COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS

UNIVERSITY OF WUPPERTAL
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
DEPARTMENT OF MATHEMATICS AND INFORMATICS

Analytical and Numerical Approximative Methods for solving Multifactor Models for pricing of Financial Derivatives

Dissertation Thesis
Binational doctoral study

Mgr. ZUZANA BUČKOVÁ
2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170412-150623-2
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20170412-150623-2]

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS

UNIVERSITY OF WUPPERTAL
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
DEPARTMENT OF MATHEMATICS AND INFORMATICS

Analytical and Numerical Approximative Methods for solving Multifactor Models for pricing of Financial Derivatives

Dissertation Thesis
Binational doctoral study

Study program:	Economic and Financial Mathematics
Field od study:	9.1.9 Applied mathematics
Department:	Department of Applied Mathematics and Statistics FMFI UK, Mlynská dolina, 842 48 Bratislava, Slovakia
Supervisor:	Prof. RNDr. Daniel ŠEVČOVIČ, CSc.
Consultant:	Doc. RNDr. Beáta STEHLÍKOVÁ, PhD.

Study program:	Applied Mathematics and Numerical Analysis
Address Department:	Bergische Universität Wuppertal, Gaussstrasse 20, 42119 Wuppertal, Germany
Supervisor:	Prof. Dr. Matthias EHRHARDT
Consultant:	Prof. Dr. Michael GÜNTHER

Author:	Mgr. Zuzana BUČKOVÁ
Date:	2016
Place:	Bratislava, Slovakia; Wuppertal, Germany

Analytické a numerické aproximačné metódy riešenia viacfaktorových modelov oceňovania finančných derivátov

Dizertačná práca

na dosiahnutie
akademického titulu
Philosophiae Doctor (PhD.)

na
Fakulte matematiky, fyziky a informatiky
Univerzity Komenského v Bratislave
v rámci binacionálneho doktorandského štúdia
(Cotutelle)
je predložená od

Mgr. ZUZANA BUČKOVÁ

rod. Zíková

narodená 23.09.1987 v Kežmarku

Analytische und numerische Approximationsmethoden zum Lösen von Mehrfaktormodellen zur Bewertung von Finanzderivaten

Dissertation

zur Erlangung
des akademischen Grades
eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

dem
Fakultät für Mathematik und Naturwissenschaften
der
Bergischen Universität Wuppertal (BUW)
im Rahmen eines gemeinsamen binationalen
Promotionsverfahrens (Cotutelle)
vorgelegt von

Mgr. ZUZANA BUČKOVÁ

geb. Zíková

geboren am 23.09.1987 in Kežmarok, Slowakei

20504548

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Mgr. Zuzana Bučková
Študijný program: aplikovaná matematika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: aplikovaná matematika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Analytical and Numerical Approximative Methods for solving Multifactor Models for pricing of Financial Derivatives
Analytické a numerické aproximačné metódy riešenia viacfaktorových modelov oceňovania finančných derivátov

Cieľ: Cieľom práce je matematická analýza viacfaktorových modelov pre oceňovanie finančných derivátov so zameraním na modelovanie úrokových mier.

Školiteľ: prof. RNDr. Daniel Ševčovič, CSc.
Školiteľ: prof. Dr. rer. nat. Matthias Ehrhardt
Konzultant: prof. Dr. rer. nat. Michael Guenther
Konzultant: doc. RNDr. Beáta Stehlíková, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Spôsob sprístupnenia elektronickej verzie práce:
bez obmedzenia

Dátum zadania: 01.09.2011

Dátum schválenia: 23.02.2011

prof. RNDr. Marek Fila, DrSc.
garant študijného programu

....
študent

....
školiteľ

Acknowledgement

I would like to thank to my supervisor at Comenius University in Bratislava, prof. RNDr. Daniel Ševčovič, CSc., for his guidance and support during my PhD. study. I am sincerely thankful to Doc. RNDr. Beáta Stehlíková, PhD. for her professional support, patience, and her enthusiasm in this long-lasting cooperation. At the University of Wuppertal, my thanks go to Prof. Dr. Michael Günther for his guidance and support. I would like to express many thanks for supervision from Prof. Dr. Matthias Ehrhardt and for involving me from the very beginning in the STRIKE project. The numerous discussions we had were always insightful and his detailed and deep corrections continuously improved the quality of my contributions. His sense for networking and cooperation with other groups opened new perspectives throughout my PhD. Thus, I had the opportunity to cooperate with Dr. Daniel Duffy about ADE methods and Prof. Igor Tsukermann about Trefftz methods. I would like to thank them for their time and enthusiasm.

Beáta, Matthias, Daniel, and Michael are experts with overview in different fields and it has been my pleasure to learn from them. I am indebted to them for guiding me within my PhD.study and for their valuable comments on this thesis.

This 5-year binational PhD. study brought me two different perspectives, two approaches from two groups of colleagues. In the last three years of my PhD. study I was involved in the STRIKE project, *Novel Methods in Computational Finance*, what is a European mathematical research training network. Experience as an early stage researcher in this project has given me an inspiration and enabled to join a strong community of people working on similar topics. I am very thankful to Matthias and Dr. Jan ter Maten for this life-lasting experience and to all my supervisors for the possibility to join this unique project.

I would like to thank my colleagues and friends in Wuppertal and Bratislava and all the people whose I met at the numerous STRIKE events. For the very nice scientific discussion we had, I am thankful to Pedro, Shih-Hau, Igor, José, Christian, Vera, Long, Radoslav. I would like to mention some others of my friends that supported the course of my research work: Janka, Majka, Magda, Darinka, Lara, Dmitry, Guillaume. For the many pieces of advice in connection with an English language that I received, I am thankful to Martinka and for the German language ones, I am indebted to Matthias, Daniel, Christoph, and Christian.

I would especially like to thank my amazing family for the love, support, and constant encouragement I have received over the years. My deepest thanks go to my kind mum and my father, who passed away, for raising me up, being such good example, and always showing me good directions. I warmly thank my supportive husband Andrej for his love and patience as well his parents and sisters Barborka and Danka. Last, but not least I thank God for health and bleesing me for all the time.

Abstract

The thesis covers different approaches used in current modern computational finance. Analytical and numerical approximative methods are studied and discussed. Effective algorithms for solving multi-factor models for pricing of financial derivatives have been developed.

The first part of the thesis is dealing with modeling of aspects and focuses on analytical approximations in short rate models for bond pricing. We deal with a two-factor convergence model with non-constant volatility which is given by two stochastic differential equations (SDEs). Convergence model describes the evolution of interest rate in connection with the adoption of the Euro currency. From the SDE it is possible to derive the PDE for bond price. The solution of the PDE for bond price is known in closed form only in special cases, e.g. Vasicek or CIR model with zero correlation. In other cases we derived the approximation of the solution based on the idea of substitution of constant volatilities, in solution of Vasicek, by non-constant volatilities. To improve the quality in fitting exact yield curves by their estimates, we proposed a few changes in models. The first one is based on estimating the short rate from the term structures in the Vasicek model. We consider the short rate in the European model for unobservable variable and we estimate it together with other model parameters. The second way to improve a model is to define the European short rate as a sum of two unobservable factors. In this way, we obtain a three-factor convergence model. We derived the accuracy for these approximations, proposed calibration algorithms and we tested them on simulated and real market data, as well.

The second part of the thesis focuses on the numerical methods. Firstly we study Fichera theory which describes proper treatment of defining the boundary condition. It is useful for partial differential equation which degenerates on the boundary. The derivation of the Fichera function for short rate models is presented. The core of this part is based on Alternating direction explicit methods (ADE) which belong to not well studied finite difference methods from 60s years of the 20th century. There is not a lot of literature regarding this topic. We provide numerical analysis, studying stability and consistency for convection-diffusion-reactions equations in the one-dimensional case. We implement ADE methods for two-dimensional call option and three-dimensional spread option model. Extensions for higher dimensional Black-Scholes models are suggested. We end up this part of the thesis with an alternative numerical approach called Trefftz methods which belong to Flexible Local Approximation MEthods (FLAME). We briefly outline the usage in computational finance.

Keywords: short rate models, convergence model of interest rate, bond pricing, Black-Scholes model, option pricing, approximate analytic solution, order of accuracy, numerical methods, calibration, simulated data, real market data, ADE, splitting schemes, Trefftz, Fichera theory

Slovak

Práca popisuje rôzne prístupy používané v súčasnom modernom oceňovaní finančných derivátov. Zaoberáme sa analytickými a numerickými aproximačnými metódami. Vyvinuli sme efektívne algoritmy riešenia viacfaktorových modelov oceňovania finančných derivátov.

Prvá časť práce sa zaoberá modelovaním rôznych aspektov a je zameraná na analytické aproximácie cien dlhopisov v modeloch krátkodobých úrokových mier. Zaoberáme sa dvojfaktorovým konvergenčným modelom s nekonštantnom volatilitou, ktorý je daný dvomi stochastickými diferenciálnymi rovnicami. Konvergenčný model opisuje vývoj úrokovej miery v súvislosti s prijatím eura. Zo stochastickej diferenciálnej rovnice je možné odvodiť parciálnu diferenciálnu rovnicu pre cenu dlhopisu. Riešenie parciálnej diferenciálnej rovnice pre cenu dlhopisu v uzavretej forme je známe iba v špeciálnych prípadoch, napr. Vašíčkov model alebo CIR model s nulovou koreláciou. V ostatných prípadoch, sme odvodili aproximáciu riešenia založenú na myšlienke substitúcie konštantných volatilít, v riešení Vašíčkovho modelu, nekonštantnými volatilitami. Z dôvodu zlepšenia kvality zhody odhadnutých a presných výnosových kriviek sme navrhli niekoľko zmien v modeloch. Prvá z nich je založená na odhade výnosových kriviek z časovej štruktúry úrokových mier vo Vašíčkovom modeli. Krátkodobú úrokovú mieru považujeme za nepozorovateľnú premennú a odhadujeme ju spolu s ostatnými parametrami modelu. Druhý spôsob ako vylepšiť model je definovanie európskej krátkodobej úrokovej miery ako súčtu dvoch nepozorovateľných faktorov. Týmto spôsobom získavame trojfaktorový konvergenčný model. Odvodili sme presnosť aproximácie, navrhli sme postup kalibrácie a testovali sme ho na simulovaných a reálnych trhových dátach.

Druhá časť práce sa zameriava na numerické metódy. Najskôr študujeme Ficherovu teóriu, ktorá popisuje správne zaobchádzanie a definovanie okrajových podmienok pre parciálne diferenciálne rovnice, ktoré degenerujú na hranici. V práci uvádzame odvodenie Ficherových podmienok pre modely krátkodobých úrokových mier. Jadrom tejto časti sú ADE (alternating direction explicit) metódy zo 60. rokov 20. storočia, ku ktorým sa nenachádza veľa literatúry. V práci je obsiahnutá numerická analýza, štúdium stability a konzistencie pre konvekčno-difúzno-reakčnú rovnicu v jednorozmernom prípade. ADE metódy implementujeme pre dvojrozmerné call opcie a trojrozmerné spread opcie. Navrhujeme rozšírenia na viacrozmerné prípady Black-Scholesovho modelu. Túto časť práce ukončujeme alternatívnou metódou nazývanou Trefftz, ktorá patrí medzi Flexible Local Approximation MEthods (FLAME).

Kľúčové slová: krátkodobé modely úrokových mier, konvergenčný model úrokovej miery, oceňovanie dlhopisov, Black-Scholesov model, oceňovanie opcií, analytická aproximácia riešenia, rád presnosti, numerické metódy, kalibrácia, simulované dáta, reálne trhové dáta, ADE, splitting schémy, Trefftz metódy, Ficherova teória

German
Die Doktorarbeit beinhaltet verschiedene Methoden, die in der heutigen modernen Finanzmathematik eingesetzt werden. Es werden analytische und numerische Approximationsmethoden analysiert und diskutiert, sowie effektive Algorithmen für Multifaktormodelle zur Bewertung von Finanzderivaten entwickelt.

Der erste Teil der Doktorarbeit behandelt Modellierungsaspekte und ist auf die analytische Approximation von Zinssatzmodellen im Anleihenmarkt fokussiert. Wir behandeln ein Zweifaktorkonvergenzmodell mit nichtkonstanter Volatilität, das durch zwei stochastische Differentialgleichungen (SDG) gegeben ist. Das Modell beschreibt die Entwicklung von Zinsraten in Verbindung mit dem Eurowechselkurs. Ausgehend von der SDG ist es möglich eine partielle Differentialgleichung (PDG) für den Anleihekurs herzuleiten. Eine Angabe der Lösung der PDG ist nur in Einzelfällen in geschlossener Form möglich, z.B. im Vasicek or CIR Modell mit Korrelation null. In anderen Fällen haben wir eine Approximation an die Lösung des CIR Modells durch Ersetzen der konstanten Volatilität duch eine flexible Volatilität erhalten. Um eine höhere Genauigkeit bei der Anpassung an die reale Zinskurve zu erhalten, haben wir einige Änderungen innerhalb des Modells vorgeschlagen. Die erste basiert dabei auf der Schätzung des Momentanzinses durch die Zinsstrukturkurse innerhalb des Vasicek-Modells. Wir betrachten den Momentanzins im europäischen Modell für eine unbeobachtbare Variable und schätzen diese zusammen mit den anderen Modellparametern. Als zweite Verbesserungsmöglichkeit des Modells betrachten wir den europäischen Momentanzins als Summe von zwei unbeobachtbaren Prozessen. Auf diesem Wege erhalten wir ein Dreifaktorkonvergenzmodell. Wir zeigen die Genauigkeit dieser Approximationen, schlagen Kalibirierungsalgorithmen vor und testen die Modelle an simulierten, sowie realen Marktdaten

Der zweite Teil der vorliegenden Arbeit beschäftigt sich mit numerischen Methoden. Zuerst erläutern wir die Fichera-Theorie, die eine systematische Untersuchung von Randbedingungen erlaubt. Sie ist bei partiellen Differentialgleichungen, die am Rand degenerieren, von großem Nutzen. Es wird die Fichera-Funktion für Zinssatzmodelle hergeleitet. Den Kern der Doktorarbeit bilden Alternating Direction Explicit (ADE) Verfahren, aus den 60er des 20. Jahrhunderts die zu den nicht ausgiebig untersuchten Verfahren zählen. Daher existiert heute nur sehr wenig Literatur zu diesem Thema. Wir führen eine numerische Analyse durch und untersuchen die Stabilitäts- und Konsistenzeigenschaften für Konvektions-Diffusions-Reaktions Gleichungen in einer Raumdimension. Wir implementieren ADE Methoden für zweidimensionale Call-Optionen und dreidimensionale Spreadoptionsmodelle. Zusätzlich werden Erweiterungen für das höherdimensionale Black-Scholes-Modell vorgeschlagen. Wir beenden diesen Abschnitt der Doktorarbeit mit einer alternativen numerischen Methode, der sogenannten Trefftz-Methode, die zu der Klasse der Flexible Local Approximation MEthods (FLAME) gehört. Wir erläutern kurz ihre Nutzung im Rahmen der Finanzmathematik.

Schlüsselwörter: short rate Modelle (Momentanzins Modelle), Konvergenzmodell für Zinssätze, Anleihebewertung, Black-Scholes Modell, Optionsbewertung, approximierte analytische Lösung, Genauigkeitsordnung, Numerische Methoden, Kalibrierung, simulierten Daten, realen Marktdaten, ADE, splitting Schemas, Trefftz, Fichera Theorie

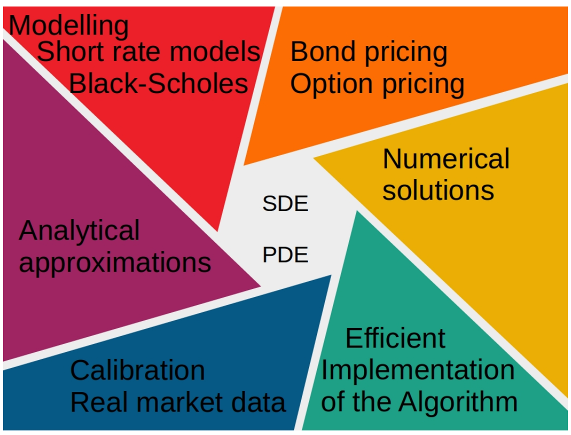
Contents

Abbreviations

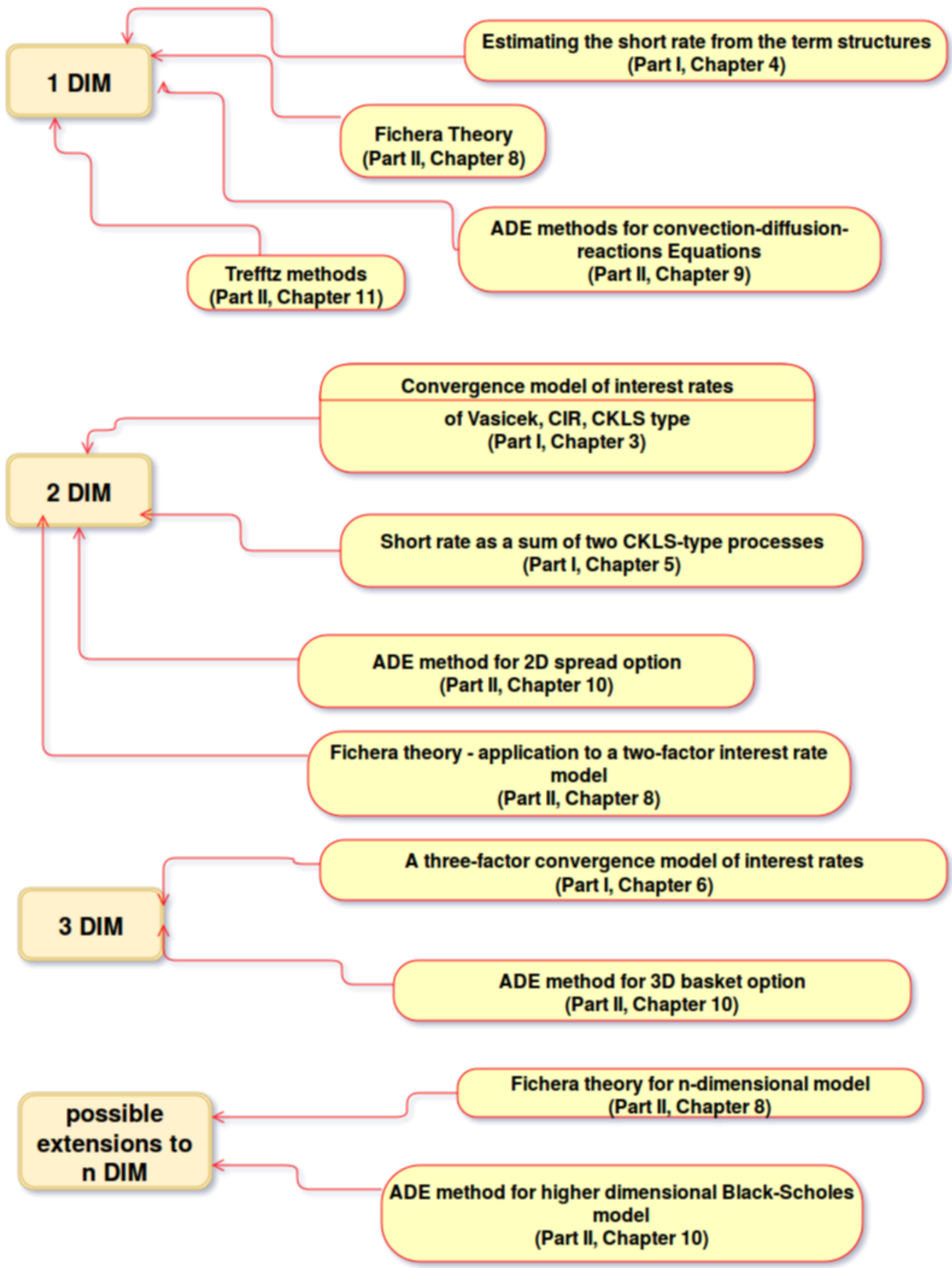
ODE	ordinary differential equation
PDE	partial differential equation
SDE	Stochastic Differential Equation
BS	Black-Scholes
CIR	Cox-Ingersoll-Ross model of short term, in which $\gamma = \frac{1}{2}$
CKLS	Chan-Karolyi-Longstaff-Sanders model of short term, in which γ is given generally
FDM	Finite Difference Methods
ADE	Alternating Direction Explicit
EURIBOR	European Interbank Offered Rate
BRIBOR	Bratislava Interbank Offered Rate
x_d	lower index d - dependence on variable x with respect to domestic data
x_e	lower index e - dependence on variable x with respect to European data
y'	derivation of function y with respect to the one unknown variable, used in ODEs
$\frac{dy}{dt}$	derivation of function y with respect to the one unknown variable t, used in ODEs
$\frac{\partial y}{\partial t}$	derivation of function y with respect to the one unknown variable t, used in PDEs

Foreword

The goal of the thesis is to provide a wide scope of techniques used in computational finance. On the one hand we see importance of the analytical techniques, on the other hand we tackle with numerical schemes. Another goal is to provide models and their solutions which are easy implementable. The better model, the better description of the reality. But the more complex model, the more troubles. Extension to higher dimensional (or nonlinear models) is necessary but our goal is to keep in mind the simpler model, the better. We do not want to deal with calibration and solving too complex models, because something it is even not possible. The suggested model or scheme must be tractable. In recent years we are witnesses of the negative interest rates in the whole European union. This fact must be considered and included to the all the models created in the way that they are capable to distinguish and cover all the situations. If the case is too complicated, we should provide an implementation in the way where it is possible and easy to provide parallelization of the algorithm. Computational finance is applied science and it requires knowledge from various fields in mathematics: SDEs, PDEs, analytical techniques, numerical analysis, optimization, programming and everything with having some knowledges from pricing of financial derivatives, such as options and bonds. Goal of this thesis is to cover all these subjects and suggest the effective methods for the given task. The aim was successfully reached and thesis is equally covering keywords from the Figure below.



In the thesis we deal with one-dimensional (or in other terminology one-factor), two-dimensional and three-dimensional models and we outline extensions to higher dimensional cases. Sorting according to dimensionality is displayed in the following diagram.



This thesis is cumulative one, the chapters are mostly based on the published results and each chapter is accesible in any order allowing a swift reading to readers. For reader interested in numerical analysis we refer to the second part of the thesis, for reader interested more in analytical techniques we recommend to read the first part of the thesis. For reader who would like to read the thesis just partially but from both parts something we recommend more valuable results can be found in the Chapter 3 and in the Chapter 9. The thesis does not have extended theoretical part but it is a collection of own research results.

1 Chapter 1

Outline of the thesis and related scientific works

1.1 Outline of the first part

In the first part of the thesis we deal with the suggestions for modeling of the interest rate in the multi-factor models which provide a good fitting to the real market data. Bonds are derivatives of the interest rate. As instantaneous interest rates (short rates) can be considered overnight rates (Eonia) or Euribor for a short time. Bond prices are specified by the parameters of the model, short rates and term structure of the interest rate. Observable data are e.g. Euribor data for 1, 2, 3, 6, 9, 12 months to the maturity. The first part of the thesis consists of four published papers.

We indicate an overview on the short rate models and solutions of given PDE in case that the solution is known. We focus on the derivation of an approximative analytical formula for the general case of models in which there is any exact explicit solution. We present a proof of accuracy of the proposed approximation. We propose a calibration algorithm based on using information from the term structures. Defining optimization tasks, all the model parameters were estimated. Using simulated and real market data, the algorithm was validated. On the one hand, suggested approximation and calibration algorithm provide reasonably accurate results which was proved, but on the other hand, by comparing accuracy of estimated and exact yield curves we did not achieve satisfactory results.

Chapter 3 is based on the following paper [14]

- *Z. Bučková (Zíková), B. Stehlíková, Convergence model of interest rates of CKLS type, Kybernetika 48(3), 2012, 567-586*

where we study special case of two-factor model: *convergence model of interest rate*, firstly defined by Corzo and Schwartz [20] in 2000 and its later generalizations. This model expresses the dependence of the evolution of the domestic short rate on the short rate of the monetary union (European monetary union). We study two-factor convergence models with different type of the volatility functions. For the two-factor convergence model with Vasicek type of the volatility (constant volatility) and CIR with zero correlation; there are known solutions in closed form formula. But there is no closed form formula for general CKLS model. The derivation of the analytical approximation of the 4th order and proof of its accuracy is given in [14]. We suggest an improvement of the approximation to the 6th order. Model is firstly tested on simulated data and approximative bond yields are compared with the reference solutions. We provide calibration procedure which is based on the derived approximative formula and usage of the real market term structure data and formulation of the optimization task, where the difference between the estimated and market yield curves is minimized. Since European parameters are not dependent on the domestic ones, we consider European bond yield parameters separately and we estimate them firstly. Secondly we take these estimates and use them as starting values for the estimates in the whole model. But still there are many parameters, so we provide an optimization algorithm in more steps: starting with the estimates of the risk-neutral drift, volatility and final adjustment of the parameters. Fitting of the bond yield curves is satisfactory on the simulated data, but not in case of real market data. But the weak point is already in the estimation of the European yield curves and this error is propagated to the domestic model. The one thing which is responsible for this behavior are 'bad data'. Overnight interest rates can be influenced by speculation on the market. Although we are not able to influence the data, we can model them. We deal with this idea in the Chapter 4 in our paper [31]

- *J. Halgašová, B. Stehlíková, Z. Bučková (Zíková): Estimating the short rate from the term structures in the Vasicek model, Tatra Mountains Mathematical Publications 61: 87-103, 2014*

To improve the quality in fitting exact yield curves by its estimates, we proposed a few changes in models. The first one is based on estimating the short rate from the term structures in the Vasicek model. We consider the short rate in the European model for unobservable variable and we estimate it together with other model parameters. The second way to improve a model is to define the European short rate as a sum of two unobservable factors (see the Chapter 5). In this way, we obtain a three-factor convergence model (see Chapter 6).

The Chapter 5 explains modeling of one interest rate as a sum of two unobservable processes. As an improvement for modeling of the European interest rate we suggest the short rate model of interest rates in which the short rate is defined as a sum of two stochastic factors. Each of these factors is modeled by SDEs with a linear drift and the volatility proportional to a power of the factor. We propose calibration methods which - under the assumption of constant volatilities – allow us to estimate the term structure of interest rate as well as the unobserved short rate, although we are not able to recover all the parameters. We apply it to real data and show that it can provide a better fit compared to a one-factor model. A simple simulated example suggests that the method can be also applied to estimate the short rate even if the volatilities have a general form. Therefore we propose an analytical approximation formula for bond prices in the model and derive the order of its accuracy.

The Chapter 5 is based on [10]

- *Z. Bučková, J. Halgašová, B. Stehlíková: Short rate as a sum of CKLS-type processes, accepted for publication in Proceedings of Numerical analysis and applications conference, Springer Verlag in LNCS, 2016.*

Separation in the stochastic interest rate to the two separable processes leads to more complex model. Instead of a one-factor short rate model, we have a two-factor model with correlation. We derive an analytical approximation with its accuracy for this model and we test it on simulated and real market data. The approximation in a two-factor model is much more better than approximation in a one-factor model. We take this advantage and we change modeled European short rate in the Chapter 3. Instead of one stochastic process for European short rate, we define two stochastic processes of CKLS-type whose sum represents European short rate. In total this modeling leads to the three SDEs, hence a three-factor convergence model. Its analytical solution, accuracy and numerical experiments can be found in [51]

- *B. Stehlíková, Z. Bučková (Zíková): A three-factor convergence model of interest rates. Proceedings of Algoritmy 2012, pp. 95-104.*

Combining two approaches from [14] and [10] we suggested a three-factor convergence model of interest rates. In all the previous models, the European rates are modeled by a one-factor model. This, however, does not provide a satisfactory fit to the market data. A better fit can be obtained using the model, where the short rate is a sum of two unobservable factors. We model European rate by 2 SDEs and the domestic interest rate by 1 stochastic differential equation. Therefore, we build the convergence model for the domestic rates based on this evolution of the European market. We study the prices of the domestic bonds in this model which are given by the solution of the PDEs. In general, it does not have an explicit solution. Hence we suggest an analytical approximative formula and derive the order of its accuracy in a particular case.

1.2 Outline of the second part

Alternatively to the analytical solutions of the PDEs there are numerical methods for pricing financial derivatives. In the second part of the thesis we deal with FDM (finite difference methods), esp. Alternative Direction Explicit (ADE) methods. We provide numerical analysis of ADE methods in one-dimensional cases and we suggest implementation algorithm for higher-dimensional models. The Chapter 8 is based on [9]

- *Z. Bučková, M. Ehrhardt, M. Günther: Fichera theory and its application to finance, Proceedings ECMI 2014, Taormina, Sicily, Italy, 2016*

we discuss theory from 1960 written by Gaetano Fichera. It is very useful for equations degenerating on the boundary, in terms of defining boundary conditions. According to the sign of the Fichera function there is a difference between the outflow boundary where we we must not supply BCs and inflow boundary where the definition of the BCs is needed.

In the Chapter 8 we apply this theory to the one-factor and two-factor interest rate model. Results of boundary decomposition to the regions with positive and negative Fichera function are displayed graphically. As a numerical example we apply FDM to the interest rate model and we display situation, where the non-respecting of the Fichera theory leads to the significant instability in the numerical solution. Results from the Fichera theory applied to the short rate models correspond to the well-known Feller condition. Since current interest rates may be negative, the Feller condition can be violated.

In two last Chapters 9 and 10 we study ADE schemes which were suggested in 1958 by Saul'ev, later developed by Larkin, Bakarat and Clark. In the last decade Leung and Osher and Daniel Duffy are dealing with these schemes. ADE schemes are efficient explicit FDM with the second order of accuracy and stability similar to the implicit schemes.
The Chapter 9 is based on the following paper [8]:

- *Z. Bučková, M. Ehrhardt, M. Günther: Alternating Direction Explicit Methods for Convection Diffusion Equations, Acta Math. Univ. Comenianae, Vol. LXXXI: 309–325, 2015*

where we investigate stability and consistency properties for one-dimensional convection-diffusion-reaction equations. The basic idea of the ADE schemes consists of combining two explicit solutions (called sweeps). Although the consistency of the single sweep is $O(k^2+h^2+\frac{k}{h})$, in the average the term $\frac{k}{h}$ is eliminated and the consistency of the final combined solution is of $O(k^2+h^2)$ order, where k is a time step and h is a space step. Stability analysis for various modifications of the ADE methods consists of proofs based on the matrix approach or von-Neumann stability approach.

The Chapter 10 is based on the paper [11]:

- *Z. Bučková, P. Pólvora, M. Ehrhardt, M. Günther: Implementation of Alternating Direction Explicit Methods to higher dimensional Black-Scholes Equation, AIP Conf. Proc. 1773, 030001; 2016*

where we suggest an algorithm for the implementation of the ADE schemes for higher-dimensional models. The number of sweeps is not increasing with dimension, hence n-dimensional model also requires for the purpose of preservation of the desired properties of the scheme two sweeps. Other literature sources state possible extension to higher-dimensional models but it is not really clear how to do it, hence it motivated us to suggest and describe algorithm properly. As test examples we consider two-dimensional spread option and three-dimensional basket call option. Numerical solution, with its accuracy and experimental order of convergence, is presented in the Chapter 10. ADE schemes are also good candidates to parallelize and they possess a good potential to succeed in the higher dimensional models, what is one of the current challenges in computational finance.

There is also some other alternative approach, e.g. the Trefftz method, based on the local approximation methods which is studied in the last chapter of the thesis.

Part I

Analytical Approximations of Interest Short Rate Models

Chapter 2

2 Introduction: Pricing of financial derivatives

Real market observable data give us information about the evolution and dynamics of interest rates, stock prices, exchange rates, ... The evolution is influenced by different factors, economic situations, investors' speculations, membership in the currency union and a lot of others. There are a lot of attempts to create models which capture the reality in the best ways. Adding more parameters, using non-constant or even stochastic variables, instead of constant variables; modeling using nonlinear equation - all these improvements lead to better models which can capture very well the market data. But how far should we go, how complex models should we construct? The more complex models, the more difficult calibration. Solutions in closed form formulas are available only for simple models.
There are various approaches to pricing. Dynamics of the stock price, interest rate, volatility are described by SDEs which can be solved using analytical methods or numerical simulations called Monte Carlo Method. From SDE using Itô formula and constructing risk-neutral portfolio we can derive the corresponding partial differential equation (PDE) which describes the price of the bond, or option. In our work we deal with the analytical and numerical PDE approaches. Interest rate modeling using short rate models, analytical approximations for bond pricing and its accuracy are discussed in detail in the first part of the thesis. Second part of the thesis is focused on the efficient numerical solutions of higher dimensional option pricing problems which are described by a parabolic PDE, also called also Black-Scholes (BS) model.

Financial derivatives are contracts, each of which value is derived are derived from the underlying assets. Interest rates, stock prices, indexes are used as underlying assets. They are derivatives of the interest rate are bonds, swaps, caps and floors. Typical derivatives of stocks are options. Financial derivatives are tools for protecting (hedging) the portfolio. Investors are looking for an optimal allocation of stocks and bonds in their portfolios, since they would like to minimize the risk and hence protect their portfolios. Stocks represent more risky assets with higher returns. Bonds bring lower risk and lower returns.

A bond is the simplest derivative of an interest rate which in the maturity time pays out its owner nominal value and in the arranged times pays out regular interest, called coupon. Bond with the nominal value equal to 1 is called a discount bond.

From the EURIBOR rates we can construct a yield curve which represents the dependence of the bond yield on the maturity of the bonds. They are usually increasing, because for the longer time, we lend with higher interest rates.

The limit value of the term structure of the interest rates (2.1, left), the $r(t) = \lim_{T \to t^+} R(t,T)$ is instantaneous interest rate, called *short rate* (2.1, right). It represents beginning of the yield curve, hence it is an interest rate for a very short time. In practice it is approximated by an interest rate with short maturity, e.g. overnight interest rates: EONIA (Euro Overnight Index Average). It is usual reference interest rate for one-day trading in European currency union.
For more details to interest rates see for example [7] , [50], [39].

Figure 2.1 shows an example of a short rate evolution and of a term structure at a given day.

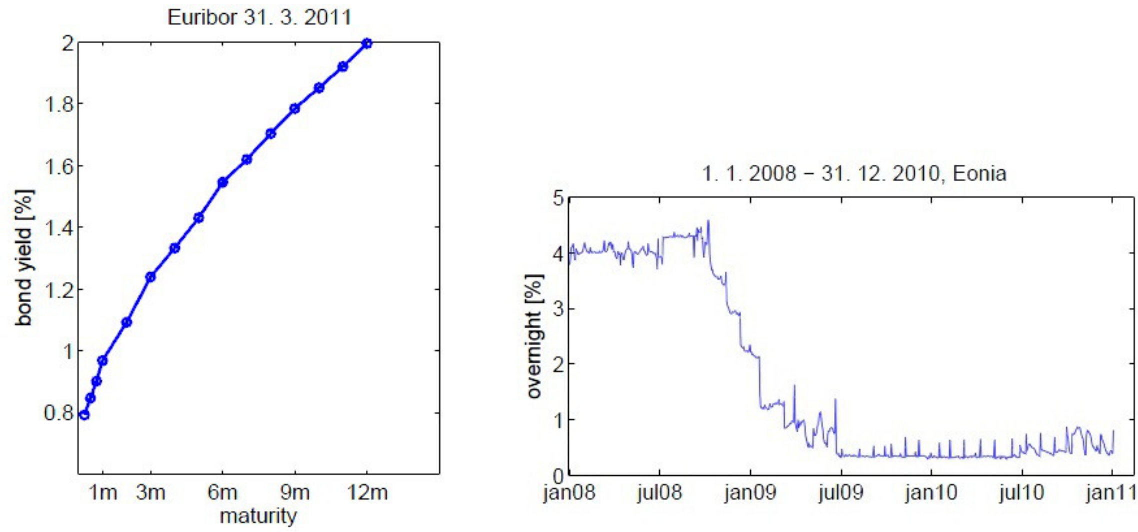


Figure 2.1: EURIBOR - term structure of interest rate (left), EONIA - short rate (right).

Good representation of the current state of the company is evolution of its stock price. Over the time they record some fluctuations around the drift. According to the drift we talk about bear market (decreasing trend) and bull market (increasing trend).

Option values are financial derivatives which are based on the stock prices. European call (put) option is a contract between share holders and its buyers which gives an opportunity to buy (sell) a stock for given price (strike price) at the maturity time. American type of options can be exercised in any time to the maturity time. There are different types of options, called exotic options, such as Asian, Barrier, Binary (Digital), Look-back, Rainbow, Russian, Bermudan and many others.

Models which are studied in this work, are described by the SDE:

$$\mathrm{d}X_t = \mu(X_t,t)\mathrm{d}t + \sigma(X_t,t)\mathrm{d}W,$$

where W is Wiener process. Function $\mu(X_t,t)$ is the trend or drift of the equation and

$\sigma(X_t, t)$ describes fluctuations around the drift. A solution of this SDE is a stochastic process X_t. For scalar X_t we have one-factor models, for vector valued X_t we deal with multi-factor models. Dynamics of the evolution of the process X_t is described by SDE, where X_t can represent an underlying asset, usually a stock price, but it can also be interest rate or volatility. In case X_t is a short rate, derived PDE represents equation for pricing of bonds. In case X_t is a stock price, derived PDE is an equation for option pricing, called also Black-Scholes model. If additionally there is given a stochastic volatility, we get the Heston model; and if there is given also a stochastic interest rate, it leads to the Heston-Hull-White model.

2.1 Bond pricing in short rate models

For modeling of the interest rate in our thesis we use short rate models. Alternative to the multi-factor short rate model represents Quadratic Gaussian, LIBOR (also known as BGM Model (Brace Gatarek Musiela Model)) and swap market models (an evolution of more forward interest rates) which are suitable for pricing of swaps and caps. There are known analytical solutions for some log-normal LIBOR and swap market models. Description of some of these models can be found in [7], [44].

A discount bond is a security which pays its holder a unit amount of money at specified time T (called maturity). $P(t,T)$ is the price of a discount bond with maturity T at time t. It defines the corresponding interest rate $R(t,T)$ by the formula

$$P(t,T) = e^{-R(t,T)(T-t)}, \text{i.e. } R(t,T) = -\frac{\ln P(t,T)}{T-t}.$$

A zero-coupon yield curve, also called term structure of interest rates, is then formed by interest rates with different maturities. Short rate (or instantaneous interest rate) is the interest rate for infinitesimally short time. It can be seen as the beginning of the yield curve: $r(t) = \lim_{t\to T^-} R(t,T)$. For a more detailed introduction to short rate modeling see e.g. [7], [34].

In short rate models, the short rate is modeled by a SDE. In particular, in Vasicek model [58], it is modeled by a mean-reverting Ornstein-Uhlenbeck process

$$dr = \kappa(\theta - r)dt + \sigma dw,$$

where κ, θ, σ are positive parameters and w is a Wiener process. It can be shown that after the specification of the so called market price of risk, the bond price $P(\tau, r)$ with maturity τ, when the current level of the short rate is r, is a solution to a parabolic PDE. In Vasicek model, it is customary to consider the constant market price of risk λ. Then, the bond price P satisfies

$$-\frac{\partial P}{\partial \tau} + (\kappa(\theta - r) - \lambda\sigma)\frac{\partial P}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 P}{\partial r^2} - rP = 0 \tag{2.1}$$

for all r and $\tau > 0$ and the initial condition $P(0,r) = 1$ for all r. This equation has an explicit solution which can be written as

$$\ln P(\tau,r) = \frac{1-e^{-\kappa\tau}}{\kappa}(R_\infty - r) - R_\infty\tau - \frac{\sigma^2}{4\kappa^3}(1-e^{-\kappa\tau})^2, \tag{2.2}$$

where $R_\infty = \frac{\kappa\theta-\lambda\sigma}{\kappa} - \frac{\sigma^2}{2\kappa^2}$ (see [58]). In Figure 2.2 we show a simulated behavior of the short rate (depicting also its equilibrium value θ) and term structures for several values of the short rate for the parameters equal to $\kappa = 5.00, \theta = 0.02, \sigma = 0.02, \lambda = -0.5$.

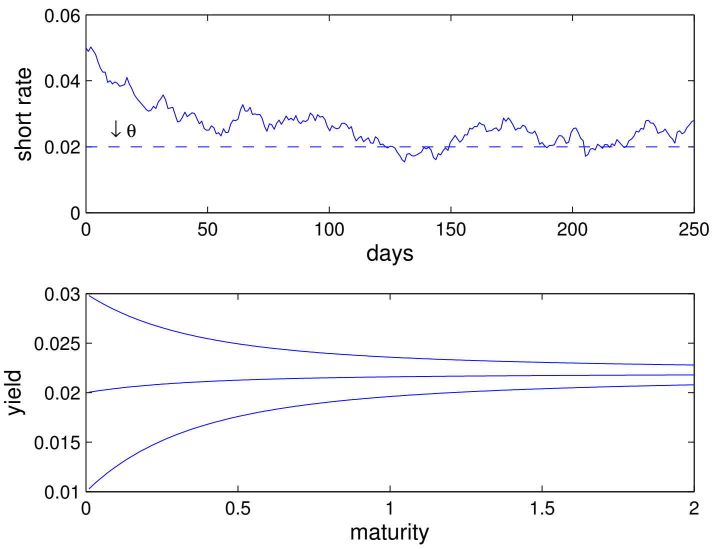


Figure 2.2: Simulated behavior of the short rate (above) and examples of term structures (below) for the parameters $\kappa = 5.00, \theta = 0.02, \sigma = 0.02, \lambda = -0.5$

Note that, although the model has four parameters - short rate parameters κ, θ, σ and market price of risk λ - parameters θ and λ enter the PDE (2.1) and hence also its solution (2.2) only through the term $\kappa\theta - \lambda\sigma$. Subsequently, it is possible to find a formula for bond price with three parameters. It is customary to do so by defining $\alpha = \kappa\theta - \lambda\sigma, \beta = -\kappa$. Parameters α, β are called risk neutral parameters, because they are related to an alternative formulation of the model in the so called risk neutral measure. For more details about risk neutral methodology see e.g. [34].

