Technische Universität Kaiserslautern
Fachbereich Mathematik

# On Numerical Pricing Methods of Innovative Financial Products

***Kalina Natcheva-Acar***

Vom Fachbereich Mathematik
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr.rer.nat)
genehmigte Dissertation

1. Gutachter: Prof. Dr. Ralf Korn

2. Gutachter: Prof. Dr. Christian Hipp

Vollzug der Promotion: 24.11.2006

*To my family.*

I would like to thank all the people who helped and supported me in the process of writing my thesis. Foremost, I would like to express my deep gratitude to my supervisor Prof. Dr. Ralf Korn not only for the interesting topic, helpful discussions and comments but also for inspiring my enthusiasm to work in the field of financial mathematics. A special thank I also owe my friend Sarp Kaya Acar for believing in me more than I do myself, for his helpful remarks and support and not on the last place for inspiring me with his love in all spheres of my life. I am also very grateful for the help of all my colleagues and for the wonderful time with them in the Fraunhofer Institute. In this place, I want to thank for the very useful discussions with Dr. Martin Krekel, Dr. Gerald Kroisandt, Prof. Dr. Holger Kraft and Dr. habil. Jörg Wenzel. Not on the last place, I want to thank my colleagues and friends Johan de Kock and Evren Baydar who helped me in many various ways. Further, I want to thank Dr. Ingo Fahrner and Dr. Jörg Zipperer from the Bayerische Hypo- und Vereinsbank AG for their useful comments and discussions. A deep heartfelt thank I owe my family - Nely, Petar and Milva - for their dear love and support.

# Contents

# Preface

The fast development of the financial markets in the last decade has lead to the creation of a variety of innovative interest rate related products that require advanced numerical pricing methods. Examples in this respect are products with a complicated strong path-dependence such as a Target Redemption Note, a Ratchet Cap, a Ladder Swap and others. On the other side, the usage of the standard in the literature one-factor Hull and White [34] type of short rate models allows only for a perfect correlation between all continuously compounded spot rates or Libor rates and thus are not suited for pricing innovative products depending on several Libor rates such as for example a "steepener" option.

One possible solution to this problem deliver the two-factor short rate models and in this thesis we consider a two-factor Hull and White [34] type of a short rate process derived from the Heath, Jarrow, Morton [30] framework by limiting the volatility structure of the forward rate process to a deterministic one. For this reason, we start the thesis with an introduction into the Heath, Jarrow, Morton [30] framework, examine the practical problems with its application and prove as a special case the derivation of the short rate dynamics for the limited volatility structure of a Cheyette [21] type .

In this place, we remark that one of the most commonly used numerical method for approximating the price of an option when no closed-form solution is attainable is the lattice method which (with a suitable choice of transition probabilities and transition states) converges weakly to the approximated continuous process (see e.g. Cox, Ross and Rubinstein [26], Hull [33] and many others). In this thesis, we shall often choose to use a variety of modified (binomial, trinomial and quadrinomial) tree constructions as a main numerical pricing tool due to their flexibility and fast convergence and (when there is no closed-form solution) compare their results with fine grid Monte Carlo simulations. We use Monte Carlo simulation as a verification tool due to its almost sure convergence ensured by the Strong Law of Large Numbers (see e.g. Reider [56], Broadie and Glasserman [13],

Glynn and Whitt [29], Korn and Korn [41] and others) but in this place we remark that although very flexible, the fine grid Monte Carlo simulation is not a very practical tool since the computational time it requires (especially in higher dimensions) is often enormously long.

For the purpose of pricing the already mentioned innovative short-rate related products, in Chapter 2 we offer and examine two different lattice construction methods for the two-factor Hull-White [34] type of a short rate process which are able to deal easily both with modeling of the mean-reversion of the underlying process and with the strong path-dependence of the priced options. Additionally, we prove that the so-called rotated lattice construction method overcomes the typical for the existing two-factor tree constructions problem with obtaining negative "risk-neutral probabilities". With a variety of numerical examples, we show that this leads to a stability in the results especially in cases of high volatility parameters and negative correlation between the base factors (which is typically the case in reality).

Further, noticing that Chan *et al* [18] and Ritchken and Sankarasubramanian [58] showed that option prices are sensitive to the level of the short rate volatility, we examine in Chapter 3 and Chapter 4 the pricing of European and American options where the short rate process has a volatility structure of a Cheyette [21] type (introduced in details in Chapter 1). In this relation, we examine the application of the two offered lattice construction methods and compare their results with the Monte Carlo simulation ones for a variety of examples. Additionally, for the pricing of American options with the Monte Carlo method we expand and implement the simulation algorithm of Longstaff and Schwartz [47]. With a variety of numerical examples we compare again the stability and the convergence of the different lattice construction methods.

Dealing with the problems of pricing strongly path-dependent options, we come across the cumulative Parisian barrier option pricing problem. We notice that in their classical form, the cumulative Parisian barrier options have been priced both analytically (in a quasi closed form) and with a tree approximation (based on the Forward Shooting Grid algorithm, see e.g. Hull and White [36], Kwok and Lau [44] and others). However,

in Chapter 5 we offer an additional tree construction method which can be seen as a direct binomial tree integration that uses the analytically calculated conditional survival probabilities. The advantage of the offered method is on one side that the conditional survival probabilities are easier to calculate than the closed-form solution itself and on the other side that this tree construction is very flexible in the sense that it allows easy incorporation of additional features such as e.g a forward starting one. The obtained results are better than the Forward Shooting Grid tree ones and are very close to the analytical quasi closed form solution.

In Chapter 6, we pay our attention to pricing another type of innovative interest rate alike products - namely the Longevity bond - whose coupon payments depend on the survival function of a given cohort. Due to the lack of a market for mortality, for the pricing of the Longevity bonds we develop (following Korn, Natcheva and Zipperer [42]) a framework that contains principles from both Insurance and Financial mathematic. Further on, we calibrate the existing models for the stochastic mortality dynamics to historical German data and additionally offer new stochastic extensions of the classical (deterministic) models of mortality such as the Gompertz and the Makeham one (see e.g. Benjamin and Pollard [5] for an extensive comparison of the existing classical models). Finally, we compare and analyze the results of the application of all considered models to the pricing of a Longevity bond on the longevity of the German males.

# Chapter 1

# On the Practical Implementation of the HJM Framework

In the general theory developed for the needs of pricing interest rate sensitive derivatives there are two main streams.

The first theoretical direction is concentrated on specifying directly the dynamics of the instantaneous short rate process as a Markov one and further (by the means of its time dependent drift term) fit the zero-coupon bond prices deduced from the model to a set of initially observed in the market zero coupon bond prices (usually called "initial term structure"). Main contributors to this theoretical stream are Hull and White [34]. The major advantage of this type of models hides in their Markovian feature which allows analytical and numerical tractability (e.g. recombining binomial/trinomial/quadrinomial trees, easy Monte Carlo simulation etc.). However, their major drawback is that by choosing the drift of the short rate process we are left with **no freedom** in selecting the drift of the forward rate since the latter can be seen as a derived quantity of the short one, see an example in Brigo and Mercurio [12].

The other stream in the literature is based on the Heath, Jarrow, Morton [30] framework which permits an arbitrary term structure of volatility (and covariance) of the forward rates across different maturities. In this framework, by specifying the volatility term structure of the forward rates, it is possible to deduce virtually any existing (or required) short rate model. Another advantage of it is that the current term structure is by construction

an input parameter for it. However, its major drawback is (as it will be explicitly shown in the next section) that in its general form it leads to a short rate process that is not a Markovian one which causes limitation of its practical applicability.

A possible way to overcome the drawbacks of the two approaches is offered by Cheyette [20] and Li, Ritchken and Sankarasubramanian [46] who introduce a reformulation of the Heath, Jarrow, Morton [30] framework (by limiting the class of volatility functions of the forward rates) which overcomes its practical difficulties while keeping the desired generality of its forward rate representation.

In the next two sections, we will first introduce the Heath, Jarrow, Morton [30] framework and explicitely state the core of its direct limited practical application and further introduce and derive its mentioned reformulation offered by Cheyette [20] and Li, Ritchken and Sankarasubramanian [46].

## 1.1 Introduction into the Heath Jarrow Morton Framework

Let us assume we have a complete filtered probability space $\left(\Omega, \{\mathfrak{F}\}_{t\in[0,T]}, P\right)$ with a flow of information given by the natural filtration $\{\mathfrak{F}\}_{t\in[0,T]}$ satistying the usual conditions and with a finite time horizon $[0,T]$. Further, we assume an arbitrage-free market. In addition, let us assume we are in a Heath, Jarrow, Morton [30] framework. This means that the instantaneous forward rate $f(t,T)$ prevailing at the market at time $t$ for a fixed maturity $T$, evolves under the physical measure $P$ according to

$$\begin{aligned} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dW(t) \\ f(0,T) &= f^M(0,T) \end{aligned}$$

where $f^M(0,T)$ denotes the market instantaneous forward interest rate curve at time $t = 0$, $W = (W_1, \ldots, W_N)$ is an $N$-dimensional Brownian motion under measure $P$, $\sigma(t,T) = (\sigma_1(t,T), \ldots, \sigma_N(t,T))$ is a vector of adapted processes and the drift $\alpha(t,T)$ is also an adapted process.

Further, Heath, Jarrow, Morton [30] (from now on abbreviated as "HJM") showed that in an arbitrage free market, the drift process of the instantaneous forward rate under the risk-neutral martingale measure[1] $Q$ cannot be arbitrarily chosen, since it is completely determined by the volatility function of the forward rate process in the following way

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds = \sum_{i=1}^N \sigma_i(t,T)\int_t^T \sigma_i(t,s)ds.$$

The integrated instantaneous forward interest rate is given by

$$f(t,T) = f(0,T) + \sum_{i=1}^N \int_0^t \sigma_i(u,T)\int_u^T \sigma_i(u,s)dsdu + \sum_{i=1}^N \int_0^t \sigma_i(s,T)dW_i^Q(s)$$

where $W_i^Q(t)$ is now a $Q$-Brownian motion and thus the instantaneous short rate process is easily found to be

$$r(t) = f(t,t) = f(0,t) + \sum_{i=1}^N \int_0^t \sigma_i(u,t)\int_u^t \sigma_i(u,s)dsdu + \sum_{i=1}^N \int_0^t \sigma_i(s,t)dW_i^Q(s).$$

Now, notice that as the time $t$ appears both as extremes of the integration and inside of the integration functions, the short rate process is not Markovian. Put in another way, the general HJM model can be seen as a joint Markov process in an **infinite** number of forward rates, i.e. the state space of the dynamics is infinite dimensional even in the case of only one random factor. Notice that even for $i = 1$, by Itô's formula

$$\begin{aligned} dr(t) = &\left[\frac{\partial f(0,t)}{\partial t} + \frac{\partial}{\partial t}\int_0^t \sigma(u,t)\int_u^t \sigma(u,s)dsdu + \int_0^t \frac{\partial}{\partial t}\sigma(u,t)dW^Q(u)\right]dt \\ &+ \sigma(t,t)dW^Q(t) \end{aligned}$$

and the drift depends on the whole history of the Brownian motion and the forward rate volatility. Therefore, approximation of its distribution via trees will involve non-recombining trees which does not allow a high number of time steps.

As the Markov property of the short rate is a desired feature, a way is needed that leads back to it. Carverhill [17] showed that there exist suitable constructions of the volatility

[1] The existence of an equivalent to the physical measure $P$ risk-neutral measure $Q$ is ensured by the assumption of an arbitrage-free market (see e.g. Börk [8], Korn and Korn [41] and others).

of the instantaneous forward rate for which the short rate is a Markov process. Thus, if we write

$$\sigma_i(t,T) = \xi_i(t)\psi_i(T), \quad i = 1, \dots, N$$

for some functions $\xi_i$ and $\psi_i$ that are deterministic and strictly positive, then the short rate process can be calculated to be

$$r(t) = f(0,t) + \sum_{i=1}^{N} \psi_i(t) \int_0^t \xi_i^2(u) \int_0^t \psi_i(s) ds du + \sum_{i=1}^{N} \psi_i(t) \int_0^t \xi_i(s) dW_i^Q(s).$$

In the case $i = 1$, applying Itô's formula yields

$$dr(t) = \left[ \alpha'(t) + \psi'(t) \frac{r(t) - \alpha(t)}{\psi(t)} \right] dt + \xi(t)\psi(t) dW^Q(t)$$

for $\alpha(t) := f(0,t) + \psi(t) \int_0^t \xi^2(u) \int_u^t \psi(s) ds du$. Notice that we are back to the Hull-White setting since we have now a Markovian short rate process with deterministic volatility. However, limiting the class of volatility functions to a deterministic one is too restricting and this leads us to the next section.

## 1.2 Markovian Short Rate Modeling of Cheyette Type

In 1992, Cheyette [20] proved that it is possible to approximate a large class of HJM models with an arbitrage-free Markov model in a **finite** number of state variables up to an arbitrary accuracy by just slightly limiting the class of volatility functions (but still keeping its stochastic nature). Independent from him, in 1995 Li, Ritchken and Sankarasubramanian [46] identified also the necessary and sufficient condition for capturing the path dependence in the short rate process by a single additional condition and further offered and investigated the lattice construction needed for the approximation of the short rate process. However, due to small difference in the time of the two articles and for sake of simplicity, in this thesis we will denote the obtained short rate model a "Cheyette" one and mention explicitly that we work with a lattice construction of a Li, Ritchken and Sankarasubramanian [46]-type.

In this thesis, we limit our attention only to the forward rate dynamics in a one-factor context but remark that Cheyette [20] showed as an example that the US Treasury term structure can be represented with three leading factors $\tilde{f}^{(i)}(t,T)$, $i = 1,2,3$ such that $f(t,T) = f(0,T) + \sum_i \tilde{f}(t,T)$ where the first two factors need five state variables and the third one needs nine state variables.

**Proposition 1.** *If we choose the following form of the volatility of the forward rate*

$$\sigma(t,T) := \sum_{i=1}^{N} \alpha_i(T) \frac{\xi_i(t)}{\alpha_i(t)} \tag{1.1}$$

*for some deterministic functions of time $\alpha_1, \ldots, \alpha_n$ and adapted processes $\xi_1, \ldots, \xi_N$, then although the interest rate itself is not a Markov process, it will be an ingredient in an $N \times (N+3)/2$-state Markov process $\chi(\cdot) = (r(\cdot), \varphi_{ij}(\cdot)_{i,j=1,\ldots,N})$ with*

$$\varphi_{ij}(t) \;=\; \varphi_{ji}(t) = \int_0^t \sigma_i(s,t)\sigma_j(s,t)ds = \int_0^t \frac{\alpha_i(t)\alpha_j(t)}{\alpha_i(s)\alpha_j(s)} \xi_i(s)\xi_j(s)ds, \quad i,j = 1,\ldots,N$$

*where we have denoted $\sigma_i(t,T) = \alpha_i(T)\xi_i(t)/\alpha_i(t)$ for $i = 1,\ldots,N$. Thus, the components of the $\chi$ process form the system*

$$\begin{aligned} dX_i(t) &= \left( X_i(t)\partial_t \log\alpha_i(t) + \sum_{k=1}^{N} \varphi_{ik}(t) \right) dt + \xi_i(t) dW_i^Q(t),\ X_i(0) = 0, \quad i = 1,\ldots N \\ d\varphi_{ij}(t) &= \left[ \xi_i(t)\xi_j(t) + \varphi_{ij}(t)\partial_t \log\left(\alpha_i(t)\alpha_j(t)\right) \right] dt \quad i,j = 1,\ldots,N. \end{aligned}$$

*where*

$$r(t) \;=\; f(0,t) + \sum_{i=1}^{N} X_i(t).$$

**Proof:** Without loss of generality, we prove this for $N = 1$, i.e. for $\sigma(t,T) = \frac{\alpha(T)}{\alpha(t)}\xi(t)$. Recall

$$f(t,T) = f(0,T) + \int_0^t \sigma(u,T) \int_u^T \sigma(u,s)dsdu + \int_0^t \sigma(s,T)dW^Q(s) \tag{1.2}$$

then denoting by $A(t) : \mathbb{R} \to \mathbb{R}$ the primitive of $\alpha(t)$, i.e. $\frac{dA(t)}{dt} = \alpha(t)$ we can write

$$\int_t^T \sigma(t,s)ds = \xi(t)\frac{A(T) - A(t)}{\alpha(t)}$$

and substituting into (1.2) we obtain

$$f(t,T) = f(0,T) + \alpha(T)\left(\int_0^t \xi^2(s)\frac{A(T)-A(s)}{\alpha^2(s)}ds + \int_0^t \frac{\xi(s)}{\alpha(s)}dW^Q(s)\right) \tag{1.3}$$

and thus also

$$r(t) = f(t,t) = f(0,t) + \alpha(t)\left(\int_0^t \xi^2(s)\frac{A(t)-A(s)}{\alpha^2(s)}ds + \int_0^t \frac{\xi(s)}{\alpha(s)}dW^Q(s)\right). \tag{1.4}$$

From (1.4) we estimate

$$\int_0^t \frac{\xi(s)}{\alpha(s)}dW^Q(s) = \frac{r(t)-f(0,t)}{\alpha(t)} - \int_0^t \frac{\xi^2(s)}{\alpha^2(s)}\left[A(t)-A(s)\right]ds \tag{1.5}$$

which we will use later on.

On the other side, denoting

$$\Gamma(t) := f(0,t) + \alpha(t)\int_0^t \xi^2(s)\frac{A(t)-A(s)}{\alpha^2(s)}ds \tag{1.6}$$

and differentiating (1.4) we obtain

$$dr(t) = \Gamma'(t)dt + \alpha'(t)\int_0^t \frac{\xi(s)}{\alpha(s)}dW(s) + \xi(t)dW^Q(t). \tag{1.7}$$

Additionally, since

$$\begin{aligned}
\Gamma'(t) : &= f'(0,t) + \alpha'(t)A(t)\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}ds + \alpha^2(t)\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}ds \\
&\quad + \alpha(t)A(t)\frac{\xi^2(t)}{\alpha^2(t)} - \alpha'(t)\int_0^t \frac{\xi^2(s)A(s)}{\alpha^2(s)}ds - \alpha(t)A(t)\frac{\xi^2(t)}{\alpha^2(t)} \\
&= f'(0,t) + \alpha'(t)A(t)\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}ds + \alpha^2(t)\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}ds - \alpha'(t)\int_0^t \frac{\xi^2(s)A(s)}{\alpha^2(s)}ds
\end{aligned}$$

we obtain together with the expression (1.5)

$$dr(t) = f'(0,t)dt + \alpha^2(t)\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}ds\,dt + \frac{\alpha'(t)}{\alpha(t)}\left[r(t)-f(0,t)\right]dt + \xi(t)dW^Q(t). \tag{1.8}$$

At this moment, we notice that the stochastic process $\xi(t)$ is in fact the volatility of the spot interest rate. Further, in that form of the short rate, since $\xi(t)$ is a stochastic process (and we need its whole history until time $t$ in order to be able to find the short rate), the

spot rate is still not Markovian. However, by adding another variable $\varphi(t)$ to account for the path-dependence such that

$$\varphi(t) := \alpha^2(t) \int_0^t \frac{\xi^2(s)}{\alpha^2(s)} ds$$

we would have that $\chi = (r(t), \varphi(t))$ is a Markov process. Thus, we can rewrite (1.8) as

$$\begin{aligned} dr(t) &= \mu(t, r, \varphi)dt + \xi(t)dW^Q(t), \quad r(0) = r_0 \\ \mu(t, r, \varphi) &:= f'(0, t) + \varphi(t) + \partial_t \log \alpha(t) \left[r(t) - f(0, t)\right]. \end{aligned}$$

Defining $X(t) := r(t) - f(0, t)$ we can write its dynamic as

$$\begin{aligned} dX(t) &= \mu(t, X, \varphi)dt + \xi(t)dW^Q(t), \quad X(0) = 0 \\ \mu(t, X, \varphi) &:= \varphi(t) + \partial_t \log \alpha(t) X(t) \end{aligned}$$

and thus have the choice to model either directly the short rate process $r(t)$ or the process $X(t)$.

Finally, we notice that by its definition, $\varphi(t)$ can be written also as a solution of the following general first order linear differential equation

$$d\varphi(t) = \xi(t)^2 dt + \varphi(t)\partial_t \log \alpha(t), \quad \varphi(0) = 0.$$

■

**Proposition 2.** *Using the same notations, the forward rate $f(t, T)$ can be written as*

$$f(t, T) = f(0, T) + \sum_{j=1}^{N} \frac{\alpha_j(T)}{\alpha_j(t)} \left[ X_j(t) + \sum_{k=1}^{N} \beta_k(t, T)\varphi_{jk}(t) \right]$$

*where*

$$\beta(t, T) := \frac{A(T) - A(t)}{\alpha(t)}.$$

**Proof:** We show this again w.l.o.g. only for $N = 1$.

Recall that substituting the Cheyette's form of the forward rate volatility (1.1) into the HJM forward rate dynamic, we obtain

$$f(t, T) = f(0, T) + \alpha(T) \left( \xi^2(s) \frac{A(T) - A(s)}{\alpha^2(s)} ds - \int_0^t \frac{\xi(s)}{\alpha(s)} dW^Q(s) \right)$$

and plugging in (1.5) yields

$$\begin{aligned} f(t,T) &= f(0,T) + \alpha(T)\left(\frac{X(t)}{\alpha(t)} - \int_0^t \frac{\xi^2(s)}{\alpha^2(s)}[A(t)-A(s)]ds + \int_0^t \frac{\xi^2(s)}{\alpha^2(s)}[A(T)-A(s)]ds\right) \\ &= f(0,T) + \frac{\alpha(T)}{\alpha(t)}X(t) + \frac{\alpha(T)}{\alpha(t)}\frac{[A(T)-A(t)]}{\alpha(t)}\int_0^t \frac{\xi^2(s)}{\alpha^2(s)}\alpha^2(t)ds \\ &= f(0,T) + \frac{\alpha(T)}{\alpha(t)}X(t) + \frac{\alpha(T)}{\alpha(t)}\beta(t,T)\varphi(t) \end{aligned}$$

■

This result shows explicitly that the entire forward rate curve $f(t,T)$ is known at any time $t$ in terms of the state variables $X_i(t)$, $i = 1,\ldots,N$.

**Proposition 3.** *The zero coupon bond price $P(t,T)$ at time $t$ with maturity at time $T$ is calculated in the Cheyette model as*

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left(-\sum_{j=1}^{N}\beta_j(t,T)X_j(t) - \sum_{i,j=1}^{N}\frac{(A_i(T)-A_i(t))\,(A_j(T)-A_j(t))}{2\alpha_i(t)\alpha_j(t)}\varphi_{ij}(t)\right)$$

**Proof:** Due to the no-arbitrage assumption, the price of a zero-coupon bond $P(t,T)$ can be written as an expectation of its discounted payoff under an equivalent martingale measure $Q$ such that

$$P(t,T) = E^Q\left(e^{-\int_t^T r(s)ds}\middle|\mathfrak{F}_t\right) = e^{-\int_t^T f(t,u)du}.$$

For simplicity, we prove the Proposition again only for $N = 1$.

Recalling that

$$f(t,T) = f(0,T) + \frac{\alpha(T)}{\alpha(t)}X(t) + \frac{\alpha(T)}{\alpha(t)}\beta(t,T)\varphi(t) \tag{1.9}$$

we can find

$$\begin{aligned} \int_t^T f(t,u)du &= \int_t^T f(0,u)du + \int_t^T \frac{\alpha(u)}{\alpha(t)}X(t)du + \int_t^T \frac{\alpha(u)}{\alpha(t)}\beta(t,u)\varphi(t)du \\ &= \int_t^T f(0,u)du + \frac{X(t)}{\alpha(t)}\int_t^T \alpha(u)du + \frac{\varphi(t)}{\alpha(t)}\int_t^T \frac{(A(u)-A(t))}{\alpha(t)}\alpha(u)du \\ &= \int_t^T f(0,u)du + X(t)\frac{A(T)-A(t)}{\alpha(t)} + \frac{\varphi(t)}{\alpha(t)^2}\int_t^T [A(u)-A(t)]dA(u) \\ &= \int_t^T f(0,u)du + X(t)\frac{A(T)-A(t)}{\alpha(t)} + \frac{\varphi(t)}{\alpha(t)^2}\frac{[A(T)-A(t)]^2}{2}. \end{aligned}$$

Using again the definition of $\beta(t,T)$ yields the short form

$$\int_t^T f(t,u)du = \int_t^T f(0,u)du + X(t)\beta(t,T) + \frac{1}{2}\varphi(t)\beta(t,T)^2.$$

Substituting $\int_t^T f(t,u)du$ into (1.9) yields the result. ■

Notice that for practical purposes, we can further restrict the volatility functions to those that permit expressions in terms only of the current time and the *constant relative* maturity $(T-t)$ as

$$\sigma(t,T) = \sum_{i=1}^N \xi_i(t)\exp\left(-\int_t^T \kappa_i(u)du\right) = \sum_{i=1}^N \xi_i(t)\exp\left(-\int_0^{T-t} \kappa_i(t+u)du\right)$$

where $\kappa_i(u)$'s are deterministic functions of time. In this case, the dynamics of the components of the $\chi$ process can be written in the slightly simplified form:

$$\begin{aligned} dX_i(t) &= \left(-\kappa_i(t)X_i(t) + \sum_{k=1}^N \varphi_{ik}(t)\right)dt + \xi_i(t)dW_i^Q(t),\ X_i(0)=0, \quad i=1,\dots N \\ d\varphi_{ij}(t) &= \left[\xi_i(t)\xi_j(t) - (\kappa_i(t)+\kappa_j(t))\,\varphi_{ij}(t)\right]dt,\ \varphi_{ij}(0)=0, \quad i,j=1,\dots,N. \end{aligned}$$

Notice that $\kappa_i(\cdot)$, $i=1,\dots,N$ take the role of the mean-reversion rate in the factors of the short rate dynamics under the risk-neutral measure $Q$.

**Proposition 4.** *The forward rate $f(t,T)$ can be written as*

$$f(t,T) = f(0,T) + \sum_{j=1}^N e^{-\int_t^T \kappa_j(x)dx}\left[X_j(t) + \sum_{k=1}^N \beta_k(t,T)\varphi_{jk}(t)\right]$$

*and the zero-coupon bond price $P(t,T)$ as*

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left(-\sum_{j=1}^N \beta_j(t,T)X_j(t) - \frac{1}{2}\sum_{i,j=1}^N \beta_i(t,T)\beta_j(t,T)\varphi_{ij}(t)\right)$$

*where*

$$\beta(t,T) := \int_t^T \exp\left(-\int_t^u \kappa(x)dx\right)du.$$

**Proof:** Follows directly from Proposition 2 and Proposition 3 by substituting with $\alpha_i(t) = e^{-\kappa_i t}$. ■

Notice here that by limiting the class of volatility functions of the instantaneous forward rate, one can derive virtually any known short-rate model. For example, as noticed by Brigo and Mercurio [12], if we restrict the volatility function of the forward rate (in a one-factor HJM framework) to a deterministic one then the model we obtain for the short rate process is equivalent to a Hull-White one (e.g. a Hull-White extension of the Vasicek model). Following that idea, in Chapter 2 we will limit ourselves to the case of $N = 2$ and deterministic functions $\xi_1$ and $\xi_2$ (i.e. deterministic volatility of the forward rate). Then, expanding the ideas of Ritchken and Sankarasubramanian [57] for the construction of a one-dimensional Cheyette-type tree we will offer and examine two different ways of constructing two-dimensional trees for the chosen process.

In addition, since the two-factor short rate model we obtain from the HJM framework (with deterministic volatility of Cheyette type) already contains the initial forward rates as an input we notice that we do not need an extra calibration to the initial term structure of the forward rates as it is done in the trinomial tree construction by Hull and White (see e.g. Hull and White [35], Hull [37] or Hull [33]).

In Chapter 3 and Chapter 4, we will consider the case of a HJM forward rate with volatility of Cheyette type. We will set $N = 1$ and $\xi(t) = \sigma r(t)^\gamma$ for some $\gamma \in \mathbb{R}^+$ and $\sigma \in \mathbb{R}^+$ and use it for pricing options on stocks of an European and American type. For the American options, we will perform a Monte Carlo simulation for which the Longstaff and Schwartz [47] approach will be generalized for the Cheyette type of interest rate process.

# Chapter 2

# Numerical Solutions for the Two-factor Hull-White Model

## 2.1 Motivation

In this Chapter, we are going to deal with a two-factor Hull-White interest rate model for which we shall offer a modified method for its binomial tree construction that manages some of the problems of the standard tree construction methods. In this moment, the reader might ask himself the following two questions:

1. Why do we need to deal with two-factor interest rate models?
2. Why do we need to modify the standard tree construction methods?

The main purpose of this introduction section is to answer the upper questions.

**Disadvantages of the One-Factor Interest Rate Models**

The evolution of the whole yield curve is characterized by the evolution of a single underlying quantity which is the interest rate $r(\cdot)$. As an example, if we consider the zero-coupon bond prices at time $t$ with maturity $T$, then using the relationship

$$P(t,T) = E^Q\left[\exp\left(-\int_t^T r(s)ds\right)\Big|\mathfrak{F}_t\right]$$

and specifying the parameters of the short rate allows us to reconstruct the bond prices for all times $t$ and maturities $T$. Therefore, choosing a poor model for the evolution of

the short rate $r(\cdot)$ will result in a poor model for the whole yield curve.

If we consider now as an example of a one-factor model for the development of the interest rate the Vasicek one[1]:

$$dr_t = \kappa\left(\theta - r_t\right)dt + \sigma dW_t, \quad r(0) = r_0$$

and notice that since it is an affine linear one we can write the zero-coupon bond prices as $P(t,T) = A(t,T)e^{-B(t,T)r_t}$ for some deterministic functions $A(\cdot,\cdot)$ and $B(\cdot,\cdot)$ then, for the continuously compounded spot rate $R(t,T)$ defined such that

$$P(t,T) := \exp\left[-\left(T-t\right)R\left(t,T\right)\right]$$

we obtain

$$R(t,T) := -\frac{\ln P(t,T)}{T-t} = -\frac{\ln A(t,T)}{T-t} + \frac{B(t,T)}{T-t}r_t.$$

Thus, if we want to price a complicated financial product depending on e.g. two continuously compounded spot rates with maturities $T_1$ and $T_2$ we would need to use in the pricing their correlation which unfortunately in a one-factor model is always calculated to be

$$corr\left(R(t,T_1), R(t,T_2)\right) = 1.$$

Since in this model the continuously compounded spot rates are always perfectly correlated for all maturities, a small change in one of them will immediately propagate though the whole yield curve. Notice that this is also approximately valid for the Libor rates in any model with only one factor for the interest rate dynamics. In this respect, the Libor rate $L(t,T)$ is the simply-compounded (in contrast to the continuously compounded) spot interest rate given at time $t$ for maturity $T$, $(t < T)$ and is defined as

$$L(t,T) := \frac{1 - P(t,T)}{(T-t)P(t,T)}$$

which means that the zero-coupon bond price expressed in terms of the Libor rate is given by

$$P(t,T) = \frac{1}{1 + (T-t)L(t,T)}$$

which directly explains the meaning of the simple compounding.

[1]The same analysis is valid for any other one-factor affine-linear short rate model.

**Remark 2.1.1.** *We can give as an example of such a complicated product an option on the so-called "steepener" whose coupons are defined as*

$$n\left[L(t_i,T_2)-L(t_i,T_1)\right]^+, \quad \textit{for } i=1,\dots,N,\ n\in\mathbb{N} \textit{ and } T_1<T_2.$$

*where $L(t,T)$ denotes the Libor rate at time $t$ with maturity $T$.*

*Notice that the core of the steepener hides in the expectation of the investors of steeper development of the yield curve since the steeper it is, the higher the coupons will be. The factor $n$ is used to multiply the effect of the different development of the Libor rates with different maturities.*

If we consider now a two-factor additive Vasicek model (denoted in Brigo and Mercurio [12] by G2++) given as

$$\begin{aligned} r_t &= x_t^1+x_t^2+\varphi(t) \\ dx_t^1 &= -\kappa_1 x_t^1 dt+\sigma dW_t^1, \quad x^1(0)=0 \\ dx_t^2 &= -\kappa_2 x_t^2 dt+\sigma dW_t^2, \quad x^2(0)=0 \end{aligned}$$

with $dW_1dW_2=\rho dt$ for which we can again write the bond prices as

$$P(t,T)=A(t,T)\exp\left(-B^1(t,T)x_t^1-B^2(t,T)x_t^2\right)$$

we obtain the following form of the continuously compounded spot rate

$$R(t,T):=-\frac{\ln P(t,T)}{T-t}=-\frac{\ln A(t,T)}{T-t}+\frac{B^1(t,T)}{T-t}x_t^1+\frac{B^2(t,T)}{T-t}x_t^2.$$

Now notice that

$$corr\left(R(t,T_1),R(t,T_2)\right)=corr\left(\frac{B^1(t,T_1)}{T_1-t}x_t^1+\frac{B^2(t,T_1)}{T_1-t}x_t^2,\frac{B^1(t,T_2)}{T_2-t}x_t^1+\frac{B^2(t,T_2)}{T_2-t}x_t^2\right)\neq 1$$

due to the correlation between the underlying processes i.e. in a two-factor model the yields with different maturities are not modeled anymore as perfectly correlated ones.

In addition, various historical analysis of the whole yield curve based on Principle Components suggest that under the objective measure two or three (but not! only one) principle components usually explain most of the total variation. For example Jamshidian

and Zhu [38] analyze JPY, USD and DEM data and show that only one factor explains $67\% - 76\%$ of the total variation while three principle components account for $93\% - 94\%$. For other examples we refer to Rebonato [55].

All examples suggest that more stochastic factors are needed in order to model the whole yield curve evolution in a sufficiently accurate way. However, one has to make here a compromise between increasing the numerical complexity of the problem and its better modeling. In this respect, we notice that considering even two-factor models poses already a lot of numerical difficulties and therefore, in this work we limit ourselves to a two-factor interest rate modeling.

**Disadvantages of the Classical Binomial Tree Construction Methods in a Two-Factor Setting**

We have already mentioned that there are some problems with the standard numerical methods used for pricing claims contingent on the interest rate in a two-factor setting. For the finite difference methods (and in particular binomial/trinomial tree construction as a form of explicit finite difference methods, see for proofs of the equivalence Heston and Zhou [31]) these are the risk-neutral probabilities that can become negative and the complicated mean-reversion modeling. An extensive overview of the influence of the negative coefficients in two-factor models is given by Zvan, Forsyth and Vetzal [67] for a variety of finite-difference/finite element constructions. And for the difficulties of modeling mean-reversion processes with a trinomial lattice we refer to Hull and White [37].

