option values in the domain $S_0 \in [50,160]$.

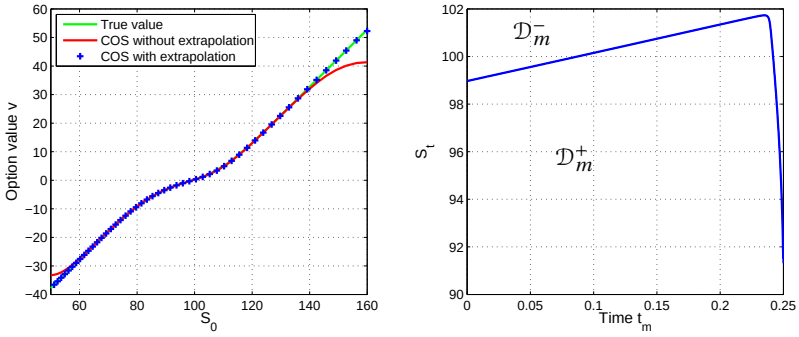


Figure 2.5.3: Bull option under uncertain volatility (GBM).

Table 2.5.4: Bull option under uncertain volatility, with extrapolation, GBM.

M	100	200	400	800	1600
$\hat{v}(t_0,x_0)$	0.2342475	0.2313446	0.2298836	0.2291501	0.2287825
$\hat{v}_R(t_0,x_0;M)$	0.2284146	0.2284142			

Remark 2.5.2. *(Incorrect control) In Result 2.4.1, we deduced an exponentially converging error in N, assuming that the correct control law was found. In practice, we may however find incorrect control values for small values of N and the convergence result therefore holds for sufficiently large value N. In the previous experiments we took N sufficiently large, so that the COS formula for the continuation values was sufficiently accurate. If we would choose N too small, the COS formula may give incorrect control values. This is demonstrated in Figure 2.5.4, where the left-hand side picture shows the continuation values for both control choices with $N=2^5$, which is apparently too small. The right-hand side picture shows the difference between the two functions for $N=2^5$ and a sufficiently large value $N=2^9$. This function is used to determine the optimal volatility. It is clear that we find incorrect control values for*

small values of N and the convergence result in Result 2.4.1 only holds for sufficiently large N.

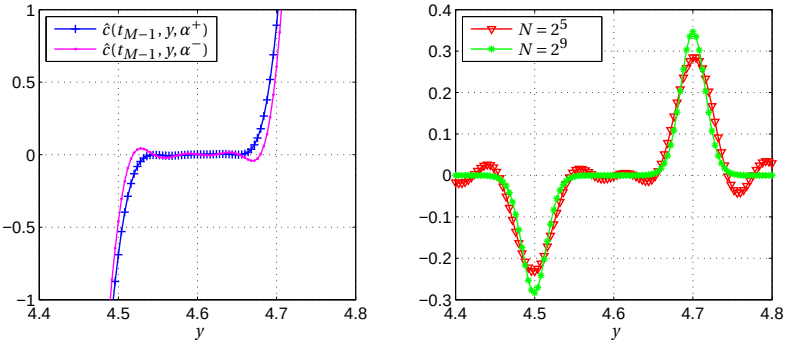


Figure 2.5.4: Left: continuation values with small $N = 2^5$, right: difference between two continuation values for small $N = 2^5$ and high $N = 2^9$ (bull option).

Remark 2.5.3. *(Convergence in M) We approximate the continuous-time stochastic control problem (2.5.3) by its discrete-time variant (2.5.9), with a piecewise constant control policy over time. Convergence to the continuous solution is obtained by increasing the number of time steps, M, and, if desired, by a Richardson extrapolation method in M. The option values converge monotonically in M. The digital call prices converge with order 1/2 in M, see Tables 2.5.2 and 2.5.3, whereas the butterfly and bull option converges with order 1 (Tables 2.5.1 and 2.5.4). Most probably the properties of the payoff function give rise to different convergence rates. However, more research is needed to understand this behavior.*

Remark 2.5.4. *(PDE methods) The COS method converges exponentially in N to the solution of the discrete-time control problem (2.5.9). On the other hand, the PDE methods in [PFV03, vdPO12, ZW09] solve the continuous-time problem and employ space and time grids, with step sizes ΔS and Δt, respectively. The fully implicit discretization schemes in [PFV03, ZW09] converge unconditionally to the viscosity solution of the corresponding PDE, as $\Delta S, \Delta t \to 0$.*

2.5.2. OPTIMAL CONSUMPTION PATH

The second example we discuss is a simplified version[5] of Merton's optimal consumption-portfolio problem [Mer69]. Here, an agent consumes a proportion of his wealth and invests the remaining part in assets, with rate of return, $\bar{\mu}$, and fixed volatility, $\bar{\sigma}$. The dynamics of the invested capital are given by

$$dK_s = \bar{\mu} K_s ds + \bar{\sigma} K_s d\omega_s. \tag{2.5.26}$$

Let $\alpha_s W_s$ denote the amount of wealth consumed at time s, with α_s the control process and wealth W_s. Taking into account equation (2.5.26) and consumption gives

$$\begin{aligned} dW_s &= \tfrac{W_s}{K_s} dK_s - \alpha_s W_s ds \\ &= (\bar{\mu} - \alpha_s) W_s ds + \bar{\sigma} W_s d\omega_s. \end{aligned} \tag{2.5.27}$$

[5] This example is also a simplified version of the two-dimensional climate-economics model in Chapter 5.

The agent chooses his consumption to maximize his expected discounted utility of consumption over a finite time horizon with terminal time T. The optimal consumption problem can be represented by the stochastic control problem:

$$v(t,W) = \max_{\boldsymbol{\alpha}\in\mathcal{A}} \mathbb{E}^{t,W}\left[\int_t^T e^{-\rho(s-t)}\mathcal{U}(\boldsymbol{\alpha}_s W_s)\,ds + e^{-\rho(T-t)}\mathcal{U}(W_T)\right], \tag{2.5.28}$$

where W is the current wealth level, and $\rho \geq 0$ is the utility discount rate [Coc05].

The *utility function*, $\mathcal{U}(.)$, measures the utility gain of consumption $\boldsymbol{\alpha}_s W_s$. We presume a constant relative risk aversion (CRRA) utility function

$$\mathcal{U}(C) = C^{1-\eta}/(1-\eta), \qquad \eta \geq 0, \eta \neq 1. \tag{2.5.29}$$

The exact solution to this continuous-time control problem can be found by the corresponding HJB equation and the verification theorem ([Mer90, Pha09]):

$$v(t,W) = \frac{b(t)}{1-\eta} W^{1-\eta}, \quad \text{with} \quad b(t) = \left(\frac{1+(\nu-1)\exp(\nu(t-T))}{\nu}\right)^{\eta}, \tag{2.5.30}$$

where $\nu = (\rho - (1-\eta)\bar{\mu} + \frac{1}{2}(1-\eta)\eta\bar{\sigma}^2)/\eta$. The optimal consumption law is then given by

$$\boldsymbol{\alpha}^*(t) = b(t)^{-1/\eta}. \tag{2.5.31}$$

Note that this optimal control process is independent of current wealth level, W. For the tests in this section we choose the following set of parameters:

$$T = 100,\ \rho = 0.03,\ \eta = 4,\ W_0 = 100,\ \bar{\mu} = 0.04,\ \bar{\sigma} = 0.1. \tag{2.5.32}$$

Figure 2.5.5 shows the optimal control law. At the terminal time, T, the remaining wealth is assumed to be completely consumed. Based on economic arguments the control process $\boldsymbol{\alpha}^*(t)$ goes to one if the time approaches the terminal time. At earlier time levels the control process has reached a steady state.

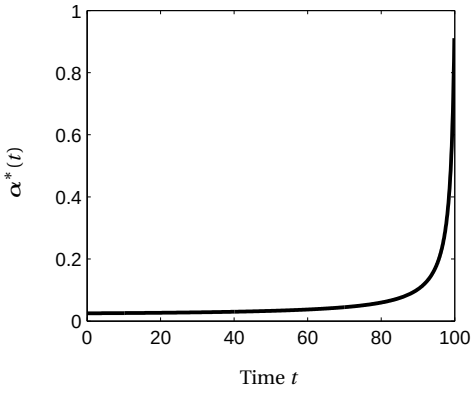


Figure 2.5.5: Optimal control law $\boldsymbol{\alpha}^*(t)$.

For the numerical approach we employ an equidistant grid of control times, $t_0, t_1, \ldots, t_M = T$, with $\Delta t := t_m - t_{m-1}$. At each control time, t_m, one can choose a constant fraction

of wealth, $\alpha_m \in A = [0,1]$, which is consumed during the time interval $[t_m, t_{m+1}]$[6]. We switch to the log-wealth process, $X_s = \log W_s$, so

$$dX_s = (\bar{\mu} - \alpha_m - \tfrac{1}{2}\bar{\sigma}^2)ds + \bar{\sigma}\,d\omega_s, \quad \text{for } s \in [t_m, t_{m+1}]. \tag{2.5.33}$$

We rewrite the value function of the discrete-time problem as

$$v(t,x) = \max_{\boldsymbol{\alpha}\in\hat{\mathscr{A}}} \mathbb{E}^{t,x}\left[\int_t^T e^{-\rho(s-t)}\mathcal{U}(\boldsymbol{\alpha}_s e^{X_s})ds + e^{-\rho(T-t)}\mathcal{U}(e^{X_T})\right], \tag{2.5.34}$$

where x is the current log-wealth level. $\hat{\mathscr{A}} \subset \mathscr{A}$ denotes the set of all possible control paths $\{\alpha_m\}_{m=0}^{M-1}$, where α_m is valued in the control set A. The dynamic programming principle gives us

$$\begin{aligned} v(t_{m-1},x) &= \max_{\alpha_{m-1}\in A} \mathbb{E}^{t_{m-1},x}\left[\int_{t_{m-1}}^{t_m} e^{-\rho(s-t)}\mathcal{U}(\alpha_{m-1}e^{X_s})ds + e^{-\rho\Delta t}v(t_m, X_{t_m})\right] \\ &:= \max_{\alpha_{m-1}\in A}[F(x,\alpha_{m-1}) + c(t_{m-1},x,\alpha_{m-1})]. \end{aligned} \tag{2.5.35}$$

For the time-independent profit function we find

$$\begin{aligned} F(x,\alpha) &= \tfrac{\alpha^{1-\eta}}{(1-\eta)\xi(\alpha)}(e^{\Delta t\xi(\alpha)} - 1)e^{(1-\eta)x}, \\ &\text{with} \quad \xi(\alpha) = -\rho + (1-\eta)(\bar{\mu} - \alpha - \tfrac{1}{2}\bar{\sigma}^2) + \tfrac{1}{2}(1-\eta)^2\bar{\sigma}^2. \end{aligned} \tag{2.5.36}$$

Applying the COS formula yields

$$\hat{c}(t_{m-1},x,\alpha) = e^{-\rho\Delta t}{\sum_{k=0}^{N-1}}'\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\middle|\alpha\right)e^{ik\pi\frac{x-a}{b-a}}\right\}\mathcal{V}_k(t_m), \tag{2.5.37}$$

where the characteristic function is given by

$$\phi_{levy}(u|\alpha) = \exp\left(i\,u(\bar{\mu} - \alpha - \tfrac{1}{2}\bar{\sigma}^2)\Delta t - \tfrac{1}{2}\bar{\sigma}^2u^2\Delta t\right). \tag{2.5.38}$$

We use Algorithm 1 to solve the discrete-time stochastic control problem (2.5.34). The coefficients at time t_M read

$$\begin{aligned} \mathcal{V}_k(t_M) &= \tfrac{2}{b-a}\int_a^b \tfrac{e^{(1-\eta)y}}{1-\eta}\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &= \tfrac{2}{b-a}\tfrac{1}{(1-\eta)^2}\chi_k(a(1-\eta),b(1-\eta),a(1-\eta),b(1-\eta)), \end{aligned} \tag{2.5.39}$$

with the analytic function χ_k in Appendix A.

We divide the interval $[a,b]$ into $Q = 200$ equally-sized subintervals $\mathcal{D}_m^q$ and determine the optimal control value for the midpoint of each subinterval by using an optimization algorithm of MATLAB, which is based on golden section search and parabolic interpolation. We assume that this value is an accurate approximation for the entire subinterval.

[6] For this problem, the true optimal control values are in the set $A = [0,1]$. A wider control set only makes the difficulties that we describe here even more severe.

The terms $\mathscr{F}_k$ are time-independent and known analytically:

$$\begin{aligned}\mathscr{F}_k(z_1,z_2,\alpha) &= \tfrac{2}{b-a}\int_{z_1}^{z_2} F(y,\alpha)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy \\ &= \tfrac{\alpha^{1-\eta}}{(1-\eta)^2\xi(\alpha)}(e^{\Delta t\xi(\alpha)}-1)\tfrac{2}{b-a}\chi_k(z_1(1-\eta),z_2(1-\eta),a(1-\eta),b(1-\eta)). \quad (2.5.40)\end{aligned}$$

The coefficients $\hat{\mathscr{C}}_k$ are recovered by the coefficients $\hat{\mathscr{V}}_k$ from the next time level and the FFT algorithm, as explained in Section 2.3.2.

We first show in Section 2.5.2.1 how the local error of the COS formula propagates backwards in time. Then the extrapolation methodology from Section 2.4.2 is applied to improve the solution (in Section 2.5.2.2). The following parameters are used for the tests in the next subsections:

$$[a,b]=[-2,8],\ N=2^{10},\ M=100. \quad (2.5.41)$$

2.5.2.1. EXAMPLE OF A PROPAGATING ERROR

Here, we show how local errors propagate when we do not apply extrapolation for the continuation values. Parameter N is chosen sufficiently large, so that error ϵ_2 can be neglected. From the error analysis in Section 2.4.1 it follows that

$$\hat{c}(t_{M-1},x,\alpha;[a,b],N) = e^{-\rho\Delta t}\int_{[a,b]} v(t_M,y)p(y|x,\alpha)dy + e^{-\rho\Delta t}\int_{\mathbb{R}\setminus[a,b]} v_N(t_M,y)p(y|x,\alpha)dy. \quad (2.5.42)$$

Function $v(t_M,y)$ resembles $v_N(t_M,y)$ on $[a,b]$ for sufficiently large values of N. However, the approximated continuation value is inaccurate when error $\epsilon_1+\epsilon_3$ in (2.4.6) is large. The inaccuracy of $\hat{c}(t_{M-1},x,\alpha)$ may give rise to a propagating error. First of all, it may result in an incorrect control value from the maximization operator in (2.3.15). Secondly, an inaccurate value function and coefficients $\hat{\mathscr{V}}_k(t_{M-1})$ give rise to inaccurate numerical continuation values at time t_{M-2}. This is demonstrated by four plots in Figure 2.5.6.

The upper-left plot presents the terminal reward function, the Fourier cosine expansion, and its extension outside the expansion interval, which is symmetric in $a=-2$ and different from the correct terminal reward function at the left-hand side of a. In the lower-left plot the continuation value is shown for two control values, $\alpha=0.7$ and $\alpha=1.0$. The exact optimal control value for the discrete-time stochastic control problem at time t_{M-1} is $\alpha_{M-1}\approx 0.7$. It is independent of the current log-wealth level. From Figure 2.5.6 it is clear that significant errors occur in the vicinity of the left-side domain boundary, as the control value $\alpha=1.0$ appears to be optimal here.

The addition of the profit function gives us the graphs in the upper-right plot. As the value function at time t_{M-1} is defined by the maximization operator over all $\alpha_{M-1}\in A$, inaccurate control values will be determined, at least for $x<-1$. Because of this, the approximated value function will be too high. The lower-left plot shows the continuation values at time t_{M-2} for two different control values. The local error at time level $M-1$ has propagated and, in addition, errors from the COS formula occur in the vicinity of the boundary. The correct optimal control value equals $\alpha_{M-2}\approx 0.4$ and, again, incorrect control values will be determined by the maximization operator, if extrapolation is not used here.

The solution for all time steps gives rise to the optimal control values in Figure 2.5.7a. The correct control law should be independent of x, which is clearly not the case.

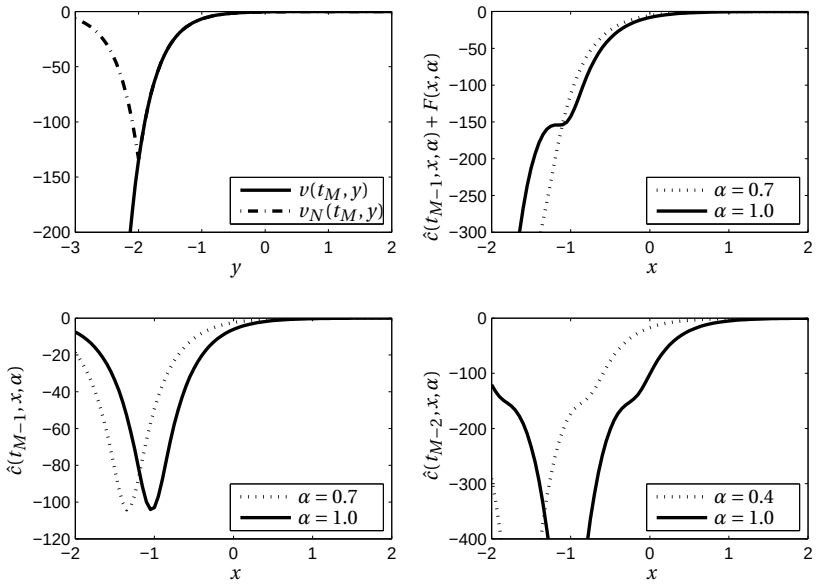
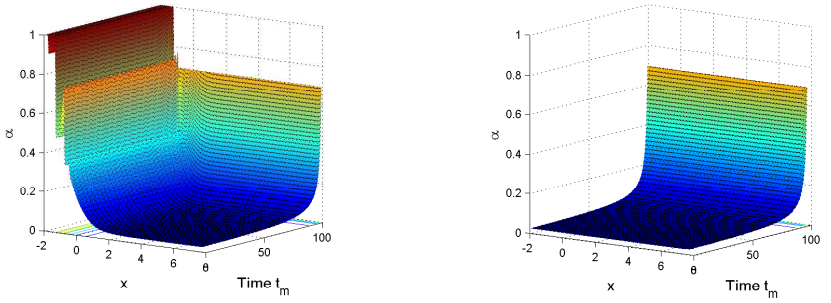


Figure 2.5.6: The propagation of local errors ($[a, b] = [-2, 8]$).



(a) Incorrect values, due to local errors and their propagation.

(b) Correct values, with extrapolation method.

Figure 2.5.7: Optimal control laws.

2.5.2.2. IMPROVEMENT BY EXTRAPOLATION

We use the extrapolation technique from Section 2.4.2 to deal with the propagating error. The function $p(y|x, \alpha)$ represents a normal density function of a random variable with distribution

$$\mathcal{N}\left(x + (\bar{\mu} - \alpha - \tfrac{1}{2}\bar{\sigma}^2)\Delta t, \bar{\sigma}^2 \Delta t\right). \tag{2.5.43}$$

We can presume that the continuation value is well approximated on $[x_L, b]$, with

$$x_L := a - (\bar{\mu} - \alpha - \tfrac{1}{2}\bar{\sigma}^2)\Delta t + 5\bar{\sigma}\sqrt{\Delta t}. \tag{2.5.44}$$

The terminal reward function, the density, and profit functions are all of exponential form. Therefore, we approximate $c(t_{m-1}, x, \alpha)$ on $[a, x_L]$ by employing an *exponential extrapolation technique*[7], as follows:

$$\hat{c}^{ex}(t_{m-1}, x, \alpha) := \hat{c}(t_{m-1}, x_L, \alpha) \exp\big(P(y - x_L)\big), \quad \text{for } x \in [a, x_L], \tag{2.5.45}$$

with

$$P := \tfrac{\hat{c}_x(t_{m-1}, x_L, \alpha)}{\hat{c}(t_{m-1}, x_L, \alpha)}. \tag{2.5.46}$$

As proposed in Section 2.4.2, we will use the improved continuation value, $\tilde{c}(t_{m-1}, x, \alpha)$, to find the optimal control values and to approximate the coefficients $\mathscr{C}_k(t_{m-1}, z_1, z_2, \alpha)$. Suppose the interval $[z_1, z_2] \subset [a, b]$ can be divided into $[z_1, x_L]$ and $[x_L, z_2]$. Then the corresponding Fourier cosine coefficients read

$$\begin{aligned}\tilde{\mathscr{C}}_k(t_{m-1}, z_1, z_2, \alpha) = {} & \tfrac{2}{b-a} \int_{z_1}^{x_L} \hat{c}^{ex}(t_{m-1}, x_L, \alpha) \cos\big(k\pi \tfrac{y-a}{b-a}\big)\, dy \\ & + e^{-\rho \Delta t} \Re\left\{ {\sum_{j=0}^{N-1}}' \phi_{levy}\left(\tfrac{j\pi}{b-a} \middle| \alpha\right) \mathscr{V}_j(t_m) \mathscr{M}_{k,j}(x_L, z_2) \right\},\end{aligned} \tag{2.5.47}$$

where

$$\tfrac{2}{b-a} \int_{z_1}^{x_L} \hat{c}^{ex}(t_{m-1}, x_L, \alpha) \cos\big(k\pi \tfrac{y-a}{b-a}\big)\, dy = \tfrac{2}{b-a} \hat{c}(t_{m-1}, x_L, \alpha) \tfrac{e^{-Px_L}}{P} \chi_k(Pz_1, Px_L, Pa, Pb). \tag{2.5.48}$$

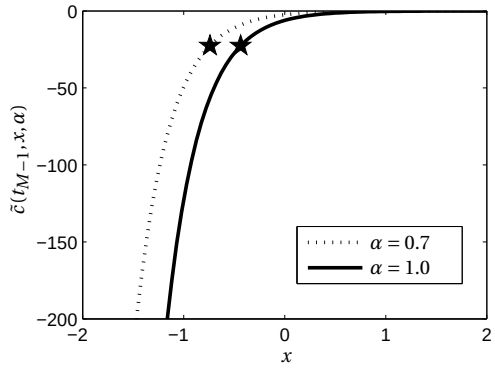


Figure 2.5.8: Extrapolation of the continuation value (★ = x_L).

Figure 2.5.8 shows the values x_L and the improved continuation values for $\alpha = 0.7$ and $\alpha = 1.0$. The values are accurate, even in the vicinity of boundary $a = -2$, and the correct optimal control values will be determined, see Figure 2.5.7b for the complete result.

The exact solution to the continuous-time stochastic control problem is $v(t_0, x_0) = -0.8419$. For the discrete-time variant we find $\hat{v}(t_0, x_0) = -0.8430$. Table 2.5.5 shows the value function for different numbers of control times, M. They converge to the true value. We can conclude that the procedure with exponential extrapolation works highly satisfactorily.

[7] Note that polynomial extrapolation will also work well here.

Table 2.5.5: Results optimal consumption-portfolio problem ($N = 2^{10}$).

M	10	25	50	100	150	200	$\hat{v}_R(t_0,x_0;25)$
$\hat{v}(t_0,x_0)$	-0.8843	-0.8548	-0.8459	-0.8430	-0.8424	-0.8422	-0.8420

Remark 2.5.5. *The dynamics of the invested capital, equation (2.5.26), can easily be extended to a jump-diffusion process. Then no analytic solution to the stochastic control problem is available, but we can apply the COS method, as the characteristic function of a jump-diffusion process is known analytically, with the extrapolation technique.*

2.6. CONCLUSION

In this chapter, we presented a general approach for solving stochastic control problems under a one-dimensional Lévy process. The method relies on the dynamic programming principle and the COS formula, which is based on Fourier cosine series expansions. A recursive algorithm has been defined, based on the recursive recovery of the series coefficients. With the use of a Fast Fourier Transform algorithm we reach a computational complexity of order $\mathcal{O}(N \log_2 N)$ per time step, where N denotes the number of terms in the series expansions.

We provided an extensive error analysis, with which we acquired knowledge about the origin and evolution of errors. We demonstrated how errors of the COS formula in the vicinity of domain boundaries may arise, and how they may propagate backwards in time. This understanding enabled us to improve the method by introducing an extrapolation method for the area in which the COS formula may give inaccurate continuation values. Extrapolation by Taylor expansion or by exponential extrapolation can easily be applied as the derivatives of approximated continuation values can be computed easily based on the COS formula. An exponentially converging error, in N, is found for a sufficiently accurate extrapolation method, $[a,b] \subset \mathbb{R}$ sufficiently wide and a probability density function in the class $C^\infty([a,b])$. A density function with discontinuity in one of its derivatives results in an algebraic convergence.

In [FO09], the COS method has been employed for pricing Bermudan and barrier options. A difference in using the COS method for stochastic control problems is the dependence of the characteristic function of the stochastic process on the control value, which may be time-consuming. Besides, we need to determine the optimal control law from a finite or continuous control set, for the entire spatial domain. Therefore, the approximated value function needs to be accurate over the complete spatial domain, which is not always the case when using the COS formula. These difficulties have been solved by the extrapolation technique.

We tested our numerical method by two examples, financial options under uncertain volatility, based on the models from [PFV03, ZW09], and an optimal consumption-portfolio problem. The COS method for stochastic control problems performed highly satisfactory.

Many other problems from finance and the real options context can be represented as a stochastic control problem. This makes our methodology applicable to various practical problems.

CHAPTER 3

On the Application of Spectral Filters in a Fourier Option Pricing Technique

3.1. INTRODUCTION

Fourier techniques have now become well-established in computational finance for efficiently pricing certain financial instruments. For instance European options, certain options with early-exercise features, and also exotic options, like Asian, multi-asset, or barrier options, have been priced by Fourier techniques. The Fourier techniques belong to the class of numerical integration option pricing methods. They are referred to as 'transform methods', because a transformation to the Fourier domain is combined with numerical integration [CM99, LK07, DI06, BL11]. Transform methods can readily be used for asset price models for which the characteristic function (i.e., the Fourier transform of the probability density function) is available. A specific Fourier pricing technique, which we consider here, is the COS method. This pricing method is based on Fourier cosine series expansions. The issues and remedies we address here will, however, be of relevance for other transform methods as well.

As long as the governing probability density function is sufficiently smooth, an exponential error convergence in the number of cosine terms is achieved by the COS method. For certain choices of the asset dynamics, however, the governing probability density function is *not* smooth everywhere. Smoothness issues are encountered, for example, when we model the asset price by the Variance Gamma process [MCC98]. This results in only algebraic convergence for the option price. For hedging purposes agents are also interested in the option's Greeks, which measure risk sensitivities of the option price. Their approximations may even suffer more from non-smoothness conditions and for these cases it is desirable to find a faster converging method for the Greeks. In the field of risk management we deal with discrete random variables representing individual obligors that may have a default problem. Portfolio loss modeling with small-sized portfolios typically produces stepwise cumulative distribution functions and inaccuracies around the discontinuities may occur.

In general, when Fourier techniques are employed to specific cases with non-smooth functions, the *Gibbs phenomenon* may become apparent. This seriously impacts the efficiency and accuracy of the financial valuation. The Gibbs phenomenon reflects the diffi-

This chapter is based on the article 'On the application of spectral filters in a Fourier option pricing technique', to appear in *Journal of Computational Finance*, 2015 [RVO15]. In this chapter, we truncate the Fourier and Fourier cosine series from $-N$ and 0, respectively, to N, in order to stay inline with the literature about spectral filters. This is in contrast with the rest of this thesis, where we sum from 0 to $N-1$ in the truncated Fourier cosine series.

culty of approximating a discontinuous function by a finite Fourier series. Although the limit of the partial sums represents the original function exactly, in the finite case there is always an overshoot at a jump discontinuity. The width of this overshoot decreases with the number of Fourier terms, but the height of the maximum does not. The Gibbs phenomenon is also related to the principle that the decay of the Fourier coefficients is governed by the smoothness of the function concerned. Functions with a discontinuity in one of the derivatives will have algebraically decaying Fourier coefficients, that result in a slowly converging Fourier series. The local effect of the Gibbs phenomenon gives rise to oscillations near the jumps. However, there also is a global effect: although the error decays away from the jumps, the decay rate is only first order. Thus, the existence of one or more discontinuities drastically reduces the convergence rate over the whole domain, and spectral accuracy is lost [Tad07, GS97].

The research field dealing with Gibbs oscillations is wide and well-established. An excellent overview into the various improvement techniques is given by Tadmor [Tad07]. We will focus on the use of *spectral filters* to deal with the Gibbs phenomenon appearing for non-smooth densities and discrete distribution functions. This is one of the very basic techniques in this field, but we will see that it fits very well to the applications at hand.

Several other techniques have been proposed in the literature to reduce or remove the Gibbs phenomenon. The optimal filter order for a function with discontinuities is an increasing function of the distance to the nearest discontinuity. The idea of adaptive filtering is then to vary the filter order so that it is close to this optimal value. Tadmor and Tanner [Tad07, Tan06, TT05] describe adaptive filters recovering root-exponential accuracy. This type of filtering is related to superconvergent extraction techniques in finite element methods [vSRV11]. It is however not trivial to efficiently employ an adaptive filter in the context of Fourier option pricing, so we stay with non-adaptive filters here. Mollifiers are a time domain equivalent to filters, in the sense that multiplication by a function in Fourier space corresponds to a convolution in physical space (which is the basis for mollifiers). Implementation of mollifiers in a Fourier option pricing technique would require reconstruction to the time-domain, which is computationally relatively expensive. The same is true for methods like Gegenbauer polynomial reconstruction, see [Tad07, GS97], Digital Total Variation (DTV) filtering [COS01, Sar06] and for the hybrid methods [Gel00], where a polynomial reconstruction is used only where needed and filtering is used elsewhere.

We start in Section 3.2 with Fourier series and Fourier cosine series and explain how we can employ the characteristic function of a random variable to approximate the corresponding density or distribution function. Also the COS method for pricing European options is briefly described. In Section 3.3 the convergence of the series and improvements by spectral filters are discussed. Extensive numerical experiments are performed in Section 3.4. Finally, Section 3.5 concludes.

3.2. FOURIER AND FOURIER COSINE SERIES EXPANSIONS

The *Fourier series* of an integrable function $f(y)$ supported on a finite interval $[a,b]$ is defined as ([SS03])

$$f(y) = \sum_{k=-\infty}^{\infty} \mathscr{F}_k^F \exp\left(ik\pi \tfrac{2y}{b-a}\right), \tag{3.2.1}$$

with Fourier coefficients given by

$$\mathscr{F}_k^F = \tfrac{1}{b-a}\int_a^b f(y)\exp\left(-ik\pi\tfrac{2y}{b-a}\right)dy. \tag{3.2.2}$$

The COS method for pricing European options is based on the *Fourier cosine series* expansion, which is defined by

$$f(y) = \sum_{k=0}^{\infty}{}' \mathscr{F}_k \cos\left(k\pi\tfrac{y-a}{b-a}\right), \tag{3.2.3}$$

with Fourier cosine coefficients given by

$$\mathscr{F}_k = \tfrac{2}{b-a}\int_a^b f(y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{3.2.4}$$

The prime $'$ in the summation indicates division of the first term by two. These cosine series can be seen as a classical Fourier series of a function $f^{ext}(y) = f(|y-a|)$, on an extended interval $[2a-b, b]$, which is mirrored around the midpoint a to make it an even function.

3.2.1. RECOVERY DENSITY AND DISTRIBUTION FUNCTION

In financial option pricing we deal with stochastic asset prices. The corresponding probability density function is usually unknown. The characteristic function is however often known [DPS00] and can be used to approximate the density and distribution function. In Section 3.4.3 we discuss a model for discrete portfolio losses. For example, the loss may be either zero or one. An approximation of the distribution function of a discrete random variable, also by using the characteristic function, is derived here as well.

Continuous Random Variable Suppose we have a continuous random variable X, with cumulative distribution function $F(.)$, probability density function $p(.)$, and characteristic function $\varphi(.)$. As presented in the previous chapter, the Fourier cosine series expansion of the density function reads

$$p(y) = \sum_{k=0}^{\infty}{}' \mathscr{P}_k \cos\left(k\pi\tfrac{y-a}{b-a}\right), \tag{3.2.5}$$

with Fourier cosine coefficients

$$\mathscr{P}_k = \tfrac{2}{b-a}\int_a^b p(y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy = \tfrac{2}{b-a}\Re\left\{\int_a^b p(y)\exp\left(ik\pi\tfrac{y-a}{b-a}\right)dy\right\}, \tag{3.2.6}$$

where $\Re\{.\}$ denotes taking the real part. If the density function $p(y)$ decays rapidly to zero for $y \to \pm\infty$, the integration range in the following equation can be truncated without loss of any significant accuracy:

$$\varphi(u) = \int_{\mathbb{R}} e^{iuy}p(y)\,dy \approx \int_a^b e^{iuy}p(y)\,dy := \hat{\varphi}(u), \tag{3.2.7}$$

and therefore the characteristic function can be used to efficiently calculate the Fourier coefficients. Combining (3.2.6) and (3.2.7) gives us:

$$\mathscr{P}_k = \tfrac{2}{b-a}\Re\left\{\hat{\varphi}\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{-a}{b-a}}\right\} \approx \tfrac{2}{b-a}\Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{-a}{b-a}}\right\} := \Phi_k. \tag{3.2.8}$$

After truncation of the series summation we end up with the approximation

$$\hat{p}(y) := \sum_{k=0}^{N}{}' \Phi_k \cos\left(k\pi \tfrac{y-a}{b-a}\right). \tag{3.2.9}$$

The distribution function can now be approximated as follows,

$$\begin{aligned} F(y) = \mathbb{P}(X \le y) &\approx \int_a^y \hat{p}(t)\,dt = \sum_{k=0}^{N}{}' \Phi_k \int_a^y \cos\left(k\pi \tfrac{t-a}{b-a}\right) dt \\ &= \tfrac{1}{2}\tfrac{2}{b-a}(y-a) + \sum_{k=1}^{N} \tfrac{2}{k\pi} \Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\right) e^{ik\pi \frac{-a}{b-a}}\right\} \sin\left(k\pi \tfrac{y-a}{b-a}\right). \end{aligned} \tag{3.2.10}$$

If the interval $[a,b]$ is chosen sufficiently wide, then the series truncation error dominates the error of the approximations (3.2.9) and (3.2.10). We refer to Section 3.3.5 for details about the series truncation error in the COS formula.

Discrete Random Variable If X is a discrete random variable, then a density function does not exist and we use the following Lévy inversion formula connecting the distribution function, $F(.)$, and characteristic function, $\varphi(.)$. For $y-h, y+h \in C(F)$, with $C(F)$ the continuity set of F, there holds ([Gut05])

$$\begin{aligned} \tfrac{F(y+h)-F(y-h)}{2h} &= \lim_{T\to\infty} \tfrac{1}{2h}\tfrac{1}{2\pi} \int_{-T}^{T} \varphi(u) \tfrac{e^{-iu(y-h)} - e^{-iu(y+h)}}{iu}\,du \\ &= \lim_{T\to\infty} \tfrac{1}{2\pi} \int_{-T}^{T} \varphi(u) e^{-iuy} \tfrac{\sin uh}{uh}\,du. \end{aligned} \tag{3.2.11}$$

Suppose F is concentrated on the interval $[a,\infty)$ and $F(a)=0$. Then, we get ($h \ge 0$)

$$\begin{aligned} F(a+h) &= \lim_{T\to\infty} \tfrac{1}{\pi} \int_{-T}^{T} \varphi(u) e^{-iua} \tfrac{\sin uh}{u}\,du \\ &= \lim_{T\to\infty} \tfrac{2}{\pi} \int_{0}^{T} \Re\left\{\varphi(u) e^{-iua}\right\} \tfrac{\sin uh}{u}\,du. \end{aligned} \tag{3.2.12}$$

Numerical integration with step size Δu results in the approximation ($y \ge a, y \in C(F)$)

$$\begin{aligned} F(y) &= \lim_{T\to\infty} \tfrac{2}{\pi} \int_0^T \Re\left\{\varphi(u) e^{-iua}\right\} \tfrac{\sin u(y-a)}{u}\,du \\ &\approx \tfrac{1}{2}\tfrac{2\Delta u}{\pi}(y-a) + \sum_{k=1}^{N} \tfrac{2}{k\pi} \Re\left\{\varphi\left(k\Delta u\right) e^{-ik\Delta u a}\right\} \sin\left(k\Delta u(y-a)\right). \end{aligned} \tag{3.2.13}$$

For $\Delta u = \frac{\pi}{b-a}$ this, in fact, corresponds to the formula for the distribution function of a continuous random variable (equation (3.2.10)).

3.2.2. COS METHOD FOR EUROPEAN OPTIONS AND GREEKS

The asset price is modeled by a stochastic process and is denoted by S_t. The risk-neutral valuation formula for a European option with payoff function $g(.)$ reads

$$v(t_0,x) = e^{-r\Delta t}\mathbb{E}_{\mathbb{Q}}[g(X_T)|X_{t_0}=x] = e^{-r\Delta t}\int_{\mathbb{R}} g(y)\,p(y|x)\,dy. \tag{3.2.14}$$

Here X_t is the state process, which can be any monotone function of the underlying asset price S_t. In this chapter it is taken to be the scaled log-asset price, $X_t := \log(S_t/K)$, where K is the option's strike price. $\mathbb{E}_{\mathbb{Q}}$ denotes the expectation under risk-neutral measure $\mathbb{Q}$; Δt is the difference between the time of maturity T and the initial time $t_0 = 0$; $p(y|x)$ is the conditional probability density of X_T, given $X_{t_0} = x$, and r is the risk-free interest rate. In other words, the option price is equal to the expected value of its discounted future payoff, under a certain probability measure.

In this chapter we consider a specific Fourier technique, the COS method, to approximate the expected value of an arbitrary function of random variables. The method has been developed in the first place for pricing financial options, see Section 2.3.4. The Fourier cosine coefficients $\mathscr{P}_k(x)$ of the density function are approximated by, similar as equation (3.2.8),

$$\begin{aligned}\mathscr{P}_k(x) &= \tfrac{2}{b-a}\int_a^b p(y|x)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy\\ &\approx \tfrac{2}{b-a}\Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{-a}{b-a}}\right\} := \Phi_k(x).\end{aligned} \tag{3.2.15}$$

The Fourier cosine series coefficients of the payoff function are denoted by

$$\mathscr{G}_k = \tfrac{2}{b-a}\int_a^b g(y)\cos\left(k\pi\tfrac{y-a}{b-a}\right)dy. \tag{3.2.16}$$

Recall the *COS pricing formula* (equation (2.3.31)):

$$\begin{aligned}\hat{v}(t_0,x) &:= \tfrac{b-a}{2}e^{-r\Delta t}\sum_{k=0}^{N}{}'\Phi_k(x)\mathscr{G}_k\\ &= e^{-r\Delta t}\sum_{k=0}^{N}{}'\Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{-a}{b-a}}\right\}\mathscr{G}_k\\ &= e^{-r\Delta t}\sum_{k=0}^{N}{}'\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}\mathscr{G}_k.\end{aligned} \tag{3.2.17}$$

The last equality holds for processes with independent, stationary increments, such as Lévy processes, which include the log-versions of geometric Brownian motion, Variance Gamma, and CGMY models. In that case, the characteristic function can be written as a product of e^{iux} and a part independent of x, that is, $\varphi(u|x) = e^{iux}\phi_{levy}(u)$.

The integration range $[a,b]$ must be chosen carefully to avoid significant errors. An interval which is too small will result in integration range truncation errors and a too wide interval may give rise to cancelation errors. For now we mention the results given by [FO08], in which a rule of thumb for choosing the integration range is given:

$$[a,b] := \left[\kappa_1 - L\sqrt{\kappa_2+\sqrt{\kappa_4}}, \kappa_1 + L\sqrt{\kappa_2+\sqrt{\kappa_4}}\right], \qquad L = 10. \tag{3.2.18}$$

$\kappa_1, \kappa_2, \ldots$ are the cumulants of the underlying stochastic process X_t (see, for example, [FO08, Table 11]).

3.2.2.1. THE OPTION'S GREEKS

The Greeks indicate the sensitivities of the option price with respect to a change in its underlying or model parameters, such as the asset price or the volatility. They are used to hedge the risks in a portfolio. The most well-known Greek parameter is the option's Delta, Δ, i.e., the first derivative of the option price with respect to underlying asset price S_t. Gamma, Γ, is the second derivative of the option price with respect to the asset price. By the COS pricing formula (3.2.17) we naturally find the following approximations ($x = \log(S/K)$),

$$\Delta = \frac{\partial v(t_0,x)}{\partial S} \approx \frac{\partial \hat{v}(t_0,x)}{\partial x}\frac{1}{S} = e^{-r\Delta t}\sum_{k=0}^{N}{}'\,\Re\left\{\phi_{levy}\left(\frac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\frac{ik\pi}{b-a}\right\}\mathscr{G}_k\frac{1}{S}, \tag{3.2.19a}$$

$$\begin{aligned}\Gamma &= \frac{\partial^2 v(t_0,x)}{\partial S^2} \approx \left(\frac{\partial^2 \hat{v}(t_0,x)}{\partial x^2} - \frac{\partial \hat{v}(t_0,x)}{\partial x}\right)\frac{1}{S^2}\\ &= e^{-r\Delta t}\sum_{k=0}^{N}{}'\,\Re\left\{\phi_{levy}\left(\frac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\left[\left(\frac{ik\pi}{b-a}\right)^2 - \frac{ik\pi}{b-a}\right]\right\}\mathscr{G}_k\frac{1}{S^2}.\end{aligned} \tag{3.2.19b}$$

3.3. CONVERGENCE AND IMPROVEMENTS BY SPECTRAL FILTERS

In this section, we use, without loss of generality, the interval $[a,b] = [0,2\pi]$ and we consider the classical Fourier series. As the Fourier cosine series of a function is equivalent to the symmetrically extended version on an extended domain, the theory here also applies to Fourier cosine series.

3.3.1. CONVERGENCE OF FOURIER SERIES

The partial sum of the Fourier series of a function $f(y)$ on $[0,2\pi]$ is given by

$$f_N(y) := \sum_{|k|\le N}\mathscr{F}_k^F e^{iky}. \tag{3.3.1}$$

To make statements about the convergence rate we look at pointwise convergence. If f is a continuous periodic function on $[0,2\pi]$, so $f(0) = f(2\pi)$, and the Fourier series of f is absolutely convergent, $\sum_{k=-\infty}^{\infty}|\mathscr{F}_k^F| < \infty$, then the Fourier series converges uniformly to f. For a non-continuous function there is no uniform convergence and its Fourier series converges to the average of the left- and right-hand limits at a jump discontinuity. In the case that discontinuities of any order are present, the Fourier series expansion only exhibits algebraic convergence. For jump discontinuities, we even encounter 0th order convergence, which leads to the Gibbs-overshoot.

The speed at which the Fourier coefficients decay depends on the smoothness of the function, as is stated in the following theorem:

Theorem 3.3.1. *(Integration-by-parts coefficient bound)[Boy01]*
If

$$f(0) = f(2\pi),\ f^{(1)}(0) = f^{(1)}(2\pi),\ldots,\ f^{(n-2)}(0) = f^{(n-2)}(2\pi) \tag{3.3.2}$$

and $f^{(n)}(y)$ is integrable, then

$$\mathscr{F}_k^F \sim \mathscr{O}\left(|k|^{-n}\right) \quad as \quad |k| \to \infty. \tag{3.3.3}$$

Here $f^{(n)}$ denotes the nth derivative of $f(y)$. The integrability of $f^{(n)}$ requires that $f(y)$, $f^{(1)}(y)$,..., $f^{(n-2)}(y)$ should be continuous.

The absolute error of truncation of the expansion after N terms is denoted by

$$E_{Tr}(N) := |f(y) - f_N(y)| \leq \sum_{k=N+1}^{\infty} |\mathscr{F}_k^F + \mathscr{F}_{-k}^F|. \tag{3.3.4}$$

In general the convergence rate of a Fourier series depends on the smoothness of the function on the expansion interval. We refer to [Boy01] for the definitions of algebraically and exponentially converging terms. The following proposition allows us to bound the series truncation error of geometrically and algebraically converging series.

Proposition 3.3.1. *(Last coefficient error estimate)[Boy01]*
The truncation error is of the same order of magnitude as the last coefficient retained in the truncation for a series with (at least) geometric convergence.

If the series has algebraic convergence index $n > 1$, i.e., if $a_k \sim \mathscr{O}(1/k^n)$ for large k, then

$$E_{Tr}(N) \sim \mathscr{O}(|Na_N|). \tag{3.3.5}$$

In the numerical experiments in Section 3.4.1 we will observe that the Fourier coefficients of an $(n-2)$-times continuously differentiable function, $f \in C^{n-2}$, will converge with an algebraic index of convergence n, see Theorem 3.3.1. According to the proposition above we would expect truncation error $E_{Tr}(N)$ to decrease with order $\mathscr{O}\left(N^{1-n}\right)$. However, we find a faster convergence, of order n. This may be due to the alternating behavior of the series.

3.3.2. SPECTRAL FILTER

In this section we explain the notion of spectral filters, by which we aim to mitigate oscillations related to the Gibbs phenomenon and achieve faster convergence for problems in financial mathematics. Filtering is carried out in Fourier space and the idea is to premultiply the expansion coefficients by a decreasing function in such a way that these decay faster. A properly chosen filter will improve the convergence rate away from discontinuities. The following definition is from [Van91, GS97]:

Definition 3.3.1. *(Fourier space filter of order p) A real and $C^\infty([0,1])$ even function $\hat{s}(\eta)$ is called a filter of order p, if*

1. $\hat{s}(0) = 1$ *and* $\hat{s}^{(\ell)}(0) = 0, \quad 1 \leq \ell \leq p-1$,
2. $\hat{s}(\eta) = 0$ *for* $|\eta| \geq 1$,
3. $\hat{s}(\eta) \in C^{p-1}, \quad \eta \in (-\infty, \infty)$.

Conditions 2 and 3 imply $\hat{s}^{(\ell)}(1) = 0, \quad 0 \leq \ell \leq p-1$.

The *filtered* partial sum of a Fourier series is simply defined by

$$f_N^{\hat{s}}(y) = \sum_{|k| \leq N} \hat{s}(k/N) \mathscr{F}_k^F e^{iky}. \tag{3.3.6}$$

We can rewrite this as a convolution in physical space:

$$f_N^{\hat{s}}(y) = \frac{1}{2\pi} \int_0^{2\pi} s(y-t) f(t)\, dt, \quad \text{with} \quad s(x) = \sum_{|k| \leq \infty} \hat{s}(k/N) e^{ikx}, \quad x \in [0, 2\pi]. \tag{3.3.7}$$

Note that $\hat{s}(k/N) = 0$ for $|k| > N$. A filter is a continuous function which only modifies high frequency modes, not the low modes. Filtering may remove the Gibbs phenomenon away from a discontinuity, the error depends on the distance to the discontinuity, as we will confirm in Section 3.3.3. Since the approximation will be smoothened, recovery in the vicinity of a 'jump' will not always improve for low values of N. Filtering does not affect the total mass of the resulting approximation (which should be one for a probability density), since the first coefficient is never altered. To be precise,

$$\int_0^{2\pi} f_N^{\hat{s}}(y)\,dy = \sum_{|k|\leq N} \hat{s}(k/N)\mathscr{F}_k^F \int_0^{2\pi} e^{iky}\,dy = \hat{s}(0/N)\mathscr{F}_0^F 2\pi = \int_0^{2\pi} f(y)\,dy. \tag{3.3.8}$$

3.3.2.1. EXAMPLES OF SPECTRAL FILTERS

The following filters are well-known from the literature [Van91, HGG07, GS97]:

- Fejér filter ([Fej00]): $\hat{s}(\eta) = 1 - |\eta|$, with order $p = 1$.
- Lanczos filter ([Lan56]): $\hat{s}(\eta) = \sin(\pi\eta)/(\pi\eta)$, with order $p = 1$.
- Raised cosine filter: $\hat{s}(\eta) = \frac{1}{2}(1 + \cos(\pi\eta))$, with order $p = 2$.

Also general pth order spectral filters exist:

- Exponential filter ([GS97]): $\hat{s}(\eta) = \exp(-\alpha\eta^p)$, where p must be even and $\alpha = -\log\epsilon_m$, with ϵ_m the machine epsilon.
- Vandeven filter ([Van91, Pey02]):

$$\hat{s}(\eta) = 1 - \frac{(2p-1)!}{((p-1)!)^2}\int_0^{|\eta|} t^{p-1}(1-t)^{p-1}\,dt. \tag{3.3.9}$$

- Erfc-Log filter ([Boy96]): Boyd showed that the Vandeven filter can be approximated quite accurately by an analytic function which satisfies all conditions, i.e., by the Erfc-Log filter:

$$\hat{s}(\eta) = \tfrac{1}{2}\mathrm{erfc}\left(2\sqrt{p}(|\eta| - \tfrac{1}{2})\sqrt{\frac{-\log(1-4(|\eta|-1/2)^2)}{4(|\eta|-1/2)^2}}\right), \tag{3.3.10}$$

  where $\mathrm{erfc}(\cdot)$ is the complimentary Gauss error function.

3.3.3. CONVERGENCE AND ERROR ANALYSIS

A higher-order filter modifies the original function in smooth regions away from a discontinuity and high-order accuracy is desirable away from a discontinuity. Low-order filtering is however desirable close to a discontinuity, because higher p values then give rise to a highly oscillatory filtered function $f_N^{\hat{s}}$. The following theorem gives a bound on the error. It can be extended in a straightforward way to a function with more points of discontinuity ξ_m.

Theorem 3.3.2. *[GS97, HGG07] Let $f(y)$ be a piecewise $C^p([0,2\pi])$ function with one point of discontinuity ξ. Let $\hat{s}(k/N)$ be a filter of order p. Let now y be a point in $[0,2\pi]$ and denote by $d(y) := \min_{k=-1,0,1}|y - \xi + 2k\pi|$. Then (if $y \neq \xi$),*

$$\begin{aligned}|f_N^{\hat{s}}(y) - f(y)| &= \left|\frac{1}{2\pi}\sum_{\ell=0}^{p-1} s_{\ell+1}(d(y))\left(f^{(\ell)}(\xi^+) - f^{(\ell)}(\xi^-)\right) + \frac{1}{2\pi}\int_0^{2\pi} s_p(y-t)f^{(p)}(t)\,dt\right| \\ &\leq cN^{1-p}d(y)^{1-p}\mathscr{K}(f) + cN^{\frac{1}{2}-p}||f^{(p)}||_{L^2},\end{aligned} \tag{3.3.11}$$

where

$$\mathcal{K}(f) = \sum_{\ell=0}^{p-1} d(y)^\ell \left(f^{(\ell)}(\xi^+) - f^{(\ell)}(\xi^-) \right) \int_{-\infty}^{\infty} |G_\ell^{(p-\ell)}(\eta)| \, d\eta, \tag{3.3.12a}$$

$$G_\ell(\eta) = \frac{\hat{s}(\eta)-1}{\eta^\ell}, \tag{3.3.12b}$$

c is a constant independent of f and N, and

$$s_0(x) = s(x), \qquad s'_\ell = s_{\ell-1}, \quad \ell \geq 1, \qquad \int_0^{2\pi} s_\ell(t)\, dt = 0, \quad \ell \geq 1. \tag{3.3.13}$$

The error bound decreases with $d(y)$, i.e., with the distance to the discontinuity. The filter order determines the rate at which the error remaining after filtering decays. If we have $f \notin C^{p-1}$, i.e., if $f(y)$ has a jump discontinuity at one or more points of order smaller than, or equal to, $p-1$, the following estimate holds: $|f_N^{\hat{s}}(y) - f(y)| \sim \mathcal{O}\left(N^{1-p}\right)$. If $f \in C^{p-1}$, i.e., if $f(y)$ is smooth in the sense of possessing at least $p-1$ continuous derivatives, then: $|f_N^{\hat{s}}(y) - f(y)| \sim \mathcal{O}\left(N^{\frac{1}{2}-p}\right)$. [1]

In the numerical experiments in Section 3.4.1 we will observe a somewhat *faster convergence* than prescribed by Theorem 3.3.2. This can be explained by the following observations. For the first part in equation (3.3.11) the authors in [Van91] prove by induction that

$$s_\ell(x) \sim \mathcal{O}\left(N^{1-p}\right), \quad x \in (0, 2\pi), 0 \leq \ell \leq p. \tag{3.3.14}$$

Table 3.3.1 shows the order of convergence for s_0 and s_1 that we observed by numerical experiments. We tested six different filters: The Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter, and the Erfc-Log filter, as described in Section 3.3.2. Besides we used different filter orders. Here "exp" denotes exponential convergence. The algebraic index of convergence given by equation (3.3.14) is thus not strict and can be higher than order p.

Table 3.3.1: Algebraic index of convergence ($x \in (0, 2\pi)$).

Filter	$s_0(x)$	$s_1(x)$
Fejér	1	1
Lanczos	1	2
Raised cosine	2	3
Exponential ($p = 2,4,6$)	exp	exp
Vandeven ($p = 1,3,5$)	p	p
Vandeven ($p = 2,4$)	p	$p+1$
Erfc-Log ($p = 1,2,3,4,5$)	p	p

[1] The Euler-accelerated partial sum ([Boy11]) gives a geometric error, i.e., of order $\mathcal{O}(\exp(-\mu(y)N))$. However, we are not able to obtain results for $N > 1024$ as the computation of the filter is then limited by our double precision computations. Besides, the computational costs of this filter are about 100 times as high as the exponential filter, for our numerical examples.

For the second part in equation (3.3.11) the authors in [Van91, GS97] use the inequalities

$$\left|\tfrac{1}{2\pi}\int_0^{2\pi} s_\ell(y-t)f^{(\ell)}(t)\,dt\right| \le \tfrac{1}{2\pi}\sqrt{\int_0^{2\pi} s_\ell^2(y-t)\,dt}\sqrt{\int_0^{2\pi}(f^{(\ell)}(t))^2\,dt} \tag{3.3.15}$$

and find the upper bound

$$\left|\tfrac{1}{2\pi}\int_0^{2\pi} s_\ell(y-t)f^{(\ell)}(t)\,dt\right| \le \sim \mathcal{O}\left(N^{\frac{1}{2}-\ell}\right). \tag{3.3.16}$$

However, for the filters in Table 3.3.1 we observe by our numerical computations that

$$\left|\tfrac{1}{2\pi}\int_0^{2\pi} s_\ell(y-t)f^{(\ell)}(t)\,dt\right| \le \sim \mathcal{O}\left(N^{\max(-p,\frac{1}{2}-\ell)}\right), \quad 1\le \ell \le 10. \tag{3.3.17}$$

3.3.4. FILTERING AND THE COS METHOD

One of the reasons why the COS method is highly efficient is because pricing formula (3.2.17) works directly with the coefficients, without a-priori recovery of the functions $p(y|x)$ or $g(y)$. Spectral filters work strictly in the Fourier domain and therefore they can be used directly in the COS pricing formula. Once a suitable filter and order have been chosen one can multiply the Fourier cosine coefficients by a factor $\hat{s}(k/N)$ and work with the COS method as before. This gives us the *filter-COS pricing formula*, simply as follows

$$\hat{v}^{filter}(t_0,x) := e^{-r\Delta t}\sum_{k=0}^{N}{}'\,\hat{s}\left(\tfrac{k}{N}\right)\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}\mathcal{G}_k, \tag{3.3.18}$$

where $\hat{s}$ can be any non-adaptive filter. It does not add significant computational costs. Unfortunately, the same does not hold for the adaptive filters [TT05, Boy11, Tan06], because if we vary coefficients depending on position we cannot use the substitution of equation (3.2.16) which leads to the COS method anymore.

3.3.5. ERROR ANALYSIS COS METHOD AND FILTER-COS METHOD

Without Filtering The error of the COS formula without filtering terms is composed of three parts: The integration range truncation error, the series truncation error, and the error related to approximating $\mathcal{P}_k$ by Φ_k. Both a density function and payoff function with a discontinuity in one of their derivatives, results in an algebraic convergence. If the computational domain $[a,b]$ is chosen sufficiently wide, then the so-called *series truncation error* $E_{Tr}^{COS}(t_0,x;N)$, i.e.,

$$\begin{aligned} E_{Tr}^{COS}(t_0,x;N) &:= e^{-r\Delta t}\tfrac{b-a}{2}\sum_{k=N+1}^{\infty}\Phi_k(x)\mathcal{G}_k \\ &= e^{-r\Delta t}\int_a^b g(y)\left[p(y|x)-p_N(y|x)\right]dy, \end{aligned} \tag{3.3.19}$$

dominates the total error. Here $p_N(y|x)$ denotes the truncated Fourier cosine series of the density on $[a,b]$. The series truncation error depends on the smoothness of underlying probability density function $p(y|x)$ and payoff function $g(y)$. With equation (3.3.19) it follows that if either Φ_k or $\mathcal{G}_k$ decreases exponentially, then the error has exponential convergence in N. If however both decay algebraically, then we end up with algebraic convergence.

If $\Phi_k(x)$ and $\mathscr{G}_k$ have algebraic index of decay $n_\varphi(x)$ and n_g, respectively, then Proposition 3.3.1 gives $E_{Tr}^{COS}(t_0,x;N) \sim \mathscr{O}\left(N^{-n_\varphi(x)-n_g+1}\right)$. In the numerical examples we will observe convergence rates $E_{Tr}^{COS}(t_0,x;N) \sim \mathscr{O}\left(N^{-n_\varphi(x)-n_g}\right)$, which is probably due to the alternating behavior of the series.

Payoff functions, like puts and calls, are in general non-smooth, which is the reason for slowly decreasing Fourier coefficients. Asset prices modeled by geometric Brownian motion, jump-diffusion, or the Heston model lead to exponential decay of the coefficients Φ_k, resulting in an exponentially converging COS formula. However, the density functions of the Variance Gamma and CGMY models may be non-smooth. Together with a non-smooth payoff function this will result in rather slow algebraic error convergence in N. Applying an appropriate spectral filter will improve the convergence rate.

With Filtering Applying spectral filter $\hat{s}$ in the COS formula leads to the following relation,

$$\begin{aligned} E_{Tr}^{filter-COS}(t_0,x;N) &:= e^{-r\Delta t}\tfrac{b-a}{2}{\sum_{k=0}^{N}}'\left(1-\hat{s}\left(k/N\right)\right)\Phi_k(x)\mathscr{G}_k + e^{-r\Delta t}\tfrac{b-a}{2}\sum_{k=N+1}^{\infty}\Phi_k(x)\mathscr{G}_k \\ &= e^{-r\Delta t}\int_a^b g(y)\left[p(y|x)-p_N^{\hat{s}}(y|x)\right]dy. \end{aligned} \tag{3.3.20}$$

The absolute value can be bounded by

$$|E_{Tr}^{filter-COS}(t_0,x;N)| \leq e^{-r\Delta t}\int_a^b |g(y)||p(y|x)-p_N^{\hat{s}}(y|x)|\,dy = \mathscr{O}\left(N^{-n_p^{\hat{s}}(x)}\right). \tag{3.3.21}$$

$n_p^{\hat{s}}(x)$ denotes the algebraic index of convergence of the error $p(y|x)-p_N^{\hat{s}}(y|x)$ of the filtered partial sum, which is discussed in Section 3.3.3 (Theorem 3.3.2). The absolute error depends on the distance to the discontinuity of the density function and may be larger for strike prices K close to the discontinuity. Only *at* the discontinuity the error will not improve, but we integrate over the whole interval $[a,b]$. We observe higher absolute errors in the option value for strikes near the discontinuity of the density function, but appropriate filters will improve the convergence *rate*. The filter-COS formula is beneficial especially when the number of terms N increases.

Option's Greeks Computing the option's Greeks was briefly described in Section 3.2.2.1. The error of the COS formula for the option's Delta, without filtering, is given by

$$E_{Tr}^{COS-\Delta}(t_0,x;N) := e^{-r\Delta t}\tfrac{b-a}{2}\sum_{k=N+1}^{\infty}\tfrac{\partial\Phi_k(x)}{\partial x}\mathscr{G}_k\tfrac{1}{S}, \tag{3.3.22a}$$

$$\tfrac{\partial\Phi_k(x)}{\partial x} = \tfrac{2}{b-a}\Re\left\{\phi_{levy}\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\tfrac{ik\pi}{b-a}\right\}. \tag{3.3.22b}$$

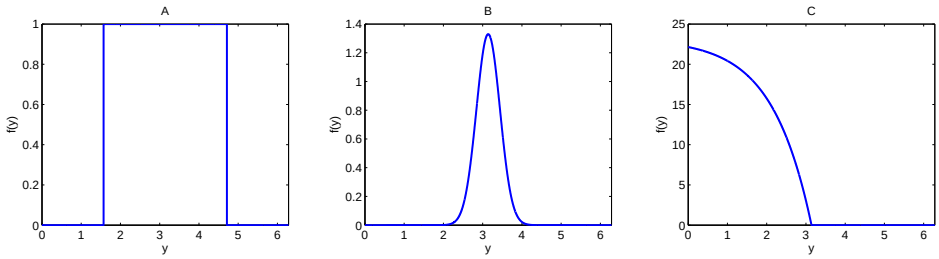
As the coefficients for the Greeks are multiplied by factors $\frac{ik\pi}{b-a}$, the algebraic index of convergence is reduced by one. So, finding a faster converging method becomes even more beneficial for the option's Greeks. With a filter-COS formula for the option's Greeks we achieve the same convergence rates as for the option prices.

3.4. NUMERICAL EXPERIMENTS

In this section we discuss several numerical experiments supporting the insights from earlier sections. MATLAB 7.11.0 R2010b is used for the computations, with double precision. We start in Section 3.4.1 with three basic test functions $f(y)$, representing the option payoff or probability density features. Subsequently, in Section 3.4.2.1 the density recovery of the Variance Gamma process is studied. Convergence of option prices, computed by the filter-COS method, is discussed in Section 3.4.2.2 for European-style options and in Section 3.4.2.3 for Bermudan-style options. An example from portfolio loss modeling, resulting in a staircase distribution function, is presented in Section 3.4.3.

3.4.1. CONVERGENCE TEST FUNCTIONS

We perform tests with three different functions $f(y)$ and different filters. The test functions are shown in Figure 3.4.1, with $[a,b] = [0,2\pi]$. Function A represents a block function with two jump discontinuities, function B is smooth, and function C has a discontinuity in the first derivative.



(a) $f(y) = \mathbf{1}_{y\in[0.5\pi,1.5\pi]}$.

(b) $f(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}\right)$ $(\mu = \pi, \sigma = 0.3)$.

(c) $f(y) = \max(e^{\pi} - e^{y}, 0)$.

Figure 3.4.1: Three test functions $f(y)$.

The left-side plot in Figure 3.4.2 shows function A (green) and its Fourier series approximation without (red) and with (blue) exponential filter of order $p = 4$ ($N = 32$). The middle plot gives the corresponding error. At the jump discontinuities the limit converges to the average of the values of the function at either side of the jump. In the right-side plot we display the convergence of the error at the point $y = \pi$, for increasing values of N (log-log plot). The local effect of the Gibbs phenomenon gives rise to oscillations near the jumps and the partial sum does not converge at the jump, thereby resulting in a lack of uniform convergence. However, there also is a global effect: although the error decays away from the jumps, the algebraic pointwise convergence is only first order, $\mathcal{O}(1/N)$. The filter improves the error significantly, especially away from the jump discontinuities. The error at point $y = \pi$ decreases exponentially in N due to the usage of the filter.

To recover the functions, we test the performance of six filters, i.e., the Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter, and the Erfc-Log filter (see Section 3.3.2). We also employ different filter orders. Table 3.4.1 presents the algebraic index of convergence observed for the three test functions. Here "exp" denotes exponential convergence. With two numbers, e.g., "2, 4", convergence is order $\mathcal{O}(N^{-2})$ on $[0,\pi]$ and $\mathcal{O}(N^{-4})$ on $[\pi,2\pi]$. A star (*) in the table indicates that the order of convergence

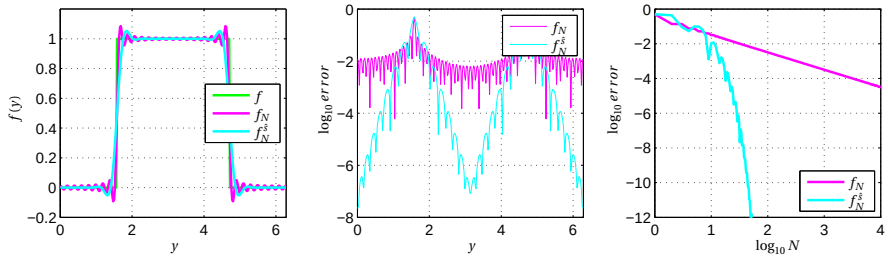


Figure 3.4.2: Recovery $f(y)$ (function A) without and with filter (exponential, order $p = 4$) ($N = 32$).

was not clearly measurable in our numerical experiments.

Table 3.4.1: Algebraic index of convergence test functions.

Filter	Function A	Function B	Function C
No filter	1	exp	1
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	3	2	2, 3
Exponential ($p = 2$)	exp	2	2, exp
Exponential ($p = 4$)	exp	4	4, exp
Exponential ($p = 6$)	exp	6	*, exp
Vandeven ($p = 1,3,5$)	p	p	p
Vandeven ($p = 2,4$)	$p+1$	p	p, $p+1$
Erfc-Log ($p = 1,2,3,4,5$)	p	p	p

Based on the error analysis of Theorem 3.3.2 and our observations in Section 3.3.3 we can explain the numbers in Table 3.4.1 as follows:

For function A we have

$$\begin{aligned} |f_N^{\hat{s}}(y) - f(y)| &= \sum_m \tfrac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-) \right) + \tfrac{1}{2\pi} \int_0^{2\pi} s_1(y-t) f^{(1)}(t)\, dt \\ &= \sum_m \tfrac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-) \right). \end{aligned} \tag{3.4.1}$$

The observed decay rates of s_1 are given in Table 3.3.1 and they correspond to the results in the second column in Table 3.4.1.

Function B is approximately smooth, so that

$$|f_N^{\hat{s}}(y) - f(y)| = \tfrac{1}{2\pi} \int_0^{2\pi} s_q(y-t) f^{(q)}(t)\, dt, \quad \forall q \geq 1. \tag{3.4.2}$$

Following the arguments in Section 3.3.3 we expect

$$\tfrac{1}{2\pi} \int_0^{2\pi} s_q(y-t) f^{(q)}(t)\, dt \sim \mathcal{O}\left(N^{\max(-p, \frac{1}{2}-q)} \right). \tag{3.4.3}$$

This gives us the convergence order $\mathcal{O}(N^{-p})$, as observed.

The error of function C corresponds to the error of function B for $y \in [0, \pi]$. And the error corresponds to the error of function A for $y \in [\pi, 2\pi]$.

3.4.2. THE VARIANCE GAMMA PROCESS

In this section, we discuss various applications of the COS method in the context of the Variance Gamma (VG) process [MCC98, Sch03]. We start with the accurate and efficient recovery of the VG density. Then European and Bermudan option prices are analyzed.