2.2 Short rate models

Short rate models are formulated by stochastic differential equation (SDE) for a variable X:

$$\mathrm{d}X_t = \mu(X_t,t)\mathrm{d}t + \sigma(X_t,t)\mathrm{d}W$$

which defines the short rate $r = r(X)$. Here W is a Wiener process, function $\mu(X,t)$ is the trend or drift part and the volatility $\sigma(X,t)$ represents fluctuations around the drift.

Choosing different drift $\mu(X,t)$ and volatility $\sigma(X,t)$ leads to various one-factor models (where X is a scalar) and multi-factor models (where X is a vector).

2.2.1 One-factor models

In one-factor models the evolution of the short rate is given by one scalar SDE:

$$\mathrm{d}r = \mu(r,t)\mathrm{d}t + \sigma(r,t)\mathrm{d}W. \tag{2.3}$$

The Table 2.1, with data taken from [7] and [50], gives an overview of one-factor models, in chronological order. We record also the equation for the dynamic of the short rate and distribution of the interest rate.

Year	**Model**	**SDE**	$r_t > 0$	$r_t \sim$	SOL
1977	Vasicek	$\mathrm{d}r_t = \kappa(\theta - r_t)\,\mathrm{d}t + \sigma \mathrm{d}W_t$	x	N	✓
1978	Dothan	$\mathrm{d}r_t = ar_t\mathrm{d}t + \sigma r_t \mathrm{d}W_t$	✓	LN	✓
1985	Cox-Ingersoll-Ross	$\mathrm{d}r_t = \kappa(\theta - r_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\mathrm{d}W_t$	✓	$NC\chi^2$	✓
1990	Hull & White	$\mathrm{d}r_t = \kappa(\theta_t - r_t)\,\mathrm{d}t + \sigma \mathrm{d}W_t$	x	N	✓
1990	Exponential Vasicek	$\mathrm{d}r_t = r_t(\mu - a\ln r_t)\mathrm{d}t + \sigma r_t \mathrm{d}W_t$	✓	LN	x
1991	Black & Karasinski	$\mathrm{d}r_t = r_t(\mu_t - a\ln r_t)\mathrm{d}t + \sigma r_t \mathrm{d}W_t$	✓	LN	x
1992	CKLS	$\mathrm{d}r_t = \kappa(\theta - r_r)\,\mathrm{d}t + \sigma r_t^{\gamma} \mathrm{d}W_t$	✓	-	x
2000	Mercurio & Moraleda	$\mathrm{d}r_t = r_t\left(\eta - \left(\lambda - \frac{\gamma}{a+\gamma_t}\right)\ln r_t\right)\mathrm{d}t + \sigma r_t \mathrm{d}W_t$	✓	LN	✓

Legend: Y=yes, N=no, SOL=existence of the explicit solution of the bond price in closed form formula, N=normal distribution, LN=log-normal distribution, $NC\chi^2$=non-central Chi-squared distribution.

Table 2.1: Overview of one-factor short rate models

If the drift of the process has the form $\mu(r,t) = \kappa(\theta - r)$, where the $\kappa, \theta > 0$ are constants, the model will have property called *mean-reversion*. It means, that short rate is pulling to the limit value θ. If the interest rate r is bigger than θ, drift $\kappa(\theta - r)$ is negative, so the interest rate is pulled down to the limit value θ. Vice-versa, if the interest rate r is smaller than the value of θ, the drift $\kappa(\theta - r)$ is positive, so it means the process is pulled up to θ. It is also the case of the Vasicek model, where $\mathrm{d}r = \kappa(\theta - r)\,\mathrm{d}t + \sigma \mathrm{d}W_t$. Its disadvantage is a normal distribution of the interest rates and in consequence it can lead to negative interest rates. A normal distribution of the interest rates is also in the Hull & White model. Other models have another distribution of interest rates which will not lead to the negative interest rates. It is also the case of the Dothan model which appeared one year after Vasicek model. But as an assumption there is a geometric Brownian motion for the short rate: $\mathrm{d}r = ar\mathrm{d}t + \sigma r \mathrm{d}W$. The explicit solution of this equation is $r(t) = r(0)e^{(a - \frac{\sigma^2}{2}t) + \sigma W(t)}$, hence $E[r(t)|r(0)] = r(0)e^{at}$. For $a \neq 0$ it is not realistic: if $a > 0$, then $E[r(t)|r(0)] \to \infty$ for $t \to \infty$, if $a < 0$, then $E[r(t)|r(0)] \to 0$. Therefore the Dothan model is sometimes given with $a = 0$ (e.g. in the book [34] , we have taken the formulation of the model from [7], where a is arbitrarily). It means, there is no trend in the evolution of the interest rate, only a random component. The positivity of the interest rates $r > 0$, is also preserved with usage of exponential functions, e.g. in the exponential Vasicek model or in the model of Black and Karasinski. In these models we suppose, that $\ln r_t$ (not directly interest rate) has a normal distribution. Then r_t has a log-normal distribution. Non-negativity of the interest rates in CIR model is preserved by the property: if r_t is close to the zero, then volatility is very small, almost zero and drift is positive. If r_t becomes zero, the volatility is zero and the drift is positive, hence r_t is getting back positive value. If the Feller condition $2\kappa\theta \geq \sigma^2$ is satisfied, then process r_t is positive with the probability equal to one [34].

In the Table 2.1 we notice, that for classic models as Vasicek, Cox-Ingersoll-Ross there is a closed form formula of the bond pricing equation. For other models, on the one hand it is not possible to express this solution, but on the other hand they are more realistic to describe the structure of the real market data. There are different approaches how to find approximative solutions of these models, by using analytical techniques to find accurate approximation or usage of numerical methods, such as finite difference methods, finite elements methods. As a generalization of the Vasicek and CIR model serves the CKLS model. We consider the CKLS model and we look for its solution. In the Table 2.1 we recognize models with time dependent drift function (e.g. Hull & White, Mercurio & Moradela, Black & Karasinski), also called non-arbitrage models. The proper choice of the function θ_t, μ_t, γ_t leads to the fit of the bond yield curve given by a yield curve observed in the market.

In short rate models, the bond prices are given as solutions to a PDE. We deal with the PDE approach in the following sections.

The model can be set in two ways:

- using SDE in the real (i.e. observed) probability measure and specifying so called market price of risk,
- using SDE in the risk-neutral probability measure.

The volatilities are the same in both measures and for the drift function holds:

$$\text{(risk-neutral drift)} = \text{(real drift)} - \text{(market price of risk)} \times \text{(volatility)}, \qquad (2.4)$$

see [39]. Considering the equation (2.3) in the real measure and market price of risk equal to $\lambda(r,t)$, the bond price $P(r,t,T)$ is a solution to the PDE (see [7]):

$$\frac{\partial P}{\partial t} + \mu(r,t)\frac{\partial P}{\partial r} + \frac{\sigma^2(r,t)}{2}\frac{\partial^2 P}{\partial r^2} - rP = 0, \qquad \text{for} \quad r > 0, \quad \in (0,T), \qquad (2.5)$$

satisfying the terminal condition $P(r,T,T) = 1$.

We present some well known models and we refer the reader to [7], [53] for more detailed treatment. The common feature of these models which will be useful later, is their linear drift in the risk-neutral measure:

$$dr = (b_1 + b_2 r)dt + \sigma(r,t)dw. \tag{2.6}$$

The simplest model is the Vasicek model which has a mean-reverting drift $\mu(r,t) = \kappa(\theta - r)$ and constant volatility $\sigma(r,t) = \sigma$ in the real measure, where $\kappa, \theta, \sigma > 0$ are constants. Assuming a constant market price of risk λ we obtain the risk-neutral form (2.6) with $b_1 = \kappa\theta - \lambda\sigma$, $b_2 = -\kappa$, $\sigma(r,t) = \sigma$. The PDE (2.5) for the bond price $P(\tau, r)$ with maturity $\tau = T - t$ has a solution in the form:

$$P(r,\tau) = e^{A(\tau) - rD(\tau)}, \tag{2.7}$$

where the functions $D(\tau), A(\tau)$ can be expressed as follows: (see, e.g. [7], [34], [39])

$$\begin{aligned} D(\tau) &= \frac{-1 + e^{b_2\tau}}{b_2}, \\ A(\tau) &= \Big(\frac{-1 + e^{b_2\tau}}{b_2} - \tau\Big) \times \Big(\frac{-b_1}{b_2} - \frac{\sigma^2}{2b_2^2}\Big) + \frac{\sigma^2}{4b_2^3}\left(1 - e^{b_2\tau}\right)^2. \end{aligned} \tag{2.8}$$

The Cox-Ingersoll-Ross (CIR) model also assumes a mean-reverting drift in the real measure, but the volatility is taken to be $\sigma(r,t) = \sigma\sqrt{r}$. If $\lambda(r,t) = \lambda\sqrt{r}$, we again obtain the risk-neutral process (2.6), this time with $b_1 = \kappa\theta, b_2 = -\kappa - \lambda\sigma, \sigma(r,t) = \sigma\sqrt{r}$. A solution of the PDE (2.5) again takes the form (2.7). Functions $A(\tau)$ and $D(\tau)$ are given by (see, e.g. ([7], [34], [39])

$$A(\tau) = -\frac{2b_1}{\sigma_e^2}\ln\frac{2\theta e^{(\theta - b_2)\frac{\tau}{2}}}{(\theta - b_2)(e^{\theta\tau} - 1) + 2\theta}, \qquad D(\tau) = \frac{-2\left(e^{\theta\tau} - 1\right)}{(\theta - b_2)(e^{\theta\tau} - 1) + 2\theta}, \tag{2.9}$$

where $\theta = \sqrt{b_2^2 + 2\sigma_e^2}$.

A convenient property of Vasicek and CIR models is the existence of explicit solutions to the bond pricing equation. However, their choice of volatility is not confirmed by analysis of real data. In their pioneering paper [16] Chan-Karolyi-Longstaff-Sanders (CKLS) considered a model with $\sigma(r,t) = \sigma r^\gamma$ with $\gamma \geq 0$. Most of the previously considered models, including Vasicek and CIR, do not provide a good fit to real market data except for a more general CKLS model which can copy reality in a better way. The CKLS model does not admit a closed form expression for bond prices. Approximate analytical solutions and their accuracy were studied in [18], [50], [51].

2.2.2 Two-factor models

Let us consider a model defined by the following system of SDEs:

$$\begin{aligned} dr &= \mu_r(r,x,t)dt + \sigma_r(r,x,t)dW_1, \\ dx &= \mu_x(r,x,t)dt + \sigma_x(r,x,t)dW_2, \end{aligned} \tag{2.10}$$

where $\rho \in (-1,1)$ is the correlation between the increments of Wiener processes W_1 and W_2, i.e. $Cov(dW_1,dW_2) = \rho dt$. The process x is a random process which is connected with an instantaneous rate. It can be a long-term interest rate, a short-term interest rate in another country, etc. Relations between real and risk-neutral parameters are analogous as in the one-factor case:

$$\begin{aligned}
&(\text{risk-neutral drift function})_r = (\text{real drift function})_r - \lambda_r(r,x,t) \times (\text{volatility})_r, \\
&(\text{risk-neutral drift function})_x = (\text{real drift function})_x - \lambda_x(r,x,t) \times (\text{volatility})_x,
\end{aligned}$$

where λ_r, λ_x are market prices of risk of the short rate and the factor x respectively.

If the short rate satisfies the SDE (2.10) in the real measure and market prices of risk are $\lambda_r(r,x,t), \lambda_x(r,x,t)$, then the bond price P satisfies the following PDE (assuming that the factor x is positive):

$$\begin{aligned}
&\frac{\partial P}{\partial t} + (\mu_r(r,x,t) - \lambda_r(r,x,t)\sigma_r(r,x,t))\frac{\partial P}{\partial r} + (\mu_x(r,x,t) - \lambda_x(r,x,t)\sigma_x(r,x,t))\frac{\partial P}{\partial x} \\
&\quad + \frac{\sigma_r(r,x,t)^2}{2}\frac{\partial^2 P}{\partial r} + \frac{\sigma_x(r,x,t)^2}{2}\frac{\partial^2 P}{\partial x} + \rho\sigma_r(r,x,t)\sigma_x(r,x,t)\frac{\partial^2 P}{\partial r \partial x} - rP = 0
\end{aligned}$$

for $r,x > 0$, $t \in (0,T)$ and the terminal condition $P(r,x,T) = 1$ for $r,x > 0$. The PDE is derived using Itô lemma and construction of risk-less portfolio, see, e.g. [34],[7].

2.2.3 Multi-factor short rate models

Considering more SDEs for the interest rate, we can capture reality much more better. The more underlying equations with more parameters, the better fitting of the real market data. The more complex model, the more difficult to calibrate it. So which model for term structures should one use? With these questions authors deal in the paper [46], where following criteria are considered.

'A practitioner wants a model which is

1. *flexible enough to cover most situations arising in practice;*
2. *simple enough that one can compute answers in reasonable time;*
3. *well-specified, in that required inputs can be observed or estimated;*
4. *realistic, in that the model will not do silly things.*

Additionally, the practitioner shares the view if an econometrician who wants

- *a good fit of the model to data;*

and a theoretical economist would also require

- *an equilibrium derivation of the model.'*

In our thesis we deal with two-factor convergence model of interest rates in the Chapter 3. A three dimensional model is considered in the Chapter 6. This way we could continue and extend the model, where dynamics of the process is described with n SDEs; which leads to the n-th order PDE. Theoretically unlimited extension of the model is possible, but we will care about simplicity of the model which can be solved in an reasonable time.

2.3 The calibration algorithm

Model calibration is certainly not a problem with a straightforward solution. In general, calibration methods can focus on statistical analysis of time series of instantaneous interest rate, match of theoretical and market yield curves, or combine these two approaches. An example of statistical analysis is the paper [16] where the form of the volatility in one-factor model is determined using the generalized method of moments applied to time series of short rate. An example of comparison of theoretical and estimated yield curves is in paper [49]. The existence of explicit formulae for bond prices in one-factor CIR model allowed the calibration of parameters in this way. The combination of these approaches can be found for example in paper [20] about the Vasicek convergence model. All parameters which can be estimated from the time series of domestic and European short rates, are estimated in this way. The remaining market prices of risk are then estimated using the yield curves.

However, using this combined approach, most parameters are estimated from the time series of the short rates. Information from the time series of interest rates with other maturities (which contain several times more data) is used only to estimate the market price of risk.

In [36] the authors consider a CIR convergence model and uses a modification of Ait-Sahalia's approximation of densities to estimate short rate parameters. Market price of risk is estimated from the yield curves. However, the author claims, that by changing some already estimated short rate parameters it is possible to obtain a significant improvement of objective function. Therefore, our aim was to propose such a calibration method that would use the information from term structures to estimate all parameters. Such an approach requires an efficient calculation of bond prices. This is achieved by using an approximative analytical formula.

Chapter 3
Convergence model of interest rates of CKLS type

This Chapter is based on the following paper [14]

- *Z. Bučková (Zíková), B. Stehlíková, Convergence model of interest rates of CKLS type, Kybernetika 48(3), 2012, 567-586*

Dynamics of the interest rate in various countries can be independent, but usually there are some connections and interaction which can be expressed as a mutual connection. Typical example is the of the interest rate before entering the monetary union.

It is worth noting that before adopting the euro, the interest rates in the country are influenced by the rates in the eurozone. We illustrate this with Figure 3.1, where we show the Slovak and eurozone instantaneous interest rates in the last quarter before Slovakia adopted the euro currency. These features in the models cause that we call them *convergence models of the interest rate*. Adoption of the new currency in the country is visible on the interest rate which converges to the interest rate of the bigger country or union. It describes an evolution of the short rate of the one small country with respect to the evolution of the short rate of the bigger country or a union of some countries. Convergence models of interest rates are studied in Chapter 3 and 6.

Convergence model of the interest rate is a special case of two-factor short rate models. Firstly it was proposed by Corzo and Schwartz in [20] with constant volatilities, it means as the two-factor Vasicek convergence model. In our Chapter 3 we suggest and study two-factor convergence model with nonconstant volatility, esp. of type of Chan-Karolyi-Longstaff-Sanders (CKLS) model. This form of volatility was first defined in [16]. Since solution of the suggested model is not given in a closed form formula (just for special cases) we are looking for a suitable approximation based on the idea of Stehlíková, Ševčovič [53]. Substituting its constant volatilities by instantaneous volatilities we obtain an approximation of the solution for a more general model. We compute the order of accuracy for this approximation, propose an algorithm for calibration of the model and we test it on the simulated and real market data.

The Chapter is organized as follows: The Section 3.1 describes convergence models as a special class of two-factor models. In particular, we present a generalization of the known models which we study in the following Sections. The closed form of the bond pricing equation is not known and hence in the Section 4 we propose an approximation formula for the domestic bond price and derive its accuracy. In Sections 3.4 - 3.8 we

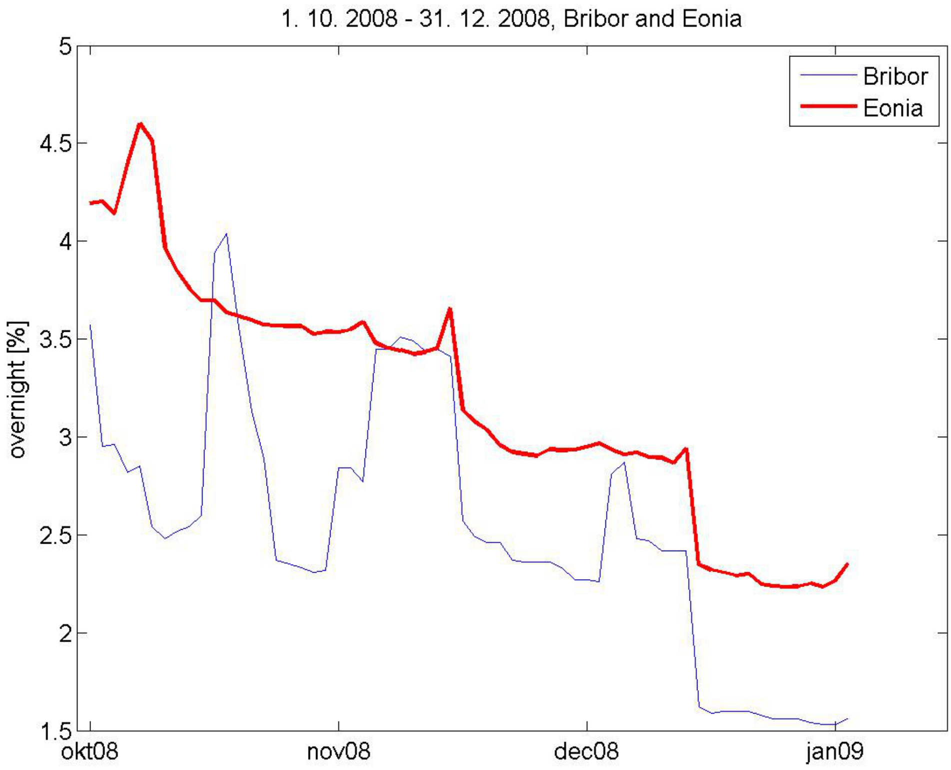


Figure 3.1: Instantaneous interest rate - Bribor and Eonia, last quarter 2008.

deal with calibration algorithm - its description, simulation study and application to real market data. In the last Section we give some concluding remarks.

3.1 Convergence models

Convergence models form a special class of two-factor models. A convergence model is used to model the entry of observed country into the European Monetary Union (EMU). It describes the behavior of two short-term interest rates, the domestic one and the instantaneous short rate for EMU countries. The european short rate is modeled using a one-factor model. It is assumed to have an influence on the evolution of the domestic short rate and hence it enters the SDE for its evolution. This kind of model was proposed for the first time in [20]. The model is based on Vasicek model, the volatilities of the short rates are constant. Analogical model of Cox-Ingersoll-Ross type, where the volatilities are proportional to the square root of the short rate, was considered in [35] and [36]. In the following sections we describe these two models and show how they price the bonds. Then we present a generalization with nonlinear volatility, which is analogous to the volatility in one-factor CKLS model.

3.1.1 The convergence model of Vasicek type

The first convergence model was proposed in the paper [20] by Corzo and Schwartz in the real probability measure:

$$\begin{aligned} \mathrm{d}r_d &= (a + b\,(r_e - r_d))\,\mathrm{d}t + \sigma_d \mathrm{d}W_d, \\ \mathrm{d}r_e &= (c\,(d - r_e))\,\mathrm{d}t + \sigma_e \mathrm{d}W_e, \end{aligned} \tag{3.1}$$

where $Cov(dW_1, dW_2) = \rho\mathrm{d}t$. They considered constant market prices of risk, i. e. $\lambda_d(r_d, r_e, \tau) = \lambda_d$ and $\lambda_e(r_d, r_e, \tau) = \lambda_e$. Hence for the European interest rate we have a one-factor Vasicek model and we can easily price European bonds. The coefficient $b > 0$ expresses the power of attracting the domestic short rate to the European one with the possibility of deviation determined by the coefficient a. Rewriting the model into risk-neutral measure we obtain:

$$\begin{aligned} \mathrm{d}r_d &= (a + b\,(r_e - r_d) - \lambda_d \sigma_d)\,\mathrm{d}t + \sigma_d \mathrm{d}W_d, \\ \mathrm{d}r_e &= (c\,(d - r_e) - \lambda_e \sigma_e)\,\mathrm{d}t + \sigma_e \mathrm{d}W_e, \end{aligned} \tag{3.2}$$

where $Cov[dW_d, dW_e] = \rho\mathrm{d}t$. We consider a more general model in risk-neutral measure, in which the risk-neutral drift of the domestic short rate is given by a general linear function of variables r_d, r_e and the risk-neutral drift of the European short rate is a general linear function of r_e. It means that the evolution of the domestic and the European short rates is given by:

$$\begin{aligned} \mathrm{d}r_d &= (a_1 + a_2 r_d + a_3 r_e)\,\mathrm{d}t + \sigma_d \mathrm{d}W_d, \\ \mathrm{d}r_e &= (b_1 + b_2 r_e)\,\mathrm{d}t + \sigma_e \mathrm{d}W_e, \end{aligned} \tag{3.3}$$

where $Cov[dW_d, dW_e] = \rho\mathrm{d}t$. Note that the system (3.3) has the form of the system (3.2) with $a_1 = a - \lambda_d \sigma_d$, $a_2 = -b$, $a_3 = b$, $b_1 = cd - \lambda_e \sigma_e$, $b_2 = -c$. The price $P(r_d, r_e, \tau)$ of a bond with time to maturity $\tau = T - t$ then satisfies the PDE

$$\begin{aligned} &-\frac{\partial P}{\partial \tau} + (a_1 + a_2 r_d + a_3 r_e)\frac{\partial P}{\partial r_d} + (b_1 + b_2 r_e)\frac{\partial P}{\partial r_e} \\ &\quad + \frac{\sigma_d^2}{2}\frac{\partial^2 P}{\partial r_d^2} + \frac{\sigma_e^2}{2}\frac{\partial^2 P}{\partial r_e^2} + \rho \sigma_d \sigma_e \frac{\partial^2 P}{\partial r_d \partial r_e} - r_d P = 0, \end{aligned} \tag{3.4}$$

for $r_d, r_e > 0$, $\tau \in (0, T)$ and the initial condition $P(r_d, r_e, 0) = 1$ for $r_d, r_e > 0$. Its solution can be found in the same way as in the original paper [20]. Assuming the solution in the form

$$P(r_d, r_e, \tau) = e^{A(\tau) - D(\tau) r_d - U(\tau) r_e}, \tag{3.5}$$

and inserting it into the equation (3.4) we obtain the system of ordinary differential equations (ODEs):

$$\begin{aligned}\dot{D}(\tau) &= 1 + a_2 D(\tau),\\ \dot{U}(\tau) &= a_3 D(\tau) + b_2 U(\tau),\\ \dot{A}(\tau) &= -a_1 D(\tau) - b_1 U(\tau) + \frac{\sigma_d^2 D^2(\tau)}{2} + \frac{\sigma_e^2 U^2(\tau)}{2} + \rho\sigma_d\sigma_e D(\tau)U(\tau)\end{aligned} \tag{3.6}$$

with initial conditions $A(0) = D(0) = U(0) = 0$. The solution of this system reads:

$$\begin{aligned}D(\tau) &= \frac{-1 + e^{a_2\tau}}{a_2},\\ U(\tau) &= \frac{a_3\left(a_2 - a_2 e^{b_2\tau} + b_2\left(-1 + e^{a_2\tau}\right)\right)}{a_2\left(a_2 - b_2\right) b_2},\\ A(\tau) &= \int_0^\tau -a_1 D(s) - b_1 U(s) + \frac{\sigma_d^2 D^2(s)}{2} + \frac{\sigma_e^2 U^2(s)}{2} + \rho\sigma_d\sigma_e D(s)U(s)\mathrm{d}s.\end{aligned} \tag{3.7}$$

Note that the function $A(\tau)$ can be easily written in the closed form without an integral. We leave it in this form for the sake of brevity. Furthermore, we consider only the case when $a_2 \neq b_2$. If $a_2 = b_2$, then $U(\tau)$ has another form, but it is a very special case and we will not consider it further.

3.1.2 Convergence model of CIR type

First we formulate the convergence model of CIR type (i.e.the volatilities are proportional to the square root of the short rates) in the real measure.

$$\begin{aligned}\mathrm{d}r_d &= (a + b\,(r_e - r_d))\,\mathrm{d}t + \sigma_d\sqrt{r_d}\mathrm{d}W_d,\\ \mathrm{d}r_e &= (c\,(d - r_e))\,\mathrm{d}t + \sigma_e\sqrt{r_e}\mathrm{d}W_e,\end{aligned} \tag{3.8}$$

where $Cov[dW_d, dW_e] = \rho\mathrm{d}t$. If we assume the market prices of risk equal to $\lambda_e\sqrt{r_e}$, $\lambda_d\sqrt{r_d}$ we obtain risk neutral processes of the form:

$$\begin{aligned}\mathrm{d}r_d &= (a_1 + a_2 r_d + a_3 r_e)\,\mathrm{d}t + \sigma_d\sqrt{r_d}\mathrm{d}W_d,\\ \mathrm{d}r_e &= (b_1 + b_2 r_e)\,\mathrm{d}t + \sigma_e\sqrt{r_e}\mathrm{d}W_e,\end{aligned} \tag{3.9}$$

where $Cov[dW_d, dW_e] = \rho\mathrm{d}t$. In what follows we consider this general risk- neutral formulation (3.9).

The European short rate is described by one-factor CIR model, so we are able to price European bonds using an explicit formula. The price of domestic bond $P(r_d, r_e, \tau)$ with maturity τ satisfies the PDE

$$
\begin{aligned}
&-\frac{\partial P}{\partial \tau}+(a_1+a_2r_d+a_3r_e)\frac{\partial P}{\partial r_d}+(b_1+b_2r_e)\frac{\partial P}{\partial r_e} \\
&+\frac{\sigma_d^2r_d^2}{2}\frac{\partial^2 P}{\partial r_d^2}+\frac{\sigma_e^2r_e^2}{2}\frac{\partial^2 P}{\partial r_e^2}+\rho\sigma_d\sqrt{r_d}\sigma_e\sqrt{r_e}\frac{\partial^2 P}{\partial r_d\partial r_e}-r_dP=0,
\end{aligned} \tag{3.10}
$$

for $r_d, r_e > 0, \tau \in (0,T)$ with the initial condition $P(r_d,r_e,0)=1$ for $r_d,r_e>0$. It was shown in [35] (in a slightly different parametrization of the model) that solution in the form (3.5) exists only when $\rho=0$. In this case we obtain system of ODEs

$$
\begin{aligned}
\dot{D}(\tau)&=1+a_2D(\tau)-\frac{\sigma_d^2D^2(\tau)}{2},\\
\dot{U}(\tau)&=a_3D(\tau)+b_2U(\tau)-\frac{\sigma_e^2U^2(\tau)}{2},\\
\dot{A}(\tau)&=-a_1D(\tau)-b_1U(\tau),
\end{aligned} \tag{3.11}
$$

with initial conditions $A(0)=D(0)=U(0)=0$ which can be solved numerically.

3.1.3 Convergence model of CKLS type

We consider a model in which the risk-neutral drift of the European short rate r_e is a linear function of r_e, risk-neutral drift of the domestic short rate r_d is a linear function of r_d and r_e and volatilities take the form $\sigma_e r_e^{\gamma_e}$ and $\sigma_d r_d^{\gamma_d}$, i.e.

$$
\begin{aligned}
\mathrm{d}r_d&=(a_1+a_2r_d+a_3r_e)\mathrm{d}t+\sigma_d r_d^{\gamma_d}dW_d,\\
\mathrm{d}r_e&=(b_1+b_2r_e)\mathrm{d}t+\sigma_e r_e^{\gamma_e}dW_e,
\end{aligned} \tag{3.12}
$$

where $Cov[dW_d,dW_e]=\rho\mathrm{d}t$. Parameters $a_1,a_2,a_3,b_1,b_2\in\mathbb{R}, \sigma_d,\sigma_e>0, \gamma_d,\gamma_e\geq 0$ are given constants and $\rho\in(-1,1)$ is a constant correlation between the increments of Wiener processes $\mathrm{d}W_d$ and $\mathrm{d}W_e$. We will refer to this model as *two-factor convergence model of Chan-Karolyi-Longstaff-Sanders (CKLS) type*. The domestic bond price $P(r_d,r_e,\tau)$ with the maturity τ satisfies the PDE:

$$
\begin{aligned}
&-\frac{\partial P}{\partial \tau}+(a_1+a_2r_d+a_3r_e)\frac{\partial P}{\partial r_d}+(b_1+b_2r_e)\frac{\partial P}{\partial r_e} \\
&+\frac{\sigma_d^2r_d^{2\gamma_d}}{2}\frac{\partial^2 P}{\partial r_d^2}+\frac{\sigma_e^2r_e^{2\gamma_e}}{2}\frac{\partial^2 P}{\partial r_e^2}+\rho\sigma_d r_d^{\gamma_d}\sigma_e r_e^{\gamma_e}\frac{\partial^2 P}{\partial r_d\partial r_e}-r_dP=0,
\end{aligned} \tag{3.13}
$$

for $r_d,r_e>0, \tau\in(0,T)$, with the initial condition $P(r_d,r_e,0)=1$ for $r_d,r_e>0$. Unlike for Vasicek and uncorrelated CIR model, in this case it is not possible to find solution in the separable form (3.5). For this reason, we are seeking for an approximative solution (3.13).

3.2 Approximation of the domestic bond price solution

The bond prices in the CKLS type convergence model are not known in a closed form. This has already been the case for the European bonds, i.e. for the one-factor CKLS model. We use the approximation from [51]. In this approximation we consider the one-factor Vasicek model with the same risk-neutral drift and we replace the constant volatility by the current volatility σr^γ in the closed form formula for the bond prices. We obtain

$$\ln P_e^{ap}(\tau,r) = \left(\frac{b_1}{b_2}+\frac{\sigma^2 r^{2\gamma}}{2b_2^2}\right)\left(\frac{1-e^{b_2\tau}}{b_2}+\tau\right)+\frac{\sigma^2 r^{2\gamma}}{4b_2^3}\left(1-e^{b_2\tau}\right)^2+\frac{1-e^{b_2\tau}}{b_2}r. \quad (3.14)$$

We use this approach to propose an approximation for the domestic bond prices. We consider the domestic bond prices in Vasicek convergence model with the same risk-neutral drift and we set $\sigma_d r_d^{\gamma_d}$ instead of σ_d and $\sigma_e r_e^{\gamma_e}$ instead of σ_e into (3.7). Hence, we have

$$\ln P^{ap} = A - Dr_d - Ur_e, \quad (3.15)$$

where

$$D(\tau) = \frac{-1+e^{a_2\tau}}{a_2},$$

$$U(\tau) = \frac{a_3\left(a_2 - a_2 e^{b_2\tau} + b_2\left(-1+e^{a_2\tau}\right)\right)}{a_2\left(a_2-b_2\right)b_2},$$

$$A(\tau) = \int_0^\tau -a_1 D(s) - b_1 U(s) + \frac{\sigma_d^2 r_d^{2\gamma_d} D^2(s)}{2} + \frac{\sigma_e^2 r_e^{2\gamma_e} U^2(s)}{2} + \rho\sigma_d r_d^{\gamma_d}\sigma_e r_e^{\gamma_e} D(s)U(s)\mathrm{d}s.$$

3.2.1 Accuracy of the approximation for CIR model with zero correlation

In CIR convergence model the domestic bond price $P^{CIR,\rho=0}$ exhibits a separable form (3.5) and functions A, D, U are characterized by a system of ODEs (3.11). This enables us to compute Taylor expansion of its logarithm around $\tau = 0$. We can compare it with the expansion of the proposed approximation $\ln P^{CIR,\rho=0,ap}$ (computed either using its closed form expression (3.15) or the system of ODEs (3.7) for the Vasicek convergence model). More detailed computation can be found in [13]. In this way we obtain the accuracy of the approximation for the CIR model with zero correlation:

$$\ln P^{CIR,\rho=0,ap} - \ln P^{CIR,\rho=0} = \frac{1}{24}\left(-a_2\sigma_d^2 r_d - a_1\sigma_d^2 - a_3\sigma_d^2 r_e\right)\tau^4 + o(\tau^4) \quad (3.16)$$

for $\tau \to 0^+$.

3.3 Numerical results for CIR model with zero correlation

Let us consider real measure parameters: $a = 0$, $b = 2$, $\sigma_d = 0.03$, $c = 0.2$, $d = 0.01$, $\sigma_e = 0.01$ and market price of risk $\lambda_d = -0.25$, $\lambda_e = -0.1$. In the risk-neutral setting (3.9) we have $a_1 = a - \lambda_d \sigma_d = 0.0075$, $a_2 = -b = -2$, $a_3 = b = 2$, $b_1 = cd - \lambda_e \sigma_e = 0.003$, $b_2 = -c = -0.2$, $\sigma_d = 0.03$, $\sigma_e = 0.01$. With the initial values for the short rates $r_d = 1.7\%$ a $r_e = 1\%$ we generate the evolution of domestic and European short rates using Euler-Maruyama discretization. In Figure 3.2 we see that this choice of parameters leads to a realistic behavior of interest rates. In Table 3.1 we compare the exact interest rate (i.e. the numerical solution of the system (2.8)) and the approximative interest rate given by (3.15). We observe very small differences. Note that the Euribor market data are recorded with the accuracy 10^{-3}. Choosing other days, with other combination of r_d, r_e, leads to very similar results. The difference between exact and approximative interest rates remains nearly the same.

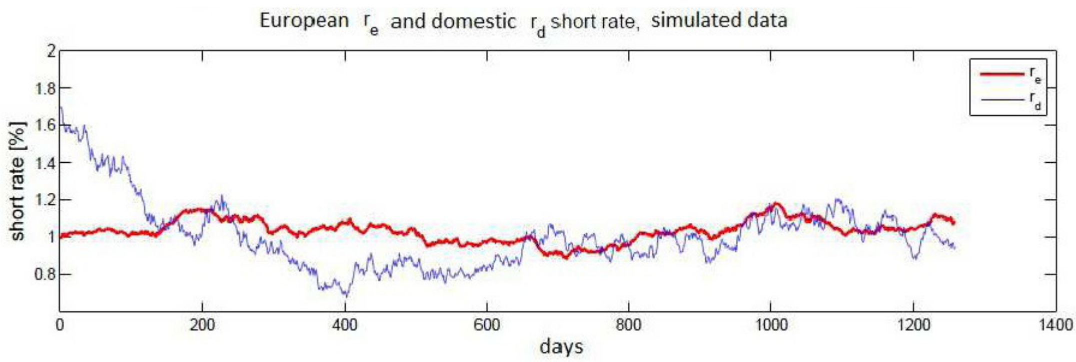


Figure 3.2: Simulation of European and domestic short rate for 1200 days.

Mat. [year]	Exact yield [%]	Aprox. yield [%]	Diff. [%]
$\frac{1}{4}$	1.63257	1.63256	7.1E-006
$\frac{1}{2}$	1.58685	1.58684	1.4E-005
$\frac{3}{4}$	1.55614	1.55614	4.8E-006
1	1.53593	1.53592	1.1E-005
5	1.56154	1.56155	-5.0E-006
10	1.65315	1.65323	-8.3E-005
20	1.74696	1.74722	-2.5E-004
30	1.78751	1.78787	-3.7E-004

Mat. [year]	Exact yield [%]	Aprox. yield [%]	Diff. [%]
$\frac{1}{4}$	1.08249	1.08250	-8.2E-006
$\frac{1}{2}$	1.15994	1.15996	-1.7E-005
$\frac{3}{4}$	1.21963	1.21964	-7.0E-006
1	1.26669	1.26671	-1.6E-005
5	1.53685	1.53691	-6.2E-005
10	1.65113	1.65127	-1.4E-004
20	1.74855	1.74884	-2.9E-004
30	1.78879	1.78918	-3.9E-004

Table 3.1: Exact and approximative domestic yield for 1st (left) observed day, $r_d = 1.7\%$, $r_e = 1\%$ and for 252nd (right) observed day, $r_d = 1.75\%$, $r_e = 1.06\%$.

3.3.1 Accuracy of the approximation for general CKLS model

The aim of this section is to derive the order of accuracy of the proposed approximation in the general case. We use an analogous method as in [51] and [53] for one-factor models and in [36] to study the influence of correlation ρ on bond prices in the convergence CIR model.

Let $f^{ex} = \ln P^{ex}$ be the logarithm of the exact price P^{ex} of the domestic bond in two factor convergence model of CKLS type. It satisfies the PDE (3.13). Let $f^{ap} = \ln P^{ap}$ be the logarithm of the approximative price P^{ap} for the domestic bond price given by (3.15). By setting f^{ap} to the left-hand side of (3.13) we obtain non-zero right-hand side which we denote as $h(r_d, r_e, \tau)$. We expand it into Taylor expansion and obtain that

$$h(r_d, r_e, \tau) = k_3(r_d, r_e)\tau^3 + k_4(r_d, r_e)\tau^4 + o(\tau^4), \tag{3.17}$$

for $\tau \to 0^+$, where

$$k_3(r_d, r_e) = \frac{1}{6}\sigma_d^2\gamma_d r_d^{2\gamma_d-2}\left(2a_1r_d + 2a_2r_d^2 + 2a_3r_dr_e - r_d^{2\gamma_d}\sigma_d^2 + 2\gamma_d r_d^{2\gamma_d}\sigma_d^2\right),$$

$$\begin{aligned}
k_4(r_d, r_e) =& \frac{1}{48}\frac{1}{r_e^2}r_d^{-2+\gamma_d}\sigma_d\Big(12a_2^2\gamma_d r_d^{2+\gamma_d}r_e^2\sigma_d - 16\gamma_d r_d^{1+3\gamma_d}r_e^2\sigma_d^3 + 6a_3b_1\gamma_e r_d^2 r_e^{1+\gamma_e}\rho\sigma_e \\
&+ 6a_3b_2\gamma_e r_d^2 r_e^{2+\gamma_e}\rho\sigma_e + 6a_3^2\gamma_d r_d r_e^{3+\gamma_e}\rho\sigma_e - 3a_3\gamma_d r_d^{2\gamma_d}r_e^{2+\gamma_e}\rho\sigma_d^2\sigma_e \\
&+ 3a_3\gamma_d^2 r_d^{2\gamma_d}r_e^{2+\gamma_e}\rho\sigma_d^2\sigma_e + 6a_3\gamma_d\gamma_e r_d^{1+\gamma_d}r_e^{1+2\gamma_e}\rho^2\sigma_d\sigma_e^2 - 3a_3\gamma_e r_d^2 r_e^{3\gamma_e}\rho\sigma_e^3 \\
&+ 3a_3\gamma_e^2 r_d^2 r_e^{3\gamma_e}\rho\sigma_e^3 + 6a_1\gamma_d r_d r_e^2\left(2a_2r_d^{\gamma_d}\sigma_d + a_3r_e^{\gamma_e}\rho\sigma_e\right) \\
&+ 6a_2\gamma_d r_e^2\left(\left(-1+2\gamma_d\right)r_d^{3\gamma_d}\sigma_d^3 + a_3r_d\left(2r_d^{\gamma_d}r_e\sigma_d + r_dr_e^{\gamma_e}\rho\sigma_e\right)\right)\Big).
\end{aligned}$$

We define the function $g(\tau, r_d, r_e) := f^{ap} - f^{ex} = \ln P^{ap} - \ln P^{ex}$ as a difference between logarithm of the approximation and the exact price. Using the PDEs satisfied by f^{ex} and f^{ap} we obtain the following PDE for the function g:

$$\begin{aligned}
&-\frac{\partial g}{\partial \tau} + (a_1 + a_2r_d + a_3r_e)\frac{\partial g}{\partial r_d} + (b_1 + b_2r_e)\frac{\partial g}{\partial r_e} + \frac{\sigma_d^2 r_d^{2\gamma_d}}{2}\left[\left(\frac{\partial g}{\partial r_d}\right)^2 + \frac{\partial^2 g}{\partial r_d^2}\right] \\
&+\frac{\sigma_e^2 r_e^{2\gamma_e}}{2}\left[\left(\frac{\partial g}{\partial r_e}\right)^2 + \frac{\partial^2 g}{\partial r_d^2}\right] + \rho\sigma_d r_d^{\gamma_d}\sigma_e r_e^{\gamma_e}\left(\frac{\partial g}{\partial r_d}\frac{\partial g}{\partial r_e} + \frac{\partial^2 g}{\partial r_d \partial r_e}\right) \\
&= h(r_d, r_e, \tau) + \frac{\sigma_d^2 r_d^{2\gamma_d}}{2}\left[\left(\frac{\partial f^{ex}}{\partial r_d}\right)^2 - \frac{\partial f^{ap}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_d}\right] + \frac{\sigma_e^2 r_e^{2\gamma_e}}{2}\left[\left(\frac{\partial f^{ex}}{\partial r_e}\right)^2 - \frac{\partial f^{ap}}{\partial r_e}\frac{\partial f^{ex}}{\partial r_e}\right] \\
&+\rho\sigma_d r_d^{\gamma_d}\sigma_e r_e^{\gamma_e}\left[2\frac{\partial f^{ex}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ap}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ex}}{\partial r_d}\frac{\partial f^{ap}}{\partial r_e}\right].
\end{aligned} \tag{3.18}$$

Suppose that $g(r_d, r_e, \tau) = \sum_{k=\omega}^{\infty} c_k(r_d, r_e)\tau^k$. For $\tau = 0$ is both the exact and approximative bond price equal to one, so $f^{ex}(r_d, r_e, 0) = f^{ap}(r_d, r_e, 0) = 0$. It means that $\omega > 0$ and on the left hand side of the equation (5.10) the term with the lowest order is $c_\omega \omega \tau^{\omega-1}$. Now we investigate the order of the right hand side of the equation 5.10.

