Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Alexander van Haastrecht

Pricing Long-Term Options with Stochastic Volatility and Stochastic Interest Rate

PhD Thesis 2010-045

Pricing Long-term Options with Stochastic Volatility and Stochastic Interest Rates

ISBN 978-90-8570-520-8

Published by Wöhrmann Print Service, Zutphen, The Netherlands

Pricing Long-term Options with Stochastic Volatility and Stochastic Interest Rates

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op donderdag 8 juli 2010, te 14:00 uur

door

Alexander van Haastrecht

geboren te Alkmaar

Promotiecommissie

Promotor:
Prof. dr. A.A.J. Pelsser

Overige leden:
Prof. dr. M.J. Goovaerts
Prof. dr. F.C.J.M. de Jong
Prof. dr. R. Kaas
Prof. dr. J.M. Schumacher
dr. P.J.C. Spreij
Prof. dr. M.H. Vellekoop

Faculteit der Economie en Bedrijfskunde

Preface

This thesis is the result of two and a half years research at the University of Amsterdam and Delta Lloyd. When doing research in actuarial and mathematical finance, the link between academia and the financial industry provides a very inspiring blend. Still, performing a PhD in such a field is a hurdle that cannot be overcome without the help of a great number of special people.
At Delta Lloyd Leven I am particularly indebted to Diederik Schouten and Aartjan Paauw for creating my position, which enabled me to combine practical work with academic research. I would like to thank Diederik for the stimulating working environment his department provided and for the freedom, and hence the confidence, he gave me to investigate new developments and incorporate them into every day's pricing and risk management processes.
At the University of Amsterdam, special thanks go out to my supervisor and co-author Antoon Pelsser. His guidance, enthusiastic mind set and challenging ideas formed the basis for performing cutting edge research. I am grateful for the great opportunities he gave me to present our work abroad at the International Symposium on Insurance and Finance in Bergen, the Quantitative Methods in Finance conference in Sydney and the Actuarial and Financial Mathematics conference in Brussel. It was an honor and a privilege to work with Antoon.
Next, I would like to thank my co-authors, Roger Lord, Richard Plat and David Schrager. Roger's never-ending enthusiasm, persistence and profound knowledge on the valuation of exotic options have increased the quality of this thesis. Richard's positive personality together with the many fruitful and interesting discussions we had on various aspects of the valuation and risk management of insurance products, made the days at the university very enjoyable. I would also like to thank to David Schrager for his valuable contributions on the pricing of long dated contracts.
In general I am grateful to Netspar for its financial support, which enabled me to attend several conferences. I would like to thank André Zegeling for putting me wise about financial mathematics and making me enthusiastic to perform scientific research. I am much obliged to several people at the University of Amsterdam and would like to thank Joost Driessen, Michiel Janssen, Rob Kaas, Antoon Pelsser and Peter Spreij for producing a pleasant and inspiring scientific atmosphere. At Delta Lloyd Leven I would like to thank my colleagues for taking an interest in my research and stimulating me to develop scientific as well as practically applicable methods. Special acknowledgements are due to Jeroen Decuypere and Gert Frijters for their feedback, amicable cooperation and in particular for enduring many of my talks and obstinate comments. Likewise, my friends are thanked for their interests and for providing the necessary distractions. Above all, I would like to thank my mother, grandparents and Suus for their invaluable love, support and patience.

Contents

Chapter 1

Introduction and Outline

The markets for long-term options have witnessed an explosive growth over the last decade. One of the recent developments is the expansion of markets for long-term European equity, exchange rate and inflation options. Currently, liquid prices for maturities up to thirty years and beyond are shown for these products. Also the markets for hybrid options, which depend on multiple underlying assets, are starting to take off. Pension contracts, for instance, incorporate options which are exposed to both equity and interest rates risks. The correlation, a measure for the dependency between the underlying assets, has a large impact on the pricing and risk management of such contracts. The simultaneous decreases in asset prices and interest rates around 2003 and during the recent credit crunch are "perfect" examples hereof and caused the funding ratios of many pension funds to drop to historically low levels. Appropriate methods for the pricing and risk management of long-term contracts, should therefore at least be able to deal with such (joint) market risks.

Though the history of option contracts is dated back to ancient Greek and Babylonian times, the trading of financial securities on exchanges arose from the 16th to 18th centuries. In Antwerp, Amsterdam and London, well organized exchanges were established dealing a range of commodities and financial options. These activities included trading for future delivery, "time bargains", as well as options. At the bourse of Amsterdam contract prices for tulip bulbs reached extraordinarily high levels around 1637 and then suddenly collapsed, which is generally considered as the first ever recorded financial bubble, referred to as "tulip mania". In the U.S. the first formal futures and options exchange, the Chicago Board of Trade, was established in 1848 and initially served to reduce the seasonality effects of grain harvesting. At the time, the distribution of grain took place in Chicago due to its central location, however its storage facilities were unable to accommodate the enormous increase in supply after the harvest, whilst the same facilities were underutilized during other periods of the year. As a consequence spot prices for grain fell and rose severely. Futures contracts, on the other hand, allowed farmers to store the grain at other places and deliver it to Chicago at a later time. In this way, farmers were able transfer the price risk associated with the grain as it could always be sold and delivered

[0]Background information for this chapter has been used from Chance (1998), Lewis (2000), Poitras (2008) and Lord (2008).

anywhere else at any time. However, it was not until 1973, marked by both the creation of the Chicago Board Options Exchange and the publication of the Black and Scholes (1973) and Merton (1973) papers, that option trading really took off.

Prior to the breakthrough of Black, Scholes and Merton, market participants would have had to rely on heuristics methods and their own views of the future to determine the prices of options. Starting with Bachelier (1900), who suggested a fair game approach using a normal distribution for the underlying asset, attempts were made to develop option pricing formulas. Nevertheless, all these approaches lacked the crucial insight of Black and Scholes (1973) and Merton (1973) that, under certain assumptions, an option can be turned into a risk-free instrument using a technique called "dynamic hedging". Under the assumption that no arbitrage opportunities exist in the financial markets, the price of the option price should equal the price of its replicating portfolio, irrespective of investor's risk attitudes and expectations. This finding together with an increasing computational power, formed the basis for an explosive revolution in the use of derivatives, making this industry as large as it is today.

Using the replication argument, more exotic structures could also be priced. In such a procedure, the prices of actively traded contracts, like futures or European options, are used to determine the hedge and hence the involved costs of the exotic option. Therefore, an option pricing model can essentially be interpreted as an extrapolation method of the prices of these simpler instrument. A necessary requirement for such an extrapolation to make sense, is that within the model the prices of simpler contracts coincide with their market price. It became clear that this was not the case within the Black and Scholes (1973) model and that some of its assumptions, like constant volatility and constant interest rates were inappropriate. A lot of research within mathematical finance, has therefore concentrated on the development of alternative models and asset price dynamics, such that the prices of liquid vanilla options and the stochastic nature of the underlying asset are matched in a more suitable way.

Generally, to price an exotic option, one first chooses a financial model for the underlying asset, which is calibrated to match the prices of traded vanilla contracts as closely as possible. Using appropriate numerical techniques, the calibrated model is then applied to price a specific exotic option. Many mathematical models may fit such a description and additional criteria have to be considered to assess the fitness of a model to price financial derivatives. First, as a pricing model can essentially be interpreted as an extrapolation tool between liquid simple contracts to complex derivatives, it should be economically plausible and parsimonious. Secondly, practitioners demand fast and accurate prices and sensitivities of financial contracts, therefore the model should be analytically tractable. Thirdly, because financial models and option contracts are becoming increasingly complex, efficient methods have to be developed to cope with these evolutions. For a realistic pricing and risk management of long-term options, it is furthermore strongly advised to incorporate empirical phenomena as heavy-tailed returns, stochastic interest rates and general correlation structures into a derivative pricing model.

All three parts of this thesis, are devoted to the valuation of long-dated derivatives. Part I is ded-

icated to the incorporation of long-term maturities into the development of new models and the derivation of closed-form pricing formulas herein. Part II considers the pricing of exotic options using Monte Carlo simulation, in particular the development of efficient discretization schemes. Finally, Part III develops pricing formulas for embedded insurance options and performs a quantitative analysis of their valuations. For each part, the remainder of this chapter discusses its motivation, scope and contribution to the literature. Furthermore, it will elaborate how the parts are interconnected with each other.

Part I: Stochastic Interest Rates and Stochastic Volatility

Many of the assumptions of the Black and Scholes (1973) model, like constant volatility and constant interest rates, do not find justification in the financial markets. In particular since the equity crash of the late eighties a battery of complex models has been proposed to improve upon these misspecifications. One class of models relaxes the deterministic volatility assumption and incorporates an empirical financial phenomenon known as volatility clustering, providing more realistic, heavy-tailed asset returns. Other popular approaches involve the use of stochastic interest rates, local volatility or jumps, depending on the specific application. For instance in the case of embedded insurance derivatives, characterized by long maturities, it would be suitable to incorporate stochastic volatility and stochastic interest rates into an option pricing model.

An overview of the literature related to the pricing of long-term options under stochastic volatility and stochastic interest rates is given in Chapter 2, which also provides an short introduction to the methods and techniques used throughout this thesis.

Chapter 3 is concerned with the pricing of long-dated exotic contracts with stochastic volatility and stochastic interest rates. To incorporate the need for stochastic interest rates, the stochastic volatility models of Stein and Stein (1991) and Schöbel and Zhu (1999) are expanded to allow for Hull and White (1993) interest rates and a general correlation structure between the stock price process, its stochastic volatility and interest rates. The resulting Schöbel-Zhu-Hull-White (SZHW) model, can be placed in the general framework for affine models of Duffie et al. (2000) and Duffie et al. (2003), and benefits greatly from the analytical tractability that is typical for this class of models. Our contribution to the existing literature is threefold. First, we derive the characteristic functions of the log-asset price, which enables an efficient closed-form pricing of European options by Fourier inversion. Secondly, since the practical relevance of any model is limited without a numerical implementation, we extensively consider the efficient implementation of the transform inversion required to price European options. In particular we derive the limiting behaviour of the characteristic function of the SZHW model which allows us to calculate the inversion integral much more accurately. Thirdly, we generalize the SZHW model to be able to value foreign exchange options in a framework where both domestic and foreign interest rate processes are stochastic.

To price cross-currency derivatives, such as foreign exchange and inflation options, most

investment banks have standardized on a three-factor modelling framework. Hereby the index has a deterministic volatility, and the interest rates of the two currencies are driven by one-factor Gaussian models. This deterministic volatility assumption, though technically very convenient, does not find justification in the financial markets for equity, FX or inflation options. In fact, the markets for these products exhibit a strong volatility skew or smile, implying log index returns that deviate from normality and suggest the use of skewed and heavy tailed distributions. Moreover, many multi-currency structures are directly exposed to the shape of the volatility structure as they often incorporate multiple strikes as well as callable and knockout components. Examples hereof include Limited Price Index (Inflation) options, Power Reverse Dual Contract (Foreign Exchange) swaps and Pension Funding Ratio (Equity-Interest Rates) options. Suitable derivative pricing models, aimed to quantify this volatility exposure and corresponding hedge costs, should therefore at least be able to incorporate the volatility shapes of the vanilla option markets. Though in the current literature various methods exist to incorporate volatility smiles, to price multi-currency options also a term-structure involving various time points of the forward index is required. The incorporation of stochastic interest rates makes the connection between the two particularly non-trivial and is by Piterbarg (2005) even dubbed as "perhaps even the most important current outstanding problems for quantitative research departments worldwide".

Chapter 4 of this thesis develops a generic multi-currency framework incorporating the need for stochastic interest rates, stochastic volatility, whilst allowing for closed-form calibration formulas for vanilla options. The modelling framework considers the pricing of inflation, foreign exchange and stock options under multi-factor Gaussian interest rates, Schöbel and Zhu (1999) and Heston (1993) stochastic volatility, hereby using a full correlation structure between all driving quantities. Relying on Fourier transform methods, we show that vanilla call and put options, forward starting options, year-on-year inflation-indexed swaps and inflation-indexed caps and floors can be valued in closed-form. We suggest a new calibration algorithm, based on a control variate technique, for the generic Heston model. This new method is compared to the Markovian projection technique of Antonov et al. (2008) and turns out to provide certain advantages over the Markovian projection approximation. Furthermore, we demonstrate that the frameworks are well able to incorporate the markets implied volatility shapes. Finally, due to the generic setup, the multi-currency framework has the additional advantages that it can be used for multi-asset purposes and is fast enough for the real life risk management of big portfolios of inflation, FX, equity, interest rates, commodities and hybrid option contracts.

Part II: Efficient Simulation Methods for Valuing Exotic Derivatives

Once a suitable financial model is selected and calibrated, the next step is to apply it in practice. Though certain models yield closed-form solutions for some contracts, the fast majority of products cannot be priced in closed-form. Monte Carlo methods provide an extremely popular and flexible pricing alternative to value such complex derivatives. Due to technical advances

such as multi-processor programming, increasing computational power and modern day variance reduction techniques, the Monte Carlo technique is expected to become even more widely applicable in the near future. By its nature, however, these methods are relatively time consuming as they are based on repetitive simulations. Many attention of both academics as practitioners is devoted to the efficient simulation schemes aiming to minimize the computational efforts whilst retaining a high degree of accuracy. In Part II of this thesis we deal with efficient discretization methods for stochastic volatility models.

Though an exact simulation method for the Heston (1993) model was developed in Broadie and Kaya (2006), its practical use is limited due to its complexity and lack of speed. For instance, to simulate the Non-central Chi-squared distributed variance process an acceptance and rejection technique is suggested, which hinders the sensitivity analysis and which cannot be used in conjunction with low-discrepancy numbers. Euler discretizations give rise to a completely different category of problems. For example, while the continuous-time variance process of the Heston (1993) model is guaranteed to be non-negative, its Euler discretization is not.

Chapter 5 deals with efficient discretizations schemes for the Heston (1993) stochastic volatility model. The considered schemes improve upon disadvantages of the exact method considered in Broadie and Kaya (2006). To overcome the acceptance and rejection sampling method of the exact scheme, one can create a large three-dimensional cache of the inverse from the non-central chi-squared distribution function for all conceivable values of the number of degrees of freedom, the non-centrality parameters and its function values. Nonetheless, as the parameter-space is potentially very large, such a brute-force caching method is not realizable for practical purposes. Using a conditioning argument, we will however show that the three-dimensional inverse of the Non-central Chi-squared distribution can effectively be reduced to a one dimensional search space for the case of the Heston (1993) model. We develop three new efficient simulation schemes, using this insight. Finally, we perform an extensive numerical comparison between the new methods with other recent schemes. Approximations of the exact scheme based on drift interpolations of the integrated variance process, are found to be several times more efficient than the recent Euler, Kahl and Jäckel (2006) and exact schemes.

A major problem signaled with Euler schemes in the simulation of stochastic volatility models is their inability to generate the proper correlation between the increments of the asset and the stochastic volatility processes. As the correlation parameter in the stochastic volatility models is an important determinant of the skew in implied volatilities, not being able to match this parameter, leads to a significant mispricing of options with strikes far away from the at-the-money level. In the Heston (1993) model, this so-called “leaking correlation” problem, is partially caused by the fact that an Euler discretization tries to approximate a square root process, by a Gaussian process. However even when the stochastic volatility itself is Gaussian, such as in the Schöbel and Zhu (1999) model, the problem of “leaking correlation” is still an issue.

In Chapter 6 discretization schemes are presented for the Schöbel and Zhu (1999) stochastic

volatility model, tailored to match correlation between the increments of the asset and the variance processes of the continuous-time process. Particular attention is given to the 'leaking correlation' issue in the simulation of Heston (1993) and Schöbel and Zhu (1999) models. Furthermore the simulation in the Schöbel-Zhu-Hull-White extension, which incorporates the need for stochastic interest rates, is considered. This is closely related to recent advances in the development of markets for long-term derivatives, described in Part I of this thesis, for which maturities the inclusion of stochastic interest rates in a derivatives pricing model is more suitable. Though the continuous time processes of the considered stochastic volatility models are guaranteed to be a martingale, and hence have finite first moments, this does not necessarily hold for their discretizations. To this end, we derive conditions for the regularity of the developed discretization schemes and investigate how to ensure the exact martingale property. Finally, we numerically compare the new simulation schemes to other recent schemes in the literature. For a special case of the Schöbel and Zhu (1999) model which coincides with the Heston (1993) model, our proposed scheme has a similar performance to the QE-M scheme of Andersen (2008), whilst being slightly more efficient in terms of computational time required. For the Schöbel-Zhu cases not coinciding with the Heston model, it is found that our scheme consistently outperforms the Euler scheme. These results affirm that Andersen's result is more widely applicable than to the Heston model alone; for the simulation of stochastic volatility models, it is of great importance to match the correlation between the asset price and its stochastic volatility process.

Part III: Applications to Insurance Markets

The third and last part of this thesis is concerned with the pricing of two popular type of contracts appearing in insurance markets. Using the methods of Part I and II of this thesis, we investigate the impact of stochastic volatility, stochastic interest rates and a general correlation structure on the valuation of insurance contracts. By developing closed-form solutions for the prices of forward starting and guaranteed annuity options, we are able to carry out a quantitative analysis on the pricing of these embedded options. The analysis performed in Chapter 7 and 8 stands out, compared to the existing literature, by taking both stochastic volatility and stochastic interest rates explicitly into account.

Forward starting options form the basis for many Unit-Linked guarantees, cliquet and ratchet options. Due to their popularity, these products recently attracted a lot of attention from both academics and practitioners. Forward starting contracts belong to the class of path-dependent European-style contracts in the sense that they not only depend on the terminal value of the underlying asset, but also on the asset price at an intermediate point. For instance, a forward starting option may provide the holder a call option with a strike equal to a fixed proportion of the underlying asset price at some intermediate date. These options are frequently used by insurance companies to hedge year on year Unit-Linked guarantees, but also many structured products, tailored for investors seeking for upside potential whilst preserving protection against downside movements, involve such forward starting optionalities.

Chapter 7 is concerned with a quantitative analysis on the valuation of forward starting options, where we explicitly account for stochastic volatility, stochastic interest rates and a general correlation structure between all underlying processes. This analysis is facilitated by the development of closed-form formulas, based on Fourier inversions, for these contracts. Compared to vanilla options, forward starting structures are much more sensitive to future interest rate movements, volatility smiles as well as their correlation structure with the underlying asset. It is found that it is important to take stochastic interest rates, volatility and a general correlation structure into account for a proper valuation and hedging of these securities: ignoring one of these aspect can lead to serious mispricings and hedge errors.

In Chapter 8 the pricing of guaranteed annuity options (GAOs) is investigated using a stochastic volatility model for equity prices. GAOs are options providing the right to convert a policyholder's accumulated funds to a life annuity at a fixed rate when the policy matures. These options were a common feature in UK retirement savings contracts issued in the 1970's and 1980's when interest rates were high, but they caused problems for insurers as the interest rates began to fall in the 1990's. Currently, these options are frequently sold in the U.S. and Japan as part of variable annuity products. Until now, for the pricing of these options generally a geometric Brownian motion for equity prices is assumed. However, given the long maturities of the insurance contracts a stochastic volatility model for equity prices, providing more realistic equity returns, would be more suitable.

The contribution of Chapter 8 is threefold. First, closed-form expressions are derived for prices of GAOs assuming stochastic volatility for equity prices and either a one-factor or two-factor Gaussian interest rate model. Secondly, we come up with a more efficient GAO pricing formula, than considered in Chu and Kwok (2007), for an equity model with constant volatility. Under two-factor Gaussian rates, these authors argue that no analytical pricing formula exists and hence propose several approximation methods for its valuation. In this chapter we do derive an exact closed-form pricing formula in terms of a single numerical integral. This method is preferable compared to these latter approaches, as it gives exact GAO prices over all strike levels whilst being computational very efficient to compute. Finally, using U.S. and EU market option data, we investigate the effects of a stochastic volatility model on the pricing of GAOs. For both markets, the results indicate that the impact of ignoring a stochastic volatility model can be significant.

Part I

Stochastic Volatility and Stochastic Interest Rates

Chapter 2

Stochastic Models

2.1 Affine Diffusions

In this section we discuss the class of affine models, as this class of models is frequently used throughout this thesis. The class of affine models, was introduced in Duffie and Kan (1996), in an interest rate framework, and was later generalized in great detail by Duffie et al. (2000) and Duffie et al. (2003). In the remainder of this section we discuss the subclass of affine diffusions, and its Fourier transform. This subclass is not as general as the affine jump-diffusion processes considered in Duffie et al. (2003), but is sufficient for our purposes. The popularity of the affine class of models arises from their flexibility and generic setup combined with a great analytical tractability, which facilitates the calibration and simulation of such models. Eminent members of this class are the term structure models of Hull and White (1993) and Cox et al. (1985), the Black and Scholes (1973) model, and also the stochastic volatility models of Heston (1993) and Schöbel and Zhu (1999).

The class of affine diffusions can be characterized as follows; let X be a real-valued n-dimensional Markov process satisfying

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t), \tag{2.1}$$

where $W(t)$ is a standard Brownian motion in $\mathbb{R}^n$, $\mu(X(t)) \in \mathbb{R}^n$ and $\sigma(X(t)) \in \mathbb{R}^{n\times n}$. Then, we call the process X affine if and only if the diffusion coefficients are of the following form

$$\begin{aligned} \mu(x) &= K_0 + K_1 x, \quad \text{for } K = (K_0, K_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}, & (2.2)\\ \left(\sigma(x)\sigma^T(x)\right)_{ij} &= (H_0)_{ij} + (H_1)_{ij} \cdot x, \quad \text{for } H = (H_0, H_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}. & (2.3) \end{aligned}$$

In addition, the short term interest rate $r(X(t))$ is also assumed to be affine:

$$r(x) = \rho_0 + \rho_1^T x, \quad \text{for } R = (\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^n. \tag{2.4}$$

In total, for a process $X(t)$ to be affine, both its instantaneous drift, variance and short rate need to be an affine combination of the factors.

The first key result in Duffie et al. (2000) is that the characteristic function (and the moment generating function, when it exists) of $X(t)$, is known in closed-form up to the solution of a system of Ordinary Differential Equations. That is, for $u \in \mathbb{C}^n$ the Fourier transform $\phi(u, X(t), t, T)$ of $X(t)$ is given by

$$\phi(u, X(t), t, T) = \mathbb{E}\left[e^{-\int_t^T r(X(s))ds} e^{uX(T)} \middle| \mathcal{F}_t\right] = e^{A(u,t,T)+B(u,t,T)\cdot X(t)}, \tag{2.5}$$

where A and B solve the following system of Riccati equations

$$\frac{dA(u,t,T)}{dt} = \rho_0 - K_0 \cdot B(u,t,T) - \frac{1}{2}B^T(u,t,T)H_0B(u,t,T), \tag{2.6}$$

$$\frac{dB(u,t,T)}{dt} = \rho_1 - K_1^T B(u,t,T) - \frac{1}{2}B^T(u,t,T)H_1B(u,t,T). \tag{2.7}$$

subject to the terminal condition $A(u,T,T) = 0$ and $B(u,T,T) = u$. In general the solutions to A and B to the system of Ordinary Differential Equations (2.6) have to be computed numerically, see e.g. Press and Flannery (1992). Of course models for which A and B can be computed in closed-form have a large advantage.

In the context of stochastic volatility models, expression (2.5) provides an expression for the characteristic function of the logarithm of the asset price. Closed-form prices of European options can be obtained by inverting this characteristic function, see Section 2.5. We will provide the characteristic functions of the considered models in this thesis when start using them. For example, the characteristic functions of the Schöbel-Zhu and Heston model, extended with stochastic interest rates, are derived in Chapter 3 and 4.

2.2 Stochastic Interest Rates

There is a vast literature on modelling the term structure of interest rates. Approaches for the pricing and risk management of interest rate derivatives have been described by Vasicek (1977), Hull and White (1993), Cox et al. (1985), Heath et al. (1992), Miltersen et al. (1977), Brace et al. (1977), Hunt et al. (2000). These interest rate models can be subdivided into short rate, Heath-Jarrow-Morton, market and Markov functional models. We refer to Brigo and Mercurio (2006), Pelsser (2000) for an extensive overview of interest rate modelling literature. It goes beyond the scope of this thesis to discuss all (dis)advantages of each of these framework, which for example can be judged in terms of its dimensionality, number of parameters versus calibration quality, correlation structures, or its analytical tractability. However, above all, it is important that the main price sensitivities of the considered contract are properly captured by the chosen interest rate model.

In this thesis we restrict ourselves most of the time to short rate models, as they are sufficiently able capture the desired interest rate characteristics and often allow for a tractable pricing of

long-term options. An additional advantage of short rate models is that they can be placed in the affine class and therefore benefit from the analytical properties typical for this class. In affine term structure models, one can for example express the price of a zero coupon bond maturing at time T into an exponential affine form of its state variable(s), i.e.

$$P(t,T) = e^{A(u,t,T)+B(u,t,T)\cdot X(t)}, \tag{2.8}$$

which can be obtained from (2.5)-(2.7). As the future term structure of interest rates, implied by the discount factors, is of crucial importance for the pricing of interest rate derivatives, this result is extremely useful. Finally, note that in many short models, for instance in the models of Vasicek (1977), Hull and White (1993) and Cox et al. (1985), closed-form solutions exists for A and B.

2.2.1 Change of Numeraire

In an arbitrage-free and complete market the value of any contingent claim can be uniquely determined as the expectation of the payoff normalized by the money market account under a unique equivalent measure, see Harrison and Kreps (1979), Harrison and Pliska (1981). Under this measure the expected return on all assets is equal to the risk-free rate, hence this measure is dubbed as "the risk-neutral measure", denoted here by Q. The normalizing asset, in these papers the money market account, is called the numeraire. In Geman et al. (1996) it is shown that not only the money market account can be used as numeraire, but every strictly positive self-financing portfolio of traded assets, can be used as numeraire. Their change of numeraire technique demonstrates how to change from one numeraire to another by switching to a different probability measure. As a byproduct every positive non-dividend paying asset divided by its numeraire, is a martingale under the measure associated with that numeraire.

For example, the measure change from the money market measure Q to the T-forward measure Q^T can be established by the following Radon-Nikodym derivative

$$\frac{dQ^T}{dQ} = \frac{\exp\left[-\int_0^T r(u)du\right]}{P(0,T)},$$

and the corresponding T-forward measure, see Geman et al. (1996), hence uses the time T zero-coupon bond price $P(0,T)$ as numeraire. In particular for the pricing time-T payoffs, it can be convenient to use the time-T discount factor as numeraire, since the numeraire is then equal to one at the time of maturity.

2.2.2 Hull-White model

To be able to exactly fit the currently observed term structure of interest, Hull and White (1993) extended the Vasicek (1977) model with a time-varying mean reversion level. That is, under

the money market measure Q (which uses the risk-free interest rate of the bank account as numeraire), the Hull and White (1993) model provides following dynamics

$$dr(t) = [\theta(t) - ar(t)]\,dt + \sigma dW_r(t), \tag{2.9}$$

with a,σ two constants, $W_r(t)$ a Brownian motion and where the deterministic function $\theta(t)$ is chosen to exactly fit the currently observed term structure of interest rates. As the short rate follows an Ornstein-Uhlenbeck process, this model implies a normal distribution for the short interest rate at each time. The Gaussian distribution allows for the derivation of analytical formulas and the construction of efficient numerical methods for the pricing of various interest rate derivatives. On the other, the theoretical possibility of r dropping below zero, due to the Gaussian distribution, is a clear drawback of the model. However, in practice, the probability of negative rates in this model is often negligible , e.g. see Brigo and Mercurio (2006), and even so, does not lead to any issues for most products.

Following Pelsser (2000), Brigo and Mercurio (2006), the Hull and White (1993) model can more conveniently be parameterized as

$$r(t) = x(t) + \beta(t), \tag{2.10}$$

which separates the interest rate process $r(t)$ into a deterministic function β and a standard Ornstein-Uhlenbeck process $x(t)$ satisfying

$$dx(t) = -ax(t)dt + \sigma dW_r(t), \tag{2.11}$$

In particular, in order to exactly fit the currently observed term structure of interest rates, we have that

$$\int_t^T \beta(u)du = \ln\frac{P(0,t)}{P(0,T)} + \frac{1}{2}\left[V(0,T) - V(0,t)\right], \tag{2.12}$$

with $V(t,T)$ defined as

$$V(t,T) := \frac{\sigma^2}{a^2}\left[T - t + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right]. \tag{2.13}$$

For the Hull and White (1993) model, one can use the normality and (2.5) to obtain the following expression for the price of a zero-coupon bond $P(t,T)$ maturing at time T

$$P(t,T) = A(t,T)e^{-B(t,T)x(t)}, \tag{2.14}$$

with:

$$A(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left[\frac{1}{2}\Big(V(t,T) - V(0,T) + V(0,t)\Big)\right] \tag{2.15}$$

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a} \tag{2.16}$$

and with $P(0, s)$ denoting the market's time zero discount factor maturing for time s.

By applying the change of numeraire technique of Section 2.2, one can change the underlying numeraire from the money market account to the time T discount bond price and evaluate the model dynamics under the T-forward measure Q^T. Girsanov's theorem implies that in the Hull and White (1993) model the process $W_r^T(t)$ defined by

$$dW_r^T(t) = dW_r(t) + \sigma B(t, T)dt, \tag{2.17}$$

is a Brownian motion under the T-forward measure. An explicit solution for $x(t)$, starting from time s, under Q^T is hence given by

$$x(t) = x(s)e^{-a(t-s)} - M^T(s, t) + \sigma \int_s^t e^{-a(t-u)} dW_r^T(u), \tag{2.18}$$

where $M^T(s, t)$ is a deterministic function satisfying

$$M^T(s, t) = \frac{\sigma^2}{a^2}\Big(1 - e^{-a(t-s)}\Big) - \frac{\sigma^2}{2a^2}\Big(e^{-a(T-t)} - e^{-a(T+t-2s)}\Big). \tag{2.19}$$

Therefore, we obtain from Itô's isometry that $r(t)$, conditional on time s under the T-forward measure, is normally distributed with mean $\mu_r(s, t)$ and variance $\sigma_r^2(s, t)$ given by

$$\mu_r(s, t) = x(s)e^{-a(t-s)} - M^T(s, t) + \beta(t), \tag{2.20}$$

$$\sigma_r^2(s, t) = \frac{\sigma^2}{2a}\Big(1 - e^{-2a(t-s)}\Big). \tag{2.21}$$

Using these dynamics, and the fact that the numeraire follows a log-normal distribution in this model, many vanilla options can be priced in closed-form using Black and Scholes (1973) style formulas, see e.g. Hull and White (1993), Pelsser (2000) or Brigo and Mercurio (2006). Also the simulation of the model can be done in exact fashion as the underlying state variables follow a (joint) normal distribution, see e.g. Glasserman (2003). In general, the Hull and White (1993) model has the advantage, and owes its popularity, from the fact that it is very tractable.

2.3 Stochastic Volatility

Since the introduction of the Black and Scholes (1973) model and in particular since the equity crash of the late eighties a battery of complex models has been proposed to relax some misspecifications of the model. Though the Black and Scholes (1973) model has theoretical and practical appealing properties, most of its assumptions, like constant volatility or constant interest rates, do not find justification in the financial markets. One class of models relaxes the constant volatility assumption and incorporates a financial phenomenon known as volatility clustering, i.e. they make volatility stochastic. Examples of models belonging to this class are the stochastic volatility models of Hull and White (1987), Stein and Stein (1991), Heston (1993) and Schöbel and

Zhu (1999). In Heston (1993) it was first shown that the price for call/put options are given in terms of numerical integrals over the characteristic function of the logarithm of the asset price, thus, provided that the characteristic function can be computed in an efficient way, this allows for efficient calibrations to market option data. Note that the option pricing formula presented in the original Heston (1993) paper requires the evaluation of two numerical integrals, whereas more recent methods only use one numerical integral, see Section 2.5.

2.3.1 Heston's and Schöbel-Zhu's Stochastic Volatility Model

In the stochastic volatility class, the models of Heston (1993) and Schöbel and Zhu (1999) stand out by allowing for flexibility between the correlation of the asset price and its stochastic volatility, whilst they also provide a closed-form formula for their characteristic function. As these two models play a prominent role in the rest of this thesis, we provide their model dynamics here and discuss their relationships.

The Heston (1993) stochastic volatility model assumes the following dynamics

$$\begin{aligned} dS(t) &= S(t)\left[\mu(t)dt + \sqrt{v(t)}\,dW_S(t)\right], && (2.22)\\ dv(t) &= \kappa(\theta - v(t))dt + \xi\sqrt{v(t)}\,dW_V(t), && (2.23) \end{aligned}$$

In this stochastic differential equation, $S(t)$ represents the asset price, with a stochastic variance $v(t)$ that follows a mean reverting, square-root/Feller/CIR process. The Brownian motions $W_S(t)$ and $W_V(t)$ are correlated with correlation coefficient ρ. In the Schöbel and Zhu (1999) stochastic volatility model, not the variance, but the volatility is modelled via a mean reverting, Ornstein-Uhlenbeck process:

$$\begin{aligned} dS(t) &= S(t)\left[\mu(t)dt + v(t)dW_S(t)\right], && (2.24)\\ dv(t) &= \kappa[\psi - v(t)]dt + \tau dW_v(t), \quad v(0) = v_0, && (2.25) \end{aligned}$$

and the Brownian motions are again correlated with coefficient ρ. We postpone the interpretation of the above model parameters to later chapters.

At first sight, one curious property of the Schöbel and Zhu (1999) model is that the volatility process $v(t)$ affects the sign of the instantaneous correlation between $v(t)$ and $\ln S(t)$. Indeed, one can show that

$$\mathrm{Corr}\Big(d\ln S(t), dv(t)\Big) = \frac{\rho_{Sv}v(t)\tau}{\sqrt{v^2(t)\tau^2}} = \rho_{Sv}\,\mathrm{sgn}\Big(v(t)\Big), \tag{2.26}$$

which will effectively cause the correlation between $v(t)$ and the stock price $S(t)$ to change sign. This effect is visualized in Figure **1**, where we have plotted a sample path of $\ln S(t)$, $v(t)$ and $|v(t)|$.

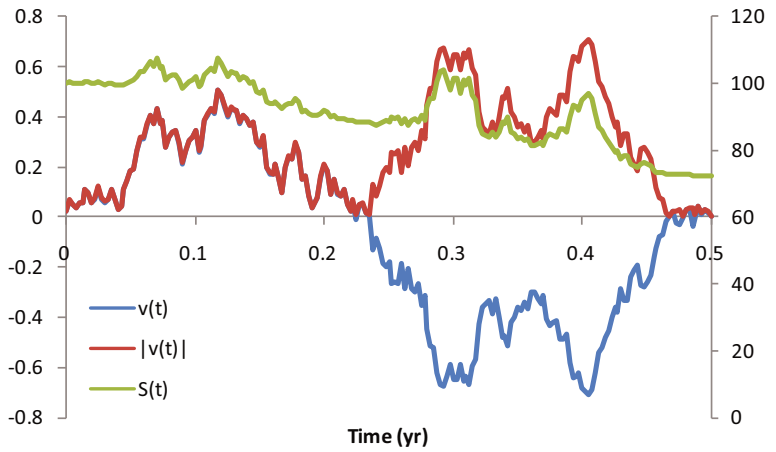

Figure 1: Sample path of $\ln S(t)$, $\nu(t)$ and $|\nu(t)|$. SZ parameters are $\kappa = \tau = 1$, $\nu(0) = \psi = 25\%$, $\rho = 1$, x(0)=100.

Indeed, when $\nu(t)$ is negative and decreasing, the asset price is increasing, contrary to what one would expect from the parameter configuration. The key lies therein that $\nu(t)$ should not be interpreted as the volatility of the underlying asset. It is merely a latent variable which drives the true volatility of the asset, the true volatility being defined as the square root of the instantaneous variance of the logarithm of the stock price. Using the Itô-Tanaka theorem (see Revuz and Yor (1999)), one can indeed show that the correlation between $\ln S(t)$, and $|\nu(t)|$ is equal to $\rho_{S\nu}$, as we would like it to be.

If we take the pair $(S(t), \nu(t))$ as state variables, the Heston (1993) and Schöbel and Zhu (1999) model are not affine as the instantaneous variance of the stock price is equal to $\nu(t)S^2(t)$. Nonetheless, if we consider $\left(\ln S(t), \nu^2(t)\right)$ as state variables for the Heston (1993) model, and $\left(\ln S(t), \nu(t), \nu^2(t)\right)$ for the Schöbel and Zhu (1999) model, we do obtain affine processes. The characteristic functions of these models will be provided when we start using these models in later chapters. For instance the characteristic function of the Schöbel and Zhu (1999) model is derived in Chapter 3.

Relationship between the Heston and Schöbel-Zhu model

It was already noted in Heston (1993), that an Ornstein-Uhlenbeck process for the volatility is closely related to a square-root process for the variance process. If the volatility $\nu(t)$ is modelled by an Ornstein-Uhlenbeck process, as in the Schöbel and Zhu (1999), then Itô's lemma shows that the variance process $\nu^2(t)$ follow the dynamics

$$d\nu^2(t) = 2\kappa\Big(\frac{\tau^2}{2\kappa} + \psi\nu(t) - \nu^2(t)\Big)dt + 2\tau\nu(t)dW_\nu(t). \tag{2.27}$$

From the variance dynamics (2.23), the relationship between the Heston and Schöbel-Zhu model can directly be observed: in case the long-term mean ψ of the volatility process in (2.25) is equal to zero, Schöbel-Zhu model is equivalent to the Heston model in which $\lambda = 2\kappa$, $\xi = 2\tau$ and $\theta = \frac{\tau^2}{2\kappa}$. The overlap of the models is restricted to this very special case. However one can use this relationship to establish a relation between the characteristic functions of the Heston and the Schöbel-Zhu(-Hull-White) model, see Lord and Kahl (2007): it turns out that the characteristic function of SZHW model can be written as the Heston characteristic multiplied by a simple extra factor.

2.4 Volatility, Correlation and Stochastic Interest Rates

It is hardly necessary to motivate the inclusion of stochastic volatility in a derivative pricing model for long-term options. Also the addition of interest rates as a stochastic factor is important when considering long-maturity derivatives and has been the subject of empirical investigations, most notably by Bakshi et al. (2000). Approaches that covered both stochastic volatility and stochastic interest rates in derivative pricing were described in Scott (1997), Bakshi et al. (1997), Amin and Ng (1993), Andreasen (2006), van der Ploeg (2007), Ahlip (2008), Antonov et al. (2008) and Grzelak and Oosterlee (2009). None of these models, however, provide both an exact closed-form call pricing formula, and also explicitly incorporate the correlation between underlying asset and the term structure of interest rates.

Scott (1997), Bakshi et al. (1997), Amin and Ng (1993) consider independent interest rates or do not derive closed-form formulas for vanilla options. Andreasen (2006) does derive closed-form Fourier expressions for vanilla options, similar to Section 2.5, but uses an indirect approach in the form of a volatility displacement parameter to correlate the independent Hull and White (1993) interest rate drivers with the underlying FX rate and which can lead to unrealistic model parameters, see e.g. Antonov et al. (2008). In van der Ploeg (2007) exact closed-form solutions are derived for various models, but in all cases only under simplifying assumptions on the correlation structure between the stochastic volatility and the corresponding stochastic interest rates model. Ahlip (2008) considers an extension of the Schöbel-Zhu model with Gaussian interest rates for the pricing of exchange rate options, but assumes a perfect correlation between all stochastic processes, considering essentially only a one factor model. Antonov et al. (2008) does consider a full correlation structure between the stochastic interest rates, the underlying FX rate and uses a Heston (1993) stochastic volatility under a full correlation structure. To price vanilla options, however, approximations are being made, which deteriorate for larger maturities or more extreme model parameters. Grzelak and Oosterlee (2009) consider the same modelling framework as Antonov et al. (2008), but follow a different approach, extending the work of Giese (2004), by using approximations for the covariance term between the volatility and the interest rates. To obtain the characteristic function of logarithm of the asset price, using the approximations, one then repeatedly has to solve for a system of ODEs using numerical methods. Analogously, this method is further generalized in Grzelak et al. (2009) in a framework where the single-factor short interest rates are replaced by their corresponding multi-factor

models.

In the Black-Scholes-Hull-White (BSHW) model, in which the underlying asset has a deterministic volatility, one can derive closed-form formulas for vanilla options and allow for stochastic interest rates, which are correlated with the underlying asset. In the literature this model is used for the pricing of long-term options. For example in the context of Guaranteed Annuity Options, see also Chapter 8, the BSHW model is being used in Ballotta and Haberman (2003), Boyle and Hardy (2003) and Chu and Kwok (2007). By many authors, e.g. see Ballotta and Haberman (2003) and Piterbarg (2005) it is however noted that, given the long maturities typical in such contracts, a stochastic volatility model for the underlying asset would be more suitable. To this end, frameworks that allow for both stochastic volatility, stochastic interest rates, a full correlation structure and an exact closed-form pricing of equity, FX and inflation options are derived in Chapter 3 and 4. As the BSHW model is standard in the literature, and serves as a useful benchmark, we discuss it briefly in the next section.

2.4.1 Black-Scholes-Hull-White model

The BSHW model combines the Black and Scholes (1973) and the Hull and White (1993) model and is applicable for the pricing of equity and hybrid options, see for instance Ballotta and Haberman (2003) and Brigo and Mercurio (2006). A trivial extension of the model is used for the pricing of cross-currency derivatives, e.g. see Sippel and Ohkoshi (2002), Jarrow and Yildirim (2003) and Piterbarg (2005). Under the risk-neutral measure Q, which uses the bank account as numeraire, the dynamics of stock price $S(t)$ and short interest rate $r(t)$ in the BSHW model are given by

$$\begin{aligned} dr(t) &= [\theta(t) - ar(t)]\, dt + \sigma dW_r(t), && (2.28)\\ dS(t) &= S(t)\left[r(t)dt + \eta dW_S(t)\right], && (2.29) \end{aligned}$$

where $W_S(t)$ and $W_r(t)$ are two Brownian motions, with correlation ρ. Here the short interest rate $r(t) = x(t) + \beta(t)$ follows a Hull and White (1993) process, see Section 2.2.2. Using Geman et al. (1996), analogously to Section 2.2, we can switch from the money market measure Q to the T-forward measure Q^T, which implies the following model dynamics:

$$\begin{aligned} dx(t) &= -ax(t)dt - \sigma^2 B(t,T)dt + \sigma dW_r^T(t), && (2.30)\\ dS(t) &= S(t)\left[r(t)dt + \eta\rho\sigma B(t,T)dt + \eta dW_S^T(t)\right]. && (2.31) \end{aligned}$$

So, conditional on time s, using Itô's lemma and direct integration, this gives us following explicit solutions under Q^T:

$$\begin{aligned} x(t) &= x(s)e^{-a(t-s)} - \int_s^t e^{-a(t-u)}\sigma^2 B(u,T)du + \int_s^t \sigma e^{-a(t-u)} dW_r^T(t) && (2.32)\\ S(t) &= S(s)\exp\left[\int_s^t (r(u) - \rho\eta\sigma B(u,T) - \frac{1}{2}\eta^2)du + \int_s^t \eta dW_S^T(u)\right]. && (2.33) \end{aligned}$$

For every derivatives pricing model it is important that efficient, preferably closed-form, formulas are available for the calibration and pricing of vanilla options. We therefore demonstrate how vanilla call and put options can be priced in closed-form in the BSHW model.

starting from time t, we can evaluate the price of a European stock option ($\omega = 1$ for a call option, $\omega = -1$ for a put option) with strike K as

$$C(S(t), K, T, \omega) = \mathbb{E}^{Q}\Big[\exp[-\int_t^T r(u)du](w(S(T) - K))^+ \Big| \mathcal{F}_t\Big]. \tag{2.34}$$

Instead of evaluating expected discounted payoff under the risk-neutral bank account measure, we have shown in Section 2.2 how to change the underlying probability measure and evaluate this expectation under the T-forward probability measure Q^T, which is equivalent to choosing the T-discount bond as numeraire.

$$C(S(t), K, T, \omega) = P(t,T)\mathbb{E}^{Q^T}\Big[\Big(\omega(S(T) - K)\Big)^+ \Big| \mathcal{F}_t\Big]. \tag{2.35}$$

As $S(T)$ follows a log-normal distribution under Q^T, we can evaluate the latter expectation by means of the Black and Scholes (1973) formula, for which we need to determine the mean of $S(T)$ and variance of $\ln S(T)$ under the T-forward measure Q^T;

To this end, first note that from (2.10), (2.30), (2.12), Fubini's theorem and some algebra, one can obtain that

$$\begin{aligned}
\int_t^T r(u)du &= \int_t^T \beta(u)du + \int_t^T x(u)du \\
&= \ln\frac{P(0,t)}{P(0,T)} + \frac{1}{2}\left[V(0,T) - V(0,t)\right] \\
&\quad + B(t,T)x(t) - V(t,T) + \int_t^T \sigma B(u,T)dW_r^T(u),
\end{aligned}$$

so that under the T-forward measure Q^T we have

$$\begin{aligned}
\frac{S(T)}{S(t)} &= e^{\int_t^T \left(r(u) - \rho\eta\sigma B(u,T) - \frac{1}{2}\eta^2\right)du + \int_s^t \eta dW_S^T(u)} \\
&= \frac{P(0,t)}{P(0,T)} e^{B(t,T)x(t) + \frac{1}{2}\left[V(0,T) - V(0,t) - V(t,T)\right] - \frac{1}{2}v_S^2(t,T) + \sigma\int_t^T B(u,T)dW_r^T(u) + \eta\int_t^T B(u,T)} \\
&= \frac{1}{P(t,T)} e^{-\frac{1}{2}v_S^2(t,T) + \int_t^T \sigma B(u,T)dW_r^T(u) + \eta\int_t^T B(u,T)dW_S^T(u)},
\end{aligned} \tag{2.36}$$

which implies that $S(T)$ follows a log-normal distribution with mean and "log variance" given

by

$$F(t,T) := \mathbb{E}^{Q^T}\left[S(T)\middle|\mathcal{F}_t\right] = \frac{S(t)}{P(t,T)}, \tag{2.37}$$

$$v_S^2(t,T) := \mathrm{Var}\left[\ln S(T)\middle|\mathcal{F}_t\right] = V(t,T) + 2\rho\eta\int_t^T \sigma B(u,T) + \eta^2(T-t). \tag{2.38}$$

Omitting the dependence on t, T in F and v_S, we can express the option price (2.35) for a European stock option with strike K as in terms of the Black and Scholes (1973) formula as

$$C(S(t),K,T,\omega) = P(t,T)\Big[F\omega N(\omega d_1(K,F,v_S)) - K\omega N(\omega d_2(K,F,v_S))\Big], \tag{2.39}$$

$$d_1(K,F,v_S) = \frac{\ln(F/K) + v_S^2/2}{v_S}, \tag{2.40}$$

$$d_2(K,F,v_S) = d_1(K,F,v_S) - v_S, \tag{2.41}$$

where $\omega = 1$ for a call and $\omega = -1$ for a put option.

Note that the difference $v_S^2(t,T) - \eta(T-t)$, typically is increasing function of the time to maturity. Consequently, the larger the maturity, the larger the impact of the stochastic interest rates on the option price. Therefore, as expected, long-term options are more sensitive to the behaviour of the stochastic interest rates than shorter dated derivatives.

2.5 Characteristic Function and Option Pricing

The use of transform analysis and inversions of characteristic functions for option pricing, was pioneered by Heston (1993). His stochastic volatility option pricing model provided closed-form expressions for the price of vanilla options in terms by inverting the characteristic function of the logarithm of underlying asset. Prior to Heston, Stein and Stein (1991) used the characteristic function to calculate the stock price distribution in their stochastic volatility model. However, whereas the approach of Stein and Stein (1991) heavily relies on the independence between the asset and its stochastic volatility, the approach of Heston (1993) is applicable to any model in which the characteristic function of the logarithm of the asset price is known in closed form.

Under the T-forward measure Q^T, starting from time t, the price of a European option on the underlying asset with strike K is given by

$$C(S(t),K,T) = P(t,T)\mathbb{E}^{Q^T}\left[\left(\omega\left[S(T)-K\right]\right)^+\middle|\mathcal{F}_t\right], \tag{2.42}$$

with $P(t,T)$ the price of a zero coupon bond, $\mathcal{F}_t$ the time-t filtration and with $\omega = 1$ for a call and $\omega = -1$ for a put option. The price of the above option can be expressed in closed-form by inverting the characteristic function of the logarithm of the underlying asset. For instance, Heston (1993) expresses the option price in terms of a generalized version of the Black and

Scholes (1973) formula, in which cumulative probabilities are expressed as integrals over the forward characteristic function $\phi(u)$,

$$\phi(u) = \mathbb{E}^{Q^T}\Big[e^{iu \ln S(T)} \Big| \mathcal{F}_t\Big]. \tag{2.43}$$

Though the approach of Heston (1993) is general and holds for every option pricing model in which the forward characteristic function of the logarithm of the asset price is known, it suffers from large numerical complications, see e.g. Lewis (2001), Lord and Kahl (2008).

To allow for more flexibility and greater numerical efficiency, Carr and Madan (1999) found an alternative representation for the European call price (2.42). These authors take the Fourier transform of the dampened call option price $c_T(k) := \exp(\alpha k) C_T(k)$ with respect the logarithm of the strike price K as:

$$\psi(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk, \tag{2.44}$$

the Fourier transform of which, after using Fubini's theorem and some algebra, can be explicitly expressed in terms of the forward characteristic function $\phi(u)$, i.e.

$$\psi(v) := \frac{\phi\Big(v - (\alpha+1)i\Big)}{(\alpha + iv)(\alpha + 1 + iv)}. \tag{2.45}$$

By inverting the Fourier transform and undampening the call option price the following result is obtained:

$$C(S(t), K, T) = P(t,T)\frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\Big(e^{-(\alpha+iv)k}\psi_T(v)\Big)dv + R\Big(F(t), K, \alpha\Big), \tag{2.46}$$

with forward price $F(t) := \frac{S(t)}{P(t,T)}$ and residue term R, see Lewis (2001) and Lord and Kahl (2008), equal to

$$R\Big(F(t), K, \alpha\Big) := F(t) \cdot 1_{\{\alpha \leq 0\}} - K \cdot 1_{\{\alpha \leq -1\}} - \frac{1}{2}\Big(F(t) \cdot 1_{\{\alpha=0\}} - K \cdot 1_{\{\alpha=-1\}}\Big). \tag{2.47}$$

The dampening parameter α is used to ensure that the dampened call price $c_T(k)$ is L^1 integrable, which is a sufficient condition for the Fourier transform to exist.

To make the existence more explicit, a sufficient condition for $c_T(k)$ to be integrable is provided by $\psi(0) = \int_{-\infty}^{\infty} c_T(k)dk$ being finite, see Carr and Madan (1999). From (2.45) one can observe that $\psi(0)$ is finite provided that

$$\phi(-(\alpha+1)i) = \mathbb{E}^{Q^T}\Big[S(T)^{\alpha+1}\Big] < \infty. \tag{2.48}$$

For the Fourier transform (2.46) to exist, it therefore suffices that the $(\alpha+1)$-th moment of the asset price is finite.

The option pricing result (2.46) is analyzed in great detail by Lee (2004), Lewis (2001) and Lord and Kahl (2008). These papers conclude that an optimal choice of α, minimizing sampling and truncation error, depends both on the chosen model and its parameters, as well as the strike of the underlying option.

Chapter 3

The Schöbel-Zhu-Hull-White Model

*This chapter is based on:

Alexander van Haastrecht, Roger Lord, Antoon Pelsser and David Schrager (2009). "Pricing Long-dated Insurance Contracts with Stochastic Interest Rates and Stochastic Volatility", *Insurance: Mathematics and Economics, vol. 45, no. 3, pp. 436-448.*.

3.1 Introduction

The derivative markets are maturing more and more. Not only are increasingly exotic structures created, the markets for plain vanilla derivatives are also growing. One of the recent advances in equity derivatives and exchange rate derivatives is the development of a market for long-maturity European options[1]. In this chapter we develop a stochastic volatility model aimed at pricing and risk managing long-maturity insurance contracts involving equity, interest rate and exchange rate risks.

We extend the models by Stein and Stein (1991) and Schöbel and Zhu (1999) to allow for Hull and White (1993) stochastic interest rates as well as correlation between the stock price process, its stochastic volatility and interest rates. The resulting model is dubbed as the Schöbel-Zhu Hull-White (SZHW) model. Our model enables to take into account two important factors in the pricing of long-maturity equity or exchange rate derivatives: stochastic volatility and stochastic interest rates, whilst also taking into account the correlation between those processes explicitly. It is hardly necessary to motivate the inclusion of stochastic volatility in a derivative pricing model. The addition of interest rates as a stochastic factor is important when considering long-maturity equity derivatives and has been the subject of empirical investigations most notably by Bakshi et al. (2000). These authors show that the hedging performance of delta hedging strategies of long-maturity options improves when taking stochastic interest rates into account, whilst the interest rate risk is less important for short maturity options. This result is also intuitively appealing since the interest rate risk of equity derivatives, the option's rho,

[1]The implied volatility service of MarkIT, a financial data provider, shows regular quotes on a large number of major equity indices for option maturities up to 10-15 years.

is increasing with time to maturity. The SZHW model can be used in the pricing and risk management for a range of insurance and exotic derivatives contracts. One can for example think of pension products, variable and guaranteed annuities (e.g. see Ballotta and Haberman (2003)), long-maturity PRDC FX contracts (e.g. see Piterbarg (2005)), rate of return guarantees in Unit-Linked contracts (e.g. see Schrager and Pelsser (2004)) and many other structures which have a long-term nature.

The model of this chapter can be placed in the derivative pricing literature on stochastic volatility models as it adds to or extends work by Stein and Stein (1991), Heston (1993), Schöbel and Zhu (1999) or, since our model can be placed in the affine class, in the more general context of Duffie et al. (2000), Duffie et al. (2003) and van der Ploeg (2006). The SZHW model benefits greatly from the analytical tractability typical for this class of models. Our work can also be viewed as an extension of the work by Amin and Jarrow (1992) to stochastic volatility. In a related work Ahlip (2008) considers an extension of the Schöbel-Zhu model to Gaussian stochastic interest rates for pricing of exchange rate options. Upon a closer look however the correlation structure considered in that work is limited to perfect correlation between the stochastic processes. The affine stochastic volatility models fall in the broader literature on stochastic volatility which covers both volatility modelling for the purpose of derivative pricing as well as real world volatility modelling. Previous work that covered both stochastic volatility and stochastic interest rates in derivative pricing include: Scott (1997), Bakshi et al. (1997), Amin and Ng (1993), Andreasen (2006), van der Ploeg (2007), Antonov et al. (2008), Grzelak and Oosterlee (2009) and Grzelak et al. (2009). The SZHW model distinguishes itself from these models by a closed form call pricing formula and/or explicit, rather than implicit, incorporation of the correlation between underlying and the term structure of interest rates.

Our contribution to the existing literature is fourfold:

- First, we derive the conditional characteristic function of the SZHW model in closed form and analyse pricing vanilla equity calls and puts using transform inversion.
- Second, since the practical relevance of any model is limited without a numerical implementation, we extensively consider the efficient implementation of the Fourier transform inversion (see Lord and Kahl (2008)) required to price European options. In particular we derive a theoretical result on the limiting behaviour of the conditional characteristic function of the SZHW model which allows us to calculate the inversion integral much more accurately.
- Third, we consider the pricing of forward starting options.
- Fourth, we generalize the SZHW model to be able to value FX options in a framework where both domestic and foreign interest rate processes are stochastic.

The outline for the remainder of the chapter is as follows. First, we introduce the model and focus on the analytical properties. Second, we consider the effect of stochastic interest rates and

correlation on the implied volatility term structure. Third, we consider the numerical implementation of the transform inversion integral. Fourth, we consider the pricing of forward starting options. Fifth, we present the extension of the model for FX options involving two interest rate processes. Finally we conclude.

3.2 The Schöbel-Zhu-Hull-White model

The model we will derive here is a combination of the famous Hull and White (1993) model for the stochastic interest rates and the Schöbel and Zhu (1999) model for stochastic volatility. The model has three key variables, which we allow to be correlated with each other: the stock price $x(t)$, the Hull-White interest rate process $r(t)$ and the stochastic stock volatility which follows an Ornstein-Uhlenbeck process cf. Schöbel and Zhu (1999). The risk-neutral asset price dynamics of the Schöbel-Zhu-Hull-White (SZHW) read:

$$dS(t) = S(t)r(t)dt + S(t)v(t)dW_S(t), \qquad S(0) = S_0, \tag{3.1}$$

$$dr(t) = (\theta(t) - ar(t))dt + \sigma dW_r(t), \qquad r(0) = r_0, \tag{3.2}$$

$$dv(t) = \kappa(\psi - v(t))dt + \tau dW_v(t), \qquad v(0) = v_0, \tag{3.3}$$

where $a, \sigma, \kappa, \psi, \tau$ are positive parameters which can be inferred from market data and correspond to the mean reversion and volatility of the short rate process, and the mean reversion, long-term volatility and volatility of the volatility process respectively. The quantity r_0 and the deterministic function $\theta(t)$ are used to match the currently observed term structure of interest rates, e.g. see Hull and White (1993). The hidden parameter $v_0 > 0$, corresponds to the current instantaneous volatility and hence should be determined directly from market (e.g. just as the non-observable short interest rate), but is in practice often (mis-)used as extra parameter for calibration. Finally, $\widetilde{W}(t) = (W_S(t), W_r(t), W_v(t))$ denotes a Brownian motion under the risk-neutral measure Q with covariance matrix:

$$\mathrm{Cov}(\widetilde{W}(t)) = \begin{pmatrix} 1 & \rho_{Sr} & \rho_{Sv} \\ \rho_{Sr} & 1 & \rho_{rv} \\ \rho_{Sv} & \rho_{rv} & 1 \end{pmatrix} t \tag{3.4}$$

Note that as $v(t)$ follows an Ornstein-Uhlenbeck process, there is a possibility that $v(t)$ becomes negative; effectively this implies that the sign of instantaneous correlation between $\ln S(t)$ and $v(t)$ changes as $v(t)$ goes through zero. However, the actual volatility is $|v(t)|$, see Section 2.3.1, and it does not have this feature.

3.3 European option pricing

General payoffs which are a function of the stock price at maturity T can be priced using the corresponding characteristic function of the log-asset price. Therefore we evaluate the probability distribution of the T-forward stock price at time T. Instead of evaluating expected discounted payoff under the risk-neutral bank account measure, we can also change the underlying proba-

bility measure to evaluate this expectation under the T-forward probability measure Q^T, which is equivalent to choosing the T-discount bond as numeraire, see Chapter 2. Hence starting from time t, we can evaluate the price of a European stock option ($w = 1$ for a call option, $w = -1$ for a put option) with strike $K = \exp(k)$ as

$$\mathbb{E}^Q\Big[\exp[-\int_t^T r(u)du](w(S(T)-K))^+\big|\mathcal{F}_t\Big] = P(t,T)\mathbb{E}^{Q^T}\Big[(w(F^T(T)-K))^+\big|\mathcal{F}_t\Big], \tag{3.5}$$

where $P(t,T)$ denotes the price of a (pure) discount bond and $F^T(t) := \frac{S(t)}{P(t,T)}$ denotes the T-forward stock price. The above expression can be numerically evaluated by means of a Fourier inversion of the log-asset price characteristic function, see equation (2.46) of Chapter 2. This characteristic function is derived in the following subsection. Section 3.5 is concerned with the numerical implementation of the Fourier transform and presents an alternative pricing equation which transforms the integration domain to the unit interval, see e.g. Lord and Kahl (2008), and hence avoids truncation errors.

The T-forward dynamics

Recall from Chapter 2, that in the Hull-White model, one write the following for the price of a discount bond:

$$P(t,T) = \exp[A_r(t,T) - B_r(t,T)r(t)], \tag{3.6}$$

where $A_r(t,T)$ contains information on the currently observed term structure of interest rates and $B_r(t,T) := (1-e^{-a(T-t)})/a$. Note that this expression is equivalent to equation (2.14) for the zero-coupon bond price, and differs only in terms of notation. The forward stock price can therefore be expressed as

$$F^T(t) = \frac{S(t)}{\exp[A_r(t,T) - B_r(t,T)r(t)]}. \tag{3.7}$$

Under the risk-neutral measure Q (where we use the money market bank account as numeraire) the discount bond price follows the process $dP(t,T) = r(t)P(t,T)dt - \sigma B_r(t,T)P(t,T)dW_r(t)$. Hence, by an application of Itô's lemma, we find the following T-forward stock price process:

$$\begin{aligned} dF^T(t) &= \Big(\sigma^2 B_r^2(t,T) + \rho_{Sr}\nu(t)\sigma B_r(t,T)\Big)F^T(t)dt \\ &\quad + \nu(t)F^T(t)dW_S(t) + \sigma B_r(t,T)F^T(t)dW_r(t) \end{aligned} \tag{3.8}$$

By definition the forward stock price will be a martingale under the T-forward measure. This is achieved by defining the following transformations of the Brownian motions:

$$\begin{aligned} dW_r(t) &\mapsto dW_r^T(t) - \sigma B_r(t,T)dt, \\ dW_S(t) &\mapsto dW_S^T(t) - \rho_{Sr}\sigma B_r(t,T)dt, \\ dW_\nu(t) &\mapsto dW_\nu^T(t) - \rho_{r\nu}\sigma B_r(t,T)dt. \end{aligned} \tag{3.9}$$

Hence under the T-forward measure the processes for $F^T(t)$ and $v(t)$ are given by

$$dF^T(t) = v(t)F^T(t)dW_S^T(t) + \sigma B_r(t,T)F^T(t)dW_r^T(t), \tag{3.10}$$

$$dv(t) = \kappa\Big((\psi - \frac{\rho_{rv}\sigma\tau}{\kappa}B_r(t,T)) - v(t)\Big)dt + \tau dW_v^T(t), \tag{3.11}$$

where $W_S^T(t), W_r^T(t), W_v^T(t)$ are now Brownian motions under the T-forward Q^T. We can simplify (3.10) by switching to logarithmic coordinates: defining $y(t) := \log(F^T(t))$ and an application of Itô's lemma yields

$$dy(t) = -\frac{1}{2}v_F^2(t)dt + v(t)W_S^T(t) + \sigma B_r(t,T)W_r^T(t), \tag{3.12}$$

$$dv(t) = \kappa(\xi(t) - v(t))dt + \tau dW_v^T(t) \tag{3.13}$$

with

$$v_F^2(t) := v^2(t) + 2\rho_{Sr}v(t)\sigma B_r(t,T) + \sigma^2 B_r^2(t,T) \tag{3.14}$$

$$\xi(t) := \Big(\psi - \frac{\rho_{rv}\sigma\tau}{\kappa}B_r(t,T)\Big). \tag{3.15}$$

Notice that we now have reduced the system (3.1) of the three variables $x(t)$, $r(t)$ and $v(t)$ under the risk-neutral measure, to the system (3.12) of two variables $y(t)$ and $v(t)$ under the T-forward measure. What remains is to find the characteristic function of the reduced system of variables.

Determining the characteristic function of the forward log-asset price

We will now determine the characteristic function of the reduced system (3.12), which we will do by means of a PDE approach. That is, we apply the Feynman-Kac theorem and reduce the problem of finding the characteristic of the forward log-asset price dynamics to solving a partial differential equation; that is, the Feynman-Kac theorem implies that the characteristic function

$$f(t,y,v) = \mathbb{E}^{Q^T}\Big[\exp(iuy(T))\Big|\mathcal{F}_t\Big], \tag{3.16}$$

is given by the solution of the following partial differential equation

$$0 = f_t - \frac{1}{2}v_F^2(t)f_y + \kappa(\xi(t) - v(t))f_v + \frac{1}{2}v_F^2(t)f_{yy} \tag{3.17}$$
$$+(\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_r(t,T))f_{yv} + \frac{1}{2}\tau^2 f_{vv},$$

$$f(T,y,v) = \exp\Big(iuy(T)\Big), \tag{3.18}$$

where the subscripts denote partial derivatives and we took into account that the covariance term $dy(t)dv(t)$ is equal to

$$dy(t)dv(t) = (v(t)dW_S^T(t) + \sigma B_r(t,T)dW_r^T(t))(\tau dW_v^T(t)) = (\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_r(t,T))dt, \tag{3.19}$$

and to ease the notation we dropped the explicit (t, y, v)-dependence for f.
Solving the defining partial differential equation (3.17) subject to the boundary condition (3.18), leads to the following proposition.

Proposition 3.3.1 *The characteristic function of T-forward log-asset price of the SZHW model is given by the following closed-form solution:*

$$f(t,y,v) = \exp\Big[A(u,t,T) + B(u,t,T)y(t) + C(u,t,T)v(t) + \frac{1}{2}D(u,t,T)v^2(t)\Big], \tag{3.20}$$

where:

$$\begin{aligned} A(u,t,T) &= -\frac{1}{2}u(i+u)V(t,T) \\ &\quad + \int_t^T \Big[\Big(\kappa\psi + \rho_{rv}(iu-1)\tau\sigma B_r(s,T)\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds \end{aligned} \tag{3.21}$$

$$B(u,t,T) = iu, \tag{3.22}$$

$$C(u,t,T) = -u(i+u)\frac{\Big((\gamma_3 - \gamma_4 e^{-2\gamma(T-t)}) - (\gamma_5 e^{-a(T-t)} - \gamma_6 e^{-(2\gamma+a)(T-t)}) - \gamma_7 e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \tag{3.23}$$

$$D(u,t,T) = -u(i+u)\frac{1 - e^{-2\gamma(T-t)}}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \tag{3.24}$$

with:

$$\begin{aligned} \gamma &= \sqrt{(\kappa - \rho_{Sv}\tau iu)^2 + \tau^2 u(i+u)}, & \gamma_1 &= \gamma + (\kappa - \rho_{Sv}\tau iu), \\ \gamma_2 &= \gamma - (\kappa - \rho_{Sv}\tau iu), & \gamma_3 &= \frac{\rho_{Sr}\sigma\gamma_1 + \kappa a\psi + \rho_{rv}\sigma\tau(iu-1)}{a\gamma}, \\ \gamma_4 &= \frac{\rho_{Sr}\sigma\gamma_2 - \kappa a\psi - \rho_{rv}\sigma\tau(iu-1)}{a\gamma}, & \gamma_5 &= \frac{\rho_{Sr}\sigma\gamma_1 + \rho_{rv}\sigma\tau(iu-1)}{a(\gamma - a)}, \\ \gamma_6 &= \frac{\rho_{Sr}\sigma\gamma_2 - \rho_{rv}\sigma\tau(iu-1)}{a(\gamma + a)}, & \gamma_7 &= (\gamma_3 - \gamma_4) - (\gamma_5 - \gamma_6), \end{aligned} \tag{3.25}$$

and:

$$V(t,T) = \frac{\sigma^2}{a^2}\Big((T-t) + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\Big). \tag{3.26}$$

Proof The model we are considering is not an affine model in $y(t)$ and $v(t)$, but it is if we enlarge

the state space to include $v^2(t)$:

$$dy(t) = -\frac{1}{2}v_F^2(t)dt + v_F(t)dW_F^T(t) \tag{3.27}$$

$$dv(t) = \kappa(\xi(t) - v(t))dt + \tau dW_v^T(t) \tag{3.28}$$

$$dv^2(t) = 2v(t)dv(t) + \tau^2 dt = 2\kappa\Big(\frac{\tau^2}{2\kappa} + \xi(t)v(t) - v^2(t)\Big)dt + 2\tau v(t)dW_v^T(t) \tag{3.29}$$

We can find the characteristic function of the T-forward log price by solving the partial differential equation (3.17) for joint characteristic function $f(t, y, v)$ with corresponding boundary condition (3.18); substituting the partial derivatives of the functional form (3.20) into (3.17) provides us with four ordinary differential equations containing the functions $A(t)$, $B(t)$, $C(t)$ and $D(t)$. Solving this system yields the above solution, see appendix 3.9. □

We note that the strip of regularity of the SZHW characteristic function, i.e. for which values the Fourier transform of the characteristic function exists, is the same as that of the Schöbel and Zhu (1999) model, for which we refer the reader to Lord and Kahl (2008).

3.4 Impact of stochastic interest rates and correlation

To gain some insights into the impact of the correlated stochastic rates and corresponding parameter sensitivities we will look at the at-the-money implied volatility structure which we compute for different parameter settings. Besides comparing different parameter settings of the SZHW model, we also make a comparison with the classical Schöbel and Zhu (1999) model to determine the impact of stochastic rates in general. The behaviour of the 'non-interest rate' parameters are similar to other stochastic volatility models like Heston (1993) and Schöbel and Zhu (1999), that is the volatility of the volatility lift the wings of the volatility smile, the correlation between the stock process and the volatility process can incorporate a skew, and the short and long-term vol determine the level of the implied volatility structure. The impact of stochastic rates and the corresponding correlation are plotted in Figure **1**.

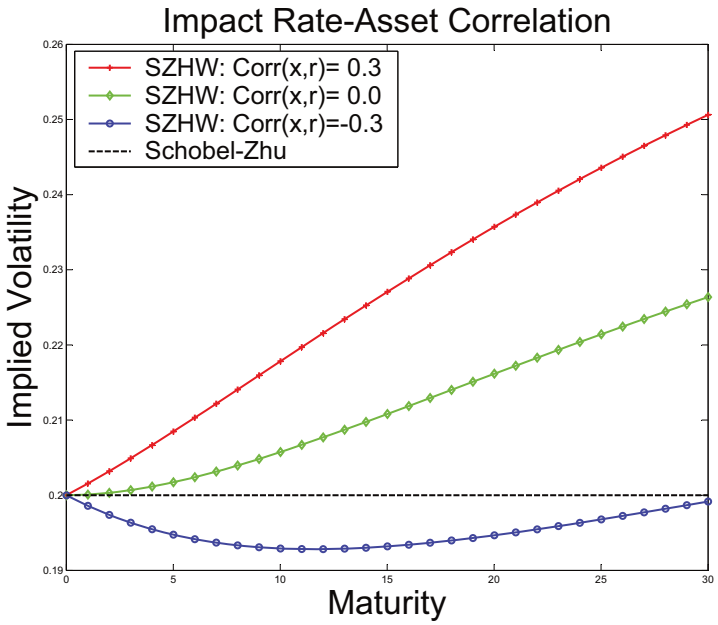

Figure 1: Impact of ρ_{Sr} on at-the-money implied volatilities. The graph corresponds to the (degenerate) Black-Scholes-Hull-White case with parameter values $r(t) = 0.05$, $a = 0.05$, $\sigma = 0.01$, $v(0) = \psi = 0.20$, $\rho_{Sv} = 0.0$ and constant volatility process.

From Figure **1**, one can see that the stochastic interest rates add extra flexibility to the modelling framework; by changing the rate-asset correlation one can create an upward (or an initially downward) sloping term structure of volatility, even in case the volatility process is constant. If we compare the case with zero correlation between the equity and interest rate drivers with the ordinary process with deterministic rates, we see that the stochastic rates make the term structure upward sloping. Note that this is in correspondence with empirical data, which shows higher at-the-money volatilities the longer the maturities go. The effect becomes more apparent for maturities larger than five years; while for one years the effect of uncorrelated stochastic rates is below a basis point, the effect on a five year option is already more than ten basis points which increases to a couple of hundred basis points for a thirty year option. These model effects also correspond with a general feature of the interest rate market: the market's view on the uncertainty of long-maturity bonds is often much higher than that of shorter bond, hence reflecting the increasing impact of stochastic interest rates for long-maturity equity options. Finally, we note that for higher positive values of linear correlation coefficient between equity and the interest rate component, the impact of stochastic rates becomes more apparent.

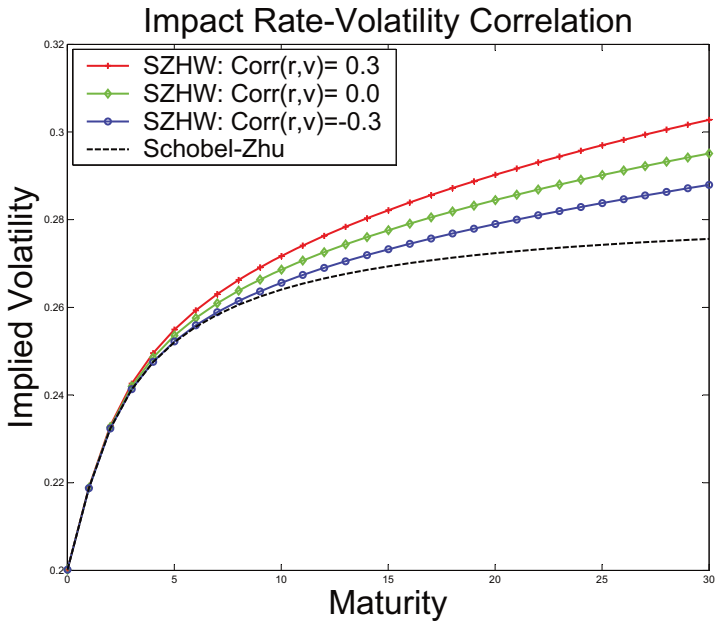

Figure 2: Impact of ρ_{Sv} on at-the-money implied volatilities. The graph corresponds to the parameter values $r(t) = 0.05$, $a = 0.05$, $\sigma = 0.01$, $v(0) = \psi = 0.20$, $\rho_{Sr} = 0.0$ and with volatility process with a mean reversion coefficient of $\kappa = 0.5$ and volatility of volatility $\tau = 0.2$.

From Figure **2**, one can see that the effect of the correlation coefficient between the drivers of the rate and volatility process is similar, however the impact on the implied volatility structure is less severe and different in sign: a positive correlation coefficient causes a dampening effect, whereas a negative correlation increases the overall volatility, which effect can also be seen from the volatility dynamics (3.11). Note hereby that the increasing term structure for $\rho_{rv} =$, in the Figure **2** is mainly caused by the Schöbel-Zhu stochastic volatility process in comparison to the deterministic volatility process used in Figure **1**. In comparison to the Schöbel and Zhu (1999) model, we can see that the stochastic interest rates increase the slope of the term structure. More importantly, the implied volatilities do not die out, but remain upward sloping, which behaviour often corresponds with implied volatility quotes in long-maturity equity (e.g. see MarkIT) or FX (e.g. see Andreasen (2006)) options. However for strong positive correlation values this might be the other way around. In contrast to the first picture, we see somewhat smaller effects: for example the increasing effect of stochastic rates is even larger than that of the dampening effect of a positive correlation of 30% between the rate and volatility drivers. Again we see that the effects of stochastic rates become more apparent for longer maturities.

In general, we can see from Figure **1** and **2** that stochastic rates have a significant impact on the backbone of the implied volatility structure and add extra flexibility to the modelling framework. The effects become more apparent for larger maturities and for larger absolute values of the correlation coefficients. Hereby the effect of correlation coefficient between equity and interest rates seems to be the most determinant factor. One can then use these degrees of freedom in

several ways: either one jointly calibrates these parameters to implied volatility surfaces (or some other options), or one can first calibrate these and then use the other parameters to calibrate the remainder of the model. In our opinion this choice has to depend on the exotic product: if the correlations are of larger impact on a exotic product (e.g. on a hybrid equity-interest rate product) than on short-dated vanilla calls, it might then be preferable to use a historical estimate for the correlation coefficient at the cost of a slightly worse calibration result. One way or the other, the SZHW stands out by the additional freedom it offers by explicitly modelling the correlation coefficient between the underlying, the stochastic volatility and the stochastic interest rates.

3.5 Calculating the inverse Fourier transform

In Lord and Kahl (2008) the practical calculation of the inverse Fourier transform (2.46), in general and for specific models, is discussed in great detail. They recommend that

- Any truncation error is avoided by appropriately transforming the range of integration to a finite interval.
- An adaptive integration algorithm is used, hereby allowing the discretization error to be of a prescribed maximum size.
- The damping parameter α is chosen such that the integrand is minimized in $v = 0$, which typically leads to much more accurate prices for options which have long maturities and/or are away from the at-the-money level.