3.4.2.1. VARIANCE GAMMA DENSITY RECOVERY

In the case of modeling asset prices by a fat-tailed density function, the exponential Variance Gamma jump process can be applied. The Variance Gamma process is obtained by evaluating a Brownian motion with drift ϑ and volatility σ at a random time given by a gamma process γ_t with mean rate one and variance rate ν, [MCC98, Sch03]:

$$L_t^{VG} = \vartheta\gamma_t + \sigma\omega_{\gamma_t}. \tag{3.4.4}$$

The risk-neutral asset price is then defined as $S_t = S_{t_0} e^{(r+w)\Delta t} e^{L_t^{VG}}$, $w = \frac{1}{\nu}\log(1-\vartheta\nu-\frac{1}{2}\sigma^2\nu)$ [FO09]. The VG process is of bounded variation, has independent, stationary increments, and is defined by an infinite arrival of jumps. The VG density can be characterized by a *fat tail*: it is suitable to model phenomena where small and relatively large asset values are more probable than would be the case for the lognormal distribution. The characteristic function, $\varphi(u|x) = \mathbb{E}[e^{iuX_T}|X_{t_0} = x] = e^{iux+iu(r+w)\Delta t}\phi_{VG}(u)$, is given by [Sch03, MCC98], with

$$\phi_{VG}(u) = \left(1 - iu\vartheta\nu + \tfrac{1}{2}\sigma^2\nu u^2\right)^{-\Delta t/\nu} \sim \mathcal{O}\left(u^{-2\Delta t/\nu}\right). \tag{3.4.5}$$

In [MCC98], the following expression for the VG density function was derived:

$$p_{VG}(y) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi z}} \exp\left(-\frac{(y-\vartheta z)^2}{2\sigma^2 z}\right) \frac{z^{\frac{\Delta t}{\nu}-1}\exp(-\frac{z}{\nu})}{\nu^{\frac{\Delta t}{\nu}}\Gamma(\frac{\Delta t}{\nu})}\, dz, \tag{3.4.6}$$

with Γ the Gamma function. It is computationally rather expensive to evaluate (3.4.6) at each point in the domain of interest. The smoothness of the density function depends on its parameters, to be more precise, with higher values of $\Delta t/\nu$ a larger number of derivatives exists.

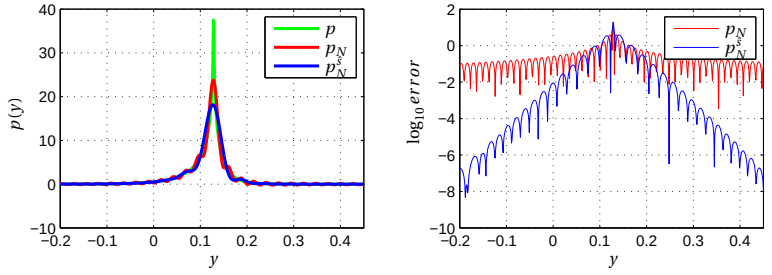


Figure 3.4.3: VG density and error for $T = 0.1$ ($N = 128$).

The parameters used for our tests here are the same as in [FO08], i.e.,:

$$K = 90, S_0 = 100, r = 0.1, \sigma = 0.12, \vartheta = -0.14, \nu = 0.2. \tag{3.4.7}$$

Figure 3.4.3 shows the VG density for terminal time $T = 0.1$ (green). For small values of T the peak in the density gets really sharp and is difficult to approximate it accurately by a Fourier cosine series. For increasing ν-values the peak sharpens, while increasing σ-values result in a smaller peak and wider tails, without altering the smoothness properties of the function around the peak significantly.

Figure 3.4.3 also shows the approximations by Fourier cosine series, i.e., equation (3.2.9), with and without exponential filter (order $p = 6$). The characteristic function exhibits an algebraic decay with order $2\Delta t/\nu$, giving rise to an algebraic decay of the Fourier coefficients and therefore slow convergence of the COS method, especially for the densities with small time interval Δt. Filtering works well away from the peak in the VG density, but right at the peak the approximation becomes somewhat worse. Note that for larger T-values, like for $T = 1$, the function is sufficiently smooth and addition of the filter does not improve its already accurate approximation.

We evaluate the performance of the filters to recover the VG density. The six different filters are used and we experiment with different filter orders. We find that lower order filters smoothen the sharp peak too much. Table 3.4.2 shows the measured algebraic index of convergence for the density recovery. We observe a significant improvement for $T = 0.025$ and $T = 0.1$, whereas for $T = 1$ the 'no filter case' is superior.

Table 3.4.2: Algebraic index of convergence VG density recovery with three values of T (The star (*) indicates that the order of convergence was not clearly measurable in our numerical experiments).

Filter	$T = 0.025$	$T = 0.1$	$T = 1$
No filter	0.25	1	10
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	2	2	2
Exponential ($p = 2,4$)	p	p	p
Exponential ($p = 6$)	*	6	6
Vandeven ($p = 1,2,3,4,5$)	p	p	p
Erfc-Log ($p = 1,2,3,4,5$)	p	p	p

3.4.2.2. EUROPEAN OPTIONS AND GREEKS UNDER VARIANCE GAMMA

We investigate here the convergence of the COS and the filter-COS method for pricing European-style options. The payoff function of a put reads $g(S) = \max(K - S, 0)$. For a digital (or binary) call option the payoff is either one or zero, with payoff function $g(S) = \mathbf{1}_{S \geq K}$. Reference values for the experiments are obtained by selecting an accurate filter and a very large number of terms in the series expansions, see Table 3.4.3.

The convergence of the COS method *without filtering* is of order $2\Delta t/\nu + 2$ for put options and $2\Delta t/\nu + 1$ for digital options. The algebraic index of convergence for the Greek Delta Δ (equation (3.2.19a)), without filtering, is one order lower and for Gamma Γ (equation (3.2.19b)) it is two orders lower, see Table 3.4.4. The absolute value of the series terms gives us $\mathcal{O}\left(1/k^{2\Delta t/\nu+1-2}\right)$ convergence for the Gamma Γ of a digital option. So, the algebraic index of convergence is even *nonpositive* for expiration times $T = 0.1$ and $T = 0.025$ and the series does not converge ("Div"), because of the Cauchy convergence criterion. The intuition behind this is that for very short expiration times, the Γ converges to the Dirac delta function.

Table 3.4.5 shows the observed order of convergence of the European option with different spectral filters. Similar convergence results are obtained for the Greeks Δ and Γ. We tend to prefer the exponential filter to the Vandeven and Erfc-Log filters. Its implementation is easiest and fastest, although the other filters are not significantly more time-consuming. Furthermore, we observe an exponential convergence for step functions with the exponential filter (see Figure 3.4.2 and Table 3.4.1), which is advantageous for recovery of a distribution function for discrete random variables. Therefore, we focus on the exponential filter in the remainder of this chapter.

For asset price processes with independent, stationary increments, like the VG model, we can employ the filter-COS method to compute option values for multiple strike prices simultaneously. For example, for $K \in [80,120]$ we obtain the same convergence results. However, we observe higher absolute errors in the option values for strike prices near the peak value in the VG density function. This can be explained by the smoothing of the peak by a filter.

Figure 3.4.4 displays the error of the COS formula for the option value and the Greeks for expiration time $T = 0.1$ in a log-log plot. Exponential filters with different orders are used. Note that for some cases the so-called roundoff-plateau, with minimal attainable accuracy due to machine precision, is reached. The use of filters improves the error and convergence order significantly, especially regarding the option's Gamma Γ.

Table 3.4.3: Reference values European options.

	$T = 0.025$		$T = 0.1$		$T = 1$	
	Put	Digital	Put	Digital	Put	Digital
Option value $v(t_0,x)$	0.02435	89.1883	0.09819	86.9759	0.53472	74.7855
Delta Δ	-0.50629	12.8417	-1.82737	40.6550	-5.50365	63.1902
Gamma Γ	11.5565	-313.402	36.5895	-843.107	56.8712	-581.247

Table 3.4.4: Algebraic index of convergence for European options (no filter).

	$T = 0.025$		$T = 0.1$		$T = 1$	
	Put	Digital	Put	Digital	Put	Digital
Option value $v(t_0,x)$	2.25	1.25	3	2	12	11
Delta Δ	1.25	0.25	2	1	11	10
Gamma Γ	0.25	Div	1	Div	10	9

Table 3.4.5: Algebraic index of convergence for European options observed by filter-COS ($T = 0.025$, $T = 0.1$, and $T = 1$).

Filter	Put	Digital
Fejér	1	1
Lanczos	2	2
Raised cosine	2	2
Exponential ($p = 2,4,6$)	p	p
Vandeven ($p = 1,2,3,4,5$)	p	p
Erfc-Log ($p = 1,2,3,4,5$)	p	p

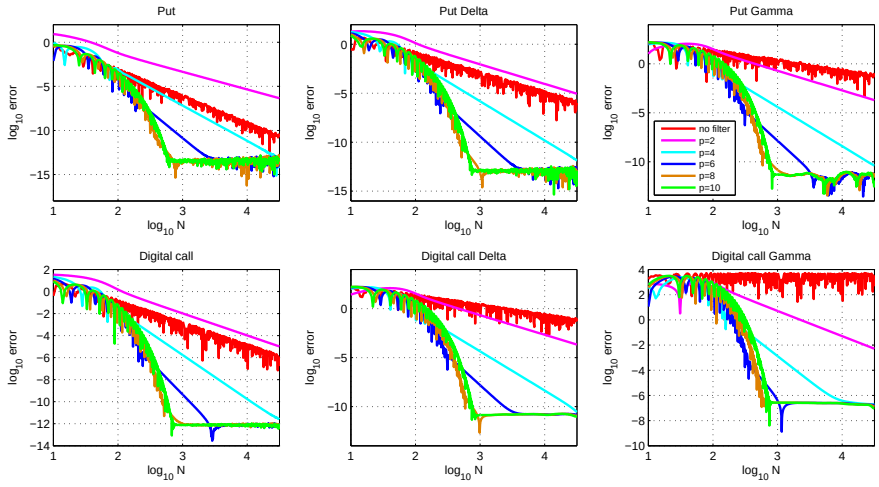


Figure 3.4.4: Error convergence for European options; put (top) and digital call (bottom) plus the Greeks, with exponential filters ($T = 0.1$).

3.4.2.3. BERMUDAN OPTION PRICE UNDER VG

A Bermudan-style option can be exercised at a *set of M early-exercise dates* prior to the expiration time T, $t_0 < t_1 < \ldots t_m < \ldots < t_M = T$, with timestep $\Delta t := t_{m+1} - t_m$. The authors in [FO09] describe a recursive algorithm, based on the COS method, for pricing Bermudan options backwards in time via Bellman's principle of optimality. This algorithm is similar to the method described in Chapter 2. We also employ the COS method for Bermudan options here, but replace the coefficients Φ_k by the filtered version, i.e., by $\hat{s}(k/N)\Phi_k$, similar as in equation (3.3.18).

In Table 3.4.6 and in Figure 3.4.5 the results for the exponential filter with $T = 1$ and different numbers of early-exercise dates are presented. The COS method with filtering becomes more and more beneficial when more early-exercise dates are used, for example $M = 20$. Then the algebraic convergence rates are, however, not clearly measurable. A larger number of exercise dates implies a smaller timestep Δt between the exercise dates, which decreases the convergence rate of the COS formula without filtering.

3.4.3. PORTFOLIO LOSS DISTRIBUTION

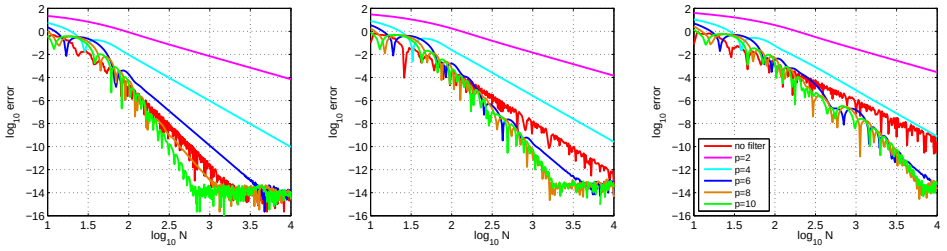
In this section we present a final example, which has a financial background in risk management. The distribution function of interest will be a staircase function, which is difficult to approximate by Fourier series. The filtering technique will however improve the results significantly.

For a bank it is important to manage the risk originating from its business activities. The credit risk underlying a credit portfolio is one of the largest risk portions of a bank. For quantifying losses in credit portfolios one often looks at the Value at Risk (VaR). The VaR of a portfolio at confidence level α is given by the smallest value x, for which the probability that loss L exceeds x is at most $(1-\alpha)$:

$$\text{VaR}_\alpha = \inf\{x \in \mathbb{R} : \mathbb{P}(L > x) \leq (1-\alpha)\} = \inf\{x \in \mathbb{R} : F_L(x) \geq \alpha\}, \tag{3.4.8}$$

Table 3.4.6: Algebraic index of convergence for Bermudan put options ($T = 1$) (The star (*) indicates that the order of convergence was not clearly measurable in our numerical experiments).

Filter	$T = 1$		
	$M = 2$	$M = 4$	$M = 8$
No filter	7	4.5	3.25
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	2	2	2
Exponential ($p = 2,4$)	p	p	p
Exponential ($p = 6$)	6	6	*
Vandeven ($p = 1,2,3,4,5$)	p	p	p
Erfc-Log ($p = 1,2,3,4,5$)	p	p	p



(a) $M = 2$, $v(t_0,x) = 0.64386$. (b) $M = 4$, $v(t_0,x) = 0.71161$. (c) $M = 8$, $v(t_0,x) = 0.75220$.

Figure 3.4.5: Error convergence for Bermudan options with an increasing number of early-exercise dates, filter-COS technique with exponential filters ($T = 1$).

where F_L is the cumulative loss distribution function.

The Vasicek model [Vas02] is often used to find an approximation to the loss distribution and to compute the VaR. Under this model losses occur when an obligor defaults in a fixed time horizon. Suppose there are $m = 1,\ldots,M$ issuers and X_m represents the individual asset return of issuer m. In accordance with the Vasicek model we then use

$$X_m = \sqrt{\rho_m}Y + \sqrt{1-\rho_m}Z_m, \quad m = 1,\ldots,M, \tag{3.4.9}$$

where Y is a common economic factor, Z_m is the idiosyncratic factor for issuer m, and ρ_m is the correlation between Y and Z_m. All random variables are assumed to follow a standard normal distribution and Y and Z_m are independent. If the asset return falls below a default threshold c_m, there is a loss λ_m. We define the default probability of issuer m by $p_m := \mathbb{P}(X_m < c_m)$. The individual credit loss is defined by $L_m = \lambda_m \mathbf{1}_{X_m<c_m}$ and the total portfolio loss reads $L = \sum_{m=1}^{M} L_m$. If there is only one issuer then the total loss is given by the discrete random variable $\lambda_1 \mathbf{1}_{X_1<c_1}$.

The COS method can be used to approximate the cumulative distribution function of the loss by means of its characteristic function, see equation (3.2.13), and thus to calculate the VaR. The probability distribution is based on a discrete set of events, resulting in a *stepwise distribution function*, which causes the COS method to suffer from the Gibbs phenomenon: with a loss distribution which is discontinuous, significant errors appear around

the points of discontinuity [Fan12].

Following [MOG11] we take $M = 20$ issuers, with default probability p_m of 1%, asset correlation ρ_m of 50%, and exposure $\lambda_m = 1$. For the COS method we take $N = 2^{10}$, $[a,b] = \left[0, \sum_{m=1}^{M} \lambda_m\right]$. The characteristic function of L can be written as

$$\varphi_L(u) = \mathbb{E}[e^{iuL}] = \mathbb{E}\left[\mathbb{E}\left[e^{iu\sum_{m=1}^{M}\lambda_m \mathbf{1}_{X_m<c_m}} \middle| Y\right]\right]. \tag{3.4.10}$$

An analytic expression is available for the inner conditional expectation. In [Fan12] an integration rule to approximate the outer expectation is employed, whereas in [GRM06] Monte Carlo simulations on Y was used. We test both approaches, with a grid, $y = [-5 : 0.1 : 5]$, for the numerical integration (denoted by COS+NI) and by 5000 simulated values for Y in the Monte Carlo experiment (denoted by COS+MC). With these choices, the computation of the characteristic function, which is the most time-consuming part, is approximately 50 times more expensive for COS+MC.

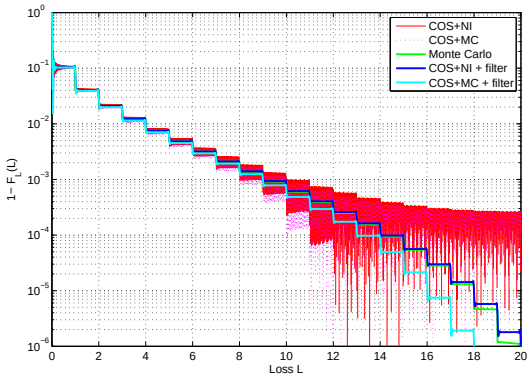


Figure 3.4.6: Recovery of portfolio loss distribution function using COS method and filtered version (exponential filter, order $p = 6$).

The loss distribution of this example portfolio is plotted in Figure 3.4.6. An accurate estimation of $1 - F_L(L)$ results in an accurate Value at Risk. The green line is the result of a full Monte Carlo simulation, with 100,000,000 replications for each X_m, and serves as our reference solution here. The red-dotted and magenta-dotted lines are the COS method approximations with numerical integration and Monte Carlo simulation for Y, respectively. The blue and cyan lines are the filtered-COS results, using an exponential filter with $p = 6$. The COS method without filtering shows a highly oscillatory behavior and does not give accurate results in the tail, where $1 - F_L(L)$ is very small. The COS method with filter *and* with numerical integration, gives however highly accurate results that correspond very well to the full Monte Carlo simulation. The difference gets smaller when the number of Monte Carlo simulations is increased. The results with the filter and Monte Carlo simulations for Y, COS+MC, are sensitive to outliers of draws for Y and these results are not very satisfactory. We would also like to mention that approximations based on Haar wavelets [OGO13, MOG11] give accurate portfolio loss VaR estimates as well.

Remark 3.4.1. *The model and computational technique can be extended to higher-dimensional systematic risk factors $Y = [Y_1, \ldots, Y_d]$. In that case, the computation of the outer expectation in equation* (3.4.10) *can be performed by adaptive integration, as in [Hua09].*

3.5. CONCLUSIONS

The COS method is an option pricing method based on Fourier cosine expansions which performs very well in general. When the underlying density function is smooth, we achieve an exponential convergence in the number of cosine coefficients. When the underlying density is not smooth, however, the method may suffer from the Gibbs phenomenon and the convergence is only of algebraic order. A filtering technique to improve the convergence rate for these cases has been discussed in this chapter. In practical cases where the COS method degrades due to discontinuities in functions, the convergence with a filter improves significantly in terms of the number of required Fourier coefficients as well as in CPU time.

Non-adaptive spectral filtering takes place in the Fourier domain and combines therefore very well with the COS option pricing formula, without adding significant computational costs. The Fourier coefficients are premultiplied by a decreasing function $\hat{s}(k/N)$ so that they decay faster and so that the convergence rate away from a discontinuity is improved. The technique can be used for one-dimensional problems, but also in higher dimensions.

In the numerical examples we tested six different filters, i.e., the Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter, and the Erfc-Log filter. Especially the exponential filter gave highly accurate results for stepwise functions.

The plain COS method for option pricing under the Variance Gamma asset price process results in algebraic convergence. Our filter-COS formula improves the algebraic index of convergence, in particular for short time horizons. Moreover, for the computation of the option's Greeks, that suffer from an even lower convergence rate without filtering, spectral filters are highly beneficial. As another example in finance, we discussed portfolio loss modeling. Discrete random variables then give rise to stepwise cumulative distribution functions. We derived a COS formula to recover the distribution, which, of course, also suffered from the Gibbs phenomenon and the resulting oscillations. The approximation drastically improves by applying spectral filters.

Improved convergence comes without additional computational costs in these applications and the fact that the filtering is easy to implement, even in multiple dimensions, makes the filter-COS method a natural solution for some of the problems described.

CHAPTER 4

Two-Dimensional Fourier Cosine Series Expansion Method for Pricing Financial Options

4.1. INTRODUCTION

In financial markets traders deal in assets and options, like the well-known call and put options. Besides these, many 'exotic' options have been defined that have more complex contract details and are not traded at regulated exchanges.

One class of exotic option contracts is called the class of multicolor *rainbow options*, whose payoff may depend on multiple assets, like on the average or the maximum of asset prices. The value of the option depends on the contract details and on the underlying asset prices.

Computational finance deals with numerical and computational questions regarding efficient option pricing and calibration. Usually, an asset price model is calibrated to liquidly available plain vanilla options (calls and puts) from a regulated exchange. For the valuation of the exotic options other computational methods are typically used. Option pricing techniques can be divided into the categories of Monte Carlo simulation, partial differential equation (PDE) methods, and Fourier-based methods. Often Monte Carlo methods are used to price high-dimensional option contracts. The method presented here can be seen as an alternative (deterministic) pricing technique, which can deal with multi-asset option problems of *medium-sized* dimensionality, meaning resolving two-dimensional (2D) to approximately five-dimensional (5D) integrals. The method we propose for pricing higher-dimensional options is based on the Fourier transform of the transitional density function and is especially suitable for asset price models in the class of Lévy processes.

The previous strain of literature on the COS method was based on the one-dimensional (1D) characteristic function of a single stochastic process. In this chapter, we extend the Fourier method to higher dimensions and price in particular two-color *rainbow options*, which are contracts written on two underlying assets. Well-known examples include the valuation of basket and call-on-maximum options. Methods for both European- and Bermudan-style rainbow options are developed here. The resulting algorithm can be applied to models such as correlated geometric Brownian motions or multidimensional processes with jumps. The method is highly efficient for asset prices in the class of Lévy processes. In the literature, mainly Monte Carlo-based methods are being used to solve higher-dimensional pricing problems, see [AB04, BS07, Boy88, BEG89]. Leentvaar and Oosterlee worked on a parallel Fourier-based method [LO08a] and parallel sparse grid methods

This chapter is based on the article 'Two-dimensional Fourier cosine series expansion method for pricing financial options', published in *SIAM Journal on Scientific Computing*, Vol. 34(5), B642-B671, 2012 [RO12b].

[LO08b] for pricing multi-asset options. The authors in [CF08] demonstrate an implicit PDE discretization method for rainbow options under jump-diffusion processes. We will compare our results to reference values in the literature mentioned above.

The methodology presented here can also be applied to pricing options with one underlying asset, for which the dynamics are governed by two or more correlated stochastic processes. For example, the popular *Heston model* describes an asset price process with a stochastic volatility [Hes93]. The instantaneous variance process follows a mean-reverting square root (CIR) process. This model is able to capture smile and skew patterns in the implied volatility surface. Besides, the closed-form univariate characteristic function of the log-asset price process makes it easy to implement the Heston model in Fourier-based methods, see [AMST07, LK10] for European calls. For the valuation of Bermudan and discrete barrier options, the authors of [FO11] combined the COS formula for the log-asset dimension and a quadrature rule in the log-variance dimension. Since the bivariate characteristic function of the log-asset price and variance is available, we can also apply the 2D-COS formula to this problem. We investigate the 2D-COS method particularly for Bermudan put options under the Heston dynamics.

The outline of this chapter is as follows. We start with the presentation of the 2D-COS formula for pricing European rainbow options (Section 4.2) and the 2D-COS method for solving Bermudan pricing problems (Section 4.3). Section 4.4 discusses option pricing under the Heston model, which is an affine diffusion process, but not in the Lévy class. The error analysis in Section 4.5 indicates an exponentially converging error for smooth density functions. A non-smooth density function results in algebraic convergence. Then, in Sections 4.6 and 4.7, numerical tests are performed. The 2D-COS method can easily be extended to higher dimensions and we give some insights into the possibilities and difficulties in Section 4.8. Section 4.9 concludes this chapter's findings.

4.2. EUROPEAN RAINBOW OPTIONS

In this section, we explain the *2D-COS formula* for approximating discounted expected payoffs. The method is based on the Fourier cosine series of the payoff function and the density. The density function of a stochastic process is usually not known, but often its characteristic function is known (see [DPS00, FO08]). This enables us to approximate the Fourier coefficients efficiently.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $T > 0$ be a finite terminal time, and $\mathbb{F} = (\mathcal{F}_s)_{0 \le s \le T}$ be a filtration satisfying the usual conditions. The process $\mathbf{X}_t = (X_t^1, X_t^2)$ denotes a 2D stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, representing the *log-asset prices*. We assume that the bivariate characteristic function of the stochastic process is known, which is the case, for example, for affine jump-diffusions [DPS00]. The value of a European rainbow option, with payoff function $g(.)$, is given by the risk-neutral option valuation formula ([Shr08])

$$v(t_0, \mathbf{x}) = e^{-r\Delta t} \mathbb{E}^{t_0, \mathbf{x}} \left[g(\mathbf{X}_T) \right] = e^{-r\Delta t} \iint_{\mathbb{R}^2} g(\mathbf{y})\, p(\mathbf{y}|\mathbf{x})\, d\mathbf{y}. \tag{4.2.1}$$

Here, $\mathbf{x} = (x_1, x_2)$ is the current state, $p(y_1, y_2|x_1, x_2)$ is the conditional density function, r is the risk-free rate, and time to expiration is denoted by $\Delta t := T - t_0$. In the derivation of the COS formula, we distinguish three different approximation steps. The errors introduced in each step are discussed in Section 4.5.1.

Step 1: We assume that the integrand is integrable, which is common for the problems we deal with. Because of that, we can, for given $\mathbf{x}$, truncate the infinite integration ranges

to some domain $[a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$ without losing significant accuracy. This gives the multidimensional Fourier cosine expansion formulation

$$\begin{aligned} v_1(t_0, \mathbf{x}) &= e^{-r\Delta t} \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(\mathbf{y}) p(\mathbf{y}|\mathbf{x}) d y_1 d y_2 \\ &= e^{-r\Delta t} \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(\mathbf{y}) \sum_{k_1=0}^{\infty}{}' \sum_{k_2=0}^{\infty}{}' \mathscr{P}_{k_1,k_2}(\mathbf{x}) \cos\left(k_1 \pi \tfrac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \tfrac{y_2-a_2}{b_2-a_2}\right) d y_1 d y_2. \end{aligned} \tag{4.2.2}$$

The notation v_i is used for the different approximations of v and keeps track of the numerical errors that set in from each step. For final approximations we also use the 'hat'-notation, like $\hat{v}$, $\hat{c}$, etc. In the second line in (4.2.2), the conditional density is replaced by its Fourier cosine expansion in $\mathbf{y}$ on $[a_1, b_1] \times [a_2, b_2]$, with series coefficients $\{\mathscr{P}_{k_1,k_2}(\mathbf{x})\}_{k_1=0}^{\infty}$ defined by

$$\mathscr{P}_{k_1,k_2}(\mathbf{x}) := \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \int_{a_2}^{b_2} \int_{a_1}^{b_1} p(\mathbf{y}|\mathbf{x}) \cos\left(k_1 \pi \tfrac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \tfrac{y_2-a_2}{b_2-a_2}\right) d y_1 d y_2. \tag{4.2.3}$$

$\sum'$ in (4.2.2) means that the first term of the summation has half weight. We interchange summation and integration and define

$$\mathscr{V}_{k_1,k_2}(T) := \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(\mathbf{y}) \cos\left(k_1 \pi \tfrac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \tfrac{y_2-a_2}{b_2-a_2}\right) d y_1 d y_2, \tag{4.2.4}$$

which are the Fourier cosine series coefficients of $v(T, \mathbf{y}) = g(\mathbf{y})$ on $[a_1, b_1] \times [a_2, b_2]$.

Step 2: Truncation of the series summations gives

$$v_2(t_0, \mathbf{x}) = \tfrac{b_1-a_1}{2} \tfrac{b_2-a_2}{2} e^{-r\Delta t} \sum_{k_1=0}^{N_1-1}{}' \sum_{k_2=0}^{N_2-1}{}' \mathscr{P}_{k_1,k_2}(\mathbf{x}) \mathscr{V}_{k_1,k_2}(T). \tag{4.2.5}$$

Step 3: Next, the coefficients $\mathscr{P}_{k_1,k_2}(\mathbf{x})$ are approximated by

$$\Phi_{k_1,k_2}(\mathbf{x}) := \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \iint_{\mathbb{R}^2} p(\mathbf{y}|\mathbf{x}) \cos\left(k_1 \pi \tfrac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2 \pi \tfrac{y_2-a_2}{b_2-a_2}\right) d y_1 d y_2. \tag{4.2.6}$$

The 2D-COS formula is based on the following goniometric relation ([Keh]):

$$2\cos(\alpha)\cos(\beta) = \cos(\alpha+\beta) + \cos(\alpha-\beta). \tag{4.2.7}$$

With this we obtain

$$2\Phi_{k_1,k_2}(\mathbf{x}) = \Phi^{+}_{k_1,k_2}(\mathbf{x}) + \Phi^{-}_{k_1,k_2}(\mathbf{x}), \tag{4.2.8}$$

where

$$\Phi^{\pm}_{k_1,k_2}(\mathbf{x}) := \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \iint_{\mathbb{R}^2} p(\mathbf{y}|\mathbf{x}) \cos\left(k_1 \pi \tfrac{y_1-a_1}{b_1-a_1} \pm k_2 \pi \tfrac{y_2-a_2}{b_2-a_2}\right) d y_1 d y_2. \tag{4.2.9}$$

Now, the coefficients $\Phi^{\pm}_{k_1,k_2}(\mathbf{x})$ can be calculated by

$$\begin{aligned} \Phi^{\pm}_{k_1,k_2}(\mathbf{x}) &= \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \Re\Big\{ \iint_{\mathbb{R}^2} p(\mathbf{y}|\mathbf{x}) \exp\left(i k_1 \pi \tfrac{y_1}{b_1-a_1} \pm i k_2 \pi \tfrac{y_2}{b_2-a_2}\right) d\mathbf{y} \\ &\qquad \cdot \exp\left(-i k_1 \pi \tfrac{a_1}{b_1-a_1} \mp i k_2 \pi \tfrac{a_2}{b_2-a_2}\right) \Big\} \\ &= \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \Re\left\{ \varphi\left(\tfrac{k_1\pi}{b_1-a_1}, \pm \tfrac{k_2\pi}{b_2-a_2} \middle| \mathbf{x}\right) \exp\left(-i k_1 \pi \tfrac{a_1}{b_1-a_1} \mp i k_2 \pi \tfrac{a_2}{b_2-a_2}\right) \right\} \\ &= \tfrac{2}{b_1-a_1} \tfrac{2}{b_2-a_2} \Re\left\{ \phi_{levy}\left(\tfrac{k_1\pi}{b_1-a_1}, \pm \tfrac{k_2\pi}{b_2-a_2}\right) \exp\left(i k_1 \pi \tfrac{x_1-a_1}{b_1-a_1} \pm i k_2 \pi \tfrac{x_2-a_2}{b_2-a_2}\right) \right\}. \end{aligned} \tag{4.2.10}$$

$\Re\{.\}$ again denotes taking the real part of the input argument. $\varphi(.,.|\mathbf{x})$ is the *bivariate conditional characteristic function* of $\mathbf{X}_T$, given $\mathbf{X}_{t_0}=\mathbf{x}$, ([Soo73])

$$\varphi(\mathbf{u}|\mathbf{x})=\mathbb{E}\Big[e^{i\mathbf{u}\cdot\mathbf{X}_T}\Big|\mathcal{F}_{t_0}\Big]=\iint_{\mathbb{R}^2}e^{i\mathbf{u}\cdot\mathbf{y}}p(\mathbf{y}|\mathbf{x})d\mathbf{y}. \tag{4.2.11}$$

Examples of these characteristic functions can be found in Section 4.6. The last equality in (4.2.10) holds particularly for Lévy processes, for which $\phi_{levy}(u_1,u_2):=\varphi(u_1,u_2|0,0)$. Inserting (4.2.10) into (4.2.5) gives us the *2D-COS formula* for approximation of $v(t_0,\mathbf{x})$:

$$\begin{aligned}\hat{v}(t_0,\mathbf{x}) &:= \tfrac{b_1-a_1}{2}\tfrac{b_2-a_2}{2}e^{-r\Delta t}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}'\tfrac{1}{2}\Big[\Phi^{+}_{k_1,k_2}(\mathbf{x})+\Phi^{-}_{k_1,k_2}(\mathbf{x})\Big]\mathcal{V}_{k_1,k_2}(T)\\ &= e^{-r\Delta t}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}'\tfrac{1}{2}\Big[\Re\Big\{\phi_{levy}\Big(\tfrac{k_1\pi}{b_1-a_1},+\tfrac{k_2\pi}{b_2-a_2}\Big)\exp\Big(ik_1\pi\tfrac{x_1-a_1}{b_1-a_1}+ik_2\pi\tfrac{x_2-a_2}{b_2-a_2}\Big)\Big\}\\ &\quad+\Re\Big\{\phi_{levy}\Big(\tfrac{k_1\pi}{b_1-a_1},-\tfrac{k_2\pi}{b_2-a_2}\Big)\exp\Big(ik_1\pi\tfrac{x_1-a_1}{b_1-a_1}-ik_2\pi\tfrac{x_2-a_2}{b_2-a_2}\Big)\Big\}\Big]\mathcal{V}_{k_1,k_2}(T).\end{aligned} \tag{4.2.12}$$

With the multidimensional-COS formula, calculation of the option's Greeks is straightforward, as explained for the 1D case in Section 3.2.2.1.

Remark 4.2.1. *If the characteristic function is not available directly or not known analytically, it may be approximated. Local volatility models, for example, typically do not yield analytic functions φ, but recent research in [PPR11] proposes a second-order approximation formula, so that an approximate characteristic function may be derived.*

Another numerical approach approximates the stochastic process by, for example the well-known Euler scheme. Other stochastic Taylor schemes, such as the Milstein and Order 2.0 weak Taylor schemes, are discussed in [KP92] and Chapter 7. In Chapter 7 we explicitly derive the characteristic function for these discrete forms of a stochastic differential equation.

4

4.3. BERMUDAN RAINBOW OPTIONS

We generalize the multidimensional-COS method to pricing Bermudan rainbow options with a 2D underlying log-asset price process, $\mathbf{X}_t=(X_t^1,X_t^2)$, that is in the class of Lévy processes. A Bermudan option can be exercised at a fixed set of M early-exercise times $t_0<t_1<\ldots t_m<\ldots<t_M=T$, with $\Delta t:=t_{m+1}-t_m$. The payoff function is denoted by $g(.)$. The problem is solved backwards in time, with

$$\begin{cases} v(t_M,\mathbf{x}) &= g(\mathbf{x}),\\ c(t_{m-1},\mathbf{x}) &= e^{-r\Delta t}\mathbb{E}\big[v(t_m,\mathbf{X}_{t_m})|\mathbf{X}_{t_{m-1}}=\mathbf{x}\big],\\ v(t_{m-1},\mathbf{x}) &= \max[g(\mathbf{x}),c(t_{m-1},\mathbf{x})], \qquad 2\le m\le M,\\ v(t_0,\mathbf{x}_0) &= c(t_0,\mathbf{x}_0).\end{cases} \tag{4.3.1}$$

Function $c(t_{m-1},\mathbf{x})$ is called the continuation value and is approximated by the 2D-COS formula

$$\hat{c}(t_{m-1},\mathbf{x}):=\tfrac{b_1-a_1}{2}\tfrac{b_2-a_2}{2}e^{-r\Delta t}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}'\tfrac{1}{2}\Big[\Phi^{+}_{k_1,k_2}(\mathbf{x})+\Phi^{-}_{k_1,k_2}(\mathbf{x})\Big]\mathcal{V}_{k_1,k_2}(t_m). \tag{4.3.2}$$

The Fourier coefficients of the value function in (4.3.2) are given by

$$\mathcal{V}_{k_1,k_2}(t_m) := \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1} v(t_m,\mathbf{y})\cos\left(k_1\pi\frac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\frac{y_2-a_2}{b_2-a_2}\right)dy_1\,dy_2. \quad (4.3.3)$$

The option function is now approximated by $\hat{v}(t_{m-1},\mathbf{x}) := \max[g(\mathbf{x}), \hat{c}(t_{m-1},\mathbf{x})]$.

4.3.1. RECURSION FORMULA FOR COEFFICIENTS $\mathcal{V}_{k_1,k_2}(t_m)$

In this section, a recursive algorithm for recovering the coefficients $\mathcal{V}_{k_1,k_2}(t_m)$, backwards in time, is derived.

In the coefficients $\mathcal{V}_{k_1,k_2}(t_M)$, the terminal condition $v(t_M,\mathbf{y}) = g(\mathbf{y})$ appears. Some payoff functions provide analytic solutions to these coefficients in (4.3.3), otherwise they can be approximated, as explained in Section 4.3.2.1.

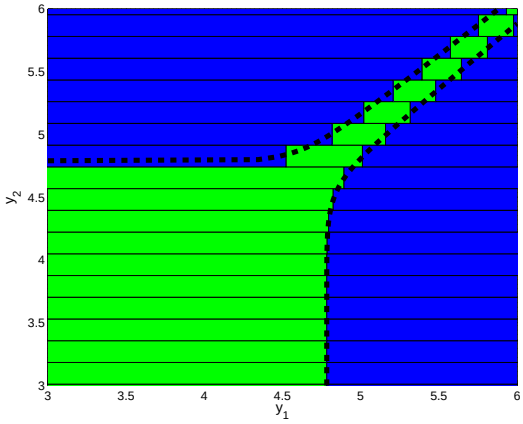


Figure 4.3.1: Rectangular regions $\mathcal{C}^q$ (green) and $\mathcal{G}^p$ (blue) and accurate boundary (dashed line).

For the coefficients that are used to approximate the continuation values at times $t_0,\ldots,t_{M-2}$, the value function, $v(t_m,\mathbf{y}) = \max\left[g(\mathbf{y}), c(t_m,\mathbf{y})\right]$, appears in the terms $\mathcal{V}_{k_1,k_2}(t_m)$ and we need to find an optimal policy for all state values $\mathbf{y} \in [a_1,b_1]\times[a_2,b_2]$. We divide the domain $[a_1,b_1]\times[a_2,b_2]$ into rectangular subdomains $\mathcal{C}^q$ and $\mathcal{G}^p$, so that approximately for all states $\mathbf{y}\in\mathcal{C}^q$ it is optimal to continue and for all $\mathbf{y}\in\mathcal{G}^p$ it is optimal to exercise the option. The concept is demonstrated in Figure 4.3.1 for a call-on-maximum option. The blue rectangles represent the early-exercise regions $\mathcal{G}^p$, the green ones are the continuation regions $\mathcal{C}^q$, and the dashed line shows an accurate boundary. We can split the integral in the definition of $\mathcal{V}_{k_1,k_2}$ into different parts:

$$\begin{aligned}\mathcal{V}_{k_1,k_2}(t_m) &= \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\sum_p\iint_{\mathcal{G}^p} g(\mathbf{y})\cos\left(k_1\pi\frac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\frac{y_2-a_2}{b_2-a_2}\right)d\mathbf{y}\\ &+\frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\sum_q\iint_{\mathcal{C}^q} c(t_m,\mathbf{y})\cos\left(k_1\pi\frac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\frac{y_2-a_2}{b_2-a_2}\right)d\mathbf{y}\\ &:= \sum_p \mathcal{G}_{k_1,k_2}(\mathcal{G}^p)+\sum_q \mathcal{C}_{k_1,k_2}(t_m,\mathcal{C}^q) \qquad (m\neq 0,M). \end{aligned} \quad (4.3.4)$$

We approximate the terms $\mathcal{C}_{k_1,k_2}(t_{M-1},[z_q,\,z_{q+1}]\times[w_q,\,w_{q+1}])$ in (4.3.4), where the variables z_q, z_{q+1}, w_q, and w_{q+1} denote the corner points of the rectangular continuation region $\mathcal{C}^q$. For the integrand of the terms $\mathcal{C}_{k_1,k_2}$ we again apply the 2D Fourier cosine expansion by inserting the COS formula for $c(t_{M-1},\mathbf{y})$, i.e., equation (4.3.2). The approximation reads as

$$
\begin{aligned}
&\hat{\mathcal{C}}_{k_1,k_2}(t_{M-1},[z_q,\,z_{q+1}]\times[w_q,\,w_{q+1}])\\
&\quad:=\tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\int_{w_q}^{w_{q+1}}\int_{z_q}^{z_{q+1}}\hat{c}(t_{M-1},\mathbf{y})\cos\Big(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\Big)\cos\Big(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\Big)\,dy_1\,dy_2\\
&\quad=\int_{w_q}^{w_{q+1}}\int_{z_q}^{z_{q+1}}\sum_{j_1=0}^{N_1-1}{}'\sum_{j_2=0}^{N_2-1}{}'e^{-r\Delta t}\tfrac{1}{2}\Big[\Phi^+_{j_1,j_2}(\mathbf{y})+\Phi^-_{j_1,j_2}(\mathbf{y})\Big]\mathcal{V}_{j_1,j_2}(t_M)\\
&\qquad\qquad\cdot\cos\Big(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\Big)\cos\Big(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\Big)\,dy_1\,dy_2\\
&\quad=\Re\Bigg\{\sum_{j_1=0}^{N_1-1}{}'\sum_{j_2=0}^{N_2-1}{}'\tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\Big(\tfrac{j_1\pi}{b_1-a_1},+\tfrac{j_2\pi}{b_2-a_2}\Big)\mathcal{V}_{j_1,j_2}(t_M)\\
&\qquad\qquad\cdot\mathcal{M}^+_{k_1,j_1}(z_q,z_{q+1},a_1,b_1)\mathcal{M}^+_{k_2,j_2}(w_q,w_{q+1},a_2,b_2)\Bigg\}\\
&\quad+\Re\Bigg\{\sum_{j_1=0}^{N_1-1}{}'\sum_{j_2=0}^{N_2-1}{}'\tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\Big(\tfrac{j_1\pi}{b_1-a_1},-\tfrac{j_2\pi}{b_2-a_2}\Big)\mathcal{V}_{j_1,j_2}(t_M)\\
&\qquad\qquad\cdot\mathcal{M}^+_{k_1,j_1}(z_q,z_{q+1},a_1,b_1)\mathcal{M}^-_{k_2,j_2}(w_q,w_{q+1},a_2,b_2)\Bigg\},
\end{aligned}
\tag{4.3.5}
$$

where the elements of square-matrices $\mathcal{M}^+$ and $\mathcal{M}^-$ are given by

$$
\mathcal{M}^+_{k,l}(z_1,z_2,a,b):=\tfrac{2}{b-a}\int_{z_1}^{z_2}e^{il\pi\frac{y-a}{b-a}}\cos\big(k\pi\tfrac{y-a}{b-a}\big)\,dy, \tag{4.3.6a}
$$

$$
\mathcal{M}^-_{k,l}(z_1,z_2,a,b):=\tfrac{2}{b-a}\int_{z_1}^{z_2}e^{-il\pi\frac{y-a}{b-a}}\cos\big(k\pi\tfrac{y-a}{b-a}\big)\,dy. \tag{4.3.6b}
$$

We thus find

$$
\begin{aligned}
&\hat{\mathcal{C}}_{k_1,k_2}(t_{M-1},[z_q,\,z_{q+1}]\times[w_q,\,w_{q+1}])\\
&\quad=\Re\Bigg\{\sum_{j_1=0}^{N_1-1}{}'\mathcal{M}^+_{k_1,j_1}(z_q,z_{q+1},a_1,b_1)\mathcal{K}_{j_1,k_2}(t_M,w_q,w_{q+1})\Bigg\},
\end{aligned}
\tag{4.3.7}
$$

where

$$
\begin{aligned}
\mathcal{K}_{j_1,k_2}(t_M,w_q,w_{q+1}):=&\sum_{j_2=0}^{N_2-1}{}'\tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\Big(\tfrac{j_1\pi}{b_1-a_1},+\tfrac{j_2\pi}{b_2-a_2}\Big)\mathcal{V}_{j_1,j_2}(t_M)\mathcal{M}^+_{k_2,j_2}(w_q,w_{q+1},a_2,b_2)\\
&+\sum_{j_2=0}^{N_2-1}{}'\tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\Big(\tfrac{j_1\pi}{b_1-a_1},-\tfrac{j_2\pi}{b_2-a_2}\Big)\mathcal{V}_{j_1,j_2}(t_M)\mathcal{M}^-_{k_2,j_2}(w_q,w_{q+1},a_2,b_2).
\end{aligned}
\tag{4.3.8}
$$

The elements of $(N_1\times N_2)$-matrix $\mathcal{K}$ are calculated in a rowwise fashion. The row-vector $\mathcal{K}_{j_1,\cdot}=\{\mathcal{K}_{j_1,k_2}\}_{k_2=0}^{N_2-1}$ can be written as two matrix-vector multiplications:

$$
\mathcal{K}_{j_1,\cdot}(t_M,w_q,w_{q+1})=\mathcal{M}^+(w_q,w_{q+1},a_2,b_2)\mathbf{w}^{q+}_{j_1,\cdot}+\mathcal{M}^-(w_q,w_{q+1},a_2,b_2)\mathbf{w}^{q-}_{j_1,\cdot}, \tag{4.3.9}
$$

where

$$\mathbf{w}^{q\pm}_{j_1,\cdot} := \{w^{q\pm}_{j_1,j_2}\}^{N_2-1}_{j_2=0}, \quad \text{with} \quad w^{q\pm}_{j_1,j_2} := \tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\left(\tfrac{j_1\pi}{b_1-a_1}, \pm\tfrac{j_2\pi}{b_2-a_2}\right)\mathcal{V}_{j_1,j_2}(t_M). \tag{4.3.10}$$

Then, the matrix $\hat{\mathcal{C}}_{k_1,k_2}$ is computed in a columnwise fashion. The column-vector $\hat{\mathcal{C}}_{\cdot,k_2} = \{\hat{\mathcal{C}}_{k_1,k_2}\}^{N_1-1}_{k_1=0}$ is calculated by one matrix-vector product,

$$\hat{\mathcal{C}}_{\cdot,k_2}(t_{M-1},[z_q,z_{q+1}]\times[w_q,w_{q+1}]) = \Re\left\{\mathcal{M}^{+}(z_q,z_{q+1},a_1,b_1).\mathcal{K}_{\cdot,k_2}(t_M,w_q,w_{q+1})\right\}, \tag{4.3.11}$$

with column-vector $\mathcal{K}_{\cdot,k_2} = \{\mathcal{K}_{j_1,k_2}\}^{N_1-1}_{j_1=0}$.

The coefficients $\mathcal{G}_{k_1,k_2}([z_p,z_{p+1}]\times[w_p,w_{p+1}])$ are defined by

$$\begin{aligned}&\mathcal{G}_{k_1,k_2}([z_p,z_{p+1}]\times[w_p,w_{p+1}])\\ &\quad= \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\int_{w_p}^{w_{p+1}}\int_{z_p}^{z_{p+1}} g(\mathbf{y})\cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)dy_1\,dy_2.\end{aligned} \tag{4.3.12}$$

These terms may admit an analytic solution, however, in some practical applications an analytic solution is not present. Methods to approximate these terms are proposed in Section 4.3.2.2.

We end up with the approximated coefficients

$$\hat{\mathcal{V}}_{k_1,k_2}(t_{M-1}) := \sum_p \mathcal{G}_{k_1,k_2}(\mathcal{G}^p) + \sum_q \hat{\mathcal{C}}_{k_1,k_2}(t_{M-1},\mathcal{C}^q). \tag{4.3.13}$$

For the other coefficients $\mathcal{V}_{k_1,k_2}(t_m)$, the approximations $\hat{c}(t_m,\mathbf{y})$ *and* $\hat{\mathcal{V}}_{j_1,j_2}(t_{m+1})$ will be used to approximate the terms $\mathcal{C}_{k_1,k_2}(t_m,[z_q,z_{q+1}]\times[w_q,w_{q+1}])$, and the elements of the corresponding matrix $\hat{\mathcal{K}}$ are

$$\begin{aligned}\hat{\mathcal{K}}_{j_1,k_2}(t_{m+1},w_q,w_{q+1}) = &\sum_{j_2=0}^{N_2-1}{}' \tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\left(\tfrac{j_1\pi}{b_1-a_1},+\tfrac{j_2\pi}{b_2-a_2}\right)\hat{\mathcal{V}}_{j_1,j_2}(t_{m+1})\mathcal{M}^{+}_{k_2,j_2}(w_q,w_{q+1},a_2,b_2)\\ &+\sum_{j_2=0}^{N_2-1}{}' \tfrac{1}{2}e^{-r\Delta t}\phi_{levy}\left(\tfrac{j_1\pi}{b_1-a_1},-\tfrac{j_2\pi}{b_2-a_2}\right)\hat{\mathcal{V}}_{j_1,j_2}(t_{m+1})\mathcal{M}^{-}_{k_2,j_2}(w_q,w_{q+1},a_2,b_2).\end{aligned} \tag{4.3.14}$$

Fast Fourier Transform The matrix-vector products $\mathcal{M}^{+}\mathbf{w}$ and $\mathcal{M}^{-}\mathbf{w}$ in the computation of matrices $\mathcal{K}$, $\hat{\mathcal{K}}$, and $\hat{\mathcal{C}}$ can be computed efficiently by a Fourier-based algorithm, as stated in Theorem 1 (Appendix B). The computation time achieved is $\mathcal{O}(N\log_2 N)$, with N the length of the vector.

Algorithm We can recover the terms $\hat{\mathcal{V}}_{k_1,k_2}(t_m)$ recursively, starting with $\mathcal{V}_{k_1,k_2}(t_M)$. The algorithm for solving the pricing problem backwards in time reads as follows.

Algorithm 2. *(2D-COS method for pricing Bermudan rainbow options)*
Initialization: *Calculate coefficients* $\mathcal{V}_{k_1,k_2}(t_M)$.
Main loop to recover $\hat{\mathcal{V}}(t_m)$***:***
For $m = M-1$ *to* 1*:*

- *Determine the optimal continuation regions* $\mathcal{C}^q$ *and early-exercise regions* $\mathcal{G}^p$*, as in Figure 4.3.1.*
- *Compute* $\hat{\mathcal{V}}(t_m)$ *from (4.3.4) with the help of the FFT algorithm.*

Final step: *Compute* $\hat{v}(t_0,\mathbf{x}_0)$ *by inserting* $\hat{\mathcal{V}}_{k_1,k_2}(t_1)$ *into equation (4.3.2).*

Computational Complexity The initialization is of order $\mathcal{O}(N_1N_2)$. In the main loop there are $M-1$ iterations in which the following computations are performed. The construction of one matrix $\mathcal{K}$ or $\hat{\mathcal{K}}$ costs $\mathcal{O}(2N_1N_2\log_2 N_2)$ operations. Computation of $\hat{\mathcal{C}}_{k_1,k_2}(t_m,[z_q,z_{q+1}]\times[w_q,w_{q+1}])$ takes $\mathcal{O}\left(N_2N_1\log_2 N_1\right)$ operations. $\mathcal{G}_{k_1,k_2}([z_p,z_{p+1}]\times[w_p,w_{p+1}])$ is of order $\mathcal{O}(N_1N_2)$. The computation time is linear in the number of continuation and early-exercise regions. The final step takes $\mathcal{O}(N_1N_2)$ operations.

4.3.2. APPROXIMATION METHODS FOR THE COEFFICIENTS $\mathcal{V}(T)$ AND $\mathcal{G}(\mathcal{G}^p)$

In this section, we propose methods for approximating the terminal coefficients $\mathcal{V}_{k_1,k_2,\ldots,k_n}(T)$ and the terms $\mathcal{G}_{k_1,k_2,\ldots,k_n}(\mathcal{G}^p)$ that are specific for the multidimensional-COS method.

In the 1D pricing problem, the terminal coefficients $\mathcal{V}_{k_1}(T)$ admit analytic solutions for several options, like put- and call-based options, digital options, and power options. Besides, in the 1D-COS method for pricing Bermudan options, the terms $\mathcal{G}_{k_1}(\mathcal{G}^p)$ are also usually known analytically.

In two dimensions, the payoff functions of, for instance, a geometric basket or a call-on-maximum option provide analytic solutions to the 2D coefficients $\mathcal{V}_{k_1,k_2}(T)$, but this is generally an exception. If no exact representation is available, then they can be approximated by using *discrete Fourier cosine transforms* (DCTs) or the *Clenshaw-Curtis quadrature rule*. The usage of DCTs is explained in Section 4.3.2.1. Also, analytic forms for the terms $\mathcal{G}_{k_1,k_2,\ldots,k_n}(\mathcal{G}^p)$ are in general not available in the multidimensional version. An approximation method, based on the Fourier cosine expansion of the payoff function, is discussed in Section 4.3.2.2.

4.3.2.1. DISCRETE FOURIER COSINE TRANSFORMS

In this section, we explain this idea of using DCTs to approximate the terminal coefficients $\mathcal{V}_{k_1,k_2}(T)$. For this, we take $Q \geq \max[N_1,N_2]$ grid-points for each spatial dimension and define

$$y_i^{n_i} := a_i + \left(n_i+\tfrac{1}{2}\right)\tfrac{b_i-a_i}{Q} \quad \text{and} \quad \Delta y_i := \tfrac{b_i-a_i}{Q}, \quad i=1,2. \tag{4.3.15}$$

The midpoint-rule integration gives us

$$\begin{aligned}\mathcal{V}_{k_1,k_2}(T) &\approx \sum_{n_1=0}^{Q-1}\sum_{n_2=0}^{Q-1}\tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}g(y_1^{n_1},y_2^{n_2})\cos\left(k_1\pi\tfrac{y_1^{n_1}-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2^{n_2}-a_2}{b_2-a_2}\right)\Delta y_1\Delta y_2\\ &= \sum_{n_1=0}^{Q-1}\sum_{n_2=0}^{Q-1} g(y_1^{n_1},y_2^{n_2})\cos\left(k_1\pi\tfrac{2n_1+1}{2Q}\right)\cos\left(k_2\pi\tfrac{2n_2+1}{2Q}\right)\tfrac{2}{Q}\tfrac{2}{Q}. \end{aligned} \tag{4.3.16}$$

The above 2D DCT (Type II) can be calculated efficiently by, for example, the function `dct2` of MATLAB. The approximated coefficients are denoted by $\mathcal{V}_{k_1,k_2}^{DCT}(T)$, with the corresponding computed European option value $\hat{v}^{DCT}(t_0,\mathbf{x})$. Now, an extra error is introduced:

$$\begin{aligned}\epsilon_{DCT}(t_0,\mathbf{x}) &:= \hat{v}(t_0,\mathbf{x}) - \hat{v}^{DCT}(t_0,\mathbf{x})\\ &= \tfrac{b_1-a_1}{2}\tfrac{b_2-a_2}{2}e^{-r\Delta t}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}'\Phi_{k_1,k_2}(\mathbf{x})[\mathcal{V}_{k_1,k_2}(T)-\mathcal{V}_{k_1,k_2}^{DCT}(T)].\end{aligned} \tag{4.3.17}$$

This error ϵ_{DCT} converges algebraically in Q, with order two. We will confirm this by an example in Section 4.6.1.1. The DCT method can be extended to higher dimensions for the approximation of coefficients $\mathcal{V}_{k_1,k_2,\ldots,k_n}(T)$.

Remark 4.3.1. *The above approximation, (4.3.16), is based on the midpoint-rule integration. Higher-order methods or adaptive quadrature rules may improve the efficiency.*

4.3.2.2. APPROXIMATION METHODS FOR $\mathcal{G}(\mathcal{Y}^p)$

The terms $\mathcal{G}_{k_1,k_2}$ are defined by

$$\begin{aligned}&\mathcal{G}_{k_1,k_2}([z_p,z_{p+1}]\times[w_p,w_{p+1}])\\ &\quad= \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\int_{w_p}^{w_{p+1}}\int_{z_p}^{z_{p+1}} g(\mathbf{y})\cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)dy_1\,dy_2.\end{aligned} \tag{4.3.18}$$

These terms may admit an analytic solution, however, in many practical applications the calculation of coefficients $\mathcal{G}_{k_1,k_2}(\mathcal{Y}^p)$ is time-consuming, or an analytic solution is not present. Then, we can use *discrete Fourier transforms* to approximate them, similarly as in Section 4.3.2.1. However, this may be time-consuming too, especially with a large number of timesteps. Another way is the usage of the Fourier cosine expansion of the payoff function. First, we explain this idea in one dimension.

The COS method for 1D Bermudan options was developed in [FO09]. We use 1D method parameters a_1, b_1, and N_1. For a put option it follows that

$$\mathcal{V}_{k_1}(t_m) = \mathcal{G}_{k_1}(\mathcal{Y}) + \mathcal{C}_{k_1}(t_m,\mathcal{C}). \tag{4.3.19}$$

The coefficients $\mathcal{C}_{k_1}$ are approximated by

$$\hat{\mathcal{C}}_{k_1}(t_m,z_1,z_2) = \Re\left\{\sum_{j_1=0}^{N_1-1}{}' e^{-r\Delta t}\phi_{levy}\left(\tfrac{j_1\pi}{b_1-a_1}\right)\mathcal{V}_{j_1}(t_{m+1})\mathcal{M}_{k_1,j_1}^{+}(z_1,z_2,a_1,b_1)\right\}. \tag{4.3.20}$$

In [FO09], the coefficients $\mathcal{G}_{k_1}(z_1,z_2)$ are assumed to be known analytically, and

$$\hat{\mathcal{V}}_{k_1}(t_m) := \mathcal{G}_{k_1}(\mathcal{Y}) + \hat{\mathcal{C}}_{k_1}(t_m,\mathcal{C}). \tag{4.3.21}$$

However, the coefficients $\mathcal{G}_{k_1}$ can also be approximated by a Fourier series expansion of the payoff function, i.e.,

$$g(x_1) \approx g_N(x_1) = \sum_{k_1=0}^{N_1-1}{}' \cos\left(k_1\pi\tfrac{x_1-a_1}{b_1-a_1}\right)\mathcal{V}_{k_1}(T) = \sum_{k_1=0}^{N_1-1}{}' \Re\left\{\exp\left(ik_1\pi\tfrac{x_1-a_1}{b_1-a_1}\right)\right\}\mathcal{V}_{k_1}(T), \tag{4.3.22}$$

with $\{\mathcal{V}_{k_1}(T)\}_{k_1=0}^{\infty}$ the terminal coefficients and with $v(T,y_1) = g(y_1)$. Inserting function g_N into the terms $\mathcal{G}_{k_1}$ gives

$$\begin{aligned}\hat{\mathcal{G}}_{k_1}(z_1,z_2) &:= \tfrac{2}{b_1-a_1}\int_{z_1}^{z_2} g_N(y_1)\cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)dy_1 \\ &= \Re\left\{\sum_{j_1=0}^{N_1-1}{}' \mathcal{V}_{j_1}(T)\mathcal{M}^{+}_{k_1,j_1}(z_1,z_2,a_1,b_1)\right\}.\end{aligned} \tag{4.3.23}$$

The computation of $\hat{\mathcal{C}}$ and $\hat{\mathcal{G}}$ can now be done simultaneously as

$$\begin{aligned}\hat{\mathcal{V}}_{k_1}(t_m) &:= \hat{\mathcal{G}}_{k_1}(\mathcal{G}) + \hat{\mathcal{C}}_{k_1}(t_m,\mathcal{C}) \\ &= \mathcal{V}_{k_1}(T) - \hat{\mathcal{G}}_{k_1}(\mathcal{C}) + \hat{\mathcal{C}}_{k_1}(t_m,\mathcal{C}) \\ &= \mathcal{V}_{k_1}(T) + \Re\left\{\sum_{j_1=0}^{N_1-1}{}' \left[-\mathcal{V}_{j_1}(T) + e^{-r\Delta t}\phi_{levy}\left(\tfrac{j_1\pi}{b_1-a_1}\right)\mathcal{V}_{j_1}(t_{m+1})\right]\mathcal{M}^{+}_{k_1,j_1}(\mathcal{C},a_1,b_1)\right\}.\end{aligned} \tag{4.3.24}$$

The error of this new approach converges algebraically as the Fourier series of the payoff function. The new approach is a little bit faster than (4.3.21) (with analytic $\mathcal{G}_{k_1}(\mathcal{G})$), however, a higher value of N_1 is needed to reach the same accuracy, see Figure 4.3.2 and Table 4.3.1. However, in multiple dimensions this approach may be beneficial and time-efficient.

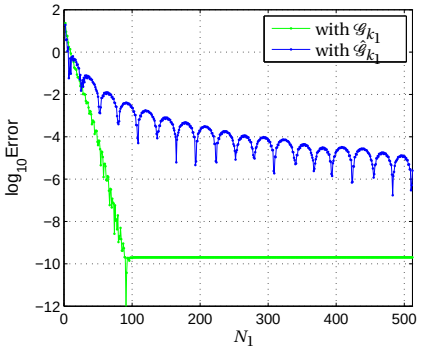


Figure 4.3.2: Comparison between method (4.3.21) and (4.3.24).

Table 4.3.1: Error, reference price is 10.479520123.

N_1	method [FO09], (4.3.21)	new method, (4.3.24)
32	-4.893e-3	-6.22e-2
64	-1.393e-6	1.14e-2
128	-2.023e-10	1.22e-3
256	-2.023e-10	-4.43e-5
512	-2.024e-10	2.52e-6
1024	-2.024e-10	-1.18e-7

Now we return to equation (4.3.18) and explain the approximation method for two di-

mensions. The Fourier cosine expansion of the payoff function can be written as

$$\begin{aligned} g_N(\mathbf{y}) &:= \sum_{k_1=0}^{N_1-1\,\prime} \sum_{k_2=0}^{N_2-1\,\prime} \cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)\mathcal{V}_{k_1,k_2}(T) \\ &= \sum_{k_1=0}^{N_1-1\,\prime} \sum_{k_2=0}^{N_2-1\,\prime} \tfrac{1}{2}\Big[\Re\left\{\exp\left(ik_1\pi\tfrac{y_1-a_1}{b_1-a_1}+ik_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)\right\} \\ &\qquad + \Re\left\{\exp\left(ik_1\pi\tfrac{y_1-a_1}{b_1-a_1}-ik_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)\right\}\Big]\mathcal{V}_{k_1,k_2}(T). \end{aligned} \tag{4.3.25}$$

With this, the coefficients $\mathcal{G}_{k_1,k_2}$ can by approximated by $\hat{\mathcal{G}}_{k_1,k_2}$, similarly as in (4.3.5):

$$\begin{aligned} &\hat{\mathcal{G}}_{k_1,k_2}([z_p,z_{p+1}]\times[w_p,w_{p+1}]) \\ &\quad = \Re\left\{\sum_{j_1=0}^{N_1-1\,\prime}\sum_{j_2=0}^{N_2-1\,\prime}\tfrac{1}{2}\mathcal{V}_{j_1,j_2}(T)\mathcal{M}^{+}_{k_1,j_1}(z_p,z_{p+1},a_1,b_1)\mathcal{M}^{+}_{k_2,j_2}(w_p,w_{p+1},a_2,b_2)\right\} \\ &\quad + \Re\left\{\sum_{j_1=0}^{N_1-1\,\prime}\sum_{j_2=0}^{N_2-1\,\prime}\tfrac{1}{2}\mathcal{V}_{j_1,j_2}(T)\mathcal{M}^{+}_{k_1,j_1}(z_p,z_{p+1},a_1,b_1)\mathcal{M}^{-}_{k_2,j_2}(w_p,w_{p+1},a_2,b_2)\right\}. \end{aligned} \tag{4.3.26}$$

We will use these approximations in Sections 4.6.2.1 and 4.6.2.2. This approximation method can be extended to higher dimensions; if the multidimensional coefficients $\mathcal{G}_{k_1,k_2,\ldots,k_n}(\mathcal{G}^p)$ are not known analytically, then we can use the terminal coefficients $\mathcal{V}_{k_1,k_2,\ldots,k_n}(T)$ to approximate them.

4.4. BERMUDAN OPTIONS UNDER THE HESTON MODEL

In this section, we explain the 2D-COS method for a Bermudan pricing problem, in which the asset price follows the Heston dynamics. This 2D stochastic process is one of the important processes in financial option pricing. In [FO11], Bermudan options under the Heston stochastic volatility model are priced using the COS formula for the log-asset dimension, combined with a quadrature rule in the log-variance dimension. For this, the closed-form density function of the transformed log-variance process is used. In our approach we employ the *bivariate characteristic function* of the log-asset price and the variance, which is also available in closed-form, and the 2D-COS formula can be applied. However, the 2D stochastic process is not in the class of Lévy processes and efficient matrix-vector multiplication using the FFT algorithm (Appendix B) within the COS method will therefore be applicable in only one dimension, similarly as in [ZGO12].

Let process $\mathbf{X}_t = (X_t^1, X_t^2) = (X_t, v_t)$ represent the Heston model. X_t represents the log-asset price process and v_t is the variance process ($v_0 \geq 0$), with dynamics

$$dX_t = (r - \tfrac{1}{2}v_t)dt + \rho\sqrt{v_t}d\omega_t^1 + \sqrt{1-\rho^2}\sqrt{v_t}d\omega_t^2, \tag{4.4.1}$$

$$dv_t = \kappa(\bar{v} - v_t)dt + \eta\sqrt{v_t}d\omega_t^1. \tag{4.4.2}$$

Here, r represents the risk free rate, $\kappa > 0$ the mean reversion rate, $\bar{v} > 0$ the long run variance, $\eta > 0$ the volatility of variance (vol-of-vol), and ρ the correlation coefficient. $\boldsymbol{\omega}_t = (\omega_t^1, \omega_t^2)$ is a 2D Wiener process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. The variance process remains strictly positive if the Feller condition, $2\kappa\bar{v} \geq \eta^2$, is satisfied, otherwise the boundary at zero is attainable and strongly reflecting [JKWW11].

The process $\mathbf{X}_t$ is affine and the *bivariate characteristic function* is of the form

$$\begin{aligned}\phi(\mathbf{u},\mathbf{X}_t,t,T) &= \mathbb{E}[e^{iu_1X_T+iu_2v_T}|X_t,v_t]\\ &= \exp\Big(B_1(T-t,\mathbf{u})X_t + B_2(T-t,\mathbf{u})v_t + A(T-t,\mathbf{u})\Big).\end{aligned} \tag{4.4.3}$$

We define

$$\beta := \kappa - i\rho\eta u_1, \tag{4.4.4a}$$

$$D := \sqrt{\beta^2+\eta^2u_1(i+u_1)}, \tag{4.4.4b}$$

$$G := (\beta - D - iu_2\eta^2)/(\beta + D - iu_2\eta^2). \tag{4.4.4c}$$

The functions A, B_1, and B_2 are solutions to a system of ordinary differential equations (ODEs) of Riccati type ([DPS00]):

$$\frac{\partial B_1(t,\mathbf{u})}{\partial t} = 0, \tag{4.4.5a}$$

$$\frac{\partial B_2(t,\mathbf{u})}{\partial t} = \tfrac{1}{2}\eta^2B_2^2(t,\mathbf{u}) - \beta B_2(t,\mathbf{u}) - \tfrac{1}{2}u_1(i+u_1), \tag{4.4.5b}$$

$$\frac{\partial A(t,\mathbf{u})}{\partial t} = iu_1r + \kappa\bar{v}B_2(t,\mathbf{u}), \tag{4.4.5c}$$

with initial conditions $B_1(0,\mathbf{u}) = iu_1$, $B_2(0,\mathbf{u}) = iu_2$, $A(0,\mathbf{u}) = 0$. Solving the ODEs gives ([Kah08])

$$B_1(t,\mathbf{u}) = iu_1, \tag{4.4.6a}$$

$$B_2(t,\mathbf{u}) = \tfrac{1}{\eta^2}\tfrac{\beta-D-(\beta+D)Ge^{-Dt}}{1-Ge^{-Dt}}, \tag{4.4.6b}$$

$$A(t,\mathbf{u}) = iu_1rt + \tfrac{\kappa\bar{v}}{\eta^2}\left[(\beta-D)t - 2\log\left(\tfrac{Ge^{-Dt}-1}{G-1}\right)\right]. \tag{4.4.6c}$$

Remark 4.4.1. *The characteristic function involves a multivalued complex logarithm in $A(t,\mathbf{u})$. Most software packages restrict the logarithm to its principal branch. Then, the characteristic function can become discontinuous in $\mathbf{u}$, which results in incorrect option prices. The same problem arises in the 2D characteristic function:*

$$\phi^{1D}(u_1,\mathbf{X}_t,t,T) = \mathbb{E}\left[e^{iu_1X_T}\middle|X_t,v_t\right]. \tag{4.4.7}$$

Solutions to the problem of choosing the correct branch here are discussed extensively in [LK06, LK10, AMST07, KJ05]. In [LK10], the function $A(t,u_1)$ ($u_2=0$) appears in the expression for the 1D characteristic function. They prove that the argument of the logarithm never crosses the negative real axis, so that the principle branch is the correct one. We have not been able to complete the proof that justifies the use of the principle branch for $A(t,\mathbf{u})$ in two dimensions. However, experiments showed that the logarithm's argument does not cross the real negative real axis for the parameter values in Section 4.7.