We know that $f^{ex}(r_d, r_e, 0) = 0$. It means that $f^{ex} = O(\tau)$ and also partial derivation $\frac{\partial f^{ex}}{\partial r_d}$ and $\frac{\partial f^{ex}}{\partial r_e}$ are of the order $O(\tau)$. From the approximation formula (3.15) we can see that $\frac{\partial f^{ap}}{\partial r_d} = O(\tau)$, $\frac{\partial f^{ap}}{\partial r_e} = O(\tau^2)$. Since $h(r_d, r_e, \tau) = O(\tau^3)$, the right hand side of the equation (5.10) is at least of the order τ^2. The left hand side of the equation (5.10) is of the order $\tau^{\omega-1}$ and hence $\omega - 1 \geq 2$, i.e. $\omega \geq 3$. It means that

$$f^{ap}(r_d, r_e, \tau) - f^{ex}(r_d, r_e, \tau) = O(\tau^3).$$

Using this expression we can improve the estimation of the derivative $\frac{\partial f^{ex}}{\partial r_e}$ as follows: $\frac{\partial f^{ex}}{\partial r_e} = \frac{\partial f^{ap}}{\partial r_e} + O(\tau^3) = O(\tau^2) + O(\tau^3) = O(\tau^2)$. We also estimate the terms on the right hand side in the equation (5.10):

$$\left(\frac{\partial f^{ex}}{\partial r_d}\right)^2 - \frac{\partial f^{ap}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_d} = \frac{\partial f^{ex}}{\partial r_d}\left(\frac{\partial f^{ex}}{\partial r_d} - \frac{\partial f^{ap}}{\partial r_d}\right) = O(\tau).O(\tau^3) = O(\tau^4), \quad (3.19)$$

$$\left(\frac{\partial f^{ex}}{\partial r_e}\right)^2 - \frac{\partial f^{ap}}{\partial r_e}\frac{\partial f^{ex}}{\partial r_e} = \frac{\partial f^{ex}}{\partial r_e}\left(\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ap}}{\partial r_e}\right) = O(\tau^2).O(\tau^3) = O(\tau^5), \quad (3.20)$$

$$\begin{aligned} 2\frac{\partial f^{ex}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ap}}{\partial r_d}\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ex}}{\partial r_d}\frac{\partial f^{ap}}{\partial r_e} &= \frac{\partial f^{ex}}{\partial r_d}\left(\frac{\partial f^{ex}}{\partial r_e} - \frac{\partial f^{ap}}{\partial r_e}\right) \\ + \frac{\partial f^{ex}}{\partial r_e}\left(\frac{\partial f^{ex}}{\partial r_d} - \frac{\partial f^{ap}}{\partial r_d}\right) &= O(\tau).O(\tau^3) + O(\tau^2).O(\tau^3) = O(\tau^4) + O(\tau^5) = O(\tau^4). \end{aligned} \quad (3.21)$$

Since $h(r_d, r_e, \tau) = O(\tau^3)$, the right hand side of the equation (5.10) is $O(\tau^3)$ and the coefficient at τ^3 is the coefficient of the function $h(r_d, r_e, \tau)$ at τ^3, i.e. $k_3(r_d, r_e)$. It means that $\omega = 4$ and comparing the coefficients at τ^3 on the left and right-hand side of (5.10) we obtain $-4c_4(r_d, r_e) = k_3(r_d, r_e)$, i.e. $c_4(r_d, r_e) = -\frac{1}{4}k_3(r_d, r_e)$. Hence we have proved the following theorem.

Theorem 3.1. *Let $P^{ex}(r_d, r_e, \tau)$ be the price of the domestic bond in two-factor CKLS convergence model, i.e. satisfying equation (3.13) and let P^{ap} be the approximative solution defined by (3.15). Then*

$$\ln P^{ap}(r_d, r_e, \tau) - \ln P^{ex}(r_d, r_e, \tau) = c_4(r_d, r_e)\tau^4 + o(\tau^4)$$

for $\tau \to 0^+$, where coefficient c_4 is given by

$$c_4(r_d, r_e) = -\frac{1}{24}\sigma_d^2 \gamma_d r_d^{2\gamma_d - 2}\left(2a_1 r_d + 2a_2 r_d^2 + 2a_3 r_d r_e - r_d^{2\gamma_d}\sigma_d^2 + 2\gamma_d r_d^{2\gamma_d}\sigma_d^2\right). \quad (3.22)$$

Note that if we substitute $\gamma_d = \frac{1}{2}$ and $\rho = 0$ into Theorem 3.1, we obtain the formula (3.16) for CIR model derived earlier in (3.16).

3.3.2 Improvement of the approximation

In some cases it is possible to improve an approximation by calculating more terms in Taylor expansion of the function $g = \ln P^{ap} - \ln P^{ex}$. It is so also in this case. Using that $f^{ap} - f^{ex} = O(\tau^4)$, we are able to improve estimates (3.19) and (3.21) and to deduce that also the coefficient at τ^4 on the right hand side of equation (5.10) comes only from the function h. Hence it is equal to $k_4(r_d, r_e)$ which is given by (3.18). Comparing coefficients at τ^4 on the left and right hand side of (5.10) we obtain:

$$-5c_5 + (a_1 + a_2 r_d + a_3 r_e)\frac{\partial c_4}{\partial r_d} + (b_1 + b_2 r_e)\frac{\partial c_4}{\partial r_e}$$
$$+\frac{\sigma_d^2 r_d^{2\gamma_d}}{2}\frac{\partial^2 c_4}{\partial r_d^2} + \frac{\sigma_e^2 r_e^{2\gamma_e}}{2}\frac{\partial^2 c_4}{\partial r_e^2} + 4\rho\sigma_d r_d^{\gamma_d}\sigma_e r_e^{\gamma_e}\frac{\partial^2 c_4}{\partial r_d \partial r_e} = k_4$$

which enables us to express c_5 using already known quantities.

Let us define an approximation $\ln P^{ap2}$ by:

$$\ln P^{ap2}(r_d, r_e, \tau) = \ln P^{ap} - c_4(r_d, r_e)\tau^4 - c_5(r_d, r_e)\tau^5.$$

Then $\ln P^{ap2} - \ln P^{ex} = O(\tau^6)$ and therefore the new approximation $\ln P^{ap2}$ is of the order $O(\tau^6)$.

3.4 Formulation of the optimization problems in the calibration algorithm

We consider the convergence model of CKLS type in the risk-neutral measure given by equation (3.12). Firstly let us define the following notation

- P_d^{ap}, P_e^{ap} are approximations of the price of domestic and European bonds,
- R_d, R_e are actual yields observed on the market,
- $\tau_d = (\tau_d^1, \ldots \tau_d^{m_d})$, $\tau_e = (\tau_e^1, \ldots \tau_e^{m_e})$ are maturities of domestic and European yields,
- the data are observed during n_d, resp. n_e days,
- index i corresponds to days and index j corresponds to maturities.

We consider estimation of the parameters of European interest rates as a separate problem. We assume that the relationship between European and domestic interest rates is not a mutual influence of two variables, but the European rates influence the domestic ones. Hence the estimated European parameters of the model can not be dependent on the choice of country for which we consider the convergence model and on the domestic interest rates in this country. This approach was also used in [35]. The calibration procedure is therefore divided into two steps:

1. Estimation of European parameters which is based on minimizing the function

$$F_e(b_1,b_2,\sigma_e,\gamma_e) = \frac{1}{m_e n_e}\sum_{i=1}^{n_e}\sum_{j=1}^{m_e} w_e(i,j)\Big(-\frac{\ln P_e^{ap}(i,j)}{\tau_e(j)} - R_e(i,j)\Big)^2.$$

2. Estimation of domestic parameters which is based on minimizing the function

$$F_d(a_1,a_2,a_3,\sigma_d,\rho,\gamma_e) = \frac{1}{m_d n_d}\sum_{i=1}^{n_d}\sum_{j=1}^{m_d} w_d(i,j)\Big(-\frac{\ln P_d^{ap}(i,j)}{\tau_d(j)} - R_d(i,j)\Big)^2,$$

   where in the computation of P_d^{ap} we use the values b_1, b_2, σ_e obtained in the first step.

Functions w_e, w_d express weights. As in [49] we choose $w_e(i,j) = \tau_e(j)^2$ a $w_d(i,j) = \tau_d(j)^2$. However, the proposed algorithm can be adapted also for a different choice of weights. For our choice of weights we have the following objective functions:

$$F_e(b_1,b_2,\sigma_e,\gamma_e) = \frac{1}{m_e n_e}\sum_{i=1}^{n_e}\sum_{j=1}^{m_e}\Big(\ln P_e^{ap}(i,j) + R_e(i,j)\tau_e(j)\Big)^2, \tag{3.23}$$

$$F_d(a_1,a_2,a_3,\sigma_d,\rho,\gamma_d) = \frac{1}{m_d n_d}\sum_{i=1}^{n_d}\sum_{j=1}^{m_d}\Big(\ln P_d^{ap}(i,j) + R_d(i,j)\tau_d(j)\Big)^2. \tag{3.24}$$

3.5 The algorithm for estimating parameters in the CIR model with zero correlation

Our first goal is to estimate parameters in the convergence CIR model with zero correlation. In this case we can calculate the exact yield curve by solving a system of ODEs (3.11). We can therefore simulate the exact data and monitor the accuracy of our calibration and its individual steps.

3.5.1 Simulated data

We choose the same parameters as in the Section 3.3. We generate the domestic and European short rates for $n_e = n_d = 1260$ days, i.e.5 years (252 days per year). Furthermore, we generate European (using the explicit formula) and domestic (numerically solving the system of ODEs) yields for maturities $\tau_e = \tau_d = \left(\frac{1}{12}, \frac{2}{12}, \dots \frac{12}{12}\right)$, i.e. $m_e = m_d = 12$.

3.5.2 Estimation of the European parameters

This step is an estimate of the CIR model parameters. Estimation method is taken from [51]. For a given value of the power γ_e ($\frac{1}{2}$ in this case) the estimation of the other three parameters can be reduced to a one-dimensional problem. The remaining two parameters can be expressed from the first-order conditions and substituted into the objective function which then becomes of one parameter. The objective function is then optimized with respect to this parameter.

We note (see again [51]) that if we estimate a model with different γ_e than the true value (in particular, if we estimate a Vasicek model parameter), the estimate of the risk-neutral drift does not change much. This feature was an inspiration for estimation of domestic parameters, described in the next section.

3.5.3 Estimation of the domestic parameters

Step 1: Estimation of the risk-neutral drift
Based on the results for the one-factor model we try to estimate the risk-neutral drift of domestic interest rates as risk-neutral drift for Vasicek convergence model. When doing so, we add an upper index *vas* to all parameters and objective function i.e. F_d^{vas}, a_1^{vas}, a_2^{vas}, etc., to emphasize that we are estimating Vasicek model. We omit this index when using them in the subsequent steps.

The first step is to estimate the one-factor Vasicek model parameters for European data for which we use algorithm from [51]. Then, to estimate the domestic parameters, we optimize the function F_d^{vas} in the form:

$$F_d^{vas}(a_1^{vas}, a_2^{vas}, a_3^{vas}, \sigma_d^{vas}) = \sum_{i=1}^{n_d} \sum_{j=1}^{m_d} \left(\ln P_d^{vas}(i,j) + R_d(i,j)\tau_d(j) \right)^2. \tag{3.25}$$

Recall that in Vasicek model

$$\ln P_d^{vas}(r_d, r_e, \tau) = A(\tau) - D(\tau) r_d - U(\tau) r_e, \tag{3.26}$$

hence the term in F_d^{vas} corresponding to the particular $j-$th maturity $\tau_d(j)$ is

$$\sum_{i=1}^{n_d} \left(A(\tau_d(j)) - D(\tau_d(j)) r_d(i) - U(\tau_d(j)) r_e(i) + R_d(i,j)\tau_d(j) \right)^2. \tag{3.27}$$

Since j is fixed, values $A(\tau_d(j))$, $-D(\tau_d(j))$, $-U(\tau_d(j))$ are constants. The sum (3.27) which should be small (to minimize sum over j), then resembles the linear regression

$$-R_d(i,j)\tau_d(j) \sim c_{0j} + c_{1j} r_d(i) + c_{2j} r_e(i) \qquad \text{for} \quad i = 1, \dots n. \tag{3.28}$$

For each j we solve this linear regression and write the results into matrix

$$C = \begin{bmatrix} c_{01} & c_{11} & c_{21} \\ c_{02} & c_{12} & c_{22} \\ \vdots & \vdots & \vdots \\ c_{0m_d} & c_{1m_d} & c_{2m_d} \end{bmatrix}.$$

Comparing (3.26) and (3.28) we see that

$$c_{0j} \sim A(\tau_d(j)), \qquad c_{1j} \sim -D(\tau_d(j)), \qquad c_{2j} \sim -U(\tau_d(j)). \tag{3.29}$$

We determine the parameters of the functions A, D, U to obtain a good match of the terms in (3.29).

- Function D depends only on the parameter a_2^{vas}. We solve one-dimensional optimization problem

$$G_1(a_2^{vas}) = \sum_{j=1}^{m_d} \Big(-D(\tau_d(j), a_2^{vas}) - c_{1j} \Big)^2 \to \min_{a_2^{vas}} \tag{3.30}$$

  and we obtain the estimate of the parameter a_2^{vas}.

- The function U depends on the parameters a_3^{vas}, b_1^{vas}. Parameter b_1^{vas} is already estimated from European interest rates. Hence we have a one-dimensional optimization problem again:

$$G_2(a_2^{vas}) = \sum_{j=1}^{m_d} \Big(-U(\tau_d(j), a_3^{vas}) - c_{2j} \Big)^2 \to \min_{a_3^{vas}} \tag{3.31}$$

  and by solving it we obtain the estimate of a_3^{vas}.

- Function A depends on all parameters a_1^{vas}, a_2^{vas}, a_3^{vas}, b_1^{vas}, b_2^{vas}, σ_d^{vas}, σ_e^{vas}, but all parameters except a_1^{vas}, σ_d^{vas} are already estimated. Note that A is a linear function of the parameters a_1^{vas} and $(\sigma_d^{vas})^2$. Therefore the optimal solution of the problem

$$G_3(a_1^{vas}, (\sigma_d^{vas})^2) = \sum_{j=1}^{m_d} \Big(A(\tau_d(j), a_1^{vas}, (\sigma_d^{vas})^2) - c_{0j} \Big)^2 \to \min_{a_1^{vas}, (\sigma_d^{vas})^2} \tag{3.32}$$

  can be calculated explicitly from the first order optimality conditions by solving system of two linear equations. However, we observed (for several sets of generated data) that these estimates are unstable because the system matrix is ill-conditioned, with the condition number between 10^{18} and 10^{21}. It turned out that a better approach is to use only the first order condition from the derivative with respect to $(\sigma_d^{vas})^2$. Hence we proceed as follows.

  The function $\ln P^{ap}$ is expressed in the form:

$$\ln P^{ap} = A(\tau) + D(\tau) r_d + U(\tau) r_e = c_0(r_d, r_e, \tau) + c_1(r_d, r_e, \tau)(\sigma_d^{vas})^2, \tag{3.33}$$

where the coefficients $c_0(r_d, r_e, \tau)$, $c_1(r_d, r_e, \tau)$ do not depend on $(\sigma_d^{vas})^2$ and can be expressed explicitly. For given values of remaining parameters, the optimal value of $(\sigma_d^{vas})^2$ is calculated.

Thus for each value a_1^{vas} we have the corresponding optimal value $(\sigma_d^{vas})^2$ and we can formulate a one-dimensional optimization problem:

$$G_4(a_1^{vas}) = \sum_{j=1}^{m_d} \Big(A(\tau_d(j), a_1^{vas}) - c_{0j} \Big)^2 \to \min_{a_1^{vas}}. \tag{3.34}$$

This procedure produces stable results.

Step 2: Estimation of the volatility
So far we have estimated the parameters b_1, b_2, σ_e for European interest rates and parameters a_1, a_2, a_3 from the drift of the domestic interest rate. Substituting all these parameters into the objective function F_d it remains a function of one parameter σ_d and it is easy to find its optimal value.

Step 3: Final modification of the parameters
In the first two steps we have sequentially estimated all the domestic parameters. However, this does not guarantee that we have achieved the global minimum of the objective function. Hence, we try to improve them by optimizing the function F_d with respect to all of them together. The current estimated values (which are expected to be nearly optimal) were taken as starting values and the optimization was performed one more time with respect to all parameters.

3.5.4 Simulation analysis

We have implemented a numerical experiment in which we generated 1000 sets of domestic and European short rates and yield curves. We have used the same parameters in risk-neutral measure as in Sections 3.3 and 3.5.1. The initial values were generated from uniform distribution on the interval [0.02, 0.04] for the domestic and from the interval [0.005, 0.025] for the European short rate.

Our aim was to check the accuracy of the proposed estimation algorithm, as well as to see the usefulness of the step 3 described above since it requires much more time that the previous steps.

Based on the results, we decide to stop the estimation after the second step. The estimation of the drift is very precise. Less precision is achieved at estimating volatility, but it is still satisfactory. Table 3.2 shows what does this precision mean for the estimated yield curves. We again recall that the market Euribor rates are quoted with three decimal places. We conclude that using our algorithm the yield curves are estimated with a high precision. The detailed descriptive statistics of the European estimates and of the domestic estimates after Step 2 and after Step 3 can be found in [13].

6 months	differences in domestic yields after 2nd step in [%]	differences in domestic yields final estimates in [%]
minimum	1.19E-08	3.83E-10
maximum	4.17E-06	2.15E-06
median	4.74E-07	1.15E-07
mean value	7.01E-07	2.58E-07
standard deviation	6.85E-07	3.85E-07

12 months	differences in domestic yields after 2nd step in [%]	differences in domestic yields final estimates in [%]
minimum	4.87E-09	1.54E-11
maximum	9.78E-06	1.67E-06
median	1.06E-06	6.56E-08
mean value	1.66E-06	1.70E-07
standard deviation	1.70E-06	2.58E-07

Table 3.2: Descriptive statistics for domestic estimates after 2nd step and final estimates for maturities 6 and 12 months.

3.6 Generalization for CKLS model with zero correlation and the known γ_e, γ_d

The generalization for CKLS model with zero correlation and the known γ_e, γ_d is straightforward. First we estimate the European model parameters. For given value of γ_e the estimation of the other three parameters can be reduced to a one-dimensional problem, as it was defined in 3.5.2. Outcome of this optimization is an estimation of parameters b_1, b_2, σ_e.

Secondly we estimate domestic parameters. The estimation of risk-neutral drift remains the same as in Section 3.5.3, because it is the estimate of the Vasicek convergence model. The estimation of volatility is realized as the minimization of the objective function over the parameter σ_d in the same way as in section 3.5.3. The only change is the calculation of the objective function where instead of $\gamma_d = \frac{1}{2}$ we consider another γ_d. Based on the simulation results for the CIR model we omit final four-dimensional parameter optimization with respect to the parameters a_1, a_2, a_3, σ_d.

3.7 Estimation of correlation ρ a parameters γ_e, γ_d

To estimate the power γ_e we can use the procedure from [51]. The estimation described above is performed over a range of γ_e. Based on the objective function, the optimal γ_e is chosen. However, trying this approach to estimate γ_d and ρ did not work. We tried to find an explanation why these strategies fail in [13], where a more detailed treatment can be found.

The approximation error and the dependence on the correlation ρ and the power γ_d is numerically about the same order. Probably, it is the reason why we can not distinguish them and therefore we can not determine their correct values. However, the errors of the consecutive steps accumulate. We refer the reader to [48] for more numerical and analytical results on this question.

3.8 Calibration of the model using real market data

We have used Bribor (Bratislava Interbank Offered Rate) and Euribor (Euro Interbank Offered Rate) data from the last three months before the Slovak Republic entered the monetary union (1. 10. 2008 - 31. 12. 2008, i.e. $n_e = n_d = 62$). As the domestic short rate we use overnight Bribor, as the European short rate we use Eonia (Euro Over Night Index Average). The yields are considered for the same set of maturities in both domestic and European case. We take $\tau_e = \tau_d = (\frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \frac{6}{12}, \frac{9}{12}, \frac{12}{12})$, i.e. $m_e = m_d = 6$.

Using these data we have estimated the convergence CIR model (3.9) with zero correlation, i.e. $\rho = 0$. The estimates of the parameters are summarized in Table 3.3.

parameter	b_1	b_2	σ_e	a_1	a_2	a_3	σ_d
estimated value	0.0227	0.5000	1.1427	0.0879	-8.2052	7.3827	5.0000

Table 3.3: Estimated parameters of the CIR convergence model (3.9) with zero correlation.

We compare the exact and the estimated yield curves for several selected days. In the Figure 3.3 we show European (left) and domestic (right) yield curves for 1st, 31st and 61st day.

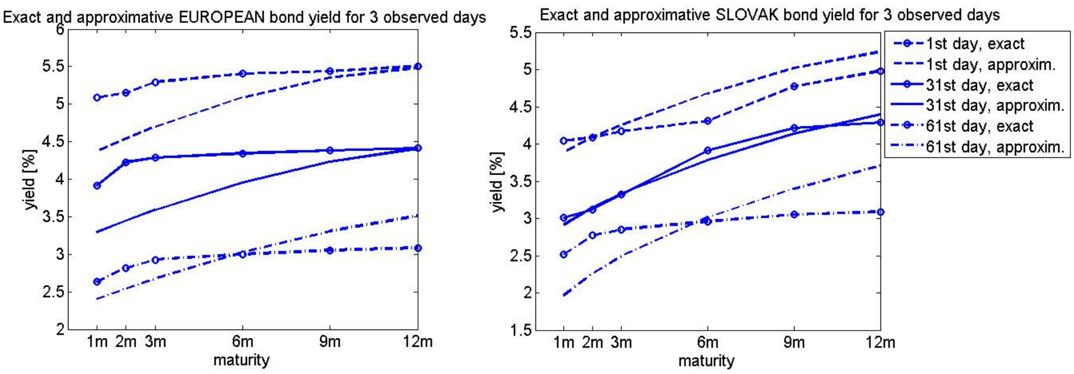


Figure 3.3: Estimated and real term structures for three observed days - European (left), Slovak (right).

In the Table 3.4 we numerically compare exact and estimated yields for one typical day. We observe much lower accuracy for the European rates. Hence an important task is to

Mat. [year]	**Market yield**	**Estim. yield**	**Diff.**	**Rel. diff.**
$\frac{1}{12}$	3.9140	3.2953	0.6187	0.158
$\frac{2}{12}$	4.2260	3.4461	0.7799	0.185
$\frac{3}{12}$	4.2860	3.5891	0.6969	0.163
$\frac{6}{12}$	4.3450	3.9604	0.3846	0.089
$\frac{9}{12}$	4.3810	4.2327	0.1483	0.034
$\frac{12}{12}$	4.4120	4.4065	0.0055	0.001

Mat. [year]	**Market yield**	**Estim. yield**	**Diff.**	**Rel. diff.**
$\frac{1}{12}$	3.0100	2.9174	0.0926	0.031
$\frac{2}{12}$	3.1200	3.1490	-0.0290	0.009
$\frac{3}{12}$	3.3200	3.3336	-0.0136	0.004
$\frac{6}{12}$	3.9100	3.7805	0.1295	0.033
$\frac{9}{12}$	4.2100	4.1358	0.0742	0.018
$\frac{12}{12}$	4.2900	4.3957	-0.1057	0.025

Table 3.4: Accuracy of the estimation of European (left) and domestic (right) yield curves.

improve the estimation of the European data, since its results are used in finding an estimate of domestic parameters. If we had chosen a different model for European interest rates, it might have also improved the estimation of domestic yield curves.

Our next aim is to propose an alternative model for the estimation of European interest rates and thus improve not only the fit of the European data but also the domestic ones, because estimated European parameters enter the estimation of domestic parameters as already known constants. Suggestions for improvement of modeling of the European data are described in the Chapter 4 and in the Chapter 5. In the Chapter 4 we suggest the approach, where the beginning of the yield curve can be estimated from Vasicek model with other model parameters, instead of taking it as a known value. In the Chapter 5 we suggest approach which can be used for modeling of the European data. Instead of considering the one-factor CKLS model for European interest rate, we can extend it to two-factor short rate model, where the sum of two modeled factor represent our European interest rate.

Chapter 4

Estimating the short rate from the term structures in the Vasicek model

In short rate models, bond prices and term structures of interest rates are determined by the parameters of the model and the current level of the instantaneous interest rate (so called short rate). The instantaneous interest rate can be approximated by the market overnight which - however - can be influenced by speculations on the market. The aim of this chapter is to propose a calibration method where we consider the short rate to be a variable unobservable on the market and estimate it together with the model parameters for the case of Vasicek model.

This Chapter is based on the following paper:

- *J. Halgašová, B. Stehlíková, Z. Bučková (Zíková): Estimating the short rate from the term structures in the Vasicek model, Tatra Mountains Mathematical Publications 61: 87-103, 2014*

Our aim is to use observable market term structures to calibrate the model, i.e. infer the values of the parameters using a certain criterion. One approach to calibration of the short rate models is based on minimizing the errors of the theoretical yields compared to the yields observed on the market. This approach was used for example in [49], [48]. Let us denote by R_{ij} the yield observed on the i-th day for j-th maturity and by $R(\tau_j, r_i)$ the yields computed using the Vasicek formula with j-th maturity τ_j and the short rate r_i realized on the i-th day. Using the weighted mean square error (the weight given to the i-th day and j-th maturity is w_{ij}), we minimize the function

$$F = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \left(R(\tau_j, r_i) - R_{ij} \right)^2, \tag{4.1}$$

where n is number of days and m is number of maturities which are observed on each of the days.

Recall that to compute the Vasicek yields, the value of the short rate is necessary. However, the short rate, defined as the beginning of the term structure of interest rates, is only a theoretical variable, not observed on the market. In practice, it can be approximated by a yield with short maturity, such as overnight in [49], [48], [14] or 1-month yields in [16], [27], etc. Using 1-month (or some other) yields is, however, not consistent with the interpretation of the short rate as limit of the yields, as maturity approaches zero. Note that in the papers [16], [27] this problem did not arise, since they considered only one time serie as an approximation of the short rate, not the whole term structure. In [49], [48], [14], when dealing with term structures, overnight was taken to approximate the short rate. However, even using the overnight which is closest to the short rate regarding the time, is questionable. The overnight rate, observed on the market, can be influenced by speculations. Hence we consider the short rate as an unobservable variable and estimate it from the term structures together with the parameters of the model.

The Chapter is organized as follows: In the following Section we present the procedure for calibrating model parameters and the evolution of the short rate. In the Section 4.2 we simulate data and test the proposed procedure. Finally, in the Section 4.3, we apply it to real market data. We end the chapter with some concluding remarks.

4.1 Calibration procedure

According to the considerations in 2.3, the objective function (4.1) will be minimized with respect to the model parameters α, β, σ^2, as well as the time series of the short rate $\mathbf{r} = (r_1, \ldots, r_n)'$.

The key observation is noting that the logarithm of the bond price in the Vasicek model (2.2) is a linear function of the parameters α and σ^2 and the short rate r:

$$\ln P(\tau, r) = c_0(\tau) r + c_1(\tau)\alpha + c_2(\tau)\sigma^2,$$

where

$$c_0 = \frac{1-e^{\beta\tau}}{\beta}, \quad c_1 = \frac{1}{\beta}\left[\frac{1-e^{\beta\tau}}{\beta} + \tau\right], \quad c_2 = \frac{1}{2\beta^2}\left[\frac{1-e^{\beta\tau}}{\beta} + \tau + \frac{(1-e^{\beta\tau})^2}{2\beta}\right].$$

Hence the objective function (4.1)

$$\begin{aligned} F(\alpha, \beta, \sigma^2, \mathbf{r}) &= \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left(R(\tau_j, r_i) - R_{ij}\right)^2 \\ &= \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} \frac{w_{ij}}{\tau_j^2}\left(\ln P(\tau_j, r_i) + \tau_j R_{ij}\right)^2 \end{aligned} \quad (4.2)$$

is quadratic in α, σ^2 and the components of $\mathbf{r}$. The optimal values for the given value of β are then easily obtained from the first order conditions which form a system of $n+2$ linear equations:

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix},$$

where

$$\boldsymbol{A} = \begin{bmatrix} \sum_{i,j} \frac{w_{i,j}}{\tau_j^2} c_1^2 & \sum_{i,j} \frac{w_{i,j}}{\tau_j^2} c_1 c_2 \\ \sum_{i,j} \frac{w_{i,j}}{\tau_j^2} c_1 c_2 & \sum_{i,j} \frac{w_{i,j}}{\tau_j^2} c_2^2 \end{bmatrix},$$

$$B = C' = \begin{bmatrix} \sum_j \frac{w_{1,j}}{\tau_j^2} c_1 c_0 & \sum_j \frac{w_{2,j}}{\tau_j^2} c_1 c_0 & \cdots & \sum_j \frac{w_{n,j}}{\tau_j^2} c_1 c_0 \\ \sum_j \frac{w_{1,j}}{\tau_j^2} c_2 c_0 & \sum_j \frac{w_{2,j}}{\tau_j^2} c_2 c_0 & \cdots & \sum_j \frac{w_{n,j}}{\tau_j^2} c_2 c_0 \end{bmatrix},$$

$$D = \begin{bmatrix} \sum_j \frac{w_{1,j}}{\tau_j^2} c_0^2 & 0 & \cdots & 0 \\ 0 & \sum_j \frac{w_{2,j}}{\tau_j^2} c_0^2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_j \frac{w_{n,j}}{\tau_j^2} c_0^2 \end{bmatrix},$$

$$x' = \left[\alpha, \sigma^2\right], \qquad y' = [r_1, r_2, \cdots, r_n], \qquad u' = \left[-\sum_{i,j} \frac{w_{i,j}}{\tau_j} R_{i,j} c_1, \quad -\sum_{i,j} \frac{w_{i,j}}{\tau_j} R_{i,j} c_2\right],$$

$$v' = \left[-\sum_j \frac{w_{1,j}}{\tau_j} R_{1,j} c_0, \quad -\sum_j \frac{w_{2,j}}{\tau_j} R_{2,j} c_0, \quad \cdots, \quad -\sum_j \frac{w_{n,j}}{\tau_j} R_{n,j} c_0\right].$$

Because of the special structure of the linear system, it is possible to reduce its dimensionality. The block D is diagonal and hence it is easy to find its inverse. Consequently, we are able to express the vector y in the following way:

$$Cx + Dy = v \qquad \Rightarrow \qquad y = D^{-1}(v - Cx).$$

From the equation $Ax + By = u$ we then obtain

$$(A - BD^{-1}C)x = u - BD^{-1}v$$

which is a system of two linear equations.

In this way we are able to find the optimal values of the parameters α and σ, and the short rate vector $\mathbf{r}$ for the given value of β. Finding the optimal β is then a one-dimensional optimization problem.

4.2 Application to simulated data

In the previous Section we have proposed a calibration procedure which estimates model parameters $\alpha, \sigma^2, \mathbf{r}$ using closed formulas from the Section 2 based on the first order conditions for minimizing the quadratic function for given parameter β. Given the optimal parameters $\alpha, \sigma^2, \mathbf{r}$ for each β, it is easy to find the optimal value of the parameter β, since it is a one dimensional optimization problem. The estimate is robust. There are no numerical problems in the calibration procedure.

The accuracy of the estimation, when tested on simulated data, is very good and there seems to be no numerical problems. We show one illustrative example here.

Using the real measure parameters from the introduction (i.e. $\kappa = 5.00, \theta = 0.02, \sigma = 0.02, \lambda = -0.5$) we simulate the time serie of the daily short rate values for 252 days (i.e. one year) and for each day we compute the yield curves with 12 maturities: 1 month, 2 months, …, 12 months. We use these yields as the input for the proposed calibration procedure. Following [48] and [49], we use the weights equal to the square of the corresponding maturity, i.e. $w_{ij} = \tau_j^2$.

Our values of real measure parameters and the market price of risk imply the following risk neutral parameters: $\alpha = 0.11$, $\beta = -5.00$, $\sigma = 0.02$. Recall that the calibration reduces to one-dimensional optimization, where the optimal value of β is found. Figure 4.1 shows the dependence of the objective function on β using a simulated set of data described above. Finding the optimal β and corresponding values of α and σ, we obtain the following estimates of the parameters: $\alpha = 0.1099999999979$, $\beta = -5.000000000000018$, $\sigma = 0.01999999943821$. As we can see, the parameters are almost exactly estimated. Also real and estimated short rates almost coincide; Figure 4.2 shows their difference which is of the order 10^{-16}. Figure 4.3 shows some of the fitted term structures.

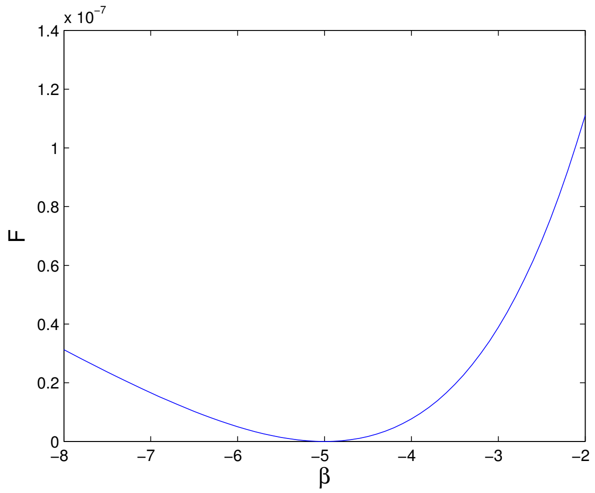


Figure 4.1: Dependence of the objective function F on parameter β using simulated data.

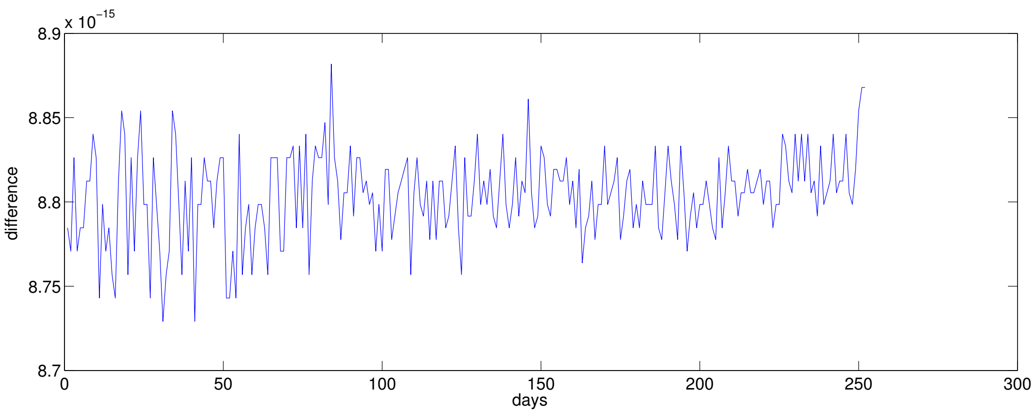


Figure 4.2: Difference between real and estimated short rate using simulated data.

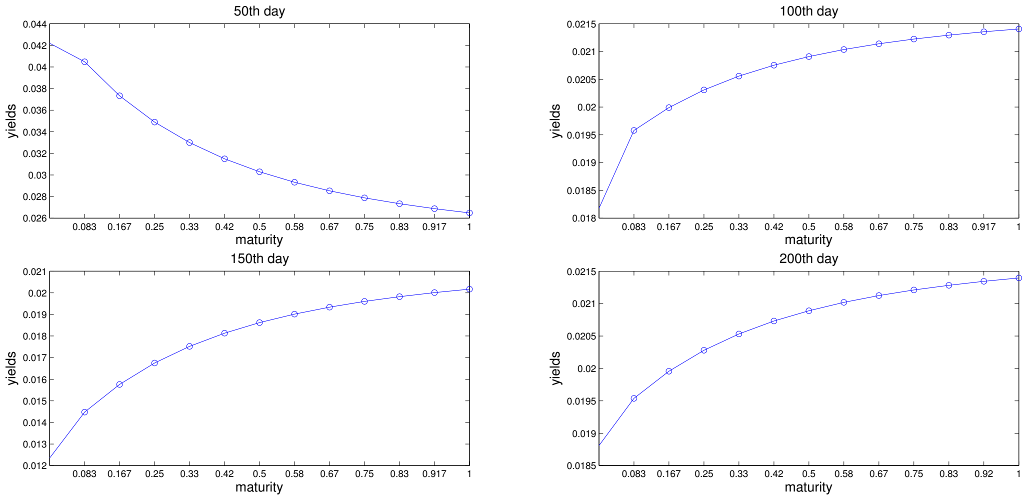


Figure 4.3: Examples of fitted term structures using simulated data.

4.3 Application to real market data

In this section we address the following two questions:

- How is the estimated short rate related to the market overnight? Can the short rate be approximated by market overnight or is it necessary to treat it as an unobservable factor which needs to be estimated?
- Is the estimated short rate robust to changing the maturities of the interest rates used for calibration?

4.3.1 Comparison between estimated short rate and overnight

One of the motivations for estimating the short rate from the market (observable) data, are the results from the paper [14], where we considered the convergence model for the Slovak interest rates before adoption of Euro currency in 2009. The first step, when building the convergence model, is specifying the one-factor model for the European rates. We have used Euribor[1] term structures and Eonia[2] as the approximation of the European short rate when calibrating the model. However, this leads to a poor fit of the term structure. The difference between the short rate as estimated from the term structures and the market overnight would explain the observed bad quality of the fit. Therefore, we use the proposed methodology for the Euribor rates in 2008. With a similar motivation in mind (Estonia adopted Euro in 2011), we do the same for the Euribor rates in 2010. For

[1]Euribor - European Interbank Offered Rate - is the rate at which euro interbank term deposits are offered by one prime bank to another prime bank; source: *http://www.euribor-ebf.eu/*

[2]Eonia - Euro OverNight Index Average - is the effective overnight reference rate for the euro and is computed from overnight unsecured lending transactions undertaken in the interbank market, source: *http://www.euribor-ebf.eu/*

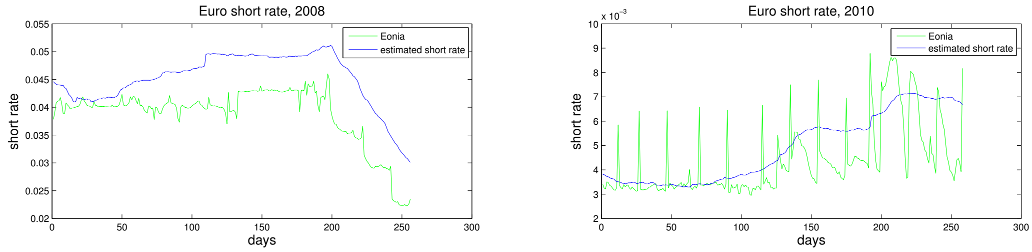


Figure 4.4: Comparison of estimated short rate and overnight for Euribor: 2008 (left), 2010 (right).

a comparison, we use also Estonian interest rates (Talibor[3]) from the same time periods. These data sets are described in Table 4.1.

data set	frequency	maturities
Euribor	daily	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 months
Talibor	daily	1, 2, 3, 4, 5, 6 months

Table 4.1: Data sets for comparing the estimated short rate with market overnight. Every data set is considered separately for the years 2008 and 2010.

The results are presented in Figure 4.4 (Euribor) and Figure 4.5 (Talibor). We see that although the estimate of the short rate for Euribor in 2008 has a similar behavior as the market overnight, it is higher and has a smaller volatility. The latter feature is especially pronounced in 2010, when the levels are approximately the same, but they are very different regarding the volatility. In the case of Talibor in 2008, there seems to be a difference between the estimated short rate and the market overnight which does not vary much in time, while their volatility is similar.