In this work, we offer a modified two-dimensional binomial lattice which ensures (by rotation of the underlying processes in a similar way as the one by Hull and White [35] and a proper definition of the jump heights, following Cheyette [21]) that the estimated risk-neutral probabilities are well-defined and in addition allows easy implementation of the mean-reversion feature. In addition, we show how the tree can be adapted for the pricing of interest rate claims with a complicated path dependence by the means of a construction similar to the Forward Shooting Grid algorithm (first offered by Hull and White [36]) so that the lattice incorporates both discrete coupon payments and discrete updates (jumps) in the value of the path-dependent variable.

## 2.2 Introduction of the Model

Let us assume that the forward rate follows a one-factor HJM model which can be approximated following Cheyette [21] as a two-state Markov process by limiting the volatility of the forward rate to

$$\sigma(t,T) = \sum_{i=1}^{2} \sigma_i(t) e^{-\int_t^T \kappa_i(s)ds}$$

with $\kappa_i : [0,T] \to \mathbb{R}^+$, $\sigma_i : [0,T] \to \mathbb{R}^+$, $i = 1,2$. The forward rate equation with the so-defined volatility can be written as a function of two state variables $X_1(t)$ and $X_2(t)$ as

$$f(t,T) - f(0,T) = \sum_{i=1}^{2} \exp\left(-\int_t^T \kappa_i(s)ds\right)\left[X_i(t) + \sum_{k=1}^{2} \beta_k(t,T)\varphi_{ik}(0,t)\right] \tag{2.1}$$

for

$$\beta_i(t,T) = \int_t^T \exp\left(-\int_t^u \kappa_i(s)ds\right) du, \quad i = 1,2.$$

The cumulative quadratic covariation of the state processes is defined as

$$\begin{aligned}
\varphi_{11}(u,t) &= \int_u^t \mathrm{cov}(X_1(s), X_1(s))ds = \int_u^t \sigma_1(s)^2 e^{-2\int_s^t \kappa_1(x)dx} ds \\
\varphi_{12}(u,t) &= \int_u^t \mathrm{cov}(X_1(s), X_2(s))ds = \varphi_{21}(u,t) = \int_u^t \rho(s)\sigma_1(s)\sigma_2(s) e^{-\int_s^t (\kappa_1(x)+\kappa_2(x))dx} ds \\
\varphi_{22}(u,t) &= \int_u^t \mathrm{cov}(X_2(s), X_2(s))ds = \int_u^t \sigma_2(s)^2 e^{-2\int_s^t \kappa_2(x)dx} ds.
\end{aligned}$$

Using $r(t) = f(t,t)$ we can find from (2.1) the short rate to be

$$r(t) = f(0,t) + X_1(t) + X_2(t) \tag{2.2}$$

where following Cheyette's model, the state variables $X_1(t)$ and $X_2(t)$ are respectively defined under an equivalent martingale measure $Q$ as

$$dX_1(t) = \left(-\kappa_1(t)X_1(t) + \sum_{k=1}^{2} \varphi_{1k}(0,t)\right) dt + \sigma_1(t) d\tilde{W}_1^Q(t), \quad X_1(0) = 0 \tag{2.3}$$

$$dX_2(t) = \left(-\kappa_2(t)X_2(t) + \sum_{k=1}^{2} \varphi_{2k}(0,t)\right) dt + \sigma_2(t) d\tilde{W}_2^Q(t), \quad X_2(0) = 0 \tag{2.4}$$

with correlation $d\langle \tilde{W}_1^Q, \tilde{W}_2^Q \rangle_t = \rho dt$.

We notice that this construction is very similar to the G2++ (offered by Brigo and Mercurio [12], see Section 2.1) and can be seen as its calibrated to the initial term structure equivalent. Further, Brigo and Mercurio [12] show the equivalence between the G2++ model and the two-factor Hull-White short-rate model. In this work, we choose the upper two-factor model due to its computational tractability and the fact that it fits the initial term structure by its construction.

As next, for sake of simplicity let us set the volatilities and the mean-reversion drifts to be constant. Thus, we can rewrite

$$dX_1(t) = (-\kappa_1 X_1(t) + \varphi_{11}(0,t) + \varphi_{12}(0,t))\, dt + \sigma_1 d\tilde{W}_1^Q(t), \quad X_1(0) = 0 \tag{2.5}$$

$$dX_2(t) = (-\kappa_2 X_2(t) + \varphi_{21}(0,t) + \varphi_{22}(0,t))\, dt + \sigma_2 d\tilde{W}_2^Q(t), \quad X_2(0) = 0 \tag{2.6}$$

with

$$\begin{aligned}
\varphi_{11}(0,t) &= \frac{\sigma_1^2\,(1 - e^{-2\kappa_1 t})}{2\kappa_1} \\
\varphi_{12}(0,t) &= \varphi_{21}(0,t) = \frac{\rho\sigma_1\sigma_2}{\kappa_1 + \kappa_2}\left(1 - e^{-(\kappa_1+\kappa_2)t}\right) \\
\varphi_{22}(0,t) &= \frac{\sigma_2^2\,(1 - e^{-2\kappa_2 t})}{2\kappa_2}.
\end{aligned}$$

Using the form of the analytical strong solution of a generalized linear stochastic differential equation, we can write for $s \leq t$

$$\begin{aligned}
X_i(t) &= e^{-\kappa_i(t-s)}\Big(X_i(s) + \int_s^t \varphi_{i1}(0,u)e^{\kappa_i(u-s)}du + \int_s^t \varphi_{i2}(0,u)e^{\kappa_i(u-s)}du + \int_s^t \sigma_i e^{\kappa_i(u-s)}d\tilde{W}_i^Q(u)\Big) \\
&= X_i(s)e^{-\kappa_i(t-s)} + \int_s^t \varphi_{i1}(0,u)e^{-\kappa_i(t-u)}du + \int_s^t \varphi_{i2}(0,u)e^{-\kappa_i(t-u)}du + \int_s^t \sigma_i e^{-\kappa_i(t-u)}d\tilde{W}_i^Q(u)
\end{aligned}$$

for $i = 1, 2$ and therefore

$$E^Q\big(r(t)|\mathfrak{F}_s\big) = f(0,t) + E^Q(X_1(t)|\mathfrak{F}_s\big) + E^Q(X_2(t)|\mathfrak{F}_s\big)$$

$$
\begin{aligned}
&= f(0,t) + X_1(s)e^{-\kappa_1(t-s)} + X_2(s)e^{-\kappa_2(t-s)} \\
&\quad + \frac{\sigma_1^2}{2\kappa_1^2}\left(1-e^{-\kappa_1(t-s)}\right)^2 + \frac{\sigma_2^2}{2\kappa_2^2}\left(1-e^{-\kappa_2(t-s)}\right)^2 \\
&\quad + \frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2}\left(1-e^{-\kappa_1(t-s)}\right)\left(1-e^{-\kappa_2(t-s)}\right) && (2.7) \\
&\quad + \varphi_{11}(0,s)e^{-\kappa_1(t-s)}\beta_1(s,t) + \varphi_{12}(0,s)e^{-\kappa_1(t-s)}\beta_2(s,t) && (2.8)
\end{aligned}
$$

$$
\mathrm{Var}^Q\big(r(t)|\mathfrak{F}_s\big) = \frac{\sigma_1^2\left(1-e^{-2\kappa_1(t-s)}\right)}{2\kappa_1} + 2\frac{\rho\sigma_1\sigma_2}{\kappa_1+\kappa_2}\left(1-e^{-(\kappa_1+\kappa_2)(t-s)}\right) + \frac{\sigma_2^2\left(1-e^{-2\kappa_2(t-s)}\right)}{2\kappa_2}. \tag{2.9}
$$

where we notice that the short rate process $r(t)$ is conditionally Gaussian.

Denoting

$$
\begin{aligned}
\mu_r := E^Q\big(r(t)|\mathfrak{F}_0\big) &= f(0,t) + \frac{\sigma_1^2}{2\kappa_1^2}\left(1-e^{-\kappa_1 t}\right)^2 + \frac{\sigma_2^2}{2\kappa_2^2}\left(1-e^{-\kappa_2 t}\right)^2 \\
&\quad + \frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2}\left(1-e^{-\kappa_1 t}\right)\left(1-e^{-\kappa_2 t}\right) \\
\sigma_r := \mathrm{Var}^Q\big(r(t)|\mathfrak{F}_0\big) &= \frac{\sigma_1^2\left(1-e^{-2\kappa_1 t}\right)}{2\kappa_1} + 2\frac{\rho\sigma_1\sigma_2}{\kappa_1+\kappa_2}\left(1-e^{-(\kappa_1+\kappa_2)t}\right) + \frac{\sigma_2^2\left(1-e^{-2\kappa_2 t}\right)}{2\kappa_2}
\end{aligned}
$$

and using that the short rate is normally distributed with mean $\mu_r$ and variance $\sigma_r$ we can easily estimate the risk-neutral probability for the short rate to become negative to be

$$
Q(r(t)<0) = E_Q\left(\mathbb{1}_{\{r(t)<0\}}\right) = E_Q\left(\mathbb{1}_{\left\{\frac{r(t)-\mu_r}{\sigma_r} < -\frac{\mu_r}{\sigma_r}\right\}}\right) = \Phi\left(-\frac{\mu_r}{\sigma_r}\right) > 0
$$

where $\Phi$ denotes the cumulative distribution function of the standard normal distribution. We remark here that although the upper probability is in most cases negligibly small, this is a definite drawback of the model. On the other side, considering models of the CIR type for the underlying $X_1$ and $X_2$ processes (which will ensure that the short rate process stays positive) requires (as noticed by Brigo and Mercurio [12]) to force their correlation to be 0 in order to maintain at least the analytical tractability of the bond prices. In this way, a compromise has to be made between the ability of the model to reproduce "humped" term structure of the forward volatility (which requires a correlation between

the base factors different from zero) and the ability of the model to ensure an interest rate dynamics that remains positive.

To explain this in detail, notice that we can write the forward rates dynamic as

$$df(t,T) = (\cdot)dt + \sum_{i=1}^{2} e^{-\kappa_i(T-t)}\sigma_i d\tilde{W}_i^Q(t)$$

and thus estimate

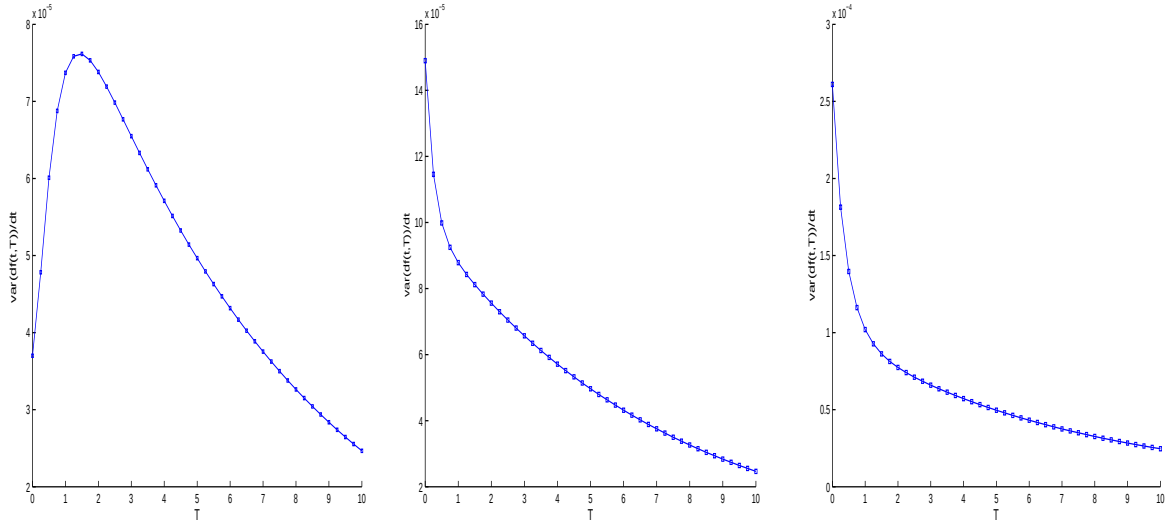
$$\frac{\mathrm{Var}\left(df(t,T)\right)}{dt} = \sum_{i=1}^{2} e^{-2\kappa_i(T-t)}\sigma_i^2 + 2\rho e^{-(\kappa_1+\kappa_2)(T-t)}\sigma_1\sigma_2.$$

Plotting at any time $t$ the volatility curve of the instantaneous forward rate such that $T \mapsto \frac{\mathrm{Var}(df(t,T))}{dt}$ gives us Figure 2.1. Now notice that a **"humped"** shape of the volatility curve can be achieved for some suitable values for $\kappa_1$ and $\kappa_2$ only in the case of $\rho < 0$. Humped shape of the term structure of volatility means that the volatilities are rising at short maturities to a maximum and then decrease for the longer maturities. Being able to reproduce such an effect is a desirable feature of any forward rate model (also possible in the Libor market model, see e.g. Brigo and Mercurio [12]) since the "humped" shape of the volatility curve is an often observed behavior of the market data.

We remark here that Cheyette [21] showed empirically using data from the US Treasury that the long-term behavior of the interest rates is governed by mean reversion parameters that are positive (for all principal components). This means that for $T \to \infty$ we would have $\sigma(t,T) \to 0$ or in specific the volatility of the forward rates with maturities far away in the future approaches to zero. Further, the limit distribution of the short rate process is the normal distribution with mean $\mu_r(\infty)$ and variance $\sigma_r(\infty)$ which we find in the case of positive mean-reversion parameters to be

$$\begin{aligned}
\mu_r \underset{t\to\infty}{\longrightarrow} \mu_r(\infty) &= \lim_{t\to\infty} f(0,t) + \frac{\sigma_1^2}{2\kappa_1^2} + \frac{\sigma_2^2}{2\kappa_2^2} + \frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2} \\
\sigma_r \underset{t\to\infty}{\longrightarrow} \sigma_r(\infty) &= \frac{\sigma_1^2}{2\kappa_1} + \frac{2\rho\sigma_1\sigma_2}{\kappa_1+\kappa_2} + \frac{\sigma_2^2}{2\kappa_2}.
\end{aligned}$$

In contrast to it, negative mean reversion coefficients would imply that the volatility of the forward rates will diverge and allow the short rate to become arbitrarily large.




(a) $\rho = -0.8$. (b) $\rho = 0$. (c) $\rho = 0.8$.

Figure 2.1: Volatility curve of the instantaneous forward rate. $\kappa_1 = 0.07$, $\kappa_2 = 2$, $\sigma_1 = 0.01$, $\sigma_2 = 0.007$.

Naturally, as next comes the question how to price financial instruments in the given framework. One way to do this is to derive the pricing partial differential equation (PDE) and then solve it for the respective boundary conditions. For this purpose, let us denote $V(t, X_1, X_2)$ to be a claim, contingent on the stochastic interest rate movement. By Itô's Lemma, follows that

$$\begin{aligned} dV(t, X_1, X_2) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial X_1}dX_1 + \frac{\partial V}{\partial X_2}dX_2 \\ &\quad + \frac{1}{2}\left(\frac{\partial^2 V}{\partial X_1^2}d\langle X_1\rangle + \frac{\partial^2 V}{\partial X_2^2}d\langle X_2\rangle + 2\frac{\partial^2 V}{\partial X_1 X_2}d\langle X_1, X_2\rangle\right) \\ &= \frac{\partial V}{\partial t}dt + \sum_{i=1}^{2}\left(-\kappa_i X_i + \sum_{j=1}^{2}\varphi_{ij}(0,t)\right)\frac{\partial V}{\partial X_i}dt + \frac{\partial V}{\partial X_1}\sigma_1 d\tilde{W}_1^Q + \frac{\partial V}{\partial X_2}\sigma_1 d\tilde{W}_2^Q \\ &\quad + \frac{1}{2}\sum_{i,j=1}^{2}\sigma_i\sigma_j\frac{\partial^2 V}{\partial X_i \partial X_j}dt \end{aligned} \tag{2.10}$$

Next, since we have an arbitrage-free market, the expected return of our tradable instrument under an equivalent martingale measure $Q$ should equal the riskless interest rate. This is equivalent to requiring that $E^Q(dV(t, X_1, X_2)) = rV(t, X_1, X_2)dt$. Applying the

no-arbitrage requirement to (2.10) yields the pricing PDE:

$$0 = \frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{2}\sigma_i\sigma_j\frac{\partial^2 V}{\partial X_i \partial X_j} + \sum_{i=1}^{2}\left(-\kappa_i X_i + \sum_{j=1}^{2}\varphi_{ij}(0,t)\right)\frac{\partial V}{\partial X_i} - rV \tag{2.11}$$

The pricing PDE can be analytically solved for some simple types of options such as the zero-coupon bond with the boundary condition $V(T, X_1, X_2) = 1$, the European bond option, Caplet, Floorlet or Swaption.

Another way to price a contingent claim in this framework is to follow the so called "Martingale Approach" which consists of calculating the expectation under an equivalent martingale measure $Q$ of its discounted future cash flows (see e.g. Börk [8], Hull [33], Korn and Korn [41] and others). Therefore, if we assume that the claim we want to price has at maturity $T > 0$ a payoff $C(T, X_1, X_2)$ then its price at time $t < T$ can be calculated as

$$V(t, X_1, X_2) = E^Q\left(e^{-\int_t^T r(s)ds}C(T, X_1, X_2)\big|\mathfrak{F}_t\right). \tag{2.12}$$

If we take for an example the zero-coupon bond $P(t,T)$, then to estimate its price instead of solving the pricing PDE, it is much easier to calculate (2.12) and thus obtain

$$\begin{aligned} P(t,T) &= E^Q\left(e^{-\int_t^T r(s)ds}\big|\mathfrak{F}_t\right) = e^{-\int_t^T f(t,u)du} \\ &= \exp\left(-\int_t^T f(0,u)du - \sum_{i=1}^{2}\int_t^T e^{-k_i(u-t)}X_i(t)du\right) \\ &\quad .\exp\left(-\sum_{i=1}^{2}\int_t^T e^{-k_i(u-t)}\sum_{j=1}^{2}\beta_j(t,u)\varphi_{ij}(0,t)du\right) \\ &= \frac{P(0,T)}{P(0,t)}\exp\left(-\beta_1(t,T)X_1(t) - \beta_2(t,T)X_2(t) - \frac{1}{2}\sum_{i,j=1}^{2}\beta_i(t,T)\beta_j(t,T)\varphi_{ij}(0,t)\right) \end{aligned} \tag{2.13}$$

with

$$\beta_1(t,T) := \begin{cases} \frac{1-\exp[-\kappa_1(T-t)]}{\kappa_1} & \text{if } \kappa_1 \neq 0, \\ T-t & \text{else.} \end{cases}$$

$$\beta_2(t,T) := \begin{cases} \frac{1-\exp[-\kappa_2(T-t)]}{\kappa_2} & \text{if } \kappa_2 \neq 0, \\ T-t & \text{else} \end{cases}$$

**Theorem 2.2.1.** *Using the already specified short rate dynamics, the price* $CB(t, T_1, T_2, K)$ *at time* $t$ *of a* ***European bond call option*** *with maturity* $T_1 > t$ *and strike* $K$ *on a zero-coupon bond with maturity* $T_2 > T_1$ *is calculated to*

$$\begin{aligned} CB(t, T_1, T_2, K) &= P(t, T_2)\Phi\left(\frac{ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right) + \frac{1}{2}\Sigma_p(t, T_1, T_2)}{\Sigma_p(t, T_1, T_2)}\right) \\ &\quad -P(t, T_1)K\Phi\left(\frac{ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right) - \frac{1}{2}\Sigma_p(t, T_1, T_2)}{\Sigma_p(t, T_1, T_2)}\right) \end{aligned}$$

*where* $\Phi(\cdot)$ *denotes the cumulative standard normal distribution function and*

$$\Sigma_p^2(t, T_1, T_2) := \beta_1(T_1, T_2)^2\varphi_{11}(t, T_1) + \beta_2(T_1, T_2)^2\varphi_{22}(t, T_1) + 2\beta_1(T_1, T_2)\beta_2(T_1, T_2)\varphi_{12}(t, T_1).$$

*The price* $PB(t, T_1, T_2, K)$ *of a European bond put option is given by*

$$\begin{aligned} PB(t, T_1, T_2, K) &= -P(t, T_2)\Phi\left(\frac{ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right) - \frac{1}{2}\Sigma_p(t, T_1, T_2)}{\Sigma_p(t, T_1, T_2)}\right) \\ &\quad +P(t, T_1)K\Phi\left(\frac{ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right) + \frac{1}{2}\Sigma_p(t, T_1, T_2)}{\Sigma_p(t, T_1, T_2)}\right). \end{aligned}$$

**Proof:** see Appendix A. ■

**Remark 2.2.1.** *Notice that the modeling of the short rate process with the Gaussian* $G2++$ *model of Brigo and Mercurio[12] is equivalent to its modeling with the two factor model we have deduced from Cheyette's [21] construction. However, the definition of the dynamics of the separate factors* $X_1(t)$ *and* $X_2(t)$ *differs and thus the proof of the closed-form solution of a bond option is also different and for this reason, it is given in Appendix A.*

**Theorem 2.2.2.** *Using the already specified short rate dynamics, the price* $C(t, T_1, T_2, K)$ *at time* $t$ *of a* ***caplet*** *resetting at time* $T_1 > t$*, with payoff at time* $T_2 > T_1$ *and strike* $K$

*is calculated to*

$$\begin{aligned} C(t,T_1,T_2,K) &= P(t,T_1)\Phi\left(\frac{ln\left(\frac{P(t,T_1)}{KP(t,T_2)}\right)+\frac{1}{2}\Sigma_p(t,T_1,T_2)}{\Sigma_p(t,T_1,T_2)}\right) \\ &\quad -K^*P(t,T_2)\Phi\left(\frac{ln\left(\frac{P(t,T_1)}{K^*P(t,T_2)}\right)-\frac{1}{2}\Sigma_p(t,T_1,T_2)}{\Sigma_p(t,T_1,T_2)}\right) \end{aligned}$$

*with* $K^* = \frac{1}{1+K(T_2-T_1)}$, $\Phi(\cdot)$ *denoting the cumulative standard normal distribution function and*

$$\Sigma_p^2(t,T_1,T_2) := \beta_1(T_1,T_2)^2\varphi_{11}(t,T_1)+\beta_2(T_1,T_2)^2\varphi_{22}(t,T_1)+2\beta_1(T_1,T_2)\beta_2(T_1,T_2)\varphi_{12}(t,T_1).$$

**Proof:** By the no-arbitrage principle, the price of the Caplet is generally written as

$$C(t,T_1,T_2,K) = E^Q\left(e^{-\int_0^{T_2} r(s)ds}(T_2-T_1)\left[L(T_1,T_2)-K\right]^+\middle|\mathfrak{F}_t\right)$$

and if we change to the $Q^{T_1}$-forward measure (recall its definition in the proof of Theorem 2.2.1) it can equivalently be written as

$$C(t,T_1,T_2,K) = P(t,T_1)E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}(T_2-T_1)\left[L(T_1,T_2)-K\right]^+\middle|\mathfrak{F}_t\right).$$

Next, recalling that the Libor rate $L(t,T)$ is defined as

$$L(T_1,T_2) = \frac{1}{T_2-T_1}\left(\frac{1}{P(T_1,T_2)}-1\right)$$

yields

$$\begin{aligned} C(t,T_1,T_2,K) &= P(t,T_1)E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}\left[(T_2-T_1)L(T_1,T_2)-(T_2-T_1)K\right]^+\middle|\mathfrak{F}_t\right) \\ &= P(t,T_1)\left(1+K(T_2-T_1)\right) \\ &\quad .E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}\left[\frac{1}{1+K(T_2-T_1)}-P(T_1,T_2)\right]^+\frac{1}{P(T_1,T_2)}\middle|\mathfrak{F}_t\right) \\ &= P(t,T_1)\frac{1}{K^*} \\ &\quad .E^{Q^{T_1}}\left(E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}\left[K^*-P(T_1,T_2)\right]^+\frac{1}{P(T_1,T_2)}\middle|\mathfrak{F}_{T_1}\right)\middle|\mathfrak{F}_t\right) \\ &= P(t,T_1)\frac{1}{K^*} \\ &\quad .E^{Q^{T_1}}\left(\left[K^*-P(T_1,T_2]\right)^+\frac{1}{P(T_1,T_2)}E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}\middle|\mathfrak{F}_{T_1}\right)\middle|\mathfrak{F}_t\right) \\ &= P(t,T_1)\frac{1}{K^*}E^{Q^{T_1}}\left(\left[K^*-P(T_1,T_2)\right]^+\middle|\mathfrak{F}_t\right) = \frac{1}{K^*}PB(t,T_1,T_2,K^*) \end{aligned}$$

where we have used that

$$P(T_1, T_2) = E^Q\left(e^{-\int_{T_1}^{T_2} r(s)ds}\middle|\mathfrak{F}_{T_1}\right) = E^{Q^{T_1}}\left(e^{-\int_{T_1}^{T_2} r(s)ds}\middle|\mathfrak{F}_{T_1}\right)$$

since at time $T_1$ the $Q$-measure and the $Q^{T_1}$-measure coincide (i.e. we have that the Radon-Nikodym derivative $\frac{dQ}{dQ^{T_1}} = 1$). ■

Apart from the very few closed-form solutions we have calculated, pricing interest rate contingent claims with complicated path dependence in the given framework cannot be done in an analytic form and therefore, we have to approximate their prices using different numerical methods. One approach (which we do not consider here) can be to use the Alternating Direction Implicit (ADI) finite difference method by Craig and Sneyd [27] to approximate the pricing PDE (2.10).

Another option is to approximate the distribution of the underlying processes using recombining trees. In this respect, we notice that in an earlier work of Krekel and Natcheva [43] (about the numerical pricing of interest rate claims using the Cheyette's model in a one-factor case), the ADI finite difference method has showed comparatively the same results as the tree one but with greater time consumption (due to the bigger grid construction).

Finally, another possibility is a direct estimation of the option value in (2.12) using a fine-grid Monte Carlo simulation. For this method, we remark that apart from its flexibility and almost sure convergence (based on the Strong Law of Large Numbers), it is also the one with the biggest computer time consumption.

Therefore, in this thesis we have chosen to use the following two numerical approaches

- ➪ Tree construction approximating the distribution of the underlying processes
- ➪ Fine grid Monte Carlo simulation for comparison of the results.

As next, we are going to present two different ways for the construction of a two-dimensional lattice for pricing interest rate claims. Theoretical and numerical comparison between the two construction methods will also be delivered.

## 2.3 Tree Construction with Rotation of the Base Processes

### 2.3.1 The Idea of Rotation

One possible way to construct a two-dimensional tree is to rotate the base random variables at each time step in such a way that they are independent. This means that we have to transform the stochastic processes $X_1$ and $X_2$ to space variables $Y_1$ and $Y_2$ which are uncorrelated and linear combinations of the base ones.

A natural question that comes at this point is why we need a more complicated procedure of rotating the basis processes rather than constructing directly a quadrinomial tree. The motivation is that in a quadrinomial tree matching the given processes we come across risk-neutral probabilities that are **not bounded** in $[0,1]^2$. However, in defense of the quadrinomial tree it can be said that for most usual input parameters both trees deliver almost the same results whereas the quadrinomial tree requires slightly less computer memory. But still, in cases of even relatively high volatility parameters, the quadrinomial tree leads to a greater estimation error and exhibits slower convergence to the true price. At the same time, although more complicated to construct, the rotation tree uses properly defined probabilities and is thus robust in the cases of high volatilities. Therefore, both methods have their advantages and disadvantages which we will present and compare, as well as observe their behavior in the pricing of a variety of options.

In this section, we will deal with the rotated tree. For this reason, we need first to make a spectral decomposition of the covariance matrix of $(dX_1(t), dX_2(t))$:

$$\frac{1}{dt}\mathbf{\Sigma} := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \mathbf{U}\Lambda\mathbf{U}^\top = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{pmatrix} \tag{2.14}$$

where we normalize the eigenvectors to length of one and thus

$$\mathbf{U} := \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \mathbf{u_1'} \\ \mathbf{u_2'} \end{pmatrix}$$

is the rotation matrix with rotation angle $\varphi$. Notice that $\mathbf{U}$ is an orthogonal matrix (i.e. $UU^\top = I$ where $I$ denotes the identity matrix) and $\Lambda$ is a matrix with the eigenvalues of

[2] It will be shown later that though not bounded, the risk-neutral probabilities of the quadrinomial tree converge with the refinement of the discretization to properly defined ones.

$\Sigma$ which are found as the solution of $|\frac{1}{dt}\Sigma - \lambda I| = 0$. This is equivalent to solving

$$\left(\sigma_1^2 - \lambda\right)\left(\sigma_2^2 - \lambda\right) = \rho^2\sigma_1^2\sigma_2^2.$$

and thus

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2}(\sigma_1^2 + \sigma_2^2) \pm \frac{1}{2}\sqrt{(\sigma_1^2 + \sigma_2^2)^2 - 4(1-\rho^2)\sigma_1^2\sigma_2^2}, \quad \rho \neq 0 \qquad (2.15) \\ \lambda_1 &= \sigma_1^2, \lambda_2 = \sigma_2^2, \quad \rho = 0. \end{aligned}$$

In addition, if we write $\cos\varphi = \frac{1}{\sqrt{1+\beta^2}}$ and $\sin\varphi = \frac{\beta}{\sqrt{1+\beta^2}}$ as functions of some unknown $\beta$, then substituting $\cos\varphi$ and $\sin\varphi$ in (2.14) and solving it for $\beta$ we obtain $\beta = \pm\frac{-\lambda_1+\sigma_1^2}{\rho\sigma_1\sigma_2}$. It is an equivalent problem whether we take $\beta = \frac{-\lambda_1+\sigma_1^2}{\rho\sigma_1\sigma_2}$ or $\beta = -\frac{-\lambda_1+\sigma_1^2}{\rho\sigma_1\sigma_2}$ (only the direction of the rotation changes), therefore we choose to work with the first one.

**Theorem 2.3.1.** *Defining* $\mathbf{Y} := \mathbf{UX}$ *where* $\mathbf{Y} = (Y_1(t), Y_2(t))^\top$ *and* $\mathbf{X} = (X_1(t), X_2(t))^\top$ *or in specific*

$$Y_1(t) := \mathbf{u_1'X} = \cos\varphi\, X_1(t) - \sin\varphi\, X_2(t) \qquad (2.16)$$

$$Y_2(t) := \mathbf{u_2'X} = \sin\varphi\, X_1(t) + \cos\varphi\, X_2(t) \qquad (2.17)$$

*yields*

$$dY_1(t) = \alpha_1(t, Y_1, Y_2)dt + \sqrt{\lambda_1}dW_1(t)$$

*and*

$$dY_2(t) = \alpha_2(t, Y_1, Y_2)dt + \sqrt{\lambda_2}dW_2(t)$$

*where* $W_1$ *and* $W_2$ *are independent Brownian motions and the new drifts are found to be*

$$\begin{aligned} \alpha_1(t, Y_1(t), Y_2(t)) &:= \cos\varphi\,\mu_1(t, X_1) - \sin\varphi\,\mu_2(t, X_2) \qquad (2.18) \\ &= -(\kappa_1\cos^2\varphi + \kappa_2\sin^2\varphi)Y_1(t) - \sin\varphi\cos\varphi(\kappa_1 - \kappa_2)Y_2(t) \\ &\quad + \cos\varphi(\varphi_{11}(t) + \varphi_{12}(t)) - \sin\varphi(\varphi_{21}(t) + \varphi_{22}(t)) \\ \alpha_2(t, Y_1(t), Y_2(t)) &:= \sin\varphi\,\mu_1(t, X_2) + \cos\varphi\,\mu_2(t, X_2) \qquad (2.19) \\ &= -(\kappa_2\cos^2\varphi + \kappa_1\sin^2\varphi)Y_2(t) - \sin\varphi\cos\varphi(\kappa_1 - \kappa_2)Y_1(t) \\ &\quad + \cos\varphi(\varphi_{21}(t) + \varphi_{22}(t)) + \sin\varphi(\varphi_{11}(t) + \varphi_{12}(t)) \end{aligned}$$

*for* $\mu_1(t, X_1) := -\kappa_1 X_1(t) + \varphi_{11}(t) + \varphi_{12}(t)$ *and* $\mu_2(t, X_2) := -\kappa_2 X_2(t) + \varphi_{21}(t) + \varphi_{22}(t)$.

**Proof:** Applying Itô's Lemma, we obtain directly

$$\begin{aligned} dY_1(t) &= \cos\varphi\, dX_1(t) - \sin\varphi\, dX_2(t) \\ &= \big(\cos\varphi\, \mu_1(t, X_1) - \sin\varphi\, \mu_2(t, X_2)\big)dt + \big(\cos\varphi\, \sigma_1 d\tilde{W}_1(t) - \sin\varphi\, \sigma_2 d\tilde{W}_2(t)\big) \\ &= \alpha_1(t, Y_1, Y_2)dt + \sqrt{\cos\varphi^2\sigma_1^2 + \sin\varphi^2\sigma_2^2 - 2\cos\varphi\sin\varphi\sigma_1\sigma_2\rho}\, dW_1(t) \end{aligned}$$

and

$$\begin{aligned} dY_2(t) &= \sin\varphi\, dX_1(t) + \cos\varphi\, dX_2(t) \\ &= \big(\sin\varphi\, \mu_1(t, X_1) + \cos\varphi\, \mu_2(t, X_2)\big)dt + \big(\sin\varphi\, \sigma_1 d\tilde{W}_1(t) + \cos\varphi\, \sigma_2 d\tilde{W}_2(t)\big) \\ &= \alpha_2(t, Y_1, Y_2)dt + \sqrt{\sin\varphi^1\sigma_1^2 + \cos\varphi^2\sigma_2^2 + 2\cos\varphi\sin\varphi\sigma_1\sigma_2\rho}\, dW_2(t) \end{aligned}$$

for the drifts as defined in (2.18) and (2.19). Notice that to find the drifts we have used that $\mathbf{X} = \mathbf{U}^\top\mathbf{Y}$ as $\mathbf{U}$ is an orthogonal matrix.