The variance process can be expressed as a time-changed squared Bessel process with dimension $\delta_B := \frac{4\kappa\bar{v}}{\eta^2}$, which has an absolutely continuous distribution if $\delta_B > 0$ (i.e., a density exists) and a probability mass at the origin if $\delta_B = 0$ [JYC09, Duf05]. The assumptions on the parameters yield $\delta_B > 0$, which justifies the use of the analytic density function in equation (4.2.1) and approximated coefficients Φ_{k_1,k_2} (equation (4.2.10)), with characteristic function given by equation (4.2.11).

We apply the 2D-COS method to price Bermudan options under the Heston dynamics. We again take fixed time steps $\Delta t := t_{m+1} - t_m$ and define

$$\varphi(u_1, u_2|x, v) := \phi(\mathbf{u}, \mathbf{x}, t_m, t_{m+1}) := e^{iu_1 x} e^{B_2(\Delta t, \mathbf{u})v} \phi_A(\mathbf{u}), \quad \text{where } \phi_A(\mathbf{u}) := e^{A(\Delta t, \mathbf{u})}. \tag{4.4.8}$$

The approximation of the coefficients $\mathscr{C}_{k_1,k_2}$ now reads as

$$\begin{aligned}
&\hat{\mathscr{C}}_{k_1,k_2}(t_{M-1}, [z_q, z_{q+1}] \times [w_q, w_{q+1}]) \\
&\quad = \Re\Bigg\{ {\sum_{j_1=0}^{N_1-1}}' {\sum_{j_2=0}^{N_2-1}}' \tfrac{1}{2} e^{-r\Delta t} \phi_A\left(\tfrac{j_1\pi}{b_1-a_1}, +\tfrac{j_2\pi}{b_2-a_2}\right) \mathscr{V}_{j_1,j_2}(t_M) \\
&\qquad\qquad \cdot \mathscr{M}^{+}_{k_1,j_1}(z_q, z_{q+1}, a_1, b_1) \mathscr{H}^{+}_{k_2,j_2}\left(w_q, w_{q+1}, a_2, b_2, \tfrac{j_1\pi}{b_1-a_1}\right) \Bigg\} \\
&\quad + \Re\Bigg\{ {\sum_{j_1=0}^{N_1-1}}' {\sum_{j_2=0}^{N_2-1}}' \tfrac{1}{2} e^{-r\Delta t} \phi_A\left(\tfrac{j_1\pi}{b_1-a_1}, -\tfrac{j_2\pi}{b_2-a_2}\right) \mathscr{V}_{j_1,j_2}(t_M) \\
&\qquad\qquad \cdot \mathscr{M}^{+}_{k_1,j_1}(z_q, z_{q+1}, a_1, b_1) \mathscr{H}^{-}_{k_2,j_2}\left(w_q, w_{q+1}, a_2, b_2, \tfrac{j_1\pi}{b_1-a_1}\right) \Bigg\} \\
&\quad := \Re\Bigg\{ {\sum_{j_1=0}^{N_1-1}}' \mathscr{M}^{+}_{k_1,j_1}(z_q, z_{q+1}, a_1, b_1) \mathscr{K}^{Heston}_{j_1,k_2}(t_M, w_q, w_{q+1}) \Bigg\},
\end{aligned} \tag{4.4.9}$$

where the elements of matrices $\mathscr{H}^{\pm}$ are given by

$$\begin{aligned}
&\mathscr{H}^{\pm}_{k_2,j_2}\left(w_q, w_{q+1}, a_2, b_2, \tfrac{j_1\pi}{b_1-a_1}\right) \\
&\quad := \tfrac{2}{b_2-a_2} \int_{w_q}^{w_{q+1}} e^{y_2 B_2\left(\Delta t, \frac{j_1\pi}{b_1-a_1}, \pm\frac{j_2\pi}{b_2-a_2}\right)} e^{\pm i j_2 \pi \frac{-a_2}{b_2-a_2}} \cos\left(k_2\pi \tfrac{y_2-a_2}{b_2-a_2}\right) dy_2.
\end{aligned} \tag{4.4.10}$$

As before, the elements of $(N_1 \times N_2)$-matrix $\mathscr{K}^{Heston}$ are calculated in a rowwise fashion and the row-vector $\mathscr{K}^{Heston}_{j_1,.} = \{\mathscr{K}^{Heston}_{j_1,k_2}\}_{k_2=0}^{N_2-1}$ can be written as two matrix-vector multiplications,

$$\begin{aligned}
\mathscr{K}^{Heston}_{j_1,.}(t_M, w_q, w_{q+1}) = &\mathscr{H}^{+}\left(w_q, w_{q+1}, a_2, b_2, \tfrac{j_1\pi}{b_1-a_1}\right) \mathbf{w}^{q+}_{j_1,.} \\
&+ \mathscr{H}^{-}\left(w_q, w_{q+1}, a_2, b_2, \tfrac{j_1\pi}{b_1-a_1}\right) \mathbf{w}^{q-}_{j_1,.},
\end{aligned} \tag{4.4.11}$$

where

$$\mathbf{w}^{q\pm}_{j_1,.} := \{w^{q\pm}_{j_1,j_2}\}_{j_2=0}^{N_2-1}, \quad \text{with} \quad w^{q\pm}_{j_1,j_2} := \tfrac{1}{2} e^{-r\Delta t} \phi_A\left(\tfrac{j_1\pi}{b_1-a_1}, \pm\tfrac{j_2\pi}{b_2-a_2}\right) \mathscr{V}_{j_1,j_2}(t_M). \tag{4.4.12}$$

The above equations are similar to equations (4.3.5) and (4.3.8) for pricing rainbow options. However, the matrices $\mathscr{H}^{\pm}$ cannot be decomposed into a Hankel and Toeplitz matrix anymore and the computation of $\mathscr{K}^{Heston}_{j_1,.}$ is therefore of order $\mathscr{O}(N_2^2)$. For the numerical tests in Section 4.7 we will take a fixed grid $\{w_q\}$. Then, the matrices $\mathscr{H}^{\pm}$ need to be computed only once and this computation is part of the algorithm's initialization.

4.5. ERROR CONVERGENCE AND COMPUTATIONAL DOMAIN

The method parameters of the 2D-COS method are the integration ranges, $[a_1, b_1]$ and $[a_2, b_2]$, and the number of Fourier cosine terms, N_1 and N_2. Convergence in the number of cosine terms is discussed in the error analysis in Section 4.5.1. Section 4.5.2 gives some suggestions for the choice of the computational domain.

4.5.1. ERROR ANALYSIS

The error analysis is similar to the analyses in Sections 2.4.1 and 2.4.3. We start with the local error of the COS formula. Then, a result for the propagating error is presented.

The local error of the COS formula is defined by

$$\epsilon(t_{m-1},\mathbf{x}) := c(t_{m-1},\mathbf{x}) - \hat{c}(t_{m-1},\mathbf{x}). \tag{4.5.1}$$

Let us assume that the terms $\mathcal{V}_{k_1,k_2}(t_m)$ are known analytically. Then, errors are introduced in three steps (see Section 4.2): the truncation of the integration range, the substitution of the density by its cosine series expansion on the truncated range, and the substitution of the series coefficients by the characteristic function approximation. The key to bounding the error lies in the decay rate of the Fourier cosine series coefficients. We discuss the three errors one after the other.

Step 1: The integration range truncation error:

$$\epsilon_1(t_{m-1},\mathbf{x}) := c(t_{m-1},\mathbf{x}) - c_1(t_{m-1},\mathbf{x}) = e^{-r\Delta t} \iint_{\mathbb{R}^2\setminus[a_1,b_1]\times[a_2,b_2]} v(t_m,\mathbf{y})\, p(\mathbf{y}|\mathbf{x})\, d\mathbf{y}. \tag{4.5.2}$$

Step 2: The series truncation error:

$$\begin{aligned}\epsilon_2(t_{m-1},\mathbf{x}) &:= c_1(t_{m-1},\mathbf{x}) - c_2(t_{m-1},\mathbf{x})\\ &= \tfrac{b_1-a_1}{2}\tfrac{b_2-a_2}{2} e^{-r\Delta t} \sum_{\substack{k_1,k_2\\ k_1\geq N_1 \text{ or } k_2\geq N_2}}^{\infty} \mathscr{P}_{k_1,k_2}(\mathbf{x})\mathcal{V}_{k_1,k_2}(t_m)\\ &= e^{-r\Delta t}\int_{a_2}^{b_2}\int_{a_1}^{b_1} v(t_m,\mathbf{y})\left[p(\mathbf{y}|\mathbf{x}) - p_N(\mathbf{y}|\mathbf{x})\right] d\mathbf{y}.\end{aligned} \tag{4.5.3}$$

The functions $v_N(t_m,\mathbf{y})$ and $p_N(\mathbf{y}|\mathbf{x})$ denote the Fourier cosine series expansions of the value function and the density function, using N_1 and N_2 terms in the series summations.

The convergence rate of Fourier cosine series depends on the properties of the approximated functions in the expansion interval. The coefficients $\mathscr{P}_{k_1,k_2}$ usually decay faster than $\mathcal{V}_{k_1,k_2}$. With [Boy01], we find that the error converges exponentially in N_1 and N_2 for density functions in the class $C^\infty([a_1, b_1]\times[a_2, b_2])$. A density function with discontinuity in one of its derivatives results in an algebraic convergence of the Fourier cosine series expansion.

Step 3: The error related to approximating $\mathscr{P}_{k_1,k_2}(\mathbf{x})$ by $\Phi_{k_1,k_2}(\mathbf{x})$ (see equation (4.2.6)):

$$\begin{aligned}
\epsilon_3(t_{m-1},\mathbf{x}) &:= c_2(t_{m-1},\mathbf{x}) - \hat{c}(t_{m-1},\mathbf{x}) \\
&= \tfrac{b_1-a_1}{2}\tfrac{b_2-a_2}{2} e^{-r\Delta t} \sum_{k_1=0}^{N_1-1\prime} \sum_{k_2=0}^{N_2-1\prime} (\mathscr{P}_{k_1,k_2}(\mathbf{x}) - \Phi_{k_1,k_2}(\mathbf{x})) \mathcal{V}_{k_1,k_2}(t_m) \\
&= -e^{-r\Delta t} \iint_{\mathbb{R}^2\setminus[a_1,b_1]\times[a_2,b_2]} \left[\sum_{k_1=0}^{N_1-1\prime} \sum_{k_2=0}^{N_2-1\prime} \cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right) \cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right) \mathcal{V}_{k_1,k_2}(t_m) \right] p(\mathbf{y}|\mathbf{x})\, d\mathbf{y} \\
&= -e^{-r\Delta t} \iint_{\mathbb{R}^2\setminus[a_1,b_1]\times[a_2,b_2]} v_N(t_m,\mathbf{y})\, p(\mathbf{y}|\mathbf{x})\, d\mathbf{y}. \qquad (4.5.4)
\end{aligned}$$

Addition of the three errors gives

$$\begin{aligned}
\epsilon(t_{m-1},\mathbf{x}) = e^{-r\Delta t} &\iint_{\mathbb{R}^2\setminus[a_1,b_1]\times[a_2,b_2]} \left[v(t_m,\mathbf{y}) - v_N(t_m,\mathbf{y}) \right] p(\mathbf{y}|\mathbf{x})\, d\mathbf{y} \\
&+ e^{-r\Delta t} \iint_{[a_1,b_1]\times[a_2,b_2]} v(t_m,\mathbf{y}) \left[p(\mathbf{y}|\mathbf{x}) - p_N(\mathbf{y}|\mathbf{x}) \right] d\mathbf{y}. \qquad (4.5.5)
\end{aligned}$$

Result 4.5.1. *Let us assume that the terms $\mathcal{V}_{k_1,k_2}(t_m)$ are exact. If the integration domain $[a_1, b_1]\times[a_2, b_2]$ is chosen sufficiently wide, then the series truncation error ϵ_2 dominates the overall local error. Then, for smooth density functions the error ϵ converges exponentially in N_1 and N_2, otherwise it converges algebraically. The extra error ϵ_{DCT}, introduced by approximation of the terms $\mathcal{V}_{k_1,k_2}(T)$ with DCTs, has been discussed in Section 4.3.2.1.*

We now discuss the error in the terms $\mathcal{V}_{k_1,k_2}(t_m)$ and define

$$\varepsilon_{k_1,k_2}(t_m,\mathcal{C}^q) := \mathscr{C}_{k_1,k_2}(t_m,\mathcal{C}^q) - \hat{\mathscr{C}}_{k_1,k_2}(t_m,\mathcal{C}^q). \qquad (4.5.6)$$

The terms $\mathscr{G}_{k_1,k_2}(\mathcal{G}^p)$ are assumed to be exact, so that the error in the Fourier coefficients is given by

$$\varepsilon_{k_1,k_2}(t_m) := \mathcal{V}_{k_1,k_2}(t_m) - \hat{\mathcal{V}}_{k_1,k_2}(t_m) = \sum_q \varepsilon_{k_1,k_2}(t_m,\mathcal{C}^q). \qquad (4.5.7)$$

Result 4.5.2. *With $[a_1, b_1]\times[a_2, b_2] \subset \mathbb{R}^2$ chosen sufficiently wide and a probability density function p in $C^\infty([a_1, b_1]\times[a_2, b_2])$, error $\varepsilon_{k_1,k_2}(t_m)$ converges exponentially in N_1 and N_2 for $1 \le m \le M-1$. If the local error converges algebraically, then so does error $\varepsilon_{k_1,k_2}(t_m)$.*

The proof of this result is similar to that for pricing Bermudan options with one underlying asset, which can be found in [FO09]. The convergence is algebraic for non-smooth density functions or if we approximate $\mathscr{G}_{k_1,k_2}(\mathcal{G}^p)$ by an algebraically converging method, see Section 4.3.2.2.

4.5.2. COMPUTATIONAL DOMAIN

The performance of the (2D-)COS method is sensitive to the choice of the computational domain $[a_1, b_1]\times[a_2, b_2]$. If the domain size is set too small, then the resulting option values may be too low. However, the larger the domain, the more terms in the series expansions that are required to reach a certain accuracy.

For the tests on rainbow options, we will take $a_1 = a_2 = a$ and $b_1 = b_2 = b$, where

$$a := \min_i \left[\kappa_1^i - L\sqrt{\kappa_2^i + \sqrt{\kappa_4^i}} \right], \quad b := \max_i \left[\kappa_1^i + L\sqrt{\kappa_2^i + \sqrt{\kappa_4^i}} \right], \qquad L = 10. \qquad (4.5.8)$$

κ_j^i denotes the jth cumulant of the stochastic variable X_T^i. For the cumulants κ_1, κ_2, and κ_4 of the Brownian motion and the log-jump-diffusion process we refer the reader to [FO08]. The choice of equal domain sizes for both dimensions facilitates the calculation of the coefficients $\mathcal{V}_{k_1,k_2}(T)$.

For the Heston dynamics, we base the interval $[a_1, b_1]$ on the cumulants κ_1 and κ_2 of X_T:

$$[a_1, b_1] := \left[\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}\right], \qquad L = 8. \tag{4.5.9}$$

The second cumulant can be approximated by $\kappa_2 \approx \bar{\nu}(1+\eta)T$. For $a_2 \geq 0$ and b_2, however, we take a tolerance level $\text{TOL} = 10^{-4}$ and determine the integration range so that

$$F_{v_T|v_0}(a_2|v_0) = \text{TOL} = 1 - F_{v_T|v_0}(b_2|v_0), \tag{4.5.10}$$

where $F_{v_T|v_0}$ represents the cumulative distribution function of the variance at the terminal time. This is a general strategy in the case of non-smooth densities.

If N_1 and N_2 are chosen sufficiently large, then a larger size of the computational domain should not affect the option price.

4.6. NUMERICAL EXPERIMENTS RAINBOW OPTIONS

In this section we perform several numerical experiments to test the 2D-COS method for pricing European and Bermudan rainbow options. We will validate the algorithm and show its efficiency. The following options are studied: geometric basket, arithmetic basket, put-on-minimum, and call-on-maximum options. The asset price paths are modeled by either correlated geometric Brownian motions (GBMs) or by Merton's jump-diffusion processes (but in principle we can use other Lévy processes). MATLAB 7.11.0 is used for the computations, with an Intel(R) Quad-Core 2.83 GHz and 8 GB RAM. For the tests in this section, we take an equal number of terms in both series expansions, that is, $N_2 = N_1$.

Geometric Brownian Motion Under GBM the risk-neutral asset prices evolve according to the following dynamics:

$$dS_t^i = (r - \delta_i)S_t^i dt + S_t^i \bar{\sigma}_i d\tilde{\omega}_t^i, \qquad i = 1,2, \tag{4.6.1}$$

where $\tilde{\boldsymbol{\omega}}_t = (\tilde{\omega}_t^1, \tilde{\omega}_t^2)$ is a 2D correlated Wiener process with correlation $d\tilde{\omega}_t^i d\tilde{\omega}_t^j = \rho_{ij} dt$, r the risk-free rate, δ_i the dividend rate, and $\bar{\sigma}_i$ the volatility of asset i. We switch to the log-processes $X_t^i := \log S_t^i$:

$$dX_t^i = (r - \delta_i - \tfrac{1}{2}\bar{\sigma}_i^2)dt + \bar{\sigma}_i d\tilde{\omega}_t^i. \tag{4.6.2}$$

The log-asset prices at time t_m, given the values at time t_{m-1}, are bivariate normally distributed,

$$\mathbf{X}_{t_m} \sim \mathcal{N}(\mathbf{X}_{t_{m-1}} + \boldsymbol{\mu}\Delta t, \boldsymbol{\Sigma}), \tag{4.6.3}$$

with $\mu_i = r - \delta_i - \frac{1}{2}\bar{\sigma}_i^2$ and covariance matrix $\boldsymbol{\Sigma}_{ij} = \bar{\sigma}_i \bar{\sigma}_j \rho_{ij} \Delta t$. The characteristic function reads as $\varphi(\mathbf{u}|\mathbf{x}) = e^{i\mathbf{x}'\mathbf{u}}\phi_{levy}(\mathbf{u})$, with

$$\phi_{levy}(\mathbf{u}) = \exp(i\boldsymbol{\mu}'\Delta t\mathbf{u} - \tfrac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}). \tag{4.6.4}$$

Jump-Diffusion Process Under jump-diffusion the asset prices follow

$$dS_t^i = (r - \lambda\kappa_i)S_t^i dt + S_t^i \bar{\sigma}_i d\tilde{\omega}_t^i + (e^{J_i} - 1)S_t^i dq_t, \qquad i = 1,2, \tag{4.6.5}$$

with $\kappa_i := \mathbb{E}[e^{J_i} - 1]$, q_t a Poisson process with intensity rate λ, and $\mathbf{J} = (J_1, J_2)$ bivariate normally distributed jumps, with mean $\boldsymbol{\mu}^J = [\mu_1^J \ \mu_2^J]'$ and covariance matrix $\boldsymbol{\Sigma}_{ij}^J = \sigma_i^J \sigma_j^J \rho_{ij}^J$. The log-processes $X_t^i := \log S_t^i$ read as

$$dX_t^i = (r - \lambda\kappa_i - \tfrac{1}{2}\bar{\sigma}_i^2)dt + \bar{\sigma}_i d\tilde{\omega}_t^i + J_i dq_t. \tag{4.6.6}$$

The characteristic function reads as $\varphi(\mathbf{u}|\mathbf{x}) = e^{i\mathbf{x}'\mathbf{u}}\phi_{levy}(\mathbf{u})$, with

$$\phi_{levy}(\mathbf{u}) = \exp\left(i\boldsymbol{\mu}'\Delta t\mathbf{u} - \tfrac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}\right)\exp\left(\lambda\Delta t(\exp(i\boldsymbol{\mu}^{J\prime}\mathbf{u} - \tfrac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}^J\mathbf{u}) - 1)\right), \tag{4.6.7}$$

where $\mu_i = r - \lambda\kappa_i - \frac{1}{2}\bar{\sigma}_i^2$ and $\boldsymbol{\Sigma}_{ij} = \bar{\sigma}_i\bar{\sigma}_j\rho_{ij}\Delta t$.

The bivariate density functions of the correlated Brownian motion and jump-diffusion process are both in C^∞, which will result in exponentially converging errors in N_1 and N_2. *Density recovery* using Fourier cosine expansions and the characteristic function gives

$$\hat{p}(\mathbf{y}|\mathbf{x}) := \sum_{k_1=0}^{N_1-1\prime}\sum_{k_2=0}^{N_2-1\prime} \Phi_{k_1,k_2}(\mathbf{x})\cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right). \tag{4.6.8}$$

The approximated density function of a log-jump-diffusion process is presented in Figure 4.6.1.

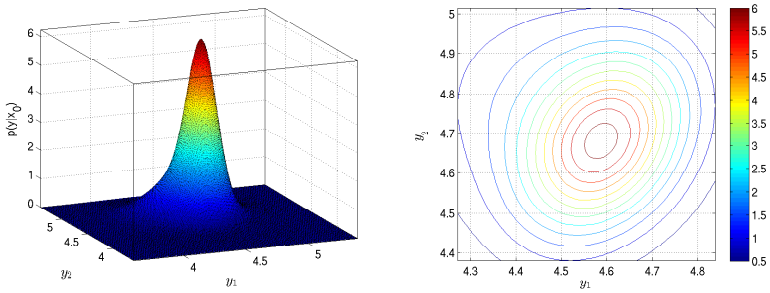


Figure 4.6.1: Density recovery jump-diffusion process $\hat{p}(\mathbf{X}_T|\mathbf{x}_0)$ (Parameter set II).

Parameter Sets The following five parameters sets are used for the rainbow options ($t_0 = 0$).

Set I: $\mathbf{S}_0 = [90 \quad 110]'$, $r = 0.04$, $\delta_i = 0$, $\bar{\sigma} = [0.2 \quad 0.3]'$, $\rho = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix}$, $T = 1$, $K = 100$.

Set II: $\mathbf{S}_0 = [90 \quad 110]'$, $r = 0.05$, $\bar{\sigma} = [0.12 \quad 0.15]'$, $\rho = \begin{bmatrix} 1 & 0.30 \\ 0.30 & 1 \end{bmatrix}$, $\lambda = 0.60$,

$\boldsymbol{\mu}^J = [-0.10 \quad 0.10]'$, $\sigma^J = [0.17 \quad 0.13]'$, $\rho^J = \begin{bmatrix} 1 & -0.20 \\ -0.20 & 1 \end{bmatrix}$, $T = 1$, $K = 100$.

Set III: $\mathbf{S}_0 = [40 \quad 40]'$, $r = 0.048790$, $\delta_i = 0$, $\bar{\sigma} = [0.2 \quad 0.3]'$, $\rho = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$, $T = 7/12$, $K = 40$.

Set IV: $\mathbf{S}_0 = [100 \quad 100 \quad 100]'$, $r = 0.04$, $\delta_i = 0$, $\bar{\sigma} = [0.3 \quad 0.35 \quad 0.4]'$,

$\rho = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$, $T = 1$, $K = 100$.

Set V: $r = 0.05$, $\delta_i = 0.1$, $\bar{\sigma} = [0.2 \quad 0.2]'$, $\rho = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $T = 3$, $K = 100$.

4.6.1. EUROPEAN RAINBOW OPTIONS

We start in Sections 4.6.1.1 and 4.6.1.2 by pricing two types of basket options with two underlying asset prices, namely,

- geometric basket call options, with payoff $g(\mathbf{x}) = \left(\sqrt{e^{x_1}}\sqrt{e^{x_2}} - K\right)^+$,
- arithmetic basket call options, with payoff $g(\mathbf{x}) = \left(\frac{1}{2}e^{x_1} + \frac{1}{2}e^{x_2} - K\right)^+$.

Then, Section 4.6.1.3 discusses call-on-maximum and put-on-minimum options.

4

4.6.1.1. GEOMETRIC BASKET CALL OPTION

The price of a geometric basket call option under GBM equals the price of a Black-Scholes call option with initial option price $\hat{S}_0 = \sqrt{S_0^1}\sqrt{S_0^2}$, dividend rate $\hat{\delta}$, and volatility $\hat{\sigma}$ ([LO08a]), where

$$\hat{\sigma} = \sqrt{(\tfrac{1}{2})^2 \textstyle\sum_{i,j} \bar{\sigma}_i \bar{\sigma}_j \rho_{ij}} \qquad \text{and} \qquad \hat{\delta} = \tfrac{1}{2}\textstyle\sum_i (\delta_i + \frac{1}{2}\bar{\sigma}_i^2) - \frac{1}{2}\hat{\sigma}^2. \tag{4.6.9}$$

So, we can compare our results with the analytic option values. The Fourier cosine coefficients of the payoff function are given by

$$\mathcal{V}_{k_1,k_2}(T) = \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1} (\sqrt{e^{y_1}}\sqrt{e^{y_2}} - K)^+ \cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right) dy_1\, dy_2. \tag{4.6.10}$$

An analytic solution is available and can be found using, for instance, Maple 14. We use parameter set I. Besides, we test the method for a deep out-of-the-money option with strike price $K = 200$. The option values are $v(t_0, \mathbf{x}_0) = 8.8808$ ($K = 100$) and $v(t_0, \mathbf{x}_0) = 2.8 \times 10^{-3}$ ($K = 200$). The results in Table 4.6.1 and Figure 4.6.2 are highly satisfactory and show exponential convergence. Convergence is reached in milliseconds. For both strike prices, the Fourier coefficients $\mathcal{V}_{k_1,k_2}$ are of the same order of magnitude. The last coefficients retained in the series expansion truncation determine the order of magnitude of the truncation error. Because of that, the error results are similar, however, the relative error is larger for the out-of-the-money option.

Table 4.6.1: Results geometric basket call (GBM).

	$N_1 (= N_2)$			
	10	20	40	80
Error ($K = 100$)	-7.60e-1	-4.07e-2	-1.42e-5	2.34e-13
Error ($K = 200$)	-7.23e-1	-4.06e-2	-1.42e-5	3.08e-13
CPU time (ms)	1.65	1.99	3.15	7.46

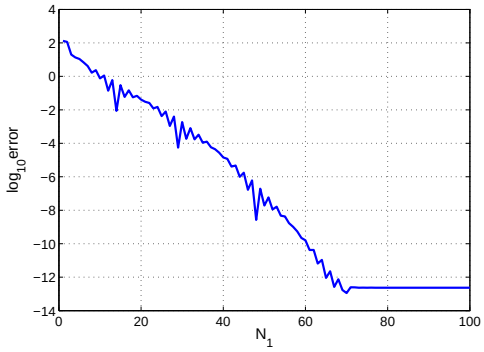


Figure 4.6.2: Error geometric basket call ($K = 100$) (GBM).

In Section 4.6.1.2, we will price arithmetic basket options, for which the terminal coefficients $\mathcal{V}_{k_1,k_2}(T)$ need to be approximated by means of 2D discrete Fourier cosine transforms, as explained in Section 4.3.2.1. Since there is an analytic solution available for the *geometric* basket options, we can analyze the error of the discretization approach for this option. For each spatial dimension we now take $Q \geq N_1$ grid-points. So we replace the payoff coefficients by a discrete approximation.

First, we analyze the approximate coefficients $\mathcal{V}^{DCT}_{k_1,k_2}(T)$ for the geometric basket option and calculate the maximum absolute error. This error converges quadratically in Q, see Figure 4.6.3. Second, we calculate the errors of the COS formula, see Section 4.5.1. The computational domain is chosen sufficiently large, so that we can neglect errors ϵ_1 and ϵ_3, in (4.5.2) and (4.5.4), respectively. Results are shown in Table 4.6.2 and Figure 4.6.4. The left-side plot shows the series truncation error ϵ_2, which indeed converges exponentially in N_1. The midplot confirms the algebraically converging error ϵ_{DCT}, and the right-side plot shows the total error ϵ. If N_1 is chosen sufficiently large, then the error of using discrete Fourier transforms dominates the total error. The computation time of the coefficients is of order $\mathcal{O}(Q^2 \log_2 Q)$ and is the most time-consuming part. The CPU time is, however, still less than one second.

4.6.1.2. ARITHMETIC BASKET CALL OPTION

For arithmetic basket options under geometric Brownian motion, there is no analytic solution to the option price. Instead, we use the reference option value $v(t_0, \mathbf{x}_0) \approx 10.173230$, obtained by using $Q = 5000$, $N_1 = N_2 = 100$ (Set I). This value is validated by a plain Monte Carlo simulation with 10^6 runs, which results in $v(t_0, \mathbf{x}_0) \approx 10.1714$ with standard deviation 0.017.

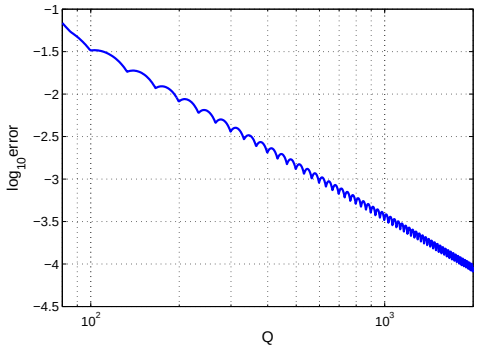


Figure 4.6.3: Maximum absolute error $\mathcal{V}_{k_1,k_2}^{DCT}(T)$ $(k_1,k_2 = 1,\ldots,80)$.

Table 4.6.2: Results geometric basket call, with payoff coefficients approximated by DCTs.

(a) Absolute error

Q	$N_1(=N_2)$			
	10	20	40	80
250	1.09e+1	7.61e-1	4.18e-2	1.14e-3
500	7.59e-1	4.02e-2	3.91e-4	4.05e-4
1000	7.60e-1	4.06e-2	4.96e-5	6.38e-5
2000	7.60e-1	4.07e-2	1.36e-5	5.75e-7

(b) CPU time (s)

Q	$N_1(=N_2)$			
	10	20	40	80
250	0.007	0.007	0.008	0.014
500	0.04	0.04	0.04	0.06
1000	0.17	0.17	0.17	0.20
2000	0.71	0.71	0.73	0.75

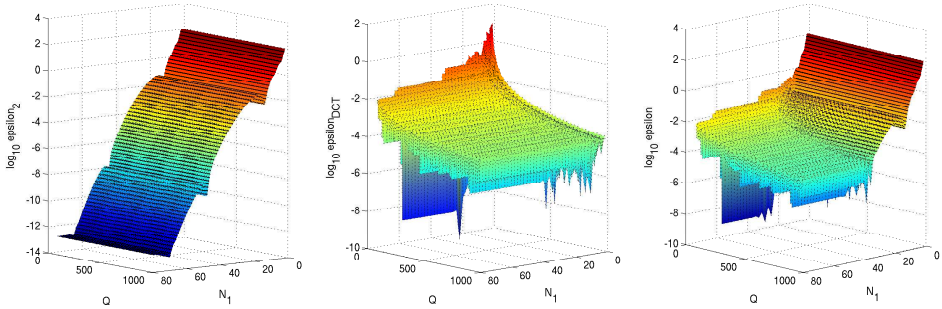


Figure 4.6.4: Error geometric basket call ($K = 100$), with payoff coefficients approximated by DCTs.

The Fourier cosine coefficients of the payoff function are given by

$$\mathcal{V}_{k_1,k_2}(T) = \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1}(\tfrac{1}{2}e^{y_1}+\tfrac{1}{2}e^{y_2}-K)^+\cos\left(k_1\pi\tfrac{y_1-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{y_2-a_2}{b_2-a_2}\right)dy_1\,dy_2. \tag{4.6.11}$$

No analytic representation is available and we approximate the coefficients by using discrete Fourier cosine transforms. The results in Table 4.6.3 are satisfactory and the error converges exponentially in N_1. The CPU times are the same as in Table 4.6.2 and conver-

gence is achieved within two seconds.

Table 4.6.3: Absolute error arithmetic basket call (GBM).

Q	$N_1 (= N_2)$			
	10	20	40	80
250	2.97e0	1.85e-1	2.65e-3	2.70e-3
500	2.98e0	1.91e-1	1.97e-5	2.97e-5
1000	2.98e0	1.91e-1	4.82e-5	1.29e-6
2000	2.98e0	1.91e-1	5.08e-5	1.28e-6

4.6.1.3. CALL-ON-MAX AND PUT-ON-MIN OPTIONS

In this section, we discuss the following two-color rainbow options:

- call-on-maximum option, with payoff $g(\mathbf{x}) = (\max(e^{x_1}, e^{x_2}) - K)^+$,
- put-on-minimum option, with payoff $g(\mathbf{x}) = (K - \min(e^{x_1}, e^{x_2}))^+$.

The Fourier coefficients of the payoff functions can be calculated analytically. For the European call-on-max and put-on-min options under GBM we use parameters set III ([Boy88]) and compare the results with the analytic solutions from [Stu82]. Figure 4.6.5 shows the results and again confirms exponential convergence.

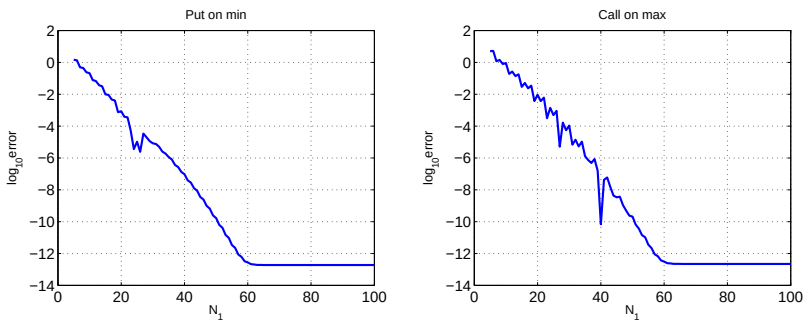


Figure 4.6.5: Error put-on-min and call-on-max option (GBM).

Next we consider a put-on-minimum option under a jump-diffusion process. Under jump-diffusion asset price processes, there is no analytic solution to the option values available. We use the model parameters from set II and set $N_1 = N_2 = 125$, for which machine precision is reached. The results in Table 4.6.4 are achieved in a few milliseconds and correspond well to the reference prices in [CF08].

4.6.2. BERMUDAN RAINBOW OPTIONS

For pricing Bermudan options we need to determine rectangular continuation and early-exercise regions. For this, we divide the domain of the second dimension, $[a_2, b_2]$, into J subintervals:

$$[a_2, b_2] = [w_0, w_1] \cup [w_1, w_2] \ldots [w_q, w_{q+1}] \ldots [w_{J-1}, w_J]. \tag{4.6.12}$$

Table 4.6.4: Put-on-min option values $\hat{v}(t_0, \mathbf{x}_0)$ (jump-diffusion).

S_0^2	S_0^1		
	90	100	110
90	15.6916	13.4073	12.1305
100	12.1918	9.1360	7.5175
110	10.3853	6.7274	4.8337

At the center of each subinterval, we determine the value(s) y^* for which the optimal exercise policy changes to the optimal continuation, i.e.,

$$g(y^*, \tfrac{1}{2}(w_q + w_{q+1})) = c(t_m, y^*, \tfrac{1}{2}(w_q + w_{q+1})). \tag{4.6.13}$$

For, for instance, a basket put option, we can then define an early-exercise region $\mathcal{G}^q = [a_1, y^*] \times [w_q, w_{q+1}]$ and a continuation region $\mathcal{C}^q = [y^*, b_1] \times [w_q, w_{q+1}]$. Therefore, the computational domain is divided into J early-exercise and continuation regions. In the case of a call-on-max option there is a multiply connected exercise region, see Figure 4.3.1, and for each subinterval $[w_q, w_{q+1}]$ we have one or two early-exercise regions $\mathcal{G}^q$. As long as we select a technique by which we can distinguish multiple continuation and early-exercise regions from each other, this does not affect the complexity of the method.

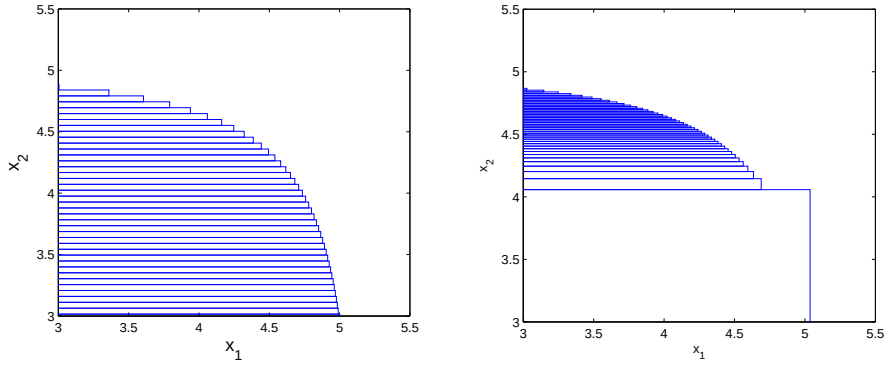
Remark 4.6.1. *We elaborate on the error in the determination of the early-exercise regions and the convergence in parameter J. We assume that N_1 and N_2 are chosen sufficiently large, so that the approximated option values are sufficiently accurate. Besides, we presume that a root-finding method is available which is able to accurately determine the value(s) y^*, see (4.6.13). For example, the Newton method with an accurate initial guess suffices. In this setting, the early-exercise regions converge, quadratically in J, to the true regions by means of the rectangle-rule-based method. However, if we would choose N_1, N_2 too small, the COS formula might give inaccurate, oscillatory options values, resulting in inaccurate early-exercise regions and the convergence result holding only for sufficiently large N_1, N_2.*

For *equidistant* intervals $[w_q, w_{q+1}]$ we take $w_q = a_2 + (b_2 - a_2)q/J$, which gives the exercise regions $\mathcal{G}^q$ shown in Figure 4.6.6a. A *nonequidistant* grid, as in Figure 4.6.6b, is based on the quantile function (the inverse distribution function) of asset price X_T^2. We use $w_0 = a_2$, $w_J = b_2$, and $w_q = F^{-1}_{X_T^2}(q/J)$. The large rectangle at the bottom of the domain is not an accurate representation of the true region. However, it is located near the boundary of the computational domain, outside the domain of interest, and therefore does not affect the option values in the middle of the domain. The usage of nonequidistant grids typically leads to more efficient pricing.

Remark 4.6.2. *If the quantile function is not known analytically, then we can approximate the random variable by a normally distributed variable with the same mean (first cumulant) and variance (second cumulant), for which the quantile function is known. We will use this approach for the jump-diffusion process in Section 4.6.2.3.*

4.6.2.1. GEOMETRIC BASKET PUT OPTION

We start by pricing Bermudan-style geometric basket put options under geometric Brownian motion. For the tests we use parameter Set I and take $M = 10$ early-exercise dates. The



(a) Equidistant grid. (b) Nonequidistant grid.

Figure 4.6.6: Rectangular early-exercise regions $\mathcal{G}^q$ (arithmetic basket put).

reference solution to the put option, $v(t_0,\mathbf{x}) = 6.95904$, equals a Bermudan put option in one dimension ([BS07]) and is computed with the 1D-COS method. The terms $\mathcal{G}_{k_1,k_2}(\mathcal{G}^p)$ admit an analytic solution. The results are presented in Table 4.6.5. Convergence is exponential in N_1 and algebraical in J, with order two, see Section 4.5.1. The use of nonequidistant grids improves the convergence significantly. The computation time is linear in J and the results for, for example, $N_1 = 80$ and $J = 160$ are obtained in about 30 seconds.

Table 4.6.5: Absolute error Bermudan geometric basket put.

(a) Equidistant grid.

J	$N_1 (= N_2)$		
	40	80	160
20	6.65e-2	7.77e-2	7.79e-2
40	1.22e-2	1.56e-2	1.56e-2
80	1.69e-4	3.54e-3	3.51e-3
160	2.51e-3	8.86e-4	8.56e-4
320	3.15e-3	2.42e-4	2.13e-4

(b) Nonequidistant grid.

J	$N_1 (= N_2)$		
	40	80	160
20	7.07e-4	2.23e-3	2.20e-3
40	2.63e-3	5.13e-4	4.83e-4
80	2.97e-3	1.46e-4	1.16e-4
160	3.17e-3	5.85e-5	2.88e-5
320	3.18e-3	3.70e-5	7.17e-6

In Section 4.6.2.2, we will price Bermudan arithmetic basket options, for which the coefficients $\mathcal{V}_{k_1,k_2}(T)$ are approximated by means of 2D discrete Fourier cosine transforms and the terms $\mathcal{G}_{k_1,k_2}$ are approximated by $\hat{\mathcal{G}}_{k_1,k_2}$, as explained in Section 4.3.2.2 (with equation (4.3.25)). We examine here the error of this approach by using the Bermudan *geometric* basket, for which an accurate reference price is available. We take $Q = 2000$. Table 4.6.6 shows satisfactorily results. The results are similar to those of Table 4.6.5, with analytic terms $\mathcal{G}_{k_1,k_2}$. However, for higher values of N_1 the usage of analytic terms is more accurate.

4.6.2.2. ARITHMETIC BASKET PUT OPTION

For the arithmetic basket option, we take the same model parameters (Set I) as in the previous section, with $M = 10$ early-exercise dates. We approximate the Fourier coefficients at

Table 4.6.6: Absolute error Bermudan geometric basket put, with $\hat{\mathcal{G}}_{k_1,k_2}$ ($Q = 2000$).

(a) Equidistant grid.

J	$N_1(=N_2)$		
	40	80	160
20	7.05e-02	7.78e-02	7.79e-02
40	1.70e-02	1.58e-02	1.56e-02
80	5.56e-03	3.69e-03	3.54e-03
160	2.96e-03	1.09e-03	9.13e-04
320	2.33e-03	4.66e-04	2.86e-04

(b) Nonequidistant grid.

J	$N_1(=N_2)$		
	40	80	160
20	4.55e-03	2.45e-03	2.25e-03
40	2.69e-03	7.39e-04	5.54e-04
80	2.37e-03	3.74e-04	1.92e-04
160	2.23e-03	2.87e-04	1.06e-04
320	2.21e-03	2.66e-04	8.52e-05

the terminal time by using the discrete Fourier cosine transform and the coefficients $\mathcal{G}_{k_1,k_2}$ are approximated by $\hat{\mathcal{G}}_{k_1,k_2}$. The results in Table 4.6.7 show converging option values.

Table 4.6.7: Bermudan arithmetic basket put option values $\hat{v}(t_0,\mathbf{x}_0)$, with $\hat{\mathcal{G}}_{k_1,k_2}$ ($Q = 2000$).

	Nonequidistant grid		
J	$N_1(=N_2)$		
	40	80	160
20	6.6086	6.6067	6.6077
40	6.6102	6.6091	6.6102
80	6.6106	6.6096	6.6108
160	6.6109	6.6097	6.6109
320	6.6108	6.6097	6.6109

4.6.2.3. CALL-ON-MAX AND PUT-ON-MIN OPTIONS

Here, we consider Bermudan call-on-maximum and put-on-minimum options. They are discussed in, among others, [AB04, CF08, JO12, HG08]. The authors in [BD97] prove properties of the continuation and early-exercise regions for the American-style option. For the 2D-COS method we apply a nonequidistant grid and use Remark 4.6.2 for the jump-diffusion process.

The convergence results for a call-on-max option under GBM, Set V with $M = 9$, are presented in Table 4.6.8a. In Table 4.6.8b, option prices for different initial asset values are shown. They correspond to the values in the third and fourth columns, which are obtained in [AB04].

At last, we price put-on-minimum options under a jump-diffusion process (parameters set II). The results in Table 4.6.9a show converging values. This verifies the applicability of the 2D-COS method to the broad class of Lévy processes. With increasing the number of early-exercise dates, M, the option prices converge to the American prices in [CF08] (Table 4.6.9b).

Remark 4.6.3. *Other rainbow options that can by priced easily and efficiently with the 2D-COS formula are, among others, double binary (digital) cash-or-nothing options, spread options, two-asset correlation options, and exchange options.*

Table 4.6.8: Bermudan call-on-max option values $\hat{v}(t_0,\mathbf{x})$ (GBM).

(a) Convergence in N_1 and J ($S_0^2 = S_0^1 = 100$).

J	$N_1 (= N_2)$		
	40	80	160
40	14.2946	13.8919	13.8921
80	14.3032	13.8993	13.8994
160	14.3025	13.9011	13.9013
320	14.3029	13.9015	13.9017
640	14.3037	13.9016	13.9018

(b) Option prices for different initial values ($S_0^2 = S_0^1$) ($N_1 = 160, J = 320$).

S_0^1	2D-COS	95% CI [AB04]	Binomial [AB04]
90	8.0727	[8.053, 8.082]	8.075
100	13.9017	[13.892, 13.934]	13.902
110	21.3437	[21.316, 21.359]	21.345

Table 4.6.9: Bermudan put-on-min option values $\hat{v}(t_0,\mathbf{x})$ (jump-diffusion).

(a) Convergence in N_1 and J ($M = 9$).

J	$N_1 (= N_2)$		
	40	80	160
40	9.5495	9.5570	9.5586
80	9.5597	9.5580	9.5597
160	9.5606	9.5582	9.5599
320	9.5621	9.5584	9.5600
640	9.5632	9.5584	9.5600

(b) Convergence in M ($J = 320$, $N_1 = 160$).

M	$\hat{v}(t_0,\mathbf{x})$
2	9.3577
4	9.4863
8	9.5526
16	9.5862
32	9.6033

4.7. NUMERICAL EXPERIMENTS UNDER THE HESTON MODEL

In this section, we test the performance of the 2D-COS method for pricing Bermudan put options under the Heston dynamics. We use three different parameter sets:

Set A: $\rho = 0.1$, $v_0 = 0.0625$, $\bar{v} = 0.16$, $S_0 = 10$, $K = 10$, $r = 0.1$, $\eta = 0.9$, $\kappa = 5$, $T = 0.25$.

Set B: $\rho = -0.64$, $v_0 = 0.0348$, $\bar{v} = 0.0348$, $S_0 = 100$, $K = 100$, $r = 0.04$, $\eta = 0.9$, $\kappa = 1.15$, $T = 0.25$.

Set C: $\rho = -0.9$, $v_0 = 0.04$, $\bar{v} = 0.04$, $S_0 = 100$, $K = 100$, $r = 0.04$, $\eta = 0.5$, $\kappa = 0.5$, $T = 1$.

Having the Feller condition satisfied is equivalent to having $2\kappa\bar{v}/\eta^2 - 1 := q_F \geq 0$. For Set A we have $q_F = 0.98$ and the variance process remains strictly positive. For sets B and C we have $q_F = -0.47$ and $q_F = -0.84$, respectively, and the Feller condition is not satisfied.

Figure 4.7.1 shows the approximated bivariate density function $\hat{p}(X_T, v_T | x_0, v_0)$, as obtained by density recovery, equation (4.6.8), for set A.

We find that

$$|\Phi^{\pm}_{k_1,k_2}(\mathbf{x})| \leq \frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\left|\varphi\left(\frac{k_1\pi}{b_1-a_1}, \pm\frac{k_2\pi}{b_2-a_2}\middle|\mathbf{x}\right)\right| \sim \mathcal{O}\left(k_2^{-(q_F+1)}\right), \tag{4.7.1}$$

based on the univariate characteristic function of the variance process. The coefficients Φ_{k_1,k_2} decrease exponentially in k_1 and algebraically in k_2 with order $q_F + 1$. For set C this yields convergence order 0.16, which will result in very slow convergence of the 2D-COS method in this case.

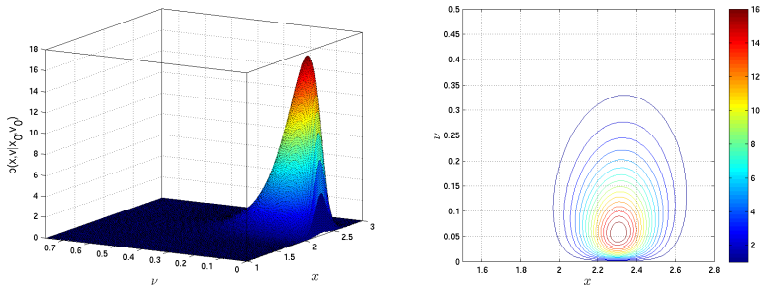


Figure 4.7.1: Bivariate density function (Set A).

4.7.1. EUROPEAN OPTIONS WITH BERMUDAN FRAMEWORK

To test the algorithm for pricing under the Heston stochastic volatility model, we calculate *European option prices with the Bermudan framework*, with $M = 12$ time steps. At every time step t_m we take only a continuation region, $\mathcal{C}^1 = [a_1, b_1] \times [a_2, b_2]$, which then corresponds to a European option and no early-exercise opportunities.

Reference prices are obtained by the 1D-COS method for European options. The error results are shown in Table 4.7.1. The error converges exponentially in N_1 and algebraically in N_2, as expected. Sets A and B give very accurate prices. However, the convergence for Set C is relatively slow and the obtained values are not highly accurate. The reason for this is the peaked density function, with coefficients Φ_{k_1,k_2} only slowly decreasing in k_2.

Table 4.7.1: Error European put option prices.

(a) Set A, $v(t_0,\mathbf{x}) = 0.5015$ ($q_F = 0.98$).

N_1	N_2		
	50	100	200
50	-1.01e-4	-1.03e-5	1.02e-6
100	-1.07e-4	-1.63e-5	-4.90e-6
200	-1.07e-4	-1.62e-5	-4.86e-6

(b) Set B, $v(t_0,\mathbf{x}) = 3.1325$ ($q_F = -0.47$).

N_1	N_2		
	50	100	200
50	3.83e-4	2.05e-4	8.94e-5
100	3.75e-4	1.95e-4	7.90e-5
200	3.75e-4	1.95e-4	7.90e-5

(c) Set C, $v(t_0,\mathbf{x}) = 6.2711$ ($q_F = -0.84$).

N_1	N_2		
	250	500	1000
50	7.8787e-02	4.3008e-02	2.6131e-02
100	6.8267e-02	3.1137e-02	1.3630e-02
200	6.7979e-02	3.0760e-02	1.3220e-02

4.7.2. BERMUDAN PUT OPTIONS UNDER HESTON DYNAMICS

In this section, we consider Bermudan put options with $M = 10$ early-exercise dates. For the first test, we select $J = 2^7$ continuation and early-exercise regions. The results in Table 4.7.2 show convergence in N_1 and N_2 and the prices for sets A and B match the results in [FO11] very well. Convergence for set C, with $q_F = -0.84$, is somewhat slow because of the slowly

decreasing Fourier coefficients in this case. The quadrature-COS method ([FO11]) gives the reference value 5.3982.

In a second test we vary the number of continuation and exercise regions. The convergence is quadratic in J, see Table 4.7.3.

The computation times are significantly longer than for the Bermudan rainbow options under Lévy processes in Section 4.6.2 because efficient matrix-vector multiplication with the FFT algorithm is applicable in only one direction, see Section 4.4. However, accuracy of four digits can still be obtained within seconds for set A and set B. PDE methods (e.g., [HitH13]) may be better for set C.

Table 4.7.2: Results Bermudan put option ($J = 2^7$).

(a) Option values, Set A ($q_F = 0.98$).

N_1	N_2			
	40	60	80	100
40	0.517765	0.517869	0.517894	0.517903
60	0.517176	0.517285	0.517311	0.517320
80	0.517021	0.517130	0.517156	0.517165
100	0.517008	0.517116	0.517143	0.517152

(b) Option values, Set B ($q_F = -0.47$).

N_1	N_2			
	40	60	80	100
40	3.20083	3.20077	3.20071	3.20066
60	3.19909	3.19903	3.19897	3.19893
80	3.19912	3.19907	3.19901	3.19897
100	3.19910	3.19905	3.19899	3.19894

(c) Option values, Set C ($q_F = -0.84$).

N_1	N_2			
	40	60	80	100
40	5.70464	5.57416	5.50068	5.45386
60	5.77720	5.67101	5.61174	5.57399
80	5.75689	5.64811	5.58898	5.55195
100	5.75808	5.64673	5.58552	5.54697

(d) CPU time (s).

N_1	N_2			
	40	60	80	100
40	6.7	10.5	14.8	22.8
60	8.8	14.0	20.2	31.2
80	11.2	17.9	25.8	42.9
100	13.5	21.4	31.2	51.2

Table 4.7.3: Bermudan put option values $\hat{v}(t_0, \mathbf{x})$, convergence in J ($N_1 = 100$, $N_2 = 100$).

	$\log_2 J$				
	3	4	5	6	7
Set A	0.515991	0.516784	0.517060	0.517133	0.517152
Set B	3.186917	3.196282	3.198380	3.198837	3.198944
Set C	5.42431	5.51038	5.53874	5.54523	5.54697
CPU time (s)	3.4	6.3	12.4	25.1	51.2

4.8. HIGHER-DIMENSIONAL COS METHOD

The 2D-COS formula can be generalized to higher dimensions easily. In this section we elaborate on the possibilities and difficulties of the higher-dimensional COS method. In Section 4.8.1, we demonstrate an example with three underlying assets.

Suppose we have an n-dimensional asset price process, with $N_1 = N_2 = \ldots = N_n = N$ terms in the series summations. Then, with the methodology that is described in Section 4.3, the computational complexity of pricing a Bermudan option is $\mathcal{O}(2^n N^n \log_2 N)$. It is clear that we cannot choose n too large, as then the curse of dimensionality sets in. Besides, the data storage grows exponentially in n. For example, in seven dimensions, with $N = 100$, the storage of the Fourier coefficients $\mathcal{V}(t_m)$ requires $8 \cdot 10^5$ GB of memory. Dimensions $n \in \{2, \ldots, 5\}$ should, however, still be fine. For higher dimensions, efficient aggregation of states or sparse grid methods are needed. However, most existing methods for a dimensionality of ten and higher are based on Monte Carlo methods.

For pricing Bermudan options we need to determine n-dimensional continuation and early-exercise hypercubes. In three-dimensional (3D) space, we get rectangular cuboids. For four or higher dimensions it gets harder to visualize the regions. Special algorithms to determine the regions efficiently, or even adaptively, may be helpful.

The terminal coefficients $\mathcal{V}_{k_1,k_2,\ldots,k_n}(T)$ can be approximated by using discrete Fourier cosine transforms, see Section 4.3.2.2. Computation or approximation of the multidimensional coefficients $\mathcal{G}_{k_1,k_2,\ldots,k_n}(\mathcal{G}^p)$ can be time-consuming or no analytic expression may be available, but we can use the terminal coefficients to approximate them. This approach was explained in Section 4.3.2.2.

4

4.8.1. 3D-COS FORMULA

The three-dimensional version of the COS formula reads

$$\hat{v}(t_0,\mathbf{x}) := \prod_{i=1}^{3} \tfrac{b_i-a_i}{2} e^{-r\Delta t} {\sum_{k_1=0}^{N_1-1}}' {\sum_{k_2=0}^{N_2-1}}' {\sum_{k_3=0}^{N_3-1}}' \tfrac{1}{4} \Big[\Phi^{++}_{k_1,k_2,k_3}(\mathbf{x}) + \Phi^{+-}_{k_1,k_2,k_3}(\mathbf{x}) + \Phi^{-+}_{k_1,k_2,k_3}(\mathbf{x}) + \Phi^{--}_{k_1,k_2,k_3}(\mathbf{x}) \Big] \mathcal{V}_{k_1,k_2,k_3}(T). \tag{4.8.1}$$

The coefficients $\Phi^{\pm\pm}_{k_1,k_2,k_3}$ are given by

$$\begin{aligned}\Phi^{\pm\pm}_{k_1,k_2,k_3}(\mathbf{x}) &= \prod_{i=1}^{3} \tfrac{2}{b_i-a_i} \Re\Big\{ \varphi\Big(\tfrac{k_1\pi}{b_1-a_1}, \pm\tfrac{k_2\pi}{b_2-a_2}, \pm\tfrac{k_3\pi}{b_3-a_3} \Big| \mathbf{x} \Big) \\ &\qquad \cdot \exp\Big(-ik_1\pi\tfrac{a_1}{b_1-a_1} \mp ik_2\pi\tfrac{a_2}{b_2-a_2} \mp ik_3\pi\tfrac{a_3}{b_3-a_3} \Big) \Big\} \\ &= \prod_{i=1}^{3} \tfrac{2}{b_i-a_i} \Re\Big\{ \phi_{levy}\Big(\tfrac{k_1\pi}{b_1-a_1}, \pm\tfrac{k_2\pi}{b_2-a_2}, \pm\tfrac{k_3\pi}{b_3-a_3} \Big) \\ &\qquad \cdot \exp\Big(ik_1\pi\tfrac{x_1-a_1}{b_1-a_1} \pm ik_2\pi\tfrac{x_2-a_2}{b_2-a_2} \pm ik_3\pi\tfrac{x_3-a_3}{b_3-a_3} \Big) \Big\}, \end{aligned} \tag{4.8.2}$$

where the last equality holds for Lévy processes, with $\phi_{levy}(u_1,u_2,u_3) := \varphi(u_1,u_2,u_3|0,0,0)$.

We test the 3D-COS formula by pricing European basket call options, with parameter set IV. The results in [LO08b] are used as reference prices. For the coefficients $\mathcal{V}_{k_1,k_2,k_3}(T)$ we use discrete Fourier cosine transformations to approximate them. The results are presented in Table 4.8.1. They show that the multidimensional-COS formula is very accurate. The CPU time for $N_1 = 40$ and $Q = 200$ is less than two seconds.

Table 4.8.1: Absolute error geometric basket call, $v(t_0,\mathbf{x}_0) = 11.9791$, and arithmetic basket call, $v(t_0,\mathbf{x}_0) = 13.2449$ ($N_1 = N_2 = N_3$).

Q	Geometric basket N_1				Arithmetic basket N_1			
	10	20	40	80	10	20	40	80
50	1.78	3.12e-2	4.03e-3	n/a	2.11e+1	4.59e-1	6.94e-3	n/a
100	1.79	4.53e-2	7.92e-3	7.91e-3	2.14e+1	4.80e-1	9.82e-5	7.82e-5
200	1.81	3.68e-2	9.88e-4	9.89e-4	2.15e+1	4.86e-1	<5.0e-5	2.13e-5
400	1.81	3.77e-2	2.46e-4	2.47e-4	2.15e+1	4.87e-1	<5.0e-5	<5.0e-5

4.9. CONCLUSION

In this chapter we presented the generalization of the COS method to higher dimensions. The recursive algorithm can be applied to, for example, pricing rainbow options, but also to pricing single-asset options under the Heston stochastic volatility model.

The 2D-COS method for valuation of Bermudan options is based on the dynamic programming principle and 2D Fourier cosine series expansions. For smooth density functions, the resulting method converges exponentially in N_1 and N_2, the number of terms in the series summations. Otherwise we achieve algebraic convergence, as in the 1D case. For multidimensional stochastic processes in the class of Lévy processes, we can apply efficient matrix-vector multiplication using the FFT algorithm. The Heston model is not in the Lévy class and the FFT algorithm will be applicable in only one dimension.

We performed extensive numerical experiments with European and Bermudan options. The algorithm for pricing two-color rainbow options can be used for various payoffs and performs highly satisfactorily. The asset prices in the tests are modeled by geometric Brownian motions or correlated jump-diffusions, but the method can be applied to a broad class of multidimensional asset price processes for which the characteristic function is available.

The density function of the Heston dynamics may be non-smooth, especially near zero variance if the Feller condition is not satisfied, and peaked densities give rise to slow algebraic convergence. However, for a wide range of parameter values, the 2D-COS method achieves accurate Bermudan put prices.

CHAPTER 5

The Social Discount Rate under a Stochastic A2 Scenario

5.1. INTRODUCTION CLIMATE CHANGE ECONOMICS

Over the past 600,000 years, until the pre-industrial revolution, the level of atmospheric CO_2 has varied from 180 to 300 parts per million (ppm) [Wei09, S$^+$07]. Nowadays, the level is at 390 ppm and is steeply increasing [NOA]. A widespread believe is that the changing climate is related to the increasing CO_2 level. According to the Intergovernmental Panel on Climate Change (IPCC) [S$^+$07], global mean surface temperatures have risen by $0.74\,^{\circ}\text{C} \pm 0.18\,^{\circ}\text{C}$ when estimated by a linear trend over the last 100 years (1906-2005). Moreover, the rate of warming over the last 50 years has almost doubled that over the last 100 years ($0.13\,^{\circ}\text{C} \pm 0.03\,^{\circ}\text{C}$ versus $0.07\,^{\circ}\text{C} \pm 0.02\,^{\circ}\text{C}$ per decade). It is considered very unlikely that climate changes of at least the seven centuries prior to 1950 were due to variability generated within the climate system alone.

Emission of greenhouse gases (GHGs) increases the radiative forcing in the atmosphere. Besides, there are so-called feedback factors which amplify, or diminish, the effect of forcing and temperature increase. Examples are ice-albedo feedback, water vapor feedback, and cloud feedback. The system of climate feedbacks is explained in [RB07, Roe09]. A conventional measure of the climate system's response to forcing is the climate sensitivity. The climate sensitivity ΔT_{2x} is defined as the equilibrium change in global and annual surface air temperature due to an increase in downward radiative flux that would result from sustained doubling of atmospheric CO_2 over its pre-industrial value. A doubling of carbon dioxide produces a radiative forcing of about 3.7 watts per square meter. The IPCC (2007) estimates a 'likely' range of ΔT_{2x} lying between 2 and 4.5 °C and 'very unlikely' to be less than 1.5 °C [S$^+$07, RB13].

In the SRES report [NAD$^+$00], the IPCC discusses several emission scenarios where they distinguish between GHGs emission, land-use, population, socioeconomic development, and technological change. They provide projections of future temperature rises. The different scenarios give a broad range of possible outcomes. For example, the A1FI scenario has a best estimate temperature rise of 4.0 °C (at 2090-2099 relative to 1980-1999) with a likely range of 2.4 to 6.4 °C, whereas the B1 scenario gives a range of only 1.1 to 2.9 °C.

Above figures demonstrate the uncertainty about future climate. The changing climate and temperature increase affect the production of food, the occurrence of flooding and other natural disasters, as well as the economic progress. Climate change economics investigates the impact of climate change on the world economy. Two leading research articles in this field are written by Stern [Ste07] and Nordhaus [Nor08]. One of the main conclusions of

This chapter is based on the article 'The social discount rate under a stochastic A2 scenario', 2014 [ARO14].

Stern is that "climate change is a serious global threat, and it demands an urgent global response". On the other hand, Nordhaus claims that despite the threats climate change poses to the global economy, it would be more equitable and efficient to invest in reproducible capital and human capital now so as to build up the productive base of economies and to put into effect controls on carbon in an increasing, but gradual manner, starting several decades from now. Dasgupta [Das08] illustrates the impact of the choice of different preference parameters (time discount factor, risk aversion,...) and explains how this leads to the two opposite conclusions of Stern and Nordhaus. For the parameter choice of Nordhaus he finds a high discount rate 4.3%, while Sterns parameter values give 1.4%.

Discount rates are used in cost benefit analysis (CBA) to convert future costs and benefits to present values. Suppose a project has high initial costs and benefits in the far future. Then a lower discount rate assigns a higher value to the future benefits and increases the profitability of the project. It is common practice to use a lower discount rate for investments with high uncertain risks. In this chapter we investigate the impact of climate change and its inherent uncertainty on discount rates.

The research in this chapter is a result of collaboration with CPB Netherlands Bureau for Economic Policy Analysis. The numerical methods from the previous chapters are applied to a specific model, but the focus here is on the modeling and economic implications. Because of that, the contents, terminology, and notation differ from the rest of this thesis. For example, T denotes temperature instead of the terminal time (t^* here) to stay in line with notation in literature about climate change economics.

5.2. INTRODUCTION SOCIAL DISCOUNT RATE AND A2 SCENARIO

Recent work shows that the social discount rate may escape the logic of exponential discounting, when shocks on capital productivity are permanent instead of transitory. Here, the social discount rate is defined as the certainty-equivalent rate of return on a take-it-or-leave-it marginal investment at time 0 with a certain payoff at time t. For example, using a simple Ramsey optimal growth model with an immediate and once-and-forever shock to capital productivity, Weitzman [Wei10] finds that the social discount rate for three centuries hence equals 0.6% even though the expected rate of return in his model equals 6% per year. Intuitively, permanent shocks imply that the riskiness of consumption is exponentially increasing over time, giving even moderately risk-averse consumers a strong incentive to increase their savings. Although the message from Weitzman's paper is crystal clear - in cost-benefit analysis, the far-distant future should be discounted far less heavily than is done by standard exponential discounting at a constant rate - any practical implementation within climate-change CBA's hinges crucially on the availability of social discount rates that incorporate actual climate risks.

To that end, this chapter employs the stochastic A2 scenario of Roe [RB13] to study the impact of climate risk on the social discount rate between 2010 and 2300. This stochastic A2 scenario emulates the climate forcing of the IPCC A2 scenario in the 21st century, after which a decline in forcing is assumed at a rate which approximately stabilizes temperature for the median trajectory. It not only recognizes that the thermal inertia of the ocean will prevent too abrupt warming of the earth's climate, but also that the earth's climate sensitivity and its response time are positively related: whereas it will take 50 years to reach two-thirds of the equilibrium impact when climate sensitivity equals 1.5 °C, it will

take 5000 years when climate sensitivity equals 15 °C. Taken together, these considerations imply that the transient response of the earth's climate system to increases in greenhouse gases is limited [RB13]. Nonetheless, it is conceivable, although unlikely, that we might see double-digit increases in temperature three centuries hence. For example, the 2-sigma temperature range for the stochastic A2 scenario employed in this chapter equals [3.2, 6.2] °C in 2100, but [3.3, 10.8] °C in 2300.