To apply these results to the SZHW model, recall from (2.46) of Chapter 2, that for general α, the following pricing equation holds

$$C_T(k) = P(t,T)\frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(e^{-(\alpha+iv)k}\psi_T(v)\Big)dv + R\Big(F^T(t), K, \alpha\Big), \tag{3.30}$$

with $P(t,T)$ the current time discount factor for maturing at time T, $F^T(t)$ the T-forward asset price and $k := \ln K$ the logarithm of the strike. By changing variables from v to $g(v)$, which maps $[0,\infty) \mapsto [0,1]$, the pricing equation (3.30) becomes

$$C_T(k) = P(t,T)\frac{1}{\pi}\int_0^1 \mathrm{Re}\Big(e^{-(\alpha+ig(v))k}\psi_T(g(v))\cdot g'(v)\Big)dv + R\Big(F^T(t), K, \alpha\Big). \tag{3.31}$$

To avoid numerical complications, one carefully has to choose the transformation function g such that the integrand remains finite over the range of integration, as it is in (3.30). To find such a transformation, we analyse the limiting behaviour of the characteristic function. In particular, suppose that the characteristic function $\phi_T(u)$ of the SZHW model for large values of u behaves as

$$\exp\Big(\phi_r(u) + i\phi_i(u)\Big), \tag{3.32}$$

with both $\phi_r(u)$ and $\phi_i(u)$ functions on the real line. The integrand in (3.30) will then have the following asymptotics

$$\mathrm{Re}\Big(e^{-i(u-i\alpha)k}\frac{\phi_T(u-(\alpha+1)i)}{(\alpha+iu)(\alpha+i+iu)}\Big) \propto \frac{e^{-\alpha k+\psi_r\left(u-(\alpha+1)i\right)}}{u^2}\cdot\cos\Big(ku-\psi_i(u-(\alpha+1)i)\Big). \tag{3.33}$$

In the remainder we will determine ψ_r, which will tell us which transformation function is suitable to use. Lord and Kahl (2008) already supply a number of intermediary results for the Schöbel and Zhu (1999) model, but as the notation we use here is slightly different, we will briefly restate these results. For large values of u, only γ,γ_1 and γ_2 in (3.25) are $O(u)$, whereas γ_3 to γ_6 tend to a constant, and γ_7 is actually $O(\frac{1}{u})$. The limits we require here are

$$\begin{aligned}
\lim_{u\to\infty}\frac{\gamma(u)}{u} &= \tau\sqrt{1-\rho_{Sv}} =: \gamma(\infty), && (3.34)\\
\lim_{u\to\infty}\frac{\gamma_1(u)}{u} &= \gamma(\infty) - i\rho_{Sv}\tau =: \gamma_1(\infty), && (3.35)\\
\lim_{u\to\infty}\gamma_3(u) &= \sigma\frac{\rho_{Sr}\gamma(\infty)+i\tau(\rho_{rv}-\rho_{Sr}\rho_{Sv})}{a\gamma(\infty)} =: \gamma_3(\infty), && (3.36)\\
\lim_{u\to\infty}\gamma_5(u) &= \sigma\frac{\rho_{Sr}\gamma(\infty)+i\tau(\rho_{rv}-\rho_{Sr}\rho_{Sv})}{a\gamma(\infty)} =: \gamma_5(\infty). && (3.37)
\end{aligned}$$

We find that the limiting behaviour for $C(u,t,T)$ in (3.23) follows from

$$\begin{aligned}
\lim_{u\to\infty}\frac{C(u,t,T)}{u} &= -\frac{\gamma_3(\infty)-\gamma_5(\infty)e^{-a(T-t)}}{\gamma_1(\infty)} = -\frac{i\rho_{rv}+\rho_{Sr}(\sqrt{1-\rho_{xv}^2}-i\rho_{Sv})}{\tau(1-\rho_{xv}^2-i\rho_{Sv}\sqrt{1-\rho_{xv}^2})}\frac{\sigma}{\tau}B_r(t,T)\\
&\equiv C(\infty)\frac{\sigma}{\tau}B_r(t,T). && (3.38)
\end{aligned}$$

From the above result,the limiting behaviour of $D(u,t,T)$ in (3.24) for large values of u follows as

$$\lim_{u\to\infty}\frac{D(u,t,T)}{u} = -\frac{1}{\gamma_1(\infty)}. \tag{3.39}$$

Finally, we need to analyse $A(t) = A(u,t,T)$ in (3.21). Its defining ODE (3.94) can be found in appendix 3.9, i.e.

$$\begin{aligned}
\frac{\partial A(u,t,T)}{\partial t} &= -\Big[\kappa\xi(t)+iu\rho_{rv}\tau\sigma B_r(t,T)\Big]C(u,t,T)+\frac{1}{2}u(i+u)\sigma^2B_r^2(t,T)\\
&\quad -\frac{1}{2}\tau^2(C^2(u,t,T)+D(u,t,T)). && (3.40)
\end{aligned}$$

The first derivative of $A(u,t,T)$ behaves as $O(u^2)$ for large values of u, as can be seen from

$$\lim_{u\to\infty}\frac{1}{u^2}\frac{\partial A(u,t,T)}{\partial t} = \frac{1}{2}\Big(1-C^2(\infty)-2i\rho_{rv}C(\infty)\Big)\sigma^2B_r^2(t,T) \tag{3.41}$$

Finally, together with the boundary condition $A(u,T,T) = 0$, we have

$$\lim_{u\to\infty}\frac{A(u,t,T)}{u^2} = -\int_t^T \lim_{u\to\infty}\frac{1}{u^2}\frac{\partial A(u,s,T)}{\partial s}ds = -\frac{1}{2}V(t,T)\cdot\Big(1-C^2(\infty)-2i\rho_{rv}C(\infty)\Big) \equiv -A(\infty), \tag{3.42}$$

where $V(t,T)$ denotes the integrated bond variance, i.e. as defined in (3.26). One can show that $\mathrm{Re}\Big(A(\infty)\Big) \geq 0$ as $V(t,T) \geq 0$ and:

$$\mathrm{Re}\Big(C^2(\infty)+2i\rho_{rv}C(\infty)\Big) = \frac{\rho_{xr}^2 - 2\rho_{rv}\rho_{Sr}\rho_{Sv} + \rho_{rv}^2(4\rho_{xv}^2-3)}{1-\rho_{xv}^2} \leq 1. \tag{3.43}$$

This follows by maximizing the right-hand side with respect to the constraint that the three correlations constitute a positive semi-definite correlation matrix. For example, the maximum is achieved when $\rho_{Sr} = -\frac{1}{2}\sqrt{3}$, $\rho_{Sv} = -\frac{1}{2}$ and $\rho_{rv} = 0$.
The above analysis determines ϕ_r as

$$\phi_r\Big(u-(\alpha+1)i\Big) = -\mathrm{Re}\Big(A(\infty)\Big)\cdot u^2. \tag{3.44}$$

One can conclude that the tail behaviour of the characteristic function of the SZHW model is quite different from that of the Schöbel and Zhu (1999) model; whereas the decay in the Schöbel-Zhu model is only exponential, the decay here resembles that of a Gaussian characteristic function, caused by the addition of a Gaussian short rate process. Clearly, if σ (the volatility of the short rate) is zero, $A(\infty) = 0$ and the decay of the characteristic function becomes exponential once again. As the tail behaviour of the characteristic function is of the same from as that of the Black and Scholes (1973) characteristic function, an appropriate transformation function is $g : [0,\infty) \mapsto [0,1]$, as in Lord and Kahl (2008),

$$g(u) = -\frac{\ln u}{\sqrt{A(\infty)}}, \tag{3.45}$$

which can be used in the pricing equation (3.31).

3.6 Forward starting options

Due to the popularity of forward starting options such as cliquets, the pricing of forward starting options recently attracted the attention of both practitioners and academics (e.g. see Lucic (2003), Hong (2004), Kruse and Nögel (2005) and Brigo and Mercurio (2006)). In this section we will show how one can price forward starting options within the SZHW framework; following Hong (2004), we consider the (forward) log return of the asset price S:

$$z(T_{i-1},T_i) := \log\Big(\frac{S(T_i)}{S(T_{i-1})}\Big). \tag{3.46}$$

Since

$$\log S(t) = y(t) + \log P(t, T_i), \tag{3.47}$$

we can express (3.46) also in terms of the T_i-forward log-asset price $y(t) = \log F^{T_i}(t)$, i.e.

$$z(T_{i-1}, T_i) = y(T_i) - y(T_{i-1}) - \log P(T_{i-1}, T_i), \tag{3.48}$$

with $t \leq T_{i-1} \leq T_i$. We are interested in the following forward starting call option with strike $K = \exp(k)$ on the return $\frac{S(T_i)}{S(T_{i-1})}$,

$$\begin{aligned} C_{T_{i-1},T_i}(k) &= \mathbb{E}^Q\Big[\exp\Big[-\int_t^{T_i} r(u)du\Big]\Big(\frac{S(T_i)}{S(T_{i-1})} - K\Big)^+\Big|\mathcal{F}_t\Big] \\ &= P(t,T_i)\mathbb{E}^{Q^T}\Big[(F^{T_i}_{T_{i-1},T_i}(T_i) - K)^+\Big|\mathcal{F}_t\Big], \end{aligned} \tag{3.49}$$

where

$$F^{T_i}_{T_{i-1},T_i}(T_i) := \exp[z(T_{i-1}, T_i)]$$

denotes the forward return between T_{i-1} and T_i. Note that the above expression is nothing more than some call option under the T-forward measure. Therefore, as noted by Hong (2004), the pricing of forward starting options can be reduced to finding the characteristic function of the log forward return under the T-forward measure; by replacing the log-asset price by the forward log-return one can directly apply the pricing equation (3.30) or (3.31), i.e. by replacing the corresponding characteristic function by $\psi_{T_{i-1},T_i}(v)$: the characteristic function (under the T_i-forward measure) of the forward log-return between T_{i-1} and T_i. What remains to be done for the pricing of forward starting options is the derivation of this forward characteristic function, which we will deal with in the following subsection.

3.6.1 Forward characteristic function

We will now derive the forward characteristic function of the forward log return $z^{T_i}_{T_{i-1},T_i} = y(T_i) - y(T_{i-1}) - \log P(T_{i-1})$ in the SZHW model. In the derivation we will use the now following lemma.

Lemma 3.6.1 *Let Z be a standard normal distributed random variable, furthermore let p and q be two positive constants. Then the characteristic function, of $Y := pZ + \frac{q}{2}Z^2$ is given by*

$$\phi_Y(u, p, q) := I\!E \exp(iuY) = \frac{\exp\Big(\frac{-p^2u^2}{2-2iuq}\Big)}{\sqrt{1 - iuq}}, \tag{3.50}$$

Proof Either by completing the square and using properties of the non-central chi-squared distribution or by direct integration of an exponential affine form against the normal distribution, e.g. see Johnson et al. (1994) or Glasserman (2003). □

Before we can apply the above lemma we first need to rewrite the characteristic function of the log-return $y(T_i) - y(T_{i-1})$ in the form of the above lemma. To simplify the notation we write

$B := iu$, $A(T_{i-1}) := A(u, T_{i-1}, T_i)$, $C(T_{i-1}) := C(u, T_{i-1}, T_i)$ and $D(T_{i-1}) := D(u, T_{i-1}, T_i)$. By using the tower law for conditional expectations and the (conditional) characteristic function of the SZHW model one can then obtain

$$\begin{aligned}
\phi_{T_{i-1},T_i}(u) &= \mathbb{E}^{Q^T}\Big\{\exp\Big(iu\Big[y(T_i) - y(T_{i-1}) - \log P(T_{i-1}, T_i)\Big]\Big)\Big|\mathcal{F}_t\Big\} \qquad (3.51)\\
&= \mathbb{E}^{Q^T}\Big\{\mathbb{E}^{Q^T}\Big[\exp\Big(iu[y(T_i) - y(T_{i-1}) - \log P(T_{i-1}, T_i)]\Big)\Big|\mathcal{F}_{T_{i-1}}\Big]\Big|\mathcal{F}_t\Big\}\\
&= \exp[A(T_{i-1}) - iuA_r(T_{i-1}, T_i)] \cdot\\
&\qquad \mathbb{E}^{Q^T}\Big\{\exp\Big[iuB_r(T_{i-1}, T_i)r(T_{i-1}) + C(T_{i-1})\nu(T_{i-1}) + \frac{1}{2}D(T_{i-1})\nu^2(T_{i-1})\Big]\Big|\mathcal{F}_t\Big\}.
\end{aligned}$$

Since the pair $\Big(r(T_{i-1}), \nu(T_{i-1})\Big)$, conditional on the information set $\mathcal{F}_t$ on time t, follows a joint Gaussian distribution with means μ_r, μ_ν (see (3.56), (3.58)) and variances σ_r^2, σ_ν^2 (see (3.57), (3.59)), we can write the sum of dependent normal variates $r(T_{i-1}), \nu(T_{i-1})$ in terms of two independent standard normal distributions Z_1 and Z_2 (e.g. by a Cholesky decomposition):

$$\begin{aligned}
iubr(T_{i-1}) + c\nu(T_{i-1}) + \frac{1}{2}d\nu^2(T_{i-1}) &\overset{d}{=} iub\Big(\mu_r + \sigma_r[\rho_{r\nu}(t, T_{i-1})Z_1 + \sqrt{1 - \rho_{r\nu}^2(t, T_{i-1})}\,Z_2]\Big)\\
&\quad + c\Big(\mu_\nu + \sigma_\nu Z_1\Big) + \frac{1}{2}d(\mu_\nu + \sigma_\nu Z_1)^2 \qquad (3.52)\\
&= iub\mu_r + c\mu_\nu + \frac{1}{2}d\mu_\nu^2 + iub\sigma_r\sqrt{1 - \rho_{r\nu}^2(t, T_{i-1})}\,Z_2\\
&\quad + \Big[c\sigma_\nu + d\mu_\nu\sigma_\nu + iub\rho_{r\nu}(t, T_{i-1})\sigma_r\Big]Z_1 + \frac{1}{2}d\sigma_\nu^2 Z_1^2,
\end{aligned}$$

where the correlation $\rho_{r\nu}(t, T_{i-1})$ between $r(T_{i-1})$ and $\nu(T_{i-1})$ over the interval $[t, T_{i-1}]$ is given by

$$\rho_{r\nu}(t, T_{i-1}) = \frac{\rho_{r\nu}\sigma\tau}{\sigma_r\sigma_\nu(a + \kappa)}\Big[1 - e^{-(a+\kappa)(T_{i-1}-t)}\Big]. \qquad (3.53)$$

Hence, using the independence of Z_1, Z_2 and equation (3.52), one can find the following expression for the forward characteristic function

$$\begin{aligned}
\phi_{T_{i-1},T_i}(u) &= \exp\Big[A(T_{i-1}) + iu(B_r(T_{i-1}, T_i)\mu_r - A_r(T_{i-1}, T_i)) + C(T_{i-1})\mu_\nu + \frac{1}{2}D(T_{i-1})\mu_\nu^2\Big]\\
&\quad \mathbb{E}^{Q^T}\Big\{\exp\Big[iuB_r(T_{i-1}, T_i)\sigma_r\sqrt{1 - \rho_{r\nu}^2(t, T_{i-1})}\,Z_2\Big]\Big|\mathcal{F}_t\Big\}\\
&\quad \cdot\mathbb{E}^{Q^T}\Big\{\exp\Big(\Big[C(T_{i-1})\sigma_\nu + D(T_{i-1})\mu_\nu\sigma_\nu + iuB_r(T_{i-1}, T_i)\rho_{r\nu}(t, T_{i-1})\sigma_r\Big]Z_1\\
&\qquad + \frac{1}{2}D(T_{i-1})\sigma_\nu^2 Z_1^2\Big)\Big|\mathcal{F}_t\Big\}. \qquad (3.54)
\end{aligned}$$

Hence we come to the following proposition

Proposition 3.6.2 *Starting from the current time t, the characteristic function of the forward log*

return $z(T_{i-1}, T_i)$ under the T_i-forward measure is given by the following closed-form solution:

$$\begin{aligned}\phi_{T_{i-1},T_i}(u) &= \exp\Big[A(T_{i-1}) + iu\Big[B_r(T_{i-1},T_i)\mu_r - A_r(T_{i-1},T_i)\Big] + C(T_{i-1})\mu_v + \frac{1}{2}D(T_{i-1})\mu_v^2\Big] \\ &\quad \cdot\phi_{Z_2}\Big(iuB_r(T_{i-1},T_i)\sigma_r\sqrt{1-\rho_{rv}^2(t,T_{i-1})}\Big)\phi_Y\Big(1, P(T_{i-1}), Q(T_{i-1})\Big) \qquad (3.55)\end{aligned}$$

with

$$\begin{aligned}P(T_{i-1}) &= C(T_{i-1})\sigma_v + D(T_{i-1})\mu_v\sigma_v + iu\rho_{rv}(t,T_{i-1})B_r(T_{i-1},T_i)\sigma_r, \\ Q(T_{i-1}) &= D(T_{i-1})\sigma_v^2, \\ \phi_{Z_2}(y) &= \exp\Big(\frac{y^2}{2}\Big),\end{aligned}$$

and where $\phi_Y\Big(-i, P(T_{i-1}), Q(T_{i-1})\Big)$ is given by Lemma 3.6.1.

Proof The result follows directly by evaluating the expectations from expression (3.54) for the moment-generating function of the standard Gaussian distribution Z_2 evaluated in the point $iuB_r(T_{i-1},T_i)\sigma_r\sqrt{1-\rho_{rv}^2(t,T_{i-1})}$, while the second expectation is the moment generating function of the random variable $Y = P(T_{i-1})Z_1 + \frac{Q(T_{i-1})}{2}Z_1^2$ evaluated in the unit point in the point $-i$, for which an analytical expression is given by Lemma 3.6.1. □

What yet remains, is to determine (conditional on the time-t) the T_i-forward mean and variance of the interest rate and volatility processes $r(T_{i-1})$ and $v(T_{i-1})$.

3.6.2 Moments of the Hull-White short interest rate

From (2.18) and Itô's isometry, we immediately have that $r(T_{i-1})$, under the T_i-forward measure (starting from time t), is normally distributed with mean μ_r and variance σ_r^2 given by

$$\mu_r = x(t)e^{-a(T_{i-1}-t)} - M^{T_i}(t,T_{i-1}) + \beta(T_{i-1}), \qquad (3.56)$$

$$\sigma_r^2 = \frac{\sigma^2}{2a}\Big(1-e^{-2a(T_{i-1}-t)}\Big), \qquad (3.57)$$

which can hence be used in Proposition 3.6.2.

3.6.3 Moments of the Schöbel-Zhu volatility process

To determine the first two moments of the Schöbel-Zhu volatility process, under the T_i-forward measure, for a certain time $T_{i-1} \leq T_i$ and conditional on the filtration at time t, one can integrate the dynamics of (3.11) to obtain

$$v(T_{i-1}) = v(t)e^{-\kappa(T_{i-1}-t)} + \int_t^{T_{i-1}} \kappa\xi(u)e^{-\kappa(T_{i-1}-u)}du + \int_t^{T_{i-1}} \tau e^{-\kappa(T_{i-1}-u)}dW_v^T(u),$$

where $\xi(u) := \psi - \frac{\rho_{rv}\sigma\tau}{a\kappa}(1 - e^{a(T_i-u)})$. Hence under the T_i-forward measure, from Itô's isometry, we have the following for the mean and standard deviation of v:

$$\begin{aligned}\mu_v &= v(t)e^{-\kappa(T_{i-1}-t)} + \Big(\psi - \frac{\rho_{rv}\sigma\tau}{a\kappa}\Big)\Big(1 - e^{-\kappa(T_{i-1}-t)}\Big) \\ &\quad - \frac{\rho_{rv}\sigma\tau}{a(\kappa+a)}\Big(e^{-a(T_i-t)-\kappa(T_{i-1}-t)} - e^{-a(T_i-T_{i-1})}\Big), \qquad (3.58)\\ \sigma_v^2 &= \frac{\tau^2}{2\kappa}\Big(1 - e^{-2\kappa(T_{i-1}-t)}\Big), \qquad (3.59)\end{aligned}$$

which can hence be used in Proposition 3.6.2.

3.7 Schöbel-Zhu-Hull-White Foreign Exchange model

In this section we present the Schöbel-Zhu-Hull-White Foreign Exchange (SZHW-FX) model. That is, we introduce a domestic and a foreign exchange currency, which are modeled by Hull-White processes. We model the exchange rate process by geometric motion where we let the volatility follow an Ornstein-Uhlenbeck process. Moreover we allow all factors to be correlated with each other.
Notation is as follows: we let $X(t)$ denote the Foreign Exchange (FX) rate, with volatility v, between the domestic currency r_1 and the foreign currency r_2. The risk-neutral FX dynamics of the Schöbel-Zhu-Hull-White (SZHW) then read:

$$\begin{aligned}dX(t) &= X(t)(r_1(t) - r_2(t))dt + X(t)v(t)dW_X(t), & X(0) &= x_0, & (3.60)\\ dr_1(t) &= (\theta_1(t) - a_1 r_1(t))dt + \sigma_1 dW_{r_1}(t), & r_1(0) &= r_0^1, & (3.61)\\ dr_2(t) &= (\theta_2(t) - a_2 r_2(t) - \rho_{Xr_2} v(t)\sigma_2)dt + \sigma_2 dW_{r_2}(t), & r_2(0) &= r_0^2, & (3.62)\\ dv(t) &= \kappa(\psi - v(t))dt + \tau dW_v(t), & v(0) &= v_0, & (3.63)\end{aligned}$$

where $a_i, \sigma_i, \kappa, \psi, \tau$ are positive parameters. Hence the domestic and the (shifted) foreign interest rate markets are modeled by Hull-White models and the exchange rate is modeled by a Schöbel-Zhu stochastic volatility model. $\widetilde{W}(t) = (W_X(t), W_{r_1}, W_{r_2}(t), W_v(t))$ denotes a Brownian motion under the risk-neutral measure Q with a positive covariance matrix:

$$\mathrm{Cov}(\widetilde{W}(t)) = \begin{pmatrix} 1 & \rho_{Xr_1} & \rho_{Xr_2} & \rho_{Xv} \\ \rho_{Xr_1} & 1 & \rho_{r_1r_2} & \rho_{r_1v} \\ \rho_{Xr_2} & \rho_{r_1r_2} & 1 & \rho_{r_2v} \\ \rho_{Xv} & \rho_{r_1v} & \rho_{r_2v} & 1 \end{pmatrix} t \qquad (3.64)$$

We will now show that the above model dynamics yield a closed-form expression for the price of an European FX-option with strike K and maturity T. Hence we consider:

$$\mathbb{E}^Q\Big[\frac{(w(X(T) - K))^+}{N_1(T)}\Big|\mathcal{F}_t\Big], \qquad (3.65)$$

where $w = \pm 1$ for a call/put option and with

$$N_1(T) = \exp\Big[\int_t^T r(u)du\Big] \tag{3.66}$$

denotes the bank-account in the domestic economy. We can also represent the expectation (3.65) in the domestic T-forward measure Q^T associated with a domestic zero-coupon bond option $P_1(t,T)$ which matures at time T, hence we obtain

$$\mathbb{E}^Q\Big[\frac{(w(X(T)-K))^+}{N_1(T)}\Big|\mathcal{F}_t\Big] = P_1(t,T)\mathbb{E}^{Q^T}\Big[(w(FFX^T(T)-K))^+\Big|\mathcal{F}_t\Big], \tag{3.67}$$

where

$$FFX_T(t) = \frac{X(t)P_2(t,T)}{P_1(t,T)} \tag{3.68}$$

denotes the forward FX-rate under the domestic T-forward measure and in the Hull and White (1993) model:

$$P_i(t,T) = \exp[A_i(t,T) - B_i(t,T)x_i(t)] \qquad \text{with: } B_i(t,T) := \frac{1-e^{-a_i(T-t)}}{a_i}, \tag{3.69}$$

where $A_i(t,T)$ is a deterministic function. We can express the forward FX-rate as

$$FFX_T(t) = \frac{X(t)\exp[A_2(t,T) - B_2(t,T)x_2(t)]}{\exp[A_1(t,T) - B_1(t,T)x_1(t)]}. \tag{3.70}$$

Note that under their own risk-neutral measures (where the money market bank account of their own currency is used as numeraire) the discount bond prices follow the SDEs:

$$\frac{dP_i(t,T)}{P_i(t,T)} = r_i(t)dt - \sigma_i B_i(t,T)dW_{r_i}(t), \tag{3.71}$$

hence, by an application of Itô's lemma, we find the following dynamics for the T-forward stock price process

$$\begin{aligned}\frac{dFFX_T(t)}{FFX_T(t)} &= (\sigma_1^2 B_1^2(t,T) + \rho_{Xr_1}\nu(t)\sigma_1 B_1(t,T) - \rho_{r_1r_2}\sigma_2 B_2(t,T)\sigma_1 B_1(t,T))dt \\ &\quad +\nu(t)dW_X(t) + \sigma_1 B_1(t,T)dW_{r_1}(t) - \sigma_2 B_2(t,T)dW_{r_2}(t).\end{aligned} \tag{3.72}$$

By definition the forward FX-rate is a martingale process under the domestic T-forward measure. This is achieved by defining the following transformations of the Brownian motion(s):

$$\begin{aligned} dW_{r_1}(t) &\mapsto dW^T_{r_1}(t) - \sigma_1 B_1(t,T)dt, \\ dW_{r_2}(t) &\mapsto dW^T_{r_2}(t) - \rho_{r_1 r_2}\sigma_1 B_1(t,T)dt, \\ dW_X(t) &\mapsto dW^T_X(t) - \rho_{X r_1}\sigma_1 B_1(t,T)dt, \\ dW_\nu(t) &\mapsto dW^T_\nu(t) - \rho_{r_1 \nu}\sigma_1 B_1(t,T)dt. \end{aligned}$$

Hence under the domestic T-forward measure the forward FX-rate and the associated volatility process are given by

$$\frac{dFFX_T(t)}{FFX_T(t)} = \nu(t)dW^T_X(t) + \sigma_1 B_1(t,T)dW^T_{r_1} - \sigma_2 B_2(t,T)dW^T_{r_2}(t) \tag{3.73}$$

$$d\nu(t) = \kappa\Big(\psi - \frac{\rho_{r_1\nu}\sigma_1\tau}{\kappa}B_1(t,T) - \nu(t)\Big)dt + \tau dW^T_\nu(t). \tag{3.74}$$

We can simplify (3.73) by switching to logarithmic coordinates by defining $y(t) := \log(FFX_T(t))$; an application of Itô's lemma yields

$$dy(t) = -\frac{1}{2}\nu^2_F(t)dt + \nu(t)dW^T_X(t) + \sigma_1 B_1(t,T)dW^T_{r_1} - \sigma_2 B_2(t,T)dW^T_{r_2}(t) \tag{3.75}$$

$$d\nu(t) = \kappa\Big(\xi(t) - \nu(t)\Big)dt + \tau dW^T_\nu(t), \tag{3.76}$$

with:

$$\begin{aligned} \nu^2_F(t) := {} & \nu^2(t) + \sigma_1^2 B_1^2(t,T) + \sigma_2^2 B_2^2(t,T) + 2\rho_{X r_1}\nu(t)\sigma_1 B_1(t,T) \\ & -2\rho_{X r_2}\nu(t)\sigma_2 B_2(t,T) - 2\rho_{r_1 r_2}\sigma_1 B_1(t,T)\sigma_2 B_2(t,T) \end{aligned} \tag{3.77}$$

$$\xi(t) := \psi - \frac{\rho_{r_1\nu}\sigma_1\tau}{\kappa}B_1(t,T). \tag{3.78}$$

Notice that we have now reduced the system (3.60) of the variables $X(t), r_1(t), r_2(t), \nu(t)$ under the domestic risk-neutral measure, to the system (3.75) of variables $y(t)$ and $\nu(t)$ under the domestic T-forward measure. What now remains is to determine the characteristic function of this reduced system.

Determining the characteristic function of the forward log-FX rate

We will now determine the characteristic function of the forward FX rate. Since this calculation goes in a similar way as the calculation of the ordinary characteristic function of the Schöbel-Zhu-Hull-White model of Section 3.2, we restrict ourselves to the most important steps. Again we apply the Feynman-Kac theorem and reduce the search for the characteristic function of the forward-FX rate dynamics to solving a partial differential equation. That is, we try to determine the Kolmogorov backward partial differential equation of the joint probability function

$f = f(t, y, v)$. To this end we need to take into account the following covariance term

$$\begin{aligned} dy(t)dv(t) &= \Big(v(t)dW_X^T(t) + \sigma_1 B_1(t,T)dW_{r_1}^T(t) - \sigma_2 B_2(t,T)dW_{r_1}^T(t)\Big)\Big(\tau dW_v^T(t)\Big) \\ &= \Big(\rho_{Xv}\tau v(t) + \rho_{r_1 v}\tau\sigma_1 B_1(t,T) - \rho_{r_2 v}\tau\sigma_2 B_2(t,T)\Big)dt. \end{aligned} \tag{3.79}$$

Hence using (3.75) and (3.79), the Feynman-Kac theorem then implies that the solution of the following PDE

$$\begin{aligned} 0 &= f_t - \frac{1}{2}v_F^2(t)f_y + \kappa\big(\xi(t) - v(t)\big)f_v + \frac{1}{2}v_F^2(t)f_{yy} \\ &\quad + \Big(\rho_{Xv}\tau v(t) + \rho_{r_1 v}\tau\sigma_1 B_1(t,T) - \rho_{r_2 v}\tau\sigma_2 B_2(t,T)\Big)f_{yv} + \frac{1}{2}\tau^2 f_{vv}, \end{aligned} \tag{3.80}$$

subject to the terminal boundary condition $f(T, y, v) = \exp(iuy(T))$, equals the characteristic function of the forward FX-rate dynamics. Solving the above system hence leads to the following proposition.

Proposition 3.7.1 *The characteristic function of domestic T-forward log SZHW-FX-rate is given by the following closed-form solution:*

$$f(t, y, v) = \exp\Big[A(t) + B(t)y(t) + C(t)v(t) + \frac{1}{2}D(t)v^2(t)\Big], \tag{3.81}$$

where:

$$\begin{aligned} A(u,t,T) &= \frac{1}{2}(B^2 - B)V_{FX}(t,T) \qquad (3.82) \\ &\quad + \int_t^T \Big[\Big(\kappa\psi + \rho_{r_1 v}(iu-1)\tau\sigma_1 B_1(s,T) - \rho_{r_2 v}iu\tau\sigma_2 B_2(s,T)\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds, \\ B &= iu, \qquad (3.83) \\ C(u,t,T) &= -u(i+u)\frac{\Big((\gamma_3 - \gamma_4 e^{-2\gamma(T-t)}) - (\gamma_5 e^{-a_1(T-t)} - \gamma_6 e^{-(2\gamma+a_1)(T-t)}) - \gamma_7 e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}} \qquad (3.84) \\ &\quad + u(i+u)\frac{\Big((\gamma_8 - \gamma_9 e^{-2\gamma(T-t)}) - (\gamma_{10} e^{-a_2(T-t)} - \gamma_{11} e^{-(2\gamma+a_2)(T-t)}) - \gamma_{12} e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \\ D(u,t,T) &= -u(i+u)\frac{1 - e^{-2\gamma(T-t)}}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}} \qquad (3.85) \end{aligned}$$

with:

$$\begin{aligned}
\gamma &= \sqrt{(\kappa - \rho_{Xv}\tau B)^2 - \tau^2(B^2 - B)}, && (3.86)\\
\gamma_1 &= \gamma + (\kappa - \rho_{Xv}\tau B), & \gamma_2 &= \gamma - (\kappa - \rho_{Xv}\tau B),\\
\gamma_3 &= \frac{\rho_{Xr_1}\sigma_1\gamma_1 + \kappa a_1\psi + \rho_{r_1v}\sigma_1\tau(iu-1)}{a_1\gamma}, & \gamma_4 &= \frac{\rho_{Xr_1}\sigma_1\gamma_2 - \kappa a_1\psi - \rho_{r_1v}\sigma_1\tau(iu-1)}{a_1\gamma},\\
\gamma_5 &= \frac{\rho_{Xr_1}\sigma_1\gamma_1 + \rho_{r_1v}\sigma_1\tau(iu-1)}{a_1(\gamma - a_1)}, & \gamma_6 &= \frac{\rho_{Xr_1}\sigma_1\gamma_2 - \rho_{r_1v}\sigma_1\tau(iu-1)}{a_1(\gamma + a_1)},\\
\gamma_8 &= \frac{\rho_{Xr_2}\sigma_2\gamma_1 + \rho_{r_2v}\sigma_2\tau B}{a_2\gamma}, & \gamma_9 &= \frac{\rho_{Xr_2}\sigma_2\gamma_2 - \rho_{r_2v}\sigma_2\tau B}{a_2\gamma},\\
\gamma_{10} &= \frac{\rho_{Xr_2}\sigma_2\gamma_1 + \rho_{r_2v}\sigma_2\tau B}{a_2(\gamma - a_2)}, & \gamma_{11} &= \frac{\rho_{Xr_2}\sigma_2\gamma_2 - \rho_{r_2v}\sigma_2\tau B}{a_2(\gamma + a_2)},\\
\gamma_7 &= (\gamma_3 - \gamma_4) - (\gamma_5 - \gamma_6), & \gamma_{12} &= (\gamma_8 - \gamma_9) - (\gamma_{10} - \gamma_{11})
\end{aligned}$$

and:

$$\begin{aligned}
V_{FX}(t,T) \;:=\; & \frac{\sigma_1^2}{a_1^2}\Big((T-t) + \frac{2}{a_1}e^{-a_1(T-t)} - \frac{1}{2a_1}e^{-2a_1(T-t)} - \frac{3}{2a_1}\Big) && (3.87)\\
& + \frac{\sigma_2^2}{a_2^2}\Big((T-t) + \frac{2}{a_2}e^{-a_2(T-t)} - \frac{1}{2a_2}e^{-2a_2(T-t)} - \frac{3}{2a_2}\Big)\\
& - 2\rho_{r_1r_2}\frac{\sigma_1\sigma_2}{a_1a_2}\Big((T-t) + \frac{e^{-a_1(T-t)} - 1}{a_1} + \frac{e^{-a_2(T-t)} - 1}{a_2} - \frac{e^{-(a_1+a_2)(T-t)} - 1}{a_1 + a_2}\Big).
\end{aligned}$$

Proof See appendix 3.9.1. □

The strip of regularity and the decay of the characteristic function can be determined analogous to the SZHW model. The function $C(u,t,T)$ once again determines the strip of regularity, whereas $A(u,t,T)$ ensures the characteristic function decays like $\exp\big(-C(u,t,T)u^2\big)$, where the exact constant follows from an analysis similar to that in Section 3.5.

3.8 Conclusion

We have introduced the SZHW model which allows for the pricing of insurance contracts under both stochastic volatility and stochastic interest rates in conjunction with an explicit incorporation of the correlation between the underlying asset and the term structure of interest rates. As insurance contracts typically involve long maturities, they are are much more sensitive to changes in the interest rates and the volatility. Therefore, having the flexibility to correlate the underlying asset price with both the stochastic volatility and the stochastic interest rates yields a more realistic model, which is of practical importance for the pricing and hedging of long-term options.
Our model incorporates the closed-form pricing of European options by Fourier transforming the

conditional characteristic function of the asset price in closed-form. We extensively considered the numerical implementation of the pricing formulas which enables a fast and accurate valuation of European options, which is a big advantage for the calibration (and sensitivity analysis) of the model to market prices. We have also derived a closed-form pricing formula for forward starting options, which allows for a calibration of the model to forward smiles.
The SZHW model will be especially useful in the pricing and risk management of insurance contracts and other long-maturity exotic derivatives. Examples include pension products, variable and guaranteed annuities, rate of return guarantees, unit-linked contracts and exotic options like PRDC FX options which have a long-term nature. For these products it is especially important to consider the risk of the underlying in conjunction with the interest rate risk of the contract. Given empirical data on option prices our model can be used to examine the pricing and especially hedging performance of stochastic volatility models while correcting for interest rate risk. An empirical study on the relative performance of the SZHW model versus other stochastic volatility models, as well as the relative benefit of the modelling of stochastic interest rates (covered earlier by Bakshi et al. (1997)), is beyond the scope of this thesis, and is left for future research.

3.9 Appendix

In this appendix we will show that the partial differential equation (3.17)

$$f_t + \kappa(\xi(t) - v(t))f_v + \frac{1}{2}v_F^2(t)(f_{yy} - f_y) + (\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_r(t,T))f_{yv} + \frac{1}{2}\tau^2 f_{vv} = 0, \qquad (3.88)$$

subject to the terminal condition

$$f(T,y,v) = \psi(y,v) := \exp\Big(iuy(T)\Big),$$

has a solution given by (3.20) - (3.24).

To ease the notation, we from here on omit the explicit dependence on u and T in the A, B, C, D terms and hence write $A(t)$ instead of $A(u,t,T)$ for these terms. Using the ansatz

$$f(t,y,v) = \exp\Big[A(t) + B(t)y(t) + C(t)v(t) + \frac{1}{2}D(t)v^2(t)\Big], \qquad (3.89)$$

we find the following partial derivatives for $f = f(t,y,v)$:

$$\begin{aligned}
f_t &= f\cdot(A'(t) + B'(t)y(t) + C'(t)v(t) + \frac{1}{2}D'(t)v^2(t)), \quad f_y = fB(t),\\
f_v &= f\cdot(C(t) + D(t)v(t)), \quad f_{yy} = fB^2(t), \quad f_{yv} = fB(t)(C(t) + D(t)v(t))\\
f_{vv} &= f\cdot(C(t) + D(t)v^2(t)) = f\cdot(C^2(t) + D(t) + 2C(t)D(t)v(t) + D^2(t)v^2(t))
\end{aligned}$$

Substituting these partial derivatives into the partial differential equation (3.88) then gives

$$\begin{aligned}
&\Big(A'(t) + B'(t)y(t) + C'(t)v(t) + \frac{1}{2}D'(t)v^2(t)\Big) + \kappa(\xi(t) - v(t))(C(t) + D(t)v(t))\\
&+\frac{1}{2}\Big(v^2(t) + 2\rho_{Sr}v(t)\sigma B_r(t,T) + \sigma^2 B_r^2(t,T)\Big)(B^2(t) - B(t))\\
&+(\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_r(t,T))B(t)(C(t) + D(t)v(t))\\
&+\frac{1}{2}\tau^2(C^2(t) + D(t) + 2C(t)D(t)v(t) + D^2(t)v^2(t)) = 0. \qquad (3.90)
\end{aligned}$$

Collecting terms for $y(t)$,$v(t)$, and $\frac{1}{2}v^2(t)$ then yields the following four ordinary differential equa-

tions for the functions $A(t), \ldots, D(t)$:

$$\begin{aligned}
0 &= B^{'}(t), && (3.91)\\
0 &= D^{'}(t) - 2(\kappa - \rho_{Sv}\tau B)D(t) + \tau^2 D^2(t) + (B^2 - B), && (3.92)\\
0 &= C^{'}(t) + (\rho_{Sv}\tau B - \kappa + \tau^2 D)C(t) + \rho_{Sr}\sigma B_r(t,T)(B^2 - B) && (3.93)\\
&\quad + (\kappa\xi(t) + \rho_{rv}\tau\sigma B_r(t,T)B)D(t),\\
0 &= A^{'}(t) + (\kappa\xi(t) + \rho_{rv}\tau\sigma B_r(t,T)B)C(t)\\
&\quad + \frac{1}{2}\sigma^2 B_r^2(t,T)(B^2 - B) + \frac{1}{2}\tau^2(C^2(t) + D(t)). && (3.94)
\end{aligned}$$

It immediately that follows $B(t) = B$ equals a constant, subject to the boundary condition (3.9) we find

$$B = iu. \tag{3.95}$$

The second equation (3.92) yields a Riccati equation with constant coefficients with boundary condition $D(T) = 0$:

$$\begin{aligned}
D^{'}(t) &= -(B^2 - B) + 2(\kappa - \rho_{Sv}\tau B)D(t) - \tau^2 D^2(t)\\
&=: q_0 + q_1 D(t) + q_2 D^2(t)
\end{aligned}$$

Making the substitution $D(t) = \frac{-v^{'}(t)}{q_2 v(t)}$ transforms the Riccati equation into the following second order linear differential equation with constant coefficients:

$$v^{''}(t) - q_1 v^{'}(t) + q_0 q_2 v(t) = 0, \tag{3.96}$$

which solution is given by

$$\begin{aligned}
v(t) &= \gamma_1 \exp[\lambda_+(T-t)] + \gamma_2 \exp[\lambda_-(T-t)],\\
\lambda_\pm &= -\frac{q_1}{2} \pm \sqrt{q_1^2 - 4q_0 q_2}
\end{aligned}$$

Hence defining $\gamma = \sqrt{q_1^2 - 4q_0 q_2}$ we find:

$$\begin{aligned}
D(t) &= \frac{-v^{'}(t)}{q_2 v(t)} = -\frac{1}{\tau^2}\frac{\gamma_1\gamma_2 e^{\gamma(T-t)} - \gamma_1\gamma_2 e^{-\gamma(T-t)}}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}}\\
&= (B^2 - B)\frac{e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}} = -u(i+u)\frac{1 - e^{-2\gamma(T-t)}}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}} && (3.97)\\
\text{with: } \gamma &= \sqrt{(\kappa - \rho_{Sv}\tau B)^2 - \tau^2(B^2 - B)}, && (3.98)\\
\gamma_1 &= \gamma + \frac{1}{2}q_1 = \gamma + (\kappa - \rho_{Sv}\tau B), && (3.99)\\
\gamma_2 &= \gamma - \frac{1}{2}q_1 = \gamma - (\kappa - \rho_{Sv}\tau B). && (3.100)
\end{aligned}$$

Here the constants in equation (3.97) are determined from the identity $(\gamma + \frac{1}{2}q_1)(\gamma - \frac{1}{2}q_1) = -(B^2 - B)\tau^2$ and the boundary condition $D(T) = 0$.

The third equation (3.93) looks pretty daunting, but is merely a first order linear ordinary differential equation of the form $C'(t) + g(t)C(t) + h(t) = 0$. Subject to the boundary condition $C(T) = 0$ and using (3.15), we can hence represent a solution for $C(t)$ as:

$$\begin{aligned} C(t) &= \int_t^T h(s) \exp[\int_t^s g(w)dw]ds, && (3.101)\\ \text{with: } g(w) &= -(\kappa - \rho_{Sv}\tau B) + \tau^2 D(w), && (3.102)\\ h(s) &= \rho_{Sr}\sigma B_r(s,T)(B^2 - B) + (\kappa\xi(s) + \rho_{rv}\tau\sigma B_r(s,T)B)D(s) \\ &= \rho_{Sr}\sigma B_r(s,T)(B^2 - B) + (\kappa\psi + \rho_{rv}(B-1)\tau\sigma B_r(s,T))D(s). && (3.103) \end{aligned}$$

We first consider the integral over g: dividing equation (3.92) by $D(t)$, rearranging terms and integrating we find the surprisingly simple solution:

$$\begin{aligned} \int g(w)dw &= \int -(\kappa - \rho_{Sv}\tau B) + \tau^2 D(w)dw \\ &= \int (\kappa - \rho_{Sv}\tau B) - \frac{(B^2 - B)}{D(w)} - \frac{D'(w)}{D(w)}dw \\ &= \log\Big(\gamma_1 e^{\gamma(T-w)} + \gamma_2 e^{-\gamma(T-w)}\Big) + C, && (3.104) \end{aligned}$$

where C denotes the integration constant. Hence taking the exponent and filling in the required integration boundaries yields

$$\exp\Big[\int_t^s g(w)dw\Big] = \frac{\gamma_1 e^{\gamma(T-s)} + \gamma_2 e^{-\gamma(T-s)}}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}}, \qquad (3.105)$$

and after a straightforward calculation we get for $C(t)$:

$$\begin{aligned} C(t) &= \frac{1}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}} \int_t^T h(s)\Big(\gamma_1 e^{\gamma(T-s)} + \gamma_2 e^{-\gamma(T-s)}\Big)ds \\ &= (B^2 - B)\frac{\Big((\gamma_3 e^{\gamma(T-t)} - \gamma_4 e^{-\gamma(T-t)}) - (\gamma_5 e^{(\gamma-a)(T-t)} - \gamma_6 e^{-(\gamma+a)(T-t)}) - \gamma_7\Big)}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}} \\ &= -u(i+u)\frac{\Big((\gamma_3 - \gamma_4 e^{-2\gamma(T-t)}) - (\gamma_5 e^{-a(T-t)} - \gamma_6 e^{-(2\gamma+a)(T-t)}) - \gamma_7 e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \qquad (3.106) \end{aligned}$$

with $\gamma, \gamma_1, \ldots, \gamma_7$ as defined in (3.25).

Finally, by solving equation (3.94), we find the following expression for $A(t)$:

$$\begin{aligned} A(t) &= \int_t^T \frac{1}{2}(B^2 - B)\sigma^2 B_r^2(s,T)ds \\ &\quad + \int_t^T \Big[\Big(\kappa\xi(t) + \rho_{rv}\tau\sigma B_r(s,T)B\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds \\ &= -\frac{1}{2}u(i+u)V(t,T) \\ &\quad + \int_t^T \Big[\Big(\kappa\psi + \rho_{rv}(iu-1)\tau\sigma B_r(s,T)\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds \end{aligned} \tag{3.107}$$

where $V(t,T)$ can be found by simple integration and is given by

$$V(t,T) = \frac{\sigma^2}{a^2}\Big((T-t) + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\Big). \tag{3.108}$$

It is possible to write a closed-form expression for the remaining integral in (3.107). As the ordinary differential equation for $D(s)$ is exactly the same as in the Heston (1993) or Schöbel and Zhu (1999) model, it will involve a complex logarithm and should therefore be evaluated as outlined in Lord and Kahl (2007) in order to avoid any (complex) discontinuities. The main problem however lies in the integrals over $C(s)$ and $C^2(s)$, which will involve the Gaussian hypergeometric function ${}_2F_1(a,b,c;z)$. The most efficient way to evaluate this hypergeometric function (according to *Numerical Recipes*, Press and Flannery (1992)) is to integrate the defining differential equation. Since all of the terms involved in $D(u)$ are also required in $C(u)$, numerical integration of the second part of (3.107) seems to be the most efficient method for evaluating $A(t)$. Hereby we conveniently avoid any computational issues regarding complex discontinuities altogether.

3.9.1 Deriving the log FX-rate characteristic function

In this appendix we will prove that the partial differential equation (3.80), i.e.

$$\begin{aligned} 0 &= f_t + \kappa(\xi(t) - v(t))f_v + \frac{1}{2}v_F^2(t)(f_{yy} - f_y) \\ &\quad + \Big(\rho_{Xv}\tau v(t) + \rho_{r_1v}\tau\sigma_1 B_1(t,T) - \rho_{r_2v}\tau\sigma_2 B_2(t,T)\Big)f_{yv} + \frac{1}{2}\tau^2 f_{vv}, \end{aligned} \tag{3.109}$$

subject to the terminal condition $f(T,y,\sigma) = \exp(iuy(T))$ has a solution given by (3.81)-(3.86); we follow the same approach as in section (3.9), that is we use the ansatz (3.81), find the corresponding partial derivatives and substitute these in the PDE (3.109).

Expanding $v_F^2(t)$ according to (3.77) and collecting the terms for $y(t)$, $v(t)$ and $\frac{1}{2}v^2(t)$ yields the

following system of ordinary differential equations for the functions $A(t), \ldots, D(t)$:

$$
\begin{aligned}
0 &= B^{'}(t), && (3.110)\\
0 &= D^{'}(t) - 2(\kappa - \rho_{Xv}\tau B)D(t) + \tau^2 D^2(t) + (B^2 - B), && (3.111)\\
0 &= C^{'}(t) + (\rho_{Xv}\tau B - \kappa + \tau^2 D)C(t) + (\rho_{Xr_1}\sigma_1 B_1(t,T) - \rho_{Xr_2}\sigma_2 B_2(t,T))(B^2 - B)\\
&\quad + \Big(\kappa\xi(t) + (\rho_{r_1v}\tau\sigma_1 B_1(t,T) - \rho_{r_2v}\tau\sigma_2 B_2(t,T))B\Big)D(t), && (3.112)\\
0 &= A^{'}(t) + (\kappa\xi(t) + \rho_{r_1v}\tau\sigma_1 B_1(t,T)B - \rho_{r_2v}\tau\sigma_2 B_2(t,T)B)C(t)\\
&\quad + \Big(\frac{1}{2}\sigma_1^2 B_1^2(t,T) + \frac{1}{2}\sigma_2^2 B_2^2(t,T) - \rho_{r_1r_2}\sigma_1 B_1(t,T)\sigma_2 B_2(t,T)\Big)(B^2 - B)\\
&\quad + \frac{1}{2}\tau^2(C^2(t) + D(t)) && (3.113)
\end{aligned}
$$

Hence we end up with an analogue system of ordinary differential equations as in section (3.9): the first two differential equations (3.110) and (3.111) for B and $D(t)$ are equivalent to (3.91) and (3.92) whose solutions are given in the equations (3.95) and (3.97)-(3.100). The third equation (3.112) for $C(t)$ looks pretty daunting, but is again merely a first order linear differential equation of the form $C'(t) + g(t)C(t) + h(t) = 0$, with associated boundary condition $C(T) = 0$. Hence expanding $\xi(t)$ according to (3.78), we can represent a solution for $C(t)$ as:

$$
\begin{aligned}
C(t) &= \int_t^T h(s)\exp[\int_t^s g(w)dw]ds, && (3.114)\\
\text{with: } g(w) &= -(\kappa - \rho_{Xv}\tau B) + \tau^2 D(w), && (3.115)\\
h(s) &= \Big(\rho_{Xr_1}\sigma_1 B_1(s,T) - \rho_{Xr_2}\sigma_2 B_2(s,T)\Big)(B^2 - B)\\
&\quad + \Big(\kappa\xi(s) + (\rho_{r_1v}\tau\sigma_1 B_1(s,T) - \rho_{r_2v}\tau\sigma_2 B_2(s,T))B\Big)D(s)\\
&= \rho_{Xr_1}\sigma_1 B_1(s,T)(B^2 - B) + (\kappa\psi + \rho_{r_1v}(B-1)\tau\sigma_1 B_1(s,T))D(s)\\
&\quad - \rho_{Xr_2}\sigma_2 B_2(s,T)(B^2 - B) - (\rho_{r_2v}B\tau\sigma_2 B_2(s,T))D(s). && (3.116)
\end{aligned}
$$

Now notice that the integral over g is equivalent to (3.104), hence its solution is given by equation (3.105), i.e.

$$
\exp\Big[\int_t^s g(w)dw\Big] = \frac{\gamma_1 e^{\gamma(T-s)} + \gamma_2 e^{-\gamma(T-s)}}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}}, \qquad (3.117)
$$

with γ, γ_1 and γ_2 defined in (3.86). Substituting this expression into (3.114) we find (after a long

but straightforward calculation) for $C(t)$:

$$\begin{aligned} C(t) &= (B^2-B)\frac{\Big((\gamma_3 e^{\gamma(T-t)} - \gamma_4 e^{-\gamma(T-t)}) - (\gamma_5 e^{(\gamma-a_1)(T-t)} - \gamma_6 e^{-(\gamma+a_1)(T-t)}) - \gamma_7\Big)}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}} \\ &\quad -(B^2-B)\frac{\Big((\gamma_8 e^{\gamma(T-t)} - \gamma_9 e^{-\gamma(T-t)}) - (\gamma_{10} e^{(\gamma-a_2)(T-t)} - \gamma_{11} e^{-(\gamma+a_2)(T-t)}) - \gamma_{12}\Big)}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}} \\ &= -u(i+u)\frac{\Big((\gamma_3 - \gamma_4 e^{-2\gamma(T-t)}) - (\gamma_5 e^{-a_1(T-t)} - \gamma_6 e^{-(2\gamma+a_1)(T-t)}) - \gamma_7 e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}} \\ &\quad +u(i+u)\frac{\Big((\gamma_8 - \gamma_9 e^{-2\gamma(T-t)}) - (\gamma_{10} e^{-a_2(T-t)} - \gamma_{11} e^{-(2\gamma+a_2)(T-t)}) - \gamma_{12} e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \end{aligned} \tag{3.118}$$

with $\gamma, \gamma_1, \ldots, \gamma_{12}$ as defined in (3.86).

Finally, by solving equation (3.113), we find the following expression for $A(t)$:

$$\begin{aligned} A(t) &= \int_t^T \frac{1}{2}(B^2-B)\Big(\sigma_1^2 B_1^2(s,T) + \sigma_2^2 B_2^2(s,T) - 2\rho_{r_1 r_2}\sigma_1 B_1(s,T)\sigma_2 B_2(s,T)\Big)ds \\ &\quad + \int_t^T \Big[\kappa\Big(\xi(s) + \rho_{r_1 v} B\tau\sigma_1 B_1(t,T) - \rho_{r_2 v} B\tau\sigma_2 B_2(t,T)\Big)\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds \\ &= \frac{1}{2}(B^2-B)V_{FX}(t,T) \\ &\quad + \int_t^T \Big[(\kappa\psi + \rho_{r_1 v}(iu-1)\tau\sigma_1 B_1(s,T) - \rho_{r_2 v} iu\tau\sigma_2 B_2(s,T))C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds, \end{aligned} \tag{3.119}$$

where $V_{FX}(t,T)$ can found by simple integration and is given by:

$$\begin{aligned} V_{FX}(t,T) &:= \frac{\sigma_1^2}{a_1^2}\Big((T-t) + \frac{2}{a_1}e^{-a_1(T-t)} - \frac{1}{2a_1}e^{-2a_1(T-t)} - \frac{3}{2a_1}\Big) \\ &\quad + \frac{\sigma_2^2}{a_2^2}\Big((T-t) + \frac{2}{a_2}e^{-a_2(T-t)} - \frac{1}{2a_2}e^{-2a_2(T-t)} - \frac{3}{2a_2}\Big) \\ &\quad - 2\rho_{r_1 r_2}\frac{\sigma_1\sigma_2}{a_1 a_2}\Big((T-t) + \frac{e^{-a_1(T-t)}-1}{a_1} + \frac{e^{-a_2(T-t)}-1}{a_2} - \frac{e^{-(a_1+a_2)(T-t)}-1}{a_1+a_2}\Big). \end{aligned} \tag{3.120}$$

Analogous to (3.107), integrating over the $C(s)$ and $C^2(s)$ terms in (3.119) seems to be the most efficient method to evaluate $A(t)$.

Chapter 4

Generic pricing of FX, Inflation and Stock Options under Stochastic Interest Rates and Stochastic Volatility

*This chapter is based on:

Alexander van Haastrecht and Antoon Pelsser (2008). "Generic pricing of FX, Inflation and Stock options under Stochastic Interest Rates and Stochastic Volatility", *working paper*.

4.1 Introduction

The markets for long maturity and hybrid derivatives are developing more and more. Not only are increasingly exotic structures created, also the markets for plain vanilla derivatives are growing. One of the recent advances is the development of long maturity option markets across various asset classes; during the last years long maturity securities, such as Target Auto Redemption Notes (TARN) equity-interest rate options (e.g. see Caps (2007)), Power-Reverse Dual-Currency (PRDC) Foreign Exchange (FX) swaps (e.g. see Piterbarg (2005)) and inflation-indexed Limited Price Indices (LPI) structures (e.g see Mercurio (2005) or Mercurio and Moreni (2006a)) have become increasingly popular. For FX, inflation and hybrid structures, which explicitly depend on future interest rates evolutions, it is apparent that the use of stochastic interest rates is crucial in any derivative pricing model. However, the addition of stochastic rates is also important for the pricing and in particularly the hedging of long maturity equity derivatives (e.g. see Bakshi et al. (2000)). First, the option's rho, which measures/hedges the interest rate risk of the derivative, is increasing with time to maturity. Secondly, the stochastic interest rates are important for exotic option pricing since the numeraire is the discount bond associated with the maturity of the option: because the long term interest rates are to a reasonable degree correlated with FX/inflation/equity indices, the rates directly influence the pricing kernel used in exotic option pricing.
Most investment banks have now standardized a three-factor modelling framework to price cross-currency (i.e. FX and inflation) options (see Piterbarg (2005), Sippel and Ohkoshi (2002) or Jar-

row and Yildirim (2003)). Here the index follows a log-normal process, and the interest rates of the two currencies are modelled with Gaussian Hull and White (1993) frameworks. The choice for Gaussian models for the interest rates and log-normality for the index has allowed for a very efficient, essentially closed-form, calibration to at-the-money options on the index, i.e. on the FX-rate or stock price. The assumption of log-normality for an index, though technically very convenient, does not find its justification in the financial equity markets (e.g. see Bakshi et al. (1997)), the FX markets (e.g see Piterbarg (2005), Caps (2007)) nor in the inflation markets (e.g. see Mercurio and Moreni (2006a), Mercurio and Moreni (2009) or Kenyon (2008)). In fact, the markets for these products exhibit a strong volatility skew or smile, implying log index returns deviating from normality and suggesting the use of skewed and heavier tailed distributions. Moreover many multi-currency structures (like LPIs or PRDCs) are particularly sensitive to volatility skews/smiles as they often incorporate multiple strikes as well as callable/knockout components. Hence appropriate exotic option pricing models, which need to quantify the volatility exposure in such structures, should at least be able to incorporate the smiles/skews in the vanilla markets. While various methods exist to incorporate volatility smiles (i.e. local volatility, stochastic volatility and/or jumps), the calibration of such models is by no means trivial. A skew-mechanism is normally applied to the forward index price (i.e. the FX-rate, CPI/Equity index), however to price multi-currency options also a term-structure involving various time points of the forward index is required. The incorporation of stochastic interest rates makes the connection between the two particularly non-trivial (e.g. see Piterbarg (2005) or Antonov et al. (2008)). Though the issue is important, Piterbarg (2005) even dubs it as 'perhaps even the most important current outstanding problems for quantitative research departments worldwide', there is remarkably little literature available on the subject even though the problem attracted both the attention of practitioners as well as from academia (e.g. see van der Ploeg (2007)).

Only very recently a few approaches have been suggested. A local volatility approach is used in Piterbarg (2005) who derives approximating formulas for calibration. Andreasen (2006) combines Heston (1993) stochastic volatility with independent stochastic interest rates drivers and derives closed-form Fourier expressions for vanilla options. To correlate the independent rate drivers with the FX-rate, Andreasen (2006) uses an indirect approach in the form of a volatility displacement parameter, which has some disadvantages as that it can lead to extreme model parameters (e.g. see Antonov et al. (2008)). This latter framework is generalized by Kainth and Saravanamuttu (2007), which authors consider the pricing of double no-touch options in a model with stochastic correlation and double Heston dynamics for the stochastic volatility. The calibration of FX options stochastic interest rates with Heston (1993) stochastic volatility under a full correlation structure is undertaken in Antonov et al. (2008) who use a so-called "Markovian" projection to derive approximation formulas. Though their projection technique is elegant, the quality of their approximation deteriorates for larger maturities or more extreme model parameters. The exact pricing of FX options under Schöbel and Zhu (1999) stochastic volatility, single-factor Gaussian rates and a full correlation structure was recently has been considered in Chapter 3. The modelling of inflation smile, has been considered by several authors, e.g. see Belgrade et al. (2004), Kenyon (2008) and Mercurio and Moreni (2009).

In this chapter, building on the results of Chapter 3, Antonov et al. (2008), Andreasen (2006) and Piterbarg (2005), we consider the pricing of foreign exchange, inflation and stock options under

Schöbel and Zhu (1999) and Heston (1993) stochastic volatility and under multi-factor Gaussian interest rates with a full correlation structure. Since stock and FX options are a special (nested) cases of inflation-indexed caps/floors[1] we will mainly focus on the pricing of inflation index derivatives. The stock and FX model option pricing formulas hence follow directly from our generalization of the foreign exchange inflation framework of Jarrow and Yildirim (2003). The setup of the chapter is as follows: in Section 4.2 we introduce our new model, Section 4.3 considers the basic vanilla derivatives and the pricing methodology used for the model. In Section 4.4 we derive the characteristic functions required for the Fourier-based pricing methods: under Schöbel and Zhu (1999) stochastic volatility we can derive the cf. of our model in closed-form, under Heston (1993) stochastic volatility it is extremely challenging derive the characteristic function of the general model in closed-form, nonetheless we demonstrate how the characteristic function of the special (uncorrelated) case can be used as a simple and efficient control variate for the general model. Finally, Section 4.6 concludes.

4.2 The model

Before introducing the general model, we first consider the Jarrow and Yildirim (2003) model which can be seen as a special (degenerate) case of our model. The Jarrow and Yildirim (2003) framework for modelling inflation and real rates is based on a foreign-exchange analogy between the real and nominal economy. That is, the real rates are seen as interest rates in the real (foreign) economy, whereas the nominal rates represent the interest rates in the nominal (domestic) economy. The inflation index then represents the exchange rate between the nominal (domestic) and real (foreign) currency. There are several assumptions that can be made with respect to the evolution of these dynamics: we first discuss the classical Jarrow and Yildirim (2003) model, before turning to generalized model setups. For clarity of exposition we will use constant model parameters in both frameworks, it is however obvious to extend this to time-dependent model parameters.

4.2.1 Special case: Jarrow-Yildirim (2003) model

Jarrow and Yildirim (2003) assume that the real-world evolution of the nominal and real instantaneous forward rates is given by HJM-dynamics, whereas the inflation index is log-normally distributed. Though several choices can be made with respect to the volatility structure within an HJM-model, Jarrow and Yildirim (2003) assume that the forward rate volatilities are given by $\sigma e^{a(T-t)}$. Using the equivalent formulation of the HJM-model in terms of instantaneous short rates then results in the following dynamics under the risk-neutral measure Q^n, see Jarrow and Yildirim (2003).

[1] In our framework an inflation option can be seen as forward-starting FX-option, hence the pricing of FX-option follows from the pricing of an inflation option by setting the forward starting date equal to the current date. A stock option can be seen as an FX-option in which (possibly deterministic) foreign interest rates represent the continuous dividend yield.

Proposition 4.2.1 *The Q_n dynamics of the instantaneous nominal rate $n(t)$, real rate $r(t)$ and the inflation index $I(t)$ are given by*

$$\begin{aligned} dn(t) &= [\vartheta_n(t) - a_n n(t)]dt + \sigma_n dW_n(t), && (4.1)\\ dr(t) &= [\vartheta_r(t) - \rho_{r,I}\sigma_I\sigma_r - a_r r(t)]dt + \sigma_r dW_r(t), && (4.2)\\ dI(t) &= I(t)[n(t) - r(t)]dt + \sigma_I I(t) dW_I(t), && (4.3) \end{aligned}$$

with $a_n, a_r, \sigma_n, \sigma_r, \sigma_I$ positive parameters (possibly time-dependent) and where (W_n, W_r, W_I) is a Brownian motion under Q^n (i.e. with the nominal bank-account as numeraire) with correlations $\rho_{n,r}$,$\rho_{n,I}$ and $\rho_{r,I}$, and with $\vartheta_n(T)$ and $\vartheta_r(t)$ deterministic functions which are used to exactly fit the term structure the nominal and real interest rates.

Note that the covariance in (4.2) between the inflation and real rate term $\rho_{r,I}\sigma_I\sigma_r$, arises due to a change of the real to the nominal risk-neutral measure, e.g. see Geman et al. (1996). With this particular volatility structure, Jarrow and Yildirim (2003) thus assumed that both the nominal as real (instantaneous) rates followed Hull and White (1993) processes under their own risk-neutral measure. Moreover they showed that the real rate still follows an Ornstein-Uhlenbeck process under the nominal risk-neutral measure Q_n and that the inflation index $I(T)$ for each $t < T$ is log-normal distributed under Q_n, in particular one can write:

$$I(T) = I(t)\exp\Big(\int_t^T [n(u) - r(u) - \frac{1}{2}\sigma_I^2]du + \int_t^T \sigma_I dW_I(u)]\Big). \tag{4.4}$$

The main advantage of the Jarrow and Yildirim (2003) model is its tractability; one for example has analytical formulas for the prices of year on years inflation-indexed swaps (see Brigo and Mercurio (2006) pp.653, formula 16.15) and closed-form Black-like formulas for the prices of inflation-indexed caplets (see Brigo and Mercurio (2006) pp.663, formula 17.4). Though one can challenge the one-factor rate models, the biggest disadvantage of the Jarrow and Yildirim (2003) model for the pricing of inflation derivatives is most often the log-normal assumption of the inflation index, which does not find its justification in the markets, e.g. see Mercurio and Moreni (2006b), Kenyon (2008) or Kruse (2007).

4.2.2 General model

In this section we will present a general model, which can be seen as an extension of the models of Jarrow and Yildirim (2003) and the SZHW model from Chapter 3. The first extension is that instead of one-factor Hull and White (1993) models for the instantaneous nominal and real rates, we let the short rate be driven by multiple (correlated) factors. We use an equivalent additive formulation for Hull-White interest rates in terms of a sum of correlated Gaussian factors plus a deterministic function, i.e. we write the model into an affine factors formulation, e.g. Duffie et al. (2000) and Duffie et al. (2003). The deterministic factor can be chosen as to exactly fit the term structure of the nominal or real interest rates, e.g. see Brigo and Mercurio (2006) or Pelsser (2000). The nominal short interest rate be driven by K correlated Gaussian factors and the real

short rate by M factors, the multi-factor Gaussian interest can hence be represented as:

$$n(t) = \varphi_n(t) + \bar{\mathbf{1}} \cdot X_n(t), \qquad r(t) = \varphi_r(t) + \bar{\mathbf{1}} \cdot X_r(t), \tag{4.5}$$

with $\bar{\mathbf{1}}$ a vector of ones and where $\varphi_n(t), \varphi_r(t)$ are the deterministic functions to fit the nominal and real term structure (in particular $\varphi_n(0) = n(0)$ and $\varphi_r(0) = r(0)$) and with $X_n(t)$, X_r Gaussian rate vectors which drive respectively the nominal and real rates, i.e. with typical elements the Gaussian factors $x_n^i(t)$, $x_r^j(t)$.

The second extension in our model is that we make the volatility σ_I stochastic. Moreover we let this stochastic volatility factor, which we from now on denote by $\nu(t)$, be correlated with the instantaneous interest rates and the inflation index. Two popular choices within the stochastic volatility literature are the models of Heston (1993) and Schöbel and Zhu (1999). In the latter the volatility is modeled as an Ornstein-Uhlenbeck process

$$d\nu(t) = \kappa[\psi - \nu(t)]dt + \tau dW_\nu(t), \quad \nu(0) = \nu_0 \tag{4.6}$$

with κ, ψ, σ_ν positive parameters and where $W_\nu(t)$ is a Brownian motion that is correlated with the other driving factors, especially the asset price. Note that we have a positive probability that $\nu(t)$ in (4.6) can become negative, which will cause the correlation between $\nu(t)$ and the other driving factors to (temporarily) change sign.
The most popular stochastic volatility model, however, is the Heston (1993) model which mainly owns its popularity due to its analytical tractability. In the Heston model, the variance is modeled by the following Feller/CIR/square-root process

$$d\nu^2(t) = \kappa[\theta - \nu^2(t)]dt + \xi\nu(t)dW_\nu(t), \quad \nu^2(0) = \nu_0^2 \tag{4.7}$$

with κ, θ, ξ positive parameters and where W_ν represents again a Brownian that is correlated with the other model factors.

With the multi-factor Gaussian rates and with stochastic volatility a la Schöbel-Zhu or Heston, we come to the following the dynamics of our model. The $\mathcal{Q}_n$ dynamics of the K-factor instantaneous nominal rate $n(t)$, M-factor real rate $r(t)$ and the inflation index $I(t)$ are given by

$$\begin{aligned}
dx_n^i(t) &= -a_n^i x_n^i(t)dt + \sigma_n^i dW_{n_i}(t) \quad i = 1, \ldots, K, && (4.8)\\
dx_r^j(t) &= [-a_r^j x_r^j(t) - \rho_{I,x_r^j}\nu(t)\sigma_r^j]dt + \sigma_r^j dW_{r_j}(t) \quad j = 1, \ldots, M, && (4.9)\\
dI(t) &= I(t)[n(t) - r(t)]dt + \nu(t)I(t)dW_I(t) && (4.10)
\end{aligned}$$

with $a_n^i(t), a_r^j, \sigma_n^i(t), \sigma_r^j$ positive parameters, $\nu(t)$ the stochastic volatility factor with dynamics given by (4.6) or (4.7), and where $W_*(t) := (W_{n_1}, \ldots, W_{n_K}, W_{r_1}, \ldots, W_{r_M}, W_\nu)$ is a Brownian motion under $\mathcal{Q}^n$ with (possibly) a full correlation structure.

The multi-factor Gaussian model is still very tractable; one, for example, has the following

analytical formulas for zero-coupon bond prices:

$$P_n(t,T) = \mathbb{E}_n\left\{e^{-\int_t^T n(u)du}\right\} = A_n(t,T)e^{-B_n(t,T)X_n(t)}, \tag{4.11}$$

$$P_r(t,T) = \mathbb{E}_r\left\{e^{-\int_t^T r(u)du}\right\} = A_r(t,T)e^{-B_r(t,T)X_r(t)}. \tag{4.12}$$

with $B_n(t,T), B_r(t,T)$ vectors with typical elements $B_n^i(t,T), B_r^i(t,T)$, and where $A_n(t,T), A_r(t,T), B_n^i(t,T), B_r^i(t,T)$ are affine functions, e.g. see Appendix 4.7.2. A useful quantity for the pricing of inflation-indexed options will turn out the be the forward inflation index $I_F(t)$ under under the nominal T-forward measure for a general maturity T, i.e.

$$I_F(t) = I(t)\frac{P_r(t,T)}{P_n(t,T)}. \tag{4.13}$$

Hence since $I_F(T) = I(T)$, we can directly substitute the forward inflation index dynamics for the inflation index, to price European time-T options. In the following subsection we will derive the dynamics of $I_F(t)$ under the nominal T-forward measure.

Dynamics under the T-forward measure

Using the change of numeraire technique of Geman et al. (1996), we will now derive the dynamics of our model under the T-forward measure for a general maturity T. Note that under their risk-neutral measures the nominal and real discount bond prices follows the processes

$$\frac{dP_n(t,T)}{P_n(t,T)} = n(t)dt + \Sigma_n(t,T)dW_n(t), \quad \frac{dP_r(t,T)}{P_r(t,T)} = r(t)dt + \Sigma_r(t,T)dW_r(t), \tag{4.14}$$

where $\Sigma_i(t,T)$, $i \in \{n,r\}$ denotes the vector of zero bond volatilities, with typical element $\sigma_i^k B_i^k(t,T)$, and with W_i a vector Brownian Motion. Hence, by an application of Itô's lemma, we come to the following proposition.

Proposition 4.2.2 *The Q_n^T dynamics for the T-forward asset price, under the T-forward Brownian Motion $W^*(t)$, has the following SDEs*

$$dx_n^k(t) = \left[-a_n^k x_n^k - R_{x_n^i,X_n}\Sigma_n(t,T)\right]dt + \sigma_n^k dW_T^*(t), \qquad k = 1,\ldots,K, \tag{4.15}$$

$$dx_r^j(t) = \left[-a_r^j x_r^j - \sigma_r^j \rho_{I,x_r^j}\nu(t) - \sigma_j^r R_{x_r^j,X_n}\Sigma_n(t,T)\right]dt + \sigma_r^j dW_T^*(t), \qquad j = 1,\ldots,M, \tag{4.16}$$

$$\frac{dI_F(t)}{I_F(t)} = \left(\nu(t) + \Sigma_n(t,T) - \Sigma_r(t,T)\right)dW_T^*(t) \tag{4.17}$$

where $R_{x_n^i,X_n}$, $R_{x_r^j,X_n}$ denote the correlation vectors, between respectively $x_n^i(t)$, $x_r^j(t)$ and the vector of nominal interest rate drivers $X_n(t)$. The stochastic volatility SDE in the Schöbel-Zhu case is

given by

$$dv(t) = \kappa\Big[\xi(t) - v(t)\Big]dt + \tau dW_T^*(t), \tag{4.18}$$

$$\xi(t) = \psi - \tau\frac{R_{v,X_n}}{\kappa}\Sigma_n(t) \tag{4.19}$$

while the Heston dynamics become

$$dv^2(t) = \kappa\Big[\zeta(t) - v^2(t)\Big]dt + \xi v(t)dW_T^*(t) \tag{4.20}$$

$$\zeta(t) = \theta - v(t)\xi\frac{R_{v,X_n}}{\kappa}\Sigma_n(t). \tag{4.21}$$

Correlations are introduced via vector volatilities in the above models. Note that this the Heston parametrization (4.20) is consistent with the dynamics as provided in (2.23) of Chapter 2. However, to make the notation between the Schöbel-Zhu and Heston model more uniform, we adopt a slightly different notation here.

We can simplify (4.17) further, by switching to logarithmic coordinates: defining

$$z(t) := \log I_F(t) = \log\left(I(t)\frac{P_r(t,T)}{P_n(t,T)}\right), \tag{4.22}$$

and an application of Itô's lemma yields

$$dz(t) = -\frac{1}{2}v_F^2(t)dt + v_F(t)dW^*(t), \tag{4.23}$$

with $v_F(t) := \Big[v(t) + \Sigma_n(t,T) - \Sigma_r(t,T)\Big]$ the instantaneous variance of the forward inflation index (explicitly defined in (4.49)). Notice that we now have transformed the system of (4.2.1) of the variables $x_n^1(t), \ldots, x_n^K(t), x_r^1(t), \ldots, x_r^M(t), I(t), v(t)$, under the nominal risk-neutral measure, to the system (4.23)-(4.18) of variables $z(t), v_F(t)$, under the T-forward measure. This latter system will be used to determine characteristic function of log inflation rate in our model, see Section 4.4.

4.3 Pricing and Applications

In this section we will briefly discuss the main vanilla inflation, FX and equity derivatives and discuss how these securities can be priced in closed-form by our model. Before turning to the market-specific structures, we first consider the general pricing methodology.

4.3.1 Pricing

Recall from Section 2.5 from Chapter 2 that under the T-forward measure Q^T, we can write the following for the price $C_T(k)$ of an European option ($\omega = 1$ for a call, $\omega = -1$ for a put) maturing

at time T, with strike $K = \exp(k)$, on an asset I:

$$C_T(k,\omega) = P_n(t,T)\mathbb{E}_n^{Q^T}\left\{\left[\omega\Big(I_F^T(T) - K\Big)\right]^+ \Big|\mathcal{F}_t\right\} \tag{4.24}$$

where $P_n(t,T)$ denotes the price of a (pure) discount bond and $I_F^T(t) := \frac{I(t)}{P_n(t,T)}$ denotes the T-forward index price. This expression expression can be numerically evaluated by means of a Fourier inversion of the characteristic function ϕ_T of the T-forward log index price $z(T) := \log I_F^T(t)$, see Chapter 2. That is, provided that the regularity conditions for the Fourier Transformations are satisfied, i.e. $\alpha > 0$ for a call ($\omega = 1$) and $\alpha > 1$ for a put ($\omega = -1$), one can write the following for the corresponding European option price:

$$C_T(k,\omega,\alpha) = \frac{P_n(t,T)}{\pi}\int_0^\infty \mathrm{Re}\Big(e^{-(\omega\alpha+iv)k}\psi_T(v,\omega,\alpha)\Big)dv, \tag{4.25}$$

with

$$\psi_T(v,\omega,\alpha) := \frac{\phi_T\Big(v-(\omega\alpha+1)i\Big)}{(\omega\alpha+iv)(\omega\alpha+1+iv)}, \tag{4.26}$$

and where $\phi_T(u) := \mathbb{E}^{Q^T}\Big[\exp\Big(iuz(T)\Big)\Big|\mathcal{F}_t\Big]$ denotes the T-forward conditional characteristic function of the log index price. The characteristic and forward characteristic function under Schöbel and Zhu (1999) volatility can respectively be found in proposition 4.4.1 and 4.4.2, where under Heston (1993) volatility these can be found in proposition 4.4.5 and 4.4.8.

4.3.2 Inflation derivatives

Before dealing with the pricing of inflation-index derivatives within the general model (4.8), we first discuss the main (vanilla) inflation-indexed securities. Hereby we adopt the notation that is used in Brigo and Mercurio (2006) and Mercurio (2005), to which authors we also refer for an excellent overview of interest rate and inflation-indexed derivatives and models.

Inflation-indexed swaps

Given a set of payment dates $T_1,\ldots,T_M$, an inflation-indexed swap (IIS) is a swap where, on each date, party A pays party B the inflation rate over a predefined period, while party B pays party A a fixed rate. This inflation rate is calculated as the percentage return of the inflation index (e.g. HICP ex Tobacco in the Eurozone) over the time interval it applies to. The two main IIS contracts that are traded in the markets are the zero-coupon inflation-indexed swap (ZCIIS) and the year-on-year inflation-indexed swap (YYIIS).
In the ZCIIS, the payoff at time T_M, assuming $T_M = M$ years, party B pays party A the fixed amount

$$N[(1+K)^M - 1], \tag{4.27}$$

where K is the strike (e.g. the break-even inflation rate) and N the nominal value of the contract. In exchange, party A pays party B, at the time final time T_M, the floating amount of

$$N\Big[\frac{I(T_M)}{I_0} - 1\Big], \tag{4.28}$$

with $I(T_M), I_0$ the inflation/CPI index respectively at time T_M and T_0. In the YYIIS, at each time T_i, party B pays party A the fixed amount

$$N\phi_i K, \tag{4.29}$$

where ϕ_i denotes the fixed-leg year fraction for the interval $[T_{i-1},T_i]$, and N the nominal value of the YYIIS. In exchange, at each time T_i, party A pays party B the floating amount

$$N\psi_i\Big[\frac{I(T_i)}{I(T_{i-1})} - 1\Big], \tag{4.30}$$

where ψ_i denotes the fixed leg year fraction for the interval $[T_{i-1},T_i]$ ($T_0 := 0$).