Next, we also want to show that the variances of the new processes are well-defined. Estimating $\mathrm{var}(dY_1(t))$ yields

$$\begin{aligned} \mathrm{var}(dY_1(t)) &= e_{11}^2\mathrm{var}(dX_1(t))dt + e_{12}^2\mathrm{var}(dX_2(t))dt + 2e_{11}e_{12}\mathrm{cov}(dX_1(t), dX_2(t)) \\ &= e_{11}^2\sigma_2^2\, dt + e_{12}^2\sigma_2^2\, dt + 2e_{11}e_{12}\sigma_1\sigma_2\rho\, dt \\ &= \cos^2\varphi\sigma_1^2\, dt + \sin^2\varphi\sigma_2^2\, dt - 2\cos\varphi\sin\varphi\sigma_1\sigma_2\rho\, dt \\ &= \frac{(\sigma_1^2 - \lambda_2)\sigma_1^2}{\lambda_1 - \lambda_2}\, dt - \frac{(\lambda_2 - \sigma_2^2)\sigma_2^2}{\lambda_1 - \lambda_2}\, dt + \frac{2(\rho\sigma_1\sigma_2)^2}{\lambda_1 - \lambda_2}\, dt \\ &= \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\, dt + \frac{1}{2}\sqrt{(\sigma_1^2 + \sigma_2^2)\, dt - 4(1-\rho^2)\sigma_1^2\sigma_2^2}\, dt \\ &= \lambda_1\, dt. \end{aligned}$$

By analogy follows that $\mathrm{var}(dY_2(t)) = \lambda_1\, dt$.

Finally, we want to show that $Y_1(t)$ and $Y_2(t)$ are independent which means that $W_1(t)$ and $W_2(t)$ should also be independent. For that purpose, we calculate the covariance

between the new processes

$$
\begin{aligned}
&\operatorname{cov}(dY_1(t), dY_2(t)) = \langle dY_1, dY_2\rangle_t = E(dY_1(t)dY_2(t)) - E(dY_1(t))E(dY_2(t)) \\
&= \alpha_1(t, Y_1, Y_2)dt\,\alpha_2(t, Y_1, Y_2)dt + E\big(\sqrt{\lambda_1}dW_1(t)\sqrt{\lambda_2}dW_2(t)\big) - \alpha_1(t, Y_1, Y_2)dt\alpha_2(t, Y_1, Y_2)dt \\
&= E\big((e_{11}\sigma_1\, d\tilde{W}_1(t) + e_{12}\sigma_2\, d\tilde{W}_2(t))(e_{21}\sigma_1\, d\tilde{W}_1(t) + e_{22}\sigma_2\, d\tilde{W}_2(t))\big) \\
&= E\big(e_{11}e_{21}\sigma_1^2 d\tilde{W}_1(t)^2 + e_{12}e_{21}\sigma_1\sigma_2 d\tilde{W}_1(t)d\tilde{W}_2(t) + e_{11}e_{22}\sigma_1\sigma_2 d\tilde{W}_1(t)d\tilde{W}_2(t) \\
&\quad + e_{12}e_{22}\sigma_2^2(t)d\tilde{W}_2(t)^2\big) \\
&= \cos\varphi\sin\varphi\sigma_1^2 dt - \sin^2\varphi\sigma_1\sigma_2\rho\, dt + \cos^2\varphi\sigma_1\sigma_2\rho\, dt - \cos\varphi\sin\varphi\sigma_2^2\, dt
\end{aligned}
$$

and using the following equalities which one can directly derive from (2.14)

$$
\begin{aligned}
\cos\varphi\sin\varphi &= \frac{\rho\sigma_1\sigma_2}{\lambda_2 - \lambda_1} \\
\cos^2\varphi &= \frac{\sigma_1^2 - \lambda_2}{\lambda_1 - \lambda_2} \\
\sin^2\varphi &= -\frac{\lambda_2 - \sigma_2^2}{\lambda_1 - \lambda_2}
\end{aligned}
$$

we find that

$$\operatorname{cov}(dY_1(t), dY_2(t)) = 0.$$

Now, we have showed that the new factors are uncorrelated and therefore also the Brownian motions $W_1(t)$ and $W_2(t)$ are uncorrelated. Since the Brownian motions are ingredients in a Gaussian two-variate process this also implies their independence. ■

**Remark 2.3.1.** *The covariance matrix of the new factors is exactly the matrix $\Lambda$ of eigenvalues of $\Sigma$ and in addition the whole transformation preserves the variance of the original processes such that*

$$var(dX_1(t)) + var(dX_2(t)) = var(dY_1(t)) + var(dY_2(t))$$

**Theorem 2.3.2.** *The covariance matrix $\Sigma$ between the base processes $X_1$ and $X_2$ is preserved after the rotation.*

**Proof:** Using that $X = U^\top Y$ we estimate

$$\begin{aligned} cov\,(X_1, X_2) &= cov\,(Y_1 \cos\varphi + Y_2 \sin\varphi,\ -Y_1 \sin\varphi + Y_2 \cos\varphi) \\ &= -\cos\varphi \sin\varphi cov\,(Y_1, Y_1) - \sin^2\varphi cov\,(Y_1, Y_2) + \cos^2\varphi cov\,(Y_1, Y_2) \\ &\quad + \cos\varphi \sin\varphi cov\,(Y_2, Y_2) \\ &= -\cos\varphi \sin\varphi \lambda_1 + \cos\varphi \sin\varphi \lambda_2 \\ &= \rho\sigma_1\sigma_2 \end{aligned}$$

and also

$$\begin{aligned} var(X_1) &= var\,(Y_1 \cos\varphi + Y_2 \sin\varphi) \\ &= \cos^2\varphi var(Y_1) + \sin^2\varphi(Y_2) + 2\cos\varphi \sin\varphi cov\,(Y_1, Y_2) \\ &= \cos^2\varphi\lambda_1 + \sin^2\varphi\lambda_2 = \sigma_1^2 \end{aligned}$$

where we have used that due to the independence of the processes $Y_1$ and $Y_2$ the covariance between them is zero. By analogy, $var(X_2) = \sigma_2^2$. ■

### 2.3.2 Tree Construction

Following the same logic as in Li, Ritchken and Sankarasubramanian [46], let us assume that at time t we are at nodes $y_1^a$ and $y_2^a$. In the next time step, the first variable moves either up to $y_1^{a^+}$ or down to $y_1^{a^-}$ and the second variable moves respectively up to $y_2^{a^+}$ or down to $y_2^{a^-}$. Next, in order to decrease the number of the unknown variables we define the up and down jump heights of $Y_1$ as functions of a new variable $J_1$ and the up and down jump heights of $Y_2$ of another new variable $J_2$.

**Definition 2.3.1.** *Denoting* $Z_1 := floor\left[\frac{\alpha_1(t,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}\right]$ *and* $Z_2 := floor\left[\frac{\alpha_2(t,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_2}}\right]$ *we define*

$$J_1(t, y_1^a, y_2^a) := \begin{cases} Z_1 & \text{if } Z_1 \text{ even,} \\ Z_1 + 1 & \text{else.} \end{cases}$$

$$J_2(t, y_1^a, y_2^a) := \begin{cases} Z_2 & \text{if } Z_2 \text{ even,} \\ Z_2 + 1 & \text{else} \end{cases}$$

*and thus define the up and down jumps of the base processes to be*

$$\begin{aligned} y_1^{a^+} &:= y_1^a + \big(J_1(t, y_1^a, y_2^a) + 1\big)\sqrt{\Delta t \lambda_1}, \quad y_1^{a^-} := y_1^a + \big(J_1(t, y_1^a, y_2^a) - 1\big)\sqrt{\Delta t \lambda_1} \\ y_2^{a^+} &:= y_2^a + \big(J_2(t, y_1^a, y_2^a) + 1\big)\sqrt{\Delta t \lambda_2}, \quad y_2^{a^-} := y_2^a + \big(J_2(t, y_1^a, y_2^a) - 1\big)\sqrt{\Delta t \lambda_2} \end{aligned}$$

**Theorem 2.3.3.** *If we denote the probability of an up-jump of the first process by $p_1$ and of the second process by $p_2$ and set*

$$p_1 = \frac{\alpha_1(t, y_1^a, y_2^a)\sqrt{\Delta t} + \big(1 - J_1(t, y_1^a, y_2^a)\big)\sqrt{\lambda_1}}{2\sqrt{\lambda_1}} \tag{2.20}$$

$$p_2 = \frac{\alpha_2(t, y_1^a, y_2^a)\sqrt{\Delta t} + \big(1 - J_2(t, y_1^a, y_2^a)\big)\sqrt{\lambda_2}}{2\sqrt{\lambda_2}} \tag{2.21}$$

*then with this choice of the probabilities and with the jump heights given in Definition 2.3.1, we have that the tree matches locally perfectly the first moments of the base processes and approximately the second moments of the base processes. In addition, the covariance between the basis processes is also perfectly matched.*

**Proof:** Denoting the approximated (with the lattice) local mean and variance respectively with $\widehat{E}(\Delta y_i)$ and $\widehat{\mathrm{var}}(\Delta y_i)$, for $i = 1, 2$ and the true local mean and variance of the underlying process by $E(\Delta y_i)$ and $\mathrm{var}(\Delta y_i)$, for $i = 1, 2$ we notice by the definition of the risk-neutral tree probabilities that

$$\begin{aligned} \widehat{E}(\Delta y_1) = p_1\left(y_1^{a^+} - y_1^a\right) + (1 - p_1)\left(y_i^{a^-} - y_1^a\right) &= \alpha_1\left(t, y_1^a, y_2^a\right)\Delta t = E(\Delta y_1) \\ \widehat{E}(\Delta y_2) = p_2\left(y_2^{a^+} - y_2^a\right) + (1 - p_2)\left(y_2^{a^-} - y_2^a\right) &= \alpha_2\left(t, y_1^a, y_2^a\right)\Delta t = E(\Delta y_2) \end{aligned}$$

which implies that the first moments are locally perfectly matched.

For the second moments, let us consider only the first process as the second one is analog to it. For $J_1$ given in Definition 2.3.1, we have that

$$\begin{aligned} \widehat{\mathrm{var}}\left(\Delta y_1\right) &= \widehat{E}\left(\Delta y_1^2\right) - \widehat{E}\left(\Delta y_1\right)^2 \\ &= p_1(J_1(t, y_1^a, y_2^a) + 1)^2\Delta t\lambda_1 + (1 - p_1)(J_1(t, y_1^a, y_2^a) - 1)^2\Delta t\lambda_1 - \alpha_1(t, y_1, y_2)^2\Delta t^2 \\ &= 4p_1 J_1(t, y_1^a, y_2^a)\Delta t\lambda_1 + (J_1(t, y_1^a, y_2^a) - 1)^2\Delta t\lambda_1 - \alpha_1(t, y_1, y_2)^2\Delta t^2 \\ &= -J_1(t, y_1^a, y_2^a)^2\Delta t\lambda_1 + 2\alpha_1(t, y_1, y_2)\Delta t\sqrt{\Delta t\lambda_1}J_1(t, y_1^a, y_2^a) - \alpha_1(t, y_1, y_2)^2\Delta t^2 \\ &\quad + \lambda_1\Delta t^2 \\ &= -(J_1(t, y_1^a, y_2^a)\sqrt{\Delta t\lambda_1} - \alpha_1(t, y_1, y_2)\Delta t)^2 + \lambda_1\Delta t \end{aligned}$$

and since $\operatorname{var}(\Delta y_1) = \lambda_1 \Delta t$ there will be an equality only if

$$J_1(t, y_1^a, y_2^a) = \frac{\alpha_1(t, y_1^a, y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}$$

which is approximately the case by the definition of $J_1(t, y_1^a, y_2^a)$.

Thus, we have obtained that $\widehat{\operatorname{var}}(\Delta y_1) \approx \operatorname{var}(\Delta y_1)$ which means that locally the variance of the original process is approximately matched.

Finally, we want to show that with the so defined jump-heights and risk-neutral probabilities the local approximated covariance between the new base processes $Y_1$ and $Y_2$ is still zero. For that reason, note that

$$\begin{aligned}
\widehat{E}(\Delta y_1 \Delta y_2) &= p_1 p_2 \left(y_1^{a^+} - y_1^a\right)\left(y_2^{a^+} - y_2^a\right) + p_1(1-p_2)\left(y_1^{a^+} - y_1^a\right)\left(y_2^{a^-} - y_2^a\right) \\
&\quad + (1-p_1)p_2\left(y_1^{a^-} - y_1^a\right)\left(y_2^{a^+} - y_2^a\right) + (1-p_1)(1-p_2)\left(y_1^{a^-} - y_1^a\right)\left(y_2^{a^-} - y_2^a\right) \\
&= p_1\left(y_1^{a^+} - y_1^a\right)\left[p_2\left(y_2^{a^+} - y_2^a\right) + (1-p_2)\left(y_2^{a^-} - y_2^a\right)\right] \\
&\quad + (1-p_1)\left(y_1^{a^-} - y_1^a\right)\left[p_2\left(y_2^{a^+} - y_2^a\right) + (1-p_2)\left(y_2^{a^-} - y_2^a\right)\right] \\
&= \alpha_2(t, y_1^a, y_2^a)\Delta t\left[p_1\left(y_1^{a^+} - y_1^a\right) + (1-p_1)\left(y_1^{a^-} - y_1^a\right)\right] \\
&= \alpha_1(t, y_1^a, y_2^a)\Delta t \alpha_2(t, y_1^a, y_2^a)\Delta t
\end{aligned}$$

and thus we obtain for the local covariance

$$\widehat{\operatorname{cov}}(\Delta y_1, \Delta y_2) = \widehat{E}(\Delta y_1 \Delta y_2) - \widehat{E}(\Delta y_1)\widehat{E}(\Delta y_2) = 0. \qquad \blacksquare$$

To see what happens with the local covariance between the original processes $X_1$ and $X_2$ we can write in a similar way as in Theorem 2.3.2

$$\begin{aligned}
\widehat{\operatorname{cov}}(\Delta x_1 \Delta x_2) &= [\Delta y_1 \cos\varphi + \Delta y_2 \sin\varphi,\ -\Delta y_1 \sin\varphi + \Delta y_2 \cos\varphi] \\
&= -\cos\varphi \sin\varphi \widehat{\operatorname{var}}(\Delta y_1) - \sin^2\varphi \widehat{\operatorname{cov}}(\Delta y_1, \Delta y_2) \\
&\quad + \cos^2\varphi \widehat{\operatorname{cov}}(\Delta y_1, \Delta y_2) + \cos\varphi \sin\varphi \widehat{\operatorname{var}}(\Delta y_2)
\end{aligned}$$

and then referring to Theorem 2.3.3 we obtain

$$\widehat{\operatorname{cov}}(\Delta x_1, \Delta x_2) = -\cos\varphi \sin\varphi \lambda_1 \Delta t + \cos\varphi \sin\varphi \lambda_2 \Delta t = \rho \sigma_1 \sigma_2 \Delta t = \operatorname{cov}(\Delta x_1, \Delta x_2).$$

The interpretation is that although we are constructing a rotated tree based on the two independent processes $Y_1$ and $Y_2$, when we transform it back to the original problem the covariance between the original processes $X_1$ and $X_2$ is preserved.

**Remark 2.3.2.** *We want to point out here that we cannot choose exactly*

$$J_1(t, y_1^a, y_2^a) = \frac{\alpha_1(t, y_1^a, y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}$$

*because of two reasons*

- *If we want to construct a recombining tree we would like that the absolute value of the up-jump and the absolute value of the down-jump are integer multiples of the same jump heights. Notice that this will not be the case for real values of $J_1(t, y_1^a, y_2^a)$ and $J_2(t, y_1^a, y_2^a)$. Therefore, we need a construction of $J_1(t, y_1, y_2)$ and $J_2(t, y_1, y_2)$ that allows them to take only integer values;*

- *On the other side, if we choose simply*

$$J_1(t, y_1^a, y_2^a) = \mathit{floor}\left[\frac{\alpha_1(t, y_1^a, y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}\right]$$

  *the tree will be allowed to expand endlessly, hence we will take no account of the mean-reversion in the underlying processes. Finally, if we set*

$$J_1(t, y_1^a, y_2^a) = \begin{cases} Z_1 & \text{if } Z_1 \text{ even,} \\ Z_1 + 1 & \text{else.} \end{cases}$$

$$J_2(t, y_1^a, y_2^a) = \begin{cases} Z_2 & \text{if } Z_2 \text{ even,} \\ Z_2 + 1 & \text{else.} \end{cases}$$

  *for $Z_1 := \mathit{floor}\left[\frac{\alpha_1(t,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}\right]$ and $Z_2 := \mathit{floor}\left[\frac{\alpha_2(t,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_2}}\right]$ we notice that in a "normal case" of not too strong mean reversion, the values of $J_1(t, y_1, y_a)$ and $J_2(t, y_1, y_2)$ will most of the time be equal to zero. Thus, we will have a product of two binomial trees in the usual sense. But as soon as $|\alpha_1(t, y_1, y_2)| > \sqrt{\lambda_1 \Delta t}$ or in other words as soon as the local mean dominates the local variance, the branching of the tree will be changed and thus $J_1(t, y_1, y_2)$ and $J_2(t, y_1, y_2)$ will take the values*

*of $-2$ in the upper part of the tree and of $2$ in the down part. Thus, we would have both required effects, namely that the local variance will be approximately matched and the effect of mean reversion will be taken into account. See Figure 2.2 for an example of the mentioned tree construction.*

**Remark 2.3.3.** *Notice that*

$$y_1^{a^+} = y_1^a + \big(J_1(t, y_1^a, y_2^a) + 1\big)\sqrt{\Delta t \lambda_1} \geq y_1^a + \alpha(t, y_1^a, y_2^a)\Delta t$$

*and*

$$y_1^{a^-} = y_1^a + \big(J_1(t, y_1^a, y_2^a) - 1\big)\sqrt{\Delta t \lambda_1} \leq y_1^a + \alpha(t, y_1^a, y_2^a)\Delta t$$

*can be interpreted as the up and down jumps bracketing the real local mean of the approximated process.*

**Lemma 2.3.1.** *In addition to matching the first two moments of the base processes, the probabilities in (2.20) and (2.21) are also well-defined, i.e. $p_1 \in [0,1]$ and $p_2 \in [0,1]$.*

**Proof:** By definition

$$J_1(t, y_1^a, y_2^a) \in \left[\frac{\alpha(t, y_1^a, y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}} - 1, \frac{\alpha(t, y_1^a, y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}} + 1\right]$$

and therefore

$$\big(1 - J_1(t, y_1^a, y_2^a)\big)\sqrt{\lambda_1} \in \left[-\alpha(t, y_1^a, y_2^a)\sqrt{\Delta t},\, 2\sqrt{\lambda_1} - \alpha_1(t, y_1^a, y_2^a)\sqrt{\Delta t}\right].$$

Applying this to the definition of the probabilities in (2.20) and (2.21) one can easily see that $p_1 \in [0,1]$ and $p_2 \in [0,1]$. ■

Once we have constructed the rotated two dimensional tree, we want to use it to price contingent claims via backward recursion. Hence, we start with calculating the values of the claim at maturity at each node of the tree and we denote them by $g_n(y_1^a, y_2^a)$[3].

[3]Notice that typically the option value $g_i(\cdot)$ at time $t_i = 1, \ldots, n$ is given as a function of the interest rate $r\,(t_i, X_1(t_i), X_2(t_i))$ (and not as we define it of $\mathbf{Y}$) but using $\mathbf{X} = \mathbf{Y}\mathbf{U}^\top$ we can easily transform the rotated tree back to the original problem and thus w.l.o.g we are allowed to define directly the option price as a function of the values of $\mathbf{Y}$.

Next, we assume we have calculated all the claim values at the $(i+1)$-th period and now we want to calculate them at the $(i)$-th one. In combining the trees, the independence of the processes allows us to build the new probabilities by simple cross product of the probabilities of the base processes. Thus, given that at time $i$ the state variables are $(y_1^a, y_2^a)$ we have that

$$\begin{aligned} g_i\left(y_1^a, y_2^a\right) &= \Big[p_1 p_2 g_{i+1}\big(y_1^{a^+}, y_2^{a^+}\big) + p_1(1-p_2) g_{i+1}\big(y_1^{a^+}, y_2^{a^-}\big) \\ &\quad +(1-p_1)p_2 g_{i+1}\big(y_1^{a^-}, y_2^{a^+}\big) + (1-p_1)(1-p_2) g_{i+1}\big(y_1^{a^-}, y_2^{a^-}\big)\Big] e^{-R(t, y_1^a, y_2^a)\Delta t} \end{aligned} \tag{2.22}$$

where the interest rate as a function of the new base factors can be written as

$$R(t, y_1, y_2) := f(0,t) + (\cos\varphi y_1 + \sin\varphi y_2) + (-\sin\varphi y_1 + \cos\varphi y_2).$$

We repeat the procedure until we find $g_0(y_1, y_2)$ which is an approximation of the price of the claim today.

## 2.4 Construction of a Quadrinomial Tree

Another way to deal with our problem consists of keeping the original underlying processes

$$dX_1(t) = (-\kappa_1 X_1(t) + \varphi_{11}(t) + \varphi_{12}(t))\, dt + \sigma_1 d\tilde{W}^1(t) \tag{2.23}$$

$$dX_2(t) = (-\kappa_2 X_2(t) + \varphi_{21}(t) + \varphi_{22}(t))\, dt + \sigma_2 d\tilde{W}^2(t) \tag{2.24}$$

and directly constructing a quadrinomial tree. For shortness of the notations, let us denote the drifts of the two processes by $\nu_1(t, X_1) := -\kappa_1 X_1(t) + \varphi_{11}(t) + \varphi_{12}(t)$ and $\nu_2(t, X_2) := -\kappa_2 X_2(t) + \varphi_{21}(t) + \varphi_{22}(t)$.

Next, if we assume that at time $t$ the tree is in state $(x_1^a, x_2^a)$ then at time $t + \Delta t$ it can move to the following four states

$$\begin{aligned} &\left(x_1^{a^+}, x_2^{a^+}\right) && \text{with probability } p_{uu} \\ &\left(x_1^{a^+}, x_2^{a^-}\right) && \text{with probability } p_{ud} \\ &\left(x_1^{a^-}, x_2^{a^+}\right) && \text{with probability } p_{du} \\ &\left(x_1^{a^-}, x_2^{a^-}\right) && \text{with probability } p_{dd} \end{aligned}$$

**Definition 2.4.1.** *Let us define the up and down jumps of both process by*

$$\begin{aligned} x_1^{a^+} &:= x_1^a + \big(J_1(t,x_1^a)+1\big)\sqrt{\Delta t}\sigma_1, \quad x_1^{a^-} := x_1^a + \big(J_1(t,x_1^a)-1\big)\sqrt{\Delta t}\sigma_1 \\ x_2^{a^+} &:= x_2^a + \big(J_2(t,x_2^a)+1\big)\sqrt{\Delta t}\sigma_2, \quad x_2^{a^-} := x_2^a + \big(J_2(t,x_2^a)-1\big)\sqrt{\Delta t}\sigma_2 \end{aligned}$$

*where*

$$J_1(t,x_1^a) := \begin{cases} Z_1 & \text{if } Z_1 \text{ even,} \\ Z_1+1 & \text{else.} \end{cases}$$

$$J_2(t,x_2^a) := \begin{cases} Z_2 & \text{if } Z_2 \text{ even,} \\ Z_2+1 & \text{else.} \end{cases}$$

*for* $Z_1 :=$ *floor* $\left[\frac{\nu_1(t,x_1^a)\sqrt{\Delta t}}{\sigma_1}\right]$ *and* $Z_2 :=$ *floor* $\left[\frac{\nu_2(t,x_2^a)\sqrt{\Delta t}}{\sigma_2}\right]$.

**Theorem 2.4.1.** *If the subsequent probabilities for the* $(y_1^a, y_2^a)$ *node are given by*

$$\begin{aligned} p_{uu} &= \frac{1}{4}\left[\frac{\rho\sigma_1\sigma_2 + (J_1-1)(J_2-1)\sigma_1\sigma_2 - \nu_1(t,x_1^a)\sqrt{\Delta t}(J_2-1)\sigma_2 - \nu_2(t,x_2^a)\sqrt{\Delta t}(J_1-1)\sigma_1}{\sigma_1\sigma_2}\right] \\ p_{dd} &= \frac{1}{4}\left[\frac{\rho\sigma_1\sigma_2 + (J_1+1)(J_2+1)\sigma_1\sigma_2 - \nu_1(t,x_1^a)\sqrt{\Delta t}(J_2+1)\sigma_2 - \nu_2(t,x_2^a)\sqrt{\Delta t}(J_1+1)\sigma_1}{\sigma_1\sigma_2}\right] \\ p_{ud} &= \frac{1}{4}\left[\frac{-\rho\sigma_1\sigma_2 - (J_1-1)(J_2+1)\sigma_1\sigma_2 + \nu_1(t,x_1^a)\sqrt{\Delta t}(J_2+1)\sigma_2 + \nu_2(t,x_2^a)\sqrt{\Delta t}(J_1-1)\sigma_1}{\sigma_1\sigma_2}\right] \\ p_{du} &= \frac{1}{4}\left[\frac{-\rho\sigma_1\sigma_2 - (J_1+1)(J_2-1)\sigma_1\sigma_2 + \nu_1(t,x_1^a)\sqrt{\Delta t}(J_2-1)\sigma_2 + \nu_2(t,x_2^a)\sqrt{\Delta t}(J_1+1)\sigma_1}{\sigma_1\sigma_2}\right] \end{aligned}$$

*then with the choice of jump heights given in Definition 2.4.1, the tree matches locally perfectly the first and approximately the second moments of the underlying processes and in addition, up to terms of order* $\Delta t^2$ *it matches the local covariance between the underlying processes.*

**Proof:** Denoting the approximated (with the lattice) local mean and variance respectively with $\widehat{E}(\Delta x_i)$ and $\widehat{\text{var}}(\Delta x_i)$, for $i = 1,2$ and the true local mean and variance of the underlying process by $E(\Delta x_i)$ and $\text{var}(\Delta x_i)$, for $i = 1,2$ we notice that matching the first

two moments implies that the risk-neutral tree probabilities have to solve

$$\widehat{E}(\Delta x_1) = (p_{uu}+p_{ud})(x_1^{a^+}-x_1^a)+(p_{du}+p_{dd})(x_1^{a^-}-x_1^a) \stackrel{!}{=} E(\Delta x_1) = \nu_1(t,x_1^a)\Delta t \tag{2.25}$$

$$\begin{aligned}\widehat{E}(\Delta x_1^2) &= (p_{uu}+p_{ud})(x_1^{a^+}-x_1^a)^2+(p_{du}+p_{dd})(x_1^{a^-}-x_1^a)^2\\ &\stackrel{!}{=} E(\Delta x_1^2) = \sigma_1^2\Delta t+\nu_1(t,x_1^a)^2\Delta t^2 \end{aligned}\tag{2.26}$$

$$\widehat{E}(\Delta x_2) = (p_{uu}+p_{du})(x_2^{a^+}-x_2^a)+(p_{ud}+p_{dd})(x_2^{a^-}-x_2^a) \stackrel{!}{=} E(\Delta x_2) = \nu_2(t,x_2^a)\Delta t \tag{2.27}$$

$$\begin{aligned}\widehat{E}(\Delta x_2^2) &= (p_{uu}+p_{du})(x_2^{a^+}-x_2^a)^2+(p_{du}+p_{dd})(x_2^{a^-}-x_2^a)^2\\ &\stackrel{!}{=} E(\Delta x_2^2) = \sigma_2^2\Delta t+\nu_2(t,x_2^a)^2\Delta t^2 \end{aligned}\tag{2.28}$$

$$\begin{aligned}\widehat{E}(\Delta x_1\Delta x_2) &= \widehat{\mathrm{cov}}(\Delta x_1\Delta x_2)+\widehat{E}(\Delta x_1)\widehat{E}(\Delta x_2)\\ &= p_{uu}(x_1^{a^+}-x_1^a)(x_2^{a^+}-x_2^a)+p_{ud}(x_1^{a^+}-x_1^a)(x_2^{a^-}-x_2^a)\\ &\quad +p_{du}(x_1^{a^-}-x_1^a)(x_2^{a^+}-x_2^a)+p_{dd}(x_1^{a^-}-x_1^a)(x_2^{a^-}-x_2^a)\\ &\stackrel{!}{=} E(\Delta x_1\Delta x_2) = \sigma_1\sigma_2\rho\Delta t+\nu_1(t,x_1^a)\nu(t,x_2^a)\Delta t^2 \end{aligned}\tag{2.29}$$

$$1 = p_{uu}+p_{ud}+p_{du}+p_{dd} \tag{2.30}$$

where $\Delta x_1$ and $\Delta x_2$ denote the changes in the respective processes $x_1$ and $x_2$ from time $t$ to time $t+\Delta t$. Ignoring the terms of order $\Delta t^2$ in the matching of the correlation and substituting with the definitions of $x_1^{a^\pm}$ and $x_2^{a^\pm}$ we solve equations (2.25), (2.27), (2.29) and (2.30) for the four unknown probabilities and obtain

$$\begin{aligned}
p_{uu} &= \frac{1}{4}\left[\frac{\rho\sigma_1\sigma_2+(J_1-1)(J_2-1)\sigma_1\sigma_2-\nu_1(t,x_1^a)\sqrt{\Delta t}(J_2-1)\sigma_2-\nu_2(t,x_2^a)\sqrt{\Delta t}(J_1-1)\sigma_1}{\sigma_1\sigma_2}\right]\\
p_{dd} &= \frac{1}{4}\left[\frac{\rho\sigma_1\sigma_2+(J_1+1)(J_2+1)\sigma_1\sigma_2-\nu_1(t,x_1^a)\sqrt{\Delta t}(J_2+1)\sigma_2-\nu_2(t,x_2^a)\sqrt{\Delta t}(J_1+1)\sigma_1}{\sigma_1\sigma_2}\right]\\
p_{ud} &= \frac{1}{4}\left[\frac{-\rho\sigma_1\sigma_2-(J_1-1)(J_2+1)\sigma_1\sigma_2+\nu_1(t,x_1^a)\sqrt{\Delta t}(J_2+1)\sigma_2+\nu_2(t,x_2^a)\sqrt{\Delta t}(J_1-1)\sigma_1}{\sigma_1\sigma_2}\right]\\
p_{du} &= \frac{1}{4}\left[\frac{-\rho\sigma_1\sigma_2-(J_1+1)(J_2-1)\sigma_1\sigma_2+\nu_1(t,x_1^a)\sqrt{\Delta t}(J_2-1)\sigma_2+\nu_2(t,x_2^a)\sqrt{\Delta t}(J_1+1)\sigma_1}{\sigma_1\sigma_2}\right]
\end{aligned}$$

Next, let us calculate the local conditional variances for the already found probabilities

first for the $x_1$ process:

$$\begin{aligned}
\widehat{\mathrm{var}}\left(\Delta x_1\right) &= \widehat{E}\left(\Delta x_1^2\right) - \widehat{E}\left(\Delta x_1\right)^2 \\
&= (p_{uu} + p_{ud})(x_1^{a^+} - x_1^a)^2 + (p_{du} + p_{dd})(x_1^{a^-} - x_1^a)^2 - \nu_1(t, x_1^a)^2 \Delta t^2 \\
&= \left(\frac{\nu_1(t, x_1^a)\sqrt{\Delta t} - (J_1 - 1)\sigma_1}{2\sigma_1}\right)(1 + J_1)^2 \Delta t \sigma_1^2 \\
&\quad + \left(1 - \frac{\nu_1(t, x_1^a)\sqrt{\Delta t} - (J_1 - 1)\sigma_1}{2\sigma_1}\right)(1 - J_1)^2 \Delta t \sigma_1^2 - \nu_1(t, x_1^a)^2 \Delta t \\
&= -J_1^2 \Delta t \sigma_1^2 + 2\nu_1(t, x_1^a)\Delta t \sqrt{\Delta t \sigma_1^2} J_1 - \nu_1(t, x_1^a)^2 \Delta t + \sigma_1^2 \Delta t \\
&= -(J_1 \sqrt{\Delta t}\sigma_1 - \nu_1(t, x_1^a)\Delta t)^2 + \sigma_1^2 \Delta t
\end{aligned}$$

and since $\mathrm{var}\left(\Delta x_1\right) = \sigma_1^2 \Delta t$ we have as in the rotated tree that the local variance is approximately matched due to the right definition of the $J_1$ process. By symmetry, the same follows for the $x_2$ process. ■

**Remark 2.4.1.** *Notice that although the sum of the probabilities is one, they are not necessarily bounded between* 0 *and* 1*. It can however be shown that when we increase the refinement, the probabilities converge to values between* 0 *and* 1*. It is also the case that the probabilities are not well-defined mostly in extreme cases of high drifts, volatilities and correlation.*

*A similar problem is encountered also in the Hull and White [37] 2-factor trinomial tree construction and also by Brigo and Mercurio [12] in their quadrinomial tree construction. There are two possible ways to proceed.*

*The first approach is to change the correlation coefficient at every node where the probabilities are not bounded in* $[0, 1]$*, until we have well-defined probabilities. Brigo and Mercurio claim that this procedure although theoretically not consistent is practically applicable as the correlation has a very negligible contribution to the result.*

*A second way, (considered by Zvan, Forsyth and Vetzal [67]) is to treat the probabilities of the tree as weights and neglect that they are not properly defined. In this respect, the*

*authors show for a variety of finite difference constructions (although for rather regular parameters) that the meshes allowing for negative coefficients satisfy discrete maximum and minimum principles as the mesh size parameter approaches zero and show no obvious oscillations in the level curves of the priced option values.*

*In this thesis, we follow the second approach in the construction of the quadrinomial tree, as at least theoretically we can increase the refinement of the tree until we have well-defined probabilities.*

**Lemma 2.4.1.** *If we let the number of time steps tend to infinity and denote* $p_{uu} \xrightarrow{\Delta t \to 0} \widetilde{p}_{uu}$, $p_{ud} \xrightarrow{\Delta t \to 0} \widetilde{p}_{ud}$, $p_{du} \xrightarrow{\Delta t \to 0} \widetilde{p}_{du}$, $p_{dd} \xrightarrow{\Delta t \to 0} \widetilde{p}_{dd}$, *then we have that* $\widetilde{p}_{uu} \in [0,1]$, $\widetilde{p}_{ud} \in [0,1]$, $\widetilde{p}_{du} \in [0,1]$, $\widetilde{p}_{dd} \in [0,1]$.