To access the impact of the 'gradual' dynamics of the earth's temperature on the social discount rate, we employ a stochastic version of a Ramsey optimal growth model, which was originally developed by Cox, Ingersoll, and Ross [CIR85a]. Our research thus fits into a small number of papers that have used this CIR-model as a motivation to establish a schedule of declining discount rates by using historical data [NP03, GKPP07, MMP13]. In contrast, our model is forward looking and not only incorporates the above-mentioned climate-change-related stylized facts, but also uses the idea of rare macro-economic disasters [Rie88, Bar09] to adequately capture key features of financial markets data. Thus, in a world unaffected by climate change, the expected rate of return on capital and the social discount rate are constant over time and equal to 6% and 1% respectively, implying an equity premium of 5%.

Our main results are fourfold. First, climate-change related risks will not have a sizeable effect on the social discount rate in this century. Intuitively, the earth's inertia prevents 'catastrophic' states of nature to be reached with sufficient probability before 2100. Second, the social discount rate may see a rapid decline after the turn of this century, where the extent of the decline primarily depends on the size of the damage in 'catastrophic' states of nature. For example, under the reactive specification of the damage function proposed by Weitzman [Wei12], the social discount rate reaches subzero levels as of 2180, after which it declines to -2.0% in 2300. The decrease of the social discount rate is more moderate under the more conventional quadratic specification of the damage function used by Nordhaus [Nor08], reaching 0.7% in 2200 and 0.6% in 2300. All in all, our simulations reveal that relatively moderate deviations from the quadratic damage function are sufficient for a subzero social discount rate by 2300. Third, the addition of interaction between economic and climate risks significantly lowers our estimate of the social discount rate by about 0.9% by 2300. Fourth, the social discount rate under the stochastic scenario deviates significantly from its deterministic counterpart, signifying that the proper treatment of far-distant uncertainty is critically important. To illustrate, by 2300 the social discount rate under the stochastic scenario is 0.1 percentage point (13%) lower for the quadratic damage function, but more than 2.3 percentage points (767%) lower for the reactive damage function.

Surprisingly, our results show that Weitzman's stylized example with an immediate and once-and-forever shock to capital productivity is actually at the upper end of our estimates of the social discount rate. Under the stochastic A2 scenario, any deviation from the frequently used quadratic damage function will result in lower, and in many cases negative, estimates of the social discount rate by 2300. Notice that negative real discount rates are relatively common and may occur over extended periods of time [RS11, MMP13]. For example, in times of a crisis, investors may have such a desire to avoid risk that they are willing to accept a negative yield on riskless investments.[1] Our results do not necessitate unrealistically low assumptions on either the pure rate of time preference or the coefficient of relative risk aversion, as, in our base case, these have values of 0.015 and 4, respectively. Instead, it

[1] In the aftermath of the financial crisis of 2007, both 2-year German government bonds and 1-month US Treasury bills have been sold at a negative real rates [U.S14].

is the *combination* of the reactive damage function and the risk of high temperatures under the A2 scenario that is crucial for obtaining negative social discount rates three centuries hence.

Interestingly, a substantial part of the difference between our results and Weitzman's stylized example can be attributed to the fact that the expected rate of return on the take-it-or-leave-it investment with a certain payoff at time-t contains a term premium. This term premium measures an investments' potential to hedge against shifts in temperature and can be traced back to the desire of non-log-utility consumers to hedge against changes in temperature (cf. Cox, Ingersoll, and Ross [CIR81]). As a result, the discount rate to be used at each instant for this take-it-or-leave-it-investment will in general not be equal to the prevailing risk-free rate. In our results, this desire to hedge results in a decrease of the social discount rate (SDR) of 0.04 percentage points for the quadratic and 0.9 percentage points for the reactive damage function.

To date, there is, to the best of our knowledge, no study tying the stochastic development of temperatures under a realistic emissions scenario, such as the IPCC A2 scenario, to the level of the social discount rate. Using the parable of a once-and-forever shock to capital productivity, Weitzman [Wei10] obtains an analytical solution for the social discount rate in the context of a simple Ramsey optimal growth model. In a standard Lucas tree economy, Gollier [Gol14] shows that efficient discount rates are decreasing with maturity, i.e., contain a term premium, when growth rates are serially correlated, but he does not explicitly model temperature. Nordhaus [Nor08] and Stern [Ste07] use assumptions on the pure rate of time preference and the coefficient of relative risk aversion (which they refer to as the elasticity of marginal utility) to obtain a social discount rate, but they do not establish a formal link between uncertainty and the social discount rate. In a study on the optimal level of environmental investment within a setting of economic and environmental disasters, Barro [Bar13] stresses the importance of gauging (fat-tailed) uncertainty based on empirical evidence. Using a database of 185 economic rare disasters for 40 countries over periods going back as far as 1870, he is able to explain the observed equity premium on financial markets without invoking unrealistically low assumptions on either the pure rate of time preference and the coefficient of relative risk aversion. He extends the model to include environmental rare disasters, but the present-day probability of an environmental disaster under the baseline calibration does not concord with the fact that the thermal inertia of the ocean will prevent too abrupt warming of the earth's climate. Cai [CJL13] extend the DICE model of Nordhaus [Nor08] to a stochastic setting, but they do not include fat-tailed uncertainty based on empirical evidence on rare macro-economic disasters.

The outline of this chapter is as follows. In Section 5.3, we present our model of choice and derive the corresponding social discount rate. Section 5.4 presents and discusses the calibration of the model's economic and physical parameters. The results are presented in Section 5.5. Section 5.6 discusses some policy implications and concludes.

5.3. THE OPTIMAL PORTFOLIO MODEL AND THE SOCIAL DISCOUNT RATE

In this section, we explain our stochastic general equilibrium model, which is based on the general equilibrium model of asset pricing by [CIR85a]. Consider an economy in which a single production sector produces the consumption good C_t. The time-t value of an initial investment in this production sector is denoted by K_t. Under continuous reinvestment of

the output, this value evolves according to the following stochastic differential equation

$$dK_t = (\alpha(T_t) - \kappa\lambda)K_t dt + K_t G(t) d\omega_t + K_t \kappa dq_t, \tag{5.3.1a}$$

$$G(t)d\omega_t = \begin{bmatrix} G_1(t) & G_2(t) \end{bmatrix} \begin{bmatrix} d\omega_t^1 \\ d\omega_t^2 \end{bmatrix}. \tag{5.3.1b}$$

Here, ω_t denotes a two-dimensional Wiener process. The increments $d\omega_t^i$ are independent normally distributed with mean zero and standard deviation $\sqrt{dt}$. This Wiener process evolves continuously over time: whereas random process ω_t^1 represents 'normal' economic fluctuations, random process ω_t^2 represents fluctuations in the earth's climate. $G(t)$ denotes the return volatility, which may be time-dependent. We extend the model of [CIR85a] with jumps, as in Ahn [AT88]. Let q_t be a Poisson process with intensity rate λ, while $\kappa < 0$ is its fixed jump size. This Poisson process is discontinuous and reflects that from time to time macroeconomic disasters, such as wars, depressions and revolutions, may hit the economy (cf. Barro [Bar09]). In our model, these investments have stochastic constant returns to scale with expected value $\alpha(T_t)$. Note that we explicitly allow the expected rate of return α to depend on the temperature T_t, which follows the following stochastic differential equation

$$dT_t = \mu(t, T_t)dt + \sigma_2(t, T_t)d\omega_t^2, \quad T_{t_0} = T_0. \tag{5.3.2}$$

Here, μ is the drift of temperature and σ_2 is the temperature volatility. Note that unexpected changes in the economy do not affect temperature as we set $\sigma_1 = 0$. Thus, equations (5.3.1) to (5.3.2) describe an economy that is confronted with exogenous temperature risk: unexpected changes in temperature and unexpected changes in the rate of return on capital are correlated with covariance $\mathrm{Cov}(dK_t, dT_t)/dt = K_t G(t)\sigma'(t, T_t)$, with $\sigma = [0, \sigma_2]$.

The consumer allocates a share a of his wealth W_t to the physical production opportunity (with expected return $\alpha(T)$) and the remaining part $1-a$ to the riskless opportunity for borrowing and lending (with a risk-free return[2] r). Thus, aW_t denotes the amount of time-t wealth invested in the physical production sector. In addition, he chooses his consumption flow C_t as to maximize the expected discounted utility over the interval $[t, t^*]$, where t^* denotes the terminal time. We use the univariate CRRA utility function

$$\mathcal{U}(C) = C^{1-\eta}/(1-\eta), \quad \eta > 0, \eta \neq 1, \tag{5.3.3}$$

where η is the constant coefficient of relative risk aversion. The value function is

$$v(t, W, T) = \max_{\{a,C\}} \mathbb{E}^{t,W,T}\left[\int_t^{t^*} e^{-\delta(\tau-t)}\mathcal{U}(C_\tau)d\tau + e^{-\delta(t^*-t)}\mathcal{U}(W_{t^*})\right], \tag{5.3.4}$$

with δ, the pure rate of time preference, capturing the consumer's impatience. The wealth dynamics evolve according to

$$\begin{aligned} dW_t &= aW_t\tfrac{dK_t}{K_t} + (1-a)rW_t dt - C_t dt \\ &= \big(a(\alpha(T_t) - \kappa\lambda - r) + r - p_t\big)W_t dt + aW_t G(t)d\omega_t + aW_t\kappa dq_t, \end{aligned} \tag{5.3.5}$$

[2] In general, the risk-free rate will depend on temperature T_t, wealth W_t and time t.

5

where $p := C/W$ denotes the proportion of wealth consumed. The corresponding Hamilton-Jacobi-Bellman equation reads [CIR85a, AT88]

$$\begin{aligned} 0 = \max_{C\geq 0, a>0} \Big[& \mathcal{U}(C) + v_W\left(a(\alpha-\kappa\lambda-r)W + rW - C\right) + \tfrac{1}{2} v_{WW} a^2 W^2 GG' \\ & + v_T\mu + \tfrac{1}{2} v_{TT}\sigma\sigma' + v_{WT} aWG\sigma' + v_t - \delta v + \left(v(t, W+a\kappa W, T) - v\right)\lambda \Big] \\ := \max_{C\geq 0, a>0} & \psi(C, a | t, W, T), \qquad \forall (t, W, T) \in [0, t^*) \times \mathbb{R}_+ \times \mathbb{R}, \end{aligned} \tag{5.3.6}$$

with boundary conditions $v(t,0,T) = 0$ and $v(t^*, W, T) = \mathcal{U}(W)$. The subscripts of the value function denote partial derivatives. The optimal values of the control variables are denoted by C^* and a^*. An economic equilibrium is defined as a set of stochastic processes (r, a^*, C^*) satisfying the first order conditions $\psi_C = 0$ and $\psi_a = 0$, and the market clearing condition $a = 1$. We get

$$0 = \mathcal{U}_C - v_W, \tag{5.3.7a}$$

$$0 = v_W(\alpha-\kappa\lambda-r)W + v_{WW} aW^2 GG' + v_{WT} WG\sigma' + v_W(t, W+a\kappa W, T) W\kappa\lambda. \tag{5.3.7b}$$

The first order condition for consumption gives $C^*(t,W,T) = v_W^{-1/\eta}(t,W,T)$. With separation of variables, it follows from Cox, Ingersoll, and Ross [CIR85b] that the value function takes the form

$$v(t,W,T) = f(t,T)\mathcal{U}(W) \tag{5.3.8}$$

and

$$-\frac{W v_{WW}}{v_W} = \eta, \quad -\frac{v_{WT}}{v_W} = -\frac{f_T}{f} = \eta\frac{C_T^*}{C^*}. \tag{5.3.9}$$

This implies that the optimal consumption policy $C^*(t,W,T)/W = f(t,T)^{-1/\eta} = p^*(t,T)$ does not depend on the wealth level. The choice of the utility function implies that the equilibrium risk-free rate $r(t,T)$ (hereafter simply denoted by the risk-free rate) is independent of the wealth level too. It can be derived as in Cox, Ingersoll, and Ross [CIR85a] and Ahn [AT88], and equals

$$r(t,T) = \alpha - \eta GG' - \eta\tfrac{C_T^*}{C^*} G\sigma' + [(1+\kappa)^{-\eta} - 1]\kappa\lambda, \tag{5.3.10}$$

where C_T^* represents the derivative of optimal consumption with respect to temperature. Inspection of equation (5.3.10) reveals that the risk free rate is equal to the expected rate of return α minus the risk premium on aggregate wealth $\eta GG' + \eta\frac{C_T^*}{C^*} G\sigma' - [(1+\kappa)^{-\eta} - 1]\kappa\lambda$. The risk premium on aggregate wealth can be associated with the presence of 'normal' economic fluctuations $\eta GG'$, temperature risk $\eta\frac{C_T^*}{C^*} G\sigma'$, and the presence of macroeconomic disasters $-[(1+\kappa)^{-\eta} - 1]\kappa\lambda$. When wealth and temperature are uncorrelated, that is $G\sigma' = 0$, the second term in the risk premium will cancel out. The risk premium associated with macroeconomic disasters is proportional to the disaster probability λ, but depends nonlinearly on the disaster size κ and the coefficient of relative risk aversion η.

The Social Discount Rate The prime interest of this chapter lies with the social discount rate, which is defined as the certainty-equivalent rate of return on a take-it-or-leave-it investment at time t_0 with a certain payoff at time t (cf. Weitzman [Wei10]). Notice that this

social discount rate is *not* equal to the risk-free rate $r(t, T_t)$ in equation (5.3.10), which is the instantaneous rate of return on an riskless investment in the interval $(t, t+dt)$. Instead, the social discount rate $R(t)$ is given by

$$R(t) = -\tfrac{1}{t-t_0}\log(\Phi(t)). \tag{5.3.11}$$

Here, the time-t expected discount factor $\Phi(t)$ denotes the current (time t_0) price of a pure discount bond promising to pay one euro at time t. According to Cox, Ingersoll, and Ross [CIR85a, Lemma 3], the value of this bond is given by

$$\Phi(t) = \mathbb{E}\left[\exp\left(-\int_{t_0}^{t}\beta(\tau, T_\tau)d\tau\right)\right], \tag{5.3.12}$$

where the expectation is taken over the actual temperature process (5.3.2) and $\beta(\tau, T_\tau)$ denotes the expected rate of return on our pure discount bond. Equation (5.3.12) states that the present value of this bond equals the expected discounted value of the promised payment of one euro at time t, where the relevant discount rate is given by its rate of return $\beta(\tau, T_\tau)$. Unfortunately, it does not provide a constructive way of finding $\Phi(t)$, since the randomly varying rate of return $\beta(\tau, T_\tau)$ is in general not known in advance (cf. Cox, Ingersoll, and Ross [CIR85a]). In particular, notice that the rate of return $\beta(\tau, T_\tau)$ will in general not be equal to the risk-free rate $r(\tau, T_\tau)$, since the rate of return on long-term pure discount bonds will in general contain a term premium. This term premium can either be positive of negative and measures the extra compensation for the bond's potential to hedge against shifts in the expected rate of return $\alpha(T_t)$ resulting from shifts in temperature T_t (cf. Cox, Ingersoll, and Ross [CIR81]). Only in the special case where we have equality of local expected rates of return on all bonds, we have $\beta(\tau, T_\tau) = r(\tau, T_\tau)$.[3] In fact, equation (5.3.12) provides a generalization of the social discount concept employed by Weitzman [Wei98]. To see this, consider his case where the risk-free rate $r_j(\tau)$ is drawn from a known distribution at time $t_0 = 0$ for $\tau \geq t_0$ and all uncertainty is resolved immediately after the draw. From (5.3.12), we have that $\Phi(t) = \mathbb{E}[e^{-\int_{t_0}^{t} r_j(\tau)d\tau}]$, which is equal to Weitzman's formula (5) on page 203 in [Wei98].

Fortunately, Lemma 4 of Cox, Ingersoll, and Ross [CIR85a] does provide us with a practical way of determining the price of a pure discount bond promising to pay one euro at time t. It reads

$$\Phi(t) = \hat{\mathbb{E}}\left[\exp\left(-\int_{t_0}^{t} r(\tau, T_\tau)d\tau\right)\right], \tag{5.3.13}$$

where $\hat{\mathbb{E}}$ denotes the expectation with respect to risk-adjusted temperature process

$$dT_t = [\mu(t, T_t) - \phi_T]dt + \sigma_2(t, T_t)d\omega_t^2, \quad T_{t_0} = T_0, \tag{5.3.14}$$

and where $\phi_T = \eta G\sigma' + \eta C_T^*/C^*\sigma\sigma'$ is the temperature risk premium.[4] Intuitively, equation (5.3.13) states that the price of a pure discount bond promising to pay one euro at time

[3] In the asset pricing literature, this condition is known as the Local Expectations Hypothesis, see Cox, Ingersoll, and Ross [CIR81]. They show that a primary requirement for the Local Expectations Hypothesis to hold is that utility is logarithmic, i.e., $\mathcal{U}(C) = \log(C)$. In that case, the value function is separable in wealth and temperature, i.e., $v_{WT} = 0$ (see Cox, Ingersoll, and Ross [CIR85b]), and consumers have no desire to hedge against changes in temperature.

[4] From Cox, Ingersoll, and Ross [CIR85a, equation (20)], we have $\phi_T = -\frac{v_{WW}}{v_W}WG\sigma' - \frac{v_{WT}}{v_W}\sigma\sigma'$. Using (5.3.9), gives $\phi_T = \eta G\sigma' + \eta C_T^*/C^*\sigma\sigma'$.

t is equal to its expected discounted value, where the discount rate is equal to the risk-free rate and the expectation is taken with respect to the risk-adjusted temperature process in (5.3.14). The risk adjustment is done by subtracting the temperature risk premium ϕ_T from the drift of the actual temperature process (5.3.2). Since the expressions in (5.3.12) and (5.3.13) measure the same value, the special case of $\phi_T = 0$ implies that the expected rate of return on our pure discount bond must be equal to the risk-free rate in that case, i.e., we have $\beta(\tau, T_\tau) = r(\tau, T_\tau)$. This confirms our claim that equation (5.3.12) provides indeed a generalization of the social discount rate concept of Weitzman [Wei98].

5.4. CALIBRATION OF THE MODEL

The basic idea underlying the calibration of our model is that it should be able to explain a number of important stylized facts regarding asset returns or, equivalently, discount rates. More specifically, we require that - given reasonable values for the pure rate of time preference and the coefficient of relative risk aversion - the model is able to get into the right ball-park regarding the historically observed values for both the risk-free rate and the expected rate of return on assets. In other words, we impose the not unreasonable assumption that at the initial time $t_0 = 2010$, climate change has not yet affected the rate of return on assets. Thus, the expected rate of return $\alpha(T_0)$ is set equal to 6% per year. This is somewhat below the expected rate of return in Barro [Bar06, Bar09], but concords with the value chosen by Weitzman [Wei10]. In addition, the risk-free rate $r(t_0, T_0)$ is set equal to 1% per year. The risk premium is then equal to 5% per year.

Table 5.4.1: Parameters in the baseline calibration.

Parameter		Value
η	: coefficient of relative risk aversion	4
G	: return volatility vector	$[0.02, 0]$
σ	: temperature volatility vector	$[0, \sigma_2]$
λ	: macroeconomic disaster probability	0.017
κ	: effective macroeconomic disaster size	-0.406
$\alpha(T_0)$	: expected rate of return absent climate change	0.06
δ	: pure rate of time preference	0.015
δ_K	: rate of depreciation	0.05

Next, we calibrate the parameters in equation (5.3.10) for the case of $G\sigma' = 0$. As noticed by Barro [Bar06], the usual view in the finance literature is that η is in the range of something like 2 to 5. Weitzman [Wei10] uses a value of 3, while Nordhaus [Nor08] uses a value of 2. However, Barro [Bar09] finds that values lower than 3 do not accord with observed equity premia and risk-free rates. In our baseline calibration, we use η equal to 4. This and subsequent calibration parameters are collected in Table 5.4.1. We take the return volatility vector G equal to $[0.02, 0]$. The value of G_1 accords with Barro [Bar06, Bar09], but notice that its value is quantitatively unimportant in the calibrations, as $\eta G_1 G_1 = 0.16\%$.[5] This is the famous risk-premium puzzle. However, the disaster probability λ and the contraction proportion κ are nontrivial in the calibrations: $-[(1+\kappa)^{-\eta} - 1]\kappa\lambda$ equals 4.85%, when $\lambda = 0.017$ per year and $\kappa = -0.406$. Here, $\lambda = 0.017$ corresponds to the observed frequency at which macroeconomic disasters, such as wars, depressions, and financial crises, have

[5] We have not been able to find any empirical information on the value of G_2. Hence, we set this parameter to zero in the baseline calibration.

occurred during the last century [Bar09]. Instead of using Barro's (2009) empirical distribution of κ with impacts ranging from 15% to 64% and a mean of 29%, we use the effective average value of a contraction, which is equal to 40.6%. As noted by Barro [Bar09], this effective average value generates about the same equity premium and welfare effects as the empirical observed frequency distribution. Figure 5.4.1 displays the iso-risk-free-rate lines for $r(t_0, T_0)$ and shows which choices of κ and η keep the level of the risk-free rate constant (given the calibrated values of the other parameters $G_1 = 0.02, \alpha = 0.06$ and $\lambda = 0.017$). In the figure, our baseline calibration is indicated by a star. In the sensitivity analysis, we will vary κ and η such that the risk-free rate $r(t_0, T_0)$ remains unchanged. Finally, we set the pure rate of time preference equal to 0.015, but notice that, in the baseline calibration, its choice is immaterial for the level of the risk-free rate, because $G\sigma' = 0$.[6]

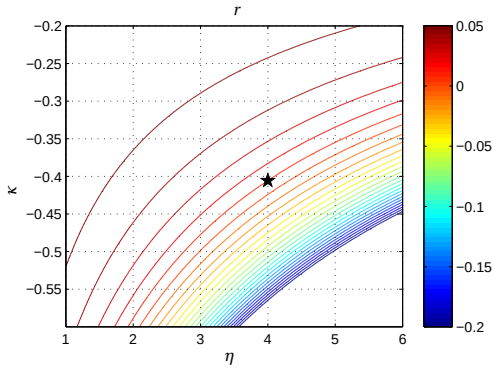


Figure 5.4.1: Risk-free rate $r(t_0, T_0)$ as function of κ and η.

To visualize the role of the return volatility G_1, the disaster probability λ and the contraction proportion κ in the level of the risk premium, Figure 5.4.2 displays twenty-five randomly drawn wealth paths for the period 2010 to 2200 satisfying equation (5.3.5) and the parameter values of Table 5.4.1.[7] The figure clearly displays the contrast between the almost negligible influence of the return volatility G_1, which is associated with the 'wiggling' behavior of the wealth paths, and the huge impact of contractions, which is associated with the large discontinuous jumps in wealth.

Next, we specify the dependence of the expected rate of return $\alpha(T)$ on temperature T. Tol [Tol09] provides a comprehensive list of the welfare effects of climate change. None of the thirteen studies in his review has looked at the welfare effects of temperature increases beyond 3 °C. Moreover, these studies show that considerable uncertainty exists regarding the welfare impact of climate change, even for 'moderate' temperature increases. For example, the estimates for a 2.5 °C warming range from -1.9 percent to +0.9 percent of GDP. The absence of more-than-three-degrees studies implies that any estimate of the welfare impacts of climate change beyond 3 °C is, by necessity, based on extrapolation. And although we may be fairly confident about a damage extrapolation for a warming of 4 °C, that

[6] Notice that this is an artifact of our specific modeling assumptions: with constant stochastic returns to scale and power utility, changes in the pure rate of time preference have no effect on the expected rate of return, the risk-free rate, and the equity premium (cf. Barro [Bar09]).

[7] Notice that in these simulations (only), we assumed that climate change has no effect on the rate of return $\alpha(T)$, i.e., $\forall T : \alpha(T) = \alpha(T_0)$. Consequently, optimal consumption is a fixed percentage of wealth.

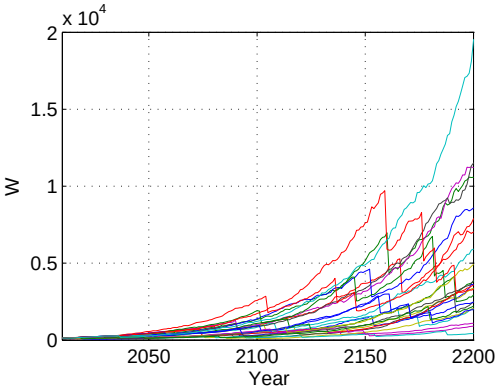


Figure 5.4.2: Simulated wealth paths.

confidence will be greatly reduced for damage extrapolations to 10 °C or even 15 °C. Essentially, we have no objective way to determine the magnitude of high-temperature damages [Wei12]. Therefore, in this study, we use a 'range' of damage functions, encompassing both the low- and high-end of the available appraisals. Specifically, the damage function employed by Nordhaus [Nor08] serves as our low-end damage function, whereas the damage function employed by Weitzman [Wei12] serves as our high-end damage function, i.e.,

$$\Omega^N(T) = 1/\left(1+(T/20.46)^2\right), \tag{5.4.1a}$$

$$\Omega^W(T) = 1/\left(1+(T/20.46)^2+(T/6.081)^{6.754}\right). \tag{5.4.1b}$$

Above, $\Omega^N(T)$ and $\Omega^W(T)$ denote the 'Nordhaus'- and 'Weitzman'-damage function respectively. Notice that these damage functions are not readily comparable, as Nordhaus defines his damage function $\Omega^N(T)$ over gross output, whereas Weitzman defines his damage function $\Omega^W(T)$ over net output. For practical reasons that will become clear below, we choose to retain these differences. Hence, the expected rates of return for our low- and high-end cases are given by

$$\alpha^N(T) = -\delta_K + (\alpha(T_0)+\delta_K)\Omega^N(T), \tag{5.4.2a}$$

$$\alpha^W(T) = \alpha(T_0)\Omega^W(T), \tag{5.4.2b}$$

where δ_K denotes the rate of depreciation in the economy and is taken to be equal to 5%. Figure 5.4.3 displays the range of our assumptions on the expected rate of return. Observe that the expected rates of return α^N and α^W are almost indistinguishable for temperature increase less than 3 °C, but differ markedly for larger increases in temperature. For example, an increase of 6 °C results in an expected rate of return of 5.1% (Nordhaus) and 3.0% (Weitzman), whereas an increase of 10 °C results in an expected rate of return of 3.9% (Nordhaus) and 0.2% (Weitzman). Finally, notice that these differences would have been much larger, had we chosen to define the Weitzman damage function $\Omega^W(T)$ over gross, instead of net, output.

We calibrate our temperature process on the A2 scenario of [RB13], who employ a simple climate model to describe the interactions between the atmosphere, the surface and deep ocean layer, thereby capturing three material characteristics of global warming. First, the

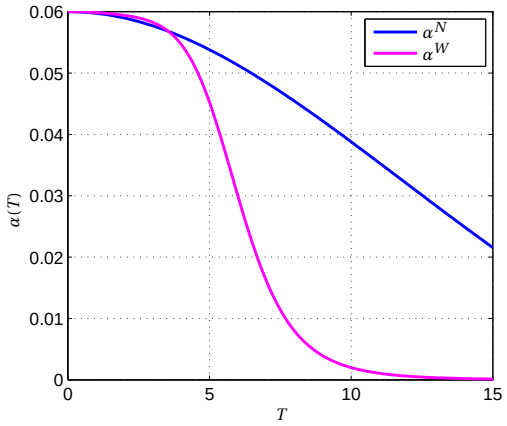


Figure 5.4.3: The low- and high-end expected rates of return $\alpha^i(T)$.

thermal inertia represented by the deep ocean slows down the response of surface temperatures on global warming. Second, climate sensitivity is assumed to be normally distributed, which results in a skewed density for the temperature response. Third, the response time, i.e., the time needed to reach the equilibrium temperature given a certain level of forcing, is positively related to climate sensitivity: whereas it will take approximately one hundred years to achieve equilibrium, when climate sensitivity is low, it will take thousands of years to achieve equilibrium when climate sensitivity is high. Figure 5.4.4 presents the simulation results of [RB13]. In their scenario, the climate forcing is chosen to emulate the IPCC A2 scenario [NAD$^+$00], where after a decline in forcing is assumed at a rate that approximately stabilizes the temperature for the median trajectory. In Figure 5.4.4, the shadings represent the 1-, 2-, and 3-sigma ranges for climate sensitivity.

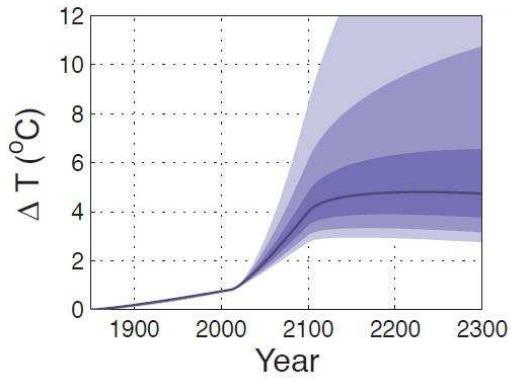
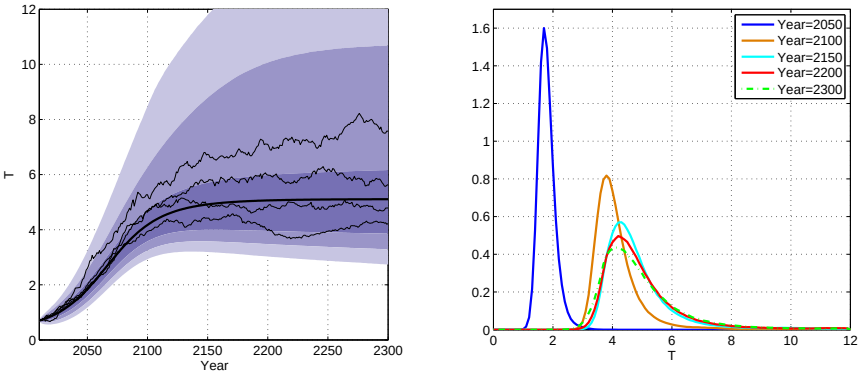


Figure 5.4.4: Quantile functions of Roe [RB13], their Fig. 4b.

We take initial time $t_0 = 2010$ and terminal time $t^* = 2300$ and calibrate functions $\mu(t,T)$ and $\sigma_2(t,T)$ in equation (5.3.2) such that the quantiles of our simulations resemble the re-

sults of [RB13].[8] Figure 5.4.5a shows a number of simulated temperature paths together with our simulated ranges for climate sensitivity, which resemble the results of [RB13] rather well. The shape of the temperature distribution at particular times is shown in Figure 5.4.5b, which clearly illustrates the time evolution of uncertainty under the stochastic A2 scenario: even though temperature increases of more than 6 °C are very unlikely to occur before 2100, such increases are much more likely to occur after 2100.



(a) Temperature paths T_t and sigma ranges.

(b) Density function T_t for different times t.

Figure 5.4.5: Calibrated temperature process.

5.5. RESULTS

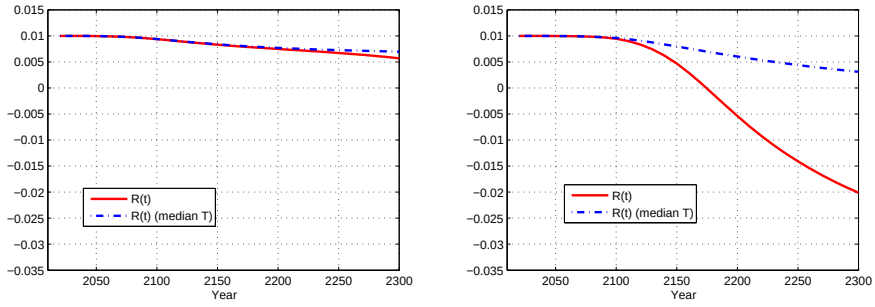
Using the parameters in the baseline calibration displayed in Table 5.4.1, we determine the social discount rate for both our low-end and high-end damage functions.[9] For sake of reference, the dashed blue lines in Figure 5.5.1 show the development of the social discount rate at the median value of temperature, i.e., the temperature on the (deterministic) A2 path of the IPCC, under both the low-end (left figure) and high-end damage function (right figure). These figures reveal immediately that the social discount rate on the deterministic A2 path does not differ markedly between these damage functions. Intuitively, this follows from the observation that on the IPCC A2 path the temperature increase is limited to 5 °C. At these temperature increases, the low- and high-end damage functions are still rather similar (see Figure 5.4.3). Finally, notice that the social discount rate on the IPCC A2 path decreases only mildly from 1% in 2010 (our starting year) to respectively 0.7% ('low-end' damage function) and 0.3% ('high-end' damage function) in 2300.

Next, the red lines in Figure 5.5.1 show the development of the social discount rate under our stochastic A2 scenario for both the low-end and high-end damage function. A number of striking results appear from this figure. First of all, comparison of Figures 5.5.1a and 5.5.1b shows that in both cases the social discount rate is almost stable over the course

[8] Details on the calibration procedure are provided in Appendix 5.A.

[9] Our procedure to obtain the social discount rate is as follows. First, we solve the optimal consumption problem of equation (5.3.4) by using the 2D-COS method, see Appendix 5.B. Substitution of the optimal consumption policy in equation (5.3.10) gives the risk-free rate $r(t,T)$. Finally, we approximate the expected value in equation (5.3.13) by means of a Monte Carlo simulation of the temperature and corresponding risk-free rate paths.

of this century. The intuition is that the earth's inertia prevents strong increases in temperature - and thereby high damages - before the end of this century. From this, however, it does not follow that climate change damages are immaterial for the social discount rate. To the contrary, Figure 5.5.1a shows that under the low-end damage function, the social discount rate slowly declines from 0.9% in 2100 to 0.6% in 2300, which is just under the SDR on the (deterministic) IPCC A2 path of 0.7%. However, under the high-end damage function the decline is much more pronounced: the social discount rate declines from 0.9% in 2100 to -2.0%(!) in 2300 (see Figure 5.5.1b). Thus, under our stochastic A2 scenario and the high-end damage function, the net present value of 1 euro in 2300 would be equal to 350 euros today. Comparison of the social discount rates in Figures 5.5.1a and 5.5.1b reveals that this 'discounting reversal' must be attributed to the choice for the high-end damage function, as the social discount rate is still positive under the low-end damage function. Notice, however, that even under the low-end damage function, the logic of exponential discounting is strongly diminished: 1 euro in 2300 would be equal to 0.18 euro today, instead of 0.13 (using the social discount rate that corresponds to the median value of temperature as the relevant discount rate) or zero (using the expected rate of return of 6% as the relevant discount rate).[10]



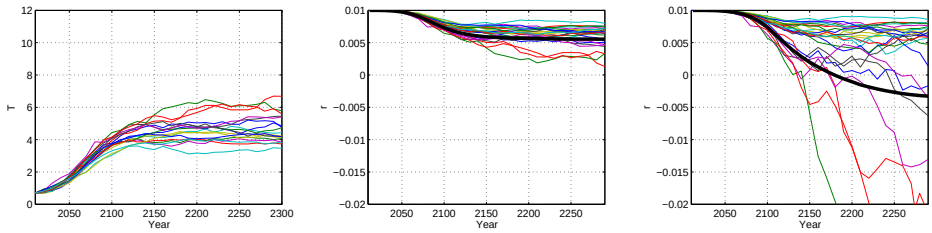
(a) low-end case, α^N (b) high-end case, α^W.

Figure 5.5.1: The social discount rate $R(t)$ in the baseline calibration.

To further apprehend our results, Figure 5.5.2a presents 25 randomly drawn temperature paths from equation (5.3.2). On these paths, the global mean temperature in 2300 varies between 3 and 8.5 °C. The corresponding risk-free rates are shown in Figures 5.5.2b and 5.5.2c for our low-end and high-end damage function, respectively. From these figures, it is immediately apparent that the negative values of the risk-free rate on high-temperature paths are driving our results. As in Weitzman [Wei10], low-probability states of the world play an important role in long-term discounting, if they are associated with a sufficiently high and persistent impact. Under our stochastic A2 scenario, the persistency of the impacts follows from the inertia of the climate system and the shape of the damage function: when today's temperature and damages are high, it is very likely that tomorrow's temperature and damages will be high as well. Surprisingly, our estimates of the social discount rate turn out to be much lower than in Weitzman's stylized example, even though in our

[10] All these present value calculations follow by taking the relevant discount rate $R(2300)$, and calculating $e^{-290R(2300)}$. We have $e^{-290\times-0.02} = 350$, $e^{-290\times0.006} = 0.18$, $e^{-290\times0.007} = 0.13$, and $e^{-290\times0.06} \approx 0$.

case - and in contrast to the immediate and once-and-forever shock considered by Weitzman [Wei10] - high-impact states of the world do not occur before the end of this century. Intuitively, this can be explained by the observation that under our stochastic A2 scenario low-probability, high impact states of the world are actually much more likely than in Weitzman's stylized example. For example, the probability that the expected rate of return is no larger than 1%, equals 0.49% in the case considered by Weitzman [Wei10],[11] but nearly 10% under the stochastic A2 scenario and the high-end damage function.



(a) T_t paths. (b) low-end damage function (c) high-end damage function

Figure 5.5.2: Simulated temperature paths (left) and the corresponding risk-free rates (middle and right).

To access the sensitivity of our results with respect to the damage function, Figure 5.5.3 shows the social discount rate for the intermediate expected rate of return $\alpha^I = 0.5(\alpha^N + \alpha^W)$. This figure reveals that - even for this intermediate case - the social discount rate becomes negative just after 2250. This can be explained by the fact that the probability that the expected rate of return is no larger than 1% is still 1.9% under the stochastic A2 scenario and the intermediate case, which remains much higher than Weitzman's 0.49%. Hence, the prime factor driving our results is not the damage function per se (although damage must surely be sizeable), but the relatively large probabilities of extreme climate change under the stochastic A2 scenario.

Table 5.5.1: The decrease in the SDR resulting from the term premium in the pure discount bond.

Damage function	2010	2050	2100	2150	2200	2250	2300
Low-end	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0
Intermediate	0.0	0.0	0.0	-0.0	-0.1	-0.2	-0.4
High-end	0.0	0.0	-0.0	-0.3	-0.9	-1.0	-0.9

Table 5.5.1 shows that the term premium in the pure discount bond underlying the SDR explains a substantial part of the SDR's decreasing term structure, at least for the intermediate and the high-end damage functions.[12] Whereas, even in 2300, the contribution of the term premium to the SDR is almost negligible for the low-end damage function, it is sizeable for both the intermediate and high-end damage functions. For example, in 2300, the impact of the term premium in the pure discount bond on the SDR is 0.9 percentage points for the high-end and 0.4 percentage points for the intermediate damage function. Thus the term premium accounts for almost 29% of the total decrease in the SDR for the high-end and

[11] See his Table 2 on page 9.
[12] In the Table, a '-0.0' denotes a negative number which is rounded of to one decimal.

25% for the intermediate damage function.[13] Intuitively, the term premium will only effect the SDR when temperature has a sizeable impact on aggregate risk, since, under those circumstances, a bond's potential to hedge against shifts in temperature has value. Obviously, this is the case under both the intermediate and the high-end damage function, but not so under the low-end damage function. This can also be confirmed directly from Figure 5.5.1a, where the narrow margin between the SDR on the deterministic and stochastic paths is a clear signal that aggregate risk is only changing slowly in that case.

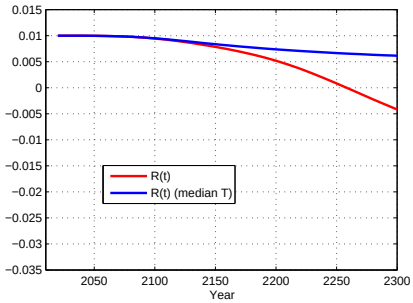


Figure 5.5.3: The social discount rate $R(t)$ under the intermediate case, α^I.

Under our baseline assumption that unexpected changes in temperature are uncorrelated with unexpected changes in wealth, changes in the coefficient of relative risk aversion η *do not* affect the social discount rate. To see why, use $G\sigma' = 0$ and rewrite (5.3.10) as

$$r(t,T) = r(t_0,T_0) + \alpha(T) - \alpha(T_0), \tag{5.5.1}$$

which states that the risk-free rate at time t equals the risk-free rate at time t_0, $r(t_0,T_0)$, plus the expected change in the rate of return on assets between t_0 and t, $\alpha(T) - \alpha(T_0)$. Subsequently, substitution of (5.5.1) into (5.3.13) and (5.3.11) gives

$$R(t) = r(t_0,T_0) - \alpha(T_0) - \tfrac{1}{t-t_0}\log\hat{\mathbb{E}}\left[\exp\left(\int_{t_0}^{t} -\alpha(T_\tau)\,d\tau\right)\right], \tag{5.5.2}$$

from which it immediately follows that the social discount rate $R(t)$ is indeed independent of the level of risk aversion η. Equation (5.5.2) states that the time-t social discount rate equals the time-zero risk-free rate $r(t_0,T_0)$ *minus* the time-0 expected rate of return on assets $\alpha(T_0)$ *plus* the time-t certainty-equivalent rate of return on assets. Intuitively, the surprising result that the social discount rate is independent from the coefficient of relative risk aversion can be understood by realizing first of all that the risk-free rate at t_0 is not affected by changes in the coefficient of relative risk aversion η: the requirement that $r(t_0,T_0) = 0.01$ implies that we are moving along the iso-risk-free-rate line indicated by a star in Figure 5.4.1. Hence, any change in the coefficient of relative risk aversion must be fully offset by an appropriate change in the disaster size κ, leaving the risk-free rate at t_0 unchanged.[14] In addition, changes in risk aversion do not affect the expected rate of return on assets $\alpha(T)$. Hence, the social discount rate is unaffected by changes in risk aversion,

[13] As measured from the no-climate-change value of the SDR of 1.0%.

[14] Notice that the derivative of $r(t,T)$ with respect to η equals $-GG' - \log(1+\kappa)\kappa\lambda(1+\kappa)^{-\eta}$. Hence, in general, higher risk aversion will lower the risk-free rate.

whenever unexpected changes in temperature are uncorrelated with unexpected changes in wealth.

Inspection of equation (5.5.2) shows that high values of the expected rate of return will discount themselves out of existence, which implies that the certainty-equivalent rate of return on assets tends to the lowest possible value when t goes to infinity (cf. Weitzman [Wei98]). In terms of our stochastic A2 scenario, this means that the expected rates of return belonging to low temperature paths become less important over time, whereas the expected rates of return belonging to high temperature paths become more important over time. Eventually, the expected rates of return on the highest temperature path will come to dominate the social discount rate $R(t)$.

Finally, we explore the sensitivity of our results with respect to $G_2(t)$, which captures the impact of unexpected changes in temperature on the rate of return on assets. Given the lack of empirical evidence on this parameter, we let $G_2(t)$ decline linearly from zero in 2010 to -0.025 in 2050, after which its value remains constant. This choice implies that, as of 2050, an unexpected increase of temperature by 0.5 °C decreases the rate of return by 12.5%. So, in case the expected rate of return would have been 0.06, an unexpected temperature increase of 0.5 °C would have resulted in a $0.06 \times 0.125 = 0.0075$ percentage point decrease of the rate of return on assets. Figure 5.5.4 shows that - under the high-end damage function and compared to our baseline calibration - such a return volatility decreases the social discount rate in 2300 from -2.0% to -2.9%.[15] Moreover, the impact of the return volatility $G_2(t)$ on the social discount rate is increasing over time until 2200, after which it roughly stabilizes. Since the return volatility itself is constant from 2050 onwards, this implies that the additional uncertainty associated with this return volatility matters more when wealth is low (and climate damages are high). Given that the risk-free rate $r(t,T)$ now depends on the level of risk aversion η, as can be seen from equation (5.3.10), it is to be expected that the social discount rate will depend on the level of risk aversion as well. Indeed, Figure 5.5.5 confirms that higher risk aversion is associated with a (somewhat) lower social discount rate.[16]

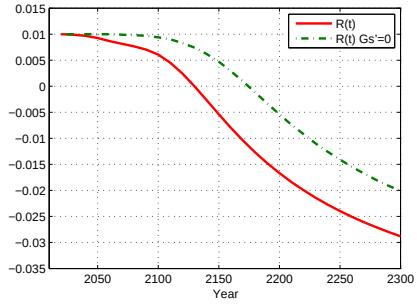


Figure 5.5.4: The social discount rate $R(t)$ under the high-end damage function for the case of time-dependent (red continuous line) and non-time-dependent (dashed green line) return volatility $G_2(t)$.

[15] Under the low-end damage function, the correlation between unexpected changes in temperature and unexpected changes in wealth has a negligible impact on the social discount rate. Results are available from the authors upon request.

[16] In order to preserve the notion that the model should be in accordance with historically observed levels of the risk-free rate, we simultaneously vary the level of risk aversion η and the effective disaster size κ, keeping the risk-free rate at time t_0 at 1%, i.e., we are moving along the iso-risk-free-rate line indicated by a star in Figure 5.4.1.

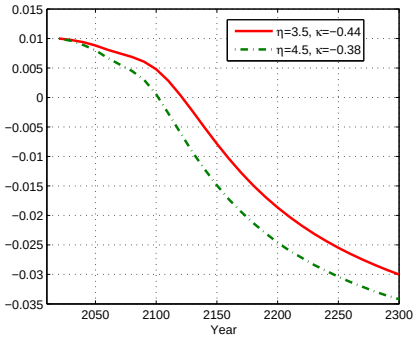


Figure 5.5.5: The social discount rate $R(t)$ under the high-end damage function and time-dependent return volatility $G_2(t)$ for the case of low risk aversion $\eta = 3.5; \kappa - 0.44$ (continuous red line) and high risk aversion $\eta = 4.5; \kappa - 0.38$ (dashed green line).

5.6. DISCUSSION AND CONCLUSION

This chapter has shown that the far-distant future may be crucially important for long-term discounting. Even though the inertia of the earth's climate prevents a too rapid increase of surface temperature, our results suggest that a small, but non-negligible probability of an extreme temperature response in the far-distant future may drive the social discount rate to subzero levels. As a result, the present value of one euro three centuries hence may be everywhere between 0.18 and 350 euros.[17] Notwithstanding the extremely wide margin in this present value, what stands out is that it is much, much larger than the present value under standard exponential discounting at a constant rate. For example, taking the discount rate three centuries hence based on our low-end (Nordhaus) damage function results in a present value that is almost ten times larger than the present value that is obtained by using the much-criticized, and supposedly too low, discount rate of 1.4% of the Stern Review.[18]

Importantly, our results suggest that the key issue determining the level of the social discount rate is not how large damages are or might be (although they surely must be substantial to invoke a sizeable response in the social discount rate), but how large the probability of extreme climate change is. Under our stochastic A2 scenario, this probability is sufficiently large to make the logic of exponential discounting in the far distant future (almost) completely irrelevant. Importantly, our limited state of knowledge regarding the high-temperature damages does not interfere with this conclusion, as the logic of exponential discounting is severely diminished under both the quadratic and reactive specification of the damage function. The shape of the damage function is, however, critically important for the occurrence of negative discount rates. Whereas the social discount rate will be strictly positive under our low-end, quadratic damage function, it may reach subzero levels under more extreme specifications of the damage function, such as the reactive damage function proposed by Weitzman [Wei12]. In that respect, it must be noticed that in the 'short' run, i.e., this century, the impact of the damage function on the social discount rate seems to be severely constrained by the inertia of the earth's climate system. In particular, our results suggest that the probability of a high-enough temperature response under our

[17] From $R(2300) = 0.006$ and -0.02, we have $e^{-290\times0.006} = 0.18$ and $e^{-290\times-0.02} = 350$, respectively.
[18] $e^{-290\times0.0061} = 0.17$ is about ten times larger than $e^{-290\times0.014} = 0.017$.

stochastic A2 scenario is insufficient to drive this century's social discount rate away from its conventional, no climate change, level.

Thus, our results seem to imply that, by and large, exponential discounting will still produce consistent valuations in *this* century. This surprising result is, however, counterbalanced by the observation that the SDR in the next century will be hugely affected by climate change unless we are more or less certain that damages from climate change will be relatively low. In many instances, it might even become negative. If it does, the social cost of carbon will explode to infinity, thus providing a clear signal that society should reduce emissions whatever the cost. Of course, the price signal of a discount rate on the IPCC A2 path is valid only for marginal mitigation projects, i.e., for investment projects that do not affect aggregate welfare, aggregate emissions and global temperature. When considering non-marginal investment projects, the probability of extreme climate change will decline, and it is likely that the SDR will increase from the levels reported in this chapter. As a result, the social cost of carbon will decline, thereby reducing the value of the marginal mitigation project on the new emissions path. The extent to which our results carry over to more moderate emission scenarios, such as the IPCC B2 scenario, is therefore crucial for cost-benefit analysis of climate change and remains a interesting topic for further research.

Appendix

5.A. CONSTRUCTION TEMPERATURE PROCESS

In this section we explain our procedure to calibrate the temperature dynamics,

$$dT_t = \mu(t, T_t)dt + \sigma_2(t, T_t)d\omega_t^2. \tag{5.A.1}$$

We consider a time horizon $[t_0, t^*] = [2010, 2300]$ years and set $T_{t_0} = 0.7\,^\circ$C. For the construction of functions $\mu(t, T)$ and $\sigma_2(t, T)$ we use so-called 'reference paths'. The reference paths are represented by the black lines in Figure 5.A.1. The slopes of the reference paths are known and for each year and temperature the value of function $\mu(t, T)$ is interpolated using the three reference paths and their slopes.

During the first years, the volatility $\sigma_2(t_0, T)$ equals 0.03. For temperature values close to the lowest reference path, volatility will stay close to this value. For temperatures along the mid reference path, volatility will increase moderately and linearly in time, up to $\sigma_2(2100, T) = 0.045$ after 90 years and will stay constant thereafter. Along the highest reference path, volatility will increase linearly in time up to the high value $\sigma_2(2100, T) = 0.2$. After 90 years the volatility is assumed to decreases to the value $\sigma_2(t^*, T) = 0.08$ at the terminal time t^*.

The lowest and mid reference paths tend to an equilibrium temperature, while equilibrium is not reached yet on the highest reference path. Again interpolation with three reference paths is used to find the value $\sigma_2(t, T)$. With this, we capture the feature of high uncertainty on high temperature paths [Roe09].

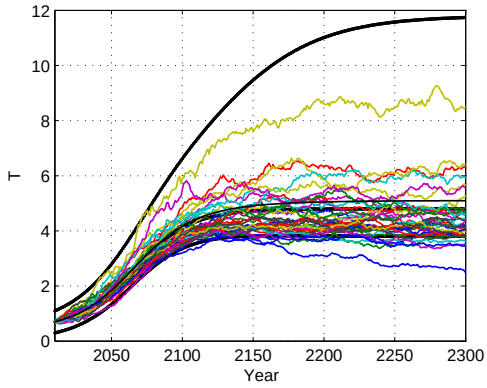


Figure 5.A.1: The three reference paths (black lines) used for construction of the temperature model together with twenty five simulated temperature paths.

5.B. NUMERICAL METHOD

Our numerical method to solve the optimal portfolio problem is based on the dynamic programming principle and Fourier cosine series expansions, called the COS method. In this section we give some details of this method.

The value function is given by

$$v(t,W,T) = \max_{\{C\geq 0\}} \mathbb{E}^{t,W,T}\left[\int_t^{t^*} e^{-\delta(\tau-t)}\mathcal{U}(C_\tau)d\tau + e^{-\delta(t^*-t)}\mathcal{U}(W_{t^*})\right]. \tag{5.B.1}$$

W and T represent the current wealth and temperature level, respectively, and the wealth dynamics evolve according to

$$dW_t = \left(\alpha(T_t) - \kappa\lambda - p_t\right)W_t dt + W_t G d\omega_t + W_t \kappa dq_t. \tag{5.B.2}$$

We switch to the log-domain, $X_t := \log W_t$, with

$$dX_t = \left(\alpha(T_t) - \kappa\lambda - p_t - \tfrac{1}{2}GG'\right)dt + G d\omega_t + \log(1+\kappa)dq_t. \tag{5.B.3}$$

The corresponding value function is

$$v(t,x,T) := \max_{\{p\geq 0\}} \mathbb{E}^{t,x,T}\left[\int_t^{t^*} e^{-\delta(\tau-t)}\mathcal{U}(e^{X_\tau}p_\tau)d\tau + e^{-\delta(t^*-t)}\mathcal{U}(e^{X_{t^*}})\right]. \tag{5.B.4}$$

x and T represent the current log-wealth and temperature level, respectively. We take an equidistant grid of control times $t_0 < t_1 < \ldots t_m < \ldots < t_M = t^*$, with timestep $\Delta t := t_{m+1} - t_m$ and approximate processes by their discrete variants. The log-wealth and temperature processes are discretized by an Euler scheme (see, for example, [PBL10, KKK10]).

We approximate the value function by (see [KD01])

$$v(t_m,x,T) = \max_{\{p_i\}} \mathbb{E}^{t_m,x,T}\left[\sum_{i=m}^{M-1} e^{-\delta(t_i-t_m)}\mathcal{U}(e^{X_{t_i}}p_i)\Delta t + e^{-\delta(t_M-t_m)}\mathcal{U}(e^{X_{t_M}})\right]. \tag{5.B.5}$$

The discrete-time value function converges in M to the continuous-time variant. The dynamic programming principle gives us

$$\begin{aligned} v(t_m,x,T) &= \max_{p_m} \mathbb{E}^{t_m,x,T}\left[\mathcal{U}(e^x p_m)\Delta t + e^{-\delta\Delta t} v(t_{m+1}, X_{t_{m+1}}, T_{t_{m+1}})\right] \\ &:= \max_{p_m}\left[\mathcal{U}(e^x p_m)\Delta t + c(t_m,x,T,p_m)\right]. \end{aligned} \tag{5.B.6}$$

The second part is called the continuation value. The continuation value can be approximated by using the 2D-COS formula (Chapter 4) and the problem is solved backwards in time (Bellman's principle of optimality), as explained in Chapter 2. We use utility function $\mathcal{U}(C) = C^{1-\eta}/(1-\eta)$, so $p^* = p^*(t,T)$ (see Section 5.3). The method parameters are

$[a_1,b_1],[a_2,b_2]$	computational domain log-wealth (1^{th} dimension) and temperature (2^{nd} dimension), respectively,
N_1, N_2	number of Fourier cosine coefficients in 1^{th} and 2^{nd} dimension, resp.,
J_1, J_2	number of control regions in 1^{th} and 2^{nd} dimension, respectively.

The continuation value is approximated by a two-dimensional COS formula

$$\begin{aligned} \hat{c}(t_m,x,T,p_m) = e^{-\delta\Delta t} \sum_{k_1=0}^{N_1-1\prime} \sum_{k_2=0}^{N_2-1\prime} \tfrac{1}{2}\Bigg[&\Re\left\{\varphi\left(\tfrac{k_1\pi}{b_1-a_1}, +\tfrac{k_2\pi}{b_2-a_2}\middle| x,T,p_m\right) e^{ik_1\pi\frac{-a_1}{b_1-a_1}} e^{ik_2\pi\frac{-a_2}{b_2-a_2}}\right\} \\ &+ \Re\left\{\varphi\left(\tfrac{k_1\pi}{b_1-a_1}, -\tfrac{k_2\pi}{b_2-a_2}\middle| x,T,p_m\right) e^{ik_1\pi\frac{-a_1}{b_1-a_1}} e^{-ik_2\pi\frac{-a_2}{b_2-a_2}}\right\}\Bigg] \mathcal{V}_{k_1,k_2}(t_{m+1}). \end{aligned} \tag{5.B.7}$$

$\varphi(.,.|x,T,p_m)$ is the bivariate conditional characteristic function of $(X_{t_{m+1}}, T_{t_{m+1}})$, given $(X_{t_m}, T_{t_m}) = (x,T)$ and value p_m. The Fourier cosine coefficients of the value function are given by

$$\mathcal{V}_{k_1,k_2}(t_m) := \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\int_{a_2}^{b_2}\int_{a_1}^{b_1} v(t_m,y,\varsigma)\cos\left(k_1\pi\tfrac{y-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{\varsigma-a_2}{b_2-a_2}\right)dyd\varsigma. \tag{5.B.8}$$

We divide the domain $[a_1,b_1]\times[a_2,b_2]$ into rectangular subdomains $\mathcal{D}^{q_1,q_2} = [z_{q_1}, z_{q_1+1}]\times[w_{q_2}, w_{q_2+1}]$. For each subdomain we determine the optimal control value p^{q_1,q_2}, then

$$\begin{aligned} \mathcal{V}_{k_1,k_2}(t_m) &= \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\sum_{q_1,q_2}\iint_{\mathcal{D}^{q_1,q_2}} \mathcal{U}(e^y p^{q_1,q_2})\Delta t\cos\left(k_1\pi\tfrac{y-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{\varsigma-a_2}{b_2-a_2}\right)dyd\varsigma \\ &\quad + \tfrac{2}{b_1-a_1}\tfrac{2}{b_2-a_2}\sum_{q_1,q_2}\iint_{\mathcal{D}^{q_1,q_2}} c(t_m,y,\varsigma,p^{q_1,q_2})\cos\left(k_1\pi\tfrac{y-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\tfrac{\varsigma-a_2}{b_2-a_2}\right)dyd\varsigma \\ &:= \sum_{q_1,q_2}\mathcal{U}_{k_1,k_2}(\mathcal{D}^{q_1,q_2}, p^{q_1,q_2})\Delta t + \sum_{q_1,q_2}\mathcal{C}_{k_1,k_2}(t_m,\mathcal{D}^{q_1,q_2}, p^{q_1,q_2}). \end{aligned} \tag{5.B.9}$$

The terms $\mathcal{U}_{k_1,k_2}$ are time-independent and known analytically. The term $\mathcal{C}_{k_1,k_2}$ are approximated by using an FFT algorithm. With this we can recover the coefficients $\mathcal{V}_{k_1,k_2}(t_m)$ backwards in time and solve the optimal portfolio problem (see Algorithm 1). All parameters are chosen such that the numerical approximations give rise to converged solutions, values and functions of interest.

CHAPTER 6

A Fourier Cosine Method for an Efficient Computation of Solutions to BSDEs

6.1. INTRODUCTION

Whereas the theory and applications of classical *forward* stochastic differential equations (FSDEs), with a prescribed initial value, is traditional and became widely known, we are concerned with *backward* stochastic differential equations (BSDEs) in this chapter. A BSDE is a stochastic differential equation for which a terminal condition, instead of an initial condition, has been specified and its solution consists of a pair of processes. The linear type of equation was introduced by Bismut in [Bis73], where linear BSDEs were used in stochastic optimal control problems as adjoint equations in the stochastic version of the Pontryagin's maximum principle. The general notion of BSDE has been introduced by Pardoux and Peng [PP90]. They proved existence and uniqueness of solutions of BSDEs under some Lipschitz conditions on the driver function. Many researchers have attempted to relax these restrictions. For example, the authors in [LSM97] show existence of a minimal solution under more general assumption for the driver function, which is assumed to be continuous with linear growth in some of its arguments. Kobylanski [Kob00] provided uniqueness and existence results for a driver with quadratic growth in one of its arguments in a Brownian filtration. These results are generalized to an infinite activity jump setting in [LS14]. For a general introduction to BSDEs we refer to [Pha09, EKPQ97].

In recent years, BSDEs have received more attention in mathematical finance and economics. For example, the Black-Scholes formula for pricing options can be represented by a system of decoupled forward-backward SDEs. Market imperfections can also be incorporated, such as different lending and borrowing rates for money, the presence of transaction costs, or short sales constraints. These imperfections give rise to more involved nonlinear BSDEs. If the asset price follows a jump diffusion process then the option cannot perfectly be replicated by assets and cash, i.e., the market is not complete. One can set up a minimum variance hedge to value the option. Another way to value and hedge options in this setting is by utility indifference pricing, where a certain utility value is assigned to the possible profits and losses of the hedging portfolio. The pricing problem can be solved by means of a BSDE with jumps.

As explained in Chapter 1, the well-known Feynman-Kac theorem gives a probabilistic representation for the solution of a linear parabolic partial differential equation (PDE) by means of the corresponding FSDE and a conditional expectation. The solution of a

This chapter is based on the article 'A Fourier cosine method for an efficient computation of solutions to BSDEs', to appear in *SIAM Journal on Scientific Computing*, 2015 [RO15].

BSDE provides a probabilistic representation for *semilinear parabolic PDEs*, see for example [PP92], which is a generalization of the Feynman-Kac theorem. Also the converse relation holds. This connection enables us to solve a semilinear PDE by probabilistic numerical methods, like Monte Carlo simulation techniques.

Probabilistic numerical methods to solve BSDEs may, for example, rely on time discretization of the stochastic process and approximations for the appearing conditional expectations. Least-squares Monte Carlo regression to approximate the conditional expectations is used in, for example, [LGW06, GLW05, BS12]. A rich literature exists on other methods, for example based on chaos decomposition formulas [BL13]. In this chapter we employ a general θ-method for the time-integration ([Ise09]) and propose a new method to approximate the solution backwards in time. This approach is based on the COS method. The method is based on Fourier cosine series expansions and relies on the characteristic function of the transitional density, which enables us to approximate the conditional expectations is a very efficient way. We call the method the *BCOS method*, short for BSDE-COS method.

We start in Section 6.2 with notation, definitions, and a further introduction to BSDEs, where also the link with semilinear partial differential equations is stated. A general time discretization of the BSDE results in expressions with conditional expectations (Section 6.3). These conditional expectations are computed by the BCOS method (Section 6.4) and the problem is then solved backwards in time. We perform extensive numerical experiments in Section 6.5. Then, in Section 6.6, utility indifference pricing and the related maximization problems are discussed. We derive a numerical scheme for the resulting BSDE with jumps in Section 6.6.3. Results in Section 6.7 show the utility indifference ask and bid prices.

6.2. BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

We start with some notation and definitions, for which we follow the survey paper [EKPQ97]. Let $\omega = (\omega_t)_{0\le t\le T}$ be a standard one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ the natural filtration of the Brownian motion ω, and T a fixed finite time horizon. We denote by $\mathbb{H}_T^2(\mathbb{R})$ the set of predictable processes $\eta : \Omega \times [0,T] \to \mathbb{R}$ such that $\mathbb{E}\left[\int_0^T |\eta_t|^2 dt\right] < \infty$ and by $\mathbb{L}_T^2(\mathbb{R})$ the set of $\mathcal{F}_T$-measurable random variables $X : \Omega \to \mathbb{R}$ that are square integrable. We consider the BSDE

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t d\omega_t, \quad Y_T = \xi, \tag{6.2.1}$$

where function $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is $\mathscr{P} \otimes \mathscr{B} \otimes \mathscr{B}$-measurable. $\mathscr{P}$ is the set of $\mathcal{F}_t$-progressively measurable scalar processes on $\Omega \times [0,T]$. $f(.)$ is the *generator* or *driver* of the process and the *terminal condition* $\xi : \Omega \to \mathbb{R}$ is an $\mathcal{F}_T$-measurable random variable. For simplicity we use one-dimensional processes, but the BSDE theory can be extended to higher dimensions, with d-dimensional processes ω_t and Y_t and an $n \times d$-dimensional Z_t process, as described in [EKPQ97].

A solution to BSDE (6.2.1) is given by a pair of processes (Y, Z), with Y a continuous real-valued adapted process and Z a real-valued predictable process satisfying $\int_0^T |Z_t|^2 dt < \infty, \mathbb{P}$ a.s., satisfying

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s d\omega_s, \quad 0 \le t \le T. \tag{6.2.2}$$

Unlike an FSDE, the solution of a BSDE is thus a *pair* of adapted processes (Y, Z). Note that

BSDEs cannot be considered as time-reversed FSDEs, because at time t the pair (Y_t, Z_t) is $\mathcal{F}_t$-measurable and the process does not yet 'know' the terminal condition.

Function f and terminal condition ξ are called *standard parameters* for the BSDE, if $\xi \in \mathbb{L}^2_T(\mathbb{R})$, $f(.,0,0) \in \mathbb{H}^2_T(\mathbb{R})$, and f is uniformly Lipschitz in y and z, with Lipschitz constant L_f. A result from [EKPQ97, Pha09, PP90] is that, given a pair of standard parameters (f,ξ), there exists a unique solution $(Y,Z) \in \mathbb{H}^2_T(\mathbb{R}) \times \mathbb{H}^2_T(\mathbb{R})$ to BSDE (6.2.1).

Markovian Case for the BSDE A *linear* parabolic PDE has a probabilistic representation by means of the Feynman-Kac theorem. Here, we consider a *semilinear* parabolic PDE of the form

$$\frac{\partial v}{\partial t}(t,x) + \mathscr{L} v(t,x) + f(t,x,v(t,x),\sigma(t,x)D_x v(t,x)) = 0, \quad (t,x) \in [0,T) \times \mathbb{R} \tag{6.2.3a}$$

$$v(T,x) = g(x), \quad x \in \mathbb{R}, \tag{6.2.3b}$$

with the differential operator of second order

$$\mathscr{L} v(t,x) = \mu(t,x) D_x v(t,x) + \tfrac{1}{2}\sigma^2(t,x) D_x^2 v(t,x). \tag{6.2.4}$$

D_x and D_x^2 denote the first and second derivative of a function with respect to the x-variable, respectively.

This PDE also has a probabilistic representation, by means of the following FSDE,

$$X_t = x, \quad dX_s = \mu(s,X_s)ds + \sigma(s,X_s)d\omega_s, \quad t \le s \le T. \tag{6.2.5}$$

and BSDE

$$-dY_s = f(s, X_s^{t,x}, Y_s, Z_s)ds - Z_s d\omega_s, \quad Y_T = g(X_T^{t,x}), \tag{6.2.6}$$

whose terminal condition is determined by the terminal value of FSDE (6.2.5). $X_s^{t,x}$ denotes the solution to (6.2.5) starting from x at time t, and $(Y_s^{t,x}, Z_s^{t,x})$ is the corresponding solution to the BSDE.

The coefficients $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\mu : [0,T] \times \mathbb{R} \to \mathbb{R}$ in (6.2.5) are assumed to be Lipschitz in x and satisfy a linear growth condition in x. Functions $f : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are assumed to be uniformly continuous with respect to x. Moreover, f satisfies a Lipschitz condition in (y,z) and there exists a constant C such that $|f(t,x,y,z)| + |g(x)| \le C(1+|x|^p + |y| + |z|)$, $p \ge 1/2$.

The conditions on f and ξ guarantee the existence of a unique solution (Y,Z) to the BSDE (6.2.6). Together with the Markov property of the process X, we notice that there exists a deterministic function $v(t,x)$ such that the solution Y of the BSDE is $Y_s^{t,x} = v(s, X_s^{t,x})$, $t \le s \le T$. The solution of the BSDE is said to be Markovian as it can be written as a function of time and the state process $X_s^{t,x}$. The following results hold

Result 6.2.1. *([PP92, Pha09]) Let $v \in C^{1,2}$ be a classical solution to* (6.2.3) *and suppose there exists a constant $C \ge 0$ such that, for all (t,x), $|v(t,x)| + |\sigma(t,x)D_x v(t,x)| \le C(1+|x|)$. Then the pair (Y,Z), defined by*

$$Y_s^{t,x} = v(s, X_s^{t,x}), \quad Z_s^{t,x} = \sigma(s, X_s^{t,x}) D_x v(s, X_s^{t,x}), \quad t \le s \le T, \tag{6.2.7}$$

is the solution to BSDE (6.2.6) *(a so-called verification result).*

The converse result states: Suppose (Y,Z) is the solution to the BSDE, then the function defined by $v(t,x) = Y_t^{t,x}$ is a viscosity solution to the PDE.