Let P_n and P_r respectively denote the (zero-coupon) discount bond prices of the real and nominal economy, then standard no-arbitrage theory and some straightforward rewriting show that the price of an ZCIIS (zero-coupon inflation-indexed swap) can be expressed as

$$\mathbf{ZCIIS}(t, T_M, I_0, N, K) = N\Big[\frac{I(t)}{I_0}P_r(t, T_M) - (1 + K)^M\Big], \tag{4.31}$$

which quantity is model-independent. That is, the above price is not based on any specific assumptions on the evolution of the (real and nominal) interest rates, but simply follows from the absence of arbitrage. This is an important fact, since it allows us to calibrate our model appropriately; more specifically, it allows us to strip, without ambiguity, real zero-coupon bond prices, from the quotes prices of ZCIIS. More specifically, given a set of market quotes of $K = K(T_M)$ at time $t = 0$, we can use equation (4.31) together with the net present value (4.27) to determine discount bonds of the real economy, i.e.

$$P_r(0, T_M) = P_n(0, T_M)(1 + K(T_M))^M. \tag{4.32}$$

A completely different story applies to the valuation of a YYIIS (year-on-year inflation-indexed swap), which in fact depends on the evolution of the underlying quantities and hence its price is model dependent; note that the value at time $t < T_{i-1}$ of the payoff (4.30) at time T_i is

$$\begin{aligned}\mathbf{YYIIS}(t, T_{i-1}, T_i, \psi_i, N) &= N\psi_i\mathbb{E}_n\Big\{e^{-\int_t^{T_i} n(u)du}\Big[\frac{I(T_i)}{I(T_{i-1})} - 1\Big]\Big|\mathcal{F}_t\Big\} \\ &= N\psi_i\mathbb{E}_n\Big\{e^{-\int_t^{T_i} n(u)du}\frac{I(T_i)}{I(T_{i-1})}\Big|\mathcal{F}_t\Big\} - N\psi_i P_n(t, T_i),\end{aligned} \tag{4.33}$$

where $\mathbb{E}_n$ denotes the expectation under the nominal risk-neutral measure. We briefly comment

on why the latter expectation is model dependent, first notice that

$$\mathbb{E}_n\left\{e^{-\int_t^{T_i} n(u)du}\frac{I(T_i)}{I(T_{i-1})}\Big|\mathcal{F}_t\right\} = \mathbb{E}_n\left\{e^{-\int_t^{T_{i-1}} n(u)du}P_r(T_{i-1},T_i)\Big|\mathcal{F}_t\right\}, \tag{4.34}$$

hence we can interpret the expectation from (4.33) as the nominal price of a derivative that payoffs off (in nominal units), the real zero-coupon bond price $P_r(T_{i-1},T_i)$ at time T_i. Alternatively we can also evaluate the latter expectation under a different measure, e.g. see Geman et al. (1996). Denote with Q_n^T as the nominal T-forward measure for some maturity T and let $\mathbb{E}_n^T$ represent the expectation taken under this measure, then we can write (4.34) as:

$$\mathbb{E}_n\left\{e^{-\int_t^{T_{i-1}} n(u)du}P_r(T_{i-1},T_i)\Big|\mathcal{F}_t\right\} = P_n(t,T_{i-1})\mathbb{E}_n^{T_{i-1}}\left\{P_r(T_{i-1},T_i)\Big|\mathcal{F}_t\right\}. \tag{4.35}$$

If the nominal or real rates are deterministic, then this expectation would reduce to the present value (in nominal units) of the forward price of the real zero-coupon bond, i.e. we would then have

$$P_n(t,T_{i-1})\mathbb{E}^{T_{i-1}}\left\{P_r(T_{i-1},T_i)\Big|\mathcal{F}_t\right\} = P_r(T_{i-1},T_i)P_n(t,T_{i-1}). \tag{4.36}$$

However for inflation-linked derivative pricing purposes it is usually desirable (if not necessary) that real rates are stochastic, and the expectation of (4.33) is model dependent. In fact, if the nominal and real rates are correlated (and hence stochastic), the change of measure will change the drift of the real rate $r(t)$ and hence also the expectation of (4.35). In interest rate terms, this effect is known under the term convexity adjustment, e.g. see Pelsser (2000) or Brigo and Mercurio (2006). For example if one assumes one-factor Gaussian rates (as in the JY model), one will see this convexity effect for any non-zero correlation coefficient between the nominal and real rates.

Finally note that we can also evaluate the expectation of (4.33) under the T_i-forward measure, i.e.

$$\mathbf{YYIIS}(t,T_{i-1},T_i,\psi_i,N) = N\psi_i P(t,T_i)\mathbb{E}_n^{T_i}\left\{\frac{I(T_i)}{I(T_{i-1})}\Big|\mathcal{F}_t\right\} - N\psi_i P_n(t,T_i). \tag{4.37}$$

This latter interpretation, which expresses the YYIIS (year-on-year inflation-indexed swap) as the T_i-forward expectation of the return on the inflation index, is very useful for our pricing methodology (see Section 4.3.1), because it expresses the price of a YYIIS in terms of the distribution of $\frac{I(T_i)}{I(T_{i-1})}$ under the T_{i-1}-forward measure.

Inflation-indexed caplets/floorlets

An inflation-indexed caplet can be seen as a call option on the inflation rate implied by the inflation (e.g. CPI) index. Analogously, an inflation-indexed floorlet can be seen as put option on the same inflation rate. In formulas we can write the following for the payoff of an IICplt (inflation-indexed caplet/floorlet) at time T_i

$$N\psi_i\left[\omega\left(\frac{I(T_i)}{I(T_{i-1})} - 1 - \kappa\right)\right]^+, \tag{4.38}$$

where N denotes the nominal value of the contract, κ the strike, ψ_i the year fraction for the interval $[T_{i-1},T_i]$ and $\omega = 1$ for a caplet and $\omega = -1$ for a floorlet. Setting $K := 1 + \kappa$, standard no-arbitrage theory implies that the value of the payoff (4.38) at time $t \leq T_{i-1}$ is

$$\begin{aligned} \mathbf{IICplt}(t, T_{i-1}, T_i, \psi_i, K, N, \omega) &= N\psi_i \mathbb{E}_n\left\{ e^{-\int_t^{T_i} n(u)du}\left[\omega\left(\frac{I(T_i)}{I(T_{i-1})} - K\right)\right]^+ |\mathcal{F}_t\right\} \\ &= N\psi_i P_n(t,T_i)\mathbb{E}_n^{T_i}\left\{\left[\omega\left(\frac{I(T_i)}{I(T_{i-1})} - K\right)\right]^+ |\mathcal{F}_t\right\}. \end{aligned} \tag{4.39}$$

Since (4.39) is equivalent to a call option on the forward return of the inflation index, the pricing of an inflation-indexed caplet/floorlet is thus very similar to that of a forward starting (cliquet) option.

Pricing

The crucial quantity for the pricing of the inflation-indexed derivatives in our model (4.8) is the log-return $y(T_{i-1}, T_i)$ of the inflation index over the interval $[T_{i-1}, T_i]$ under the T_i-forward measure $Q_n^{T_i}$, i.e.

$$y(T_{i-1}, T_i) = \log\left(\frac{I(T_i)}{I(T_{i-1})}\right), \tag{4.40}$$

and henceforth we assume that we explicitly know the characteristic function ϕ_{T_{i-1},T_i} of $y(T_{i-1}, T_i)$,

$$\phi_{T_{i-1},T_i}(u) := \mathbb{E}_n^{T_i}\left[\exp\left(iuy(T_{i-1}, T_i)\right)|\mathcal{F}_t\right]. \tag{4.41}$$

The derivations and explicit formulas of the characteristic function(s) are discussed in Section 4.4.

Pricing of inflation-indexed swaps

The main two inflation-indexed swaps are the ZCIIS and the YYIIS. Recall that the zero-coupon swap is model independent and is simply given by no-arbitrage arguments, i.e. by (4.31). Given the characteristic function $\phi_{T_{i-1},T_i}(u)$ from (4.41) of the log-inflation return under the T_i-forward measure, the pricing of a YYIIS is extremely simple. In fact recall from (4.37) that we have the following expression for the price of a YYIIS:

$$\mathbf{YYIIS}(t, T_{i-1}, T_i, \psi_i, N) = N\psi_i P(t,T_i)\mathbb{E}_n^{T_i}\left\{\frac{I(T_i)}{I(T_{i-1})}|\mathcal{F}_t\right\} - N\psi_i P_n(t,T_i), \tag{4.42}$$

and then note that the expectation in the above expression is nothing more than the characteristic function of the log-return evaluated in the complex-valued point $-i$,

$$\mathbb{E}_n^{T_i}\left\{\frac{I(T_i)}{I(T_{i-1})}|\mathcal{F}_t\right\} = \mathbb{E}_n^{T_i}\left\{\exp\left[i(-i)\log\left(\frac{I(T_i)}{I(T_{i-1})}\right)\right]|\mathcal{F}_t\right\} = \phi_{T_{i-1},T_i}(-i). \tag{4.43}$$

Hence the price of a YYIIS is just given by the following simple expression:

$$\mathbf{YYIIS}(t, T_{i-1}, T_i, \psi_i, N) = N\psi_i P(t, T_i)\phi_{T_{i-1},T_i}(-i) - N\psi_i P_n(t, T_i). \tag{4.44}$$

Pricing of inflation-indexed caplets/floorlets

The pricing of forward starting options like cliquets, attracted recent attention of both practitioners as well as from academia (e.g. see Lucic (2003), Hong (2004) and Brigo and Mercurio (2006)). In this section we will show how one can price inflation call options in the framework of Carr and Madan (1999). Working under the T_i-forward measure, we are in particular interested in the T_i-forward log return $y(T_{i-1}, T_i)$ on the inflation index between the times T_{i-1} and T_i, i.e. as defined in 4.40. From (4.39) we know that we can express an inflation caplet as a call option on the forward return of the index. We can then place this directly in the Carr and Madan (1999) methodology of Section 4.3.1. Provided that $\alpha > 0$ for a caplet and $\alpha > 1$ for a floorlet, we can write the following for the price of an IICplt (inflation-indexed caplet):

$$\begin{aligned}
\mathbf{IICplt}(t, T_{i-1}, T_i, \psi_i, K, N, \omega) &= N\psi_i P_n(t, T_i)\mathbb{E}_n^{T_i}\left\{\left[\omega\left(\frac{I(T_i)}{I(T_{i-1})} - K\right)\right]^+ |\mathcal{F}_t\right\} \\
&= N\psi_i P_n(t, T_i)\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[e^{-(\omega\alpha+iv)\log K}\Psi_{T_{i-1},T_i}(v, \omega, \alpha)\right]dv, \qquad (4.45)
\end{aligned}$$

with $\Psi_{T_{i-1},T_i}(v, \omega, \alpha)$ in (4.26) a function of the characteristic function $\phi_{T_{i-1},T_i}(u)$ of (4.41). Alternatively the price of a floorlet can be expressed in terms of the corresponding caplet price (and vice versa) by means of a put-call parity, e.g. see Mercurio (2005). Given that we know the characteristic function, formula (4.45) provides an efficient and accurate way for determining the prices of inflation-indexed caps/floors. What remains is the derivation of this forward characteristic function, which we will discuss in Section 4.4. The corresponding characteristic functions can be found in Propositions 4.4.2 (for Schöbel and Zhu (1999) volatility) and 4.4.8 (for Heston (1993) volatility).

4.3.3 FX and stock derivatives

The pricing of FX and stock derivatives within the general model (4.8) can be done using similar techniques as in the previous section with inflation-indexed derivatives. The main difference is that inflation-indexed derivatives are usually forward-starting options, whereas the vanilla FX and stock options do not share this feature. In a way, one can therefore treat FX and stock options within the FX setup of our (4.8) as nested (degenerate) cases of inflation derivatives by choosing the forward-starting date equal to the current date and normalising the stock/index price by $I(0)$, i.e. in accordance with (4.40). In a similar spirit, one can see a stock option as a FX option in which the foreign instantaneous interest rate represents the stochastic (or deterministic) continuous dividend rate of the stock.

For clarity we provide the pricing formulas for FX and stock options: working under the T-forward measure, the pricing formulas require the conditional characteristic function

$$\phi_T(u) := \mathbb{E}^{Q^T}\Big[\exp\Big(iuz(T)\Big)\Big|\mathcal{F}_t\Big] \tag{4.46}$$

of the log index/FX-rate/stock price $z(T) := \log I(T)$. Equipped with this characteristic function, the time-T forward FX-rate FFX(T) (i.e. with convexity adjustment when the foreign interest rates are stochastic) is given by

$$\text{FFX}(T) = \mathbb{E}^{Q^T}\left[I(T)\right] = \phi_T(-i). \tag{4.47}$$

Provided with the log-asset price characteristic function, one can immediately price a call/put option on the stock or FX-rate within 'Fourier-inversion' framework of Section 4.3.1. More specifically, one can directly substitute the characteristic function for ϕ_T into the pricing formulas (4.25)-(4.26). Completely analogously to inflation-indexed options, one can price forward-starting (cliquet) options on the forward return of the FX-rate/stock index by substituting the characteristic function $\phi_{T_{i-1},T_i}(u)$ of the forward log return (4.40) into the pricing equations (4.25)-(4.26). We will discuss the derivation of both these characteristic functions in the next section.

4.4 Characteristic function of the model

In this section we will turn to the derivation of the characteristic function of the log inflation return under the nominal T-forward measure Q_n^T. For an inflation index which is driven by a Schöbel-Zhu stochastic volatility process, we are able to derive a closed-form expression, whereas for the Heston stochastic volatility case we are able to approximate this characteristic function. Before turning to these derivations, we first turn to a volatility aspect of the inflation index and to the Gaussian interest rates, which treatment is common for both volatility choices.

Volatility driver and multi-factor Gaussian rates

To ease notation we introduce some matrix notation: let $\Sigma(t,T)$ denote the $1 \times (1+K+M)$ column vector of 'volatilities' driving the Brownian motion of the T-forward inflation index, with corresponding $(1+K+M)\times(1+K+M)$ correlation matrix R, i.e.

$$\Sigma(t,T) = \begin{bmatrix} \nu(t) \\ \sigma_n^1(t)B_n^1(t,T) \\ \vdots \\ \sigma_n^K(t)B_n^K(t,T) \\ -\sigma_r^1(t)B_r^1(t,T) \\ \vdots \\ -\sigma_r^M(t)B_r^M(t,T) \end{bmatrix}, R = \begin{pmatrix} 1 & \rho_{x_n^1,\nu} & \cdots & \rho_{x_r^K,\nu} & \rho_{x_r^1,\nu} & \cdots & \rho_{x_r^M,\nu} \\ \rho_{x_n^1,\nu} & 1 & \cdots & \rho_{x_n^1,x_n^K} & \rho_{x_n^1,x_r^1} & \cdots & \rho_{x_n^1,x_r^M} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \rho_{x_K^1,\nu} & \rho_{x_n^1,x_n^K} & \cdots & 1 & \rho_{x_n^K,x_r^1} & \cdots & \rho_{x_n^K,x_r^M} \\ \rho_{x_r^1,\nu} & \rho_{x_n^1,x_r^1} & \cdots & \rho_{x_n^K,x_r^1} & 1 & \cdots & \rho_{x_r^1,x_r^M} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{\nu,x_r^M} & \rho_{x_n^1,x_r^M} & \ddots & \rho_{x_n^K,x_r^M} & \rho_{x_r^1,x_r^M} & \cdots & 1 \end{pmatrix}, \tag{4.48}$$

Hence we can write the following for the instantaneous variance $\nu_F(t)$ of the inflation index under the T-forward measure:

$$\nu_F^2(t) = \Sigma'(t,T)R\Sigma(t,T), \tag{4.49}$$

with Σ' the transpose of Σ. Another useful expression is the integrated variance of the multi-factor Gaussian rate process; we can write the following for the instantaneous variance $\nu_{K,M}(t)$ of the sum of the rate processes:

$$\nu_{K,M}^2(t,T) = \sum_{i=2}^{K+M+1} \left(\Sigma^{(i)}(t,T)\right)^2 + 2\sum_{i=2}^{K+M+1}\sum_{j=i+1}^{K+M+1} R^{(i,j)}\Sigma^{(i)}(t,T)\Sigma^{(j)}(t,T) \tag{4.50}$$

with $\Sigma^{(i)}$ is the i-th element of the vector $\Sigma(t,T)$ and where $R^{(i,j)}$ denotes the element at row i and column j of the matrix R. Note the shift in index, which is due the presence of the volatility driver $\nu(t)$ in (4.48). For the integrated rate variance $V_{K+M}(t,T)$ one has the following expression

$$V_{K,M}(t,T) := \int_t^T \nu_{K,M}^2(u,T)du = \sum_{i=2}^{K+M+1} C^{(i,i)} + 2\sum_{i=2}^{K+M+1}\sum_{j=i+1}^{K+M+1} R^{(i,j)}C^{(i,j)}, \tag{4.51}$$

where $C^{(i,j)}$ denotes the integrated covariance between the i-th and the the j-th element of the vector of rate volatilities $\Sigma(t)$.

For the covariance $C^{(2,K+M+1)}$ between the first and the $K+M$-th element, one for example has[2]

$$C^{(2,K+M+1)} := -\frac{\sigma_n^1\sigma_M^r}{a_n^1 a_r^M}\left((T-t) + \frac{e^{-a_n^1(T-t)}-1}{a_n^1} + \frac{e^{-a_r^M(T-t)}-1}{a_r^M} - \frac{e^{-(a_n^1+a_r^M)(T-t)}-1}{a_n^1+a_r^M}\right). \tag{4.53}$$

4.4.1 Schöbel-Zhu stochastic volatility

In this section we will determine the characteristic function (under the T-forward measure) of the forward log-inflation return $y(T_{i-1},T_i)$ between times T_{i-1} and T_i. For this we first need to determine the characteristic function of the T-forward log-inflation rate $z(T)$ for a general maturity T. Building forth on the results of Chapter 3, in which the characteristic function for the one-factor Schöbel-Zhu-Hull-White model is derived, we will derive the characteristic function of the multi-factor model in the following subsection.

[2]It is indeed possible to consider time-dependent parameters, in which case the covariance $C^{(2,K+M+1)}$ is given by the following time-dependent integral expression

$$C^{(2,K+M+1)} := \int_t^T \left(\sigma_n^1(u)B_n^1(u,T)\right)\left(-\sigma_M^r(u)B_r^M(u,T)\right)du. \tag{4.52}$$

We can do this for all formulas in this chapter. However as the resulting integral expressions become obscure, whilst the generalization is obvious, we use constant parameters for clarity of exposition.

Characteristic function of the log-index price

We will now determine the characteristic function of the reduced system (4.23), for which we shall use a partial differential equation approach. Recall from (4.22) that $z(t) := \log I_F(t)$ is defined as the T-forward log-asset price; subject to the terminal condition

$$f(T, z, \nu) = \exp\Big(iuz(T)\Big). \tag{4.54}$$

The Feynman-Kac theorem implies that the expected value of $\exp(iuz(T))$ equals the solution of the Kolmogorov backward partial differential equation for the joint probability distribution function $f(t, z, \nu)$, i.e.

$$f := f(t, z, \nu) = \mathbb{E}^{Q^T}\Big[\exp(iuz(T))\Big|\mathcal{F}_t\Big]. \tag{4.55}$$

Thus the solution for f equals the characteristic function of the forward asset price dynamics (starting from z at time t). To obtain the Kolmogorov backward partial differential equation for the joint probability distribution function $f = f(t, y, \nu)$, we need to take into account the covariance term $dz(t)d\nu(t)$ which equals

$$\begin{aligned} dz(t)d\nu(t) &= \Big(\nu(t) + \Sigma_n(t,T) - \Sigma_r(t,T)\Big)dW^*(t)\Big(\tau dW_\nu^T(t)\Big) \\ &= \Big(\rho_{I\nu}\tau\nu(t) + \tau R_{\nu,X_n}\Sigma_n(t,T) - \tau\rho_{\nu,X_r}\Sigma_r(t,T)\Big)dt. \end{aligned} \tag{4.56}$$

The model we are considering is not an affine model in $z(t)$ and $\nu(t)$, but it is if we enlarge the state space to include $\nu^2(t)$:

$$dz(t) = -\frac{1}{2}\nu_F^2(t)dt + \nu_F(t)dW^*(t) \tag{4.57}$$

$$d\nu(t) = \kappa\Big[\xi(t) - \nu(t)\Big]dt + \tau dW_\nu^T(t) \tag{4.58}$$

$$d\nu^2(t) = 2\nu(t)d\nu(t) + \tau^2 dt = 2\kappa\Big(\frac{\tau^2}{2\kappa} + \xi(t)\nu(t) - \nu^2(t)\Big)dt + 2\tau\nu(t)dW_\nu(t). \tag{4.59}$$

Using (4.57) and (4.56), we obtain the following partial differential equation for $f(t, z, \nu)$:

$$\begin{aligned} 0 &= f_t - \frac{1}{2}\nu_F^2(t)f_z + \kappa(\xi(t) - \nu(t))f_\nu + \frac{1}{2}\nu_F^2(t)f_{zz} \\ &\quad + \Big(\rho_{I\nu}\tau\nu(t) + \tau R_{\nu,X_n}\Sigma_n(t,T) - \tau\rho_{\nu,X_r}\Sigma_r(t,T)\Big)f_{z\nu} + \frac{1}{2}\tau^2 f_{\nu\nu}. \end{aligned} \tag{4.60}$$

Solving this partial differential equation, subject to the terminal boundary condition (4.54), provides us with the characteristic function of the forward asset price dynamics (starting from z at time t). Due to the affine structure of our model, we come to the following proposition.

Proposition 4.4.1 *The characteristic function of the domestic T-forward log inflation-rate of the model with Schöbel and Zhu (1999) stochastic volatility is given by the following closed-form*

solution:

$$f(t,z,v) = \exp\Big[A(u,t,T) + B(u,t,T)z(t) + C(u,t,T)v(t) + \frac{1}{2}D(u,t,T)v^2(t)\Big], \tag{4.61}$$

where:

$$\begin{aligned}
A(u,t,T) &= -\frac{1}{2}u(i+u)V_{K,M}(t,T) \\
&\quad + \int_t^T \Bigg\{\Big[\kappa\psi + (iu-1)\sum_{i=1}^{K}\rho_{x_n^i,v}\tau\sigma_n^i B_n^i(t,T) - iu\sum_{j=1}^{M}\rho_{x_r^j,v}\tau\sigma_r^j B_r^j(t,T)\Big]C(u,s,T) \\
&\qquad + \frac{1}{2}\tau^2\Big(C^2(u,s,T) + D(u,s,T)\Big)\Bigg\}ds, \qquad (4.62)\\
B(u,t,T) &= B := iu, \qquad (4.63)\\
C(u,t,T) &= -\frac{u(i+u)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}\Bigg\{\gamma_0\Big(1+e^{-2\gamma(T-t)}\Big) \qquad (4.64)\\
&\quad + \sum_{i=1}^{K}\Big[\Big(\gamma_3^i - \gamma_4^i e^{-2\gamma(T-t)}\Big) - \Big(\gamma_5^i e^{-a_n^i(T-t)} - \gamma_6^i e^{-(2\gamma+a_n^i)(T-t)}\Big) - \gamma_7^i e^{-\gamma(T-t)}\Big] \\
&\quad - \sum_{j=1}^{M}\Big[\Big(\gamma_8^j - \gamma_9^j e^{-2\gamma(T-t)}\Big) - \Big(\gamma_{10}^j e^{-a_r^j(T-t)} - \gamma_{11}^j e^{-(2\gamma+a_r^j)(T-t)}\Big) - \gamma_{12}^j e^{-\gamma(T-t)}\Big]\Bigg\} \\
D(u,t,T) &= -u(i+u)\frac{1-e^{-2\gamma(T-t)}}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \qquad (4.65)
\end{aligned}$$

with $V_{K,M}(t,T)$, as defined in (4.51), the integrated variance of the Gaussian rate processes. The constants are:

$$\begin{aligned}
\gamma &= \sqrt{(\kappa - \rho_{I,v}\tau B)^2 - \tau^2(B^2 - B)}, & \gamma_0 &= \frac{\kappa\psi}{\gamma} \qquad (4.66)\\
\gamma_1 &= \gamma + (\kappa - \rho_{I,v}\tau B), & \gamma_2 &= \gamma - (\kappa - \rho_{I,v}\tau B),\\
\gamma_3^i &= \frac{\rho_{I,x_n^i}\sigma_n^i\gamma_1 + \rho_{x_n^i,v}\sigma_n^i\tau(iu-1)}{a_n^i\gamma}, & \gamma_4^i &= \frac{\rho_{I,x_n^i}\sigma_n^i\gamma_2 - \rho_{x_n^i,v}\sigma_n^i\tau(iu-1)}{a_n^i\gamma},\\
\gamma_5^i &= \frac{\rho_{I,x_n^i}\sigma_n^i\gamma_1 + \rho_{x_n^i,v}\sigma_n^i\tau(iu-1)}{a_n^i(\gamma - a_n^i)}, & \gamma_6^i &= \frac{\rho_{I,x_n^i}\sigma_n^i\gamma_2 - \rho_{x_n^i,v}\sigma_n^i\tau(iu-1)}{a_n^i(\gamma + a_n^i)},\\
\gamma_8^j &= \frac{\rho_{I,x_r^j}\sigma_r^j\gamma_1 + \rho_{x_r^j,v}\sigma_r^j\tau B}{a_r^j\gamma}, & \gamma_9^j &= \frac{\rho_{I,x_r^j}\sigma_r^j\gamma_2 - \rho_{x_r^j,v}\sigma_r^j\tau B}{a_r^j\gamma},\\
\gamma_{10}^j &= \frac{\rho_{I,x_r^j}\sigma_r^j\gamma_1 + \rho_{x_r^j,v}\sigma_r^j\tau B}{a_r^j(\gamma - a_r^j)}, & \gamma_{11} &= \frac{\rho_{I,x_r^j}\sigma_r^j\gamma_2 - \rho_{x_r^j,v}\sigma_r^j\tau B}{a_r^j(\gamma + a_r^j)},\\
\gamma_7^i &= (\gamma_3^i - \gamma_4^i) - (\gamma_5^i - \gamma_6^i), & \gamma_{12}^j &= (\gamma_8^j - \gamma_9^j) - (\gamma_{10}^j - \gamma_{11}^j)
\end{aligned}$$

Proof See appendix 4.7.1.

In the following section we are able to derive the forward starting characteristic of the log-inflation index return, using the above characteristic function of the log-inflation index under the T-forward measure.

Characteristic function of the log index return

Recently the pricing of forward starting options attracted the attention of both practitioners as well as from academia e.g. see Lucic (2003), Hong (2004), Chapter 3 and in an inflation context Mercurio and Moreni (2006a) and Kruse (2007). In this section we will consider the pricing of forward starting options like inflation caplets within the general model setup combined with Schöbel-Zhu volatility. In particular, using the framework of Carr and Madan (1999), as described in Section 4.3.1, it suffices to know the characteristic function of the following log-inflation index return under the T_i-forward measure:

$$y(T_{i-1}, T_i) := \log\Big(\frac{I(T_i)}{I(T_{i-1})}\Big) = \log I(T_i) - \log I(T_{i-1}). \tag{4.67}$$

Since $I(t) := I_F(t)\frac{P_n(t,T_i)}{P_r(t,T_i)}$, we can also express this return in terms of the T_i-forward log inflation rate $z(t) := \log\Big(I_F(t)\Big)$, i.e.

$$y(T_{i-1}, T_i) = z(T_i) - z(T_{i-1}) - \log P_n(T_{i-1}, T_i) + \log P_r(T_{i-1}, T_i). \tag{4.68}$$

We are interested in the characteristic function $\phi_{T_{i-1},T_i}(u)$ of the log-inflation index return $y(T_{i-1}, T_i)$ under the T_i forward measure, i.e.

$$\phi_{T_{i-1},T_i}(u) := \mathbb{E}^{Q^T}\Big[\exp\Big(iu(y(T_{i-1}, T_i))\Big)\Big|\mathcal{F}_t\Big]. \tag{4.69}$$

First define,

$$\Lambda := \exp\Big(iu\Big[z(T_i) - z(T_{i-1}) - \log P_n(T_{i-1}, T_i) + \log P_r(T_{i-1}, T_i)\Big]\Big) \tag{4.70}$$

hence by using the tower law for conditional expectations and the (conditional) characteristic function of our model (4.61), we obtain the following expression for the characteristic function

of the (forward) log-return:

$$\begin{aligned}
\phi_{T_{i-1},T_i}(u) &= \mathbb{E}_n^{T_i}\Big\{\Lambda\Big|\mathcal{F}_t\Big\} = \mathbb{E}_n^{T_i}\Big\{\mathbb{E}_n^{T_i}\Big[\Lambda\Big|\mathcal{F}_{T_{i-1}}\Big]\Big|\mathcal{F}_t\Big\} \\
&= \mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[-z(T_{i-1}) - \log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i)\Big]\Big) \\
&\qquad \cdot\mathbb{E}_n^{T_i}\Big[\exp[iuz(T_i)]\Big|\mathcal{F}_{T_{i-1}}\Big]\Big|\mathcal{F}_t\Big\} \\
&= \exp\Big(iu\Big[A_r(T_{i-1},T_i) - A_n(T_{i-1},T_i)\Big] + A(u,T_{i-1},T_i)\Big) \qquad (4.71) \\
&\quad \times\mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[B_n(T_{i-1},T_i)X_n(T_{i-1}) - B_r(T_{i-1},T_i)X_r(T_{i-1})\Big] \\
&\qquad +C(u,T_{i-1},T_i)\nu(T_{i-1}) + \frac{1}{2}D(u,T_{i-1},T_i)\nu^2(T_{i-1})\Big)\Big|\mathcal{F}_t\Big\}
\end{aligned}$$

Though the latter expectation depends only on the (correlated) Gaussian variates $x_n^k(T_{i-1})$, $x_r^j(T_{i-1})$,$\nu(T_{i-1})$, we also have that the integrated volatility process $\int_t^{T_{i-1}}\nu(u)du$ arises in the real rate processes $x_r^j(T_{i-1})$, e.g. see Proposition 4.2.2. To this end, we decompose $x_r^j(T_{i-1})$ into

$$\begin{aligned}
x_r^j(T_{i-1}) &= V^j(T_{i-1}) + \widetilde{x_r^j}(T_{i-1}) \qquad (4.72) \\
V^j(T_{i-1}) &:= \rho_{I,x_r^j}\sigma_r^j\int_t^{T_{i-1}} e^{-a_r^j(T_{i-1}-u)}\nu(u)du \\
&\sim N\Big(\mu_V^j(t,T_{i-1}),\sigma_V^j(t,T_{i-1})\Big) \qquad (4.73) \\
\widetilde{x_r^j}(T_{i-1}) &= \mu_r^j(t,T_{i-1}) + \sigma_r^j\int_t^{T_{i-1}} e^{-a_r^j(T_{i-1}-u)}dW_{r_j}^{T_i}(u), \\
&\sim N\Big(\mu_r^j(t,T_{i-1}),\sigma_r^j(t,T_{i-1})\Big), \qquad (4.74)
\end{aligned}$$

where $\mu_r^j(t,T_{i-1})$, $\sigma_r^j(t,T_{i-1})$, $\mu_V^j(t,T_{i-1})$ and $\sigma_V^j(t,T_{i-1})$ as defined in (4.136), (4.137), (4.145) and (4.146), see appendix 4.7.2.

Hence we find that the characteristic function (4.71) is of the following Gaussian-quadratic form:

$$\begin{aligned}
&\exp\Big(iu\Big[A_r(T_{i-1},T_i) - A_n(T_{i-1},T_i)\Big] + A(u,T_{i-1},T_i)\Big) \\
&\times\mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[B_n(T_{i-1},T_i)X_n(T_{i-1}) - B_r(T_{i-1},T_i)\Big(\bar{V}(T_{i-1}) + \widetilde{X}_r(T_{i-1})\Big)\Big] \\
&\qquad +C(u,T_{i-1},T_i)\nu(T_{i-1}) + \frac{1}{2}D(u,T_{i-1},T_i)\nu^2(T_{i-1})\Big)\Big|\mathcal{F}_t\Big\} \\
=:\ &\mathbb{E}_n^{T_i}\Big\{\exp\Big[a_0 + a'Z + Z'BZ\Big]\Big\}, \qquad (4.75)
\end{aligned}$$

with a_0 a constant, $a^{'}$ a row-vector, B a matrix and where Z follows a multivariate standard normal distribution with correlation matrix S. Thus the random vector Z consists of the $1 + K + 2M$ driving elements $\nu, x_n^1, \ldots, x_n^K, x_r^1, \ldots, x_r^M, V^1, \ldots, V^M$. Note that since we are only dealing with one quadratic term (i.e. $\nu^2(T_{i-1})$), we can reduce the quadratic form (4.75) of the random vector X to

$$\mathbb{E}_n^{T_i}\Big\{\exp\Big[a_0 + a^{'}Z + b_0 Z^{(1)^2}\Big]\Big\}, \tag{4.76}$$

where the constants a_0, b_0, the column-vector a and the correlation matrix S of the standard Gaussian vector Z, can be easily be deduced from (4.75) and are explicitly defined in appendix 4.7.2.

Using standard theory on Gaussian-quadratic forms (e.g. see Glasserman (2003) or Feuerverger and Wong (2000)) we can now easily find an explicit solution for (4.75). Recalling that (4.75) is equivalent to the characteristic function (4.71) of the forward return on the log inflation index, we come to the following proposition.

Proposition 4.4.2 *Let C be a matrix (with typical element $c_{i,j}$) satisfying $C^{'}C = S$ (e.g. by a Cholesky decomposition), define*

$$p_j := \sum_{i=1}^{1+K+2M} c_{i,j} a^{(i)}, \tag{4.77}$$

$$q_1 := \sum_{i=1}^{1+K+2M} c_{i,1}^2 b_0, \tag{4.78}$$

with correlation matrix S, column-vector a and constant b_0 as defined in Appendix 4.7.2. The characteristic function of the forward log return $y(T_{i-1}, T_i)$ (4.67) under the T_i-forward measure is given by the following closed-form solution:

$$\phi_{T_{i-1},T_i}(u) = \frac{\exp\Big(a_0 + \frac{p_1^2}{2q_1(1-2q_1)} - \frac{p_1^2}{4q_1} + \sum_{j=2}^{1+K+2M} \frac{p_j^2}{2}\Big)}{\sqrt{1-2q_1}}. \tag{4.79}$$

Proof Since (4.75) is equivalent to (4.71), the characteristic function of the forward return on the log inflation index is given by an explicit solution of the Gaussian-quadratic form (4.75), which is given by standard theory on quadratic forms, e.g. see Glasserman (2003) or Feuerverger and Wong (2000). □

Equipped with the characteristic function of the log-inflation index return, the prices of year-on-year inflation-indexed swaps and inflation-indexed caps/floors are directly obtained by the formulas (4.44) and (4.45).

4.4.2 Heston stochastic volatility

The characteristic function-based pricing method in our model with Heston (1993) stochastic volatility will turn out to be somewhat more involved than under Schöbel and Zhu (1999) stochastic volatility. In fact for the general model Heston (1993) stochastic volatility we need to resort to approximate methods for the pricing of inflation-indexed options.
Recall from (4.17) and (4.20) that the general model dynamics with Heston (1993) volatility under the T-forward measure Q_n^T are given by

$$\frac{dI_F(t)}{I_F(t)} = \Big(\nu(t) + \Sigma_n(t,T) - \Sigma_r(t,T)\Big)dW_*^T(t), \tag{4.80}$$

$$d\nu^2(t) = \kappa\Big[\zeta(t) - \nu^2(t)\Big]dt + \xi\nu(t)dW_\nu^T(t). \tag{4.81}$$

To derive the characteristic function of the log-inflation rate, one can in principle then pursue the same steps as in the model with Schöbel and Zhu (1999) volatility, that is solving the Kolmogorov backward equation for the log-inflation rate with a certain boundary condition. However, due to the square-root volatility process, the Heston partial differential equation in combination with correlated Gaussian rates is unfortunately not affine any more. Hence, contrary to the previous model, there is (as far as we know) no exact closed-form expression for the characteristic function for this model. Nevertheless, in case we make the simplifying assumption that the rate processes are perpendicular to the stochastic volatility and the asset price processes, one can easily find an closed-form solution for its characteristic function. For the general case, we consider two alternative pricing methods

1. A projection of the characteristic function in the general model onto the uncorrelated case.

2. A control variate based Monte Carlo pricing technique that uses an exact result from the uncorrelated model.

The setup of the following section is therefore as follows: we first discuss the pricing for the log-inflation rate and the log-inflation index return in the model with uncorrelated Heston (1993) stochastic volatility. Then we show a projection technique of the general case onto the uncorrelated model. Finally, though the projection already works quite well, we also discuss the use of the approximate model as control variate in a Monte Carlo pricing procedure of the exact model.

Characteristic function of the log-index price: uncorrelated case

For the derivation of the characteristic function of the uncorrelated model (i.e. with rate processes perpendicular to the variance and asset price process), we will use two propositions.
First of all, let $z(t) = \log \frac{I(t)P_r(t,T)}{P_n(t,T)}$ denote the T-forward log-asset price, with dynamics

$$dz(t) = -\frac{1}{2}\nu^2(t) + \nu(t)dW_I^T(t), \tag{4.82}$$

$$d\nu^2(t) = \kappa\Big[\theta - \nu^2(t)\Big]dt + \xi\nu(t)dW_\nu^T(t), \tag{4.83}$$

i.e. with stochastic interest rate dynamics. One then has following proposition regarding the characteristic function of $z(t)$.

Proposition 4.4.3 *starting from time t, the characteristic function $\phi_{HE}(u)$ of the T-forward log-asset price $z(T)$ of the classical Heston (1993) model is given by*

$$\phi_{HE}(u) := \exp\Big[iuz(t) + A_{HE}(u,t,T) + B_{HE}(u,t,T)v^2(t)\Big] \tag{4.84}$$

where:

$$A_{HE}(u,t,T) := \theta\kappa\xi^{-2}\Big((\kappa - \rho\xi iu - d)T - 2\log(\frac{1-g_2e^{-dT}}{1-g_2})\Big), \tag{4.85}$$

$$B_{HE}(u,t,T) := \xi^{-2}(\kappa - \rho\xi iu - d)\frac{1-e^{-dT}}{1-g_2e^{-dT}} \tag{4.86}$$

and with:

$$d := \sqrt{(\rho\xi iu - \kappa)^2 + \xi^2(iu + u^2)}, \tag{4.87}$$

$$g_2 := \frac{\kappa - \rho\xi iu - d}{\kappa - \rho\xi iu + d}. \tag{4.88}$$

Proof For the proof we refer to Heston (1993) or Gatheral (2005). □

Note that in the literature one can find two (mathematically) equivalent formulations for the Heston characteristic function: the one presented above can for example be found in Albrecher et al. (2005) or Gatheral (2005) and is free of a numerical difficulty called branch cutting, while another representation can be found in the original Heston paper Heston (1993) or Kahl and Jäckel (2005), which may cause some numerical difficulties for certain model parameters, see Albrecher et al. (2005).

The second proposition concerns the interest rates part of the inflation dynamics. To this end, define

$$R_{K,M}(t,T) := -\frac{1}{2}V_{K,M}(t,T) + \int_t^T \Big[\Sigma_n(u,T)dW_n^T(u) - \Sigma_r(u,T)dW_r^T(u)\Big]du, \tag{4.89}$$

we then come to the following proposition of the characteristic function of $R_{K,M}(t,T)$.

Proposition 4.4.4 *The characteristic function of $\phi_R(u)$ of $R_{K,M}(t,T)$ (4.89) is given by*

$$\phi_{K,M}(u) := \exp\Big[-\frac{1}{2}u(i+u)V_{K,M}(t,T)\Big]. \tag{4.90}$$

Proof As $\int_t^T \Sigma_i(u,T)du$, $i \in n, r$ follows a Gaussian distribution with mean 0, $R_{K,M}(t,T)$ as sum of Gaussian variates is also Gaussian with mean $-\frac{1}{2}V_{K,M}(t,T)$. From Fubini and Itô's isometry,

it follows that $R_{K,M}(t,T)$ is normally distributed with mean $-\frac{1}{2}V_{K,M}(t,T)$ and variance $V_{K,M}(t,T)$ as explicitly given by (4.51). Moreover, the characteristic function $\phi_{K,M}(u)$ of $R_{K,M}(t,T)$ follows directly as consequence of this normality. □

With the results from Propositions 4.4.3 and 4.4.4, we can now easily determine the characteristic function of the log-inflation index in the uncorrelated model, which results in the following proposition.

Proposition 4.4.5 *The characteristic function $\phi_F(u)$ for the log-inflation index $\log I_F(t)$ of the uncorrelated Heston dynamics (4.80) is given by the following closed-form expression:*

$$\begin{aligned}\phi_F(u) &= \phi_{HE}(u)\cdot\phi_{K,M}(u) \\ &= \exp\Big[iuz(t) + A_{HE}(u,t,T) + B_{HE}(u,t,T)v^2(t) - \frac{1}{2}u\Big(i+u\Big)V_{K,M}(t,T)\Big]\end{aligned} \tag{4.91}$$

Proof Since the Brownian motions driving the Heston dynamics $z(t)$, i.e. $W_I^T(t)$ and $W_v^T(t)$, are uncorrelated with the Brownian motions that drive the rate process $R_{K,M}(t,T)$, i.e. $W_{n_i}^T(u)$ and $W_{r_j}^T(u)$, we have that we can write for the log-inflation index dynamics $\log I_F(t)$ of the dynamics of (4.23) (or equivalently of (4.80)) as the sum of the above two processes, i.e.

$$\log I_F(t) = z(t) + R_{K,M}(t,T).$$

Since the driving Brownian motions are uncorrelated, we then have that $z(t)$ is independent of $R_{K,M}(t,T)$ and furthermore that the characteristic function $\phi_F(u)$ of $\log I_F(t)$ is given by the product of the characteristic functions of $z(t)$ and $R_{K,M}(t,T)$. □

Characteristic function of the log index return: uncorrelated case

We will now derive the (forward-starting) characteristic function of the log-inflation index return. Just as in our model from Section 4.4.1, we follow Hong (2004) and Chapter 3; that is, we consider the characteristic function $\phi_{T_{i-1},T_i}(u)$ of the log-inflation index return

$$y(T_{i-1},T_i) := \log\Big(\frac{I(T_i)}{I(T_{i-1})}\Big) = z(T_i) - z(T_{i-1}) - \log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i). \tag{4.92}$$

In particular we want to solve the characteristic function $\phi_{T_{i-1},T_i}(u)$ of $y(T_{i-1},T_i)$ under the T_i-forward measure; using similar arguments (e.g. the tower law for conditional expectations) as in (4.71), we can obtain the following expression of the forward-starting characteristic function in

our (uncorrelated) model:

$$\begin{aligned}\phi_{T_{i-1},T_i}(u) &= \mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[-z(T_{i-1}) - \log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i)\Big]\Big)\\ &\qquad \cdot\mathbb{E}_n^{T_i}\Big[\exp[iuz(T_i)]\Big|\mathcal{F}_{T_{i-1}}\Big]\Big|\mathcal{F}_t\Big\}\\ &= \exp\Big(A_{HE}(u,T_{i-1},T_i) - \frac{1}{2}u\big(i+u\big)V_{K,M}(T_{i-1},T_i)\Big)\\ &\qquad \cdot\mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[-\log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i)\Big]\\ &\qquad\qquad + B_{HE}(u,T_{i-1},T_i)v^2(T_{i-1})\Big)\Big|\mathcal{F}_t\Big\}. \end{aligned} \tag{4.93}$$

Hence since the rate processes $x_n^k(T_{i-1})$ and $x_r^j(T_{i-1})$ are independent of the variance process $v^2(T_{i-1})$, we have

$$\begin{aligned}\phi_{T_{i-1},T_i}(u) &= \exp\Big(A_{HE}(u,T_{i-1},T_i) - \frac{1}{2}u\big(i+u\big)V_{K,M}(T_{i-1},T_i)\Big)\\ &\qquad \cdot\mathbb{E}_n^{T_i}\Big\{\exp\Big(iu\Big[-\log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i)\Big]\Big|\mathcal{F}_t\Big\}\\ &\qquad \cdot\mathbb{E}_n^{T_i}\Big\{\exp\Big[B_{HE}(u,T_{i-1},T_i)v^2(T_{i-1})\Big]\Big|\mathcal{F}_t\Big\}. \end{aligned} \tag{4.94}$$

Hence its remains to evaluate the expectations in the latter expression; since the first expectation can be seen as the characteristic function of the log-bond prices, we have the following proposition.

Proposition 4.4.6 *The characteristic function $\phi_{K,M}(u)$ of the log bond prices in (4.94) under the T_i-forward measure is given by*

$$\phi_{K,M}(u) = \exp\Big[iuh_0 - \frac{u^2}{2}h' S_R h\Big], \tag{4.95}$$

with the constant h_0, column vector h and correlation matrix S_R respectively as defined in (4.159), (4.160) and (4.161).

Proof Note that one can write

$$-\log P_n(T_{i-1},T_i) + \log P_r(T_{i-1},T_i) =: h_0 + h' Z_R, \tag{4.96}$$

with Z_R the random Gaussian vector consisting of the normalized stochastic parts of the Gaussian factors $x_n^1,\ldots,x_n^K, x_r^1,\ldots,x_r^M$. Therefore (4.96) is nothing more than the characteristic function of the normal distribution $h_0 + h' Z_R$, which is given by expression (4.95).
Alternatively, one can see this expectation as a special case of the Gaussian-quadratic form (4.75) of the model in Proposition 4.4.2, i.e. without the volatility components $v(t)$ and $V^j(t)$). □

For the calculation of the second expectation of (4.94) we will use the following property of the square root process $\nu^2(T_{i-1})$.

Proposition 4.4.7 *The characteristic function $\phi_{\nu^2}(y)$ of $\nu^2(T_{i-1})$ is given by*

$$\phi_{\nu^2}(y) = I\!E\Big[\exp(iy\nu^2(T_{i-1}))\Big] = \frac{\exp(\frac{ciy\lambda}{1-2ciy})}{(1-2ciy)^{\frac{2\kappa\theta}{\xi^2}}}, \tag{4.97}$$

where

$$c := \frac{\xi^2(1-e^{-\kappa(T_{i-1}-t)})}{4\kappa}, \tag{4.98}$$

$$\lambda := \frac{4\kappa e^{-\kappa(T_{i-1}-t)}v(s)}{\xi^2(1-e^{-\kappa(T_{i-1}-t)})}. \tag{4.99}$$

Proof The proposition follows directly from the fact that variance process $\nu^2(T_{i-1})$ is distributed as a constant c times a non-central chi-squared distribution with $\frac{4\kappa\theta}{\xi^2}$ degrees of freedom and non-centrality parameter λ, e.g. see Cox et al. (1985). □

Hence we come to the following proposition for the characteristic function $\phi_{T_{i-1},T_i}(u)$ as in expression (4.94).

Proposition 4.4.8 *The forward-starting characteristic function $\phi_{T_{i-1},T_i}(u)$ of the model (4.8) with uncorrelated Heston (1993) stochastic volatility is given by the following closed-form expression:*

$$\phi_{T_{i-1},T_i}(u) = \exp\Big(A_{HE}(u,T_{i-1},T_i) - \frac{1}{2}u\big(i+u\big)V_{K,M}(T_{i-1},T_i)\Big)\phi_{K,M}(u)\phi_{\nu^2}\Big(-iB_{HE}(u,T_{i-1},T_i)\Big) \tag{4.100}$$

with $A_{HE}(u,t,T_{i-1})$ and $B_{HE}(u,t,T_{i-1})$ as defined in equations (4.85) and (4.86), and with $\phi_{K,M}(u)$ and $\phi_{\nu^2}(u)$ as in proposition (4.4.6) and 4.4.7.

Proof The characteristic function (4.100) of the forward log-inflation index return follows directly by evaluating the two expectations of (4.94). The first expectation of (4.94) equals the characteristic-generating function $\phi_{K,M}(u)$ of the log bond prices (4.96). The second expectation equals the moment-generating function ϕ_{ν^2} of the shifted non-central chi-squared distributed random variable $\nu^2(T_{i-1})$, evaluated in the point $B(u,t,T_{i-1})$. □

Projection of the general case onto the uncorrelated model

Since in the general Heston model setup (i.e. with a full correlation structure) the affine structure is destroyed, it is challenging to find the characteristic function of the log-inflation index; we are not aware of an exact closed-form expression for characteristic function in the Heston model with correlated Gaussian rates. Nevertheless one can try to approximate the general dynamics by a simpler process for which a closed-form pricing expression does exists. Where a heuristic approach based on moment-matching techniques was suggested by van Haastrecht (2007), a

more rigorous projection method was recently suggested by Antonov et al. (2008), which uses a Markovian projection technique of the general model onto the (affine) uncorrelated model. After the projected parameters are determined, one can just use the uncorrelated model and corresponding pricing formulas to price stock, foreign exchange and inflation derivatives. Though the Markovian projection technique is fast and works well for mild parameter settings and short maturities (i.e. when the 'distance' between the models is relatively small), the projection is rather involved and deteriorates for longer maturities and more extreme model parameters (i.e. when the 'distance' is relatively large), in particular for a large index-rate correlation in combination with a high volatility of the rates. For details on the Markovian projection and numerical results of the approximation, we refer the reader to Antonov et al. (2008).

Monte Carlo pricing method for the general model

Instead of approximating the prices of vanilla options in the general Heston setup, e.g. by a projection technique as touched upon in subsection 4.4.2, one can also entail a Monte Carlo procedure to price these options. Where the approximation formulas can be rather biased for certain model settings (e.g. see the discussion in subsection 4.4.2), a Monte Carlo estimate has the advantage that it converges to the true option price in the limit for the number of sample paths. Moreover a Monte Carlo procedure is generic (i.e. is suitable for a whole range of exotic options) and is straightforward to implement (if not already implemented for exotic option pricing). The main practical disadvantage of a Monte Carlo calibration procedure, is the speed with which vanilla option prices can be calculated within some error measure. Since one repeatedly needs to update an error function of the 'distance' between model and market vanilla prices, the speed to calculate these model option prices is rather important. Even though one can price multiple options (e.g. on different times) with one Monte Carlo run, the use of closed-form option pricing formulas is often much faster. Nevertheless, with the use modern-day variance reduction techniques and the ever-growing computational power (in particular the fact that Monte Carlo procedure can be easily parallelized over multiple processers), we expect Monte Carlo techniques to become even more popular in the near future.
In this section we present an very effective control variate estimator for the pricing of vanilla options the general Heston dynamics. To demonstrate its efficiency, we take the pricing of a vanilla call option as example. To benchmark the numerical results against the Markovian projection, we consider the same hybrid equity-interest rate (stock) example as in Antonov et al. (2008). The setup of this section is as follows: we first discuss the control variate technique for the general model, after which we demonstrate which variance reductions can be expected and discuss its numerical efficiency.

Uncorrelated price as control variate estimator

As discussed in Section 4.4.2, Monte Carlo pricing procedures might be easy to implement and quite generic, but often lack of speed and are hence sometimes being considered as 'brute-force'. Nowadays, however, a whole variety of variance reductions techniques are available to boost the computational efficiency of the Monte Carlo run, e.g. see Glasserman (2003) or Jäckel (2002)

for an overview of such methods. One of these variance reductions techniques is the control variate estimator. The key idea behind this technique is that we can use the error in estimating a similar quantity (from which we know the theoretical value) as a control to correct for the Monte Carlo error for the unknown quantity, see Glasserman (2003). The effectiveness of such a control variate depends explicitly on the correlation between the control and the to be estimated price. Thus if the control contains many information of the estimated price, it can correct quite a lot of Monte Carlo noise in the resulting estimator (and vice versa). Mathematically, it can be shown that, if the correlation between control and the standard Monte Carlo estimator are correlated with correlation coefficient ρ in combination with an optimal control parameter, one obtains (on average) a variance reduction of

$$\mathbf{VR}(\rho) = \frac{1}{1-\rho^2}, \tag{4.101}$$

which is enormous for ρ close to one, e.g. see Glasserman (2003).

Before turning to the control variate estimator, we first introduce some notation. Let $\overline{C^0}$,$\overline{C^\rho}$ and C_i^0, C_i^ρ respectively denote the expected (European) call option price and the simulated call option prices for the general (superscript ρ) and the uncorrelated (superscript 0) dynamics. Since we know the call option price C^0 of the uncorrelated price in closed-form by inverting (4.100), and usually this price is largely correlated with the call option price C^ρ of the general model, we propose to use C^0 as a control for C^ρ. Since the prices are highly correlated, we expect to see large variance reductions of the control variate estimator $\widetilde{C^\rho}(b)$ over the ordinary estimator C^ρ, i.e. from formula (4.101). The resulting control variate estimator $\widetilde{C^\rho}(b)$ is given by

$$\widetilde{C^\rho}(b) = \frac{1}{n}\sum_{i=1}^{n}(C_i^\rho - b(C_i^0 - \mathbb{E}[C^0])), \tag{4.102}$$

where b is a real coefficient. The optimal coefficient b^* that minimizes the variance of (4.102) can easily by calculated and is explicitly given by

$$b^* = \frac{\sigma_{C^\rho}}{\sigma_{C^0}}\rho_{C^0,C^\rho} = \frac{\mathrm{Cov}[C^0, C^\rho]}{\mathrm{Var}[C^0]}. \tag{4.103}$$

Note that one often also needs to estimate b^* from the simulations and this might induce some bias in the effectiveness (4.101) of the control variate. However, as discussed in Glasserman (2003), this bias is often very small; in case ρ_{C^0,C^ρ} is close to one and $\sigma_{C^\rho} \approx \sigma_C$ (which more than often is the case), it might even be a more efficient to just set b^* equal to one (since one does not have to estimate b^*, see Glasserman (2003). In Section 4.5.1 the quality of the control variate estimator is investigated.

4.5 Applications and Numerical Results

In this section, we look at two applications of the model; first, for an equity example and with Heston (1993) stochastic volatility, we test the quality of the control variate estimator $\widetilde{C^{\rho}}$ of (4.102), compare it to the Markovian projection technique of Antonov et al. (2008) and discuss its practical applicability in a Monte Carlo calibration and/or pricing procedure. Secondly, we consider two applications (one with Schöbel and Zhu (1999) and one with Heston (1993) stochastic volatility) in which we calibrate our model to FX (option) market data. The example explicitly takes into account the pronounced long-term FX implied volatility skew/smile that is present in the markets. Finally the results are compared and analyzed.

4.5.1 Quality of the control variate estimator

To test the numerical quality of the control variate estimator $\widetilde{C^{\rho}}$ of (4.102), we turn to the pricing of (European) call options under the general hybrid Heston dynamics. To this end we consider two different parameter settings, listed in Table **1** below.

Example	κ	ξ	ρ_{I,v^2}	$v(0)$	θ	y_r	y_q	a_n	σ_n	ρ_{I,x_n^1}	ρ_{v^2,x_n^1}
case I	2.0	1.0	-0.3	0.09	0.09	0.04	0.0	0.03	0.007	$(*)$	0.0
case II	0.25	0.625	-0.4	0.0625	0.0625	0.05	0.02	0.05	0.01	0.30	0.15

Table 1: Numerical test cases for the Control Variate estimator (4.102). y_r denotes the continuous (constant) interest rate yield, y_q the continuous (constant) dividend yield, the $(*)$ indicates that we vary this parameter during the experiments and in all cases $I(0) = 100$.

Both test cases roughly correspond to parameter settings that are likely to be encountered in medium to long maturity equity markets. The first test case is prevalent in the existing literature: similar Heston parameter settings, in a pure equity context, are considered in Broadie and Kaya (2006), Lord et al. (2008) and Andersen (2008). The second test case is taken from Antonov et al. (2008) wherein it serves to test their Markovian projection approximation, i.e. as touched upon in Section 4.4.2. Using these test cases, we first look at the quality of the control as function of the equity rates correlation coefficient and secondly we investigate the efficiency the control variate estimator (4.102) as function of the option maturity and compare it with the Markovian Projection technique of Antonov et al. (2008).

Results for case I

Though the uncorrelated price is often highly correlated with the price of the general model, the efficiency is dependent on the specific model parameters. For example notice that for $\rho_{I,x_1^n} = \rho_{v^2,x_1^n} = 0\%$, the control variate estimator is exact, because in that case the uncorrelated price equals the required estimate. Though the effectiveness depends on both correlation parameters, the impact of the correlation rate-vol is usually much smaller than the impact of the rate-stock correlation, e.g. see Antonov et al. (2008) or Chapter 3. Moreover, from a practical point of

view, the rate-stock parameter is most important for the pricing and hedging of hybrid equity-interest rate securities. We therefore restrict ourselves to investigate the impact of the rate-stock parameter on the quality of the control variate estimator: we look at the (empirical) variance reductions for a three year call option with an ATMF (at-the-money-forward) strike level of 100% for different ρ_{I,x_1^n}. The results can be found in Table **2** below,

ρ_{I,x_1^n}	$\widehat{\rho_{C^0,C^\rho}}$	$\widehat{b}$	**Var. Red.**	ρ_{I,x_1^n}	$\widehat{\rho_{C^0,C^\rho}}$	$\widehat{b}$	**Var. Red.**
-0.9	99.859%	0.960	**356**	0.9	99.864%	1.034	**367**
-0.8	99.911%	0.965	**562**	0.8	99.913%	1.031	**574**
-0.7	99.940%	0.970	**839**	0.7	99.941%	1.027	**852**
-0.6	99.960%	0.974	**1 254**	0.6	99.961%	1.024	**1 268**
-0.5	99.974%	0.979	**1 937**	0.5	99.974%	1.020	**1 950**
-0.4	99.984%	0.983	**3 188**	0.4	99.984%	1.016	**3 202**
-0.3	99.992%	0.987	**5 888**	0.3	99.992%	1.012	**5 902**
-0.2	99.996%	0.992	**13 597**	0.2	99.996%	1.008	**13 614**
-0.1	99.999%	0.996	**55 209**	0.1	99.999%	1.004	**55 252**
0	100%	1	∞	-	-	-	-

Table 2: Expected variance reductions when using the Control variate estimator of (4.102) instead of the standard Monte Carlo estimator. For various values of ρ_{I,x_1^n} the expected reduction for a three-year call option with an at-the-money strike is calculated using the estimates $\widehat{b}$ and $\widehat{\rho_{C^0,C^\rho}}$ respectively for the optimal control coefficient and correlation between the control and the estimated quantity. Parameter settings from **case I** of Table **1**. Results based on 50.000 pseudo-random paths.

From the above table (the case $\rho_{I,x_1^n} = 1$ does not constitute in a valid correlation matrix and is hence omitted), we can see that the control is in all cases very effective, i.e. resulting in large to huge variance reductions. As expected, the variance reductions become larger for smaller absolute values of ρ_{I,x_1^n}; for the case $\rho_{I,x_1^n} = 0$, the control is perfect and results in a zero variance control variate estimator, whereas for larger values of $|\rho_{I,x_1^n}|$ the correlation between the 'uncorrelated' and 'correlated' option prices is smaller and therefore reduces the effectiveness of the control, as is being theoretically underpinned by formula (4.101). Thus from Table **2** we can see that the effectiveness of the control, i.e. the resulting variance reduction, depends to a large extent on the absolute value of the correlation 'between' interest rates and equity underlying. Finally, it is worthwhile mentioning that because the ρ_{C^0,C^ρ} and $\frac{\sigma_{C^\rho}}{\sigma_{C^0}}$ the (estimated) optimal coefficients $\widehat{b}$ are also close to one. In such a situation it might be more efficient to just set $b^* = 1$ and consequently save the computational effort in estimating $\widehat{\rho_{C^0,C^\rho}}$, see Glasserman (2003).

Results for case II

The second test case of Table **1**, consists of an experiment where investigate the variance reductions of (4.102) over the standard Monte Carlo estimator for European call options of different maturities and strikes. Furthermore, since the same parameters are being used in Antonov et al.

(2008), we can use these results to draw a comparison between the Monte Carlo control variate estimator and the Markovian projection technique. The numerical results can be found in Table **3**.

From that table, we can see that the control variate estimator by far outperforms the ordinary Monte Carlo estimator; for short to moderate maturity options the control variate shows large to huge variance reduction factors varying from 629 to 7938. For middle to long term option options, the variance reductions are smaller, but still quite reasonable with reductions from 54 to 371. If we look at the variance reductions over different strike levels, the differences are somewhat smaller. It is worthwhile to notice that, for a fixed maturity, the control variate is most effective for out-of-money options, which are usually the hardest options to value by (plain) Monte Carlo.
We conclude the analysis, by comparing the Monte Carlo control variate estimator (4.102) with the Markovian Projection technique. The results of the best projection technique of Antonov et al. (2008) is denoted by Heston DV (displaced volatility) and can be found in the fourth column of Table **3**.
A crucial difference between the simulation and MP method is that the MP technique is in principle a biased approximation, whereas the control variate is unbiased and converges to the true price. However, in practice one often only has a limited available computational budget and one will also note bias in the Monte estimates as a consequence of the limited number of simulations; this bias could be larger than the error in the approximation. Essentially the choice between both methods therefore constitutes of a tradeoff between speed and accuracy, which might differ across applications and products.

Maturity in years	Strike	Sim. Vol. (std. dev.)* (Ordinary MC)	MP error*	**Var. Red. (CV variate MC)**
1	86.07	24.45 (0.06)	0.04	**6 381**
1	92.77	22.25 (0.05)	0.02	**5 884**
1	100.00	20.36 (0.05)	-0.04	**5 717**
1	107.79	19.42 (0.05)	-0.08	**6 549**
1	116.18	19.67 (0.06)	-0.03	**7 938**
3	77.12	22.61 (0.08)	0.03	**661**
3	87.82	20.05 (0.08)	0.01	**622**
3	100.00	17.95 (0.09)	-0.04	**629**
3	113.87	17.23 (0.13)	-0.09	**763**
3	129.67	18.02 (0.18)	-0.09	**985**
5	71.50	21.89 (0.06)	0.06	**250**
5	84.56	19.43 (0.05)	0.02	**240**
5	100.00	17.49 (0.06)	-0.05	**246**
5	118.26	16.83 (0.08)	-0.11	**295**
5	139.85	17.55 (0.12)	-0.13	**371**
10	62.23	21.55 (0.07)	0.06	**98**
10	78.89	19.52 (0.07)	0.00	**100**
10	100.00	18.01 (0.08)	-0.10	**106**
10	126.77	17.41 (0.11)	-0.19	**124**
10	160.70	17.75 (0.16)	-0.24	**152**
20	51.13	22.28 (0.06)	0.03	**54**
20	71.50	20.91 (0.06)	-0.05	**55**
20	100.00	19.94 (0.06)	-0.17	**57**
20	139.85	19.44 (0.09)	-0.27	**63**
20	195.58	19.40 (0.13)	0.35	**72**

Table 3: Variance reductions for **case I** of Table **1** using 50000 pseudo-random paths. Reported is the variance reduction factor ('Var. Red.'), i.e. the fraction between the variance of the control variate (4.102) and the standard Monte Carlo estimator. The starred results, i.e. from the simulated volatility and standard deviations ('Sim. Vol. (std. dev.)*') and the errors of the Heston DV Markovian projection ('MP error*'), were taken from Antonov et al. (2008).

Using the numerical results of Table **3**, let us consider the concrete example of pricing a ten-year ATMF call option. For the sake of the argument we assume here that the Monte Carlo volatility of 18.01 is in fact the true volatility and hence the Markovian Projection error is 0.10. We can then ask ourselves how many simulations are needed to improve the error of this approximation in at least 90% of the cases. By definition, 90% of all the spanned confidence intervals should contain the 'true' price of 18.01, hence to improve the MP error, we should aim to get the standard deviation of the Monte Carlo estimated volatility smaller than $\frac{0.10}{\Psi^{-1}(95\%)} = 0.061$ with Ψ^{-1} the inverse of a standard Gaussian distribution. Using the fact that the Black and Scholes (1973) ATMF price is close to linear as a function of the volatility, taking the standard deviation 0.08

of the simulated volatility and the variance reduction factor 108 of the above table and assuming a convergence rate of the Monte Carlo of one over square root of the number of simulations N, one can find that one needs

$$M = \frac{\text{Var}_N}{\text{Var}_{\text{REQ}}} \frac{N}{\text{VR}} = \frac{0.08^2}{0.061^2} \frac{50000}{108} = 802$$

simulations to improve upon the MP error in 90% of the cases, with VR the variance reduction factor and where Var_{REQ} represents the required variance corresponding to a confidence level $1 - \alpha = 90\%$. Should we for example take $\alpha = 50\%$, one can find that on average one only has to use 134 simulations to perform 'equally well' as the MP projection. Hence due to the large variance reductions, only a very moderate amount of simulations is needed to come up with a good estimate. Though the above analysis is too small to draw very strong conclusions about the comparison between the MP projection technique and the control variate, the main conclusion we do like to draw is that only a moderate amount of simulations is required to obtain reliable price/volatility estimates for the above call options: in most situations a couple of thousand paths will suffice to obtain prices that lie within typical bid-ask spreads.
Finally we would also like to point out that the MP projection might also be used in conjunction with the control variate estimator (4.102) in a model calibration procedure; a first point (in future research) could be to investigate the quality of the MP projection as control for the exact dynamics. Secondly, in a practical implementation one might first use the MP approximation to calibrate the model (which consists of most of the iterations) and consecutively use the control variate to refine the (near) optimal parameters found in the previous minimization. Please note hereby that (for each new parameter guess) one only needs a single Monte Carlo run to price all options simultaneously. In this way (assuming one uses a sufficiently large number of paths in the last few optimization steps using the Monte Carlo) one can get the best of both worlds, i.e. the speed of an approximating formula combined with the accuracy of the control variate estimator.

4.5.2 Calibration to FX market

We will test our model by calibrating it to FX option market data. To this end, we consider the same vanilla FX data (see appendix 4.7.3) as is considered in Piterbarg (2005) who uses this set for the calibration of his local volatility model. In an elegant paper, Piterbarg (2005) concludes that for the pricing and managing of exotic FX derivatives (i.e. PRDCs), it is essential to take the FX implied volatility skew/smile into account; hence though FX model setups may differ, i.e. local volatility in Piterbarg (2005), Heston (1993) stochastic volatility with independent stochastic interest rate drivers in Andreasen (2006) and our stochastic volatility model with multi-factor Gaussian rates and Heston (1993) or Schöbel and Zhu (1999) volatility under a full correlation structure, all these models share the essential feature of explicitly accounting for the FX skew/smile.
For the calibration results of our model we consider the same interest rate and correlation parameters as in Piterbarg (2005); that is, the interest curves in the domestic (Japanese yen) and foreign

(US dollar) economies are given by

$$\begin{aligned} P_n(0,T) &= \exp\Big(-0.02 \cdot T\Big), \\ P_r(0,T) &= \exp\Big(-0.05 \cdot T\Big), \end{aligned}$$

and the one-factor Hull and White (1993) interest rate parameters for the interest rate evolutions in both currencies are given by

$$\begin{aligned} a_n(t) &:= 0.0\%, \quad \sigma_n(t) := 0.70\%, \\ a_r(t) &:= 5.0\%, \quad \sigma_r(t) := 1.2\%. \end{aligned}$$

The correlation parameters are given by

$$\rho_{n,r} = 25.00\%, \quad \rho_{I,n} = \rho_{I,r} = -15.00\%, \quad \rho_{n,\nu} = \rho_{r,\nu} = 0.00\%.$$

Note that our stochastic volatility model has the additional flexibility of correlating the domestic of foreign exchanges with the volatility drivers (i.e. through $\rho_{n,\nu}$ or $\rho_{r,\nu}$), however for simplicity we fix them on zero here. The initial spot FX rate (yen per dollar) is set at 105.00. The ten expiry dates and the seven strikes that are considered for the calibration, are given in Table **4** of appendix 4.7.3. For each maturity T_n, the strikes $K_i(T_n)$ are being computed using the formula

$$K_i(T_n) = F(0,T_n) \exp\Big(0.1 \cdot \delta_i \sqrt{T_n}\Big), \qquad \delta_i \in \{-1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5\}. \quad (4.104)$$

In particular, note that the fourth strike level corresponds to the forward FX rate for that date. The implied volatilities corresponding to the above strikes and maturities can be found in Table **5** of appendix 4.7.3. With the above setup, we consider in the next section how well the models (4.8), i.e. with Schöbel and Zhu (1999) and Heston (1993) stochastic volatility, fit the market implied volatilities of Table **5**.

Calibration results

We calibrate the models (4.8) with Schöbel and Zhu (1999) and Heston (1993) stochastic volatility to the various maturities by minimising the differences between model and market implied volatilities by using a local optimization method. The differences are reported in Table **6** and **7** below of Appendix 4.7.3. For visual comparison, we represent the calibration results for a few maturities in the graphs below.

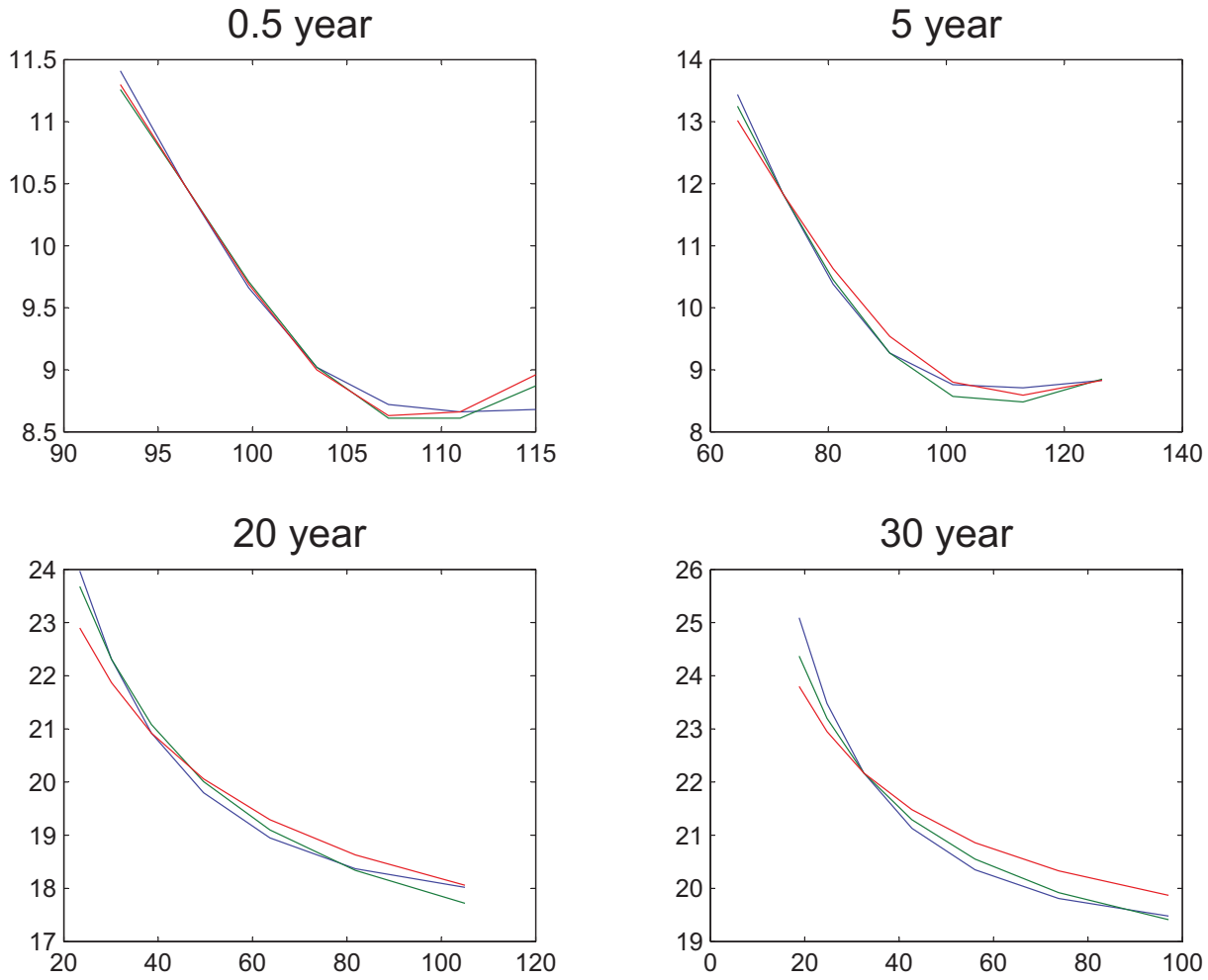

Figure 1: Calibration results for the model with Schöbel-Zhu and Heston stochastic volatility. For the maturities 0.5, 5, 20 and 30 the implied volatilities (vertical axis) are plotted against the corresponding strikes (horizontal axis). The market data is represented in blue, the model (4.8) with Schöbel-Zhu volatility in red and the model (4.8) with Heston volatility in green.

We first consider the model (4.8) with Schöbel and Zhu (1999) stochastic volatility. The model produces a good fit to the market, as can be seen from Table **6** and Figure **1**, with differences smaller than 0.50% in most points and with a good fit around the at-the-money-forward volatilities and the slope of the volatility skews for each maturity. The model produces similar calibration results as the models of Piterbarg (2005) and Andreasen (2006). The low-strike (in-the-money call) options are underestimated by the model, which seems to have slight difficulties in fitting the tails of the implied volatility structure, suggesting the addition of an extra factor, e.g. a trivial extension including Poisson-type jumps. Nonetheless, the smiles produced by the model are much closer to the market than a log-normal model would indicate, in particular for in- and out-the-money options.
Secondly, we consider the model (4.8) with Heston (1993) stochastic volatility. For simplicity

we have considered uncorrelated stochastic volatility, as we can then directly price the required FX options in closed form. Nonetheless, the calibration results to call option prices should be very similar as it is shown in Antonov et al. (2008), that the parameters of the general model can often be projected onto parameters of the uncorrelated model, while to a large extent preserving option prices and model characteristics. The calibration results can be found in Figure **1** and in Table **7** of appendix 4.7.3. From this, we can see that the model again produces a very good fit to the market, with differences now even smaller than 0.30% in most points and with excellent fits across moneyness and maturities. It seems that Heston (1993) model is slightly better in fitting extreme/convex FX skew we calibrating against; in a way is able to capture both the volatility part of the at-the-money prices, as well as the extremes of the in- and out-the-money prices. Alternatively, one can argue that the addition of an extra factor is still needed for the pricing of certain exotic options (e.g. see van der Ploeg (2006) and Fouque et al. (2000)), which discussion is however beyond the scope of this article.
As shown in Piterbarg (2005) and Andreasen (2006), it is of crucial importance to take the FX skew into account for the pricing and managing of exotic FX structures like PRDCs (power reverse dual contracts) or cliquets. Therefore, since the skews/smiles generated by our stochastic volatility models are much closer to the market than produced by a log-normal model, we can conclude our stochastic volatility model(s) (4.8) is better suited to price and manage these exotic FX structures. Finally, though the models of Piterbarg (2005) and Andreasen (2006) account for the FX skew, our model stands out as we model stochastic volatility (versus local volatility used in Piterbarg (2005)) and stochastic interest rates, whilst we allow all driving model factors to be instantaneously correlated with each other (versus independent Gaussian rates used in Andreasen (2006)). Having this flexibility yields a realistic model, which is of practical importance for the pricing and hedging of options with a long-term FX exposure.
Given data on FX prices our model can also be used to examine the pricing and hedging performance of products which explicitly depend on future volatility smiles, such as barrier FX options. An empirical study on the relative performance of the stochastic volatility models discussed here versus other models for FX, as well as the calibration and pricing of liquid vol-sensitive instruments such as double-no-touch options (e.g. see Kainth and Saravanamuttu (2007)), is beyond the scope of this thesis, and is left for future research.

4.5.3 Calibration to Inflation Markets

In a recent paper, Mercurio and Moreni (2009) consider the pricing of inflation-indexed year-on-year and zero-coupon caps/floors, using a market model with SABR Hagan et al. (2002) stochastic volatility dynamics for year-on-year inflation rates and a lognormal Libor Market model for Nominal interest rates. Other market model approaches for inflation can for instance be found in Belgrade et al. (2004), Kenyon (2008) and Brigo and Mercurio (2006). Compared to the latter models, the approach considered in Mercurio and Moreni (2009) stands out by reconciling both zero-coupon and year-on-year quotes. Similar to the framework considered in this paper, these authors consider a full correlation structure between the stochastic quantities underlying the model, whilst preserving closed-form and flexible calibration methods for calibration to market option data.