We present also the fitted term structures from 2010: Figure 4.6 and Table 4.2 show Euribor term structures; Figure 4.7 and Table 4.3 show Talibor term structures. In Figures 4.6 and 4.7 we can observe good fit of term structures compared to Figure 4 in [14] where the short rate was identified with the market values of the overnight rates. This observation is confirmed also by Tables 4.6 and 4.7 (differences between exact and estimated yields are 10^{-4}-10^{-5}) in contrast with Table 4 in [14] (differences are about 10^{-1}). To sum it up, we have achieved much higher estimation accuracy using estimated short rate values in our models.

[3]Talibor - Tallinn Interbank Offered Rate - was based on the interest rates at which banks offered to lend unsecured funds to other banks in the Estonian wholesale money market or interbank market in Estonian croons, source: *http://www.eestipank.ee/*

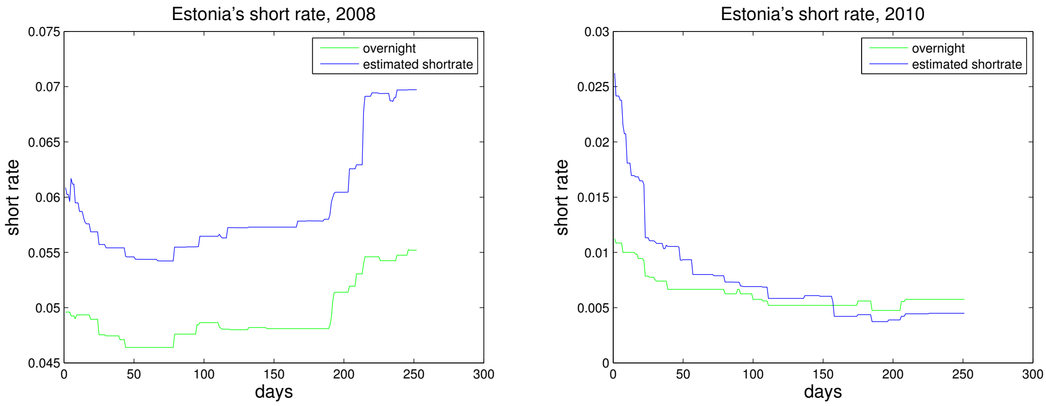


Figure 4.5: Comparison of estimated short rate and overnight for Talibor: 2008 (left), 2010 (right).

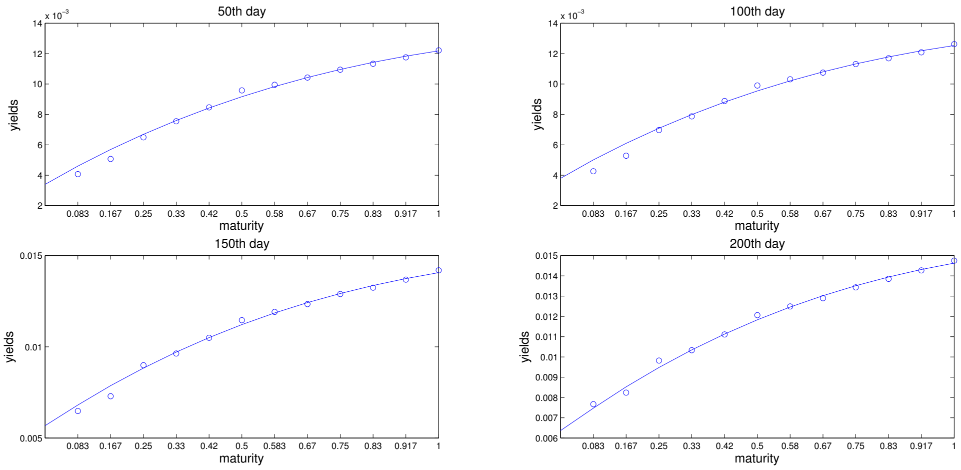


Figure 4.6: Accuracy of estimated yield curves. Euribor 2010.

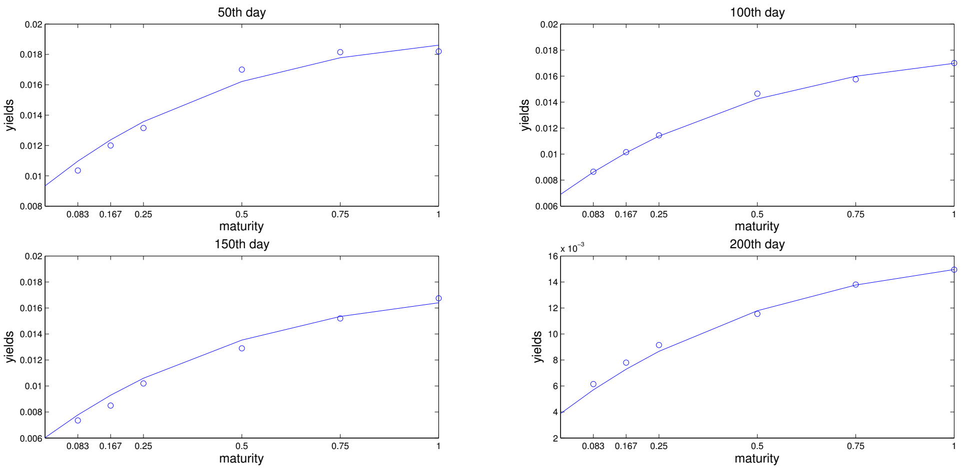


Figure 4.7: Accuracy of estimated yield curves. Talibor 2010.

Maturity [years]	50th day	100th day	150th day	200th day
0.083	5.30E-04	7.49E-04	3.29E-04	1.96E-04
0.167	6.22E-04	8.16E-04	5.79E-04	2.83E-04
0.25	2.00E-04	1.17E-04	1.56E-04	3.41E-04
0.33	4.73E-05	1.19E-04	8.04E-05	1.55E-05
0.42	3.98E-05	7.38E-05	1.23E-05	1.81E-05
0.5	4.17E-04	3.57E-04	2.45E-04	2.28E-04
0.58	1.19E-04	1.04E-04	5.73E-05	2.96E-05
0.67	7.25E-06	4.67E-05	8.01E-05	1.19E-04
0.75	1.68E-05	1.10E-05	3.10E-05	8.14E-05
0.83	9.34E-05	9.23E-05	1.20E-04	9.14E-05
0.917	8.05E-05	1.04E-04	5.92E-05	4.28E-05
1	2.82E-05	8.94E-05	1.27E-04	1.21E-04

Table 4.2: Accuracy of estimated yield curves - absolute values of differences between the real and estimated rates. Euribor 2010.

Maturity [years]	50th day	100th day	150th day	200th day
0.083	6.14E-04	3.12E-05	4.27E-04	4.43E-04
0.167	3.66E-04	4.73E-05	7.90E-04	5.08E-04
0.25	4.20E-04	6.57E-05	4.00E-04	4.80E-04
0.5	7.85E-04	4.09E-04	6.32E-04	2.40E-04
0.75	3.72E-04	2.39E-04	1.47E-04	3.30E-05
1	4.12E-04	1.61E-05	3.51E-04	1.21E-05

Table 4.3: Accuracy of estimated yield curves - absolute values of differences between the real and estimated rates. Talibor 2010.

4.3.2 Estimated short rates using different sets of maturities

Canadian interest rates[4] are available for a wide range of maturities up to 30 years which allows us to test the robustness of the short rate estimates to the choice of maturities used in calibration. We used three sets of parameters: the first one includes equally spaced maturities up to 30 years, the second one consists of shorter maturities up to 2 years and the third one goes up to 10 years. Details are given in Table 4.4.

We estimate the model separately for each of the years and in Figure 4.8 we record the different estimates of short rate depending on the input data.

Having in mind the high precision of the method on the simulated data (although we have presented a simulation example only with maturities from 1 month to 12 months, the procedure is very precise also for other choices of maturities), we would expect to

[4] yield curves for zero-coupon bonds, generated using pricing data for Government of Canada bonds and treasury bills, source: *http://www.bankofcanada.ca*

data set	frequency	maturities
Canada 1	daily	0.25, 2.5, 5, 7.5, 12.5, 10, 15, 17.5, 20, 22.5, 25, 27.5, 30 years
Canada 2	daily	0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2 years
Canada 3	daily	1, 2, 3, 4, 5, 6, 7, 8, 9, 10 years

Table 4.4: Data sets for comparing the estimated short rate using different maturities. Every data set is considered separately for the years 2007, 2008, 2009, 2010 and 2011.

obtain almost identical estimates of the short rate behavior. Hence the differences, such as those observed in Figure 4.8 would suggest the inadequacy of the Vasicek model.

On the other hand, we are also interested in the impact on accuracy of estimation of yield curves. We present the results from the year 2011 and for each data set we compare real and estimated yield curves for selected days in Figures 4.9, 4.10, 4.11 and Tables 4.5, 4.6, 4.7. In general, the fit can be considered to be good. Note that on the 150th day for data set 2 we observe a term structure shape (firstly decreasing and then increasing) that is not possible to obtain in Vasicek model which allows only monotone and humped (firstly increasing and then decreasing) term structures (c.f.[58]). These shapes are estimated well; recall that in the construction of the objective function we have put more weight to estimate the longer maturities.

Maturity [years]	50th day	100th day	150th day	200th day
0.25	7.90E-03	4.53E-03	8.12E-05	1.39E-03
2.5	4.38E-03	1.51E-03	3.87E-03	8.14E-04
5	2.13E-03	3.83E-04	3.76E-03	1.12E-03
7.5	1.12E-03	3.95E-04	1.66E-03	6.38E-04
12.5	4.09E-04	7.01E-04	3.89E-04	3.10E-04
10	9.43E-05	3.63E-04	1.97E-04	2.20E-04
15	2.58E-04	1.89E-04	2.41E-04	1.61E-04
17.5	1.19E-04	5.07E-04	1.36E-04	2.04E-05
20	9.45E-05	4.63E-04	3.47E-05	1.49E-04
22.5	1.51E-04	1.92E-04	1.15E-04	2.16E-04
25	9.22E-06	8.63E-05	9.36E-05	1.18E-04
27.5	2.10E-04	2.18E-04	9.77E-05	4.65E-05
30	8.09E-05	1.74E-04	2.98E-04	4.42E-06

Table 4.5: Accuracy of estimated yield curves - absolute values of differences between the real and estimated rates. Canada 2011, set 1.

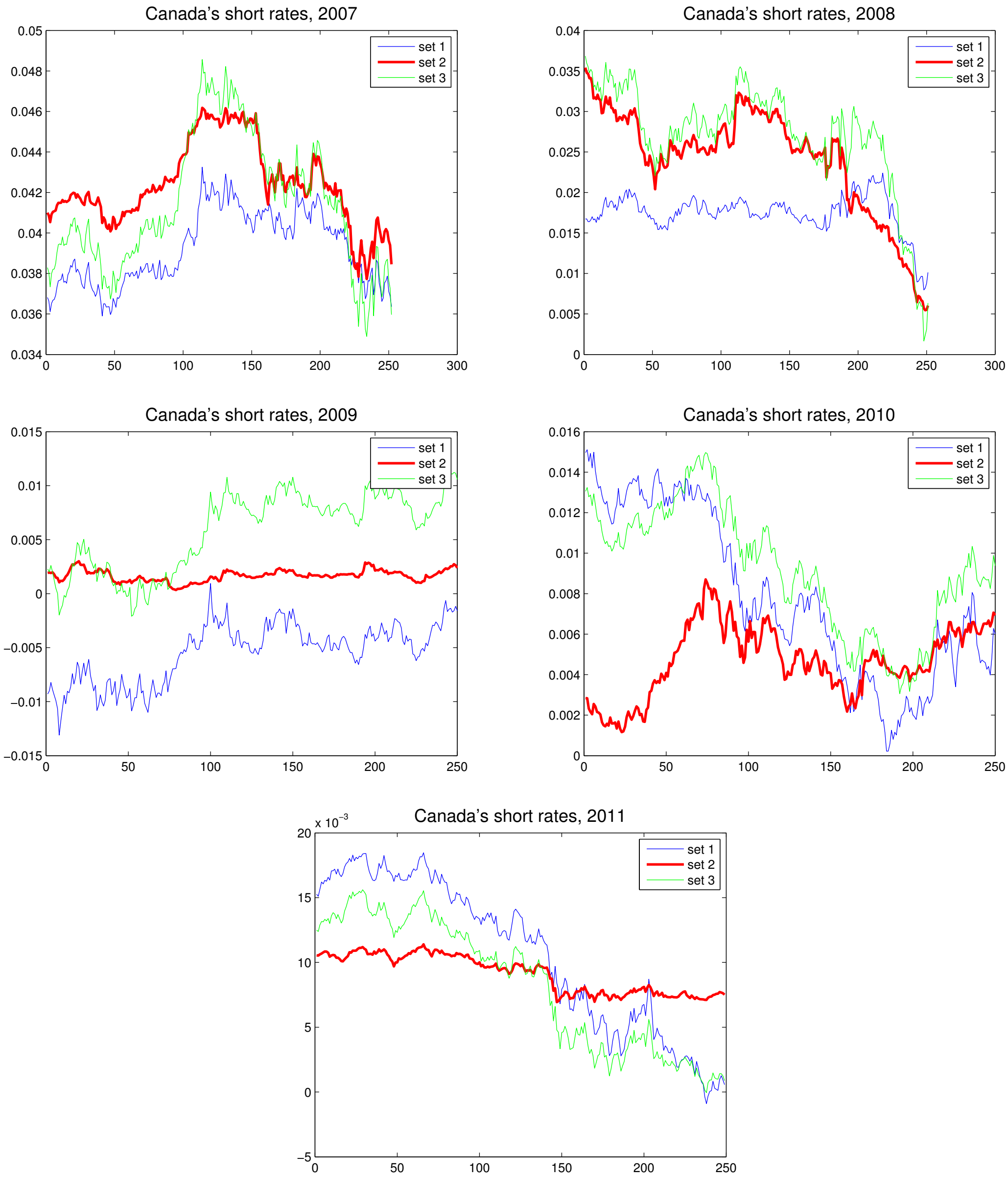


Figure 4.8: Estimated short rates for Canada, estimated separately for each of the years from 2007 to 2011.

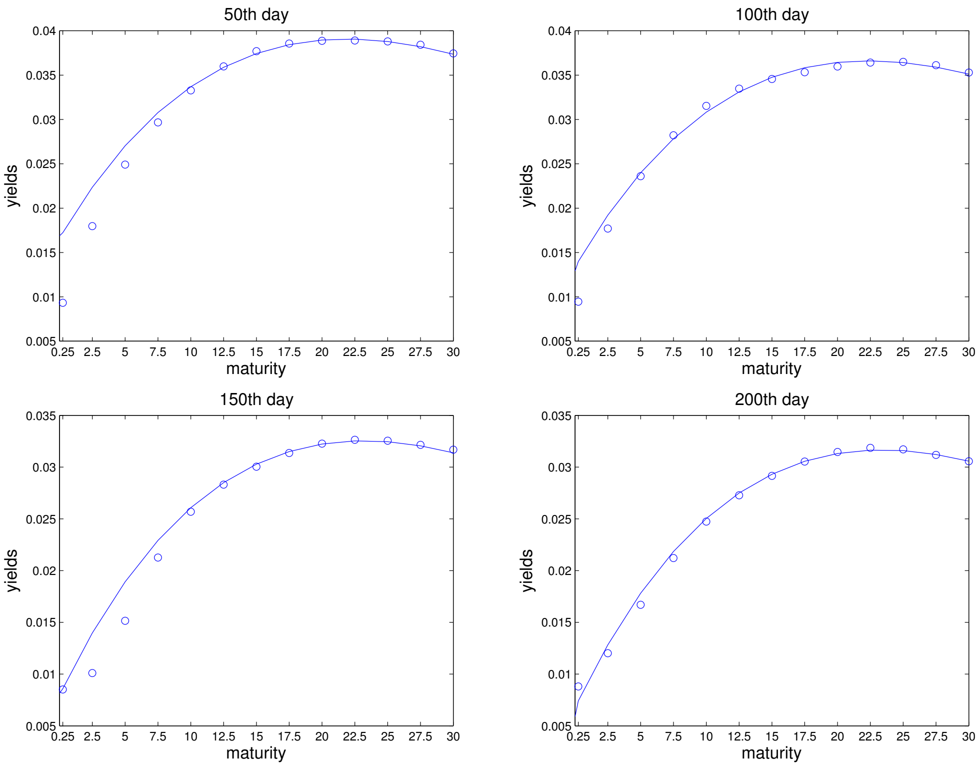


Figure 4.9: Accuracy of estimated yield curves. Canada 2011, set 1.

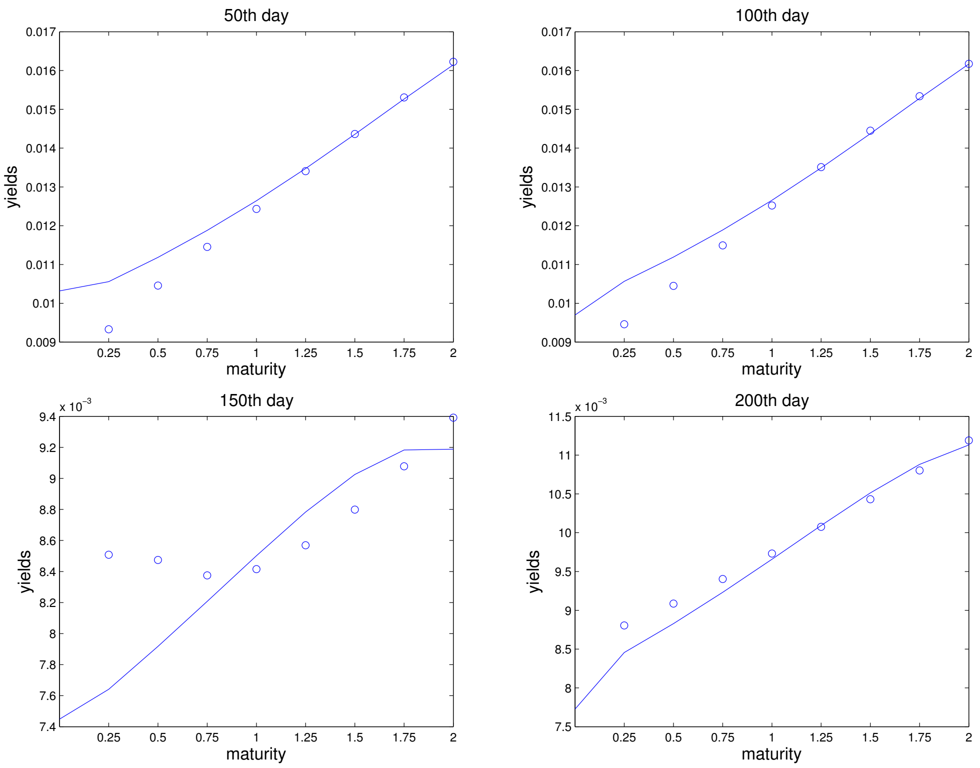


Figure 4.10: Accuracy of estimated yield curves. Canada 2011, set 2.

Maturity [years]	50th day	100th day	150th day	200th day
0.25	1.23E-03	1.10E-03	8.67E-04	3.50E-04
0.5	7.26E-04	7.42E-04	5.57E-04	2.58E-04
0.75	4.24E-04	3.96E-04	1.68E-04	1.71E-04
1	2.10E-04	1.35E-04	8.67E-05	7.44E-05
1.25	6.92E-05	2.19E-05	2.14E-04	1.74E-05
1.5	8.48E-06	7.66E-05	2.26E-04	8.19E-05
1.75	4.30E-05	5.49E-05	1.05E-04	7.80E-05
2	7.77E-05	2.98E-06	2.03E-04	5.84E-05

Table 4.6: Accuracy of estimated yield curves - absolute values of differences between the real and estimated rates. Canada 2011, set 2.

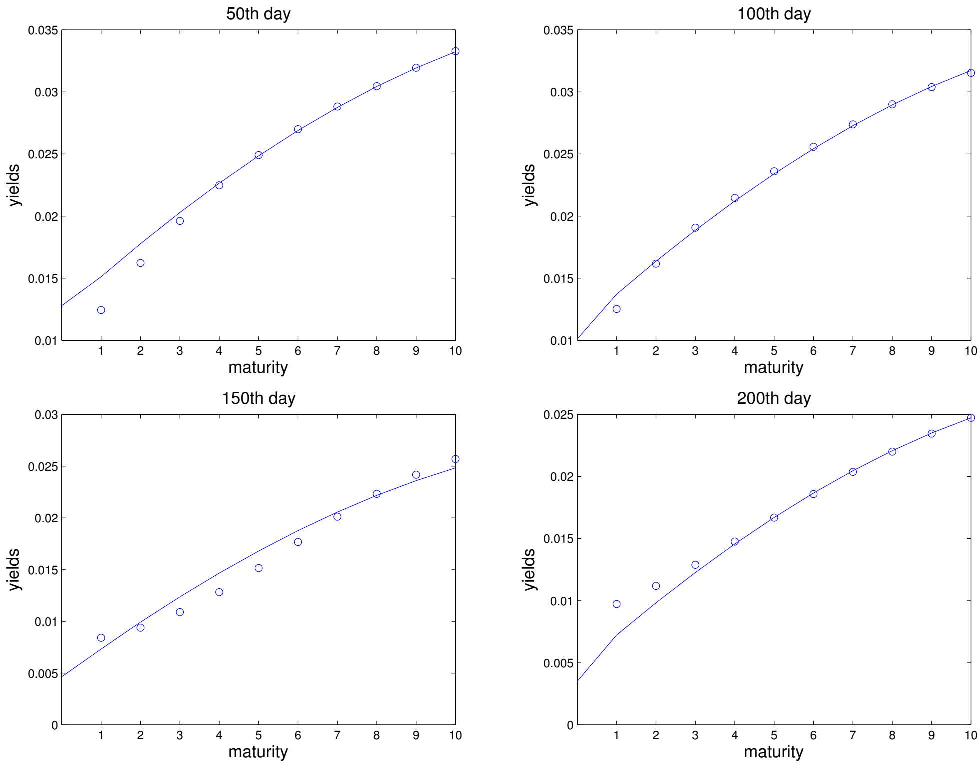


Figure 4.11: Accuracy of estimated yield curves. Canada 2011, set 3.

Maturity [years]	**50th day**	**100th day**	**150th day**	**200th day**
1	2.68E-03	1.19E-03	1.09E-03	2.50E-03
2	1.55E-03	1.90E-04	5.30E-04	1.37E-03
3	6.81E-04	2.02E-04	1.47E-03	6.15E-04
4	1.67E-04	2.55E-04	1.84E-03	1.92E-04
5	5.87E-05	2.04E-04	1.65E-03	6.98E-06
6	1.05E-04	1.52E-04	1.10E-03	7.77E-05
7	7.23E-05	1.10E-04	4.43E-04	8.79E-05
8	2.83E-05	5.23E-05	1.43E-04	7.46E-05
9	1.44E-05	4.59E-05	5.74E-04	4.82E-05
10	5.13E-05	1.88E-04	8.53E-04	6.40E-07

Table 4.7: Accuracy of estimated yield curves - absolute values of differences between the real and estimated rates. Canada 2011, set 3.

Chapter 5

Short rate as a sum of two CKLS-type processes

We study the short rate model of interest rates, in which the short rate is defined as a sum of two stochastic factors. Each of these factors is modeled by a stochastic differential equation with a linear drift and the volatility proportional to a power of the factor. We show a calibration methods which - under the assumption of constant volatilities - allow us to estimate the term structure of interest rate as well as the unobserved short rate, although we are not able to recover all the parameters. We apply it to real data and show that it can provide a better fit compared to a one-factor model. A simple simulated example suggests that the method can also be applied to estimate the short rate even if the volatilities have a general form. Therefore we propose an analytical approximation formula for bond prices in such a model and derive the order of its accuracy.

The Chapter 5 is based on:

- *Z. Bučková, J. Halgašová, B. Stehlíková: Short rate as a sum of CKLS-type processes, accepted for publication in Proceedings of Numerical analysis and applications conference, Springer Verlag in LNCS, 2016.*

5.1 Model

In particular, we are concerned with a model where the short rate r is given by $r = r_1 + r_2$ and the risk neutral dynamics of the factors r_1 and r_2 is as follows:

$$\begin{aligned} dr_1 &= (\alpha_1 + \beta_1 r_1)dt + \sigma_1 r_1^{\gamma_1} dw_1, \\ dr_2 &= (\alpha_2 + \beta_2 r_2)dt + \sigma_2 r_2^{\gamma_2} dw_2, \end{aligned} \quad (5.1)$$

where the correlation between increments of Wiener processes is ρ, i.e., $\mathbb{E}(dw_1 dw_1) = \rho dt$. In particular we note that by taking $\gamma_1 > 0$ and $\gamma_2 = 0$ we are able to model negative interest rates (both instantaneous short rate and interest rates with other maturities) which were actually a reality recently in Eurozone (see historical data at *www.euribor.org*). This can also be accomplished by a simple one-factor Vasicek model. However, a consequence of Vasicek model is the same variance of short rate, regardless of its level. On the other hand, the real data suggest that volatilities of interest rates decrease as interest rates themselves decrease. The model with $\gamma_1 > 0$ and $\gamma_2 = 0$ has the variance dependent on the level of factor r_1.

Before using a certain model we need to calibrate it, i.e., estimate its parameters from the available data. One approach to calibration of interest rate models is based on minimizing the weighted squared differences between theoretical yields and the real market ones, see, e.g., [48], [49]. Let R_{ij} be the yield observed at i-th day for j-th maturity τ_j and $R(\tau_j, r_{1j}, r_{2j})$ the yield computed from the two factor model, where r_{1i} and r_{2i} are factors of the short rate at i-th day. We denote by w_{ij} the weight of the i-th day and j-th maturity observation in the objective function. In general, we look for the values of the parameters and the decomposition of the short rate to the factors, which minimize the objective function

$$F(r_{1i}, r_{2i}, \alpha_i, \beta_i, \gamma_i, \sigma_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \Big(R(\tau_j, r_{1i}, r_{2i}) - R_{ij} \Big)^2. \tag{5.2}$$

In order to solve this optimization problem, we need to evaluate the yields given by the model which is equivalent to solving the PDE for bond prices $P(\tau, r_1, r_2)$, which reads as

$$\begin{aligned} & -\frac{\partial P}{\partial \tau} + [\alpha_1 + \beta_1 r_1] \frac{\partial P}{\partial r_1} + [\alpha_2 + \beta_2 r_2] \frac{\partial P}{\partial r_2} \\ & + \frac{\sigma_1^2 r_1^{2\gamma_1}}{2} \frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2 r_2^{2\gamma_2}}{2} \frac{\partial^2 P}{\partial r_2^2} + \rho \sigma_1 \sigma_2 r_1^{\gamma_1} r_2^{\gamma_2} \frac{\partial^2 P}{\partial r_1 \partial r_2} - (r_1 + r_2) P = 0 \end{aligned} \tag{5.3}$$

for any r_1, r_2 from their domain and any time to maturity $\tau \in [0, T)$, with initial condition $P(0, r_1, r_2) = 1$ for any r_1, r_2, see [34]. Closed form solutions are available only in special cases. For the model (5.1) , c.f. [7], it is only the Vasicek case $\gamma_1 = \gamma_2 = 0$ and the CIR case $\gamma_1 = \gamma_2 = 1/2$ but only with zero correlation $\rho = 0$ and a mixed model $\gamma_1 = 0$, $\gamma_2 = 1/2$ again with $\rho = 0$. In the remaining cases we need some approximation, which can be obtained using a certain numerical method, Monte Carlo simulation of an approximate analytical solution.

The paper is formulated as follows: In the following section we consider the uncorrelated case of the two-factor Vasicek model, i.e., the model (5.1) with $\gamma_1 = \gamma_2 = 0$ and $\rho = 0$, and the possibility to estimate its parameters and the short rate factors using the objective function (5.2). In the Section 5.3 we apply this algorithm to real data and we note its advantage in fitting the market interest rates, compared to one-factor Vasicek model. The Section 5.4 present a simulated example which shows a performance of this algorithm when estimating the short rate from a general model (5.1), i.e., a robustness to misspecified volatility. This motivates us to develop an analytical approximation formula for the bond prices for the model (5.1) and derive the order of its accuracy which we do in the Section 5.5. We end this chapter with concluding remarks.

5.2 Two-factor Vasicek model: singularity and transformation

In this section we consider the model (5.1) with $\gamma_1 = \gamma_2 = 0$, in which case the formulae for the bond prices are known, see for example [7]. Moreover we assume that $\rho = 0$, so the increments of the Wiener processed determining the factors of the short rate are uncorrelated. We write the bond price P as

$$\log P(\tau, r_1, r_2) = c_{01}(\tau)r_1 + c_{02}(\tau)r_2 + c_{11}(\tau)\alpha_1 + c_{12}(\tau)\alpha_2 + c_{21}(\tau)\sigma_1^2 + c_{22}(\tau)\sigma_2^2,$$

where, for $k = 1$ and $k = 2$,

$$c_{0k} = \frac{1-e^{\beta_k \tau}}{\beta_k}, c_{1k} = \frac{1}{\beta_k}\left(\frac{1-e^{\beta_k \tau}}{\beta_k} + \tau\right), c_{2k} = \frac{1}{2\beta_k^2}\left(\frac{1-e^{\beta_k \tau}}{\beta_k} + \tau + \frac{(1-e^{\beta_k \tau})^2}{2\beta_k}\right)$$

We fix the values of β_1 and β_2. Then the objective function (5.2) can be written as

$$\begin{aligned} F &= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{w_{ij}}{\tau_j^2}\left(\log P(\tau_j, r_{1i}, r_{2i}) + R_{ij}\tau_j\right)^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{w_{ij}}{\tau_j^2}\left(c_{01}(\tau_j)r_{1i} + c_{02}(\tau)r_{2i} + c_{11}(\tau_j)\alpha_1 + c_{12}(\tau_j)\alpha_2 + \right. \\ & \quad \left. c_{21}(\tau_j)\sigma_1^2 + c_{22}(\tau_j)\sigma_2^2 + R_{ij}\tau_j\right)^2, \end{aligned}$$

which can be represented as a weighted linear regression problem without intercept, with parameters $r_{1i}, r_{2i}, \alpha_1, \alpha_2, \sigma_1^2, \sigma_2^2$ to be estimated. However, the regressors are linearly dependent and hence the estimates minimizing the objective function are not uniquely determined. In the context of calibrating the yield curves, this means that different sets of parameter values and factor evolutions lead to the same optimal fit of the term structures. In particular, we have

$$-\frac{1}{\beta_2}c_{01}(\tau) + \frac{1}{\beta_2}c_{02}(\tau) + \frac{\beta_1}{\beta_2}c_{11}(\tau) = c_{12}(\tau).$$

Substituting this into the formula for the logarithm of the bond price we get

$$\begin{aligned} \log P(\tau, r_1, r_2) &= c_{01}(\tau)r_1 + c_{02}(\tau)r_2 + c_{11}(\tau)\alpha_1 + c_{12}(\tau)\alpha_2 + c_{21}(\tau)\sigma_1^2 + c_{22}(\tau)\sigma_2^2 \\ &= \left(r_{1i} - \frac{\alpha_2}{\beta_2}\right)c_{01}(\tau_j) + \left(r_{2i} + \frac{\alpha_2}{\beta_2}\right)c_{02}(\tau_j)\left(\alpha_1 + \frac{\alpha_2\beta_1}{\beta_2}\right)c_{11}(\tau_j) \\ & \quad + c_{21}(\tau_j)\sigma_1^2 + c_{22}(\tau_j)\sigma_2^2. \end{aligned}$$

The objective function of the regression problem then reads as

$$\begin{aligned} F &= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{w_{ij}}{\tau_j^2}\left(\left(r_{1i} - \frac{\alpha_2}{\beta_2}\right)c_{01}(\tau_j) + \left(r_{2i} + \frac{\alpha_2}{\beta_2}\right)c_{02}(\tau_j)\right. \\ & \quad \left. + \left(\alpha_1 + \frac{\alpha_2\beta_1}{\beta_2}\right)c_{11}(\tau_j) + c_{21}(\tau_j)\sigma_1^2 + c_{22}(\tau_j)\sigma_2^2 + R_{ij}\tau_j\right)^2, \end{aligned} \tag{5.4}$$

which is already regular. Note that we are not able to estimate all the parameters, nor the separate factors r_1 and r_2. However, the sum of the parameters corresponding to c_{01} and c_{02} is the sum of r_1 and r_2, i.e., the short rate r.

Thus, for a given pair (β_1, β_2) we find the optimal values of the regression problem above and note the attained value of the objective function. Then, we optimize for the values of β_1, β_2. For these optimal β_1, β_2 we note the coefficients corresponding to c_{01} and c_{02}. These are estimated shifted factors and their sum is the estimate of the short rate.

5.3 Application to real market data

We use this algorithm to the two data sets considered in paper [31] dealing with estimating the short rate using one-factor Vasicek model: Euribor data from last quarter of 2008 and last quarter of 2011. We note that in the first case, the fit of the one-factor Vasicek was much better then in the second case.

It can be expected that in the case when already a one-factor model provides a good fit, estimating a two-factor model does not yield much change into the results. However, if the fit of a one-factor model is not satisfactory, the estimates from the two-factor model can be more substantially different. From Figure 5.1 we can see that the fit of the term structures has significantly improved by adding the second factor in the last quarter of 2011.

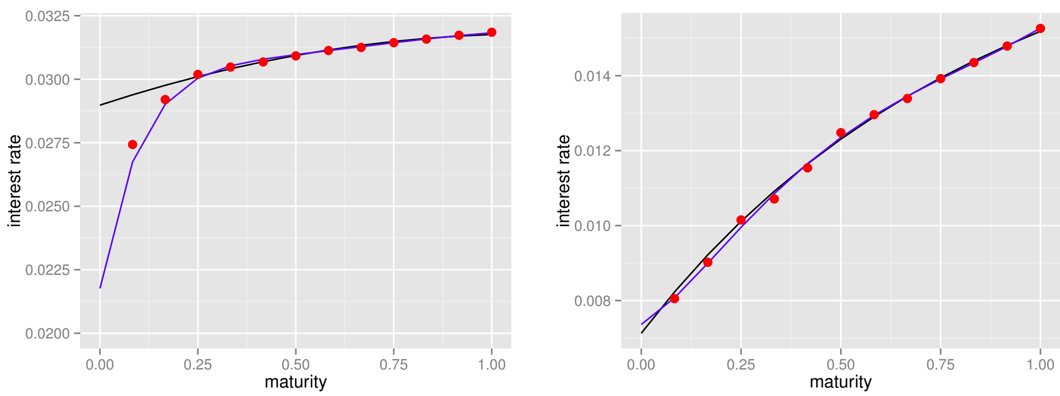


Figure 5.1: Fitted yield curves using real data - a selected day in 2008 (left) and 2010 (right): blue lines show the fit from the 2-factor model, black lines from the 1-factor model, red circles are market data

5.4 Robustness of the short rate estimates

Naturally, the algorithm described in the previous section works well in case of data simulated from the two-factor Vasicek model. However, we noted the estimate of the short rate is remarkable accurate even when the volatility is misspecified. In particular, since we are able to compute exact bond prices from the two-factor CIR model with uncorrelated factors and test the algorithm on these data.

We simulate two factor CIR model with the parameters taken from [17]: $\kappa_1 = 1.8341, \theta_1 = 0.05148, \sigma_1 = 0.1543,\ \kappa_2 = 0.005212, \theta_2 = 0.03083, \sigma_2 = 0.06689$. We simulate daily data from one quarter (assuming 252 trading days in a year). Then, we consider market prices of risk $\lambda)1 = -0.1253, \lambda_2 = -0.06650$ from [17] and compute the term structures for maturities $1, 2, , \ldots, 12$ months for each day using the exact formulae. These data are used as inputs to estimation of the two-factor Vasicek model. A sample result, comparing the simulated short rate and its estimate is presented in Figure 5.2.

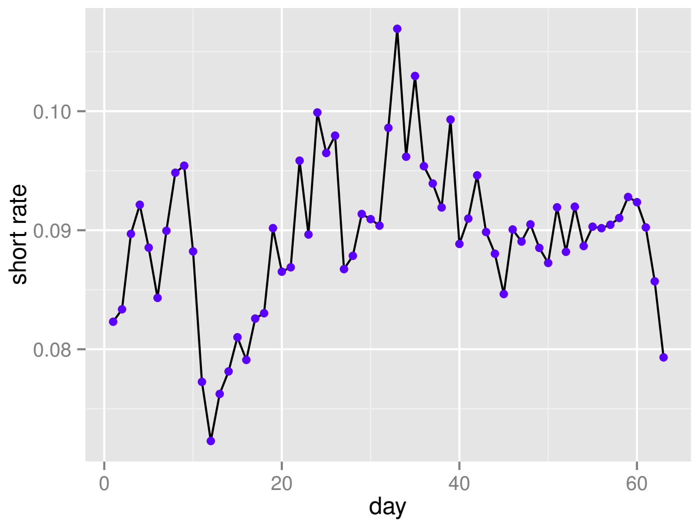


Figure 5.2: Estimating short rate using data simulated from the two-factor CIR model: simulated (points) and estimated (line) short rate.

In spite of misspecification of the model, the terms corresponding to $\left(r_{1i} - \frac{\alpha_2}{\beta_2}\right)$ and $\left(r_{2i} + \frac{\alpha_2}{\beta_2}\right)$ indeed estimate the factors up to a constant shift. This is displayed in Figure 5.3; note the vertical axis for each pair of the graphs.

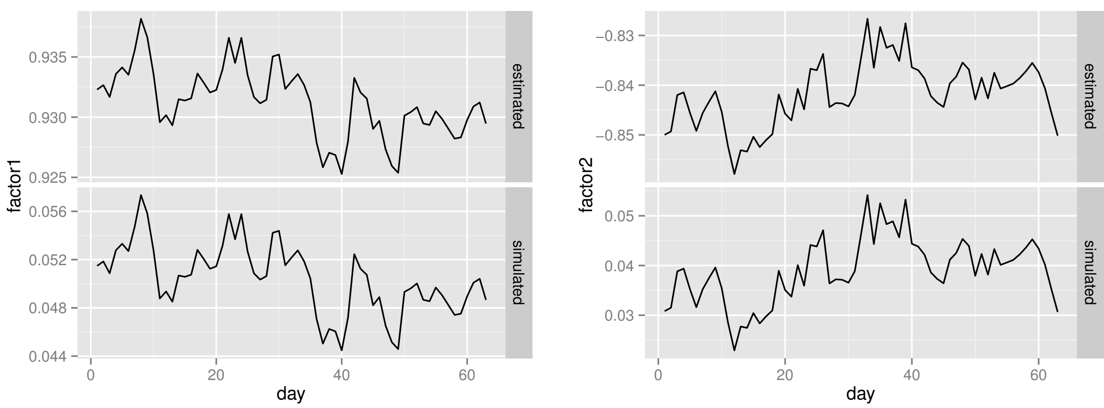


Figure 5.3: Estimating factors up to an additive constant using data simulated from the two-factor CIR model.

5.5 Approximation of the bond prices in the CKLS model

Based on the example in the previous section, we might want to estimate the short rate by application of the algorithms for the two-factor Vasicek model, even though we expect the volatility to have a more general form. Estimates of the short rate factors, up to an additive constant, might be a valuable results, since their knowledge greatly reduced the dimension of the optimization problem (5.2). However, we need to compute the bond prices in a CKLS general model - either their exact values or a sufficiently accurate approximation. Since they are going to be used in a calibration of a certain kind, they should be calculated quickly and without numerical problems. The aim of this section is to provide an analytical approximation formula for these bond prices and to derive order of its accuracy.

The motivation comes from the paper [51] where an approximation of bond prices for a one-factor CKLS model was proposed. Note that if the correlation in the two-factor CKLS model is zero, the bond price is equal to the sum of two terms corresponding to solutions to bond pricing PDE originating from one factor CKLS models, with factors r_1 and r_2 taking the role of a short rate. Therefore, the bond price could be approximated as a sum of the approximations corresponding to these one-factor models. They are obtained from the Vasicek bond price formula, by substituting its constant volatility by instantaneous volatility from the CKLS model. It is shown in [51] that the error of logarithm of the bond price is then $O(\tau^4)$ as $\tau \to 0^+$. We generalize this idea to the two-factor case and suggest the following approximation.