The verification result follows from application of Itô's lemma to $v(t,X_t)$ ([Pha09]):

$$\begin{aligned} dv(t,X_t) &= (v_t(t,X_t) + \mathscr{L}v(t,X_t))\,dt + \sigma(t,X_t)D_x v(t,X_t)d\omega_t \\ &= -f(t,X_t,v(t,X_t),\sigma(t,X_t)D_x v(t,X_t))dt + \sigma(t,X_t)D_x v(t,X_t)d\omega_t. \end{aligned} \tag{6.2.8}$$

So, solving the semilinear PDE or the corresponding BSDE results in the same solution. A PDE can be solved by applying numerical discretization techniques and for BSDEs probabilistic numerical methods are available. For example, Picard methods for Y, see [BD07, GL10], give rise to a sequence of 'easy' linear BSDEs. Another class of methods focuses on dynamic programming equations, see [BT04, Zha04, GT13, CM12]. Our probabilistic solution method to the BSDE is in this class and consists of two steps: First of all, the FSDE is simulated by a discretization scheme and the general θ-time discretization of the BSDE then results in expressions with conditional expectations (see Section 6.3). Secondly, the conditional expectations are computed by the BCOS method (see Section 6.4) and the problem is solved backwards in time.

6.3. DISCRETIZATION OF THE BSDE

We wish to discretize the forward stochastic process,

$$X_0 = x_0 \text{ given}, \quad X_t = X_0 + \int_0^t \mu(s,X_s)ds + \int_0^t \sigma(s,X_s)d\omega_s, \tag{6.3.1}$$

and the backward process

$$Y_t = \xi + \int_t^T f(s,\mathbf{\Lambda}_s)ds - \int_t^T Z_s d\omega_s, \quad \xi = g(X_T), \tag{6.3.2}$$

with $\mathbf{\Lambda}_s := (X_s, Y_s, Z_s)$. For this, we define a partition $\Delta : 0 = t_0 < t_1 < t_2 < \ldots < t_m < \ldots < t_M = T$, with fixed time steps $\Delta t := t_{m+1} - t_m$. For notational convenience we write $X_m = X_{t_m}$, $Y_m = Y_{t_m}$, $Z_m = Z_{t_m}$ and define $\Delta\omega_{m+1} := \omega_{t_{m+1}} - \omega_{t_m}$. With ω_t a Wiener process, the increments $\Delta\omega_{m+1} \sim \mathcal{N}(0,\Delta t)$ are normally distributed. The classical Euler discretization X^Δ of the FSDE reads, in this case,

$$X_0^\Delta = x_0, \quad X_{m+1}^\Delta = X_m^\Delta + \mu(t_m,X_m^\Delta)\Delta t + \sigma(t_m,X_m^\Delta)\Delta\omega_{m+1}, \quad m = 0,\ldots,M-1. \tag{6.3.3}$$

For the BSDE, we then start with

$$Y_m = Y_{m+1} + \int_{t_m}^{t_{m+1}} f(s,\mathbf{\Lambda}_s)ds - \int_{t_m}^{t_{m+1}} Z_s d\omega_s. \tag{6.3.4}$$

By a basic Euler discretization, backwards in time, we would require the unknown value Y_{m+1} to approximate Y_m. This scheme does hence not suffice, as it would not take into account the adaptability constraints on Y and Z. To obtain a computationally viable backward induction scheme we should take conditional expectations, which will result in a similar approximation scheme to the BSDE as used in [ZWP09]. For the $\mathcal{F}_{t_m}$-measurable random variables Y_m and Z_m it holds that $\mathbb{E}_m[Y_m] = Y_m$ and $\mathbb{E}_m[Z_m] = Z_m$, where $\mathbb{E}_m[.]$ represents the conditional expectation $\mathbb{E}[.|\mathcal{F}_{t_m}]$. Taking conditional expectations at both sides of equation (6.3.4) then results in

$$\begin{aligned} Y_m &= \mathbb{E}_m[Y_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m[f(s,\mathbf{\Lambda}_s)]ds \\ &\approx \mathbb{E}_m[Y_{m+1}] + \Delta t\theta_1 f(t_m,\mathbf{\Lambda}_m) + \Delta t(1-\theta_1)\mathbb{E}_m[f(t_{m+1},\mathbf{\Lambda}_{m+1})], \quad \theta_1 \in [0,1]. \end{aligned} \tag{6.3.5}$$

The integrand in the above equation (6.3.5) is a deterministic continuous function of time s, so that we can use the well-known θ-time discretization method to approximate the integral [Ise09]. Multiplying both sides of equation (6.3.4) by $\Delta\omega_{m+1}$, taking the conditional expectation, and applying the θ-method also gives us

$$\begin{aligned}0 &= \mathbb{E}_m[Y_{m+1}\Delta\omega_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m[f(s,\mathbf{\Lambda}_s)(\omega_s-\omega_{t_m})]ds - \int_{t_m}^{t_{m+1}} \mathbb{E}_m[Z_s]ds \\ &\approx \mathbb{E}_m[Y_{m+1}\Delta\omega_{m+1}] + \Delta t(1-\theta_2)\mathbb{E}_m[f(t_{m+1},\mathbf{\Lambda}_{m+1})\Delta\omega_{m+1}] \\ &\quad - \Delta t\theta_2 Z_m - \Delta t(1-\theta_2)\mathbb{E}_m[Z_{m+1}], \quad \theta_2 \in [0,1]. \end{aligned} \tag{6.3.6}$$

Note that for equations (6.3.5) and (6.3.6) we use two different time discretization parameters θ_1 and θ_2, respectively. The above equations lead to a discrete-time approximation (Y^Δ, Z^Δ) for (Y,Z):

$$Y_M^\Delta = g(X_M^\Delta), \quad Z_M^\Delta = \sigma(t_M, X_M^\Delta)D_x g(X_M^\Delta), \tag{6.3.7a}$$

for $m = M-1,\dots,0$:

$$\begin{aligned} Z_m^\Delta &= -\theta_2^{-1}(1-\theta_2)\mathbb{E}_m[Z_{m+1}^\Delta] + \tfrac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_m[Y_{m+1}^\Delta\Delta\omega_{m+1}] \\ &\quad + \theta_2^{-1}(1-\theta_2)\mathbb{E}_m[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)\Delta\omega_{m+1}], \end{aligned} \tag{6.3.7b}$$

$$Y_m^\Delta = \mathbb{E}_m[Y_{m+1}^\Delta] + \Delta t\theta_1 f(t_m,\mathbf{\Lambda}_m^\Delta) + \Delta t(1-\theta_1)\mathbb{E}_m[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)]. \tag{6.3.7c}$$

The use of $\theta_1 = 0$ gives us an *explicit* scheme for Y_m^Δ, whereas $\theta_1 \in (0,1]$ results in an *implicit* scheme. To solve for Z_m^Δ, we should have obviously $\theta_2 \neq 0$ in equation (6.3.7b), which gives an explicit scheme for Z_m^Δ. For the terminal value Z_M^Δ we use the relation from Result 6.2.1. At the points where g is not continuously differentiable, we consider a one-sided derivative[1].

The terminal condition is a deterministic function of X_M^Δ and X^Δ is a Markov process. Then it is easily seen, using an induction argument, that there are deterministic functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$, so that

$$Y_m^\Delta = y_m^\Delta(X_m^\Delta), \quad Z_m^\Delta = z_m^\Delta(X_m^\Delta). \tag{6.3.8}$$

So, the random variables Y_m^Δ and Z_m^Δ are functions of X_m^Δ and the conditional expectations can be replaced by $\mathbb{E}_m^x[.] \equiv \mathbb{E}[.|X_m^\Delta = x]$. Note that functions y_m^Δ and z_m^Δ depend on the discretization partition Δ.

Equations (6.3.7) provide us with a scheme to solve the BSDE backwards in time, starting at terminal time T. One could use least squares Monte Carlo methods, like the Longstaff-Schwartz method or stochastic grid method [JO12], to approximate the conditional expectations, see for example [LGW06, GLW05, BS12]. The authors of [BT04] apply a Malliavin-based algorithm to solve them, whereas [MPSMT02] employs a binomial tree method. In the next section, we introduce a Fourier method to solve the BSDE.

6.4. BCOS METHOD

In this section we explain our method of choice to compute the conditional expectations in (6.3.7) and solve the problem recursively, backwards in time. Our method is an extension

[1] It is also possible to take $\theta_1 = \theta_2 = 1$ in the first iteration with time step $(\Delta t)^2$, which gives the same convergence results.

of the COS method. The COS method for computing Bermudan options also consists of a backwards-in-time scheme to find the conditional expectations of the continuation value, see [FO09]. The method for solving BSDEs with a COS method is named the *BCOS method* here. First, in Section 6.4.1 we derive the COS formulas and define the Fourier cosine coefficients. Then, Sections 6.4.2 and 6.4.3 are devoted to the approximation of functions z_m^Δ and y_m^Δ. Section 6.4.4 discusses the recursive recovery of the Fourier coefficients and Section 6.4.5 the error components.

6.4.1. COS FORMULAS AND FOURIER COSINE COEFFICIENTS

Suppose we wish to approximate the expectation

$$\mathbb{E}_m^x[v(t_{m+1}, X_{m+1}^\Delta)] = \int_{\mathbb{R}} v(t_{m+1},\zeta)\, p(\zeta|x)\, d\zeta, \tag{6.4.1}$$

where v represents a general functional and $p(\zeta|x) = \mathbb{P}(X_{m+1}^\Delta = \zeta | X_m^\Delta = x)$ denotes the continuous transitional density function. The *COS formula* (see Section 2.3.1) gives the approximation

$$\mathbb{E}_m^x[v(t_{m+1}, X_{m+1}^\Delta)] \approx \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1}) \Re\left\{\phi\left(\tfrac{k\pi}{b-a}\middle|x\right) e^{ik\pi\frac{x-a}{b-a}}\right\} = \tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1})\Phi_k(x). \tag{6.4.2}$$

Here the Fourier cosine series coefficients of function $v(t_{m+1},\zeta)$ are given by

$$\mathcal{V}_k(t_{m+1}) := \tfrac{2}{b-a}\int_a^b v(t_{m+1},\zeta)\cos\left(k\pi\tfrac{\zeta-a}{b-a}\right) d\zeta. \tag{6.4.3}$$

The Fourier cosine coefficients of the transitional density function $p(\zeta|x)$ are approximated as follows

$$\begin{aligned}\mathscr{P}_k(x) &\approx \tfrac{2}{b-a}\int_{\mathbb{R}} p(\zeta|x)\cos\left(k\pi\tfrac{\zeta-a}{b-a}\right) d\zeta = \tfrac{2}{b-a}\Re\left\{\varphi\left(\tfrac{k\pi}{b-a}\middle|x\right) e^{ik\pi\frac{-a}{b-a}}\right\} \\ &= \tfrac{2}{b-a}\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\middle|x\right) e^{ik\pi\frac{x-a}{b-a}}\right\} := \Phi_k(x).\end{aligned} \tag{6.4.4}$$

$\Re\{.\}$ again denotes taking the real part of the input argument and $\varphi(.|x)$ is the conditional *characteristic function* of X_{m+1}^Δ, given $X_m^\Delta = x$. The characteristic function encountered here is written as

$$\begin{aligned}&\varphi(u|x) = \varphi(u|0)e^{iux} = \phi(u|x)e^{iux}, \\ \text{with}\quad &\phi(u|x) := \exp\left(iu\mu(t_m,x)\Delta t - \tfrac{1}{2}u^2\sigma^2(t_m,x)\Delta t\right).\end{aligned} \tag{6.4.5}$$

For solving the BSDE we need to deal with expectations of the form $\mathbb{E}_m^x[v(t_{m+1}, X_{m+1}^\Delta)\Delta\omega_{m+1}]$. With the help of the equality (6.A.3) in Appendix 6.A.I, they

can be computed by

$$\begin{aligned}\mathbb{E}_m^x[v(t_{m+1},X_{m+1}^\Delta)\Delta\omega_{m+1}] &\approx \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1})\mathbb{E}_m^x\left[\cos\left(k\pi\tfrac{X_{m+1}^\Delta-a}{b-a}\right)\Delta\omega_{m+1}\right]\\ &= \sigma(t_m,x)\Delta t\sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1})\mathbb{E}_m^x\left[\tfrac{-k\pi}{b-a}\sin\left(k\pi\tfrac{X_{m+1}^\Delta-a}{b-a}\right)\right]\\ &\approx \sigma(t_m,x)\Delta t\sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1})\Re\left\{i\tfrac{k\pi}{b-a}\phi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{x-a}{b-a}}\right\}\\ &:= \sigma(t_m,x)\Delta t\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1})\Phi_k'(x). \end{aligned} \tag{6.4.6}$$

Now we return to the BSDE problem (6.3.7), where we defined the deterministic functions $y_m^\Delta(X_m^\Delta)=Y_m^\Delta$ and $z_m^\Delta(X_m^\Delta)=Z_m^\Delta$. Let $\mathcal{Y}_k^\Delta(t_{m+1})$ be the Fourier cosine coefficients of $y_{m+1}^\Delta(x)$ in (6.3.7c), i.e.,

$$\mathcal{Y}_k^\Delta(t_{m+1}) = \tfrac{2}{b-a}\int_a^b y_{m+1}^\Delta(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \tag{6.4.7}$$

$\mathcal{Z}_k^\Delta(t_{m+1})$ the Fourier cosine coefficients of function $z_{m+1}^\Delta(x)$ in (6.3.7b), i.e.,

$$\mathcal{Z}_k^\Delta(t_{m+1}) = \tfrac{2}{b-a}\int_a^b z_{m+1}^\Delta(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \tag{6.4.8}$$

and $\mathcal{F}_k^\Delta(t_{m+1})$ the Fourier cosine coefficients of driver $f(t_{m+1},x,y_{m+1}^\Delta(x),z_{m+1}^\Delta(x))$, i.e.,

$$\mathcal{F}_k^\Delta(t_{m+1}) = \tfrac{2}{b-a}\int_a^b f(t_{m+1},x,y_{m+1}^\Delta(x),z_{m+1}^\Delta(x))\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx. \tag{6.4.9}$$

In Sections 6.4.2 and 6.4.3, we will assume that the above coefficients are given. In Section 6.4.4 the algorithm to recover these coefficients recursively, backwards in time, will be discussed.

6

6.4.2. COS APPROXIMATION OF FUNCTION $z_m^\Delta(x)$

For the computation of $z_m^\Delta(x)$ in (6.3.7b), we need to compute three expectations, $\mathbb{E}_m^x[Z_{m+1}^\Delta]$, $\mathbb{E}_m^x[Y_{m+1}^\Delta\Delta\omega_{m+1}]$, and $\mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)\Delta\omega_{m+1}]$. With the help of COS formulas we can derive the following approximations for these expectations:

$$\mathbb{E}_m^x[Z_{m+1}^\Delta] \approx \sum_{k=0}^{N-1}{}' \mathcal{Z}_k^\Delta(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{x-a}{b-a}}\right\}, \tag{6.4.10a}$$

$$\mathbb{E}_m^x[Y_{m+1}^\Delta\Delta\omega_{m+1}] \approx \sum_{k=0}^{N-1}{}' \mathcal{Y}_k^\Delta(t_{m+1})\sigma(t_m,x)\Delta t\Re\left\{\tfrac{ik\pi}{b-a}\phi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{x-a}{b-a}}\right\}, \tag{6.4.10b}$$

$$\mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)\Delta\omega_{m+1}] \approx \sum_{k=0}^{N-1}{}' \mathcal{F}_k^\Delta(t_{m+1})\sigma(t_m,x)\Delta t\Re\left\{\tfrac{ik\pi}{b-a}\phi\left(\tfrac{k\pi}{b-a}\middle|x\right)e^{ik\pi\frac{x-a}{b-a}}\right\}. \tag{6.4.10c}$$

We then find as COS approximation

$$z_m^\Delta(x) \approx -\tfrac{1-\theta_2}{\theta_2}\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \mathcal{Z}_k^\Delta(t_{m+1})\Phi_k(x) + \tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \left(\tfrac{1}{\Delta t\theta_2}\mathcal{Y}_k^\Delta(t_{m+1}) + \tfrac{1-\theta_2}{\theta_2}\mathcal{F}_k^\Delta(t_{m+1})\right)\sigma(t_m,x)\Delta t\Phi_k'(x), \tag{6.4.11}$$

with Φ_k and Φ_k' as defined in (6.4.4) and (6.4.6), respectively.

6.4.3. COS APPROXIMATION OF FUNCTION $y_m^\Delta(x)$

For the computation of function $y_m^\Delta(x)$ in equation (6.3.7c) there are two explicit parts, $\mathbb{E}_m^x[Y_{m+1}^\Delta]$ and $\mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)]$, that are approximated by the following COS formulas

$$\mathbb{E}_m^x[Y_{m+1}^\Delta] \approx \sum_{k=0}^{N-1}{}' \mathcal{Y}_k^\Delta(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}, \tag{6.4.12a}$$

$$\mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)] \approx \sum_{k=0}^{N-1}{}' \mathcal{F}_k^\Delta(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}. \tag{6.4.12b}$$

Besides, when $\theta_1 > 0$, we also have an implicit part, for which we define

$$h_m^\Delta(x) := \mathbb{E}_m^x[Y_{m+1}^\Delta] + \Delta t(1-\theta_1)\mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^\Delta)] \approx \tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \mathcal{Y}_k^\Delta(t_{m+1})\Phi_k(x) + \Delta t(1-\theta_1)\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}' \mathcal{F}_k^\Delta(t_{m+1})\Phi_k(x), \tag{6.4.13}$$

with Φ_k from (6.4.4). Now we can write

$$y_m^\Delta(x) = \Delta t\theta_1 f(t_m,x,y_m^\Delta(x),z_m^\Delta(x)) + h_m^\Delta(x). \tag{6.4.14}$$

In order to determine function $y_m^\Delta(x)$ in equation (6.4.14), we will perform *P Picard iterations* (see also [GLW05]), starting with an initial guess, $y_m^{\Delta,0}(x) := \mathbb{E}_m^x[Y_{m+1}^\Delta]$ (see equation (6.4.12a)). The convergence properties of the Picard iterations to the 'true' values $y_m^\Delta(x)$ are discussed in Section 6.4.5.

6.4.4. RECOVERY OF COEFFICIENTS AND ALGORITHM

The computation of functions $z_m^\Delta(x)$ and $y_m^\Delta(x)$ at time-point t_m requires the Fourier cosine coefficients $\mathcal{Z}_k^\Delta(t_{m+1})$, $\mathcal{Y}_k^\Delta(t_{m+1})$, and $\mathcal{F}_k^\Delta(t_{m+1})$ at time-point t_{m+1}. For the next time step in the BCOS method we wish to compute functions $z_{m-1}^\Delta(x)$ and $y_{m-1}^\Delta(x)$ at time-point t_{m-1}, for which we need the Fourier cosine coefficients of time-point t_m. The coefficients can be computed recursively backwards in time, as we explain in this section.

We assume a constant drift μ and volatility σ here, and

$$X_{m+1}^\Delta = X_m^\Delta + \mu\Delta t + \sigma\Delta\omega_{m+1}. \tag{6.4.15}$$

Now, function $\phi(u)$ does not depend on x. In Remark 6.4.1 we will comment on the use of more general functions $\mu(t,x)$ and $\sigma(t,x)$.

Firstly, the computation of the coefficients $\mathcal{Z}_k^{\Delta}(t_m)$ can be divided into three parts, similar as equation (6.4.10). We then use the following approximations

$$\frac{2}{b-a}\int_a^b \mathbb{E}_m^x[Z_{m+1}^{\Delta}]\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx \approx \Re\left\{\sum_{j=0}^{N-1}{}' \mathcal{Z}_j^{\Delta}(t_{m+1})\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}, \tag{6.4.16a}$$

$$\frac{2}{b-a}\int_a^b \mathbb{E}_m^x[Y_{m+1}^{\Delta}\Delta\omega_{m+1}]\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx \approx \Re\left\{\sum_{j=0}^{N-1}{}' \tfrac{ij\pi}{b-a}\sigma\Delta t\mathcal{Y}_j^{\Delta}(t_{m+1})\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}, \tag{6.4.16b}$$

$$\frac{2}{b-a}\int_a^b \mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1}^{\Delta})\Delta\omega_{m+1}]\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx \approx \Re\left\{\sum_{j=0}^{N-1}{}' \tfrac{ij\pi}{b-a}\sigma\Delta t\mathcal{F}_j^{\Delta}(t_{m+1})\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}, \tag{6.4.16c}$$

with matrix elements

$$\mathcal{M}_{k,j} := \tfrac{2}{b-a}\int_a^b e^{ij\pi\frac{x-a}{b-a}}\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx. \tag{6.4.17}$$

These approximations can be found by inserting COS formulas (6.4.10). Note that the approximation signs "≈" are due to the errors of the COS formulas, i.e., truncation of the integration range to a finite interval $[a,b]$, truncation of the infinite sums to a finite number of terms N, and the substitution of the series coefficients by the characteristic function approximation. The coefficients $\mathcal{Z}_k^{\Delta}(t_m)$ are then computed as follows

$$\mathcal{Z}_k^{\Delta}(t_m) \approx \Re\left\{\sum_{j=0}^{N-1}{}' \left[-\tfrac{1-\theta_2}{\theta_2}\mathcal{Z}_j^{\Delta}(t_{m+1}) + \tfrac{ij\pi}{b-a}\sigma\Delta t\left(\tfrac{1}{\Delta t\theta_2}\mathcal{Y}_j^{\Delta}(t_{m+1}) + \tfrac{1-\theta_2}{\theta_2}\mathcal{F}_j^{\Delta}(t_{m+1})\right)\right]\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}. \tag{6.4.18}$$

Secondly, the coefficients $\mathcal{H}_k^{\Delta}(t_m)$ of function $h_m^{\Delta}(x)$ in (6.4.13) are computed by

$$\mathcal{H}_k^{\Delta}(t_m) = \tfrac{2}{b-a}\int_a^b h_m^{\Delta}(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx \approx \Re\left\{\sum_{j=0}^{N-1}{}' \left[\mathcal{Y}_j^{\Delta}(t_{m+1}) + \Delta t(1-\theta_1)\mathcal{F}_j^{\Delta}(t_{m+1})\right]\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}. \tag{6.4.19}$$

The Fourier cosine coefficients $\mathcal{Z}_k^{\Delta}(t_m)$ in (6.4.18) and $\mathcal{H}_k^{\Delta}(t_m)$ in (6.4.19), for $k = 0,1,\ldots,N-1$, can thus be computed by one matrix-vector multiplication. These matrix-vector multiplications $\mathcal{M}\mathbf{w}$ can be done efficiently with the use of an FFT algorithm, see Appendix B. With this the computational complexity is reduced from order $\mathcal{O}(N^2)$ to $\mathcal{O}(N\log N)$, with N the number of terms in the summations.

At last, the coefficients $\mathcal{F}_k^{\Delta,P-1}(t_m)$ of function $f(t_m,x,y_m^{\Delta,P-1}(x),z_m^{\Delta}(x))$ are given by

$$\mathcal{F}_k^{\Delta,P-1}(t_m) := \tfrac{2}{b-a}\int_a^b f(t_m,x,y_m^{\Delta,P-1}(x),z_m^{\Delta}(x))\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx. \tag{6.4.20}$$

They are approximated by a *discrete Fourier cosine transform (DCT)*. For this we need to compute the integrand $f(t_m, x, y_m^{\Delta,P-1}(x), z_m^{\Delta}(x))$ on an *equidistant x-grid* with N grid points, as explained in Appendix C. With a converging Picard method, we have $\mathscr{F}_k^{\Delta}(t_m) \approx \mathscr{F}_k^{\Delta,P-1}(t_m)$ for sufficiently many iterations P. Then,

$$\mathscr{Y}_k^{\Delta}(t_m) \approx \Delta t \theta_1 \mathscr{F}_k^{\Delta,P-1}(t_m) + \mathscr{H}_k^{\Delta}(t_m). \tag{6.4.21}$$

With the aforementioned formulas we approximate the Fourier cosine coefficients $\mathscr{Z}_k^{\Delta}(t_m)$, $\mathscr{Y}_k^{\Delta}(t_m)$, and $\mathscr{F}_k^{\Delta}(t_m)$ by using the coefficients of time-point t_{m+1}. Starting with the coefficients at the terminal time we can solve them recursively, backwards in time. The evolution of the extra error introduced by approximation of the coefficients has been discussed in detail in Section 2.4.3. The final approximations of the functions $y_m^{\Delta}(x)$ and $z_m^{\Delta}(x)$ by the BCOS method are denoted by $\hat{y}_m^{\Delta}(x)$ and $\hat{z}_m^{\Delta}(x)$, respectively. The overall algorithm to solve the BSDE (6.3.7) backwards in time can be summarized as:

Algorithm 3. *(BCOS method)*
Initialization: *Compute, or approximate, the terminal coefficients* $\mathscr{Y}_k^{\Delta}(t_M)$, $\mathscr{Z}_k^{\Delta}(t_M)$, *and* $\mathscr{F}_k^{\Delta}(t_M)$.
Main loop to recover $\mathscr{Z}_k^{\Delta}(t_m)$, $\mathscr{F}_k^{\Delta}(t_m)$, ***and*** $\mathscr{Y}_k^{\Delta}(t_m)$***:***
For $m = M-1$ *to* 1*:*

- *Compute the functions* $\hat{z}_m^{\Delta}(x)$, $f(t_m, x, \hat{y}_m^{\Delta}(x), \hat{z}_m^{\Delta}(x))$, *and* $\hat{y}_m^{\Delta}(x)$.
- *Determine the corresponding Fourier cosine coefficients* $\mathscr{Z}_k^{\Delta}(t_m)$, $\mathscr{F}_k^{\Delta}(t_m)$, *and* $\mathscr{Y}_k^{\Delta}(t_m)$, *as described in Sections 6.4.2, 6.4.3, and 6.4.4.*

Final step: *Compute* $\hat{z}_0^{\Delta}(x_0)$ *and* $\hat{y}_0^{\Delta}(x_0)$.

Remark 6.4.1. *For general drift* $\mu(t,x)$ *and volatility* $\sigma(t,x)$ *in* (6.4.15) *we need to compute the following integrals to recover the Fourier cosine coefficients*

$$\tfrac{2}{b-a}\int_a^b \phi\left(\tfrac{j\pi}{b-a}\middle| x\right) e^{ij\pi\frac{x-a}{b-a}} \cos\left(k\pi\tfrac{x-a}{b-a}\right) dx, \tag{6.4.22}$$

which is not equal to $\phi\left(\frac{j\pi}{b-a}\right)\mathscr{M}_{k,j}$ *(as in equations* (6.4.16)*). As the integration kernel is smooth, we can approximate the integrals efficiently by, for example, a Clenshaw-Curtis quadrature rule [Gen72]. Another way is to approximate the coefficients* $\mathscr{Z}_k^{\Delta}(t_m)$ *by using a DCT too.*

The Euler discretization for general drift and volatility terms exhibits only first order weak convergence, which may hamper the convergence of the discretized BSDE. The usage of the simplified second order weak Taylor scheme may improve the convergence rate and for some processes one can use an exact simulation scheme[2].

6.4.5. ERRORS AND COMPUTATIONAL COMPLEXITY

In the BCOS method when solving BSDEs several approximation errors are encountered. In the first place there are discretization errors, due to the discrete-time approximation of the

[2] This is discussed in Chapter 7.

stochastic processes. Moreover, errors are introduced by the COS formulas and the Picard method. These error components and the computational complexity are discussed in this section.

Discretization Error of the BSDE We perform an error analysis[3] for the scheme with $\theta_1 = \theta_2 = \frac{1}{2}$ and *assume constant* μ *and* σ (see equation (6.4.15)), so that $X_m^\Delta = X_m$. We define the *local* θ-discretization errors in equations (6.3.5) and (6.3.6) by

$$R_m^y(x) := \int_{t_m}^{t_{m+1}} \mathbb{E}_m^x[f(s,\mathbf{\Lambda}_s)]ds - \tfrac{1}{2}\Delta t f(t_m,\mathbf{\Lambda}_m) - \tfrac{1}{2}\Delta t \mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1})], \tag{6.4.23a}$$

$$\begin{aligned} R_m^z(x) := &\int_{t_m}^{t_{m+1}} \mathbb{E}_m^x[f(s,\mathbf{\Lambda}_s)(\omega_s-\omega_{t_m})]ds - \tfrac{1}{2}\Delta t \mathbb{E}_m^x[f(t_{m+1},\mathbf{\Lambda}_{m+1})\Delta\omega_{m+1}] \\ &- \int_{t_m}^{t_{m+1}} \mathbb{E}_m^x[Z_s]ds + \tfrac{1}{2}\Delta t Z_m + \tfrac{1}{2}\Delta t \mathbb{E}_m^x[Z_{m+1}]. \end{aligned} \tag{6.4.23b}$$

The orders of these errors depend on the smoothness of the integrands with respect to time s. If functions f and g are sufficiently smooth and bounded, with bounded derivatives, then the absolute values of the terms $R_m^y(x)$, $R_m^z(x)$, $\frac{1}{\Delta t}\mathbb{E}_m^x[R_{m+1}^y(X_{m+1})\Delta\omega_{m+1}]$ and $(\frac{1}{\Delta t}R_m^z(x) - \mathbb{E}_m^x[R_{m+1}^z(X_{m+1})])$ can be bounded by $C(\Delta t)^3$, with C a constant only depending on T, functions g and f, and μ, σ (similar as in Section 7.4.3).

The *global* errors due to the θ-time discretization in (6.3.7c) and (6.3.7b) are denoted by

$$\begin{aligned} &\epsilon_m^y(X_m) := y_m(X_m) - y_m^\Delta(X_m), \quad \epsilon_m^z(X_m) := z_m(X_m) - z_m^\Delta(X_m), \\ &\epsilon_m^f(X_m) := f(t_m,\mathbf{\Lambda}_m) - f(t_m,\mathbf{\Lambda}_m^\Delta). \end{aligned} \tag{6.4.24}$$

We omit the dependency of the local and global errors on the state of the FSDE for notational convenience. For the y-component we have ($m \le M-1$)

$$\epsilon_m^y = \mathbb{E}_m^x[\epsilon_{m+1}^y] + \tfrac{1}{2}\Delta t \epsilon_m^f + \tfrac{1}{2}\Delta t \mathbb{E}_m^x[\epsilon_{m+1}^f] + R_m^y. \tag{6.4.25}$$

With the Lipschitz assumption on driver function f, this error can be bounded, for $\frac{1}{2}\Delta t L_f < 1$, by

$$\begin{aligned} |\epsilon_m^y| \le &\frac{1+\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}\mathbb{E}_m^x[|\epsilon_{m+1}^y|] + \frac{\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}|\epsilon_m^z| + \frac{\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}\mathbb{E}_m^x[|\epsilon_{m+1}^z|] \\ &+ \frac{1}{1-\frac{1}{2}\Delta t L_f}C(\Delta t)^3. \end{aligned} \tag{6.4.26}$$

For the z-component we have

$$\epsilon_m^z = \tfrac{2}{\Delta t}\mathbb{E}_m^x[\epsilon_{m+1}^y\Delta\omega_{m+1}] + \mathbb{E}_m^x[\epsilon_{m+1}^f\Delta\omega_{m+1}] - \mathbb{E}_m^x[\epsilon_{m+1}^z] + \tfrac{2}{\Delta t}R_m^z. \tag{6.4.27}$$

Substituting the similar equations for ϵ_{m+1}^y and ϵ_{m+1}^z as in (6.4.25) and (6.4.27) gives ($m \le M-2$)

$$\begin{aligned} \epsilon_m^z = &\tfrac{2}{\Delta t}\mathbb{E}_m^x[\epsilon_{m+2}^y\Delta\omega_{m+1}] + \mathbb{E}_m^x[\epsilon_{m+1}^f\Delta\omega_{m+1}] + \mathbb{E}_m^x[\epsilon_{m+2}^f\Delta\omega_{m+1}] \\ &+ \mathbb{E}_m^x[\epsilon_{m+1}^f\Delta\omega_{m+1}] - \tfrac{2}{\Delta t}\mathbb{E}_m^x[\epsilon_{m+2}^y\Delta\omega_{m+2}] - \mathbb{E}_m^x[\epsilon_{m+2}^f\Delta\omega_{m+2}] + \mathbb{E}_m^x[\epsilon_{m+2}^z] \\ &+ \tfrac{2}{\Delta t}\mathbb{E}_m^x[R_{m+1}^y\Delta\omega_{m+1}] - \tfrac{2}{\Delta t}\mathbb{E}_m^x[R_{m+1}^z] + \tfrac{2}{\Delta t}R_m^z. \end{aligned} \tag{6.4.28}$$

[3] The error analyses for other processes and other discretization schemes for the FSDE, such as the Milstein scheme, are discussed in Chapter 7.

Error ϵ^y_{m+2} is a function of X_{m+2}. The equalities (6.A.4) and (6.A.5) in Appendix 6.A.I then give us

$$\begin{aligned}\mathbb{E}^x_m[\epsilon^y_{m+2}\Delta\omega_{m+1}] &= \mathbb{E}^x_m[y_{m+2}(X_{m+2})\Delta\omega_{m+1} - y^\Delta_{m+2}(X_{m+2})\Delta\omega_{m+1}]\\ &= \sigma\Delta t\mathbb{E}^x_m[D_x y_{m+2}(X_{m+2}) - D_x y^\Delta_{m+2}(X_{m+2})]\\ &= \mathbb{E}^x_m[y_{m+2}(X_{m+2})\Delta\omega_{m+2} - y^\Delta_{m+2}(X_{m+2})\Delta\omega_{m+2}]\\ &= \mathbb{E}^x_m[\epsilon^y_{m+2}\Delta\omega_{m+2}]. \end{aligned} \tag{6.4.29}$$

We can also write error ϵ^f_{m+2} as a function of X_{m+2}, as

$$\begin{aligned}\epsilon^f_{m+2} &= f(t_{m+2},\mathbf{\Lambda}_{m+2}) - f(t_{m+2},\mathbf{\Lambda}^\Delta_{m+2})\\ &= f(t_{m+2},X_{m+2},y_{m+2}(X_{m+2}),z_{m+2}(X_{m+2}))\\ &\quad - f(t_{m+2},X_{m+2},y^\Delta_{m+2}(X_{m+2}),z^\Delta_{m+2}(X_{m+2})). \end{aligned} \tag{6.4.30}$$

The equalities in Appendix 6.A.I result in

$$\begin{aligned}\mathbb{E}^x_m[\epsilon^f_{m+2}\Delta\omega_{m+1}] &= \sigma\Delta t\mathbb{E}^x_m\left[\tfrac{d}{dx}f(t_{m+2},\mathbf{\Lambda}_{m+2}) - \tfrac{d}{dx}f(t_{m+2},\mathbf{\Lambda}^\Delta_{m+2})\right]\\ &= \mathbb{E}^x_m[\epsilon^f_{m+2}\Delta\omega_{m+2}]. \end{aligned} \tag{6.4.31}$$

Here $\frac{d}{dx}f$ denotes the total derivative of driver f to state x, where y and z also depend on x. With the two equalities (6.4.29) and (6.4.31) we find

$$\epsilon^z_m = 2\mathbb{E}^x_m[\epsilon^f_{m+1}\Delta\omega_{m+1}] + \mathbb{E}^x_m[\epsilon^z_{m+2}] + \tfrac{2}{\Delta t}\mathbb{E}^x_m[R^y_{m+1}\Delta\omega_{m+1}] + \tfrac{2}{\Delta t}\left(R^z_m - \mathbb{E}^x_m[R^z_{m+1}]\right). \tag{6.4.32}$$

We can bound the absolute value of the first term by

$$\begin{aligned}|2\mathbb{E}^x_m[\epsilon^f_{m+1}\Delta\omega_{m+1}]| &\le 2\mathbb{E}^x_m[|\epsilon^f_{m+1}||\Delta\omega_{m+1}|] \le 2\sup|\epsilon^f_{m+1}|\sqrt{\Delta t}\\ &\le 2L_f(\sup|\epsilon^y_{m+1}| + \sup|\epsilon^z_{m+1}|)\sqrt{\Delta t}\\ &:= 2L_f(|\epsilon^y_{m+1}|_\infty + |\epsilon^z_{m+1}|_\infty)\sqrt{\Delta t}, \end{aligned} \tag{6.4.33}$$

where the suprema are taken under the condition $X_m = x$.

We can now bound the absolute error by

$$|\epsilon^z_m| \le 2L_f\sqrt{\Delta t}(|\epsilon^y_{m+1}|_\infty + |\epsilon^z_{m+1}|_\infty) + \mathbb{E}^x_m[|\epsilon^z_{m+2}|] + 2C(\Delta t)^3. \tag{6.4.34}$$

Next we sum up the errors. For $\frac{1}{2}\Delta t L_f < \frac{1}{2}$, $\Delta t \le 1$ there exists constants C_1 and C_2, depending on L_f, with ($m \le M-3$)

$$\begin{aligned}\mathbb{E}^x_m[|e_m|] &:= \mathbb{E}^x_m[|\epsilon^y_m|_\infty + \sqrt{\Delta t}|\epsilon^z_{m+1}|_\infty + \sqrt{\Delta t}|\epsilon^z_m|_\infty]\\ &\le A\mathbb{E}^x_m[|\epsilon^y_{m+2}|_\infty + \sqrt{\Delta t}|\epsilon^z_{m+2}|_\infty + \sqrt{\Delta t}|\epsilon^z_{m+3}|_\infty] + B = A\mathbb{E}^x_m[|e_{m+2}|] + B,\\ \text{with}\quad A &= \tfrac{1}{1-L_f\Delta t}(1 + C_1\Delta t),\\ B &= \tfrac{1}{1-L_f\Delta t}C_2(\Delta t)^3. \end{aligned} \tag{6.4.35}$$

Theorem 6.4.1. *Given*

$$\mathbb{E}^x_{M-1}[|\epsilon^z_M|] \sim \mathcal{O}\left((\Delta t)^3\right), \quad \mathbb{E}^x_{M-1}[|\epsilon^y_M|] \sim \mathcal{O}\left((\Delta t)^3\right), \tag{6.4.36}$$

then

$$\mathbb{E}^x_0\left[|\epsilon^y_m| + \sqrt{\Delta t}|\epsilon^z_m|\right] \leq Q(\Delta t)^2, \quad 1 \leq m \leq M, \tag{6.4.37}$$

with Q a constant only depending on T, functions g and f, and μ, σ.

Proof. With equations (6.4.25), (6.4.27), (6.4.34) it is straightforward to show that

$$\mathbb{E}^x_{M-1}[|e_{M-1}|] \sim \mathcal{O}\left((\Delta t)^2\right) \text{ and } \mathbb{E}^x_{M-2}[|e_{M-2}|] \sim \mathcal{O}\left((\Delta t)^2\right). \tag{6.4.38}$$

By induction we find

$$\mathbb{E}^x_m[|e_m|] \leq A^{\frac{1}{2}(M-m)}\mathbb{E}^x_m[|e_{M-1}| + |e_{M-2}|] + \frac{A^{\frac{1}{2}(M-m)}-1}{A-1}B, \quad \text{for } m \leq M-3. \tag{6.4.39}$$

It follows that ($1 \leq k \leq M$, $\Delta t M = T$)

$$\begin{aligned} A^k - 1 \leq A^k &\leq \left(\tfrac{1+C_1\Delta t}{1-L_f\Delta t}\right)^k = \left(1 + \tfrac{(C_1+L_f)\Delta t}{1-L_f\Delta t}\right)^k \\ &\leq \exp\left(\tfrac{(C_1+L_f)\Delta t k}{1-L_f\Delta t}\right) \leq \exp\left(\tfrac{(C_1+L_f)T}{1-\Delta t L_f}\right) \end{aligned} \tag{6.4.40}$$

is bounded and

$$\tfrac{B}{A-1} \leq \tfrac{C_2}{C_1+L_f}(\Delta t)^2. \tag{6.4.41}$$

□

The authors in [ZLJ13] obtain second order convergence in both Y and Z terms for the case that the FSDE equals the Wiener process. Convergence of (Y^Δ, Z^Δ) to (Y, Z) is discussed in [BT04, Zha04, LGW06, GLW05, BE08] for the special case $\theta_1 = \theta_2 = 1$. Under certain conditions on functions f and g, error convergence of order $\mathcal{O}\left((\Delta t)^{1/2}\right)$ in the L^2-sense was found. The authors in [BDM01] prove convergence of a discrete scheme with a scaled random walk using a Donsker-type theorem. For the error analysis of other schemes and L^p-errors we refer to [WLZ09, LZ10].

Error in COS Formulas In Section 6.3 we encountered deterministic functions y^Δ_m and z^Δ_m, such that

$$y^\Delta_m(X^\Delta_m) = Y^\Delta_m, \quad z^\Delta_m(X^\Delta_m) = Z^\Delta_m. \tag{6.4.42}$$

These functions are approximated by COS formulas and the corresponding Fourier coefficients are recovered backwards in time, resulting in the approximations

$$\hat{y}^\Delta_m(X^\Delta_m) \quad \text{and} \quad \hat{z}^\Delta_m(X^\Delta_m). \tag{6.4.43}$$

The errors of these numerical approximations are denoted by

$$\epsilon^y_{COS}(t_m, X^\Delta_m) := y^\Delta_m(X^\Delta_m) - \hat{y}^\Delta_m(X^\Delta_m), \tag{6.4.44a}$$

$$\epsilon^z_{COS}(t_m, X^\Delta_m) := z^\Delta_m(X^\Delta_m) - \hat{z}^\Delta_m(X^\Delta_m). \tag{6.4.44b}$$

Fourier series expansions and their convergence properties have been discussed in [Boy01]. Errors of the COS method are introduced in three steps (see Section 2.3.1): the truncation of the integration range, the substitution of the density by its cosine series expansion on the truncated range, and the substitution of the series coefficients by the characteristic function approximation. A detailed error analysis was given in [FO08, FO09] and Section 2.4.1. For a sufficiently wide computational domain $[a,b]$ the integration range truncation error in our domain of interest can be neglected, because the truncated mass of the density function is negligible. The local error converges exponentially in the number of terms in the series expansions for smooth density functions and a sufficiently wide integration interval. The transitional density that is related to the Euler scheme is smooth and results in exponential convergence in N. A density function with a discontinuity in one of its derivatives gives rise to an algebraic convergence in N. We refer to Chapter 3 for more information on the convergence of discontinuous functions. Algorithm 3 explains how to recover the coefficients $\mathcal{Z}_k^{\Delta}(t_m)$, $\mathcal{Y}_k^{\Delta}(t_m)$, and $\mathcal{F}_k^{\Delta}(t_m)$ backwards in time. This introduces an additional error, similar as in Section 2.4.3. The use of discrete Fourier cosine transforms (DCTs) to approximate the Fourier cosine coefficients gives an error with algebraic index of convergence two in N, as we demonstrate by an example in Section 6.5.3.

Convergence of Picard Iterations With P Picard iterations we find the fixed-point y of the equation

$$y = \Delta t \theta_1 f(t_m, x, y, z_m^{\Delta}(x)) + h_m^{\Delta}(x). \tag{6.4.45}$$

The driver function f is assumed to be Lipschitz in y and z, with Lipschitz constant L_f. For Δt small enough, i.e., $L_f \Delta t \theta_1 < 1$, a unique fixed-point exists, and the Picard iterations converge towards that point for any initial guess. The fixed-point technique converges to the true solution at the geometric rate $\Delta t \theta_1 L_f$, which depends on the Lipschitz constant of the driver function.

Remark 6.4.2. *Problems with a penalization term, see for example [BK08], may result in driver functions with a large Lipschitz constant, which gives slower convergence in the number of Picard iterations. The number of timesteps should be chosen sufficiently large, such that convergence of the iterations is guaranteed.*

Total Error The absolute value of the total errors can be bounded by

$$|\epsilon_m^y(X_m, X_m^{\Delta})| := |y_m(X_m) - \hat{y}_m^{\Delta}(X_m^{\Delta})| \leq |y_m(X_m) - y_m^{\Delta}(X_m^{\Delta})| + |\epsilon_{COS}^y(t_m, X_m^{\Delta})|, \tag{6.4.46a}$$

$$|\epsilon_m^z(X_m, X_m^{\Delta})| := |z_m(X_m) - \hat{z}_m^{\Delta}(X_m^{\Delta})| \leq |z_m(X_m) - z_m^{\Delta}(X_m^{\Delta})| + |\epsilon_{COS}^z(t_m, X_m^{\Delta})|. \tag{6.4.46b}$$

For the numerical experiments in Section 6.5 we take N sufficiently high. Then we can neglect the errors ϵ_{COS} and are able to investigate the error of the discretization scheme.

Computational Complexity The computation time of the BCOS method is linear in the number of timesteps M. For each discrete time-point t_m we perform the following operations:

- Computation of $\hat{z}_m^{\Delta}(x)$ and $\hat{h}_m^{\Delta}(x)$ on an x-grid, in $\mathcal{O}(N^2)$ operations.

- Initialization Picard method: Computation of $\hat{y}_m^{\Delta,0}(x)$ on an x-grid, in $\mathcal{O}(N^2)$ operations.
- Computation of $\hat{y}_m^{\Delta,P}(x)$ on an x-grid by P Picard iterations, in $\mathcal{O}(PN)$ operations.
- Computation of $\mathcal{Z}_k^{\Delta}(t_m)$ and $\mathcal{H}_k^{\Delta}(t_m)$ by the FFT algorithm, in $\mathcal{O}(N\log N)$ operations.
- Computation of $\mathcal{F}_k^{\Delta}(t_m) \approx \mathcal{F}_k^{\Delta,P-1}(t_m)$ by DCT (see Appendix C), in $\mathcal{O}(N\log N)$ operations.
- Computation of $\mathcal{Y}_k^{\Delta}(t_m) \approx \mathcal{Y}_k^{\Delta,P}(t_m)$, in $\mathcal{O}(N)$ operations.

For the approximation of the coefficients $\mathcal{F}_k^{\Delta,P-1}(t_m)$ in (6.4.20) by a DCT we first need to compute $\hat{z}_m^{\Delta}(x)$, $\hat{h}_m^{\Delta}(x)$, and $\hat{y}_m^{\Delta,0}(x)$ on an x-grid with N equidistant points, which is of order $\mathcal{O}(N^2)$. This is the most time-consuming part of the algorithm. However, these functions can be computed in parallel. In total, the complexity of the BCOS method, Algorithm 3, is $\mathcal{O}\left(M\left(N^2 + PN + N\log N + N\log N + N\right)\right)$.

6.5. NUMERICAL EXPERIMENTS

In this section we discuss three numerical experiments. MATLAB 7.11.0 is used for the computations, with an Intel(R) Core(TM) i5-2520M CPU @ 2.50GHz and 7.7 GB RAM. To test the general θ-method we distinguish between four discretization schemes:

Scheme A:	$\theta_1 = 0$,	$\theta_2 = 1$,	Scheme C:	$\theta_1 = 1$,	$\theta_2 = 1$,
Scheme B:	$\theta_1 = 0.5$,	$\theta_2 = 1$,	Scheme D:	$\theta_1 = 0.5$,	$\theta_2 = 0.5$.

For all four schemes, $z_m^{\Delta}(x)$ can be solved explicitly and $y_m^{\Delta}(x)$ is solved explicitly for scheme A and implicitly with $P = 5$ Picard iterations for the other schemes.

As before, we prescribe a computational domain $[a, b]$ by

$$[a, b] = \left[\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}\right], \tag{6.5.1}$$

with cumulants $\kappa_1 = x_0 + \mu T$ and $\kappa_2 = \sigma^2 T$, and $L = 10$. Furthermore, we set the number of terms in the Fourier cosine series expansions equal to $N = 2^9$. For these values the BCOS method has converged in N to machine precision.

6.5.1. EXAMPLE 1

The first example is taken from [ZLZ12]. The underlying process is the Wiener process, i.e., $X_t = \omega_t$. The BSDE reads

$$dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t d\omega_t, \tag{6.5.2a}$$

$$f(t, X_t, Y_t, Z_t) = Y_t Z_t - Z_t + 2.5Y_t - \sin(t + X_t)\cos(t + X_t) - 2\sin(t + X_t), \tag{6.5.2b}$$

$$Y_T = g(X_T) = \sin(X_T + T). \tag{6.5.2c}$$

The exact solution is given by

$$(Y_t, Z_t) = (\sin(X_t + t), \cos(X_t + t)). \tag{6.5.3}$$

We take terminal time $T = 1$, which gives $(Y_0, Z_0) = (0, 1)$. Note that driver $f(.)$ depends also on time t and state X_t. For the results of the BCOS method, we refer to Figure 6.5.1. We

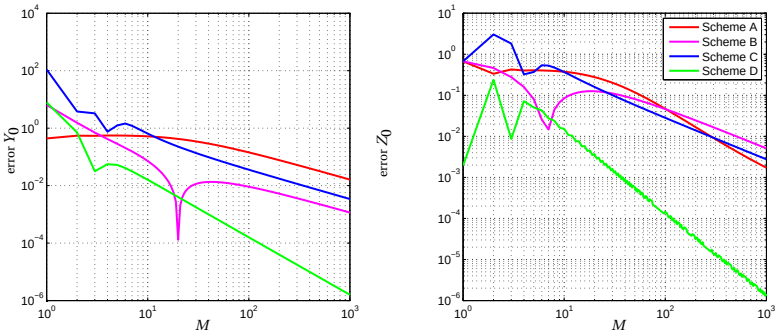


Figure 6.5.1: Results example 1 ($N = 2^9$), left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

observe that the approximated value $\hat{y}_0^{\Delta}(x_0)$ converges with $\mathcal{O}(\Delta t)$ for the schemes A, B, and C and $\mathcal{O}((\Delta t)^2)$ for scheme D. The approximated value $\hat{z}_0^{\Delta}(x_0)$ converges with $\mathcal{O}((\Delta t)^2)$ for scheme D and with $\mathcal{O}(\Delta t)$ for the other three schemes, which is in accordance with the error analysis in Section 6.4.5.

Table 6.5.1 shows CPU times, for scheme D, for different values of M and N. Each test required less than one second. Computation of the functions $\hat{z}_m^{\Delta}(x)$, $\hat{h}_m^{\Delta}(x)$, and $\hat{y}_m^{\Delta,0}(x)$ on an x-grid and the DCTs are the most time-consuming part of the algorithm. The computation time is linear in the number of timesteps M and of approximately order $\mathcal{O}(N \log N)$ in the number of terms in the Fourier cosine series expansions.

Table 6.5.1: CPU time (s).

M	4	8	16	32	64	128	256	512
$N = 2^9$	0.0301	0.0304	0.0412	0.0639	0.1071	0.1966	0.3736	0.7292

N	2^6	2^7	2^8	2^9
$M = 256$	0.0940	0.1109	0.1552	0.3736

6.5.2. EXAMPLE 2: BLACK-SCHOLES CALL OPTION

In this example we compute the price $v(t, S_t)$ of a call option by a BSDE where the underlying asset follows a geometric Brownian motion,

$$dS_t = \bar{\mu} S_t dt + \bar{\sigma} S_t d\omega_t. \tag{6.5.4}$$

The exact solution is given by the Black-Scholes price, which is known analytically [BS73]. For the derivation of the Black-Scholes PDE we set up a self-financing portfolio Y_t with a_t assets and bonds with risk-free return rate r. Markets are assumed to be complete in this model, there are no trading restrictions, and the option can be exactly replicated by the hedging portfolio, that is, $Y_T = \max(S_T - K, 0)$. Then, the option value at initial time should

be equal to the initial value of the portfolio. The portfolio evolves according to the SDE

$$dY_t = r(Y_t - a_t S_t)dt + a_t dS_t = \left(rY_t + \tfrac{\bar{\mu}-r}{\bar{\sigma}}\bar{\sigma}a_t S_t\right)dt + \bar{\sigma}a_t S_t d\omega_t. \tag{6.5.5}$$

If we set $Z_t = \bar{\sigma}a_t S_t$, then (Y,Z) solves the BSDE,

$$dY_t = -f(t,S_t,Y_t,Z_t)dt + Z_t d\omega_t, \tag{6.5.6a}$$

$$f(t,S_t,Y_t,Z_t) = -rY_t - \tfrac{\bar{\mu}-r}{\bar{\sigma}}Z_t, \tag{6.5.6b}$$

$$Y_T = \max(S_T - K, 0). \tag{6.5.6c}$$

Y_t corresponds to the value of the portfolio and Z_t is related to the hedging strategy. In this case, the driver function $f(.)$ is Lipschitz continuous and linear with respect to y and z. The option value is given by $v(t,S_t) = Y_t$ and $\bar{\sigma}S_t v_S(t,S_t) = Z_t$. For the tests, we use the following parameter values

$$S_0 = 100,\ K = 100,\ r = 0.1,\ \bar{\mu} = 0.2,\ \bar{\sigma} = 0.25,\ T = 0.1, \tag{6.5.7}$$

with the exact solutions $Y_0 = v(t_0,S_0) = 3.65997$ and $Z_0 = \bar{\sigma}S_0 v_S(t_0,S_0) = 14.14823$. For the numerical approximation, we switch to the log-asset domain $X_t = \log S_t$, with

$$dX_t = (\bar{\mu} - \tfrac{1}{2}\bar{\sigma}^2)dt + \bar{\sigma}d\omega_t. \tag{6.5.8}$$

Results of the BCOS method for all four schemes are presented in Figure 6.5.2. The approximated values $\hat{y}_0^\Delta(x_0)$ and $\hat{z}_0^\Delta(x_0)$ both converge with $\mathcal{O}(\Delta t)$ for schemes A, B, and C and with $\mathcal{O}((\Delta t)^2)$ for scheme D, as expected.

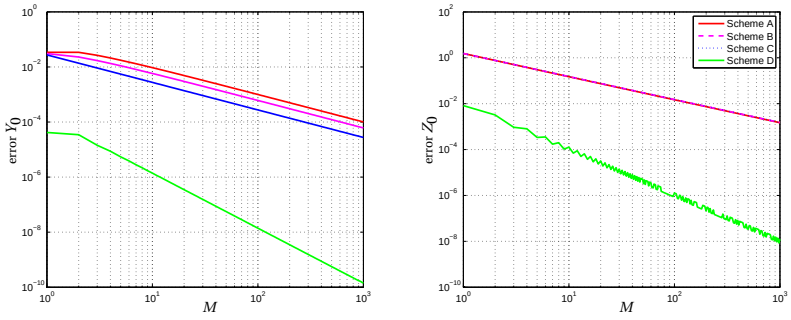


Figure 6.5.2: Results example 2 ($N = 2^9$), left: error in $\hat{y}_0^\Delta(x_0)$, right: error in $\hat{z}_0^\Delta(x_0)$.

We would like to emphasize that solving the BSDE is done under the historical, real-world $\mathbb{P}$-measure. However, the exact Black-Scholes solution does not depend on $\bar{\mu}$. In Figure 6.5.3 we see results for different values of drift $\bar{\mu}$. The convergence rates in M are the same, but a higher value of $\bar{\mu}$ gives a larger error for the same number of timesteps M. This is due to the Lipschitz constant $L_f = \max(\frac{\bar{\mu}-r}{\bar{\sigma}}, r)$, which is increasing in $\bar{\mu}$.

6.5.3. EXAMPLE 3: BID-ASK SPREAD FOR INTEREST RATES

For the pricing problem in the previous section, the driver $f(.)$ was linear and the initial option value Y_0 reduced to the expectation of the discounted option payoff under the risk-neutral measure, i.e., the Black-Scholes price. We now consider a model introduced by

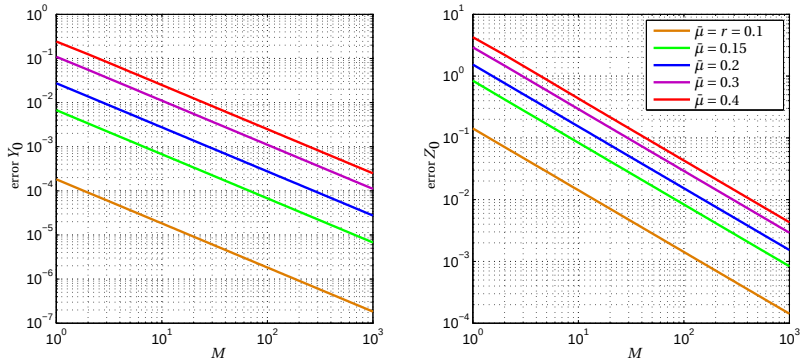


Figure 6.5.3: Results example 2 for different values of $\bar{\mu}$ (Scheme C).

Bergman [Ber95] with different interest rates for lending and borrowing a bond. This market imperfection results in a driver function which is nonlinear.

Suppose that an agent can invest in bonds at risk-free rate r and borrow money at rate $R > r$. The amount invested at time t is equal to $\max(Y_t - a_t S_t, 0)$ and the amount borrowed is $\min(Y_t - a_t S_t, 0)$. Then, the replication portfolio follows the following dynamics

$$\begin{aligned} dY_t &= r\max(Y_t - a_t S_t, 0)dt + R\min(Y_t - a_t S_t, 0)dt + a_t dS_t \\ &= \left(rY_t + \tfrac{\bar{\mu}-r}{\bar{\sigma}}\bar{\sigma}a_t S_t + (R-r)\min(Y_t - a_t S_t, 0)\right)dt + \bar{\sigma}a_t S_t d\omega_t. \end{aligned} \tag{6.5.9}$$

Again, with $Z_t = \bar{\sigma}a_t S_t$, we obtain the BSDE

$$dY_t = -f^{r,R}(t, S_t, Y_t, Z_t)dt + Z_t d\omega_t, \tag{6.5.10a}$$

$$f^{r,R}(t, S_t, Y_t, Z_t) := -rY_t - \tfrac{\bar{\mu}-r}{\bar{\sigma}}Z_t - (R-r)\min(Y_t - Z_t/\bar{\sigma}, 0), \tag{6.5.10b}$$

$$Y_T = g(S_T). \tag{6.5.10c}$$

The driver is now a nonlinear Lipschitz function and the dependence on the different return rates is emphasized by the superscripts. The Lipschitz constant of the driver function is given by $L_f = \max((\bar{\mu}-r)/\bar{\sigma}, (R-r)/\bar{\sigma}, r, R-r)$. The corresponding semilinear PDE for $v(t, S_t) = Y_t$ reads

$$\frac{\partial v}{\partial t}(t,S) + \mathscr{L}v(t,S) + f^{r,R}(t, S, v(t,S), \bar{\sigma}SD_S v(t,S)) = 0, \quad (t,S) \in [0,T)\times\mathbb{R}_+, \tag{6.5.11a}$$

$$v(T,S) = g(S), \quad S \in \mathbb{R}_+. \tag{6.5.11b}$$

This semilinear PDE is solved in [FL07] by using a PDE discretization method. According to [Gob10] the following lower bound is valid

$$Y_t^{r,R} \geq \max(Y_t^{r,r}, Y_t^{R,R}), \quad \forall t \in [0,T], \tag{6.5.12}$$

where $Y_t^{r_1,r_2}$ denotes the solution with driver f^{r_1,r_2}. In other words, the option price with different interest rates is larger than the price with fixed interest rates. For a call option it follows that $Y_t^{r,R} = Y_t^{R,R}$, in other words, the exact solution is given by the Black-Scholes call price with interest rate R and the BSDE becomes linear, as in Section 6.5.2. The authors

in [BS12] examine a combination of a long call with strike $K_1 = 95$ and two short calls with strike $K_2 = 105$, with payoff function

$$g(S) = (S - K_1)^+ - 2(S - K_2)^+, \tag{6.5.13}$$

and parameter values

$$S_0 = 100,\ r = 0.01,\ \bar{\mu} = 0.05,\ \bar{\sigma} = 0.2,\ T = 0.25. \tag{6.5.14}$$

An analytic solution is not available then. We perform tests for borrow rates $R = 0.06$ and $R = 3.01$ with corresponding Lipschitz constants $L_f = 0.25$ and $L_f = 15$, respectively. The authors in [BS12] agree that an interest rate of 301% is not relevant for financial applications, but they propose it in order to test their algorithm under an extreme situation, as we will do.

For $R = 0.06$, the reference value obtained in [BS12] is $Y_0^{r,R} = 2.96$. Error results of the BCOS method are shown in Figure 6.5.4. For this we used reference values $Y_0^{r,R} = 2.9584544$ and $Z_0^{r,R} = 0.55319$, obtained with a large number of timesteps M. The approximated value $\hat{y}_0^{\Delta}(x_0)$ converges with $\mathcal{O}(\Delta t)$ for schemes A, B, and C and approximately with $\mathcal{O}((\Delta t)^{3/2})$ for scheme D. The approximated value $\hat{z}_0^{\Delta}(x_0)$ converges with $\mathcal{O}(\Delta t)$ for schemes B and C and with a higher convergence rate for schemes A and D.

For $R = 3.01$, the authors of [BS12] find $Y_0^{r,R} = 6.4$. For large M the BCOS method gives the reference values $Y_0^{r,R} = 6.3748$ and $Z_0^{r,R} = -4.690$. The results of our numerical approximations are shown in Figure 6.5.5. Convergence to accurate values is slower and the errors are larger compared to $R = 0.06$, because the driver function has a larger Lipschitz constant. The convergence rates are not clearly readable, however for large M the orders seem to correspond to the case with $R = 0.06$. Scheme D, with $\theta_1 = \theta_2 = 0.5$, gives the best error results.

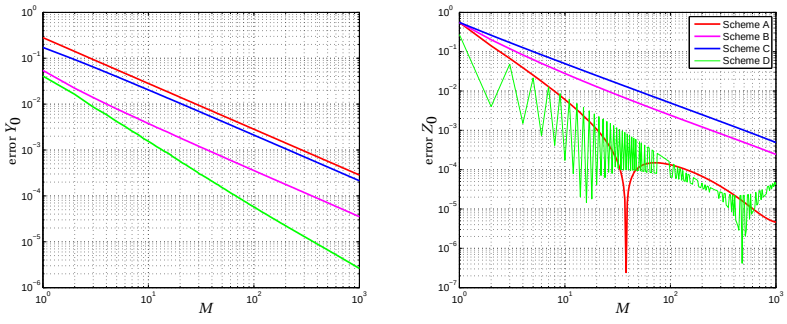


Figure 6.5.4: Results example 3, $R = 0.06$ ($N = 2^9$), left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

At last we test the convergence of the BCOS method in N, the number of Fourier cosine terms. Figure 6.5.6 presents the errors for different values of N, with $R = 3.01$ and $M = 100$ timesteps, where we computed our reference values by taking a large value of N. The convergence in N, the number of Fourier coefficients, is of second order, due to the use of discrete Fourier cosine transforms (see Appendix C).

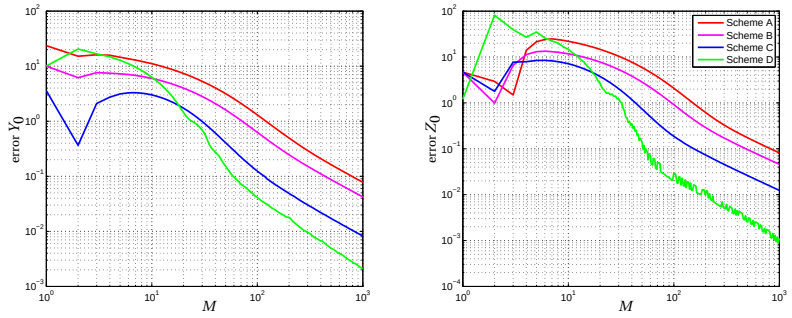


Figure 6.5.5: Results example 3, $R = 3.01$ ($N = 2^9$), left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

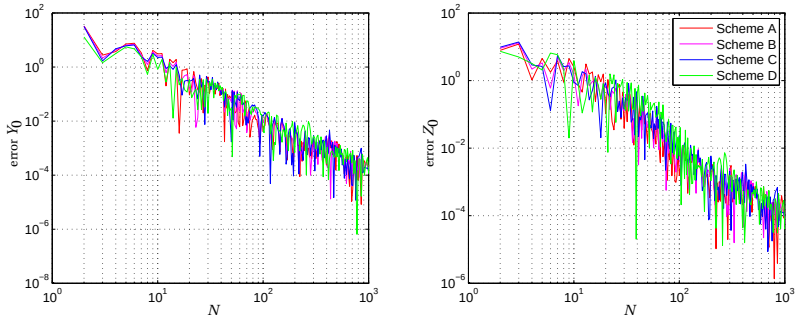


Figure 6.5.6: Convergence in N, ($R = 3.01$ and $M = 100$), left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

6.6. EXPONENTIAL UTILITY MAXIMIZATION AND UTILITY INDIFFERENCE PRICE

In a financial market with jumps or with constrained hedging strategies it is usually not possible to perform a perfect hedge which attains exactly the option payoff as the final value, there is a so-called replication error. If markets are not complete there are different ways to value options ([EKPQ97]), for example by

- *Super-strategies* are strategies with a positive replication error. The super-replicating option price is the minimal initial investment to find a strategy that always dominates the payoff of the option [EKQ95].
- *Risk-minimizing strategies*, where the problem is to find a strategy with minimal variance for the replication error [KFV09]. They were first introduced by Föllmer and Sondermann in [FS86].
- *Utility indifference pricing*, where the utility of the replication error is maximized. The corresponding price makes an agent indifferent in terms of expected utility between selling the option or not selling it. Utility indifference pricing was introduced by Hodges and Neuberger in [HN89].

We focus on utility indifference pricing, which basically consists of solving two utility maximization problems, one with and one without an option liability. In the next section

we consider a general utility maximization problem. We employ the model of Morlais in [Mor10], making use of an exponential utility function and jumps in the asset price. The problem can be defined by a BSDE including jumps. We refer to [REK00, Sek06, HIM05, MS05, Pha09] for the setting where asset prices follow only a diffusion process. This model is generalized by jumps in [Bec06, Mor10].