**Proof:** Recall the definition of $J_1$ and $J_2$

$$J_1(t, x_1^a) \quad := \quad \begin{cases} Z_1 & \text{if } Z_1 \text{ even,} \\ Z_1 + 1 & \text{else.} \end{cases}$$

$$J_2(t, x_2^a) \quad := \quad \begin{cases} Z_2 & \text{if } Z_2 \text{ even,} \\ Z_2 + 1 & \text{else.} \end{cases}$$

and notice that as 0 is an even number and

$$\begin{aligned} Z_1 &= \text{floor}\left[\frac{\nu_1(t, x_1^a)\sqrt{\Delta t}}{\sigma_1}\right] \xrightarrow{\Delta t \to 0} 0 \\ Z_2 &= \text{floor}\left[\frac{\nu_2(t, x_2^a)\sqrt{\Delta t}}{\sigma_2}\right] \xrightarrow{\Delta t \to 0} 0 \end{aligned}$$

then we have that

$$\begin{aligned} J_1(t, x_1^a) &\xrightarrow{\Delta t \to 0} 0 \\ J_2(t, x_2^a) &\xrightarrow{\Delta t \to 0} 0. \end{aligned}$$

Letting the number of time steps go to infinity and using that the correlation $\rho \in [-1, 1]$

we find that

$$
\begin{aligned}
p_{uu} &\xrightarrow{\Delta t \to 0} \frac{1}{4}(\rho + 1) \in [0, \frac{1}{2}] \\
p_{dd} &\xrightarrow{\Delta t \to 0} \frac{1}{4}(\rho + 1) \in [0, \frac{1}{2}] \\
p_{ud} &\xrightarrow{\Delta t \to 0} \frac{1}{4}(-\rho + 1) \in [0, \frac{1}{2}] \\
p_{du} &\xrightarrow{\Delta t \to 0} \frac{1}{4}(-\rho + 1) \in [0, \frac{1}{2}]
\end{aligned}
$$

■

**Remark 2.4.2.** *Finally, let us examine the convergence of the probabilities for some special cases of correlation. At first, let the two processes be perfectly correlated ($\rho = 1$). Then we obtain that*

$$
\begin{aligned}
p_{uu} &\xrightarrow{\Delta t \to 0} \frac{1}{2} \quad , \qquad p_{dd} \xrightarrow{\Delta t \to 0} \frac{1}{2} \\
p_{ud} &\xrightarrow{\Delta t \to 0} 0 \quad , \qquad p_{du} \xrightarrow{\Delta t \to 0} 0
\end{aligned}
$$

*which logically excludes the events where we have a jump-up in one of the trees and a jump-down of the other one due to the perfect correlation.*

*Next, we consider the perfect negative correlation case. Here we find*

$$
\begin{aligned}
p_{uu} &\xrightarrow{\Delta t \to 0} 0 \quad , \qquad p_{dd} \xrightarrow{\Delta t \to 0} 0 \\
p_{ud} &\xrightarrow{\Delta t \to 0} \frac{1}{2} \quad , \qquad p_{du} \xrightarrow{\Delta t \to 0} \frac{1}{2}
\end{aligned}
$$

*which by analogy means that the trees are allowed to move only in opposite directions.*

*The last case is when the trees are uncorrelated ($\rho = 0$). Here we obtain*

$$
\begin{aligned}
p_{uu} &\xrightarrow{\Delta t \to 0} \frac{1}{4} \quad , \qquad p_{dd} \xrightarrow{\Delta t \to 0} \frac{1}{4} \\
p_{ud} &\xrightarrow{\Delta t \to 0} \frac{1}{4} \quad , \qquad p_{du} \xrightarrow{\Delta t \to 0} \frac{1}{4}
\end{aligned}
$$

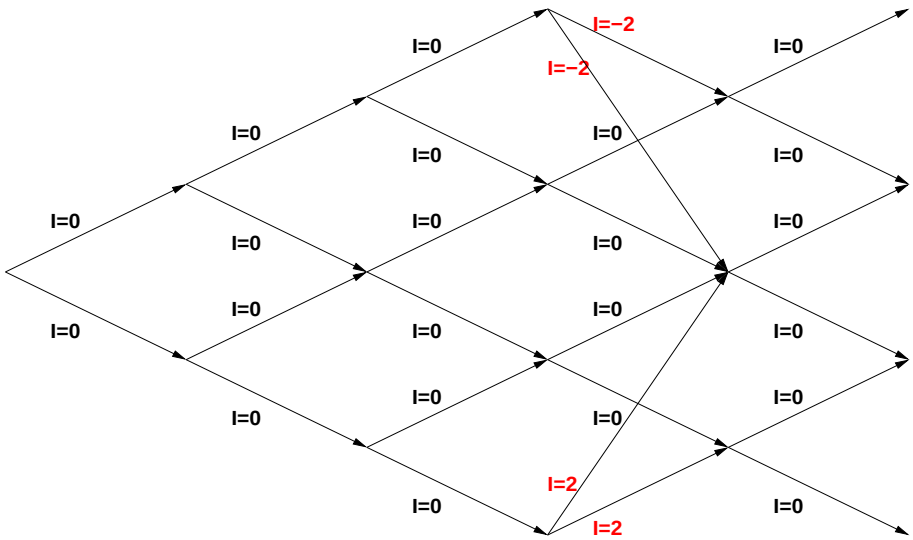



Figure 2.2: Mean Reversion in an Adaptive Binomial Tree. Results are taken from a rotated tree in the direction of the $Y_1$ process where we have set $\kappa_1 = \kappa_2 = 60\%$, $\sigma_1 = \sigma_2 = 0.3\%$, $T = 1$, $\rho = 0$.

*which gives equal weights to all four jumps.*

*Notice that in the previous method for the tree construction we would always have convergence of $p_1$ and $p_2$ to 0.5 as the one-dimensional trees are independent.*

Finally, we can conclude that the probabilities are asymptotically well-defined and increasing the number of time steps, we can obtain the right convergence. However, the rate of this convergence is so slow that practically, the computer memory required to store the refined tree is often insufficient. This is one reason to prefer the first method where we have proved that the probabilities are well-defined.

**Remark 2.4.3.** *A very interesting feature of both trees is that the jump-heights are node-dependent and therefore (at each node) locally determined. This property of the tree allows self-integration of the mean-reversion feature of the approximated process. For a comparison, in the Hull and White [37] trinomial tree a certain choice of a "stretch parameter" $j_{max}$ has to be done before implementing the algorithm. This parameter is used to change the branching of the tree (and the respective probabilities) when it expands too much so that the mean-reversion is reflected. In the typical one-dimensional case, Hull and White [37] showed that in their trinomial tree construction, the indices of the nodes should not get*

*greater than $j_{max}$ which is calculated to be the smallest integer greater than $0.184/\kappa\Delta t$ and should not get smaller than $-j_{max}$ since otherwise the risk-neutral probabilities will become negative.*

*We point out here that in the lattice construction methods we have presented so far, no such external check is needed as this is done internally in the tree construction procedure by the dependence of the jump heights on the drift parameters. See Figure 2.2.*

## 2.5 Pricing of Path-Dependent Claims via Two - Dimensional Trees

### 2.5.1 General Idea

Let us assume that we are pricing options in a finite time horizon $[0, T]$ and the options can have payouts at the discrete time points $0 \leq T_0 < T_1 < \ldots < T_m = T$, where notice that $T_0$ is not necessarily 0.

Let us further assume that the claims we are pricing have "**strong path-dependence**" in the sense that their value is a function of at least one more independent variable which accounts for the path-dependence. A typical example for options with strong path-dependence is the Asian option whose payoff depends on the average value of the underlying asset until expiry. For the pricing of the Asian option, Rogers and Shi [60] showed that the so called "running average" contains all the extra information needed and they capture it by adding in the pricing PDE a new independent random variable to account for it. The option value is then a function not only of the underlying processes but also of this additional random variable.

Pioneers in the pricing of path-dependent options with strong path dependence with a lattice-based method (called the **forward shooting grid**(FSG)) are Hull and White [36] and Ritchken, Sankarasubramanian and Vijh [57]. They used the FSG algorithm in specific for pricing European or American types of Asian and look-back options. A further and more systematic framework was presented by Barraquand and Pudet [4] and an extension of the FSG for pricing Parisian-style options was offered by Avellaneda and

Wu [3] and Kwok and Lau [44]. The FSG approach is generally characterized by adding the path-dependence as an auxiliary state vector at each node of the lattice and in contrast to the finite-difference method, it does not need to deal with the corresponding governing pricing differential equation. Notice that for some types of complicated path-dependence, deriving a pricing PDE is not straightforward and in these cases the FSG has an advantage over the finite-difference method (examples here are the cumulative or consecutive Parisian barrier options).

Applied to our model, we notice that if we denote the extra variable that accounts for the path dependence by $Z$, then for a contingent claim $V$ with strong path dependence we have

$$E^Q\left(V(T_{i+1})|\mathfrak{F}_{T_i}\right) = E^Q\left(V(T_{i+1})|X_1(T_i), X_2(T_i), Z(T_i)\right), \quad \text{for} \quad i = 0, \ldots, m$$

which means that although an option with strong path-dependence is not a Markov process w.r.t. the underlying processes, it becomes Markovian when we add the information contained in the last state of the variable accounting for the path-dependence.

This variable $Z$ could be for example the running maximum or minimum of the past Libor rates, the number of times a barrier has been hit or a variety of other functions of the Libor rate which we are going to present in the next sections.

**Remark 2.5.1.** *For a comparison, notice that options with "**weak dependence**" have a pricing PDE that can be written only in the underlying variables and time with some proper boundary conditions. Therefore, using the tree construction method or the finite difference method does not require the addition of any extra variables. Typical examples of options with weak path-dependence are the simple Barrier options of the type "down-and-in", "down-and-out", "up-and-in" and "up-and-out" or the American options.*

Let us assume that the values of the state process $Z(\cdot)$ change discretely at the payout time points $0 \leq T_0 < T_1 < \ldots < T_m = T$. In order to be able to estimate the value of $Z$, we need an initializing function $Z_{init}(t, X_1, X_2)$ that gives us the start value of the process and a change function $Z_{next}(t, X_1, X_2, Z_{past})$ that gives us the jump of the $Z$-process at

each payout point. Referring to Wilmott [63], we define the evolution of the $Z$-process by the following steps

- Initialize $Z$

$$Z(T_0) := Z_{init}(T_0, X_1(T_0), X_2(T_0)) \tag{2.31}$$

- Change $Z$ at each time point $T_1, \ldots, T_m$

$$Z(T_i) := Z_{next}\left(T_i, X_1(T_i), X_2(T_i), Z(T_{i-1})\right), \quad i = 1, \ldots, m \tag{2.32}$$

- Keep the value of $Z$ for $t \in [T_i, T_{i+1})$, $i = 0, \ldots, m$

$$Z(t) := Z(T_i) \tag{2.33}$$

With the latter construction of the state process $Z$ we can easily price path-dependent claims with a Monte Carlo simulation. In the next Section, we shall present several financial instruments with strong path dependence contingent on the interest rate and for each of them we shall define the appropriate path-dependent variable $Z$ and its initializing and renewing functions.

As the tree construction is an approximation of a continuous distribution, due to no-arbitrage reasons we need the following requirement:

$$V\left(T_i^-, X_1(T_i), X_2(T_i), Z(T_i^-)\right) = V\left(T_i^+, X_1(T_i), X_2(T_i), Z_{next}(T_i, X_1(T_i), X_2(T_i), Z(T_i^-)\right) \tag{2.34}$$

for $i = 1, \ldots, m$ and where $T_i^-$ denotes the time infinitesimally before $T_i$ and $T_i^+$ denotes the time infinitesimally after $T_i$.

This requirement is called a "**Jump Condition due to an update of** $Z$" and is a simple reflection of the fact that as we get closer to a point where the value of the $Z$-process will be updated, due to the continuity of the underlying process (between two payoff points)

and the deterministic rule of an update of $Z$, we become aware of the value of $Z$ after an update. Thus, the information about the future value of $Z$ will already be incorporated in the price of the contingent claim and since no money will really "change hands" at an update point, there can also be no jump in the price of the option (see Zhu and Stokes [66] or Wilmott [63] for more details).

Notice that in (2.34) we have implicitly assumed that the value of the basis processes does not change in the interval $[T_i^-, T_i^+]$ for all $i = 1, \ldots, m$ and thus we can write $X_1(T_i^-) = X_1(T_i^+) = X_1(T_i)$ and $X_2(T_i^-) = X_2(T_i^+) = X_2(T_i)$ for $i = 1, \ldots, m$. We keep this assumption until the end of this work.

On the other side, due to simple no-arbitrage considerations notice that the value of the option does change (or jump) in the case of payoffs at discrete time points (see Forsyth, Vertzal and Zvan [28] or again Wilmott [63]). Thus, we require that

$$\begin{aligned} V\big(T_i^-, X_1(T_i), X_2(T_i), Z(T_i^+)\big) &= V\big(T_i^+, X_1(T_i), X_2(T_i), Z(T_i^+)\big) \\ &\quad + \text{Payout}\big(T_i^+, X_1(T_i), X_2(T_i), Z(T_i^+)\big) \end{aligned} \tag{2.35}$$

which is usually also called a "**Jump condition due to Payout**".

Now comes the question to apply both conditions to the tree procedure.

To answer it, let us first assume that the discretization of the time interval $[0, T]$ needed for the tree construction consists of $n$ time steps $0 \leq t_0 < t_1 < \ldots < t_n = T$ such that $n$ is generally much bigger than $m$. Let us also assume for simplicity (which is relaxed in the practical application) that there is always a time step of the tree on all payout points i.e. we assume that the time slices of the tree can be written as $T_0 = t_0 < \ldots < t_k = T_1 < \ldots < t_{mp} = T_m$ where $p = \frac{n}{m}$ with $p \in \mathbb{N}$.

**Range of Z** :

While constructing the tree, we notice that various different paths lead to each node and to each of them corresponds a different value of the $Z$ variable that accounts for the path-dependence of the option. Notice that after $n$ time steps the different values of $Z$ in a single node will be equal to the total number of paths that lead to the specific node. Since it is impossible to keep track of all of these values for a big number of discretization steps

$n$, we choose to keep at each node only the range of the $Z$-values (i.e. the smallest and the biggest value of $Z$ attained in a given node) and then partition the obtained interval into a finite number of points in which the option value will be computed.

### 2.5.2 Forward Construction of the Tree

We calculate the range of the $Z$ variable in the forward construction of the tree where we assume, without loss of generality (as this will be the case of all path dependent options we will present in the next section and later price) that $Z_{next}(t, X_1(t), X_2(t), Z_{old})$ is monotonously increasing in $Z_{old}$. In addition, we make a difference between its values before an update at time $t$ which we denote again with $Z(t^-)$ and after an update at time $t$ which we denote with $Z(t^+)$ where $t^-$ is a time instance infinitesimally before time $t$ and $t^+$ is a time instance infinitesimally after time $t$. Hence, we construct the range of the $Z$-variable for all $t_j = 0, \ldots, n$ in the following way:

- Initialize $Z$ in $t_0$:

$$\begin{aligned} Z_{min}(t_0^-) &= Z_{max}(t_0^-) = Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right) \\ Z_{min}(t_0^+) &= Z_{max}(t_0^+) = Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right). \end{aligned}$$

- Point of a **no-update** of the value of $Z$, i.e. when $t_j \neq T_i$, for $i = 1, \ldots, m, j = 1, \ldots, n$, then the range of $Z$ before a jump in node $(*)$ is given by $[Z_{min}(t_j^-)^*, Z_{max}(t_j^-)^*]$ such that

$$\begin{aligned} Z_{max}(t_j^-)^* &= \max\left(Z_{max}(t_{j-1}) \text{ of all nodes in } t_{j-1} \text{ that lead to node } (*)\right) \\ Z_{min}(t_j^-)^* &= \min\left(Z_{min}(t_{j-1}) \text{ of all nodes in } t_{j-1} \text{ that lead to node } (*)\right) \end{aligned}$$

  and its range after an update is given by $[Z_{min}(t_j^+), Z_{max}(t_j^+)]$ such that

$$\begin{aligned} Z_{max}(t_j^+)^* &= Z_{max}(t_j^-)^* \\ Z_{min}(t_j^+)^* &= Z_{min}(t_j^-)^*. \end{aligned}$$

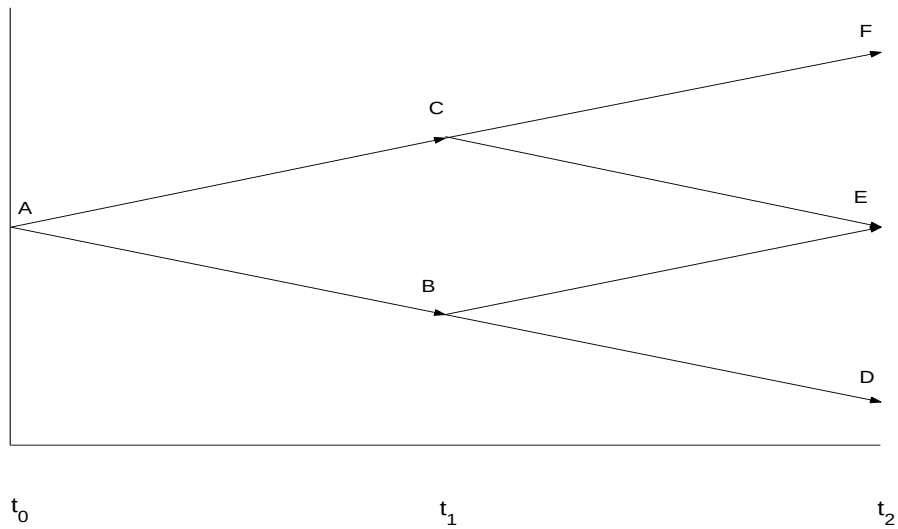



Figure 2.3: One-dimensional Binomial Tree Example.

➪ Point of an **update** in the value of $Z$, i.e. when $t_j = T_i$, for $i = 1, \ldots, m, j = 1, \ldots, n$. First, we have to again find the range of $Z$ before an update in node $(*)$ which is given by $[Z_{min}(t_j^-)^*, Z_{max}(t_j^-)^*]$ using

$$\begin{aligned} Z_{max}(t_j^-)^* &= \max\big(Z_{max}(t_{j-1}^+) \text{ of all nodes in } t_{j-1} \text{ that that lead to node } (*)\big) \\ Z_{min}(t_j^-)^* &= \min\big(Z_{min}(t_{j-1}^+) \text{ of all nodes in } t_{j-1} \text{ that that lead to node } (*)\big) \end{aligned}$$

but then the range of $Z$ in node $(*)$ after an update is given by $[Z_{min}(t_i^+)^*, Z_{max}(t_i^+)^*]$ such that

$$\begin{aligned} Z_{min}(t_i^+)^* &= Z_{next}\big(t_i, X_1^*(t_i), X_2^*(t_i), Z_{min}(t_i^-)^*\big) \\ Z_{max}(t_i^+)^* &= Z_{next}\big(t_i, X_1^*(t_i), X_2^*(t_i), Z_{max}(t_i^-)^*\big). \end{aligned}$$

Thus, at each node we keep track of the maximum and minimum value of the $Z$ process that has been attained at that node.

**Example 2.5.1.** *To illustrate the method of estimating the range of a path-dependent variable $Z$ we consider a simple one-dimensional binomial tree with discretization $n = 2$ and the specific nodes as denoted in Figure 2.3. Assume further that there is a change in the $Z$ process only at time $t_2$. We proceed as follows:*

*1). At time $t = t_0$ node A:*

$$Z_{max}(t_0^-) = Z_{min}(t_0^-) = Z_{max}(t_0^+) = Z_{min}(t_0^+) = Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right)$$

*2). At time $t = t_1$ node B:*

$$\begin{aligned}
Z_{max}(t_1^-)^B &= Z_{max}(t_0^-)^A \\
Z_{min}(t_1^-)^B &= Z_{min}(t_0^-)^A \\
Z_{max}(t_1^+)^B &= Z_{max}(t_1^-)^B \\
Z_{min}(t_1^+)^B &= Z_{min}(t_1^-)^B
\end{aligned}$$

*by analogy for node C.*

*3). At time $t = t_2$ node D:*

$$\begin{aligned}
Z_{max}(t_2^-)^D &= Z_{max}(t_1)^B \\
Z_{min}(t_2^-)^D &= Z_{min}(t_1)^B \\
Z_{max}(t_2^+)^D &= Z_{next}\left(t_2, X_1(t_2)^D, X_2(t_2)^D, Z_{max}(t_2^-)^D\right) \\
Z_{min}(t_2^+)^D &= Z_{next}\left(t_2, X_1(t_2)^D, X_2(t_2)^D, Z_{min}(t_2^-)^D\right)
\end{aligned}$$

*and by analogy for node F. Notice that for node E:*

$$\begin{aligned}
Z_{max}(t_2^-)^E &= max\left(Z_{max}(t_1)^B, Z_{max}(t_1)^C\right) \\
Z_{min}(t_2^-)^E &= min\left(Z_{min}(t_1)^B, Z_{min}(t_1)^C\right) \\
Z_{max}(t_2^+)^E &= Z_{next}\left(t_2, X_1(t_2)^E, X_2(t_2)^E, Z_{max}(t_2^-)^E\right) \\
Z_{min}(t_2^+)^E &= Z_{next}\left(t_2, X_1(t_2)^E, X_2(t_2)^E, Z_{min}(t_2^-)^E\right).
\end{aligned}$$

**Remark 2.5.2.** *Notice that if we were keeping the values of $Z$, then for a single node we would have*

$$Z(t_j^-) = Z(t_{j-1}^+)^* \quad \textit{for all} \quad j = 1, \ldots, n \tag{2.36}$$

*for all the four states at time $t_j$ stemming from node $(*)$ due to the construction of $Z$ (recall (2.33)) and also*

$$Z(t_j^+) = Z_{next}\left(t_j, X_1(t_j), X_2(t_j), Z(t_j^-)\right) \quad \text{if} \quad t_j = T_i \quad \text{for} \quad j = 1, \ldots, n, i = 1, \ldots, m \tag{2.37}$$

$$Z(t_j^+) = Z(t_j^-) \quad \text{if} \quad t_j \neq T_i \quad \text{for} \quad j = 1, \ldots, n, i = 1, \ldots, m. \tag{2.38}$$

### 2.5.3 Backward Recursion

In the backward procedure, at each node we discretize the intervals $[Z_{max}, Z_{min}]$ for both before and after jump values into $K \in \mathbb{N}$ equidistant subintervals and denote the $k$-th point by $Z_k$ with

$$Z_k = Z_{min} + k\frac{Z_{max} - Z_{min}}{K}, \quad k = 0, \ldots, K.$$

Adopting the notation $V\left(t^-, X_1(t^-), X_2(t^-), Z(t^-)\right)$ for the value of the claim at time $t^-$ infinitesimally before $t$ and $V\left(t^+, X_1(t^+), X_2(t^+), Z(t^+)\right)$ for its value at time $t^+$ infinitesimally after $t$, we notice that in order to take into account both "Jump Conditions" (2.34) and (2.35) we need to combine the tree construction schemes of Figure 2.4 a.) and b.) and thus obtain Figure 2.4 c.).

More precisely, in the backward tree construction we go through the following steps:

**A**. Find the value $V\left(t_n^-, X_1(t_n), X_2(t_n), Z_k(t_n^+)\right)$ of the option at maturity $t_n = T_m = T$ for all $k = 0, \ldots, K$ using

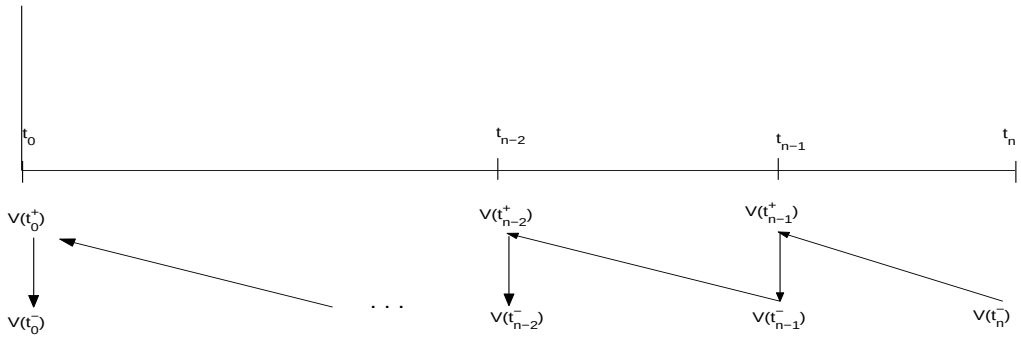
$$V\left(t_n^-, X_1(t_n), X_2(t_n), Z_k(t_n^+)\right) = \text{Terminal Payout}\left(t_n, X_1(t_n), X_2(t_n), Z_k(t_n^+)\right), \; k = 0, \ldots, K$$

if there is a final payout at maturity or using

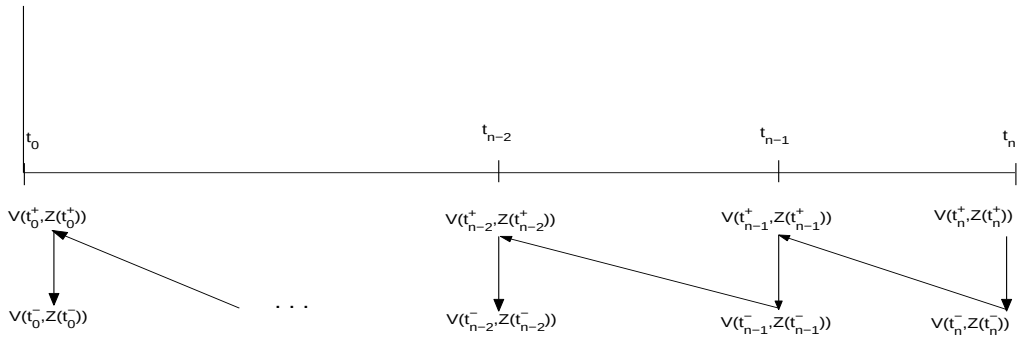
$$V\left(t_n^-, X_1(t_n), X_2(t_n), Z_k(t_n^+)\right) = 0, \quad k = 0, \ldots, K$$

if there is no payout at maturity.

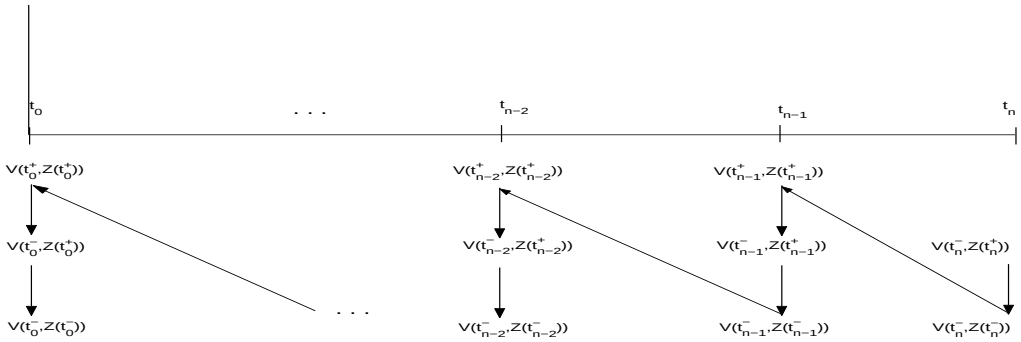
**B**. For $j = n, \ldots, 0$




(a) Backward tree construction procedure in the case of a "Jump Condition due to Payout" (2.34).




(b) Backward tree construction procedure in the case of a "Jump Condition due to an update of Z" (2.35).




(c) Backward tree construction procedure in the case of both "Jump Condition due to an update of Z" (2.35) and "Jump Condition due to Payout" (2.34).

Figure 2.4: Backward Tree Construction Scheme.

**1.** **a)** If $t_j$ is a point of an update of $Z$, due to (2.34) we have that

$$V\left(t_j^-, X_1(t_j^-), X_2(t_j), Z_k(t_j^-)\right) \\ = V\left(t_j^-, X_1(t_j), X_2(t_j), Z_{next}\left(t_j, X_1(t_j), X_2(t_j), Z_k(t_j^-)\right)\right)$$

for all $k = 0, \ldots, K$ and thus we can find the option values before a change in the $Z$-process from the option values after an update of the $Z$-process (which is found in step A.) by interpolation (see Remark 2.5.3).

**b)** If $t_j$ is not an update point of $Z$, then $Z_k(t_j^+) = Z_k(t_j^-)$, $k = 1, \ldots, K$ and $V\left(t_j^-, X_1(t_j), X_2(t_j), Z_k(t_j^-)\right) = V\left(t_j^-, X_1(t_j), X_2(t_j), Z_k(t_j^+)\right)$, $k = 0, \ldots, K$.

**2.** Assuming that at time $t_{j-1}$ we are in node $(*)$, due to (2.36) we have

$$Z_k(t_j^-)^{*uu} = Z_k(t_j^-)^{*ud} = Z_k(t_j^-)^{*du} = Z_k(t_j^-)^{*dd} = Z_k(t_{j-1}^+)^*,\ k = 0, \ldots, K$$

for all four subsequent nodes and thus

$$\begin{aligned} V\left(t_{j-1}^+, X_1(t_{j-1})^*, X_2(t_{j-1})^*, Z_k(t_{j-1}^+)^*\right) \\ = \ & p_{uu} V\left(t_j^-, X_1(t_j)^{*uu}, X_2(t_j)^{*uu}, Z_k(t_{j-1}^+)^*\right) \\ & + p_{ud} V\left(t_j^-, X_1(t_j)^{*ud}, X_2(t_j)^{*ud}, Z_k(t_{j-1}^+)^*\right) \\ & + p_{du} V\left(t_j^-, X_1(t_j)^{*du}, X_2(t_j)^{*du}, Z_k(t_{j-1}^+)^*\right) \\ & + p_{dd} V\left(t_j^-, X_1(t_j)^{*dd}, X_2(t_j)^{*dd}, Z_k(t_{j-1}^+)^*\right), \quad k = 0, \ldots, K. \end{aligned} \tag{2.39}$$

Recall that we have only estimated in step A.1) the option value $V\left(t_j^-, \cdot, \cdot, \cdot\right)$ only for the equidistant points $Z_k(t_j^-) = Z_{min}(t_j^-) + k\frac{Z_{max}(t_j^-) - Z_{min}(t_j^-)}{K}$, $k = 0, \ldots, K$. It may thus happen that we do not have $V\left(t_j^-, \cdot, \cdot, \cdot\right)$ exactly for the value of $Z_k(t_{j-1}^+)^*$ and we will need to linearly interpolate between the available values in order to calculate (2.39).

**3.** Next,

**a)** If $t_{j-1}$ is a payout point, due to (2.35) we have that

$$\begin{aligned} &V\left(t_{j-1}^-, X_1(t_{j-1}), X_2(t_{j-1}), Z_k(t_{j-1}^+)\right) \\ &\quad = V\left(t_{j-1}^+, X_1(t_{j-1}), X_2(t_{j-1}), Z_k(t_{j-1}^+)\right) \\ &\quad +\text{Payout}\big(t_{j-1}^+, X_1(t_{j-1}), X_2(t_{j-1}), Z_k(t_{j-1}^+))\big), \quad k = 0, \ldots, K. \end{aligned}$$

Move to the previous timeslice (step B.).

**b)** If $t_{j-1}$ is not a payout point, then

$$\begin{aligned} &V\left(t_{j-1}^-, X_1(t_{j-1}), X_2(t_{j-1}), Z_k(t_{j-1}^+)\right) \\ &\quad = V\left(t_{j-1}^+, X_1(t_{j-1}), X_2(t_{j-1}), Z_k(t_{j-1}^+)\right), \quad k = 0, \ldots, K. \end{aligned}$$

Move to the previous timeslice (step B.).

The final result (or the approximated price of the option) is given by $V\left(t_0^-, X_1(t_0), X_2(t_0), Z_k(t_0^-)\right)$, $k = 0, \ldots, K$ where since $Z_{max}(t_0^-) = Z_{min}(t_0^-)$ we have that also $Z_0(t_0^-) = \ldots = Z_K(t_0^-)$ and therefore

$$V\left(t_0^-, X_1(t_0), X_2(t_0), Z_0(t_0^-)\right) = \ldots = V\left(t_0^-, X_1(t_0), X_2(t_0), Z_K(t_0^-)\right).$$

**Remark 2.5.3.** *In point B1.a) in the backward recursion, let us assume that at time $t_j$ we are at node $\left(x_1^*, x_2^*, Z(t_j^+)^*\right)$ for which the Z-process has the range before a change $\left[Z_{min}(t_j^-)^*, Z_{min}(t_j^-)^*\right]$ and there is also a jump of Z at that point. Then, due to the "Jump condition due to an update of Z" in (2.34) we require that*

$$V\left(t_j^-, X_1^*, X_2^*, Z_k(t_j^-)^*\right) = V\left(t_j^-, X_1^*, X_2^*, Z_{next}\left(t_j, X_1^*, X_2^*, Z_k(t_j^-)^*\right)\right), \; k = 1, \ldots, K. \tag{2.40}$$

*It is now possible that the value of the option after a change in Z exactly for $Z_{next}\left(t_j, X_1^*, X_2^*, Z_k(t_j^-)^*\right)$ has not been calculated. However, as a consequence of the monotonicity of the renewing function $Z_{next}(\cdot)$ in $Z_{old}$ we would have that*

$$Z_{next}\left(t_j, X_1^*, X_2^*, Z_k(t_j^-)^*\right) \in \left[Z_{next}\left(t_j, X_1^*, X_2^*, Z_{min}(t_j^-)^*\right), \; Z_{next}\left(t_j, X_1^*, X_2^*, Z_{max}(t_j^-)^*\right)\right]$$

*and due to the right construction of the range of $Z$ we would also have that*

$$\left[Z_{next}\left(t_j, X_1^*, X_2^*, Z_{min}(t_j^-)^*\right), Z_{next}\left(t_j, X_1^*, X_2^*, Z_{max}(t_j^-)^*\right)\right] \subset \left[Z_{min}(t_j^+)^*, Z_{max}(t_j^+)^*\right].$$

*As a result, the price of the contingent claim for values of $Z$ slightly smaller and slightly bigger than $Z_{next}\left(t_j, X_1^*, X_2^*, Z_k(t_j^-)^*\right)$ will be available*[4]*. Linearly interpolating between then, using the normalized distances as weights, we can get the required value of the option before a jump in the $Z$-process at time $t_j$.*

**Remark 2.5.4.** *For analysis of the propagation of the interpolation error and its convergence as the time step tends to zero, we refer to Barraquand and Pudet [4] or Forsyth, Vetzal and Zvan [28].*

### 2.5.4 Example 1: AutoCap

An AutoCap, also called a "flexible cap" or a "limit cap" is similar to a **Cap**, with the additional feature that it allows at most $\gamma \leq m$, $\gamma \in \mathbb{N}$ Caplets to be exercised over the lifetime of the option and they have to be **automatically** exercised when in the money. Notice that $m$ denotes the total number of possible Caplets in the lifetime of the option.