Theorem 5.1. *Let P^{ap} be the approximative and P^{ex} be the exact price of the bond in CKLS model.Then for $\tau \to 0^+$*

$$lnP^{ap}(\tau,r_1,r_2) - lnP^{ex}(\tau,r_1,r_2) = c_4(r_1,r_2)\tau^4 + o(\tau^4) \tag{5.5}$$

where the coefficient c_4 is given by

$$c_4(r_1,r_2) = -\frac{1}{24r_1^2r_2^2}\Big((2\gamma_1^2-\gamma_1)(r_1^{4\gamma_1}r_2^2\sigma_1^4) + (2\gamma_2^2-\gamma_2)(r_1^2r_2^{4\gamma_2}\sigma_2^4) \tag{5.6}$$

$$+\rho\gamma_1(\gamma_1-1)r_1^{3\gamma_1}r_2^{\gamma_2+2}\sigma_1^3\sigma_2 + \rho\gamma_2(\gamma_2-1)r_1^{\gamma_1+2}r_2^{3\gamma_2}\sigma_1\sigma_2^3 \tag{5.7}$$

$$+2\gamma_2(\alpha_2+\beta_2r_2)(\rho\sigma_1\sigma_2r_1^{2+\gamma_1}r_2^{1+\gamma_2} + \sigma_2^2r_1^2r_2^{1+2\gamma_2}) + 2\gamma_1\gamma_2\rho^2\sigma_1^2\sigma_2^2r_1^{2\gamma_1+1}r_2^{2\gamma_2+1} \tag{5.8}$$

$$+2\gamma_1r_1r_2^2\sigma_1(\alpha_1+\beta_1r_1)\big(r_1^{2\gamma_1}\sigma_1 + \rho\sigma_2r_1^{\gamma_1}r_2^{\gamma_2}\big)\Big). \tag{5.9}$$

Remark 5.2. *From the above considerations it follows that* $\log P^{ap} - \log P^{ex}$ *is $O(\tau^4)$ in the case of zero correlation ρ. What needs to be done is showing that the same order of accuracy is achieved also in the case of general ρ.*

Proof. Let us define function $f^{ex}(\tau,r_1,r_2) = \ln P^{ex}(\tau,r_1,r_2)$, where P^{ex} is the exact solution of the equation (5.3) Then the partial differential equation (5.3) for f^{ex} is given by:

$$
\begin{aligned}
&-\frac{\partial f^{ex}}{\partial \tau}+[\alpha_1+\beta_1 r_1]\frac{\partial f^{ex}}{\partial r_1}+[\alpha_2+\beta_2 r_2]\frac{\partial f^{ex}}{\partial r_2}\\
&+\frac{\sigma_1^2 r_1^{2\gamma_1}}{2}\left[\left(\frac{\partial f^{ex}}{\partial r_i}\right)^2+\frac{\partial^2 f^{ex}}{\partial r_i^2}\right]+\frac{\sigma_2^2 r_2^{2\gamma_2}}{2}\left[\left(\frac{\partial f^{ex}}{\partial r_i}\right)^2+\frac{\partial^2 f^{ex}}{\partial r_i^2}\right]\\
&+\rho\sigma_1\sigma_2 r_1^{\gamma_1} r_2^{\gamma_2}\left[\frac{\partial f^{ex}}{\partial r_1}\frac{\partial f^{ex}}{\partial r_2}+\frac{\partial^2 f^{ex}}{\partial r_1\partial r_2}\right]-(r_1+r_2)=0.
\end{aligned}
$$

For the approximation $f^{ap}(\tau,r_1,r_2)=\ln P^{ap}(\tau,r_1,r_2)$ we obtain from the former PDE with nontrivial right-hand side $h(\tau,r_1,r_2)$:

$$
\begin{aligned}
&-\frac{\partial f^{ap}}{\partial \tau}+[\alpha_1+\beta_1 r_1]\frac{\partial f^{ap}}{\partial r_1}+[\alpha_2+\beta_2 r_2]\frac{\partial f^{ap}}{\partial r_2}\\
&+\frac{\sigma_1^2 r_1^{2\gamma_1}}{2}\left[\left(\frac{\partial f^{ap}}{\partial r_i}\right)^2+\frac{\partial^2 f^{ap}}{\partial r_i^2}\right]+\frac{\sigma_2^2 r_2^{2\gamma_2}}{2}\left[\left(\frac{\partial f^{ap}}{\partial r_i}\right)^2+\frac{\partial^2 f^{ap}}{\partial r_i^2}\right]\\
&+\rho\sigma_1\sigma_2 r_1^{\gamma_1} r_2^{\gamma_2}\left[\frac{\partial f^{ap}}{\partial r_1}\frac{\partial f^{ap}}{\partial r_2}+\frac{\partial^2 f^{ap}}{\partial r_1\partial r_2}\right]-(r_1+r_2)=h(\tau,r_1,r_2).
\end{aligned}
$$

In the next step we substitute to the previous equation approximation of the bond price and make Taylor expansion of all the terms with respect to τ:

$$h(\tau,r_1,r_2)=k_3(r_1,r_2)\tau^3+o(\tau^3),$$

where k_3 reads as

$$
\begin{aligned}
k_3(r_1,r_2)=\frac{1}{6r_1^2r_2^2}\Big(&(2\gamma_1^2-\gamma_1)(r_1^{4\gamma_1}r_2^2\sigma_1^4)+(2\gamma_2^2-\gamma_2)(r_1^2r_2^{4\gamma_2}\sigma_2^4)\\
&+\rho\gamma_1(\gamma_1-1)r_1^{3\gamma_1}r_2^{\gamma_2+2}\sigma_1^3\sigma_2+\rho\gamma_2(\gamma_2-1)r_1^{\gamma_1+2}r_2^{3\gamma_2}\sigma_1\sigma_2^3\\
&+2\gamma_2(\alpha_2+\beta_2 r_2)(\rho\sigma_1\sigma_2 r_1^{2+\gamma_1}r_2^{1+\gamma_2}+\sigma_2^2r_1^2r_2^{1+2\gamma_2})+2\gamma_1\gamma_2\rho^2\sigma_1^2\sigma_2^2r_1^{2\gamma_1+1}r_2^{2\gamma_2+1}\\
&+2\gamma_1 r_1 r_2^2\sigma_1(\alpha_1+\beta_1 r_1)\big(r_1^{2\gamma_1}\sigma_1+\rho\sigma_2 r_1^{\gamma_1}r_2^{\gamma_2}\big)\Big).
\end{aligned}
$$

Let us consider function $g(\tau,r_1,r_2)=f^{ap}-f^{ex}$. It satisfies the equation

$$
\begin{aligned}
&-\frac{\partial g}{\partial \tau}+[\alpha_1+\beta_1 r_1]\frac{\partial g}{\partial r_1}+[\alpha_2+\beta_2 r_2]\frac{\partial g}{\partial r_2}+\frac{\sigma_1^2 r_1^{2\gamma_1}}{2}\left[\left(\frac{\partial^2 g}{\partial r_1^2}\right)^2+\frac{\partial^2 g}{\partial r_1}\right]\\
&\quad+\frac{\sigma_2^2 r_2^{2\gamma_2}}{2}\left[\left(\frac{\partial^2 g}{\partial r_2^2}\right)^2+\frac{\partial^2 g}{\partial r_2}\right]+\rho\sigma_1\sigma_2 r_1^{\gamma_1}r_2^{\gamma_2}\left[\frac{\partial g}{\partial r_1}\frac{\partial g}{\partial r_2}+\frac{\partial^2 g}{\partial r_1\partial r_2}\right]\\
&=h(\tau,r_1,r_2)-\sigma_1^2 r_1^{2\gamma_1}\frac{\partial f^{ex}}{\partial r_1}\frac{\partial g}{\partial r_1}-\sigma_2^2 r_2^{2\gamma_2}\frac{\partial f^{ex}}{\partial r_2}\frac{\partial g}{\partial r_2}-\rho\sigma_1\sigma_2 r_1^{\gamma_1}r_2^{\gamma_2}\left[\frac{\partial g}{\partial r_1}\frac{\partial f^{ex}}{\partial r_2}-\frac{\partial g}{\partial r_2}\frac{\partial f^{ex}}{\partial r_1}\right].
\end{aligned}
\tag{5.10}
$$

Taylor expansion of this equation with respect to τ is given by:

$$g(\tau,r_1,r_2)=\sum_{i=0}^{\infty}c_i(r_1,r_2)\tau^i=\sum_{i=\omega}^{\infty}c_i(r_1,r_2)\tau^i,$$

where coefficient $c_\omega(r_1,r_2)\tau^\omega$ is the first non-zero term.
Thus we have $\partial_\tau g=\omega c_\omega(r_1,r_2)\tau^{\omega-1}+o(\tau^{\omega-1})$. Note that $\omega\neq 0$. Coefficient c_0 can not be the first non-zero term in the expansion, because it represents value of the function g in the maturity time of the bond and hence it equals zero (since both f^{ap} and f^{ex} are equal to 1 at maturity). Except for function $h(\tau,r_1,r_2)=k_3(r_1,r_2)\tau^3+o(\tau^3)$, all the terms in the equation (5.10) are multiplied by at least one of the derivatives $\partial_{r_1}g$, $\partial_{r_2}g$, which are of order $O(\tau)$. Hence all the terms, except $h(\tau,r_1,r_2)$, are of the order $o(\tau^{\omega-1})$ for $\tau\to 0^+$. Equation (5.10) then implies

$$-\omega c_\omega(r_1,r_2)\tau^{\omega-1}=k_3(r_1,r_2)\tau^3.$$

We get $\omega=4$, which means that

$$g(\tau,r_1,r_2)=\ln P^{ap}(\tau,r_1,r_2)-\ln P^{ex}(\tau,r_1,r_2)=-\frac{1}{4}k_3(r_1,r_2)\tau^4+o(\tau^4).$$

$\square$

Note that considering a difference of the logarithms of the bond prices is convenient because of calculation of the relative error and the differences in the term structures.

In the next Chapter it is outlined how to implement this approach to two-factor convergence model from The Chapter 3. Combing the approach from these two Chapters leads to the three-factor convergence interest rate model. Details are explained in The Chapter 6.

Chapter 6

6 A three-factor convergence model of interest rates

We propose the three-factor convergence model of CKLS, in which the European short rate is given as a sum of two unobservable factors on the market. The evolution of these two-factors is described by two SDEs. The third SDE describes the evolution of the domestic short rates. We derived the PDE for bond prices and we proposed the approximative analytical formula for the solution of this model.

- *B. Stehlíková, Z. Bučková (Zíková): A three-factor convergence model of interest rates. Proceedings of Algoritmy 2012, pp. 95-104.*

A convergence model of interest rates explains the evolution of the domestic short rate in connection with the European rate. The first model of this kind was proposed by Corzo and Schwartz in 2000 and its generalizations were studied later. In all these models, the European rates are modeled by a one-factor model. This, however, does not provide a satisfactory fit to the market data. A better fit can be obtained using the model, where the short rate is a sum of two unobservable factors. Therefore, we build the convergence model for the domestic rates based on this evolution of the European market. We study the prices of the domestic bonds in this model which are given by the solution of the PDEs. In general, it does not have an explicit solution. Hence we suggest an analytical approximative formula and derive the order of its accuracy in a particular case.

Generalization of the convergence model by Corzo and Schwarz has been studied in the thesis [35] and the paper [14]. The European short rate is assumed to follow CIR and general CKLS models respectively. It is shown in [36] that in the case of uncorrelated Wiener processes governing the evolution of the European and domestic short rate, the pricing of a domestic bond can be reduced to solving a system of ordinary differential equations. For the general case, an analytical approximation formula has been suggested in [14]. This model was then fit to real Euro area and Slovak data in the last quarter before Slovakia joined the monetary union. The resulting fit is, however, not satisfying. This is true for the modeling European rates by the CIR model in the first place. Hence the first question when building a convergence model is the suitable model for the European interest rates. This has been found in the thesis [30]. The instantaneous interest rate is modeled as a sum of two unobservable mean reverting factors. Their sum is also considered to be unobservable, instead of identifying it with an overnight rate, to prevent the possible effect of speculations on the market affecting the overnight. This model achieves a much better fit, see the comparison in the Figure 6.1.

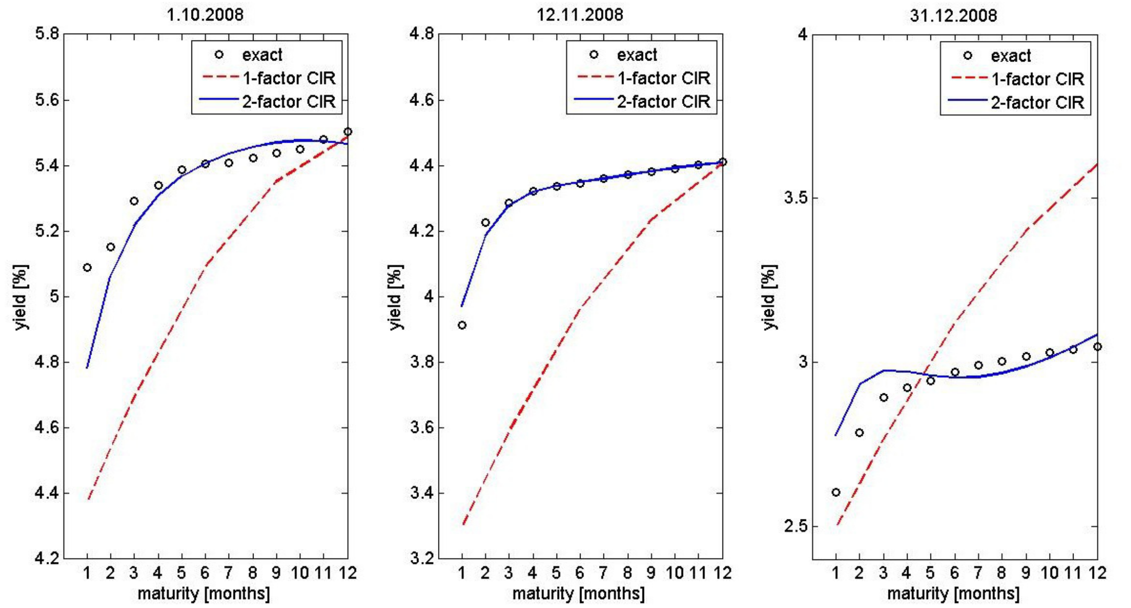


Figure 6.1: Fitting the European term structures from the last quarter of 2008 using the 1-factor CIR and 2-factor CIR models, selected days. Source: [14], [30].

In this Chapter we propose the convergence model, where the European short rate is modeled as the sum of two factors of the CKLS type. Pricing European bonds is derived in the cited work [30]. Here we focus on pricing the domestic bonds - finding explicit solutions, proposing an analytical approximation for the general case and its preliminary analysis.

The Chapter is organized as follows: In the Section 6.1 we define the model in terms of a system of stochastic differential equations. The Section 6.2 deals with bond pricing which is firstly considered in the general case and then in the special cases which will be needed in the rest of the Chapter. In particular, we derive a closed form solution for the Vasicek-type of a model and a reduction to a system of ODEs for a special case of the CIR-type model. Based on the Vasicek closed form solution, we propose an analytical approximation formula for the general CKLS-type model. Using the ODE representation of the exact solution of the CIR model, we derive the order of accuracy of the approximation formula in this case. In the Section 6.3 we test the proposed approximation numerically.

6.1 Formulation of the model

We propose the following model for the joint dynamics of the European r_e and domestic r_d instantaneous interest rate. The European rate $r_e = r_1 + r_2$ is modeled as the sum of the two mean-reverting factors r_1 and r_2, while the domestic rate r_d reverts to the European rate. Volatilities of the processes are assumed to have a general CKLS form. Hence

$$\begin{aligned} dr_1 &= \kappa_1(\theta_1 - r_1)dt + \sigma_1 r_1^{\gamma_1} dw_1 \\ dr_2 &= \kappa_2(\theta_2 - r_2)dt + \sigma_2 r_2^{\gamma_2} dw_2 \\ dr_d &= \kappa_d((r_1 + r_2) - r_d)dt + \sigma_d r_d^{\gamma_d} dw_d \end{aligned}$$

with $Cor(dw) = \mathscr{R}dt$, where $dw = (dw_1, dw_2, dw_d)^T$ is a vector of Wiener processes with correlation matrix $\mathscr{R}$, whose elements (i.e., correlations between r_i and r_j) we denote by ρ_{ij}.

Figure 6.2 and Figure 6.3 show the evolution of the factors and the interest rates for the following set of parameters: $\kappa_1 = 3, \theta_1 = 0.02, \sigma_1 = 0.05, \gamma_1 = 0.5, \kappa_2 = 10, \theta_2 = 0.01; \sigma_2 = 0.05, \gamma_2 = 0.5, \kappa_d = 1, \sigma_d = 0.02, \gamma_d = 0.5, \rho_{ij} = 0$ for all i, j.

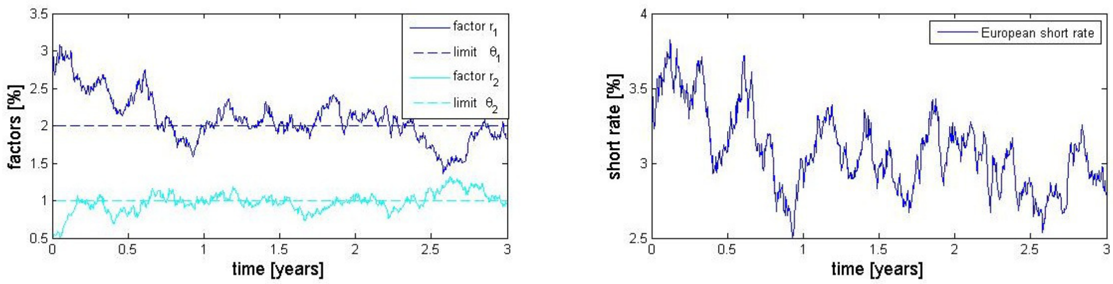


Figure 6.2: Simulation of the factors r_1, r_2 (left) and the European short rate $r_e = r_1 + r_2$ (right).

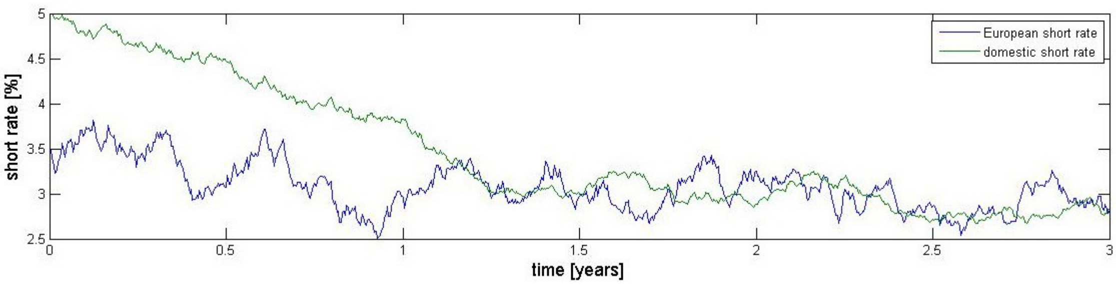


Figure 6.3: Simulation of the European short rate r_e and the domestic short rate r_d.

6.2 Bond prices

To compute the bond prices, it is necessary to specify the so called market prices of risk for each factor, in addition to the SDEs for the short rates. Denoting the market prices of risk as $\lambda_1 = \lambda_1(t, r_1, r_2, r_d)$, $\lambda_2 = \lambda_2(t, r_1, r_2, r_d)$, $\lambda_d = \lambda_d(t, r_1, r_2, r_d)$ we obtain the following PDE for the price $P = P(\tau, r_1, r_2, r_d)$ of the bond with time to maturity

$\tau = T - t$ (c.f. [34]):

$$
\begin{aligned}
-\frac{\partial P}{\partial \tau} + \left[\kappa_d((r_1+r_2)-r_d) - \lambda_d \sigma_d r_d^{\gamma_d}\right] \frac{\partial P}{\partial r_d} + \left[\kappa_1(\theta_1 - r_1) - \lambda_1 \sigma_1 r_1^{\gamma_1}\right] \frac{\partial P}{\partial r_1} \\
+ \left[\kappa_2(\theta_2 - r_2) - \lambda_2 \sigma_2 r_2^{\gamma_2}\right] \frac{\partial P}{\partial r_2} + \frac{\sigma_d^2 r_d^{2\gamma_d}}{2} \frac{\partial^2 P}{\partial r_d^2} + \frac{\sigma_1^2 r_1^{2\gamma_1}}{2} \frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2 r_2^{2\gamma_2}}{2} \frac{\partial^2 P}{\partial r_2^2} \\
+\rho_{1d}\sigma_d r_d^{\gamma_d} \sigma_1 r_1^{\gamma_1} \frac{\partial^2 P}{\partial r_d \partial r_1} + \rho_{2d}\sigma_d r_d^{\gamma_d} \sigma_2 r_2^{\gamma_2} \frac{\partial^2 P}{\partial r_d \partial r_2} + \rho_{12}\sigma_1 r_1^{\gamma_1} \sigma_2 r_2^{\gamma_2} \frac{\partial^2 P}{\partial r_1 \partial r_2} - r_d P &= 0.
\end{aligned}
$$

The PDE holds for all $r_d, r_1, r_2 > 0$ and $\tau \in [0, T]$ and it satisfies the initial condition $P(0, r_d, r_1, r_2) = 1$ for all $r_d, r_1, r_2 > 0$.

6.2.1 Vasicek and CIR type convergence models

We define the Vasicek-type convergence model as the model, where the volatilities of the factors are all constant (i.e., $\gamma_1 = \gamma_2 = \gamma_d = 0$), as a generalization of the one-factor model [58]. Similarly as in this one-factor model, we consider constant market prices of risk, i.e., $\lambda_1(t, r_1, r_2, r_d) = \lambda_1, \lambda_2(t, r_1, r_2, r_d) = \lambda_2, \lambda_d(t, r_1, r_2, r_d) = \lambda_d$, where λ_1, λ_2 and λ_d are constants.

Similarly, as in one-factor and two-factor models proposed in [21] we define the CIR-type convergence model as the model with $\gamma_1 = \gamma_2 = \gamma_d = 1/2$ and the market prices of risk proportional to the square roots of the corresponding factors, i.e., $\lambda_1(t, r_1, r_2, r_d) = \lambda_1\sqrt{r_1}, \lambda_2(t, r_1, r_2, r_d) = \lambda_2\sqrt{r_2}, \lambda_d(t, r_1, r_2, r_d) = \lambda_d\sqrt{r_d}$, where λ_1, λ_2 and λ_d are constants.

The PDE for the bond price then reads as

$$
\begin{aligned}
-\frac{\partial P}{\partial \tau} + \mu_d \frac{\partial P}{\partial r_d} + \mu_2 \frac{\partial P}{\partial r_1} + \mu_3 \frac{\partial P}{\partial r_2} + + \frac{\sigma_d^2 r_d^{2\gamma_d}}{2} \frac{\partial^2 P}{\partial r_d^2} + \frac{\sigma_1^2 r_1^{2\gamma_1}}{2} \frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2 r_2^{2\gamma_2}}{2} \frac{\partial^2 P}{\partial r_2^2} \\
+\rho_{1d}\sigma_d r_d^{\gamma_d} \sigma_1 r_1^{\gamma_1} \frac{\partial^2 P}{\partial r_d \partial r_1} + \rho_{2d}\sigma_d r_d^{\gamma_d} \sigma_2 r_2^{\gamma_2} \frac{\partial^2 P}{\partial r_d \partial r_2} + \rho_{12}\sigma_1 r_1^{\gamma_1} \sigma_2 r_2^{\gamma_2} \frac{\partial^2 P}{\partial r_1 \partial r_2} - r_d P &= 0,
\end{aligned}
$$

where

$$\mu_d = a_1 + a_2 r_d + a_3 r_1 + a_4 r_2, \mu_2 = b_1 + b_2 r_1, \mu_3 = c_1 + c_2 r_2$$

(note that they are in fact the so called risk neutral drifts, c.f. [7] for the relation between bond pricing and the risk neutral measure) with

- $a_1 = -\lambda_d \sigma_d$, $a_2 = -\kappa_d$, $a_3 = \kappa_d$, $a_4 = \kappa_d$, $b_1 = \kappa_1\theta_1 - \lambda_1\sigma_1$,$b_2 = -\kappa_1$, $c_1 = \kappa_2\theta_2 - \lambda_2\sigma_2$, $c_2 = -\kappa_2$ in the Vasicek-type model,
- $a_1 = 0$ $a_2 = -\kappa_d - \lambda_d\sigma_d$, $a_3 = \kappa_d$, $a_4 = \kappa_d$, $b_1 = \kappa_1\theta_1$, $b_2 = -\kappa_1 - \lambda_1\sigma_1$, $c_1 = \kappa_2\theta_2$, $c_2 = -\kappa_2 - \lambda_2\sigma_2$ in the CIR-type model.

We show that in the Vasicek case and the uncorrelated version (i.e., if the Wiener processes w_1, w_2, w_d are uncorrelated) of the CIR case, the solution of the PDE, can be written in a separable form

$$P(r_d, r_1, r_2, \tau) = e^{A(\tau)r_d + B(\tau)r_1 + C(\tau)r_2 + D(\tau)}. \tag{6.1}$$

Furthermore, in the Vasicek model the functions A, B, C, D can be written in the closed form. In the CIR model, they are solutions to the system of ODEs which can be solved numerically much easier than the original PDE.

To prove the claim about the Vasicek model we insert the expected form of the solution (6.1) into the PDE with $\gamma_i = 0$. We obtain

$$\begin{aligned} r_d(-\dot{A} + a_2 A - 1) + r_1(-\dot{B} + a_3 A + b_2 B) + r_2(-\dot{C} + a_4 A + c_2 C) \\ +(-\dot{D} + a_1 A + b_1 B + c_1 C + \frac{\sigma_d^2}{2} A^2 + \frac{\sigma_1^2}{2} B^2 + \frac{\sigma_2^2}{2} C^2 \\ +\rho_{1d}\sigma_d\sigma_1 AB + \rho_{2d}\sigma_d\sigma_2 AC + \rho_{12}\sigma_1\sigma_2 BC) &= 0 \end{aligned}$$

which implies the following system of ODEs:

$$\begin{aligned} \dot{A} &= a_2 A - 1, \\ \dot{B} &= a_3 A + b_2 B, \\ \dot{C} &= a_4 A + c_2 C, \\ \dot{D} &= a_1 A + b_1 B + c_1 C + \frac{\sigma_d^2}{2} A^2 + \frac{\sigma_1^2}{2} B^2 + \frac{\sigma_2^2}{2} C^2 + \rho_{1d}\sigma_d\sigma_1 AB \\ &\quad +\rho_{2d}\sigma_d\sigma_2 AC + \rho_{12}\sigma_1\sigma_2 BC, \end{aligned} \tag{6.2}$$

with initial conditions $A(0) = B(0) = C(0) = D(0) = 0$. Functions A, B, C are easily found to be equal to (here and in the subsequent analysis we assume that $a_2 \neq b_2$ and $a_2 \neq c_2$, and we omit the very special case when the coefficients are equal)

$$A(\tau) = \frac{1 - e^{a_2\tau}}{a_2},$$

$$B(\tau) = \frac{a_3\left(b_2(1 - e^{a_2\tau}) - a_2(1 - e^{b_2\tau})\right)}{a_2 b_2(a_2 - b_2)},$$

$$C(\tau) = \frac{a_4\left(c_2(1 - e^{a_2\tau}) - a_2(1 - e^{c_2\tau})\right)}{a_2 c_2(a_2 - c_2)}.$$

The function D can be found by integration. For the sake of brevity we omit the details.

Now we consider the uncorrelated CIR case. Substituting $\gamma_i = 1/2$ and zero correlations $\rho_{ij} = 0$; and inserting the expected form of the solution (6.1) into the PDE we obtain

$$\begin{aligned} r_d(-\dot{A} + a_2 A + \frac{\sigma_d^2}{2} A^2 - 1) + r_1(-\dot{B} + a_3 A + b_2 B + \frac{\sigma_1^2}{2} B^2) \\ +r_2(-\dot{C} + a_4 A + c_2 C \frac{\sigma_2 1^2}{2} C^2) + (-\dot{D} + a_1 A + b_1 B + c_1 C) &= 0 \end{aligned}$$

which implies the system of ODEs

$$\begin{aligned}
\dot{A} &= a_2 A + \frac{\sigma_d^2}{2} A^2 - 1, \\
\dot{B} &= a_3 A + b_2 B + \frac{\sigma_1^2}{2} B^2, \\
\dot{C} &= a_4 A + c_2 C + \frac{\sigma_2^2}{2} C^2, \\
\dot{D} &= a_1 A + b_1 B + c_1 C,
\end{aligned} \tag{6.3}$$

with initial conditions $A(0) = B(0) = C(0) = D(0) = 0$. Firstly, we find the function A by separation of variables. Then, we independently numerically solve the ODEs for B and C, and finally by numerical integration we obtain the function D.

Figure 6.4 shows the examples of term structures from the CIR-type model, where we have taken $\lambda_d = \lambda_1 = \lambda_2 = 0$. The remaining parameters are the same as in the section 6.1. Note the variety of the term structure shapes which can be obtained for the same values of both the domestic short rate r_d and the European short rate r_e, depending on the decomposition of r_e into the factors r_1 and r_2.

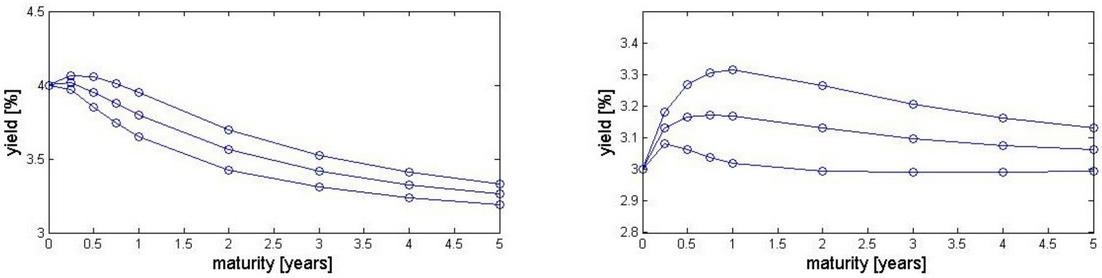


Figure 6.4: Examples of term structures in the CIR-type convergence model. Domestic short rate r_d equals 4% (left) and 3% (right). European short rate r_e equals 5%, the term structures correspond to its different decompositions into factors: $r_1 = 4\%, r_2 = 1\%$; $r_1 = 2.5\%, r_2 = 2.5\%$; $r_1 = 1\%, r_2 = 4\%$.

6.2.2 Analytical approximation formula for general convergence model

In the general case of the convergence model the assumption (6.1) does not lead to a solution. We use the idea of finding an approximative formula which has been successfully used in simpler models (one-factor models in [53], two-factor models in [30] and [14]). We consider the closed form solution from the model of the Vasicek type and replace its constant volatilities $\sigma_1, \sigma_2, \sigma_d$ by instantaneous volatilities $\sigma_1 r_1^{\gamma_1}$, $\sigma_2 r_2^{\gamma_2}$, $\sigma_d r_d^{\gamma_d}$. In this way we obtain the approximation $P^{ap} = P^{ap}(\tau, r_1, r_2, r_d)$.

6.2.3 Order of accuracy in the case of uncorrelated CIR model

Recall that we have the separated form of the solution (6.1) for the bond price in CIR model with zero correlations ρ_{ij} and the system of ODEs (6.3). The system (6.3) enables us to compute the derivatives of the functions A,B,C,D at $\tau = 0$ (see Table 6.1) and consequently the Taylor series expansion of $\ln P(\tau, r_1, r_2, r_d)$ around $\tau = 0$.

Table 6.1: Calculation of the derivatives of functions A,B,C,D from the CIR model with zero correlations

i	**0**	**1**	**2**	**3**	**4**
$A^i(0)$	0	-1	$-a_2$	$-a_2^2+\sigma_d^2$	$-a_2^3+4a_2\sigma_d^2$
$B^i(0)$	0	0	$-a_3$	$-a_3a_2-a_3b_2$	$-a_2^2a_3+a_3\sigma_d^2-a_2a_3b_2-a_3b_2^2$
$C^i(0)$	0	0	$-a_4$	$-a_4a_2-a_4c_2$	$-a_2^2a_4+a_4\sigma_d^2-a_2a_4c_2-a_4c_2^2$
$D^i(0)$	0	0	$-a_1$	$-a_1a_2-b_1a_3-c_1a_4$	$-a_1a_2^2+a_1\sigma_d^2-a_2a_3b_1$ $-a_3b_1b_2-a_2a_4c_1-a_4c_1c_2$

The approximation formula P^{ap} is given in the closed form, hence the Taylor series can be computed also for $\ln P^{ap}(\tau, r_1, r_2, r_d)$. (Alternatively, we can use the system of ODEs similarly as in the case of the exact solution.) The derivatives needed in the expansion are shown in Table 6.2.

Table 6.2: Calculation of the derivatives of functions A,B,C,D from the approximation of the CIR model with zero correlations.

i	**0**	**1**	**2**	**3**	**4**
$A^i(0)$	0	-1	$-a_2$	$-a_2^2$	$-a_2^3$
$B^i(0)$	0	0	$-a_3$	$-a_3a_2-a_3b_2$	$-a_2^2a_3-a_2a_3b_2-a_3b_2^2$
$C^i(0)$	0	0	$-a_4$	$-a_4a_2-a_4c_2$	$-a_2^2a_4-a_2a_4c_2-a_4c_2^2$
$D^i(0)$	0	0	$-a_1$	$-a_1a_2-b_1a_3-c_1a_4+\sigma_d^2 r_d$	$-a_1a_2^2-a_2a_3b_1-a_3b_1b_2$ $-a_3b_1b_2-a_2a_4c_1-a_4c_1c_2$ $+3a_2\sigma_d^2 r_d$

Comparing the expressions in Table (6.1) and Table (6.2) we obtain the order of the difference $\ln P^{ap}(\tau, r_1, r_2, r_d) - \ln P(\tau, r_1, r_2, r_d)$ which can be interpreted in terms of the relative error in bond prices and the absolute error in term structures, as stated in the following theorem and its corollary.

Theorem 6.1. *Let $P^{CIR,\rho=0}$ be the bond price in the CIR-type convergence model with zero correlations and let $P^{CIR,\rho=0,ap}$ be its approximation proposed in the Section 6.2.2. Then*

$$\ln P^{CIR,\rho=0,ap} - \ln P^{CIR,\rho=0} = -\frac{1}{24}\sigma_d^2\left(a_1 + a_2 r_d + a_3 r_1 + a_4 r_2\right)\tau^4 + o(\tau^4)$$

for $\tau \to 0^+$.

Note that the form of the leading term of the approximation error (i.e., $-\frac{1}{24}\sigma_d^2$ times the risk neutral domestic drift) is the same as in the two-factor convergence model [14], where the analogical strategy of forming the approximative formula has been used.

Corollary 6.2. *1. The relative error of the bond price satisfies*

$$\frac{P^{CIR,\rho=0,ap}-P^{CIR,\rho=0}}{P^{CIR,\rho=0}}=-\frac{1}{24}\sigma_d^2\,(a_1+a_2r_d+a_3r_1+a_4r_2)\,\tau^4+o(\tau^4)$$

for $\tau\to 0^+$.

2. The error in interest rates R can be expressed as

$$R^{CIR,\rho=0,ap}-R^{CIR,\rho=0}=\frac{1}{24}\sigma_d^2\,(a_1+a_2r_d+a_3r_1+a_4r_2)\,\tau^3+o(\tau^3)$$

for $\tau\to 0^+$.

Proof. The first corollary is a consequence of the Taylor expansion of the exponential function $e^x=1+x+o(x)$ for $x\to 0^+$. The second corollary follows from the formula $R(\tau,r)=-\frac{\ln P(\tau,r)}{\tau}$ for calculating the interest rates R from the bond prices P (c.f. [7], [34]). □

6.3 Numerical experiment

We consider the term structures presented in Figure 6.4 and compare them with the approximate values obtained by the proposed formula. The results are summarized in Table 6.3 and Table 6.4. The accuracy is very high (note that Euribor is quoted to three decimal places) even for higher maturities.

Table 6.3: Exact interest rates and their approximations obtained by the proposed formula. The domestic short rate is 4%, the European short rate is 5%, the columns correspond to the different values of the factors: $r_1=4\%, r_2=1\%$ (left), $r_1=2.5\%, r_2=2.5\%$ (middle), $r_1=1\%, r_2=4\%$ (right).

maturity	exact	approx.	exact	approx.	exact	approx.
0	4.00000	4.00000	4.00000	4.00000	4.00000	4.00000
0.25	4.06607	4.06607	4.01638	4.01638	3.96668	3.96668
0.5	4.05591	4.05591	3.95219	3.95219	3.84847	3.84847
0.75	4.00932	4.00931	3.87493	3.87493	3.74055	3.74054
1	3.94734	3.94733	3.7995	3.79949	3.65166	3.65165
2	3.69802	3.69796	3.56221	3.56217	3.4264	3.42638
3	3.52184	3.52171	3.41487	3.41479	3.30791	3.30788
4	3.40688	3.40669	3.32208	3.32196	3.23728	3.23724
5	3.32995	3.32972	3.26077	3.26062	3.19158	3.19153

Table 6.4: Exact interest rates and their approximations obtained by the proposed formula. The domestic short rate is 3%, the European short rate is 5%, the columns correspond to the different values of the factors: $r_1 = 4\%, r_2 = 1\%$ (left), $r_1 = 2.5\%, r_2 = 2.5\%$ (middle), $r_1 = 1\%, r_2 = 4\%$ (right).

maturity	**exact**	**approx.**	**exact**	**approx.**	**exact**	**approx.**
0	3.00000	3.00000	3.00000	3.00000	3.00000	3.00000
0.25	3.18127	3.18127	3.13158	3.13158	3.08189	3.08189
0.5	3.26898	3.26898	3.16526	3.16526	3.06154	3.06154
0.75	3.30582	3.30583	3.17144	3.17144	3.03705	3.03705
1	3.31524	3.31524	3.1674	3.16741	3.01957	3.01957
2	3.26573	3.2657	3.12992	3.12991	2.99411	2.99412
3	3.20515	3.20508	3.09818	3.09816	2.99122	2.99124
4	3.1615	3.1614	3.0767	3.07667	2.9919	2.99194
5	3.13134	3.13121	3.06215	3.06211	2.99296	2.99301

Part II

Alternating direction explicit methods, Fichera theory and Trefftz methods

Chapter 7

Introduction to the numerical solutions, ADE schemes, Fichera theory and option pricing

We focus on the ADE methods, as an efficient scheme, which can be used for a wide range of financial problems. Originally, we had planned to implement special meshes, such as *Shishkin's mesh*, but analysis showed that usage of this mesh decreases the convergence order of our scheme. Instead of a second order scheme, we would just obtain a first order scheme.
Hence, we use a *uniform mesh*, in all space directions, and both in time and space. According to the consistency proofs and experimental convergence study of the ADE schemes, we can confirm that they are suitable for the uniform grid, because usage of a nonuniform grid would ruin the second order accuracy, as well.

Designing the numerical scheme, we not only need to take care for the choice of a mesh, but we also have to choose the boundary conditions carefully, as well. Because of the issue with the boundary conditions we have studied the *Fichera theory*, which helped us to distinguish how to define boundary conditions for PDEs degenerating on the boundary. According to the *sign of the Fichera function*, we chose which kind of boundary conditions needs to be supplied.
The second issue about boundary conditions is the influence of the stability of the numerical scheme. Since the matrix approach also includes boundary conditions, we prefer to use it for the stability analysis instead of the von Neumann stability analysis.

We have considered the ADE method, that strongly uses boundary data in the solution algorithm and hence it is very sensible to incorrect treatment of boundary conditions. We have implemented the ADE scheme for solving linear and nonlinear BS equations by treating the nonlinearity explicitly. ADE scheme consists of two steps (sweeps). In the first step an upward sweeping is used and in the second step on downward sweeping is used and they are combined after each time step. To our knowledge, the ADE scheme has not been applied to nonlinear PDEs before.
It can compete to the Crank-Nicolson scheme, Alternating Direction Implicit (ADI) and locally one-dimensional LOD splitting method. Applying the ADE method to linear models leads to an explicit scheme with unconditional stability. Applying the ADE to nonlinear models does not lead to an explicit scheme any more. In each time step we

need to solve a scalar nonlinear equation, but no any more nonlinear system of equation. Hence, the computational effort using ADE instead of an implicit scheme is reduced significantly. For nonlinear cases we obtain only conditional stability. ADI methods and Splitting methods are examples of the Multiplicative Operator Scheme (MOS), which is difficult to parallelize. Methods from the family of Additive Operator Scheme (AOS) can be parallelized. ADE methods also belong to this group of methods. The ADE scheme consists of two explicit sweeps. The sweeping procedure is done from one boundary to another and vice versa.

7.1 Proper treatment of boundary conditions, using Fichera theory

The Fichera theory was first proposed in 1960 by Gaetano Fichera and later developed by Olejnik and Radkevič in 1973. It turned out to be very useful for establishing the well-posedness of initial boundary value problems for PDEs degenerating to hyperbolic PDEs at the boundary.

The *Fichera theory* focuses on the question of appropriate *boundary conditions* (BCs) for parabolic PDEs degenerating at the boundary. According to the sign of the *Fichera function* one can separate the outflow or inflow part of the solution at the boundary. Thus, this classical theory indicates whether one has to supply a BC at the degenerating boundary.

In this paper we illustrate the application of the Fichera theory to the Cox-Ingersoll-Ross (CIR) interest rate model and its generalization, the *Chan-Karolyi-Longstaff-Sanders* (CKLS) model [16]. Here, at the left boundary the interest rate tends to zero and thus the parabolic PDE degenerates to a hyperbolic one. For further applications of Fichera theory to other current models in financial mathematics we refer the interested reader to [22].

7.2 Option pricing with Black-Scholes model

In 1973 Fischer Black and Myron Scholes in the paper [5] derived the well-known Black-Scholes formula (BS), which has the form of the PDE for the option price. Starting with SDE for an underlying stock price, using Itô formula, constructing risk-free portfolio, using no-arbitrage principe they derived the derived formula, whose solution is near to a fair price in the market. Later the paper *Black-Scholes options pricing model* by Robert Merton was published. In 1997 Scholes and Merton were awarded by the Nobel prize for their work.

The Black-Scholes equation is a parabolic PDE with space dependent coefficients:

$$v_t = \frac{1}{2}\sigma^2 S^2 v_{SS} + rSv_S - rv, \qquad t \geq 0, \forall S \in \mathbb{R}, \tag{7.1}$$

where the solution $v(S,t)$ stands for a European option price. A European call (put) option is a contract between its buyer and holder, to buy (sell) a stock at the maturity time T (final time) for the fixed price K, called also strike price. Solution of the linear equation (7.1) is given in closed form formula and it is known as a Black-Scholes formula. Black-Scholes equation is derived under strict assumptions in the market, such as no transaction costs, illiquidity, etc. Modeling this phenomena in a more realistic way, it leads to the nonlinear BS model which does not have any more analytical solution.