6.6.1. EXPONENTIAL UTILITY MAXIMIZATION UNDER JUMP-DIFFUSION WITH OPTION PAYOFF

Following the notation in [Mor10], the probability space $(\Omega, \mathbb{F}, \mathbb{P})$ is now equipped with two independent stochastic processes: the standard Brownian motion ω and a real-valued Poisson point process defined on $\Omega \times [0,T] \times E$, with $E := \mathbb{R}/\{0\}$. We denote by $N(dJ, dt)$ the associated Poisson random measure whose compensator is assumed to be of the form $\nu(dJ)dt$, where $\nu(dJ)$ stands for the Lévy measure, which is positive and satisfies

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_E (1 \wedge |J|)^2 \nu(dJ) < \infty. \tag{6.6.1}$$

$N(B,t)$, $B \subset \mathbb{R}$, is the number of jumps with size in set B which occur before or at time t and $\nu(B)$ counts the expected number of jumps in a unit time interval. $\mathcal{F}$ is the completed filtration generated by both processes ω and N. The so-called compensated Poisson random measure, $\tilde{N}$, is given by

$$\tilde{N}(dJ, dt) = N(dJ, dt) - \nu(dJ)dt. \tag{6.6.2}$$

The asset price is supposed to follow the following jump-diffusion process,

$$dS_t / S_{t-} = b(t)dt + \sigma(t)d\omega_t + \int_E \beta(t,J)\tilde{N}(dJ, dt), \tag{6.6.3}$$

S_{t-} represents the value of the asset just before a possible jump occurs. The jumps may model the occurrence of, for example, market crashes or default losses. An agent sells a bounded $\mathcal{F}_T$-measurable option payoff $\xi = g(S_T)$ at time $t = 0$. He is endowed with some initial capital w and then invests $\boldsymbol{\alpha}_t$, $t \in [0,T]$ of his portfolio W_t in assets. The remaining part is invested in a risk-free opportunity with zero rate of return, i.e., $r = 0$. The dynamics of this self-financing portfolio read

$$dW_t = \boldsymbol{\alpha}_t \frac{dS_t}{S_{t-}} = \boldsymbol{\alpha}_t b(t)dt + \boldsymbol{\alpha}_t \sigma(t)d\omega_t + \boldsymbol{\alpha}_t \int_E \beta(t,J)\tilde{N}(dJ, dt), \quad W_0 = w. \tag{6.6.4}$$

At terminal time T there is an uncertain claim ξ and the agent is able to reduce the risk by his trading strategy. The attitude of the agent towards possible profits and losses is measured by an exponential utility,

$$\mathcal{U}(x) = -\exp(-\eta x), \quad \eta > 0. \tag{6.6.5}$$

The utility function is monotonically increasing and concave; η is the coefficient of absolute risk aversion and represents the degree of risk aversion. A higher value of η corresponds to a higher level of risk aversion. A negative amount of final wealth has a higher weight than a positive amount, in other words, more weight is given to unfavorable losses. $\eta = 0$ corresponds to risk neutrality and $\eta = \infty$ to absolute risk aversion. The agent wants to maximize his expected utility at time T and his *objective function* now reads

$$V(w) = \max_{\boldsymbol{\alpha} \in \mathcal{A}} \mathbb{E}\left[\mathcal{U}(W_T - \xi)\right] = \max_{\boldsymbol{\alpha} \in \mathcal{A}} \mathbb{E}\left[\mathcal{U}\left(w + \int_0^T \boldsymbol{\alpha}_t \frac{dS_t}{S_{t-}} - \xi\right)\right], \tag{6.6.6}$$

where we maximize over the investment opportunities $\boldsymbol{\alpha}$ in the constraint set $\mathscr{A}$ with admissible strategies. Possible trading strategies may be restricted, for example an agent may be forced not to hold a negative amount of assets. For the tests in Section 6.7 we will take $A = [\alpha^-, \alpha^+]$.

The objective function can also be characterized by a backward stochastic differential equation with jumps (BSDEJ), as follows

$$V(w) = \mathcal{U}(w - Y_0) = -e^{-\eta w} e^{\eta Y_0}, \tag{6.6.7}$$

where (Y, Z, U) is the solution to a BSDEJ, given by ([Mor10])

$$dY_t = -f(t, Z_t, U_t)dt + Z_t d\omega_t + \int_E U_t(J)\tilde{N}(dJ, dt), \quad Y_T = \xi, \tag{6.6.8a}$$

$$f(t,z,u) = -z\frac{b(t)}{\sigma(t)} - \frac{1}{2\eta}\left|\frac{b(t)}{\sigma(t)}\right|^2 + \min_{\alpha\in A}\left[\frac{\eta}{2}\left(\alpha\sigma(t) - \left(z + \frac{1}{\eta}\frac{b(t)}{\sigma(t)}\right)\right)^2 + |u(.) - \alpha\beta(t,.)|_\eta\right], \tag{6.6.8b}$$

$$|u(.)|_\eta = \int_E \frac{\exp(\eta u(J)) - \eta u(J) - 1}{\eta}\nu(dJ), \tag{6.6.8c}$$

The solution of the above BSDEJ consists of a triplet of processes (Y, Z, U) in $\mathscr{S}^\infty(\mathbb{R}) \times L^2(\omega) \times L^2(\tilde{N})$ [4]. Existence and uniqueness results for this BSDEJ are provided in [Mor10]. For more information about existence and uniqueness for BSDEJs, we refer to [TL94, BBP97, Roy06]. Furthermore, there exists an optimal predictable strategy $\boldsymbol{\alpha}^* \in \mathscr{A}$ that attains the minimum in (6.6.8b) for $(t, z, u) = (t, Z_t, U_t)$.

6.6.2. UTILITY INDIFFERENCE PRICE

Now we start with the utility indifference price, where the idea is the following. The seller of an option receives the option premium and hedges the option with an optimal strategy that maximizes the utility of the portfolio value at the terminal time minus the payoff. We also determine the expected utility without the option trade. The *utility indifference price* of the option is defined as the additional initial wealth with which the seller can achieve the same utility as without the option.

Let $u_0(w)$ denote the utility maximization value without the option payoff

$$u_0(w) = \max_{\boldsymbol{\alpha}\in\mathscr{A}} \mathbb{E}\left[\mathcal{U}(W_T)\right], \tag{6.6.9}$$

and $u_\xi(w)$ denotes the utility maximization value in presence of the option,

$$u_\xi(w) = \max_{\boldsymbol{\alpha}\in\mathscr{A}} \mathbb{E}\left[\mathcal{U}(W_T - \xi)\right]. \tag{6.6.10}$$

The *seller's indifference price (ask price)* v^a satisfies

$$u_0(w) = u_\xi(w + v^a). \tag{6.6.11}$$

[4] Following [Mor10], $\mathscr{S}^\infty(\mathbb{R})$ is the set of all adapted processes Y with càdlàg paths such that $\sup_\Omega(\sup_{t\in[0,T]}|Y_t|) < \infty$. $L^2(\omega)$ is the set of all predictable processes Z such that $\mathbb{E}[\int_0^T |Z_s|^2 ds] < \infty$ $L^2(\tilde{N})$ is the set of all $\mathscr{P}\otimes\mathscr{B}(E)$-measurable processes U such that $\mathbb{E}[\int_0^T \int_E |U_s(J)|^2 \nu(dJ)ds] < \infty$. $\mathscr{P}$ stands for the σ-field of all predictable sets of $[0,T]\times\Omega$ and $\mathscr{B}(E)$ the Borel field of E.

In other words, it is the price at which a seller is indifferent, in the sense that the expected utility under optimal trading remains the same, between selling the option for price v^a and not selling any option. We need to solve for v^a and with the theory in Section 6.6.1 we find

$$\mathcal{U}(w - Y_0^0) = \mathcal{U}(w + v^a - Y_0^\xi) \quad \Longrightarrow \quad v^a = Y_0^\xi - Y_0^0, \tag{6.6.12}$$

where Y_t^0 and Y_t^ξ follow BSDEJ (6.6.8) with terminal conditions $Y_T = 0$ and $Y_T = \xi$, respectively. With this we can value an option under jump-diffusion and when the trading strategies are constraint, for example $A = [\alpha^-, \alpha^+]$, with the help of BSDEJs.

The *buyer's indifference price (bid price)* v^b is defined in a similar way and satisfies

$$u_0(w) = u_{-\xi}(w - v^b). \tag{6.6.13}$$

Again with the BSDE approach we find

$$\mathcal{U}(w - Y_0^0) = \mathcal{U}(w - v^b - Y_0^{-\xi}) \quad \Longrightarrow \quad v^b = Y_0^0 - Y_0^{-\xi}. \tag{6.6.14}$$

Below is a list of properties of utility indifference prices (see, for example, [HH09, REK00]). We here denote by $v(\eta, \xi)$ the option price with coefficient of absolute risk aversion η and option ξ.

- Prices v^b and v^a are independent of initial wealth w.
- Bid and ask prices are related via $v^b(\eta, \xi) = -v^a(\eta, -\xi)$.
- The ask price is larger than the bid price: $v^a \geq v^b$.
- If the market is complete, i.e., there are no jumps and $A = \mathbb{R}$, then the option is perfectly replicable. The driver function reduces to $f(t, z, u) = -z\frac{b}{\sigma} - \frac{1}{2\eta}|\frac{b}{\sigma}|^2$ and the utility indifference prices reduce to the Black-Scholes prices.

6.6.3. DISCRETIZATION AND BCOS METHOD FOR BSDEJS

In this section, we explain the BCOS method to solve BSDEJ (6.6.8). We suppose that the asset price follows the following FSDE:

$$dS_t / S_{t-} = b\,dt + \sigma\,d\omega_t + \int_E \beta(J)\tilde{N}(dJ, dt), \quad \text{with } \beta(J) = e^J - 1. \tag{6.6.15}$$

Moreover, E is assumed to be a finite set, $E = \{j_1, j_2, \ldots j_{n_j}\}$, with Lévy measure $\nu(\{j_\ell\}) = \lambda p_\ell$, where $\lambda = \nu(\mathbb{R})$ is the intensity rate. In other words, p_ℓ is the probability of jump size j_ℓ and $\nu(dJ) = \lambda \sum_{\ell=1}^{n_j} p_\ell \delta_{j_\ell}(dJ)$. So,

$$\int_E \beta(J)\tilde{N}(dJ, dt) = \sum_{\ell=1}^{n_j} \beta(j_\ell)\tilde{N}(\{j_\ell\}, dt). \tag{6.6.16}$$

We define $\mu := b - \int_E \beta(J)\nu(dJ)$ and switch to the log-asset domain $X_t = \log S_t$, i.e.,

$$dX_t = \left(\mu - \tfrac{1}{2}\sigma^2 + \int_E J\nu(dJ)\right)dt + \sigma\,d\omega_t + \int_E J\tilde{N}(dJ, dt). \tag{6.6.17}$$

The Euler discretization of the FSDE (6.6.17) reads

$$X_{m+1}^\Delta = X_m^\Delta + \left(\mu - \tfrac{1}{2}\sigma^2 + \int_E J\nu(dJ)\right)\Delta t + \sigma\Delta\omega_{m+1} + \int_E J\tilde{N}(dJ, \Delta t), \tag{6.6.18}$$

where we defined $\tilde{N}(dJ,\Delta t) := \tilde{N}(dJ,(t_m,t_{m+1}]) = \tilde{N}(dJ,t_{m+1}) - \tilde{N}(dJ,t_m)$. The characteristic function of X^{Δ}_{m+1}, given $X^{\Delta}_m = x$, reads

$$\varphi(u|x) = \varphi(u|0)e^{iux} = \phi(u)e^{iux}, \quad \text{with} \tag{6.6.19}$$
$$\phi(u) := \exp\left(iu(\mu - \tfrac{1}{2}\sigma^2)\Delta t - \tfrac{1}{2}u^2\sigma^2\Delta t\right)e^{\lambda\Delta t(\varphi_J(u)-1)}, \quad \varphi_J(u) = \sum_{\ell=1}^{n_j} p_\ell e^{iuj_\ell}.$$

For the discretization of the BSDEJ, we start from

$$Y_m = Y_{m+1} + \int_{t_m}^{t_{m+1}} f(s,Z_s,U_s)ds - \int_{t_m}^{t_{m+1}} Z_s d\omega_s - \int_{t_m}^{t_{m+1}}\int_E U_s(J)\tilde{N}(dJ,ds). \tag{6.6.20}$$

Both processes ω and $\tilde{N}$ are independent. Taking conditional expectations of both sides of (6.6.20) and applying the θ-method results, similar as equation (6.3.5), in

$$\begin{aligned} Y_m \approx{}& \mathbb{E}_m[Y_{m+1}] + \Delta t\theta_1 f(t_m,Z_m,U_m) + \Delta t(1-\theta_1)\mathbb{E}_m[f(t_{m+1},Z_{m+1},U_{m+1})], \\ & \theta_1 \in [0,1]. \end{aligned} \tag{6.6.21}$$

Multiplying both sides of equation (6.6.20) by $\Delta\omega_{m+1}$ and taking conditional expectations gives us, similar as equation (6.3.6),

$$\begin{aligned} 0 \approx{}& \mathbb{E}_m[Y_{m+1}\Delta\omega_{m+1}] + \Delta t(1-\theta_2)\mathbb{E}_m[f(t_{m+1},Z_{m+1},U_{m+1})\Delta\omega_{m+1}] \\ & - \Delta t\theta_2 Z_m - \Delta t(1-\theta_2)\mathbb{E}_m[Z_{m+1}], \quad \theta_2 \in [0,1]. \end{aligned} \tag{6.6.22}$$

Multiplying both sides of equation (6.6.20) by $\tilde{N}(\{j_\ell\},\Delta t)$ and taking conditional expectations gives

$$\begin{aligned} 0 ={}& \mathbb{E}_m[Y_{m+1}\tilde{N}(\{j_\ell\},\Delta t)] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m\left[f(s,Z_s,U_s)\tilde{N}(\{j_\ell\},s-t_m)\right]ds \\ & - \int_{t_m}^{t_{m+1}} p_\ell\lambda\mathbb{E}_m\left[U_s(j_\ell)\right]ds, \end{aligned} \tag{6.6.23}$$

where we used the Itô isometry for

$$\begin{aligned} & \mathbb{E}_m\left[\int_{t_m}^{t_{m+1}}\int_E U_s(J)\tilde{N}(dJ,ds)\tilde{N}(\{j_\ell\},\Delta t)\right] \\ & = \mathbb{E}_m\left[\int_{t_m}^{t_{m+1}}\int_E U_s(J)\tilde{N}(dJ,ds)\int_{t_m}^{t_{m+1}}\int_E \delta_{j_\ell}(J)\tilde{N}(dJ,ds)\right] = \mathbb{E}_m\left[\int_{t_m}^{t_{m+1}} p_\ell\lambda U_s(j_\ell)ds\right]. \end{aligned} \tag{6.6.24}$$

By the θ-discretization we get

$$\begin{aligned} 0 \approx{}& \mathbb{E}_m[Y_{m+1}\tilde{N}(\{j_\ell\},\Delta t)] + \Delta t(1-\theta_3)\mathbb{E}_m[f(t_{m+1},Z_{m+1},U_{m+1})\tilde{N}(\{j_\ell\},\Delta t)] \\ & - p_\ell\lambda\Delta t\theta_3 U_m(j_\ell) - p_\ell\lambda\Delta t(1-\theta_3)\mathbb{E}_m[U_{m+1}(j_\ell)], \quad \theta_3 \in [0,1], \text{ for } \ell = 1,\ldots,n_j. \end{aligned} \tag{6.6.25}$$

The above equations lead to a time discretization $(Y^\Delta, Z^\Delta, U^\Delta)$ for (Y, Z, U), as follows

$$Y_M^\Delta = g(X_M^\Delta), \tag{6.6.26a}$$

for $m = M-1, \ldots, 0$:

$$Z_m^\Delta = -\theta_2^{-1}(1-\theta_2)\mathbb{E}_m[Z_{m+1}^\Delta] + \tfrac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_m[Y_{m+1}^\Delta \Delta\omega_{m+1}] + \theta_2^{-1}(1-\theta_2)\mathbb{E}_m[f(t_{m+1}, Z_{m+1}^\Delta, U_{m+1}^\Delta)\Delta\omega_{m+1}], \tag{6.6.26b}$$

$$U_m^\Delta(j_\ell) = -\theta_3^{-1}(1-\theta_3)\mathbb{E}_m[U_{m+1}^\Delta(j_\ell)] + \tfrac{1}{p_\ell\lambda\Delta t}\theta_3^{-1}\mathbb{E}_m[Y_{m+1}^\Delta \tilde{N}(\{j_\ell\}, \Delta t)] + \tfrac{1}{p_\ell\lambda}\theta_3^{-1}(1-\theta_3)\mathbb{E}_m[f(t_{m+1}, Z_{m+1}^\Delta, U_{m+1}^\Delta)\tilde{N}(\{j_\ell\}, \Delta t)], \quad \ell = 1, \ldots, n_j \tag{6.6.26c}$$

$$Y_m^\Delta = \mathbb{E}_m[Y_{m+1}^\Delta] + \Delta t\theta_1 f(t_m, Z_m^\Delta, U_m^\Delta) + \Delta t(1-\theta_1)\mathbb{E}_m[f(t_{m+1}, Z_{m+1}^\Delta, U_{m+1}^\Delta)]. \tag{6.6.26d}$$

As the terminal condition is a deterministic function of X_M^Δ and because X^Δ is a Markov process, it is easily seen that there are deterministic functions $y_m^\Delta(x)$, $z_m^\Delta(x)$, and $u_m^\Delta(x, j_\ell)$ so that

$$Y_m^\Delta = y_m^\Delta(X_m^\Delta), \quad Z_m^\Delta = z_m^\Delta(X_m^\Delta), \quad U_m^\Delta(j_\ell) = u_m^\Delta(X_m^\Delta, j_\ell), \quad \ell = 1, \ldots, n_j. \tag{6.6.27}$$

So, the random variables Y_m^Δ, Z_m^Δ, and $U_m^\Delta(j_\ell)$ are functions of X_m^Δ, for each $m = 0, \ldots, M$. The functions $y_m^\Delta(x)$, $z_m^\Delta(x)$, and $u_m^\Delta(x, j_\ell)$ are obtained in a backward manner. Similar as in Section 6.4, the Fourier cosine coefficients of the functions $z_m^\Delta(x)$, $f(t_m, z_m^\Delta(x), u_m^\Delta(x, .))$, and $y_m^\Delta(x)$ are denoted by $\mathcal{Z}_k^\Delta(t_m)$, $\mathcal{F}_k^\Delta(t_m)$, and $\mathcal{Y}_k^\Delta(t_m)$, respectively. Let $\mathcal{U}_k^{\Delta,\ell}(t_m)$ be the Fourier cosine coefficients of $u_m^\Delta(x, j_\ell)$, i.e.,

$$\mathcal{U}_k^{\Delta,\ell}(t_m) = \tfrac{2}{b-a}\int_a^b u_m^\Delta(x, j_\ell)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \quad \ell = 1, \ldots, n_j. \tag{6.6.28}$$

We obtain the following COS formulas to approximate the conditional expectations in equation (6.6.26c), see Appendix 6.A.II for details,

$$\mathbb{E}_m^x[U_{m+1}^\Delta(j_\ell)] \approx \sum_{k=0}^{N-1}{}' \mathcal{U}_k^{\Delta,\ell}(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right\}, \tag{6.6.29a}$$

$$\mathbb{E}_m^x\left[Y_{m+1}^\Delta \tilde{N}(\{j_\ell\}, \Delta t)\right] \approx \sum_{k=0}^{N-1}{}' \mathcal{Y}_k^\Delta(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}} p_\ell\lambda\Delta t\left[\exp\left(i\tfrac{k\pi j_\ell}{b-a}\right)-1\right]\right\}, \tag{6.6.29b}$$

$$\mathbb{E}_m^x\left[f(t_{m+1}, Z_{m+1}^\Delta, U_{m+1}^\Delta)\tilde{N}(\{j_\ell\}, \Delta t)\right] \approx \sum_{k=0}^{N-1}{}' \mathcal{F}_k^\Delta(t_{m+1})\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}} p_\ell\lambda\Delta t\left[\exp\left(i\tfrac{k\pi j_\ell}{b-a}\right)-1\right]\right\}. \tag{6.6.29c}$$

Furthermore we use the COS formulas (6.4.10) and (6.4.12) from Sections 6.4.2 and 6.4.3,

obtained with the equality in Appendix 6.A.I, to find

$$z_m^{\Delta}(x) \approx -\tfrac{1-\theta_2}{\theta_2}\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}'\mathcal{Z}_k^{\Delta}(t_{m+1})\Phi_k(x) + \tfrac{1}{\Delta t\theta_2}\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}'\mathcal{Y}_k^{\Delta}(t_{m+1})\sigma\Delta t\Phi_k'(x)$$
$$+ \tfrac{1-\theta_2}{\theta_2}\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}'\mathcal{F}_k^{\Delta}(t_{m+1})\sigma\Delta t\Phi_k'(x), \tag{6.6.30a}$$

$$u_m^{\Delta}(x,j_\ell) \approx -\tfrac{1-\theta_3}{\theta_3}\tfrac{b-a}{2}\sum_{k=0}^{N-1}{}'\mathcal{U}_k^{\Delta,\ell}(t_{m+1})\Phi_k(x) \tag{6.6.30b}$$
$$+ \sum_{k=0}^{N-1}{}'\left(\tfrac{1}{\Delta t\theta_3}\mathcal{Y}_k^{\Delta}(t_{m+1}) + \tfrac{1-\theta_3}{\theta_3}\mathcal{F}_k^{\Delta}(t_{m+1})\right)\Re\left\{\phi\left(\tfrac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\Delta t\left[\exp\left(i\tfrac{k\pi j_\ell}{b-a}\right)-1\right]\right\},$$

$$y_m^{\Delta}(x) \approx \tfrac{b-a}{2}\sum_{k=0}^{N-1}{}'\left(\mathcal{Y}_k^{\Delta}(t_{m+1}) + \Delta t(1-\theta_1)\mathcal{F}_k^{\Delta}(t_{m+1})\right)\Phi_k(x)$$
$$+ \Delta t\theta_1 f(t_m, z_m^{\Delta}(x), u_m^{\Delta}(x,.)). \tag{6.6.30c}$$

The coefficients $\mathcal{Z}_k^{\Delta}(t_m)$, $\mathcal{F}_k^{\Delta}(t_m)$, and $\mathcal{Y}_k^{\Delta}(t_m)$ are recovered in a similar way as explained in Section 6.4.4. The computation of the Fourier cosine coefficients $\mathcal{U}_k^{\Delta,\ell}(t_m)$ of function $u_m^{\Delta}(x,j_\ell)$ can be decomposed into three parts. In summary, this results in

$$\mathcal{U}_k^{\Delta,\ell}(t_m) \approx \Re\left\{\sum_{j=0}^{N-1}{}'\left[-\tfrac{1-\theta_3}{\theta_3}\mathcal{U}_j^{\Delta,\ell}(t_{m+1}) + \tfrac{1}{\Delta t\theta_3}\Delta t\left[\exp\left(i\tfrac{k\pi j_\ell}{b-a}\right)-1\right]\mathcal{Y}_j^{\Delta}(t_{m+1})\right.\right.$$
$$\left.\left. + \tfrac{1-\theta_3}{\theta_3}\Delta t\left[\exp\left(i\tfrac{k\pi j_\ell}{b-a}\right)-1\right]\mathcal{F}_j^{\Delta}(t_{m+1})\right]\phi\left(\tfrac{j\pi}{b-a}\right)\mathcal{M}_{k,j}\right\}. \tag{6.6.31}$$

With the above equations we can recover the Fourier cosine coefficients recursively and solve the BSDEJ backwards in time.

6.6.4. REFERENCE VALUES

We first explain briefly how we can use the COS method, in a completely different way, to obtain reference values for the numerical tests in Section 6.7. The utility maximization problem,

$$V(w) = \max_{\boldsymbol{\alpha}\in\mathcal{A}}\mathbb{E}\left[\mathcal{U}(W_T - g(S_T))\right], \tag{6.6.32}$$

is a two-dimensional stochastic control problem with the following underlying processes

$$dS_t/S_{t-} = b\,dt + \sigma\,d\omega_t + \int_E \beta(J)\tilde{N}(dJ,dt), \tag{6.6.33a}$$

$$dW_t = \boldsymbol{\alpha}_t b\,dt + \boldsymbol{\alpha}_t\sigma\,d\omega_t + \boldsymbol{\alpha}_t\int_E \beta(J)\tilde{N}(dJ,dt). \tag{6.6.33b}$$

We can solve this problem by a combination of the 1D-COS method for stochastic control problems (Chapter 2) and the 2D-COS method (Chapter 4). The 2D-COS method was developed for pricing rainbow options, for which the payoff depends on two or more asset price processes, and can also be applied to stochastic control problems.

If it is not possible to invest in assets and to hedge the risky option, i.e., $\alpha_t = 0$, $\forall\, t \in [0,T]$, then the portfolio value $W_t = w$ is constant and the problem reduces to

$$V(w) = \mathbb{E}\left[\mathcal{U}(w - g(S_T))\right] = -e^{-\eta w}\mathbb{E}\left[e^{\eta g(S_T)}\right]. \tag{6.6.34}$$

We can approximate this one-dimensional expectation by using the one-dimensional COS formula.

6.7. NUMERICAL EXPERIMENTS BSDEJ

In this section we use the BCOS method to value a put option under jump-diffusion asset prices by using utility indifference pricing, as explained in Section 6.6.2. For the numerical tests, we use the following parameter values

$$S_0 = 1,\ K = 1,\ b = 0.1779,\ \sigma = 0.2,\ T = 0.1. \tag{6.7.1}$$

The jumps occurring are assumed to be bivariate distributed with possible jump sizes j_1 and j_2, with

$$j_1 = -0.1338,\ j_2 = -0.9838,\ p_1 = p_2 = 0.5,\ \lambda = 0.0228, \tag{6.7.2}$$

so that the expected value is -0.5588 and the standard deviation is 0.4250. These values correspond to the real-world $\mathbb{P}$-measure for the jump-diffusion asset price in [KFV09].

We choose the computational domain

$$[a,b] = \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4}},\ \kappa_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4}}\right], \qquad L = 12. \tag{6.7.3}$$

The cumulants κ_1, κ_2, and κ_4 of the Brownian motion and the Merton jump-diffusion process are, for example, given in [FO08]. Again we set the number of terms in the Fourier cosine series expansions equal to $N = 2^9$.

We distinguish between three θ-discretization schemes:

Scheme E: $\theta_1 = 1$, $\theta_2 = 1$, $\theta_3 = 1$,
Scheme F: $\theta_1 = 0.5$, $\theta_2 = 1$, $\theta_3 = 1$,
Scheme G: $\theta_1 = 0.5$, $\theta_2 = 0.5$, $\theta_3 = 0.5$.

For solving equations (6.6.26) in the first time iteration, $m = M-1$, we set $\theta_1 = \theta_2 = \theta_3 = 1$, because the driver function $f(.)$ depends on the unprescribed values $z_M^{\Delta}(x)$ and $u_M^{\Delta}(x,.)$.

No Hedge We start with the setting where it is not possible to invest in assets and to hedge the risky option, i.e., $\alpha_t = 0$, $\forall\, t \in [0,T]$. In Figure 6.7.1 results of the BCOS method are shown. The left-side plot shows the initial values of the BSDEs, Y_0^{ξ}, $Y_0^{-\xi}$, and Y_0^0, for different values of η, and the right-side plot gives the bid and ask prices. The dots are the values obtained by the BCOS method, while the black circles give the reference value obtained by the COS method as described in Section 6.6.4. The approximated values correspond to the reference values.

Restricted Hedging Strategy For the second test we assume that the set of admissible strategies is given by $A = [-15, 15]$. In other words, a maximum of 15 Euro is used to buy or sell assets. We use the Newton's method to find the optimal strategy in equation (6.6.8b). Figure 6.7.2 presents the results of the BCOS method.

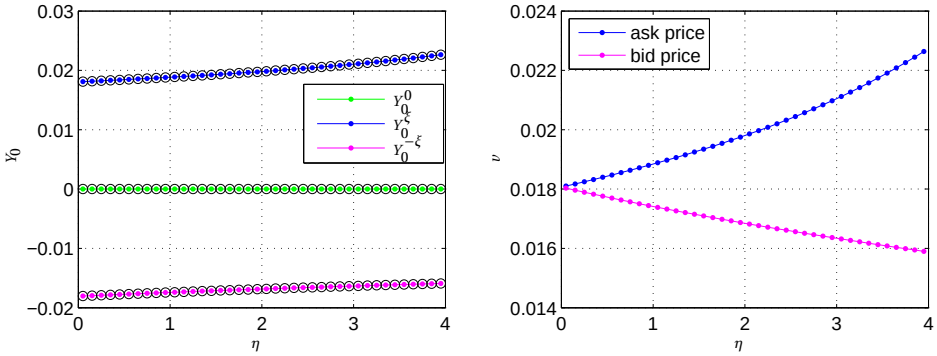


Figure 6.7.1: Results Y_0 and utility indifference prices (Scheme G, $N = 2^9$, $M = 64$, $\Delta t = 0.1/64$) (Reference values (black circles) are obtained by the (2D-)COS method).

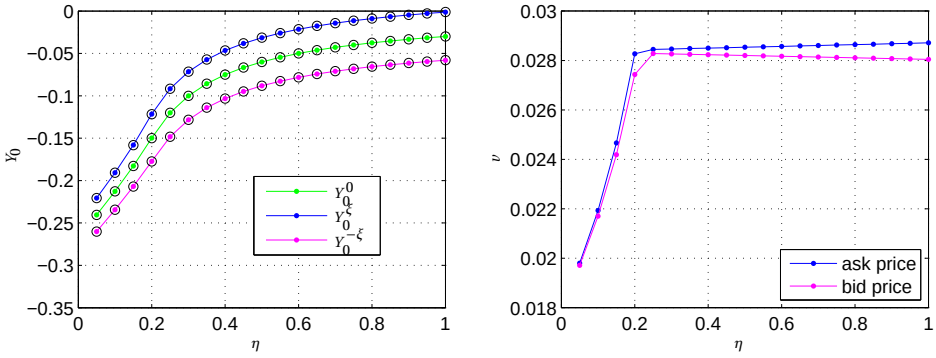


Figure 6.7.2: Results Y_0 and utility indifference prices (Scheme G, $N = 2^9$, $M = 64$, $\Delta t = 0.1/64$) (Reference values (black circles) are obtained by the 2D-COS method).

6

Convergence in M For the last test we investigate the convergence of the error in the number of timesteps M for $\eta = 1$ and with terminal conditions ξ and $-\xi$. Reference values are obtained by choosing a large number of timesteps M. The results are shown in Figure 6.7.3. The approximated value $\hat{y}_0^\Delta(x_0)$ converges with $\mathcal{O}(\Delta t)$ for schemes E and F and with $\mathcal{O}((\Delta t)^2)$ for scheme G, as expected. The values $\hat{z}_0^\Delta(x_0)$, $\hat{u}_0^\Delta(x_0, j_1)$, and $\hat{u}_0^\Delta(x_0, j_2)$ converge with $\mathcal{O}(\Delta t)$ for all three schemes. Again the scheme with $\theta_i = 1/2$, $i = 1,2,3$ gives the best convergence rate. The CPU times for different values of N and M are shown in Table 6.7.1.

Table 6.7.1: CPU time (s).

M	4	8	16	32	64	128	256	512
$N = 2^9$	0.0694	0.1086	0.1908	0.3358	0.6428	1.2555	2.4931	4.9387

N	2^6	2^7	2^8	2^9
$M = 256$	0.7897	1.0745	1.5204	2.4931

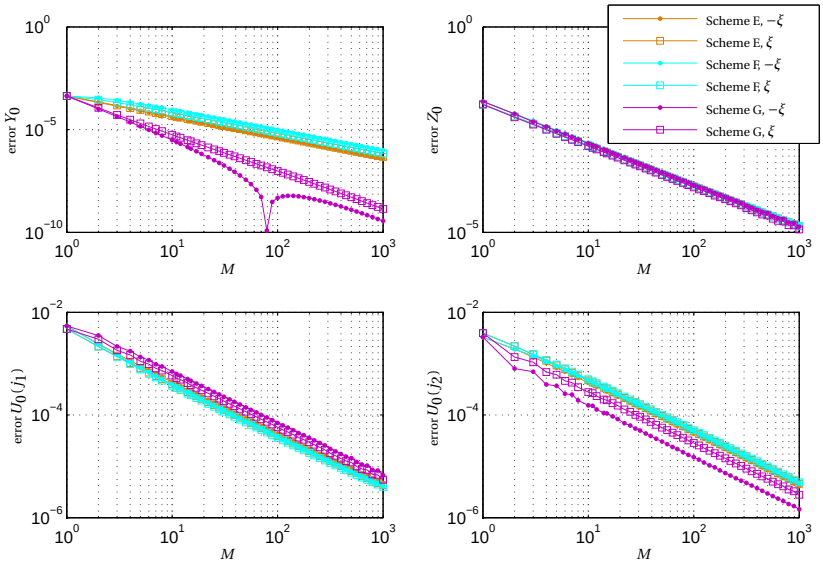


Figure 6.7.3: Convergence in M ($N = 2^9$), upper left: error in $\hat{y}_0^{\Delta}(x_0)$, upper right: error in $\hat{z}_0^{\Delta}(x_0)$, lower left: error in $\hat{u}_0^{\Delta}(x_0, j_1)$, lower right: error in $\hat{u}_0^{\Delta}(x_0, j_2)$.

6.8. CONCLUSIONS AND OUTLOOK

In this chapter we proposed a probabilistic numerical method for solving backward stochastic differential equations (BSDEs). The first step consists of discretizing the BSDE by taking conditional expectations and applying a general θ-discretization for the time-integrals. Then, the BCOS method solves the problem backwards in time by approximating the conditional expectations with the help of COS formulas. The Fourier cosine coefficients are recovered recursively in an efficient way by using discrete Fourier cosine transforms and an FFT algorithm.

Numerical tests demonstrate the applicability of the BCOS method for BSDEs in economic and financial problems. In the tests we observed different convergence results for Z_0 and Y_0. The convergence of the error in the number of timesteps depends on the smoothness and Lipschitz constant of the driver function and the terminal condition. In general, we achieve the highest convergence rate for the θ-scheme with $\theta_1 = \theta_2 = 1/2$.

Utility indifference pricing is used to value options in an incomplete market under a jump-diffusion asset price process, possibly with a restricted hedging portfolio. The bid and ask prices are represented by BSDEs with jumps. We extended our BCOS method to solving these BSDEJs under jump-diffusion with a finite number of jump sizes. Numerical experiments show highly satisfactorily and efficient pricing results. The θ-scheme with $\theta_1 = \theta_2 = \theta_3 = 1/2$ gives the fastest convergence.

The COS method is applicable for all Lévy processes and is especially efficient for the affine class. BSDEs driven by Lévy processes are discussed in [NS01] and they are a challenging extension for the BCOS method. Another interesting extension are second order BSDEs [FTW11].

Appendix

6.A. COS FORMULAS

In this section we explain how to approximate several conditional expectations under the discrete process

$$X^{\Delta}_{m+1} = X^{\Delta}_m + \mu(t_m,x)\Delta t + \sigma(t_m,x)\Delta\omega_{m+1} + \int_E JN(dJ,\Delta t), \tag{6.A.1}$$

with characteristic function

$$\begin{aligned}&\varphi(u|x) = \varphi(u|0)e^{iux} = \phi(u|x)e^{iux}, \quad \text{with}\\ &\phi(u|x) := \exp\left(iu\mu(t_m,x)\Delta t - \tfrac{1}{2}u^2\sigma^2(t_m,x)\Delta t\right)e^{\lambda\Delta t(\varphi_J(u)-1)},\end{aligned} \tag{6.A.2}$$

where $\varphi_J(u) = \sum_{\ell=1}^{n_j} p_\ell e^{iuj_\ell}$ denotes the characteristic function of jump size J.

6.A.I. COMPUTATION OF EXPECTATION $\mathbb{E}^x_m[\cdot\,\Delta\omega_{m+1}]$

Integration by parts gives us, for sufficiently smooth v,

$$\begin{aligned}&\mathbb{E}^x_m\left[v(t_{m+1},X^{\Delta}_{m+1})\Delta\omega_{m+1}\right]\\ &= \mathbb{E}^x_m\left[v\left(t_{m+1},x+\mu(t_m,x)\Delta t+\sigma(t_m,x)\Delta\omega_{m+1}+\int_E JN(dJ,\Delta t)\right)\Delta\omega_{m+1}\right]\\ &= \mathbb{E}^x_m\left[\tfrac{1}{\sqrt{2\pi}\sqrt{\Delta t}}\int_{\mathbb{R}} v\left(t_{m+1},x+\mu(t_m,x)\Delta t+\sigma(t_m,x)\zeta+\int_E JN(dJ,\Delta t)\right)\zeta e^{-\frac{1}{2}\left(\frac{\zeta}{\sqrt{\Delta t}}\right)^2}d\zeta\right]\\ &= \mathbb{E}^x_m\left[\tfrac{\sigma(t_m,x)\Delta t}{\sqrt{2\pi}\sqrt{\Delta t}}\int_{\mathbb{R}} D_x v(t_{m+1},x+\mu(t_m,x)\Delta t+\sigma(t_m,x)\zeta+\int_E JN(dJ,\Delta t))e^{-\frac{1}{2}\left(\frac{\zeta}{\sqrt{\Delta t}}\right)^2}d\zeta\right]\\ &= \sigma(t_m,x)\Delta t\mathbb{E}^x_m\left[D_x v(t_{m+1},X^{\Delta}_{m+1})\right].\end{aligned} \tag{6.A.3}$$

For the error analysis in Section 6.4.5 we assume constant μ and σ terms, then iterated conditioning gives

$$\mathbb{E}^x_m\left[v(t_{m+2},X^{\Delta}_{m+2})\Delta\omega_{m+2}\right] = \sigma\Delta t\mathbb{E}^x_m\left[D_x v(t_{m+2},X^{\Delta}_{m+2})\right] \tag{6.A.4}$$

and

$$\begin{aligned}&\mathbb{E}^x_m\left[v(t_{m+2},X^{\Delta}_{m+2})\Delta\omega_{m+1}\right]\\ &\quad= \mathbb{E}^x_m\left[v\left(t_{m+2},x+\mu 2\Delta t+\sigma\Delta\omega_{m+1}+\sigma\Delta\omega_{m+2}+\int_E JN(dJ,2\Delta t)\right)\Delta\omega_{m+1}\right]\\ &\quad= \sigma\Delta t\mathbb{E}^x_m\left[D_x v(t_{m+2},X^{\Delta}_{m+2})\right].\end{aligned} \tag{6.A.5}$$

The derivation for diffusion processes can be found by omitting the jump part in the derivation.

6.A.II. COMPUTATION OF EXPECTATION $\mathbb{E}_m^x\left[\cdot\,\tilde{N}(\{j_\ell\},\Delta t)\right]$

Equation (6.6.29b), and similarly equation (6.6.29c), require the computation of

$$\begin{aligned}\mathbb{E}_m^x\left[v(t_{m+1},X_{m+1}^\Delta)\tilde{N}(\{j_\ell\},\Delta t)\right] &= \mathbb{E}_m^x\left[v(t_{m+1},X_{m+1}^\Delta)N(\{j_\ell\},\Delta t)\right]\\ &\quad-\mathbb{E}_m^x\left[v(t_{m+1},X_{m+1}^\Delta)\right]\nu(\{j_\ell\})\Delta t.\end{aligned}\tag{6.A.6}$$

The first part in (6.A.6) can be written as

$$\mathbb{E}_m^x\left[v(t_{m+1},X_{m+1}^\Delta)N(\{j_\ell\},\Delta t)\right]\approx{\sum_{k=0}^{N-1}}'\mathcal{V}_k(t_{m+1})\Re\left\{\mathbb{E}_m^x\left[N(\{j_\ell\},\Delta t)\exp\left(iu(X_{m+1}^\Delta-a)\right)\right]\right\},\tag{6.A.7}$$

with

$$\begin{aligned}\mathbb{E}_m^x\left[N(\{j_\ell\},\Delta t)\exp\left(iu(X_{m+1}^\Delta-a)\right)\right] &= \mathbb{E}_m^x\left[\exp\left(iu(x+\mu(t_m,x)\Delta t+\sigma(t_m,x)\Delta\omega_{m+1}-a)\right)\right]\\ &\quad\cdot\mathbb{E}_m^x\left[N(\{j_\ell\},\Delta t)\exp\left(iu\int_E JN(dJ,\Delta t)\right)\right].\end{aligned}\tag{6.A.8}$$

Now let $\tau_q, q=1,2,\ldots,N_{\Delta t}$ denote the jump times between t_m and t_{m+1}, with jump sizes J_{τ_q}. Then, we find by the law of iterated expectations the following equality

$$\begin{aligned}\mathbb{E}\left[N(\{j_\ell\},\Delta t)\exp\left(iu\int_E JN(dJ,\Delta t)\right)\right] &= \mathbb{E}\left[\sum_{q=1}^{N_{\Delta t}}\mathbf{1}_{j_\ell}(J_{\tau_q})\exp\left(iu\sum_{l=1}^{N_{\Delta t}}J_{\tau_l}\right)\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[\sum_{q=1}^{N_{\Delta t}}\mathbf{1}_{j_\ell}(J_{\tau_q})\exp\left(iu\sum_{l=1}^{N_{\Delta t}}J_{\tau_l}\right)\Big|N_{\Delta t}\right]\right]\\ &= \sum_{n=0}^{\infty}e^{-\lambda\Delta t}\tfrac{(\lambda\Delta t)^n}{n!}\mathbb{E}\left[\sum_{q=1}^{n}\mathbf{1}_{j_\ell}(J_{\tau_q})\exp\left(iu\sum_{l=1}^{n}J_{\tau_l}\right)\right]\\ &= \sum_{n=0}^{\infty}e^{-\lambda\Delta t}\tfrac{(\lambda\Delta t)^n}{n!}\sum_{q=1}^{n}\mathbb{E}\left[\mathbf{1}_{j_\ell}(J_{\tau_q})\exp\left(iuJ_{\tau_q}\right)\right]\mathbb{E}\left[\exp\left(iu\sum_{l=1,l\neq q}^{n}J_{\tau_l}\right)\right]\\ &= \sum_{n=0}^{\infty}e^{-\lambda\Delta t}\tfrac{(\lambda\Delta t)^n}{n!}np_\ell e^{iuj_\ell}\left(\varphi_J(u)\right)^{n-1}\\ &= e^{iuj_\ell}p_\ell\lambda\Delta t e^{\lambda\Delta t(\varphi_J(u)-1)}.\end{aligned}\tag{6.A.9}$$

We end up with the approximation

$$\mathbb{E}_m^x\left[v(t_{m+1},X_{m+1}^\Delta)\tilde{N}(\{j_\ell\},\Delta t)\right]\approx{\sum_{k=0}^{N-1}}'\mathcal{V}_k(t_{m+1})\Re\left\{\phi(u)e^{iu(x-a)}\left[\exp\left(iuj_\ell\right)-1\right]p_\ell\lambda\Delta t\right\}.\tag{6.A.10}$$

CHAPTER 7

Numerical Fourier Method and Second-Order Taylor Scheme for Backward SDEs in Finance

7.1. INTRODUCTION

Backward stochastic differential equations (BSDEs) form an interesting recent concept in financial mathematics. Their range of applicability has increased, for example, by counterparty credit exposure and also insurance applications [Cré12]. The asset dynamics have also been generalized, for example to jump diffusion processes (see [BE08] and the previous chapter). Recently, several advanced probabilistic numerical methods have been developed for FBSDEs, like advanced Monte Carlo methods [LGW06, GLW05, BS12], integration methods [ZCP06] and also Fourier methods [HO14]. Of course, a natural aim for these solution methods is to make them highly efficient so that they can compete with semilinear PDE discretization and solution methods. With the BCOS method (Backward Stochastic Differential Equation COS method), in Chapter 6, we developed an efficient FBSDE solution method for asset dynamics for which a characteristic function of the continuous process can easily be derived.

In this chapter we extend the BCOS method. We apply the well-known Euler and Milstein schemes and an Order 2.0 weak Taylor scheme to the forward stochastic differential equation. Then we explicitly derive the characteristic function for the *discrete form* of a forward stochastic differential equation (FSDE), which will give us interesting opportunities and generalizations. Traditionally, we know that a characteristic function can be derived, as the Fourier transform of the probability density function for models from the class of regular affine processes of [DFS03], which also includes the exponentially affine jump-diffusion class of and some stochastic volatility models, and for exponentially Lévy models. When FSDE dynamics are not affine, like in the case of local volatility dynamics of Constant Elasticity of Variance (CEV) process, on which the Stochastic Alpha Beta Rho (SABR) model is based, we cannot traditionally employ Fourier techniques. The characteristic function corresponding to the discrete process can be however used, as we will show, to price options by a Fourier technique. This feature forms a natural generalization of our BCOS method. In Chapter 6 we focused on normal, lognormal and jump processes for which we could relatively easily achieve second-order convergence, in the number of timesteps, in an FBSDE context by means of a θ-time-discretization scheme ([Ise09]) and the COS Fourier cosine expansion technique. In this chapter the second-order accuracy is achieved, in combination with a fast and efficient scheme, also for the CEV model, for time- and spatially-dependent

This chapter is based on the article 'Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance', submitted for publication, 2014 [RO14].

processes, as well as for the Cox-Ingersoll-Ross (CIR) process. In the traditional context, it has been shown in [FO08] that Fourier cosine expansions will lead to exponentially convergent computational methods for smooth density functions. Here, in the discrete case, we will not observe exponential convergence in the number of Fourier coefficients, but second-order convergence, due to the use of discrete Fourier cosine transforms.

We start in Section 7.2 with definitions of forward and backward SDEs and present the link between coupled FBSDEs and semilinear partial differential equations. Section 7.2.2 discusses discrete schemes for the FSDE, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes and exact simulation schemes. Our numerical algorithm is described in Section 7.3, where we start with the characteristic function of the discretized FSDE and an example of Bermudan options (under the risk-neutral measure). A general time-discretization of a BSDE results in expressions with conditional expectations (Section 7.3.2). These conditional expectations are computed by the COS formulas (Section 7.3.3) and the problem is then solved backwards in time. Section 7.3.4 presents the overall algorithm. An error analysis is performed in Section 7.4. We demonstrate the numerical method by extensive experiments in Section 7.5. Section 7.6 concludes.

7.2. BACKWARD AND FORWARD STOCHASTIC DIFFERENTIAL EQUATIONS

For notation and definitions for BSDEs, for which we follow the survey paper [EKPQ97], we refer to Section 6.2.

7.2.1. DECOUPLED FBSDES

Here, we consider a *semilinear* parabolic PDE of the form[1]

$$\frac{\partial v}{\partial t}(t,x)+\mu(x)D_x v(t,x)+\tfrac{1}{2}\sigma^2(x)D_x^2 v(t,x)+f(t,x,v(t,x),\sigma(x)D_x v(t,x))=0,$$

$$(t,x)\in[0,T)\times\mathbb{R}, \tag{7.2.1a}$$

$$v(T,x)=g(x),\quad x\in\mathbb{R}. \tag{7.2.1b}$$

This PDE has a probabilistic representation, by means of the following FSDE,

$$X_t=x,\quad dX_s=\mu(X_s)ds+\sigma(X_s)d\omega_s,\quad t\le s\le T. \tag{7.2.2}$$

and BSDE

$$dY_s=-f(s,X_s,Y_s,Z_s)ds+Z_s d\omega_s,\quad Y_T=g(X_T), \tag{7.2.3}$$

whose terminal condition is determined by the terminal value of FSDE (7.2.2). Notice that also the driver function now depends on the state of the FSDE. Let $X_s^{t,x}$ denote the solution to (7.2.2), starting from x at time t, and $(Y_s^{t,x},Z_s^{t,x})$ be the corresponding solution to the BSDE. For assumptions on the functions σ, μ, f and g, to guarantee the existence of a unique solution (Y,Z) to the BSDE (7.2.3), we refer to Section 6.2. The following results hold

Result 7.2.1. *([PP92, Pha09]) Let $v\in C^{1,2}$ be a classical solution to* (7.2.1a) *and suppose there exists a constant $C\ge 0$ such that, for all (t,x), $|v(t,x)|+|\sigma(x)D_x v(t,x)|\le C(1+|x|)$. Then the pair (Y,Z), defined by*

$$Y_s^{t,x}=v(s,X_s^{t,x}),\quad Z_s^{t,x}=\sigma(X_s^{t,x})D_x v(s,X_s^{t,x}),\quad t\le s\le T, \tag{7.2.4}$$

[1] In Chapter 6 the drift $\mu(x)$ and diffusion term $\sigma(x)$ also depend on time. Here we omit this time-dependency for ease of notation. In Section 7.5.4 we include an example with time-dependent functions.

is the solution to BSDE (7.2.3) *(a so-called verification result).*

The converse result states: Suppose (Y,Z) *is the solution to the BSDE, then the function defined by* $v(t,x) = Y_t^{t,x}$ *is a viscosity solution to the PDE.*

For our numerical method, we wish to discretize both the forward and backward stochastic processes by schemes higher than order one and aim to find a flexible and efficient solution method, competitive in performance. In the following subsection we recap the Itô-Taylor expansion and discretization schemes for the FSDEs.

7.2.2. ITÔ-TAYLOR EXPANSION AND DISCRETIZATION SCHEMES

For an extensive survey on stochastic Taylor expansions we refer to [KP92]. We consider the integral form of the FSDE,

$$X_t = X_{t_0} + \int_{t_0}^{t} \mu(X_s)ds + \int_{t_0}^{t} \sigma(X_s)d\omega_s, \quad t_0 \le t \le T. \tag{7.2.5}$$

Drift function $\mu(.)$ and volatility function $\sigma(.)$ are assumed to be sufficiently smooth. Itô's formula for a general sufficiently smooth function $h(t,x)$ gives

$$h(t,X_t) = h(t_0,X_{t_0}) + \int_{t_0}^{t} \mathscr{L}^0 h(s,X_s)ds + \int_{t_0}^{t} \mathscr{L}^1 h(s,X_s)d\omega_s, \quad t_0 \le t \le T, \tag{7.2.6}$$

with diffusion operators

$$\mathscr{L}^0 := \tfrac{\partial}{\partial t} + \mu\tfrac{\partial}{\partial x} + \tfrac{1}{2}\sigma^2\tfrac{\partial^2}{\partial x^2}, \quad \mathscr{L}^1 := \sigma\tfrac{\partial}{\partial x}. \tag{7.2.7}$$

By applying Itô's formula to the functions $\mu(X_s)$ and $\sigma(X_s)$ in (7.2.5) we find

$$\begin{aligned} X_t = X_{t_0} + \mu(X_{t_0})I_{(0),t_0,t} + \sigma(X_{t_0})I_{(1),t_0,t} + \int_{t_0}^{t}\int_{t_0}^{s} \mathscr{L}^0\mu(X_u)duds \\ + \int_{t_0}^{t}\int_{t_0}^{s} \mathscr{L}^1\mu(X_u)d\omega_u ds + \int_{t_0}^{t}\int_{t_0}^{s} \mathscr{L}^0\sigma(X_u)dud\omega_s + \int_{t_0}^{t}\int_{t_0}^{s} \mathscr{L}^1\sigma(X_u)d\omega_u d\omega_s, \end{aligned} \tag{7.2.8}$$

with Itô integrals

$$I_{(0),\rho,\tau} := \int_{\rho}^{\tau} ds = \tau - \rho, \quad I_{(1),\rho,\tau} := \int_{\rho}^{\tau} d\omega_s = \omega_\tau - \omega_\rho. \tag{7.2.9}$$

The general multiple Itô integral is defined recursively by (see [KP92], Chapter 5.2)

$$I_\alpha[h(.,X_.)]_{\rho,\tau} := \begin{cases} h(\tau,X_\tau), & l(\alpha) = 0, \\ \int_\rho^\tau I_{\alpha-}[h(.,X_.)]_{\rho,s}ds, & l(\alpha) \ge 1, \alpha_l = 0, \\ \int_\rho^\tau I_{\alpha-}[h(.,X_.)]_{\rho,s}d\omega_s, & l(\alpha) \ge 1, \alpha_l = 1, \end{cases} \tag{7.2.10}$$

with $l(\alpha)$ the size of multi-index vector $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_{l(\alpha)})$, $\alpha- := (\alpha_1,\alpha_2,\ldots,\alpha_{l(\alpha)-1})$, and $-\alpha := (\alpha_2,\ldots,\alpha_l)$. $\alpha_\emptyset$ denotes the multi-index of length zero. For notational convenience we write $I_{\alpha,\rho,\tau} = I_\alpha[1]_{\rho,\tau}$. The Itô coefficient functions are defined by $h_\alpha = \mathscr{L}^{\alpha_1}\mathscr{L}^{\alpha_2}\ldots\mathscr{L}^{\alpha_l}h$. Let $\mathscr{M}$ denote the set of all multi-indices. Subset $\mathscr{A} \subset \mathscr{M}$ is called a hierarchical set if $\mathscr{A} \neq \emptyset$,

$\sup_{\alpha\in\mathscr{A}} l(\alpha) < \infty$, and $-\alpha \in \mathscr{A}$ for all $\alpha \in \mathscr{A}\backslash\alpha_\emptyset$. For any hierarchical set $\mathscr{A}$ the corresponding remainder set $\mathscr{B}(\mathscr{A})$ is defined by $\mathscr{B}(\mathscr{A}) = \{\alpha \in \mathscr{M}\backslash\mathscr{A} : -\alpha \in \mathscr{A}\}$. By iterating Itô's formula we obtain the Itô-Taylor expansion, as follows

$$h(t,X_t) = \sum_{\alpha\in\mathscr{A}} h_\alpha(t_0,X_{t_0})I_{\alpha,t_0,t} + \sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[h_\alpha(.,X_.)]_{t_0,t}. \tag{7.2.11}$$

Notice that for the Itô-Taylor expansion of (7.2.5) we use $h(t,x) = x$.

We now briefly discuss some discretization schemes for FSDE (7.2.5). We define a time-grid $t_0, t_1, \ldots, t_m, \ldots, t_M = T$, with fixed timesteps $\Delta t := t_{m+1} - t_m$. For notational convenience we write $X_m = X_{t_m}$, $\omega_m = \omega_{t_m}$, $\Delta\omega_{m+1} := \omega_{m+1} - \omega_m$, and $I_{\alpha,m+1} := I_{\alpha,t_m,t_{m+1}}$. The approximated process is denoted by $X_m^\Delta = X_{t_m}^\Delta$. We start with $X_0^\Delta = X_0$ and one of the following forward schemes is used to determine the values X_{m+1}^Δ, for $m = 0,\ldots,M-1$. The strong convergence rate γ_s and weak convergence rate γ_w satisfy, for sufficiently small Δt, ([KP92], Chapter 9.6 and 9.7)

$$\mathbb{E}_0[|X_T - X_T^\Delta|] \le C(\Delta t)^{\gamma_s}, \quad |\mathbb{E}_0[P(X_T) - P(X_T^\Delta)]| \le C(\Delta t)^{\gamma_w}, \tag{7.2.12}$$

with $C > 0$ a constant, which does not depend on Δt, and $P(.)$ any $2(\gamma_w+1)$ times continuously differentiable function of polynomial growth.

Well-known schemes include the *Euler scheme* ($\gamma_s = \frac{1}{2}$, $\gamma_w = 1$), i.e.,

$$\begin{aligned} X_{m+1}^\Delta &= X_m^\Delta + \mu(X_m^\Delta)I_{(0),m+1} + \sigma(X_m^\Delta)I_{(1),m+1} \\ &= X_m^\Delta + \mu(X_m^\Delta)\Delta t + \sigma(X_m^\Delta)\Delta\omega_{m+1} \end{aligned} \tag{7.2.13}$$

and the *Milstein scheme* ($\gamma_s = 1$, $\gamma_w = 1$), i.e.,

$$\begin{aligned} X_{m+1}^\Delta &= X_m^\Delta + \mu(X_m^\Delta)I_{(0),m+1} + \sigma(X_m^\Delta)I_{(1),m+1} + \mathscr{L}^1\sigma(X_m^\Delta)I_{(1,1),m+1} \\ &= X_m^\Delta + \mu(X_m^\Delta)\Delta t + \sigma(X_m^\Delta)\Delta\omega_{m+1} + \sigma(X_m^\Delta)\sigma_x(X_m^\Delta)\tfrac{1}{2}[(\Delta\omega_{m+1})^2 - \Delta t]. \end{aligned} \tag{7.2.14}$$

Here we use the short-hand notation $\sigma_x = D_x\sigma$. We also consider the *Order 2.0 weak Taylor scheme* ($\gamma_s = 1$, $\gamma_w = 2$), defined as,

$$\begin{aligned} X_{m+1}^\Delta &= X_m^\Delta + \mu(X_m^\Delta)I_{(0),m+1} + \sigma(X_m^\Delta)I_{(1),m+1} + \mathscr{L}^1\sigma(X_m^\Delta)I_{(1,1),m+1} \\ &\quad + \mathscr{L}^1\mu(X_m^\Delta)I_{(1,0),m+1} + \mathscr{L}^0\sigma(X_m^\Delta)I_{(0,1),m+1} + \mathscr{L}^0\mu(X_m^\Delta)I_{(0,0),m+1} \\ &= X_m^\Delta + \mu(X_m^\Delta)\Delta t + \sigma(X_m^\Delta)\Delta\omega_{m+1} + \sigma(X_m^\Delta)\sigma_x(X_m^\Delta)\tfrac{1}{2}[(\Delta\omega_{m+1})^2 - \Delta t] \\ &\quad + \mu_x(X_m^\Delta)\sigma(X_m^\Delta)\Delta\bar{\omega}_{m+1} \\ &\quad + \left(\mu(X_m^\Delta)\sigma_x(X_m^\Delta) + \tfrac{1}{2}\sigma_{xx}(X_m^\Delta)\sigma^2(X_m^\Delta)\right)[\Delta\omega_{m+1}\Delta t - \Delta\bar{\omega}_{m+1}] \\ &\quad + \left(\mu(X_m^\Delta)\mu_x(X_m^\Delta) + \tfrac{1}{2}\mu_{xx}(X_m^\Delta)\sigma^2(X_m^\Delta)\right)\tfrac{1}{2}(\Delta t)^2, \end{aligned} \tag{7.2.15}$$

with

$$\Delta\bar{\omega}_{m+1} := I_{(1,0),m+1} \sim \mathscr{N}(0, \tfrac{1}{3}(\Delta t)^3), \tag{7.2.16}$$

$$\mathbb{E}[\Delta\bar{\omega}_{m+1}] = 0,\ \mathrm{Var}(\Delta\bar{\omega}_{m+1}) = \tfrac{1}{3}(\Delta t)^3,\ \mathrm{Cov}(\Delta\omega_{m+1},\Delta\bar{\omega}_{m+1}) = \tfrac{1}{2}(\Delta t)^2. \tag{7.2.17}$$

If we replace the Wiener process ω_t by a trinomial tree with increments $\Delta\hat{\omega}_{m+1} \in \{-\sqrt{3}\sqrt{\Delta t}, 0, +\sqrt{3}\sqrt{\Delta t}\}$ and replace $\Delta\bar{\omega}_{m+1}$ by $\frac{1}{2}\Delta\hat{\omega}_{m+1}\Delta t$, this leads to the so-called Order

2.0 *simplified* weak Taylor scheme [KP92]. However, we do not replace $\Delta\omega_{m+1}$ by a discrete tree, but keep the continuous random variable. We replace $\Delta\omega_{m+1}$ by $\Delta\tilde{\omega}_{m+1} := \frac{1}{2}\Delta\omega_{m+1}\Delta t$. The moments do not change up to order two:

$$\mathbb{E}[\Delta\tilde{\omega}_{m+1}] = 0,\ \mathrm{Var}(\Delta\tilde{\omega}_{m+1}) = \tfrac{1}{2}(\Delta t)^3,\ \mathrm{Cov}(\Delta\omega_{m+1}, \Delta\tilde{\omega}_{m+1}) = \tfrac{1}{2}(\Delta t)^2. \tag{7.2.18}$$

This gives us the *Order 2.0 'continuous' simplified weak Taylor scheme*

$$\begin{aligned} X^{\Delta}_{m+1} &= X^{\Delta}_m + \mu(X^{\Delta}_m)\Delta t + \sigma(X^{\Delta}_m)\Delta\omega_{m+1} \\ &\quad + \sigma(X^{\Delta}_m)\sigma_x(X^{\Delta}_m)\tfrac{1}{2}[(\Delta\omega_{m+1})^2 - \Delta t] \\ &\quad + \left(\mu_x(X^{\Delta}_m)\sigma(X^{\Delta}_m) + \mu(X^{\Delta}_m)\sigma_x(X^{\Delta}_m) + \tfrac{1}{2}\sigma_{xx}(X^{\Delta}_m)\sigma^2(X^{\Delta}_m)\right)\tfrac{1}{2}\Delta\omega_{m+1}\Delta t \\ &\quad + \left(\mu(X^{\Delta}_m)\mu_x(X^{\Delta}_m) + \tfrac{1}{2}\mu_{xx}(X^{\Delta}_m)\sigma^2(X^{\Delta}_m)\right)\tfrac{1}{2}(\Delta t)^2. \end{aligned} \tag{7.2.19}$$

We abbreviate the Order 2.0 'continuous' simplified weak Taylor scheme by *2.0-weak-Taylor scheme*. With the theory in [KP92] it can be proved that $\gamma_w = 2$.

For, for example, the Cox-Ingersoll-Ross (CIR) process and Heston model *exact simulation schemes* exist to simulate

$$X_{m+1} = X_m + \int_{t_m}^{t_{m+1}} \mu(X_s)\,ds + \int_{t_m}^{t_{m+1}} \sigma(X_s)\,d\omega_s, \tag{7.2.20}$$

by sampling directly the explicit transitional density function. For the CIR process we have

$$\mu(x) = \kappa(\bar{x} - x), \quad \sigma(x) = \eta\sqrt{x}. \tag{7.2.21}$$

There holds $X_t \geq 0$. If the Feller condition, $2\kappa\bar{x} \geq \eta^2$, is satisfied, then the process is strictly positive. If the Feller condition is not satisfied then the process can become zero. We define $q_F := 2\kappa\bar{x}/\eta^2 - 1$ and $\zeta := 2\kappa/\left((1-e^{-\kappa\Delta t})\eta^2\right)$. Then $2\zeta X_{m+1}$ is noncentral chi-squared distributed with degrees of freedom $2(q_F+1)$ and noncentrality parameter $2\zeta X_m e^{-\kappa\Delta t}$ (see [FO11, JYC09]). The characteristic function of X_{m+1} is known, i.e.,

$$\varphi_{X_{m+1}}(u|X_m = x) = \exp\left(\frac{iuxe^{-\kappa\Delta t}}{1 - \frac{iu(1-e^{-\kappa\Delta t})\eta^2}{2\kappa}}\right)\left(1 - \tfrac{iu(1-e^{-\kappa\Delta t})\eta^2}{2\kappa}\right)^{-2\kappa\bar{x}/\eta^2}. \tag{7.2.22}$$

7.3. BCOS METHOD

In Section 7.3.1 we derive the characteristic function of the underlying discretized FSDE. In Section 7.3.2 we discuss the θ- and Δ-time-discretization schemes for the coupled FBSDE and in Section 7.3.3 we derive COS formulas to approximate the occurring expectations by using the characteristic function. Section 7.3.4 presents the overall BCOS algorithm.

7.3.1. CHARACTERISTIC FUNCTION DISCRETIZATION SCHEMES FSDE

We can write the Euler, Milstein, and 2.0-weak-Taylor discretization schemes from the previous section in the following general form

$$X^{\Delta}_{m+1} = x + m(x)\Delta t + s(x)\Delta\omega_{m+1} + \kappa(x)(\Delta\omega_{m+1})^2, \quad X^{\Delta}_m = x. \tag{7.3.1}$$

For the Euler scheme we have

$$m(x) = \mu(x), \quad s(x) = \sigma(x), \quad \kappa(x) = 0, \tag{7.3.2}$$

for the Milstein scheme

$$m(x) = \mu(x) - \tfrac{1}{2}\sigma(x)\sigma_x(x), \quad s(x) = \sigma(x), \quad \kappa(x) = \tfrac{1}{2}\sigma(x)\sigma_x(x), \tag{7.3.3}$$

and for the 2.0-weak-Taylor scheme

$$m(x) = \mu(x) - \tfrac{1}{2}\sigma(x)\sigma_x(x) + \tfrac{1}{2}\left(\mu(x)\mu_x(x) + \tfrac{1}{2}\mu_{xx}(x)\sigma^2(x)\right)\Delta t, \tag{7.3.4a}$$

$$s(x) = \sigma(x) + \tfrac{1}{2}\left(\mu_x(x)\sigma(x) + \mu(x)\sigma_x(x) + \tfrac{1}{2}\sigma_{xx}(x)\sigma^2(x)\right)\Delta t, \tag{7.3.4b}$$

$$\kappa(x) = \tfrac{1}{2}\sigma(x)\sigma_x(x). \tag{7.3.4c}$$

Lemma 7.3.1. *The characteristic function of X^Δ_{m+1}, given $X^\Delta_m = x$, in equation* (7.3.1) *is given by*

$$\begin{aligned}\varphi_{X^\Delta_{m+1}}(u|X^\Delta_m = x) &= \mathbb{E}\left[\exp\left(iuX^\Delta_{m+1}\right)\middle| X^\Delta_m = x\right] \\ &= \exp\left(iux + ium(x)\Delta t - \frac{\frac{1}{2}u^2 s^2(x)\Delta t}{1-2iu\kappa(x)\Delta t}\right)(1-2iu\kappa(x)\Delta t)^{-1/2}.\end{aligned} \tag{7.3.5}$$

For $\kappa(x) = 0$ it follows that[2]

$$\varphi_{X^\Delta_{m+1}}(u|X^\Delta_m = x) = e^{iux + ium(x)\Delta t - \frac{1}{2}u^2 s^2(x)\Delta t}. \tag{7.3.6}$$

Proof. With polynomial factorization we can rewrite equation (7.3.1) as (for $\kappa(x) \neq 0$)

$$\begin{aligned}X^\Delta_{m+1} &= x + m(x)\Delta t + \kappa(x)\left(\Delta\omega_{m+1} + \tfrac{1}{2}\tfrac{s(x)}{\kappa(x)}\right)^2 - \tfrac{1}{4}\tfrac{s^2(x)}{\kappa(x)} \\ &\overset{d}{=} x + m(x)\Delta t - \tfrac{1}{4}\tfrac{s^2(x)}{\kappa(x)} + \kappa(x)\Delta t\left(\varepsilon^{\mathcal{N}}_{m+1} + \sqrt{\lambda(x)}\right)^2,\end{aligned} \tag{7.3.7}$$

with $\lambda(x) := \frac{1}{4}\frac{s^2(x)}{\kappa^2(x)\Delta t}$ and $\Delta\omega_{m+1} \overset{d}{=} \sqrt{\Delta t}\varepsilon^{\mathcal{N}}_{m+1}$, $\varepsilon^{\mathcal{N}}_{m+1} \sim \mathcal{N}(0,1)$. Random variable $\left(\varepsilon^{\mathcal{N}}_{m+1} + \sqrt{\lambda(x)}\right)^2 \sim \chi'^2_1(\lambda(x))$ is governed by a noncentral chi-squared distribution with degrees of freedom $\nu = 1$ and $\lambda(x)$ the noncentrality parameter. The characteristic function of a noncentral chi-squared distributed random variable reads

$$\varphi_{\chi'^2_\nu(\lambda)}(u) = \exp\left(\frac{i\lambda u}{1-2iu}\right)(1-2iu)^{-\nu/2}. \tag{7.3.8}$$

The characteristic function of X^Δ_{m+1}, given $X^\Delta_m = x$, is then given by

$$\begin{aligned}\varphi_{X^\Delta_{m+1}}(u|X^\Delta_m = x) &= \mathbb{E}\left[\exp\left(iuX^\Delta_{m+1}\right)\middle| X^\Delta_m = x\right] \\ &= \exp\left(iux + ium(x)\Delta t - iu\tfrac{1}{4}\tfrac{s^2(x)}{\kappa(x)}\right)\varphi_{\chi'^2_1(\lambda(x))}(u\kappa(x)\Delta t) \\ &= \exp\left(iux + ium(x)\Delta t - iu\tfrac{1}{4}\tfrac{s^2(x)}{\kappa(x)}\right)\exp\left(\frac{iu\frac{1}{4}\frac{s^2(x)}{\kappa(x)}}{1-2iu\kappa(x)\Delta t}\right)(1-2iu\kappa(x)\Delta t)^{-1/2} \\ &= \exp\left(iux + ium(x)\Delta t - \frac{\frac{1}{2}u^2 s^2(x)\Delta t}{1-2iu\kappa(x)\Delta t}\right)(1-2iu\kappa(x)\Delta t)^{-1/2}.\end{aligned} \tag{7.3.9}$$

□

[2] In Section 7.5.1 we discuss an example with $\kappa(0) = 0$.