**Definition 2.5.1.** *At the end of a period $i$, an AutoCap Option is defined to have the following payoff*

$$N_i(t_i - t_{i-1})r_i$$

*for a given deterministic notional value $N_i$, $i = 1, \ldots, m$ and where*

$$r_i := \left(L\left(t_{i-1}, t_i\right) - K_i\right)^+ 1_{\{A_i < \gamma\}}, \quad i = 0, \ldots, m$$

*with $L\left(t_{i-1}, t_i\right)$ denoting the Libor rate, set at time $t_{i-1}$ for a delivery at time $t_i$ and $K_i$, $i = 1, \ldots, m$ are given deterministic strike prices. The number of Caplets that have been exercised before period $i$ is defined as*

$$A_i := \sum_{j<i} 1_{\{L(t_j, t_{j+1}) > K_j\}}.$$

[4]Since we have divided the range of $Z$ into $K$ equidistant intervals, we interpolate between the closest values in the discretization. We will notice later that for some of the options with complicated path-dependence, the refinement of the discretization of $Z$ will play an important role in the convergence of the approximated option value to its true price.

Thus, the total discounted (until time 0) payoff of the AutoCap can be written as

$$\sum_{i=1}^{m} N_i(t_i - t_{i-1})\big(L(t_{i-1}, t_i) - K_i\big)^+ D(0, t_i) 1_{\{A_i < \gamma\}}.$$

for $D(0, t_i)$ denoting the proper discount rate for a payoff at time $t_i$.

**Remark 2.5.5.** *Recall for comparison that the discounted payoff of a simple Cap is given by*

$$\sum_{i=1}^{m} N(t_i - t_{i-1})\big(L(t_{i-1}, t_i) - K\big)^+ D(0, t_i).$$

*from where one can easily notice the resemblance with the AutoCap.*

**The construction of the Z-variable is as follows:**

$$\begin{aligned} Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right) &:= 0 \\ Z_{next}\left(t_i, X_1(t_i), X_2(t_i), Z_{old}\right) &:= \begin{cases} Z_{old} + 1 & \text{if} \quad L\left(t_{i-1}, t_i; X_1(t_i), X_2(t_2)\right) > s_i, \\ Z_{old} & \text{else} \end{cases} \end{aligned}$$

where the Libor rate is given by

$$L(t_{i-1}, t_i; X_1(t_i), X_2(t_2)) = \frac{1}{t_i - t_{i-1}} \left( \frac{1}{Bond\left(t_{i-1}, t_i, X_1(t_i), X_2(t_i)\right)} - 1 \right)$$

with the bond price as calculated in (2.13). Notice that we have chosen for a state $Z$-variable the number of exercised Caplets. Afterwards, the payoff of the AutoCap can easily be reconstructed with the help of the value of $Z$.

### 2.5.5 Example 2: Ladder Swap

**Definition 2.5.2.** *A Ladder Swap is a financial product that has at the end of each period $i$ the following payoff*

$$N_i(t_i - t_{i-1}) r_i$$

*for some given notional values $N_0, \ldots, N_m$ and where the path-dependent random variable $r_i$ is defined as follows:*

$$\begin{aligned} r_0 &:= s_0 \\ r_i &:= r_{i-1} + s_i - L(t_{i-1}, t_i), \quad i = 1, \ldots, m \end{aligned}$$

*for given rates $s_0, \ldots, s_m$ fixed in advance and where $L(t_{i-1}, t_i)$ denotes the Libor rate, set at time $t_{i-1}$ for a delivery at time $t_i$.*

The sum of all discounted (until time 0) payouts of the Ladder Swap can be written as

$$\sum_{i=1}^{m} N_i(t_i - t_{i-1}) \left( s_0 + \sum_{j=1}^{i} \left( s_j - L\left(t_{j-1}, t_j\right) \right) \right) D(0, t_i)$$

or equivalently as

$$\sum_{i=1}^{m} N_i(t_i - t_{i-1}) \left( \sum_{j=0}^{i} s_j - \sum_{j=1}^{i} L\left(t_{j-1}, t_j\right) \right) D(0, t_i)$$

where we can notice that at the end of each period $i$ there is an exchange of the accumulated until time $t_i$ fixed rate $\sum_{j=0}^{i} s_j$ for the accumulated until time $t_i$ Libor rates $\sum_{j=0}^{i} L\left(t_{j-1}, t_j\right)$. Notice for comparison that the discounted payoff of a "Receiver Forward Swap" is given by

$$\sum_{i=1}^{m} N(t_i - t_{i-1}) \left( S - L\left(t_{i-1}, t_i\right) \right) D(0, t_i)$$

where we have again an exchange of a fix for floating rate.

**The construction of the Z-variable is as follows:**

$$\begin{aligned} Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right) &:= s_0 \\ Z_{next}\left(t_i, X_1(t_i), X_2(t_i), Z_{old}\right) &:= Z_{old} + s_i - L(t_{i-1}, t_i; X_1(t_i), X_2(t_2)). \end{aligned}$$

which means that we choose the value of the $r_i$ to be our additional state variable $Z$.

### 2.5.6 Example 3: Ratchet Cap

A Ratchet option (or cliquet option) is a series of consecutive forward start options. The first is active immediately. The second becomes active when the first expires, etc. Each option is struck at-the-money when it becomes active. The effect of the entire instrument is an option that periodically "locks in" profits. Ratchet features can be incorporated into different structures. For example, there are Ratchet caps or Ratchet floors. In this thesis, we will deal only with pricing of a Ratchet cap.

**Definition 2.5.3.** *A Ratchet Cap is a financial product that has at the end of each period $i$ the following payoff*

$$N_i(t_i - t_{i-1})r_i$$

*for some given notional values $N_0, \ldots, N_n$. The path-dependent random variable $r_i$ is defined as*

$$\begin{aligned} r_0 &:= d_0 \\ r_i &:= min\left(L(t_{i-1}, t_i),\, r_{i-1} + d_i\right), \quad i = 1, \ldots, m \end{aligned}$$

*for given steps $d_0, \ldots, d_m$ fixed in advance and where $L(t_{i-1}, t_i)$ denotes the Libor rate set at time $t_{i-1}$ for a delivery at time $t_i$.*

**The construction of the Z-variable is as follows:**

$$\begin{aligned} Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right) &:= d_0 \\ Z_{next}\left(t_i, X_1(t_i), X_2(t_i), Z_{old}\right) &:= \min\left\{Z_{old} + d_i,\, L\left(t_{i-1}, t_i; X_1(t_i), X_2(t_2)\right)\right\}. \end{aligned}$$

where again we have chosen for a state variable the $r_i$-process.

### 2.5.7 Example 4: Target Redemption Note

The target redemption note is an index linked note that provides a guaranteed sum of coupons (target cap) with the possibility of early termination. In a typical structure, the first coupon payment is fixed. The subsequent coupons are calculated based on an inverse floating Libor / Euribor formula. Once the accumulated coupons have reached the pre-specified target cap, the note will be terminated with final payment of the par.

This kind of structured products enjoyed great popularity among Asian investors when the interest rates were at a low level in the early 2000's. The major source of risk stems from the uncertainty of the termination date of the note, which is the earlier of the pre-specified note's maturity date or the coupon payment date when the accumulated coupons reach the target sum.

Respectively, the pricing model of the Target Redemption Note is characterized by the uncertain knock-out feature, where the knock-out criterion depends on a path dependent state variable defined by the running accumulated coupon sum. In this sense, the Target Redemption Note resembles an Asian barrier option where the knock-out is also determined by a path dependent variable defined by the weighted sum of the underlying equity values.

In some simplified cases, such as a one-factor model for the interest rate and single coupon date, it is possible to obtain closed form valuation formula for the note value by decomposing its value into a sum of discount bonds and a European put option (see Chu and Kwok [22]).

**Definition 2.5.4.** *A Target Redemption Note is a financial product that has at the end of each period $i$ the following payoff*

$$N_i(t_i - t_{i-1})r_i$$

*for some given notional values $N_0, \ldots, N_m$ and where the path-dependent random variable $r_i$ is defined as follows:*

$$\begin{aligned} r_0 &:= s_0 \\ r_i &:= \begin{cases} s_i - L(t_{i-1}, t_i) & \text{if} \quad \sum_{j=0}^{i} r_j < F, \\ 0 & \text{else} \end{cases} \end{aligned}$$

*for given steps $s_0, \ldots, s_m$ fixed in advance, given target value $F \in \mathbb{R}$ and where $L(t_{i-1}, t_i)$ denotes the Libor rate set at time $t_{i-1}$ for a delivery at time $t_i$.*

Thus, the total discounted payoff of the target redemption note can be written as

$$\sum_{i=1}^{m} N_i(t_i - t_{i-1})r_i D(0, t_i).$$

**The construction of the Z-variable is as follows:**

If we denote $Y_i := s_i - L\left(t_{i-1}, t_i; X_1(t_i), X_2(t_i)\right)$ then

$$\begin{aligned} Z_{init}\left(t_0, X_1(t_0), X_2(t_0)\right) &:= s_0 \\ Z_{next}\left(t_i, X_1(t_i), X_2(t_i), Z_{old}\right) &:= Z_{old} + Y_i, \quad i = 1, \ldots, m \end{aligned}$$

which means that we choose for a variable $Z$ the running accumulated coupon sum. With the help of it, we can construct also the $r_i$-value as

$$\begin{aligned} r\left(t_0, X_1(t_0), X_2(t_0)\right) &:= s_0 \\ r\left(t_i, X_1(t_i), X_2(t_i)\right) &:= \begin{cases} Y_i & \text{if} \quad Z\left(t_i, X_1(t_i), X_2(t_i)\right) < F, \\ 0 & \text{else} \end{cases} \end{aligned}$$

and the discounted until time 0 payoff at time $t_i$ of the Target Redemption Note can then be written as

$$N_i(t_i - t_{i-1}) r\left(t_i, X_1(t_i), X_2(t_i)\right) D(0, t_i)$$

## 2.6 Numerical Results

In this section, we will price several financial instruments. We assume an initial term structure such that $P(0,t) = e^{-r_0 t}$ with $r_0 = 0.04$. First, we will price with the lattice a simple coupon bond in order to check whether the tree discounts correctly a simple claim. The coupon bond has a maturity of 4 years and yearly coupons of 1 unit. Its price at time 0 can be written as

$$\sum_{i=1}^{4} P(0, t_i), \quad where \quad t_i = 1, 2, 3, 4.$$

Assuming a flat initial term structure, its price can be estimated to be 3.62297.

Next, we price a European bond option with maturity $t_0 > 0$ on a bond with maturity $t_1$ and having a payoff at maturity $t_0$ of $\left(P(t_0, t_1) - K\right)^+, \quad t_1 > t_0$ for some strike $K$. We set $t_0$ to 1 year, $t_1$ to 5 years and the strike $K = 0.04$. We compare the result with the closed-form solution in Theorem 2.2.1.

Further, we price a Caplet resetting at time $t_0$ with maturity $t_1 > t_0$ and having a payoff at time $t_1$ of $\left(L(t_0, t_1) - K\right)^+ (t_1 - t_0), \quad t_1 > t_0$ for some strike $K$. We set $t_0$ to 1 year, $t_1$ to 5 years and the strike $K = 0.04$. We compare the result with the closed-form solution in Theorem 2.2.2

Finally, we price all types of path-dependent options presented in the previous section where we set their maturity to be 5 years, $m = 5$, $t_i - t_{i-1} = 1$ for all $i = 1, \ldots, 5$ and for

the notional values we set $N_0 = N_1 = 0$ and $N_1 = \ldots = N_5 = 1$ which mean that their first payout is at time $T_2$ .

In addition, the option specific parameters are set for now as follows:

For the AutoCap:

$$\begin{aligned} \gamma &= 3 \\ K_0 &= K_1 = 0 \\ K_2 &= \ldots = K_5 = 0.04. \end{aligned}$$

For the Ladder Swap:

$$\begin{aligned} s_0 &= s_1 = 0 \\ s_2 &= \ldots = s_5 = 0.04. \end{aligned}$$

For the Ratchet Cap:

$$\begin{aligned} d_0 &= d_1 = d_2 = 0 \\ d_3 &= \ldots = d_5 = 0.01. \end{aligned}$$

For the Target Redemption Note:

$$\begin{aligned} s_0 &= s_1 = 0 \\ s_2 &= \ldots = s_5 = 0.04 \\ F &= 0.10 \end{aligned}$$

For comparison purposes we shall perform Monte Carlo simulations with 500 time steps and 200000 paths for both processes.

The errors are cited in basic points (i.e. multiplied with 10000) and thus are defined as *(MC(Analytical)Price - Tree Price)*10000*. We denote by $N$ the number of time steps in the tree construction and by $Z$ the number of points in the discretization of the additional variable $Z$. We change the volatilities, the mean-reversion rates and the correlation of the base processes.

**Example 1 : Very Small Volatility Parameters:**

| **Coupon Bond with** $\rho = -0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Analytical | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 3.62297001 | 3.6229710 | -0.00989 | 3.6229713 | -0.01289 |
| 100 | 0.04 | 2 | 3.62297001 | 3.6229705 | -0.00489 | 3.6229707 | -0.00689 |
| 150 | 0.0266 | 2 | 3.62297001 | 3.6229703 | -0.00289 | 3.6229705 | -0.00489 |
| 200 | 0.02 | 2 | 3.62297001 | 3.6229702 | -0.00238 | 3.6229703 | -0.00289 |

| **Coupon Bond with** $\rho = 0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Analytical | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 3.62297001 | 3.62297317 | -0.0316 | 3.62297315 | -0.0315 |
| 100 | 0.04 | 2 | 3.62297001 | 3.62297159 | -0.0158 | 3.62297158 | -0.0157 |
| 150 | 0.0266 | 2 | 3.62297001 | 3.62297106 | -0.0105 | 3.62297105 | -0.0104 |
| 200 | 0.02 | 2 | 3.62297001 | 3.62297079 | -0.0078 | 3.62297079 | -0.0078 |

| **European Call Bond Option with** $\rho = -0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Analytical | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 0.780299176 | 0.780299196 | -0.000200 | 0.780299222 | -0.00046 |
| 100 | 0.04 | 2 | 0.780299176 | 0.780299185 | -0.000090 | 0.780299199 | -0.00023 |
| 150 | 0.0266 | 2 | 0.780299176 | 0.780299182 | -0.000060 | 0.780299191 | -0.00015 |
| 200 | 0.02 | 2 | 0.780299176 | 0.780299181 | -0.000048 | 0.780299188 | -0.00012 |

| **Caplet with** $\rho = -0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Analytical | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 1.1082836e-002 | 1.108239306e-002 | 0.00443 | 1.108275882e-002 | 0.00078 |
| 100 | 0.04 | 2 | 1.1082836e-002 | 1.108268437e-002 | 0.00152 | 1.108279144e-002 | 0.00045 |
| 150 | 0.0266 | 2 | 1.1082836e-002 | 1.108271527e-002 | 0.00121 | 1.108278169e-002 | 0.00055 |
| 200 | 0.02 | 2 | 1.1082836e-002 | 1.108274682e-002 | 0.00090 | 1.108280284e-002 | 0.00034 |

| **Caplet with** $\rho = 0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Analytical | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 1.1213899e-002 | 1.121669148e-002 | -0.02792 | 1.121430021e-002 | -0.00401 |
| 100 | 0.04 | 2 | 1.1213899e-002 | 1.121540741e-002 | -0.01508 | 1.121413883e-002 | -0.00239 |
| 150 | 0.0266 | 2 | 1.1213899e-002 | 1.121344029e-002 | 0.00459 | 1.121401645e-002 | -0.00117 |
| 200 | 0.02 | 2 | 1.1213899e-002 | 1.121456900e-002 | - 0.0067 | 1.121404101e-002 | -0.00142 |

| **AutoCap (Path-dependent) with $\rho = -0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 3.90479e-003 | 3.8411939e-003 | 0.64 | 3.9001881e-003 | 0.0460 |
| 100 | 0.04 | 2 | 3.90479e-003 | 3.8971497e-003 | 0.08 | 3.9010158e-003 | 0.0374 |
| 150 | 0.0266 | 2 | 3.90479e-003 | 3.9048011e.003 | -0.00011 | 3.8961161e-003 | 0.085 |

| **AutoCap (Path-dependent) with $\rho = 0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 5.4499854e-003 | 5.5561257e-003 | -1.0612 | 5.46149314e-003 | -0.1150 |
| 100 | 0.04 | 2 | 5.4499854e-003 | 5.4336662e-003 | 0.1632 | 5.43083113e-003 | 0.1915 |
| 150 | 0.0266 | 2 | 5.4499854e-003 | 5.4089610e.003 | 0.4102 | 5.43404719e-003 | 0.1594 |

| **Target Redemption Note (Path-dependent) with $\rho = -0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | -2.837077e-003 | -2.82245635e-003 | -0.1462 | -2.82207981e-003 | -0.1499 |
| 100 | 0.04 | 2 | -2.837077e-003 | -2.82250247e-003 | -0.1457 | -2.82231464e-003 | -0.1476 |
| 150 | 0.0266 | 2 | -2.837077e-003 | -2.82230841e-003 | -0.1477 | -2.83252200e-003 | -0.0456 |

| **Target Redemption Note (Path-dependent) with $\rho = 0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | -2.8665385e-003 | -2.82232619e-003 | -0.4421 | -2.82227836e-003 | -0.4426 |
| 100 | 0.04 | 2 | -2.8665385e-003 | -2.82338856e-003 | -0.4315 | -2.82282161e-003 | -0.4372 |
| 150 | 0.0266 | 2 | -2.8665385e-003 | -2.92704178e-003 | 0.6050 | -2.893116347e-003 | 0.2658 |

| **Ratchet Cap (Path-dependent) with $\rho = -0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 5.047297e-002 | 5.0474014e-002 | -0.01 | 5.04744031-002 | -0.01 |
| 100 | 0.04 | 2 | 5.047297e-002 | 5.0474006e-002 | -0.01 | 5.0474014e-002 | -0.01 |
| 150 | 0.0266 | 2 | 5.047297e-002 | 5.0474006e-002 | -0.01 | 5.0474012e-002 | -0.01 |

| **Ratchet Cap (Path-dependent) with $\rho = 0.7$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 5.0470901e-002 | 5.04740498e-002 | -0.03148 | 5.04740507e-002 | -0.03149 |
| 100 | 0.04 | 2 | 5.0470901e-002 | 5.04740103e-002 | -0.03119 | 5.04740207e-002 | -0.03119 |
| 150 | 0.0266 | 2 | 5.0470901e-002 | 5.04740182e-002 | -0.03117 | 5.04740185e-002 | -0.03117 |

| **Ladder Swap Option (Path-dependent) with** $\rho = -0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.04 | 2 | 1.351349e-001 | 1.3515803e-001 | -0.2313 | 1.3515881e-001 | -0.2391 |
| 100 | 0.08 | 2 | 1.351349e-001 | 1.3515797e-001 | -0.2307 | 1.3515835e-001 | -0.2345 |
| 150 | 0.0266 | 2 | 1.351349e-001 | 1.3515821e-001 | -0.2331 | 1.3515847e-001 | -0.2357 |

| **Ladder Swap Option (Path-dependent) with** $\rho = 0.7$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.2\%, \sigma_2 = 0.3\%, \quad \kappa_1 = 90\%, \kappa_2 = 30\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.04 | 2 | 1.3507154e-001 | 1.3517148e-001 | -0.999 | 1.3517159e-001 | -1.000 |
| 100 | 0.08 | 2 | 1.3507154e-001 | 1.3517102e-001 | -0.994 | 1.3517107e-001 | -0.995 |
| 150 | 0.0266 | 2 | 1.3507154e-001 | 1.3517151e-001 | -0.999 | 1.3517155e-001 | -1.000 |

And for some other parameter values

| **AutoCap (Path-dependent) with** $\rho = 0.98$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.5\%, \sigma_2 = 0.4\%, \quad \kappa_1 = 80\%, \kappa_2 = 60\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 7.7397869e-003 | 7.9309435e-003 | -1.91 | 7.8554489e-003 | -1.14 |
| 100 | 0.04 | 2 | 7.7397869e-003 | 7.759374e-003 | -0.20 | 7.7681009e-002 | -0.28 |
| 150 | 0.0266 | 2 | 7.7397869e-003 | 7.8000845e-003 | -0.60 | 7.7833874e-002 | -0.43 |

| **Auto Cap (Path-dependent) with** $\rho = -0.98$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 0.5\%, \sigma_2 = 0.4\%, \quad \kappa_1 = 80\%, \kappa_2 = 60\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 2.3047699e-003 | 2.2990304e-003 | 0.06 | 2.3392384e-003 | -0.32 |
| 100 | 0.04 | 2 | 2.3047699e-003 | 2.3022782e-003 | 0.02 | 2.3255449e-003 | -0.21 |
| 150 | 0.0266 | 2 | 2.3047699e-003 | 2.3004871e-003 | 0.0428 | 2.3064832e-003e | -0.017 |

The results exhibit that both methods perform almost equally well for very small volatility parameters with errors in most of the cases smaller than one base point. We remark here that for the upper parameter values, the probabilities in the quadrinomial tree are well-defined, i.e. in $[0, 1]$ and the differences between the two tree constructions are mostly due to numerical errors. Further, we notice only a negligible difference in the performance of the methods due to a positive or negative correlation between the basis factors $X_1$ and $X_2$.

**Example 2: Higher Volatility Parameters:**

On the other side, it can also be observed that for higher volatility parameters the performance of the quadrinomial tree is worse than the rotated one. This effect is observed exclusively in the

case of negative correlation between the base factor. To demonstrate it, we give as an example the Target Redemption Note for $\sigma_1 = 7\%$, $\sigma_2 = 5\%, \kappa_1 = 40\%, \kappa_2 = 50\%, \rho = -0.9$ where we obtain the following results:

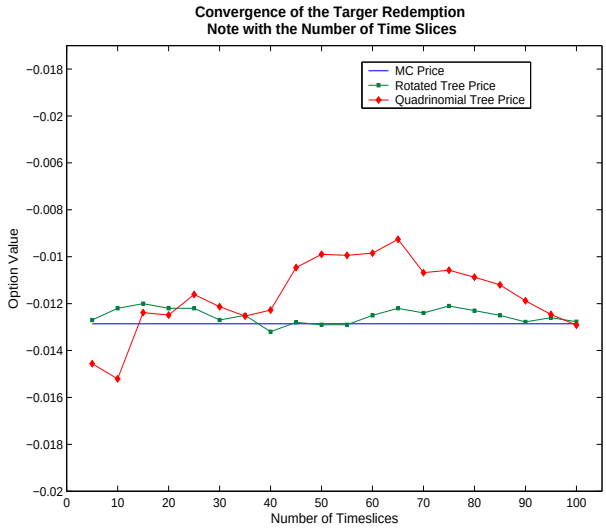
| **Target Redemption Note (Path-dependent) with $\rho = -0.9$** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 7\%, \sigma_2 = 5\%, \quad \kappa_1 = 40\%, \kappa_2 = 50\%$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 25 | 0.16 | 2 | -1.2859698e-002 | -1.21752837e-002 | 6.844 | -1.13128957e-002 | 12.468 |
| 50 | 0.08 | 2 | -1.2859698e-002 | -1.29474606e-002 | -0.878 | -9.89736977e-003 | 29.623 |
| 75 | 0.053 | 2 | -1.2859698e-002 | -1.20955741e-002 | 7.641 | -1.05767920e-002 | 22.829 |
| 100 | 0.04 | 2 | -1.2859698e-002 | -1.27701130e-002 | 0.893 | -1.29176814e-002 | 0.579 |

which are also plotted in Figure 2.5.

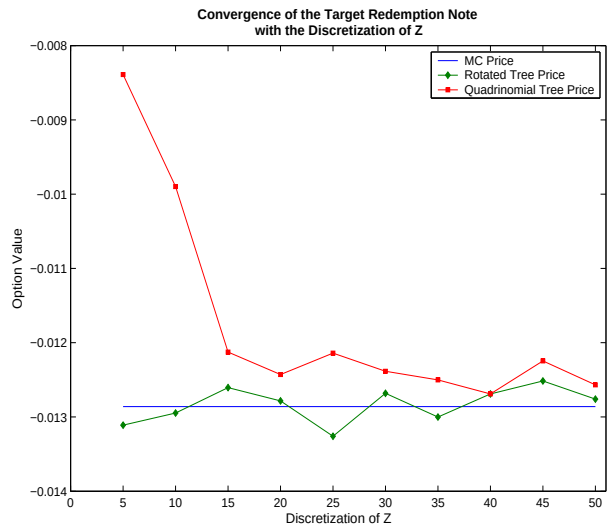
Notice that the results of the Target Redemption Note are very sensitive to the values of the extra variable $Z$ and its discretization and (especially in the case of the quadrinomial tree) converge to the MC price with the increase of its refinement (Figure 2.5b)). Additionally, in Figure 2.7 a) and b) we have plotted a section of the two-dimensional tree for both tree construction methods, taken at the 25-th time slice with $n = 50$ (in the pricing of a Target Redemption Note with the upper parameters). For the quadrinomial tree, it is easy to see on Figure 2.7 a) that there is a strong oscillation of the option price for small values of $X_1$ and $X_2$ due to unbounded probabilities. Whereas the solution surface of the rotated tree remains well-behaved.

In addition, we have plotted in Figure 2.7 c) and d) the projection of the respective solution surfaces in 2.7 a) and b) on the $X_1/X_2$ plane. Notice that Figure 2.7 d) clearly reflects the idea of rotation of the base processes. Whereas the grid of the quadrinomial tree is very regular (equidistant). Note that due to the small mean-reversion parameters there is no "contraction" of the quadrinomial tree (i.e. we have always $I = J = 0$).

Further, we have plotted a projection of the payoff from Figure 2.7 b) on the plane of the new independent variables $Y_1$ and $Y_2$ in Figure 2.8. We can conclude that due to the resulted scaling, the mean-reversion plays a role and as a result there are no points in the grid for very high values of $Y_1$ and very low of $Y_2$ and vice versa. This causes on its turn in the $X_1/X_2$ plane projection of the Rotated tree "fatter" long sides for middle value of $X_1$ and $X_2$ and "thinner" long sides for the rest of their values.

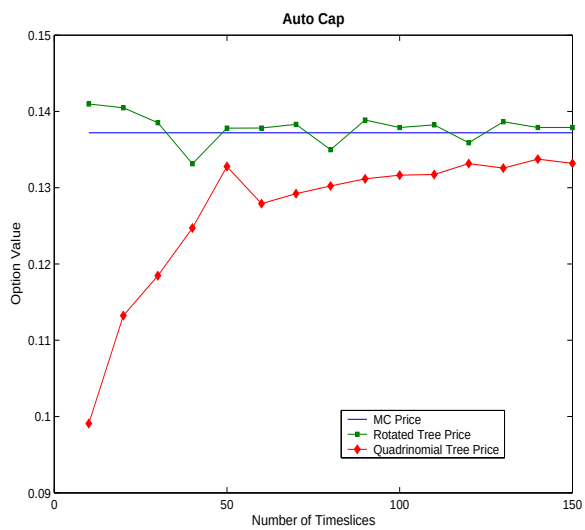



(a) Convergence with the increase of the number of time slices. Discretization of $Z$ is 10.

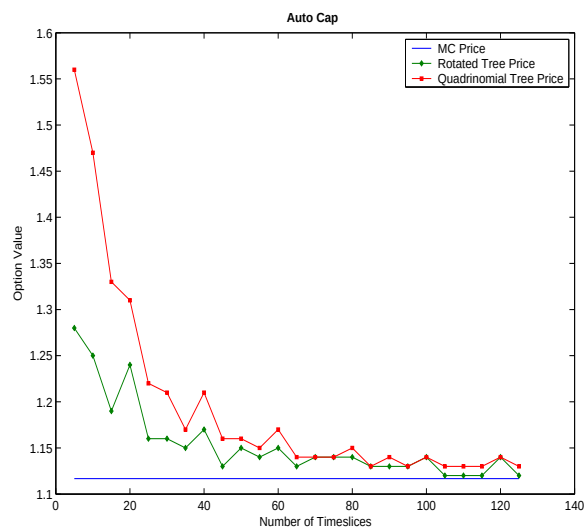



(b) Convergence with the increase of the discretization of $Z$. Number of time slices is 50.

Figure 2.5: Target Redemption Note (approximated) prices for $\sigma_1 = 7\%$, $\sigma_2 = 5\%, \kappa_1 = 40\%, \kappa_2 = 50\%, \rho = -0.9$.




(a) $\sigma_1 = \sigma_2 = 30\%, \kappa_1 = 70\%, \kappa_2 = 80\%, \rho = -0.8$..




(b) $\sigma_1 = 70\%, \sigma_2 = 40\%, \kappa_1 = 0.5\%, \kappa_2 = 80\%, \rho = -0.8$.

Figure 2.6: AutoCap (approximated) prices. Discretization of $Z$=2.

As an additional example, we price the AutoCap option for the parameter cases $\sigma_1 = \sigma_2 = 30\%, \kappa_1 = 70\%, \kappa_2 = 80\%, \rho = -0.8$ and $\sigma_1 = 70\%, \sigma_2 = 40\%, \kappa_1 = 0.5\%, \kappa_2 = 80\%, \rho = -0.8$. The results are plotted in Figure 2.6. The AutoCap tree results (in both cases) show no sensitivity to the refinement of the $Z$-discretization for which reason we have set its refinement to 2. In these two additional examples, we observe again a smaller error in the rotated tree compared to the quadrinomial one.

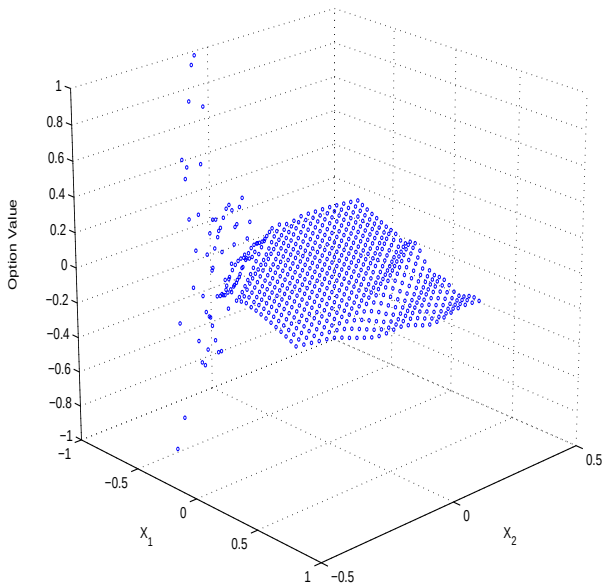
**Example 3: One Extreme Case of High Volatilities and High Mean Reversion Parameters**

We present also a very extreme case of high volatilities and high mean reversion parameters where the quadrinomial tree strongly deteriorates while the rotated one still delivers a good performance

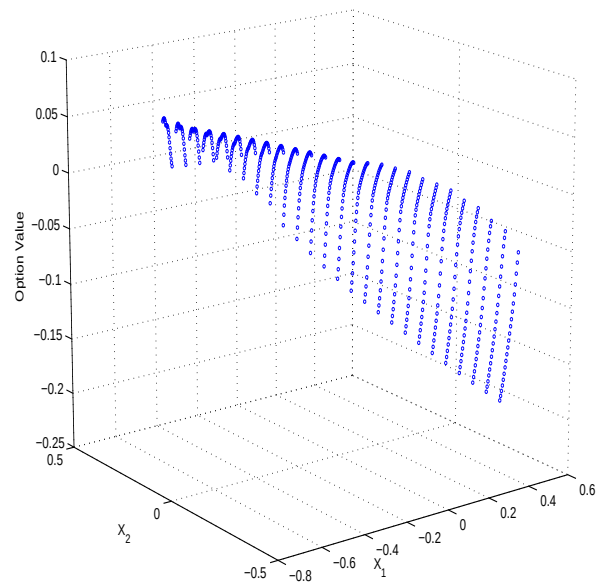
| **Auto Cap (Path-dependent)** | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\sigma_1 = 30\%, \sigma_2 = 30\%, \quad \kappa_1 = 99\%, \kappa_2 = 99\%, \rho = -0.99$ | | | | | | | |
| N | dt | Z | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 50 | 0.08 | 2 | 2.554497e-002 | 2.5682228e-002 | -1.38 | 9.6793597e-003 | 158.67 |
| 100 | 0.04 | 2 | 2.554497e-002 | 2.522369e-002 | 3.21 | 1.2283712e-003 | 132.61 |
| 150 | 0.0266 | 2 | 2.554497e-002 | 2.539574e-002 | 1.49 | 1.3981513e-003 | 115.63 |

**Example 4: Low Mean Reversion, Varying Volatility and Correlation**

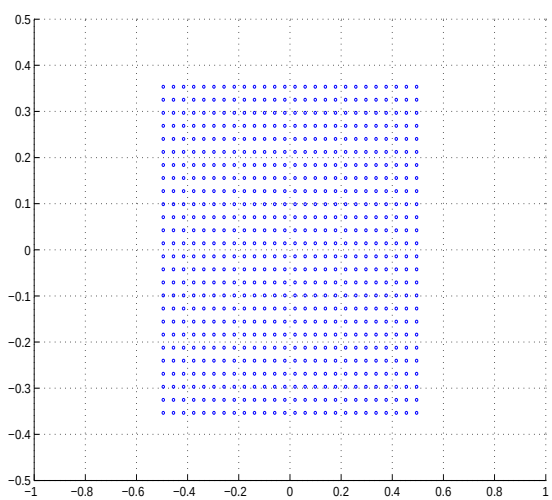
The aim of the following examples is to examine the behavior of both tree constructions for some cases of high to very high volatility parameters but low mean-reversion ones. In addition we switch between two values of the correlation parameter namely $-0.9$ and $0.9$. The errors are now cited in real terms and not in base points. The errors of both tree constructions in pricing a Caplet (with the same specifications, i.e. $t_1 = 5$, $t_0 = 1$, strike $K = 0.04$) are plotted in Figure 2.9 and Figure 2.10.



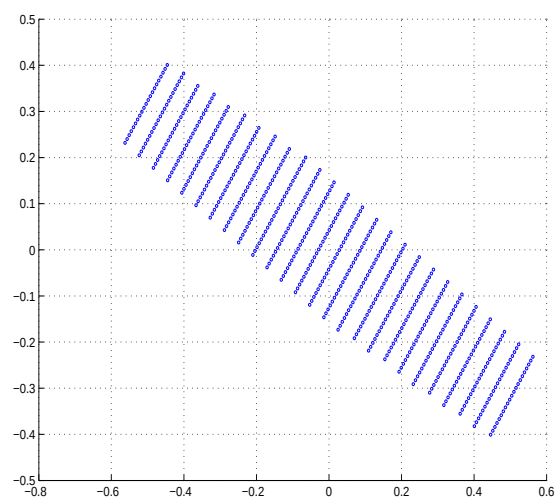

(a) Payoff of a Target Redemption Note, Quadrinomial Tree




(b) Payoff of a Target Redemption Note, Rotated Tree



(c) Projection of the Payoff (on the $X_1/X_2$ plane) of a Target Redemption Note, Quadrinomial Tree



(d) Projection of the Payoff (on the $X_1/X_2$ plane) of a Target Redemption Note, Rotated Grid

Figure 2.7: Payoff and Projection of the Payoff (on the $X_1/X_2$ plane) of a Target Redemption Note for the case $\sigma_1 = 7\%$, $\sigma_2 = 5\%, \kappa_1 = 40\%, \kappa_2 = 50\%, \rho = -0.9$ taken at the 25-th timeslice with total number of timeslices 50.