Differences between market models and low dimensional Markov models, as considered in this chapter, have been described by several authors, e.g. see Brigo and Mercurio (2006), Pelsser (2000); market models explicitly model observable quantities (e.g. Year-on-Year inflation rates), and due to their dimensionality provide a larger calibration flexibility compared to low-dimensional Markov models. On the other hand, the dimensionality of such models can also be disadvantageous. For instance, due to a lack of calibration instruments in less liquid markets (such as inflation options), hedges and calibrations may become unstable when using market models, e.g. see Jäckel and Bonneton (2010). These models can also be relatively slow compared to low-dimensional Markov models, e.g. see Glasserman (2003). In this sense, both market and low dimensional market models show (dis)advantages and the model choice ultimately depends on the exotic product one wants to price.
One clear objective to judge the quality of a model is of course its calibration flexibility. We will therefore look at calibrations of the model (4.8) with Schöbel and Zhu (1999) stochastic volatility, with one-factor rates for the nominal economies, to inflation-indexed caplets and floorlets. To test the calibration of this model, we use the same market data and zero correlation assumptions as in the first case of Mercurio and Moreni (2009). Adopting this setup has the additional advantage that it enables us to draw a comparison between these methods, for a further description of the market data we refer the reader to that paper. Calibration results are shown in Figure **2**.

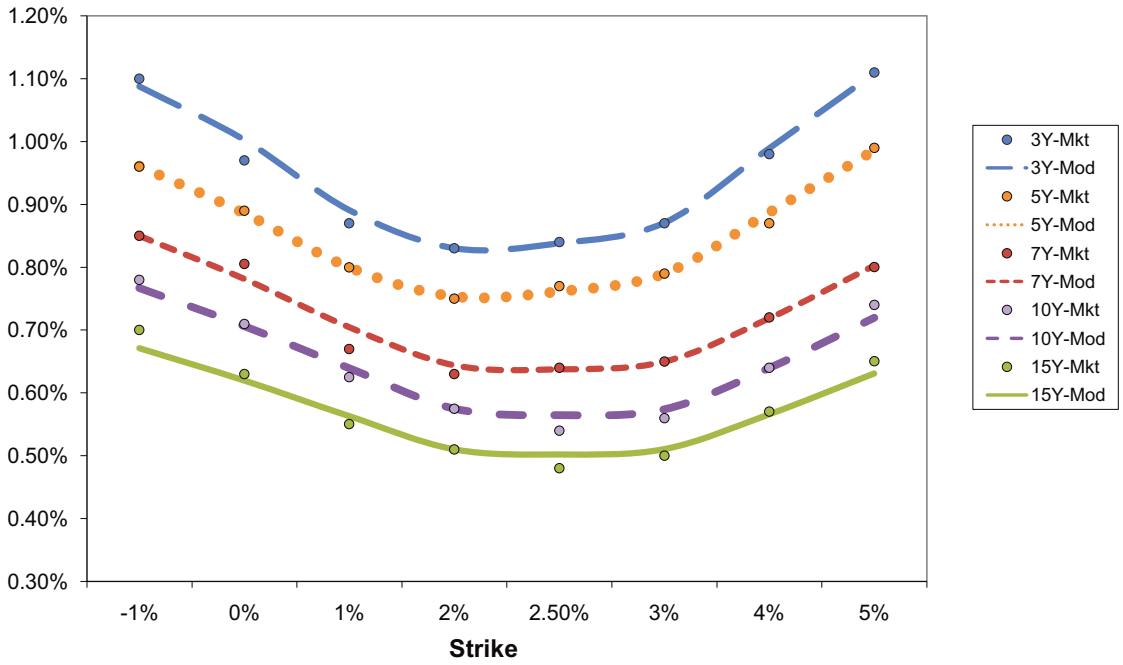

Figure 2: Calibration results: market and model implied volatilities for caplets/floorlets maturing in 3,5, 7, 10, 15 years. Calibration results for the model the model (4.8) with Schöbel-Zhu volatility. Data for September 4th 2008, corresponding with the uncorrelated case of Figure 1 in Mercurio and Moreni (2009).

We can see from Figure **2**, in which market and model implied volatilities are reported, that the fit is accurate. We note that the market data displays small non-smooth behaviours where cap and

floor quotes meet (strikes $2\% - 2.5\%$) or on single strikes (e.g. the 7-years 0%-floorlet). Similar to Mercurio and Moreni (2009), we consider these discrepancies as being essentially bound to liquidity reasons and stress that a parameterized models also provide useful smoothing tools for such market data. The calibration results of the model with Schöbel-Zhu and SABR stochastic volatility are very similar: both models are well able to qualitatively fit the shape of the implied volatility, whilst they are also capable to detect small market anomalies. Because the liquidity of inflation options is not that large, as can be seen from wide bid/ask spreads for inflation caps and floors, such smoothing of market data might be very useful, as indicated Mercurio and Moreni (2009) and Jäckel and Bonneton (2010). We therefore conclude that whilst the Schöbel-Zhu stochastic volatility model has all advantages of a low dimensional Markov model, it is also flexible enough to fit prices of vanilla inflation-indexed options in an accurate way.

4.6 Conclusion

We have introduced a generic model incorporating stochastic interest rates and stochastic volatility under a full correlation structure of all driving model factors, with closed-form pricing formulas for vanilla options and which is able to incorporate the markets implied volatility structures. Having the flexibility to correlate the underlying FX/Inflation/Stock-index with both the stochastic volatility and the stochastic interest rates yields a realistic model, which is of practical importance for the pricing and hedging of options with a long-term exposure. Furthermore, closed-form pricing of vanilla FX, Inflation and stock options is a big advantage for the calibration (and sensitivity analysis) of the model; using Fourier methods, we have shown how vanilla call/put options, forward starting options, year-on-year inflation-indexed swaps and inflation-indexed caps/floors can be valued in closed-form. Hereby it must be noted that our model can cover Poisson type jumps with a trivial extension. Under Schöbel and Zhu (1999) stochastic volatility, using its affine properties, we were able to derive the corresponding characteristic functions in closed-form. Under Heston (1993) stochastic volatility, the characteristic functions can only be derived explicitly under special zero correlation assumptions. Nonetheless, we have demonstrated that one can still use these pricing formulas: either by using a projection of the general model onto the uncorrelated case, or by using it as a control variate for the general model. The latter method even results in such large variance reductions that its incorporation in the calibration becomes a more than viable option. Our model can be used for multi-asset purposes (e.g. interest rates, FX, inflation, equity, commodities) and is fast enough for the real life risk management of big portfolios of such products. We think it is particularly suitable for the pricing and hedging of long-dated multi-currency structures (e.g. hybrid TARN options, variable annuities, inflation LPI options and PRDC FX swaps) which are sensitive to both future interest rates evolutions as well as movements from the underlying index and/or corresponding volatility smiles.

4.7 Appendix

4.7.1 Deriving the characteristic function of the log 'Schöbel-Zhu' inflation rate

In this appendix we will prove that the partial differential equation (4.60), i.e.

$$\begin{aligned} 0 &= f_t - \frac{1}{2}\nu_F^2(t) f_z + \kappa(\xi(t) - \nu(t)) f_\nu + \frac{1}{2}\nu_F^2(t) f_{zz} \\ &\quad + \Big(\rho_{I\nu}\tau\nu(t) + \sum_{i=1}^{K} \rho_{x_n^i\nu}\tau\sigma_n^i B_n^i(t,T) - \sum_{j=1}^{M} \rho_{x_r^j\nu}\tau\sigma_r^j B_r^j(t,T)\Big) f_{z\nu} + \frac{1}{2}\tau^2 f_{\nu\nu}, \end{aligned} \tag{4.105}$$

subject to the boundary condition $f(T, y, \sigma) = \exp(iuy(T))$ has a solution given by (4.61)-(4.66); to solve this differential equation, we use the ansatz (4.61), find the corresponding partial derivatives and substitute these in (4.60). We then obtain a system of ordinary differential equations that is similar to the one-factor model as in van Haastrecht et al. (2009) and which can be solved using similar methods.

Expanding $\nu_F^2(t)$ according to (4.49) and collecting the terms for $y(t)$, $\nu(t)$ and $\frac{1}{2}\nu^2(t)$ yields the following system of ordinary differential equations for the functions $A(u,t,T), \ldots, D(u,t,T)$:

$$0 = \frac{\partial B(u,t,T)}{\partial t}, \tag{4.106}$$

$$0 = \frac{\partial D(u,t,T)}{\partial t} - 2(\kappa - \rho_{x\nu}\tau B) D(u,t,T) + \tau^2 D^2(t) + (B^2 - B), \tag{4.107}$$

$$\begin{aligned} 0 &= \frac{\partial C(u,t,T)}{\partial t} + (\rho_{x\nu}\tau B - \kappa + \tau^2 D) C(u,t,T) \\ &\quad + \Big\{ \sum_{i=1}^{K} [\rho_{Ix_n^i}\sigma_n^i B_n^i(t,T)] - \sum_{j=1}^{M} [\rho_{Ix_r^j}\sigma_r^j B_r^j(t,T)] \Big\} (B^2 - B) \\ &\quad + \Big\{ \kappa\xi(t) + \Big(\sum_{i=1}^{K} [\rho_{x_n^i\nu}\tau\sigma_n^i B_n^i(t,T)] - \sum_{j=1}^{M} [\rho_{x_r^j\nu}\tau\sigma_r^j B_r^j(t,T)\Big) B \Big\} D(u,t,T), \end{aligned} \tag{4.108}$$

$$\begin{aligned} 0 &= \frac{\partial A(u,t,T)}{\partial t} + \Big[\kappa\xi(t) + \sum_{i=1}^{K} \rho_{x_n^i\nu}\tau\sigma_n^i B_n^i(t,T) B - \sum_{j=1}^{M} \rho_{x_r^j\nu}\tau\sigma_r^j B_r^j(t,T) B \Big] C(u,t,T) \\ &\quad + \frac{1}{2}\tau^2 (C^2(u,t,T) + D(u,t,T)) + \frac{1}{2}(B^2 - B)\nu_{K,M}^2(t,T), \end{aligned} \tag{4.109}$$

with $\nu_{K,M}^2(t,T)$ the instantaneous variance of the Gaussian rate processes, see (4.50). It immediately that follows $B(u,t,T) = B$ equals a constant since its derivative is zero, subject to the boundary condition (4.54) we have

$$B = iu. \tag{4.110}$$

The second equation (4.107) yields a Riccati equation with constant coefficients and boundary condition $D(u,T,T)=0$ which is equivalent to the PDE for the D-term in the SZHW model (see Chapter 3) and has the following solution:

$$\begin{aligned} D(u,t,T) &= -u(i+u)\frac{1-e^{-2\gamma(T-t)}}{\gamma_1+\gamma_2 e^{-2\gamma(T-t)}}, && (4.111)\\ \text{with: } \gamma &= \sqrt{(\kappa-\rho_{xv}\tau B)^2-\tau^2(B^2-B)}, && (4.112)\\ \gamma_1 &= \gamma+\frac{1}{2}q_1=\gamma+(\kappa-\rho_{xv}\tau B), && (4.113)\\ \gamma_2 &= \gamma-\frac{1}{2}q_1=\gamma-(\kappa-\rho_{xv}\tau B). && (4.114)\end{aligned}$$

The third equation (4.108) for $C(u,t,T)$ is a first order linear differential equation of the form $\frac{\partial C(u,t,T)}{\partial t}+g(t)C(u,t,T)+h(t)=0$, with associated boundary condition $C(u,T,T)=0$. Hence we can represent a solution for $C(u,t,T)$ as:

$$\begin{aligned} C(u,t,T) &= \int_t^T h(s)\exp[\int_t^s g(v)dv]ds, && (4.115)\\ \text{with: } g(v) &= -(\kappa-\rho_{xv}\tau B)+\tau^2 D(u,v,T), && (4.116)\\ h(s) &= \Big(\kappa\xi(u)+\sum_{i=1}^{K}(\rho_{x_n^i v}\tau\sigma_n^i B_n^i(s,T)-\sum_{j=1}^{M}\rho_{x_r^j v}\tau\sigma_r^j B_r^j(s,T))B\Big)D(u,s,T)\\ &\quad+\Big(\sum_{i=1}^{K}\rho_{Ix_n^i}\sigma_n^i B_n^i(s,T)-\sum_{j=1}^{M}\rho_{Ix_r^j}\sigma_r^j B_r^j(s,T)\Big)(B^2-B)\\ &= \kappa\psi D(u,s,T)\\ &\quad+\sum_{i=1}^{K}\Big\{\rho_{Ix_n^i}\sigma_n^i B_n^i(s,T)(B^2-B)+\Big[\rho_{x_n^i v}(B-1)\tau\sigma_n^i B_n^i(s,T)\Big]D(u,s,T)\Big\}\\ &\quad-\sum_{j=1}^{M}\Big\{\rho_{Ix_r^j}\sigma_r^j B_r^j(s,T)(B^2-B)+\Big[\rho_{x_r^j v}B\tau\sigma_r^j B_r^j(s,T)\Big]D(u,s,T)\Big\}. && (4.117)\end{aligned}$$

We first consider the integral over g: dividing equation (4.116) by $D(u,t,T)$, rearranging terms and integrating we find the surprisingly simple solution:

$$\begin{aligned} \int g(v)dv &= \int -(\kappa-\rho_{xv}\tau B)+\tau^2 D(u,v,T)dv\\ &= \int(\kappa-\rho_{xv}\tau B)-\frac{(B^2-B)}{D(u,v,T)}-\frac{\partial D(u,v,T)}{\partial v}\frac{1}{D(u,v,T)}dv\\ &= \log\Big(\gamma_1 e^{\gamma(T-t)}+\gamma_2 e^{-\gamma(T-t)}\Big)+c, && (4.118)\end{aligned}$$

where γ,γ_1 and γ_2 are defined in (4.66) and with c the integration constant. Hence taking the

exponent and filling in the required integration boundaries yields

$$\exp\Big[\int_t^s g(v)dv\Big] = \frac{\gamma_1 e^{\gamma(T-s)} + \gamma_2 e^{-\gamma(T-s)}}{\gamma_1 e^{\gamma(T-t)} + \gamma_2 e^{-\gamma(T-t)}}, \tag{4.119}$$

Substituting this expression into (4.115) we find (after a long but straightforward calculation) for $C(u,t,T)$:

$$\begin{aligned} C(u,t,T) &= -\frac{u(i+u)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}\Big\{\gamma_0\Big(1+e^{-2\gamma(T-t)}\Big) \\ &\quad + \sum_{i=1}^{K}\Big[\Big(\gamma_3^i - \gamma_4^i e^{-2\gamma(T-t)}\Big) - \Big(\gamma_5^i e^{-a_n^i(T-t)} - \gamma_6^i e^{-(2\gamma+a_n^i)(T-t)}\Big) - \gamma_7^i e^{-\gamma(T-t)}\Big] \\ &\quad - \sum_{j=1}^{M}\Big[\Big(\gamma_8^j - \gamma_9^j e^{-2\gamma(T-t)}\Big) - \Big(\gamma_{10}^j e^{-a_r^j(T-t)} - \gamma_{11}^j e^{-(2\gamma+a_r^j)(T-t)}\Big) - \gamma_{12}^j e^{-\gamma(T-t)}\Big]\Big\} \end{aligned} \tag{4.120}$$

with the constants $\gamma, \gamma_0, \ldots, \gamma_{12}^j$ as defined in (4.66).

Finally, by integration equation (4.110), we find the following expression for $A(u,t,T)$:

$$\begin{aligned} A(u,t,T) &= \int_t^T \Big[\frac{1}{2}(B^2 - B)v_{K,M}^2(t,T) + \kappa\xi(s)C(u,s,T) + \frac{1}{2}\tau^2\big(C^2(u,s,T) + D(u,s,T)\big)\Big]ds \\ &= -\frac{1}{2}u(i+u)V_{K,M}(t,T) \\ &\quad + \int_t^T \Big\{\Big[\kappa\psi + (iu-1)\sum_{i=1}^{K}\rho_{x_n^i\nu}\tau\sigma_n^i B_n^i(t,T) - iu\sum_{j=1}^{M}\rho_{x_r^j\nu}\tau\sigma_r^j B_r^j(t,T)\Big]C(u,s,T) \\ &\qquad + \frac{1}{2}\tau^2\Big(C^2(u,s,T) + D(u,s,T)\Big)\Big\}ds, \end{aligned} \tag{4.121}$$

where $V_{K,M}(t,T)$ is the integrated variance of the multi-factor Gaussian rates which can found by simple integration, see (4.51). It is possible to write a closed-form expression for the remaining integral in (4.121). As the ordinary differential equation for $D(u,s,T)$ is exactly the same as in the Heston (1993) or Schöbel and Zhu (1999) model, it will involve a complex logarithm and should therefore be evaluated as outlined in Lord and Kahl (2007) in order to avoid any discontinuities. The main problem however lies in the integrals over $C(u,s,T)$ and $C^2(u,s,T)$, which will involve the Gaussian hypergeometric ${}_2F_1(a,b,c;z)$. The most efficient way to evaluate this hypergeometric function (according to *Numerical Recipes*, Press and Flannery (1992)) is to integrate the defining differential equation. Since all of the terms involved in $D(u,s,T)$ are also required in $C(u,s,T)$, numerical integration of the second part of (4.121) seems to be the most efficient method for evaluating $A(u,t,T)$. Hereby we conveniently avoid any issues regarding

complex discontinuities altogether.

4.7.2 Analytical properties of the Gaussian factors driving the asset price process

In this section we will discuss some properties of the processes that drive the asset price dynamics. That is, we discuss the pricing of bonds under multi-factor Gaussian interest rates (Section 4.7.2) and the moments of the Gaussian interest rates processes and the Ornstein-Uhlenbeck distributed volatility process under the T-forward measure (Section 4.7.2).

Zero-coupon bond prices under multi-factor Gaussian rates

In this appendix we briefly review zero-coupon bond prices of the Gaussian multi-factor rate model, i.e. one has the following analytical formulas for the zero-coupon bond prices (e.g. see Brigo and Mercurio (2006) for the two-factor model, which easily extends to the multi-factor case):

$$P_n(t,T) = \mathbb{E}_n\left\{e^{-\int_t^T n(u)du}\right\} = A_n(t,T)e^{-B_n(t,T)X_n(t)} = A_n(t,T)e^{-\sum_{i=0}^{K} B_n^i(t,T)x_n^i(t)}, \tag{4.122}$$

$$A_n(t,T) = \frac{P_n^M(0,T)}{P_n^M(0,t)}\exp\left\{\frac{1}{2}\Big[V_n(t,T) - V_n(0,T) + V_n(0,t)\Big]\right\}, \tag{4.123}$$

here $B_n^i(t,T) = \frac{1-e^{-a_n^i(T-t)}}{a_n^i}$. It is straightforward to generalize this to the case of time-dependent model parameters, i.e. in that case $B_n^i(t,T) := \int_t^T e^{-a(u)\cdot(T-u)}du$. Expressions for the real bond prices $P_r(t,T)$ and affine terms $A_r(t,T)$, $B_r(t,T)$ are completely analogous.

For the integrated rate variances $V_i(t,T)$, one also has closed-form expressions. To this end we let (just as in Section 4.4) $C^{(i,j)}$ and $R^{(i,j)}$ respectively denote the integrated covariance and correlation between the i-th and j-th element of the vector of rate volatilities $\Sigma(t)$ of (4.48). One can then express the integrated rate variances as

$$V_n(t,T) = \sum_{i=2}^{K+1} C^{(i,i)} + 2\sum_{i=2}^{K+1}\sum_{j=i+1}^{K+1} R^{(i,j)}C^{(i,j)}, \tag{4.124}$$

$$V_r(t,T) = \sum_{i=K+2}^{K+M+1} C^{(i,i)} + 2\sum_{i=K+2}^{K+M+1}\sum_{j=i+1}^{K+M+1} R^{(i,j)}C^{(i,j)}. \tag{4.125}$$

Expressions for these covariances are provided in Section 4.4.

Moments of the interest rate and volatility processes

In this appendix, we will derive the moments of the stochastic factors that drive the nominal, real and volatility rate. Since all factors follow Ornstein-Uhlenbeck processes, the moments can be found relatively easy.

Moments of the volatility process

By integrating the T_i forward dynamics of (4.18) conditional on $\nu(t)$, we obtain

$$\nu(T_{i-1}) = \nu(t)e^{-\kappa(T_{i-1}-t)} + \int_t^{T_{i-1}} \kappa\xi(u)e^{-\kappa(T_{i-1}-u)}du + \tau\int_t^{T_{i-1}} e^{-\kappa(T_{i-1}-u)}dW_\nu^T(u), \tag{4.126}$$

where $\xi(u) := \psi - \sum_{i=1}^K \frac{\rho_{x_n^i,\nu}\sigma_n^i\tau}{a_n^i\kappa}[1-e^{-a_n^i(T_i-u)}]$. From Itô's isometry, we then have that the mean and variance of ν under the T_i-forward measure are given by:

$$\begin{aligned}
\mu_\nu &= \nu(t)e^{-\kappa(T_{i-1}-t)} + \Big(\psi - \sum_{i=1}^K \frac{\rho_{x_n^i,\nu}\sigma_n^i\tau}{a_n^i\kappa}\Big)\Big(1-e^{-\kappa(T_i-t)}\Big) \\
&\quad - \sum_{i=1}^K \frac{\rho_{x_n^i,\nu}\sigma_n^i\tau}{a_n^i(\kappa+a_n^i)}\Big[e^{-a_n^i(T_i-t)-\kappa(T_{i-1}-t)} - e^{-a_n^i(T_i-T_{i-1})}\Big], \qquad (4.127)\\
\sigma_\nu^2 &= \frac{\tau^2}{2\kappa}\Big(1-e^{-2\kappa(T_{i-1}-t)}\Big). \qquad (4.128)
\end{aligned}$$

Moments of the rate processes

Starting from time t, one can integrate the rate dynamics of $x_n^i(T_{i-1})$ and $x_r^j(T_{i-1})$, from time t to T_{i-1}, to obtain the following following explicit solutions (see also Pelsser (2000) or Brigo and Mercurio (2006))

$$\begin{aligned}
x_n^k(T_{i-1}) &= x_n^k(t)e^{-a_n^k(T_{i-1}-t)} - M_{n_k}^{T_i}(t,T_{i-1}) + \sigma_n^k\int_t^{T_{i-1}} e^{-a_n^k(T_{i-1}-u)}dW_{n_k}^{T_i}(u), \qquad (4.129)\\
x_r^j(T_{i-1}) &= x_r^j(t)e^{-a_r^j(T_{i-1}-t)} - M_{r_j}^{T_i}(t,T_{i-1}) + \sigma_r^j\int_t^{T_{i-1}} e^{-a_r^j(T_{i-1}-u)}dW_{r_j}^{T_i}(u), \qquad (4.130)
\end{aligned}$$

where

$$\begin{aligned} M_{n_k}^{T_i}(t,T_{i-1}) &= \int_t^{T_{i-1}} \Big[\sigma_n^k \sum_{i=1}^{K} \rho_{x_n^i,x_n^k}\sigma_n^i B_n^i(u,T_i)\Big] e^{-a_n^k(T_{i-1}-u)}du \\ &= \sigma_n^k \frac{1-e^{-a_n^k(T_{i-1}-t)}}{a_n^k} \sum_{i=1}^{K} \rho_{x_n^i,x_n^k}\frac{\sigma_n^i}{a_n^i} \\ &\quad - \sigma_n^k \sum_{i=1}^{K} \rho_{x_n^i,x_n^k}\frac{\sigma_n^i}{a_n^i(a_n^k+a_n^i)}\Big[e^{-a_n^i(T_i-T_{i-1})} - e^{-a_n^i(T_i-t)-a_n^k(T_{i-1}-t)}\Big], \end{aligned} \tag{4.131}$$

$$\begin{aligned} M_{r_j}^{T_i}(t,T_{i-1}) &= \int_t^{T_{i-1}} \Big[\rho_{I,x_r^j}\nu(u)\sigma_r^j + \sigma_r^j \sum_{i=1}^{K} \rho_{x_n^i,x_r^j}\sigma_n^i B_n^i(u,T_i)\Big] e^{-a_r^j(T_{i-1}-u)}du \\ &=: \widetilde{M_\nu^{T_i}}(t,T_{i-1}) + \widetilde{M_{r_j}^{T_i}}(t,T_{i-1}). \end{aligned} \tag{4.132}$$

In the last step we decompose $M_{r_j}^{T_i}(t,T_{i-1})$ into a deterministic part, denoted by $\widetilde{M_{r_j}^{T_i}}(t,T_{i-1})$ and a stochastic part depending on $\nu(u)$, denoted by $\widetilde{M_V^{T_i}}(t,T_{i-1})$. The calculation of the $\widetilde{M_{r_j}^{T_i}}(t,T_{i-1})$-term is similar to the nominal interest rate case and results in the following expression:

$$\begin{aligned} \widetilde{M_{r_j}^{T_i}}(t,T_{i-1}) &= \sigma_r^j \frac{1-e^{-a_r^j(T_{i-1}-t)}}{a_r^j} \sum_{i=1}^{K} \rho_{x_n^i,x_r^j}\frac{\sigma_n^i}{a_n^i} \\ &\quad - \sigma_r^j \sum_{i=1}^{K} \rho_{x_n^i,x_r^j}\frac{\sigma_n^i}{a_n^i(a_r^j+a_n^i)}\Big[e^{-a_n^i(T_i-T_{i-1})} - e^{-a_n^i(T_i-t)-a_r^j(T_{i-1}-t)}\Big]. \end{aligned} \tag{4.133}$$

Hence from Itô's isometry we then have that the mean and variance of $x_n^k(T_{i-1})$ and $\widetilde{x_r^j}(T_{i-1})$ (starting from time t) are respectively given by

$$\mu_n^k(t,T_{i-1}) = x_n^k(t)e^{-a_n^k(T_{i-1}-t)} - M_{n_k}^{T_i}(t,T_{i-1}) \tag{4.134}$$

$$\Big(\sigma_n^k(t,T_{i-1})\Big)^2 = \frac{(\sigma_n^k)^2}{2a_n^k}\Big(1-e^{-2a_n^k(T_{i-1}-t)}\Big) \tag{4.135}$$

$$\mu_r^j(t,T_{i-1}) = x_r^j(t)e^{-a_r^j(T_{i-1}-t)} - \widetilde{M_{r_j}^{T_i}}(t,T_{i-1}) \tag{4.136}$$

$$\Big(\sigma_r^j(t,T_{i-1})\Big)^2 = \frac{(\sigma_r^j)^2}{2a_r^j}\Big(1-e^{-2a_r^j(T_{i-1}-t)}\Big). \tag{4.137}$$

It remains to determine the moments of $\widetilde{M_\nu^{T_i}}(t,T_{i-1})$, i.e. of

$$\widetilde{M_V^{T_i}}(t,T_{i-1}) = \sigma_r^j\rho_{I,x_r^j}\int_t^{T_{i-1}} \nu(u)e^{-a_r^j(T_{i-1}-u)}du. \tag{4.138}$$

By substituting the explicit solution (4.126) for $\nu(u)$ one obtains the following three integrals:

$$\sigma_r^j \rho_{I,x_r^j} \nu(t) \int_t^{T_{i-1}} e^{-\kappa(u-t)} e^{-a_r^j (T_{i-1}-u)} du \tag{4.139}$$

$$\sigma_r^j \rho_{I,x_r^j} \int_t^{T_{i-1}} \Big[\int_t^u \xi(s) e^{-\kappa(u-s)} ds \Big] e^{-a_r^j (T_{i-1}-u)} du \tag{4.140}$$

$$\sigma_r^j \rho_{I,x_r^j} \tau \int_t^{T_{i-1}} \Big[\int_t^u e^{\kappa(u-s)} dW_\nu^{T_i}(s) \Big] e^{-a_r^j (T_{i-1}-u)} du. \tag{4.141}$$

The integral of (4.139) can be calculated exactly and becomes

$$\nu(t) \frac{\sigma_r^j \rho_{I,x_r^j}}{(a_r^j - \kappa)} \Big[e^{-\kappa(T_{i-1}-t)} - e^{-a_r^j (T_{i-1}-t)} \Big]. \tag{4.142}$$

By using Fubini's theorem to interchange the order of integration, one can find that the integral of (4.140) becomes

$$\begin{aligned} &\sigma_r^j \rho_{I,x_r^j} \frac{\kappa e^{-a_r^j (T_{i-1}-t)} + (a_r^j - \kappa) - a_r^j e^{-\kappa(T_{i-1}-t)}}{(a_r^j - \kappa) \kappa a_r^j} \Big[\psi - \sum_{i=1}^{K} \frac{\rho_{x_n^i,\nu} \sigma_n^i \tau}{a_n^i \kappa} \Big] \\ &- \frac{\sigma_r^j \rho_{I,x_r^j}}{\kappa (a_r^j - \kappa)} \sum_{i=1}^{K} \frac{\rho_{x_n^i,\nu} \sigma_n^i \tau}{(\kappa + a_n^i) a_n^i (a_r^j + a_n^i)} \Big\{ (a_r^j + a_n^i) e^{-\kappa(T_{i-1}-1) - a_n^i (T_i - t)} \\ &\qquad - (\kappa + a_n^i) e^{-a_r^j (T_{i-1}-t) - a_n^i (T_i - t)} - (a_r^j - \kappa) e^{-a_n^i (T_i - T_{i-1})} \Big\}. \end{aligned} \tag{4.143}$$

Again by changing the integration order , we find that the following expression holds for the stochastic integral of (4.141):

$$\frac{\sigma_r^j \rho_{I,x_r^j} \tau}{(a_r^j - \kappa)} \int_t^{T_{i-1}} \Big[e^{-\kappa(T_{i-1}-s)} - e^{-a_r^j (T_{i-1}-s)} \Big] dW_\nu^{T_i}(s). \tag{4.144}$$

Hence from Itô's isometry, we have that $\widetilde{M_\nu^{T_i}}(t, T_{i-1})$ of (4.138) is normally distributed with mean

$\mu_V^j(t,T_{i-1})$ and variance $\left(\sigma_V^j(t,T_{i-1})\right)^2$ given by

$$\begin{aligned}\mu_V^j(t,T_{i-1}) &= \sigma_r^j\rho_{I,x_r^j}\nu(t)\Big[e^{-\kappa(T_{i-1}-t)}-e^{-a_r^j(T_{i-1}-t)}\Big] \\ &+\sigma_r^j\rho_{I,x_r^j}\frac{\kappa e^{-a_r^j(T_{i-1}-t)}+(a_r^j-\kappa)-a_r^j e^{-\kappa(T_{i-1}-t)}}{(a_r^j-\kappa)\kappa a_r^j}\Big[\psi-\sum_{i=1}^K\frac{\rho_{x_n^i,\nu}\sigma_n^i\tau}{a_n^i\kappa}\Big] \\ &-\frac{\sigma_r^j\rho_{I,x_r^j}}{\kappa(a_r^j-\kappa)}\sum_{i=1}^K\frac{\rho_{x_n^i,\nu}\sigma_n^i\tau}{(\kappa+a_n^i)a_n^i(a_r^j+a_n^i)}\Big\{(a_r^j+a_n^i)e^{-\kappa(T_{i-1}-1)-a_n^i(T_i-t)} \\ &-(\kappa+a_n^i)e^{-a_r^j(T_{i-1}-t)-a_n^i(T_i-t)}-(a_r^j-\kappa)e^{-a_n^i(T_i-T_{i-1})}\Big\}.\end{aligned} \tag{4.145}$$

$$\begin{aligned}\left(\sigma_V^j(t,T_{i-1})\right)^2 &= \left(\frac{\sigma_r^j\rho_{I,x_r^j}\tau}{(a_r^j-\kappa)}\right)^2\Big\{\frac{1}{2\kappa}+\frac{1}{2a_r^j}-\frac{2}{(\kappa+a_r^j)} \\ &-\frac{e^{-2\kappa(T_{i-1}-t)}}{2\kappa}-\frac{e^{-2a_r^j(T_{i-1}-t)}}{2a_r^j}+\frac{2e^{-(\kappa+a_r^j)(T_{i-1}-t)}}{(\kappa+a_r^j)}\Big\}\end{aligned} \tag{4.146}$$

Terminal correlations between the driving factors

In this section we provide simple analytical expressions for the (terminal) correlations between the driving model factors, $\nu, x_n^1,\ldots,x_n^K,x_r^1,\ldots,x_r^K,V^1,\ldots,V^M$, from the current time t to time T_{i-1}. To this end, we consider the following explicit solutions for these Gaussian processes:

$$\nu(T_{i-1}) = O(dt)+\tau\int_t^{T_{i-1}}e^{-\kappa(T_{i-1}-u)}dW_\nu^{T_i}(u), \tag{4.147}$$

$$x_n^k(T_{i-1}) = O(dt)+\sigma_n^k\int_t^{T_{i-1}}e^{-a_n^k(T_{i-1}-u)}dW_{n_k}^{T_i}(u), \tag{4.148}$$

$$x_r^j(T_{i-1}) = O(dt)+\sigma_r^j\int_t^{T_{i-1}}e^{-a_r^j(T_{i-1}-u)}dW_{r_j}^{T_i}(u), \tag{4.149}$$

$$V^j(T_{i-1}) = O(dt)+\frac{\sigma_r^j\rho_{I,\nu}\tau}{(a_r^j-\kappa)}\int_t^{T_{i-1}}\Big[e^{-\kappa(T_{i-1}-u)}-e^{-a_r^j(T_{i-1}-u)}\Big]dW_\nu^{T_i}(u). \tag{4.150}$$

All of the above processes can be written in the form

$$y_m(T_{i-1}) = O(dt)+c_m\int_t^{T_{i-1}}a_m(u)dW_m(u),$$

hence by Itô's isometry the correlation can be easily calculated; in general, we have that the correlation between, say $y_1(T_{i-1})$ and $y_2(T_{i-1})$, is given by

$$\begin{aligned}\rho_{y_1,y_2}(t,T_{i-1}) &= \frac{\mathrm{Cov}\Big(y_1(T_{i-1}),y_2(T_{i-1})\Big)}{\sqrt{\mathrm{Var}\Big(y_1(T_{i-1})\Big)\cdot\mathrm{Var}\Big(y_2(T_{i-1})\Big)}} \\ &= \rho_{y_1,y_2}\int_t^{T_{i-1}} c_1a_1(u)c_2a_2(u)du\Big/\sqrt{\int_t^{T_{i-1}}[c_1a_1(u)]^2du\cdot\int_t^{T_{i-1}}[c_2a_2(u)]^2du}\,.\end{aligned} \tag{4.151}$$

After identification in (4.147)-(4.150), one has that $a_m(u)$ takes two particular forms

$$a_m(u)=\begin{cases} e^{-b_m(T_{i-1}-u)} & \text{for } \nu, x_n^1,\ldots,x_n^K,x_r^1,\ldots,x_r^K,\\ & b_m\in\{\kappa,a_n^1,\ldots,a_n^k,a_r^1,\ldots,a_r^M\},\\ e^{-\kappa(T_{i-1}-u)}-e^{-b_m(T_{i-1}-u)} & \text{for } V^1,\ldots,V^M,\\ & b_m\in\{a_r^1,\ldots,a_r^M\}\end{cases}$$

Hence by combining the above two forms and using formula (4.151), one has that the resulting correlations take one of the three forms below; to ease notation, we first define the following two integral expressions:

$$\begin{aligned} I_1(b_m) &= \int_t^{T_{i-1}} c_m^2\Big[e^{-b_m(T_{i-1}-u)}\Big]^2du \\ &= c_m^2\left\{\frac{1-e^{-2b_m(T_{i-1}-t)}}{2b_m}\right\} \\ I_2(b_m) &= \int_t^{T_{i-1}} c_m^2\Big[e^{-\kappa(T_{i-1}-u)}-e^{-b_m(T_{i-1}-u)}\Big]^2du \\ &= c_m^2\left\{\frac{1}{2\kappa}+\frac{1}{2b_m}-\frac{2}{(\kappa+b_m)}-\frac{e^{-2\kappa(T_{i-1}-t)}}{2\kappa}-\frac{e^{-2b_m(T_{i-1}-t)}}{2b_m}+\frac{2e^{-(\kappa+b_m)(T_{i-1}-t)}}{(\kappa+b_m)}\right\}.\end{aligned}$$

If $a_1(u)$ and $a_2(u)$ are both of the first form, then the correlation between $y_1(T_{i-1})$ and $y_2(T_{i-1})$ is given by

$$\rho_{y_1,y_2}\frac{c_1c_2}{\sqrt{I_1(b_1)I_1(b_2)}}\frac{1-e^{-(b_1+b_2)(T_{i-1}-t)}}{(b_1+b_2)}. \tag{4.152}$$

If $a_1(u)$ is of the first form and $a_2(u)$ of the second, then the correlation between $y_1(T_{i-1})$ and $y_2(T_{i-1})$ is given by

$$\rho_{y_1,y_2}\frac{c_1c_2}{\sqrt{I_1(b_1)I_2(b_2)}}\left[\frac{1-e^{-(b_1+\kappa)(T_{i-1}-t)}}{(b_1+\kappa)}-\frac{1-e^{-(b_1+b_2)(T_{i-1}-t)}}{(b_1+b_2)}\right]. \tag{4.153}$$

Finally, if $a_1(u)$ and $a_2(u)$ are both of the second form then the correlation between $y_1(T_{i-1})$ and $y_2(T_{i-1})$ is given by

$$\rho_{y_1,y_2}\frac{c_1c_2}{\sqrt{I_2(b_1)I_2(b_2)}} \quad \Big[\frac{1-e^{-2\kappa(T_{i-1}-t)}}{2\kappa}+\frac{1-e^{-(b_1+b_2)(T_{i-1}-t)}}{(b_1+b_2)} - \frac{1-e^{-(b_1+\kappa)(T_{i-1}-t)}}{(b_1+\kappa)}-\frac{1-e^{-(b_2+\kappa)(T_{i-1}-t)}}{(b_2+\kappa)}\Big]. \tag{4.154}$$

Constants in the Quadratic form (4.76)

The constants a_0, b_0 and vector a of the quadratic form (4.76) can be directly extracted from equation (4.75) and are given by

$$\begin{aligned} a_0 &:= iu\Big[A_r(T_{i-1},T_i)-A_n(T_{i-1},T_i)\Big]+A(u,T_{i-1},T_i) \\ &\quad +C(T_{i-1})\mu_v(t,T_{i-1})+\frac{1}{2}D(T_{i-1})\mu_v^2(t,T_{i-1}) \\ &\quad +iu\sum_{k=1}^{K}B_n^k(T_{i-1},T_i)\mu_n^k(t,T_{i-1}) \\ &\quad -iu\sum_{j=1}^{M}B_r^j(T_{i-1},T_i)\Big[\mu_V^j(t,T_{i-1})+\mu_r^j(t,T_{i-1})\Big], \end{aligned} \tag{4.155}$$

$$b_0 := \frac{1}{2}D(u,T_{i-1},T_i)\sigma_v^2(t,T_{i-1}), \tag{4.156}$$

$$a := iu\begin{bmatrix} \sigma_v(t,T_{i-1})\Big[C(T_{i-1})+D(T_{i-1})\mu_v(t,T_{i-1})\Big] \\ \sigma_n^1(t,T_{i-1})B_n^1(T_{i-1},T_i) \\ \vdots \\ \sigma_n^K(t,T_{i-1})B_n^K(T_{i-1},T_i) \\ -\sigma_r^1(t,T_{i-1})B_r^1(T_{i-1},T_i) \\ \vdots \\ -\sigma_r^M(t,T_{i-1})B_r^M(T_{i-1},T_i) \\ \sigma_V^1(t,T_{i-1})B_r^1(T_{i-1},T_i) \\ \vdots \\ \sigma_V^M(t,T_{i-1})B_r^M(T_{i-1},T_i) \end{bmatrix}, \tag{4.157}$$

and with the $(1+K+2M)\times(1+K+2M)$ correlation matrix S given by

$$S := \begin{pmatrix} 1 & \rho_{x_n^1,v}(t,T_{i-1}) & \dots & \rho_{V^M,v}(t,T_{i-1}) \\ \rho_{x_n^1,v}(t,T_{i-1}) & 1 & \dots & \rho_{x_n^1,V^M}(t,T_{i-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{V^M,v}(t,T_{i-1}) & \rho_{x_n^1,V^M}(t,T_{i-1}) & \dots & 1 \end{pmatrix}. \tag{4.158}$$

Constants in proposition 4.4.6

The constant h_0 and vector h and correlation matrix S_R be extracted from equation (4.96) and are given by:

$$\begin{aligned} h_0 &:= [A_r(T_{i-1},T_i) - A_n(T_{i-1},T_i)] \\ &\quad + \sum_{k=1}^{K} B_n^k(T_{i-1},T_i)\mu_n^k(t,T_{i-1}) - \sum_{j=1}^{M} B_r^j(T_{i-1},T_i)\mu_r^j(t,T_{i-1}), \end{aligned} \tag{4.159}$$

$$h := \begin{bmatrix} \sigma_n^1(t,T_{i-1})B_n^1(T_{i-1},T_i) \\ \dots \\ \sigma_n^K(t,T_{i-1})B_n^K(T_{i-1},T_i) \\ -\sigma_r^1(t,T_{i-1})B_r^1(T_{i-1},T_i) \\ \dots \\ -\sigma_r^M(t,T_{i-1})B_r^M(T_{i-1},T_i) \end{bmatrix}, \tag{4.160}$$

with $(K+M)\times(K+M)$ correlation matrix S_R given by

$$S_R := \begin{pmatrix} 1 & \dots & \rho_{x_n^1,x_n^K}(t,T_{i-1}) & \rho_{x_n^1,x_r^1}(t,T_{i-1}) & \dots & \rho_{x_n^1,x_r^M}(t,T_{i-1}) \\ \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ \rho_{x_n^1,x_n^K}(t,T_{i-1}) & \dots & 1 & \rho_{x_n^K,x_r^1}(t,T_{i-1}) & \dots & \rho_{x_n^K,x_r^M}(t,T_{i-1}) \\ \rho_{x_n^1,x_r^1}(t,T_{i-1}) & \dots & \rho_{x_n^K,x_r^1}(t,T_{i-1}) & 1 & \dots & \rho_{x_r^1,x_r^M}(t,T_{i-1}) \\ \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ \rho_{x_n^1,x_r^M}(t,T_{i-1}) & \dots & \rho_{x_n^K,x_r^M}(t,T_{i-1}) & \rho_{x_r^1,x_r^M}(t,T_{i-1}) & \dots & 1 \end{pmatrix}. \tag{4.161}$$

4.7.3 FX Calibration

We briefly describe the used FX market data set, which can be found in Piterbarg (2005): the set consists of ten maturities, each with seven strikes. The strikes are computed according to formula (4.104). These strikes and corresponding Black and Scholes (1973) implied volatilities can be found in Table **4** and **5** below. Note in Table **5** the increasing term structure of implied volatility and the pronounced implied volatility skew/smile, which both do not die out for long-maturities.

	strike 1	strike 2	strike 3	strike 4	strike 5	strike 6	strike 7
0.5	93.03	96.38	99.84	103.44	107.16	111.02	115.01
1	87.70	92.20	96.93	101.90	107.12	112.61	118.39
3	74.01	80.70	88.00	95.96	104.64	114.11	124.43
5	64.62	72.27	80.81	90.37	101.06	113.02	126.39
7	57.23	65.33	74.57	85.11	97.15	110.89	126.57
10	48.41	56.70	66.41	77.79	91.11	106.72	125.00
15	37.45	45.45	55.16	66.95	81.26	98.62	119.69
20	29.46	36.85	46.08	57.63	72.06	90.12	112.71
25	23.43	30.08	38.63	49.60	63.69	81.77	105.00
30	18.77	24.69	32.46	42.69	56.14	73.82	97.08

Table 4: Strikes.

	strike 1	strike 2	strike 3	strike 4	strike5	strike 6	strike 7
0.5	11.41%	10.49%	9.66%	9.02%	8.72%	8.66%	8.68%
1	12.23%	10.98%	9.82%	8.95%	8.59%	8.59%	8.65%
3	12.94%	11.35%	9.89%	8.78%	8.34%	8.36%	8.46%
5	13.44%	11.84%	10.38%	9.27%	8.76%	8.71%	8.83%
7	14.29%	12.68%	11.23%	10.12%	9.52%	9.37%	9.43%
10	16.43%	14.79%	13.34%	12.18%	11.43%	11.07%	10.99%
15	20.93%	19.13%	17.56%	16.27%	15.29%	14.65%	14.29%
20	22.96%	21.19%	19.68%	18.44%	17.50%	16.84%	16.46%
25	23.97%	22.31%	20.92%	19.80%	18.95%	18.37%	18.02%
30	25.09%	23.48%	22.17%	21.13%	20.35%	19.81%	19.48%

Table 5: Market implied vols.

We then report the detailed calibration results of the model (4.8 to the above market data.
In Table **6** and **7**, we report the calibration differences, in implied volatilities for the model (4.8), respectively with Schöbel and Zhu (1999) and Heston (1993) stochastic volatility. For an analysis of these results, see Section 4.5.2.

	strike 1	strike 2	strike 3	strike 4	strike5	strike 6	strike 7
0.5	-0.11%	0.00%	0.03%	-0.02%	-0.09%	0.00%	0.28%
1	-0.18%	0.00%	0.08%	0.00%	-0.18%	-0.14%	0.22%
3	-0.47%	0.00%	0.29%	0.30%	0.00%	-0.16%	0.02%
5	-0.42%	0.00%	0.25%	0.27%	0.04%	-0.12%	0.00%
7	-0.74%	0.00%	0.57%	0.80%	0.56%	-0.07%	-0.81%
10	-0.64%	0.00%	0.44%	0.62%	0.46%	0.00%	-0.61%
15	-0.45%	0.00%	0.22%	0.22%	0.02%	-0.35%	-0.82%
20	-0.83%	-0.27%	0.07%	0.22%	0.18%	0.00%	-0.33%
25	-1.07%	-0.44%	0.00%	0.26%	0.34%	0.26%	0.04%
30	-1.29%	-0.53%	0.00%	0.35%	0.51%	0.52%	0.39%

Table 6: Differences, in implied Black volatilities, between market and model values using Schöbel-Zhu stochastic volatility.

	strike 1	strike 2	strike 3	strike 4	strike5	strike 6	strike 7
0.5	-0.15%	0.00%	0.05%	0.00%	-0.11%	-0.05%	0.19%
1	-0.20%	0.00%	0.09%	0.00%	-0.13%	0.00%	0.41%
3	-0.25%	0.00%	0.09%	0.00%	-0.18%	-0.06%	0.43%
5	-0.19%	0.00%	0.07%	0.00%	-0.19%	-0.23%	0.02%
7	-0.11%	0.00%	0.03%	0.00%	-0.09%	-0.14%	0.05%
10	-0.04%	0.01%	0.01%	0.01%	-0.02%	-0.05%	0.00%
15	-0.01%	0.01%	0.00%	-0.01%	0.00%	-0.01%	0.00%
20	-0.30%	0.00%	0.13%	0.13%	0.00%	-0.24%	-0.60%
25	-0.29%	0.00%	0.16%	0.21%	0.15%	-0.03%	-0.30%
30	-0.72%	-0.28%	0.00%	0.16%	0.20%	0.11%	-0.07%

Table 7: Differences, in implied Black volatilities, between market and model values using Heston stochastic volatility.

Part II

Simulation Methods for Valuing Exotic Derivatives

Chapter 5

Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model

*This chapter is based on:

Alexander van Haastrecht and Antoon Pelsser (2010). "Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model", *International Journal of Theoretical and Applied Finance, vol. 13, no. 1, pp. 1-43.*.

5.1 Introduction

The behaviour of financial derivatives is usually modeled by stochastic differential equations that (jointly) describe the movements of the underlying financial assets such as the stock prices, stock variances, interest rates or currencies. Though some models yield closed-form solutions for certain derivatives, the great majority of the exotic options cannot be priced in closed-form. Especially for (forward) path-dependent options, the Monte Carlo approach yields a popular and flexible pricing alternative. Because of the increasingly computational power combined with the use of modern day variance reduction techniques, Monte Carlo techniques are expected to become even more widely applicable in the near future.
Since the introduction of the Black and Scholes (1973) model and in particular since the equity crash of the late eighties a battery of complex models has been proposed to relax some misspecifications of the model. Though the Black and Scholes (1973) model has theoretical and practical appealing properties, most of its assumptions, like constant volatility or constant interest rates, do not find justification in the financial markets; one class of models relaxes the constant volatility assumption and incorporates a financial phenomena know as volatility clustering, i.e. they make volatility stochastic. Within this class are the stochastic volatility models of Hull and White (1987), the Stein and Stein (1991) and the Schöbel and Zhu (1999) model. However by far the most popular stochastic volatility model is the Heston (1993) model, mainly caused by the fact that this model, until the introduction of the Schöbel and Zhu (1999) model, was the only stochastic volatility model that allowed for flexibility over the leverage effect, yet

also yielded a closed-form solution for call/put options in terms of one numerical integral[1]. With such a closed-form solution the computation of vanilla European option prices can be done in a fast and stable fashion, hence allowing for efficient calibrations to market option data.

Literature review

Despite the fact that the Heston model was already introduced in 1993, there has been relatively little research on efficient discretization methods of its continuous time dynamics. This is in particularly remarkable if one considers that most practical applications of such models, e.g. the pricing and hedging of path-dependent derivatives, practically always involve Monte carlo methods. Only recently a few papers on efficient discretization methods appeared.
A bias-free (discretization) method was introduced in Broadie and Kaya (2006), who developed a scheme that could simulate the Heston process (i.e. stock and variance) from its exact distribution. Though the paper is elegant, its practical use is limited: first of all, the algorithm requires Fourier inversion of the conditional characteristic function of the integrated variance process. Next to the fact the inversion is time-consuming, it is also complex and can lead to numerical errors (e.g. truncation). Secondly, the variance process has to be simulated using an acceptance and rejection technique, which will scramble random paths when parameters are perturbed, resulting in a low correlation in pre- and post perturbation paths and hence introduces a large Monte Carlo bias in sensitivity analysis (e.g. see Glasserman (2003)). For the same reason also low-discrepancy numbers cannot not be applied in conjunction with the Broadie and Kaya (2006) (BK) scheme.
Several Euler schemes are considered in Lord et al. (2008), in particular they investigate how to deal with negative values of the variance process that occur when one uses a direct discretization. The fix that empirically seems to work best is known as the Full Truncation (FT) scheme. Though the fix is highly heuristic and uses no known analytical properties of the variance process, the scheme seems to work surprisingly well in comparison to other more complicated schemes. Nonetheless it should be noted that for many relevant parameter configurations[2] the discretization error is still quite high for a practical number of time steps. For practical cases, the discretization grid therefore still needs to be rather small to obtain an accurate scheme without significant bias. Approximations to the exact schemes are considered in Smith (2008) and Andersen (2008). Smith (2008) approximates the Fourier inversions required to simulate the integrated variance process, where Andersen (2008) focuses on the variance process and develops two efficient schemes based on moment-matching techniques.
Though we are aware of the fact that the schemes presented so far certainly do not contain a comprehensive list of the all the available schemes, we feel that the schemes mentioned so far

[1]The method of the original Heston paper required the calculation of two numerical integrals, whereas some more recent methods require only the evaluation of one numerical integral, e.g. see Carr and Madan (1999), Lord and Kahl (2008) or Lee (2004).

[2]Heston models which are calibrated to main derivative markets, usually have parameter configurations such that variance process has a relatively high probability of reaching the origin. This is often needed to incorporate the level of skew or kurtosis that is present in market option prices.

stand out for particular reasons: the BK scheme for its exactness, the first order Euler scheme with FT fix for its simplicity and the QE-M scheme for its efficiency. For alternative schemes we refer to the log-normal moment-matching schemes of Andersen and Brotherton-Ratcliffe (2001), and for second and third order schemes to Alfonsi (2005), Alfonsi (2010), Glasserman (2003) and the references therein.
As noted by Broadie, one way to overcome the acceptance and rejection sampling method of the BK scheme (which seems to be the most important bottleneck of the scheme) would be to use a direct inversion of the non-central chi-squared distribution to generate a sample of the variance process. Unfortunately, however, no analytical expression exists for this inverse and one has to use a (time-consuming) root finding procedure to numerically invert the distribution. As direct inversion is too slow for practical purposes, Broadie and Andersen note that another solution could be to create a large three-dimensional[3] cache of the inverse from the non-central chi-squared distribution function for all conceivable values of the number of degrees of freedom and the non-centrality parameters. However, as the parameter-space is potentially very large, Andersen (2008) comments that such 'brute-force' caching would have its own challenges (e.g. the dimensioning of the cache and design of inter- and extrapolation rules). Using a conditioning argument, we will however show that the three-dimensional inverse of the Non-central Chi-squared distribution can effectively be reduced to a one dimensional search space for the case of the Heston (1993) model.
The setup of this chapter as follows. In Section 5.2 we first discuss Euler and Milstein schemes after which we explain the exact scheme of Broadie and Kaya (2006) as well as its bottlenecks. In Section 5.3 we discuss the projection of the three-dimensional inverse of non-central chi-squared distribution onto a one dimensional cache for the case of the Heston (1993) model and we go into the details on how to apply the caching technique as efficiently as possible. This analysis hence results in three new schemes (NCI, NCI-QE and BK-DI), for which we derive martingale properties and regularity conditions of the discretized asset price in Section 5.4. In Section 5.5 we perform an extensive numerical study involving these new schemes and the exact scheme of Broadie and Kaya (2006), the almost exact scheme of Smith (2008), the Kahl-Jäckel scheme, the Full Truncation scheme of Lord et al. (2008) and the Quadratic Exponential scheme of Andersen (2008). To strengthen this numerical analysis we use four different test cases (including European-style and path-dependent options) and a high number of sample paths in conjunction with a variance reduction technique, which enables us to obtain highly accurate results. Using this setup, we are able to make a comprehensive (and differentiated) numerical comparison of the efficiency of the considered schemes, leading to the conclusion of the chapter in Section 5.6.

[3] In principle a three-dimensional cache is needed as the inverse of the non-central chi-squared distribution has three variables: its functions value, the number of degrees of freedom and the non-centrality parameter.

5.2 Heston simulation schemes: Euler, Milstein and exact method

To be clear about notations, we shortly formulate the dynamics of the Heston (1993) model

$$\frac{dS(t)}{S(t)} = r(t)dt + \sqrt{v(t)}dW^S(t), \qquad S(0) := S_0 \geq 0, \tag{5.1}$$

$$dv(t) = \kappa(\theta - v(t))dt + \xi\sqrt{v(t)}dW^V(t), \qquad v(0) := v_0 \geq 0, \tag{5.2}$$

with (W_1, W_2) a two-dimensional Brownian motion under the risk-neutral measure Q with instantaneous correlation ρ, i.e.

$$dW^S(t)dW^V(t) = \rho dt. \tag{5.3}$$

The model parameters are the initial variance $v_0 > 0$, the long run variance $\theta \geq 0$, the mean reversion rate $\kappa \geq 0$, the volatility of the variance $\xi \geq 0$ and the leverage parameter $|\rho_{sv}| \leq 1$. Typically, one finds $-1.0 < \rho_{sv} < -0.6$ implying that the Heston dynamics correlate a down move in the stock with an up move in the volatility (a phenomenon known as leverage effect). For simplicity we here assume that $r(t)$ is constant, hence from now on we will write $r(t) \equiv r$. Since the characteristic function of the log-asset price is known in closed-form for the Heston model, the calibration to vanilla call options can be done efficiently using Fourier inversion, e.g. see Carr and Madan (1999). Please note that in the literature there exists two (theoretically equivalent) formulations of the Heston characteristic function, however as shown in Albrecher et al. (2005) one formulation (as in Heston (1993)) leads to a numerical difficulty called branch cutting, while the other formulation does not have this problem.

5.2.1 Analytical properties of the variance process

The square root variance dynamics of the Heston was first introduced in a finance (i.e. interest rates) context in Cox et al. (1985); there exist several analytical results for the Feller/CIR/square-root process of (5.2), for example the variance process is guaranteed to always be greater or equal to zero. Specifically, if $2\kappa\theta \geq \xi^2$ the Feller condition states that the process can never reach zero (a condition which is however hardly ever satisfied in calibrations to real market data) and for $2\kappa\theta < \xi^2$ we have that the origin is accessible and strongly reflecting.
The distribution of the variance process is also known; conditional on $v(s)$ $(s < t)$, we have that the variance process is distributed as a constant C_0 times a non-central chi-squared distribution with d degrees of freedom and non-centrality parameter λ, i.e.

$$\mathbb{P}(v(t) \leq x | v(s)) = F_{\chi^2_d(\lambda)}\left(\frac{x}{C_0}\right), \tag{5.4}$$

where $F_{\chi^2_d(\lambda)}(\frac{x}{C_0})$ represents the cumulative distribution of the non-central chi-squared distribution, i.e.

$$F_{\chi^2_d(\lambda)}(z) = \sum_{i=0}^{\infty} \frac{e^{-\frac{\lambda}{2}}(\frac{\lambda}{2})^i}{i!} \frac{\int_0^z z^{\frac{d}{2}} e^{-\frac{u}{2}} du}{\Gamma(i+\frac{k}{2})}, \tag{5.5}$$

with

$$C_0 := \frac{\xi^2(1-e^{-\kappa\Delta t})}{4\kappa}, \quad d := \frac{4\kappa\theta}{\xi^2}, \quad \lambda := \frac{4\kappa e^{-\kappa\Delta t} v(s)}{\xi^2(1-e^{-\kappa\Delta t})} \quad \text{and: } \Delta t := t-s. \tag{5.6}$$

Hence note from (5.5) that the non-central chi-squared distribution is equivalent to an ordinary chi-squared G with $d+2N$ degrees of freedom, where N is a Poisson-distribution with mean $\frac{1}{2}\lambda$. The cumulative distribution of (5.5) thus can be written in the following form

$$F_{\chi^2_d(\lambda)}(z) = \sum_{i=0}^{\infty} \mathbb{P}(N=i) G_{\chi^2}(z, d+2i), \tag{5.7}$$

which will be an important expression in the remainder of this chapter. From known properties of the non-central chi-squared distribution (e.g. see Cox et al. (1985) or Abramowitz and Stegun (1964)) we then have that the mean m and variance s^2 of $v(t)$ conditional on $v(s)$ are given by

$$m := \theta + (v(s)-\theta)e^{-\kappa\Delta t}, \tag{5.8}$$

$$s^2 := \frac{v(s)\xi^2 e^{-\kappa\Delta t}}{\kappa}(1-e^{-\kappa\Delta t}) + \frac{\theta\xi^2}{2\kappa}(1-e^{-\kappa\Delta t})^2. \tag{5.9}$$

While some discretization schemes of the Heston dynamics heavily rely on these properties (e.g. see Broadie and Kaya (2006),Andersen (2008) and Smith (2008)), other schemes do not incorporate the specific distributional properties (e.g. see the Euler and Milstein schemes of Lord et al. (2008) and Kahl and Jäckel (2006)).

5.2.2 (Log-)Euler scheme

Probably the simplest way to discretize the variance dynamics is by using a first-order Euler scheme. One should however take care on how to fix negative values of the variance process; handling negative values in the wrong way leads to extremely biased schemes, whereas using the right fix leads to an Euler scheme that outperforms almost all existing schemes in terms of computational efficiency, e.g. see Lord et al. (2008). Since not all literature sources use the proper fix when comparing their scheme with an Euler scheme and the scheme provides a good intuition behind the difficulties of the simulation of the Heston model, we explicitly discuss the Euler scheme here.

Starting from time s a naive Euler discretization of the variance process for $t > s$ (with $\Delta t := t-s$) reads

$$v(t) = v(s) + \kappa\Delta t(\theta - v(s)) + \xi\sqrt{v(s)}Z_v\sqrt{\Delta t}, \tag{5.10}$$

with Z_v standard normal distributed. The main source of difficulty in above scheme is that the

variance can become negative, explicitly the probability of $v(t)$ becoming negative is

$$\mathbb{P}(v(t) < 0) = \mathbb{P}\Big(Z_v < \frac{-v(s) - \kappa\Delta t(\theta - v(s))}{\xi\sqrt{v(s)}\sqrt{\Delta t}}\Big) = \Phi\Big(\frac{-v(s) - \kappa\Delta t(\theta - v(s))}{\xi\sqrt{v(s)\,\Delta t}}\Big). \tag{5.11}$$

Notice that though this probability decays to zero as Δt becomes smaller, it will be strictly positive for any choice of the time step Δt (unless $\xi = 0$). Hence if one does not want the root of the variance to cross over to the imaginary domain, one has to decide what to do if the variance process turns negative in an Euler scheme. Several ad-hoc fixes for this exist in the literature, for example by making zero an absorbing or reflecting boundary for the variance process. Lord et al. (2008) unify several Euler schemes in the following framework:

$$v(t) = f_1(v(s)) + \kappa\Delta t(\theta - f_2(v(s))) + \xi\sqrt{f_3(v(s))}\,Z_v\sqrt{\Delta t}, \tag{5.12}$$

where all schemes should satisfy $f_i(x) = x$ for $x \geq 0$ and $f_3(x) \geq 0$ for all $x \in \mathbb{R}$. This translates into the natural conditions that for positive values of the variance the regular Euler scheme should be employed, and that strictly negative values of the discretized variance process should be transformed into positive ones. The most sensible choices for $f_i(x)$ are the identity function ($f(x) = x$), absorption ($f(x) = x^+$) or reflection ($f(x) = |x|$). Since all schemes coincide and are bias-free as $\Delta t \to 0$, the choice of the fix seems innocent and almost indifferent. The contrary is true: while some schemes are extremely biased for practical sizes of the time step, others turn out to be almost bias-free not too extreme parameter configurations. The fix that seems to work the best is produced by the so-called Full Truncation (2007) scheme and chooses $f_1(x) := x$, $f_2(x) = f_3(x) := \max(x, 0) = x^+$, see Lord et al. (2008). The resulting Euler scheme reads

$$v(t) = v(s) + \kappa\Delta t(\theta - v(s)^+) + \xi\sqrt{v(s)^+}\,Z_V\sqrt{\Delta t}\,. \tag{5.13}$$

Hence provided with a discretization scheme for the variance process, we also need to specify the simulation schemes of the corresponding asset price process. The most straightforward choices would be to either directly apply an Euler discretization scheme to the stock price process of equation (5.1) or to simulate the stock price from its exact (conditional) distribution. Direct discretization yields the following **Euler scheme**

$$S(t) = S(s)\Big(1 + r\Delta t + \sqrt{f_5(v(t))}\,Z_S\sqrt{\Delta t}\Big) \tag{5.14}$$

and does entails some discretization error of the exact process. Here Z_S is a normal distributed random variable (with correlation ρ to Z_V) and $f_5(x)$ should be chosen non-negative.

Alternatively one can also use the exact solution of the stock price dynamics (5.1), which by an application of Itô's lemma is given by

$$S(t) = S(s)\exp\Big[\int_s^t [r - \frac{1}{2}v(u)]du + \int_s^t \sqrt{v(u)}\,dW_S(u)\Big]. \tag{5.15}$$

Hence taking logarithms and discretising in an Eulerly fashion, one obtains the following **log-Euler scheme**

$$\log(S(t)) = \log(S(s)) + \left[r - \frac{1}{2} f_4(v(s))\right]\Delta t + \sqrt{f_5(v(s))}\, Z_S \sqrt{\Delta t}. \tag{5.16}$$

The above described log-Euler scheme does not entail any discretization error in the stock price direction, of course the scheme usually does show biases in the Euler discretization the variance process (and thus in resulting stock prices). Following Lord et al. (2008) we choose to set $f_3(x) = f_4(x) = f_5(x) = x^+$, which seem to be the most logical choices, since the Ito correction term of equation (5.16) is then consistent with the corresponding the volatility of the stock price, hence this implies the martingale condition of the stock price process is preserved in the discretization. In an implementation the correlated standard normal random variables Z_V and Z_S can (for example) be generated with the use of a Cholesky decomposition: with a (instantaneous) correlation of ρ between the driving Brownian motions this can be done by setting $Z_V := Z_1$ and $Z_S := \rho Z_V + \sqrt{1-\rho^2}\, Z_2$, where Z_1 and Z_2 are two independent draws from the standard normal distribution.

Note that the pure Euler scheme (5.14) can be seen as a first order approximation of above log-Euler scheme. Since the log-Euler scheme entails no discretization error in the stock price direction, we prefer to work under this log transformation when employing an Euler scheme, e.g. see also Lord et al. (2008)). Additionally since the full truncation scheme seems to have the smallest bias among all Euler schemes, we adopt this fix for possible negative values of the variance process when using an Euler scheme. The main advantage of the Euler scheme lies in its simplicity and speed: little code and computing time is needed to compute one iteration in the scheme. Additionally the use of the scheme is not restricted to the Heston model, but can also be applied to all kind models, for example to the family of CEV-processes of Lord et al. (2008). Its generality also implies its weakness: the Euler scheme doesn't use any information of known analytical properties of the square root variance process.

Full truncation algorithm

Using a log-Euler scheme for the stock price process, the full truncation scheme for the Heston can be summarized by the following algorithm:

1. Generate a random sample Z_1 from the standard normal distribution[4] and set $Z_v := Z_1$.
2. Given $v(s)$, compute $v(t)$ from equation (5.13).
3. Generate a random sample Z_2 from the standard normal distribution and set

$$Z_S := \rho_{\text{sv}} Z_V + \sqrt{1-\rho_{\text{sv}}^2}\, Z_2. \tag{5.17}$$

[4]It may be advisable to use an inversion method for generating of normal samples, since then also a quasi random generator can be used. This inversion over an uniform random variable with the ('approximated') inverse standard normal distribution function can for example be done using Wichura's method, see Wichura (1998).

4. Given $\log(S(s))$, compute $\log(S(t))$ using equation (5.16).

5.2.3 Kahl-Jäckel Scheme

A generic implicit Milstein scheme for the variance process in combination with an alternative discretization for the stock price was suggested in Kahl and Jäckel (2006), i.e. the following discretization scheme was proposed:

$$v(t) = \frac{v(s) + \kappa\theta\Delta t + \xi\sqrt{v(s)}Z_V\sqrt{\Delta t} + \frac{1}{4}\xi^2\Delta t(Z_V^2 - 1)}{1 + \kappa\Delta t} \tag{5.18}$$

$$\begin{aligned}\log(S(t)) = {} & \log(S(s)) + \Big[r - \frac{(v(s) + v(t))}{4}\Big]\Delta t + \rho\sqrt{v(s)}Z_V\sqrt{\Delta t} \\ & + \frac{1}{2}(\sqrt{v(s)} + \sqrt{v(t)})(Z_S + \rho Z_V)\sqrt{\Delta t} + \frac{\rho\xi\Delta t}{2}(Z_V^2 - 1)\end{aligned} \tag{5.19}$$

Kahl-Jäckel show that this scheme results in positive paths for the variance process for $2\kappa\theta > \xi^2$, a condition which is hard to meet in practice. Hence in many practical implementations of the above dynamics, one has to decide on how to fix negative values of the variance process. Since Kahl and Jäckel (2006) do not specifically tackle this issue, we follow Andersen (2008) who adopts the same fix as Lord et al. (2008) use in the full truncation Euler scheme. That is, whenever the variance process drops below zero, we use (5.13) rather than (5.18) and take $v(s)^+$ and $v(t)^+$ rather than using $v(s)$ and $v(t)$ in (5.19). The resulting algorithm is similar to the ft-algorithm 5.2.2: one just replaces the variance and asset process from (2) and (4) with the above defined discretizations for the variance and asset process.

5.2.4 Exact scheme of Broadie and Kaya

In an elegant paper, Broadie and Kaya (2006) worked out an exact simulation scheme for the Heston model. Though theoretically the method is exact, its practical use is limited; the scheme suffers from a lack of speed, it is complex and sensitivity calculations (often used for risk management) are hard since the scheme relies on acceptance and rejection sampling techniques. For example, the numerical tests in Lord et al. (2008) show that for most practical situations even a simple Euler scheme outperforms the exact scheme in terms of computational efficiency[5]. In most practical situations a direct implementation of the exact scheme is probably not the best available option (see section 5.2.5). There are however some approximations or computational

[5]Note that in the numerical results of Broadie and Kaya (2006) and Smith (2008), an Euler scheme is used that handles negative values of the variance in a suboptimal way. However as shown in Lord et al. (2008) the choice on how to cope with negative values of the variance process is extremely important for the quality (i.e. bias) of the scheme. Because the (semi-)exact schemes in Broadie and Kaya (2006) and Smith (2008) are benchmarked against a suboptimal Euler scheme, this leads them to a false conclusion in comparing their schemes against 'the' Euler scheme. This was point was first noted in Lord et al. (2008) and later on in Andersen (2008). From their results in can for example be seen that the Euler scheme (with the 'right' fix) outperforms the exact and Kahl-Jäckel scheme in terms of computational efficiency, whereas in Broadie and Kaya (2006) and in Smith (2008) (who use suboptimal fixes) this is just the other way around.

tricks that can be made to improve upon the computational efficiency. For example, Andersen (2008) and Smith (2008) both use the exact scheme as starting point and from there on try to improve upon some of the incorporated bottlenecks.

We will first discuss the exact method and its incorporated difficulties: by using the explicit solution (5.15) of the asset price process and consecutively using Itô's lemma and using a Cholesky decomposition one obtains

$$\begin{aligned}\log(S(t)) &= \log(S(s)) - \frac{1}{2}\int_s^t v(u)du \\ &\quad +\rho\int_s^t \sqrt{v(u)}\,dW_v(u) + \sqrt{1-\rho^2}\int_s^t \sqrt{v(u)}\,dW(u),\end{aligned} \tag{5.20}$$

where $W(u)$ is a Brownian motion independent of $W_v(u)$. Integrating the square-root variance process of equation (5.2) gives the following solution:

$$v(t) = v(s) + \int_s^t \kappa(\theta - v(u))du + \xi\int_s^t \sqrt{v(u)}\,dW_v(u), \tag{5.21}$$

or equivalently

$$\int_s^t \sqrt{v(u)}\,dW_v(u) = \xi^{-1}\big(v(t) - v(s) - \kappa\theta\Delta t + \kappa\int_s^t v(u)du\big). \tag{5.22}$$

In Broadie and Kaya (2006), it is then noticed that one can substitute equation (5.22) into the solution (5.20) to arrive at

$$\begin{aligned}\log(S(t)) &= \log(S(s)) + \frac{\kappa\rho}{\xi}\int_s^t v(u)du - \frac{1}{2}\int_s^t v(u)du \\ &\quad +\frac{\rho}{\xi}\big(v(t) - v(s) - \kappa\theta\Delta t\big) + \sqrt{1-\rho^2}\int_s^t \sqrt{v(u)}\,dW(u),\end{aligned} \tag{5.23}$$

hence an exact simulation involves sampling from the following three stochastic quantities:

1. $v(t)|v(s)$: from (5.4) and (5.6) we have that $v(t)|v(s)$ is distributed as a constant C_0 times a non-central chi-squared distribution with d degrees of freedom and non-centrality parameter λ.

2. $\int_s^t v(u)du|v(s), v(t)$: Broadie and Kaya (2006) derive the characteristic function

$$\begin{aligned}\Psi\Big(a, v(s), v(t)\Big) &= \mathbb{E}\Big[\exp(ia\int_s^t v(u)du\Big)\Big|v(s), v(t)] \\ &= \frac{\gamma(a)e^{\frac{1}{2}(\gamma(a)-\kappa)(t-s))}(1-\exp(-\kappa(t-s))}{\kappa(1-e^{-\gamma(a)(t-s)})} \\ &\quad \times \exp\Big[\frac{v(s)+v(t)}{\xi^2}\Big(\frac{\kappa(1+e^{-\kappa(t-s)}}{1-e^{-\kappa(t-s)}} - \frac{\gamma(a)(1+e^{-\gamma(a)(t-s)})}{1-e^{-\gamma(a)(t-s)}}\Big)\Big] \\ &\quad \times \frac{I_{\frac{1}{2}d-1}\Big[\sqrt{v(s)v(t)}4\gamma(a)e^{-\frac{\gamma(a)}{2}(t-s)}/\xi^2(1-e^{-\gamma(a)(t-s)})\Big]}{I_{\frac{1}{2}d-1}\Big[\sqrt{v(s)v(t)}4\gamma(a)e^{-\frac{\kappa}{2}(t-s)}/\xi^2(1-e^{-\kappa(t-s)})\Big]}, \end{aligned} \tag{5.24}$$

with $\gamma(a) := \sqrt{\kappa^2 - 2\xi^2 ia}$, $d := \frac{4\kappa\theta}{\xi^2}$ and where $I_\nu(x)$ is the modified Bessel function of the first kind. Hence the characteristic function (5.24) can numerically be inverted to obtain the value of the distribution function $G(x)$ for a certain point $x \in \Omega$, i.e.

$$G(x, v(s), v(t)) = \frac{2}{\pi}\int_0^\infty \frac{\sin(ax)}{x}\mathrm{Re}\Big[\Psi(a, v(s), v(t))\Big]da. \tag{5.25}$$

Finally to generate sample from $\int_s^t v(u)du|v(s), v(t)$ one can use

$$G(\int_s^t v(u)du|v(s), v(t)) = U, \tag{5.26}$$

and invert G over a uniform random sample U to find $x_i : x_i = G^{-1}(U, v(s), v(t))$, e.g. by a Newton-Raphson root search of $G(x_i, v(s), v(t)) - U = 0$. Note that such a root finding procedure involves multiple Fourier inversions: one for each evaluation of $G(x_i, v(s), v(t))$.

3. $\int_s^t \sqrt{v(u)}dW(u)|\int_s^t v(u)du$: since $v(u)$ is independent of $W(u)$, it directly follows that the this expression is distributed as $N(0, \int_s^t v(u))$. Hence this sampling can be done easily and efficient by sampling from a normal distribution.

Broadie and Kaya algorithm

Exact simulation of (5.23) is feasible and can be performed by the following algorithm:

1. Conditional on $v(s)$, use the definitions of (5.6) to generate a sample of $v(t)$ by sampling from a constant times a non-central chi-squared distribution with d degrees of freedom and non-centrality parameter λ.

2. Conditional on $v(s)$ and $v(t)$, generate a sample of $\int_s^t v(u)du$ by a numerical inversion of the distribution function G of $\Big(\int_s^t v(u)du\,\Big|\,v(s), v(t)\Big)$ over a uniform sample U, for example

by a root search $G(x_i, v(s), v(t)) - U = 0$. Since the distribution function G is not known in closed form, $G(x_i, v(s), v(t))$ has to be obtained by Fourier inversion of the characteristic function of $\int_s^t v(u)du|v(s), v(t)$.

3. Use (5.22) to set:

$$\int_s^t \sqrt{v(u)}\, dW_v(u) = \xi^{-1}(v(t) - v(s) - \kappa\theta\Delta t + \kappa \int_s^t v(u)du) \tag{5.27}$$

4. Generate an independent random sample Z_S from the standard normal distribution and use the fact that $\int_s^t \sqrt{V(u)}\, dW(u)$ is normally distributed with mean zero and variance $\int_s^t V(u)du$ and thus can be generated as

$$\int_s^t \sqrt{V(u)}\, dW(u) \sim Z_S \sqrt{\Delta t \int_s^t V(u)du}\,, \tag{5.28}$$

5. Given $\log(S(s)), \int_s^t \sqrt{v(u)}\, dW_v(u)$, $\int_s^t \sqrt{V(u)}\, dW(u)$ and $\int_s^t V(u)du$ use (5.23) to compute $\log(S(t))$.

5.2.5 Disadvantages of the exact scheme

Though the Broadie and Kaya scheme is theoretically appealing (and this was probably also the primary objective of their paper), we will discuss in the following section why its practical use might be limited. That is, we discuss some practical implementation issues that are incorporated with the use of the exact scheme.

First of all, (5.2.4, step 1) requires that the variance process $v(t)|v(s)$ has to be sampled from a constant C_0 times a non-central chi-squared distribution with d degrees of freedom and non-centrality parameter λ (see (5.4)):

$$v(t) \overset{d}{=} C_0 \chi_d^2(\lambda). \tag{5.29}$$

For simulation purposes one can use the following representations of the non-central chi-squared distribution (see Johnson et al. (1994) and Glasserman (2003)):

$$\chi_d^2(\lambda) \overset{d}{=} \begin{cases} (Z + \sqrt{\lambda})^2 + \chi_{d-1}^2 & \text{for } d > 1, \\ \chi_{d+2N}^2 & \text{for } d > 0, \end{cases} \tag{5.30}$$

with $Z \sim N(0,1)$, χ_ν^2 an ordinary chi-squared distribution with ν degrees of freedom and where N is Poisson distributed with mean $\mu := \frac{1}{2}\lambda$. Since in most practical applications $d << 1$, one is usually forced to work with the second representation of the non central chi-squared distribution[6];

[6]Otherwise, if $d > 1$, one might want to use the first representation, since sampling from the normal distribution is usually more efficient than sampling from a Poisson distribution.

hence exact sampling from the variance process can be done by first conditioning on a Poisson variate and consecutively generating a sample from a chi-squared or gamma distribution[7]. Since direct inversion of the gamma distribution is relatively slow, Broadie and Kaya (2006) suggest to use an acceptance and rejection method to generate gamma variates. Though such sampling can be done fairly quick (e.g. by making use of some recent advances of Marsaglia and Tsang (2000)), the methods are still relatively slow in comparison to sampling methods for normal variates.
Moreover the main disadvantage of acceptance and rejection techniques is that the (number of) samples depend on the specific Heston parameters. As a consequence the total drawings of random numbers cannot be predetermined and sample paths will show a rather small correlation coefficient for different parameter inputs. These properties are usually inconvenient in financial applications, since both perturbation analysis[8] (to calculate model sensitivities with respect to different parameters) as well as the use of quasi random numbers generator becomes extremely hard, not to say almost practically impossible.
Another practical difficulty of the scheme lies in step (5.2.4-1), where one has to generate a sample of $\int_s^t v(u)du|v(s), v(t)$ by numerically inverting the distribution function of $(\int_s^t v(u)du|v(s), v(t))$ over an uniform random variable u, by a root search of $G(x_i, v(s), v(t)) - U = 0$. However because the distribution function G is not known is closed form, it has to be obtained by Fourier inverting the characteristic function (5.24), which contains two modified Bessel functions that are both represented by an infinite series. The root-finding procedure (and the involved Fourier inversions) has to repeated several times until a tolerance level ε is reached for a guess x_i, such that $G(x_i, v(s), v(t)) - U < \varepsilon$. Next to the fact that both in the evaluation of (5.24) as well as the required Fourier inversions require a great computational effort, the implementation of this step also has to be done with great care to avoid noticeable biases from the numerical inversions.