7.3.1.1. INTERMEZZO: BERMUDAN PUT OPTION - CEV - $\mathbb{Q}$-MEASURE

To test the discretization schemes of the FSDE and the discrete characteristic function, we perform an option pricing experiment with the COS method for a Bermudan put option. We take ten early-exercise dates τ_j, $j = 1, \ldots, 10$, with fixed time intervals $T/10$. The underlying asset price under the risk-neutral $\mathbb{Q}$-measure follows a Constant Elasticity of Variance (CEV) process,

$$dX_s = rX_s ds + \bar{\sigma} X_s^{\gamma} d\omega_s. \tag{7.3.10}$$

The option price $v(t, X_t)$ is given by the risk-neutral valuation formula

$$v(t,x) = \sup_{\tau_j} e^{-r(\tau_j - t)} \mathbb{E}^{\mathbb{Q}} \left[g(X_{\tau_j}) | X_t = x \right], \tag{7.3.11}$$

with payoff function $g(x) = \max(K - x, 0)$. This problem can also be represented by a linear parabolic PDE variational inequality by means of the Feynman-Kac theorem. We take the number of timesteps of the time-discretization schemes, that is M, equal to a multiple of the number of early-exercise dates. The dynamic programming principle gives

$$v(t_m, x) = \begin{cases} e^{-r\Delta t} \mathbb{E}\left[v(t_{m+1}, X_{m+1}) \middle| X_m = x \right], & \text{for } t_m \neq \tau_j, \\ \max\left[e^{-r\Delta t} \mathbb{E}\left[v(t_{m+1}, X_{m+1}) \middle| X_m = x \right], g(x) \right], & \text{for } t_m = \tau_j. \end{cases} \tag{7.3.12}$$

The COS formula gives us

$$\begin{aligned} \mathbb{E}\left[v(t_{m+1}, X_{m+1}) \middle| X_m = x \right] &\approx \mathbb{E}\left[v(t_{m+1}, X^{\Delta}_{m+1}) \middle| X^{\Delta}_m = x \right] \\ &\approx \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1}) \mathbb{E}\left[\cos\left(k\pi \tfrac{X^{\Delta}_{m+1} - a}{b-a} \right) \middle| X^{\Delta}_m = x \right] \\ &= \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1}) \Re\left\{ \mathbb{E}\left[\exp\left(ik\pi \tfrac{X^{\Delta}_{m+1}}{b-a} \right) \middle| X^{\Delta}_m = x \right] \exp\left(ik\pi \tfrac{-a}{b-a} \right) \right\} \\ &= \sum_{k=0}^{N-1}{}' \mathcal{V}_k(t_{m+1}) \Re\left\{ \varphi_{X^{\Delta}_{m+1}}\left(\tfrac{k\pi}{b-a} \middle| X^{\Delta}_m = x \right) \exp\left(ik\pi \tfrac{-a}{b-a} \right) \right\}, \end{aligned} \tag{7.3.13}$$

where $\mathcal{V}_k(t_{m+1})$ denote the Fourier cosine coefficients of $v(t_{m+1}, x)$, i.e.,

$$\mathcal{V}_k(t_{m+1}) = \tfrac{2}{b-a} \int_a^b v(t_{m+1}, x) \cos\left(k\pi \tfrac{x-a}{b-a} \right) dx. \tag{7.3.14}$$

The coefficients are recovered recursively backwards in time (similar as in Section 7.3.4). The characteristic function of the discretized FSDE, X^{Δ}_{m+1}, is known, whereas the characteristic function of the FSDE, X_{m+1}, is only available in closed form for $\gamma \in \{0, 0.5, 1, 1.5\}$ [Lew00]. For the tests, we use the following parameter values

$$X_0 = 100,\ K = 100,\ r = 0.1,\ T = 0.1. \tag{7.3.15}$$

We take the elasticity of variance equal to $\gamma = 0.2$ and $\gamma = 0.8$ and choose $\bar{\sigma}$ so that $\sigma(X_0) = 25$. The exact solutions for corresponding European options ([Hul09]) are given by $v(t_0, X_0) = 2.6650$ and $v(t_0, X_0) = 2.6655$, respectively. The Bermudan values are $v(t_0, X_0) =$

2.7353 and $v(t_0, X_0) = 2.7373$, where we computed our reference values by taking high values of M and N.

The results for the Euler, Milstein, and 2.0-weak-Taylor schemes are shown in Figure 7.3.1 for $N = 2^9$. The 2.0-weak-Taylor scheme results in second-order convergence in M. It is interesting to see that the 2.0-weak-Taylor does not only have a better convergence rate, but also the absolute value of the error is lower even for small M. Because of that, we only need a small number of timesteps to achieve a small error. For comparison, with the 2.0-weak-Taylor scheme we need only 20 timesteps to get errors smaller than 10^{-5}, whereas the Euler scheme requires approximately 900 timesteps.

Table 7.3.1 presents the errors for the 2.0-weak-Taylor scheme for different values of N, with $M = 1000$ timesteps. The convergence in N, the number of Fourier coefficients, is of second order, due to the use of discrete Fourier cosine transforms (see Section 7.3.4).

The CPU time is shown in the last row. The computation time for the Euler scheme is only slightly shorter. The computation of the expected values, with the COS formula, on an x-grid with length N is $\mathcal{O}(N^2)$, due to matrix-vector multiplications. The usage of discrete Fourier cosine transforms (DCTs) for the recovery of the coefficients is of order $\mathcal{O}(N \log N)$. As the use of the DCTs is the most time-consuming part we do not observe quadratic complexity for these values of N.

We would like to mention that this method can also be applied to time-dependent drift and diffusion terms and to other local volatility models.

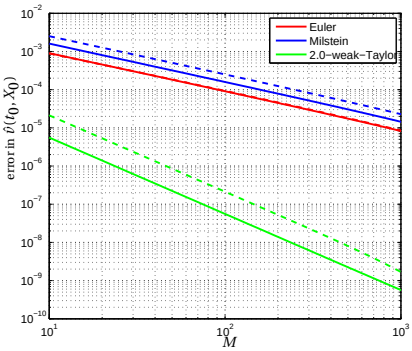


Figure 7.3.1: Results Bermudan put, $\gamma = 0.2$ (lines) and $\gamma = 0.8$ (dashes) ($N = 2^9$).

Table 7.3.1: Error and CPU time for 2.0-weak-Taylor scheme ($\gamma = 0.2$, $M = 1000$).

N	2^4	2^5	2^6	2^7	2^8	2^9
2.0-weak-Taylor	3.2235e-02	2.5416e-03	2.9661e-04	1.1736e-04	7.9437e-05	1.6482e-05
CPU time (s)	0.0127	0.0128	0.0132	0.0143	0.0169	0.0352

7.3.2. Δ-TIME-DISCRETIZATION SCHEME

In this section, we focus on the full discretization scheme of the decoupled FBSDE system[3]. First, we consider the integral form of the BSDE,

$$Y_t = g(X_T) + \int_t^T f(s, \mathbf{\Lambda}_s) ds - \int_t^T Z_s d\omega_s, \tag{7.3.16}$$

with $\mathbf{\Lambda}_s := (X_s, Y_s, Z_s)$. For the discretization of the BSDE at time-point t_m, we then start with

$$Y_m = Y_{m+1} + \int_{t_m}^{t_{m+1}} f(s, \mathbf{\Lambda}_s) ds - \int_{t_m}^{t_{m+1}} Z_s d\omega_s. \tag{7.3.17}$$

Taking conditional expectations at both sides of equation (7.3.17) and applying the θ-method ([Ise09]) results in

$$\begin{aligned} Y_m &= \mathbb{E}_m [Y_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m \left[f(s, \mathbf{\Lambda}_s) \right] ds \\ &\approx \mathbb{E}_m [Y_{m+1}] + \Delta t \theta f(t_m, \mathbf{\Lambda}_m) + \Delta t (1-\theta) \mathbb{E}_m \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}) \right], \end{aligned} \tag{7.3.18}$$

where $\mathbb{E}_m[.]$ represents the conditional expectation $\mathbb{E}[.|\mathcal{F}_{t_m}]$. Multiplying both sides of equation (7.3.17) by $\Delta\omega_{m+1}$, taking the conditional expectations, and applying the θ-method gives us

$$\begin{aligned} 0 &= \mathbb{E}_m [Y_{m+1} \Delta\omega_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m \left[f(s, \mathbf{\Lambda}_s)(\omega_s - \omega_{t_m}) \right] ds - \int_{t_m}^{t_{m+1}} \mathbb{E}_m [Z_s] ds \\ &\approx \mathbb{E}_m [Y_{m+1} \Delta\omega_{m+1}] + \Delta t (1-\theta) \mathbb{E}_m \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}) \Delta\omega_{m+1} \right] \\ &\quad - \Delta t \theta Z_m - \Delta t (1-\theta) \mathbb{E}_m [Z_{m+1}]. \end{aligned} \tag{7.3.19}$$

For the approximation of the FSDE we will use the Euler, Milstein, or 2.0-weak-Taylor schemes, as described in Section 7.2.2. Then the following Δ-*time-discretization scheme* algorithm is used to approximate the BSDE

$$Y_M^\Delta = g(X_M^\Delta), \quad Z_M^\Delta = \sigma(X_M^\Delta) D_x g(X_M^\Delta), \tag{7.3.20a}$$

for $m = M-1, \ldots, 0$:

$$\begin{aligned} Z_m^\Delta &= \tfrac{1}{\Delta t} \theta^{-1} \mathbb{E}_m \left[Y_{m+1}^\Delta \Delta\omega_{m+1} \right] + \theta^{-1}(1-\theta) \mathbb{E}_m \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta) \Delta\omega_{m+1} \right] \\ &\quad - \theta^{-1}(1-\theta) \mathbb{E}_m \left[Z_{m+1}^\Delta \right], \end{aligned} \tag{7.3.20b}$$

$$Y_m^\Delta = \mathbb{E}_m \left[Y_{m+1}^\Delta \right] + \Delta t \theta f(t_m, \mathbf{\Lambda}_m^\Delta) + \Delta t (1-\theta) \mathbb{E}_m \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta) \right]. \tag{7.3.20c}$$

The values Y_m^Δ and Z_m^Δ depend on the value of the forward process. Then it is easily seen, using an induction argument, that there are deterministic functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$, so that

$$Y_m^\Delta = y_m^\Delta(X_m^\Delta), \quad Z_m^\Delta = z_m^\Delta(X_m^\Delta). \tag{7.3.21}$$

X_m denotes the exact solution of the FSDE (7.2.2) at time t_m and X_m^Δ is its discrete approximation. Let $X_k^{m,x}$ denote the value of X_k, given $X_m = x$. Similarly, $X_k^{\Delta,m,x}$ denotes the value of X_k^Δ, given $X_m^\Delta = x$. We use the following notation

[3]The scheme is similar to the discretization scheme in Section 6.3. However, we here use $\theta_1 = \theta_2 = \theta$.

- $(y_m(x), z_m(x))$ is the exact solution of BSDE (7.3.16) at time t_m, given $X_m = x$.

- $(y_m^\theta(x), z_m^\theta(x))$ is the discrete approximation of the BSDE with the θ-time-discretization scheme, given exact solution $X_m = x$, and, for $m = M-1, \dots, 0$,

$$\begin{aligned} z_m^\theta(x) &= \tfrac{1}{\Delta t}\theta^{-1}\mathbb{E}\left[y_{m+1}^\theta(X_{m+1}^{m,x})\Delta\omega_{m+1}\right] \\ &\quad + \theta^{-1}(1-\theta)\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\theta(X_{m+1}^{m,x}))\Delta\omega_{m+1}\right] - \theta^{-1}(1-\theta)\mathbb{E}\left[z_{m+1}^\theta(X_{m+1}^{m,x})\right], \end{aligned} \tag{7.3.22a}$$

$$\begin{aligned} y_m^\theta(x) &= \mathbb{E}\left[y_{m+1}^\theta(X_{m+1}^{m,x})\right] + \Delta t\theta f(t_m, \mathbf{\Lambda}_m^\theta(x)) \\ &\quad + \Delta t(1-\theta)\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\theta(X_{m+1}^{m,x}))\right]. \end{aligned} \tag{7.3.22b}$$

- $(y_m^\Delta(x), z_m^\Delta(x))$ is the discrete approximation with the Δ-time-discretization scheme, given discrete approximation $X_m^\Delta = x$, and, for $m = M-1, \dots, 0$,

$$\begin{aligned} z_m^\Delta(x) &= \tfrac{1}{\Delta t}\theta^{-1}\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] \\ &\quad + \theta^{-1}(1-\theta)\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\Delta\omega_{m+1}\right] - \theta^{-1}(1-\theta)\mathbb{E}\left[z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right], \end{aligned} \tag{7.3.23a}$$

$$\begin{aligned} y_m^\Delta(x) &= \mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right] + \Delta t\theta f(t_m, \mathbf{\Lambda}_m^\Delta(x)) \\ &\quad + \Delta t(1-\theta)\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\right]. \end{aligned} \tag{7.3.23b}$$

The values $y_m^\theta(x)$ and $y_m^\Delta(x)$ are implicit for $\theta > 0$ and are determined by performing $P = 10$ *Picard iterations*, starting with initial guesses $\mathbb{E}\left[y_{m+1}^\theta(X_{m+1}^{m,x})\right]$ and $\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right]$, respectively.

In the BCOS method we here use the characteristic function of the underlying discretized FSDE, as discussed in Section 7.3.1, to approximate the appearing conditional expectations $\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right]$, $\mathbb{E}\left[z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right]$, and $\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\right]$. Besides, we need to approximate the expected values $\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right]$ and $\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\Delta\omega_{m+1}\right]$. Details are discussed in the following subsection.

7.3.3. EXPECTED VALUES Δ-TIME-DISCRETIZATION SCHEME FBSDE

In this section we derive an equation for the conditional expectations $\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{\Delta,m,x})\right]$ and $\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right]$ under the general discrete dynamics (7.3.1). Here $h(t,x)$ is a general function.

Let $\mathcal{H}_k(t_{m+1})$ denote the Fourier cosine coefficients of $h(t_{m+1}, x)$, i.e.,

$$\mathcal{H}_k(t_{m+1}) = \tfrac{2}{b-a}\int_a^b h(t_{m+1}, x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \tag{7.3.24}$$

With the COS formula we get

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{\Delta,m,x})\right] \approx \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1})\mathbb{E}\left[\cos\left(k\pi\tfrac{X_{m+1}^{\Delta,m,x}-a}{b-a}\right)\right]$$
$$= \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1})\Re\left\{\varphi_{X_{m+1}^{\Delta}}\left(\tfrac{k\pi}{b-a}\middle| X_m^{\Delta}=x\right)\exp\left(ik\pi\tfrac{-a}{b-a}\right)\right\} \tag{7.3.25}$$

and Fourier cosine series give

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] \approx \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1})\mathbb{E}\left[\cos\left(k\pi\tfrac{X_{m+1}^{\Delta,m,x}-a}{b-a}\right)\Delta\omega_{m+1}\right]$$
$$= \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1})\Re\left\{\mathbb{E}\left[\exp\left(ik\pi\tfrac{X_{m+1}^{\Delta,m,x}}{b-a}\right)\Delta\omega_{m+1}\right]\exp\left(ik\pi\tfrac{-a}{b-a}\right)\right\}. \tag{7.3.26}$$

Integration by parts gives

$$\begin{aligned}
&\mathbb{E}\left[\exp(iuX_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right]\\
&= \mathbb{E}\left[\exp\left(iux+ium(x)\Delta t+ius(x)\Delta\omega_{m+1}+iu\kappa(x)(\Delta\omega_{m+1})^2\right)\Delta\omega_{m+1}\right]\\
&= \tfrac{1}{\sqrt{2\pi}\sqrt{\Delta t}}\int_{\mathbb{R}}\exp\left(iux+ium(x)\Delta t+ius(x)\zeta+iu\kappa(x)\zeta^2\right)\zeta e^{-\frac{1}{2}\left(\frac{\zeta}{\sqrt{\Delta t}}\right)^2}d\zeta\\
&= \tfrac{1}{\sqrt{2\pi}\sqrt{\Delta t}}\int_{\mathbb{R}}\{s(x)+2\kappa(x)\zeta\}\Delta tD_x\exp\left(iux+ium(x)\Delta t+ius(x)\zeta+iu\kappa(x)\zeta^2\right)e^{-\frac{1}{2}\left(\frac{\zeta}{\sqrt{\Delta t}}\right)^2}d\zeta\\
&= s(x)\Delta t\mathbb{E}\left[D_x\exp(iuX_{m+1}^{\Delta,m,x})\right]+2\kappa(x)\Delta t\mathbb{E}\left[D_x\exp(iuX_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right].
\end{aligned} \tag{7.3.27}$$

Using the same procedure for the last term in (7.3.27) and iterating recursively gives us

$$\begin{aligned}
\mathbb{E}\left[\exp(iuX_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] &= s(x)\Delta t\mathbb{E}\left[D_x\exp(iuX_{m+1}^{\Delta,m,x})\right]\\
&+ s(x)\Delta t(2\kappa(x)\Delta t)\mathbb{E}\left[D_x^2\exp(iuX_{m+1}^{\Delta,m,x})\right]\\
&+ s(x)\Delta t(2\kappa(x)\Delta t)^2\mathbb{E}\left[D_x^3\exp(iuX_{m+1}^{\Delta,m,x})\right]\\
&+ s(x)\Delta t(2\kappa(x)\Delta t)^3\mathbb{E}\left[D_x^4\exp(iuX_{m+1}^{\Delta,m,x})\right]+\ldots.
\end{aligned} \tag{7.3.28}$$

It holds that

$$\begin{aligned}
\mathbb{E}\left[\frac{\partial^{(\ell)}}{\partial(X_{m+1}^{\Delta,m,x})^{(\ell)}}\exp\left(iuX_{m+1}^{\Delta,m,x}\right)\right] &= (iu)^{\ell}\mathbb{E}\left[\exp\left(iuX_{m+1}^{\Delta,m,x}\right)\right]\\
&= (iu)^{\ell}\varphi_{X_{m+1}^{\Delta}}(u|X_m^{\Delta}=x).
\end{aligned} \tag{7.3.29}$$

For the numerical experiments, it appears sufficient to take only the first two terms in (7.3.28), as the other terms are at least of order $\mathscr{O}((\Delta t)^3)$, and we end up with

$$\begin{aligned}
\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] \approx \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1})\Re\Big\{&\Big(s(x)\Delta t\left(i\tfrac{k\pi}{b-a}\right)\varphi_{X_{m+1}^{\Delta}}\left(\tfrac{k\pi}{b-a}\middle| X_m^{\Delta}=x\right)\\
&+ s(x)\Delta t2\kappa(x)\Delta t\left(i\tfrac{k\pi}{b-a}\right)^2\varphi_{X_{m+1}^{\Delta}}\left(\tfrac{k\pi}{b-a}\middle| X_m^{\Delta}=x\right)\Big)\exp\left(ik\pi\tfrac{-a}{b-a}\right)\Big\},
\end{aligned} \tag{7.3.30}$$

which enables us to approximate the conditional expectations in (7.3.20).

7.3.3.1. EXACT SIMULATION SCHEME

For, for example, the CIR process and Heston stochastic volatility models an *exact simulation scheme* exists based on directly sampling from the available transitional density function. In this section, we explain how to use exact simulation schemes to solve the discrete problem (7.3.22). The characteristic function of $X_{m+1}^{m,x}$ can be found as the Fourier transform of the density function and is denoted by $\varphi_{X_{m+1}}(u|X_m = x)$. With the COS formula we get

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x})\right] \approx \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1}) \Re\left\{\varphi_{X_{m+1}}\left(\tfrac{k\pi}{b-a}\middle| X_m = x\right) \exp\left(ik\pi \tfrac{-a}{b-a}\right)\right\}. \tag{7.3.31}$$

The question is how to approximate the expected value $\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}\right]$. Notice that $X_{m+1}^{m,x}$ and $\Delta\omega_{m+1}$ are correlated. We first use Fourier cosine series to get

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}\right] \approx \sum_{k=0}^{N-1}{}' \mathscr{H}_k(t_{m+1}) \Re\left\{\mathbb{E}\left[\exp\left(ik\pi \tfrac{X_{m+1}^{m,x}}{b-a}\right)\Delta\omega_{m+1}\right] \exp\left(ik\pi \tfrac{-a}{b-a}\right)\right\}, \tag{7.3.32}$$

and then with the theory in Section 7.2.2 and [KP92] we find

$$\begin{aligned}
&\mathbb{E}\left[\exp(iuX_{m+1}^{m,x})\Delta\omega_{m+1}\right] \\
&= \mathbb{E}\left[\sum_{\alpha\in\mathscr{A}} \mathscr{L}^\alpha \exp(iux) I_{\alpha,m+1}\Delta\omega_{m+1}\right] + \mathbb{E}\left[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[\mathscr{L}^\alpha \exp(iuX_\cdot^{m,x})]_{m+1}\Delta\omega_{m+1}\right] \\
&= \mathscr{L}^1 \exp(iux)\Delta t + \mathscr{L}^1\mathscr{L}^0 \exp(iux)\tfrac{1}{2}(\Delta t)^2 + \mathscr{L}^0\mathscr{L}^1 \exp(iux)\tfrac{1}{2}(\Delta t)^2 + \mathscr{O}((\Delta t)^3).
\end{aligned} \tag{7.3.33}$$

Notice that $\exp(iux)$ does not depend on time and

$$\mathscr{L}^1 \exp(iux) = \sigma(x) iu \exp(iux), \tag{7.3.34a}$$

$$\begin{aligned}
\mathscr{L}^1\mathscr{L}^0 \exp(iux) = \Big[&\sigma(x)\mu_x(x) iu + (\sigma(x)\mu(x) + \sigma^2(x)\sigma_x(x))(iu)^2 \\
&+ \tfrac{1}{2}\sigma(x)^3(iu)^3\Big] \exp(iux),
\end{aligned} \tag{7.3.34b}$$

$$\begin{aligned}
\mathscr{L}^0\mathscr{L}^1 \exp(iux) = \Big[&(\mu(x)\sigma_x(x) + \tfrac{1}{2}\sigma^2(x)\sigma_{xx}(x)) iu + (\mu(x)\sigma(x) \\
&+ \sigma^2(x)\sigma_x(x))(iu)^2 + \tfrac{1}{2}\sigma(x)^3(iu)^3\Big] \exp(iux).
\end{aligned} \tag{7.3.34c}$$

This enables us to approximate the conditional expectations in (7.3.22).

7.3.4. RECOVERY OF COEFFICIENTS AND ALGORITHM

Now we return to FBSDE problem (7.3.23), where we obtained deterministic functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$. Let $\mathscr{Y}_k^\Delta(t_m)$ be the Fourier cosine coefficients of $y_m^\Delta(x)$ in (7.3.23b), i.e.,

$$\mathscr{Y}_k^\Delta(t_m) = \tfrac{2}{b-a}\int_a^b y_m^\Delta(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \tag{7.3.35}$$

$\mathscr{Z}_k^\Delta(t_m)$ the Fourier cosine coefficients of function $z_m^\Delta(x)$ in (7.3.23a), i.e.,

$$\mathscr{Z}_k^\Delta(t_m) = \tfrac{2}{b-a}\int_a^b z_m^\Delta(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx, \tag{7.3.36}$$

and $\mathscr{F}_k^\Delta(t_m)$ the Fourier cosine coefficients of driver function $f(t_m, \mathbf{\Lambda}_m^\Delta(x))$, i.e.,

$$\mathscr{F}_k^\Delta(t_m) = \tfrac{2}{b-a}\int_a^b f(t_m, \mathbf{\Lambda}_m^\Delta(x)) \cos\left(k\pi \tfrac{x-a}{b-a}\right) dx. \tag{7.3.37}$$

The computation of functions $z_m^\Delta(x)$ and $y_m^\Delta(x)$ at time-point t_m requires the Fourier cosine coefficients $\mathscr{Z}_k^\Delta(t_{m+1})$, $\mathscr{Y}_k^\Delta(t_{m+1})$, and $\mathscr{F}_k^\Delta(t_{m+1})$ at time-point t_{m+1}, with COS formulas (7.3.25) and (7.3.30). The coefficients can be computed recursively backwards in time, as we explain in this section.

We start with the coefficients at the terminal time

$$\mathscr{Y}_k^\Delta(t_M) = \tfrac{2}{b-a}\int_a^b g(x) \cos\left(k\pi \tfrac{x-a}{b-a}\right) dx, \tag{7.3.38a}$$

$$\mathscr{Z}_k^\Delta(t_M) = \tfrac{2}{b-a}\int_a^b \sigma(x) D_x g(x) \cos\left(k\pi \tfrac{x-a}{b-a}\right) dx, \tag{7.3.38b}$$

$$\mathscr{F}_k^\Delta(t_M) = \tfrac{2}{b-a}\int_a^b f(t_m, x, g(x), \sigma(x) D_x g(x)) \cos\left(k\pi \tfrac{x-a}{b-a}\right) dx. \tag{7.3.38c}$$

For some problems the above integrals can be computed analytically. Otherwise we may approximate them, for example by computing the function on the x-grid and using the discrete Fourier cosine transform (see Appendix C) or another numerical integration method.

We then compute functions $z_{M-1}^\Delta(x)$, $f(t_{M-1}, \mathbf{\Lambda}_{M-1}^\Delta(x))$, and $y_{M-1}^\Delta(x)$, see equations (7.3.23), on the equidistant x-grid with N grid points. For this we use the Fourier cosine coefficients at time t_M and COS formulas (7.3.25) and (7.3.30). Subsequently the Fourier cosine coefficients $\mathscr{Z}_k^\Delta(t_{M-1})$, $\mathscr{F}_k^\Delta(t_{M-1})$, and $\mathscr{Y}_k^\Delta(t_{M-1})$ are approximated by a discrete Fourier cosine transform.

We repeat this procedure for all times t_m. So, we estimate the Fourier cosine coefficients $\mathscr{Z}_k^\Delta(t_m)$, $\mathscr{F}_k^\Delta(t_m)$, and $\mathscr{Y}_k^\Delta(t_m)$ by using the Fourier cosine coefficients at time t_{m+1} and the COS formulas. The approximation of the Fourier cosine coefficients introduces an additional error, which has been discussed in detail in [FO09]. The final approximations of the functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$ by the BCOS method are denoted by $\hat{y}_m^\Delta(x)$ and $\hat{z}_m^\Delta(x)$, respectively. The overall algorithm to solve the FBSDE backwards in time by Δ-time-discretization scheme (7.3.20) can be summarized as:

Algorithm 4. *(BCOS method)*
Initialization: *Compute, or approximate, the terminal coefficients* $\mathscr{Y}_k^\Delta(t_M)$, $\mathscr{F}_k^\Delta(t_M)$, *and* $\mathscr{Z}_k^\Delta(t_M)$, *equations* (7.3.38a), (7.3.38b), *and* (7.3.38c).
Main loop to recover $\mathscr{Z}_k^\Delta(t_m)$, $\mathscr{F}_k^\Delta(t_m)$, ***and*** $\mathscr{Y}_k^\Delta(t_m)$***:***
For $m = M-1$ *to* 1*:*

- *Compute the functions* $\hat{z}_m^\Delta(x)$, $f(t_m, \hat{\mathbf{\Lambda}}_m^\Delta(x))$, *and* $\hat{y}_m^\Delta(x)$, *see equations* (7.3.23), *with COS formulas* (7.3.25) *and* (7.3.30).
- *Determine the corresponding Fourier cosine coefficients* $\mathscr{Z}_k^\Delta(t_m)$, $\mathscr{F}_k^\Delta(t_m)$, *and* $\mathscr{Y}_k^\Delta(t_m)$ *by using a DCT (Appendix C).*

Final step: *Compute* $\hat{z}_0^\Delta(X_0)$ *and* $\hat{y}_0^\Delta(X_0)$.

The algorithm for the exact simulation scheme, with final values $\hat{z}_0^\theta(X_0)$ and $\hat{y}_0^\theta(X_0)$, is similar, except that we use COS formulas (7.3.31) and (7.3.32)-(7.3.33).

Remark 7.3.1. *For constant drift μ and volatility σ terms we can apply the efficient Fast Fourier transform (FFT) algorithm to recover the Fourier cosine coefficients, as explained in Section 6.4.4.*

7.4. ERROR ANALYSIS

In this section we perform an error analysis of the discretization with timesteps Δt for the 2.0-weak-Taylor scheme and parameter $\theta = 1/2$. The convergence in N, the number of Fourier coefficients, is of second order, due to the use of discrete Fourier cosine transforms. The error of the Fourier cosine formulas has been analyzed in the previous chapters and we refer to them for more details. We start with Itô-Taylor expansions and related expected values in Section 7.4.1. In Section 7.4.2 we look at the local errors related to the discretization scheme for the FSDE. In Section 7.4.3 we discuss the local errors of the θ-time-discretization. A global error result is presented in Section 7.4.4.

7.4.1. ITÔ-TAYLOR EXPANSION

The Itô-Taylor expansion of a general sufficiently smooth function $h(t,x)$ reads ([KP92])

$$h(t_{m+1}, X_{m+1}^{m,x}) = \sum_{\alpha\in\mathscr{A}} h_\alpha(t_m,x) I_{\alpha,m+1} + \sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[h_\alpha(.,X_.^{m,x})]_{m+1}, \tag{7.4.1}$$

$$\begin{aligned} h(t_{m+2}, X_{m+2}^{m,x}) &= \sum_{\beta\in\mathscr{A}}\sum_{\alpha\in\mathscr{A}} h_{\alpha\beta}(t_m,x) I_{\alpha,m+1} I_{\beta,m+2} \\ &\quad + \sum_{\beta\in\mathscr{A}}\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[h_{\alpha\beta}(X_.^{m,x})]_{m+1} I_{\beta,m+2} + \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[h_\beta(.,X_.^{m,x})]_{m+2}, \end{aligned} \tag{7.4.2}$$

with $\mathscr{A}$ a hierarchical set.

Lemma 7.4.1. *For a sufficiently smooth function $h(t,x)$ we have the following conditional expectations of the Itô-Taylor expansion*

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x})\right] = h(t_m,x) + h_{(0)}(t_m,x)\Delta t + h_{(0,0)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + \mathscr{O}((\Delta t)^3), \tag{7.4.3a}$$

$$\begin{aligned} &\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}\right] \\ &\quad = h_{(1)}(t_m,x)\Delta t + \left[h_{(1,0)}(t_m,x) + h_{(0,1)}(t_m,x)\right]\tfrac{1}{2}(\Delta t)^2 + \mathscr{O}((\Delta t)^3), \end{aligned} \tag{7.4.3b}$$

$$\mathbb{E}\left[h(t_{m+2}, X_{m+2}^{m,x})\right] = h(t_m,x) + h_{(0)}(t_m,x)2\Delta t + h_{(0,0)}(t_m,x)2(\Delta t)^2 + \mathscr{O}((\Delta t)^3), \tag{7.4.3c}$$

$$\begin{aligned} &\mathbb{E}\left[h(t_{m+2}, X_{m+2}^{m,x})\Delta\omega_{m+1}\right] \\ &\quad = h_{(1)}(t_m,x)\Delta t + h_{(0,1)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + h_{(1,0)}(t_m,x)\tfrac{3}{2}(\Delta t)^2 + \mathscr{O}((\Delta t)^3), \end{aligned} \tag{7.4.3d}$$

$$\begin{aligned} &\mathbb{E}\left[h(t_{m+2}, X_{m+2}^{m,x})\Delta\omega_{m+2}\right] \\ &\quad = h_{(1)}(t_m,x)\Delta t + h_{(0,1)}(t_m,x)\tfrac{3}{2}(\Delta t)^2 + h_{(1,0)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + \mathscr{O}((\Delta t)^3), \end{aligned} \tag{7.4.3e}$$

$$\begin{aligned} &\tfrac{2}{\Delta t}\mathbb{E}\left[h(t_{m+2}, X_{m+2}^{m,x})\left(\Delta\omega_{m+1} - \Delta\omega_{m+2}\right)\right] \\ &\quad = \tfrac{2}{\Delta t}\left[h_{(1,0)}(t_m,x) - h_{(0,1)}(t_m,x)\right](\Delta t)^2 + \mathscr{O}((\Delta t)^2). \end{aligned} \tag{7.4.3f}$$

This lemma can be proved with the help of [KP92], Chapter 5.7. The definitions of the operators $\mathscr{L}^0$ and $\mathscr{L}^1$, equation (7.2.7), give

$$h_{(1,0)}(t_m,x) - h_{(0,1)}(t_m,x) = D_x h(t_m,x)\left[\sigma(x)\mu_x(x) - \mu(x)\sigma_x(x) - \tfrac{1}{2}\sigma^2(x)\sigma_{xx}(x)\right]. \tag{7.4.4}$$

7.4.2. LOCAL ERROR Δ-TIME-DISCRETIZATION FSDE

The weak convergence rate of the 2.0-weak-Taylor scheme is denoted by $\gamma_w^{2T} = 2$.

Lemma 7.4.2. *For a sufficiently smooth function $h(t,x)$ we have the following local weak errors of the 2.0-weak-Taylor scheme*

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x})\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.4.5a}$$

$$\mathbb{E}\left[\left(h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x})\right)\Delta\omega_{m+1}\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.4.5b}$$

$$\mathbb{E}\left[h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta,m,x})\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.4.5c}$$

$$\mathbb{E}\left[\left(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta,m,x})\right)\Delta\omega_{m+1}\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.4.5d}$$

$$\mathbb{E}\left[\left(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta,m,x})\right)\Delta\omega_{m+2}\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.4.5e}$$

$$\tfrac{2}{\Delta t}\mathbb{E}\left[(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta,m,x}))(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{7.4.5f}$$

For a proof we refer to Appendix 7.A.

7.4.3. LOCAL ERROR θ-TIME-DISCRETIZATION SCHEME

The equation for $y_m(x)$ is given by (see (7.3.18))

$$y_m(x) = \mathbb{E}\left[y_{m+1}(X_{m+1}^{m,x})\right] + \Delta t\tfrac{1}{2} f(t_m, \mathbf{\Lambda}_m(x)) + \Delta t\tfrac{1}{2}\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))\right] + R_m^y(x), \tag{7.4.6}$$

and the equation for $z_m(x)$ reads (see (7.3.19))

$$\begin{aligned} z_m(x) = {} & \tfrac{2}{\Delta t}\mathbb{E}\left[y_{m+1}(X_{m+1}^{m,x})\Delta\omega_{m+1}\right] + \mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))\Delta\omega_{m+1}\right] \\ & - \mathbb{E}\left[z_{m+1}(X_{m+1}^{m,x})\right] + \tfrac{2}{\Delta t} R_m^z(x), \end{aligned} \tag{7.4.7}$$

with θ-discretization errors

$$\begin{aligned} R_m^y(x) = {} & \int_{t_m}^{t_{m+1}} \mathbb{E}\left[f(s, \mathbf{\Lambda}_s(X_s^{m,x}))\right] ds \\ & - \left\{\Delta t\tfrac{1}{2} f(t_m, \mathbf{\Lambda}_m(x)) + \Delta t\tfrac{1}{2}\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))\right]\right\}, \end{aligned} \tag{7.4.8}$$

$$\begin{aligned} R_m^z(x) = {} & \int_{t_m}^{t_{m+1}} \mathbb{E}\left[f(s, \mathbf{\Lambda}_s(X_s^{m,x}))(\omega_s - \omega_{t_m})\right] ds - \left\{\Delta t\tfrac{1}{2}\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))\Delta\omega_{m+1}\right]\right\} \\ & - \int_{t_m}^{t_{m+1}} \mathbb{E}\left[z_s(X_s^{m,x})\right] ds + \left\{\Delta t\tfrac{1}{2} z_m(x) + \Delta t\tfrac{1}{2}\mathbb{E}\left[z_{m+1}(X_{m+1}^{m,x})\right]\right\}. \end{aligned} \tag{7.4.9}$$

Here x denotes the exact solution of the FSDE (7.2.2) at time t_m, in other words, $x = X_m$.

Lemma 7.4.3. *For sufficiently smooth functions $f(.)$ and $g(.)$ the θ-discretization errors are of order*

$$R_m^y(x) = \mathscr{O}((\Delta t)^3), \tag{7.4.10a}$$

$$R_m^z(x) = \mathscr{O}((\Delta t)^3). \tag{7.4.10b}$$

Besides,

$$\mathbb{E}\left[R^y_{m+1}(X^{m,x}_{m+1})\Delta\omega_{m+1}\right] = \mathcal{O}((\Delta t)^4), \tag{7.4.11a}$$
$$\mathbb{E}\left[R^z_{m+1}(X^{m,x}_{m+1})\right] - R^z_m(x) = \mathcal{O}((\Delta t)^4). \tag{7.4.11b}$$

Proof. For equations (7.4.10a) and (7.4.10b) we use that for a general sufficiently smooth function $h(t,x)$, with bounded derivatives, it holds that

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_m}^{t_{m+1}} h(s,X^{m,x}_s)ds - \Delta t\tfrac{1}{2}\left(h(t_m,x)+h(t_{m+1},X^{m,x}_{m+1})\right)\right]\\
&\quad= h(t_m,x)\Delta t + h_{(0)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + h_{(0,0)}(t_m,x)\tfrac{1}{3!}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right)\\
&\quad- \Delta t\tfrac{1}{2}\left(h(t_m,x)+h(t_m,x)+h_{(0)}(t_m,x)\Delta t + h_{(0,0)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + \mathcal{O}\left((\Delta t)^3\right)\right)\\
&\quad= h_{(0,0)}(t_m,x)\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right) = \mathcal{O}\left((\Delta t)^3\right).
\end{aligned} \tag{7.4.12}$$

For equation (7.4.10b) we use:

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_m}^{t_{m+1}} h(s,X^{m,x}_s)\Delta\omega_{m+1}ds - \Delta t\tfrac{1}{2}h(t_{m+1},X^{m,x}_{m+1})\Delta\omega_{m+1}\right]\\
&\quad= h_{(1)}(t_m,x)\tfrac{1}{2}(\Delta t)^2 + \left[h_{(1,0)}(t_m,x)+h_{(0,1)}(t_m,x)\right]\tfrac{1}{3!}(\Delta t)^3 + \mathcal{O}((\Delta t)^4)\\
&\quad- \Delta t\tfrac{1}{2}\left(h_{(1)}(t_m,x)\Delta t + \left[h_{(1,0)}(t_m,x)+h_{(0,1)}(t_m,x)\right]\tfrac{1}{2}(\Delta t)^2 + \mathcal{O}((\Delta t)^3)\right)\\
&\quad= \left[h_{(1,0)}(t_m,x)+h_{(0,1)}(t_m,x)\right]\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right) = \mathcal{O}\left((\Delta t)^3\right).
\end{aligned} \tag{7.4.13}$$

For equation (7.4.11a) we use:

$$\begin{aligned}
&\mathbb{E}\left[\Delta\omega_{m+1}\left(\int_{t_{m+1}}^{t_{m+2}} h(s,X^{m,x}_s)ds - \tfrac{1}{2}\Delta t\left(h(t_{m+1},X^{m,x}_{m+1})+h(t_{m+2},X^{m,x}_{m+2})\right)\right)\right]\\
&\quad= \mathbb{E}\left[\Delta\omega_{m+1}\mathbb{E}\left[\int_{t_{m+1}}^{t_{m+2}} h(s,X^{m+1,X^{m,x}_{m+1}}_s)ds - \Delta t\tfrac{1}{2}\left(h(t_{m+1},X^{m,x}_{m+1})+h(t_{m+2},X^{m+1,X^{m,x}_{m+1}}_{m+2})\right)\right]\right]\\
&\quad= \mathbb{E}\left[\Delta\omega_{m+1}h_{(0,0)}(t_{m+1},X^{m,x}_{m+1})\right]\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right)\\
&\quad= h_{(1,0,0)}(t_m,x)\tfrac{-1}{12}(\Delta t)^4 + \mathcal{O}\left((\Delta t)^4\right) = \mathcal{O}\left((\Delta t)^4\right).
\end{aligned} \tag{7.4.14}$$

For equation (7.4.11b) we use:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\int_{t_{m+1}}^{t_{m+2}} h(s,X^{m+1,X^{m,x}_{m+1}}_s)ds - \Delta t\tfrac{1}{2}\left(h(t_{m+1},X^{m,x}_{m+1})+h(t_{m+2},X^{m+1,X^{m,x}_{m+1}}_{m+2})\right)\right]\right]\\
&\quad- \mathbb{E}\left[\int_{t_m}^{t_{m+1}} h(s,X^{m,x}_s)ds - \Delta t\tfrac{1}{2}\left(h(t_m,x)+h(t_{m+1},X^{m,x}_{m+1})\right)\right]\\
&\quad= \mathbb{E}\left[h_{(0,0)}(t_{m+1},X^{m,x}_{m+1})\right]\tfrac{-1}{12}(\Delta t)^3 - h_{(0,0)}(t_m,x)\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right)\\
&\quad= \left(h_{(0,0)}(t_m,x) + h_{(0,0,0)}(t_m,x)\Delta t + \mathcal{O}\left((\Delta t)^2\right)\right)\tfrac{-1}{12}(\Delta t)^3\\
&\quad- h_{(0,0)}(t_m,x)\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right)\\
&\quad= h_{(0,0,0)}(t_m,x)\tfrac{-1}{12}(\Delta t)^4 + \mathcal{O}\left((\Delta t)^4\right) = \mathcal{O}\left((\Delta t)^4\right)
\end{aligned} \tag{7.4.15}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\int_{t_{m+1}}^{t_{m+2}} h(s, X_s^{m+1,X_{m+1}^{m,x}})\Delta\omega_{m+2}ds - \Delta t\tfrac{1}{2}h(t_{m+2}, X_{m+2}^{m+1,X_{m+1}^{m,x}})\Delta\omega_{m+2}\right]\right] \\
&\quad - \mathbb{E}\left[\int_{t_m}^{t_{m+1}} h(s, X_s^{m,x})\Delta\omega_{m+1}ds - \Delta t\tfrac{1}{2}h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}\right] \\
&\quad = \mathbb{E}\left[h_{(1,0)}(t_{m+1}, X_{m+1}^{m,x}) + h_{(0,1)}(t_{m+1}, X_{m+1}^{m,x})\right]\tfrac{-1}{12}(\Delta t)^3 \\
&\quad - \left[h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)\right]\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right) \\
&\quad = \left(\left[h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)\right] + \left[h_{(0,1,0)}(t_m, x) + h_{(0,0,1)}(t_m, x)\right]\Delta t + \mathcal{O}\left((\Delta t)^2\right)\right)\tfrac{-1}{12}(\Delta t)^3 \\
&\quad - \left[h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)\right]\tfrac{-1}{12}(\Delta t)^3 + \mathcal{O}\left((\Delta t)^4\right) \\
&\quad = \left[h_{(0,1,0)}(t_m, x) + h_{(0,0,1)}(t_m, x)\right]\tfrac{-1}{12}(\Delta t)^4 + \mathcal{O}\left((\Delta t)^4\right) = \mathcal{O}\left((\Delta t)^4\right). \qquad (7.4.16)
\end{aligned}$$

□

7.4.4. GLOBAL ERROR Δ-TIME-DISCRETIZATION SCHEME FBSDE

The equation for $y_m^\Delta(x)$ is now given by

$$y_m^\Delta(x) = \mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right] + \Delta t\tfrac{1}{2}f(t_m, \mathbf{\Lambda}_m^\Delta(x)) + \Delta t\tfrac{1}{2}\mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\right], \qquad (7.4.17)$$

and the equation for $z_m^\Delta(x)$ reads

$$z_m^\Delta(x) = \tfrac{2}{\Delta t}\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] + \mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\Delta\omega_{m+1}\right] - \mathbb{E}\left[z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right]. \qquad (7.4.18)$$

We define the following global errors,

$$\epsilon_m^y(X_m, X_m^\Delta) := y_m(X_m) - y_m^\Delta(X_m^\Delta), \qquad (7.4.19a)$$
$$\epsilon_m^z(X_m, X_m^\Delta) := z_m(X_m) - z_m^\Delta(X_m^\Delta), \qquad (7.4.19b)$$
$$\epsilon_m^f(X_m, X_m^\Delta) := f(t_m, \mathbf{\Lambda}_m(X_m)) - f(t_m, \mathbf{\Lambda}_m^\Delta(X_m^\Delta)), \qquad (7.4.19c)$$

and

$$\epsilon_m^y(x) := y_m(x) - y_m^\Delta(x), \qquad (7.4.20a)$$
$$\epsilon_m^z(x) := z_m(x) - z_m^\Delta(x), \qquad (7.4.20b)$$
$$\epsilon_m^f(x) := f(t_m, \mathbf{\Lambda}_m(x)) - f(t_m, \mathbf{\Lambda}_m^\Delta(x)). \qquad (7.4.20c)$$

For the error at time t_m we find

$$\begin{aligned}
\epsilon_m^y(X_m, X_m^\Delta) &= y_m(X_m) - y_m^\Delta(X_m) + y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta) \\
&= \epsilon_m^y(X_m) + y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta), \qquad (7.4.21a) \\
|\mathbb{E}_0[\epsilon_m^y(X_m, X_m^\Delta)]| &\le |\mathbb{E}_0[\epsilon_m^y(X_m)]| + |\mathbb{E}_0[y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta)]|. \qquad (7.4.21b)
\end{aligned}$$

The second part is the weak error and, assuming that $y_m^\Delta(x)$ and $z_m^\Delta(x)$ are $2(\gamma_w^{2T}+1)$-times continuously differentiable with polynomial growth, we have

$$|\mathbb{E}_0[\epsilon_m^y(X_m, X_m^\Delta)]| \le |\mathbb{E}_0[\epsilon_m^y(X_m)]| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}). \qquad (7.4.22)$$

Similarly, we get

$$|\mathbb{E}_0[\epsilon_m^z(X_m, X_m^\Delta)]| \le |\mathbb{E}_0[\epsilon_m^z(X_m)]| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}). \tag{7.4.23}$$

We find the following bounds on the errors $\epsilon_m^y(x)$ and $\epsilon_m^z(x)$.

Theorem 7.4.1. *Given*

$$\mathbb{E}_{M-1}[|\epsilon_M^y(X_M)|] \sim \mathcal{O}\left((\Delta t)^3\right), \quad \mathbb{E}_{M-1}^x[|\epsilon_M^z(X_M)|] \sim \mathcal{O}\left((\Delta t)^3\right), \tag{7.4.24}$$

then

$$\mathbb{E}_0\left[|\epsilon_m^y(X_m)| + \Delta t|\epsilon_m^z(X_m)|\right] \le \mathcal{O}\left((\Delta t)^2\right). \tag{7.4.25}$$

Proof. **Error Y** For the Y-component we have with (7.4.6) and (7.4.17)

$$\begin{aligned}\epsilon_m^y(x) = y_m(x) - y_m^\Delta(x) &= \mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\right] + \Delta t\tfrac{1}{2}\epsilon_m^f(x) \\ &+ \Delta t\tfrac{1}{2}\mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\right] + R_m^y(x).\end{aligned} \tag{7.4.26}$$

It follows, with equality (7.4.5a), that

$$\begin{aligned}\mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\right] &= \mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x})\right] + \mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{m,x}) - y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right] \\ &= \mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x})\right] + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \qquad (7.4.27a) \\ \mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\right] &= \mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x})\right] \\ &+ \mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{m,x})) - f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\right] \\ &= \mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x})\right] + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \qquad (7.4.27b)\end{aligned}$$

Driver function $f(t,x,y,z)$ is Lipschitz in x, y, and z, so that

$$|\epsilon_m^f(x)| \le L_f\left(|\epsilon_m^y(x)| + |\epsilon_m^z(x)|\right), \tag{7.4.28a}$$

$$\left|\mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x})\right]\right| \le L_f\mathbb{E}\left[|\epsilon_{m+1}^y(X_{m+1}^{m,x})| + |\epsilon_{m+1}^z(X_{m+1}^{m,x})|\right]. \tag{7.4.28b}$$

Then,

$$\begin{aligned}|\epsilon_m^y(x)| \le{}& \frac{1+\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}|\mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x})\right]| + \frac{\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}|\epsilon_m^z(x)| \\ &+ \frac{\frac{1}{2}\Delta t L_f}{1-\frac{1}{2}\Delta t L_f}\mathbb{E}\left[|\epsilon_{m+1}^z(X_{m+1}^{m,x})|\right] + \frac{1}{1-\frac{1}{2}\Delta t L_f}\mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}) + \frac{1}{1-\frac{1}{2}\Delta t L_f}\mathcal{O}((\Delta t)^3).\end{aligned} \tag{7.4.29}$$

Error Z For the Z-component we have, with (7.4.7) and (7.4.18)

$$\begin{aligned}\epsilon_m^z(x) &= z_m(x) - z_m^\Delta(x) \\ &= \tfrac{2}{\Delta t}\mathbb{E}\left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] + \mathbb{E}\left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] \\ &- \mathbb{E}\left[\epsilon_{m+1}^z(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x})\right] + \tfrac{2}{\Delta t}R_m^z(x).\end{aligned} \tag{7.4.30}$$

Substituting the similar equations for ϵ^y_{m+1} and ϵ^z_{m+1} as in (7.4.26) and (7.4.30) gives,

$$\begin{aligned}\epsilon^z_m(x) = \tfrac{2}{\Delta t}\mathbb{E}\left[\epsilon^y_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+1}\right] + \mathbb{E}\left[\epsilon^f_{m+1}(X^{m,x}_{m+1}, X^{\Delta,m,x}_{m+1})\Delta\omega_{m+1}\right] \\ + \mathbb{E}\left[\epsilon^f_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+1}\right] + \tfrac{2}{\Delta t}\mathbb{E}\left[R^y_{m+1}(X^{m,x}_{m+1})\Delta\omega_{m+1}\right] \\ + \mathbb{E}\left[\epsilon^f_{m+1}(X^{m,x}_{m+1}, X^{\Delta,m,x}_{m+1})\Delta\omega_{m+1}\right] - \tfrac{2}{\Delta t}\mathbb{E}\left[\epsilon^y_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+2}\right] \\ - \mathbb{E}\left[\epsilon^f_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+2}\right] + \mathbb{E}\left[\epsilon^z_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\right] \\ - \tfrac{2}{\Delta t}\mathbb{E}\left[R^z_{m+1}(X^{m,x}_{m+1})\right] + \tfrac{2}{\Delta t}R^z_m(x). \end{aligned} \tag{7.4.31}$$

The different terms in the above equation can be bounded as follows:

- With equality (7.4.5c),

$$\left|\mathbb{E}\left[\epsilon^z_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\right]\right| \leq \left|\mathbb{E}\left[\epsilon^z_{m+2}(X^{m,x}_{m+2})\right]\right| + \mathcal{O}((\Delta t)^{\gamma^{2T}_w+1}). \tag{7.4.32}$$

- With equality (7.4.5b),

$$\left|2\mathbb{E}\left[\epsilon^f_{m+1}(X^{m,x}_{m+1}, X^{\Delta,m,x}_{m+1})\Delta\omega_{m+1}\right]\right| \leq C\sqrt{\Delta t}|\epsilon^f_{m+1}(X^{m,x}_{m+1})|_\infty + \mathcal{O}((\Delta t)^{\gamma^{2T}_w+1}), \tag{7.4.33}$$

  where $|.|_\infty$ denotes the infinite norm and $C > 0$ a constant.

- With equality (7.4.5d),

$$\left|\mathbb{E}\left[\epsilon^f_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+1}\right]\right| \leq C\sqrt{\Delta t}|\epsilon^f_{m+2}(X^{m,x}_{m+2})|_\infty + \mathcal{O}((\Delta t)^{\gamma^{2T}_w+1}). \tag{7.4.34}$$

- With equality (7.4.5e),

$$\left|\mathbb{E}\left[\epsilon^f_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})\Delta\omega_{m+2}\right]\right| \leq C\sqrt{\Delta t}|\epsilon^f_{m+2}(X^{m,x}_{m+2})|_\infty + \mathcal{O}((\Delta t)^{\gamma^{2T}_w+1}). \tag{7.4.35}$$

- With[4] equality (7.4.5f),

$$\begin{aligned}&\left|\tfrac{2}{\Delta t}\mathbb{E}\left[\epsilon^y_{m+2}(X^{m,x}_{m+2}, X^{\Delta,m,x}_{m+2})(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right]\right| \\ &\qquad \leq \tfrac{2}{\Delta t}\left|\mathbb{E}\left[\epsilon^y_{m+2}(X^{m,x}_{m+2})(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right]\right| + \mathcal{O}((\Delta t)^{\gamma^{2T}_w+1}). \end{aligned} \tag{7.4.36}$$

  We use the following inequality (see equation (7.4.3f))

$$\tfrac{2}{\Delta t}\left|\mathbb{E}\left[\epsilon^y_{m+2}(X^{m,x}_{m+2})(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right]\right| \leq C\Delta t|\sigma(x)\epsilon^{D_x y}_m(x)|, \tag{7.4.37}$$

  with

$$\begin{aligned}\sigma(x)\epsilon^{D_x y}_m(x) &:= \sigma(x)D_x y_m(x) - \sigma(x)D_x y^\Delta_m(x) \\ &= z_m(x) - z^\Delta_m(x) + z^\Delta_m(x) - \mathscr{L}^1 y^\Delta_m(x) \\ &= \epsilon^z_m(x) + z^\Delta_m(x) - \mathscr{L}^1 y^\Delta_m(x) \end{aligned} \tag{7.4.38}$$

[4] For constant drift $\mu(.)$ and volatility $\sigma(.)$ this term cancels, see Chapter 6.

and

$$\begin{aligned} z_m^\Delta(x) - \mathscr{L}^1 y_m^\Delta(x) &= \tfrac{2}{\Delta t}\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\Delta\omega_{m+1}\right] \\ &\quad + \mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))\Delta\omega_{m+1}\right] - \mathbb{E}\left[z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})\right] - \mathscr{L}^1 y_m^\Delta(x) \\ &= \tfrac{2}{\Delta t}\mathbb{E}\left[y_{m+1}^\Delta(X_{m+1}^{m,x})\Delta\omega_{m+1}\right] + \mathbb{E}\left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{m,x}))\Delta\omega_{m+1}\right] \\ &\quad - \mathbb{E}\left[z_{m+1}^\Delta(X_{m+1}^{m,x})\right] + \mathscr{O}((\Delta t)^{\gamma_w^{2T}}) - \mathscr{L}^1 y_m^\Delta(x). \end{aligned} \tag{7.4.39}$$

With equations (7.4.3a) and (7.4.3b) we find

$$\begin{aligned} z_m^\Delta(x) - \mathscr{L}^1 y_m^\Delta(x) &\approx \tfrac{2}{\Delta t}\mathscr{L}^1 y_m^\Delta(x)\Delta t + \tfrac{2}{\Delta t}[\mathscr{L}^1\mathscr{L}^0 y_m^\Delta(x) + \mathscr{L}^0\mathscr{L}^1 y_m^\Delta(x)]\tfrac{1}{2}(\Delta t)^2 \\ &\quad + \mathscr{L}^1 f(t_m, \mathbf{\Lambda}_m^\Delta(x))\Delta t + [\mathscr{L}^1\mathscr{L}^0 f(t_m, \mathbf{\Lambda}_m^\Delta(x)) + \mathscr{L}^0\mathscr{L}^1 f(t_m, \mathbf{\Lambda}_m^\Delta(x))]\tfrac{1}{2}(\Delta t)^2 \\ &\quad - z_m^\Delta(x) - \mathscr{L}^0 z_m^\Delta(x)\Delta t + \mathscr{O}((\Delta t)^{\gamma_w^{2T}}) - \mathscr{L}^1 y_m^\Delta(x). \end{aligned} \tag{7.4.40}$$

We deduce that $z_m^\Delta(x) - \mathscr{L}^1 y_m^\Delta(x) = \mathscr{O}(\Delta t)$ and

$$\tfrac{2}{\Delta t}\left|\mathbb{E}\left[\epsilon_{m+2}^y(X_{m+2}^{m,x})(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right]\right| \le C\Delta t|\epsilon_m^z(x)| + \mathscr{O}\left((\Delta t)^2\right). \tag{7.4.41}$$

We can now bound the absolute error by

$$\begin{aligned} (1 - C\Delta t)|\epsilon_m^z(x)| &\le |\mathbb{E}\left[\epsilon_{m+2}^z(X_{m+2}^{m,x})\right]| + c\sqrt{\Delta t}|\epsilon_{m+2}^f(X_{m+2}^{m,x})|_\infty \\ &\quad + c\sqrt{\Delta t}|\epsilon_{m+1}^f(X_{m+1}^{m,x})|_\infty + \tfrac{2}{\Delta t}\left|\mathbb{E}\left[\epsilon_{m+2}^y(X_{m+2}^{m,x})(\Delta\omega_{m+1} - \Delta\omega_{m+2})\right]\right| \\ &\quad + \mathscr{O}((\Delta t)^2) + \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \end{aligned} \tag{7.4.42}$$

Total error Summing up the errors gives us, for $\Delta t \le 1$, $(m \le M-3)$

$$\begin{aligned} &\mathbb{E}\left[|\epsilon_m^y(x)|_\infty + \Delta t|\epsilon_m^z(x)|_\infty + \Delta t|\epsilon_{m+1}^z(X_{m+1}^{m,x})|_\infty\right] \\ &\quad \le A\mathbb{E}\left[|\epsilon_{m+2}^y(X_{m+2}^{m,x})|_\infty + \Delta t|\epsilon_{m+2}^z(X_{m+2}^{m,x})|_\infty + \Delta t|\epsilon_{m+3}^z(X_{m+3}^{m,x})|_\infty\right] + B \\ &\text{with} \quad A = 1 + Q\Delta t, \\ &\quad B = \mathscr{O}((\Delta t)^3) + \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \end{aligned} \tag{7.4.43}$$

with $Q > 0$ a constant. Iterating this equality results in Theorem 7.4.1. □

7

Theorem 7.4.1 states second-order convergence for Y and first-order convergence for Z. However, in our numerical experiments we also find second-order convergence for Z if we apply the 2.0-weak-Taylor scheme and $\theta = 1/2$. The authors of [ZZJ14] apply an Order 2.0 weak Taylor scheme and a slightly different θ-scheme and they also obtain second-order convergence in their numerical experiments, for which they use a Gauss-Hermite quadrature rule to approximate the conditional expectation. In our case a Fourier-based method results in a very efficient numerical scheme. For the scheme with $\theta = 1$ the θ-discretization errors are one order lower and we find first-order convergence. For the Euler and Milstein schemes, the weak convergence rate $\gamma_w^{2T} = 2$ should be replaced by their weak convergence rates $\gamma_w^E = 1$ and $\gamma_w^M = 1$, respectively, and we find, for both $\theta = 1/2$ and $\theta = 1$, first order convergence.

7.5. NUMERICAL EXPERIMENTS FBSDE

In this section we discuss numerical experiments. MATLAB 7.11.0 is used for the computations. In the experiments we use $\theta = 1/2$ and $\theta = 1$. We prescribe a computational domain $[a,b]$ by

$$[a,b] = \left[\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}\right], \tag{7.5.1}$$

with cumulants of one Euler step $\kappa_1 = X_0 + \mu(X_0)T$ and $\kappa_2 = \sigma^2(X_0)T$, and $L = 10$. Furthermore, we set the number of terms in the Fourier cosine series expansions equal to $N = 2^9$. For these values the BCOS method has converged in N to machine precision.

7.5.1. EXAMPLE 1

The first example is derived from [MT06, MSZ08]. We take drift and diffusion term

$$\mu(x) = \tfrac{x(1+x^2)}{(2+x^2)^3}, \quad \sigma(x) = \tfrac{1+x^2}{2+x^2}. \tag{7.5.2}$$

The driver function and terminal condition are given by

$$f(t,x,y,z) = \tfrac{1}{t+1}\exp\left(-\tfrac{x^2}{t+1}\right)\left(\tfrac{4x^2(1+x^2)}{(2+x^2)^3} + \tfrac{(1+x^2)^2}{(2+x^2)^2}\left(1-\tfrac{2x^2}{t+1}\right) - \tfrac{x^2}{t+1}\right)$$
$$- \tfrac{zx}{(2+x^2)^2}\sqrt{\tfrac{1+y^2+\exp\left(-\frac{2x^2}{t+1}\right)}{1+2y^2}}, \tag{7.5.3a}$$

$$g(x) = \exp\left(-\tfrac{x^2}{T+1}\right). \tag{7.5.3b}$$

The exact solution reads $v(t,x) = \exp\left(-\frac{x^2}{t+1}\right)$. For the experiment we use parameter values $T = 10$ and $X_0 = 1$. The results are shown in Figure 7.5.1. Only with the 2.0-weak-Taylor scheme and $\theta = 1/2$ we achieve second-order convergence in M, as expected. Notice that $\kappa(x) = 0$ for $x = 0$, and then the characteristic function is given by equation (7.3.6). The characteristic function decreases at a lower exponential rate in u for $x \approx 0$. This gives rise to an unstable behavior when N is not sufficiently high. In the figure this is visible for values $M < 4$.

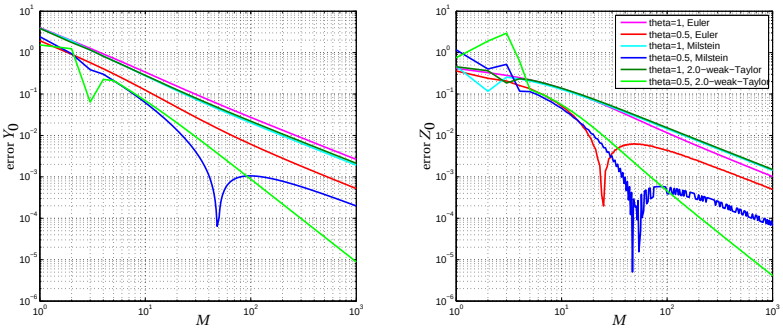


Figure 7.5.1: Results example 1, left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

Table 7.5.1 shows CPU times, where the small values demonstrate the efficiency of the BCOS method. Computation of the characteristic function and the function f on an x-grid

are the most time-consuming parts of the algorithm. The computation time is linear in the number of timesteps M and of $\mathcal{O}(N^2)$ in the number of terms in the Fourier cosine series expansions. The latter is not clearly measurable for these values of N as a significant part of the computation time is spend on discrete Fourier cosine transforms, which is of order $\mathcal{O}(N \log N)$.

Table 7.5.1: CPU time (s).

$(N = 2^9)$ M	4	8	16	32	64	128	256	512
Euler $\theta = 1$	0.0376	0.0507	0.0563	0.0852	0.1442	0.2620	0.5005	0.9958
2.0-weak-Taylor $\theta = \frac{1}{2}$	0.0976	0.1247	0.1249	0.1738	0.2738	0.4698	0.8668	1.6545

$(M = 256)$ N	2^6	2^7	2^8	2^9
Euler $\theta = 1$	0.1760	0.2028	0.2661	0.5005
2.0-weak-Taylor $\theta = \frac{1}{2}$	0.2068	0.2438	0.4711	0.8668

7.5.2. EXAMPLE 2: EUROPEAN CALL OPTION - CEV - $\mathbb{P}$-MEASURE

In the second example we compute the price $v(t, X_t)$ of a European call option where the underlying asset follows the CEV asset price process,

$$dX_s = \bar{\mu} X_s ds + \bar{\sigma} X_s^{\gamma} d\omega_s, \tag{7.5.4}$$

with $\gamma > 0$. The exact solution is given by the CEV price [Hul09]. For the derivation of the corresponding semilinear PDE we set up a self-financing portfolio Y_s with a_s assets and bonds with risk-free return rate r. Markets are assumed to be complete in this model, there are no trading restrictions, and the option can be exactly replicated by the hedging portfolio, that is, $Y_T = \max(X_T - K, 0)$. Then, the option value at initial time should be equal to the initial value of the portfolio. The hedge portfolio evolves according to the SDE

$$dY_s = r(Y_s - a_s X_s)ds + a_s dX_s = \big(rY_s + (\mu(X_s) - rX_s)a_s\big)\, ds + \sigma(X_s) a_s d\omega_s. \tag{7.5.5}$$

If we set $Z_s = \sigma(X_s) a_s$, then (Y, Z) solve the FBSDE

$$dY_s = -f(s, X_s, Y_s, Z_s)ds + Z_s d\omega_s, \quad Y_T = \max(X_T - K, 0), \tag{7.5.6a}$$

$$f(t, x, y, z) = -ry - \tfrac{\mu(x) - rx}{\sigma(x)} z = -ry - \tfrac{\bar{\mu} - r}{\bar{\sigma}} x^{1-\gamma} z. \tag{7.5.6b}$$

Y_t corresponds to the value of the portfolio and Z_t is related to the hedging strategy. The option value is given by $v(t, X_t) = Y_t$ and $\sigma(X_t) D_x v(t, X_t) = Z_t$.

For the tests, we use the following parameter values

$$X_0 = 100,\ K = 100,\ r = 0.1,\ \bar{\mu} = 0.2,\ T = 0.1. \tag{7.5.7}$$

We take the elasticity of variance equal to $\gamma = 0.2$ and $\gamma = 0.8$ and choose $\bar{\sigma}$ so that $\sigma(X_0) = 25$.

As the terminal coefficients $\mathcal{Z}_k^{\Delta}(t_M)$ (equation (7.3.38b)) and $\mathcal{F}_k^{\Delta}(t_M)$ (equation (7.3.38c)) are not known analytically and the corresponding functions are not smooth we

take $\theta = 1$ in the first iteration with time step $(\Delta t)^2$. With this choice we do not need to compute these terminal coefficients but keep second-order convergence in the first iteration. We use a very large number of timesteps $M = 10^4$ to get reference values Z_0. The results for the Euler, Milstein, and 2.0-weak-Taylor schemes are shown in Figure 7.5.2. Again we achieve second-order convergence in M with the 2.0-weak-Taylor scheme and $\theta = 1/2$.

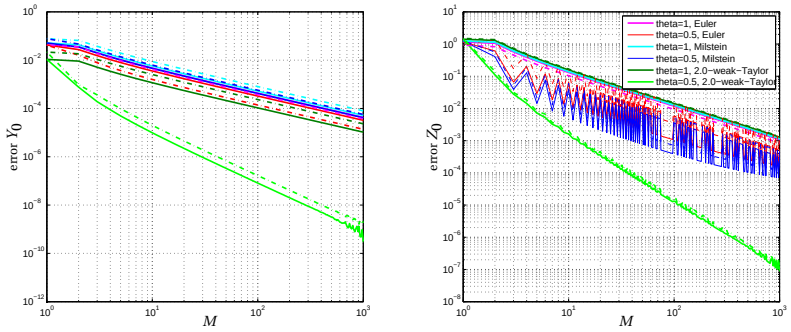


Figure 7.5.2: Results example 2, left: error in $\hat{y}_0^\Delta(x_0)$, right: error in $\hat{z}_0^\Delta(x_0)$, $\gamma = 0.2$ (lines) and $\gamma = 0.8$ (dashes).