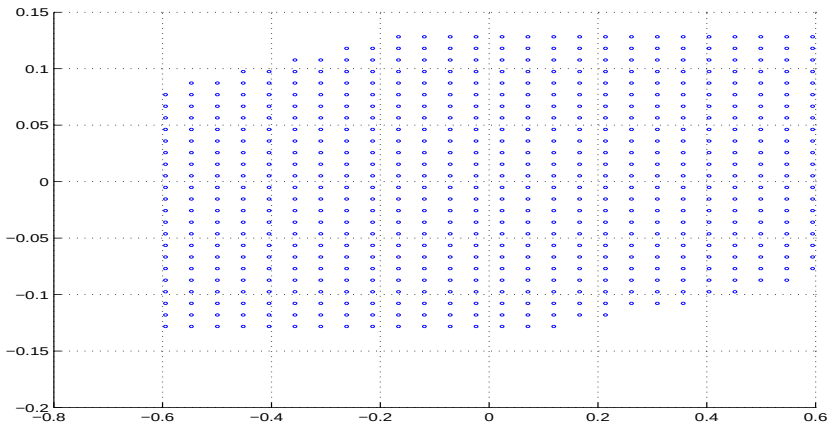



Figure 2.8: Projection of the Payoff of a Target Redemption Note on the $Y_1/Y_2$ plane for the case $\sigma_1 = 7\%$, $\sigma_2 = 5\%, \kappa_1 = 40\%, \kappa_2 = 50\%, \rho = -0.9$ taken at the 25-th timeslice with total number of timeslices 50.

Further, we want to investigate the observed effects for the pricing of two path dependent options - the Target Redemption Note and the Ladder Swap. For that purpose, for the next examples we have increased the maturity of the options to 7 years and have changed some of their parameters such that we have

$$\begin{aligned} s_0 &= s_1 = 0 \\ s_2 &= 0.02,\ s_3 = s_4 = 0.03,\ s_5 = s_6 = 0.04,\ s_7 = 0.05. \end{aligned}$$

where we have also set for the Target Redemption Note $F = 0.12$. We obtain:

| **Ladder Swap with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 90\%, \sigma_2 = 90\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | 2.86763e+00 | 6.400748e-01 | 2.2275552 | -8.428768e+11 | 842876792002.868000 |
| 20 | 2.86763e+00 | 1.489964e+00 | 1.3776664 | -3.422984e+08 | 342298354.167630 |
| 30 | 2.86763e+00 | 1.796071e+00 | 1.0715593 | 6.016619e+00 | -3.148989 |
| 40 | 2.86763e+00 | 2.283599e+00 | 0.5840313 | -7.076765e-01 | 3.575307 |
| 50 | 2.86763e+00 | 2.501697e+00 | 0.3659333 | -4.628053e-01 | 3.330435 |
| 60 | 2.86763e+00 | 2.543226e+00 | 0.3244041 | -2.832525e-01 | 3.150883 |
| 70 | 2.86763e+00 | 2.748769e+00 | 0.1188606 | 1.402770e-02 | 2.853602 |
| 80 | 2.86763e+00 | 2.844098e+00 | 0.0235318 | 2.388298e-01 | 2.628800 |
| 90 | 2.86763e+00 | 2.841081e+00 | 0.0265491 | 3.911545e-01 | 2.476475 |
| 100 | 2.86763e+00 | 2.961832e+00 | -0.0942025 | 6.281544e-01 | 2.239476 |

| **Ladder Swap with** $\rho = 0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 90\%, \sigma_2 = 90\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -7.84978e-01 | -1.471202e+00 | 0.6862248 | -1.157618e+00 | 0.372640 |
| 20 | -7.84978e-01 | -8.700793e-01 | 0.0851017 | -8.181388e-01 | 0.033161 |
| 30 | -7.84978e-01 | -8.490842e-01 | 0.0641066 | -8.246925e-01 | 0.039715 |
| 40 | -7.84978e-01 | -7.976934e-01 | 0.0127158 | -7.840143e-01 | -0.000963 |
| 50 | -7.84978e-01 | -7.626637e-01 | -0.0223139 | -7.529692e-01 | -0.032008 |
| 60 | -7.84978e-01 | -8.113270e-01 | 0.0263494 | -8.041798e-01 | 0.019202 |
| 70 | -7.84978e-01 | -7.719267e-01 | -0.0130509 | -7.688833e-01 | -0.016094 |
| 80 | -7.84978e-01 | -7.383356e-01 | -0.0466419 | -7.376970e-01 | -0.047281 |
| 90 | -7.84978e-01 | -8.023587e-01 | 0.0173811 | -8.015865e-01 | 0.016609 |
| 100 | -7.84978e-01 | -7.709323e-01 | -0.0140452 | -7.706375e-01 | -0.014340 |

| **Ladder Swap with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 50\%, \sigma_2 = 50\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | 1.50316e+00 | 1.217355e+00 | 0.285809 | -6.006052e-01 | 2.1037692 |
| 20 | 1.50316e+00 | 1.271119e+00 | 0.232045 | 1.899037e-01 | 1.3132603 |
| 30 | 1.50316e+00 | 1.244211e+00 | 0.258953 | 4.739671e-01 | 1.0291969 |
| 40 | 1.50316e+00 | 1.310946e+00 | 0.192218 | 7.296757e-01 | 0.7734883 |
| 50 | 1.50316e+00 | 1.320741e+00 | 0.182423 | 8.546492e-01 | 0.6485148 |
| 60 | 1.50316e+00 | 1.299980e+00 | 0.203184 | 9.083278e-01 | 0.5948362 |
| 70 | 1.50316e+00 | 1.330615e+00 | 0.172549 | 9.969131e-01 | 0.5062509 |
| 80 | 1.50316e+00 | 1.334704e+00 | 0.168460 | 1.043537e+00 | 0.4596274 |
| 90 | 1.50316e+00 | 1.319388e+00 | 0.183776 | 1.059590e+00 | 0.4435740 |
| 100 | 1.50316e+00 | 1.338949e+00 | 0.164215 | 1.106336e+00 | 0.3968281 |

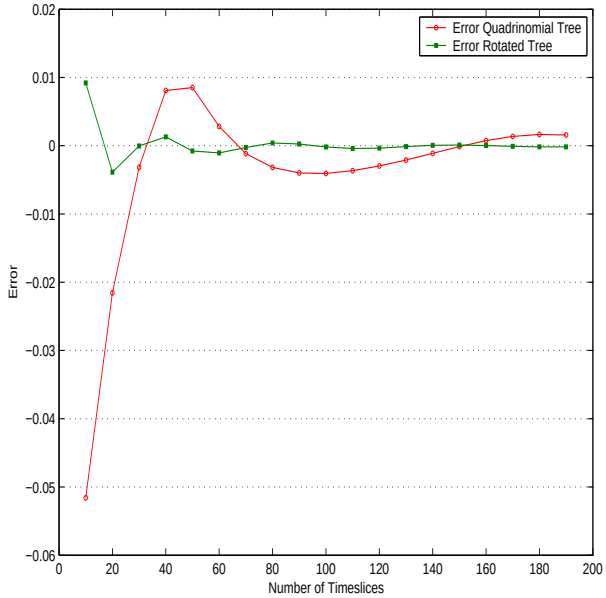
| **Ladder Swap with** $\rho = 0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 50\%, \sigma_2 = 50\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -0.10373173 | -1.445031e+00 | 1.341299 | -1.362594e+00 | 1.2588625 |
| 20 | -0.10373173 | -8.922175e-01 | 0.788486 | -8.855572e-01 | 0.7818255 |
| 30 | -0.10373173 | -9.449847e-01 | 0.841253 | -9.461692e-01 | 0.8424374 |
| 40 | -0.10373173 | -7.110663e-01 | 0.607335 | -7.190635e-01 | 0.6153318 |
| 50 | -0.10373173 | -5.506210e-01 | 0.446889 | -5.581335e-01 | 0.4544018 |
| 60 | -0.10373173 | -5.400103e-01 | 0.436279 | -5.481474e-01 | 0.4444157 |
| 70 | -0.10373173 | -3.325943e-01 | 0.228863 | -3.435943e-01 | 0.2398626 |
| 80 | -0.10373173 | -2.098351e-01 | 0.106103 | -2.215388e-01 | 0.1178071 |
| 90 | -0.10373173 | -1.931371e-01 | 0.089405 | -2.048920e-01 | 0.1011603 |
| 100 | -0.10373173 | 1.133653e-02 | -0.115068 | -3.109587e-03 | -0.1006221 |

| **Ladder Swap with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 30\%, \sigma_2 = 30\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | 6.93384e-01 | 5.7845165e-01 | 0.1149327 | 4.0053174e-01 | 0.29285261 |
| 20 | 6.93384e-01 | 5.5491769e-01 | 0.1384667 | 4.4936433e-01 | 0.24402002 |
| 30 | 6.93384e-01 | 5.3235637e-01 | 0.1610280 | 4.5559076e-01 | 0.23779359 |
| 40 | 6.93384e-01 | 5.4647404e-01 | 0.1469103 | 4.8895455e-01 | 0.20442980 |
| 50 | 6.93384e-01 | 5.4430036e-01 | 0.1490840 | 4.9786969e-01 | 0.19551466 |
| 60 | 6.93384e-01 | 5.3460807e-01 | 0.1587763 | 4.9492574e-01 | 0.19845861 |
| 70 | 6.93384e-01 | 5.4243168e-01 | 0.1509527 | 5.0861261e-01 | 0.18477174 |
| 80 | 6.93384e-01 | 5.4165415e-01 | 0.1517302 | 5.1201308e-01 | 0.18137127 |
| 90 | 6.93384e-01 | 5.3544992e-01 | 0.1579344 | 5.0872533e-01 | 0.18465902 |
| 100 | 6.93384e-01 | 5.4085501e-01 | 0.1525293 | 5.1692471e-01 | 0.17645964 |

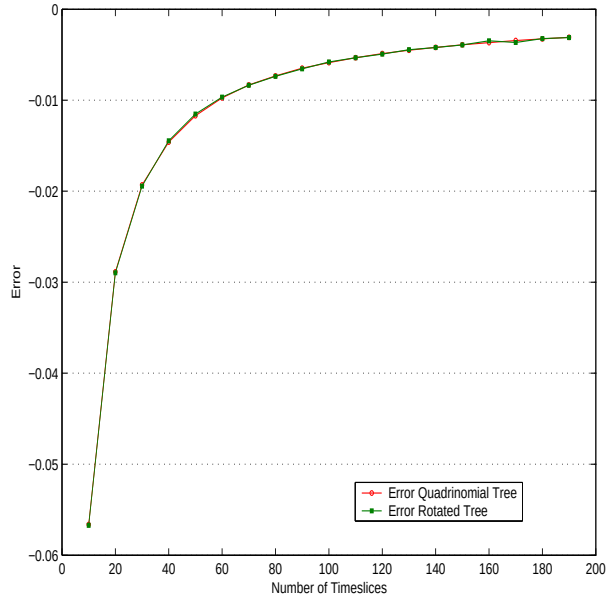
| **Ladder Swap with** $\rho = 0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 30\%, \sigma_2 = 30\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=5 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | 2.5563909 | -6.295121e-01 | 3.1859030 | -7.020275e-01 | 3.25841838 |
| 20 | 2.5563909 | - 5.208313e-02 | 2.6084740 | -8.983768e-02 | 2.64622858 |
| 30 | 2.5563909 | 2.528229e-01 | 2.3035680 | 2.276696e-01 | 2.32872130 |
| 40 | 2.5563909 | 1.002572e+00 | 1.5538191 | 9.731592e-01 | 1.58323170 |
| 50 | 2.5563909 | 1.442116e+00 | 1.1142749 | 1.411112e+00 | 1.14527843 |
| 60 | 2.5563909 | 1.623284e+00 | 0.9331071 | 1.593369e+00 | 0.96302192 |
| 70 | 2.5563909 | 2.100554e+00 | 0.4558374 | 2.070354e+00 | 0.48603660 |
| 80 | 2.5563909 | 2.381355e+00 | 0.1750357 | 2.351806e+00 | 0.20458454 |
| 90 | 2.5563909 | 2.469916e+00 | 0.0864748 | 2.441972e+00 | 0.11441940 |
| 100 | 2.5563909 | 2.794737e+00 | -0.2383462 | 2.767436e+00 | -0.21104542 |

Notice firstly that both tree constructions show similar results in the case of positive correlation parameter with the quadrinomial one converging with a slightly faster rate. On the other side, for volatility parameters starting already from around 30% and correlation $-0.9$, the Rotated tree construction exhibits better performance than the standard Quadrinomial one. The difference between the trees increases fast with the increase of the volatility. For volatility level of 90%, the Quadrinomial tree shows strongly inconsistent results whereas the Rotated one although also clearly deteriorating its performance with the increase of volatility has still a fast convergence.

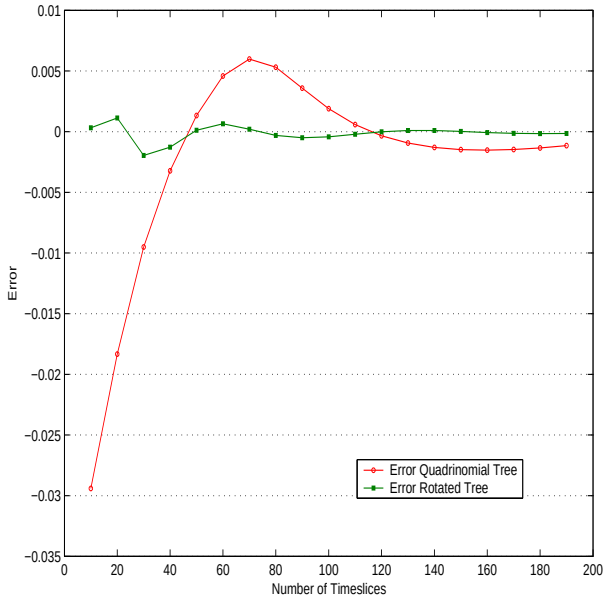
In Figure 2.11, we have plotted a section of both trees at time 25-th time slice for $n = 50$ for the price of the Ladder Swap. Notice that in the quadrinomial lattice, the highest option value is given for extremely low values of $X_1$ and $X_2$ and has an order of $10^{12}$. This means that the effect of the not properly defined probabilities cannot be spread throughout the option value surface



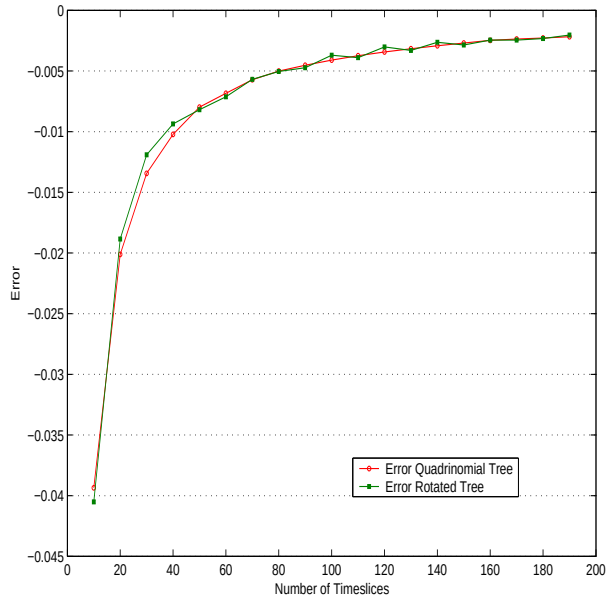

(a) $\sigma_1 = 90\%$, $\sigma_2 = 90\%, \rho = -0.9$




(b) $\sigma_1 = 90\%$, $\sigma_2 = 90\%, \rho = 0.9$

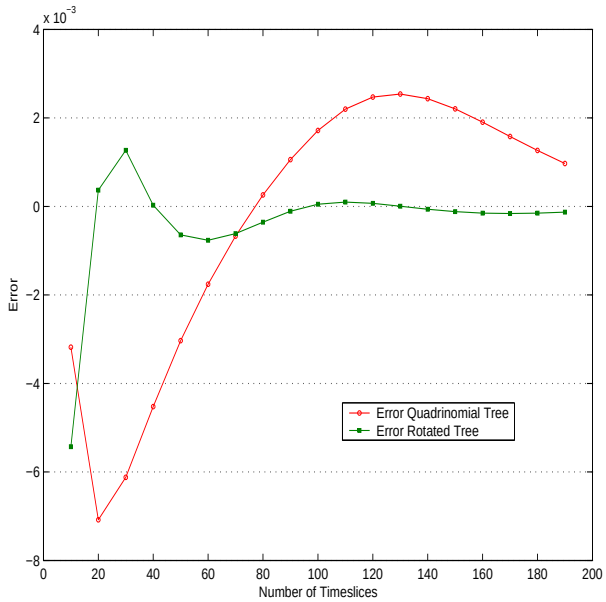



(c) $\sigma_1 = 70\%$, $\sigma_2 = 70\%, \rho = -0.9$

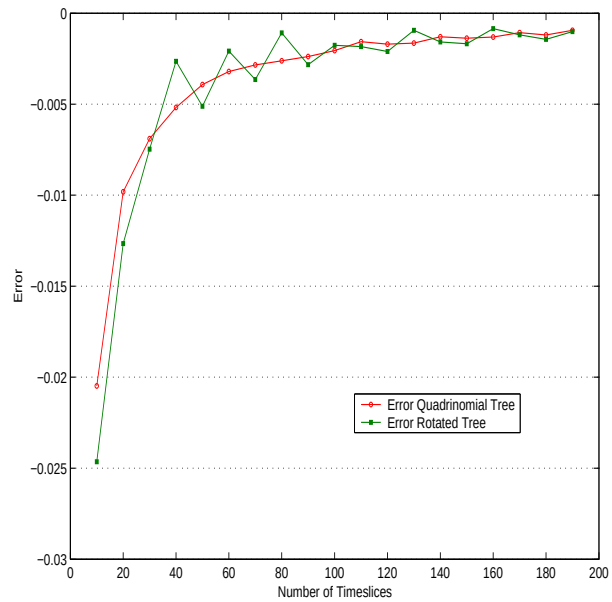



(d) $\sigma_1 = 70\%$, $\sigma_2 = 70\%, \rho = 0.9$

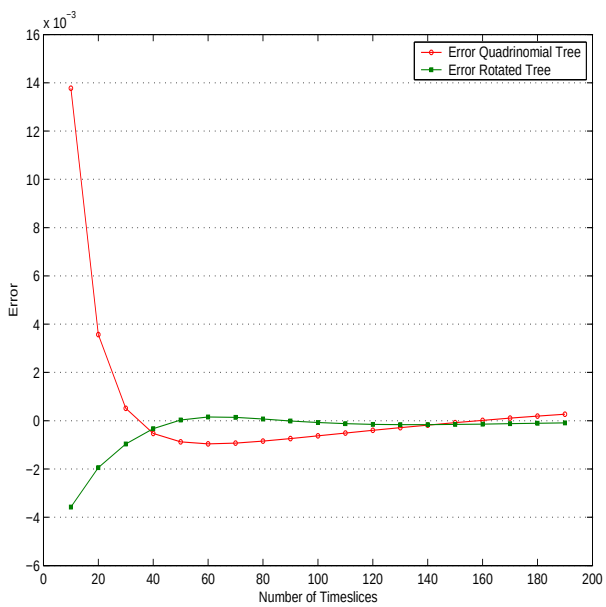
Figure 2.9: Errors of both tree method in the pricing a Caplet for the case $\kappa_1 = 7\%, \kappa_2 = 8\%$.



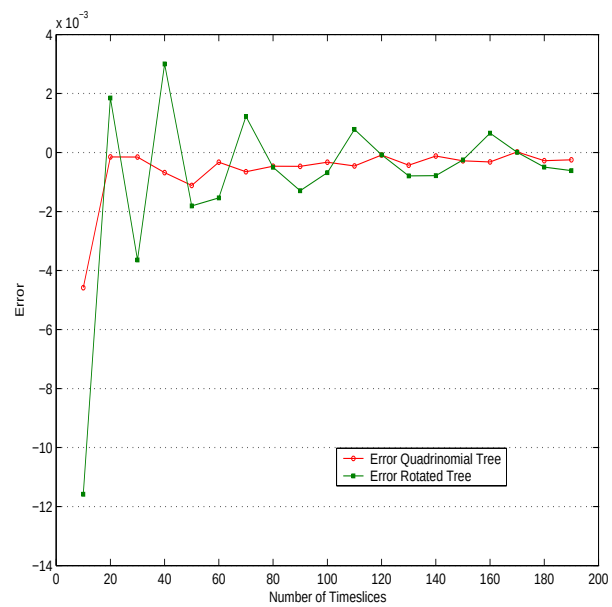

(a) $\sigma_1 = 50\%$, $\sigma_2 = 50\%$, $\rho = -0.9$




(b) $\sigma_1 = 50\%$, $\sigma_2 = 50\%$, $\rho = 0.9$

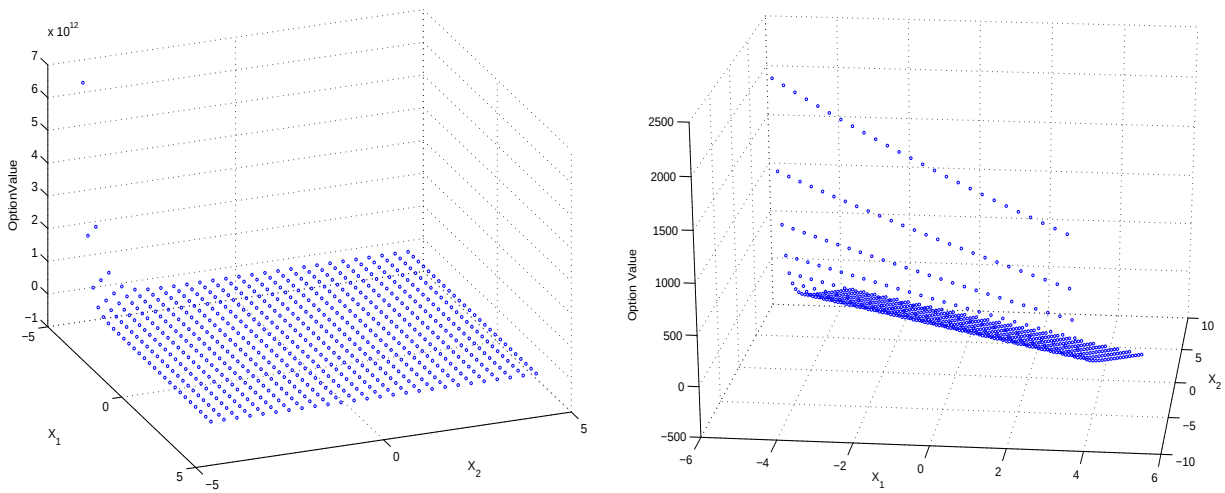



(c) $\sigma_1 = 30\%$, $\sigma_2 = 30\%$, $\rho = -0.9$




(d) $\sigma_1 = 30\%$, $\sigma_2 = 30\%$, $\rho = 0.9$

Figure 2.10: Errors of both tree method in the pricing a Caplet for the case $\kappa_1 = 7\%$, $\kappa_2 = 8\%$.




(a) Quadrinomial Tree (b) Rotated Tree

Figure 2.11: Payoff of a Ladder Swap for the case $\sigma_1 = 50\%$, $\sigma_2 = 50\%, \kappa_1 = 7\%, \kappa_2 = 8\%, \rho = -0.9$ taken at the 25-th timeslice with total number of time slices 50.

(and be equalized) but will be scaled with values around $10^{12}$, leading to a strong misprizing of the option value. On the other side, the option value surface in the Rotated tree is again well-behaved and shows no clear outliers.

The same effects can also be observed for the upper parameters in the pricing of the Target Redemption Note. We present only the results for $\rho = -0.9$, since again for positive correlation we have almost the same results for both tree construction methods:

| **Target Redemption Note with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 90\%, \sigma_2 = 90\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=15 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -1.0071863 | -1.221873e+00 | 0.214687 | -1.143659e+04 | 11435.582124 |
| 20 | -1.0071863 | -1.077609e+00 | 0.070423 | 1.519367e+06 | -1519368.055186 |
| 30 | -1.0071863 | -1.095081e+00 | 0.087895 | -2.243462e+07 | 22434615.292814 |
| 40 | -1.0071863 | -1.029860e+00 | 0.022674 | -6.872019e+06 | 6872018.460814 |
| 50 | -1.0071863 | -1.001604e+00 | -0.005583 | -8.495306e+03 | 8494.298428 |
| 60 | -1.0071863 | -9.752971e-01 | -0.031889 | 1.854531e+01 | -19.552497 |
| 70 | -1.0071863 | -9.101032e-01 | -0.097083 | -3.987027e-01 | -0.608484 |
| 80 | -1.0071863 | -8.710932e-01 | -0.136093 | -3.303453e-01 | -0.676841 |
| 90 | -1.0071863 | -8.511218e-01 | -0.156064 | -3.635602e-01 | -0.643626 |
| 100 | -1.0071863 | -7.705619e-01 | -0.236624 | -2.968952e-01 | -0.710291 |

| **Target Redemption Note with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 70\%, \sigma_2 = 70\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=15 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -0.811372 | -0.907822 | 0.096451 | -1.236249 | 0.424877 |
| 20 | -0.811372 | -0.810489 | -0.000883 | 23.539616 | -24.350988 |
| 30 | -0.811372 | -0.841646 | 0.030275 | -27.469004 | 26.657633 |
| 40 | -0.811372 | -0.823041 | 0.011669 | 0.530456 | -1.341828 |
| 50 | -0.811372 | -0.797412 | -0.013960 | -0.226540 | -0.584831 |
| 60 | -0.811372 | -0.776865 | -0.034506 | -0.192494 | -0.618878 |
| 70 | -0.811372 | -0.736652 | -0.074719 | -0.175243 | -0.636128 |
| 80 | -0.811372 | -0.733268 | -0.078104 | -0.177795 | -0.633577 |
| 90 | -0.811372 | -0.712344 | -0.099028 | -0.189708 | -0.621663 |
| 100 | -0.811372 | -0.667772 | -0.143600 | -0.197093 | -0.614279 |

| **Target Redemption Note with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 50\%, \sigma_2 = 50\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=15 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -0.599460 | -0.619322 | 0.019862 | -0.547591 | -0.051869 |
| 20 | -0.599460 | -0.601981 | 0.002521 | -0.338054 | -0.261406 |
| 30 | -0.599460 | -0.616508 | 0.017048 | -0.339077 | -0.260383 |
| 40 | -0.599460 | -0.569402 | -0.030058 | -0.244711 | -0.354749 |
| 50 | -0.599460 | -0.559740 | -0.039720 | -0.216601 | -0.382859 |
| 60 | -0.599460 | -0.560428 | -0.039032 | -0.162388 | -0.437072 |
| 70 | -0.599460 | -0.554784 | -0.044676 | -0.142801 | -0.456659 |
| 80 | -0.599460 | -0.541449 | -0.058011 | -0.124179 | -0.475281 |
| 90 | -0.599460 | -0.521677 | -0.077783 | -0.114069 | -0.485391 |
| 100 | -0.599460 | -0.508636 | -0.090824 | -0.106292 | -0.493168 |

| **Target Redemption Note with** $\rho = -0.9$ | | | | | |
|---|---|---|---|---|---|
| $\sigma_1 = 30\%, \sigma_2 = 30\%, \quad \kappa_1 = 7\%, \kappa_2 = 8\%$, discretization $Z$=15 | | | | | |
| N | Monte Carlo | Tree Rotation | error | Quadrinomial Tree | error |
| 10 | -0.361349 | -0.352093 | -0.009256 | -0.286948 | -0.074401 |
| 20 | -0.361349 | -0.364224 | 0.002875 | -0.289876 | -0.071473 |
| 30 | -0.361349 | -0.374245 | 0.012896 | -0.253331 | -0.108018 |
| 40 | -0.361349 | -0.348053 | -0.013296 | -0.233170 | -0.128179 |
| 50 | -0.361349 | -0.345188 | -0.016161 | -0.198542 | -0.162807 |
| 60 | -0.361349 | -0.346147 | -0.015202 | -0.178248 | -0.183101 |
| 70 | -0.361349 | -0.327695 | -0.033654 | -0.157891 | -0.203458 |
| 80 | -0.361349 | -0.323829 | -0.037520 | -0.124031 | -0.237318 |
| 90 | -0.361349 | -0.321888 | -0.039461 | -0.095669 | -0.265680 |
| 100 | -0.361349 | -0.304775 | -0.056574 | -0.106369 | -0.254980 |

We can conclude that also in the case of the Target Redemption Note, the rotated tree strongly outperforms the quadrinomial one (for high volatility parameters and correlation of $-0.9$). The

difference between both tree constructions is clearly visible for volatility starting around 30% and increases with the increase of volatility.

**Remark 2.6.1.** *Notice that for the well-known trinomial two-factor Hull-White [37] tree construction, first the base factors are assumed to be independent and the risk-neutral probabilities are calculated and then the probabilities are adjusted to reflect the correlation between the base factors. As a result, (similar to the quadrinomial tree case) it may happen that the probabilities are not bounded between zero and one. In such a case, the authors choose every time the probabilities are negative or greater than one the biggest/smallest correlation that forces the probabilities to be well-defined. Although in such a case it is obvious that a bias will be induces, Hull and White [37] prove that this bias will converge to zero as the refinement of the tree increases. Recall that we have proved that this is also the case with the probabilities in the quadrinomial tree.*

*On the other hand, if we are constructing a two-dimensional tree for processes with high volatilities, we have showed that the effect of the not well-defined probabilities induces a significant bias and although in such a case theoretically the tree converges, practically a proper refinement cannot be reached. Although, the bias in the trinomial tree of Hull and White was not examined here, a similar effect in cases of high volatilities is to be expected.*

*In addition, the standard trinomial two-factor Hull-White [37] tree is much harder to construct due to the fact that first a tree with no mean-reversion has to be constructed and then node by node shifted with a drift chosen so that the tree is consistent with the initial term structure. In contrast to it, the new binomial tree constructions, by definition exactly fits the current term structure.*

## 2.7 Final Remarks

At the end, we want to answer the question when one of the tree constructions is superior to the other. Specifying two very general cases for the volatility - **high volatility** (higher than 30%) and **low volatility**(lower than 30%) and also two cases for the correlation $\rho > \mathbf{0}$ and $\rho < \mathbf{0}$ we can then summarize:

| | **High Volatility** | **Low Volatility** |
|---|---|---|
| $\rho > \mathbf{0}$ | Rotated Tree≈Quadrinomial Tree | Quadrinomial Tree Slightly Better |
| $\rho < \mathbf{0}$ | Rotated Tree Definitely Better | No Clear Preference |

Notice that although the rotated tree construction delivers slightly worse results than the quadrinomial tree construction for the case of very low volatilities and positive (and some in cases for negative) correlation between the base factors, due to its good results in the other cases and the well-defined tree probabilities we can conclude that it is a robuster (than the quadrinomial tree) lattice construction method.

However, by calibration of the $G2++$ model to market caps or Swaption prices, Brigo and Mercurio [12] show that the obtained calibrated correlation is usually very negative (e.g. they obtain $\rho = -0.9914$ when calibrating to caps, which naturally means that in that case a one-factor model might be sufficient and $\rho = -0.7019$ when calibrating to Swaptions, which clearly exhibits the need for a two-factor model, see [12] p.157 and p.158) but the calibrated volatilities are very small which leads to the case in which it is hard to give a practical preference to any of the models. In that specific case, for different values of the mean-reversion parameters we can often observe that either one of the tree constructions outperforms the other, but it is hard to specify clear distinction values for these parameters. In this respect, a future work in that direction might help the study of the different lattice construction methods.

As next, we mention that both offered tree construction methods are easy to implement and program. Additionally, they are also easy to adapt for pricing different path-dependent options and are much faster than other numerical methods such as the Monte Carlo simulation or the finite-difference grid.

Finally, one should mention that the offered tree construction methods require a significant amount of computer RAM memory since it is not known in advance when the contraction point of the tree (due to the mean reversion of underlying process) will be reached and thus a standard tree construction procedure cannot be used. This is not the case in the trinomial two-factor

Hull-White [37] trees where one has to estimate in advance the maximal allowed expansion of the lattice. While from a theoretical point of view, the former feature is an advantage of our offered tree construction to the trinomial two-factor Hull-White [37] one, from a practical point of view this causes difficulties and does not allow (with the present standard computer power) a tree construction with more than 300 timesteps.

## 2.8 Suggestions for Further Research Topics

A possible topic to investigate can be a comparison of the convergence of our tree constructions to the two-factor Hull and White trinomial lattice construction. It is can also be interesting to implement the pricing of a "steepener" and investigate the influence of the correlation between the base factors on its value as well as their mean reversion parameters. Calibration to market steepener prices is also an interesting research point.

# Chapter 3

# Pricing European Options with an Interest Rate of Cheyette Type

## 3.1 Introduction

There is a wide range of controversial studies about the influence of the volatility structure of spot and forward interest rates on the prices of interest-rate sensitive claims. The oldest one is done by Hull and White [34], who argue on the bases of several numerical examples that models with a simple Vasicek type volatility can be used as an approximation even for a true volatility structure of a square root form as this would not induce a significant error in the prices of interest-rate sensitive claims.

The first numerical results opposing Hull and White [34] are delivered by Chan *et al* [18] who show after testing several interest rate models (with different volatility specifications) that option prices are sensitive to the level of the short rate volatility. However, in the models they have tested, the initial term structure of the zero coupon bonds also changes with the volatility specification and additionally influence the option prices. As a result, one cannot argue how strong the option prices are influenced only by a change in the volatility term structure.

The latest studies by Ritchken and Sankarasubramanian [58] overcome this additional influence by initializing all tested models to the same term structure and the same initial set of forward volatilities. On the bases of a variety of numerical examples, they prove that the use of generalized Vasicek type volatility structure models as an approximation in the cases when the true

volatility structure is not of that type can lead to a serious misprizing error. For that purpose, they have used as a comparison a volatility structure of a Cheyette [21] type (introduced in details in Chapter 1) which allows fluctuation of the volatility of the spot interest rate with its level and can be used as an approximation of a big set of possible volatility structures. However, due to the complicated dynamic of the corresponding (to the Cheyette volatility structure) short rate process, their numerical examples are based only on Monte Carlo simulation which is unfortunately time consuming and thus limits the practical application of the model.