5.3 Approximations to the exact scheme

As elaborated in Section 5.2.5 the exact scheme has some practical disadvantages. However it does provide an excellent base to construct approximate schemes which might be computationally more efficient. A few authors have already tried to improve the bottlenecks in simulating the variance and/or integrated variance process, e.g. see Andersen (2008) and Smith (2008). In the remainder of this section we will unify and discuss the two methodologies that can improve upon the performance of the Broadie and Kaya scheme. That is, we consider approximations of:

1. The integrated variance process.

2. The variance process itself.

Moreover we will look at schemes that combine the latter approximations.

[7]The Chi-squared distribution is a special case of the gamma distribution, $\chi^2_\nu \overset{d}{=} \text{gamma}(\frac{\nu}{2}, 2)$, where $\text{gamma}(k, \theta)$ is a gamma distribution with shape k and scale θ.

[8]The efficiency in the calculation of model sensitivities crucially depends on the size of the correlation coefficient between pre- and post perturbation paths.

Approximating the integrated variance distribution

As elaborated in section 5.2.5, a huge bottle neck of the simulation scheme is the sampling of the conditional integrated variance process. There are however several ways to approximate a sample from the integrated variance process $\int_s^t v(u)du \,\big|\, v(s), v(t)$.

1. **Drift interpolation:** Without using any specific information of the integrated variance process, one can use a drift interpolation method to approximate the integrated variance process, i.e.
$$\int_s^t v(u)du \,\big|\, v(s), v(t) \;\approx (\gamma_1 v(s) + \gamma_2 v(t))\,(t-s), \tag{5.31}$$
for some constants γ_1, γ_2, which can be set in several ways: an Euler-like setting would read $\gamma_1 = 1$, $\gamma_2 = 0$ (e.g. see Kloeden and Platen (1999) or Lord et al. (2008)), while a mid-point rule corresponds to the predictor-corrector setting $\gamma_1 = \gamma_2 = \frac{1}{2}$ (e.g. see Hunter et al. (2001) or Andersen (2008)).

2. **Approximate the Fourier inversion:** One can also try to approximate the Fourier inverted sampling of the integrated variance process. For example Smith Smith (2008) tries to speed up the inversion of the characteristic function (5.24) by caching values of a projected version hereof. Though such a method might speed up the inversion, one still has to use a rather time-consuming Fourier inversion combined with a root finding procedure to draw a sample of the integrated process. Alternatively one can try to use the first moments of the conditional integrated variance process (which can be obtained by differentiating the characteristic function of (5.24)) to develop a moment-matched sampling method.

Approximating the variance process

Another (practical) disadvantage of the exact scheme is the use of acceptance rejection sampling method for the non-central chi-squared distributed variance process (see section 5.2.5). Hence we consider two methods that can be used to approximate the variance distribution.

1. **Moment-matching:** Andersen (2008) suggests to approximate the variance process by related distributions whose first two moments are (locally) matched with those of the true variance distribution. Moreover, since the distributions can be analytically inverted, the methods can be directly used in conjunction with perturbation and low-discrepancy methods by straightforward inversions a uniform random variates.

2. **Direct inversion:** To overcome the acceptance and rejection sampling method, one can also use direct inversion of the non-central chi-squared distribution to generate a sample of the variance process. Unfortunately, however, no analytical expression exists for this inverse and one has to use a (time-consuming) root finding procedure to numerically invert the distribution. As direct inversion is too slow for practical purposes, another solution is to design a three-dimensional (i.e. with its functions value, the degrees of freedom and the non-centrality parameters) cache of the inverse from the non-central chi-squared distribution function. This suggestion was initially put forward by Broadie and Andersen (see

Andersen (2008)). However these authors commented that the cache would be of too a high dimension to be practical, hence other directions are undertaken.
In Section 5.3.4 we will however tackle this dimensionality problem and show that the initial three-dimensional parameter space of the inverses of the non-central chi-squared distribution can effectively be projected onto a one dimensional search space for the case of the Heston (1993) model. The key insights behind this projection is discussed in Section 5.3.4, where we also go into the details on how to apply the caching technique as efficiently as possible. The derivation of the exact martingale property and the regularity conditions of the schemes applying a direct inversion can be found in Section 5.4.1.

5.3.1 Broadie and Kaya Drift Interpolation scheme

Probably the easiest way to give the exact scheme a performance boost is to approximate the Fourier inverted sampling of the integrated variance process by the simple drift interpolation method of equation (5.31). Moreover since the sampling of the integrated variance process is most time-consuming step of the exact scheme, one can expect a large efficiency gain. The simulation of the Broadie and Kaya Drift Interpolation (BK-DI) scheme is straightforward; in the exact scheme of 5.2.4, one only has to replace the sampling of the integrated variance process in step 2 by the drift interpolation rule (5.31).
Hence though the resulting method is simple and reasonably efficient, sampling from the variance process is still performed by an acceptance-rejection method, which (as discussed in section 5.2.5) is rather inconvenient for most financial applications. We also like to note that though for reasonable time-spacings the drift approximation error is usually rather small, one does slightly violate the discrete-time no-arbitrage condition, i.e. the discretized stock price is not exactly a martingale. In section 5.4.1 we show how one can enforce this condition with the above discussed discretization method.

5.3.2 Almost Exact Fourier inversion scheme

Smith (2008) tries to speed up the inversion of the characteristic function (5.24) by caching values of a projected version hereof in the almost exact simulation method (AESM). The core of this method to project the exact characteristic function $\Psi(a, v(s), v(t))$, which depends on $v(s)$ and $v(t)$ via the arithmetic and geometric mean $\frac{1}{2}(v(s) + v(t))$ and $\sqrt{v(s)v(t)}$, onto a function $\widetilde{\Psi}(a, z)$ in which the dependency on the means is approximated by the combination

$$z = \omega\frac{1}{2}(v(s) + v(t)) + (1 - \omega)\sqrt{v(s)v(t)}, \quad 0 \leq \omega \leq 1, \tag{5.32}$$

for a suitable choice of ω. Hence the arithmetic and geometric mean, which are similar in expectation, are replaced by a weighted average of the two. In this way the three-dimensional

function $\Psi(a, v(s), v(t))$ is approximated by the two-dimensional function

$$\begin{aligned}\widetilde{\Psi}(a,z) &= \frac{\gamma(a)e^{\frac{1}{2}(\gamma(a)-\kappa)(t-s))}(1-\exp(-\kappa(t-s))}{\kappa(1-e^{-\gamma(a)(t-s)})} \\ &\times\exp\Big[\frac{2z}{\xi^2}\Big(\frac{\kappa(1+e^{-\kappa(t-s)}}{1-e^{-\kappa(t-s)}}-\frac{\gamma(a)(1+e^{-\gamma(a)(t-s)})}{1-e^{-\gamma(a)(t-s)}}\Big)\Big] \\ &\times\frac{I_{\frac{1}{2}d-1}\Big[z4\gamma(a)e^{-\frac{\gamma(a)}{2}(t-s)}/(\xi^2(1-e^{-\gamma(a)(t-s)}))\Big]}{I_{\frac{1}{2}d-1}\Big[z4\gamma(a)e^{-\frac{\kappa}{2}(t-s)}/(\xi^2(1-e^{-\kappa(t-s)}))\Big]},\end{aligned} \tag{5.33}$$

which can then be cached on a sufficiently small discrete (two-dimensional) grid of a and z-points. Though Smith claims that the approximation works well, the implementation still requires a time-consuming root search of Fourier inversions for each time step. Hence though the evaluation of the characteristic can be approximated in an computationally efficient way, the root search and inversion are still rather time-consuming in comparison with a simple drift interpolation method. Additionally the total algorithm has to be implemented with great care to avoid numerical truncation and discretization errors.

5.3.3 Quadratic Exponential scheme

In the Quadratic Exponential (QE) scheme, Andersen (2008) suggests to approximate the sampling from the non-central chi-squared distribution by drawing from a related distribution, which is moment-matched with the first two (conditional) moment of non-central chi-squared distribution. The choice of distribution is split up into two parts, which are based on the following observations with respect to the size of the non-centrality parameter (e.g. see Abramowitz and Stegun (1964)).

1. For a moderate non-centrality parameter the non-central chi-squared can be well represented by a power function applied to a Gaussian variable (which is equivalent to a non-central chi-squared distribution with one degree of freedom). For sufficiently high values of $v(s)$, a sample of $v(t)$ hence can be generated by

$$v(t) = a(b + Z_v)^2, \tag{5.34}$$

   where Z_v is standard normal distributed random variable and a and b are constants to be determined by moment-matching.

2. For sufficiently low values of $v(s)$, the density of $v(t)$ can (asymptotically) be approximated by

$$\mathbb{P}\Big(v(t) \in [x, x+dx]\Big) \approx (p\delta(0) + \beta(1-p)e^{-\beta x})dx, \quad x \geq 0, \tag{5.35}$$

   where δ represents Dirac's delta function, and p and β are non-negative constants.

The constants a,b,p,β can (locally) be chosen such that the first two moments of the approximate distribution matches those of the exact one. These constants depend on the size of the time-step

Δt,$v(s)$, as well as on Heston's model parameters. Sampling from these distributions can be done in a simple and efficient way:

- From the first distribution one only has to draw a standard normal random variable and apply the quadratic transformation of equation (5.34).
- Sampling according to equation (5.35) can be done by inversion of the distribution function; the distribution function is obtained by integrating the probability density function, and can subsequently be inverted to obtain the following inverse distribution function:

$$L^{-1}(u) = \begin{cases} 0 & \text{if } 0 \leq u \leq p, \\ \beta^{-1} \log(\frac{1-p}{1-u}) & \text{if } p < u \leq 1. \end{cases} \tag{5.36}$$

  Using the inverse distribution function sampling method, one obtains an easy and efficient sampling scheme by first generating a uniform random number U_v and then setting

$$v(t) = L^{-1}(U_v) \tag{5.37}$$

Together, the equations (5.34) and (5.37) define the Quadratic Exponential (QE) discretization scheme. What yet remains is the determination of the moment-matching constants a,b,p and β, as well as a rule that defines 'high and low values' of the non-centrality parameter, i.e. a rule that determines when to switch from (5.34) and (5.37). We first discuss the latter.
Recall that the conditional mean and variance of the square-root process are given by m and s^2 as defined in equations (5.8) and (5.9). Andersen bases the switching rule on the value of ψ

$$\psi := \frac{s^2}{m^2} = \frac{\frac{v(s)\xi^2 e^{-\kappa\Delta t}}{\kappa}(1 - e^{-\kappa\Delta t}) + \frac{\theta\xi^2}{2\kappa}(1 - e^{-\kappa\Delta t})^2}{\left(\theta + (v(s) - \theta)e^{-\kappa\Delta t}\right)^2} = \frac{C_1 v(s) + C_2}{\left(C_3 v(s) + C_4\right)^2}, \tag{5.38}$$

with

$$C_1 = \frac{\xi^2 e^{-\kappa\Delta t}}{\kappa}(1 - e^{-\kappa\Delta t}), \qquad C_2 = \frac{\theta\xi^2}{2\kappa}(1 - e^{-\kappa\Delta t})^2, \qquad C_3 = e^{-\kappa\Delta t}, \qquad C_4 = \theta(1 - e^{-\kappa\Delta t}).$$

Note that ψ is inversely related to the size of the non-centrality parameter. It can be shown that for $\psi \leq 2$ the quadratic transformation (5.34) can be moment-matched with the exact distribution and for $\psi \geq 1$ the exponential one of (5.37). Thus for $\psi \leq 2$, we can moment match the quadratic sampling scheme (5.34) and for $\psi \geq 1$ and we can moment match the exponential scheme (5.37). Since these domains overlap, at least one of the two methods is applicable. A natural procedure is then to introduce some critical level $\psi_c \in [1, 2]$, and use (5.34) if $\psi \leq \psi_c$ and (5.37) otherwise. Following Andersen, who notes that the exact choice of ψ_c has a relatively small impact on the quality of the overall simulation scheme, we use $\psi_c = 1.5$ in the numerical tests.
Notice though ψ (locally) has to be calculated for every step in a simulation and contains 'computationally expensive' components (e.g. the exponent $\exp(-\kappa\Delta t)$) many of these terms only depend on the size of time step. From efficiency considerations it is therefore advisable to pre-cache the static constants $C_1, \ldots, C_4$ before the Monte Carlo simulation starts. In the case

one uses a non-equidistant time grid different constants of course need to be cached for every applicable size of the time step.

The moment-matching constants a,b,p and β of the just defined sampling schemes still have to be specified, and should be chosen such that the first two (conditional) moments are matched with the first and second central moment m and s^2 of the exact non-central chi-squared distribution. The following statements hold regarding the conditional moments of the schemes (5.34) and (5.37).

1. For $\psi \leq 2$, setting

$$b^2 = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1} - 1} \geq 0, \tag{5.39}$$

$$a = \frac{m}{1 + b^2}, \tag{5.40}$$

   assures that the first two moments of the sampling scheme (5.34) are matched with the exact moments non-central chi-squared distribution, see Andersen (2008), proposition 5, pp.14.

2. For $\psi \geq 1$, setting

$$p = \frac{\psi - 1}{\psi + 1} \in [0, 1), \tag{5.41}$$

$$\beta = \frac{1 - p}{m} = \frac{2}{m(\psi + 1)} > 0, \tag{5.42}$$

   assures that the first two moments of the sampling scheme (5.37) are matched with the exact moments non-central chi-squared distribution, see Andersen (2008), proposition 6, p.15.

QE Algorithm

Assuming that some critical switching level $\psi_c \in [1, 2]$ and values for $\gamma_1, \gamma_2 \in [0, 1]$ have been selected, the Quadratic Exponential variance sampling can be summarized by the following algorithm:

1. Given $v(s)$, compute m and s^2 and $\psi = \frac{m^2}{s^2}$ using equations (5.8) and (5.9).

2. **If** $\psi \leq \psi_c$:

   (a) Compute a and b from equations (5.40) and (5.39).

   (b) Generate a random sample Z_v from the standard normal distribution.

   (c) Use (5.34), i.e. set $v(t) = a(b + Z_v)^2$.

   Otherwise, if $\psi > \psi_c$:

(a) Compute β and p according to equations (5.41) and (5.42).

(b) Draw a uniform random number U_v.

(c) Use (5.37), i.e. set $v(t) = L^{-1}(U_v)$ where L^{-1} is given in (5.36).

5.3.4 Non-central Chi-squared Inversion scheme

Instead of using moment-matched schemes, another way to circumvent the acceptance and rejection technique is to use a direct inversion of the Non-central Chi-Squared distribution. We will call this new scheme the Non-central Chi-squared Inversion (NCI) scheme; since direct inversion is too slow, another solution could be to design a three-dimensional cache over the parameter space of the inverses from the non-central chi-squared distribution function $F^{-1}(x, d, \lambda)$, which can be created by a root finding procedure of the distribution function. This suggestion was already put forward by Broadie and Andersen (e.g. see Andersen (2008), however these authors commented that the parameter space is of too high dimension (i.e. three) to create a cache. Such a three-dimensional caching, even referred to as 'brute-force' by Andersen (2008), will have its own challenges, like its dimension and the design of inter- and extrapolation rules. For this reason, the idea of the cache is not pursued further.
In the following, however, we explain an important insight which enables us to project the three-dimensional parameter space onto an effectively one dimensional search space. As a one-dimensional cache is of such a low dimension, the computational overhead of its creation is small, whilst the simulation of the variance process can be done fast and by simple interpolation method of two values of the cache over an uniform random variable. Moreover, as the total number of uniform draws is fixed (and independent of the Heston parameters), the caching technique can be applied directly in conjunction with perturbation analysis and low-discrepancy numbers, which are important techniques for the pricing and risk management of a book with exotic derivatives.

A Poisson conditioned caching method

Recall from (5.4) and (5.6) that the exact distribution of the variance process is a constant times a non-central chi-squared distribution, for which representation (5.30) can be used, i.e.

$$v(t)|v(s) \quad \stackrel{d}{=} \quad C_0 \chi^2_{d+2N} \quad \text{for } d > 0, \tag{5.43}$$

with and N a Poisson distribution[9] with mean $\mu = \frac{1}{2}\lambda$ and with (see (5.6)):

$$d = \frac{4\kappa\theta}{\xi^2}, \quad \mu = \frac{1}{2} C_5 \, v(s), \quad \text{and } C_5 := \frac{2\kappa e^{-\kappa\Delta t}}{\xi^2(1 - e^{-\kappa\Delta t})}. \tag{5.44}$$

[9]Recall: $\mathbb{P}(N = n) = \frac{\mu^n e^{-\mu}}{n!}, \quad n = 0, 1, 2, \ldots$.

Thus sampling from the non-central chi-squared distributed variance process is equivalent to sampling from a 'Poisson-conditioned' chi-squared distribution. Though this observation was already being used in the Broadie and Kaya scheme, our (yet to be described) sampling method is different. We claim our method is more efficient and better applicable in financial applications; moreover our sampling method for the variance scheme can either be used on its own or can be used as drop in for the variance sampling of the exact or almost exact scheme of Broadie and Kaya (2006) or Smith (2008). In the following sections we first describe the Poisson-sampling method and consecutively show how one can exploit a property of this distribution, which enables us to create an efficient cache (and corresponding sampling method) of the non-central chi-squared distribution.

Poisson sampling

Notice that the mean μ of the Poisson distribution depends on the size of the time step (through C_5) as well as on the current state of the variance process $v(s)$; for almost all practical model configurations one finds $\mathbb{E}[\mu] << 10$, for which the corresponding Poisson-distribution decays quite rapidly and has rather 'thin' tails: for all $\mu < 10$, $\mathbb{P}(N > 35) < 4.462 \cdot 10^{-11}$, see Figure **1**.

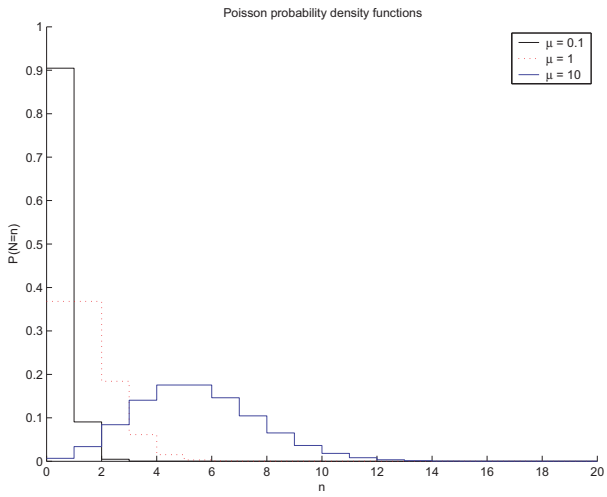

Figure 1: Poisson probability density functions for different μ.

This implies that we can (efficiently) draw a sample N_j from a Poisson distribution with a relatively small mean μ by just inverting its distribution function over an uniform random variable (e.g. see Knuth (1981) and Ahrens and Dieter (1982)):

1. Draw a uniform random number U_P, set $N_j = 0$ and $P_{N_j} = P_C = \exp(-\mu)$.
2. **While** $P_C \leq U_P$:
   $N_j \to N_j + 1$, $P_{N_j} \to P_{N_j} \cdot \frac{\mu}{N_j}$ and $P_C \to P_C + P_{N_j}$.

For small μ, where most probability mass lies within the first values of the distribution (e.g. see Figure **1**), the above inversion algorithm is very efficient. Moreover, it also implies that

we can create a cache for the non-central chi-squared distribution by precomputing chi-squared distributions for a truncated set Poisson-outcomes; since there is a little probability mass in the tail of the Poisson-distributions that one encounters during a simulation, the truncation error usually is negligible.

Caching the Chi-squared distributions

We first introduce some notation: let $N_{\max}$ represent a certain threshold level (e.g. such that $\mathbb{P}(N > N_{\max}) < \varepsilon$) and let

$$\mathcal{N} := \left\{0, \ldots, N_j, \ldots, N_{\max}\right\} \tag{5.45}$$

represent the set of Poisson-values for which we will cache the inverse of the corresponding conditional chi-squared distributions (i.e. according to 5.43). Since the inverse lives on the uniform domain, we let $\mathcal{U}_{N_j}$ represent a corresponding sorted set of uniform variables for which we the inverse $(\chi^2_{d+2N_j})^{-1}(\cdot)$ is calculated[10], i.e.

$$\mathcal{U}_{N_j} := \{0, \ldots, 1 - \delta\}\,. \tag{5.46}$$

Thus we suggest to create a cache of the values of the inverse of the non-central chi-squared distribution function by means of conditioning on a truncated range of Poisson-values and precomputing the corresponding chi-squared distribution functions, i.e. we precompute

$$H^{-1}_{N_j}(U_i) := G^{-1}_{\chi^2_{d+2N_j}}(U_i), \quad \forall N_j \in \mathcal{N}\ \forall U_i \in \mathcal{U}_{N_j}, \tag{5.47}$$

with $G^{-1}_{\chi^2_{d+2N_j}}$ the inverse chi-squared distribution with $d + 2N_j$ degrees of freedom.

Generating a sample from the variance process

From (5.6) we know that $v(t)|v(s)$ is distributed as a constant C_0 times a non-central chi-squared distribution, we can use the results of the previous subsection and sample from the variance process by first conditioning on a Poisson variable N_j and consecutively inverting the corresponding chi-squared distribution. To invert the chi-squared distribution for $N_j \leq N_{\max}$, we just draw a uniform variate and interpolate between the two values of the inverse distribution cache that surround the uniform numbers. In case $N_j > N_{\max}$ we use the distribution corresponding to $N_{\max}$ and moment-matching techniques which we explain below.

The caching method (5.47) and the following sampling rule form the core of the NCI scheme. That is, draw a Poisson number N_j and a uniform random number U_V (e.g. $U_i < U_V < U_{i+1}$) then

[10] Since $\lim_{U\to 1} G^{-1}_{\chi^2_{d+2N_j}}(U) = \infty$, one should use $1 - \delta$ instead of 1 to avoid numerical difficulties. Here δ is defined as a small machine number: in C one can for example set $\delta =$ `DBL_EPSILON` which is defined in the header `float.h`

a sample of $v(t)|v(s)$ is generated by

$$v(t) = F^{-1}_{N_j}(U_V), \tag{5.48}$$

with

$$F^{-1}_{N_j}(U_V) := \begin{cases} C_0 J(U_V) & \text{for } N_j \leq N_{\max}, \\ C_0 F^{-1}_{\chi^2_{d+2N_j}}(U_V) & \text{for } N_j > N_{\max}, \end{cases} \tag{5.49}$$

where C_0 is defined in (5.6) and where $J(\cdot)$ represents an interpolation rule. The NCI sampling scheme thus consists of an inversion of the non-central chi-squared distribution for the low and most frequent Poisson-outcomes and of a moment-matching scheme based on the chi-squared distribution for the rare and high Poisson outcomes: though the probability of $\{N > N_{\max}\}$ is usually small, it will be strictly greater than zero for all $N_{\max}$ and we decide to use direct inversion[11].

Design of the cache: a practical example

As an example we describe a method to implement the Non-Central Chi-Squared Inversion (NCI) scheme, specifically we will explore how to design the cache. A few details still has to be filled in: which value should be chosen for $N_{\max}$ in (5.45), how should the uniform numbers for a $\mathcal{U}_{N_j}$ in (5.46) be aligned and which interpolation rule J should be chosen to interpolate between two values of the inverse chi-squared distribution.

For expositional purposes we use the parameter values $v(0) = 0.09$, $\theta = 0.09$, $\kappa = 1.0$, $\xi = 1.0$ for the variance process and we use an equidistant time grid with $\Delta t = 0.25$ and maturity $T = 5$. Using (5.43), (5.44) and (5.48) this then implies that the exact distribution of the variance process equals a constant C_0 ordinary chi-squared distribution with $d + 2N = \frac{4\kappa\theta}{\xi^2} + 2N = 0.36 + 2N$ degrees of freedom, where N (cf. (5.44)) is Poisson distributed with mean $v(s)C_5 = v(s)7.042$. Using this setting as example we comment on the choice of $N_{\max}$. As shown in Table **1**, this choice mainly depends on the mean of Poisson distribution: for the case $s = 0$, we have $v(0) = 0.09$, hence the scheme implies that we need to sample from a Poisson-distribution with mean $\mu = 0.634$ and we could easily use this mean to set a bound for $N_{\max}$. Unfortunately one then ignores the randomness of the mean: even though the means of the stationary and non-stationary distribution can be approximately equal, the randomness significantly increase the mass in the tails of non-stationary distribution function (i.e. based on all Poisson-draws in the simulation) when compared to the 'stationary' Poisson-distribution at time 0. An example of this behaviour can be seen in the empirical distribution function as reported in Table **1**.

[11] Since for most parameter configurations and a reasonable choice of $N_{\max}$, the probability of $N_{\max}$ we only need to use a direct inversion a very limited amount of times, the computational overhead of direct inversion will be relatively small. Alternatively one can also opt to use an approximation for a chi-squared distribution with moderate to large degrees of freedom, e.g. see Abramowitz and Stegun (1964).

n	$\mathbb{P}(N > n)$	n	$\mathbb{P}(N > n)$	n	$\mathbb{P}(N > n)$
-1	1	4	0.0361	9	0.0075
0	0.2585	5	0.0253	10	0.0058
1	0.1362	6	0.0182	20	0.0007
2	0.0816	7	0.0133	40	$1.26 \cdot 10^{-5}$
3	0.0530	8	0.0099	80	$9.00 \cdot 10^{-7}$

Table 1: Empirical distribution function based on the Poisson samples that were drawn in 10^7 simulations with the parameters: $v(0) = 0.09$, $\theta = 0.09$, $\kappa = 1.0$, $\xi = 1.0$, $\Delta t = 0.25$ and maturity $T = 5$.

One then has to decide which size, alignment and interpolation rule one uses on the grid of uniform numbers in the cache(s) of the inverse of the chi-squared distribution with $d+2N_j$ degrees of freedom; first notice that, as shown in Table **1**, the number of draws from the corresponding chi-squared distributions differs significantly across the bins, e.g. in the example more than 74% of the drawings end up in the first bin. Another point to take into account for interpolation rule, is that the inverse distribution function is a monotone function, hence we would like the interpolation rule to preserve the monotonicity in the cached data points. A third point might be that some areas (i.e. in the tails of the distribution) are 'more difficult' to interpolate, which might plea for using a non-equidistant alignment. To keep the mapping of a uniform sample to the corresponding cached value straightforward, we simply opt to use a equidistant grid. Hence given a uniform drawing U_V and say $1, \ldots, K$-values in the cache, we the lower index i can then be easily located by evaluating $l := \text{floor}(U_V * K)$.
For the interpolation rules, we suggest two rules that preserve the monotonicity of the date and are relatively easy to implement. The first one is linear interpolation: given $U_i < U_V < U_{i+1}$ the linear interpolation rule $J(\cdot)$ is given by

$$J(U_V) := \frac{U_{i+1} - U_V}{U_{i+1} - U_i} H_j^{-1}(U_i) + \frac{U_V - U_i}{U_{i+1} - U_i} H_j^{-1}(U_{i+1}). \tag{5.50}$$

This rule has the advantage that is fast to execute, but in our experience some more points has to be cached in comparison to higher order interpolation rules to obtain full accuracy. Hence it might depend on the situation (e.g. on the number of paths) whether linear interpolation is suited. Alternatively we therefore suggest a monotone cubic Hermite spline interpolation which rule is defined as follows.

$$J(U_V) := h_{00}(t)H_j^{-1}(U_i) + h_{01}(t)H_j^{-1}(U_{i+1}) + \Delta_i(m_i h_{10}(t) + m_{i+1}h_{11}(t)), \tag{5.51}$$

with $t := \frac{U_V - U_i}{U_{i+1} - U_i}$ and where the corresponding definitions of the Hermite polynomials $h_{00}, h_{01}, h_{10}, h_{11}$ and the weights m_i are given in appendix 5.7.1. Though the spline interpolation rule requires a few more (elementary) operations in each step, one can significantly reduce the required number of points in the cache. Finally a suitable number of points and the choice of $N_{\max}$ is both case (e.g. the number of simulations) as well as parameter dependent and basically constitutes a efficiency trade-off between bias and variance. We do mind the reader to be careful with too coarse grids, since $\mathbb{E}[N] = \text{Var}[N] \to \infty$ as $\Delta t \to 0$ and one might end up with many

exceedings of $N_{\max}$ implying a lose of efficiency since in such case direct inversion[12] instead of a look-up value from the cache is being used. It is not very likely that this will cause a problem, since from the numerical results of 5.5.2 it follows that using only a few steps per year already produces schemes which have a negligible bias. As for a certain product small time steps (e.g. weekly steps) are required and one can for example retain the efficiency of the cache by using the modification as described in section 5.3.5.

NCI Algorithm

Assuming that a certain threshold level $N_{\max}$, an alignment of the grid $\mathcal{U}_{N_j}$ (5.46) and interpolation rule $J(\cdot)$ (5.49) has been chosen, the Non-central Chi-Squared Inversion (NCI) algorithm can be summarized by the following algorithm:

1. For $N_j = 0, \ldots, N_{\max}$, use (5.47) to precompute the inverses of the chi-squared distribution on the grid $\mathcal{U}_{N_j}$ and compute d using (5.44)
2. Given $v(s)$ and μ using (5.44).
3. Generate a sample from a Poisson distribution with mean μ according to (5.3.4):
   (a) Draw a uniform random number U_P.
   (b) Set $N_j = 0$ and $P_{N_j} = P_C = \exp(-\mu)$.
   (c) **While** $P_C \leq U_P$[13]:
   Set $N_j = N_j + 1$, $P_{N_j} = P_{N_j} \cdot \frac{\mu}{N_j}$ and $P_C = P_C + P_{N_j}$.
4. Generate a sample from a chi-squared distribution with $d + 2N_j$ degrees of freedom:
   (a) Draw a uniform random variable U_V.
   (b) Use (5.48), i.e. set $v(t) = F^{-1}_{N_j}(U_V)$ where $F^{-1}_{N_j}$ is defined in (5.49).

5.3.5 The NCI-QE scheme

An alternative sampling method for the non-central chi-squared distributed variance process can be to combine the NCI and QE approximations of section 5.3.4 and 5.3.3. This combination is motivated by the fact that while NCI-scheme is particularly efficient for small non-centrality parameters (i.e. implying a small Poisson-mean and hence a small cache, e.g. see 5.44), the QE-scheme works especially well for moderate to high non-centrality parameters (i.e. corresponding

[12] Since $\chi^2_{d+2N_j} \overset{d}{=} \gamma_{\frac{\nu}{2},2} \overset{d}{=} 2\gamma_{\frac{\nu}{2},1}$, the inverse of a chi-squared distribution with ν degrees of freedom can be obtained by taking twice the inverse of gamma distribution with shape $k = \frac{\nu}{2}$ and scale 1. For the inversion of the gamma distribution we use an implementation of Maddock et al. (2008). Their (root search) algorithm first uses the method of Didonato and Morris (1986) to find an initial approximation for the inverse and consecutively applies a few Halley iterations to cut off some extra digits. In many cases only a few iterations are needed for the complete algorithm, i.e. in the evaluation of the function `2.0*gamma_p_inv(k,u)`.

[13] To avoid eventual numerical errors, one might define a maximum number of loops for safety.

to the quadratic normal-like approximation (5.34), see Abramowitz and Stegun (1964) or Andersen (2008)). To see which circumstances correspond to low/high non-centrality parameters, we use the parameter settings of Table **1** and show the impact of instantaneous variance and the size of the time step on the value of λ (5.6). The results are given in Figure **2**.

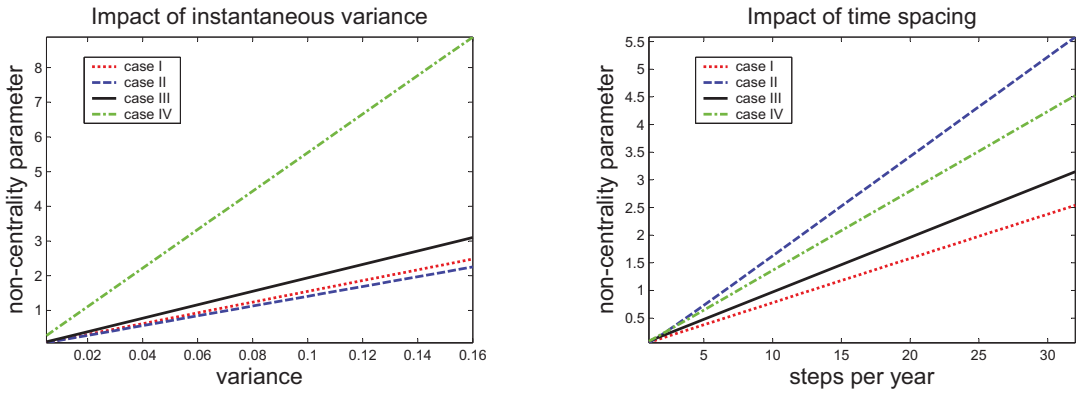

Figure 2: The non-centrality parameter λ (5.6) as function of the instantaneous variance and the size of the time step for the parameter settings of Table **1**. In the left graph we fixed the number of time steps at 8 steps per year.

From equation (5.6) and the figure and we can hence see that the non-centrality parameter grows linearly as function of the instantaneous variance and (almost) linear as function of the number of time steps. Hence recalling that the NCI/QE works particularly well for low/high non-centrality, a logical switching rule would try to combine the best of both worlds by using a critical switching value λ_c of the non-centrality parameter. That is, for $\lambda \leq \lambda_c$ one uses the NCI sampling scheme and else the quadratic moment-matched approximation (5.34) of the QE scheme. The exact choice of the switching value λ_c (provided that the moment-matching condition 5.38 is satisfied) appears to have a relatively small effect on the quality of the simulation scheme, indicating that the transition between both schemes is rather smooth for moderate degrees of freedom. In our numerical test we use $\lambda_c = 4$ as critical switching value between the schemes; first, for $\lambda > 4$ the quadratic moment-matched scheme is always applicable since then $\psi < 2$ and secondly, $\lambda < 4$ implies for the NCI scheme that we need to draw a sample from a Poisson-conditioned chi-squared distribution with mean $\mu = \frac{1}{2}\lambda < 2$. Since there is little probability mass in the Poisson distribution for small μ[14] we only need to cache a small number of chi-squared distributions (e.g. see (5.46) and (5.49)), and at the same time the use of the direct inversion (i.e. when the Poisson-draw exceeds the maximum cache size) will be limited.
By using $\lambda_c = 4$ as critical value, Figure **2** implies that for the considered parameter configurations, the scheme uses the NCI part the majority of the time for a number of time steps that is smaller than $20 - 50$ steps per year. The other way around: for larger values of the instantaneous variance and a smaller time-grid the QE scheme will used more extensively. Since in practice

[14]For all $\mu < 2$ one has $\mathbb{P}(N > 10) < 8.3 \cdot 10^{-6}$

4 – 8 time steps a year is already sufficient to produce a scheme with a negligible bias (see section 5.5), one will probably apply NCI part the majority of the time. Do note that as we make the size of the time step smaller and smaller, the use of the NCI becomes more and more remote.
The simulation algorithm for the NCI-QE scheme works straightforward: in step (2) of the NCI-algorithm 5.3.4 one evaluates μ: if $\mu < \frac{1}{2}\lambda_c$, one applies (3)-(4) or else one uses the quadratic approximation in step (2) of the QE-algorithm 5.3.3.

5.4 Asset price sampling in combination with drift interpolation

Though a sample of the integrated variance process can be generated using the exact method of Broadie and Kaya, or the almost exact method of Smith, it is probably computationally more efficient to use the simple drift interpolation method of (5.31), for example with the corrector-predictor setting $\gamma_1 = \gamma_2 = 0.5$. Hence by applying the drift interpolation method for the integrated variance process, one can modify the 'Broadie and Kaya' asset price sampling scheme of (5.23) as follows:

$$\begin{aligned}\log(S(t)) &= \log(S(s)) + r\Delta t + \frac{\kappa\rho}{\xi}(\gamma_1 v(s) + \gamma_2 v(t))\Delta t - \frac{1}{2}(\gamma_1 v(s) + \gamma_2 v(t))\Delta t \\ &\quad + \frac{\rho}{\xi}(v(t) - v(s) - \kappa\theta\Delta t) + \sqrt{1-\rho^2}\,\sqrt{\gamma_1 v(s) + \gamma_2 v(t)}\cdot Z_s\sqrt{\Delta t}, \\ &= \log(S(s)) + r\Delta t + K_0 + K_1 v(s) + K_2 v(t) + \sqrt{K_3 v(s) + K_4 v(t)}\cdot Z_S \qquad (5.52)\end{aligned}$$

where Z_s is a standard normal distributed random variable, independent of v, and with

$$\begin{aligned}K_0 &= -\frac{\rho\kappa\theta}{\xi}\Delta t, & K_1 &= \gamma_1\Delta t\Big(\frac{\kappa\rho}{\xi} - \frac{1}{2}\Big) - \frac{\rho}{\xi}, & K_2 &= \gamma_2\Delta t\Big(\frac{\kappa\rho}{\xi} - \frac{1}{2}\Big) + \frac{\rho}{\xi}, \\ K_3 &= \gamma_1\Delta t\Big(1-\rho^2\Big), & K_4 &= \gamma_2\Delta t\Big(1-\rho^2\Big).\end{aligned}$$

Hence this asset price sampling scheme can be used in conjunction with different methods to simulate the variance process. One can for example use this method in conjunction with variance sampling that we have previously considered, that is with the BK-DI, the QE, the NCI and the NCI-QE scheme of sections 5.3.1-5.3.5. Please note that the above discretization scheme usually is not completely arbitrage-free, though the bias induced from this approximation is usually rather small and controllable by reducing the size of the time step. However with little effort one can also (locally) enforce an exact no-arbitrage condition, leading to the martingale corrected BK-DI-M, QE-M, NCI-M and NCI-QE-M schemes, see Andersen (2008) and section 5.4.1.

Asset price sampling scheme

The stock price discretization scheme of the previous section can be summarized as follows:

1. Conditional on $v(s)$, use the BK-DI scheme 5.3.1, the QE scheme 5.3.3, the NCI scheme 5.3.4 or the NCI-QE scheme of 5.3.5 to generate a sample for $v(t)$.

2. Generate a random sample Z_S from the standard normal distribution.[15]

3. Given $\log(S(s))$, $v(s)$, $v(t)$ and Z_S, compute $\log(S(t))$ from (5.52).

5.4.1 Martingale correction

As discussed in Andersen and Piterbarg (2007), the continuous-time asset price process $S(t)$ might not have finite higher moments, but the discounted stock price will always be a martingale under the risk-neutral measure, i.e.

$$\mathbb{E}^Q\left[e^{-r\Delta t}S(t)\middle|\mathcal{F}_s\right] = S(s) < \infty. \tag{5.53}$$

If we however takes the exponent of the scheme (5.52), one usually has that the discounted discretized stock price process (from here on denoted by $\widetilde{S}(t)$)

$$\widetilde{S}(t) = \widetilde{S}(s)\exp\left[r\Delta t + K_0 + K_1 v(s) + K_2 v(t) + \sqrt{K_3 v(s) + K_4 v(t)}\cdot Z_S\right], \tag{5.54}$$

does not satisfy the martingale condition, i.e. with $\Delta t = t - s$,

$$\mathbb{E}^Q\left[e^{-r\Delta t}\widetilde{S}(t)\middle|\mathcal{F}_s\right] \neq \widetilde{S}(s). \tag{5.55}$$

As noted by Andersen (2008) the practical relevance of this is often minor, because the net drift away from the martingale is typically very small and controllable by reducing the size of the time step. Moreover the ability to hit the mean of the distribution does not necessarily lead to better option prices. Following Glasserman (2003) and Andersen (2008), we do discuss the martingale correction method, that is we investigate whether it is possible to modify the NCI scheme such that the discretized option price becomes martingale. Additionally we look at the regularity of the discretization scheme, e.g. we look whether there might parameter values where the $\widetilde{S}$-process might blow up in the sense that $\mathbb{E}[\widetilde{S}(t)|S(s)] = \infty$.

By the tower law of conditional expectations, we have

$$\mathbb{E}^Q\left[e^{-r\Delta t}\widetilde{S}(t)\middle|\mathcal{F}_s\right] = \mathbb{E}^Q\left[\mathbb{E}^Q\left(e^{-r\Delta t}\widetilde{S}(t)\middle|\mathcal{F}_s \vee \sigma(v(t))\right)\right], \tag{5.56}$$

with $\sigma(v(t))$ the sigma-algebra generated by $v(t)$ and where $\mathcal{F}_s \vee \sigma(v(t))$ denotes the smallest sigma-algebra containing both $\mathcal{F}_s$ and $\sigma(v(t))$. For the martingale condition (5.53) to hold, we

[15] It might be advisable to use an inversion method to generate normal samples, since then a quasi random generator is also applicable. The algorithm described in Wichura (1998) gives an approximation $\widetilde{\Psi}^{-1}$ of Ψ^{-1} with a relative error of

$$\frac{|\widetilde{\Psi}^{-1} - \Psi^{-1}|}{1 - |\Psi^{-1}|} < 10^{-15}.$$

thus need $\widetilde{S}(s)$ to equal the above expectation. Using the moment-generating function of the normal distribution, we can write the following for the discretized stock price $\widetilde{S}(t)$:

$$\begin{aligned}\widetilde{S}(t) &= \widetilde{S}(s)\exp[K_0^* + K_1 v(s)]\mathbb{E}\Big[\exp[K_2 v(t)]\mathbb{E}^Q(\exp[\sqrt{K_3 v(s) + K_4 v(t)}\cdot Z_S]\big|\mathcal{F}_s, v(t))\Big] \\ &= \widetilde{S}(s)\exp[K_0^* + K_1 v(s)]\mathbb{E}\Big[\exp[K_2 v(t) + \frac{1}{2}(K_3 v(s) + K_4 v(t))]\Big] \\ &= \widetilde{S}(s)\exp[K_0^* + (K_1 + \frac{1}{2}K_3)v(s)]\Psi_{v(t)}(A) \qquad (5.57)\end{aligned}$$

with

$$A := K_2 + \frac{1}{2}K_4, \qquad (5.58)$$

and where $\Psi_{v(t)}(x)$ denotes the moment-generating function of the (discretized) variance process $v(t)$ evaluated in the point x. Hence for the martingale condition to hold we need

$$\exp(K_0^* + (K_1 + \frac{1}{2}K_2)v(s))\Psi_{v(t)}(A) = 1, \qquad (5.59)$$

which (assuming the regularity condition $\Psi_{v(t)}(A) < \infty$) is satisfied by replacing the constant K_0 in (5.52) by

$$K_0^* := -\log(\Psi_{v(t)}(A)) - (K_1 + \frac{1}{2}K_2)v(s). \qquad (5.60)$$

Hence what remains is to determine the moment-generating function of the simulated variance processes. Additionally we look at the regularity of the schemes: we check whether there might be parameter configurations for which the moment-generating function of the simulated variance process does not exist.

5.4.2 Moment generating function of $v(t)$, Regularity

QE(-M) scheme

Using properties of the non-central chi-squared distribution with one degrees of freedom on (5.34) and direct integration on (5.37) Andersen (2008) derives the moment-generating functions of his variance schemes. For the regularity conditions to hold and the moment-generating functions to exist, the conditions $aA < \frac{1}{2}$ and $A < \beta$ have be satisfied. As motivated in Andersen (2008) these conditions are often (but not always) satisfied in practice: the most restrictive condition is that for $\rho > 0$ the size of the time step has to be sufficiently small. Assuming regularity, one can define

$$K_0^* = \begin{cases} -\frac{Ab^2 a}{1-2Aa} + \frac{1}{2}\log(1-2Aa) - (K_1 + \frac{1}{2}K_3)v(s) & \text{if } \psi \le \psi_c, \\ -\log(p + \frac{\beta(1-p)}{\beta - A}) - (K_1 + \frac{1}{2}K_3)v(s) & \text{if } \psi > \psi_c. \end{cases} \qquad (5.61)$$

and replace K_0 by K_0^* in the QE scheme of (5.52) to obtain the martingale-corrected QE(-M) scheme for the discretized stock price process S (see Andersen (2008) proposition 9, pp. 21-22).

NCI(-M) and BK-DI(-M) scheme

Since we know from (5.6) that the variance process $v(t)$ is distributed as a constant C_0 times a non-central chi-squared distribution with d degrees of freedom and non-centrality parameter λ, we find following expression for $\Psi_{v(t)}(A)$ for the exact [16] variance process $v(t)$.

$$\Psi_{v(t)}(A) = \mathbb{E}\Big[\exp(Av(t))\Big] = \frac{\exp(\frac{C_0 A\lambda}{1-2C_0A})}{(1-2C_0A)^{\frac{d}{2}}}. \tag{5.62}$$

For this expectation to exist we need $C_0A < \frac{1}{2}$, i.e.

$$\frac{\rho}{\xi}(1+\kappa\gamma_2\Delta t) - \frac{1}{2}\gamma_2\Delta T\rho^2 < \frac{2\kappa}{\xi^2(1-e^{-\kappa\Delta T})}. \tag{5.63}$$

To get a grasp at the restrictiveness of this condition, notice that the right hand side is always positive; hence it follows that for $\rho \leq 0$ the condition will always be satisfied. In contrast, for $\rho > 0$, equation (5.63) imposes a limit on the size of the time step ΔT, roughly $\rho\xi\Delta T < 2$. Assuming that the regularity conditions are satisfied, we can apply (6.72) and set

$$K_0^* = -\frac{C_0A\lambda}{1-2C_0A} + \frac{d}{2}\log(1-2C_0A) - (K_1 + \frac{1}{2}K_3)v(s) =: C_6\lambda + C_7 + C_8v(s) \tag{5.64}$$

to enforce the martingale in the NCI(-M) and BK-DI(-M) discretization scheme. Note that the constants

$$C_6 = -\frac{C_0A}{1-2C_0A}, \quad C_7 = \frac{d}{2}\log(1-2C_0A), \quad C_8 = -(K_1 + \frac{1}{2}K_3), \tag{5.65}$$

can be precomputed before the Monte Carlo run.

5.5 Numerical results

To test our new scheme we consider the pricing of an Asian and European call options in the Heston model. Since European call prices belong to the vanilla options and can be calculated with a great accuracy (e.g. see Carr and Madan (1999)), they form a standard test case. Additionally we use an Asian option as test case for a more path-dependent option.
To investigate the efficiency and discretization bias of the NCI scheme, we benchmark it to the full truncation Lord et al. (2008) and the QE-M scheme Andersen (2008). These schemes are to our knowledge the most efficient scheme in most practical situations. Note that we do not incorporate the Broadie and Kaya (2006) scheme in our comparison; first, by definition, the Broadie and Kaya scheme is exact and thus free of discretization bias. Therefore theoretically, Monte Carlo simulations are not needed to test the quality of the Broadie and Kaya scheme.

[16] Since by setting $N_{\max}$ sufficiently large we can come as close as we want to the true inverse of the variance process by the inverse as defined in (5.49), we use the above expression as moment generating function of the NCI variance process $v(t)$

However in practice, some numerical error will certainly be introduced in the required numerical inversions of the algorithm (see section 5.2.4); moreover Lord et al. even show (Lord et al. (2008) pp.16 table 5) that for equal computational budgets[17] both the right Euler scheme (i.e. the full truncation scheme) as well as a moment-matched scheme completely outperform the exact scheme. Though the almost exact simulation scheme of Smith (2008) without doubt delivers a speed to the exact scheme, we still argue that a moment-matched scheme and the full truncation Euler scheme are more efficient in most situations[18].
As we are interested in efficient discretization schemes, we follow Lord et al. (2008) and Andersen (2008), and prefer to initially compare the accuracy our schemes with the computationally most efficient schemes instead of computationally inefficient (almost) exact schemes. A comparison in terms of computational efficiency between all the schemes, including the (almost) exact methods of Broadie and Kaya (2006) and Smith (2008), can be found in Table 5.5.3.

Though our benchmark setup is similar to Lord et al. (2008) and Andersen (2008), we add a control variate to get rid of some extra Monte Carlo noise[19]: both in the analysis of Lord et al. (2008) as well as in Andersen (2008) the 'finite sample noise' of the Monte Carlo plays still plays such a big role, that even a number of 10^6 paths in some cases sometimes is not enough to draw a good comparison both schemes, let alone say something about the true quality and bias of the schemes. Though this certainly says something about the overall quality of the full truncation and QE schemes, the produced estimates still seem somewhat inconsistent and this motivates the use of the control variate. With this simple variance reduction technique we are then able to draw stronger conclusions about the overall and the relative quality of the schemes. Though low-discrepancy numbers also form a natural candidate for faster convergence behaviour, we do not use them in the benchmark to keep the comparability with the numerical results in the existing literature. We do want to emphasize that in practice, all the three schemes can be used in conjunction with low-discrepancy numbers, e.g. with the Sobol numbers as described in Jäckel (2002) and Press and Flannery (1992).

[17] With computational budget we mean CPU time. Note that because different schemes require different computational effort for the same time grid, fixing the computational budget implies that the size time step of the required schemes is adjusted to match the computational budget. Since for a fixed time step an Euler and moment-matched scheme require a smaller computational budget than the exact Broadie and Kaya scheme (that requires two numerical inversion procedures), this implies that with a equal budget, smaller time steps can be taken in the Euler and moment-matched schemes in comparison with the exact scheme.

[18] For example compare the performance of the almost exact scheme with the Kahl-Jäckel scheme in Smith (2008) to the Kahl-Jäckel scheme in Andersen (2008) or Lord et al. (2008). Whereas the almost exact scheme in Smith (2008) performs comparable to the Kahl-Jäckel scheme Kahl and Jäckel (2006), the full truncation scheme as well as the QE-M scheme outperform this scheme in Lord et al. (2008) and Andersen (2008).

[19] For an overview of other variance reduction techniques we refer to Jäckel (2002) and Glasserman (2003). For stochastic volatility models the method of Willard Willard (1997), which entirely eliminates the noise of the stock price, can be rather attractive for options that have a closed form solution under the Black and Scholes (1973) model. We do not use the latter method since we are interested in the bias of the joint simulation of stock and variance.

5.5.1 Benchmark setup

To test our discretization schemes, we consider the parameter configurations of table (**1**). These settings correspond to some different parameter settings which are likely to be encountered in equity, FX or interest rate markets.

Example	Type	κ	ξ	ρ	$v(0)$	θ	r
case I	Call-10Y	0.5	1.0	-0.9	0.04	0.04	0.00
case II	Call-5Y	1.0	1.0	-0.3	0.09	0.09	0.05
case III	Call-15Y	0.3	0.9	-0.5	0.04	0.04	0.00
case IV	Asian-4Y	1.0407	0.5196	-0.6747	0.0194	0.0586	0.00

Table 2: Test cases for the Heston schemes, in all cases $S(0) = 100$.

Notice that for all cases we have $\xi^2 >> 2\kappa\theta$, implying that the origin is accessible. The first and third example are taken from Andersen (2008) and serve to represent long-dated FX options. We can expect that such a setting is difficult, since the corresponding parameters combine a low mean-reversion κ with a high volatility of the variance process ξ; hence the variance process has a relatively high probability of reaching the troubled region near zero. The second case stems from Broadie and Kaya (2006) and can correspond to an equity setting. The last setting is taken from Smith (2008) who considers an Asian option with yearly fixings to test some models for a path-dependent equity option.

In the test cases we consider the pricing of a (vanilla/asian) call option maturing at time T and with strike K. Hence we are interested in a Monte Carlo valuation of the call option price C and Asian option price A, i.e.

$$C = e^{-rT}\mathbb{E}\Big[(S(T) - K)^+\Big], \quad \text{and:} \quad A = e^{-rT}\mathbb{E}\Big[(\sum_{j=1}^{N} \frac{S(T_j)}{N} - K)^+\Big] \tag{5.66}$$

Hence with a discretization scheme for the stock price $S(t)$, we can use Monte Carlo methods to approximate these prices by $\widehat{C}$ and $\widehat{A}$. Due to errors introduced by discretization and the Monte Carlo one in general finds that the estimated values are not equal the theoretical values. Hence we define the bias $\widehat{b}$ of the discretization scheme as this difference, i.e.

$$\widehat{b} := C - \widehat{C}. \tag{5.67}$$

In the following subsection we hence specify how the Monte Carlo estimates for $\widehat{C}$ and $\widehat{A}$ can be obtained.

Control variate estimators

The ordinary Monte Carlo estimator would consist of simulating n independent samples of the required (discounted) payoff(s) and averaging over all paths, which (by the strong law of large

numbers) converges to the expected option price $\mathbb{E}[\widehat{C}]$ of the discretized asset price dynamics. Though this estimator is generic and simple, there exist various other estimators (e.g. see Glasserman (2003) or Jäckel (2002)) which are more efficient, c.q. which have a lower variance. To obtain some variance reduction over the ordinary estimator, we therefore suggest to use a control variate estimator which is also quite generic and involves little computational overhead; For $i = 1, \ldots, n$ we generate Monte Carlo samples of the stock price(s) S_i and the corresponding option prices C_i. We can then use the following control variate estimates to estimate the vanilla/Asian call option prices:

$$\overline{C}(b_C) = \frac{1}{n}\sum_{i=1}^{n}(C_i - b_C(S_i - \mathbb{E}[S])), \qquad C_i = e^{-rT}(S_i(T) - K)^+, \tag{5.68}$$

$$\overline{A}(b_A) = \frac{1}{n}\sum_{i=1}^{n}(A_i - b_A(S_i - \mathbb{E}[S])), \qquad A_i = e^{-rT}(\sum_{j=1}^{N}\frac{S_i(T_j)}{N} - K)^+, \tag{5.69}$$

which estimators (again by the strong law of large numbers) also converge with probability one to the expected option prices. To see the effectiveness, i.e. variance reductions, of these control variate estimators over the ordinary Monte Carlo estimators we refer to appendix 5.7.2. To get more reliable estimates, we use the control variate estimators rather than the ordinary Monte Carlo estimator for the first three test cases. Though the technique is also quite effective for the considered Asian option (as can be seen from Table **5**), we rather use the ordinary Monte Carlo estimator to make our numeric results comparable with those of Smith (2008).

5.5.2 Numerical tests

For the numerical results of the first three test cases we use the following discretization schemes: the Euler-FT scheme of (5.13)-(5.16), the Kahl-Jäckel 'IM-IJK' scheme of (5.18)-(5.19), the QE-M scheme of section 5.3.3, the NCI-M scheme of section 5.3.4, the combined NCI-QEM scheme of 5.3.5 and the Broadie and Kaya 'drift interpolated-martingale corrected' scheme (denoted by "BK-DI-M"). For the latter four schemes we use the corrector-predictor scheme of (5.52) with the a mid-point rule (i.e. $\gamma_1 = \gamma_2 = \frac{1}{2}$) combined with the martingale corrections of section 5.4.1 for the discretized stock price. Then in the fourth test case we compare the results of the Euler-FT, the IM-IJK, the QE-M, the NCI-M, the NCI-QE-M and the BK-DI-M scheme against the results of the exact and almost exact scheme for a path-dependent option that can be found in Smith (2008).
To obtain accurate estimates we performed all tests using 10^6 Monte Carlo[20] paths. Moreover in the first three tests we use control variate estimators of (5.68), for all the schemes for which the control variate is applicable, to reduce the standard error even further. That is, for all the schemes that satisfy the discrete time martingale condition, i.e. all schemes except the IM-IJK. Please note that the control variate only affects the variance and not the (expected) bias of the Monte Carlo estimates. To keep our test comparable with the results in Smith (2008) we just use

[20]We use the Mersenne Twister as pseudo random number generator in combination with incremental path generations.

	Euler-FT	IM-IJK	QE-M	NCI-M	NCI-QE-M	BK-DI-M
Δt			$K = 100$			
1	-6.359 (±0.043)	-57.574 (±0.275)	-0.222 (±0.020)	0.246 (±0.022)	0.241 (±0.022)	0.234 (±0.022)
1/2	-3.718 (±0.033)	-32.828 (±0.179)	-0.110 (±0.021)	0.075 (±0.022)	0.073 (±0.022)	0.079 (±0.022)
1/4	-2.048 (±0.027)	-18.457 (±0.120)	-0.008* (±0.022)	0.015* (±0.022)	0.029 (±0.022)	0.013* (±0.022)
1/8	-1.036 (±0.024)	-10.094 (±0.084)	0.023 (±0.022)	0.008* (±0.022)	-0.002* (±0.022)	-0.001* (±0.022)
1/16	-0.537 (±0.023)	-5.325 (±0.063)	0.018* (±0.022)	-0.002* (±0.022)	-0.011* (±0.022)	0.010* (±0.022)
1/32	-0.255 (±0.023)	-2.678 (±0.050)	0.015* (±0.022)	0.005* (±0.022)	-0.007* (±0.022)	0.001* (±0.022)
			$K = 140$			
1	-4.269 (±0.041)	-51.544 (±0.242)	0.084 (±0.006)	0.029 (±0.006)	0.031 (±0.006)	0.031 (±0.006)
1/2	-1.942 (±0.024)	-28.038 (±0.146)	0.025 (±0.006)	0.011 (±0.006)	0.013 (±0.006)	0.014 (±0.006)
1/4	-0.761 (±0.013)	-14.810 (±0.086)	0.001* (±0.006)	0.002* (±0.006)	0.002* (±0.006)	0.004* (±0.006)
1/8	-0.265 (±0.009)	-7.529 (±0.051)	-0.002* (±0.006)	0.002* (±0.006)	0.003* (±0.006)	0.000* (±0.006)
1/16	-0.100 (±0.007)	-3.617 (±0.030)	0.000* (±0.006)	-0.005* (±0.006)	0.001* (±0.006)	-0.001* (±0.006)
1/32	-0.039 (±0.007)	-1.639 (±0.018)	-0.002* (±0.006)	-0.002* (±0.006)	-0.004* (±0.006)	-0.001* (±0.006)
			$K = 60$			
1	-3.120 (±0.029)	-50.708 (±0.304)	-0.027 (±0.019)	0.138 (±0.019)	0.127 (±0.019)	0.114 (±0.019)
1/2	-1.747 (±0.025)	-26.749 (±0.208)	0.047 (±0.020)	0.062 (±0.020)	0.059 (±0.020)	0.057 (±0.020)
1/4	-0.938 (±0.022)	-13.603 (±0.148)	0.039 (±0.020)	0.006* (±0.020)	0.017* (±0.020)	0.009* (±0.020)
1/8	-0.460 (±0.021)	-6.720 (±0.112)	0.030 (±0.020)	0.013* (±0.020)	-0.009* (±0.020)	0.009* (±0.020)
1/16	-0.235 (±0.020)	-3.305 (±0.091)	0.023 (±0.020)	0.002* (±0.020)	-0.011* (±0.020)	0.000* (±0.020)
1/32	-0.110 (±0.020)	-1.616 (±0.078)	0.011* (±0.020)	-0.002* (±0.020)	0.007* (±0.020)	0.001* (±0.020)

Table 3: Estimated call option prices biases for I. Numbers in parentheses are the widths of the confidence interval (5.70) at a 99% confidence level: we starred the biases that were not significantly different from zero. Exact prices respectively are 13.085, 0.296 and 44.330.

an ordinary Monte Carlo estimator for the fourth test case, i.e. the Asian option.
In tables **2**-**5** we then report the biases of the Monte Carlo estimates and the corresponding standard errors. We star the biases that are statistically insignificant at a 99% confidence level. That is, when the exact price lies within the 99% confidence window of the Monte Carlo price:

$$\text{exact value } \in \left[\overline{C} - z\frac{\sigma_C}{\sqrt{n}}, \overline{C} + z\frac{\sigma_C}{\sqrt{n}}\right], \tag{5.70}$$

with $z = 2.576 = \Phi^{-1}(1 - \frac{0.01}{2})$ the corresponding quantile-point.

Results for case I-III

The results for the estimated call option price bias of **case I** can be found below in Table **2**: we report the Monte Carlo estimates of the bias (5.67) as function of the time step Δt, for an at-, out- and in-the-money strike. The first thing to notice are the enormous differences in the magnitude of the biases between the schemes: whilst the first and second order discretization schemes (Euler/IM-IJK) are still quite biased for a 32 time steps a year, all the schemes that are based on approximating the non-central chi-squared distribution are already bias-free for a time-spacing of just 4 steps a year: that is, the biases of QE-M, NCI-M, NCI-QE-M and the BK-DI-M scheme are not significantly different from zero at a 99% confidence level on using 4 time steps a year. When we consider the bias over various strike levels, we see that the bias of the Euler and the martingale-corrected scheme decreases if the strikes goes more into the money. This is expected

since all these schemes are constructed to be bias-free for $K = 0$ (i.e. by the martingale construction).
All in all, we can conclude from the table that for practical sizes of the time step, the biases of the QE-M, the NCI-M, the NCI-QE-M and the BK-DI-M are substantially lower than those of the Euler and especially the IM-IJK scheme. Before we can conclude that the former schemes are also more efficient that the latter ones, we of course also need to look at the required computational effort of each discretization schemes, which will be addressed in section 5.5.3.

Results for case II and III

The numerical results for **case III** and **case III** can be found in the Table **4** and **5**. The results are similar to those of **case III**, but a little less severe. Still the almost exact variance schemes by far outperform the considered Euler (FT) and Milstein (KJ) scheme. The tables can be found in appendix 5.7.3.

5.5.3 Results for case IV and computational times

In the fourth test case we consider a path-dependent (Asian) option; to make the results comparable with the results as reported by Smith (2008) (table 3) we consider the same parameter settings (see Table **1**) and use the root mean square error (RMSE) as error measure for our estimates. Moreover since Smith (2008) also considers an Euler-scheme (Reflection), we can use the computational time from this scheme to scale our computational times with the ones as reported in Smith (2008). Hence we are in particular interested in the relative efficiency of the AESM in comparison to the schemes considered here; we are in particular interested in the time it takes the different schemes to reach a certain RMSE. Supported by the previous numerical results, we use 8 steps a year for the considered drift interpolation schemes and 100 time steps a year for the Euler and Milstein schemes. To obtain accurate estimates for standard Monte Carlo estimates, we use a high number of 2 560 000 paths (using pseudo-random Mersenne-Twister numbers). The results are given in Table **4**.

	Euler-R	Euler-FT	IM-IJK	AESM	QE-M	NCI-M	NCI-QE-M	BK-DI-M
Δt	1/100	1/100	1/100	1	1/8	1/8	1/8	1/8
RMSE	0.237	0.010*	0.018*	0.008*	0.009*	0.009*	0.009*	0.011*
Time (in sec.)	158.3	157.5	178.6	1134.4	16.3	22.7	17.2	25.6
Relative time	9.7	9.7	11.0	69.7	1.0	1.4	1.1	1.6

Table 4: Root mean squared errors (RMSE) for case IV. 'Exact' price 9.712: we starred the root mean squared errors with biases that were not significantly different from zero at a 99%-level. Results for the AESM are taken from Smith (2008), table 3.

The first thing to notice from Table **4** are the enormous differences in the required computational budgets; though the AESM method only requires a single step per year and we choose to use 100 time steps per year for the Euler schemes, the latter schemes are still approximately 7 times faster! This difference is even more stunning, if we compare the results for the drift interpolation

schemes with the AESM method; whilst the drift interpolation schemes show statistically indifferent RMSEs as the AESM method, they obtain this result approximately 40 to 70 times faster.
Secondly we would like to comment on the differences between the considered Euler schemes; in Broadie and Kaya (2006) and Smith (2008) some benchmarks are performed against a Euler scheme which fixes negative values in a non-optimal way (i.e. by absorbing of reflection negative values). As shown above (and extensively discussed in Lord et al. (2008)) the choice of the fix is extremely important for the overall quality of the simulation scheme. Hence when using such a non-optimal fix, one cannot conclude that "the" Euler schemes are less efficient than for example the AESM method: in Smith (2008) the Euler-R scheme is outperformed by the AESM method, while we can see from the above result that the Euler-FT by far outperforms the AESM method in terms of computational efficiency.
From a efficiency viewpoint we conclude that the QE-M scheme performs the best in the above test, followed closely by the other drift-interpolation schemes. The Euler-FT and the IM-IJK scheme follow on some distance, though these schemes share the advantage that they are simpler and more generically applicable, e.g. in other CEV models. The Euler-R scheme (though frequently used in the literature) performs extremely worse, with a large RMSE. Though the RMSE of the AESM is small and indifferent from the drift-interpolation schemes, the time required to obtain this estimate is by far the largest and except for the heavily biased non-optimal Euler-R scheme, the AESM scheme is outperformed by all the considered schemes.

5.6 Conclusion

Though the exact simulation method of Broadie and Kaya (2006) is theoretically appealing, its practical use might be limited due to its complexity and lack of speed. Moreover the suggested acceptance and rejection technique for the variance process hinders perturbation analysis, let alone the use of low-discrepancy numbers. Nevertheless the method also provides an excellent starting point for the development of some more efficient approximate schemes. In fact, several quite different methods can be considered to approximate the the sampling methods of the variance and integrated variance process.
Almost in line with the exact method, Smith (2008) recently suggested to speed up the sampling of the integrated variance with the AESM method. His results indeed indicate a significant speed up with respect to the exact scheme, nonetheless one still has to perform a Fourier inversion in each simulation step which is time-consuming. Moreover in the numerical analysis in the other discretization methods, we found that the largest variations were caused by biases in the simulation of the variance process, rather than the integrated variance process. For instance, we found with the new BK-DI scheme that it is more efficient to just use a simple drift interpolation method for the variance integral instead of an (approximate) Fourier inversion. However the prime disadvantage for financial applications for all the just mentioned schemes is the inconvenient acceptance and rejection sampling of the variance process, which heavily troubles sensitivity analysis and does not allow for the use of low-discrepancy number.
To this end, instead of an exact joint simulation of the variance and the integrated variance process, we looked at methods that approximate the variance process and just use a simple drift

approximation for its integral. The simplest scheme in this category is the Euler Full Truncation[21] scheme of Lord et al. (2008): remarkably, we found in our numerical test cases that this scheme by far outperforms many more complex schemes like the Kahl and Jäckel (2006) scheme and the AESM scheme of Smith (2008) in terms of computational efficiency. Though the Euler Full Truncation method is simple and straightforward to implement, it unfortunately still produces rather biased estimates for the moderate sizes of the time step: for our parameter settings we found that one at least has to use 32 time steps per year to obtain reasonably small biases. In a way this does not come as a surprise since the Euler scheme uses no analytical properties of the non-central chi-squared distribution variance process.

Last we considered a category of discretization methods which explicitly relies on the analytical properties of the variance process, whilst using sampling methods which are based on inversion methods rather than on acceptance rejection techniques, making the methods straightforward to use in sensitivity analysis and low-discrepancy numbers. For instance, Andersen (2008) uses moment-matching techniques to approximate the Non-central Chi-squared distributed variance process in the Quadratic Exponential (QE) scheme, where we suggested to use a one dimensional caching technique of the latter distribution in the NCI and NCI-QE schemes; though at first sight creating a cache with the inverses of this distribution might seem straightforward, this is in fact rather complicated as the parameter space of the inverse of the Non-central Chi-squared distribution is in fact three-dimensional and one has to design a potentially extremely large cache for all conceivable values of the number of degrees of freedom and non-centrality parameters, i.e. as pointed out by Broadie and Andersen (see Andersen (2008)). However, we tackled this dimensionality problem and showed that the inverses of the non-central chi-squared distribution can effectively be reduced to a one dimensional cache for the case of the Heston model. The crucial insight underlying this dimensionality reduction is that the variance process can be represented by chi-squared distribution whose degrees of freedom are given by a shifted Poisson random variable which decays extremely fast for typical Heston parameters. By conditioning on this shifted Poisson distribution, we have demonstrated that one can create an efficient cache of the variance process at relatively small computational costs, which results in the new NCI (Non-Central Chi-squared Inversion) and NCI-QE schemes. In a further analysis of these schemes, we investigated the regularity conditions and the enforcement of a local martingale property for discretized asset price, which hence lead to the NCI-M and NCI-QE-M schemes.

In the last section we performed a extensive numerical study with these new schemes and the exact scheme of Broadie and Kaya (2006), the almost exact scheme of Smith (2008), the Kahl-Jäckel scheme, the Full Truncation scheme of Lord et al. (2008) and the QE-M scheme of Andersen (2008). To strengthen this numerical analysis we used four different test cases (including European-style and a path-dependent Asian option) and a high number of sample paths in conjunction a variance reduction technique, which enables us to obtain highly accurate results. Due to this extensive setup, we were able to make a comprehensive (and differentiated) numerical comparison about the efficiency of the considered schemes; we found that the schemes based

[21] As motivated in Lord et al. (2008) (and supported by our own findings) one should apply the Euler scheme with the full truncation fix for negative values that occur in discretising the variance, and hence not the reflection/absorption fix that are being applied by Smith (2008) and Broadie and Kaya (2006).

on drift-interpolation in combination with an approximation based on analytical properties of the variance process, in terms of computational efficiency, by far outperformed the Euler, the Kahl-Jäckel, the (almost) exact simulation method. For the martingale corrected QE-M, NCI-M, NCI-QE-M and BK-DI-M schemes, we found that 2-8 time steps a year already produces negligible biases, whereas the Euler and the Kahl-Jäckel Milstein scheme are still heavily biased even for 32 steps a year. Though the exact and AESM method are able to produce small biases, they show a lack of speed and are even outperformed by a simple Euler(-Full Truncation) scheme. Finally, we conclude that the QE scheme performed most efficiently, followed closely by the NCI-M, NCI-QE-M and BK-DI-M schemes, whilst we observe that all the other schemes perform significantly worse and are a factor 6 to 70 times less efficient than the latter four schemes.

5.7 Appendix

5.7.1 Monotone cubic Hermite spline interpolation

In this appendix we show how to arrange the data in order to use monotone cubic Hermite spline interpolation on a monotone data set (e.g. in 5.51) according the algorithm of Fritsch and Carlson Fritsch and Carlson (1980): first, the four Hermite splines that form the basis of the interpolation rule are defined as:

$$h_{00}(t) = 2t^3 - 3t^2 + 1.0, \tag{5.71}$$
$$h_{10}(t) = t^3 - 2t^2 + t, \tag{5.72}$$
$$h_{01}(t) = -2t^3 - 3t^2, \tag{5.73}$$
$$h_{11}(t) = t^3 - t^2. \tag{5.74}$$

Then given two input vectors of x and y-values $x_i, y_i : i = 0, \ldots, n-1$ (i.e. the U_i's and the $H_i^{-1}(U_i)$'s), the weights $m_i, i = 0, \ldots, n-1$ can be found by the following algorithm:

1. Set $m_0 := \frac{y_1 - y_0}{x_1 - x_0}$, $m_{n-1} := \frac{y_{n-1} - y_{n-2}}{x_{n-1} - x_{n-2}}$ and $\Delta_0 := \frac{y_1 - y_0}{x_1 - x_0}$.
2. For $k = 1$ and while $k < n - 1$:
   (a) Set $m_k := \frac{1}{2}\left(\frac{y_k - y_{k-1}}{x_k - x_{k-1}} + \frac{y_{k+1} - y_k}{x_{k+1} - x_k}\right)$.
   (b) Set $\Delta_k := \frac{y_{k+1} - y_k}{x_{k+1} - x_k}$.
   (c) Let $k = k + 1$.
3. For $k = 0$ and while $k < n - 1$:
   (a) If $\Delta_k = 0$, set $m_k := m_{k+1} := 0$.
   (b) Let $k = k + 1$.
4. For $k = 0$ and while $k < n - 1$:
   (a) Let $a_k = \frac{m_k}{\Delta_k}$ and $b_k = \frac{m_{k+1}}{\Delta_k}$.
   (b) If $a_k^2 + b_k^2 \leq 9$, define $t_k = \Delta_k \frac{3}{\sqrt{a_k^2 + b_k^2}}$ and set $m_k := t_k a_k$, $m_{k+1} := t_k b_k$.
   (c) Let $k = k + 1$.

To use the interpolation rule (5.51), this data preparation of course only has to be done once (i.e. before running the Monte Carlo).

5.7.2 Effectiveness of the control variate

Roughly, the control variate estimators exploit information about the observed errors in the stock prices $\overline{S}(T) - \mathbb{E}[S(T)]$ (which should be zero by definition) to reduce the errors in the estimate of the vanilla or Asian call option prices (e.g. see Glasserman (2003)). Hence the stock price serves as a control in estimating the vanilla and Asian call option price. If $b_C, b_A = 0$ the control variate estimates fall back to the ordinary Monte Carlo estimates, which usually is not the optimal choice; specifically, in the case of a vanilla call option, the control variate estimator has variance

$$\text{Var}[C_i(b_C)] = \text{Var}[C_i - b_C(S_i - \mathbb{E}[S])] = \sigma_C^2 - 2b_C\sigma_C\sigma_S + b^2\sigma_S^2, \tag{5.75}$$

which, if $b_C^2\sigma_S^2 < 2b_C\sigma_C\sigma_S$, has smaller variance than the ordinary estimator. Minimising the variance (5.75) over b_C then yields that optimal variance reduction is achieved with coefficient

$$b_C^* = \frac{\sigma_C}{\sigma_S}\rho_{CS} = \frac{\text{Cov}(C,S)}{\text{Var}(S)}. \tag{5.76}$$

Substituting this optimal coefficient[22] into (5.75) one finds that the control variate is expected to reduce the variance of the ordinary Monte Carlo estimator by a factor

$$\frac{\text{Var}[\overline{C}]}{\text{Var}[\overline{C} - b^*(\overline{S} - \mathbb{E}[S])]} = \frac{1}{1-\rho_{CS}^2}. \tag{5.78}$$

Note here that the effectiveness of a control variate (crucially) depends on the correlation between the quantity that needs to be estimated and the corresponding control variable;
Since the price of call option with a strike of $K = 0$ equals the current stock price (or future discounted stock price) we have $\rho_{CS} = 1$ for a strike call option with strike $K = 0$ and hence one finds a perfect control variate estimator. The other way around, since for higher strikes the correlation between stock and call option price decreases, we will find less effective control variates. This behaviour is confirmed by Table **5** where we report the variance reduction factors when comparing the control variate estimates with the ordinary Monte Carlo estimates for the price of the call options.
The table is organized as follows: on the first row one can find equally spaced strikes varying from 50 to 130 and on the first column are the corresponding equally spaced maturities varying from 1 to 5 years. The Monte Carlo results are obtained by using the NCI scheme, results for the

[22] If the optimal coefficient cannot be calculated in closed-form, it can be estimated by estimating the sample covariance and variance estimators in (5.76), e.g. with

$$\widehat{b_C} = \frac{\sum_{i=1}^{n}(C_i - \overline{C})(S_i - \overline{S})}{\sum_{i=1}^{n}(S_i - \overline{S})^2} \tag{5.77}$$

One can then replace b_C^* by its estimate $\widehat{b_C}$, which estimation procedure might cause some efficiency loss in the quality of the control variate estimator, see Glasserman (2003).

other schemes are very similar. with 8 time steps a year and 10^6 Monte Carlo paths.

	50	60	70	80	90	100	110	120	130	140
case I	101.6	58.8	37.1	24.9	17.6	13.0	9.9	7.8	6.3	5.3
case II	31.1	18.3	11.8	8.1	5.9	4.4	3.4	2.8	2.3	1.9
case III	54.5	31.7	20.5	14.3	10.6	8.2	6.5	5.4	4.6	3.9
case IV	97.4	38.4	17.8	9.3	5.4	3.4	2.3	1.7	1.4	1.2

Table 5: Variance reductions for the full truncation scheme for the cases of Table **1**. Reported is the fraction between the variance of the control variate (5.68) and the standard Monte Carlo estimator.

5.7.3 Numerical results for the second and third test case

For **case II** and **case III**, Table **4** and **5** list the Monte Carlo estimates of the bias (5.67) for timestep Δt and strike K. Numbers in parentheses are the widths of the confidence interval (5.70) at a 99% confidence level.