7.5.3. EXAMPLE 3: BOND PRICE - CIR

In this section we consider the CIR interest rate process

$$dX_s = \kappa(\bar{x} - X_s)ds + \eta\sqrt{X_s}d\omega_s. \tag{7.5.8}$$

The PDE for the zero-coupon bond price is given by [Shr08, p.275]

$$\frac{\partial v}{\partial t}(t,x) + \kappa(\bar{x}-x)D_x v(t,x) + \tfrac{1}{2}\eta^2 x D_x^2 v(t,x) - xv(t,x) = 0, \quad (t,x) \in [0,T)\times\mathbb{R}_+, \tag{7.5.9a}$$

$$v(T,x) = 1, \quad x \in \mathbb{R}_+. \tag{7.5.9b}$$

This problem is related to the FBSDE

$$dY_s = -f(s,X_s,Y_s,Z_s)ds + Z_s d\omega_s, \quad Y_T = 1, \tag{7.5.10a}$$

$$f(t,x,y,z) = -xy. \tag{7.5.10b}$$

The exact solution is given by

$$h = \sqrt{\kappa^2 + 2\eta^2}, \tag{7.5.11a}$$

$$A(t,T) = \left(\frac{2he^{\frac{1}{2}(\kappa+h)(T-t)}}{2h + (\kappa+h)(e^{h(T-t)}-1)}\right)^{2\kappa\bar{x}/\eta^2}, \tag{7.5.11b}$$

$$B(t,T) = \frac{2(e^{(T-t)h}-1)}{2h + (\kappa+h)(e^{h(T-t)}-1)}, \tag{7.5.11c}$$

$$v(t,x) = A(t,T)\exp(-B(t,T)x). \tag{7.5.11d}$$

For the tests, we use the following parameter values

$$X_0 = 0.04,\ \bar{x} = 0.01,\ \kappa = 0.2,\ \eta = 0.1,\ T = 0.25. \tag{7.5.12}$$

The Feller condition is not satisfied and the process may reach zero. This is a nontrivial situation for Monte Carlo methods but for BCOS it is no problem, as we take the left-hand side of the computational domain equal to $a = 0$. For the exact simulation approach (Section 7.3.3.1) we use the characteristic function in equation (7.2.22). The results for the Euler, Milstein, 2.0-weak-Taylor and exact simulation schemes are shown in Figure 7.5.3. The 2.0-weak-Taylor and exact simulation scheme give almost the same results, with second-order convergence for $\theta = 1/2$.

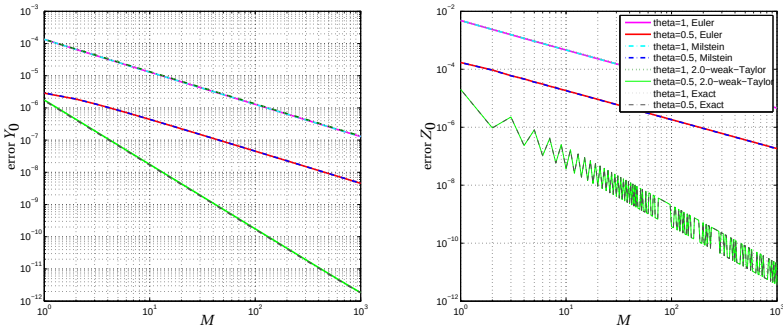


Figure 7.5.3: Results example 3, left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

7.5.4. EXAMPLE 4: TIME-DEPENDENT DRIFT AND DIFFUSION

The drift and diffusion terms of the FSDE, see equation (7.2.2), did not depend on time. In this last example we generalize the method by a time-dependent drift $\mu(t,x)$ and diffusion $\sigma(t,x)$. The numerical schemes in Sections 7.2 and 7.3 remain valid and the only change is due to the time-derivative in the diffusion operator $\mathscr{L}^0$, equation (7.2.7). The functions $m(x)$ and $s(x)$ in Section 7.3.1 for the 2.0-weak-Taylor scheme become time-dependent:

$$\begin{aligned} m(t,x) &= \mu(t,x) - \tfrac{1}{2}\sigma(t,x)\sigma_x(t,x) \\ &\quad + \tfrac{1}{2}\left(\mu_t(t,x) + \mu(t,x)\mu_x(t,x) + \tfrac{1}{2}\mu_{xx}(t,x)\sigma^2(t,x)\right)\Delta t, \qquad (7.5.13a) \\ s(t,x) &= \sigma(t,x) \qquad (7.5.13b) \\ &\quad + \tfrac{1}{2}\left(\mu_x(t,x)\sigma(t,x) + \sigma_t(t,x) + \mu(t,x)\sigma_x(t,x) + \tfrac{1}{2}\sigma_{xx}(t,x)\sigma^2(t,x)\right)\Delta t. \end{aligned}$$

We perform a test for the price $v(t,X_t)$ of a call option where the underlying asset follows a geometric Brownian motion with time-dependent drift and volatility. We choose the following periodic functionals:

$$\mu(t,x) := \bar{\mu}(t)x, \quad \bar{\mu}(t) := \bar{\mu}_0 + \bar{\mu}_1 \sin\left(\frac{2\pi t}{\zeta_1}\right) + \bar{\mu}_2 \sin\left(\frac{2\pi t}{\zeta_2}\right), \qquad (7.5.14a)$$

$$\sigma(t,x) := \bar{\sigma}(t)x, \quad \bar{\sigma}(t) := \bar{\sigma}_0 + \bar{\sigma}_1 \sin\left(\frac{2\pi t}{\zeta_1}\right) + \bar{\sigma}_2 \sin\left(\frac{2\pi t}{\zeta_2}\right). \qquad (7.5.14b)$$

The hedge portfolio, see Section 7.5.2, evolves according to the SDE

$$dY_s = \left(rY_s + \tfrac{\bar{\mu}(s)-r}{\bar{\sigma}(s)}\bar{\sigma}(s)a_s X_s\right)ds + \sigma(s,X_s)a_s d\omega_s. \qquad (7.5.15)$$

If we set $Z_s = \sigma(s, X_s)a_s$, then (Y, Z) solve the FBSDE

$$dY_s = -f(s, X_s, Y_s, Z_s)ds + Z_s d\omega_s, \quad Y_T = \max(X_T - K, 0), \tag{7.5.16a}$$

$$f(t, x, y, z) = -ry - \frac{\bar{\mu}(t) - r}{\bar{\sigma}(t)}z. \tag{7.5.16b}$$

Again the option value is given by $v(t, X_t) = Y_t$ and $\sigma(t, X_t)D_x v(t, X_t) = Z_t$. The exact solution of this local volatility model is given by the Black-Scholes price and Delta with volatility parameter $\sqrt{\frac{1}{T-t}\int_t^T \bar{\sigma}^2(s)ds}$ [Pas11]. For the tests, we use the following parameter values

$$\begin{aligned} &X_0 = 100,\ K = 100,\ r = 0.1,\ \bar{\mu}_0 = 0.2,\ \bar{\sigma}_0 = 0.25,\ T = 0.25, \\ &\bar{\mu}_1 = 0.1,\ \bar{\mu}_2 = 0.02,\ \bar{\sigma}_1 = 0.125,\ \bar{\sigma}_2 = 0.025,\ \zeta_1 = 1,\ \zeta_2 = \tfrac{1}{4}, \end{aligned} \tag{7.5.17}$$

with the exact solutions $Y_0 = v(t_0, X_0) = 7.8159$ and $Z_0 = \sigma X_0 D_x v(t_0, X_0) = 14.8115$. The results are shown in Figure 7.5.4. The computation time has increased because the characteristic function (7.3.5) changes with each timestep. The computation time for $M = 1000$ is about eight times longer.

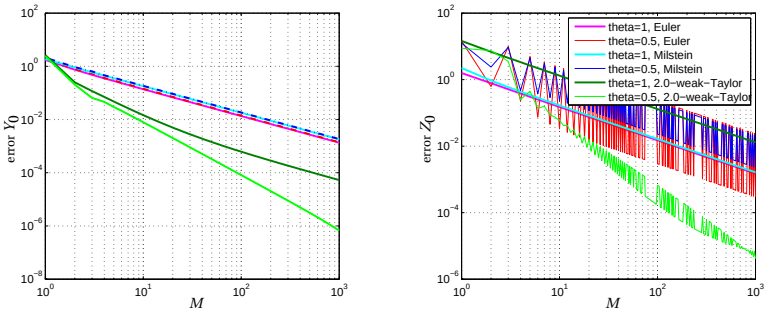


Figure 7.5.4: Results example 4, left: error in $\hat{y}_0^{\Delta}(x_0)$, right: error in $\hat{z}_0^{\Delta}(x_0)$.

7.6. CONCLUSION

In this chapter we extended the probabilistic numerical BCOS method from Chapter 6, for solving decoupled forward-backward stochastic differential equations. The underlying forward stochastic differential equation is now approximated by different Taylor discretization schemes, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes, or by exact simulation. The discretization of the FBSDE with the θ-method results in a backward induction scheme with conditional expectations. The expected values are approximated by a Fourier cosine method and relies on the availability of the characteristic function for these discrete Taylor schemes. In this way we generalize the applicability of the BCOS method to FSDEs for which the 'continuous' characteristic function is not available. The Fourier cosine coefficients are recovered recursively in an efficient way by using discrete Fourier cosine transforms and an FFT algorithm. Numerical tests demonstrate the applicability of the BCOS method for BSDEs in financial problems. In the tests we observed different convergence results for Z_0 and Y_0. The 2.0-weak-Taylor and exact simulation scheme with $\theta = 1/2$ result

in second-order convergence in the number of timesteps of the BCOS method, as expected from the error analysis.

Appendix

7.A. PROOF LEMMA 7.4.2

The Itô-Taylor expansion of FSDE X_t reads (with $c(x) = x$)

$$X_{m+1}^{m,x} = \sum_{\alpha\in\mathscr{A}} c_\alpha(x) I_{\alpha,m+1} + \sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[c_\alpha(X_\cdot^{m,x})]_{m+1}, \tag{7.A.1a}$$

$$X_{m+2}^{m,x} = \sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2}. \tag{7.A.1b}$$

For the Order 2.0 weak Taylor scheme we have

$$\begin{aligned}\mathscr{A} &= \{v,(0),(1),(1,1),(1,0),(0,1),(0,0)\} \quad \text{and} \\ \mathscr{B}(\mathscr{A}) &= \{(1,1,1),(0,0,1),(0,1,0),(1,0,0),(0,1,1),(1,0,1),(1,1,0),(0,0,0)\}.\end{aligned} \tag{7.A.2}$$

The discrete approximations read

$$X_{m+1}^{\Delta,m,x} = \sum_{\alpha\in\mathscr{A}} c_\alpha(x) I_{\alpha,m+1}, \qquad X_{m+2}^{\Delta,m,x} = \sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{\Delta,m,x}) I_{\beta,m+2}. \tag{7.A.3}$$

With the theory in Chapter 5.7 of [KP92] we find

$$\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[h(.,X_\cdot)]_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.4a}$$

$$\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[h(.,X_\cdot)]_{m+1}\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.4b}$$

$$\tfrac{2}{\Delta t}\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_{(0,\alpha)}[h_{(0,\alpha)}(.,X_\cdot)]_{m+1}\Delta\omega_{m+1}\Big] = 0, \tag{7.A.4c}$$

$$\tfrac{2}{\Delta t}\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_{(1,\alpha)}[h_{(1,\alpha)}(.,X_\cdot)]_{m+1}\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.4d}$$

$$\tfrac{2}{\Delta t}\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_0[h_{(0,\alpha)}(.,X_\cdot)]_{m+1}\Big] = \mathscr{O}((\Delta t)^{0}), \tag{7.A.4e}$$

$$\tfrac{2}{\Delta t}\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_1[h_{(1,\alpha)}(.,X_\cdot)]_{m+1}\Big] = 0, \tag{7.A.4f}$$

$$\tfrac{2}{\Delta t}\mathbb{E}_m\Big[\sum_{\alpha\in\mathscr{A}\setminus\alpha_\varnothing} I_{\alpha,m+1}\Big] + \tfrac{2}{\Delta t}\mathbb{E}_{m+1}\Big[\sum_{\alpha\in\mathscr{A}\setminus\alpha_\varnothing} I_{\alpha,m+1}\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{0}). \tag{7.A.4g}$$

With the above equations we find

$$\mathbb{E}\Big[X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x}\Big] = \mathbb{E}\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[c_\alpha(X_\cdot^{m,x})]_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.5a}$$

$$\mathbb{E}\Big[\Big(X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x}\Big)\Delta\omega_{m+1}\Big] = \mathbb{E}\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[c_\alpha(X_\cdot^{m,x})]_{m+1}\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{7.A.5b}$$

The Taylor series of function $h(t_{m+1}, x)$ in x gives

$$h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) = (-1)^{\ell+1} \sum_{\ell=1}^{\infty} \tfrac{1}{\ell!} h^{(\ell)}(t_{m+1}, X_{m+1}^{m,x}) \left(X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right)^{\ell}. \quad (7.A.6)$$

By using the strong error result ([KP92], Chapter 5.9)

$$\mathbb{E}\left[\left| X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right|^{\ell} \right] = \mathscr{O}((\Delta t)^{1.5\ell}) \quad (7.A.7)$$

for $\ell \geq 2$ and formula (7.4.1) for $\ell = 1$, we find

$$\mathbb{E}\left[h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (7.A.8a)$$

$$\mathbb{E}\left[\left(h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) \right) \Delta\omega_{m+1} \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (7.A.8b)$$

We can rewrite the difference between $X_{m+2}^{m,x}$ and $X_{m+2}^{\Delta,m,x}$ as

$$\begin{aligned} X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} &= \left(\sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} \right) \\ &\quad - \sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{m,x}) I_{\beta,m+2} - \sum_{\beta\in\mathscr{A}} c_\beta(X_{m+1}^{\Delta,m,x}) I_{\beta,m+2} \\ &= \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} + \sum_{\beta\in\mathscr{A}} \left(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \\ &= \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} \\ &\quad + \sum_{\beta\in\mathscr{A}\setminus\alpha_\emptyset} \left(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} + \sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[c_\alpha(X_\cdot^{m,x})]_{m+1}. \end{aligned} \quad (7.A.9)$$

Then we find

$$\begin{aligned} \mathbb{E}\left[X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right] &= \mathbb{E}\left[\sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} \right] \\ &\quad + \mathbb{E}\left[\sum_{\beta\in\mathscr{A}} \left(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \end{aligned} \quad (7.A.10a)$$

$$\begin{aligned} \mathbb{E}\left[\left(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right) \Delta\omega_{m+1} \right] &= \mathbb{E}\left[\sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} \Delta\omega_{m+1} \right] \\ &\quad + \mathbb{E}\left[\sum_{\beta\in\mathscr{A}} \left(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \Delta\omega_{m+1} \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \end{aligned} \quad (7.A.10b)$$

$$\begin{aligned} \mathbb{E}\left[\left(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right) \Delta\omega_{m+2} \right] &= \mathbb{E}\left[\sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2} \Delta\omega_{m+2} \right] \\ &\quad + \mathbb{E}\left[\sum_{\beta\in\mathscr{A}} \left(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \Delta\omega_{m+2} \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \end{aligned} \quad (7.A.10c)$$

Furthermore,

$$\tfrac{2}{\Delta t}\mathbb{E}\left[(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x}) (\Delta\omega_{m+1} - \Delta\omega_{m+2}) \right] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (7.A.11)$$

For the above result we divide the expected value into three parts with equation (7.A.9). Firstly

$$\begin{aligned}
&\tfrac{2}{\Delta t}\mathbb{E}\Big[\sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2}\Delta\omega_{m+1}\Big] \\
&\quad= \tfrac{2}{\Delta t}\mathbb{E}\Big[\Big(\sum_{\beta\in\mathscr{B}(\mathscr{A})} c_\beta(X_{m+1}^{m,x})I_{\beta,m+2} + \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_{(0,\beta)}[c_{(0,\beta)}(X_\cdot^{m,x})]_{m+2} \\
&\quad+ \sum_{\beta\in\mathscr{B}(\mathscr{A})} I_{(1,\beta)}[c_{(1,\beta)}(X_\cdot^{m,x})]_{m+2}\Big)\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}),
\end{aligned} \tag{7.A.12}$$

secondly

$$\tfrac{2}{\Delta t}\mathbb{E}\Big[\sum_{\beta\in\mathscr{A}\setminus\alpha_\emptyset}\Big(c_\beta(X_{m+1}^{m,x}) - c_\beta(X_{m+1}^{\Delta,m,x})\Big)I_{\beta,m+2}\,(\Delta\omega_{m+1}-\Delta\omega_{m+2})\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.13}$$

and thirdly

$$\begin{aligned}
&\tfrac{2}{\Delta t}\mathbb{E}\Big[\sum_{\alpha\in\mathscr{B}(\mathscr{A})} I_\alpha[c_\alpha(X_\cdot^{m,x})]_{m+1}\Delta\omega_{m+1}\Big] - \tfrac{2}{\Delta t}\mathbb{E}\Big[\sum_{\beta\in\mathscr{B}(\mathscr{A})} I_\beta[c_\beta(X_\cdot^{m,x})]_{m+2}\Delta\omega_{m+2}\Big] \\
&\quad= \tfrac{2}{\Delta t}\mathbb{E}\Big[\Big(\sum_{\alpha\in\mathscr{B}(\mathscr{A})} c_\alpha(x)I_{\alpha,m+1} + I_{(0,\alpha)}[c_{(0,\alpha)}(X_\cdot^{m,x})]_{m+1} + I_{(1,\alpha)}[c_{(1,\alpha)}(X_\cdot^{m,x})]_{m+1}\Big)\Delta\omega_{m+1}\Big] \\
&\quad- \tfrac{2}{\Delta t}\mathbb{E}\Big[\Big(\sum_{\beta\in\mathscr{B}(\mathscr{A})} c_\beta(x)I_{\beta,m+2} + I_0[c_{(0,\beta)}(X_\cdot^{m,x})]_{m+1}I_{\beta,m+2} + I_1[c_{(1,\beta)}(X_\cdot^{m,x})]_{m+1}I_{\beta,m+2} \\
&\qquad\qquad + I_{(0,\beta)}[c_{(0,\beta)}(X_\cdot^{m,x})]_{m+2} + I_{(1,\beta)}[c_{(1,\beta)}(X_\cdot^{m,x})]_{m+2}\Big)\Delta\omega_{m+2}\Big] \\
&\quad= \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}).
\end{aligned} \tag{7.A.14}$$

Note that in the above equation the terms $c_\alpha(x)I_{\alpha,m+1}$ and $c_\beta(x)I_{\beta,m+2}$ cancel out.

The Taylor series of function $h(t_{m+2},x)$ in x gives

$$h(t_{m+2},X_{m+2}^{m,x}) - h(t_{m+2},X_{m+2}^{\Delta,m,x}) = (-1)^{\ell+1}\sum_{\ell=1}^{\infty}\tfrac{1}{\ell!}h^{(\ell)}(t_{m+2},X_{m+2}^{m,x})\Big(X_{m+2}^{m,x}-X_{m+2}^{\Delta,m,x}\Big)^\ell. \tag{7.A.15}$$

By using the strong error result

$$\mathbb{E}\Big[\Big|X_{m+2}^{m,x}-X_{m+2}^{\Delta,m,x}\Big|^\ell\Big] = \mathscr{O}((\Delta t)^{1.5\ell}) \tag{7.A.16}$$

for $\ell\geq 2$ and formula (7.4.2) for $\ell=1$, we find

$$\mathbb{E}\Big[h(t_{m+2},X_{m+2}^{m,x}) - h(t_{m+2},X_{m+2}^{\Delta,m,x})\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.17a}$$

$$\mathbb{E}\Big[\Big(h(t_{m+2},X_{m+2}^{m,x}) - h(t_{m+2},X_{m+2}^{\Delta,m,x})\Big)\Delta\omega_{m+1}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.17b}$$

$$\mathbb{E}\Big[\Big(h(t_{m+2},X_{m+2}^{m,x}) - h(t_{m+2},X_{m+2}^{\Delta,m,x})\Big)\Delta\omega_{m+2}\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{7.A.17c}$$

$$\tfrac{2}{\Delta t}\mathbb{E}\Big[(h(t_{m+2},X_{m+2}^{m,x}) - h(t_{m+2},X_{m+2}^{\Delta,m,x}))\,(\Delta\omega_{m+1}-\Delta\omega_{m+2})\Big] = \mathscr{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{7.A.17d}$$

Remark 7.A.1. *For the Euler scheme we have*

$$\mathscr{A} = \{v, (0), (1)\} \quad \textit{and} \quad \mathscr{B}(\mathscr{A}) = \{(1,1), (1,0), (0,1), (0,0)\} \tag{7.A.18}$$

and for the Milstein scheme

$$\mathscr{A} = \{v, (0), (1), (1,1)\} \quad \textit{and} \quad \mathscr{B}(\mathscr{A}) = \{(1,0), (0,1), (0,0), (1,1,1), (0,1,1)\}. \tag{7.A.19}$$

For the Euler and Milstein scheme the weak convergence rate $\gamma_w^{2T} = 2$ is replaced by their weak convergence rates $\gamma_w^E = 1$ and $\gamma_w^M = 1$, respectively.

CHAPTER 8

Conclusions and Outlook

8.1. CONCLUSIONS

In this thesis we have presented efficient numerical methods for classes of stochastic problems. Our methods are based on the COS method, which is developed in [FO08] and [FO09]. This method approximates expected values and is based on Fourier cosine series expansions and the characteristic function of the underlying stochastic process.

In Chapter 2, we have presented a general approach for solving stochastic control problems under one-dimensional Lévy processes. The method relies on the dynamic programming principle and the COS formula. A recursive algorithm has been defined, based on the recursive recovery of the series coefficients. With the use of a Fast Fourier Transform (FFT) algorithm for Lévy processes we achieve a computational complexity of order $\mathcal{O}(N \log_2 N)$ per time step, where N denotes the number of terms in the series expansions. With an extensive error analysis we have acquired knowledge about the origin and evolution of errors. This understanding enabled us to improve the method by introducing an extrapolation method in the vicinity of domain boundaries, in which the COS formula may give inaccurate continuation values. Extrapolation by Taylor expansion or by exponential extrapolation can easily be applied as the derivatives of approximated continuation values can be computed easily based on the COS formula.

With the COS method we achieve an exponential convergence in the number of cosine coefficients for smooth density functions. If the underlying density is not smooth, however, the method may suffer from the Gibbs phenomenon and the convergence is only of algebraic order. In finance, the Variance Gamma asset process and stepwise cumulative distribution functions from portfolio loss modeling are examples resulting in slower convergence. A filtering technique improves the convergence rate significantly in terms of the number of required Fourier coefficients as well as in CPU time, see Chapter 3. The Fourier coefficients are premultiplied by a decreasing function $\hat{s}(k/N)$ so that the convergence rate is greatly improved.

In Chapter 4 we have presented the generalization of the COS method to higher dimensions. The developed recursive 2D-COS algorithm can be applied to, for example, pricing two-color European and Bermudan options with various payoffs and performs highly satisfactorily. For multidimensional stochastic processes in the class of Lévy processes, we can apply efficient matrix-vector multiplication using an FFT algorithm. The 2D-COS method can also be used for pricing single-asset options under the Heston stochastic volatility model. The Heston model is not in the Lévy class and an FFT algorithm is then applicable in only one dimension.

The climate-economics model in Chapter 5 is solved by combining the COS method for stochastic control problems and the 2D-COS formula. Economic equilibrium conditions provide an equation for the social discount rate. The results show that the far-distant future may be crucially important for long-term discounting. The results suggest that the key issue in determining the level of the social discount rate is not how large damages are or might be, although they surely must be substantial to invoke a sizeable response in the social discount rate, but how large the probability of extreme climate change is.

In Chapter 6 we have proposed a probabilistic numerical method for solving backward stochastic differential equations (BSDEs), which is called the BCOS method. The first step consists of discretizing the BSDE by taking conditional expectations and applying a general θ-discretization for the time-integrals. Then, the BCOS method solves the problem backwards in time by approximating the conditional expectations with the help of COS formulas. The Fourier cosine coefficients are recovered recursively in an efficient way by using discrete Fourier cosine transforms and an FFT algorithm. Besides, we have extended our BCOS method to solve BSDEs with jumps under jump-diffusion with a finite number of jump sizes. Numerical experiments have demonstrated the applicability of the BCOS method for FBSDEs in economic and financial problems and have shown highly satisfactorily and efficient results.

In Chapter 7 we extended the probabilistic numerical BCOS method. Now the underlying forward stochastic differential equation is approximated by different Taylor discretization schemes, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes, or by exact simulation. The concerning expected values are approximated by the Fourier cosine method and relies on the availability of the characteristic function for these discrete Taylor schemes. In this way we have generalized the applicability of the BCOS method to FSDEs for which the 'continuous' characteristic function is not available. Numerical tests demonstrate the applicability of the BCOS method for BSDEs in financial problems, such as pricing under the CEV model. The 2.0-weak-Taylor scheme with $\theta = 1/2$ results in second-order convergence, as expected from the error analysis.

8.2. OUTLOOK

Completing the research in this thesis we came up with the following interesting ideas for future research.

The multidimensional-COS method can deal with multi-asset option problems of medium-sized dimensionality, but suffers from the curse of dimensionality for high-dimensional option contracts. Combining the COS method with, for example sparse grid methods, projection methods, or Monte Carlo simulations may result in promising hybrid numerical methods.

The theory of decoupled FBSDEs has been extended to coupled FBSDEs [PT99] and second-order-BSDEs [CSTV07], which are related to quasilinear and fully nonlinear parabolic PDEs, respectively. Extensions of the BCOS method to these classes provide a way to solve nonlinear PDEs. A first step in this direction is provided in [HRO14]. BSDEs driven by Lévy processes are discussed in [NS01]. They are also a challenging extension for the BCOS method.

In portfolio valuation, the risk of counterparty's default should be taken into account. This is known as credit valuation adjustment. Relations with BSDEs are discussed in, among others, [Cré12] and open a broad class of applications for numerical methods.

The convergence of the partial sum of Fourier series is not necessarily monotone. The same is true for the COS formula, where the error in the number of Fourier cosine coefficients generally shows oscillatory behavior. Finding a monotone COS scheme is an interesting thought. Possibly a methodology similar to the filter-COS formula can deal with this.

Monotonicity together with ℓ_∞ stability and local consistency are sufficient conditions for convergence to the viscosity solution of the HJB equation. It is worthwhile to work out these conditions for the COS method. Discontinuities like the Gibbs phenomenon may require challenging technicalities.

CHAPTER 9

Appendix

A. FUNCTIONS χ_k, ψ_k, ξ_k, AND ξ_k^2

The functions χ_k and ψ_k are given by:

$$\chi_k(z_1,z_2,a,b) = \int_{z_1}^{z_2} e^y \cos\left(k\pi\tfrac{y-a}{b-a}\right) dy$$
$$\text{and} \quad \psi_k(z_1,z_2,a,b) = \int_{z_1}^{z_2} \cos\left(k\pi\tfrac{y-a}{b-a}\right) dy \tag{A.1}$$

and admit the following analytic solutions (for example with Maple 14):

$$\chi_k(z_1,z_2,a,b) = \tfrac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\Big[\cos\left(k\pi\tfrac{z_2-a}{b-a}\right)e^{z_2} - \cos\left(k\pi\tfrac{z_1-a}{b-a}\right)e^{z_1} + \tfrac{k\pi}{b-a}\sin\left(k\pi\tfrac{z_2-a}{b-a}\right)e^{z_2} - \tfrac{k\pi}{b-a}\sin\left(k\pi\tfrac{z_1-a}{b-a}\right)e^{z_1}\Big], \tag{A.2}$$

$$\psi_k(z_1,z_2,a,b) = \begin{cases} \left[\sin\left(k\pi\tfrac{z_2-a}{b-a}\right) - \sin\left(k\pi\tfrac{z_1-a}{b-a}\right)\right]\tfrac{b-a}{k\pi}, & \text{for } k\neq 0, \\ z_2 - z_1, & \text{for } k = 0. \end{cases} \tag{A.3}$$

The functions ξ_k and ξ_k^2 are given by:

$$\begin{aligned}\xi_k(z_1,z_2,a,b) &= \int_{z_1}^{z_2} y\cos\left(k\pi\tfrac{y-a}{b-a}\right) dy \\ &= \tfrac{b-a}{(k\pi)^2}\Big[\cos\left(k\pi\tfrac{z_2-a}{b-a}\right)(b-a) - \cos\left(k\pi\tfrac{z_1-a}{b-a}\right)(b-a) \\ &\quad + k\pi\sin\left(k\pi\tfrac{z_2-a}{b-a}\right)z_2 - k\pi\sin\left(k\pi\tfrac{z_1-a}{b-a}\right)z_1\Big]\end{aligned} \tag{A.4}$$

and

$$\begin{aligned}\xi_k^2(z_1,z_2,a,b) &= \int_{z_1}^{z_2} y^2\cos\left(k\pi\tfrac{y-a}{b-a}\right) dy \\ &= 2\tfrac{b-a}{(k\pi)^3}\Big[-k\pi z_2(b-a)\cos\left(k\pi\tfrac{z_2-a}{b-a}\right) + k\pi z_1(b-a)\cos\left(k\pi\tfrac{z_1-a}{b-a}\right) \\ &\quad + \left((b-a)^2 - \tfrac{1}{2}(k\pi z_2)^2\right)\sin\left(k\pi\tfrac{z_2-a}{b-a}\right) - \left((b-a)^2 - \tfrac{1}{2}(k\pi z_1)^2\right)\sin\left(k\pi\tfrac{z_1-a}{b-a}\right)\Big].\end{aligned} \tag{A.5}$$

B. FAST FOURIER TRANSFORM (FFT) ALGORITHM

In this thesis, the notation of the matrices $\mathscr{M}$, $\mathscr{M}^+$, and $\mathscr{M}^-$ differs between the different chapters. In Chapter 4 we use the notation $\mathscr{M}^+(z_1,z_2,a,b)$ and $\mathscr{M}^-(z_1,z_2,a,b)$, with different values for a and b (depending on the dimension). In Chapter 2 only matrix $\mathscr{M}^+$ with the same a and b values appears. Therefore, the notation in that chapter is simplified to $\mathscr{M}(z_1,z_2)$. In Chapter 6 the arguments z_1 and z_2 equal a and b, respectively, and we shorten notation to $\mathscr{M}$.

Theorem 1. *(Efficient computation of $\hat{\mathscr{C}}$, $\mathscr{K}$, and $\hat{\mathscr{K}}$)*
The matrix-vector product $\mathscr{M}^+(z_1,z_2,a,b)\mathbf{w}$ and $\mathscr{M}^-(z_1,z_2,a,b)\mathbf{w}$, can be computed in $\mathscr{O}(N\log_2 N)$ operations with the help of the Fast Fourier Transform algorithm, with N the size of square-matrices $\mathscr{M}^+$ and $\mathscr{M}^-$.

The key insights of this efficient computation are the equalities

$$\mathscr{M}^+_{k,j}(z_1,z_2,a,b) = -\tfrac{i}{\pi}\left(\mathscr{M}^{+c}_{k,j}(z_1,z_2,a,b)+\mathscr{M}^{+s}_{k,j}(z_1,z_2,a,b)\right), \tag{B.1a}$$

$$\mathscr{M}^-_{k,j}(z_1,z_2,a,b) = -\tfrac{i}{\pi}\left(\mathscr{M}^{-c}_{k,j}(z_1,z_2,a,b)+\mathscr{M}^{-s}_{k,j}(z_1,z_2,a,b)\right), \tag{B.1b}$$

where

$$\mathscr{M}^{\pm c}_{k,j}(z_1,z_2,a,b) = \frac{\exp\left(i(\pm j+k)\frac{(z_2-a)\pi}{b-a}\right)-\exp\left(i(\pm j+k)\frac{(z_1-a)\pi}{b-a}\right)}{\pm j+k}, \tag{B.2a}$$

$$\mathscr{M}^{\pm s}_{k,j}(z_1,z_2,a,b) = \frac{\exp\left(i(\pm j-k)\frac{(z_2-a)\pi}{b-a}\right)-\exp\left(i(\pm j-k)\frac{(z_1-a)\pi}{b-a}\right)}{\pm j-k}, \tag{B.2b}$$

with special cases

$$\mathscr{M}^{+c}_{k,j}(z_1,z_2,a,b) = \tfrac{(z_2-z_1)\pi i}{b-a}, \qquad \text{for } k=j=0, \tag{B.3a}$$

$$\mathscr{M}^{-s}_{k,j}(z_1,z_2,a,b) = \tfrac{(z_2-z_1)\pi i}{b-a}, \qquad \text{for } k=j=0, \tag{B.3b}$$

$$\mathscr{M}^{-c}_{k,j}(z_1,z_2,a,b) = \tfrac{(z_2-z_1)\pi i}{b-a}, \qquad \text{for } k=j, \tag{B.3c}$$

$$\mathscr{M}^{+s}_{k,j}(z_1,z_2,a,b) = \tfrac{(z_2-z_1)\pi i}{b-a}, \qquad \text{for } k=j. \tag{B.3d}$$

The matrices $\mathscr{M}^{+c}$ and $\mathscr{M}^{-s}$ are Hankel matrices ($\mathscr{M}_{i,j}=\mathscr{M}_{i-1,j+1}$) and $\mathscr{M}^{+s}$ and $\mathscr{M}^{-c}$ are Toeplitz matrices ($\mathscr{M}_{i,j}=\mathscr{M}_{i+1,j+1}$). These special Hankel and Toeplitz structures can be embedded in a circulant matrix form and the matrix-vector products can be written as a circular convolution of two vectors. Therefore, an FFT algorithm can be applied to achieve $\mathscr{O}(N\log_2 N)$ complexity, as described in [FO09].

C. DISCRETE FOURIER COSINE TRANSFORM (DCT)

In this appendix, we explain the idea of using discrete Fourier cosine transforms to approximate the Fourier cosine coefficients $\mathscr{F}_k$ of function $f(x)$, i.e.,

$$\mathscr{F}_k = \tfrac{2}{b-a}\int_a^b f(x)\cos\left(k\pi\tfrac{x-a}{b-a}\right)dx. \tag{C.1}$$

For this, we take N grid-points and define an equidistant x-grid

$$x_n := a + \left(n + \tfrac{1}{2}\right) \tfrac{b-a}{N} \quad \text{and} \quad \Delta x := \tfrac{b-a}{N}. \tag{C.2}$$

We determine the value of function $f(x)$ on the N grid-points. The midpoint-rule integration gives us

$$\mathscr{F}_k \approx \sum_{n=0}^{N-1} \tfrac{2}{b-a} f(x_n) \cos\left(k\pi \tfrac{x_n - a}{b-a}\right) \Delta x = \sum_{n=0}^{N-1} f(x_n) \cos\left(k\pi \tfrac{2n+1}{2N}\right) \tfrac{2}{N}. \tag{C.3}$$

The appearing DCT (Type II) can be calculated efficiently by, for example, the function `dct` of MATLAB. The error of the numerical integration is second order in N.

References

[AA00] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(3):231–262, 2000.

[AB04] L. Andersen and M. Broadie. Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science*, 50(9):1222–1234, 2004.

[AMST07] H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert. The little Heston trap. *Wilmott Magazine*, pages 83–92, January 2007.

[ARO14] R. F. T. Aalbers, M. J. Ruijter, and C. W. Oosterlee. The social discount rate under a stochastic A2 scenario. CPB Discussion Paper 296, 2014.

[AT88] C. M. Ahn and H. E. Thompson. Jump-diffusion processes and the term structure of interest rates. *The Journal of Finance*, 43(1):155–174, 1988.

[Bar06] R. J. Barro. Rare disasters and asset markets in the twentieth century. *The Quarter Journal of Economics*, 121(3):823–866, 2006.

[Bar09] R. J. Barro. Rare disasters, asset prices, and welfare costs. *American Economic Review*, 99(1):243–264, 2009.

[Bar13] R. J. Barro. Environmental protection, rare disasters, and discount rates. NBER Working Paper 19258, 2013.

[BBP97] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 60(1-2):57–83, 1997.

[BD97] M. Broadie and J. Detemple. The valuation of American options on multiple assets. *Mathematical Finance*, 7(3):241–286, 1997.

[BD07] C. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Processes and their Applications*, 117(12):1793–1812, 2007.

[BDM01] P. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for BSDEs. *Electronic Communications in Probability*, 6:1–14, 2001.

[BE08] B. Bouchard and R. Elie. Discrete-time approximation of decoupled forward–backward SDE with jumps. *Stochastic Processes and their Applications*, 118(1):53–75, 2008.

[Bec06] D. Becherer. Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging. *The Annals of Applied Probability*, 16(4):2027–2054, 2006.

[BEG89] P. P. Boyle, J. Evnine, and S. Gibbs. Numerical evaluation of multivariate contingent claims. *Review of Financial Studies*, 2(2):241–250, 1989.

[Ber95] Y. Z. Bergman. Option pricing with differential interest rates. *Review of Financial Studies*, 8(2):475–500, 1995.

[Bis73] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[BK08] C. Bender and M. Kohlmann. Optimal superhedging under nonconvex constraints - a BSDE approach. *International Journal of Theoretical and Applied Finance*, 11(4):363–380, 2008.

[BL11] S. Boyarchenko and S. Levendorskii. New efficient versions of Fourier

transform method in applications to option pricing. Available at SSRN: http://ssrn.com/abstract=1846633, 2011.

[BL13] P. Briand and C. Labart. Simulation of BSDEs by Wiener chaos expansion. HAL : hal-00688523, version 2, 2013.

[Boy88] P. P. Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 23(1):1–12, 1988.

[Boy96] J. P. Boyd. The Erfc-Log filter and the asymptotics of the Euler and Vandeven sequence accelerations. In A. V. Ilin and L. R. Scott, editors, *Proceedings of the Third International Conference on Spectral and High Order Methods*, pages 267–276. Houston Journal of Mathematics, 1996.

[Boy01] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Courier Dover Publications, New York, 2001.

[Boy11] J. P. Boyd. A proof, based on the Euler sum acceleration, of the recovery of an exponential (geometric) rate of convergence for the Fourier series of a function with Gibbs phenomenon. In J. S. Hesthaven and E. M. Rønquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 131–139. Springer Berlin Heidelberg, 2011.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[BS07] S. J. Berridge and J. M. Schumacher. Pricing high-dimensional American options using local consistency conditions. In J. Miller, D Edelman, and J. Appleby, editors, *Numerical Methods for Finance*, pages 13–52. Chapman & Hall/CRC, London, 2007.

[BS12] C. Bender and J. Steiner. Least-squares Monte Carlo for backward SDEs. *Numerical Methods in Finance*, 12:257–289, 2012.

[BT04] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[CF08] S. S. Clift and P. A. Forsyth. Numerical solution of two asset jump diffusion models for option valuation. *Applied Numerical Mathematics*, 58(6):743–782, 2008.

[CIL92] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

[CIR81] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A re-examination of traditional hypotheses about the term structure of interest rates. *The Journal of Finance*, 36(4):769–799, 1981.

[CIR85a] J. C. Cox, J. E. Ingersoll, and S. A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384, 1985.

[CIR85b] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[CJL13] Y. Cai, K. L. Judd, and T. S. Lontzek. The social cost of stochastic and irrreversible climate change. NBER Working Paper 18704, 2013.

[CM99] P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4), 1999.

[CM12] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method: application to nonlinear pricing. *SIAM Journal on Financial Mathematics*, 3(1):534–571, 2012.

[Coc05] J. H. Cochrane. *Asset pricing*. Princeton University Press, Princeton, 2005.

[COS01] T. F. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denoising. *IEEE Transactions on Image Processing*, 10(2):231–241, 2001.

[Cré12] S. Crépey. Bilateral counterparty risk under funding constraints part I: Pricing. *Mathematical Finance*, 2012.

[CSTV07] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic

differential equations and fully nonlinear parabolic pdes. *Communications on Pure and Applied Mathematics*, 60(7):1081–1110, 2007.

[CT03] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series, 2003.

[Das08] P. Dasgupta. Discounting climate change. *Journal of Risk and Uncertainty*, 37:141–169, 2008.

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

[dFV05] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25(1):87–112, 2005.

[DI06] P. Den Iseger. Numerical transform inversion using Gaussian quadrature. *Probability in the Engineering and Informational Sciences*, 20(1):1–44, 2006.

[DP94] A. K. Dixit and R. S. Pindyck. *Investment under uncertainty*. Princeton University Press, Princeton, 1994.

[DPS00] D. Duffie, J. Pan, and K. J. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[Duf05] D. Dufresne. Bessel processes and Asian options. In H. Ben-Ameur and M. Breton, editors, *Numerical Methods in Finance*, pages 35–57. Springer-Verlag, New-York, 2005.

[DW05] P. Dupuis and H. Wang. On the convergence from discrete to continuous time in an optimal stopping problem. *The Annals of Applied Probability*, 15(2):1339–1366, 2005.

[EKPQ97] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

[EKQ95] N. El Karoui and M. C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66, 1995.

[Fan12] F. Fang. Use COS method to calculate portfolio credit loss. Internal report and private communication, 2012.

[Fej00] L. Fejér. Sur les fonctions bornées et intégrables. *Comptes Rendus de l'Académie des Sciences*, 131:984–987, 1900.

[FL07] P. A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *Journal of Computational Finance*, 11(2):1–43, 2007.

[FMM11] G. Fusai, D. Marazzina, and M. Marena. Pricing discretely monitored Asian options by maturity randomization. *SIAM Journal on Financial Mathematics*, 2(1):383–403, 2011.

[FO08] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.

[FO09] F. Fang and C. W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik*, 114(1):27–62, 2009.

[FO11] F. Fang and C. W. Oosterlee. A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, 2(1):439–463, 2011.

[FS86] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 205–223. Elsevier Science, North-Holland, Amsterdam, 1986.

[FTW11] A. Fahim, N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic pdes. *The Annals of Applied Probability*, 21(4):1322–1364, 2011.

[Gel00] A. Gelb. A hybrid approach to spectral reconstruction of piecewise smooth functions. *Journal of Scientific Computing*, 15(3):293–322, 2000.

[Gen72] W. M. Gentleman. Implementing Clenshaw-Curtis quadrature, I methodology and experience. *Communications of the ACM*, 15(5):337–342, 1972.

[GKPP07] B. Groom, P. Koundori, K. Panipoulou, and T. Pantelides. Discounting the distant future:

How much does model selection affect the certainty equivalent discount rate? *Journal of Applied Econometrics*, 22(3):641–656, 2007.

[GL10] E. Gobet and C. Labart. Solving BSDE with adaptive control variate. *SIAM Journal on Numerical Analysis*, 48(1):257–277, 2010.

[GLW05] E. Gobet, J. P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005.

[Gob10] E. Gobet. Numerics of Backward SDEs. Presentation slides, 2010. Summer School in Probability Theory - Disentis - 26-30 July 2010.

[Gol14] C. Gollier. Gamma discounters are short-termist. TSE Working Paper Series 14-499, 2014.

[GRM06] P. Glasserman and J. Ruiz-Mata. Computing the credit loss distribution in the Gaussian copula model: a comparison of methods. *Journal of Credit Risk*, 2(4):33–66, 2006.

[GS97] D. Gottlieb and C.-W. Shu. On the Gibbs phenomenon and its resolution. *SIAM Review*, 39(4):644–668, 1997.

[GT13] E. Gobet and P. Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. HAL : hal-00642685, version 3, 2013.

[Gut05] A. Gut. *Probability: A Graduate Course*. Springer Texts in Statistics. Springer Science+Business Media, Inc., 2005.

[Han05] S. L. Hansen. A Malliavin-based Monte-Carlo approach for numerical solution of stochastic control problems: Experiences from Merton's problem. Working paper, 2005.

[Hes93] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[HG08] S.-F. Huang and M. Guo. Valuation of multidimensional Bermudan options. In H. K. Wolfgang, N. Hautsch, and L. Overbeck, editors, *Applied Quantitative Finance*, pages 295–309. Springer, Berlin Heidelberg, 2008.

[HGG07] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral Methods for Time-Dependent Problems*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2007.

[HH09] V. Henderson and D. Hobson. Utility indifference pricing - an overview. In R. Carmona, editor, *Indifference Pricing: Theory and Applications*, pages 44–74. Princeton University Press, 2009.

[HIM05] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15(3):1691–1712, 2005.

[HitH13] T. Haentjens and K. J. in 't Hout. ADI schemes for pricing American options under the Heston model. arXiv:1309.0110, 2013.

[HN89] S. Hodges and A. Neuberger. Optimal replication of contingent claims under transactions costs. *Review of Futures Markets*, 8(2):222–239, 1989.

[HO14] C. B. Hyndman and P. Oyono Ngou. A convolution method for numerical solution of backward stochastic differential equations. arXiv:1304.1783, 2014.

[HRO14] T. P. Huijskens, M. J. Ruijter, and C. W. Oosterlee. Numerical methods for coupled FBSDEs. Working paper, submitted for publication, 2014.

[Hua09] X. Huang. *Credit Portfolio Losses*. PhD thesis, Delft University of Technology, the Netherlands, 2009.

[Hul09] J. C. Hull. *Options, Futures and Other Derivatives*. Pearson/Prentice-Hall, New Jersey, 2009.

[Ise09] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2009.

[JKWW11] A. Janek, T. Kluge, R. Weron, and U. Wystup. FX smile in the Heston model. In P. Čížek, W. K. Härdle, and R. Weron, editors, *Statistical Tools for Finance and Insurance*, pages

133–162. Springer Verlag, 2011.

[JO12] S. Jain and C. W. Oosterlee. Pricing high-dimensional Bermudan options using the stochastic grid method. *International Journal of Computer Mathematics*, 89(9):1186–1211, 2012.

[JYC09] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Finance, 2009.

[Kah08] C. Kahl. *Modelling and Simulation of Stochastic Volatility in Finance*. Universal-Publishers, 2008.

[KD01] H. J. Kushner and P. G. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2001.

[Keh] S. Kehtari. private communication. ETH Zürich, Switzerland.

[KFV09] J. S. Kennedy, P. A. Forsyth, and K. R. Vetzal. Dynamic hedging under jump diffusion with transaction costs. *Operations Research*, 57(3):541–559, 2009.

[KJ05] C. Kahl and P. Jäckel. Not-so-complex logarithms in the Heston model. *Wilmott Magazine*, 19(9):94–103, 2005.

[KKK10] R. Korn, E. Korn, and G. Kroisandt. *Monte Carlo Methods and Models in Finance and Insurance*. Financial Mathematics. CRC Press/Taylor & Francis, 2010.

[Kob00] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28(2):558–602, 2000.

[KP92] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics Series. Springer, 1992.

[Kry99] N. V. Krylov. Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies. *Electronic Journal of Probability*, 4(2):1–19, 1999.

[Kry00] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellmans equations with variable coefficients. *Probability Theory and Related Fields*, 117(1):1–16, 2000.

[Lan56] C. Lanczos. *Applied analysis*. Prentice-Hall mathematics series. Prentice-Hall, 1956.

[Lel85] H. E. Leland. Option pricing and replication with transactions costs. *The Journal of Finance*, 40(5):1283–1301, 1985.

[Lew00] A. L. Lewis. *Option valuation under stochastic volatility: with Mathematica code*. Finance Press, 2000.

[LFBO08] R. Lord, F. Fang, F. Bervoets, and C. W. Oosterlee. A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes. *SIAM Journal on Scientific Computing*, 30(4):1678–1705, 2008.

[LGW06] J. P. Lemor, E. Gobet, and X. Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12(5):889–916, 2006.

[Lio83] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations part 2: viscosity solutions and uniqueness. *Communications in Partial Differential Equations*, 8(11):1229–1276, 1983.

[LK06] R. Lord and C. Kahl. Why the rotation count algorithm works. Tinbergen Institute Discussion Paper No. 2006-065/2, 2006.

[LK07] R. Lord and C. Kahl. Optimal Fourier inversion in semi-analytical option pricing. *Journal of Computational Finance*, 10(4):1–30, 2007.

[LK10] R. Lord and C. Kahl. Complex logarithms in Heston-like models. *Mathematical Finance*, 20(4):671–694, 2010.

[LO08a] C. C. W. Leentvaar and C. W. Oosterlee. Multi-asset option pricing using a parallel Fourier-based technique. *Journal of Computational Finance*, 12(1):1–26, 2008.

[LO08b] C. C. W. Leentvaar and C. W. Oosterlee. On coordinate transformation and grid stretching for sparse grid pricing of basket options. *Journal of Computational and Applied Mathe-*

matics, 222(1):193–209, 2008.

[LS14] R. J. A. Laeven and M. Stadje. Robust portfolio choice and indifference valuation. *Mathematics of Operations Research*, 39(4):1109–1141, 2014.

[LSM97] J. P. Lepeltier and J. San Martin. Backward stochastic differential equations with continuous coefficient. *Statistics & Probability Letters*, 32(4):425–430, 1997.

[LZ10] Y. Li and W. Zhao. Lp-error estimates for numerical schemes for solving certain kinds of backward stochastic differential equations. *Statistics & probability letters*, 80(21):1612–1617, 2010.

[MCC98] D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[Mer69] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics & Statistics*, 51(3):247–257, 1969.

[Mer90] R. C. Merton. *Continuous-time finance*. Blackwell, 1990.

[MMP13] J. Masoliver, M. Montero, and J. Perelló. Uncertain growth and the value of the future. Cowles Foundation Discussion Papers 1930, yale university, 2013.

[MOG11] J. J. Masdemont and L. Ortiz-Gracia. Haar wavelet-based approach for quantifying credit portfolio losses. *Quantitative Finance*, 2011.

[Mor10] M. A. Morlais. A new existence result for quadratic BSDEs with jumps with application to the utility maximization problem. *Stochastic Processes and their Applications*, 120(10):1966–1995, 2010.

[MPSMT02] J. Ma, P. Protter, J. San Martín, and S. Torres. Numerical method for backward stochastic differential equations. *The Annals of Applied Probability*, 12(1):302–316, 2002.

[MS05] M. Mania and M. Schweizer. Dynamic exponential utility indifference valuation. *The Annals of Applied Probability*, 15(3):2113–2143, 2005.

[MSZ08] J. Ma, J. Shen, and Y. Zhao. On numerical approximations of forward-backward stochastic differential equations. *SIAM Journal on Numerical Analysis*, 46(5):2636–2661, 2008.

[MT06] G. N. Milstein and M. V. Tretyakov. Numerical algorithms for forward-backward stochastic differential equations. *SIAM Journal on Scientific Computing*, 28(2):561–582, 2006.

[NAD+00] N. Nakićenović, J. Alcamo, G. Davis, B. De Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grübler, T. Jung, T. Kram, et al. IPCC special report on emissions scenarios (SRES), 2000.

[NOA] NOAA Earth System Research Laboratory, Global Monitoring division. Trends in atmospheric carbon dioxide.

[Nor08] W. D. Nordhaus. *A Question of Balance: Weighing the Options on Global Warming Policies*. Yale University Press, New Haven & London, 2008.

[NP03] R. Newell and W. Pizer. Discounting the benefits of climate change mitigation: How much do uncertain rates increase valuations? *Journal of Environmental Economics and Management*, 46(1):52–71, 2003.

[NS01] D. Nualart and W. Schoutens. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. *Bernoulli*, 7(5):761–776, 2001.

[OGO13] L. Ortiz-Gracia and C. W. Oosterlee. Robust pricing of European options with wavelets and the characteristic function. *SIAM Journal on Scientific Computing*, 35(5):B1055–B1084, 2013.

[O'S05] C. O'Sullivan. Path dependent option pricing under Lévy processes. In *European Finance Association 2005 Moscow Meetings Paper*, 2005.

[ØS07] B. K. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Universitext. Springer, 2007.

[Pas11] A. Pascucci. *PDE and Martingale Methods in Option Pricing*. Bocconi & Springer Series. Springer, 2011.

[PBL10] E. Platen and N. Bruti-Liberati. *Numerical Solution of Stochastic Differential Equations*

with Jumps In Finance. Stochastic Modelling and Applied Probability. Springer-Verlag, 2010.

[Pey02] R. Peyret. *Spectral Methods for Incompressible Viscous Flow*. Volume 148 in Applied Mathematical Sciences. Springer, 2002.

[PFV03] D. M. Pooley, P. A Forsyth, and K. R Vetzal. Numerical convergence properties of option pricing PDEs with uncertain volatility. *IMA Journal of Numerical Analysis*, 23(2):241–267, 2003.

[Pha09] H. Pham. *Continuous-time stochastic control and optimization with financial applications*. Springer-Verlag, Berlin, 2009.

[PP90] E. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, volume 176 of *Lecture Notes in Control and Information Sciences.*, pages 200–217. Springer, Berlin, 1992.

[PPR11] S. Pagliarani, A. Pascucci, and C. Riga. Adjoint expansions in local Lévy models. Available at SSRN: `http://ssrn.com/abstract=1937149`, 2011.

[PT99] E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields*, 114(2):123–150, 1999.

[RB07] G. H. Roe and M. B. Baker. Why is climate sensitivity so predictable? *Science*, 318:629–632, 2007.

[RB13] G. H. Roe and Y. Bauman. Climate sensitivity: Should the climate tail wag the policy dog? *Climatic Change*, 117(4):647–662, 2013.

[REK00] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.

[Rie88] T. A. Rietz. The equity risk premium: a solution. *Journal of Monetary Economics*, 22(1):117–131, 1988.

[RO12a] M. J. Ruijter and C. W. Oosterlee. The COS method for pricing options under uncertain volatility. In M. Cummins, F. Murphy, and J. J. H. Miller, editors, *Topics in Numerical Methods for Finance*, volume 19 of *Springer Proceedings in Mathematics & Statistics*, pages 95–113. Springer US, 2012.

[RO12b] M. J. Ruijter and C. W. Oosterlee. Two-dimensional Fourier cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing*, 34(5):B642–B671, 2012.

[RO14] M. J. Ruijter and C. W. Oosterlee. Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance. Working paper available at SSRN: `http://ssrn.com/abstract=2501686`, submitted for publication, 2014.

[RO15] M. J. Ruijter and C. W. Oosterlee. A Fourier cosine method for an efficient computation of solutions to BSDEs. *To appear in SIAM Journal on Scientific Computing*, 2015.

[ROA13] M. J. Ruijter, C. W. Oosterlee, and R. F. T. Aalbers. On the Fourier cosine series expansion method for stochastic control problems. *Numerical Linear Algebra with Applications*, 20(4):598–625, 2013.

[Roe09] G. H. Roe. Feedbacks, timescales, and seeing red. *Annual Review of Earth and Planetary Sciences*, 37:93–115, 2009.

[Roy06] M. Royer. Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic processes and their applications*, 116(10):1358–1376, 2006.

[RS11] C. M. Reinhart and M. B. Sbrancia. The liquidation of government debt, 2011.

[Rui10] M. J. Ruijter. Numerical treatment of stochastic control problems by Fourier-cosine series expansions: The dike height problem. Master's thesis, Delft University of Technology, the Netherlands, 2010.

[RVO15] M. J. Ruijter, M. Versteegh, and C. W. Oosterlee. On the application of spectral filters in a

Fourier option pricing technique. *To appear in Journal of Computational Finance*, 2015.

[S+07] S. D. Solomon et al. IPCC climate change 2007: The physical science basis. intergovernmental panel on climate change. 2007.

[Sar06] S. A. Sarra. Digital total variation filtering as postprocessing for Chebyshev pseudospectral methods for conservation laws. *Numerical Algorithms*, 41(1):17–33, 2006.

[Sch03] W. Schoutens. *Lévy processes in finance: pricing financial derivatives*. Wiley Series in Probability and Statistics. Wiley, 2003.

[Sek06] J. Sekine. On exponential hedging and related quadratic backward stochastic differential equations. *Applied Mathematics & Optimization*, 54(2):131–158, 2006.

[Shr08] S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, Berlin, 2008.

[Soo73] T. T. Soong. *Random Differential Equations in Science and Engineering*. Academic Press, 1973.

[SS03] E. M. Stein and R. Shakarchi. *Fourier analysis: an introduction*. Princeton Lectures in Analysis. Princeton University Press, 2003.

[Ste07] N. Stern. *The Economics of Climate Change: The Stern Review*. Cambridge University Press, Cambridge, UK, 2007.

[Stu82] R. M. Stulz. Options on the minimum or the maximum of two risky assets: Analysis and applications. *Journal of Financial Economics*, 10(2):161–185, 1982.

[Sur09] V. Surkov. *Option Pricing using Fourier Space Time-stepping Framework*. PhD thesis, University of Toronto, 2009.

[Tad07] E. Tadmor. Filters, mollifiers and the computation of the Gibbs phenomenon. *Acta Numerica*, 16:305–378, 2007.

[Tan06] J. Tanner. Optimal filter and mollifier for piecewise smooth spectral data. *Mathematics of computation*, 75(254):767–790, 2006.

[TL94] S. Tang and X. Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM Journal on Control and Optimization*, 32(5):1447–1475, 1994.

[Tol09] R. S. J. Tol. The economic effects of climate change. *Journal of Economic Perspectives*, 23(2):29–51, 2009.

[TT05] E. Tadmor and J. Tanner. Adaptive filters for piecewise smooth spectral data. *IMA journal of numerical analysis*, 25(4):635–647, 2005.

[U.S14] U.S. Department of the Treasury. `http://www.treasury.gov/resource-center/data-chart-center/i` Jan. 2014.

[Van91] H. Vandeven. Family of spectral filters for discontinuous problems. *Journal of Scientific Computing*, 6(2):159–192, 1991.

[Vas02] O. Vasicek. The distribution of loan portfolio value. *Risk*, 15(12):160–162, 2002.

[vdPO12] S. P. van der Pijl and C. W. Oosterlee. An ENO-based method for second-order equations and application to the control of dike levels. *Journal of Scientific Computing*, 50(2):462–492, 2012.

[vSRV11] P. van Slingerland, J. K. Ryan, and C. Vuik. Position-dependent smoothness-increasing accuracy-conserving (SIAC) filtering for improving discontinuous Galerkin solutions. *SIAM Journal on Scientific Computing*, 33(2):802–825, 2011.

[Wei98] M. L. Weitzman. Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management*, 36(3):201–208, 1998.

[Wei09] M. L. Weitzman. Additive damages, fat-tailed climate dynamics, and uncertain discounting. *Economics: The Open-Access, Open-Assessment E-Journal*, 3(39):1–29, 2009.

[Wei10] M. L. Weitzman. Risk-adjusted gamma discounting. *Journal of Environmental Economics and Management*, 60(1):1–13, 2010.

[Wei12] M. L. Weitzman. GHG targets as insurance against catastrophic states climate damages. *Journal of Public Economic Theory*, 14(2):221–244, 2012.

[WLZ09] J. Wang, C. Luo, and W. Zhao. Crank-Nicolson scheme and its error estimates for back-

ward stochastic differential equations. *Acta Mathematicae Applicatae Sinica, English Series*, pages 1–10, 2009.

[ZCP06] W. Zhao, L. Chen, and S. Peng. A new kind of accurate numerical method for backward stochastic differential equations. *SIAM Journal on Scientific Computing*, 28(4):1563–1581, 2006.

[ZGO12] B. Zhang, L. A. Grzelak, and C. W. Oosterlee. Efficient pricing of commodity options with early-exercise under the Ornstein-Uhlenbeck process. *Applied Numerical Mathematics*, 62(2):91–111, 2012.

[Zha04] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14(1):459–488, 2004.

[ZLJ13] W. Zhao, Y. Li, and L. Ju. Error estimates of the Crank-Nicolson scheme for solving backward stochastic differential equations. *International Journal of Numerical Analysis and Modeling*, 10(4):876–898, 2013.

[ZLZ12] W. Zhao, Y. Li, and Zhang. A generalized θ-scheme for solving backward stochastic differential equations. *Discrete and Continuous Dynamical Systems - Series B (DCDS-B)*, 17(5):1585–1603, 2012.

[ZO13a] B. Zhang and C. W. Oosterlee. An efficient pricing algorithm for swing options based on Fourier cosine expansions. *Journal of Computational Finance*, 16(4):3–34, 2013.

[ZO13b] B. Zhang and C. W. Oosterlee. Efficient pricing of European-style Asian options under exponential Lévy processes based on Fourier cosine expansions. *SIAM Journal on Financial Mathematics*, 4(1):399–426, 2013.

[ZW09] K. Zhang and S. Wang. A computational scheme for uncertain volatility model in option pricing. *Applied Numerical Mathematics*, 59(8):1754–1767, 2009.

[ZWP09] W. Zhao, J. Wang, and S. Peng. Error estimates of the theta-scheme for backward stochastic differential equations. *Discrete and Continuous Dynamical Systems - Series B*, 12(4):905–924, 2009.

[ZZJ14] W. Zhao, W. Zhang, and L. Ju. A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations. *Communications in Computational Physics*, 15(3):618–646, 2014.

Curriculum Vitæ

Maria Johanna RUIJTER

03-05-1986	Born in Venhuizen, the Netherlands.

EDUCATION

1998–2004	Gymnasium Scholengemeenschap Tabor, locatie Werenfridus, Hoorn, the Netherlands
2004–2008	Bachelor Technische Wiskunde Delft University of Technology, Delft, the Netherlands
2006-2007	Board member of study association W.I.S.V. 'Christiaan Huygens' Delft, the Netherlands
2008–2010	Master Applied Mathematics Delft University of Technology, Delft, the Netherlands
2010-2011	Economics courses Tinbergen Institute, Amsterdam, the Netherlands
2010–2015	PhD Researcher Centrum Wiskunde & Informatica, Amsterdam, the Netherlands Centraal Planbureau, Den Haag, the Netherlands
2015	PhD Applied Mathematics Delft University of Technology, Delft, the Netherlands *Thesis:* Fourier Methods for Multidimensional Problems and Backward SDEs in Finance and Economics *Promotor:* Prof. dr. ir. C. W. Oosterlee

List of Publications

9. **M. J. Ruijter** and **C. W. Oosterlee**. Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance. Working paper, submitted for publication, 2014.

8. **M. J. Ruijter** and **C. W. Oosterlee**. A Fourier cosine method for an efficient computation of solutions to BSDEs. To appear in SIAM Journal on Scientific Computing, 2015.

7. **T. P. Huijskens**, **M. J. Ruijter** and **C. W. Oosterlee**. Numerical methods for coupled FBSDEs. Working paper, submitted for publication, 2014.

6. **M. Pou**, **M. J. Ruijter** and **C. W. Oosterlee**. Extension of a Fourier-cosine method to solve BSDEs with higher dimensions. Proceedings that will be published by Springer Heidelberg as part of the ECMI book subseries of Mathematics in Industry, 2015.

5. **R. F. T. Aalbers**, **M. J. Ruijter** and **C. W. Oosterlee**. The social discount rate under a stochastic A2 scenario. CPB Discussion Paper 296, 2014.

4. **M. J. Ruijter**, **M. Versteegh** and **C. W. Oosterlee**. On the application of spectral filters in a Fourier option pricing technique. To appear in Journal of Computational Finance, 2015.

3. **M. J. Ruijter**, **C. W. Oosterlee**, and **R. F. T. Aalbers**. On the Fourier cosine series expansion method for stochastic control problems. Numerical Linear Algebra with Applications, 20(4):598-625, 2013.

2. **M. J. Ruijter** and **C. W. Oosterlee**. Two-dimensional Fourier cosine series expansion method for pricing financial options. SIAM Journal on Scientific Computing, 34(5):B642-B671, 2012.

1. **M. J. Ruijter** and **C. W. Oosterlee**. The COS method for pricing options under uncertain volatility. In M. Cummins, F. Murphy, and J. J. H. Miller, editors, Topics in Numerical Methods for Finance, volume 19 of Springer Proceedings in Mathematics & Statistics, pages 95-113. Springer US, 2012.

List of Attended Conferences with Presentation

8. QMF Quantitative Methods in Finance, Sydney, Australia, December 2013.
7. EAERE 20th Annual Conference of the European Association of Environmental and Resource Economists, Toulouse, France, June 2013.
6. ICCS International Conference on Computational Science, Workshop on Computational and Algorithmic Finance, Barcelona, Spain, June 2013.
5. Winter School on Mathematical Finance, Lunteren, the Netherlands, January 2013.
4. FM12 SIAM Conference on Financial Mathematics & Engineering, Minneapolis, United States of America, July 2012.
3. Spring Meeting WSC, Antwerp, Belgium, May 2012.
2. Workshop CWI-EUR Backward Stochastic Differential Equations, Eindhoven, the Netherlands, January 2012.
1. 3rd International Conference on Numerical Methods for Finance, Limerick, Ireland, June 2011.

Acknowledgment

Foremost, my most sincere gratitude goes to my supervisor and promotor prof. Kees Oosterlee for his guidance during my research. Kees always makes time for all his PhD students, regardless his busy agenda, and it has been a great pleasure to work with him and learn from his experience.

I enjoyed working at Centrum Wiskunde & Informatica (CWI). There is a stimulating working atmosphere and I would like to thank all my colleagues for giving me a great time. From the Scientific Computing department: Joost Batenburg, Jeroen Bedorf, Debarati Bhaumik, Rob Bisseling, Folkert Bleichrodt, Bin Chen, Daan Crommelin, Jesse Dorrestijn, Svetlana Dubinkina, Bram van Es, Qian Feng, Wagner Fortes, Jason Frank, Lech Grzelak, Willem Haverkort, Shashi Jain, Barry Koren, Álvaro Leitao Rodríguez, Keith Myerscough, Zsolt Nika, Margreet Nool, Luis Ortiz Gracia, Willem Jan Palenstijn, Daan Pelt, Linda Plantagie, Marta Pou Bueno, Benjamin Sanderse, Anton van der Stoep, María Suárez Taboada, Nick Verheul, Wander Wadman, Huiqing Wang, Jeroen Witteveen, Paul de Zeeuw, Zhichao Zhong, Xiaodong Zhuge, Nada Mitrovic, and Duda Tepsic and the runners of the CWI running group.

I would also like to thank the former developers of the COS method, Fang Fang and Bowen Zhang. They shared their knowledge about the algorithm and provided me with Matlab codes. This helped me with continuing the work on this numerical Fourier method.

For the climate-economics project we worked together with CPB Netherlands Bureau for Economic Policy Analysis. I would like to thank my colleagues from Sector 4 and Sector 5, in particular my co-supervisor Rob Aalbers, for patiently explaining the economics, and dear colleague Victoria Shestalova.

Now and then I worked at the TU Delft, Numerical Analysis department. As I finished my Master's degree there it still feels like coming home for me. Special thanks to Hans van der Weide, Fei Cong, Jing Zhao, Thomas Huijskens, Mark Versteegh, Dennis den Ouden, Martijn de Jong, Thea Vuik, and Deborah Dongor.

Na alle bovengenoemde wiskunde- en economieonderzoekers wil ik graag familie en vrienden bedanken. In het bijzonder mijn ouders, voor het vertrouwen en alles wat me tot hier heeft gebracht, Marleen, Bo, Jan-Joris en mijn betrokken schoonfamilie.
Jeroen, voor alle liefde en plezier.
En kleine, lieve Tijmen, dat één plus één drie is geworden, daar kan geen wiskundeformule tegenop!

Marjon Ruijter
January 2015