7.2.1 Multi-dimensional Black-Scholes models

One of the simplest financial derivative pricing models is the Black-Scholes model, which has the form of a one dimensional PDE with one space dimension and one time dimension.

Considering more complex models, that include a variety of market effects such as stochastic volatility or correlation among financial assets can increase the dimensionality of the pricing PDE. Also, pricing financial derivatives with more than one underlying asset yields PDEs that have at least as many spatial dimensions as the number of underlying assets.

Since a closed form formula can be only found in very special cases, determining solutions for these models has to be done in general using numerical methods, but the higher the dimension of the PDE models the bigger the overall complexity of the implementation of these methods.

Since the ADE scheme is explicit, stable and thus efficient, it represents a good candidate to compute the numerical solution of these multi-dimensional models in finance.

Here we present the implementation of the ADE schemes to two and three dimensional models appearing in finance, esp. the multi-dimensional linear Black-Scholes model. One of the advantages of this approach is that its fundamental implementation set-up can be transferred to higher dimensions.

We study a financial derivative that can be exercised only at a pre-fixed maturity time T (commonly referred as 'European' option) and whose payoff depends on the value of N financial assets with prices $S_1,\dots,S_N$. We assume a financial market with the standard Black-Scholes assumptions, explained in details e.g. in [60]. Although this is very restrictive from the modeling point of view, it is enough to illustrate the implementation of the ADE schemes in a high-dimensional setting. Under this model the price of a derivative $V(S_1,\dots,S_N,\tau)$ is given by the following $N-$dimensional linear parabolic PDE:

$$\frac{\partial V}{\partial \tau} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\Gamma_{ij}S_iS_j}{2}\frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{N} rS_i\frac{\partial V}{\partial S_i} - rV, \qquad \tau \geq 0, \forall S \in \mathbb{R}_0^+, \tag{7.2}$$

where r denotes the risk-free interest rate, $\tau = T - t$ is the remaining time to the maturity time T and we have the covariance matrix Γ,

$$\Gamma_{ij} \equiv \rho_{ij}\sigma_i\sigma_j, \qquad i,j = 1,\ldots,N, \tag{7.3}$$

with ρ_{ij} being the correlation between asset i and j and σ_i the standard deviation of the asset i. Additionally we have an initial condition which is defined by the payoff of the option,

$$V(S_1,\ldots,S_N,0) = \Phi(S_1,\ldots,S_N). \tag{7.4}$$

We obtain different models by choosing different numbers of underlying assets (i.e. the number of spatial variables) and defining different payoff functions with corresponding initial conditions. Here we consider both spread options and call options, which have payoffs given by:

$$2-D \text{ Spread option: } V(S_1,S_2,0) = \max(S_1 - S_2 - K, 0),$$
$$N-D \text{ Call option: } V(S_1,\ldots,S_N,0) = \max(\max(S_1,\ldots,S_N) - K, 0)).$$

7.3 Alternating Direction Explicit Schemes

ADE schemes are efficient finite-difference schemes to solve PDEs where the discretization of the spatial derivatives is made using available information of both the current and the previous time-steps such that the solution can be determined without solving a linear system of equations.

ADE schemes were proposed by Saul'ev [47] in 1957, later developed by Larkin [37], Bakarat and Clark [3] in 1964-66. More recently, these schemes have received some attention by Duffy [24], [23] 2013 and Leung and Osher [38] 2005 who have studied and applied these schemes in both financial modeling and other applications.

Some advantages of the ADE methods are that they can be implemented in a parallel framework and are very fast due to their explicitness; for a complete survey on the advantages and the motivation to use them in a wide range of problems we refer the reader to [22], [23].

Numerical analysis results focusing on stability and consistency considerations are described in [38] and [8]. In [8] a numerical analysis of convection-diffusion-reaction equation with constant coefficients and smooth initial data is provided. The authors proved that the ADE method applied to the one-dimensional reaction-diffusion equation on a uniform mesh with the discretization of the diffusion according to Saul'ev [47] and the discretization of the convection term following Towler and Yang [55] is unconditionally stable. If a convection term is added to the equation and upwind discretization for this term is used, the ADE scheme is also unconditionally stable c.f. [8].

In the ADE schemes one computes for each time level two different solutions which are referred to as sweeps. Hereby the number of sweeps does not depend on the dimension. It has been shown [8, 24, 38] that for the upward and downward sweep the consistency is of order $O((d\tau)^2 + h^2 + \frac{d\tau}{h})$ where $d\tau$ is the time step and h denotes the space step. An

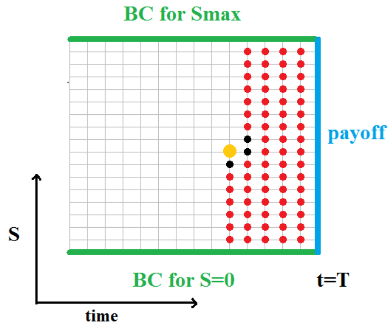


Figure 7.1: Upward sweep

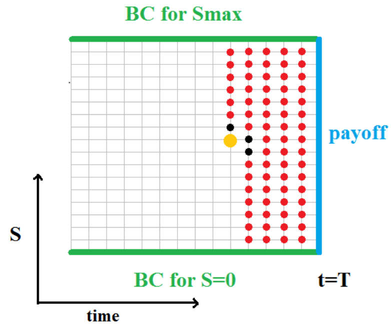


Figure 7.2: Downward sweep

exceptionality of the ADE method is that the average of upward and downward solutions has consistency of order $O((d\tau)^2+h^2)$. For linear models, unconditional stability results and the $O((d\tau)^2+h^2)$ order of consistency lead to the $O((d\tau)^2+h^2)$ convergence order.

Stability, consistency and convergence analysis can be extended to higher dimensional models.

The straightforward implementation also to nonlinear cases with preserving good stability and consistency properties of the scheme is also a strong advantage. In this paper we show how one can implement this scheme for higher dimensional models by focusing on a linear model. However, one could use this procedure for non-linear models as well. One way how to do it is to solve nonlinear equation in each time level, instead of system of nonlinear equations in case of implicit schemes. Another way is to keep nonlinearity in the explicit form and solve it directly. Powerful tool for nonlinear equations represents also the Alternating segment explicit-implicit and the implicit-explicit parallel difference method [63].

7.3.1 The Idea of the ADE scheme

The ADE scheme consists of two explicit sub steps, called sweeps. A sweeping step is constructed from one boundary to another and vice versa. Figure 7.1 is an illustrative example of an upward sweep (analogous to the downward sweep in Figure 7.2).

Figures 7.1 and 7.2 display the grid for the calculating the price of call option in the Black-Scholes model. The blue line represents the payoff as an initial condition and the green lines are given by Dirichlet boundary conditions for small and big asset values. Calculation is provided backward in time.

To calculate the value of the yellow point we use the black values. We can see that we do not use only values from the previous time level but also already known values from the current time level, which preserve explicitness of the scheme. After each time level we combine the solutions from the upward and downward sweep by averaging.

To introduce the ADE method systematically we follow the lines of Leung and Osher [38], and Duffy [22]. The computational spatial interval $(x_{\min}, x_{\max})$, or $(0, S_{\max})$, respectively, is divided into J subintervals, i.e. the space step is $h = (x_{\max} - x_{\min})/J$ and the grid points $x_j = jh$, or $h = S_{\max}/J$, $S_j = jh$, respectively. Thus we get for the coefficients of the BS equation (7.1) $a(S_j) = \frac{1}{2}\sigma^2(jh)^2$, $b(S_j) = rjh$, $c(S_j) = r$.

We consider the resulting spatial semidiscretization to the PDE (7.1), i.e. the following system of ODEs

$$v' = A(v)v, \qquad t > 0, \tag{7.5}$$

with $v(t) \in \mathbb{R}^{J-1}$. Let us consider for simplicity a uniform grid; the time interval $[0,T]$ is divided uniformly into N sub-intervals, with the step size $k = T/N$, i.e. we have the grid points $t_n = nk$. Applying the trapezoidal rule to (7.5) leads to the Crank-Nicolson scheme

$$v^{n+1} = \big[I - kA(v^n)\big]^{-1}\big[I + kA(v^n)\big]v^n, \tag{7.6}$$

where $v^n \approx v(t_n)$. While this classical scheme (7.6) is unconditionally stable and of second order in time and space, it becomes computationally expensive to invert the operator $I - kA(v^n)$ especially in higher space dimensions. In order to obtain an efficient scheme while keeping the other desirable properties, this operator is split additively by the matrix decomposition $A = L + D + U$, where L is lower diagonal, D is diagonal and U denotes an upper-diagonal matrix. Next, following the notation of [38] we further define the *symmetric splitting*

$$B = L + \frac{1}{2}D, \qquad C = U + \frac{1}{2}D. \tag{7.7}$$

Then we can formulate the three steps of the ADE scheme with its upward/downward sweeps and the combination (also for higher dimensions) as

$$\text{UP} \qquad u^{n+1} = \big[I - kB(v^n)\big]^{-1}\big[I + kC(v^n)\big]v^n, \tag{7.8}$$

$$\text{DOWN} \qquad d^{n+1} = \big[I - kC(v^n)\big]^{-1}\big[I + kB(v^n)\big]v^n, \tag{7.9}$$

$$\text{COMB} \qquad v^{n+1} = \frac{1}{2}\big[u^{n+1} + d^{n+1}\big]. \tag{7.10}$$

In other words, in the two sweeps above we assign the solution values that are already computed on the new time level to the operator to be inverted. Hence, the resulting scheme is *explicit*, i.e. efficient. There remain the questions, if we could preserve the unconditional stability and second order accuracy. This will be our main topic in the sequel.

Let us summarize the procedure for one space dimension. The approximation to the solution $v(x,t)$ at the grid point (x_j, t_n) is $c(x_j, t_n) =: c_j^n$ given as an average of upward sweep u_j^n and downward sweep d_j^n. This combination c_j^n contains the initial data at the beginning. For $n = 0, 1, \dots, N-1$ we repeat the following steps:

1. Initialization: $u_j^n = c_j^n, \quad d_j^n = c_j^n, \qquad j = 1, \dots, J-1$

2. **U**pward sweep: $u_j^{n+1}, \qquad j = 1, \dots, J-1$

3. **D**ownward sweep: $d_j^{n+1}, \qquad j = J-1, \ldots, 1$

4. **C**ombination: $\quad c^{n+1} = (u^{n+1} + d^{n+1})/2$

Using different approximation strategies for the convection, diffusion and reaction terms we obtain different variations of the ADE schemes, which were proposed by Saul'ev [47].

7.3.2 Solving PDEs with the ADE method

We start considering the partial differential equation (PDE)

$$v_t = a\,v_{xx} + b\,v_x - c\,v, \qquad t \geq 0, \forall x \in \mathbb{R}, \tag{7.11}$$

with the constant coefficients $a \geq const. > 0$, $b \geq 0$, $c \geq 0$ and supplied with smooth initial data. We denote the analytical classical solution of (7.11) by $v := v(x,t)$ and use subscripts to abbreviate partial differentiation, e.g. $v_{xx} := \partial^2 v / \partial x^2$.

Secondly, we will consider the classical linear Black-Scholes (BS) equation

$$v_t = \frac{1}{2}\sigma^2 S^2 v_{SS} + rSv_S - rv, \qquad t \geq 0, \forall S \in \mathbb{R}, \tag{7.12}$$

which is a generalization of the PDE (7.11) to space dependent coefficients. In computational finance a solution $v(S,t)$ of the PDE (7.12) represents a European option price. A European option is a contract between the holder of the option and the future buyer, that at a time instance T, the expiration time, the underlying asset (stock) can be sold or bought (call or put option) for a fixed strike price K. Using the Black-Scholes formula the option price is calculated for the corresponding underlying asset price S (stock price) in a time interval $t \in (0,T)$.

Let us note that the BS equation (7.12) is derived under quite restrictive market assumptions, which are not very realistic. Relaxing these assumptions leads to new models (e.g. including transaction costs, illiquidity on the market) that are strongly nonlinear BS equations that can only be solved analytically in very simple cases.

While there exist analytical tools to solve explicitly (7.11) and (7.12), the interest in studying the ADE method for these simple 1D cases is the fact that we want to extend this approach in a subsequent work to nonlinear PDEs and to higher dimensions. Applying the ADE to the nonlinear BS equations we need to solve only a scalar nonlinear equation (instead of a nonlinear system of equations for a standard implicit method). Thus, the computational effort using ADE instead of an implicit scheme is highly reduced. Also, for higher space dimensions the number of ADE sweeps does not increase, it remains two. These facts make the ADE methods an attractive candidate to study them in more detail.

8

Chapter 8

Fichera theory and its application to finance

Firstly, we discuss the Fichera theory that helps us to determine in which cases boundary conditions are needed and in which cases they are not allowed.

In this Chapter we outline the application of the Fichera theory to interest rates models of Cox-Ingersoll-Ross (CIR) and Chan-Karolyi-Longstaff-Sanders (CKLS) type. For the one-factor CIR model the obtained results are consistent with the corresponding Feller condition.

Chapter is based on

- *Z. Bučková, M. Ehrhardt, M. Günther: Fichera theory and its application to finance, Proceedings ECMI 2014, Taormina, Sicily, Italy, 2016*

8.1 The Boundary Value Problem for the Elliptic PDE

We consider an elliptic second order linear differential operator

$$Lu = \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + cu, \qquad x \in \Omega \subset \mathbb{R}^n, \tag{8.1}$$

where $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ is symmetric and induces a semi-definite quadratic form $\xi^\top A\xi \geq 0$ for all $\xi \in \mathbb{R}^n$. Σ denotes a piecewise smooth boundary of the domain Ω. The subset of Σ where the quadratic form vanishes, $\xi^\top A\xi = 0$, will be denoted as Σ_h (hyperbolic part) and the set of points of Σ where the quadratic form remains positive, $\xi^\top A\xi > 0$, is denoted as a Σ_p (parabolic) part. For Σ_h, the hyperbolic part of the boundary Σ_h, we introduce the *Fichera function*

$$b = \sum_{i=1}^{n} \left(b_i - \sum_{k=1}^{n} \frac{\partial a_{ik}}{\partial x_k} \right) \nu_i, \tag{8.2}$$

where ν_i is the direction cosine of the inner normal to Σ, i.e. it is $\nu_i = \cos(x_i, \vec{n}_i)$, where $\vec{n}_i$ is the inward normal vector at the boundary.

On the *hyperbolic part* of the boundary Σ_h we define according to the sign of the Fichera function the three subsets Σ_0 ($b = 0$ tangential flow), Σ_+ ($b > 0$, outflow) and Σ_- ($b < 0$, inflow), i.e. the boundary $\Sigma = \Sigma_p \cup \Sigma_h$ can be written as a unification of four boundary parts: $\Sigma = \Sigma_p \cup \Sigma_0 \cup \Sigma_+ \cup \Sigma_-$.

Olejnik and Radkevič [41, Lemma 1.1.1] showed that the sign of the Fichera function b at the single points Σ_h does not change under smooth non-degenerate changes of independent variables in a given elliptic operator (8.1). In [41, Theorem 1.1.1] it is stated that the subsets Σ_0, Σ_+, Σ_- remain invariant under a smooth nonsingular changes of independent variables in the elliptic operator (8.1).

The *parabolic boundary* Σ_p can be rewritten as a unification of two sets Σ_p^D (Dirichlet BC) and Σ_p^N (Neumann BC). Let us state one simple example.

Example 8.1. *[28] The boundary value problem for an elliptic PDE reads*

$$\begin{aligned} Lu &= f \qquad \text{on} \quad \Omega \subset \mathbb{R}^n, \\ u &= g \qquad \text{on} \quad \Sigma_- \cup \Sigma_p^D, \\ a_{ij}\frac{\partial u}{\partial x_i} n_j &= h \qquad \text{on} \quad \Sigma_p^N. \end{aligned}$$

If Σ_p^N is an empty set, we obtain a Dirichlet problem; if Σ_p^D is an empty set, a Neumann problem; if Σ_p^D and Σ_p^N are not empty, the problem is of mixed Dirichlet-Neumann type. Recall that for hyperbolic PDEs one must not supply BCs for outflow boundaries (Σ_+) or boundaries where the characteristics are tangential to the boundary (Σ_0), since this may violate the information that is transported from the interior of the domain.

8.2 Application to one-factor interest rate Models of CKLS type

We deal with the models from the first part of the thesis. We start with an interest rate model in the form of a stochastic differential equation

$$dr = \kappa(\theta - r)\,dt + \sigma r^\gamma\,dW, \tag{8.3}$$

where κ, θ are positive constants, and γ non-negative. This CKLS model [16] is a mean-reversion process with non-constant volatility σr^γ. Using the Itô formula for a duplicating portfolio in a risk neutral world one can derive a PDE for the zero-coupon bond price $P(r,\tau)$:

$$\frac{\partial P}{\partial \tau} = \alpha(r,\tau)\frac{\partial^2 P}{\partial r^2} + \beta(r,\tau)\frac{\partial P}{\partial r} - rP, \qquad r > 0,\ \tau > 0, \tag{8.4}$$

where $\alpha(r,\tau) = \frac{1}{2}\sigma^2 r^{2\gamma}$, $\beta(r,\tau) = \kappa(\theta - r)$. A closed form formula for this model can be given in special cases, c.f. [7]:

a) if $\gamma = 0$, this is the classical Vašíček model with constant volatility.

b) for $\gamma = 0.5$, we get the Cox-Ingersoll-Ross (CIR) model (CIR), [21].

For general γ (CKLS model) there is no closed form formula for the bond price $P(r,\tau)$ and the PDE (8.4) has to be solved numerically.

The volatility term in (8.4), for a short rate r tending to zero, is $\alpha(0,\tau) = 0$. Thus the parabolic PDE (8.4) reduces at $r = 0$ to the hyperbolic PDE

$$\frac{\partial P}{\partial \tau} = \kappa\theta \frac{\partial P}{\partial r}, \qquad \tau > 0. \tag{8.5}$$

Next, the Fichera function (8.2) for our model reads

$$b(r) = \beta(r,\tau) - \frac{\partial \alpha(r,\tau)}{\partial r}, \tag{8.6}$$

and we check the sign of (8.6) for $r \to 0+$:

- if $\lim_{r\to 0^+} b(r) \geq 0$ (outflow boundary) we must not supply any BCs at $r = 0$.
- if $\lim_{r\to 0^+} b(r) < 0$ (inflow boundary) we have to define BCs at $r = 0$.

Especially for the proposed model we get $b(r) = \kappa(\theta - r) - \sigma^2\gamma r^{2\gamma-1}$ and we we can distinguish the following situations:

a) for $\gamma = 0.5$ (CIR model) $\Rightarrow$ if $\kappa\theta - \sigma^2/2 \geq 0$, we do not need any BCs.

b) for $\gamma > 0.5 \Rightarrow$ if $\kappa\theta \geq 0$, we do not need any BCs.

c) for $\gamma \in (0, 0.5) \Rightarrow$ if $\lim_{r\to 0^+} b(r) = -\infty$, we need BCs.

Remark 8.1 (Feller condition)**.** *The Feller condition guaranteeing a positive interest rate defined by* (8.3) *for the one-factor CIR model is* $2\kappa\theta > \sigma^2$ *and is equivalent with the condition derived from the Fichera theory. If the Feller condition holds, then the Fichera theory states that one must not supply any BC at* $r = 0$.

8.3 A two-factor interest rate Model

We consider a general two-factor model given by the set of two SDEs

$$\mathrm{d}x_1 = (a_1 + a_2 x_1 + a_3 x_2)\,\mathrm{d}t + \sigma_1 x_1^{\gamma_1}\,\mathrm{d}W_1, \tag{8.7}$$

$$\mathrm{d}x_2 = (b_1 + b_2 x_1 + b_3 x_2)\,\mathrm{d}t + \sigma_2 x_2^{\gamma_2}\,\mathrm{d}W_2, \tag{8.8}$$

$$Cov[\mathrm{d}W_1, \mathrm{d}W_2] = \rho\,\mathrm{d}t, \tag{8.9}$$

containing as special cases the Vašíček model ($\gamma_1 = \gamma_2 = 0$) and the CIR model ($\gamma_1 = \gamma_2 = 0.5$). The drift functions are defined as linear functions of the two variables x_1 and x_2. Choosing $a_1 = b_1 = b_2 = 0$ we get two-factor convergence model of CKLS type (in case of general $\gamma_1,\ \gamma_2 \geq 0$). The variable x_1 models the interest rate of a small country (e.g. Slovakia) before entering the monetary EURO union and the variable x_2 represents the interest rate of the union of the countries (such as the EU).

Applying the standard Itô formula one can easily derive a parabolic PDE

$$\frac{\partial P}{\partial \tau} = \tilde{a}_{11}\frac{\partial^2 P}{\partial x_1^2} + \tilde{a}_{22}\frac{\partial^2 P}{\partial x_2^2} + \tilde{a}_{12}\frac{\partial^2 P}{\partial x_1 \partial x_2} + \tilde{a}_{21}\frac{\partial^2 P}{\partial x_2 \partial x_1} + \tilde{b}_1\frac{\partial P}{\partial x_1} + \tilde{b}_2\frac{\partial P}{\partial x_2} + \tilde{c}P, \tag{8.10}$$

where $P(x,y,\tau)$ represents the bond price at time τ for interest rates x and y, and

$$\tilde{a}_{11} = \frac{\sigma_1^2 x_1^{2\gamma_1}}{2}, \qquad \tilde{a}_{22} = \frac{\sigma_2^2 x_2^{2\gamma_2}}{2}, \qquad \tilde{a}_{12} = \tilde{a}_{21} = \frac{1}{2}\rho\sigma_1 x_1^{\gamma_1}\sigma_2 x_2^{\gamma_2}$$
$$\tilde{b}_1 = a_1 + a_2 x_1 + a_3 x_2, \qquad \tilde{b}_2 = b_1 + b_2 x_1 + b_3 x_2, \qquad \tilde{c} = -x_1,$$

for $x_1, x_2 \geq 0$, $\tau \in (0,T)$, with initial condition $P(x_1,x_2,0) = 1$ for $x_1,\ x_2 \neq 0$.

Now, the Fichera function (8.2) in general reads

$$b(x_1,x_2) = \left[a_1 + a_2x_1 + a_3x_2 - \left(\sigma_1^2\gamma_1 x_1^{2\gamma_1-1} + \frac{1}{2}\rho\sigma_1 x_1^{\gamma_1}\sigma_2\gamma_2 x_2^{\gamma_2-1}\right)\right]\frac{x_1}{\sqrt{1+x_1^2}}$$
$$+ \left[b_1 + b_2x_1 + b_3x_2 - \left(\frac{1}{2}\rho\sigma_1\gamma_1 x_1^{\gamma_1-1}\sigma_2 x_2^{\gamma_2} + \sigma_2^2\gamma_2 x_2^{2\gamma_2-1}\right)\right]\frac{x_2}{\sqrt{1+x_2^2}}.$$

Depending on γ_1 and γ_2, we get the following results:

- For $\gamma_1 = \gamma_2 = 0$ (classical Vašíček model), the Fichera function simplifies to

$$b(x_1,x_2) = (a_1+b_1) + (a_2+b_2)x_1 + (a_3+b_3)x_2,$$

  and boundary conditions must be supplied, if

$$\begin{cases} x_1 \leq -\frac{a_1+b_1+(a_3+b_3)x_2}{a_2+b_2} & \text{for} \quad a_2+b_2 \neq 0 \\ x_2 \leq -\frac{a_1+b_1}{a_3+b_3} & \text{for} \quad a_2+b_2 = 0, a_3+b_3 \neq 0 \\ a_1+b_1 \leq 0 & \text{for} \quad a_2+b_2 = 0, a_3+b_3 = 0 \end{cases}.$$

- For $\gamma_1 = \gamma_2 = 0.5$ (CIR model), the Fichera function simplifies to

$$b(x_1,x_2) = \left[a_1 + a_2x_1 + a_3x_2 - \left(\sigma_1^2\gamma_1 + \frac{1}{4}\rho\sigma_1\sigma_2\sqrt{\frac{x_1}{x_2}}\right)\right]\frac{x_1}{\sqrt{1+x_1^2}}$$
$$+ \left[b_1 + b_2x_1 + b_3x_2 - \left(\frac{1}{4}\rho\sigma_1\sigma_2\sqrt{\frac{x_2}{x_1}} + \sigma_2^2\gamma_2\right)\right]\frac{x_2}{\sqrt{1+x_2^2}}$$

We must supply boundary conditions for $\rho > 0$, and must not for $\rho < 0$. For $\rho = 0$, BCs at $x_2 = 0$ must be posed if $x_1 \leq \sigma_1^2 \gamma_1/(2a_2) - a_1/a_2$ (assuming $a_2 > 0$, and for $x_1 = 0$, if $x_2 \leq \sigma_2^2 \gamma_2/(2b_2) - b_1/b_2$ (assuming $b_2 > 0$), otherwise not.

- For the general case γ_1, $\gamma_2 > 0$, we discuss the boundary $x_2 = 0, x_1 > 0$; due to symmetry, the case $x_2 = 0, x_1 > 0$ follows then by changing the roles of x_1 and x_2, as well as γ_1 and γ_2. For $x_2 = 0$ the Fichera function simplifies to

$$\lim_{x_2 \to 0^+} b(x_1, x_2) = \left[a_1 + a_2 x_1 - \sigma_1^2 \gamma_1 x_1^{2\gamma_1 - 1} - \frac{1}{2} \rho \sigma_1 x_1^{\gamma_1} \sigma_2 \gamma_2 0^{\gamma_2 - 1} \right] \frac{x_1}{\sqrt{1 + x_1^2}}$$

$$= \begin{cases} \left[a_1 + a_2 x_1 - \sigma_1^2 \gamma_1 x_1^{2\gamma_1 - 1} \right] \frac{x_1}{\sqrt{1+x_1^2}} & \rho = 0 \\ -\infty & 0 < \gamma_2 < 1, \rho \neq 0 \\ \left[a_1 + a_2 x_1 - \sigma_1^2 \gamma_1 x_1^{2\gamma_1 - 1} - \frac{1}{2} \rho \sigma_1 x_1^{\gamma_1} \sigma_2 \right] \frac{x_1}{\sqrt{1+x_1^2}} & \gamma_2 = 1, \rho \neq 0 \\ \left[a_1 + a_2 x_1 - \sigma_1^2 \gamma_1 x_1^{2\gamma_1 - 1} \right] \frac{x_1}{\sqrt{1+x_1^2}} & \gamma_2 > 1, \rho \neq 0 \end{cases}$$

For $0 < \gamma_2 < 1$ and $\rho \neq 0$, BCs are needed, if ρ is positive, and BCs must not be posed, if ρ is negative. In all other cases, the sign of b, which defines whether BCs must be supplied or not, depends on a_1, a_2, σ_1, σ_2 and γ_1, see Fig. 8.1.

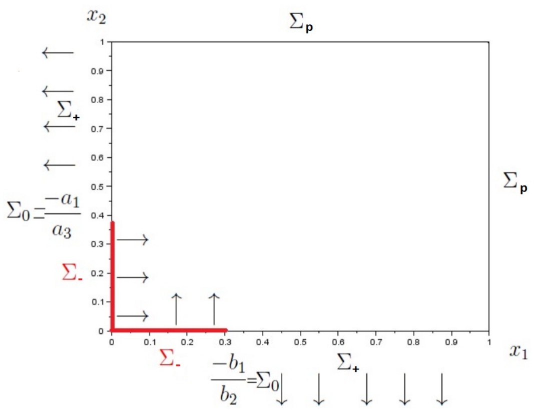


Figure 8.1: Boundary decomposition in two-factor CIR model.

8.4 Numerical Results

Choosing set of parameters $\kappa = 0.5$, $\theta = 0.05$, $\sigma = 0.1$, $\gamma = 0.5$ (CIR), we get at $r = 0$ a positive Fichera function $b = \kappa\theta - \sigma^2/2 = 0.02 > 0$. This is equivalent with the statement that the Feller condition is satisfied. According to the Fichera theory, as soon as it is outflow part of boundary, we must not supply BCs. In this example in Fig. 8.2 and Fig. 8.4 and Table 8.1, we intentionally supplied BCs in an 'outflow' situation when we should not in order to illustrate what might happen if one disregards the Fichera theory. In the evolution of the solution we can observe a peak and oscillations close to the boundary. In Fig. 8.4 we plot the relative error, which is reported also in the Table 8.1.

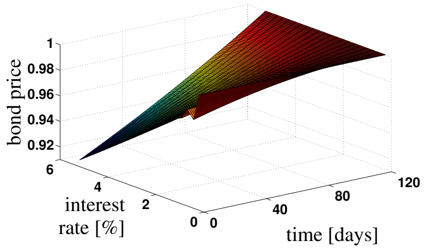


Figure 8.2: Numerical solution, Dirichlet BC

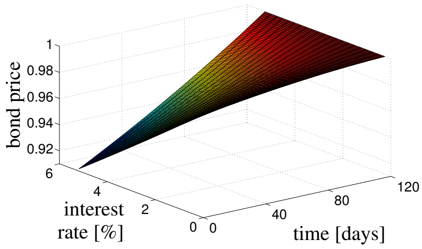


Figure 8.3: Numerical solution, without BC

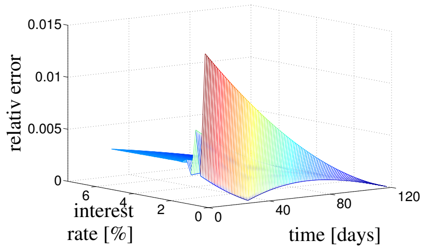


Figure 8.4: Relative error, case with Dirichlet BC

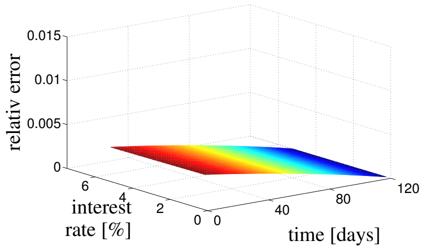


Figure 8.5: Relative error, case without BC

Table 8.1: Relative error, case with BC

time[days]	relative error
1	0.0147
40	0.0079
80	0.0029
120 (maturity)	0

Table 8.2: Relative error, case without BC

time[days]	relative error
1	0.0039
40	0.0029
80	0.0015
120 (maturity)	0

In our example we used the same parameters, but with or without defining Dirichlet BC. Here, “without BC” means that we used for the numerical BC the limit of the interior PDE for $r \to 0$. The corresponding results are shown on the right hand side, in Fig. 8.3, Fig. 8.5 and the relative errors are recorded in Table 8.1.

For the numerical solution we used the implicit finite difference method from [26]. The reference solution is obtained either as the analytic solution for the CIR model ($\gamma = 0.5$, if Feller condition is satisfied), c.f. [7] or in all other cases using a very fine resolution (and suitable BCs). The conditions at outflow boundaries are obtained by studying the limiting behavior of the interior PDE or simply by horizontal extrapolation of appropriate order. Recall that negative values of the Fichera function (i.e. an inflow boundary) corresponds to a not satisfied Feller condition and may destroy the uniqueness of solutions to the PDE.

9 Chapter 9

Alternating Direction Explicit Methods for one-dimensional Convection Diffusion Equations

Numerical analysis for one dimensional convection-diffusion-reaction equation: stability, consistency and convergence results.

- *Z. Bučková, M. Ehrhardt, M. Günther: Alternating Direction Explicit Methods for Convection Diffusion Equations, Acta Math. Univ. Comenianae, Vol. LXXXI: 309–325, 2015*

In this chapter we investigate the stability and consistency properties of *alternating direction explicit* (ADE) finite difference schemes applied to convection-diffusion-reaction equations. Employing different discretization strategies of the convection term we obtain various ADE schemes and study their stability and consistency properties. An ADE scheme consists of two sub steps (called upward and downward sweeps) where already computed values at the new time level are used in the discretization stencil. For linear convection-diffusion-reaction equations the consistency of the single sweeps is of order $O\big(k^2+h^2+k/h\big)$, but the average of these two sweeps has a consistency of order (k^2+h^2), where k, h denote the step size in time and space.

The structure of this chapter is as follows: In Section 7.3.2 we present the considered PDEs and explain the basic idea of the ADE scheme and its modified difference quotients. Next, the numerical analysis studying stability and consistency of the method is presented in Sections 9.1 and 9.2, respectively.

9.0.1 The modified difference quotients for the ADE method

In this subsection we want to illustrate the outcome of the previous Section 7.3.1. Thus, we select some spatial discretization and investigate which ADE scheme will result.

For the discretization of the *diffusion term* we use, c.f. [47]

$$\begin{aligned} \frac{\partial^2 v(x_j,t_n)}{\partial x^2} &\approx \frac{u^n_{j+1}-u^n_j-u^{n+1}_j+u^{n+1}_{j-1}}{h^2}, & j&=1,\dots,J-1 \\ \frac{\partial^2 v(x_j,t_n)}{\partial x^2} &\approx \frac{d^{n+1}_{j+1}-d^{n+1}_j-d^n_j+d^n_{j-1}}{h^2}, & j&=J-1,\dots,1. \end{aligned} \tag{9.1}$$

In order to obtain a symmetric scheme we use the following approximations of the *reaction term*, the same for the upward and downward sweep

$$\begin{aligned} v(x_j,t_n) &\approx \frac{u_j^{n+1}+u_j^n}{2}, && j=1,\dots,J-1, \\ v(x_j,t_n) &\approx \frac{d_j^{n+1}+d_j^n}{2}, && j=J-1,\dots,1. \end{aligned} \tag{9.2}$$

Different approximations of the *convection term* are possible [38], [15]. In the following we state three of them. First, Towler and Yang [55] used special kind of centered differences

$$\begin{aligned} \frac{\partial v(x_j,t_n)}{\partial x} &\approx \frac{u_{j+1}^n-u_{j-1}^{n+1}}{2h}, && j=1,\dots,J-1, \\ \frac{\partial v(x_j,t_n)}{\partial x} &\approx \frac{d_{j+1}^{n+1}-d_{j-1}^n}{2h}, && j=J-1,\dots,1. \end{aligned} \tag{9.3}$$

More accurate approximations were proposed by Roberts and Weiss [45], Piacsek and Williams [43]

$$\begin{aligned} \frac{\partial v(x_j,t_n)}{\partial x} &\approx \frac{u_{j+1}^n-u_j^n+u_j^{n+1}-u_{j-1}^{n+1}}{2h}, && j=1,\dots,J-1, \\ \frac{\partial v(x_j,t_n)}{\partial x} &\approx \frac{d_{j+1}^{n+1}-d_j^{n+1}+d_j^n-d_{j-1}^n}{2h}, && j=J-1,\dots,1. \end{aligned} \tag{9.4}$$

As a third option we will use *upwind* approximations combined with the ADE technique. Since we have in mind financial applications we will focus on left going waves, i.e. $b>0$ in (7.11). Right going waves $b<0$ are treated analogously.

The well-known first order approximation reads

$$\frac{\partial v(x_j,t)}{\partial x} \approx \frac{v_{j+1}(t)-v_j(t)}{h} \qquad j=J-1,\dots,1, \tag{9.5}$$

and the forward difference of second order [59]

$$\frac{\partial v(x_j,t)}{\partial x} \approx \frac{-v_{j+2}(t)+4v_{j+1}(t)-3v_j(t)}{2h}, \qquad j=J-1,\dots,1. \tag{9.6}$$

Applying the ADE time splitting idea of Section 7.3.1 we obtain for the upwind strategy (9.5)

$$\begin{aligned} \frac{\partial v(x_j,t_{n+1})}{\partial x} &\approx \frac{u_{j+1}^n-u_j^n}{h}, && j=1,\dots,J-1, \\ \frac{\partial v(x_j,t_{n+1})}{\partial x} &\approx \frac{d_{j+1}^n-d_j^n+d_{j+1}^{n+1}-d_j^{n+1}}{2h}, && j=J-1,\dots,1, \end{aligned} \tag{9.7}$$

and for the second order approximation

$$\frac{\partial v(x_j,t_{n+1})}{\partial x} \approx \frac{-u^n_{j+2}+4u^n_{j+1}-3u^n_j}{2h}, \qquad j=1,\dots,J-1,$$
$$\frac{\partial v(x_j,t_{n+1})}{\partial x} \approx \frac{-d^n_{j+2}+4d^n_{j+1}-3d^n_j-d^{n+1}_{j+2}+4d^{n+1}_{j+1}-3d^{n+1}_j}{4h}, \quad j=J-1,\dots,1. \tag{9.8}$$

We will show that this upwind approximation (9.7) leads to a stable scheme.

9.1 Stability of the ADE method

In this section we investigate the stability of the proposed ADE method using the matrix approach in Section 9.1.1 and the classical von-Neumann method in Section 9.1.2. For the convection-diffusion-reaction equation (7.11) we obtain unconditional stability using the matrix approach. This stability analysis can be extended by adding homogeneous BCs, without affecting the stability results. This is our motivation to deal with the matrix approach.

9.1.1 Stability analysis using the Matrix approach

We are motivated by [38], where the authors claim and proof that "if A is symmetric negative definite, the ADE scheme is unconditionally stable". We have to define symmetric discretization quotients to get symmetric discrete operators. For reaction-diffusion equation applying central difference quotients we get symmetric operator A and we can follow the ideas for the proof for the heat equation from [38].

Using upwind discretization formulas instead of central differencing leads also to an unconditionally stable scheme. "If A is lower-triangular with all diagonal elements negative, the ADE scheme is unconditionally stable" is generally claimed and proved in [38]. In the following we choose suitable differentiating approximations, we formulate theorems about stability properties and prove it.

Theorem 9.1. *The ADE scheme applied to the reaction-diffusion PDE* (7.11) *(with $b=0$) is unconditionally stable.*

Proof. Without loss of generality we focus on the upward sweep

$$\frac{u^{n+1}_j-u^n_j}{k} = a\frac{u^n_{j+1}-u^n_j-u^{n+1}_j+u^{n+1}_{j-1}}{h^2} - c\frac{u^{n+1}_j+u^n_j}{2}.$$

Let us denote the parabolic mesh ratio $\alpha := a\frac{k}{h^2}$, $\gamma := ck$; where a, c are constants.

$$u_j^{n+1} = u_j^n + \alpha\left(u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}\right) - \frac{\gamma}{2}\left(u_j^{n+1} + u_j^n\right)$$
$$\left(1+\alpha+\frac{\gamma}{2}\right)u_j^{n+1} + (-\alpha)u_{j-1}^{n+1} = \left(1-\alpha-\frac{\gamma}{2}\right)u_j^n + \alpha u_{j+1}^n \quad (9.9)$$

We follow roughly the train of thoughts of Leung and Osher [38] and write the upward sweep (9.9) with homogeneous BCs in matrix notation

$$A_u u^{n+1} = B_u u^n, \qquad n \geq 0,$$

with $A_u, B_u \in R^{(J-1)\times(J-1)}$ given by

$$A_u = \begin{pmatrix} 1+\alpha+\frac{\gamma}{2} & 0 & \dots & 0 \\ -\alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & -\alpha & 1+\alpha+\frac{\gamma}{2} \end{pmatrix} = I + \begin{pmatrix} \alpha+\frac{\gamma}{2} & 0 & \dots & 0 \\ -\alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & -\alpha & \alpha+\frac{\gamma}{2} \end{pmatrix}$$
$$A_u =: I + E,$$
$$B_u = \begin{pmatrix} 1-\alpha-\frac{\gamma}{2} & \alpha & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \alpha \\ 0 & \dots & 0 & 1-\alpha-\frac{\gamma}{2} \end{pmatrix} = I - \begin{pmatrix} \alpha+\frac{\gamma}{2} & -\alpha & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha \\ 0 & \dots & 0 & \alpha+\frac{\gamma}{2} \end{pmatrix}$$
$$B_u =: I - E^\top.$$

Next, we consider the matrices

$$A_u^\top + A_u = 2I + D,$$

$$\text{where} \qquad D := E + E^\top = \begin{pmatrix} 2\alpha+\gamma & -\alpha & \dots & 0 \\ -\alpha & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha \\ 0 & \dots & -\alpha & 2\alpha+\gamma \end{pmatrix}.$$

The matrix D is positive definite and thus we can define the induced D-norm as

$$||C||_D^2 := \sup_{x\neq 0} \frac{||C_x||_D^2}{||x||_D^2} = \sup_{x\neq 0} \frac{x^\top C^\top DCx}{x^\top Dx},$$

and the upward sweep can be written as

$$U^{n+1} = A_u^{-1} B_u U^n.$$

Next, we consider the D-norm for the upward sweep matrix $A_u^{-1}B_u$

$$||A_u^{-1}B_u||_D^2 := \sup_{x\neq 0} \frac{x^\top B_u^\top A_u^{-\top} D A_u^{-1} B_u x}{x^\top Dx}$$

The numerator $B_u^\top A_u^{-\top} D A_u^{-1} B_u$ can be easily rewritten after a few algebraic steps as $D - 2\gamma(A_u^{-1}D)^\top(A_u^{-1}D)$. From our notation $A_u = I + E$ and $B_u = I - E^\top$ follows

$$B_u^\top A_u^{-\top} D A_u^{-1} B_u = (I - E^\top)^\top A_u^{-\top} D A_u^{-1}(I - E^\top)$$

where $E^\top = D - E$. An expression in terms of matrices A_u and D gets the following form:

$$\begin{aligned}
&(A_u^\top - D)^\top A_u^{-\top} D A_u^{-1}(A_u^\top - D) \\
&\quad = D - DA_u^{-1}D - D^\top A_u^{-\top} D + DA_u^{-\top} D A_u^{-1} D \\
&\quad = D - DA_u^{-\top} A_u^\top A_u^{-1} D - DA_u^{-\top} A_u A_u^{-1} D + DA_u^{-\top} D A_u^{-1} D \\
&\quad = D + DA_u^{-\top}\left[-A_u^{-\top} - A_u + D\right] A_u^{-1} D \\
&\quad = D - 2(A_u^{-1}D)^\top (A_u^{-1}D)
\end{aligned}$$

and hence it follows

$$||A_u^{-1}B_u||_D^2 = 1 - 2\sup_{x \neq 0} \frac{||A_u^{-1}Dx||_2^2}{||x||_D^2}.$$

Thus the spectral radius of the upward sweep matrix $A_u^{-1}B_u$ reads

$$\rho(A_u^{-1}B_u) \leq ||A_u^{-1}B_u||_D < 1$$

and we can conclude that the upward sweep is unconditionally stable.