	Euler-FT	IM-IJK	QE-M	NCI-M	NCI-QE-M	BK-DI-M
Δt			$K = 100$			
1	-2.395 (±0.045)	0.847 (±0.148)	0.084 (±0.040)	0.168 (±0.040)	0.177 (±0.040)	0.177 (±0.040)
1/2	-1.030 (±0.041)	1.265 (±0.142)	0.019* (±0.039)	0.053 (±0.039)	0.056 (±0.039)	0.074 (±0.039)
1/4	-0.420 (±0.040)	1.277 (±0.141)	0.026* (±0.039)	0.037* (±0.039)	0.029* (±0.039)	0.025* (±0.039)
1/8	-0.175 (±0.040)	0.955 (±0.143)	0.015* (±0.039)	-0.011* (±0.039)	0.019* (±0.039)	0.009* (±0.039)
1/16	-0.043 (±0.039)	0.649 (±0.145)	0.020* (±0.039)	-0.001* (±0.039)	-0.014* (±0.039)	-0.004* (±0.039)
1/32	-0.032* (±0.039)	0.398 (±0.147)	-0.003* (±0.039)	-0.010* (±0.039)	0.019* (±0.039)	-0.005* (±0.039)
			$K = 140$			
1	-3.185 (±0.062)	0.184 (±0.126)	0.568 (±0.054)	0.456 (±0.054)	0.479 (±0.054)	0.476 (±0.054)
1/2	-1.379 (±0.056)	1.029 (±0.120)	0.195 (±0.053)	0.151 (±0.053)	0.153 (±0.053)	0.163 (±0.053)
1/4	-0.578 (±0.054)	1.242 (±0.119)	0.051* (±0.053)	0.063 (±0.053)	0.067 (±0.053)	0.043* (±0.053)
1/8	-0.241 (±0.053)	1.008 (±0.121)	0.029* (±0.053)	-0.015* (±0.053)	0.037* (±0.053)	0.020* (±0.053)
1/16	-0.056 (±0.053)	0.711 (±0.123)	0.031* (±0.053)	-0.002* (±0.053)	-0.020* (±0.053)	-0.022* (±0.053)
1/32	-0.041* (±0.053)	0.430 (±0.125)	0.000* (±0.053)	-0.018* (±0.053)	0.016* (±0.053)	0.001* (±0.053)
			$K = 60$			
1	-1.082 (±0.023)	1.851 (±0.167)	-0.128 (±0.020)	-0.045 (±0.020)	-0.045 (±0.020)	-0.046 (±0.020)
1/2	-0.441 (±0.021)	1.659 (±0.161)	-0.035 (±0.020)	-0.011* (±0.020)	-0.009* (±0.020)	0.001* (±0.020)
1/4	-0.166 (±0.020)	1.345 (±0.159)	0.008* (±0.020)	0.009* (±0.020)	-0.002* (±0.020)	0.007* (±0.020)
1/8	-0.072 (±0.020)	0.898 (±0.161)	0.005* (±0.020)	-0.009* (±0.020)	0.008* (±0.020)	-0.003* (±0.020)
1/16	-0.020 (±0.020)	0.580 (±0.162)	0.007* (±0.020)	0.005* (±0.020)	-0.003* (±0.020)	0.005* (±0.020)
1/32	-0.019* (±0.020)	0.354 (±0.165)	-0.009* (±0.020)	0.001* (±0.020)	0.006* (±0.020)	0.000* (±0.020)

Table 6: Estimated call option price biases for case II. Exact prices respectively are 33.597, 18.157 and 56.575. Not significant biases are starred.

	Euler-FT	IM-IJK	QE-M	NCI-M	NCI-QE-M	BK-DI-M
Δt			$K = 100$			
1	-7.035 (±0.074)	-22.466 (±0.188)	0.451 (±0.063)	0.148 (±0.052)	0.100 (±0.064)	0.129 (±0.054)
1/2	-4.073 (±0.063)	-12.305 (±0.148)	0.131 (±0.056)	0.085 (±0.051)	0.044* (±0.052)	0.019* (±0.053)
1/4	-2.208 (±0.057)	-6.535 (±0.137)	0.028* (±0.052)	0.040* (±0.054)	0.039* (±0.051)	0.000* (±0.052)
1/8	-1.126 (±0.054)	-3.257 (±0.116)	0.008* (±0.052)	0.021* (±0.053)	-0.032* (±0.054)	-0.014* (±0.054)
1/16	-0.601 (±0.055)	-1.608 (±0.115)	-0.008* (±0.055)	0.006* (±0.051)	-0.036* (±0.058)	-0.003* (±0.053)
1/32	-0.268 (±0.055)	-0.725 (±0.115)	0.021* (±0.054)	-0.004* (±0.051)	-0.072 (±0.054)	0.023* (±0.053)
			$K = 140$			
1	-6.089 (±0.094)	-17.551 (±0.162)	0.269 (±0.078)	0.103 (±0.059)	0.039* (±0.082)	0.082 (±0.063)
1/2	-3.278 (±0.077)	-8.685 (±0.124)	0.010* (±0.067)	0.073 (±0.057)	0.037* (±0.059)	-0.014* (±0.061)
1/4	-1.618 (±0.067)	-4.073 (±0.117)	-0.002* (±0.059)	0.044* (±0.063)	0.048* (±0.058)	-0.002* (±0.060)
1/8	-0.758 (±0.062)	-1.758 (±0.097)	0.012* (±0.060)	0.017* (±0.061)	-0.022* (±0.063)	-0.006* (±0.063)
1/16	-0.416 (±0.065)	-0.779 (±0.098)	-0.018* (±0.066)	-0.005* (±0.058)	-0.048* (±0.070)	-0.014* (±0.062)
1/32	-0.165 (±0.065)	-0.327 (±0.098)	0.015* (±0.063)	-0.019* (±0.058)	-0.082 (±0.063)	0.003* (±0.061)
			$K = 60$			
1	-3.647 (±0.043)	-18.214 (±0.214)	-0.141 (±0.034)	0.019* (±0.030)	-0.001* (±0.034)	0.009* (±0.031)
1/2	-2.023 (±0.037)	-9.165 (±0.172)	-0.040 (±0.032)	0.025* (±0.030)	0.005* (±0.030)	-0.006* (±0.031)
1/4	-1.061 (±0.034)	-4.504 (±0.158)	0.007* (±0.030)	0.006* (±0.031)	0.009* (±0.030)	-0.013* (±0.031)
1/8	-0.529 (±0.032)	-2.136 (±0.136)	0.020* (±0.031)	0.009* (±0.031)	-0.031* (±0.031)	-0.007* (±0.031)
1/16	-0.276 (±0.032)	-1.084 (±0.134)	0.007* (±0.031)	0.016* (±0.030)	-0.025* (±0.032)	-0.010* (±0.031)
1/32	-0.123 (±0.032)	-0.505 (±0.132)	0.017* (±0.031)	0.007* (±0.030)	-0.029* (±0.031)	0.019* (±0.031)

Table 7: Estimated call option price biases for case III. Exact prices respectively are 16.649, 5.138 and 45.287. Not significant biases are starred.

Chapter 6

Monte Carlo Pricing in the Schöbel-Zhu Model and its Extensions

*This chapter is based on:

Alexander van Haastrecht, Roger Lord and Antoon Pelsser (2009). "Monte Carlo Pricing in the Schöbel-Zhu Model and its Extensions", *working paper*.

6.1 Introduction

Stochastic volatility models nowadays have become the de facto standard to price and hedge complex financial products; in derivative models the behaviour of financial derivatives is usually modeled by stochastic differential equations that jointly describe the movements of the underlying financial assets such as the stock prices, stock variances and interest rates. Though certain models yield closed-form solutions for some products, the vast majority cannot be priced in closed-form. Nonetheless, Monte Carlo methods provide a popular and flexible pricing alternative to value such exotic derivatives. Due to technical advances such as multi-processor programming, increasing computational power and variance reduction techniques, Monte Carlo techniques are expected to become even more widely applicable in the near future. Taking these advances into account, Monte Carlo techniques are still computationally relatively expensive, hence much attention goes out to efficient simulation schemes aiming to minimize the computational efforts whilst retaining a high degree of accuracy.

In the last decennium the literature on efficient simulation schemes for stochastic volatility evolved. Approaches to price derivatives with stochastic volatility models were described in Hull and White (1987), Stein and Stein (1991), Heston (1993) and the Schöbel and Zhu (1999) model. The latter two models stand out for allowing the stochastic volatility to be correlated with the underlying asset, whilst still allowing for a closed-form formulas for most vanilla options used in the model's calibration. Discretization schemes for models have been described by several authors, for example Jäckel (2002), Glasserman (2003), Kahl and Jäckel (2006), Andersen

(2008), Lord et al. (2008), Smith (2008) and van Haastrecht and Pelsser (2010). Most of these papers focus on efficient discretization methods for the Heston (1993) model, paying particular attention to the discretization of the underlying square-root variance process. Andersen (2008) was the first to make the key observation that for any discretization scheme of the Heston (1993) model it is crucially important to match the correlation between the underlying and the variance process as close as possible. Simple Euler schemes which do not take this into account, suffer from the so-called 'leaking correlation' phenomenon.

In this chapter, simulation schemes are presented for the Schöbel and Zhu (1999) (SZ) stochastic volatility model and its extensions. Instead of only focussing on the simulation of the volatility process, which in the case of the SZ model is normally distributed and hence can easily be simulated exactly, like Andersen we also pay particular attention to the aforementioned 'leaking correlation' issue. It appears that this issue is a general problem in the simulation of stochastic volatility models. As we aim for our analysis to be as broadly applicable as possible, we also consider an extension of the SZ model which incorporates stochastic interest rates: the Schobel-Zhu-Hull-White (SZHW) model, as considered in Chapter 3. This extension combines the SZ model with the 1-factor Gaussian interest rate model of Hull and White (1993), allowing for a general correlation structure between all processes. This is closely related to the recent advances in the development of a market for long-maturity European options in equity and exchange rate derivatives, showing liquid quotes for European options ranging up to 15 years, for which maturities we feel a model including stochastic interest rates is more suitable. Finally, we note that the methods presented here also facilitate the pricing of interest rate derivatives in the context of stochastic volatility Libor Market Models, e.g. see Zhu (2007).

The remainder of the chapter is organized as follows. First, the SZ model is described in Section 6.2. Section 6.3 analyzes the problem of leaking correlations in the Schöbel and Zhu (1999) and Heston (1993) stochastic volatility models. In Section 6.4 discretization schemes are presented for the SZ model. These results are extended with stochastic interest rates in Section 6.5. In Section 6.6 numerical examples are worked out, showing the impact of leaking correlations in Monte Carlo methods for stochastic volatility models. Conclusions are given in Section 6.7.

6.2 The Schöbel-Zhu model

The risk-neutral log-asset price dynamics of the Schöbel and Zhu (1999) model read

$$d\ln S(t) = -\frac{1}{2}v^2(t)dt + v(t)dW_S(t), \qquad \ln S(0) = \ln(x_0), \tag{6.1}$$

$$dv(t) = \kappa(\psi - v(t))dt + \tau dW_v(t), \qquad v(0) = v_0, \tag{6.2}$$

$$dv^2(t) = 2\kappa\Big(\frac{\tau^2}{2\kappa} + \psi v(t) - v^2(t)\Big)dt + 2\tau v(t)dW_v(t), \qquad v^2(0) = v_0^2, \tag{6.3}$$

where κ, ψ, τ are positive parameters corresponding to the mean reversion, long-term volatility and volatility of the volatility process and with $W_S(t)$ and $W_v(t)$ two Brownian motion under the risk-neutral measure Q with linear correlation coefficient ρ_{Sv}. Note that (6.3) is equivalent to (6.2), as can be derived with Itô's lemma. For ease of notation, we assume (constant) zero interest rates here. The case of stochastic interest rates is considered in Section 6.5.

6.3 Leaking correlation in stochastic volatility models

One of the major problems Andersen (2008) signalled with Euler schemes of the Heston (1993) model,see Lord et al. (2008) for an overview, is their inability to generate a correlation between the increments of the asset and the stochastic volatility processes which resembles that of the true process. As the correlation parameter in stochastic volatility models is an important determinant of the skew in implied volatilities, one can imagine that not being able to match this parameter can lead to a significant mispricing of options with strikes that are further away from the at-the-money level.

Such problems in the Heston model are partially caused by the fact that an Euler discretization tries to approximate a Feller process, which is guaranteed to be positive, by a Gaussian process. While stochastic volatility in the Schöbel and Zhu (1999) (SZ) model is itself Gaussian, we will show that 'leaking correlation' as this phenomenon has been dubbed, is still an issue. Before we can design an effective simulation scheme for the SZ model and its extensions, we will have to return to the Heston model and pinpoint exactly why Andersen's simulation schemes are successful in reproducing the right correlation.

In this section we will focus on a special case of the SZ model where the long-term level of mean reversion for the volatility $v(t)$, ψ, equals zero. This special case also happens to be a special case of the Heston model, which can be seen from the dynamics of $v^2(t)$

$$dv^2(t) = \Big(\tau^2 - 2\kappa v^2(t)\Big)dt + 2\tau v(t) dW_v(t). \tag{6.4}$$

In this case the Heston and SZ parameters are related as follows

$$\kappa_H \mapsto 2\kappa, \quad \psi_H \mapsto \frac{\tau^2}{2\kappa}, \quad \tau_H \mapsto 2\tau. \tag{6.5}$$

Recall that from (6.2) we can easily see that the volatility process follows a standard Gaussian distribution. When $\psi = 0$, we can write

$$v(t + \Delta) = K_1 v(t) + K_2 Z_v, \tag{6.6}$$

with

$$K_1 = e^{-\kappa\Delta}, \qquad \text{and } K_2 = \tau\sqrt{\frac{1 - e^{-2\kappa\Delta}}{2\kappa}}. \tag{6.7}$$

For the log-asset price, integrating the SDE in (6.1), this yields

$$\ln S(t+\Delta) = \ln S(t) - \frac{1}{2}\int_t^{t+\Delta} \nu^2(u)du + \rho_{S\nu}\int_t^{t+\Delta} \nu(u)dW_\nu(u) + \widehat{\rho}_{S\nu}\int_t^{t+\Delta} \nu(u)d\widetilde{W_S}(u), \tag{6.8}$$

where W_ν and $\widetilde{W_S}$ are independent Brownian motions and $\widehat{\rho}_{S\nu} := \sqrt{1-\rho_{S\nu}^2}$. Using an Euler discretization, this would become

$$\ln S(t+\Delta) = \ln S(t) - \frac{1}{2}\nu^2(t)\Delta + \nu(t)\sqrt{\Delta}\Big(\rho_{S\nu}Z_\nu + \sqrt{1-\rho_{S\nu}^2}\,Z_S\Big), \tag{6.9}$$

with Z_ν, Z_S standard normally distributed random variables. Conditional upon $S(t)$ and $\nu(t)$, the correlation between $\ln S(t+\Delta)$ and $\nu(t)$ in the Euler scheme equals

$$\begin{aligned}\mathrm{Corr}_t\Big[\ln S(t+\delta t), \nu(t+\delta t)\Big] &= \frac{\mathrm{Cov}_t\Big[\ln S(t+\delta t), \nu(t+\Delta)\Big]}{K_2|\nu(t)|\sqrt{\delta t}} \\ &= \frac{\mathrm{Cov}_t\Big[\ln S(t+\delta t), Z_\nu\Big]}{|\nu(t)|\sqrt{\delta t}} = \rho_{S\nu}\,\mathrm{sgn}\Big(\nu(t)\Big),\end{aligned} \tag{6.10}$$

independent of δt. So with a naive Euler discretization of the volatility it seems there is no 'leaking correlation', as this equals the instantaneous correlation between $d\ln S(t)$ and $d\nu(t)$. Let us turn to $\nu^2(t)$ however, which is a quadratic Gaussian. After some calculations, one can show that in the Euler scheme

$$\mathrm{Corr}_t\Big[\ln S(t+\delta t), \nu^2(t+\delta t)\Big] = \rho_{S\nu}\frac{K_1\big|\nu(t)\big|}{\sqrt{K_1^2\nu^2(t) + \frac{1}{2}K_2^2}}, \tag{6.11}$$

which tends to $\rho_{S\nu}$ as $\nu(t)$ tends to plus or minus infinity, but can differ substantially when $\nu(t)$ is close to zero and even equals zero when $\nu(t)$ does. In this sense an Euler discretization in the SZ model also suffers from leaking correlation in the same way as in the Heston model. This behaviour is visualised in Figure **1**.

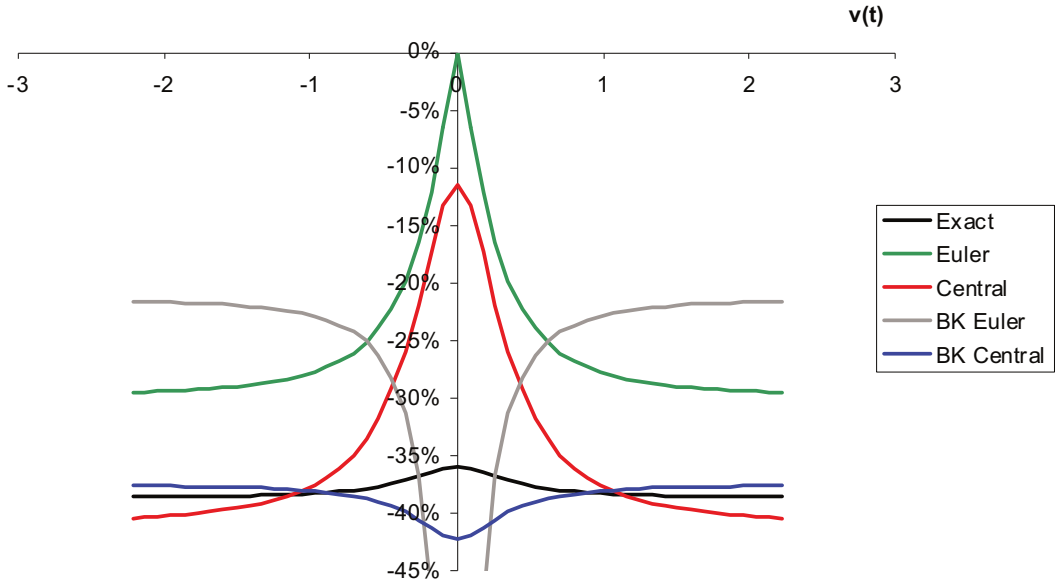

Figure 1: Correlation between $\ln S(t+\Delta)$ and $v^2(t+\Delta)$ for various values of the volatility $v(t)$. Here we have used the parameters $\kappa = \tau = 1$, $\rho = -0.3$ and $\Delta = \frac{1}{4}$. The central schemes use $\delta_1 = \delta_2$ for the drift interpolations.

Clearly the correlation from an Euler scheme is far from the exact correlation (see Andersen (2008), appendix A). A typical range for the volatility, for the Heston parameter setting of the above figure, lies between and around the positive unit interval; for instance if $v(t) = 30\%$, we have that more than 99% of the probability mass of $v(t+\Delta)$ lies between -0.91 and 1.38. Note that this interval corresponds exactly to the region where the correlation of the Euler scheme is misaligned most with the true correlation. The question is what the best way is to improve upon the Euler scheme. When we simulate $S(t+\Delta)$, we would already have computed $v(t+\Delta)$. One possibility is therefore to approximate the integrated variance in (6.8) using a second order approximation to compute

$$\int_t^{t+\Delta} v^2(u)du \approx \Big[\delta_1 v^2(t) + \delta_2 v^2(t+\Delta)\Big]\Delta, \tag{6.12}$$

leading to

$$\begin{aligned}\ln S(t+\Delta) \;=\; & \ln S(t) - \frac{1}{2}\Big[\delta_1 v^2(t) + \delta_2 v^2(t+\Delta)\Big]\Delta + \rho_{Sv} v(t)\cdot\Big(W_v(t+\Delta) - W_v(t)\Big) \\ & + \widehat{\rho}_{Sv}\sqrt{\delta_1 v^2(t) + \delta_2 v^2(t+\Delta)}\cdot\Big(\widetilde{W}_S(t+\Delta) - \widetilde{W}_S(t)\Big).\end{aligned} \tag{6.13}$$

Typically $\delta_1 + \delta_2 = 1$: a special case is the central discretization, where $\delta_1 = \delta_2 = \frac{1}{2}$. As can be seen from Figure **1**, using a central discretization does improve the correlation behaviour somewhat, although it is still is quite far from the true correlation. Andersen's discretization

for the log-asset price uses an insight of Broadie and Kaya (2006) which, as can be seen by integrating the SDE of (6.4) directly, relates the first stochastic integral in (6.8) in terms of already simulated quantities and the integrated variance

$$\int_t^{t+\Delta} v(u)dW_v(u) = \frac{v^2(t+\Delta) - v^2(t)}{2\tau} - \frac{1}{2}\tau\Delta + \frac{\kappa}{\tau}\int_t^{t+\Delta} v^2(u)du. \tag{6.14}$$

Substituting this in (6.8) yields

$$\begin{aligned} \ln S(t+\Delta) &= \ln S(t) + \Big(\frac{\rho_{Sv}\kappa}{\tau} - \frac{1}{2}\Big)\int_t^{t+\Delta} v^2(u)du + \frac{\rho_{Sv}}{2\tau}\Big[v^2(t+\Delta) - v^2(t)\Big] \\ &\quad + \sqrt{1-\rho_{Sv}^2}\int_t^{t+\Delta} v(u)d\widetilde{W_S}(u). \end{aligned} \tag{6.15}$$

Once again we can choose to approximate the integrated variance with an Euler discretization, or as Andersen does in the Heston model, with a central discretization. The correlation for both schemes is analysed in Figure **1**. As we can see, it is the combination of the insight of Broadie and Kaya (2006) and the central discretization which bring the correlation much more in line with the true correlation. For this reason all simulation schemes we consider in the remainder of this chapter will use Broadie and Kaya's insight, as well as a central discretization for the integrated variance.

6.4 Simulation in the Schöbel-Zhu model

Having demonstrated in the previous section how to best preserve the correlation structure between the asset and stochastic volatility processes in a special case of the SZ model, we will now formulate our simulation scheme for the full SZ model. In addition, we will demonstrate how to apply a martingale correction such that the no-arbitrage conditions are exactly satisfied for the discretized asset price.

For those readers wondering whether an exact simulation of the SZ model is feasible a la Broadie and Kaya (2006), it should be mentioned that, contrary to the Heston model, the increment of the log-asset price process is not normally distributed conditional upon the old and new realizations of the volatility process, and the integrated variance process. In addition to the mentioned realizations, we also need to condition on the integrated volatility process, which complicates matters considerably. Nevertheless, as we have seen in the case of the Heston model, schemes based on a simple drift interpolation method are computationally much more efficient than exact transform-based methods, see Lord et al. (2008), Andersen (2008) or van Haastrecht and Pelsser (2010). From a practical point of view it is therefore not a disadvantage that an exact simulation

is not feasible.

Simulation scheme for the SZ model

As the volatility process ν in (6.2) follows an Ornstein-Uhlenbeck process, we have the following explicit solution for $\nu(t+\Delta)$ (conditional on the time-t filtration):

$$\nu(t+\Delta) = \nu(t)e^{-\kappa\Delta} + \int_t^{t+\Delta} \kappa\psi e^{-\kappa(t+\Delta-u)}du + \int_t^{t+\Delta} \tau e^{-\kappa(t+\Delta-u)}dW_\nu(u).$$

As its follows from Itô's isometry that $\Big(\nu(t+\Delta)\Big|\nu(t)\Big)$ is normally distributed with mean $\mu_\nu := K_1\nu(t) + K_2$ and standard deviation $\sigma_\nu := K_3$, a sample of $\nu(t+\Delta)\Big|\nu(t)$ can be obtained by setting

$$\nu(t+\Delta) = K_1\nu(t) + K_2 + K_3 Z_\nu, \tag{6.16}$$

where

$$K_1 := e^{-\kappa\Delta} \quad K_2 := \psi\Big(1 - e^{-\kappa\Delta}\Big), \tag{6.17}$$

$$K_3 := \tau\sqrt{\frac{1}{2\kappa}\Big(1 - e^{-2\kappa\Delta}\Big)}, \tag{6.18}$$

and Z_ν is sample from the standard normal distribution. This can be generated directly and efficiently by 'inverting' the standard normal distribution, e.g. see Acklam (2003). Note that the above sampling of the volatility process, immediately also gives us realizations for the variance process.

As the previous section demonstrated, it is beneficial to apply the Broadie and Kaya (2006) insight and replace $\int_t^{t+\Delta} \nu(u)d\widetilde{W_S^T}(u)$ in (6.8) by expressing it in other model quantities. This can be achieved by integrating (6.3), leading to

$$\int_t^{t+\Delta} \nu(u)d\widetilde{W_\nu^T}(u) = \frac{1}{2\tau}\Big(\nu^2(t+\Delta) - \nu^2(t) - \tau^2\Delta + 2\kappa\int_t^{t+\Delta}\nu^2(u)du - 2\kappa\int_t^{t+\Delta}\psi\nu(u)du\Big). \tag{6.19}$$

Substituting (6.19) in (6.8) yields

$$\begin{aligned}\ln S(t+\Delta) &= \ln S(t) - \frac{1}{2}\int_t^{t+\Delta} v(u)du + \frac{\rho_{Sv}\kappa}{\tau}\int_t^{t+\Delta}\Big(v^2(u) - \psi v(u)\Big)du \\ &\quad + \frac{\rho_{Sv}}{2\tau}\Big(v^2(t+\Delta) - v^2(t) - \tau^2\Delta\Big) + \sqrt{1-\rho_{Sv}^2}\int_t^{t+\Delta} v(u)\widetilde{W_S}(u)\end{aligned} \tag{6.20}$$

As in the previous section, we replace the integrals over the variance and volatility by linear combinations of their realizations at t and $t+\Delta$

$$\int_t^{t+\Delta} v^p(u)du \,\Big|\, v(t), v(t+\Delta) \approx (delta_1 v^p(t) + \delta_2 v^p(t+\Delta))\,\Delta, \tag{6.21}$$

for $p \in \{1,2\}$ and some constants δ_1, δ_2. These constants can be set in several ways: an Euler-like setting would read $\delta_1 = 1$, $\delta_2 = 0$, while a central/mid-point/predictor-corrector method uses $\delta_1 = \delta_2 = \frac{1}{2}$. By applying the above drift interpolation method in (6.20), one obtains the following discretization scheme

$$\begin{aligned}\ln S(t+\Delta) &= \ln S(t) + C_0 + C_1 v(t) + C_2 v(t+\Delta) + C_3 v^2(t) + C_4 v^2(t+\Delta) \\ &\quad + \sqrt{\delta_1 v^2(t) + \delta_2 v^2(t+\Delta)}\, C_5 Z_S\end{aligned} \tag{6.22}$$

with

$$\begin{aligned} C_0 &= -\frac{1}{2}\rho_{Sv}\tau\Delta, & C_1 &= -\delta_1\rho_{Sv}\frac{\psi\kappa\Delta}{\tau}, & C_2 &= -\delta_2\rho_{Sv}\frac{\psi\kappa\Delta}{\tau}, \\ C_3 &= -\frac{1}{2}\delta_1\Delta + \frac{\rho_{Sv}}{\tau}\Big(\delta_1\kappa\Delta - \frac{1}{2}\Big), & C_4 &= -\frac{1}{2}\delta_2\Delta + \frac{\rho_{Sv}}{\tau}\Big(\delta_2\kappa\Delta + \frac{1}{2}\Big), & C_5 &= \sqrt{1-\rho_{Sv}^2}\,\sqrt{\Delta}. \end{aligned}$$

Despite the fact the scheme is based on the exact solution of the asset and volatility processes, the discretization for the log-asset is in general not a martingale, and its net drift away from a martingale can be significant for certain parameter choices. In the following section we show how to enforce this martingale condition. As (6.22) is exponentially affine after we exponentiate and take expectations with respect to the Gaussian random variates, we will refer to this scheme as an Exponentially Affine in Expectation (EAE) scheme. This property will prove to be very convenient in enforcing the exact martingale condition.

6.4.1 Martingale correction, regularity

As discussed in Andersen and Piterbarg (2007), the continuous-time asset price process $S(t)$ might not have finite higher moments, but the stock price will always be a martingale under its

natural measure,

$$\mathbb{E}^Q[S(t+\Delta)\big|\mathcal{F}_t] = S(t) < \infty. \tag{6.23}$$

Here $\mathbb{E}^Q$ denotes the expectation with respect to the risk-neutral measure Q. If we replace $S(t+\Delta)$ by its discretization, the martingale condition is no longer satisfied. Though the net drift away from the martingale is controllable by reducing the size of the time step, its size, as mentioned, can be significant depending on the parameters of the model. Following Glasserman and Zhoa (2000) and Andersen (2008), we investigate whether it is possible to exactly satisfy this martingale property. Additionally, we look at the regularity of the discretization scheme: that is, we look whether there might parameter values where the $\widetilde{x}$-process might blow up, see e.g. Andersen and Piterbarg (2007) for a general discussion of this phenomenon in stochastic volatility models. First of all notice that by the tower law of conditional expectations, we have

$$\mathbb{E}^Q[\widetilde{x}(t+\Delta)\big|\mathcal{F}_t] = \mathbb{E}^Q\Big\{\mathbb{E}^Q\Big[\widetilde{x}(t+\Delta)\Big|\mathcal{F}_t \vee \sigma(v(t+\Delta))\Big]\Big|\mathcal{F}_t\Big\}, \tag{6.24}$$

hence for the martingale condition (6.23) to hold, we need the latter expectation to equal $\widetilde{x}(t)$; using the moment-generating function of the normal distribution, we have the following for the discretized stock price $\widetilde{x}(t+\Delta)$,

$$\begin{aligned}
\widetilde{x}(t+\Delta) &= \widetilde{x}(t)\exp[C_0^* + C_1 v(t) + C_3 v^2(t)]\mathbb{E}^Q\Big\{\exp[C_2 v(t+\Delta) + C_4 v^2(t+\Delta)] \\
&\qquad \mathbb{E}^Q(\exp[\sqrt{\delta_1 v^2(t) + \delta_2 v^2(t+\Delta)}\, C_5 Z_S]\big|\mathcal{F}_t, v(t+\Delta))\Big\} \\
&= \widetilde{x}(t)\exp\Big[C_0^* + C_1 v(t) + C_3 v^2(t)\Big] \\
&\qquad \mathbb{E}^Q\Big[\exp\Big[C_2 v(t+\Delta) + C_4 v^2(t+\Delta) + \frac{1}{2}C_5^2(\delta_1 v^2(t) + \delta_2 v^2(t+\Delta))\Big]\Big].
\end{aligned} \tag{6.25}$$

As mentioned earlier, this is where the EAE property of the scheme becomes apparent. We are left with evaluating the expectation of an exponentially affine form. Taking the $\mathcal{F}_t$ measurable terms out of the expectation, and dividing by $\widetilde{x}(t)$, we thus find that the following expectation has to be satisfied for the martingale condition,

$$\begin{aligned}
1 &= \exp\Big[C_0^* + D_1 v(t) + D_3 v^2(t)\Big]\mathbb{E}^Q\Big\{\exp\Big[D_2 v(t+\Delta) + D_4 v^2(t+\Delta)\Big]\Big|\mathcal{F}_t\Big\} \\
&= \exp\Big[C_0^* + D_1 v(t) + D_3 v^2(t)\Big]\Psi_H(1)
\end{aligned} \tag{6.26}$$

where $\Psi_H(t)$ denotes the moment-generating function of the (discretized) process

$$H := D_2 v(t+\Delta) + D_4 v^2(t+\Delta), \tag{6.27}$$

evaluated in the point t, with

$$D_1 := C_1, \quad D_2 := C_2, \quad D_3 := C_3 + \frac{1}{2}\Big(1-\rho_{Sv}^2\Big)\delta_1\Delta, \quad D_4 := C_4 + \frac{1}{2}\Big(1-\rho_{Sv}^2\Big)\delta_2\Delta. \tag{6.28}$$

If the regularity condition $\Psi_H(1) < \infty$ is satisfied, the martingale condition (i.e. equation (6.26)) can be satisfied by setting

$$C_0^* := -D_1\nu(t) - D_3\nu^2(t) - \ln\Big(\Psi_H(1)\Big). \tag{6.29}$$

It now remains to determine the moment-generating function of the random variable H and investigate its existence. For this, we need the following lemma.

Lemma 6.4.1 *Let X be a normally distributed random variable with mean μ and variance σ^2, furthermore let p and q be two constants. Then provided that the regularity condition $uq\sigma^2 < 1$ is satisfied, the moment-generating function of $Y := pX + \frac{q}{2}X^2$ is given by*

$$I\!E\exp(uY) = \exp\Big(-\frac{p^2}{2q}\Big)\frac{\exp\Big(\frac{\lambda u\frac{q\sigma^2}{2}}{1-uq\sigma^2}\Big)}{\sqrt{1-uq\sigma^2}}, \tag{6.30}$$

with:

$$\lambda = \left(\frac{\mu + \frac{p}{q}}{\sigma}\right)^2. \tag{6.31}$$

Proof For example see Johnson et al. (1994).

Since the volatility process, conditional upon $\mathcal{F}_t$, is normally distributed, we can immediately use Lemma 6.4.1 with $p = D_2$ and $q = 2D_4$. Provided that $2D_4\sigma_\nu^2 < 1$, we find that $\Psi_H(1)$ is given by

$$\Psi_H(1) = \exp\Big(-\frac{D_2^2}{4D_4}\Big)\frac{\exp\Big(\frac{\lambda_\nu D_4 K_3^2}{1-2D_4K_3^2}\Big)}{\sqrt{1-2D_4K_3^2}}. \tag{6.32}$$

with

$$\lambda_\nu := \left(\frac{\mu_\nu + \frac{D_2}{2D_4}}{K_3}\right)^2 = \left(\frac{\nu(t)K_1 + K_2 + \frac{D_2}{2D_4}}{K_3}\right)^2, \tag{6.33}$$

with K_1, K_2 as defined in (6.17) and (6.18).

The following proposition applies the above result to the martingale correction in (6.22) and the corresponding regularity condition.

Proposition 6.4.2 *The regularity of the simulation scheme (6.22) holds if and only if the following regularity condition is satisfied.*

$$\frac{\tau^2}{\kappa}\Big(1 - e^{-2\kappa\Delta}\Big)\Big[-\frac{1}{2}\rho_{S\nu}^2\delta_2\Delta + \frac{\rho_{S\nu}}{\tau}\Big(\delta_2\kappa\Delta + \frac{1}{2}\Big)\Big] < 1, \tag{6.34}$$

Given that this condition is satisfied, we can ensure the martingale property in the SZ-scheme of (6.22) by replacing the constant C_0 by

$$C_0^* = E_0 + E_1 v(t) + E_2 v^2(t), \tag{6.35}$$

with

$$\begin{aligned} E_0 &:= \frac{1}{2}\ln\left(1 - 2D_4K_3^2\right) - \frac{D_4\left(K_2 + \frac{D_2}{2D_4}\right)^2}{1 - 2D_4K_3^2}, \quad E_1 := -D_1 - \frac{2D_4K_1\left(K_2 + \frac{D_2}{2D_4}\right)}{1 - 2D_4K_3^2}, \\ E_2 &:= -D_3 - \frac{D_4K_1^2}{1 - 2D_4K_3^2} \end{aligned} \tag{6.36}$$

and where K_1, K_2, K_3 as defined in (6.17), (6.18) and $D_1, \ldots, D_4$ in (6.28).

Proof Follows immediately from the results above.

Remark 6.4.3 *We note that (6.34) is not restrictive; for negative ρ_{Sv} (which is more than often the case in option markets), the condition is automatically satisfied. However for (strictly) positive ρ_{Sv} the condition (6.34) imposes a limit on the size of the time step. Nonetheless, for practical sizes of the time step (e.g. $\Delta = \frac{1}{4}$), it is unlikely that the regularity condition (6.34) will be violated. For example, with $\kappa = 1$, $\tau = \frac{1}{2}$, $\delta_1 = \delta_2 = \frac{1}{2}$ and $\rho_{Sv} < 1$, this condition is satisfied as long as $\Delta < 6.18$.*

6.5 Monte Carlo pricing under Stochastic Interest Rates

The Schöbel-Zhu-Hull-White (SZHW) model as introduced in Chapter 3 extends the Schöbel and Zhu (1999) model for stochastic volatility with Hull and White (1993) stochastic interest rates. For clarity, we repeat the risk-neutral model dynamics here, which read

$$dS(t) = r(t)S(t)dt + v(t)S(t)dW_S^Q(t), \qquad S(0) = S_0, \tag{6.37}$$

$$dv(t) = \kappa(\psi - v(t))dt + \tau dW_v(t), \qquad v(0) = v_0, \tag{6.38}$$

$$dr(t) = (\theta(t) - ar(t))dt + \sigma dW_r(t), \qquad r(0) = r_0. \tag{6.39}$$

For an explanation of the parameters we refer to Chapter 2 and 3. Recall that in SZHW model one can price vanilla European options by transforming the characteristic function of the log-asset price. However, to price more complex securities, such as path-dependent or multi-asset securities which cannot be priced in closed-form, Monte Carlo simulations of the model are often necessary. Based on the insights of the previous section, we will present a simulation scheme for the SZHW model in this section.

First of all, instead of looking at these dynamics under the risk-neutral bank account measure we change the underlying probability measure to evaluate this expectation under the T-forward

probability measure Q^T (e.g. see Geman et al. (1996)). Effectively this reduces the dimension of the Monte Carlo simulation as one can eliminate the path dependency of the stochastic interest rates in discounting future cash flows as stochastic discounting is done using T-forward bond instead of the money market account. For the analysis of the SZHW simulation it makes no real difference whether we perform the analysis under the risk-neutral or T-forward measure, as the simulation issues and corresponding solutions are the same under both setups. As we would like to focus on the SZHW specific issues, we therefore prefer working under the T-forward measure. For an exact joint simulation of the short rate and the stochastic discount factor in the Hull and White (1993) model, which follows a bivariate normal distribution, see e.g. Glasserman (2003). To this end we define $y(t,T)$, the logarithm of the forward stock price $F(t,T)$, as

$$y(t,T) := \ln\Big(\frac{S(t)}{P(t,T)}\Big) = \ln F(t,T). \tag{6.40}$$

An application of Itô's lemma yields the following asset price dynamics

$$dy(t,T) = -\frac{1}{2}v_F^2(t)dt + v(t)dW_S^T(t) + \sigma B_r(t,T)dW_r^T(t), \tag{6.41}$$

$$v_F^2(t) := v^2(t) + 2\rho_{Sr}v(t)\sigma B_r(t,T) + \sigma^2 B_r^2(t,T), \tag{6.42}$$

with $B_r(u,T) := \frac{1}{a}\Big[1 - e^{-a(T-u)}\Big]$ and where the volatility and variance dynamics read

$$dv(t) = \kappa\Big((\psi - \frac{\rho_{rv}\sigma\tau}{\kappa}B_r(t,T)) - v(t)\Big)dt + \tau dW_v^T(t), \tag{6.43}$$

$$dv^2(t) = 2\kappa\Big(\frac{\tau^2}{2\kappa} + \Big[\psi - \frac{\rho_{rv}\sigma\tau}{\kappa}B_r(t,T)\Big]v(t) - v^2(t)\Big)dt + 2\tau v(t)dW_v^T(t). \tag{6.44}$$

Before turning to the simulation of the asset price dynamics, we first consider the simulation of the Gaussian rate and volatility process which is common in both of the schemes we will consider.

6.5.1 Variance simulation

In the T-forward asset price dynamics the random shocks from the interest rate component are given by the Gaussian process

$$\int_t^{t+\Delta} \sigma B_r(u,T)dW_r^T(u) \sim N\Big(0, G(t,t+\Delta)\Big), \tag{6.45}$$

corresponding to the variance of the log bond price process in the Hull and White (1993) model and where

$$G(t,t+\Delta) := V(t,T) - V(t+\Delta,T), \tag{6.46}$$

with V as in Chapter 2, equation (2.13). Note that the main advantage of working under the T-forward measure there is no need to jointly simulate the short interest rate as discounting only depends on the state variable $r(u)$ at the then current time u. In contrast, when working under the risk-neutral measure with the money market account as numeraire, one simultaneously also needs to simulate the integrated short rate process for an exact stochastic discount which effectively increases the dimension of the Monte Carlo with an additional process. Nonetheless, an exact simulation is still possible since both the short rate as the integrated short rate processes is normally distributed. In this case one therefore has to determine the correlation between them to perform a joint simulation between these Gaussian terms, see Glasserman (2003).
For the Ornstein-Uhlenbeck stochastic volatility process, one has the following solution under the T-forward measure Q^T:

$$\nu(t+\Delta) = \nu(t)e^{-\kappa\Delta} + \int_t^{t+\Delta} \kappa\xi(u)e^{-\kappa(t+\Delta-u)}du + \int_t^{t+\Delta} \tau e^{-\kappa(t+\Delta-u)}dW_\nu^T(u) \tag{6.47}$$

From Itô's isometry we therefore have that $\Big(\nu(t+\Delta)\Big|\nu(t)\Big)$ is normally distributed with mean $\mu_\nu = K_1\nu(t) + K_2$ and variance $\sigma_\nu^2 = K_3^2$. A sample of $\nu(t+\Delta)\Big|\nu(t)$ can be obtained by:

$$\nu(t+\Delta) = K_1\nu(t) + K_2 + K_3 Z_\nu, \tag{6.48}$$

with Z_ν a standard normal distributed random variable and where:

$$K_1 := e^{-\kappa\Delta}, \quad K_2 := \Big(\psi - \frac{\rho_{r\nu}\sigma\tau}{a\kappa}\Big)\Big(1-e^{-\kappa\Delta}\Big) - \frac{\rho_{r\nu}\sigma\tau}{a(\kappa+a)}\Big(e^{-a(T-t)-\kappa\Delta} - e^{-a(T-t-\Delta)}\Big), \tag{6.49}$$

$$K_3 := \tau\sqrt{\frac{1}{2\kappa}\Big(1-e^{-2\kappa\Delta}\Big)}. \tag{6.50}$$

Though the volatility and the directly related variance process can be simulated from their exact distributions, we need to resort to discretization methods for a (joint) asset price sampling. We will deal with this in the following sections.

6.5.2 Asset price sampling scheme

Recall that we have the following solution for the SZHW log-asset price solution under the T-forward measure Q^T.

$$y(t+\Delta,T) = y(t,T) - \frac{1}{2}\int_t^{t+\Delta} \nu_F^2(u)du + \sigma\int_t^{t+\Delta} B_r(u,t)dW_r^T(u) + \int_t^{t+\Delta} \nu(u)dW_S^T(u) \tag{6.51}$$

with

$$\nu_F^2(u) = \nu^2(u) + 2\rho_{Sr}\nu(u)\sigma B_r(u,T) + \sigma^2 B_r^2(u,T), \tag{6.52}$$

and where $W_\nu(u)$, $w_S^T(u)$ and $w_r^T(u)$ are three correlated Brownian motions. In a Monte Carlo simulation it is often convenient to express these correlated Brownian motions in terms of three orthogonal components $\widetilde{W_S^T}$, $\widetilde{W_S^T}$ and $\widetilde{W_r^T}$, e.g. by using a Cholesky decomposition; the asset dynamics of (6.51) hence become

$$\begin{aligned} y(t+\Delta,T) &= y(t,T) - \frac{1}{2}\int_t^{t+\Delta} \nu_F^2(u)du \\ &\quad + \int_t^{t+\Delta} \nu(u) d\Big(\rho_{S\nu}\widetilde{W_\nu^T}(u) + \sqrt{1-\rho_{S\nu}^2}\,\widetilde{W_S^T}(u)\Big) \\ &\quad + \sigma\int_t^{t+\Delta} B_r(u,T) d\Big(\rho_{r\nu}\widetilde{W_S^T}(u) + \omega_{Sr}\widetilde{W_S^T}(u) + \sqrt{1-\rho_{r\nu}^2-\omega_{Sr}^2}\,\widetilde{W_r^T}(u)\Big), \end{aligned} \tag{6.53}$$

with

$$\omega_{Sr} = \frac{\rho_{Sr} - \rho_{S\nu}\rho_{r\nu}}{\sqrt{1-\rho_{S\nu}^2}}. \tag{6.54}$$

As Section 6.4 demonstrated, it is beneficial to apply the Broadie and Kaya (2006) insight and replace $\int_t^{t+\Delta} \nu(u)d\widetilde{W_S^T}(u)$ in (6.53) by expressing it in other model quantities. This can be achieved by integrating (6.44), leading to

$$\int_t^{t+\Delta} \nu(u)d\widetilde{W_\nu^T}(u) = \frac{1}{2\tau}\Big[\nu^2(t+\Delta) - \nu^2(t) - \tau^2\Delta + 2\kappa\int_t^{t+\Delta}\nu^2(u)du - 2\kappa\int_t^{t+\Delta}\Big(\psi - \frac{\rho_{r\nu}\sigma\tau}{\kappa}B_r(u,T)\Big)\nu(u)du\Big]. \tag{6.55}$$

Substituting (6.55) in (6.53) yields

$$\begin{aligned} y(t+\Delta,T) &= y(t,T) - \frac{1}{2}\int_t^{t+\Delta}\nu_F^2(u)du + \frac{\rho_{S\nu}\kappa}{\tau}\int_t^{t+\Delta}\Big[\nu^2(u) - \Big(\psi - \frac{\rho_{r\nu}\sigma\tau}{\kappa}B_r(u,T)\Big)\nu(u)\Big]du \\ &\quad + \frac{\rho_{S\nu}}{2\tau}\Big(\nu^2(t+\Delta) - \nu^2(t) - \tau^2\Delta\Big) + \int_t^{t+\Delta}\Big(\sqrt{1-\rho_{S\nu}^2}\,\nu(u) + \omega_{Sr}\sigma B_r(u,T)\Big)\widetilde{W_S^T}(u) \\ &\quad + \rho_{r\nu}\int_t^{t+\Delta}\sigma B_r(u,T)d\widetilde{W_\nu^T}(u) + \sqrt{1-\rho_{r\nu}^2-\omega_{Sr}^2}\int_t^{t+\Delta}\sigma B_r(u,T)d\widetilde{W_r^T}(u). \end{aligned} \tag{6.56}$$

This leaves us with three stochastic integrals, which we tackle in order of complexity. We start with last, which follows directly from Itô's isometry

$$\int_t^{t+\Delta} \sigma B_r(u,T) d\widetilde{W_r^T}(u) \sim \sqrt{\int_t^{t+\Delta} \sigma^2 B_r^2(u,T) du} \cdot Z_r, \tag{6.57}$$

with Z_r an independent (of all random variables) standard normal distributed random variable.

The first integral in (6.56) follows similarly as:

$$\begin{aligned} &\int_t^{t+\Delta} \left(\sqrt{1-\rho_{Sv}^2}\, v(u) + \omega_{Sr} \sigma B_r(u,T) \right) \widetilde{W_S^T}(u) \\ &\sim \sqrt{\int_t^{t+\Delta} \left((1-\rho_{Sv}^2) v^2(u) + 2\sqrt{1-\rho_{Sv}^2}\, \omega_{Sr} \sigma B_r(u,T) v(u) + \omega_{Sr}^2 \sigma^2 B_r^2(u,T) \right) du} \cdot Z_S, \end{aligned} \tag{6.58}$$

with Z_S an independent standard normal distributed random variable. Finally the second integral can be obtained from the fact that the pair $\left(\int_t^{t+\Delta} \sigma B_r(u,T) d\widetilde{W_v^T}(u), \int_t^{t+\Delta} d\widetilde{W_v^T}(u) \right)$ follows a bivariate normal distribution with correlation $\rho_{vv2}(t, t+\Delta)$ and a conditioning argument

$$\int_t^{t+\Delta} \sigma B_r(u,T) d\widetilde{W_v^T}(u) \;\Big|\; \int_t^{t+\Delta} d\widetilde{W_v^T}(u) \sim \sqrt{G(t,t+\Delta)} \left(\rho_{vv2}(t,t+\Delta) Z_v + \sqrt{1-\rho_{vv2}^2(t,t+\Delta)}\, Z_{v2} \right), \tag{6.59}$$

$$\rho_{vv2}(t,t+\Delta) := \frac{\int_t^{t+\Delta} \sigma B_r(u,T) du}{\sqrt{\Delta \cdot G(t,t+\Delta)}}, \tag{6.60}$$

with Z_{v2} an independent standard normal random variable.

Having eliminated all stochastic integrals, we are left with deterministic integrals over $\sigma B_r(u,T)$, $v(u)$ and powers thereof; for the deterministic integrals over $\sigma B_r(u,T)$ we use the following

explicit solutions:

$$\int_t^{t+\Delta} \sigma B_r(u,T)du = \frac{\sigma}{a}\Big[\Delta - \frac{1}{a}e^{-a(T-t-\Delta)} + \frac{1}{a}e^{-a(T-t)}\Big] =: H(t,t+\Delta), \tag{6.61}$$

$$\begin{aligned}\int_t^{t+\Delta} \sigma^2 B_r^2(u,T)du &= \frac{\sigma^2}{a^2}\Big[\Delta + \frac{1}{2a}e^{-2a(T-t-\Delta)} - \frac{2}{a}e^{-a(T-t-\Delta)} - \frac{1}{2a}e^{-2a(T-t)} + \frac{2}{a}e^{-a(T-t)}\Big],\\ &=: G(t,t+\Delta),\end{aligned} \tag{6.62}$$

whereas we will approximate all integrals over $v(u)$ by using the predictor-corrector method:

$$\int_t^{t+\Delta} v^p(u)du, \approx \Big(\delta_1 v^p(t) + \delta_2 v^p(t+\Delta)\Big)\Delta, \tag{6.63}$$

for $p \in \{1,2\}$ and some constants δ_1, δ_2.

Collecting all terms once again yields an Exponentially Affine in Expectation (EAE) scheme for the SZHW model

$$y(t+\Delta,T) = y(t,T) + C_0 + C_1 v(t) + C_2 v(t+\Delta) + C_3 v^2(t) + C_4 v^2(t+\Delta) + C_5 Z_S + C_6 Z_v + C_7 Z_{v2} + C_8 Z_r \tag{6.64}$$

where:

$$C_0 = -\frac{1}{2}\Big[G(t,t+\Delta) + \rho_{Sv}\tau\Delta\Big],$$

$$C_1 = -\delta_1\Big(\rho_{Sr}H(t,t+\Delta) + \rho_{Sv}\Big[\frac{\psi\kappa\Delta}{\tau} - \rho_{rv}H(t,t+\Delta)\Big]\Big),$$

$$C_2 = -\delta_2\Big(\rho_{Sr}H(t,t+\Delta) + \rho_{Sv}\Big[\frac{\psi\kappa\Delta}{\tau} - \rho_{rv}H(t,t+\Delta)\Big]\Big),$$

$$C_3 = -\frac{1}{2}\delta_1\Delta + \frac{\rho_{Sv}}{\tau}\Big(\delta_1\kappa\Delta - \frac{1}{2}\Big), \qquad C_4 = -\frac{1}{2}\delta_2\Delta + \frac{\rho_{Sv}}{\tau}\Big(\delta_2\kappa\Delta + \frac{1}{2}\Big),$$

$$C_5 = \sqrt{C_{50} + C_{51}v(t) + C_{52}v(t+\Delta) + C_{53}v^2(t) + C_{54}v^2(t+\Delta)}.$$

$$C_{50} = \omega_{Sr}^2 G(t,t+\Delta), \qquad C_{51} = 2\delta_1\omega_{Sr}\sqrt{1-\rho_{Sv}}H(t,t+\Delta),$$

$$C_{52} = 2\delta_2\omega_{Sr}\sqrt{1-\rho_{Sv}^2}H(t,t+\Delta), \qquad C_{53} = \delta_1\Delta\Big(1-\rho_{Sv}^2\Big),$$

$$C_{54} = \delta_2\Delta\Big(1-\rho_{Sv}^2\Big), \qquad C_6 = \rho_{rv}\sqrt{G(t,t+\Delta)}\rho_{vv2}(t,t+\Delta),$$

$$C_7 = \rho_{rv}\sqrt{G(t,t+\Delta)}\sqrt{1-\rho_{vv2}^2(t,t+\Delta)}, \qquad C_8 = \sqrt{1-\rho_{rv}^2-\omega_{Sr}^2}\sqrt{G(t,t+\Delta)}.$$

Similar to the SZ scheme (6.22), the above simulation scheme might have a net drift away from

the martingale and violate the (no-arbitrage) martingale property. Nonetheless, in the following section, we show that one can easily enforce this martingale condition by replacing the constant C_0 with C_0^* of equation (6.72).

6.5.3 Martingale correction, regularity

In this section, using similar techniques as in Section 6.4.1, we will investigate how to enforce the martingale property of the discretized asset price process $\widetilde{x}$ in predictor-corrector scheme (6.64); furthermore, we investigate the regularity of the proposed discretization scheme, i.e. we look whether there are parameter values where the $\widetilde{x}$-process blows up.

By the tower law of conditional expectations, we have the following

$$\mathbb{E}^{Q^T}[\widetilde{F}(t+\Delta)|\mathcal{F}_t] = \mathbb{E}^{Q^T}\left\{\mathbb{E}^{Q^T}\left[\widetilde{F}(t+\Delta)\middle|\mathcal{F}_t \vee \sigma\left(\nu(t+\Delta T)\right)\right]\middle|\mathcal{F}_t\right\}, \tag{6.65}$$

with $\sigma\left(\nu(t+\Delta T)\right)$ the sigma-algebra generated by $\nu(t+\Delta T)$ and where $\mathcal{F}_t, \sigma\left(\nu(t+\Delta T)\right)$ denotes the smallest sigma-algebra containing both $\mathcal{F}_s$ and $\sigma\left(\nu(t+\Delta T)\right)$. For the martingale condition (6.23) to hold, we hence need $\widetilde{F}(t)$ to equal the latter expectation under the T-forward measure Q^T. We express the inner expectation completely in terms of $\nu(t)$ and $\nu(t+\Delta)$ by exponentiating (6.64), taking the expectation over the independent normal distributions Z_S, Z_{v2} and Z_r, and noting from (6.49) that $Z_\nu := \frac{\nu(t+\Delta)-\nu(t)K_1-K_2}{K_3}$. We obtain the following expression

$$\mathbb{E}^{Q^T}\left[\widetilde{F}(t+\Delta)\middle|\mathcal{F}_t, \nu(T)\right] = \widetilde{F}(t)\exp\left[D_0 + D_1\nu(t) + D_2\nu(t+\Delta) + D_3\nu^2(t) + D_4\nu^2(t+\Delta)\right] \tag{6.66}$$

where

$$\begin{aligned} D_0 &:= C_0^* + \frac{1}{2}C_8^2 + \frac{1}{2}C_7^2 + \frac{1}{2}C_{50} - C_6\frac{K_2}{K_3}, \quad D_1 := C_1 + \frac{1}{2}C_{51} - \frac{K_1}{K_3}C_6, \\ D_2 &:= C_2 + \frac{1}{2}C_{52} + \frac{1}{K_3}C_6, \qquad D_3 := C_3 + \frac{1}{2}C_{53}, \qquad D_4 := C_4 + \frac{1}{2}C_{54}. \end{aligned} \tag{6.67}$$

Once again, due to the EAE property of our scheme, this term is exponentially affine. By substituting (6.66) in (6.65), we find that the following condition has to be satisfied for the martingale condition to hold:

$$1 = \mathbb{E}^{Q^T}\left\{\exp\left[D_0 + D_1\nu(t) + D_2\nu(t+\Delta) + D_3\nu^2(t) + D_4\nu^2(t+\Delta)\right]\middle|\mathcal{F}_t\right\}. \tag{6.68}$$

Taking the $\mathcal{F}_t$ measurable terms out of the expectation and collecting terms, we obtain

$$\begin{aligned} 1 &= \exp\left[D_0 + D_1\nu(t) + D_3\nu^2(t)\right]\mathbb{E}^{Q^T}\left\{\exp\left[D_2\nu(t+\Delta) + D_4\nu^2(t+\Delta)\right]\middle|\mathcal{F}_t\right\} \\ &= \exp\left[D_0 + D_1\nu(t) + D_3\nu^2(t)\right]\Psi_H(1) \end{aligned} \tag{6.69}$$

where $\Psi_H(t)$ denotes the moment-generating function of the (discretized) process

$$H := D_2 v(t+\Delta) + D_4 v^2(t+\Delta), \tag{6.70}$$

evaluated in the point t. Hence expanding D_0, we have that for the martingale condition to hold we need

$$1 = \exp\Big[C_0^* + \frac{1}{2}C_8^2 + \frac{1}{2}C_7^2 + \frac{1}{2}C_{50} - C_6\frac{K_2}{K_3} + D_1 v(t) + D_3 v^2(t)\Big]\Psi_H(1) \tag{6.71}$$

which (assuming the regularity condition $\Psi_H(1) < \infty$ is satisfied) can be established by setting

$$C_0^* := -\frac{1}{2}C_8^2 - \frac{1}{2}C_7^2 - \frac{1}{2}C_{50} + C_6\frac{K_2}{K_3} - D_1 v(t) - D_3 v^2(t) - \ln\Big(\Psi_H(1)\Big). \tag{6.72}$$

As $v(t+\Delta)$ is still Gaussian under Q^T, $\Psi_H(1)$ and its regularity can be determined in a similar fashion to Section 6.4.1. The following proposition applies the above result to the martingale correction and the regularity of the simulation scheme (6.64).

Proposition 6.5.1 *The regularity of the simulation scheme (6.64) holds if and only if the regularity condition (6.34) is satisfied. Given that this condition is satisfied, we can ensure the martingale property in the SZHW-scheme of (6.64) by replacing the constant C_0 by*

$$C_0^* \;=\; E_0 + E_1 v(t) + E_2 v^2(t), \tag{6.73}$$

where:

$$E_0 \;:=\; \frac{1}{2}\ln\Big(1 - 2D_4K_3^2\Big) - \frac{D_4\Big(K_2 + \frac{D_2}{2D_4}\Big)^2}{1 - 2D_4K_3^2} - \frac{1}{2}C_{50} + C_6\frac{K_2}{K_3} - \frac{1}{2}C_7^8 - \frac{1}{2}C_8^2 + \frac{D_2^2}{4D_4}, \tag{6.74}$$

$$E_1 \;:=\; -D_1 - \frac{2D_4K_1\Big(K_2 + \frac{D_2}{2D_4}\Big)}{1 - 2D_4K_3^2}, \quad E_2 := -D_3 - \frac{D_4K_1^2}{1 - 2D_4K_3^2}, \tag{6.75}$$

with K_1, K_2, K_3 as defined in (6.49), (6.50) and $D_1, \ldots, D_4$ in (6.67).

Proof Follows directly from the above results.

6.6 Numerical results

Any simulation scheme has to be tested, as they say the proof of the pudding is in the eating. In this section our goal is to test our proposed simulation schemes and compare them to alternate schemes. In our comparisons we focus on the bias of European call prices, where by bias we mean $\mathbb{E}[\widehat{\alpha}] - \alpha$, where α is the true price of the European call and $\widehat{\alpha}$ is its Monte Carlo estimator. It is of high importance for practitioners to have a bias as small as possible for reasonable sizes of the time step. Ideally one would like to be able to simulate the relevant quantities only at those points in time which are relevant to the option contract that is being priced. Unfortunately,

that is not always possible as has certainly become clear from several papers on the simulation of the Heston model, e.g. see Lord et al. (2008) or van Haastrecht and Pelsser (2010).

Table **1** contains the parameter configurations for our tests cases. Section 6.6.1 deals with the simulation scheme for the special case of the SZ model that collapses to a Heston model, see Section 6.3. Volatilities for Case I are similar to those in the equity market at the time of writing. In this test case we not only compare our scheme to an Euler scheme, but also to the best-performing scheme considered in Andersen (2008), the QE-M scheme. Finally, we also compare to a recently proposed scheme for the Heston model in Zhu (2008). In Case II we consider a setting of the SZ model which does not collapse to the Heston model. Here, the volatilities correspond to levels seen at the end of 2008 and beginning of 2009. Finally, we also test the scheme which was proposed in Section 6.5 for the SZHW model: case III deals with normal, perhaps slightly excited, long-term market volatilities.

Example	**Type**	κ	τ	$v(0)$	θ	ρ_{Sv}	r	a	σ	ρ_{Sr}	ρ_{rv}
case I	Call-5Y	0.1	0.3	0.0	0.0	-0.6	0.00	-	-	-	-
case II	Call-10Y	0.4	0.4	0.2	0.2	-0.9	0.04	-	-	-	-
case III	Call-15Y	0.4	0.4	0.2	0.2	-0.7	0.04	0.03	0.01	0.2	0.15

Table 1: Test cases for the Schöbel-Zhu/Heston, Schöbel-Zhu, and Schöbel-Zhu-Hull-White simulation scheme, in all cases $S(0) = 100$.

All numerical examples are based on a million simulation paths, where we used the stock price as a control variate and the Mersenne Twister to generate pseudo-random uniform numbers.

6.6.1 Results for the Heston/Schöbel-Zhu model

Starting with Case I, in which we consider a special case of the SZ model which corresponds to the Heston model, we can not only compare our Exponential Affine scheme (denoted by EAE) to a simple Euler scheme, but we can of course also compare it to the best-performing scheme of Andersen (2008), the QE-M scheme. Whereas our EAE scheme samples from the exact distributions of $v(t)$ and $v^2(t)$, Andersen's QE scheme uses:

- a Quadratic Gaussian distribution when $\mathrm{Var}[v^2(t+\Delta)] \,/\, \mathbb{E}[v^2(t+\Delta)] \leq 1.5$,
- a mixture of zero and an exponential function, otherwise.

Our scheme and the QE-M scheme therefore differ for low values of $v(t)$, to be precise when:

$$v^2(t) \in \Big[0, \frac{(e^{\kappa\Delta}-1)\tau^2}{4\kappa}\Big]. \tag{6.76}$$

Finally, we also compare to Zhu (2008) scheme. In this scheme, the SDE for the square root of the stochastic variance is derived. As the square root is not differentiable in zero, Itô's lemma is

incorrectly applied here, e.g. see Kahl and Jäckel (2006) or Lord et al. (2008). Luckily, Zhu's best-performing method, a moment-matching method, does not depend too much on this premise. The numerical examples in his paper suggest that this method is comparable to Andersen's QE scheme for low values of the volatility of variance parameter, but is outperformed for realistic levels of the volatility of variance parameter.

	Euler	Zhu	QE-M	EAE
Δ		$K = 100$		
1	-1.914 (±0.062)	0.479 (±0.075)	-0.059 (±0.057)	-0.130 (±0.057)
1/2	-0.781 (±0.060)	0.291 (±0.064)	-0.023* (±0.057)	-0.020* (±0.057)
1/4	-0.348 (±0.058)	0.106 (±0.060)	0.013* (±0.057)	-0.039* (±0.057)
1/8	-0.136 (±0.058)	0.049* (±0.059)	0.024* (±0.057)	-0.003* (±0.057)
1/16	0.002* (±0.058)	0.041* (±0.058)	-0.008* (±0.057)	0.013* (±0.057)
1/32	-0.015* (±0.057)	0.037* (±0.057)	-0.015* (±0.057)	-0.010* (±0.057)
		$K = 140$		
1	0.058* (±0.074)	3.537 (±0.099)	-0.524 (±0.066)	-0.354 (±0.066)
1/2	0.226 (±0.071)	2.068 (±0.078)	-0.147 (±0.066)	-0.080 (±0.066)
1/4	0.160 (±0.068)	1.049 (±0.071)	-0.006* (±0.067)	-0.051* (±0.066)
1/8	0.137 (±0.067)	0.503 (±0.068)	0.020* (±0.066)	-0.012* (±0.066)
1/16	0.138 (±0.067)	0.251 (±0.067)	-0.016* (±0.066)	0.003* (±0.066)
1/32	0.053* (±0.066)	0.138 (±0.066)	-0.038* (±0.066)	-0.015* (±0.066)
		$K = 60$		
1	-1.707 (±0.036)	-0.991 (±0.041)	0.066 (±0.036)	-0.005* (±0.036)
1/2	-0.793 (±0.037)	-0.541 (±0.038)	0.022* (±0.036)	0.004* (±0.036)
1/4	-0.381 (±0.037)	-0.254 (±0.037)	0.004* (±0.037)	-0.017* (±0.036)
1/8	-0.170 (±0.037)	-0.114 (±0.037)	0.013* (±0.037)	0.003* (±0.037)
1/16	-0.044 (±0.037)	-0.034* (±0.037)	0.005* (±0.037)	0.004* (±0.036)
1/32	-0.024* (±0.037)	-0.004* (±0.037)	-0.003* (±0.037)	-0.001* (±0.036)

Table 2: Estimated call option prices biases for case I. Numbers in parentheses are the widths of the confidence interval at a 99% confidence level: we starred the biases that were not significantly different from zero. Exact prices respectively are 27.90, 14.23 and 50.34.

In Table **2**, we have displayed estimated call option price biases for Case I, as a function of the strike level ($K = 60, 100$ or 140) and the time step Δ (1 through 1/32). Numbers in parentheses are the widths of the confidence interval at a 99% confidence level: we starred the biases that were not significantly different from zero. From the results it is clear that - at least for this parameter configuration - Zhu's method is better than a simple Euler discretization for lower strikes, though for higher strikes the Euler scheme wins. The QE-M and EAE methods however are much better in terms of bias. Both methods are too close to tell apart.

As one eventually wants to judge a scheme based on its efficiency, one should look its accuracy in combination with the computational effort of the methods. To this end we also report the computational times for the four simulation methods which are provided in Table **3** below.

Δ	Euler	Zhu	QE-M	EAE
1	0.9	1.0	1.2	1.1
1/2	1.8	2.0	2.4	2.2
1/4	3.6	4.0	5.0	4.3
1/8	7.0	7.9	10.2	8.6
1/16	13.9	15.7	20.8	17.1
1/32	28.1	31.3	41.9	33.5

Table 3: Computational times in seconds for case I for the Euler, Zhu, QE-M and EAE scheme, all with one million sample paths.

From Table **3** we can see the Euler scheme takes the least time to compute, followed by Zhu's method, the EAE scheme and the QE-M scheme. Still, the efficiency of the QE-M scheme and the EAE method by far outperforms those of the Euler and Zhu's method as can be seen if we take a look at the accuracy of the methods in Table **2**. From that table, we can see that the EAE and QE-M only need 2 or 4 time steps a year to produce of scheme with no significant bias, whereas the Euler and Zhu's scheme in most cases need at least 16 time steps a year to produce a scheme negligible bias. Though the QE-M scheme and the EAE method produce a similar accuracy, the EAE method is more efficient. This can be explained to the fact that the exact Gaussian volatility distribution of the Schöbel and Zhu (1999) model is explicitly utilized in the EAE method, whereas the variance simulation of QE-M method is tailored for Heston (1993) model.

6.6.2 Results for the Schöbel-Zhu and Schöbel-Zhu-Hull-White model

We move on to Case II and III, which are slightly more benign due to a non-zero value of ψ[1] in Case II and the inclusion of stochastic interest rates for Case III. The numerical results for these cases can be found in Table **4** and **5**. Computional times for both cases behave very similar to those reported in Table **3** and are hence omitted.

In Case II, a non-Heston SZ model, the differences between the Euler and EAE methods are indeed closer, though still noticeably in favour of the EAE method. From Table **5** we can see that especially for in-and out-of-the-money options, the EAE scheme significantly outperforms the Euler scheme.

Finally, we take a look at the performance of the simulation schemes for the SZHW model, where in addition to the SZ model we have stochastic interest rates which are correlated with both the underlying and the stochastic volatility process. While the addition of stochastic interest rates complicates the scheme slightly, the picture is similar to before as can be seen from Table 6. Again the EAE method produces a much smaller discretization error than the Euler scheme, allowing the user to utilise bigger time steps instead of the smaller ones one would be confined to when using the Euler method. For example, for the strikes considered one could safely use a

[1]This makes the distribution less fat-tailed.

timestep of a quarter of a year for the EAE method, and have a bias which is not significantly different from zero. In the Euler method, this is only achieved with a timestep equal to 1/16.

	Euler	EAE
Δ	$K = 100$	
1	-0.828 (±0.059)	-0.389 (±0.050)
1/2	-0.314 (±0.055)	-0.165 (±0.050)
1/4	-0.145 (±0.052)	-0.034* (±0.050)
1/8	-0.068 (±0.051)	0.014* (±0.050)
1/16	-0.043* (±0.051)	0.005* (±0.050)
1/32	0.017* (±0.050)	0.003* (±0.050)
	$K = 140$	
1	-0.495 (±0.080)	-0.457 (±0.063)
1/2	-0.110 (±0.072)	-0.204 (±0.064)
1/4	-0.032* (±0.068)	-0.041* (±0.064)
1/8	-0.003* (±0.066)	0.013* (±0.064)
1/16	0.002* (±0.065)	0.005* (±0.064)
1/32	0.034* (±0.065)	0.005* (±0.064)
	$K = 60$	
1	-0.699 (±0.035)	-0.278 (±0.031)
1/2	-0.298 (±0.033)	-0.110 (±0.032)
1/4	-0.151 (±0.033)	-0.023* (±0.032)
1/8	-0.070 (±0.032)	0.007* (±0.032)
1/16	-0.046 (±0.032)	0.004* (±0.032)
1/32	0.005* (±0.032)	0.004* (±0.032)

Table 4: Results for case II, exact prices: 56.77, 45.34 and 70.89.

	Euler	EAE
Δ	$K = 100$	
1	-0.323 (±0.032)	-0.069 (±0.032)
1/2	-0.151 (±0.032)	-0.036 (±0.032)
1/4	-0.087 (±0.032)	0.008* (±0.032)
1/8	-0.029* (±0.032)	0.005* (±0.032)
1/16	-0.006* (±0.032)	-0.022* (±0.032)
1/32	0.001* (±0.032)	0.001* (±0.032)
	$K = 140$	
1	-0.299 (±0.047)	-0.085 (±0.046)
1/2	-0.127 (±0.046)	-0.046* (±0.046)
1/4	-0.077 (±0.046)	0.012* (±0.046)
1/8	-0.014* (±0.046)	0.006* (±0.046)
1/16	-0.002* (±0.046)	-0.022* (±0.046)
1/32	0.014* (±0.046)	-0.008* (±0.046)
	$K = 60$	
1	-0.190 (±0.016)	-0.041 (±0.017)
1/2	-0.095 (±0.017)	-0.017 (±0.017)
1/4	-0.052 (±0.017)	0.005* (±0.017)
1/8	-0.019 (±0.017)	0.005* (±0.017)
1/16	0.005* (±0.017)	-0.007* (±0.017)
1/32	-0.001* (±0.017)	0.004* (±0.017)

Table 5: Results for case III, exact prices: 53.75, 40.69 and 69.97.

Estimated call option prices biases for case II and III. Numbers in parentheses are the widths of the confidence interval at a 99% confidence level: we starred the biases that were not significantly different from zero.

Finally, we take a look at the performance of the simulation schemes for the SZHW model, where in addition to the SZ model we have stochastic interest rates which are correlated with both the underlying and the stochastic volatility process. While the addition of stochastic interest rates complicates the scheme slightly, the picture is similar to before as can be seen from Table **5**. Again the EAE method produces a much smaller discretization error than the Euler scheme, allowing the user to utilize bigger time steps instead of the smaller ones one would be confined to when using the Euler method.

6.7 Conclusion

A major problem signaled with Euler schemes in the simulation of stochastic volatility models is their inability to generate the proper correlation between the increments of the asset and the stochastic volatility processes. As the correlation parameter in the stochastic volatility models is an important determinant of the skew in implied volatilities, not being able to match this parameter leads to a significant mispricing of options with strikes far away from the at-the-money level. In the Heston (1993) model, this so-called "leaking correlation" problem, is partially caused by the fact that an Euler discretization tries to approximate a square root process, by a Gaussian process. However even when the stochastic volatility itself is Gaussian, such as in Schöbel and Zhu (1999)-like models, we have shown that the problem of "leaking correlation" is still an issue. In this chapter we have proposed simulation algorithms for the SZ model and its extensions. By analyzing the lessons learned on how to avoid the so-called leaking correlation phenomenon in the simulation of the Heston (1993) model, we formulated a simulation scheme for the SZ model which is tailored to match the correlation between the increments of the asset price and the variance processes of the continuous-time dynamics. A simulation scheme for the Schöbel-Zhu-Hull-White model considered in Chapter 3, which incorporates the need for stochastic interest rates, was derived as well. This is closely related to the recent advances in the development of markets for long-term derivatives, for which maturities the inclusion of stochastic interest rates in a derivatives pricing model is more appropriate.
All introduced schemes have carefully been chosen to be Exponentially Affine in Expectation (EAE), which greatly facilitates the derivation of a martingale correction. The regularity of both schemes has also been studied. Finally, we numerically compared the new simulation schemes to other recent schemes in the literature. For a special case of the SZ model which coincides with the Heston model, our proposed scheme has a similar performance to the QE-M scheme of Andersen (2008), whilst being slightly more efficient in terms of computational time required. For all non-Heston SZ and the SZHW model, it has been demonstrated that our scheme consistently outperforms the Euler scheme. These results affirm that Andersen's result is more widely applicable than to the Heston model alone; for the simulation of stochastic volatility models, it is of great importance to match the correlation between the asset price and its stochastic volatility process.

Part III

Applications to Insurance Markets

Chapter 7

Accounting for Stochastic Interest Rates, Stochastic Volatility and a General Correlation Structure in the Valuation of Forward Starting Options

*This chapter is based on:

Alexander van Haastrecht and Antoon Pelsser (2010). "Accounting for Stochastic Interest Rates, Stochastic Volatility and a General Correlation Structure in the Valuation of Forward Starting Options", to appear in: *Journal of Futures Markets*.

7.1 Introduction

Forward starting options belong to the class of path-dependent European-style contracts in the sense that they not only depend on the terminal value of the underlying asset, but also on the asset price at an intermediate point (often dubbed as 'strike determination date'). Typically, a forward starting contract gives the holder a call (or put) option with a strike that is set equal to a fixed proportion of the underlying asset price at this intermediate date. A special form of these options are those on the (future) return of the underlying, which can be seen as a call option on the ratio of the stock price at maturity and the intermediate date. The latter form is often being used by insurance companies to hedge unit-linked guarantees embedded in life insurance products. Additionally, structured products involving forward starting options (like cliquet and ratchet structures) are often tailored for investors seeking for upside potential, while keeping protection against downside movements.
Though forward starting options seem quite simple exotic derivatives, their valuation can be demanding, depending on the underlying model. Our pricing takes into account two important factors in the pricing of forward starting options: stochastic volatility and stochastic interest rates, whilst also taking into account the correlation between those processes explicitly. It is

hardly necessary to motivate the inclusion of stochastic volatility in a derivative pricing model. Stochastic interest rates are crucial for the pricing of forward starting options because securities with forward starting features often have a long-dated maturities and are therefore much more interest rate sensitive, e.g. see Guo and Hung (2008) or Kijima and Muromachi (2001). The addition of interest rates as a stochastic factor has been the subject of empirical investigations most notably by Bakshi et al. (2000). These authors show that the hedging performance of delta hedging strategies of long-maturity options improves when taking stochastic interest rates into account.

The pricing of forward starting options was first considered by Rubinstein (1991) who provides a closed-form solution for the pricing of forward starting options based on the assumptions of the Black and Scholes (1973) model. Lucic (2003), Hong (2004) and Kruse and Nögel (2005) relax the constant volatility assumption and consider the pricing of forward starting options under Heston (1993) stochastic volatility. The pricing of forward starting options under stochastic volatility with independent stochastic interest rates was considered by Guo and Hung (2008), Ahlip and Rutkowski (2009) and Nunes and Alcaria (2009). The framework employed in this chapter distinguishes itself from these models by a closed form pricing formula and an explicit, rather than implicit, incorporation of the correlation between underlying and the term structure of interest rates. The flexibility of stochastic volatility model with (correlated) stochastic rates and the pricing of vanilla call options in such a framework was covered in Ahlip (2008) and Chapter 3.

The main goal of this work is performing a quantitative analysis on the pricing of forward starting options under stochastic volatility and stochastic interest rates. In particular we want to investigate the impact of stochastic volatility, stochastic interest rates as well as a realistic dependency structure between all the underlying processes on the valuation of these securities. The analysis is made possible by developing a closed-form solution for the price of a forward starting option in a model in which the instantaneous stochastic volatility is given by the Schöbel and Zhu (1999) model and the interest rates follow Hull and White (1993) dynamics. We explicitly incorporate the correlation between underlying stock and the term structure of interest rates, which is an important empirical characteristic that needs to be taken into account for the pricing and hedging of long-term options, e.g. see Bakshi et al. (2000) or Piterbarg (2005). The setup of this chapter is as follows: we discuss the modelling framework and the corresponding forward starting option problem in Section 7.2 and 7.3. Using the characteristic function of the log-asset price under the stock price measure (derived in Section 7.4), we derive in Section 7.5 the main pricing formulas of the chapter. In Section 7.6 we consider the implementation of these formulas and analyze the valuation and risk management of forward starting option under stochastic volatility, stochastic interest rates and a general correlation structure. Finally, we conclude in Section 7.7.