In this Chapter, adopting a volatility structure of a Cheyette [21] type, we contribute to the practical application of the model by offering two fast and easy to implement lattice construction methods for the pricing of interest rate sensitive claim. Finally, we perform several numerical examples in order to compare and examine their advantages and disadvantages and also offer methods to improve their convergence.

## 3.2 Definition of the Underlying Processes

In the following section, we assume that we have a complete probability space $\left(\Omega, \{\mathfrak{F}\}_{t\in[0,T]}, P\right)$ and the HJM setting where the volatility of the forward rate is limited to the following form

$$\sigma(t,T) := \sigma_f(t,t)e^{-\int_t^T \kappa(x)dx} \tag{3.1}$$

where $\sigma_f(t,t)$ is the instantaneous short rate volatility and $\kappa$ is a deterministic integrable function[1]. In such a framework, (recall Chapter 1) the price of any interest rate derivative is completely determined by a two-state Markov process $\chi(\cdot) = (r(\cdot), \phi(\cdot))$ where $r(\cdot)$ represents the short rate and $\phi(t)$ denotes the cumulative quadratic variation process

$$\phi(t) = \int_0^t \sigma(s,t)^2 ds.$$

Assume also that we have an arbitrage-free market which implies that there exists at least one equivalent (to the physical measure $P$) martingale measure we denote $Q$ under which the stock price has the following dynamics

$$dS(t) = S(t)\left[r(t)dt + \sigma_1 dW_1(t)\right], \quad S(0) = S_0 \tag{3.2}$$

[1]Notice that if we limit $\kappa$ to a positive constant, this would reflect the notion that distant forward rates are less volatile than near-term rates.

where $W_1(t)$ is a standard Brownian motion under $Q$ and $\sigma_1 > 0$ is the deterministic volatility of the stock price dynamic.

Recall further that under the class of volatilities of the forward rate given in (3.1), the instantaneous short rate process $r(\cdot)$ and the accumulated until time $t$ variance of the forward rate, evolve according to

$$\begin{aligned} dr(t) &= \mu(t,\phi,r)dt + \sigma_f(t,t)dW_2(t), \quad r(0) = r_0 \\ d\phi(t) &= \left[\sigma_f^2(t,t) - 2\kappa(t)\phi(t)\right] dt, \quad \phi(0) = 0 \\ \mu(t,\phi,r) &= \kappa(t)\left[f(0,t) - r(t)\right] + \phi(t) + \frac{d}{dt}f(0,t) \end{aligned} \tag{3.3}$$

where $\sigma_f(t,t) = \sigma_2 r(t)^\gamma$, $\gamma \geq 0$, $\sigma_2 \in \mathbb{R}^+$, $f(0,t)$ denotes the instantaneous forward interest rate set at time 0 for time $t$ and $W_2(t)$ is a standard Brownian motion process under the risk-neutral measure $Q$. The instantaneous correlation between the two processes is given by $d\langle W_1(\cdot), W_2(\cdot)\rangle_t = \rho dt$. In this case, notice that

$$\phi(t) = \int_0^t \sigma(s,t)^2 ds = \int_0^t \sigma_2^2 r(u)^2 e^{-\int_u^t \kappa(x)dx} du.$$

**Remark 3.2.1.** *The interest rate process (3.3) allows us by changing the elasticity parameter $\gamma$ to choose between a variety of driving processes. For example, setting $\gamma = 0$, we obtain the generalized Vasicek [61] model*

$$dr(t) = \mu(t,\phi,r)dt + \sigma_2 dW_2(t), \quad r(0) = r_0$$

*or by setting it to 0.5, we obtain the square root structure of Cox, Ingersoll and Ross [25]*

$$dr(t) = \mu(t,\phi,r)dt + \sigma_2\sqrt{r(t)}dW_2(t), \quad r(0) = r_0.$$

*Notice that if we want the volatility of the instantaneous interest rate to fluctuate with the level of the short rate, we require that $\gamma \geq 0$. Further, studies by Richken and Sankarasubramania [58] reveal in a variety of numerical examples that while setting the elasticity parameter $\gamma$ to zero might lead to pricing errors, the choice of different spot rate volatility structure (i.e. different $\gamma \neq 0$) may not be that important. For that reason, in this Chapter we focus our attention only on the case $\gamma = 1$.*

Assuming we want to price a claim $V(t, r(t), \phi(t), S(t))$ contingent on both the short rate and the stock price and with payoff at time $T$ given by $F(T, r(T), \phi(T), S(T))$, we can write its price by the martingale method as an expectation under a risk-neutral measure $Q$ of its discounted payoff such that

$$V(t, r(t), \phi(t), S(t)) = E^Q\left(e^{-\int_t^T r(s)ds} F(T, r(T), \phi(T), S(T)) \left|\mathfrak{F}_t\right.\right)$$

or equivalently derive its pricing PDE. First, by means of Itô's Lemma, we obtain

$$\begin{aligned}
dV(t, r(t), \phi(t), S(t)) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial \phi}d\phi + \frac{\partial V}{\partial S}dS \\
&\quad + \frac{1}{2}\left(\frac{\partial^2 V}{\partial r^2}d\langle r\rangle + \frac{\partial^2 V}{\partial S^2}d\langle S\rangle + 2\frac{\partial^2 V}{\partial rS}d\langle r, S\rangle\right) \\
&= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial r}\left(\mu(t,\phi,r)dt + \sigma_2 r(t)^\gamma dW_2(t)\right) \\
&\quad + \frac{\partial V}{\partial \phi}\left(\sigma_2^2 r(t)^{2\gamma} - 2\kappa(t)\phi(t)\right)dt + \frac{\partial V}{\partial S}S(t)\left(r(t)dt + \sigma_1 dW_1(t)\right) \\
&\quad + \frac{1}{2}\left(\frac{\partial^2 V}{\partial r^2}\sigma_2^2 r(t)^{2\gamma} + \frac{\partial^2 V}{\partial S^2}\sigma_1^2 S(t)^2 + 2\frac{\partial^2 V}{\partial rS}\sigma_1\sigma_2 r(t)^\gamma S(t)\rho\right)dt.
\end{aligned}$$

By eliminating uncertainty, we require due to the no-arbitrage principle that the expected return of our tradable instrument is the riskless interest rate, i.e.

$$E^Q\left(dV(t, r(t), \phi(t), S(t))\right) = r(t)V(t, r(t), \phi(t), S(t))\,dt.$$

This leads to the **pricing PDE**

$$\begin{aligned}
\frac{\partial V}{\partial t} &+ \frac{\partial V}{\partial r}\mu(t,\phi,r) + \frac{\partial V}{\partial \phi}\left(\sigma_2^2 r(t)^{2\gamma} - 2\kappa(t)\phi(t)\right) + \frac{\partial V}{\partial S}S(t)r(t) \\
&+ \frac{\partial^2 V}{\partial r^2}\sigma_2^2 r(t)^{2\gamma} + \frac{\partial^2 V}{\partial S^2}\sigma_1^2 S(t)^2 + 2\frac{\partial^2 V}{\partial rS}\sigma_1\sigma_2 r(t)^\gamma S(t)\rho - r(t)V(t, r(t), \phi(t), S(t)) = 0
\end{aligned}$$

with the boundary condition

$$V(T, r(T), \phi(T), S(T)) = F(T, r(T), \phi(T), S(T)).$$

Now, due to the complexity of the pricing PDE, we cannot solve it analytically even for the most simple vanilla options. Therefore, numerical methods are required in order to price contingent claims in the given framework. As in the previous Chapter, we will concentrate on lattice methods and Monte Carlo Simulation.

## 3.3 Two-Dimensional Tree Construction

Our aim, after choosing the dynamics of the underlying processes, is to be able to price claims contingent on the stock price and the interest rate. Due to the complicated nature of the underlying processes, this cannot be done in a closed form. Therefore, we need to apply numerical methods. In this section, we will present two possible lattice construction methods and compare their results with fine-grid Monte Carlo simulations.

For the purposes of the future tree construction, we need to transform the underlying stochastic processes into processes with a deterministic volatility. This can be achieved, by taking the natural logarithm of the stock price as one of the underlying processes of the two-dimensional tree:

$$\begin{aligned} d\big(\log(S(t))\big) &= (\log S(t))'\, dS(t) + \frac{1}{2}\,(\log S(t))''\, d\langle S\rangle_t \\ &= \left(r(t) - \frac{1}{2}\sigma_1^2\right) dt + \sigma_1 dW_1(t). \end{aligned}$$

Denote it by $S^*(t) := \log(S(t))$ and rewriting the expression above, we have

$$dS^*(t) = \left(r(t) - \frac{1}{2}\sigma_1^2\right) dt + \sigma_1 dW_1(t) \tag{3.4}$$

$$S^*(0) = \log(S_0). \tag{3.5}$$

Next, since the stochastic volatility of the interest rate process is also an obstacle for the tree construction, we define a new process

$$Y(t) := \int \frac{1}{\sigma_f(t,t)} dr(t) = \int \frac{1}{\sigma_2 r(t)^\gamma}\, dr(t) \tag{3.6}$$

and show that it has a unit diffusion coefficient. Integrating (3.6), we obtain

$$Y(t) = f(r(t)) = \begin{cases} \frac{r(t)^{1-\gamma}}{\sigma_2(1-\gamma)} & \text{for } \gamma \neq 1\ , \\ \frac{\log(r(t))}{\sigma_2} & \text{else.} \end{cases}$$

and we can also transform $Y(t)$ to obtain the short rate $r(t)$, by using the inverse function $f^{-1}(Y(t))$ and thus

$$r(t) = \begin{cases} [\sigma_2 Y(t)(1-\gamma)]^{\frac{1}{1-\gamma}} & \text{for } \gamma \neq 1\ , \\ e^{\sigma_2 Y(t)} & \text{else.} \end{cases}$$

Using Ito's Lemma, we can now calculate the dynamics of the new process $Y$

$$\begin{aligned}
dY(t,r(t)) &= \frac{\partial Y\,(t,r(t))}{\partial t}dt + \frac{\partial Y(t)}{\partial r(t)}dr(t) + \frac{1}{2}\frac{\partial^2 Y(t,r(t))}{\partial r(t)^2}d\langle r\rangle_t \\
&= \frac{\partial Y(t,r(t))}{\partial t}dt + \frac{\partial Y(t,r(t))}{\partial r(t)}\mu(t,\phi,r)dt + \frac{\partial Y(t,r(t))}{\partial r(t)}\sigma_2 r(t)^\gamma dW_2(t) \\
&\quad + \frac{1}{2}\sigma_2^2 r(t)^{2\gamma}\frac{\partial^2 Y(t,r(t))}{\partial r(t)^2}dt \\
&= \frac{\partial Y(t,r(t))}{\partial t}dt + \frac{\partial Y(t,r(t))}{\partial r(t)}\mu(t,\phi,r)dt + dW_2(t) \\
&\quad + \frac{1}{2}\sigma_2^2 r(t)^{2\gamma}\frac{\partial^2 Y(t,r(t))}{\partial r(t)^2}dt \\
&= \left(Y_t + Y_r\mu(t,\phi,r) + \frac{1}{2}\,\sigma_2^2 r^{2\gamma}Y_{rr}\right)dt + dW_2(t).
\end{aligned}$$

We plug in the dynamics of $r(\cdot)$ for the two cases of $\gamma$ and obtain

1.) $\gamma \neq 1$

$$\begin{aligned}
dY(t,r(t)) &= Y_t dt + \frac{r(t)^{-\gamma}}{\sigma_2}\mu(t,\phi,r)dt + \frac{1}{2}\sigma_2^2 r^{2\gamma}\left(-\frac{\gamma}{\sigma}r(t)^{-\gamma-1}\right)dt + dW_2(t) \\
&= m_r(t,Y,\phi)dt + dW_2(t), \quad \text{where} \\
m_r(t,Y,\phi) &= Y_t + \frac{r(t)^{-\gamma}}{\sigma_2}\mu(t,\phi,r) + \frac{1}{2}\sigma_2^2 r^{2\gamma}\left(-\frac{\gamma}{\sigma}r(t)^{-\gamma-1}\right) \\
&= Y_t + \frac{[\sigma_2 Y(t)(1-\gamma)]^{-\frac{\gamma}{1-\gamma}}}{\sigma_2}\mu(t,\phi,r) - \frac{\gamma}{2(1-\gamma)Y(t)} \\
&= Y_t + \frac{[\sigma_2 Y(t)(1-\gamma)]^{-\frac{\gamma}{1-\gamma}}}{\sigma_2}\left(kf(0,t) + \phi(t) + \frac{d}{dt}f(0,t)\right) \\
&\quad -kY(t)(1-\gamma) - \frac{\gamma}{2(1-\gamma)Y(t)}
\end{aligned}$$

with $Y_0 = \frac{r_0^{1-\gamma}}{\sigma_2(1-\gamma)}$.

2.) $\gamma = 1$

$$\begin{aligned}
dY(t,r(t)) &= Y_t dt + Y_r\mu(t,\phi,r)dt + \frac{1}{2}\sigma_2^2 r(t)^2 Y_{rr}dt + dW_2(t) \\
&= m_r(t,Y,\phi)dt + dW_2(t), \quad \text{where} \\
m_r(t,Y,\phi) &= Y_t + \frac{1}{\sigma_2 r(t)}\mu(t,\phi,r) - \frac{1}{2}\sigma_2 \\
&= Y_t + \frac{1}{\sigma_2 r(t)}\left(kf(0,t) + \phi(t) + \frac{d}{dt}f(0,t)\right) - \frac{k}{\sigma_2} - \frac{\sigma_2}{2}
\end{aligned}$$

with $Y_0 = \frac{\log r_0}{\sigma_2}$.

Thus, we have transformed the original problems (3.2) and (3.3) for $\gamma \neq 1$ to

$$\begin{aligned} dS^*(t) &= m_s(Y)dt + \sigma_1 dW_1(t), \quad S^*(0) = S_0^* \\ dY(t) &= m_r(t, Y, \phi)dt + dW_2(t), \quad Y(0) = Y_0 \\ d\phi(t) &= \left[\sigma_2^2 \big(\sigma_2 Y(t)(1-\gamma)\big)^{\frac{2\gamma}{1-\gamma}} - 2\kappa\phi(t)\right] dt, \quad \phi(0) = 0 \end{aligned}$$

where

$$m_s(Y) := \big(\sigma_2 Y(t)(1-\gamma)\big)^{\frac{1}{1-\gamma}} - \frac{1}{2}\sigma_1^2$$

and respectively for $\gamma = 1$ to

$$\begin{aligned} dS^*(t) &= m_s(Y)dt + \sigma_1 dW_1(t), \quad S^*(0) = S_0^* \\ dY(t) &= m_r(t, Y, \phi)dt + dW_2(t), \quad Y(0) = Y_0 \\ d\phi(t) &= \left[\sigma_2^2 e^{2\sigma_2 Y(t)} - 2\kappa\phi(t)\right] dt, \quad \phi(0) = 0 \end{aligned}$$

where

$$m_s(Y) := \sigma_2 e^{\sigma_2 Y(t)} - \frac{1}{2}\sigma_1^2.$$

### 3.3.1 Tree Construction with Rotation

In this section, we will present a method to price claims contingent on the underlying two processes by using the method of rotated two-dimensional tree, presented in Section 2.3. For that purpose, we need to find two independent new processes that are linear combination of the basic ones. Following the same steps as before, we decompose the covariance matrix between $dS^*(t)$ and $dY(t)$

$$\frac{1}{dt}\mathbf{\Sigma} := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_1 & 1 \end{pmatrix} = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{pmatrix} \tag{3.7}$$

and find the eigenvalues

$$\lambda_{1,2} = \frac{1}{2}(\sigma_1^2 + 1) \pm \frac{1}{2}\sqrt{(\sigma_1^2+1)^2 - 4(1-\rho^2)\sigma_1^2}.$$

The rotation matrix is denoted by $\mathbf{U}$ and

$$\mathbf{U} := \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

where $\varphi$ is the rotation angle and $\cos\varphi = \frac{1}{\sqrt{1+\beta^2}}$, $\sin\varphi = \frac{\beta}{\sqrt{1+\beta^2}}$, $\beta = \frac{-\lambda_1+\sigma_1^2}{\rho\sigma_1}$. The new independent processes $Y_1$ and $Y_2$ are thus written as

$$Y_1(t) = \cos\varphi\, S^*(t) - \sin\varphi\, Y(t) \tag{3.8}$$

$$Y_2(t) = \sin\varphi\, S^*(t) + \cos\varphi\, Y(t). \tag{3.9}$$

Referring to Theorem 2.3.1, we obtain that $Y_1(t)$ and $Y_2(t)$ have the following dynamics

$$\begin{aligned} dY_1(t) &= \alpha_1(t,\phi,Y_1,Y_2)dt + \sqrt{\lambda_1}d\widetilde{W}_1 \\ dY_2(t) &= \alpha_2(t,\phi,Y_1,Y_2)dt + \sqrt{\lambda_2}d\widetilde{W}_2 \\ d\phi(t) &= \left[\sigma_2^2 R(Y_1,Y_2)^{2\gamma} - 2\kappa\phi(t)\right] dt \end{aligned}$$

with $\widetilde{W}_1$ and $\widetilde{W}_2$ denoting independent Brownian motions under a risk-neutral measure $Q$ and where

$$R(Y_1,Y_2) := \begin{cases} \left[\sigma_2\big(\cos\varphi Y_2(t) - \sin\varphi Y_1(t)\big)(1-\gamma)\right]^{\frac{1}{1-\gamma}} & \text{for } \gamma \neq 1\ , \\ \exp\Big(\sigma_2\big(\cos\varphi Y_2(t) - \sin\varphi Y_1(t)\big)\Big) & \text{else.} \end{cases}$$

The drifts of the new processes are found as

$$\begin{aligned} \alpha_1(t,\phi,Y_1,Y_2) &= \cos\varphi\, m_s(Y_1,Y_2) - \sin\varphi\, m_r(t,\phi,Y_1,Y_2) \\ \alpha_2(t,\phi,Y_1,Y_2) &= \sin\varphi\, m_s(Y_1,Y_2) + \cos\varphi\, m_r(t,\phi,Y_1,Y_2). \end{aligned}$$

Notice that using $Y = g(Y_1,Y_1)$, we have transformed the drifts $m_s(\cdot)$ and $m_r(\cdot)$ as functions of the new independent processes $Y_1$ and $Y_2$

$$\begin{aligned} m_s(Y) &= m_s\left(g(Y_1,Y_2)\right) := m_s(Y_1,Y_2) \\ m_r(t,Y,\phi) &= m_r\left(t, g(Y_1,Y_2), \phi\right) := m_r(t,Y_1,Y_2,\phi) \end{aligned}$$

where the transformation function $g(Y_1, Y_2)$ is found from (3.8) and (3.9) as

$$Y(t) = g(Y_1, Y_2) = \cos\varphi Y_2(t) - \sin\varphi Y_1(t).$$

In addition, we have

$$S^*(t) = \cos\varphi Y_1(t) + \sin\varphi Y_2(t).$$

Finally, we can now approximate the distribution of $Y_1$ and $Y_2$ by two independent trees following the procedure of the Chapter 2.

**Forward Construction:**

Assume that at time $t$, the tree is in state $(y_1^a, y_2^a)$. At time $t + \Delta t$, the new states of the underlying processes $Y_1$ and $Y_2$ are given by

$$\begin{aligned} y_1^{a^+} &:= y_1^a + \big(J_1(t, \phi^a, y_1^a, y_2^a) + 1\big)\sqrt{\Delta t \lambda_1}, \quad y_1^{a^-} := y_1^a + \big(J_1(t, \phi^a, y_1^a, y_2^a) - 1\big)\sqrt{\Delta t \lambda_1} \\ y_2^{a^+} &:= y_2^a + \big(J_2(t, \phi^a, y_1^a, y_2^a) + 1\big)\sqrt{\Delta t \lambda_2}, \quad y_2^{a^-} := y_2^a + \big(J_2(t, \phi^a, y_1^a, y_2^a) - 1\big)\sqrt{\Delta t \lambda_2}. \end{aligned}$$

Notice that since the process $\phi$ is locally deterministic, the value $\phi^{a^+}$ and $\phi^{a^-}$ will be equal and entirely determined by the current state variables $(y_1^a, y_2^a, \phi^a)$. Thus, we have

$$\phi^{a^+} = \phi^{a^-} = \phi^* = \phi^a + \left[\sigma_2^2 R(y_1^a, y_2^a)^{2\gamma} - 2\kappa\phi^a\right]\Delta t.$$

Since these values are completely determined by the preceding node, after some time steps the total number of distinct values $\phi$ at each node will depend on the total number of paths leading to that node. Therefore, we use (as in Li, Ritchken and Sankarasubramanian [46]) two values $\phi_{max}$ and $\phi_{min}$ to keep track of the biggest and smallest value of all $\phi$ variables at that node. In particular, at the $\phi^a$ node we denote with $\phi_{max}^a$ and $\phi_{min}^a$ respectively the maximal and minimal value of $\phi$ and then partition the interval $[\phi_{max}^a, \phi_{min}^a]$ into $m$ equidistant subintervals. We denote by $\phi_k^a$ the $k$-th point where

$$\phi_k^a = \phi_{min}^a + \frac{1}{m}\big(\phi_{max}^a - \phi_{min}^a\big), \quad k = 0, \ldots, m$$

and

$$\phi_{min}^a = \phi_1^a < \phi_2^a < \ldots < \phi_m^a = \phi_{max}^a.$$

Let us denote the probability of an up-jump of the first process by $p_1$ and of the second process by $p_2$. By the requirement that the approximating tree matches the first moment of the underlying process, we obtain that the risk-neutral probabilities have to satisfy

$$\begin{aligned} p_1(y_1^{a^+} - y_1^a) + (1-p_1)(y_1^{a^-} - y_1^a) &= \alpha_1(t, \phi^a, y_1^a, y_2^a)\Delta t \\ p_2(y_2^{a^+} - y_2^a) + (1-p_2)(y_2^{a^-} - y_2^a) &= \alpha_2(t, \phi^a, y_1^a, y_2^a)\Delta t \end{aligned}$$

from where we find

$$p_1 = \frac{\alpha_1(t, \phi^a, y_1^a, y_2^a)\sqrt{\Delta t} + \big(1 - J_1(t, \phi^a, y_1^a, y_2^a)\big)\sqrt{\lambda_1}}{2\sqrt{\lambda_1}} \tag{3.10}$$

$$p_2 = \frac{\alpha_2(t, \phi^a, y_1^a, y_2^a)\sqrt{\Delta t} + \big(1 - J_2(t, \phi^a, y_1^a, y_2^a)\big)\sqrt{\lambda_2}}{2\sqrt{\lambda_2}}. \tag{3.11}$$

Now, let us say more about the choice of $J_1$ and $J_2$. We denote $Z_1 :=$ floor $\left[\frac{\alpha_1(t,\phi^a,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_1}}\right]$ and $Z_2 :=$ floor $\left[\frac{\alpha_2(t,\phi^a,y_1^a,y_2^a)\sqrt{\Delta t}}{\sqrt{\lambda_2}}\right]$. And then define

$$J_1(t, \phi^a, y_1^a, y_2^a) := \begin{cases} Z_1 & \text{if } Z_1 \text{ even,} \\ Z_1 + 1 & \text{else.} \end{cases}$$

$$J_2(t, \phi^a, y_1^a, y_2^a) := \begin{cases} Z_2 & \text{if } Z_2 \text{ even,} \\ Z_2 + 1 & \text{else.} \end{cases}$$

Notice that with this definition of the jump heights and the risk-neutral probabilities, we can prove in the same way as in Theorem 2.3.3 that the approximating tree matches the first and second moment of the underlying processes.

**Backward Recursion:**

Finally, once we have constructed the rotated two dimensional tree, we want to use it to price contingent claims via backward recursion. Hence, we start with calculating the values of the claim at each node of the tree at maturity, by substituting in the payoff functions the respective states of the underlying processes. We denote the option values at maturity by $g_n(y_1^a, y_2^a, \phi^*)$. Next, we come to the backward recursion. Therefore, we assume we have calculated all the claim values until the $(i+1)$-th period and we want to calculate them at the $(i)$-th period. Given that at time $i$, the state variables are $(y_1^a, y_2^a, \phi^{a*})$ we have that

$$\begin{aligned} g_i\left(y_1^a, y_2^a, \phi^a\right) = &\Big[p_1 p_2 g_{i+1}\big(y_1^{a^+}, y_2^{a^+}, \phi^{a*}\big) + p_1(1-p_2) g_{i+1}\big(y_1^{a^+}, y_2^{a^-}, \phi^{a*}\big) \\ &+(1-p_1)p_2 g_{i+1}\big(y_1^{a^-}, y_2^{a^+}, \phi^{a*}\big) + (1-p_1)(1-p_2) g_{i+1}\big(y_1^{a^-}, y_2^{a^-}, \phi^{a*}\big)\Big] e^{-R(y_1^a, y_2^a)\Delta t} \end{aligned} \tag{3.12}$$

As the value $\phi^{a*}$ is completely determined by the predecessor node $(y_1^a, y_2^a, \phi^a)$, it can happen that the value of the claim at the successor nodes exactly for $\phi^{a*}$ is not available. At the same time, the values of the claim for example at the node $(y_1^{a*}, y_2^{a*}, \phi_+^{a*})$ for a slightly bigger value of $\phi^{a*}$ and at the node $(y_1^{a*}, y_2^{a*}, \phi_-^{a*})$ for a slightly smaller value of $\phi^{a*}$ will be available, where we have that

$$\phi_-^{a*} \leq \phi^{a*} \leq \phi_+^{a*}.$$

In such a case, the values of the claim at the node $(y_1^{a*}, y_2^{a*}, \phi^{a*})$ can be found by a linear interpolation between the available surrounding values. Thus, if we denote

$$\Delta\phi := \frac{\phi_{max}(y_1^{a^+}, y_2^{a^+}) - \phi_{min}(y_1^{a^+}, y_2^{a^+})}{m}$$

and

$$l := \text{floor}\left[\frac{\phi^{a*} - \phi_{min}(y_1^{a^+}, y_2^{a^+})}{\Delta\phi}\right].$$

where $\phi_{max}(y_1^{a^+}, y_2^{a^+})$ and $\phi_{min}(y_1^{a^+}, y_2^{a^+})$ denote the biggest and smallest value of $\phi$ at the $(y_1^{a*}, y_2^{a*}, \phi^{a*})$ node, then we notice that we have found exactly the surrounding values

$$\phi_l^{a*} = \phi_-^{a*} \leq \phi^{a*} \leq \phi_+^{a*} = \phi_{l+1}^{a*}.$$

And therefore, if we denote by $q := \frac{\phi^{a*} - \phi_l^{a*}}{\Delta\phi}$ the relative distance (or weight), we can find the required value of the claim by the following interpolation

$$g_{i+1}\left(y_1^{a^+}, y_2^{a^+}, \phi^{a*}\right) = q\, g_{i+1}\left(y_1^{a^+}, y_2^{a^+}, \phi_{l+1}^{a*}\right) + (1+q)\, g_{i+1}\left(y_1^{a^+}, y_2^{a^+}, \phi_l^{a*}\right)$$

**Remark 3.3.1.** *Notice that at the edges of the tree, we would always have $\phi_{max} = \phi_{min}$ as there is only one possible path leading to each of the edges. Hence, there is no need for interpolation at those nodes.*

Finally, we repeat the same procedure until we obtain also $q_{i+1}(y_1^{a^+}, y_2^{a^-}, \phi^{a*})$ , $g_{i+1}(y_1^{a^-}, y_2^{a^+}, \phi^{a*})$ and $q_{i+1}(y_1^{a^-}, y_2^{a^-}, \phi^{a*})$. Thus, we have all the ingredients we need to calculate (3.12) and letting $i = n, \ldots, 0$ we can find the approximated price of the claim at time 0 as $q_0(y_1(0), y_2(0), \phi(0))$.

### 3.3.2 Quadrinomial Tree

For the construction of the quadrinomial tree, we choose as base processes the natural logarithm of the stock price $S^*$ and the $Y$ process. Thus, assuming that at time $t$ the tree is in state $(S^a, y^a, \phi^a)$ then, at time $(t + \Delta t)$ the new states of the $S^*$ and $Y$ processes are denoted respectively by $S^{a^+}, S^{a^-}$ and $y^{a^+}, y^{a^-}$. The new states are reached with the following probabilities:

$$\begin{array}{ll} \left(S^{a^+}, y^{a^+}\right) & \text{with probability } p_{uu} \\ \left(S^{a^+}, y^{a^-}\right) & \text{with probability } p_{ud} \\ \left(S^{a^-}, y^{a^+}\right) & \text{with probability } p_{du} \\ \left(S^{a^-}, y^{a^-}\right) & \text{with probability } p_{dd}. \end{array}$$

and are given by

$$S^{a^+} := S^a + u, \quad S^{a^-} := S^a + d$$

for the $S^*$ process where $u$ is the height of its up-jump and $d$ the height of its down-jump and

$$y^{a^+} := y^a + (J(t, y^a, \phi^a) + 1)\sqrt{\Delta t}, \quad y^{a^-} := y^a + (J(t, y^a, \phi^a) - 1)\sqrt{\Delta t}.$$

For the $Y$ process we have

$$J(t, y^a, \phi^a) \quad := \quad \begin{cases} Z & \text{if } Z \text{ even,} \\ Z + 1 & \text{else.} \end{cases}$$

for $Z :=$ floor $\left[m_r(t, y^a, \phi^a)\sqrt{\Delta t}\right]$. Notice that the jump heights of the $Y$ process are chosen as in the one-dimensional Ritchken and Sankarasubramanian [58] lattice and therefore we do not comment about them. On the other side, the jump heights of the $S^*$ process require more explanation. If we recall that the jump heights and the probabilities of the tree should be chosen in such a way that the tree matches the local mean and variance of the $S^*$ process, we require that

$$E(\Delta S) = m_s(t, y^a)\Delta t = (p_{uu} + p_{ud})\left(S^a - S^{a^+}\right) + (p_{du} + p_{dd})\left(S^a - S^{a^+}\right) \tag{3.13}$$

$$E\left(\Delta S^2\right) = \sigma_1^2 \Delta t + m_s(t, y^a)^2 \Delta t^2 = (p_{uu} + p_{ud})\left(S^a - S^{a^+}\right)^2 + (p_{du} + p_{dd})\left(S^a - S^{a^+}\right)^2. \tag{3.14}$$

If we assume (in order to decrease the number of unknown variables) that $u = -d$ and use

$$1 = p_{uu} + p_{ud} + p_{du} + p_{dd} \tag{3.15}$$

we obtain

$$\begin{aligned} (p_{uu} + p_{ud}) &= \frac{m_s(t, y^a)\,\Delta t - u}{2u} \\ u^2 &= \sigma_1^2 \Delta t + m_s(t, y^a)^2\,\Delta t^2. \end{aligned}$$

Ignoring the terms of order higher than $\Delta t$, yields a tree known in the one-dimensional case as a Cox, Ross and Rubinstein [26] lattice where $u = \sigma_1\sqrt{\Delta t}$. Adding to equations (3.13) , (3.14) and (3.15) the new equations referring to the $Y$-process and the correlation between the processes

$$\begin{aligned} E(\Delta y) &= m_r(t, y^a)\Delta t = (p_{uu} + p_{du})\left(y^a - y^{a^+}\right) + (p_{du} + p_{dd})\left(y^a - y^{a^-}\right) \\ E\left(\Delta y^2\right) &= \Delta t + m_r(t, y^a)^2 \Delta t^2 = (p_{uu} + p_{du})\left(y^a - y^{a^+}\right)^2 + (p_{du} + p_{dd})\left(y^a - y^{a^-}\right)^2 \\ E(\Delta S \Delta y) &= \sigma_1 \rho \Delta t = p_{uu}\left(S^a - S^{a^+}\right)\left(y^a - y^{a^+}\right) + p_{du}\left(S^a - S^{a^-}\right)\left(y^a - y^{a^+}\right) \\ &\quad + p_{ud}\left(S^a - S^{a^+}\right)\left(y^a - y^{a^-}\right) + p_{dd}\left(y^a - y^{a^-}\right)\left(S^a - S^{a^-}\right) \end{aligned}$$

we solve for the risk-neutral probabilities and obtain

$$\begin{aligned} p_{uu} &= \frac{1}{4}\left[\frac{\rho\sigma_1 - (J-1)\sigma_1 - m_s(t, y^a)\sqrt{\Delta t}(J-1) + m_r(t, y^a, \varphi)\sqrt{\Delta t}\sigma_1}{\sigma_1}\right] \\ p_{dd} &= \frac{1}{4}\left[\frac{\rho\sigma_1 + (J+1)\sigma_1 - m_s(t, y^a)\sqrt{\Delta t}(J+1) - m_r(t, y^a, \varphi)\sqrt{\Delta t}\sigma_1}{\sigma_1}\right] \\ p_{ud} &= \frac{1}{4}\left[\frac{-\rho\sigma_1 + (J+1)\sigma_1 + m_s(t, y^a)\sqrt{\Delta t}(J+1) - m_r(t, y^a, \varphi)\sqrt{\Delta t}\sigma_1}{\sigma_1}\right] \\ p_{du} &= \frac{1}{4}\left[\frac{-\rho\sigma_1 - (J-1)\sigma_1 + m_s(t, y^a)\sqrt{\Delta t}(J-1) + m_r(t, y^a, \varphi)\sqrt{\Delta t}\sigma_1}{\sigma_1}\right]. \end{aligned}$$

The $\phi$ process is locally deterministic and we calculate it for $\gamma \neq 1$ by

$$\phi^{a^+} = \phi^{a^+} = \phi^{a*} = \phi^a + \left[\sigma_2^2\left(\sigma_2 y^a(1-\gamma)\right)^{\frac{2\gamma}{1-\gamma}} - 2\kappa\phi^a\right]\Delta t,$$

and for $\gamma = 1$ by

$$\phi^{a^+} = \phi^{a^+} = \phi^{a*} = \phi^a + \left[\sigma_2^2 e^{2\sigma_2 y^a} - 2\kappa\phi^a\right]\Delta t.$$

Pricing claims with the quadrinomial lattice follows the same steps as in the two dimensional rotated tree and thus we have again

$$\begin{aligned} g_i\left(S^a, y^a, \phi^a\right) &= \Big[p_{uu} g_{i+1}\big(S^{a^+}, y^{a^+}, \phi^{a*}\big) + p_{ud} g_{i+1}\big(S^{a^+}, y^{a^-}, \phi^{a*}\big) \\ &\quad + p_{du} g_{i+1}\big(S^{a^-}, y^{a^+}, \phi^{a*}\big) + p_{dd} g_{i+1}\big(S^{a^-}, y^{a^-}, \phi^{a*}\big)\Big] e^{-f^{-1}(y^a)\Delta t} \end{aligned} \tag{3.16}$$

where the values of the claim for $\phi^{*a}$ are interpolated as in the previous section.