An analogous result holds for the downward step. In the corresponding equation

$$A_d d^{n+1} = B_d d^n, \qquad n \geq 0 \tag{9.10}$$

the matrices A_d and B_d are defined as $A_d = A_u^\top$ and $B_d = B_u^\top$. The analysis is done analogously: we can define a positive definite matrix and follow again the steps from the previous proof of the Theorem 9.1. Consequently also the combination, as an arithmetic average of these two sub steps, is also unconditionally stable. □

The stability analysis using the matrix approach according to [38] worked for reaction-diffusion equations with constant coefficients. However, this proof is not transferable for the stability analysis of methods with non-symmetric terms, e.g. the difference quotients for the convection term proposed by Towler and Yang (eq. 9.3), or Roberts and Weiss (eq. 9.4), c.f. Section 9.1.2.

As a remedy we can apply a modified upwind discretization of the convection term. The resulting structure of the matrices A_u, B_u is different but we can perform a similar proof.

Theorem 9.2. *ADE scheme, using upwind discretization in convection term, applied to the reaction-diffusion-convection equation* (7.11) *is unconditionally stable in the upward sweep and unconditionally stable in the downward one.*

Proof. Again, without loss of generality, we focus on the upward sweep and consider an upwind discretization for a left-going wave, i.e. $b \geq 0$ (since later we would like to extend this approach for the Black-Scholes model, where $b \geq 0$). In the upward sweep we use difference quotients using values just from the old time level (9.5)

$$\frac{u_j^{n+1}-u_j^n}{k} = a\frac{u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}}{h^2} + b\frac{u_{j+1}^n - u_j^n}{h} - c\frac{u_j^{n+1}+u_j^n}{2}.$$

Using the abbreviations $\alpha := a\frac{k}{h^2}$, $\beta := b\frac{k}{h} \geq 0$, $\gamma := ck$, we can write

$$-\alpha u_{j-1}^{n+1} + \Big(1+\alpha+\frac{\gamma}{2}\Big)u_j^{n+1} = \Big(1-\alpha-\beta-\frac{\gamma}{2}\Big)u_j^n + \Big(\alpha+\beta\Big)u_{j+1}^n \tag{9.11}$$

We follow again roughly the ideas of Leung and Osher [38] and consider the upward sweep (9.11) with homogeneous BCs

$$A_u u^{n+1} = B_u u^n, \qquad n \geq 0,$$

with the system matrices $A_u, B_u \in R^{(J-1)\times(J-1)}$ given by

$$\begin{aligned} A_u &= \begin{pmatrix} 1+\alpha+\frac{\gamma}{2} & 0 & \dots & 0 \\ -\alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & -\alpha & 1+\alpha+\frac{\gamma}{2} \end{pmatrix} \\ &= I + \begin{pmatrix} \alpha+\frac{\gamma}{2} & 0 & \dots & 0 \\ -\alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & -\alpha & \alpha+\frac{\gamma}{2} \end{pmatrix} =: I+E, \end{aligned}$$

$$\begin{aligned} B_u &= \begin{pmatrix} 1-\alpha-\beta-\frac{\gamma}{2} & \alpha+\beta & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \alpha+\beta \\ 0 & \dots & 0 & 1-\alpha-\beta-\frac{\gamma}{2} \end{pmatrix} \\ &= I - \begin{pmatrix} \alpha+\beta+\frac{\gamma}{2} & -\alpha-\beta & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha-\beta \\ 0 & \dots & 0 & \alpha+\beta+\frac{\gamma}{2} \end{pmatrix} =: I-F. \end{aligned}$$

$$\text{where} \qquad D := E+F = \begin{pmatrix} 2\alpha+\beta+\gamma & -\alpha-\beta & \dots & 0 \\ -\alpha & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha-\beta \\ 0 & \dots & -\alpha & 2\alpha+\beta+\gamma \end{pmatrix}.$$

The matrix D is not symmetric but obviously positive definite.

In the sequel we have just outlined the steps which differ from the previous proof. The numerator $B_u^\top A_u^{-\top} D A_u^{-1} B_u$ can be easily rewritten after a few algebraic steps as $D-$

$2\gamma(A_u^{-1}D)^\top(A_u^{-1}D)$.

From our notation $A_u = I + E$ and $B_u = I - F$ follows

$$B_u^\top A_u^{-\top} D A_u^{-1} B_u = (I-F)^\top A_u^{-\top} D A_u^{-1}(I-F)$$

where $F := D - E$. An expression in terms of matrices A_u and D gets the following form:

$$(I+E-D)^\top A_u^{-\top} D A_u^{-1}(I+E-D) = (A_u - D)^\top A_u^{-\top} D A_u^{-1}(A_u - D)$$

and we proceed the same way as in the previous proof.

For the *downward sweep* we have:

$$\frac{d_j^{n+1}-d_j^n}{k} = a\frac{d_{j+1}^{n+1}-d_j^{n+1}-d_j^n+d_{j-1}^n}{h^2} + b\frac{d_{j+1}^n+d_{j+1}^{n+1}-d_j^n-d_j^{n+1}}{2h} - c\frac{d_j^{n+1}+d_j^n}{2}.$$

Using the abbreviations $\alpha := a\frac{k}{h^2}$, $\beta := b\frac{k}{h} \le 0$, $\gamma := ck$, we can write

$$\Big(1+\alpha+\frac{\beta}{2}+\frac{\gamma}{2}\Big)d_j^{n+1} + \Big(-\alpha-\frac{\beta}{2}\Big)d_{j+1}^{n+1} = \alpha d_{j-1}^n + \Big(1-\alpha-\frac{\beta}{2}-\frac{\gamma}{2}\Big)d_j^n + \frac{\beta}{2}d_{j+1}^n \tag{9.12}$$

$$A_D d^{n+1} = B_D d^n, \qquad n \ge 0,$$

with A_D, $B_D \in R^{(J-1)\times(J-1)}$ given by matrices A_D, B_D. The matrix A_D is upper-diagonal $A_D = \mathrm{diag}(1+\alpha+\frac{\beta}{2}+\frac{\gamma}{2}, -\alpha-\frac{\beta}{2})$. The matrix B_D is tridiagonal with diagonal terms: $B_D = \mathrm{diag}(\alpha, 1-\alpha-\frac{\beta}{2}-\frac{\gamma}{2}, \frac{\beta}{2})$. Likewise we construct matrices $D = \mathrm{diag}(-\alpha, 2\alpha+\beta+\gamma, -\alpha-\beta)$ as a tridiagonal positive definite matrix. We can follow the same way of proof and thus we conclude the unconditional stability of the downward sweep. □

9.1.2 Von Neumann stability analysis for the convection-diffusion-reaction equation

Since analysis using matrix approach was suitable for upwind kind of approximation in convection term, here we investigate stability properties of the ADE schemes, where discretization of convection term is provided according to [55] and [45].

We consider the convection-diffusion-reaction equation (7.11) and focus on the sequel on the upward sweep of the ADE procedure. An appropriate choice for the approximation of the convection term is the one due to Roberts and Weiss [45], since performing just a downward sweep leads to the unconditionally stable solution.

Theorem 9.3. *The ADE scheme with the Roberts and Weiss approximation in the convection term, applied to the PDE* (7.11) *is conditionally stable in the upward sweep and unconditionally stable for the downward one.*

Proof. Using Roberts and Weiss discretization in convection term we get

$$\frac{u_j^{n+1}-u_j^n}{k} = a\frac{u_{j+1}^n-u_j^n-u_j^{n+1}+u_{j-1}^{n+1}}{h^2} + b\frac{u_{j+1}^n-u_j^n+u_j^{n+1}-u_{j-1}^{n+1}}{2h} - c\frac{u_j^{n+1}+u_j^n}{2}.$$

Let us denote the parabolic mesh ratio $\alpha := a\frac{k}{h^2}$, the hyperbolic mesh ratio $\beta := b\frac{k}{h}$ and $\gamma := ck$; where a, b, c are nonnegative constants.

$$u_j^{n+1} = u_j^n + \alpha\left(u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}\right) + \frac{\beta}{2}\left(u_{j+1}^n - u_j^n + u_j^{n+1} - u_{j-1}^{n+1}\right) - \frac{\gamma}{2}\left(u_j^{n+1} + u_j^n\right)$$

Applying von Neumann ansatz $u_j^n := e^{\xi tn}e^{i\lambda xj}$ the amplification factor A_1 reads:

$$A_1 = \frac{A+Be^{i\lambda h}}{C+De^{-i\lambda h}}$$

where $A = 1-\alpha-\beta/2-\gamma/2$; $B = \alpha+\beta/2$; $C = 1+\alpha-\beta/2+\gamma/2$; $D = -\alpha+\beta/2$.

For stability we require $|A_1| \le 1$, i.e.

$$|A_1|^2 = A_1\overline{A_1} = \frac{\left(A+Be^{i\lambda h}\right)\left(A+Be^{-i\lambda h}\right)}{\left(C+De^{-i\lambda h}\right)\left(C+De^{i\lambda h}\right)} \le 1$$
$$A^2+B^2+2AB\cos(\lambda h) \le C^2+D^2+2CD\cos(\lambda h)$$
$$2(AB-CD)\cos(\lambda h) \le C^2+D^2-A^2-B^2$$

$$(4\alpha-4\alpha\beta-\beta\gamma)\cos(\lambda h) \le 4\alpha-4\alpha\beta-\beta\gamma+2\gamma. \tag{9.13}$$

We need to check two cases with respect to the sign of $(4\alpha-4\alpha\beta-\beta\gamma)$.

- **Case 1:** By substituting α, β, γ into $4\alpha-4\alpha\beta-\beta\gamma > 0$ we get following condition:

$$\alpha < \frac{a}{2Pe} - \frac{ck}{4} \tag{9.14}$$

  where $Pe = \frac{bh}{2}$ is the so-called Peclet number. In this case equation (9.13) can be rewritten as

$$\cos(\lambda h) \le \frac{4\alpha-4\alpha\beta-\beta\gamma+2\gamma}{4\alpha-4\alpha\beta-\beta\gamma} \qquad \forall\lambda h \tag{9.15}$$

  i.e.

$$1 \le 1+\frac{2\gamma}{4\alpha-4\alpha\beta-\beta\gamma}$$

or

$$0 \le \frac{2\gamma}{4\alpha - 4\alpha\beta - \beta\gamma}. \tag{9.16}$$

We can notice that condition (9.16) is satisfied for all the possible values of parameters, since $\gamma > 0$ and $4\alpha - 4\alpha\beta - \beta\gamma > 0$.

- **Case 2:** We consider $4\alpha - 4\alpha\beta - \beta\gamma < 0$, what is equivalent with the condition

$$\alpha > \frac{a}{2Pe} - \frac{ck}{4}. \tag{9.17}$$

In this case equation (9.13) can be rewritten as

$$\cos(\lambda h) \ge \frac{4\alpha - 4\alpha\beta - \beta\gamma + 2\gamma}{4\alpha - 4\alpha\beta - \beta\gamma} \qquad \forall \lambda h \tag{9.18}$$

i.e.

$$-1 \ge 1 + \frac{2\gamma}{4\alpha - 4\alpha\beta - \beta\gamma},$$

or

$$2 \le \frac{-2\gamma}{4\alpha - 4\alpha\beta - \beta\gamma} \tag{9.19}$$

or

$$\frac{1}{2}\beta\gamma + 2\alpha\beta - 2\alpha \le \gamma \tag{9.20}$$

or

$$\alpha \le \frac{\alpha}{\beta} - \frac{\gamma}{4} + \frac{\gamma}{2\beta} \tag{9.21}$$

After substituting α, β, γ and after elementary algebraic steps we get

$$\alpha \le \frac{a}{2Pe} - \frac{ck}{4} + \frac{ch}{4b}. \tag{9.22}$$

Case 2 leads to conditions (9.17) and (9.22) what means that

$$\alpha \in \left(\frac{a}{2Pe} - \frac{ck}{4}, \frac{a}{2Pe} - \frac{ck}{4} + \frac{ch}{4b} \right] \tag{9.23}$$

To sum up case 1 and case 2 we can claim that conditions (9.14) and (9.23) and also considering the situation where $(4\alpha - 4\alpha\beta - \beta\gamma) = 0$ we get

$$\alpha \le \frac{a}{2Pe} - \frac{ck}{4} + \frac{ch}{4b}. \tag{9.24}$$

For the downward sweep we get the following amplification factor:

$$A_2 = \frac{\left[1-\alpha+\frac{\beta}{2}-\frac{\gamma}{2}\right]+[\alpha-\frac{\beta}{2}]e^{-i\lambda h}}{\left[1+\alpha+\frac{\beta}{2}+\frac{\gamma}{2}\right]+[-\alpha-\frac{\beta}{2}]e^{i\lambda h}} \tag{9.25}$$

Stability condition $|A_2|^2 \leq 1$ leads to the formula:

$$\cos(\lambda h) \leq \frac{4\alpha+4\alpha\beta+\beta\gamma+2\gamma}{4\alpha+4\alpha\beta+\beta\gamma}.$$

Let us note that the last condition can be simplified to the condition:

$$\frac{2\gamma}{4\alpha+4\alpha\beta+\beta\gamma} \geq 0. \tag{9.26}$$

The coefficients α, β, γ are positive, i.e. the condition (9.26) is satisfied and thus we have the unconditional stability for the downward sweep using the Roberts and Weiss approximation, which completes the proof. □

In case of the Roberts and Weiss approximation we propose to use only the unconditional stable downward sweep.

Theorem 9.4. *ADE scheme, using Towler and Yang approximation in the convection term, applied to the PDE* (7.11) *is conditionally stable in both sweeps.*

Proof. For the Towler and Yang approximation the stability condition for the upward sweep reads

$$(4\alpha-2\alpha\beta-\beta\gamma)\cos(\lambda h) \leq 4\alpha-2\alpha\beta+2\gamma, \tag{9.27}$$

where again we can distinguish 2 cases with respect to the sign of left hand side of the equation (9.27).

- **Case 1:** If $(4\alpha-2\alpha\beta-\beta\gamma) > 0$, it means

$$\alpha < \frac{a}{Pe} - \frac{ck}{2}. \tag{9.28}$$

  In this case equation (9.27) can be rewritten as

$$\cos(\lambda h) \leq \frac{4\alpha-4\alpha\beta-\beta\gamma+2\gamma}{4\alpha-2\alpha\beta-\beta\gamma} \tag{9.29}$$

$$1 \leq 1 + \frac{\gamma(2+\beta)}{4\alpha-2\alpha\beta-\beta\gamma}$$

$$0 \leq \frac{\gamma(2+\beta)}{4\alpha - 2\alpha\beta - \beta\gamma}. \tag{9.30}$$

We can notice that condition (9.30) is satisfied for all the possible values of parameters, since $\gamma \geq 0$ and $(2+\beta) > 0$ and $4\alpha - 2\alpha\beta - \beta\gamma > 0$.

- **Case 2:** We consider $(4\alpha - 2\alpha\beta - \beta\gamma) < 0$, what is equivalent with the condition

$$\alpha > \frac{a}{Pe} - \frac{ck}{2}. \tag{9.31}$$

In this case equation (9.27) can be rewritten as

$$\cos(\lambda h) \geq \frac{4\alpha - 4\alpha\beta - \beta\gamma + 2\gamma}{4\alpha - 2\alpha\beta - \beta\gamma} \tag{9.32}$$

$$-2 \geq \frac{\gamma(2+\beta)}{4\alpha - 2\alpha\beta - \beta\gamma}$$

After substituting α, β, γ and simplification it leads to the condition

$$\alpha \leq \frac{a}{Pe} - \frac{ck}{2} + \frac{ch}{2b} + \frac{1}{2}. \tag{9.33}$$

In case 2 we obtain two conditions (9.31) and (9.33), namely:

$$\alpha \in \left(\frac{a}{2Pe} - \frac{ck}{2}, \frac{a}{Pe} - \frac{ck}{2} + \frac{ch}{2b} + \frac{1}{2} \right] \tag{9.34}$$

From case 1 condition (9.28) and case 2 condition (9.34) in Towler and Yang case and considering also possibility of $(4\alpha - 2\alpha\beta - \beta\gamma) = 0$ we can sum up

$$\alpha \leq \frac{a}{Pe} - \frac{ck}{2} + \frac{ch}{2b} + \frac{1}{2}. \tag{9.35}$$

For the downward sweep the stability condition is

$$\cos(\lambda h) \leq \frac{4\alpha + 2\alpha\beta + 2\gamma}{4\alpha + 2\alpha\beta + \beta\gamma},$$

which leads to the condition:

$$\frac{k}{h^2} \leq \frac{1}{Pe} \tag{9.36}$$

Both sweeps in Towler and Yang discretization of convection term in reaction-diffusion-convection equation are conditionally stable under the conditions (9.35) and (9.36) □

9.2 Consistency Analysis of the ADE methods

In this section we provide a consistency analysis of the ADE methods for solving the convection-diffusion-reaction equation (7.11) and for the BS model.

9.2.1 Consistency of the ADE scheme for convection-diffusion-reaction equations

We study the consistency of the following ADE discretization

$$\frac{u_j^{n+1}-u_j^n}{k} = a\frac{u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}}{h^2} + b\frac{u_{j+1}^n - u_j^n + u_j^{n+1} - u_{j-1}^{n+1}}{2h} - c\frac{u_j^{n+1}+u_j^n}{2}$$

to the convection-diffusion-reaction equation (7.11). The local truncation error (LTE) of the upward sweep is given by

$$\begin{aligned} LTE_{\text{up}} = {} & k\Big(-\frac{1}{2}v_{tt} + \frac{1}{2}av_{xxt} + \frac{1}{2}bv_{xt}\Big) \\ & + k^2\Big(-\frac{1}{6}v_{ttt} + \frac{1}{4}av_{xxtt} + \frac{1}{4}bv_{xtt}\Big) + h^2\Big(\frac{1}{12}av_{xxxx}\frac{1}{6}bv_{xxx}\Big) \\ & - kh\Big(\frac{1}{6}av_{xxxt} + \frac{1}{4}bv_{xxt}\Big) - \frac{k}{h}av_{xt} - \frac{k^2}{h}\Big(\frac{1}{2}av_{xtt}\Big) - \frac{k^3}{h}\Big(\frac{1}{6}av_{xttt}\Big), \end{aligned}$$

and analogously the LTE for the downward sweep reads

$$\begin{aligned} LTE_{\text{down}} = {} & k\Big(-\frac{1}{2}v_{tt} + \frac{1}{2}av_{xxt} + \frac{1}{2}bv_{xt}\Big) \\ & + k^2\Big(-\frac{1}{6}v_{ttt} + \frac{1}{4}av_{xxtt} + \frac{1}{4}bv_{xtt}\Big) + h^2\Big(\frac{1}{12}av_{xxxx}\frac{1}{6}bv_{xxx}\Big) \\ & + kh\Big(\frac{1}{6}av_{xxxt} + \frac{1}{4}bv_{xxt}\Big) + \frac{k}{h}av_{xt} + \frac{k^2}{h}\Big(\frac{1}{2}av_{xtt}\Big) + \frac{k^3}{h}\Big(\frac{1}{6}av_{xttt}\Big). \end{aligned}$$

Thus we end up for the LTE for the combined sweep

$$LTE_{\text{ADE}} = k\Big(-\frac{1}{2}v_{tt} + \frac{1}{2}av_{xxt} + \frac{1}{2}bv_{xt}\Big) + k^2\Big(-\frac{1}{6}v_{ttt} + \frac{1}{4}av_{xxtt} + \frac{1}{4}bv_{xtt}\Big) + h^2\Big(\frac{1}{12}av_{xxxx}\frac{1}{6}bv_{xxx}\Big)$$

Assuming a constant parabolic mesh ratio k/h^2, the first order term in k can be written in the form $\mathrm{O}(k) = \mathrm{O}(h^2)$ and hence we get

$$LTE_{\text{ADE}} = k^2\Big(-\frac{1}{6}v_{ttt} + \frac{1}{4}av_{xxtt} + \frac{1}{4}bv_{xtt}\Big) + h^2\Big(\frac{1}{12}av_{xxxx}\frac{1}{6}bv_{xxx} - \frac{1}{2}v_{tt} + \frac{1}{2}av_{xxt} + \frac{1}{2}bv_{xt}\Big)$$

Hence, the order of consistency of the ADE method for the PDE (7.11) is $\mathrm{O}(k^2+h^2)$.

9.2.2 The Consistency of the ADE method for the linear BS model

As an extension of the PDE (7.11) we consider now the linear BS equation.

Theorem 9.5. *The order of consistency of the ADE method for the linear BS equation is* $O(k^2+h^2)$ *in both sweeps and in the final combined solution.*

Proof. The linear BS PDE is a special case of (7.11) with the space-dependent coefficients $a(S)=\frac{1}{2}\sigma^2 S^2$, $b(S)=rS$, $c(S)=r$. The LTE for the upward sweep reads:

$$\begin{aligned} LTE_{\mathrm{BS}} &= k\Big(-\frac{1}{2}v_{tt}+\frac{1}{2}av_{xxt}+\frac{1}{2}bv_{xt}\Big)\\ &+k^2\Big(-\frac{1}{6}v_{ttt}+\frac{1}{4}av_{xxtt}+\frac{1}{4}bv_{xtt}\Big)+h^2\Big(\frac{1}{12}av_{xxxx}\frac{1}{6}bv_{xxx}\Big)\\ &+kh\Big(-\frac{1}{6}av_{xxxt}-\frac{1}{4}bv_{xxt}\Big)+\frac{k}{h}\Big(-av_{xt}\Big)\\ &+\frac{k^2}{h}\Big(-\frac{1}{2}av_{xtt}\Big)+\frac{k^3}{h}\Big(-\frac{1}{6}av_{xttt}\Big) \end{aligned}$$

If we assume a constant parabolic mesh ratio $\alpha=k/h^2$, then we get

$$LTE=k\Big(-\frac{1}{2}v_{tt}\Big)+k^2\Big(-\frac{1}{6}v_{ttt}\Big)=\alpha h^2\Big(-\frac{1}{2}v_{tt}\Big)+k^2\Big(-\frac{1}{6}v_{ttt}\Big),$$

where we neglected higher order terms. A similar result holds for the downward sweep. We have shown that consistency for the linear BS model is $O(k^2+h^2)$ in downward, upward and hence also in the combination. □

9.2.3 Application and numerical experiments with the linear model

We apply the ADE method and calculate a price for a vanilla European call option in a classic linear BS model with constant coefficients. Choosing the following set of parameters $r=0.03$ (interest rate); $q=0$ (continuous dividend yield); $\sigma=0.2$ (volatility); $T=1$ (maturity time in years); $S_{\max}=90$ (maximal stock price); $K=30$ (strike price); and defining a grid with $N=50$ time steps; $J=200$ space steps we get an option price, which is shown in Figure 3.

In this subsection we analyze the computational and theoretical order of convergence. In Table 9.3 it is recorded an error as a difference between numerical solution using ADE method and the closed form BS formula for different meshes with fixed mesh ratio 0.23. In Table 9.4 ratios of errors from the Table 9.3 are calculated. One can observe that using double space steps, ratio of errors converges to the number 4, what confirms that the theoretical order of convergence is 2.

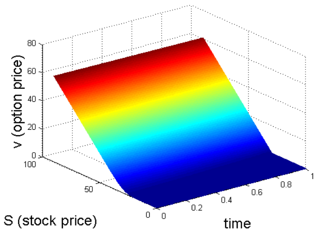


Figure 9.1: Option price

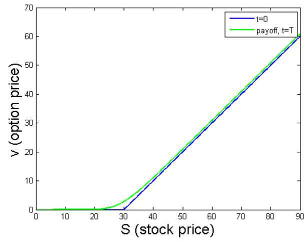


Figure 9.2: Solution at time $t = 0$ and $t = T$

N	J	mesh ratio	error
3	50	0.23	0.2458
12	100	0.23	0.0855
50	200	0.23	0.0208
200	400	0.23	0.0052
800	800	0.23	0.0013

Figure 9.3: Error as a difference between exact solution and approximation

ratio of errors	
error50/error100	2.87
error100/error200	4.11
error200/error400	4
error400/error800	4

Figure 9.4: Ratio of errors

Figures 9.5–9.9 show an error on different grids, as a difference between numerical solution and the exact one (from the BS formula). Table 9.3 records the maximum value of the error from the time $t = 0$, it means that we observe the maximal value of the errors whole calculation in the current time. At the beginning of the calculation (nearby maturity time) we can observe the highest error, which is caused by the non-smooth initial data. This error decreases during the calculation. The finer the mesh, the faster the decrease of the error (9.5)–(9.9).

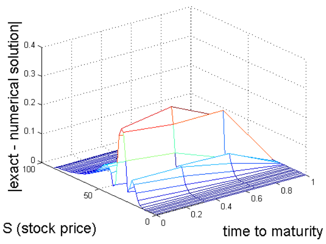


Figure 9.5: Error, $N = 3$, $J = 50$

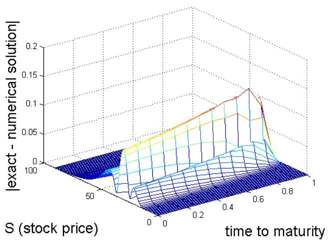


Figure 9.6: Error, $N = 12$, $J = 100$

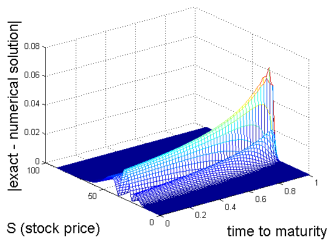


Figure 9.7: Error, $N = 50$, $J = 200$

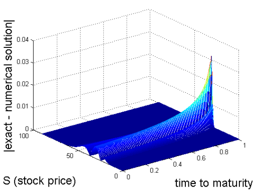


Figure 9.8: Error, $N = 200$, $J = 400$

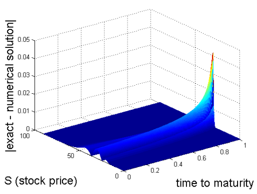


Figure 9.9: Error, $N = 800$, $J = 800$

Chapter 10
Implementation of Alternating Direction Explicit Methods for higher dimensional Black-Scholes Equation

In this work we propose Alternating Direction Explicit (ADE) schemes for the two and three dimensional linear Black-Scholes pricing model. Our implemented methodology can be easily extended to higher dimensions. The main advantage of ADE schemes is that they are explicit and exhibit good stability properties. Results concerning the experimental order of convergence are included.

This Chapter is based on:

- *Z. Bučková, P. Pólvora, M. Ehrhardt, M. Günther: Implementation of Alternating Direction Explicit Methods to higher dimensional Black-Scholes Equation, AIP Conf. Proc. 1773, 030001; 2016*

The Chapter is structured as follows. After the introduction of the ADE schemes and the multi-dimensional Black-Scholes models, we focus on the details of the ADE scheme in the second section. In the third section we introduce the numerical scheme with difference quotients for the ADE. The fourth section consists of the numerical results focusing on the experimental study of convergence for two examples using different payoff structures: two dimensional spread option model and three dimensional call option model. The last section sums up results and presents the outlook.

10.1 ADE Schemes for Multi-Dimensional Models

In this section we introduce the ADE scheme for multi-dimensional PDE models. We first consider the 2D case and then we proceed to higher dimensional cases.

10.1.1 ADE Schemes for Two-dimensional Models

We now explain in detail how to construct the ADE scheme for two-dimensional PDE models, i.e. $N = 2$ in (7.2).

The first key aspect of this scheme is choosing the difference quotients approximating the partial derivatives of our equation in a way that we use the information from both time levels without the need to solve a linear system of equations. In particular, in a two dimensional setting, we would use the points as exemplified in Figure 10.1: we wish to compute the value in black, at time level $n+1$, and we use the information from the neighbor points with an empty filling, from both time level n and $n+1$.

The second key aspect is that in order to improve the accuracy of this scheme, for each time-level two different calculations of the grid points are done using different difference quotients, these are referred to as the *downward sweep* and the *upward sweep*. Then, the solution at that time level is taken as the average of both sweeps. From Figure 10.1 right and Figure 10.2 right the difference between the two sweeps is apparent.

Figure 10.1: Downward sweep. Left figure: time level n. Right figure: time level $n+1$. We depict the spatial grid for two different time-steps. The empty circles represent the points used in the computation of the value at the location of the black circle. S_1 and S_2 denote the spatial dimensions.

Figure 10.2: Upwards sweep. As in Figure 10.1, the empty circles represent the points used in the computation of the value at the location of the black circle. S_1 and S_2 denote the spatial dimensions. Left figure: time level n. Right figure: time level $n+1$.

The final key aspect is that the structure imposed by the stencil illustrated in Figures 10.1 and 10.2 is not by itself enough to guarantee that the scheme is explicit, we must make sure that the empty filling points in the time level $n+1$ have been computed before we compute the black point. This imposes a structure on the algorithm to compute the points as illustrated in Figure 10.3.

For a fixed time level and starting from the boundary we see that in the first step we can only compute the points numbered as 1, since, our stencil is as described in the Figures 10.1 right, 10.2 right. After computing these points we have a total of four points that can be now computed, these are numbered as 2. Hence, we chose any of those points which in turn allows new points to be computed, and so forth. As long as we respect this order, our algorithm is fully explicit.

As we can see in the second step, we have more than one possibility per step as to what point to compute, hence, there are different sequences of points. A natural choice is to choose the sequence of points as shown in Figure 10.4. We called this approach of numbering as a *jumping approach* or *house numbering approach*. We are moving from one corner of the square to another where diagonal points are computed and the others. We could do the same strategy in higher dimensions, but it is not straightforward and yields no advantage in comparison with the next approach. The approach we have implemented is a *row-wise ordering* and it is displayed in Figure 10.5. It is just more straightforward way of ordering grid points. It is also more convenient to use this approach in hypercubes.

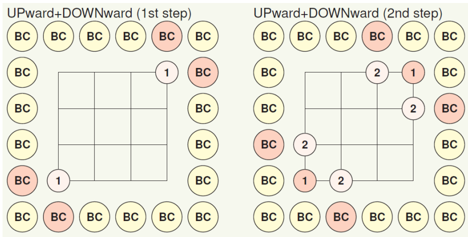


Figure 10.3: First steps of the algorithm in the 2D case. Elements numbered 1 correspond to the step 1 from both sweeps UP and DOWN. Elements numbered 2 correspond to the elements that can be computed as the second step also for both sweeps.

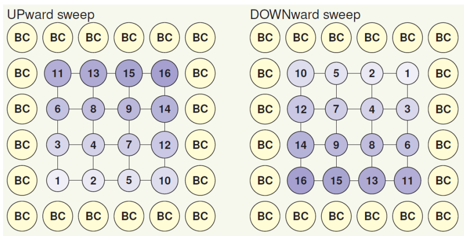


Figure 10.4: The complete algorithm in the 2D case. The points are computed in the order of the numbering. The left part of the figure refers to the UP sweep and the right to the DOWN sweep. Approach of numbering is called jumping approach or house numbering.

10.1.2 ADE Schemes for Three and Higher Dimensional Models

In this section we describe how to extend the two-dimensional ADE scheme introduced before to three and higher dimensional models. We suggest an algorithm which can be extended to higher dimensional models quite easily.

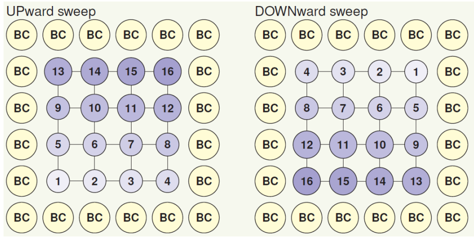


Figure 10.5: The complete algorithm in the 2D case. The points are computed in the order of the numbering. The left part of the figure refers to the UP sweep and the right to the DOWN sweep. Approach of numbering is called line approach or sequence approach.

As for the two dimensional case, a key part of the ADE in higher dimensions is to choose the proper difference quotients such that we keep good stability and consistency properties and explicitness of the scheme. Solely for simplicity we will use a uniform grid.

Consider the three dimensional case where we are solving the PDE of the price of an option under the linear Black-Scholes model introduced before, with three underlying assets. The PDE's solution will be a four-dimensional function where one dimension represents time and the other three are spatial dimensions representing the values of the underlying assets. For each time level, we have a three dimensional solution which can be illustrated as a three-dimensional grid. Recall that the initial condition is given for $V(S_1,S_2,S_3,0)$ and step by step we calculate the values for the new time layer.

As before, we retain the explicitness of the scheme by using only values that have already been computed at the current time level. Specifically, this explicit (as in the lower dimensional case) is obtained by computing the value of points in a particular sequence that only uses points that either arise from the previous time level or that have been already computed for the current time level.

For illustration purposes we depict a two-dimensional slice of the domain in Figure 10.6. We see that we move in a straight line in one dimension until we hit the boundary and then we proceed to the next point in the second dimension and so forth. By using this approach, the extension to higher dimensional models is straightforward.

10.1.3 Boundary Conditions

In higher dimensional models we also have to deal with the issue of boundary conditions. Just as in the three dimensional model 8 boundary conditions are required (each edge of the cube), for a N-dimensional model 2^N boundary conditions have to be prescribed. In an ideal case we prescribe values for the maximum values of the assets prices (truncated values) as Dirichlet boundary conditions. Alternatively one could also consider Neumann boundary conditions or Robin type boundary conditions.

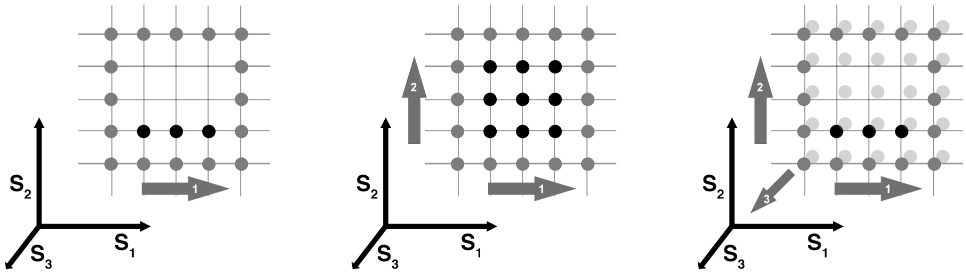


Figure 10.6: Algorithm for computing all the points in the upward sweep solution of the three dimensional implementation of the ADE, the grey dots represent the boundary conditions and the black dots represent the computed values. The arrows represent the direction and sequential order of the computation. First, second and third direction in the pictures, respectively.

10.2 Numerical Scheme

The discretization of the PDE (7.2) is done on a uniform grid. In the time domain we have N_τ subintervals of the interval $[0,T]$, thus the time step size is defined as $d\tau = T/N_t$. As we have N different underlying assets our spatial space is N-dimensional. In our numerical studies we consider both $N = 2$ and $N = 3$.

For the 3-dimensional model we have 3 spatial intervals $[x_{min}, x_{max}]$, $[y_{min}, y_{max}]$, $[z_{min}, z_{max}]$, specifically $[0, S_{1max}]$, $[0, S_{2max}]$, $[0, S_{3max}]$ as all stocks have non-negative values.

The space steps on the uniform grid are defined by the following $h_\alpha = S_{\alpha max}/N_\alpha$ for $\alpha = 1,\ldots,3$, where $S_{\alpha max}$ denotes the maximal value for the asset α and N_α denotes the number of points for the direction of the α asset.

A point on the spatial grid is then given by $[x_i, y_j, z_k]$ with $x_i = (i-1)h_1,\ y_j = (j-1)h_2, z_k = (k-1)h_3$; where $i = 1,\ldots,N_1+1,\ j = 1,\ldots,N_2+1,\ k = 1,\ldots,N_3+1$.

The discrete numerical solution of the 3-dimensional Black-Scholes equation at $[x_i, y_j, z_k]$ and time level n for the upward sweep is denoted by $u^n_{ijk} = u(x_i, y_j, z_k, n)$ and for the downward sweep is denoted by $u^n_{ijk} = d(x_i, y_j, z_k, n)$,

Since this notation would easily become very cumbersome we will introduce some abbreviations: $u(x_i, y_j, z_k, n)$ and $d(x_i, y_j, z_k, n)$ will be shortened to u^n and d^n. When we consider u at a point shifted from the point indexed by (i, j, k) we will introduce a subscript $u^n_{\beta+}$ where β denotes the direction where we're performing the shift. For example,

$$u(x_i, y_{j+1}, z_k, n) =: u^n_{2+} \qquad u(x_i, y_j, z_{k-1}, n) =: u^n_{3-}. \tag{10.1}$$

In the case that we have shifts in multiple directions we simply introduce another subscript, for example,

$$u(x_{i-1}, y_{j+1}, z_k, n) =: u^n_{1^-2^+} \qquad u(x_i, y_{j+1}, z_{k-1}, n) =: u^n_{3^-2^+}. \tag{10.2}$$

Note that this notation would not be suitable if we denote a point such as $u(x_i, y_{j+3}, z_k, n)$ but since we are considering only a first-order scheme we will not have shifts of more than 1 unit, therefore this notation is appropriate.

10.2.1 Algorithm of the Scheme

We can construct the upward sweep and the downward sweep separately for each time step and then combine them, this bring opportunities for the parallelization of the scheme. The upward sweep is calculated in a way that we are moving from one corner of the hypercube to the opposite. The downward sweep is constructed in the opposite way. This procedure can be done in different ways, but it is important to keep the explicitness of the scheme in each of the sweeps. In the following we outline the algorithms. As an illustration the upward sweep of this algorithm is represented in Figure 10.6.
According to the described procedure we construct upward and downward sweep of the solution and after each time level we calculated its average. This way we get final numerical solution c^n.
For $n = 0, 1, \dots, N_t - 1$ we repeat
1. Initialization: $u^n = c^n$; $d^n = c^n$
2. Upward: u^{n+1}_{ijk}; $i = 1, \dots, N_1 - 1$; $j = 1, \dots, N_2 - 1$; $k = 1, \dots, N_3 - 1$
3. Downward: d^{n+1}_{ijk}; $i = N_1 - 1, \dots, 1$; $j = N_2 - 1, \dots, 1$; $k = N_3 - 1, \dots, 1$
4. $c^n = (u^{n+1} + d^{n+1})/2$

10.2.2 Upward Finite Difference Quotients and Its Numerical Scheme

Finite difference quotients using the upward sweep in the ADE scheme are introduced. Exact continuous solution of the PDE (7.2) in the point $x_i, y_j, z_k, \tau_{n+\frac{1}{2}}$ is denoted as: $V := V(x,y,z,\tau)|_{(x_i,y_j,z_k,\tau_{n+\frac{1}{2}})}$ and e.g. in the time level n it is denoted as: $V^n := V(x,y,z,\tau)|_{(x_i,y_j,z_k,\tau_n)}$, For derivatives it holds as follow: $\frac{\partial V}{\partial \tau} := \frac{\partial V(x,y,z,\tau)}{\partial \tau}|_{(x_i,y_j,z_k,\tau_{n+\frac{1}{2}})}$. Approximation of the V^n is denoted as u^n for an upward sweep.

$$V \simeq \frac{V^n + V^{n+1}}{2}. \tag{10.3}$$

For the time derivative the explicit Euler discretization is used:

$$\frac{\partial V}{\partial \tau} = \frac{V^{n+1} - V^n}{d\tau} + \mathcal{O}(\tau_n^2). \tag{10.4}$$

In the convection term we choose the Robert and Weiss approximation [45]

$$\frac{\partial V}{\partial S_\alpha} = \frac{V^n_{\alpha^+} - V^n + V^{n+1} - V^{n+1}_{\alpha^-}}{2h_\alpha} + \mathscr{O}(h^2_\alpha), \qquad \forall \alpha = 1,2,3 \tag{10.5}$$

and the diffusion term is approximated by a special kind of central difference,

$$\frac{\partial^2 V}{\partial S^2_\alpha} = \frac{V^n_{\alpha^+} - V^n - V^{n+1} + V^{n+1}_{\alpha^-}}{h^2_\alpha} + \mathscr{O}(h^2_\alpha), \qquad \forall \alpha = 1,2,3. \tag{10.6}$$

Note that in all the above mentioned difference quotients we use values from two time layers in the fashion that we can use all the values from the previous time layer, but due to the algorithm explained in Figure 10.5 only known values from the current time layer are used to keep the explicitness of the algorithm.
We approximate mixed term derivatives in an explicit way, as well:

$$\frac{\partial^2 V}{\partial S_\alpha S_\beta} = \frac{V^n_{\alpha^+\beta^+} - V^n_{\alpha^+\beta^-} - V^n_{\alpha^-\beta^+} + V^n_{\alpha^-\beta^-}}{4h_\alpha h_\beta} + \mathscr{O}(h^2_\alpha + h^2_\beta) \qquad \forall \alpha,\beta = 1,2,3. \tag{10.7}$$

We now use the difference quotients introduced above to discretize the 3-dimensional Black-Scholes PDE (7.2). Let us define,

$$\gamma^{ij}_1(x_1,x_2) \equiv \frac{dt}{2h_i h_j}\Gamma_{ij} S_i(x_1) S_j(x_2), \qquad \gamma^i_2(x_1) \equiv \frac{dt}{2h_i} r S_i(x_1),$$

with $S_i(p) = (p-1)h_i$. The discretized equation for the 3D model becomes,

$$\begin{aligned} u^{n+1} - u^n = \quad & \sum_{i=1}^{3} \gamma^{ii}_1 \left[u^n_{i^+} - u^n - u^{n+1} + u^{n+1}_{i^-}\right] \\ + \quad & \sum_{i=1}^{3} \sum_{j=1, i\neq j}^{3} \frac{\gamma^{ij}_1}{4} \left[u^n_{i^+j^+} - u^n_{i^+j^-} - u^n_{i^-j^+} + u^n_{i^-j^-}\right] \\ + \quad & \sum_{i=1}^{3} \gamma^i_2 \left[u^n_{i^+} - u^n + u^{n+1} - u^{n+1}_{i^-}\right] - r\frac{u^n + u^{n+1}}{2} \end{aligned} \tag{10.8}$$

The resulting algorithm is fully explicit, if we follow the procedure illustrated in Figure 10.3. From equation (10.8) we express u^{n+1} and we realize an explicit formula for the scheme.