7.2 The modelling framework

To analyze the pricing of forward starting options under stochastic volatility and stochastic interest rates, we use the SZHW model of see Chapter 3. For clarity we will here repeat these

dynamics: in the SZHW model, the stock price $S(t)$ is governed by the following dynamics

$$dS(t) = r(t)S(t)dt + \nu(t)S(t)dW_S^Q(t), \qquad S(0) = S_0, \tag{7.1}$$

$$\nu(t) = \kappa\Big(\psi - \nu(t)\Big)dt + \tau dW_\nu^Q(t), \qquad \nu(0) = \nu_0, \tag{7.2}$$

where r follows an Hull and White (1993) process, see Chapter 2. For an explanation of the model parameters, see Chapter 3. The simulation of the SZHW model is discussed in Chapter 6.

7.3 Forward starting options

Forward starting options are contracts which not only depend on their terminal value of the underlying asset, but also on the asset price at an intermediate time between the current time and its expiry time. Kruse and Nögel (2005) consider two types of forward starting options under the Heston (1993) model: European forward starting call options on the underlying asset and on the underlying return. The first structure is prevalent in Employee stock option schemes, while the second category forms a building block for cliquet, ratchet and Unit-Linked insurance options. In both contracts a premium is paid on the purchase date, however the option's life will only start on an intermediate date (in between the purchase and expiry date, dubbed as the strike determination time). Thus, the terminal payoff of these options depends on the underlying asset price at both the maturity and the start date of the underlying option. The next definition formalizes these option types.

Definition The terminal payoff of a European forward starting call option on the underlying asset price S, with a percentage strike of K, strike determination time T_{i-1} and maturity T_i is given by

$$\Big[S(T_i) - KS(T_{i-1})\Big]^+. \tag{7.3}$$

The terminal payoff of a European forward starting call option on the return of the underlying asset price S, with an absolute strike of K, determination time T_{i-1} and maturity T_i is given by

$$\left[\frac{S(T_i)}{S(T_{i-1})} - K\right]^+. \tag{7.4}$$

7.3.1 The option pricing framework

We can express the price of the forward starting call option price $C_F(T_{i-1}, T_i)$ on the underlying asset, $t \leq T_{i-1} \leq T_i$ and with terminal payoff (7.3), in the following expectation under the risk-neutral measure Q

$$C_F(T_{i-1}, T_i) = \mathbb{E}^Q\left[e^{-\int_t^{T_i} r(u)du}\Big(S(T_i) - KS(T_{i-1})\Big)^+ \Big| \mathcal{F}_t\right]. \tag{7.5}$$

Instead of evaluating the expected discounted payoff under the risk-neutral bank account measure, we can also change the underlying probability measure to evaluate this expectation under

the stock price probability measure Q^S (e.g. see Geman et al. (1996)), i.e. with the stock price S as numeraire. Hence, starting from time t, we can evaluate the price of the forward starting option (7.5) as

$$\begin{aligned} C_F(T_{i-1}, T_i) &= S(t)\mathbb{E}^{Q^S}\left[\frac{1}{S(T_i)}\Big(S(T_i) - KS(T_{i-1})\Big)^+\Big|\mathcal{F}_t\right] \\ &= S(t)\mathbb{E}^{Q^S}\left[\Big(1 - K\frac{S(T_{i-1})}{S(T_i)}\Big)^+\Big|\mathcal{F}_t\right] \\ &= S(t)K\mathbb{E}^{Q^S}\left[\Big(\frac{1}{K} - \frac{S(T_{i-1})}{S(T_i)}\Big)^+\Big|\mathcal{F}_t\right], \end{aligned} \tag{7.6}$$

where the last line can be interpreted as put option with strike $\frac{1}{K}$ on the ratio $\frac{S(T_{i-1})}{S(T_i)}$.

In principle it also possible, following the lines of Rubinstein (1991), Guo and Hung (2008) and Ahlip and Rutkowski (2009), to express the forward starting option price as the expected value of a future call option price, i.e.

$$C_F(T_{i-1}, T_i) = S(t)\mathbb{E}^{Q^S}\left\{\frac{1}{S(T_{i-1})}\mathbb{E}^{Q^S}\left[\Big(S(T_i) - KS(T_{i-1})\Big)^+\Big|\mathcal{F}_{T_{i-1}}\right]\Big|\mathcal{F}_t\right\}. \tag{7.7}$$

The above expectation can be evaluated using similar techniques as the evaluation of formula (7.6), and results in a pricing formula containing two integrals. On the other hand, working out the equivalent expectation (7.6) results in a pricing formula which only contains one integral. Not only does this make the corresponding implementation more efficient, but even more importantly it has been shown in Andersen and Andreasen (2002) and Lord and Kahl (2008) that the double integral formulation suffers from numerical instabilities whereas the single integral can be implemented in a numerically very stable way. Hence though both approaches are mathematically equivalent, we prefer to work with expectation (7.6) over the expression in formula (7.7).

We therefore express the option (7.6) with log strike $k := \ln\frac{1}{K}$, in terms of the (T-forward) characteristic function $\phi_F(T_{i-1}, T_i, v)$ of the log ratio $\ln\frac{S(T_{i-1})}{S(T_i)}$, i.e.

$$C_F(T_{i-1}, T_i, k) = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(e^{(\alpha - iv)k}\psi_F(T_{i-1}, T_i, v)\Big)dv, \tag{7.8}$$

with

$$\psi_F(T_{i-1}, T_i, v) := \frac{\phi_F\Big(T_{i-1}, T_i, v + (\alpha - 1)i\Big)}{(iv - \alpha)(iv - \alpha + 1)},$$

with $\phi_F(T_{i-1}, T_i, v) := \mathbb{E}^{Q^S}\left[\exp\Big(iu\ln\frac{S(T_{i-1})}{S(T_i)}\Big)\Big|\mathcal{F}_t\right]$ and where $\alpha > 1$ has been introduced for Fourier

Transform regularization, see Section 2.5.

Remark 7.3.1 *For the pricing of the forward starting option on the underlying asset, it suffices to know the characteristic function $\phi_F(T_{i-1}, T_i, v)$ of $\ln \frac{S(T_{i-1})}{S(T_i)}$ under the stock price probability measure Q^S. For the derivation of this characteristic function, see Section 7.5.1.*

For the price $C_R(T_{i-1}, T_i)$ of the forward starting call option on the return of the underlying asset, i.e. with terminal payoff (7.3), the following expectation expectation under the T_i-forward measure holds

$$C_R(T_{i-1}, T_i) = P(t, T_i)\mathbb{E}^{Q^{T_i}}\Big[\Big(\frac{S(T_i)}{S(T_{i-1})} - K\Big)^+ \Big|\mathcal{F}_t\Big], \tag{7.9}$$

i.e. the corresponding numeraire is now the (pure) discount bond $P(t, T_i)$ maturing at time T_i. One can again write the option (7.6) with log strike $k := \ln K$, in terms of the (T-forward) characteristic function $\phi(T_{i-1}, T_i, v)$ of the log ratio $\ln \frac{S(T_i)}{S(T_{i-1})}$, i.e.

$$C_R(T_{i-1}, T_i, k) = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(e^{-(\alpha+iv)k}\psi(T_{i-1}, T_i, v)\Big)dv, \tag{7.10}$$

with

$$\psi_R(T_{i-1}, T_i, v) \quad := \quad \frac{\phi_R\Big(T_{i-1}, T_i, v - (\alpha+1)i\Big)}{(\alpha + iv)(\alpha + 1 + iv)},$$

with $\phi_R(T_{i-1}, T_i, v) := \mathbb{E}^{Q^{T_i}}\Big[\exp\Big(iu \ln \frac{S(T_i)}{S(T_{i-1})}\Big)\Big|\mathcal{F}_t\Big]$ and where $\alpha \in \mathbb{R}^+$ has been introduced for Fourier Transform regularization.

Remark 7.3.2 *For the pricing of the forward starting option on the return of the underlying asset, it suffices to know the characteristic function $\phi_R(T_{i-1}, T_i, v)$ of $\ln \frac{S(T_i)}{S(T_{i-1})}$ under the T_i-forward probability measure Q^{T_i}. For the derivation of the characteristic function, see Section 7.5.2.*

The remainder of the chapter hence focusses on the derivation of the above characteristic functions.

7.4 Characteristic function of the log asset price

As a preliminary step towards the general valuation results presented in Section 7.5, we derive in this section the characteristic function of the log asset price $F(t, T)$ under the stock price measure Q^S and under the T-forward measure Q^T. To this end, define the T-forward asset price at time t as

$$F(t, T) = \frac{S(t)}{P(t, T)}, \tag{7.11}$$

where $P(t,T)$ denotes the the price of a (pure) discount bond at time t maturing at time T, hence note that $F(T,T) = S(T)$. Under the risk-neutral measure Q (where we use the money market bank account as numeraire) the discount bond price follows the process $dP(t,T) = r(t)P(t,T)dt - \sigma B_{\text{HW}}(t,T)P(t,T)dW_r(t)$, where $B_{\text{HW}}(t,T) := \frac{1}{a}\Big(1 - e^{-a(T-t)}\Big)$. Hence, by an application of Itô's lemma, one has the following result for the T-forward stock price process:

$$\begin{aligned} dF(t,T) &= \Big(\rho_{Sr}\nu(t)\sigma B_{\text{HW}}(t,T) + \sigma^2 B^2_{\text{HW}}(t,T)\Big)F(t,T)dt \\ &\quad + \nu(t)F(t,T)dW_S^Q(t) + \sigma B_{\text{HW}}(t,T)F(t,T)dW_r^Q(t). \end{aligned} \tag{7.12}$$

We will use these dynamics in the following two sections to determine the characteristic function of $\ln F(T)$ under respecitvely the stock price measure and the T-forward measure.

7.4.1 Characteristic function under the stock price measure Q^S

To determine the dynamics of the forward asset price under the stock price measure, we need to change from the money market account numeraire to the stock price numeraire, see Chapter 2. The corresponding Radon-Nikodým derivative is given by

$$\frac{dQ^S}{dQ} = \frac{S(T)B(0)}{S(0)B(T)} = \exp\Big[-\frac{1}{2}\int_0^T \nu^2(u)du + \int_0^T \nu(u)dW_S^Q(u)\Big]. \tag{7.13}$$

The multi-dimensional version of Girsanov's theorem (e.g. see Brigo and Mercurio (2006)) implies that in our model

$$dW_S^{Q^S}(t) \;\mapsto\; dW_S^Q(t) - \nu(t)dt, \tag{7.14}$$

$$dW_r^{Q^S}(t) \;\mapsto\; dW_r^Q(t) - \rho_{Sr}\nu(t)dt, \tag{7.15}$$

$$dW_\nu^{Q^S}(t) \;\mapsto\; dW_\nu^Q(t) - \rho_{S\nu}\nu(t)dt, \tag{7.16}$$

are Q^S Brownian motions. Hence under Q^S we have the following model dynamics for the volatility and interest rate process

$$\begin{aligned} dF(t,T) &= \Big(\nu^2(t) + 2\rho_{Sr}\nu(t)\sigma B_{\text{HW}}(t,T) + \sigma^2 B^2_{\text{HW}}(t,T)\Big)F(t,T)dt \\ &\quad + \nu(t)F(t,T)dW_S^{Q^S}(t) + \sigma B_{\text{HW}}(t,T)F(t,T)dW_r^{Q^S}(t) \end{aligned} \tag{7.17}$$

$$dx(t) = \Big(-ax(t) + \rho_{Sr}\sigma\nu(t)\Big)dt + \sigma dW_r^{Q^S}(t), \tag{7.18}$$

$$d\nu(t) = \Big(\kappa(\psi - \nu(t)) + \rho_{S\nu}\tau\nu(t)\Big)dt + \tau dW_\nu^{Q^S}(t). \tag{7.19}$$

We can simplify (7.17) by switching to logarithmic coordinates and rotating $W_S^{Q^S}(t)$ and $W_r^{Q^S}(t)$ to a Brownian motion $W_F^{Q^S}(t)$. Defining $y(t,T) := \ln(F(t,T))$ and an application of Itô's lemma

yields the following dynamics:

$$dy(t) = +\frac{1}{2}v_F^2(t)dt + v_F(t)dW_F^{Q^S}(t), \tag{7.20}$$

$$dv(t) = \widetilde{\kappa}\Big(\widetilde{\psi} - v(t)\Big)dt + \tau dW_v^{Q^S}(t) \tag{7.21}$$

where $\widetilde{\kappa} := \kappa - \rho_{Sv}\tau$, $\widetilde{\psi} := \frac{\kappa\psi}{\widetilde{\kappa}}$ and with

$$v_F^2(t) := v^2(t) + 2\rho_{Sr}v(t)\sigma B_{\text{HW}}(t,T) + \sigma^2 B_{\text{HW}}^2(t,T). \tag{7.22}$$

Note that we now have reduced the system (7.17) of the three variables $S(t)$, $x(t)$ and $v(t)$ under the risk-neutral measure, to the system (7.20) of two variables $y(t)$ and $v(t)$ under the stock price measure. It remains to find the corresponding characteristic function in the reduced system of variables, which is the subject of the now following lemma.

Lemma 7.4.1 *Under the stock price measure Q^S, the characteristic function of the T-forward asset price $\ln F(T,T) = \ln\frac{S(T)}{P(T,T)} = \ln S(T)$ conditional on the σ-algebra $\mathcal{F}_t$ is given by the following closed-form solution:*

$$\begin{aligned}&I\!E^{Q^S}\Big[\exp\Big(iu\ln F(T,T)\Big)\Big|\mathcal{F}_t\Big]\\ &= \exp\Big[A(u,t,T) + B(u,t,T)\ln F(t,T) + C(u,t,T)v(t) + \frac{1}{2}D(u,t,T)v^2(t)\Big],\end{aligned} \tag{7.23}$$

where:

$$\begin{aligned}A(u,t,T) &= \frac{1}{2}u(i-u)V(t,T) \\ &\quad + \int_t^T\Big[\Big(\widetilde{\kappa\psi} + \rho_{rv}iu\tau\sigma B_{HW}(s,T)\Big)C(s) + \frac{1}{2}\tau^2\Big(C^2(s) + D(s)\Big)\Big]ds,\end{aligned} \tag{7.24}$$

$$B(u,t,T) = iu, \tag{7.25}$$

$$C(u,t,T) = u(i-u)\frac{\Big((\gamma_3 - \gamma_4 e^{-2\gamma(T-t)}) - (\gamma_5 e^{-a(T-t)} - \gamma_6 e^{-(2\gamma+a)(T-t)}) - \gamma_7 e^{-\gamma(T-t)}\Big)}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}} \tag{7.26}$$

$$D(u,t,T) = u(i-u)\frac{1-e^{-2\gamma(T-t)}}{\gamma_1 + \gamma_2 e^{-2\gamma(T-t)}}, \tag{7.27}$$

with:

$$\gamma = \sqrt{(\widetilde{\kappa} - \rho_{Sv}\tau iu)^2 - \tau^2 u(i-u)}, \qquad \gamma_1 = \gamma + (\widetilde{\kappa} - \rho_{Sv}\tau iu), \tag{7.28}$$

$$\gamma_2 = \gamma - (\widetilde{\kappa} - \rho_{Sv}\tau iu), \qquad \gamma_3 = \frac{\rho_{Sr}\sigma\gamma_1 + \widetilde{\kappa a}\widetilde{\psi} + \rho_{rv}\sigma\tau iu}{a\gamma},$$

$$\gamma_4 = \frac{\rho_{Sr}\sigma\gamma_2 - \widetilde{\kappa a}\widetilde{\psi} - \rho_{rv}\sigma\tau iu}{a\gamma}, \qquad \gamma_5 = \frac{\rho_{Sr}\sigma\gamma_1 + \rho_{rv}\sigma\tau iu}{a(\gamma - a)},$$

$$\gamma_6 = \frac{\rho_{Sr}\sigma\gamma_2 - \rho_{rv}\sigma\tau iu}{a(\gamma + a)}, \qquad \gamma_7 = (\gamma_3 - \gamma_4) - (\gamma_5 - \gamma_6),$$

and:

$$V(t,T) := \frac{\sigma^2}{a^2}\Big((T-t) + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\Big). \tag{7.29}$$

Proof To determine the characteristic function of $\ln F(T,T)$, we can apply the Feynman-Kac theorem and reduce the problem of finding the characteristic of the forward log-asset price dynamics to solving a partial differential equation. The Feynman-Kac theorem implies that the characteristic function

$$f(t,y,v) = \mathbb{E}^{Q^T}\Big[\exp(iuy(T))\Big|\mathcal{F}_t\Big], \tag{7.30}$$

is given by the solution of the following partial differential equation,

$$\begin{aligned} 0 &= f_t + \frac{1}{2}v_F^2(t)\Big(f_{yy} + f_y\Big) + \kappa(\xi(t) - v(t))f_v \qquad (7.31)\\ &\quad + (\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_{\text{HW}}(t,T))f_{yv} + \frac{1}{2}\tau^2 f_{vv},\\ f(T,y,v) &= \exp\Big(iuy(T)\Big), \qquad (7.32)\end{aligned}$$

where the subscripts denote partial derivatives and to ease the notation we dropped the explicit (t,y,v)-dependence for f. Furthermore we have taken into account that the covariance term $dy(t)dv(t)$ is equal to

$$dy(t)dv(t) = (v(t)dW_r^T(t) + \sigma B_{\text{HW}}(t,T)dW_r^T(t))(\tau dW_v^T(t)) = (\rho_{Sv}\tau v(t) + \rho_{rv}\tau\sigma B_{\text{HW}}(t,T))dt. \tag{7.33}$$

Some tedious algebra shows that direct substitution of the ansatz

$$f(t,y,v) = \exp\Big[A(u,t,T) + B(u,t,T)y(t) + C(u,t,T)v(t) + \frac{1}{2}D(u,t,T)v^2(t)\Big], \tag{7.34}$$

solves the partial differential equation (7.31) and hence proves the theorem. □

7.4.2 Characteristic function under the T-forward measure Q^T

For the derivation of the characteristic function of $\ln S(T)$ under the T-forward measure we refer the reader to Chapter 3. For clarity, we here repeat this result.

Lemma 7.4.2 *Under the T-forward measure Q^T, the characteristic function of the T-forward asset price* $\ln F(T,T) = \ln \frac{S(T)}{P(T,T)} = \ln S(T)$ *conditional on the time t filtration* $\mathcal{F}_t$ *is given by the following closed-form solution:*

$$f(t,y,v) = \exp\Big[L(u,t,T) + M(u,t,T)y(t) + N(u,t,T)v(t) + \frac{1}{2}O(u,t,T)v^2(t)\Big], \tag{7.35}$$

where:

$$\begin{aligned} L(u,t,T) &= -\frac{1}{2}u(i+u)V(t,T) \qquad (7.36)\\ &\quad + \int_t^T \Big[\Big(\kappa\psi + \rho_{rv}(iu-1)\tau\sigma B_{HW}(s,T)\Big)N(s) + \frac{1}{2}\tau^2\Big(N^2(s,t,T) + O(s,t,T)\Big)\Big]ds,\\ M(u,t,T) &= iu, \qquad (7.37)\\ N(u,t,T) &= -u(i+u)\frac{\Big((\delta_3 - \delta_4 e^{-2\delta(T-t)}) - (\delta_5 e^{-a(T-t)} - \delta_6 e^{-(2\delta+a)(T-t)}) - \delta_7 e^{-\delta(T-t)}\Big)}{\delta_1 + \delta_2 e^{-2\delta(T-t)}}, \qquad (7.38)\\ O(u,t,T) &= -u(i+u)\frac{1-e^{-2\delta(T-t)}}{\delta_1 + \delta_2 e^{-2\delta(T-t)}}, \qquad (7.39) \end{aligned}$$

with:

$$\begin{aligned} \delta &= \sqrt{(\kappa - \rho_{Sv}\tau iu)^2 + \tau^2 u(i+u)}, & \delta_1 &= \delta + (\kappa - \rho_{Sv}\tau iu), \qquad (7.40)\\ \delta_2 &= \delta - (\kappa - \rho_{Sv}\tau iu), & \delta_3 &= \frac{\rho_{Sr}\sigma\delta_1 + \kappa a\psi + \rho_{rv}\sigma\tau(iu-1)}{a\delta},\\ \delta_4 &= \frac{\rho_{Sr}\sigma\delta_2 - \kappa a\psi - \rho_{rv}\sigma\tau(iu-1)}{a\delta}, & \delta_5 &= \frac{\rho_{Sr}\sigma\delta_1 + \rho_{rv}\sigma\tau(iu-1)}{a(\delta-a)},\\ \delta_6 &= \frac{\rho_{Sr}\sigma\delta_2 - \rho_{rv}\sigma\tau(iu-1)}{a(\delta+a)}, & \delta_7 &= (\delta_3 - \delta_4) - (\delta_5 - \delta_6), \end{aligned}$$

and with $V(t,T)$ as in (7.29).

7.5 Valuation of forward starting call options

With the preliminary work of the previous sections, we can now present the general valuation results for the forward starting characteristic functions. The results are provided in the theorems of the following two sections.

7.5.1 Forward starting characteristic function under the stock price measure

With the help of lemma 7.4.1, we are now ready to derive the characteristic function of $\ln \frac{S(T_{i-1})}{S(T_i)}$. This characteristic function, provided by the following theorem, can then directly be plugged into the Fourier inversion formula (7.8) to price the forward starting call option (7.6) in closed-form.

Theorem 7.5.1 *Under the stock price measure $\mathbb{Q}^S$, the characteristic function $\phi_F(T_{i-1}, T_i, u)$ of $\ln \frac{S(T_{i-1})}{S(T_i)}$ is given by the following closed-form solution:*

$$\begin{aligned}\phi_F(T_{i-1}, T_i, u) = & \quad \exp\left[a_0 + a_1\mu_x + \frac{1}{2}a_1^2\sigma_x^2\left(1 - \rho_{xv}^2(t, T_{i-1})\right)\right] \\ & \times \frac{\exp\left[a_2\mu_v + a_3\mu_v^2 + \frac{\left(a_1\sigma_x\rho_{xv}^2(t,T_{i-1}) + a_2\sigma_v + 2a_3\mu_v\sigma_v\right)^2}{2(1-2a_3\sigma_v^2)}\right]}{\sqrt{1 - 2a_3\sigma_v^2}},\end{aligned} \tag{7.41}$$

where:

$$a_0 := iu \ln A_{HW}(T_{i-1}, T_i) + A(-u, T_{i-1}, T_i), \quad a_1 := -iuB_{HW}(T_{i-1}, T_i), \tag{7.42}$$

$$a_2 := C(-u, T_{i-1}, T_i) \quad a_3 := \frac{1}{2}D(-u, T_{i-1}, T_i). \tag{7.43}$$

Proof Recalling the definition (7.11) for the forward asset price and using Lemma 7.4.1, one can write the following for the characteristic function $\phi_F(T_{i-1}, T_i, u)$ of $\ln \frac{S(T_{i-1})}{S(T_i)}$:

$$\begin{aligned}\phi_F(T_{i-1}, T_i, u) &= \mathbb{E}^{\mathbb{Q}^S}\left[e^{iu \ln \frac{S(T_{i-1})}{S(T_i)}}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}^S}\left[e^{iu \ln S(T_{i-1}) - iu \ln S(T_i)}\middle|\mathcal{F}_t\right] \\ &= \mathbb{E}^{\mathbb{Q}^S}\left[e^{iu \ln P(T_{i-1},T_i) + iu \ln F(T_{i-1},T_i) - iu \ln F(T_i,T_i)}\middle|\mathcal{F}_t\right].\end{aligned}$$

Using the tower law of conditional expectations, i.e. conditioning on the time T_{i-1} filtration $\mathcal{F}_{T_{i-1}}$, we have that

$$\phi_F(T_{i-1}, T_i, u) = \mathbb{E}^{\mathbb{Q}^S}\left[e^{iu \ln P(T_{i-1,T_i}) + iu \ln F(T_{i-1},T_i)}\mathbb{E}^{\mathbb{Q}^S}\left\{e^{i(-u) \ln F(T_i,T_i)}\middle|\mathcal{F}_{T_{i-1}}\right\}\middle|\mathcal{F}_t\right],$$

and note that the inner expectation is the characteristic function of $\ln F(T_{i-1}, T_i)$ evaluated in the point $-u$, i.e. given by lemma 7.4.1. Hence substituting for this the characteristic function in the above expression, we obtain:

$$\phi_F(T_{i-1}, T_i, u) = \mathbb{E}^{\mathbb{Q}^S}\left[e^{iu \ln P(T_{i-1},T_i) + A(-u,T_{i-1},T_i) + C(-uT_{i-1},T_i)v(T_{i-1}) + D(-u,T_{i-1},T_i)v^2(T_{i-1})}\middle|\mathcal{F}_t\right]. \tag{7.44}$$

In the Gaussian rate model, one has the following expression for the time-T_{i-1} price of a zero-coupon bond $P(T_{i-1}, T_i)$ maturing at time T_i (e.g. see Brigo and Mercurio (2006)):

$$P(T_{i-1}, T_i) = A_{\text{HW}}(T_{i-1}, T_i)e^{-B_{\text{HW}}(T_{i-1},T_i)x(T_{i-1})}, \tag{7.45}$$

where

$$A_{\text{HW}}(T_{i-1},T_i) = \frac{P^M(t,T_i)}{P^M(t,T_{i-1})}\exp\Big[\frac{1}{2}\Big(V(T_{i-1},T_i)-V(t,T_i)+V(t,T_{i-1})\Big)\Big] \tag{7.46}$$

$$B_{\text{HW}}(T_{i-1},T_i) = \frac{1-e^{-a(T_i-T_{i-1})}}{a} \tag{7.47}$$

$$V(T_{i-1},T_i) = \frac{\sigma^2}{a^2}\Big((T_i-T_{i-1})+\frac{2}{a}e^{-a(T_i-T_{i-1})}-\frac{1}{2a}e^{-2a(T_i-T_{i-1})}-\frac{3}{2a}\Big). \tag{7.48}$$

Hence we can express the characteristic function $\phi_F(T_{i-1},T_i,u)$ completely in terms of the Gaussian factors $x(T_{i-1})$ and $\nu^2(T_{i-1})$, i.e.

$$\begin{aligned}\phi_F(T_{i-1},T_i,u) &= \mathbb{E}^{Q^S}\Big[\exp\Big\{iu\ln A_{\text{HW}}(T_{i-1},T_i)-B_{\text{HW}}(T_{i-1},T_i)x(T_{i-1})+A(-u,T_{i-1},T_i)\\ &\qquad +C(-u,T_{i-1},T_i)\nu(T_{i-1})+D(-u,T_{i-1},T_i)\nu^2(T_{i-1})\Big\}\Big|\mathcal{F}_t\Big]\\ &=: \mathbb{E}^{Q^S}\Big[\exp\Big\{a_0+a_1x(T_{i-1})+a_2\nu(T_{i-1})+a_3\nu^2(T_{i-1})\Big\}\Big|\mathcal{F}_t\Big],\end{aligned} \tag{7.49}$$

where the last line defines the constants $a_0,\ldots,a_3$. Because the above expression is a Gaussian quadratic form of the variables $x(T_{i-1})$ and $\nu(T_{i-1})$, one can evaluate this expectation completely in terms of the means μ_x,μ_ν, variances σ_x^2,σ_ν^2 and correlation $\rho_{x\nu}(t,T_{i-1})$ of these two state variables, e.g. see Feuerverger and Wong (2000) or Glasserman (2003). A straightforward evaluation (e.g. by completing the square or by integration the exponential affine function against the bivariate normal distribution) of this Gaussian quadratic expectation results in the characteristic function $\phi_F(T_{i-1},T_i,u)$ of (7.41) and hence proves the theorem. □

7.5.2 Forward starting characteristic function under the T-forward measure

Using lemma 7.4.2 and similar arguments as in the previous section, we can now also derive the characteristic function of $\ln\frac{S(T_i)}{S(T_{i-1})}$ under the T_i-forward probability measure. This characteristic function can directly be used in the Fourier inversion formula (7.10) to price the forward starting call option (7.9) on the return of the underlying asset in closed-form.

Theorem 7.5.2 *Under the T_i-forward measure Q^{T_i}, the characteristic function $\phi_R(T_{i-1},T_i,u)$ of $\ln\frac{S(T_i)}{S(T_{i-1})}$ is given by the following closed-form solution:*

$$\begin{aligned}\phi_R(T_{i-1},T_i,u) = &\quad \exp\Big[b_0+b_1\mu_x+\frac{1}{2}b_1^2\sigma_x^2\Big(1-\rho_{x\nu}^2(t,T_{i-1})\Big)\Big]\\ &\times\ \frac{\exp\Big[b_2\mu_\nu+b_3\mu_\nu^2+\frac{\left(b_1\sigma_x\rho_{x\nu}^2(t,T_{i-1})+b_2\sigma_\nu+2b_3\mu_\nu\sigma_\nu\right)^2}{2(1-2b_3\sigma_\nu^2)}\Big]}{\sqrt{1-2b_3\sigma_\nu^2}},\end{aligned} \tag{7.50}$$

where:

$$b_0 := -iu \ln A_{HW}(T_{i-1}, T_i) + L(u, T_{i-1}, T_i), \quad b_1 := iuB_{HW}(T_{i-1}, T_i), \tag{7.51}$$

$$b_2 := N(u, T_{i-1}, T_i) \quad b_3 := \frac{1}{2}O(u, T_{i-1}, T_i). \tag{7.52}$$

Proof Using analogous arguments as in the proof of theorem 7.5.1, one can obtain the following expression for the characteristic function ϕ_R of $\ln \frac{S(T_i)}{S(T_{i-1})}$ under the T_i-forward probability measure. Using the tower law of conditional expectations, i.e. conditioning on the time T_{i-1} filtration $\mathcal{F}_{T_{i-1}}$, we have that

$$\phi_R(T_{i-1}, T_i, u) = \mathbb{E}^{Q^{T_i}}\left[e^{-iu \ln P(T_{i-1},T_i) - iu \ln F(T_{i-1},T_i)} \mathbb{E}^{Q^{T_i}}\left\{e^{iu \ln F(T_i,T_i)}\middle|\mathcal{F}_{T_{i-1}}\right\}\middle|\mathcal{F}_t\right].$$

As the inner expectation is just the characteristic function of $\ln F(T_i, T_i)$ evaluated in the point u, we can substitute the closed-form expression of lemma 7.4.1 for this characteristic function in the above expression, i.e.

$$\begin{aligned}\phi_R(T_{i-1}, T_i, u) &= \mathbb{E}^{Q^{T_i}}\left[e^{-iu \ln P(T_{i-1},T_i) + L(u,T_{i-1},T_i) + M(uT_{i-1},T_i)v(T_{i-1}) + O(u,T_{i-1},T_i)v^2(T_{i-1})}\middle|\mathcal{F}_t\right] \\ &=: \mathbb{E}^{Q^{T_i}}\left[\exp\left\{b_0 + b_1 x(T_{i-1}) + b_2 v(T_{i-1}) + b_3 v^2(T_{i-1})\right\}\middle|\mathcal{F}_t\right] \end{aligned} \tag{7.53}$$

Note that the only difference with the Gaussian quadratic form (7.49) are the dynamics of the processes $x(T_{i-1})$ and $v(T_{i-1})$, which now instead need to be evaluated under the T_i-forward measure. Hence it can be evaluated in an analogous way as in the proof of theorem 7.5.1 resulting in the closed-form expression (7.50) for the characteristic function $\phi_R(T_{i-1}, T_i, u)$ and hence proving the theorem. □

7.6 Numerical results

To investigate the impact of stochastic volatility and stochastic interest rates on the prices of forward starting options, we will consider the following numerical test cases. As the prices of forward starting options can be calculated in closed-form, a Monte Carlo benchmark against the pricing formulas (7.8)-(7.10) forms a standard test case for their implementation. We then explicitly investigate the impact and parameter sensitivities of stochastic interest rates and stochastic volatility on the prices of forward starting options. Finally, we tackle the issue of model risk and compare our framework with the Black and Scholes (1973) and Heston (1993) model, respectively considered in Rubinstein (1991) and Guo and Hung (2008) for the valuation of forward starters.

7.6.1 Implementation of the option pricing formulas

In this section we consider the practical implementation of the pricing formulas (7.8) and (7.10); both the implementation of the inverse Fourier transform, as well as the calculation of the char-

acteristic function underlying this transform, deserve some attention. For the calculation of the inverse Fourier transform we refer the reader to Lord and Kahl (2008), Kilin (2006) and Chapter 3, where this topic is covered in great detail. Instead we focus on the application specific calculation of the characteristic functions (7.41) and (7.50). The calculation of the characteristic functions (7.41) and (7.50) is trivial up to the calculation of the constants $A(u,t,T)$ of (7.24) and $L(u,t,T)$ of (7.36), which involves the calculation of a numerical integral. Hence we focus on the calculation of $A(u,t,T)$, but a completely analogous reasoning holds for the calculation of $L(u,t,T)$.

It is possible to write a closed-form expression for the remaining integral in (7.24). As the ordinary differential equation for $D(u,t,T)$ is exactly the same as in the Heston (1993) or Schöbel and Zhu (1999) model, it will involve a complex logarithm and should therefore be evaluated as outlined in Lord and Kahl (2007) in order to avoid any discontinuities. The main problem however lies in the integrals over $C(u,t,T)$ and $C^2(u,t,T)$, which will involve the Gaussian hypergeometric ${}_2F_1(a,b,c;z)$. The most efficient way to evaluate this hypergeometric function (according to Press and Flannery (1992)) is to integrate the defining differential equation. Since all of the terms involved in $D(u,t,T)$ are also required in $C(u,t,T)$, numerical integration of the second part of (7.24) seems to be the most efficient method for evaluating $A(u,t,T)$. Note that we hereby conveniently avoid any issues regarding complex discontinuities altogether. It remains to have a closer look at the implementation of the numerical integral of $A(u,t,T)$ and $L(u,t,T)$.

We compute the prices for short and long-term forward starting option for a range of strikes and where we use fixed-point Gaussian-Legendre quadrature to compute the numerical integral in (7.24) and (7.36). Hereby we vary the number of quadrature points to determine how many points are needed in the test cases to obtain a certain accuracy. The numerical results together with the corresponding Monte Carlo estimates (using 10^6 sample paths) can be found in Table **1** and **2** below.

strike level	**CF(4)**	**CF(8)**	**CF(16)**	**CF(1024)**	**MC** (±95% **interval**)
50%	65.31	65.26	65.26	65.26	65.30 (±0.31)
75%	53.94	53.85	53.85	53.85	53.89 (±0.29)
100%	44.97	44.85	44.85	44.85	44.90 (±0.27)
125%	37.80	37.65	37.65	37.65	37.71 (±0.25)
150%	32.00	31.82	31.82	31.82	31.89 (±0.24)

Table 1: Closed-form solution prices (CF(N)) using N quadrature points for $A(u,T_1,T_2)$ in (7.24) and Monte Carlo prices (MC) of the forward starting call option (7.6) for $t = 0$, $T_1 = 5$, $T_2 = 15$ and $P(t,T_1) = P(t,T_2) = 1.0$ and model parameters $\kappa = 1.00$, $v(0) = \psi = 0.20$, $a = 0.02$, $\sigma = 0.01$, $\tau = 0.50$, $\rho_{Sv} = -0.70$, $\rho_{Sr} = 0.30$ and $\rho_{rv} = 0.15$.

strike level	**CF(1)**	**CF(2)**	**CF(4)**	**CF(1024)**	**MC** (±95% **interval**)
50%	50.23	50.24	50.24	50.24	50.27 (±0.05)
75%	26.77	26.79	26.79	26.79	26.80(±0.04)
100%	8.56	8.39	8.39	8.39	8.39 (±0.03)
125%	2.07	2.04	2.04	2.04	2.05 (±0.02)
150%	0.69	0.69	0.69	0.69	0.69 (±0.01)

Table 2: Closed-form solution prices (CF(N)) using N quadrature points for $L(u, T_1, T_2)$ in (7.36) and Monte Carlo prices (MC) of the forward starting return call option (7.9) for $t = 0$, $T_1 = 1$, $T_2 = 2$ and $P(t, T_1) = P(t, T_2) = 1.0$ and model parameters $\kappa = 0.30$, $v(0) = \psi = 0.15$, $\tau = 0.20$, $a = 0.05$, $\sigma = 0.01$, $\rho_{Sv} = -0.40$, $\rho_{Sr} = 0.20$ and $\rho_{rv} = 0.10$.

From the tables we see that the characteristic functions (7.41) and (7.50) underlying the option price formulas can be calculated very accurately, using only a small number of quadrature points; the prices of short term options (Table **1**) and long-term options (Table **2**) can be calculated within a base points accuracy by using respectively just two and eight quadrature points for the calculation of the integral in $A(u, t, T)$ and $L(u, t, T)$. Note hereby that the corresponding Monte Carlo confidence interval is also larger in test case of Table **2**, due to the longer dated maturity. Combining the efficient calculation of characteristic functions (7.41) and (7.50) with the efficient Fourier inversion techniques, we can all in all conclude the pricing of forward starting options can be done fast, highly accurate and in closed-form using the latter methods.

7.6.2 Impact of stochastic interest rates and stochastic volatility

In this section we will cover the impact of stochastic volatility and (correlated) stochastic interest rates on the prices of forward starting options. That is, we investigate qualitative aspects of our extended framework in comparison to deterministic (or independent) interest rates and volatility assumptions. Rubinstein (1991) considered the pricing of a vanilla forward starting option in the Black and Scholes (1973) framework; as both interest rates and volatilities are deterministic in this model, the prices of a forward starting options are (up to deterministic discounting effects) equal for all forward starting dates. The constant volatility assumption has been relaxed by Lucic (2003), Hong (2004) and Kruse and Nögel (2005), who consider the pricing of forward starting options under Heston (1993) stochastic volatility. The impact of stochastic volatility is graphically shown in the graphs of Figure **1**.

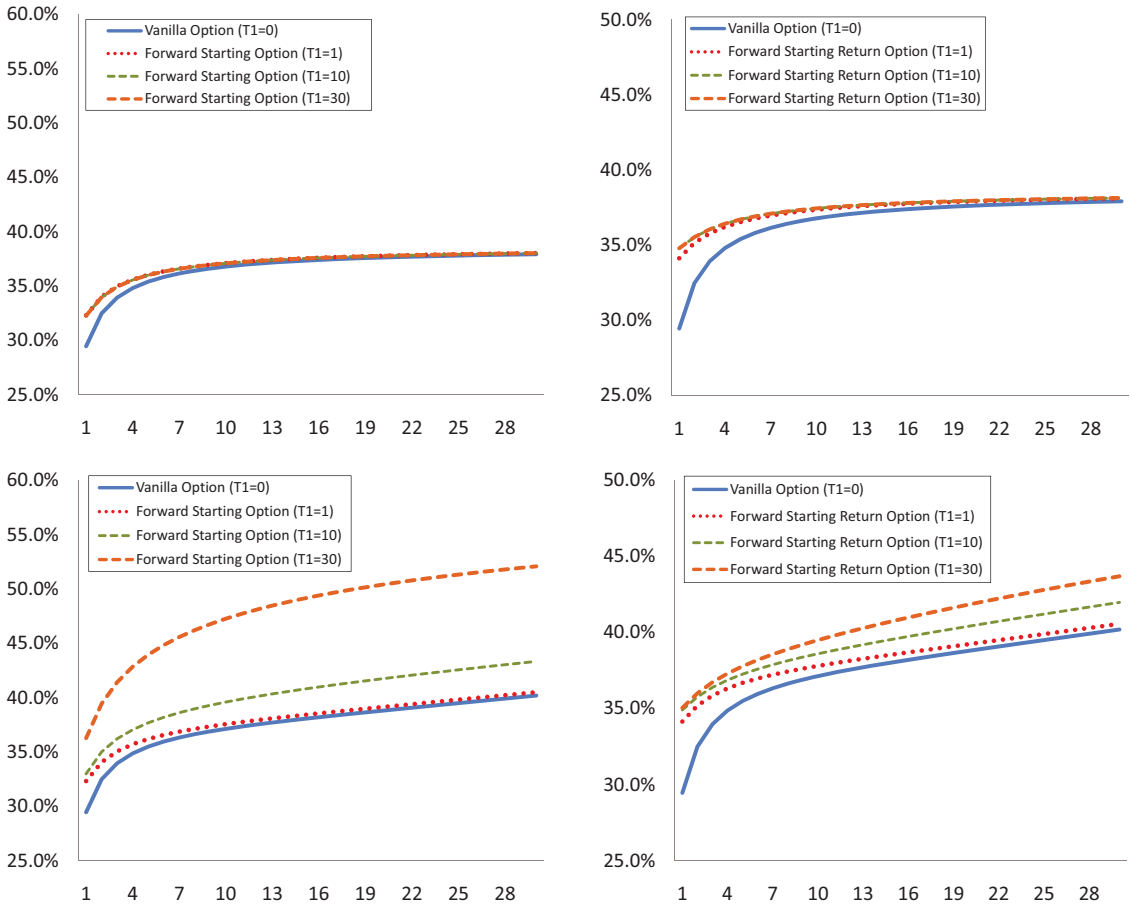

Figure 1: The figures plot, for different option maturities, the impact of stochastic interest rates on the forward implied volatility structure of an underlying call (left pictures) and return call option (right pictures). Parameters are $\kappa = 1.0$, $\nu(t) = \psi = 0.20$, $\tau = 0.5$, $\rho_{Sv} = -0.70$, $\rho_{Sr} = \rho_{rv} = 0$ and $P(t,s) = \exp(-0.04(s-t))$ for all $s > t$. The top figures plot the volatility structure for deterministic interest rates, whilst the bottom figures plot the volatility structure for stochastic interest rates case with parameters $a = 0.02$ and $\sigma = 0.01$.

Compared to constant volatility, the addition of stochastic volatility increases the future uncertainty about the underlying option price which is hence reflected in higher implied volatilities for longer forward starting dates. Intuitively this effect is rather appealing as this coincides with market prices for forward starting structures where the writer of such an option wants to be compensated for the extra (future) volatility risk he is exposed to. Furthermore it is interesting to note from the figures that these effects are more apparent where the underlying option has a short maturity, which effect may be explained by the mean reverting property of stochastic volatility that is less severe for a short term option hence increasing the future volatility risk. Finally note from the top two graphs of Figure **1** that with deterministic rates the long-term uncertainty approaches a limit (or a stationary state) as the forward starting date or the underlying option maturity increase. For example the implied volatilities for forward starting options with a forward date of ten and thirty years are exactly equal, which is counterintuitive as the term structure of implied

volatilities remains increasing for long-dated options and in general does not flatten out nor approaches a limit, for instance see the implied volatility quotes in long-maturity equity markets (readily available from MarkIT or Bloomberg) or the over-the-counter FX quotes in Piterbarg (2005) or Andreasen (2006).
The inconsistency in the way the market and the latter models look at long-dated implied volatility structures, more likely suggests that these models lack an extra factor in their pricing frameworks; this conjecture is supported by Guo and Hung (2008) and Kijima and Muromachi (2001), who claim that stochastic interest rates are crucial for the pricing of forward starting options as these securities are often much more interest rate sensitive due to their long-term nature. In fact if we look at the bottom graphs of Figure **1**, where we add stochastic interest rates to the framework with stochastic volatility, we see that the implied volatilities increase for longer forward starting and maturity dates. These model effects also correspond with a general feature of the interest rate market: the market's view on the uncertainty of long-maturity bonds is often much higher than that of shorter bond, reflecting the increasing impact of stochastic interest rates for long-dated structures. In this sense stochastic interest rates do seem to incorporate the larger uncertainty the writers of the forward starting options are exposed to.
The addition of stochastic interest rates as independent factor for the pricing of forward starting options has been investigated in Guo and Hung (2008) and Nunes and Alcaria (2009). Though one step in the right direction, the independency assumption is certainly not supported by empirical analysis (e.g. see Baur (2009)) nor do the exotic option markets (such as hybrid equity-interest rate options) price these derivatives in this way, e.g. see Andreasen (2007) or Antonov et al. (2008); from Figure **2** and **3** of Appendix 7.8.3, we see that correlated stochastic interest rates can have a big impact on the prices of forward starting options. From Figure **2** we can see that for a positive rate-asset correlation coefficient the prices of forward starting options increase and vice versa for a negative correlation coefficient. In particular note from Figure **2** that, though the correlation coefficient between the interest rates and the stock also affects the implied volatility structure of the current time vanilla options, the effects on the prices of forward starting options are much more pronounced. Forward starting options are thus not only more interest rate and volatility sensitive, but are also much more exposed to correlation risks. This is not surprising as a joint movement in both the interest rates as the asset price not only affects the future discounting, but more importantly also the (joint) asset price distribution. All in all, we can conclude that because forward starting options are very sensitive to future interest rate movements, volatility smiles as well as their dependency structure with the underlying asset, it is very important to take all these stochastic quantities into account for a proper pricing and risk management of these derivatives.

7.7 Conclusion

We performed a quantitative analysis on the valuation of forward starting options, where we explicitly accounted for stochastic volatility, stochastic interest rates as well as a general dependency structure between all underlying processes. The analysis was made possible by the development of closed-form formulas involving the pricing of the two main forward starting

structures, currently present in the literature and the financial markets. Using a probabilistic approach, we derived closed-form expressions for the characteristic functions of the assets underlying the forward starting options. We then demonstrated how forward starting options can be priced efficiently and in closed-form by Fourier inverting these forward starting characteristic functions. An additional advantage of this technique is that our modelling framework can include jumps as a trivial extension, since we already work in the Fourier option pricing domain.
Our results are of great practical importance as the derivative markets for long-dated dynamic securities such as forward starting options have grown very rapidly over the last decade; compared to vanilla options, these structures directly depend on future volatility smiles, the term structure of interest rates as well as their dependency structure with the underlying asset. Moreover, as these contracts often incorporate long-dated maturities, we found that it is of crucial importance to take stochastic interest rates, volatility and a general correlation structure into account for a proper valuation and hedging of these securities: not doing so leads to serious mispricings, not to mention the potential hedge errors. Compared to other models, the analysis performed in our framework stands out by modelling both the stochastic volatility and interest rates, as well as taking a general correlation structure between all underlying drivers explicitly into account.
Besides investigating the behaviour of these dynamic derivatives, our formulas can also be used to directly price or hedge financial contracts. For instance unit-linked guarantees embedded in life insurance products, being sold in large amounts by insurance companies, can be priced in closed-form relying on our formulas. The same applies for cliquet options, which are heavily traded in over-the-counter markets, and CEO/employee stock option plans. Furthermore, there is a big interaction between forward starting options considered here and over-the-counter exotic structures such as ratchet options and pension contracts, as these form the natural building blocks and hedge instruments for such contracts. Finally, as all the above-mentioned products explicitly depend on the future volatility smiles, the term structure of interest rates as well as their dependency structure with the underlying asset, we judge that a proper valuation framework should account for all these characteristics.

7.8 Appendix

7.8.1 Calculation of the moments for the rate processes under different measures

Stock price measure

For the computation of the characteristic functions from Theorem 7.5.1 one needs the first two moments of $x(T_{i-1})$ and $v(T_{i-1})$ (conditional on the σ-algebra $\mathcal{F}_t$) under the stock price measure Q^S. For completeness, we will therefore explicitly provide the analytical expressions for these moments: integrating the dynamics (7.17) and using Fubini's theorem, results (after some algebra) in the following explicit solutions:

$$\begin{aligned}
v(T_{i-1}) &= \widetilde{\psi} + \Big(v(t) - \widetilde{\psi}\Big)e^{-\widetilde{\kappa}(T_{i-1}-t)} + \tau \int_t^{T_{i-1}} e^{-\widetilde{\kappa}(T_{i-1}-u)} dW_v^{Q^S}(u), \\
x(T_{i-1}) &= \rho_{Sr}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-a(T_{i-1}-t)}] + \frac{v(t) - \widetilde{\psi}}{a - \widetilde{\kappa}}[e^{-\widetilde{\kappa}(T_{i-1}-t)} - e^{-a(T_{i-1}-t)}]\Big) \\
&\quad + \frac{\rho_{Sr}\sigma\tau}{(a - \widetilde{\kappa})} \int_t^{T_{i-1}} [e^{-\widetilde{\kappa}(T_{i-1}-u)} - e^{-a(T_{i-1}-u)}] dW_v^{Q^S}(u) + \sigma \int_t^{T_{i-1}} e^{-a(T_{i-1}-u)} dW_r^{Q^S}(u).
\end{aligned}$$

Using Itô's isometry, one therefore has that the pair $\Big(v(T), x(T)\Big)$, under the stock price measure and conditional on $\mathcal{F}_t$, follow a bivariate normal distribution with means μ_v,μ_x, variances σ_v^2, σ_x^2 and correlation $\rho_{xv}(t, T_{i-1})$ respectively given by

$$\mu_v = \widetilde{\psi} + \Big(v(t) - \widetilde{\psi}\Big)e^{-\widetilde{\kappa}(T_{i-1}-t)} \tag{7.54}$$

$$\sigma_v^2 = \frac{\tau^2}{2\widetilde{\kappa}}\Big(1 - e^{-2\widetilde{\kappa}(T_{i-1}-t)}\Big), \tag{7.55}$$

$$\mu_x = \rho_{Sr}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-a(T_{i-1}-t)}] + \frac{v(0) - \widetilde{\psi}}{(a - \widetilde{\kappa})}[e^{-\widetilde{\kappa}(T_{i-1}-t)} - e^{-a(T_{i-1}-t)}]\Big), \tag{7.56}$$

$$\sigma_x^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \tag{7.57}$$

$$\rho_{xv}(t, T_{i-1}) = \frac{\rho_{rv}\sigma\tau}{\sigma_x\sigma_v(a + \widetilde{\kappa})}\Big[1 - e^{-(a+\widetilde{\kappa})(T_{i-1}-t)}\Big], \tag{7.58}$$

where

$$\begin{aligned}
\sigma_1 &= \sigma\sqrt{\frac{1-e^{-2a(T_{i-1}-t)}}{2a}},\\
\sigma_2 &= \frac{\rho_{Sr}\sigma\tau}{a-\widetilde{\kappa}}\sqrt{\frac{1}{2\widetilde{\kappa}}+\frac{1}{2a}-\frac{2}{(\widetilde{\kappa}+a)}-\frac{e^{-2\widetilde{\kappa}(T_{i-1}-t)}}{2\widetilde{\kappa}}-\frac{e^{-2a(T_{i-1}-t)}}{2a}+\frac{2e^{-(\widetilde{\kappa}+a)(T_{i-1}-t)}}{(\widetilde{\kappa}+a)}},\\
\rho_{12} &= \rho_{rv}\frac{\sigma^2\rho_{Sr}\tau}{\sigma_1\sigma_2(a-\widetilde{\kappa})}\Big[\frac{1-e^{-(a+\widetilde{\kappa})(T_{i-1}-t)}}{(a+\widetilde{\kappa})}-\frac{1-e^{-2a(T_{i-1}-t)}}{2a}\Big].
\end{aligned}$$

T-forward measure

For computing the characteristic functions from theorem 7.5.2, the first moments of $x(T_{i-1})$ and $v(T_{i-1})$ under the T_i-forward measure $\mathcal{Q}^{T_i}$ are needed; one can obtain the following explicit solutions for $x(T_{i-1})$ and $v(T_{i-1})$ by direct integration of the corresponding T_i-forward dynamics, i.e.

$$\begin{aligned}
x(T_{i-1}) &= x(t)e^{-a(T_{i-1}-t)} - M^{T_i}(t,T_{i-1}) + \sigma\int_t^{T_{i-1}} e^{-a(T_{i-1}-u)}dW_r^{T_i}(u),\\
v(T_{i-1}) &= v(t)e^{-\kappa(T_{i-1}-t)} + \int_t^{T_{i-1}}\kappa\xi(u)e^{-\kappa(T_{i-1}-u)}du + \int_t^{T_{i-1}}\tau e^{-\kappa(T_{i-1}-u)}dW_v^{T}(u),
\end{aligned}$$

where

$$\xi(u) := \psi - \frac{\rho_{rv}\sigma\tau}{a\kappa}(1-e^{a(T_i-u)}), \tag{7.59}$$

$$M^{T_i}(t,T_{i-1}) := \frac{\sigma^2}{a^2}\Big(1-e^{-a(T_{i-1}-t)}\Big) - \frac{\sigma^2}{2a^2}\Big(e^{-a(T_i-T_{i-1})} - e^{-a(T_i+T_{i-1}-2t)}\Big). \tag{7.60}$$

Hence, from Itô's isometry, we immediately have that the pair $\Big(v(T_{i-1}), x(T_{i-1})\Big)$, under the T_i-forward measure and conditional on $\mathcal{F}_t$, follows a bivariate normal distribution, respectively with means μ_v, μ_x, variances $\sigma_n u^2$, σ_x^2 and correlation $\rho_{xv}(t,T_{i-1})$ given by

$$\begin{aligned}
\mu_v &= v(t)e^{-\kappa(T_{i-1}-t)} + \Big(\psi - \frac{\rho_{rv}\sigma\tau}{a\kappa}\Big)\Big(1-e^{-\kappa(T_{i-1}-t)}\Big)\\
&\quad -\frac{\rho_{rv}\sigma\tau}{a(\kappa+a)}\Big(e^{-a(T_i-t)-\kappa(T_{i-1}-t)} - e^{-a(T_i-T_{i-1})}\Big), &(7.61)\\
\sigma_v^2 &= \frac{\tau^2}{2\kappa}\Big(1-e^{-2\kappa(T_{i-1}-t)}\Big), &(7.62)\\
\mu_x &= x(t)e^{-a(T_{i-1}-t)} - M^{T_i}(t,T_{i-1}), &(7.63)\\
\sigma_x^2 &= \frac{\sigma^2}{2a}\Big(1-e^{-2a(T_{i-1}-t)}\Big), &(7.64)\\
\rho_{xv}(t,T_{i-1}) &= \frac{\rho_{rv}\sigma\tau}{\sigma_x\sigma_v(a+\kappa)}\Big[1-e^{-(a+\kappa)(T_{i-1}-t)}\Big]. &(7.65)
\end{aligned}$$

7.8.2 Impact of the rate-asset correlation coefficient on the forward starting options

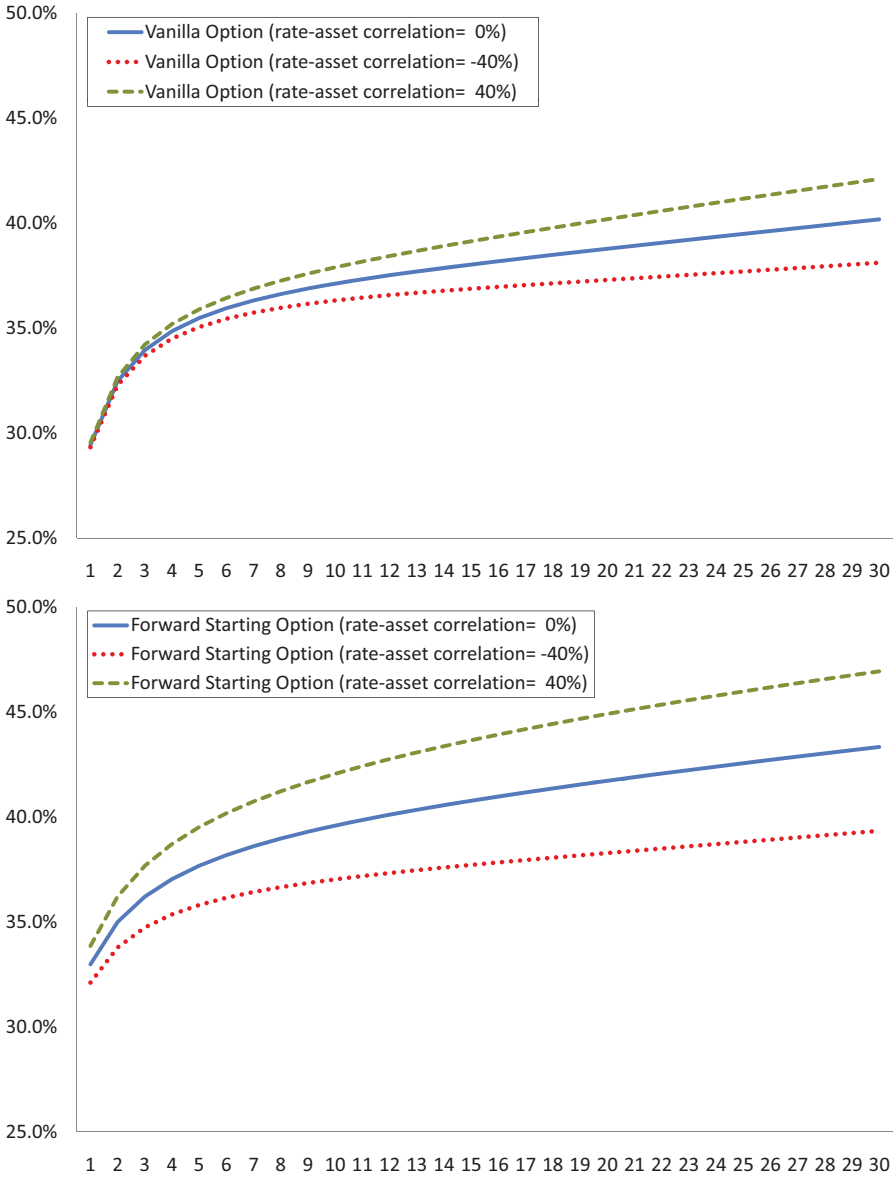

Figure 2: Impact of the rate-asset correlation ρ_{Sr} on the (forward) implied volatility structure. for different underlying call option maturities. Parameters are $\kappa = 1.0$, $\nu(t) = \psi = 0.20$, $\tau = 0.5$, $a = 0.02$, $\sigma = 0.01$, $\rho_{S\nu} = -0.70$, $\rho_{r\nu} = 0$ and $P(t,s) = \exp(-0.04(s-t))$ for all $s > t$. The top figure shows the impact of this correlation on the volatilities of the current time (vanilla) options, whereas the bottom figure plots these volatility structures for forward starting call options with strike determination date $T_1 = 10$ year.

7.8.3 Impact of the rate-volatility correlation coefficient on the forward starting options

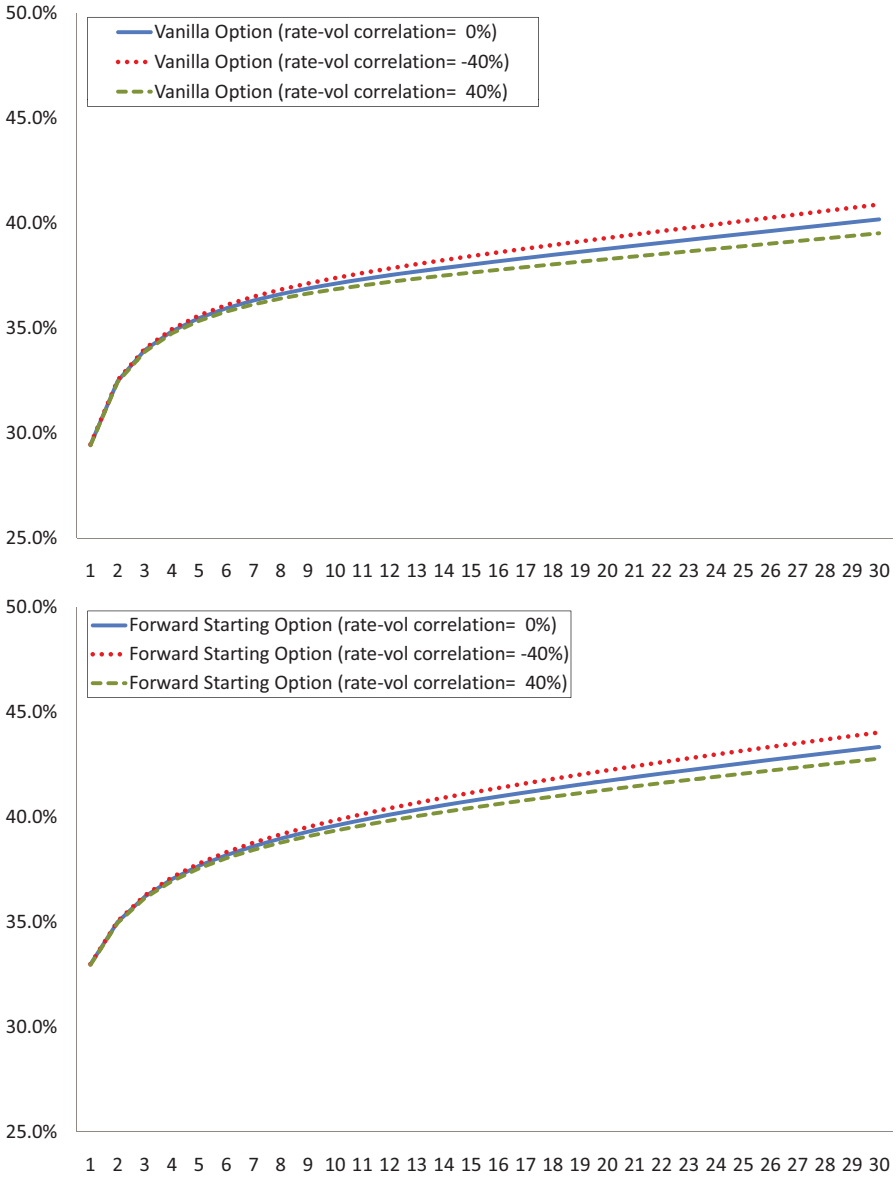

Figure 3: Impact of the rate-volatility correlation ρ_{rv} on the (forward) implied volatility structure. for different underlying call option maturities. Parameters are $\kappa = 1.0$, $v(t) = \psi = 0.20$, $\tau = 0.5$, $a = 0.02$, $\sigma = 0.01$, $\rho_{Sv} = -0.70$, $\rho_{Sr} = 0$ and $P(t,s) = \exp(-0.04(s-t))$ for all $s > t$. The top figure graphs the impact of this correlation on the volatilities of the current time (vanilla) options, whereas the bottom figure plots these volatility structures for forward starting call options with strike determination date $T_1 = 10$ year.

Chapter 8

Valuation of Guaranteed Annuity Options using a Stochastic Volatility Model for Equity Prices

*This chapter is based on:

Alexander van Haastrecht, Richard Plat and Antoon Pelsser (2009). "Valuation of Guaranteed Annuity Options using a Stochastic Volatility Model for Equity Prices", *working paper*.

8.1 Introduction

Life insurers often include embedded options in the terms of their products. One of the most familiar embedded options is the Guaranteed Annuity Option (GAO). A GAO provides the right to convert a policyholder's accumulated funds to a life annuity at a fixed rate when the policy matures. These options were a common feature in retirement savings contracts issued in the 1970's and 1980's in the United Kingdom (UK). According to Bolton et al. (1997) the most popular guaranteed conversion rate was about 11%. Due to the high interest rates at that time, the GAOs were far out of the money. However, as the interest rate levels decreased in the 1990's and the (expected) mortality rates improved, the value of the GAOs increased rapidly and amongst others led to the downfall of Equitable Life in 2000. Currently, similar options are frequently sold under the name Guaranteed Minimum Income Benefit (GMIB) in the U.S. and Japan as part of variable annuity products. The markets for variable annuities in the U.S. and Japan have grown explosively over the past years, and a growth in Europe is also expected, see Wyman (2007).

During the last decade the literature on pricing and risk management of these options has evolved. Approaches for risk management and hedging of GAOs were described in Dunbar (1999), Yang (2001), Wilkie et al. (2003) and Pelsser (2003). The pricing of GAOs and GMIBs has been described by several authors, for example van Bezooyen et al. (1998), Boyle and Hardy

(2001), Ballotta and Haberman (2003), Boyle and Hardy (2003), Biffis and Millossovich (2006), Chu and Kwok (2007), Bauer et al. (2008) and Marshall et al. (2009). In most of these papers, the focus is on unit linked deferred annuity contracts purchased originally by a single premium. Generally a standard geometric Brownian motion is assumed for equity prices. However, Ballotta and Haberman (2003) and Chu and Kwok (2007) noted that, given the long maturities of the insurance contracts, a stochastic volatility model for equity prices would be more suitable.

In this chapter closed form expressions are derived for prices of GAOs, assuming stochastic volatility for equity prices and (of course) stochastic interest rates. The model used for this is the Schöbel-Zhu Hull-White (SZHW) model, introduced in Chapter 3. The model combines the stochastic volatility model of Schöbel and Zhu (1999) with the 1-factor Gaussian interest rate model of Hull and White (1993), taking the correlation structure between those processes explicitly into account. Furthermore, this is extended to the case of a 2-factor Gaussian interest rate model.

The remainder of the chapter is organized as follows. First, in Section 8.2 the characteristics of the GAO are given. Section 8.3 describes the SZHW model to be used for the pricing of the GAO. In Section 8.4 closed-form pricing formulas are derived for the GAOs given an underlying SZHW model. These results are extended to a 2-factor Hull-White model in Section 8.5. In Section 8.6 two numerical example are worked out: the first shows the impact of stochastic volatility on the pricing of the GAO, whilst the second example deals with a comparison of the efficiency of our closed-form formula for the 2-factor model with existing methods in the literature. Conclusions are given in Section 8.7.

8.2 Guaranteed Annuity Contract

A GAO gives the holder the right to receive at the retirement data T either a cash payment equal to the investment in the equity fund $S(T)$ or a life annuity of this investment against the guaranteed rate g. A rational policy holder would choose the greater of the two assets. In other words, if at inception, the policy holder is aged x and the normal retirement date is at time T, then the annuity value at maturity is $S(T) + H(T)$, with GAO payoff $H(T)$ equal to

$$H(T) := \left(gS(T) \sum_{i=0}^{n} c_i P(T, t_i) - S(T) \right)^+, \tag{8.1}$$

provided that the policy holder is still alive at that time. Here g is the guaranteed rate, $P(T, t_i)$ the zero-coupon bond at time T maturing at t_i and c_i the insurance amounts for time i multiplied by the probability of survival from time T until time t_i for the policyholder. Without loss of generality, we will use unit insured amounts in the remainder of this chapter. Furthermore, we assume that the survival probabilities are independent of the equity prices and interest rates. Note

that

$$H(T) = gS(T)\Big(\sum_{i=0}^{n} c_i P(T,t_i) - K\Big)^+, \tag{8.2}$$

where $K := 1/g$ and $(x)^+ := \max(x,0)$. This last equality shows that one can interpret the guaranteed annuity option as a quanto call option with strike K on the zero-coupon bond portfolio $\sum_{i=0}^{n} c_i P(T,t_i)$ which is payed out using the exchange rate/currency $S(T)$, e.g. see Boyle and Hardy (2003). Under the risk-neutral measure Q, which uses the money market account $B(T)$,

$$B(T) := \exp\Big(\int_0^T r(u)du\Big) \tag{8.3}$$

as numeraire, the price of this option can be expressed as

$$C(T) = {}_xp_r\, \mathbb{E}^Q\Big[\exp\Big(-\int_0^T r(u)du\Big) gS(T)\Big(\sum_{i=0}^{n} c_i P(T,t_i) - K\Big)^+\Big], \tag{8.4}$$

where ${}_xp_r$ denotes the probability that the policy holder aged x survives until retirement age r at time T. To derive a closed-form expression for the GAO of (8.4), it is more convenient to measure payments in terms of units of stock instead of money market values. Mathematically, we can establish this by using the equity price $S(T)$ as numeraire and changing from the risk-neutral measure to the equity-price measure Q^S, see Geman et al. (1996). Under the equity-price measure Q^S, the GAO price is then given by

$$C(T) = {}_xp_r\, gS(0)\mathbb{E}^{Q^S}\Big[\Big(\sum_{i=0}^{n} c_i P(T,t_i) - K\Big)^+\Big]. \tag{8.5}$$

To evaluate this expectation we need to take into account the dynamics of the zero-coupon bonds prices $P(T,t_i)$ under the equity price measure.

Apart from the guaranteed rate, the drivers of the GAO price are the interest rates, the equity prices, the correlation between those, and the survival probabilities. The combined model for interest rates and equity prices is explained in Section 8.3. This model needs an assumption for the correlation, which could be derived from historical data. Note that if it is assumed that equity prices and interest rates are independent, expression (8.4) can be simplified to:

$$C(T) = {}_xp_r\, \mathbb{E}^Q\big[S(T)\big]\mathbb{E}^Q\Big[\exp\Big(-\int_0^T r(u)du\Big) g\Big(\sum_{i=0}^{n} c_i P(T,t_i) - K\Big)^+\Big]. \tag{8.6}$$

This means that under the assumption of independence between interest rates and equity prices, it does not matter which model is assumed for equity prices.[1] Both from historical data as well

[1] Explicit pricing formulas, for this case, under one and two-factor Gaussian interest rates are provided in appendix 8.8.3.

from market quotes, one however rarely finds that the equity prices and interest rates behave in an independent fashion. As this dependency structure is one of the main driver for the GAO price and its sensitivities, a non-trivial structure therefore has to be taken into account for a proper pricing and risk management of these derivatives, e.g. see Boyle and Hardy (2003), Ballotta and Haberman (2003) or Baur (2009).

8.3 The Schöbel-Zhu-Hull-White model

The model used for the applications in this chapter is the Schöbel-Zhu Hull-White (SZHW) model, see Chapter 3. The model combines the stochastic volatility model of Schöbel and Zhu (1999) with the 1-factor Gaussian interest rate model of Hull and White (1993), taking explicitly into account the correlation between those processes. Having the flexibility to correlate the equity price with both stochastic volatility and stochastic interest rates yields a realistic model, which is of practical importance for the pricing and hedging of options with long-term exposures such as guaranteed annuities, e.g. see Boyle and Hardy (2003).

It is hardly necessary to motivate the inclusion of stochastic volatility in a pricing model for GAOs, or long-term derivatives in general. First, compared to constant volatility models, stochastic volatility models are significantly better able to fit the market's option data, e.g. see Andreasen (2006) or Andersen and Brotherton-Ratcliffe (2001). Secondly, as stochastic interest rates and stochastic volatility are empirical phenomena, the addition of these factors yields a more realistic model, which becomes important for the pricing and especially the hedging of long-term derivatives. The addition of stochastic volatility and stochastic interest rates as stochastic factors is important when considering long-maturity equity derivatives and has been the subject of empirical investigations most notably by Bakshi et al. (2000). These authors show that the hedging performance of delta hedging strategies of long-maturity options improves when stochastic volatility and stochastic interest rates into account.
Stochastic volatility models have been described by several others, for example Stein and Stein (1991), Heston (1993), Schöbel and Zhu (1999), Duffie et al. (2000), Duffie et al. (2003), van der Ploeg (2006) and Chapter 3. Also regime-switching models are suggested in the literature for the pricing of equity-linked insurance policies, e.g. see Hardy (2001) and Brigo and Mercurio (2006). In the limit of an infinite number of regimes these models again converge to a continuous-time stochastic volatility model, however in discrete time they can benefit from a greater analytical tractability. A proper model assessment, greatly depends on the properties of the embedded options in the insurance contract. To investigate the impact of using a stochastic volatility model on the pricing of GAOs, note that the GAO directly depends on the stochastic interest rates, the underlying equity fund and the correlation between the rates and the equity. For the pricing of GAOs we therefore choose to use the SZHW model over other stochastic volatility models, as this model distinguishes itself by an explicit incorporation of the correlation between underlying equity fund and the term structure of interest rates, whilst maintaining a high degree of analytical tractability.

8.3.1 Calibration of the SZHW and BSHW model

In Section 8.6 the impact of stochastic volatility on the pricing of GAOs is analyzed. That is, we compare the pricing of GAOs in the SZHW stochastic volatility model with the Black-Scholes Hull-White (BSHW) constant volatility model, described in Chapter 2. This section is devoted to the calibration of these models and a short analysis of these calibration results.
The BSHW process for equity prices $S(t)$ under the risk neutral measure Q is:

$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma_S dW_S^Q(t), \qquad S(0) = S_0, \tag{8.7}$$

where the interest rate $r(t)$ follows Hull and White (1993) dynamics, see Chapter 2, and with the instantaneous correlation between Brownian motions of the interest rate and the equity price equal to

$$dW_S^Q(t)dW_x^Q(t) = \rho_{xS}dt. \tag{8.8}$$

To come up with a fair analysis of the impact of stochastic volatility on the pricing of GAOs, we first calibrate the BSHW and SZHW model to market's option data per end July 2007. First the Hull and White (1993) interest rate models are respectively calibrated to the EU and U.S. swaption markets. Secondly for calibration of the equity price specific model parameters, data on the Eurostoxx50 index (EU) and the S&P500 (U.S.) is used. The effective (10 years) correlation between the stock and the interest rates in the BSHW process, was hereby determined using time series analysis of the interest rates and the Eurostoxx50 (EU) and S&P500 (U.S.) index over the period from February 2002 to July 2007. For the EU and the U.S. we respectively found a correlation coefficient of 34.65% and 14.64% between the interest rates and the equity price. Note that for the aid of a fair comparison between the models, the SZHW model is calibrated in such a way that the effective correlation between interest rates and equity prices is equal to that of the BSHW process. Finally, as the considered GAO in Section 8.6 has a 10 year maturity, we need to calibrate the equity specific to the terminal distributions of the equity price at that time. To this end, we calibrate the equity models to market's options maturing in 10 years time. The calibration results to the Eurostoxx50 and S&P500 can be found in Table **1** below.

strike	Market	SZHW	BSHW
80	27.8%	27.9%	26.4%
90	27.1%	27.1%	26.4%
95	26.7%	26.7%	26.4%
100	26.4%	26.4%	26.4%
105	26.0%	26.0%	26.4%
110	25.7%	25.7%	26.4%
120	25.1%	25.1%	26.4%

Implied volatility, 10-year call options, EU.

strike	Market	SZHW	BSHW
80	27.5%	27.5%	25.8%
90	26.6%	26.6%	25.8%
95	26.2%	26.2%	25.8%
100	25.8%	25.8%	25.8%
105	25.4%	25.4%	25.8%
110	25.0%	25.0%	25.8%
120	24.3%	24.4%	25.8%

Implied volatility, 10-year call options, US

Table 1: Comparison of the calibration results for the SZHW and BSHW model for 10-year call options with different strikes. Calibrations are performed on market data for options of major indices at the end of July 2007: for EU index the EuroStoxx50 is used, whereas for US index this is the S&P500.

Notice from the tables we can see that SZHW is significantly better in capturing the market's implied volatility structure and provides an extremely good fit. The fits of the BSHW model are relatively poor. Furthermore, a direct consequence of the log-normal distribution of the BSHW model, it that the asset returns have thin tails, which does not correspond to historical data nor to the market's view on long-term asset returns. In this way, the SZHW model provides a more realistic picture on the market's view on long-term asset returns as it can incorporate heavy-tailed returns. The latter can be made especially clear by looking at the risk-neutral densities of the log-asset price of the SZHW and BSHW model. These are plotted in Figure **1** below for the BSHW and SZHW model, calibrated to EU option prices.

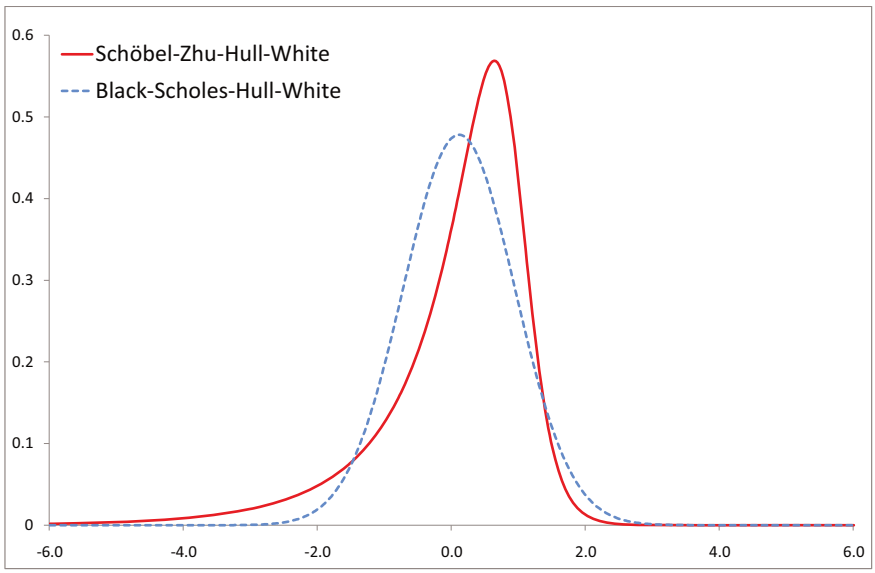

Figure 1: Risk-neutral density of the log-asset price for the SZHW and BSHW model, calibrated to EU market option data.

Clearly, the SZHW incorporates the skewness and heavy-tails seen in option markets (e.g. see Bakshi et al. (1997)) a lot more realistically than the BSHW model. The effects of these log-asset price distributions on the pricing of GAOs, combined with correlated interest rates, are extensively analyzed in Section 8.6.