10.2.3 Difference Quotients and Numerical Scheme for the Downward Sweep

Let V be the exact continuous solution in point $x_i, y_j, z_k, \tau_{n+\frac{1}{2}}$. Approximation of the V^n obtained by downward sweep is d^n, where the following difference quotients are used:

$$V \simeq \frac{V^n + V^{n+1}}{2}, \tag{10.9}$$

$$\frac{\partial V}{\partial \tau} = \frac{V^{n+1} - V^n}{dt} + \mathscr{O}(\tau_n^2), \tag{10.10}$$

$$\frac{\partial V}{\partial S_\alpha} = \frac{V_{\alpha^+}^{n+1} - V^n + V^{n+1} - V_{\alpha^-}^n}{2h_\alpha} + \mathscr{O}(h_\alpha^2), \qquad \forall \alpha = 1,2,3, \tag{10.11}$$

$$\frac{\partial^2 V}{\partial S_\alpha^2} = \frac{V_{\alpha^+}^{n+1} - V^n - V^{n+1} + V_{\alpha^-}^n}{h_\alpha^2} + \mathscr{O}(h_\alpha^2), \qquad \forall \alpha = 1,2,3, \tag{10.12}$$

$$\frac{\partial^2 V}{\partial S_\alpha S_\beta} = \frac{V_{\alpha^+\beta^+}^n - V_{\alpha^+\beta^-}^n - V_{\alpha^-\beta^+}^n + V_{\alpha^-\beta^-}^n}{4h_\alpha h_\beta} + \mathscr{O}(h_\alpha^2 + h_\beta^2), \qquad \forall \alpha, \beta = 1,2,3. \tag{10.13}$$

In the same manner we get the discretized equation for the downward sweep,

$$\begin{aligned} d^{n+1} - d^n = \quad & \sum_{i=1}^{3} \gamma_1^{ii} \left[d_{i^+}^{n+1} - d^n - d^{n+1} + d_{i^-}^n \right] \\ + \quad & \sum_{i=1}^{3} \sum_{j=1, i \neq j}^{3} \frac{\gamma_1^{ij}}{4} \left[d_{i^+j^+}^n - d_{i^+j^-}^n - d_{i^-j^+}^n + d_{i^-j^-}^n \right] \\ + \quad & \sum_{i=1}^{3} \gamma_2^i \left[d_{i^+}^{n+1} - d^n + d^{n+1} - d_{i^-}^n \right] - r \frac{d^n + d^{n+1}}{2}. \end{aligned} \tag{10.14}$$

10.3 Numerical Results and Experimental Study of Convergence

We now present numerical results for two particular cases of the implementation of the ADE scheme to Black-Scholes pricing models. In particular, we show the results for the price of a Spread option depending on two underlying assets S_1 and S_2 and a three-dimensional European Call Option on three underlying assets S_1, S_2 and S_3. For both cases we show illustrations of the obtained price surfaces and experimental convergence rates.

10.3.1 Two Dimensional Black-Scholes Model

We denote the Black-Scholes price for a spread option by $V(S_1, S_2, \tau)$ where $\tau = T - t$ is the time to maturity T. Recall that the payoff of a spread option is

$$V(S_1, S_2, 0) = \max(S_1 - S_2 - K, 0)$$

where $K \in \mathbb{R}^+$ denotes the strike price. The boundary conditions are given by:

$$V(S_1, 0, \tau) = BS_{1d}(S_1, \tau), \quad S_1, \tau \in \mathbb{R}^+,$$

$$V(0, S_2, \tau) = 0, \quad S_2, \tau \in \mathbb{R}^+,$$

$$V(S_1^{max}, S_2, \tau) = e^{-q_1\tau} S_1 - e^{-r\tau}(S_2 + K), \quad S_1^{max} := S_1 \gg S_2 + K,$$

$$V(S_1, S_2^{max}, \tau) = V_{kirk}(S_1, S_2^{max}, \tau),$$

where $BS_{1d}(S_1, \tau)$ denotes the Black-Scholes price formula for a call option on a stock with price S and time to maturity τ and $V_{kirk}(S_1, S_2^{max}, \tau)$ denotes the approximation in [2].

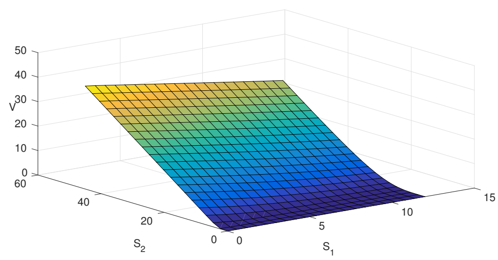


Figure 10.7: Numerical solution at the final time for two dimensional spread option on the grid with $N_1 = N_2 = 20$ space steps and $N_t = 50$ time steps.

We choose the parameters given by Table 10.1 and the different grid configurations displayed in Table 10.2.

As an example we display the numerical solution for the option price at $\tau = T$ (or equivalently $t = 0$) with a grid of $N_1 = N_2 = 20$ spatial points and $N_t = 50$ temporal points in Figure 10.7.

Parameter	**Value**
volatility of S_2 σ_2	0.3
volatility of S_1, σ_1	0.4
correlation of S_1 and S_2, ρ	0.5
maturity time T	1 (in years)
strike price K	3
maximal stock price for S_1 $S_{1\max}$	12
maximal stock price for S_2 $S_{2\max}$	45

Table 10.1: Parameters in two dimensional BS model

	N_1	N_2	N_t	$d\tau/h_1^2$	$d\tau/h_2^2$
solution 1	5	5	3	0.0578	0.004
solution 2	10	10	12	0.0578	0.004
solution 3	20	20	50	0.0578	0.004
solution 4	40	40	200	0.0578	0.004
solution 5	80	80	800	0.0578	0.004
solution 6	160	160	3200	0.0578	0.004

Table 10.2: Specifications of different grids.

In the Figure 10.8 we display a log-log plot of the errors in the L_2 norm (solid line) and the theoretical second order of convergence (dashed line).

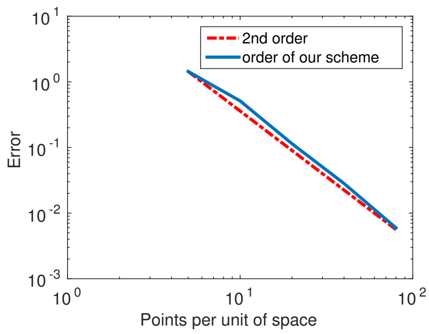


Figure 10.8: Experimental convergence analysis

10.3.2 Three Dimensional Black-Scholes Model

We now show the results of the implementation of the ADE to the three dimensional Black-Scholes model for the price $V(S_1,S_2,S_3,\tau)$ of a call option, where $\tau = T - t$ denotes the time to maturity T and S_i denotes the value of the underlying asset i. Recall the payoff for a call option:

$$V(S_1,S_2,S_3,0) = \max\left((\max(S_1,S_2,S_3) - K, 0)\right).$$

with $K \in \mathbb{R}^+$ denoting the strike price. The boundary conditions are taken from the numerical solution of the 2D Black-Scholes model, BS_{2d}, implemented as outlined in 10.3.1 but for a call-option payoff,

$$V(S_i = 0, t) = BS_{2d}(S_j, S_k, t), \quad i, j, k = 1, 2, 3, \quad i \neq j \neq k$$

$$V(S_i = S_i^{max}, t) = \max(S_i^{max} - K, 0), \quad i, j, k = 1, 2, 3.$$

In Figure 10.9 we show the price of the call option for a fixed value of S_3. The model parameters are in Table 10.3 and the grid parameters are as follows: $N_1 = N_2 = N_3 = 20$; $N_t = 50$.

Parameter	**Value**
volatility of S_1 σ_1	0.4
volatility of S_2 σ_2	0.3
volatility of S_3, σ_3	0.2
correlation of S_1 and S_2, ρ	0.0
correlation of S_2 and S_3, ρ	0.0
correlation of S_1 and S_3, ρ	0.0
maturity time T	1 (in years)
strike price K	3
maximal stock price for S_1 $S_{1\max}$	12
maximal stock price for S_2 $S_{2\max}$	12
maximal stock price for S_2 $S_{2\max}$	12

Table 10.3: Parameters in three dimensional BS model

Note that in this case we have a symmetric solution with respect to the underlying assets and hence fixing S_3 or any other asset would be identical.

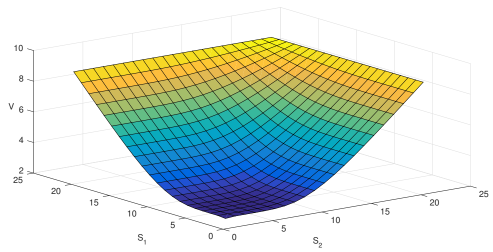


Figure 10.9: Solution of the three dimensional call option model for the fixed $S_3 = N_3/2$.

Analogously to the two dimensional case, for the three dimensional case we've computed the experimental order of convergence using different grid settings c.f. Table 10.4. Experimental results (Figure 10.10) confirm that we keep second order of convergence also in the three dimensional model.

N_1	N_2	N_3	N_t	$d\tau/h_1^2$	$d\tau/h_2^2$	$d\tau/h_3^2$
5	5	5	3	0.004	0.004	0.004
10	10	10	12	0.004	0.004	0.004
20	20	20	50	0.004	0.004	0.004
40	40	40	200	0.004	0.004	0.004
80	80	80	800	0.004	0.004	0.004
160	160	160	3200	0.004	0.004	0.004

Table 10.4: Usage of different grids.

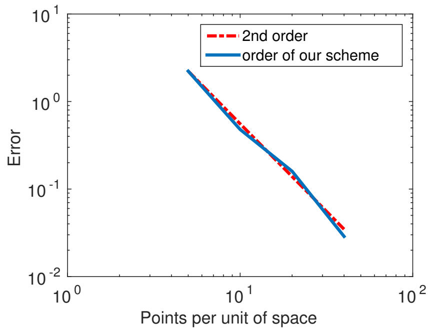


Figure 10.10: Experimental order of convergence in three dimensional call option model

10.4 Influence of dimensionality on computational complexity of the scheme

In this section we highlight the fact, where the ADE scheme has a good potential to be an effective scheme in higher dimensions. We compare it with the behavior of the classical Cranck-Nicolson scheme.

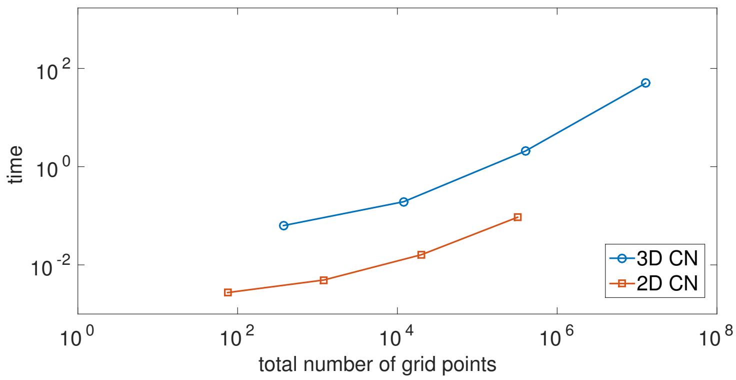


Figure 10.11: Computational complexity with respect to the total number of points in the grid for Cranck-Nicolson scheme.

Solution of the option price for Cranck-Nicolson (CN) scheme is implemented with a lot of optimization steps, so we do not compare real time for the calculation. We focus on the fact observed in the Figure 10.11 for CN scheme is growing with dimension. It means for

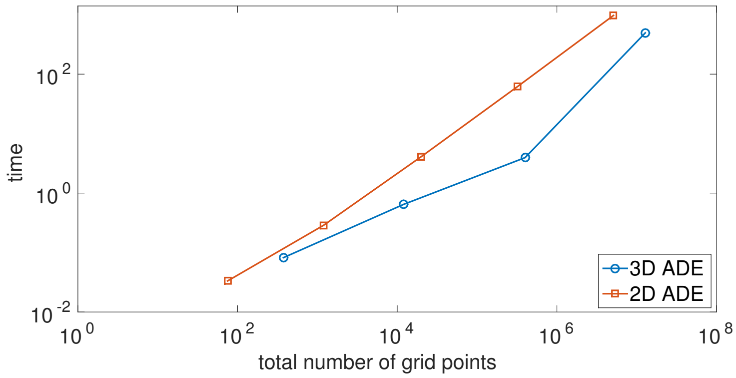


Figure 10.12: Computational complexity with respect to the total number of points in the grid for ADE scheme.

the same number of total points in a grid we need more time in the 3D model as in the 2D model. The explanation is coming from the construction of the scheme. Although for the same number of total points in a grid the size of the matrix is the same, but its structure is different. For 3D more non-diagonal terms are present and to compute solution in the implicit scheme is becoming costly for higher-dimensional models.
Costs for the ADE schemes in Figure 10.12 for higher dimensions are not growing, even opposite, since the calculation of the explicit scheme depends only on the total number of grid points and size of the stencil.

Chapter 11
Trefftz methods for the Black-Scholes equation, Flexible Local Approximation Methods

This chapter is based on the cooperation with Prof.Tsukerman and his longlasting experiences and results in this area. Goal of this chapter is to present a short overview on alternative methods for solving the Black-Scholes model, as a 'proof of the concept'.

Trefftz methods are represented by Flexible Local Approximation Methods (FLAME). They were applied in different areas, but not in finance yet. Trefftz schemes are an alternative to traditional methods solving the for Black-Scholes equation. The Trefftz approach may lead to new finite difference schemes.

Trefftz functions by definition satisfy the underlying differential equation. Examples for basis functions are exponentials, plane waves, harmonic polynomials, etc. There is a lot of study for stationary problems but how it works for time-dependent problems like the Black-Scholes equation. One example is given in [56] pp.7-8. Here, the time is considered as an additional coordinate. Basis functions are chosen as dependent functions on space and time.

11.1 How Trefftz methods work?

Trefftz methods are based on flexible local approximative functions. Approach is different from FEM and from FDM. It is an alternative approach how to define the coefficients of the FDM scheme in another way. There is defined a mesh on the computational domain Ω. On Ω we define subdomains Ω_i which have the stencil size. **Basis functions** ψ_α^i satisfy the differential equation locally. The solution of the differential equation in Ω_i is denoted by u^i and it is a linear combination of the basis functions over the subdomains.

$$u^i = \sum_\alpha c_\alpha^i \psi_\alpha^i \qquad in \quad \Omega_i$$

$$u^i = N^i c^i,$$

where c^i is vector of coefficients c_α^i and N^i represents a matrix in the $i-$th subdomain. In other notation we can specify that entries of the matrix $N_{\alpha\beta}$

$$N^T_{\alpha\beta} = \psi_\alpha(\mathbf{r}_\beta)$$

consists of the basis functions expressed in points from stencil.

There is vector s^i from nullspace of matrix N^i. Vector s^i directly specifies weights, coefficients of the numerical scheme on the given stencil:

$$s^i \in \text{Null } (N^i)^T.$$

It holds the following:

$$(s^i)^T u^i = 0,$$

i.e.

$$(s^i)^T N^i c^i = 0.$$

The idea is simple and derivation of this method with more technical details and properties and applications in the fields from physics can be found in [57].

For better understanding we illustrate this idea by the following examples:

Example 1: Laplace equation

$$u_{xx} = 0$$

Trefftz basis functions:

$$\Psi = \{1, x\}$$

Solution as a linear combination of basis functions:

$$u(x,t) = c_1\psi_1(x) + c_2\psi_2(x)$$

A three-point stencil with the following nodes:

$$x_1 = -h, \quad x_2 = 0, \quad x_3 = h$$

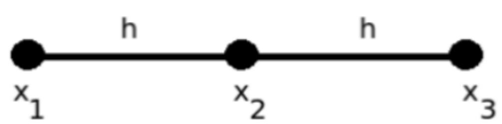


Figure 11.1: Three point stencil.

Matrix:

$$N = \begin{pmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -h & 0 & h \end{pmatrix}$$

Nullspace of the matrix N yields the weights

$$(1,-2,1)^T,$$

i.e. we obtain the standard stencil.

Example 2: Wave equation

$$c^2 u_{xx} - u_{tt} = 0 \qquad [0,L]$$

Basis functions:

$$\Psi = \{1, x-ct, x+ct, (x-ct)^2, (x+ct)^2\}$$

Stencil:

$$x_1 = [0,0], \quad x_2 = [k,0], \quad x_3 = [-k,0], \quad x_4 = [0,-h], \quad x_5 = [0,h]$$

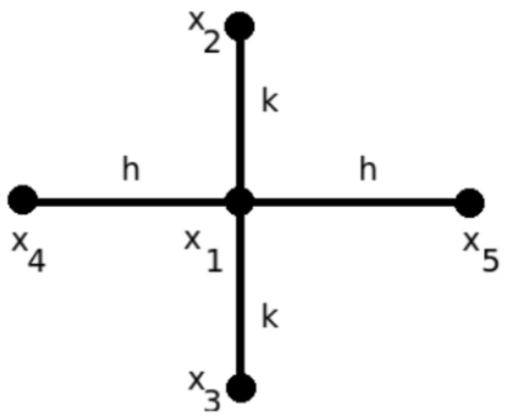


Figure 11.2: Five point stencil.

Matrix:

$$N = \begin{pmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) & \psi_1(x_4) & \psi_1(x_5) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) & \psi_2(x_4) & \psi_2(x_5) \\ \psi_3(x_1) & \psi_3(x_2) & \psi_3(x_3) & \psi_3(x_4) & \psi_3(x_5) \\ \psi_4(x_1) & \psi_4(x_2) & \psi_4(x_3) & \psi_4(x_4) & \psi_4(x_5) \\ \psi_5(x_1) & \psi_5(x_2) & \psi_5(x_3) & \psi_5(x_4) & \psi_5(x_5) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -ck & ck & -h & h \\ 0 & ck & -ck & -h & h \\ 0 & c^2k^2 & c^2k^2 & h^2 & h^2 \\ 0 & c^2k^2 & c^2k^2 & h^2 & h^2 \end{pmatrix}$$

Nullspace of matrix N is set of weights:

$$(-2+2\frac{c^2k^2}{h^2}, 1, 1, -\frac{c^2k^2}{h^2}, -\frac{c^2k^2}{h^2})$$

what leads to the symmetric stencil.

Example 3: Heat equation

$$u_{xx} - u_t = 0$$

General solution:

$$u(x,t) = ax^2 + bx + 2at + d$$

Basis functions:

$$\Psi = \{1, 0.5x^2 + t, e^{-\lambda^2 t}\sin(\lambda x), e^{-\lambda^2 t}\cos(\lambda x)\}$$

Stencil:

$$x_1 = [0,0], x_2 = [k,0], x_3 = [-k,0], x_4 = [0,-h], x_5 = [0,h]$$

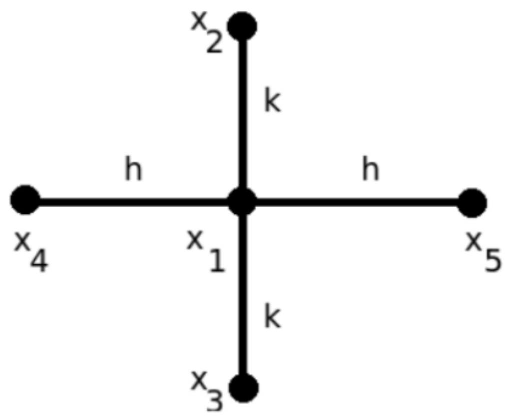


Figure 11.3: Five point stencil.

Matrix:

$$N = \begin{pmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) & \psi_1(x_4) & \psi_1(x_5) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) & \psi_2(x_4) & \psi_2(x_5) \\ \psi_3(x_1) & \psi_3(x_2) & \psi_3(x_3) & \psi_3(x_4) & \psi_3(x_5) \\ \psi_4(x_1) & \psi_4(x_2) & \psi_4(x_3) & \psi_4(x_4) & \psi_4(x_5) \\ \psi_5(x_1) & \psi_5(x_2) & \psi_5(x_3) & \psi_5(x_4) & \psi_5(x_5) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -h & h \\ 0 & k & -k & 0 & 0 \\ 1 & e^{-\lambda^2 k} & e^{\lambda^2 k} & \cos(\lambda h) & \cos(\lambda h) \end{pmatrix}$$

with the entries of the nullspace.

$$x_1 = \frac{\cos(\lambda h)}{\cos(\lambda h) - 1}$$

$$x_2 = \frac{\cos(\lambda h)}{-2\cos(\lambda h) + e^{-\lambda^2 k} + e^{\lambda^2 k}}$$

$$x_4 = \frac{-x_1 - x_2(e^{-\lambda^2 k} + e^{\lambda^2 k})}{2\cos(\lambda h)}$$

$$x_2 = x_3, \quad x_4 = x_5$$

what leads to the symmetric stencil with a parameter λ to be chosen.

Example 4:Linear Black-Scholes equation with transformation
A transformation to the heat equation is possible in case of linear Black-Scholes model with constant coefficients and we refer to the Example 3.

Example 5:Linear Black-Scholes equation, 'cheating mode'
Here the basis functions are chosen functions from the closed form formula solution of Black-Scholes. That's why we denote this approach as a 'cheating mode' since the solution is known in advance. We just want to highlight that using this functions, it leads to the generating the coefficients of the finite difference scheme.

Example 6:Linear Black-Scholes equation, without cheating mode, generating basis funtions using Taylor expansion
First step is to remove the 'cheating mode' with the exact solution as basis function. We can use functions with some financial interpretation: discount factor: $e^{-r(T-t)}$ or especially $Ke^{-r(T-t)}$ which satisfies the Black-Scholes equation locally. Another posssiblity is choice of the basis function as a stock price S.
Another possibility is to generate basis functions 'automatically', using Taylor expansions to an arbitrary order. High-order schemes can be generated by replacing the Taylor expansions with Trefftz approximations which typically have much higher accuracy. We can do it by using expansion of $V(S,t)$ around any given point into the Taylor series in S, t to an arbitrary order, substituting this expansion into the Black-Scholes equation and eliminating as many low-order terms as possible, to obtain approximate Trefftz functions.

11.2 Numerical results with Six-Point FLAME Scheme

In this section we provide numerical results of solving linear Black-Scholes equation. There is displayed exact solution of Black-Scholes in the Figure 11.4 upper left, upper right is an option price which we get using FLAME method. There was used six-point FLAME scheme. Trefftz basis function are generated in an 'automatic way' by using Taylor expansion. There is recorded error in the FLAME scheme as in the Figure 11.4 down as a difference between exact and numerical solution.

11.3 Comparison of FLAME and Crank-Nicolson scheme

In the Figure 11.5 we compare Crank-Nicolson (CN) scheme and 6 point FLAME scheme. There is displayed the ratio of errors of these 2 schemes for different space and time steps. Most of the time the error is lower for CN scheme. There are some regions with approximately the same error for both schemes.

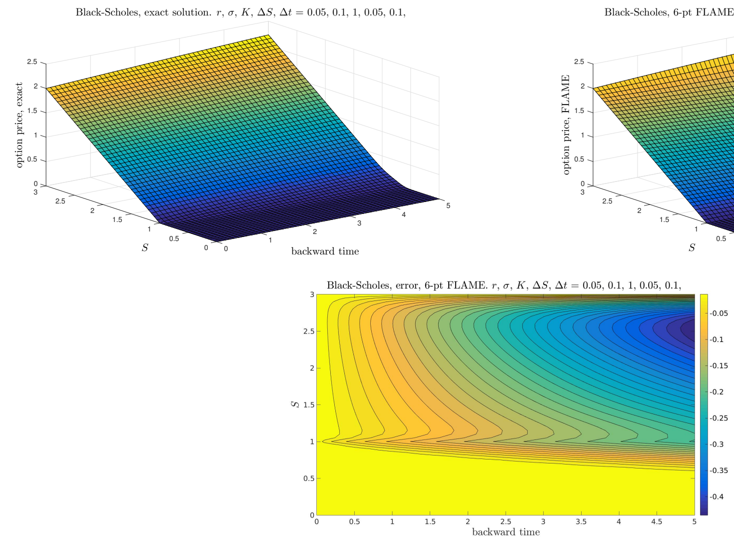


Figure 11.4: Exact (left) and FLAME (right) solutions. Trefftz basis functions obtained via Taylor expansions. Solution is accurate up to round-off. Six-point FLAME scheme: 2 levels in t $\times$ 3 layers in S. (with permission of Igor Tsukerman)

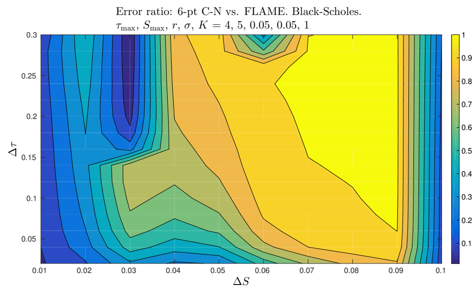


Figure 11.5: Ratio of errors - FLAME vs. CN scheme (with permission of Igor Tsukerman)

11.4 Further potential of the Trefftz schemes

This chapter serves as a proof of concepts that Trefftz methods can be used in different fields. However there is a lot of scope for improvement and suggesting the FLAME scheme with good properties. Big potential of FLAME methods is to generate as good exact solution as possible based on the choice of the basis functions. In a series of experiments we observe some numerical instabilities which are the subject for a deeper study. Suggestions of using Trefftz basis function in another approaches, e.g. Discontinuous Galerkin method are challenging, as well. FLAME has a great deal of flexibility which makes this method competitive. The application to nonlinear equations and usage of nonuniform meshes can be a nice enrichment of these approaches (it can save a lot of computational time, it is convenient to use nonuniform mesh for financial problems; e.g. a mesh according to [32]).

Conclusion and Outlook

This dissertation thesis has fulfilled its goal capturing a wide overview of handling multi-dimensional models. On the one hand it studies analytical approximations, on the other hand it focuses on the numerical analysis of the ADE methods.

The first part of the thesis is focused on searching for a suitable approximative solution of the convergence CKLS model. Crucial part is in the Chapter 3 where approximation for domestic and European bond is suggested and its accuracy derived. We suggest an improved approximation with higher accuracy order. A complete analysis of the model and its approximation, testing on simulated and providing calibration using real market data are included. Since the results with real market were not perfectly satisfactory, we implemented a few improvements. One of them is the estimation of the overnight interest rates based on the modeling from the term structures of the interest rate in Vasicek model in the Chapter 4. We have proposed and tested a procedure for estimating the short rates together with the parameters of the Vasicek model. Simulations show that the procedure exhibits high precision. When applying it to the real data, we obtain a good fit of the term structures. However, when taking different sets of maturities as inputs to the calibration, we often obtain quite different estimated evolutions of the short rate. Nevertheless, the fit of the term structures is good. We would like to study this phenomenon more deeply, find its financial interpretation and possible explanation.

Another possibility how to improve modeling of the stochastic interest rate model is to suggest the alternative model, where the one-dimensional stochastic process is modelled as a sum of two unobservable processes. Since calibration of the bond yields is dependent on the European data, improvement in fitting of the bond yield in European model will also influence the accuracy of the domestic bond yield curves.

In the Chapter 5 we studied a particular class of two-factor models of interest rates in which the short rate is defined as a sum of two CKLS-type processes. We developed a method of estimating the short rate and fitting the term structures for the special Vasicek case model and showed its usefulness by applying it to fitting Euribor interest rates. An example from the simulated data where the procedure gave a very precise estimate of the short rate even if applied to data generated from a model with nonconstant volatilities, motivated us to propose an approximation of bond prices in such a model and prove its order of accuracy. We note that besides a precise estimate of the short rate, we have also its decomposition into the factors, but these are shifted by a constant. Still, it provides a lot of information about the process and hence our future work will be concerned with using this information together with the approximation of the bonds which were derived to obtain estimates for all the parameters of the model.

We end up this modeling and analytical part of the thesis with three-factor convergence model of interest rate with the Chapter 6. Combining two-factor convergence short rate model and improvement in modeling of one interest rate as a sum of two CKLS-type processes lead to the three-factor model. We have numerically tested the proposed approximation on the CIR model with zero correlations, for which the exact solution can be expressed in a simpler form and also analytically derived its accuracy in this case. The difference of logarithms of the exact solution and the proposed approximation is of order $O(\tau^4)$. Our next aim is to derive the order of accuracy in the general case. The special form of the solution in the case considered in this thesis makes the analysis more direct, however, it is possible to study the accuracy of the approximation of the bond prices also without this structure (see the Chapter 3 for the analysis of a two-factor convergence model). Furthermore, we will look for a suitable calibration algorithm and calibrate the model to the real data to see, whether the increase complexity leads to a significant improvement in fitting the market data.

In the second part of the thesis we discuss one and two factor interest rate models and apply the classical Fichera theory to the resulting degenerate parabolic PDEs. This theory provides highly relevant information how to supply BCs in these applications.

We provided a numerical analysis for ADE methods solving linear convection-diffusion-reaction equations. The stability was investigated by two different approaches. The matrix approach yields unconditional stability in the downward sweep using upwind discretization. The von-Neumann analysis yields unconditional stability of the downward sweep using the Roberts and Weiss approximation. It turned out that the order of consistency is $\mathrm{O}(k^2+h^2+k/h)$ for the upward or downward sweeps, but its combination exhibits an increase order of consistency $\mathrm{O}(k^2+h^2)$. Next, for the BS model, as an application in computational finance, we obtained an order of consistency $\mathrm{O}(k^2+h^2)$ for both downward and upward sweeps.

We suggest the usage of ADE methods to numerically solve higher-dimensional PDEs. We implemented it for the linear 2D and 3D Black-Scholes pricing equation. The order of consistency of the implemented ADE method is $\mathbf{O}(k^2+h^2)$ and this was verified experimentally.

Further studies will be made on the implementation of this scheme to higher-dimensional, non-linear Black-Scholes models, e.g. of Zakamouline [62]. Also, since the ADE approach is quite suitable to parallelization, an implementation using a parallel computing environment will be envisaged.

In the last Chapter 11 we briefly introduce an alternative approach of solving the Black-Scholes equation based on the flexible local approximative schemes, also called Trefftz methods. The results are very preliminary and there is a lot of room for improvements.

The thesis deals with broad scope of numerical and analytical techniques. It brings unique results in form of approximations in closed form formula in short-rate models and brushes up forgotten ADE schemes, brings its numerical analysis and implements it in higher dimensional models. Also some other side results appeared as a surprise which had not been really planned (e.g. Chapter 11 or Chapter 8) what is of course a positive finding.

The thesis deals with wide scope of numerical and analytical techniques in computational finance. Unique results in form of approximations in closed form formula in short-rate models are included and forgotten ADE schemes have been reminded and studied. Implementation for one, two, three dimensional models is provided and extensions to higher dimensional models have been outlined.

Bibliography

[1] Flexible Local Approximation MEthods (FLAME). pages 189–238. 2008.

[2] C. Alexander and A. Venkatramanan. Analytic approximations for spread options. *SSRN 1012521*, 2007.

[3] H. Z. Barakat and J. A. Clark. On the solution of the diffusion equations by numerical methods. *Journal of Heat transfer*, 88(4):421–427, 1966.

[4] G. Barles and H. M. Soner. Option pricing with transaction costs and a nonlinear Black-Scholes equation. *Finance and Stochastics*, 2(4):369–397, 1998.

[5] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[6] L. A. Bordag and A. Y. Chmakova. Explicit solutions for a nonlinear model of financial derivatives. *International Journal of Theoretical and Applied Finance*, 10(01):1–21, 2007.

[7] D. Brigo and F. Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*, volume 82. 2006.

[8] Z. Bučková, M. Ehrhardt, and M. Günther. Alternating Direction Explicit Methods for Convection Diffusion Equations. *Acta Math. Univ. Comenianae*, Vol. LXXXI:309–325, 2015.

[9] Z. Bučková, M. Ehrhardt, and M. Günther. Fichera theory and its application to finance. *Proceedings ECMI 2014, Taormina, Sicily, Italy*, 2016.

[10] Z. Bučková, J. Halgašová, and B. Stehlíková. Short rate as a sum of two CKLS-type processes. *accepted for publication in Proceedings of Numerical analysis and applications conference, Springer Verlag in LNCS*, 2016.

[11] Z. Bučková, P. Pólvora, M. Ehrhardt, and M. Günther. Implementation of Alternating Direction Explicit Methods for higher dimensional Black-Scholes Equations. *AIP Conf. Proc. 1773, 030001*, 2016.

[12] Z. Bučková, B. Stehlíková, and D. Sevčovič. Numerical and analytical methods for bond pricing in short rate convergence models of interest rates. *Chapter in Advances in Mathematics Research. Volume 21, Editors: Albert R. Baswell*, 2016.

[13] Z. Bučková (Zíková). Convergence models of interest rates (in Slovak: Konvergenčné modely úrokových mier). *Master thesis*, 2011.

[14] Z. Bučková (Zíková) and B. Stehlíková. Convergence model of interest rate of CKLS type. *Kybernetika*, 8(5):7–5, 2012.

[15] L. J. Campbell and B. Yin. On the stability of alternating-direction explicit methods for advection-diffusion equations. *Numerical Methods for Partial Differential Equations*, 23(6):1429–1444, 2007.

[16] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders. An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance*, 47(3):1209–1227, 1992.

[17] R. R. Chen and L. Scott. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *The Journal of Fixed Income*, 3.3:14–31, 1993.

[18] Y. Choi and T. S. Wirjanto. An analytic approximation formula for pricing zero-coupon bonds. *Finance Research Letters*, 4(2):116–126, 2007.

[19] R. Company, L. Jódar, E. Ponsoda, and C. Ballester. Numerical analysis and simulation of option pricing problems modeling illiquid markets. *Computers & Mathematics with Applications*, 59(8):2964–2975, 2010.

[20] T. Corzo Santamaria and E. S. Schwartz. Convergence within the EU: Evidence from Interest Rates. *Economic Notes*, 29(2):243, 2000.

[21] C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–408, 1985.

[22] D. J. Duffy. Unconditionally Stable and Second-Order Accurate Explicit Finite Difference Schemes Using Domain Transformation: Part I One-Factor Equity Problems. *SSRN Electronic Journal*, 2009.

[23] D. J. Duffy. *Finite Difference methods in financial engineering: a Partial Differential Equation approach*. John Wiley & Sons, 2013.

[24] D. J. Duffy and G. Andrea. *C# for Financial Markets*. John Wiley & Sons, 2013.

[25] M. Ehrhardt. *Nonlinear models in mathematical finance: new research trends in option pricing*. Nova Science Publishers, 2008.

[26] E. Ekström, P. Lötstedt, and J. Tysk. Boundary values and finite difference methods for the single factor term structure equation. *Appl. Math. Finance*, (16):253 – 259, 2009.

[27] A. Episcopos. Further evidence on alternative continuous time models of the short-term interest rate. *Journal of International Financial Markets, Institutions and Money*, 10:199–212, 2000.

[28] G. Fichera. On a Unified theory of boundary value problems for elliptic-parabolic equations of second order in boundary value problems. *ed. R.E. Langer, University of Wisconsin Press*, 1960.

[29] R. Frey and P. Patie. Risk management for derivatives in illiquid markets: A simulation study. In *Advances in finance and stochastics*, pages 137–159. Springer, 2002.

[30] J. Halgašová. Approximation of bond prices in two-factor models of interest rates (in Slovak: Aproximácia cien dlhopisov v dvojfaktorových modeloch úrokových mier). *master thesis*, 2011.

[31] J. Halgašová, B. Stehlíková, and Z. Bučková (Zíková). Estimating the short rate from the term structures in the Vasicek model. *Tatra Mt. Math. Publ*, 61:87–103, 2014.

[32] K. J. In 't Hout and C. Mishra. Stability of ADI schemes for multidimensional diffusion equations with mixed derivative terms. *Applied Numerical Mathematics*, 74:83–94, 2013.

[33] M. Jandačka and D. Ševčovič. On the risk-adjusted pricing-methodology-based valuation of vanilla options and explanation of the volatility smile. *Journal of Applied Mathematics*, 2005(3):235–258, 2005.

[34] Y. K. Kwok. *Mathematical Models of Financial Derivatives*, volume 97. 2008.

[35] V. Lacko. Two-Factor Convergence Model Of Cox-Ingersoll-Ross Type. *Master thesis*, 2010.

[36] V. Lacko and B. Stehlíková. Two-factor Convergence Model of Cox-Ingersoll-Ross Type. *Proceedings of the 17th Forecasting Financial Markets Conference*, Hannover,, 2010.

[37] B. K. Larkin. Some stable explicit difference approximations to the diffusion equation. *Mathematics of Computation*, 18(86):196–202, 1964.

[38] S. Leung and S. Osher. An Alternating Direction Explicit (ADE) Scheme for Time-Dependent Evolution Equations. 2005.

[39] I. Melicherčík, L. Olšarová, and V. Úradníček. *Chapters from financial mathematics (in Slovak: Kapitoly z finančnej matematiky)*. Epos, 2005.

[40] R. C. Merton. Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1):141, jan 1973.

[41] O. A. Olejnik and E. V. Radkevič. Second order equations with nonnegative characteristic form. *American Mathematical Society*, 1973.

[42] G. Pealat and D. J. Duffy. The Alternating Direction Explicit (ADE) Method for One-Factor Problems. *Wilmott magazine*, 2011(54):54–60, 2011.

[43] S. A. Piacsek and G. P. Williams. Conservation properties of convection difference schemes. *Journal of Computational Physics*, 6(3):392–405, 1970.

[44] N. Privault. *An Elementary Introduction to Stochastic Interest Rate Modeling*, volume 16 of *Advanced Series on Statistical Science and Applied Probability*. World Scientific, 2012.

[45] K. V. Roberts and N. O. Weiss. Convective difference schemes. *Mathematics of Computation*, pages 272–299, 1966.

[46] L. C. G. Rogers. Which model for Term-Structure of Interest Rates Should One Use. *Mathematical Finance, IMA*, 1995.

[47] V. K. Saul'yev. A method of numerical integration of diffusion equations. *Dokl. Akad. Nauk SSSR*, 115:1077–1079, 1957.

[48] D. Ševčovič and A. U. Csajková. Calibration of one factor interest rate models. *Journal of Electrical Engineering*, 55(No. 12/s):46–50, 2004.

[49] D. Ševčovič and A. U. Csajková. On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model. *CEJOR*, 13:169–188, 2005.

[50] D. Ševčovič, K. Mikula, and B. Stehlíková. *Analytical and numerical methods for pricing of financial derivatives (in Slovak: Analytické a numerické metódy oceňovania finančných derivátov)*. Nakladateľstvo STU, 2009.

[51] B. Stehlíková. A simple analytical approximation formula for the bond price in the Chan-Karolyi-Longstaff-Sanders model. *International Journal of Numerical Analysis and Modeling*, 4(3):224–234, 2013.

[52] B. Stehlíková and Z. Bučková (Zíková). A three-factor convergence model of interest rates. *Proceedings of ALGORITMY*, pages 95–104, 2012.

[53] B. Stehlíková and D. Ševčovič. Approximate formulae for pricing zero-coupon bonds and their asymptotic analysis. *International Journal of Numerical Analysis and Modeling*, 6(2):274–283, 2009.

[54] D. Tavella and C. Randall. *Pricing financial instruments: The finite difference method*, volume 13. John Wiley & Sons, 2000.

[55] B. F. Towler and R. Y. K. Yang. Numerical stability of the classical and the modified Saul'yev's finite-difference methods. *Computers & Chemical Engineering*, 2(1):45–51, 1978.

[56] I. Tsukerman. Electromagnetic applications of a new finite-difference calculus. *IEEE Transactions on Magnetics*, 41(7):2206–2225, 2005.

[57] I. Tsukerman. *Computational Methods for Nanoscale Applications: Particles, Plasmons and Waves*. Nanostructure Science and Technology, 2007.

[58] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[59] R. Vichnevetsky and J. B. Bowles. *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*. Society for Industrial and Applied Mathematics, jan 1982.

[60] P. Wilmott, J. Dewynne, and S. Howison. *Option pricing: mathematical models and computation*. Oxford financial press, 1993.

[61] N. N. Yanenko. *The method of fractional steps*. Springer, 1971.

[62] V. I. Zakamouline. *Option Pricing and Hedging in the Presence of Transaction Costs and Nonlinear Partial Differential Equations*. Nova Science Publishers, 2008.

[63] W. Zhao, X. Yang, and L. Wu. Alternating segment explicit-implicit and implicit-explicit parallel difference method for the nonlinear Leland equation. *Advances in Difference Equations*, 2016(1):1–18, 2016.