8.4 Pricing the GAO under stochastic volatility and stochastic interest rates

For the pricing of the GAO in the SZHW model, i.e. the evaluation of (8.5), we need to consider the pricing of a zero-coupon bonds in the Gaussian rate model. In the Hull and White (1993) model, as demonstrated in Chapter 2, one has the following expression for the time-T price of a zero-coupon bond $P(T, t_i)$ maturing at time t_i :

$$P(T, t_i) = A(T, t_i)e^{-B(T,t_i)x(T)}, \tag{8.9}$$

where

$$A(T, t_i) = \frac{P^M(0, t_i)}{P^M(0, T)} \exp\Big[\frac{1}{2}\Big(V(T, t_i) - V(0, t_i) + V(0, T)\Big)\Big] \tag{8.10}$$

$$B(T, t_i) = \frac{1 - e^{-a(t_i - T)}}{a} \tag{8.11}$$

$$V(T, t_i) = \frac{\sigma^2}{a^2}\Big((t_i - T) + \frac{2}{a}e^{-a(t_i-T)} - \frac{1}{2a}e^{-2a(t_i-T)} - \frac{3}{2a}\Big), \tag{8.12}$$

and with $P^M(0, s)$ denoting the market's time zero discount factor maturing at time s. Using (8.9), we have for the GAO price (8.5) under the equity price measure Q^S:

$$C(T) = {}_xp_r\, gS(0)\mathbb{E}^{Q^S}\Big[\Big(\sum_{i=0}^{n} c_i A(T, t_i)e^{-B(T,t_i)x(T)} - K\Big)^+\Big]. \tag{8.13}$$

To further evaluate this expression, we first have to consider the dynamics of $x(T)$ under the equity price measure Q^S in the SZHW model.

8.4.1 Taking the equity price as numeraire

To change the money market account numeraire into the equity price numeraire, we need to change the underlying probability measure, see Chapter 2. The associated Radon-Nikodým derivative is given by

$$\frac{dQ^S}{dQ} = \frac{S(T)B(0)}{S(0)B(T)} = \exp\Big[-\frac{1}{2}\int_0^T \nu^2(u)du + \int_0^T \nu(u)dW_S^Q(u)\Big]. \tag{8.14}$$

The multi-dimensional version of Girsanov's theorem hence implies that

$$dW_S^{Q^S}(t) \mapsto dW_S^Q(t) - \nu(t)dt, \tag{8.15}$$
$$dW_x^{Q^S}(t) \mapsto dW_x^Q(t) - \rho_{xS}\nu(t)dt, \tag{8.16}$$
$$dW_\nu^{Q^S}(t) \mapsto dW_\nu^Q(t) - \rho_{\nu S}\nu(t)dt, \tag{8.17}$$

are Q^S Brownian motions. Hence under Q^S one has the following model dynamics for the volatility and interest rate process

$$dx(t) = -ax(t)dt + \rho_{xS}\sigma\nu(t)dt + \sigma dW_x^{Q^S}(t), \qquad x(0) = 0, \tag{8.18}$$
$$\begin{aligned} d\nu(t) &= \kappa\Big(\psi - \nu(t)\Big)dt + \rho_{\nu S}\tau\nu(t)dt + \tau dW_\nu^{Q^S}(t), \\ &= \widetilde{\kappa}\Big(\widetilde{\psi} - \nu(t)\Big)dt + \tau dW_\nu^{Q^S}(t), \qquad \nu(0) = \nu_0, \end{aligned} \tag{8.19}$$

where $\widetilde{\kappa} := \kappa - \rho_{\nu S}\tau$, $\widetilde{\psi} := \frac{\kappa\psi}{\widetilde{\kappa}}$. The case $\tilde{\kappa} \equiv 0$ the volatility $\nu(t)$ follows a standard Wiener process, can trivially be dealt using the techniques for a general $ka\tilde{p}pa \neq 0$, in which the dynamics of $\nu(t)$ follow a mean reverting Ornstein Uhlenbeck process; That is, conditional on the current time σ-algebra $\mathcal{F}_0$, after some calculations, one can show that for $ka\tilde{p}pa \neq 0$:

$$\nu(T) = \widetilde{\psi} + \Big(\nu(0) - \widetilde{\psi}\Big)e^{-\widetilde{\kappa}T} + \tau\int_0^T e^{-\widetilde{\kappa}(T-u)}dW_\nu^{Q^S}(u), \tag{8.20}$$

$$\begin{aligned} x(T) &= \rho_{xS}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-aT}] + \frac{\nu(0) - \widetilde{\psi}}{a - \widetilde{\kappa}}[e^{-\widetilde{\kappa}T} - e^{-aT}]\Big) \\ &\quad + \frac{\rho_{xS}\sigma\tau}{(a - \widetilde{\kappa})}\int_0^T [e^{-\widetilde{\kappa}(T-u)} - e^{-a(T-u)}]dW_\nu^{Q^S}(u) + \sigma\int_0^T e^{-a(T-u)}dW_x^{Q^S}(u). \end{aligned} \tag{8.21}$$

Using Itô's isometry and Fubini's theorem, we have that $x(T)$ (conditional on $\mathcal{F}_0$) is normally distributed with mean μ_x and variance σ_x^2 given by

$$\mu_x = \rho_{xS}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-aT}] + \frac{\nu(0) - \widetilde{\psi}}{(a - \widetilde{\kappa})}[e^{-\widetilde{\kappa}T} - e^{-aT}]\Big), \tag{8.22}$$
$$\sigma_x^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2. \tag{8.23}$$

where

$$\sigma_1 = \sigma\sqrt{\frac{1-e^{-2aT}}{2a}}, \tag{8.24}$$

$$\sigma_2 = \frac{\rho_{xS}\sigma\tau}{a-\widetilde{\kappa}}\sqrt{\frac{1}{2\widetilde{\kappa}}+\frac{1}{2a}-\frac{2}{(\widetilde{\kappa}+a)}-\frac{e^{-2\widetilde{\kappa}T}}{2\widetilde{\kappa}}-\frac{e^{-2aT}}{2a}+\frac{2e^{-(\widetilde{\kappa}+a)T}}{(\widetilde{\kappa}+a)}}, \tag{8.25}$$

$$\rho_{12} = \rho_{xv}\frac{\sigma^2\rho_{xS}\tau}{\sigma_1\sigma_2(a-\widetilde{\kappa})}\Big[\frac{1-e^{-(a+\widetilde{\kappa})T}}{(a+\widetilde{\kappa})}-\frac{1-e^{-2aT}}{2a}\Big]. \tag{8.26}$$

8.4.2 Closed-form formula for the GAO price

Using the results from the previous paragraph, we can now further evaluate the expression (8.13) for the GAO price in the SZHW model: as the zero-coupon bond price is a monotone function of one state variable, $x(T)$, one can use the Jamshidian (1989) result and write the call option (8.13) on the sum of zero-coupon bonds as a sum of zero-coupon bond call options: let x^* solve

$$\sum_{i=0}^{n} c_i A(T,t_i)e^{-B(T,t_i)x^*} = K, \tag{8.27}$$

and let

$$K_i := A(T,t_i)e^{-B(T,t_i)x^*}. \tag{8.28}$$

Using Jamshidian (1989), we can then write GAO as a sum of zero-coupon bond options, i.e.

$$C(T) = {}_xp_r\, gS(0)\mathbb{E}^{Q^S}\Big[\sum_{i=0}^{n} c_i\Big(A(T,t_i)e^{-B(T,t_i)x(T)}-K_i\Big)^+\Big]. \tag{8.29}$$

As $x(T)$ is normally distributed, we have that $P(T,t_i) = A(T,t_i)e^{-B(T,t_i)x(T)}$ is log-normally distributed. Provided that we know the mean M_i and variance V_i of $\ln P(T,t_i)$ under Q^S, one can directly express the above expectation in terms of the Black and Scholes (1973) formula, i.e.

$$C(T) = {}_xp_r\, gS(0)\sum_{i=0}^{n} c_i\Big[F_iN(d_1^i)-K_iN(d_2^i)\Big], \tag{8.30}$$

$$F_i = e^{M_i+\frac{1}{2}V_i}, \tag{8.31}$$

$$d_1^i = \frac{\ln(F_i/K_i)+\frac{1}{2}V_i}{\sqrt{V_i}}, \tag{8.32}$$

$$d_2^i = d_1^i-\sqrt{V_i}. \tag{8.33}$$

To determine M_i and V_i, recall from (8.22) and (8.23) that $x(T)$ is normally distributed with mean μ_x and variance σ_x. Hence with $P(T,t_i) = A(T,t_i)e^{-B(T,t_i)x(T)}$, one can directly obtain that

the mean M_i and variance V_i of $\ln P(T, t_i)$ are given by

$$M_i = \ln A(T, t_i) - B(T, t_i)\mu_x, \tag{8.34}$$
$$V_i = B^2(T, t_i)\sigma_x^2. \tag{8.35}$$

We have thus derived a closed-form formula for the price of a GAO under stochastic volatility and correlated stochastic interest rates, i.e in the SZHW model of Chapter 2. With this result, we are able to investigate the impact of stochastic volatility on the pricing of GAOs, which will be the subject of Section 8.6.1.

8.5 Extension to two-factor interest rates

In this section, we generalize the setting of the previous section from one to two-factor Gaussian interest rates. That is under the risk-neutral measure Q, we assume the following dynamics for the short interest rate:

$$r(t) = \varphi(t) + x(t) + y(t), \qquad r(0) = r_0, \tag{8.36}$$
$$dx(t) = -ax(t)dt + \sigma dW_x^Q(t), \qquad x(0) = 0, \tag{8.37}$$
$$dy(t) = -by(t)dt + \eta dW_y^Q(t), \qquad y(0) = 0, \tag{8.38}$$
$$dW_x^Q(t)dW_y^Q(t) = \rho_{xy}dt \tag{8.39}$$

Here a, b (mean reversion) and σ, η (volatility) are the positive parameters of the model and $|\rho_{xy}| \leq 1$, and $\varphi(t)$ can be used to exactly fit the current term structure of interest rates, e.g. see Brigo and Mercurio (2006). Much of the analytical structure of the one-factor Gaussian is preserved in this two-factor setting. For example time T zero-coupon bond prices maturity at time t_i are given by

$$P(T, t_i) = A(T, t_i)e^{-B(a,T,t_i)x(T)-B(b,T,t_i)y(T)}, \tag{8.40}$$

where

$$A(T, t_i) = \frac{P^M(0, t_i)}{P^M(0, T)} \exp\left[\frac{1}{2}\Big(V(T, t_i) - V(0, t_i) + V(0, T)\Big)\right] \tag{8.41}$$
$$B(z, T, t_i) = \frac{1 - e^{-z(t_i-T)}}{z} \tag{8.42}$$
$$\begin{aligned} V(T, t_i) = {} & \frac{\sigma^2}{a^2}\left[(t_i - T) + \frac{2}{a}e^{-a(t_i-T)} - \frac{1}{2a}e^{-2a(t_i-T)} - \frac{3}{2a}\right] \\ & + \frac{\eta^2}{b^2}\left[(t_i - T) + \frac{2}{b}e^{-b(t_i-T)} - \frac{1}{2b}e^{-2b(t_i-T)} - \frac{3}{2b}\right] \\ & + 2\rho_{xy}\frac{\sigma\eta}{ab}\left[(t_i - T) + \frac{e^{-a(t_i-T)} - 1}{a} + \frac{e^{-b(t_i-T)} - 1}{b} - \frac{e^{-(a+b)(t_i-T)} - 1}{a + b}\right]. \end{aligned} \tag{8.43}$$

Substituting the zero-coupon bond expression (8.40) into the pricing equation (8.5) and evaluating this expectation, results in the following closed-form expression for the GAO price:

$$C(T) = {}_xp_r\, gS(0) \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}} \Big[F_i(x)N(h_2(x)) - KN(h_1(x))\Big]dx, \tag{8.44}$$

where N denotes the cumulative standard normal distribution function and with

$$h_1(x) := \frac{y^* - \mu_y}{\sigma_y\sqrt{1-\rho_{xy}^2}} - \frac{\rho_{xy}(x-\mu_x)}{\sigma_x\sqrt{1-\rho^2}}, \tag{8.45}$$

$$h_2(x) := h_1(x) + B(b,T,t_i)\sigma_y\sqrt{1-\rho_{xy}^2}, \tag{8.46}$$

$$\lambda_i(x) := c_i A(T,t_i)e^{-B(a,T,t_i)x}, \tag{8.47}$$

$$\kappa_i(x) := -B(b,T,t_i)\Big[\mu_y - \frac{1}{2}\sigma_y^2(1-\rho_{xy}^2)B(b,T,t_i) + \rho_{xy}\sigma_y\frac{(x-\mu_x)}{\sigma_x}\Big], \tag{8.48}$$

$$F_i(x) := \sum_{i=0}^{n} \lambda_i(x)e^{\kappa_i(x)} \tag{8.49}$$

and with y^* the unique solution of

$$\sum_{i=0}^{n} \lambda_i(x)e^{-B(b,T,t_i)y^*} = K. \tag{8.50}$$

The proof of (8.44) is given in Appendix 8.8.1.

In the pricing formula (8.44 it remains to determine the first two moments of $x(T)$ and $y(T)$ and the (terminal) correlation between $x(T)$ and $y(T)$, under the equity price measure Q^S. These are given in Appendix 8.8.2. Note that in the pricing formula (8.44), one is integration a Gaussian probability density function against a bounded function. Because the Gaussian density functions decays very rapidly[2], one can therefore truncate the integration domain in an implementation of (8.44) to a suitable number of standard deviations σ_x around the mean μ_x.

[2]For instance, 99.9999% of the probability mass of a Gaussian density function lies within five standard deviations around its mean.

8.6 Numerical examples

In this section two numerical examples are given. In paragraph 8.6.1 the values of the GAO using the stochastic volatility model described in Section 8.3 are compared with values that result when a geometric Brownian motion is assumed for equity prices. In paragraph 8.6.2 our approach for two-factor interest rate models is compared with the methods described in Chu and Kwok (2007).

8.6.1 Comparison results SZHW model and Black-Scholes Hull-White model

In this section the impact of stochastic volatility of equity prices is shown for an example policy. The results for the SZHW model are compared with a model that combines a Black-Scholes process for equity prices with a one-factor Hull White model for interest rates, the so-called Black-Scholes-Hull-White (BSHW) model given in (8.7) - (8.8). The SZHW and BSHW models are both calibrated to market information (implied volatilities and interest rates) per end July 2007, see Section 8.3.1.

In the example, the policyholder is 55 years old, the retirement age is 65, giving the maturity T of the GAO option of 10 years. Furthermore, $S(0)$ is assumed to be 100. The survival rates are based on the PNMA00 table of the Continuous Mortality Investigation (CMI) for male pensioners.

In Table **2** the prices for the GAO are given for different guaranteed rates g for both models. The results for the SZHW model are obtained using the closed form expression given in (8.30) - (8.35). The pricing formula for the BSHW is a special case of this, and is also derived in Ballotta and Haberman (2003). The results are determined for EU data and U.S. data with an equity-interest rate correlation of respectively 0.347 and 0.146 (see Section 8.2). The table presents the total value of the GAO as well as the time value. While the total value gives the impact on the total prices, the time value gives more insight in the relative impact of the models (since those only have impact the time value). Also, the time value of the GAO is often reported separately, for example within Embedded Value reporting of insurers.

strike g	SZHW	BSHW	Rel. Diff
8.23%	3.82	3.07	+ 24.5%
7%	0.59	0.39	+ 50.7%
8%	2.89	2.26	+28.0%
9%	8.40	7.25	+15.8%
10%	17.02	15.53	+9.6%
11%	27.37	25.69	+6.5%
12%	38.30	36.47	+5.0%
13%	49.35	47.37	+4.2%

Total value, EU.

strike g	SZHW	BSHW	Rel. Diff
8.23%	3.82	3.07	+24.5%
7%	0.59	0.39	+50.7%
8%	2.89	2.26	+28.0%
9%	2.43	1.29	+88.9%
10%	-0.11	-1.60	-93.0%
11%	-0.93	-2.60	-64.4%
12%	-1.17	-2.99	-61.0%
13%	-1.28	-3.26	-60.7%

Time value, EU.

strike g	SZHW	BSHW	Rel. Diff
8.44%	5.43	4.84	+12.0%
7%	1.04	0.88	+18.0%
8%	3.54	3.11	+13.6%
9%	8.53	7.74	+10.3%
10%	16.06	14.90	+7.8%
11%	25.42	23.96	+6.1%
12%	35.73	34.06	+4.9%
13%	46.43	44.58	+4.1%

Total value, U.S.

strike g	SZHW	BSHW	Rel. Diff
8.44%	5.43	4.84	+ 12.0%
7%	1.04	0.88	+18.0%
8%	3.54	3.11	+13.6%
9%	7.27	6.47	+12.3%
10%	4.15	2.99	+39.1%
11%	2.86	1.40	+104.2%
12%	2.53	0.86	+195.5%
13%	2.58	0.74	+250.1%

Time value, U.S.

Table 2: Comparison of GAO total values and time values of the SZHW and BSHW model for different guaranteed rates g. In the examples: at-the-money guaranteed rate g is 8.21% (EU) and 8.44% (U.S.), effective correlation between the stock price and the interest rates is 37.3% (EU) and 25.7% (U.S.).

The table shows that the use of a stochastic volatility model such as the SZHW model has a significant impact on the total value of the GAO. The value increases with 4% - 50% for a EU data and 4% - 17% for a U.S. data, depending on the level of the guarantee.

These price differences are not caused by a volatility effect as both models are calibrated to the same market data in Section 8.3. Figure **1** of Section 8.3, however showed that the distribution of equity prices under a SZHW process has a heavy left tail, but also relatively more mass on the right of the distribution compared to the BSHW process. Given a positive correlation between equity prices and interest rates, and the fact that the GAO pays off when interest rates are low, this means that for the SZHW model there will be some very low payoffs for equity prices in the left tail, but relatively higher payoffs for the remaining scenarios. This is illustrated in Table **3**. For the EU data and $g = 8.23\%$, 50 000 Monte Carlo simulations are generated for both models and the discounted payoffs are segmented in specific intervals.

Payoff	SZHW	BSHW	Diff
0	29173	29251	-78
$(0, 1]$	3727	2621	1106
$(1, 10]$	10983	13134	-2151
$(10, 20]$	3576	3390	186
$(20, 30]$	1347	928	419
$(30, 40]$	582	366	216
$(40, 50]$	257	154	103
$(50, 60]$	130	69	61
$(60, 70]$	92	34	58
$(70, 80]$	44	25	19
$(80, 90]$	26	12	14
$(90, 100]$	16	9	7
$(100, 110]$	14	4	10
> 110	33	3	30

Table 3: Comparison distribution of discounted payoffs for SZHW and BSHW model.

The table shows that indeed:

- SZHW has relatively much payoffs in the interval (0,1) due to the heavy left tail.
- For the remaining intervals, SZHW has more mass to the right, illustrated by the less frequent payoffs in the interval $(1, 10)$ and more frequent payoffs in the intervals greater than 10.

Since the models only have impact on the time value, the relative changes in time value for in-the-money GAOs are higher, which is also illustrated in Table **2**. One might wonder why the time values for the EU data are negative for high levels of g. The reason for this is that due to the positive correlation between interest rates and equity prices, higher equity volatility means that there is a higher chance of lower payoffs, leading to a lower total option value compared to the intrinsic value. For the U.S. data no negative time values are reported. Reason for this is that due to the lower correlation between interest rates and equity prices, the effect described above is less significant than the positive impact of interest rates on the time value.

8.6.2 Comparison results of the two-factor model with Chu and Kwok (2007)

A special case of our modelling framework is considered in Chu and Kwok (2007), namely a equity model with constant volatility. Chu and Kwok (2007) argue that for a two-factor Gaussian interest rate model no analytical pricing formulas exist. Therefore they propose three approximation methods for the valuation of GAOs:

1. *Method of minimum variance duration*: This method approximates the annuity with a single zero-coupon bond and minimizes the approximation error by choosing the maturity of the zero-coupon bond to be equal to the stochastic duration.

2. *Edgeworth expansion*: This method makes use of the Edgeworth approximation of the probability distribution of the value of the annuity (see Collin-Dufresne and Goldstein (2002)).

3. *Affine approximation*: approach: This method approximates the conditional distributions of the risk factors in affine diffusions.

In the chapter the runtimes and approximation errors are compared with benchmark results using Monte Carlo simulations and the method of minimum variance duration comes out most favourably. The other approximation methods do have very long runtime, the Edgeworth expansion method requires even more time then Monte Carlo simulation.

However, as we shown in Section 8.5, it is possible to derive a closed form expression where only a single numerical integration is needed for the case of a two-factor Gaussian interest rate model. It takes hardly any runtime (a couple of hundreds of seconds) to do this numerical integration, whilst it provides exact results. The used parameter setting is the same as in Chu and Kwok (2007) and is given in Appendix 8.8.4. Table **4** shows a comparison of the results for the different methods and a Monte Carlo simulation with 1 000 000 sample paths, whereas Table **6** of Appendix 8.8.4 reports the relative differences of the various methods, compared to the exact GAO prices obtained by the closed form expression in (8.44).

r0	**Strike Level**	**Closed-form Exact**	**Min. Var. Duration**	**Edgeworth Expansion**	**Affine Approx.**	**Monte Carlo (±95% interval)**
0.5%	127%	**11.8000***	**11.8100***	**11.8161***	**11.7913***	11.7921 (±0.0366)
1.0%	123%	**9.7556***	**9.7714***	**9.7502***	**9.7412***	9.7487 (±0.0329)
1.5%	118%	**7.8741***	**7.8958***	**7.8479***	**7.8529***	7.8678 (±0.0294)
2.0%	114%	**6.1690***	6.1946	6.1293	**6.1418***	6.1633 (±0.0260)
2.5%	110%	**4.6612***	4.6860	4.6199	4.6313	4.6555 (±0.0226)
3.0%	106%	**3.3732***	3.3911	3.3408	3.3464	3.3678 (±0.0192)
3.5%	103%	**2.3217***	**2.3273***	2.2999	**2.3044***	2.3174 (±0.0159)
4.0%	99%	**1.5095***	**1.5008***	1.4897	**1.5057***	1.5065 (±0.0126)
4.5%	96%	**0.9214***	0.9008	0.8942	0.9310	0.9198 (±0.0097)
5.0%	93%	**0.5249***	0.4984	0.4922	0.5439	0.5244 (±0.0071)
5.5%	90%	**0.2778***	0.2517	-	-	0.2775 (±0.0050)
6.0%	88%	**0.1360***	0.1150	-	-	0.1354 (±0.0033)
6.5%	85%	**0.0614***	0.0471	-	-	0.0609 (±0.0021)
7.0%	83%	**0.0254***	0.0171	-	-	0.0251 (±0.0013)

Table 4: Comparison between the exact closed-form formula in (8.44), the method of minimum variance duration, the Edgeworth expansion, the affine approximation and a Monte Carlo simulation. Values inside the 95% confidence interval of the Monte Carlo estimates are starred.

The results from Tables **4** and **6** show that the approximation methods considered by Chu and Kwok (2007) break down for higher interest rates, where the guarantee is out-the-money. Note hereby that the first moment of the underlying distribution is main driving factor for option price, while for the price of out-of-the-money options the higher moments play a more important role, e.g. see Brigo and Mercurio (2006). Taking into account that the mean of the underlying annuity is determined exactly in the approximations, this implies that these methods have severe difficulties in estimating the higher moments of the underlying distribution, resulting in poor an approximation quality of the out-of-money GAOs, see Table **4** and **6**.

The exact closed-form formula (8.44) does give highly accurate prices for GAOs across for all strike levels. Differences between the Monte Carlo method and the exact formula are sampling errors as we can see that the 95% confidence interval of the Monte Carlo method is overlapping with the price of the exact closed-form formula for all cases. Typically such Monte Carlo noise increases for out-of-the-money options (e.g. see Glasserman (2003)) as can also be seen from Table **4** for the considered GAOs. Where the Affine approximation method and the Edgeworth expansion method take require a very long runtime (according to Chu and Kwok (2007), the runtime of the Edgeworth expansion is even larger than of the Monte Carlo method with 100 000 sample paths), the runtime the closed-form exact method is comparable to the method of minimum variance duration and takes only a few hundreds of a second. The closed-form exact approach proposed in Section 8.5 is preferable compared to the approaches described in Chu and Kwok (2007), as it gives exact GAO prices over all strike levels whilst being computational very efficient.

8.7 Conclusion

In this chapter closed-form formulas for the pricing of GAOs using a stochastic volatility model for equity prices. The considered framework further allows for 1-factor and 2-factor Gaussian interest rates, whilst taking the correlation between the equity, the stochastic volatility and the stochastic interest rates explicitly into account. The basis for the closed-form formulas for GAOs lies in the fact that under the equity price measure, the GAO can be written in terms of an option on a sum of coupon bearing bonds: after some calculations the Jamshidian (1989) result can be used that expresses the latter option on a sum into a sum of options which can be priced in closed-form. For 1-factor interest rates the price of a GAO can be expressed as sum of Black and Scholes (1973) options, whereas a closed form expression using a single integral can be established for the case of a two-factor Gaussian interest rate model.

A special case of our modelling framework, that is a equity model with constant volatility, is considered in Chu and Kwok (2007). These authors argue that for a two-factor Gaussian interest rate model no analytical pricing formulas exist and propose several approximation methods for the valuation of GAOs. In this chapter we did derive an exact closed-form pricing formula in terms of a single numerical integral, which called for a comparison between these valuation methods. The numerical results show that the use of the exact closed-form exact approach is

preferable compared to the approaches described in Chu and Kwok (2007), as it gives exact GAO prices over all strike levels whilst being computational very efficient to compute.

Because GAOs generally involve long-dated maturities and the annuity payoff is directly linked to the performance of an equity fund, it is for a proper pricing and risk management of such products important to consider realistic returns for the equity fund combined with a non-trivial dependency structure with the underlying interest rates. Using U.S. and the EU market option data, we investigated the effects of a stochastic volatility model for the pricing of GAOs. Time-series analysis between the considered equity funds (S&P500 for U.S. and EuroStoxx50 for EU) and the long-term interest rates revealed a substantial positive correlation. We then calibrated the stochastic and the constant volatility model to the market's options and this correlation, making sure that the implied correlation between the terminal asset price and the interest rates is equal both frameworks for a fair comparison. For both markets, the results indicate that the impact of ignoring a stochastic volatility model is significant; in the considered empirical test cases we found that, the prices for the GAOs using a stochastic volatility model for equity prices are considerably higher in comparison to the constant volatility model, especially for GAOs with out of the money strikes.

8.8 Appendix

8.8.1 Pricing of a coupon bearing option under two-factor interest rates

Let the pair $(x(T), y(T))$ follow a bivariate normal distribution, i.e. with means μ_x, μ_y, variances σ_x^2, σ_y^2 and correlation ρ. The probability density function $f(x, y)$ of $(x(T), y(T))$ is hence given by

$$f(x,y) = \frac{\exp\left\{-\frac{1}{2(1-\rho_{xy}^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho_{xy}\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}}. \tag{8.51}$$

Furthermore, let the time T price of the zero-coupon bond $P(T, t_i)$ maturing at time t_i be given by

$$P(T,t_i) = A(T,t_i)e^{-B(a,T,t_i)x(T)-B(b,T,t_i)y(T)}. \tag{8.52}$$

We then come to the following proposition.

Proposition 8.8.1 *The expected value of the coupon-bearing option maturing at time T, paying coupons c_i at times $i = 0, \ldots, n$ and with strike K is given by a one-dimensional integral, i.e.*

$$\begin{aligned}
&\mathbb{E}\left\{\left(\sum_{i=0}^{n} c_i P(T,t_i) - K\right)^+\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sum_{i=0}^{n} c_i A(T,t_i)e^{-B(a,T,t_i)x(T)-B(b,T,t_i)y(T)} - K\right)^+ f(x,y)dydx \\
&= \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}}\Big[F_i(x)N(h_2(x)) - KN(h_1(x))\Big]dx \\
&=: G(\mu_x,\mu_y,\sigma_x,\sigma_y,\rho_{xy}),
\end{aligned} \tag{8.53}$$

where N denotes the cumulative standard normal distribution function, and

$$\begin{aligned}
h_1(x) &:= \frac{y^* - \mu_y}{\sigma_y\sqrt{1-\rho_{xy}^2}} - \frac{\rho_{xy}(x-\mu_x)}{\sigma_x\sqrt{1-\rho_{xy}^2}}, \\
h_2(x) &:= h_1(x) + B(b,T,t_i)\sigma_y\sqrt{1-\rho_{xy}^2}, \\
F_i(x) &:= \sum_{i=0}^{n}\lambda_i(x)e^{\kappa_i(x)} \\
\lambda_i(x) &:= c_i A(T,t_i)e^{-B(a,T,t_i)x}, \\
\kappa_i(x) &:= -B(b,T,t_i)\Big[\mu_y - \frac{1}{2}\sigma_y^2(1-\rho_{xy}^2)B(b,T,t_i) + \rho_{xy}\sigma_y\frac{(x-\mu_x)}{\sigma_x}\Big],
\end{aligned}$$

and y^* the unique solution of

$$\sum_{i=0}^{n} \lambda_i(x) e^{-B(b,T,t_i)y^*} = K.$$

Proof The result is analogous to the derivation of the swaption price under the G2++ model, we therefore refer to equation (4.31) in Brigo and Mercurio (2006) on pages 158-159 and the corresponding proof on 173-175.

8.8.2 Moments and terminal correlation of the two-factor Gaussian interest rates

To determine the moments of $x(T)$ and $y(T)$ under the equity price measure, we need to consider the dynamics of (8.36), there stated under the risk-neutral measure Q, under the equity price measure Q^S. To change the underlying numeraire (e.g. see Geman et al. (1996)), we calculate the corresponding Radon-Nikodým derivative which is given by

$$\frac{dQ^S}{dQ} = \frac{S(T)B(0)}{S(0)B(T)} = \exp\Big[-\frac{1}{2}\int_0^T \nu^2(u)du + \int_0^T \nu(u)dW_S^Q(u)\Big]. \tag{8.54}$$

The multi-dimensional version of Girsanov's theorem hence implies that

$$\begin{aligned}
dW_S^{Q^S}(t) &\mapsto dW_S^Q(t) - \nu(t)dt, && (8.55)\\
dW_x^{Q^S}(t) &\mapsto dW_x^Q(t) - \rho_{xS}\nu(t)dt, && (8.56)\\
dW_y^{Q^S}(t) &\mapsto dW_y^Q(t) - \rho_{yS}\nu(t)dt, && (8.57)\\
dW_\nu^{Q^S}(t) &\mapsto dW_\nu^Q(t) - \rho_{\nu S}\nu(t)dt, && (8.58)
\end{aligned}$$

are Q^S Brownian motions. Hence under Q^S one has the following model dynamics for the volatility and interest rate process

$$\begin{aligned}
dx(t) &= -ax(t)dt + \rho_{xS}\sigma\nu(t)dt + \sigma dW_x^{Q^S}(t), \qquad x(0) = 0, && (8.59)\\
dy(t) &= -ay(t)dt + \rho_{yS}\eta\nu(t)dt + \eta dW_y^{Q^S}(t), \qquad x(0) = 0, && (8.60)\\
d\nu(t) &= \widetilde{\kappa}\Big(\widetilde{\psi} - \nu(t)\Big)dt + \tau dW_\nu^{Q^S}(t), \qquad \nu(0) = \nu_0, && (8.61)
\end{aligned}$$

where $\widetilde{\kappa} := \kappa - \rho_{\nu S}\tau$, $\widetilde{\psi} := \frac{\kappa\psi}{\widetilde{\kappa}}$. Integrating the latter dynamics (conditional on the current time filtration $\mathcal{F}_0$) yields the following explicit solutions:

$$\nu(T) = \widetilde{\psi} + \Big(\nu(0) - \widetilde{\psi}\Big)e^{-\widetilde{\kappa}T} + \tau \int_0^T e^{-\widetilde{\kappa}(T-u)} dW_\nu^{Q^S}(u), \tag{8.62}$$

$$\begin{aligned} x(T) &= \rho_{xS}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-aT}] + \frac{\nu(0) - \widetilde{\psi}}{a - \widetilde{\kappa}}[e^{-\widetilde{\kappa}T} - e^{-aT}]\Big) \\ &\quad + \frac{\rho_{xS}\sigma\tau}{(a - \widetilde{\kappa})} \int_0^T [e^{-\widetilde{\kappa}(T-u)} - e^{-a(T-u)}] dW_\nu^{Q^S}(u) + \sigma \int_0^T e^{-a(T-u)} dW_x^{Q^S}(u), \end{aligned} \tag{8.63}$$

$$\begin{aligned} y(T) &= \rho_{yS}\sigma\Big(\frac{\widetilde{\psi}}{b}[1 - e^{-bT}] + \frac{\nu(0) - \widetilde{\psi}}{b - \widetilde{\kappa}}[e^{-\widetilde{\kappa}T} - e^{-bT}]\Big) \\ &\quad + \frac{\rho_{yS}\eta\tau}{(b - \widetilde{\kappa})} \int_0^T [e^{-\widetilde{\kappa}(T-u)} - e^{-b(T-u)}] dW_\nu^{Q^S}(u) + \sigma \int_0^T e^{-b(T-u)} dW_x^{Q^S}(u). \end{aligned} \tag{8.64}$$

Using Itô's isometry, one has that the $x(T), y(T)$ (starting from time 0) is normally distributed with means μ_x, μ_y and variance σ_x^2, σ_y^2 and correlation $\rho_{xy}(T)$ given by

$$\mu_x := \rho_{xS}\sigma\Big(\frac{\widetilde{\psi}}{a}[1 - e^{-aT}] + \frac{\nu(0) - \widetilde{\psi}}{(a - \widetilde{\kappa})}[e^{-\widetilde{\kappa}T} - e^{-aT}]\Big), \tag{8.65}$$

$$\mu_y := \rho_{yS}\sigma\Big(\frac{\widetilde{\psi}}{b}[1 - e^{-bT}] + \frac{\nu(0) - \widetilde{\psi}}{(b - \widetilde{\kappa})}[e^{-\widetilde{\kappa}T} - e^{-bT}]\Big), \tag{8.66}$$

$$\sigma_x^2 := \sigma_1^2(\sigma, a) + \sigma_2^2(\sigma, a, \rho_{xS}) + 2\rho_{12}(\sigma, a, \rho_{x\nu}, \rho_{xS})\sigma_1(\sigma, a)\sigma_2(\sigma, a, \rho_{xS}), \tag{8.67}$$

$$\sigma_y^2 := \sigma_1^2(\eta, b) + \sigma_2^2(\eta, b, \rho_{yS}) + 2\rho_{12}(\eta, b, \rho_{y\nu}, \rho_{yS})\sigma_1(\eta, b)\sigma_2(\eta, b, \rho_{yS}), \tag{8.68}$$

$$\rho_{xy} := \frac{\mathrm{Cov}(x(T), y(T))}{\sigma_x\sigma_y}, \tag{8.69}$$

where

$$
\begin{aligned}
\sigma_1(\lambda, z) &:= \lambda \sqrt{\frac{1-e^{-2zT}}{2z}}, \\
\sigma_2(\lambda, z, \rho) &:= \frac{\rho\lambda\tau}{z-\widetilde{\kappa}} \sqrt{\frac{1}{2\widetilde{\kappa}} + \frac{1}{2z} - \frac{2}{(\widetilde{\kappa}+z)} - \frac{e^{-2\widetilde{\kappa}T}}{2\widetilde{\kappa}} - \frac{e^{-2zT}}{2z} + \frac{2e^{-(\widetilde{\kappa}+z)T}}{(\widetilde{\kappa}+z)}}, \\
\rho_{12}(\lambda, z, \rho_1, \rho_2) &:= \rho_1 \frac{\lambda^2\rho_2\tau}{\sigma_1(\lambda, z)\sigma_2(\lambda, z, \rho_2)(z-\widetilde{\kappa})}\left[\frac{e^{-(z+\widetilde{\kappa})T}}{(z+\widetilde{\kappa})} - \frac{e^{-2zT}}{2z}\right], \\
\mathrm{Cov}(x(T), y(T)) &:= \rho_{xy}\sigma\eta\left[\frac{1-e^{-(a+b)T}}{(a+b)}\right] \\
&\quad + \rho_{xv}\sigma\frac{\rho_{yS}\eta\tau}{(b-\widetilde{\kappa})}\left[\frac{1-e^{-(a+\widetilde{\kappa})T}}{(a+\widetilde{\kappa})} - \frac{1-e^{-(a+b)T}}{(a+b)}\right] \\
&\quad + \rho_{yv}\eta\frac{\rho_{xS}\sigma\tau}{(a-\widetilde{\kappa})}\left[\frac{1-e^{-(b+\widetilde{\kappa})T}}{(b+\widetilde{\kappa})} - \frac{1-e^{-(a+b)T}}{(a+b)}\right] \\
&\quad + \frac{\rho_{xS}\sigma\tau}{(a-\widetilde{\kappa})}\frac{\rho_{yS}\eta\tau}{(b-\widetilde{\kappa})}\left[\frac{1-e^{-2\widetilde{\kappa}T}}{2\widetilde{\kappa}} + \frac{1-e^{-(a+b)T}}{(a+b)} - \frac{1-e^{-(a+\widetilde{\kappa})T}}{(a+\widetilde{\kappa})} - \frac{1-e^{-(b+\widetilde{\kappa})T}}{(b+\widetilde{\kappa})}\right].
\end{aligned}
$$

8.8.3 Special case: Pricing formulas with an independent equity price process or pure interest rate guaranteed annuities

If one does not link the guaranteed annuity option to the performance of the equity (e.g. seen in the Netherlands), one has that the guaranteed annuity option price is given by

$$
C(T) = {}_xp_r\, gP(0,T)\mathbb{E}^{Q^T}\left[\left(\sum_{i=0}^{n} c_i P(T, t_i) - K\right)^+\right], \tag{8.70}
$$

where the above expectation is taken with respect to the T-forward measure Q^T, which uses the zero-coupon bond price maturing at time T as numeraire. Moreover, also in case one assumes the equity price is independent from the annuity, e.g. according to Boyle and Hardy (2003) and Pelsser (2003), one ends up with the same expectation as (8.70); one only has to multiply the currency $P(0,T)$ with the expectation future equity price, i.e. in (8.70) one has to replace resulting in replacing $P(0,T)$ by $P(0,T)\mathbb{E}^{Q^T}[S(T)] = S(0)$. In the following sections we will derive closed-form expressions for the guaranteed annuity option price under both one-factor and two-factor Gaussian interest rates.

Hull-White model

Under the T-forward dynamics of the Hull and White (1993) model, see equation (2.18), we have $x(T)$ is normally distributed with mean μ_x and variance σ_x^2 given by

$$\mu_x^T := -\frac{\sigma^2}{a^2}\Big[1 - e^{-aT}\Big] + \frac{\sigma^2}{2a^2}\Big[1 - e^{-2aT}\Big], \tag{8.71}$$

$$\sigma_x^T := \sigma\sqrt{\frac{1 - e^{-2aT}}{2a}}\,. \tag{8.72}$$

Just as in Section 8.4, we have that $x(T)$ is normally distributed, i.e. with the same variance σ_x^2, but with a different mean μ_x^T. Hence completely analogous to Section 8.4, one can use the Jamshidian (1989) result and write the call option on the sum of zero-coupon bonds as a sum of zero-coupon bond call options: let x^* solve

$$\sum_{i=0}^{n} c_i A(T, t_i) e^{-B(T,t_i)x^*} = K, \tag{8.73}$$

and let

$$K_i := A(T, t_i) e^{-B(T,t_i)x^*}. \tag{8.74}$$

Using Jamshidian (1989), we can then write GAO as a sum of zero-coupon bond options, i.e.

$$C(T) = {}_xp_r\, gS(0)\mathbb{E}^{Q^S}\left[\sum_{i=0}^{n} c_i\Big(A(T, t_i)e^{-B(T,t_i)x(T)} - K_i\Big)^+\right]. \tag{8.75}$$

As the bond price again follows a log-normal distribution in the Gaussian model, one can express GAO price in terms of the Black and Scholes (1973) formula, i.e.

$$C(T) = gP(0,T)\sum_{i=0}^{n} c_i\Big[F_i N(d_1^i) - K_i N(d_2^i)\Big], \tag{8.76}$$

$$F_i = e^{M_i + \frac{1}{2}V_i}, \tag{8.77}$$

$$d_1^i = \frac{\ln(F_i/K_i) + \frac{1}{2}V_i}{\sqrt{V_i}}, \tag{8.78}$$

$$d_2^i = d_1^i - \sqrt{V_i}, \tag{8.79}$$

where

$$M_i = \ln A(T, t_i) - B(T, t_i)\mu_x^T, \tag{8.80}$$

$$V_i = B^2(T, t_i)(\sigma_x^T)^2, \tag{8.81}$$

and note that the above expression only deviates from (8.30) in the different means and variances for the $x(T)$ process.

Gaussian Two-factor model

Under Q^T, one has the following expression for the stochastic factors $x(T), y(T)$ that drive the short interest rate (e.g. see Brigo and Mercurio (2006)):

$$x(T) = \mu_x^T + \sigma \int_0^T e^{-a(T-u)} dW_x^{Q^T}(u), \qquad y(T) = \mu_y^T + \sigma \int_0^T e^{-b(T-u)} dW_y^{Q^T}(u), \tag{8.82}$$

hence $x(T), y(T)$ is normally distributed with means μ_x^T, μ_y^T, variances σ_x^2, σ_y^2 and correlation $\rho_{xy}(T)$ given by

$$\mu_x^T := -\left(\frac{\sigma^2}{a^2} + \rho_{xy}\frac{\sigma\eta}{ab}\right)\left[1 - e^{-aT}\right] + \frac{\sigma^2}{2a^2}\left[1 - e^{-2aT}\right] + \frac{\rho_{xy}\sigma\eta}{b(a+b)}\left[1 - e^{-(a+b)T}\right], \tag{8.83}$$

$$\mu_y^T := -\left(\frac{\eta^2}{b^2} + \rho_{xy}\frac{\sigma\eta}{ab}\right)\left[1 - e^{-bT}\right] + \frac{\eta^2}{2b^2}\left[1 - e^{-2bT}\right] + \frac{\rho_{xy}\sigma\eta}{a(a+b)}\left[1 - e^{-(a+b)T}\right], \tag{8.84}$$

$$\sigma_x := \sigma\sqrt{\frac{1 - e^{-2aT}}{2a}}, \tag{8.85}$$

$$\sigma_y := \eta\sqrt{\frac{1 - e^{-2bT}}{2b}}, \tag{8.86}$$

$$\rho_{xy}(T) := \rho_{xy}\frac{\sigma\eta}{\sigma_x\sigma_y}\left[\frac{1 - e^{-(a+b)T}}{(a+b)}\right]. \tag{8.87}$$

Hence analogously to Section 8.5, one has that the GAO price is given by

$$C(T) = gP(0,T)G(\mu_x^T, \mu_y^T, \sigma_x, \sigma_y, \rho_{xy}(T)), \tag{8.88}$$

where G is a closed-form expression, i.e. defined by equation (8.53) of appendix 8.8.1.

8.8.4 Model setup and relative differences in the Chu and Kwok (1999) example

In this appendix we describe the numerical input of the example being used in Chu and Kwok (2007). We also report the relative differences between the GAO price obtained by their methods and the exact closed form expression in (8.44) for that example; note that as the Black-Scholes G2++ model, used in Chu and Kwok (2007), is special case of the Schöbel-Zhu G2++ considered in 8.5, we can one on one translate their parameters into our modelling framework. As the notation is slightly different, we explicitly provide this translation into our modelling framework.

The GAO is specified using the guaranteed rate $g = 9$, the current age of the policy holder $x = 50$ and his retirement age is $r = 65$, with corresponding probability of survival ${}_xp_r = 0.9091$ and time to expiry for the GAO equal to $T = 15$ years. The equity price is modeled by the Black and Scholes (1973) model with parameters:

$$q = 5\%, \quad S(0) = 100 \exp(-q \cdot T) = 47.24, \quad \sigma_S = 10\%, \tag{8.89}$$

where q denotes the continuous dividend rate and σ_S the constant volatility of the equity price. The current (continuous) yield curve is given by (8.91) and for the G2++ interest rate model (e.g. see Brigo and Mercurio (2006)) the following parameters are used:

$$a = 0.77, \quad b = 0.08, \quad \sigma = 2\%, \quad \eta = 1\%, \quad \rho_{xy} = -0.7. \tag{8.90}$$

where the correlations between equity and interest rate drivers given by $\rho_{xS} = 0.5$ and $\rho_{yS} = 0.0071$. Finally, the i-year survival probabilities c_i from policy holder's retirement age 65 are provided in the following table:

c_0	1.0000	c_9	0.8304	c_{18}	0.4889	c_{27}	0.0998
c_1	0.9871	c_{10}	0.8018	c_{19}	0.4414	c_{28}	0.0725
c_2	0.9730	c_{11}	0.7708	c_{20}	0.3934	c_{29}	0.0503
c_3	0.9578	c_{12}	0.7374	c_{21}	0.3454	c_{30}	0.0330
c_4	0.9411	c_{13}	0.7015	c_{22}	0.2981	c_{31}	0.0203
c_5	0.9229	c_{14}	0.6632	c_{23}	0.2523	c_{32}	0.0115
c_6	0.9029	c_{15}	0.6226	c_{24}	0.2088	c_{33}	0.0059
c_7	0.8808	c_{16}	0.5798	c_{25}	0.1684	c_{34}	0.0027
c_8	0.8567	c_{17}	0.5351	c_{26}	0.1319	c_{35}	0.0011

Table 5: i-year survival probabilities c_i from policy holder's retirement age 65. A maximum age of 100 is assumed, that is for all $j > 35$: $c_j = 0$.

In Section 8.6.2 we compared the prices of the exact closed-form solution (8.44) and estimates obtained using 1 000 000 Monte Carlo simulations with the Minimum Variance, the Edgeworth and Affine Approximation method which are used in Chu and Kwok (2007). These results can be found in Table **4**, where a comparison is given for different levels r_0 of the yield curve provided

by the (continuous) yields

$$Y(T) = r_0 + 0.04(1 - e^{-0.2T}). \tag{8.91}$$

To shed more light on the relative performance of these methods compared to the exact closed-form formula, we report in Table **6** the relative differences of these methods to this formula.

r0	Strike Level	Min. Var. Duration	Edgeworth Expansion	Affine Approx.	Monte Carlo Simulation
0.5%	127%	0.1%	0.1%	-0.1%	-0.1%
1.0%	123%	0.2%	-0.1%	-0.1%	-0.1%
1.5%	118%	0.3%	-0.3%	-0.3%	-0.1%
2.0%	114%	0.4%	-0.6%	-0.4%	-0.1%
2.5%	110%	0.5%	-0.9%	-0.6%	-0.1%
3.0%	106%	0.5%	-1.0%	-0.8%	-0.2%
3.5%	103%	0.2%	-0.9%	-0.7%	-0.2%
4.0%	99%	-0.6%	-1.3%	-0.3%	-0.2%
4.5%	96%	-2.2%	-2.9%	1.0%	-0.2%
5.0%	93%	-5.1%	-6.2%	3.6%	-0.1%
5.5%	90%	-9.4%	-	-	-0.1%
6.0%	88%	-15.4%	-	-	-0.4%
6.5%	85%	-23.3%	-	-	-0.7%
7.0%	83%	-32.8%	-	-	-1.1%

Table 6: The relative differences compared to the closed-form formula of the exact GAO price, for different strike levels around the at-the-money point, can be found in the table.

An analysis of the above results is provided in Section 8.6.2.

References

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (9th version)*. Dover (New York), 1964.

P. Acklam. An algorithm for computing the inverse normal cumulative distribution function. `http://home.online.no/~pjacklam/notes/invnorm/index.html`, 2003.

R. Ahlip. Foreign exchange options under stochastic volatility and stochastic interest rates. *International Journal of Theoretical and Applied Finance*, 11(3):277–294, 2008.

R. Ahlip and M. Rutkowski. Forward start options under stochastic volatility and stochastic interest rates. *International Journal of Theoretical and Applied Finance*, 12(2):209–225, 2009.

J.H. Ahrens and U. Dieter. Computer generation of poisson deviates from modified normal distributions. *ACM Transactions on Mathematical Software*, 8(2):163–179, 1982.

H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert. The little Heston trap. `http://perswww.kuleuven.be/%7eu0009713/{Heston}Trap.pdf`, 2005.

Aurélien Alfonsi. High order discretization schemes for the cir process: Application to affine term structure and heston models. *Journal of Mathematcal Computation*, 79:209–237, 2010.

Aurélien Alfonsi. On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

K. Amin and R. Jarrow. Pricing options on risky assets in a stochastic interest rate economy. *Mathematical Finance*, 2:217–237, 1992.

K. Amin and V. Ng. Option valuation with systematic stochastic volatility. *Journal of Finance*, 48:881–910, 1993.

L. Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance*, 11(3), 2008.

L. Andersen and J. Andreasen. Volatile volatilities. *Risk*, 15(2):163–168, 2002.

L. Andersen and R. Brotherton-Ratcliffe. Extended libor market models with stochastic volatility. Working paper, Gen Re Securities, 2001.

L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1), 2007. http://ssrn.com/abstract=685084.

J. Andreasen. Closed form pricing of fx options under stochastic rates and volatility. 2006.

J. Andreasen. A lego approach to hybrid modeling: The baest. Global Derivatives Conference, 2007.

A. Antonov, M. Arneguy, and N. Audet. Markovian projection to a displaced volatility Heston model. http://ssrn.com/abstract=1106223, 2008.

L.J.A. Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.

S. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52:2003–2049, 1997.

S. Bakshi, C. Cao, and Z. Chen. Pricing and hedging long-term options. *Journal of Econometrics*, 94:277–318, 2000.

L. Ballotta and S. Haberman. Valuation of guaranteed annuity conversion options. *Insurance: Mathematics and Economics*, 33:87–108, 2003.

D. Bauer, A. Kling, and J. Russ. A universal pricing framework for guaranteed minimum benefits in variable annuities. 2008. paper presented at AFIR Colloquium in Stockholm.

D.G. Baur. Stock-bond co-movements and cross-country linkages. http://ssrn.com/abstract=978249, 2009.

N. Belgrade, E. Benhamou, E. Koehler, and M. Mohammed. A market model for inflation. http://ssrn.com/abstract=576081, 2004.

E. Biffis and P. Millossovich. The fair value of guaranteed annuity options. *Scandinavian Actuarial Journal*, 1:23–41, 2006.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 1973.

M.J. Bolton, D.H. Carr, P.A. Collis, C.M. George, V.P. Knowles, and A.J. Whitehouse. Reserving for annuity guarantees. The Report of the Annuity Guarantees Working Party, 1997.

P.P. Boyle and M. Hardy. Guaranteed annuity options. *Astin Bulletin*, 33(2):125–152, 2003.

P.P. Boyle and M. Hardy. Mortality derivatives and the option to annuitize. *Insurance: Mathematics and Economics*, 29(3), 2001.

A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7:127–155, 1977.

D. Brigo and F. Mercurio. *Interest rate models - theory and practice*. Springer Finance, 2006.

M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion models. *Operations Research*, 54(2), 2006.

O. Caps. On the valuation of power-reverse duals and equity-rates hybrids. 2007. `www.mathfinance.com/workshop/2007/papers/caps/handouts.pdf`.

P. Carr and D.B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.

D. Chance. *Essays in Derivatives*. Wiley, 1998.

C.C. Chu and Y. K. Kwok. Valuation of guaranteed annuity options in affine term structure models. *International Journal of Theoretical and Applied Finance*, 10(2):363–387, 2007.

P. Collin-Dufresne and R.S. Goldstein. Pricing swaptions within an affine framework. *Journal of Derivatives, Fall issue*, pages 1–18, 2002.

J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

A. R. Didonato and A. H. Morris. Computation of the incomplete gamma function ratios and their inverse. *ACM Trans. Math. Software*, 12:377–393, 1986.

D. Duffie and R. Kan. A yield factor model of interest rates. *Mathematical Finance*, 6(4): 379–406, 1996.

D. Duffie, J. Pan, and K. Singleton. Transform analysis for affine jump diffusions. *Econometrica*, 68:13431376, 2000.

D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13, 2003.

N. Dunbar. Sterling swaptions under new scrutiny. *Risk*, 1999.

A. Feuerverger and A.C.M. Wong. Computation of value-at-risk for nonlinear portfolios. *Journal of Risk*, 3(1), 2000.

J. Fouque, G. Papanicolau, and R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, 2000.

F.N. Fritsch and R.E. Carlson. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysisa*, 17(2):238–246, 1980. `http://en.wikipedia.org/wiki/Monotone_cubic_spline`.

J. Gatheral. *The volatility surface: A practioner's guide*. Wiley Finance, 2005.

H. Geman, E. Karoui, and J.C. Rochet. Changes of numeraire, changes of probability measures and pricing of options. *Journal of Applied Probability*, 32:443–548, 1996.

A. Giese. On the pricing of auto-callable equity securities in the presence of stochastic volatility and stochastic interest rates. presentation at MathFinance workshop, Frankfurt, 2006, `http://www.mathfinance.de/workshop/2006/papers/giese/slides.pdf`, 2004.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer Verlag, 2003.

P. Glasserman and Z. Zhoa. Arbitrage-free discretization of lognormal forward libor and swap rate models. *Finance and Stochastics*, 4(1):35–68, 2000.

L.A. Grzelak and C.W. Oosterlee. On the heston model with stochastic interest rates. `http://ssrn.com/abstract=1382902`, 2009.

L.A. Grzelak, C.W. Oosterlee, and S. v. Weeren. Efficient option pricing with multi-factor equity interest rate hybrid models. `http://ssrn.com/abstract=1434829`, 2009.

J. Guo and M. Hung. A generalization of Rubinstein's "pay now, choose later". *The Journal of Futures Markets*, 28(5):488–515, 2008.

P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woorward. Managing smile risk. *Wilmott Magazine*, pages 84–108, 2002.

M.R. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 2001.

J.M. Harrison and D. Kreps. Martingale and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

J.M. Harrison and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their applications*, 11:215–260, 1981.

D. Heath, R.A. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992.

S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

G. Hong. Forward smile and derivatives pricing, 2004. `http://www-cfr.jbs.cam.ac.uk/archive/PRESENTATIONS/seminars/2004/hong.pdf`.

J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, 1987.

J. Hull and A. White. One factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28(2), 1993.

P.H. Hunt, J. Kennedy, and A. Pelsser. Markov functional interest rate models. *Finance and Stochastics*, 4:391–408, 2000.

C.J. Hunter, P. Jackel, and M.S. Joshi. Getting the drift - drift approximations in a forward-rate-based libor market model. *Risk*, 2001.

P. Jäckel. *Monte Carlo Methods in Finance*. Wiley Finance, 2002.

P. Jäckel and J. Bonneton. *Inflation products and inflation models (Encyclopedia of Quantitative Finance)*. John Wiley and Sons, 2010.

F. Jamshidian. An exact bond option pricing formula. *Journal of Finance*, 44(205-209), 1989.

R. Jarrow and Y. Yildirim. Pricing treasury inflation protected securities and related derivatives using an hjm model. *Journal of Financial and Quantitative Analysis*, 38(2):409–430, 2003.

N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*, volume 2. Wiley, second edition, 1994.

C. Kahl and P. Jäckel. Not-so-complex logarithms in the Heston model. *Wilmott Magazine*, pages 94–103, 2005.

C. Kahl and P. Jäckel. Fast strong approximation monte-carlo schemes for stochastic volatility models. *Quantative Finance*, 6(6), 2006.

D. Kainth and N. Saravanamuttu. Modeling the fx skew. 2007. `www.quarchome.org/FXSkew2.ppt`.

C. Kenyon. Inflation is normal. *Risk*, 2008.

F. Kilin. Accelerating the calibration of stochastic volatility models. `http://mpra.ub.uni-muenchen.de/2975/1/MPRA_paper_2975.pdf`, 2006.

P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations: Applications of Mathematics 23*. Berlin: Springer-Verlag, 1999.

D.E. Knuth. *The Art of Computer Programming, Vol.2: Seminumerical Algorithms*. Addison-Wesley, Reading, Mass., 1981.

S. Kruse. Pricing of inflation-indexed options under the assumption of a lognormal inflation index as well as under stochastic volatility. `http://ssrn.com/abstract=948399`, 2007.

S. Kruse and U. Nögel. On the pricing of forward starting options in Hestons model on stochastic volatility. *Finance and Stochastics*, 9(2):233–250, 2005.

R. Lee. Option pricing by transform methods: extension, unification and error control. *Journal of Computational Finance*, 7:51–86, 2004.

A. Lewis. A simple option formula for general jump-diffusion and other exponential lévy processes. http://ssrn.com/abstract=282110, 2001.

A. Lewis. *Option Valuation under Stochastic Volatility*. Finance Press, 2000.

R. Lord. *Efficient pricing algorithms for exotic derivatives*. PhD thesis, Erasmus Universiteit Rotterdam, 2008.

R. Lord and J. Kahl. Optimal fourier inversion in semi-analytical option pricing. *Journal of Computational Finance*, 10(4), 2008.

R. Lord and J. Kahl. Complex logarithms in Heston-like models. http://ssrn.com/abstract=1105998, 2007.

R. Lord, R. Koekkoek, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance, forthcoming*, 2008.

V. Lucic. Forward-start options in stochastic volatility models. *Wilmott Magazine*, 2003.

J. Maddock, P. A. Bristow, H. Holin, and X. Zhang. Boost math toolkit, 2008. freely available from Boost's C++ library, version 1.35: http://www.boost.org.

G. Marsaglia and W.W. Tsang. A simple method for generating gamma variables. *Transactions on Mathematical Software*, 26(3):363–372, 2000.

C. Marshall, M.R. Hardy, and D. Sanders. Static hedging strategies for guaranteed minimum income benefits. 2009.

M. Kijima and Muromachi. Pricing equity swaps in a stochastic interest rate economy. *Journal of Derivatives*, 2001.

F. Mercurio. Pricing inflation-index derivatives. *Quantitative Finance*, 5(3):289–302, 2005.

F. Mercurio and N. Moreni. A multi-factor sabr model for forward inflation rates. http://ssrn.com/abstract=1337811, 2009.

F. Mercurio and N. Moreni. Inflation with a smile. *Risk*, 2006a.

F. Mercurio and N. Moreni. Pricing inflation-indexed options with stochastic volatility. 2006b. http://www.fabiomercurio.it/stochinf.pdf.

R. C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

K. Miltersen, K. Sandmann, and D. Sondermann. Closed-form solutions for term structure derivatives with lognormal interest rates. *Journal of Finance*, 52:409–430, 1977.

J.P.V. Nunes and T.R.V. Alcaria. Two extensions to forward start options valuation. http://ssrn.com/abstract=1329370, 2009.

A.A.J. Pelsser. *Efficient Methods For Valuing Interest Rate Derivatives*. Springer Finance, 2000.

A.A.J. Pelsser. Pricing and hedging guaranteed annuity options via static option replication. *Insurance, Mathematics and Economics*, 2003.

V. Piterbarg. A multi-currency model with fx volatility skew. `http://ssrn.com/abstract=685084`, 2005.

G. Poitras. The early history of option contracts. `http://www.sfu.ca/~poitras`, 2008.

W.H. Press and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, second edition, 1992.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion (third edition)*. Springer, 1999.

M. Rubinstein. Pay now, choose later. *Risk*, 4(13), 1991.

R. Schöbel and J. Zhu. Stochastic volatility with an Ornstein-Uhlenbeck process: An extension. *European Finance Review*, 4:23–46, 1999.

D. Schrager and A.A.J. Pelsser. Pricing rate of return guarantees in regular premium unit linked insurance. *Insurance: Mathematics and Economics*, 35:369–398, 2004.

L. Scott. Option pricing when the variance changes randomly: theory, estimators and applications. *Journal of Financial and Quantitative Analysis*, 22:419–438, 1997.

J. Sippel and S. Ohkoshi. All power to prdc notes. *Risk*, 15(11), 2002.

R.D. Smith. An almost exact simulation method for the Heston model. *Journal of Computational Finance*, 11(1):115–125, 2008.

E.M. Stein and J.C. Stein. Stock-price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:727–752, 1991.

J.T.S. van Bezooyen, C.J.E. Exley, and S.J.B. Mehta. Valuing and hedging guaranteed annuity options. 1998.

A.P.C. van der Ploeg. *Stochastic Volatility and the Pricing of Financial Derivatives*. PhD thesis, Tinbergen institute/ University of Amsterdam, 2006.

A.P.C. van der Ploeg. Mathematics with industry (swi): The ing problem (hybrid Heston-Hull-White model). `http://www.math.uu.nl/swi2007/problems-ing`, 2007.

A. van Haastrecht. Efficient calibration and simulation of hybrid equity-interest rate models. Master's thesis, Free University Amsterdam,, 2007.

A. van Haastrecht and A.A.J. Pelsser. Efficient, almost exact simulation of the Heston stochastic volatility model. *International Journal of Theoretical and Applied Finance*, 13(1), 2010.

A. van Haastrecht, R. Lord, A.A.J. Pelsser, and D. Schrager. Pricing long-dated insurance contracts with stochastic volatility and stochastic interest rates. *Insurance: Mathematics and Economics*, 2009.

O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

M.J. Wichura. Algorithm as 241: The percentage points of the normal distribution. *Applied Statistics*, 37:477–484, 1998.

D. Wilkie, H. Waters, and S. Yang. Reserving, pricing and hedging for policies with guaranteed annuity options. *British Actuarial Journal*, 2003.

G.A. Willard. Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *The Journal of Derivatives*, 5(1):45–61, 1997.

O. Wyman. Va va voom. `http://www.mmc.com/knowledgecenter/OliverWymanVariableAnnuities.pdf`, 2007.

S. Yang. *Reserving, Pricing and Hedging for Guaranteed Annuity Options*. PhD thesis, Department of Actuarial Mathematics and Statistics, Heriot Watt University, 2001.

J. Zhu. A simple and exact simulation approach to Heston model. `http://ssrn.com/abstract=1153950`, 2008.

J. Zhu. An extended libor market model with nested stochastic volatility dynamics. `http://ssrn.com/abstract=955352`, 2007.

Samenvatting (Dutch Summary)

Waarderen van Langlopende Opties met Stochastische Volatiliteit en Stochastische Rentes

De markten voor langlopende opties hebben het afgelopen decennium een enorme groei doorgemaakt. Tegenwoordig worden veel van dit soort contracten evenals pensioen- en verzekeringsproducten beïnvloed door gezamenlijke bewegingen van aandelenkoersen, rentestanden en inflatieverwachtingen. Een consequentie hiervan is dat de afhankelijkheden tussen de onderliggende risico's een grote impact hebben op het risicomanagement van zulke contracten. De simultane daling van aandelenkoersen en lange rentes, in 2003 en gedurende de kredietcrisis, is hiervan een "perfect" voorbeeld en had als gevolg dat de dekkingsgraden van pensioenfondsen en solvabiliteitsratio's van financiële instellingen daalden tot historisch lage niveaus. Met name verzekeringsbedrijven en pensioenfondsen hebben in grote volumes te maken met dit soort contracten en ze beginnen zich steeds meer bewust te worden van de ingebedde opties en bijbehorende risico's. Geëigende methoden voor de waardering en het risicomanagement van langlopende opties dienen ten minste op een realistische manier met de betrokken marktrisico's om te kunnen gaan.

Hoewel optiecontracten reeds voorkwamen in de Oudheid, ten tijde van de oude Grieken en Babyloniërs, ontstond de handel van financiële producten op professioneel georganiseerde beurzen pas echt vanaf de 16de eeuw. In Antwerpen, Amsterdam en Londen ontstond in die tijd een georganiseerde handel in grondstoffen en verscheidene financiële opties. Op de beurs in Amsterdam werden contracten en opties op tulpenbollen lange tijd verhandeld. In 1637 stegen de prijzen voor tulpenbollen tot ongekende hoogten, waarna ze plotseling volledig instortten. Amsterdam heeft derhalve met de dubieuze eer mogen strijken om de eerste officiële financiële bubbel te noteren. In de boeken is deze vereeuwigd onder de naam "tulpenmania" of "bollengekte". In de Verenigde Staten werd de eerste formele beurs opgericht in 1848, "The Chicago Board of Trade". Deze beurs diende met name om seizoenseffecten in de graanoogst te reduceren. Net na de oogst konden de opslagfaciliteiten van Chicago, waar de handel en distributie van graan plaatsvond, niet op tegen de enorme toename in de aanlevering van graan. Dezelfde ruimten werden echter tijdens andere perioden van het jaar niet of nauwelijks gebruikt. Dit zorgde voor een enorme variatie en onzekerheid in de graanprijs. De introductie van

zogenaamde “futures” contracten, die worden gebruikt om een prijs vast te zetten voor levering op een toekomstig tijdstip, zorgden ervoor dat boeren het graan ook op andere plekken konden opslaan en het op een later tijdstip naar Chicago konden transporteren. Op deze manier waren boeren in staat om het prijsrisico van het graan te elimineren, aangezien het product vrijwel altijd en overal verhandeld kon worden, tegen een gegarandeerde opbrengst in Chicago. Niet alleen leidde dit tot een verbetering van de productiviteit, ook elimineerde dit de prijsonzekerheid voor boeren en graanverwerkers. Het duurde echter tot 1973, het jaar van de oprichting van de “Chicago Board Options Exchange” en de publicatie van de Black, Scholes en Merton artikelen, dat de optiehandel echt volledig op stoom kwam.

Voordat Black, Scholes en Merton met hun doorbraak kwamen, gebruikten investeerders en speculanten heuristische methoden en hun eigen visies op de toekomst om tot prijzen van opties te komen. Beginnend met Bachelier in 1900 zijn verscheidene pogingen ondernomen om optieformules te ontwikkelen. Aan al deze pogingen ontbrak echter het cruciale inzicht van Black, Scholes en Merton dat, onder bepaalde aannamen, het risico van een optie omgezet kan worden in een risicovrij instrument door een techniek genaamd “dynamisch hedgen”. Ervan uitgaande dat er geen arbitragemogelijkheden bestaan, zou de prijs van de optie derhalve gelijk moeten zijn aan de kosten van deze replicerende portefeuille, onafhankelijk van het risicogedrag of de verwachtingen van investeerders. Deze uitvinding in combinatie met een toenemende computerkracht hebben de basis gevormd voor een explosieve groei in het gebruik van derivaten en hebben deze financiële sector tot de triljoenenindustrie gemaakt die zij vandaag de dag vertegenwoordigt.

Met behulp van dit replicatieargument konden ook ingewikkeldere opties geprijsd en gehedged worden. Om een complexe optie te waarderen, worden de prijzen van actief verhandelde contracten, zoals futures en Europese opties, gebruikt om de prijs van de bijbehorende hedge te bepalen. In essentie kan een waarderingsmodel dus geïnterpreteerd worden als een extrapolatie over de prijzen van simpelere instrumenten. Om tot realistische prijzen voor ingewikkeldere producten te komen is het noodzakelijk dat de prijzen van veelverhandelde contracten in het model overeenkomen met hun marktprijs. Het werd echter duidelijk dat dit niet het geval was in het Black-Scholes model en dat een aantal van de aannamen zoals een constante volatiliteit of constante rentes hier niet geschikt voor zijn. Veel onderzoek binnen de financiële wiskunde heeft zich daarom gericht op de ontwikkeling van alternatieve modellen en prijsdynamieken, die de prijzen van de onderliggende risico’s op een betere manier meenemen.

Om een complexe optie te prijzen wordt doorgaans een drietal stappen doorlopen. Eerst wordt een stochastisch financieel model voor de onderliggende prijs gekozen. Vervolgens wordt het model geijkt (gecalibreerd) aan de hand van prijzen voor veelverhandelde contracten. Dit wordt gedaan door de modelprijzen voor deze contracten zo goed mogelijk te laten overeenkomen met hun marktprijs. Met behulp van geschikte numerieke technieken wordt het gecalibreerde model gebruikt om de risico’s en hedgekosten van een specifieke ingewikkeldere optie in te schatten. Gebruikers van deze modellen in de financiële sector eisen snelle en nauwkeurige methoden om prijzen en gevoeligheden voor financiële contracten te bepalen. Omdat financiële

modellen, optiecontracten en risicomanagement methoden steeds geavanceerder worden, dienen efficiënte methoden doorontwikkeld te worden om met deze innovaties mee te kunnen gaan. In dit proefschift, bestaande uit drie delen, staat de waardering en het risicomanagement van langlopende optiecontracten centraal. Het eerste deel richt zich op het inbedden van lange looptijden in nieuwe waarderingsmodellen. Het tweede gedeelte ontwikkelt efficiënte simulatiemethoden die gebruikt worden voor het prijzen van complexe opties. In het derde en laatste deel wordt de waardering en het risicomanagement van twee specifieke, veelvoorkomende verzekeringsopties onderzocht.

Deel I: Stochastische Volatiliteit en Stochastische Rentes

In Deel I van dit proefschrift worden nieuwe methoden voor de waardering en het risicomanagement van contracten met lange looptijden geïntroduceerd. Traditionele aannamen uit het Black-Scholes model blijken veelal niet geschikt te zijn voor het waarderen van complexe opties. In het geval van verzekeringsopties, die gekenmerkt worden door hun lange looptijden, is het bijvoorbeeld nodig om zowel stochastische rentes als dikstaartige rendementen te veronderstellen. Op deze manier ontstaat een realistischer beeld van de onderliggende risico's in deze contracten, kan de prijs nauwkeuriger bepaald worden en kan het bijbehorende risicomanagement beter worden uitgevoerd.

Een overzicht van de relevante literatuur wordt besproken in hoofdstuk 2. Daarnaast wordt er in dit hoofdstuk een korte introductie gegeven van methoden en technieken die veelvuldig in deze dissertatie terugkomen.

Hoofdstuk 3 houdt zich bezig met het waarderen van langlopende contracten onder stochastische rentes en stochastische volatiliteiten. Hierbij wordt bovendien een volledige correlatiestructuur verondersteld om de afhankelijkheid tussen de rentes, aandelen en de volatiliteit op een realistische manier te beschrijven. Dit is belangrijk voor het managen van contracten die afhangen van meerdere onderliggende risico's. Pensioenfondsen bijvoorbeeld, investeren zowel in aandelenfondsen als in vastrentende waarden waardoor gezamenlijke fluctuaties in beide categorieën van grote invloed zijn op de waarde van hun beleggingen en verplichtingen.

In Hoofdstuk 4 wordt een generiek raamwerk ontwikkeld voor de prijsstelling en het risicomanagement van grote productportefeuilles met inflatie, wisselkoers, aandelen, rente, grondstoffen en hybride opties. Speciale aandacht gaat hierbij uit naar de toevoeging van empirische marktverschijnselen zoals stijgende volatiliteitsstructuren in de prijzingsmethoden. We demonstreren dat de modellen op een efficiënte manier gecalibreerd kunnen worden en tonen aan dat het raamwerk rijk genoeg is om aandelen-, wisselkoers- en inflatieprijzen uit de markt op een nauwkeurige manier te omvatten.

Deel II: Efficiënte Simulatie Methoden voor het prijzen van Complexe Opties

Op het moment dat een financieel model is geselecteerd en op een juiste manier gecalibreerd is, is de vervolgstap om het in de praktijk toe te passen. Deel II behandelt het prijzen van complexe opties aan de hand van simulatietechnieken. Ondanks dat bepaalde modellen voor sommige contracten een simpele prijsformule kennen, kan het leeuwendeel van de financiële producten niet met een gesloten formule geprijsd worden. Monte Carlo methoden, gebaseerd om het herhaaldelijk doorlopen van stochastische scenario's, bieden echter een extreem populair en flexibel alternatief om dit soort producten te prijzen. Door technische ontwikkelingen, zoals computers met meerdere processoren en slimme simulatiemethoden, is de verwachting dat deze technieken in toekomst nog wijder toegepast zullen worden. Per definitie zijn Monte Carlo methoden relatief tijdrovend, omdat een groot aantal simulaties nodig is om tot betrouwbare schattingen te komen. Veel aandacht van academici en beoefenaars gaat daarom uit naar efficiënte manieren om de computerinspanningen van de simulaties te minimaliseren, terwijl de nauwkeurigheid behouden blijft.

Hoofdstuk 5 bestudeert efficiënte simulatieschema's voor het Heston stochastische volatiliteit model. Ondanks dat er een exacte simulatiemethode voor het Heston model is ontwikkeld door Broadie en Kaya, is de praktische toepassing hiervan beperkt door complexiteit en een gebrek aan snelheid. Euler discretisaties geven echter totaal andere problemen. Waar het werkelijke variantieproces gegarandeerd niet-negatief is, geldt dit bijvoorbeeld niet voor de Euler discretisatie hiervan. In dit hoofdstuk ontwikkelen we efficiënte benaderingsmethoden die nadelen uit het exacte schema oplossen. Daarnaast voeren we een uitgebreide numerieke vergelijking uit tussen deze nieuwe methoden en andere recent verschenen methoden. We concluderen dat discretisaties gebaseerd op benaderingen van het exacte schema vele malen efficiënter werken dan recente Euler, Kahl-Jäckel en (bijna) exacte discretisaties.

In Hoofdstuk 6 wordt een ander probleem in het simuleren van stochastische volatiliteit modellen besproken. Euler schema's blijken namelijk niet goed in staat te zijn de juiste correlatie te genereren tussen de incrementen van het prijsproces en het onderliggende volatiliteitsproces. Omdat de correlatieparameter een belangrijke rol speelt in stochastische volatiliteit modellen, kan een foute benadering van deze parameter tot significante afwijkingen in optieprijzen leiden. In het Heston model wordt dit zogenaamde "lekkende correlatieprobleem", gedeeltelijk veroorzaakt doordat een Euler discretisatie een wortelproces probeert te benaderen met een normale verdeling. Echter zelfs wanneer de volatiliteit zelf normaal verdeeld is, zoals in het Schöbel-Zhu model, blijkt dit probleem nog steeds een grote rol te spelen. We ontwikkelen nieuwe simulatie schema's voor het Schöbel-Zhu en Schöbel-Zhu-Hull-White model die speciaal ontworpen zijn om de correlatie tussen de prijs en de volatiliteit nauwkeurig te beschrijven. Het op een juiste manier meenemen van deze correlatie blijkt van cruciaal belang te zijn voor de kwaliteit van de simulatiemethoden.

Deel III: Toepassingen op Verzekeringsmarkten

Het derde en laatste gedeelte van dit proefschrift houdt zich bezig met kwantitatieve analyses over het prijzen van twee specifieke verzekeringscontracten. Gebruikmakende van de methoden uit Deel I en II van dit proefschrift, onderzoeken we de impact van stochastische volatiliteit, stochastische rentes en een algemene correlatiestructuur op de waardering van verzekeringscontracten. Door efficiënte formules af te leiden voor gegarandeerde annuïteit opties en contracten die op een toekomstig tijdstip ingaan, zijn we in staat een uitgebreid onderzoek te verrichten naar de waardering en het risicomanagement van ingebedde verzekeringsopties.

Hoofdstuk 7 richt zich op het prijzen van opties die op een toekomstige datum ingaan. Deze voorwaarts ingaande opties behoren tot de klasse van padsafhankelijke Europese opties in de zin dat ze niet alleen afhangen van de eindwaarde van het onderliggende aandeel, maar ook van de waarde op een tussenliggende datum. Dit type opties is ingebed in verschillende verzekeringscontracten, zoals in garanties van Unit-Linked contracten. Ook vormen deze contracten een onderdeel van gestructureerde producten waarbij investeerders op zoek zijn naar opwaarts potentieel, terwijl ze beschermd willen blijven tegen mogelijke koersdalingen. We onderzoeken de impact van stochastische volatiliteit, stochastische rentes en verschillende correlatiestructuren op de prijsstelling en het risicomanagement van zulke contracten. We komen tot de conclusie dat het belangrijk is om al deze aspecten mee te nemen, aangezien het negeren van één van deze aspecten tot serieuze prijs- en hedgefouten kan leiden.

Hoofdstuk 8 onderzoekt de waardering van Gegarandeerde Annuïteit Opties (GAOs), waarbij gebruik wordt gemaakt van een model met stochastische volatiliteit voor aandelenprijzen. Deze GAOs zijn opties die polishouders het recht bieden om op de pensioendatum opgebouwde waarden te converteren naar een levenslange annuïteit tegen een vaste rente. Deze opties waren een gebruikelijke eigenschap van pensioencontracten in het Verenigd Koninkrijk in de jaren ’70 en ’80 toen de rentes hoog waren, maar veroorzaakten problemen voor verzekeraars toen rentes daalden in de jaren ’90. Momenteel worden deze opties veelvuldig verkocht in de Verenigde Staten en Japan, als een onderdeel van contracten met variabele annuïteiten en recentelijk in Nederland in de vorm van gegarandeerde minimale inkooprentes. Tot nu toe werd voor het prijzen van deze opties doorgaans een model met een deterministische volatiliteit voor aandelenprijzen verondersteld. Gezien de lange looptijden van de verzekeringscontracten zou een model met stochastische volatiliteiten voor aandelenprijzen, die realistischere rendementen produceert, echter geschikter zijn. Aan de hand van marktdata van de Verenigde Staten en de Europese Unie, onderzoeken we de effecten van stochastische volatiliteit op het waarderen van GAOs. Uit de resultaten blijkt dat het meenemen van stochastische volatiliteit een grote invloed heeft op de prijsstelling en het risicomanagement van deze producten.