## 3.4 Numerical Results

In this section, we construct the already described rotated and quadrinomial lattices and use them to price simple Vanilla Call and Put options. Further, we compare the obtained results with a Monte Carlo Simulation, performed with the following Euler-type discretization

$$\begin{aligned} r(t_{i+1}) &= r(t_i) + \Big[\kappa\big(f(0,t_i) - r(t_i)\big) + \phi(t_i) + \frac{d}{dt} f(0,t_i)\Big]\Delta t_i + \sigma_2 r(t_i)^\gamma \Delta \widetilde{W}_i^2 \\ \phi(t_{i+1}) &= \phi(t_i) + \left[\sigma_2^2 r(t_i)^{2\gamma} - 2\kappa\phi(t_i)\right]\Delta t_i \\ S(t_{i+1}) &= S(t_i) + S(t_i)\left[r(t_i)\Delta t_i + \sigma_1\left(\rho\Delta\widetilde{W}_i^2 + \sqrt{1-\rho^2}\Delta\widetilde{W}_i^1\right)\right] \end{aligned}$$

with $S(t_0) = S_0$, $r(t_0) = r_0$ , $\varphi(t_0) = 0$, $0 = t_0 < t_1 < \ldots < t_n = T$ and where we have denoted $\Delta t_i = t_{i+1} - t_i$ and $\Delta W_i := W(t_{i+1}) - W(t_i)$. Notice that we have decomposed the correlated Brownian motions $W_1$ and $W_2$ into a sum of independent Brownian motions $\widetilde{W}^1$ and $\widetilde{W}^2$ using Cholesky decomposition. As a result, we can easily simulate (3.3) and (3.2) using the given forward simulation scheme.

**Black-Scholes Smoothing:**

Small improvement of the convergence of a lattice for pricing European options can be done using the method of **Black-Scholes trees** proposed by Broadie and Detemple [13]. This method consists of substituting the option value (at the period exactly before maturity) with the Black-Scholes formula (see Appendix C). The motivation is that at time $t_{n-1}$, the value of the European claim is given by

$$E^Q\left(\exp\left(-r_{t_{n-1}}\Delta_{n-1}\right)(K - S_{t_n})^+ \Big| \mathfrak{F}_{t_{n-1}}\right) = \mathrm{BS}\left(S_{t_{n-1}},\, r_{t_{n-1}},\, K,\, \Delta_{n-1}\right)$$

where BS$(\cdot)$ denotes the Black-Scholes price of an European option with time to maturity $\Delta_{n-1}$, spot price $S_{t_{n-1}}$, strike $K$ and constant interest rate $r_{t_{n-1}}$. Although the calculation of the value

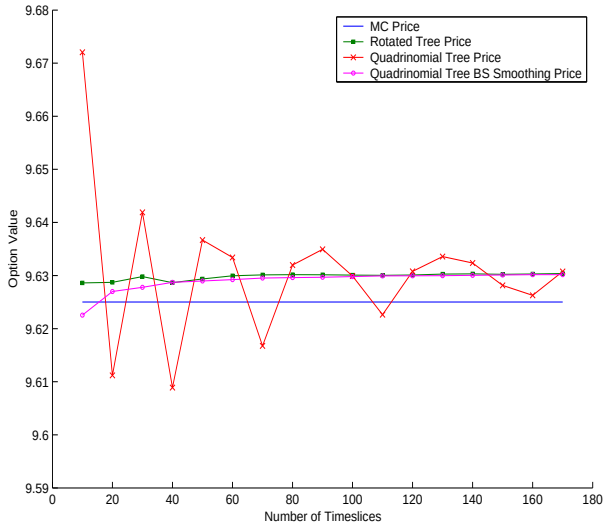
of the cumulative standard normal distribution function requires some extra time, since this is done only at the last time slice, the computational effort will not be significantly increased. As a result the method is easy to implement, even in a more complicated tree than the simple binomial one and does not cost a lot of extra time. The method is also easy to adapt for the pricing of American type of options, by substituting the continuation value of the American claim with the Black-Scholes formula. This will be explained in more details in the next chapter. However, the BS smoothing method is not very flexible since it can be applied only for the pricing of options for which the required closed-form solution one time step before maturity exists. In this respect, it cannot be applied for pricing options with complicated path dependence (e.g. some of the examples from the previous chapter).

Setting $\sigma_r = 23\%, \gamma = 1, \kappa = 0.083, S_0 = 36, \sigma_s = 30\%$ and $r_0 = 4\%$ we price a European Put Option with strike $K = 40$ and maturity $T = 1$. For the Monte Carlo simulation, we set the number of time steps to 500 and the number of paths to 200000. The results for correlation $\rho = 0.8$ and $S_0 = 30$ are given in Figure 3.1 a) and for $\rho = 0.8$ and $S_0 = 40$ in Figure 3.1 b). In addition, the results for correlation $\rho = -0.8$, $K = 40$ and $S_0 = 30$ are plotted in Figure 3.2 a) and for $\rho = -0.8$, $K = 36$ and $S_0 = 40$ in Figure 3.2 b).

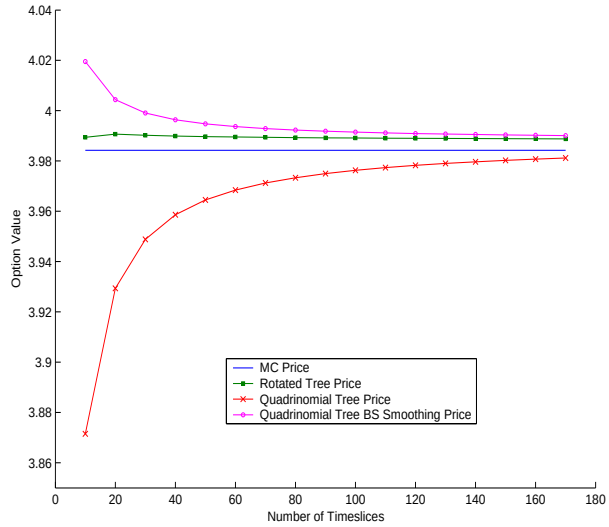
From the plots, one can easily notice that the rotated tree exhibits significantly smoother convergence than the quadrinomial tree. For that reason, we have applied the BS smoothing method only to the quadrinomial lattice. We can observe that the smoothed quadrinomial tree almost coincides with the rotated lattice. In addition, the correlation between the underlying processes has small influence on the estimated option value, in contrast to the two-factor interest rate case we investigate in Chapter 1.

Further, we price a European Call option with $K = 40$, $S_0 = 36$, $\rho = -0.8$ and the rest of the parameters set as before. We plot the results in Figure 3.3 a). In addition, the results for a European Call option with $K = 40$, $S_0 = 30$ and $\rho = 0.8$ are plotted in Figure 3.3 b).

**Example 3.4.1.** *To demonstrate the advantage of the rotated tree construction methods, we consider a simplified example of pricing options on two stocks which can be priced with a closed-form solution. For this purpose, let us assume that the dynamics of the stock prices are given*

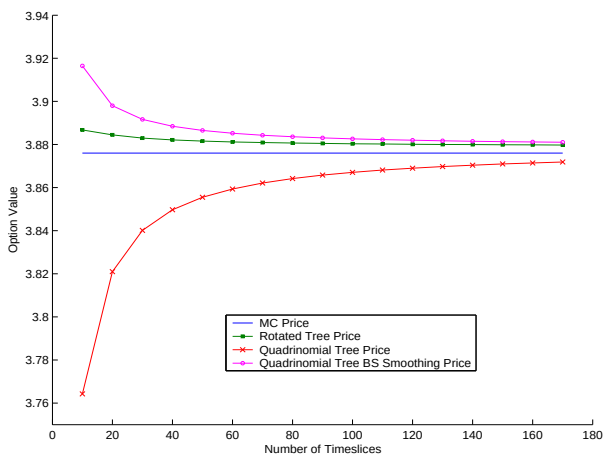



(a) European Put Option Price with Strike $K = 40$ and $S_0 = 30$.

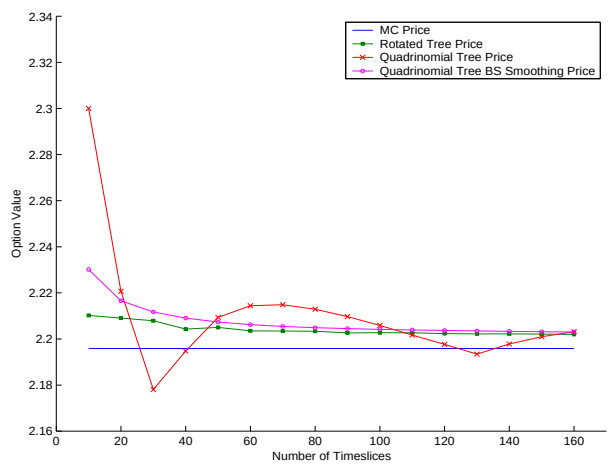



(b) European Put Option Price with Strike $K = 40$ and $S_0 = 40$.

Figure 3.1: Comparison of the Monte Carlo simulation, rotated tree or quadrinomial tree approximated prices of a European Put option. The correlation is set to $\rho = 0.8$.

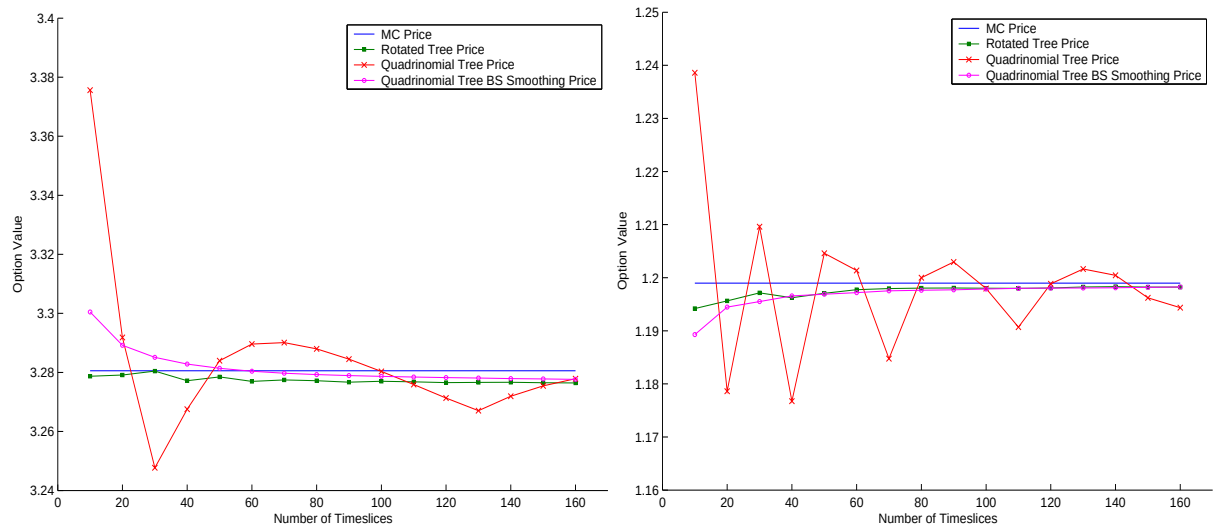



(a) European Put Option Price with Strike $K = 40$ and $S_0 = 40$.




(b) European Put Option Price with Strike $K = 36$ and $S_0 = 40$.

Figure 3.2: Comparison of the Monte Carlo simulation, rotated tree or quadrinomial tree approximated prices of a European Put option. The correlation is set to $\rho = -0.8$.




(a) European Call Option Price with $K = 40$, $S_0 = 36$ and $\rho = -0.8$.

(b) European Call Option Price with $K = 30$, $S_0 = 40$ and $\rho = 0.8$.

Figure 3.3: Comparison of the convergence of the numerical price of a European Call option delivered by Monte Carlo simulation, Rotated tree or Quadrinomial tree.

*directly under the risk-neutral measure $Q$ as*

$$\begin{aligned} dS_1(t) &= S_1(t)\left[rdt + \sigma_1 dW_1(t)\right], \quad S_1(0) = S_1^0 \\ dS_2(t) &= S_2(t)\left[rdt + \sigma_2 dW_2(t)\right], \quad S_2(0) = S_2^0 \end{aligned}$$

*where $W_1(t)$ and $W_2(t)$ are Brownian motions under the risk-neutral measure $Q$ with correlation $\rho$, $r$ is the risk-neutral interest rate and $\sigma_1 > 0$ and $\sigma_2 > 0$ are the volatilities of the stock price processes.*

*Further, for our demonstration purposes we price two type of options:*

1. *A European Call Option with strike $K$, on the maximum of the two stocks and with payoff at maturity $T$ given by*

$$(max(S_1(T), S_2(T)) - K)^+ .$$

2. *A digital Call Option with strike $K$ and payoff at maturity $T$ given by*

$$\mathbb{1}_{\{S_1(T) \geq K,\, S_2(T) \geq K\}}.$$

*In addition to the offered rotated and quadrinomial (Cox, Ross and Rubinstein [26] type) tree constructions, we examine a quadrinomial tree with alternative construction procedure (offered by Hull [33]), which consists of fixing* $p_{uu} + p_{ud} = \frac{1}{2}$ *and* $p_{uu} + p_{du} = \frac{1}{2}$*, relaxing the Cox, Ross and Rubinstein [26] assumption* $u = \frac{1}{d}$ *and finding the new up and down jumps to be*

$$\begin{aligned} u_1 &= \left(r - \frac{1}{2}\sigma_1^2\right)\Delta t + \sigma_1\sqrt{\Delta t}, \quad d_1 = \left(r - \frac{1}{2}\sigma_1^2\right)\Delta t - \sigma_1\sqrt{\Delta t} \\ u_2 &= \left(r - \frac{1}{2}\sigma_2^2\right)\Delta t + \sigma_1\sqrt{\Delta t}, \quad d_2 = \left(r - \frac{1}{2}\sigma_2^2\right)\Delta t - \sigma_1\sqrt{\Delta t}. \end{aligned}$$

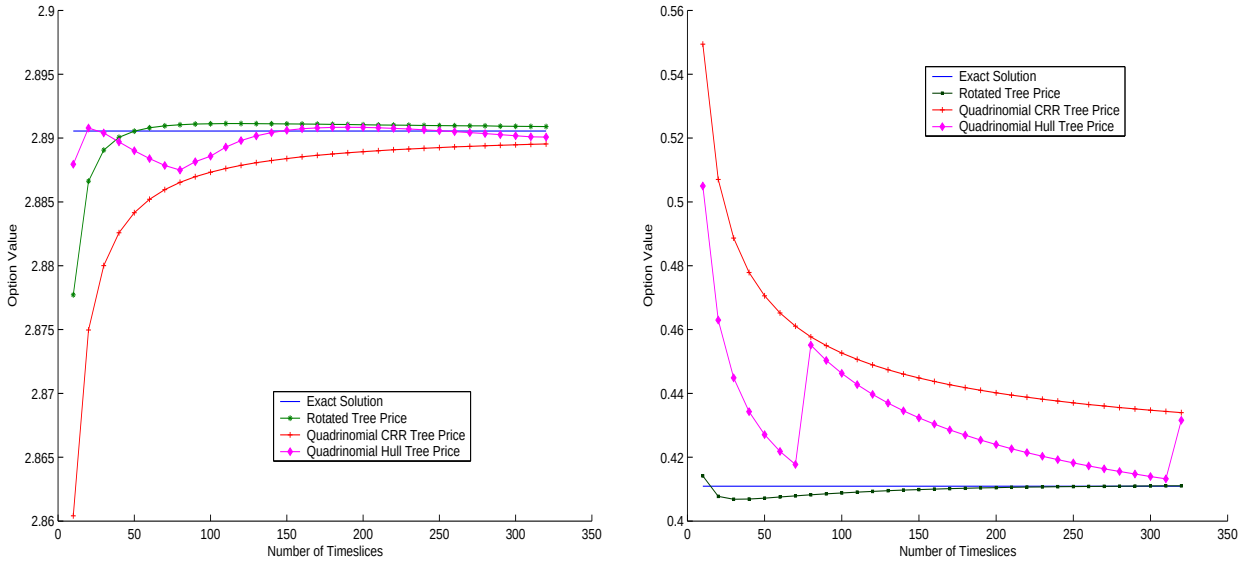
*We have denoted by* $u_1$ *and* $d_1$ *respectively the up and down jumps of* $\log(S_1(t))$ *and by* $u_2$ *and* $d_2$ *the up and down jumps of* $\log(S_2(t))$*. The new probabilities are found as*

$$\begin{aligned} p_{uu} &= \frac{1}{4}(1+\rho) \quad p_{ud} = \frac{1}{4}(1-\rho) \\ p_{dd} &= \frac{1}{4}(1+\rho) \quad p_{du} = \frac{1}{4}(1-\rho). \end{aligned}$$

*Notice that the alternative quadrinomial tree construction procedure yields well-defined probabilities, but the values of stocks at the central nodes at all time slices after* $2\Delta t$ *will be different than their values at time* $0$*. In contrast, the Cox, Ross and Rubinstein [26] (CRR) construction procedure produces a symmetrical around the central node lattice. In addition, in the case of mean-reversion in the approximated processes we cannot use (at least directly) such an alternative construction in the adaptive way we have presented in this Chapter 1 (in the sense of using node-dependent jump heights).*

*We set* $r = 5\%$*,* $\sigma_1 = 10\%$*,* $\sigma_2 = 30\%$*,* $T = 0.25$*,* $S_1(0) = S_2(0) = 40$*,* $K = 40$*,* $\rho = 0.7$ *and obtain for the European Call option on the maximum of two assets Figure 3.4 a) (with exact solution* $2.89055$ *for which we refer to Zvan, Forsyth and Vetzal [67]) and for the European digital Call option Figure 3.4 b) (with exact solution* $0.410929$ *for which we refer again to Zvan, Forsyth and Vetzal [67]).*

The results we obtain for Example 3.4.1 (in the case of a simple vanilla option), exhibit much smoother convergence of the rotated tree in comparison to both the quadrinomial CRR and Hull lattices. It can also be concluded that the pricing error of the rotated lattice is of a similar order as the Hull lattice.




(a) European Call option on the maximum of two assets.

(b) European digital Call option.

Figure 3.4: Comparison of the convergence of the numerical prices delivered by the Rotated and the Quadrinomial tree.

By the digital option pricing, the rotated tree results clearly outperform both quadrinomial tree constructions. The Hull lattice again shows better results than the CRR one. However, it also exhibits a saw-tooth convergence effect due to the non-symmetric lattice. We notice that the same results were also observed by Zvan, Forsyth and Vetzal [67] for a variety of finite-difference construction although they claim that a rotated mesh is not appropriate when interpolation is needed, e.g.in cases of discrete dividends. In this respect, we have seen in the previous chapter that there might be a small deterioration of the rotated tree in the case of discrete payments and positive correlation coefficient, however the deterioration of the quadrinomial lattice in cases of high volatility and correlation parameters is much stronger.

Finally, this allows us to claim that rotation of a two-dimensional lattice delivers faster and smoother convergence, independent of the underlying processes (but as long as they allow constant coefficients representation or transformation), compared to the direct quadrinomial lattice. However, in this Chapter, we have considered pricing options only of European type and in the next one we will further investigate the behavior of the two-dimensional lattices for pricing

American options.

# Chapter 4

# Pricing American Option with an Interest Rate Process of Cheyette Type

## 4.1 Introduction

The American options are contingent claims that allow exercise at any instant until the maturity of the option. The problem of their pricing can be transformed to a *free boundary* problem (see McKean [48]) or equivalently be written as an *optimal stopping* problem (see McKean [48], Bensoussan [6], Karatzas [39]).

Regardless of the various representations of the American option value, due to the free boundary it is not possible to find its closed form solution and thus numerical methods are required. Apart from the well-known finite difference (or in specific lattice) methods for which an extensive literature is available, there is a relatively new theoretical stream which consists of transformation of the American option pricing problem to a forward-backward stochastic differential equation with reflection and its simulation with the help of Monte Carlo methods (see Touzi and Buchard [11], Zhang [65], Longstaff and Schwartz [47], Clément, Lamberton and Protter [23], Andersen [2] etc.). All these methods are used to estimate an exercise strategy (not necessarily the optimal one!) and thus can be seen as lower bounds of the American option price. As an only reference of an upper bound of the price of an American option, we cite Rogers [59].

In this Chapter, our main aim is to price an American put option on a stock price with stochastic

interest rate of Cheyette type, using the tree construction procedures from the previous Chapter and compare the results to Monte Carlo simulation using the adapted (for the stochastic interest rate of Cheyette type) algorithm of Longstaff and Schwartz [47].

## 4.2 Transformation of the Problem as a FBSDE with Reflection

The goal of this section is to present the American option pricing problem for the given stochastic processes as a Forward-Backward stochastic differential equation (FBSDE) with reflection and to comment on the difficulties of its Monte Carlo simulation.

Let us assume again we have a complete probability space $\left(\Omega, \{\mathfrak{F}\}_{t\in[0,T]}, P\right)$ and the stock price process and the riskless interest rate process follow the same dynamic under an equivalent martingale measure $Q$ as given in the Chapter 3.

Let us further consider the case of a simple American put option with strike $K$ and maturity $T$. Following Touzi and Buchard [11], we can rewrite the pricing problem as a Forward-Backward stochastic differential equation given by

$$
\begin{aligned}
dS(t) &= S(t)\left[r(t)dt + \sigma_1 dW^1(t)\right], \quad S(0) = S_0 && (4.1)\\
dr(t) &= \mu(t,\phi,r)dt + \sigma_f(t,t)dW^2(t), \quad r(0) = r_0 && (4.2)\\
d\phi(t) &= \left[\sigma_f^2(t,t) - 2\kappa(t)\phi(t)\right]dt \\
\mu(t,\phi,r) &= \kappa(t)\left[f(0,t) - r(t)\right] + \phi(t) + \frac{d}{dt}f(0,t) \\
dY(t) &= r(t)Y(t)dt - Z_t dW^1(t), \quad Y(T) = (K - S(T))^+ && (4.3)\\
Y(t) &> (K - S(t))^+ && (4.4)
\end{aligned}
$$

where $\sigma_f(t,t) = \sigma_2 r(t)^\gamma$, $\gamma \geq 0$, $\sigma_2 \in \mathbb{R}^+$ and the instantaneous correlation between the two processes is given by $d\langle W^1(\cdot), W^2(\cdot)\rangle_t = \rho dt$. Thus, one can notice that the forward components are the stock price $S(t)$ and the interest rate $r(t)$ and the backward component is the price of

the American option which is denoted by $Y(t)$. The system can also be rewritten as

$$S(t) = S(0) + \int_0^t r(u)S(u)du + \int_0^t S(u)dW^1(u) \tag{4.5}$$

$$Y(t) = g(S_T) + \int_t^T r(u)Y(u)du - \int_t^T dW^1(u) + A_T - A_t \tag{4.6}$$

$$Y(t) \geq g(S_t), \quad t \in [0,T] \tag{4.7}$$

where $(S_t, Y_t, r_t, Z_t)$ should be adapted with respect to the complete Brownian filtration $\{\mathfrak{F}\}_{t\in[0,T]}$, $Y_t \in L^2$ for $\forall t \in [0,T]$, $Z_t \in L^2([0,T])$, $g(S_t) := (K - S_t)^+$ and $A$ is a non-decreasing càdlàg process such that

$$\int_0^T (Y_t - (K - S_t)^+)dA_t = 0, \quad A_0 = 0.$$

The intuitive meaning of the $A$-process is to "push" the solution of the BSDE upwards, so that it remains above the obstacle which in our case is the early exercise boundary. It has also the financial meaning of the cost for early exercise of the American option.

Let us now assume that $\pi$ is an arbitrary discretization of $[0,T]$ such that $0 = t_0 < t_1 \ldots < t_n = T$ and $|\pi| := \max_{1\leq i\leq n} |t_i - t_{i-1}|$. Let us also denote the appropriate discrete filtration by

$$\mathfrak{F}_i^\pi := \sigma\left(S_{t_j}^\pi, r_{t_j}^\pi\right)_{j\leq i}, \quad i = 1, \ldots, n.$$

Then, for the **forward components** of the reflected FBSDE we use the classical Euler scheme of discretization. Thus, the discretized Stock price and interest rate process become

$$\begin{aligned}
r_{t_{i+1}}^\pi &= r_{t_i}^\pi + \left[\kappa\big(f(0,t_i) - r_{t_i}^\pi\big) + \phi_{t_i}^\pi + \frac{d}{dt}f(0,t_i)\right]\Delta t_i^\pi + \sigma_2\left(r_{t_i}^\pi\right)^\gamma \Delta^\pi \widetilde{W}_i^2 \\
\phi_{t_{i+1}}^\pi &= \phi_{t_i}^\pi + \left[\sigma_2^2\left(r_{t_i}^\pi\right)^{2\gamma} - 2\kappa\phi_{t_i}^\pi\right]\Delta^\pi t_i^\pi \\
S_{t_{i+1}}^\pi &= S_{t_i}^\pi + S_{t_i}^\pi\left[r_{t_i}^\pi \Delta^\pi t_i + \sigma_1\left(\rho\Delta^\pi \widetilde{W}_i^2 + \sqrt{1-\rho^2}\Delta^\pi \widetilde{W}_i^1\right)\right]
\end{aligned}$$

with $S(t_0) = S_0$, $r(t_0) = r_0$ , $\varphi(t_0) = 0$, $\Delta_i = t_i - t_{i-1}$ and $\Delta^\pi \widetilde{W}_i := \widetilde{W}(t_{i+1})^\pi - \widetilde{W}(t_i)^\pi$, $i = 1, \ldots, n$. Notice that we have decomposed the correlated Brownian motions $W^1$ and $W^2$ into a sum of the independent Brownian motions $\widetilde{W}^1$ and $\widetilde{W}^2$ using Cholesky decomposition. In this form, we notice that it is straightforward to simulate the forward components using Monte Carlo methods.

On the other side, we cannot directly use the Euler discretization scheme for the **backward component**

$$Y^{\pi}_{t_i} - Y^{\pi}_{t_{i-1}} = r^{\pi}_{t_{i-1}} Y^{\pi}_{t_{i-1}} \Delta^{\pi}_i - Z^{\pi}_{t_{i-1}} \Delta^{\pi}_i W^1_i \tag{4.8}$$

since then, the $Y^{\pi}_{t_{i-1}}$ will be given by

$$Y^{\pi}_{t_{i-1}} = \frac{1}{\left(1 + r^{\pi}_{t_{i-1}\Delta^{\pi}_i}\right)} Y^{\pi}_{t_i} - \frac{1}{\left(1 + r^{\pi}_{t_{i-1}\Delta^{\pi}_i}\right)} Z^{\pi}_{t_{i-1}} \Delta^{\pi}_i W^1_i \tag{4.9}$$

which is obviously not measurable w.r.t. $\mathfrak{F}^{\pi}_{t_{i-1}}$. Therefore, in order to obtain a measurable backward simulation scheme, we need to take an expectation of (4.9) conditioned on $\mathfrak{F}^{\pi}_{t_{i-1}}$. If we consider in addition the reflection boundary, we obtain the following recursion

$$\begin{aligned} Y^{\pi}_T &= (K - S_T)^+ \\ Y^{\pi}_{t_{i-1}} &= \max\left( \frac{1}{1 + r^{\pi}_{t_{i-1}} \Delta^{\pi}_i} E^Q\left( Y^{\pi}_{t_i} \middle| \mathfrak{F}^{\pi}_{t_{i-1}} \right), \left( K - S^{\pi}_{t_{i-1}} \right) \right) \\ &\approx \max\left( E^Q\left( \exp\left( -r^{\pi}_{t_{i-1}} \Delta^{\pi}_i \right) Y^{\pi}_{t_i} \middle| \mathfrak{F}^{\pi}_{t_{i-1}} \right), \left( K - S^{\pi}_{t_{i-1}} \right)^+ \right). \end{aligned} \tag{4.10}$$

To obtain the $Z$-component, we multiply (4.8) with $\Delta^{\pi} W^1_i$ and take again expectation with respect to $\mathfrak{F}_{t_{i-1}}$. Thus,

$$Z^{\pi}_{t_{i-1}} = \frac{1}{\Delta^{\pi}_i} E\left( Y^{\pi}_{t_i} \Delta^{\pi} W^1_i \middle| \mathfrak{F}_{t_{i-1}} \right).$$

However, since we are only interested in the simulation of the American option price, we can ignore the $Z$-component as it has only connection to the hedging strategy. The interesting question now is how to estimate the conditional expectation, we need for simulating the $Y$-process. There are generally three approaches in the literature:

- Nonparametric Regression (Kernal estimator) : Carrièr [16];
- Maliavin Calculus : Nizar Touzi and Bruno Buchard [11], Zhang [65];
- Linear Regression : Longstaff and Schwartz [47] which we abbreviate to (LS), Clément, Lamberton and Protter [23], Andersen [2];

We are not going to comment here about the first method and for the second one we are going to give only a motivation for the case of zero interest rates. Notice that

$$\begin{aligned} E^Q(Y_{t_i}|\mathfrak{F}_{t_{i-1}}) &= E^Q(v(W_{t_i})|W_{t_{i-1}} = \omega_{t_{i-1}}) \\ &= \frac{E^Q\left(v(W_{t_i})\mathbb{1}_{\{W_{t_{i-1}}=\omega_{t_{i-1}}\}}\right)}{E^Q\left(\mathbb{1}_{\{W_{t_{i-1}}=\omega_{t_{i-1}}\}}\right)} \end{aligned}$$

where we have denoted $Y_{t_i} := v(W_{t_i})$ and have used w.l.g. that due to the Markovian property of the stock prices $\mathfrak{F}_{t_{i-1}} = S_{t_{i-1}} = W_{t_{i-1}}$. A straightforward Monte Carlo simulation of the conditional expectation using $N$ paths will have the form

$$E^Q\left(Y_{t_i}|\mathfrak{F}_{t_{i-1}}\right) \approx \frac{1}{N}\sum_{j=1}^{N}\frac{v\left(W_{t_i}^j\right)\mathbb{1}_{\left\{W_{t_{i-1}}^j=\omega_{t_{i-1}}\right\}}}{\mathbb{1}_{\left\{W_{t_{i-1}}^j=\omega_{t_{i-1}}\right\}}}$$

with the simulated Dirac measure always taking values of zero and thus causing the Monte Carlo approximation to fail. A way to overcome this problem is offered by Bouchard, Ekeland and Touzi [10] and consists of transforming the conditional expectation using Malliavin Calculus to

$$E^Q\left(Y_{t_i}|\mathfrak{F}_{t_{i-1}}\right) = \frac{E^Q\left(v(W_{t_i})\,h\left(W_{t_{i-1}},\omega_{t_{i-1}}\right)\right)}{E^Q\left(h\left(W_{t_{i-1}},\omega_{t_{i-1}}\right)\right)}$$

where $h(,)$ denotes the Heaviside function defined as $h(x,y) = \mathbb{1}_{x\geq y}$ and which allows (in a more sophisticated way) the usage of Monte Carlo simulation (see Touzi and Buchard [11]).

## 4.3 Simulation of a FBSDE with Reflection with the LS Algorithm

The core of the Monte Carlo simulation of American Options is the estimation of the conditional expectation of its continuation value, needed for the simulation of the backward component of the pricing FBSDE with reflection. Finding an algorithm that is able to do this allows us to completely specify the optimal exercise strategy of an American option for each simulated path of the underlying security.

In the following section, we are going to present the algorithm of Longstaff and Schwartz [47] for the estimation of the conditional expectation in (4.10). The very idea of it is to estimate the required conditional expectation via least squares minimization using the cross-sectional

information of the simulated paths. The method is generally easy to implement and also to adapt for the case of Cheyette interest rates.

### 4.3.1 Motivation

In the process of presenting the method of Logstaff and Schwartz [47], we need first to give a general motivation for using the least-squares method for an approximation of a conditional expectation. For that reason, referring to Brockwell and Davis [14], p.63, we give the following definition of a conditional expectation in the $L^2$ sense:

**Definition 4.3.1.** *If $Z_1, \ldots, Z_m$ are random variables on $(\Omega, \mathfrak{F}, P)$ and $Y \in L^2(\Omega, \mathfrak{F}, P)$, then the conditional expectation of $Y$ given $Z_1, \ldots, Z_m$ is defined by the projection*

$$E(Y|Z_1, \ldots, Z_m) = P_{\mathcal{M}(Z_1,\ldots,Z_m)}Y$$

*where $\mathcal{M}(Z_1, \ldots, Z_m)$ is the closed subspace of $L^2$ consisting of all random variables in $L^2$ which can be written in the form $\phi(Z_1, \ldots, Z_m)$ for some Borel function $\phi : \mathbb{R}^m \to \mathbb{R}$.*

On the other side, since the conditional expectation of $Y$ is an element of the space of square-integrable functions $L^2([0, T])$, which itself is a separable Hilbert space, there exists a countable orthonormal basis of $\mathcal{M}(Z_1, \ldots, Z_m)$ and the conditional expectation can be presented as a linear combination of the elements of the basis. Thus,

$$E(Y|Z_1, \ldots, Z_m) = \sum_{j=0}^{\infty} a_j f_j, \quad a_j \in \mathbb{R}, \quad f_j(\cdot) : \mathbb{R} \to \mathbb{R}$$

where the basis functions $f_1(\cdot), \ldots, f_M(\cdot)$ can be taken to be polynomials in $Z_1, \ldots, Z_m$ such as the Laguerre, Hermite, Chebyshev, Jacobi or other polynomials.

Since the spanning $\overline{sp}\,(1, f_1(\cdot), \ldots, f_\infty(\cdot)) = \mathcal{M}(Z_1, \ldots, Z_m)$, we would obtain from both presentations of the conditional expectation that

$$E(Y|Z_1, \ldots, Z_m) = P_{\mathcal{M}(Z_1,\ldots,Z_m)}Y = P_{\overline{sp}\,(1, f_1(Z_1,\ldots,Z_m),\ldots,f_\infty(Z_1,\ldots,Z_m))}Y$$

which transforms the problem to an easier one since now we are looking for a projection on a set of countable basis.

For practical purposes, we use only the first $M$ elements of the latter sum. We can find the estimators $\widehat{a}_1, \ldots, \widehat{a}_M$ of the needed parameters using the following least squares minimization

$$\min_{a_1,\ldots,a_M} \sum_{i=1}^{N} \left( Y_i - \sum_{j=0}^{M} a_j f_j(Z_1^j, \ldots, Z_m^j) \right)^2$$

for some observations $(Y_1, \ldots, Y_N)$, $\left(\left(Z_1^1, \ldots, Z_m^1\right), \ldots, \left(Z_1^N, \ldots, Z_m^N\right)\right)$, $N \in \mathbb{N}$. Thus, we obtain the following approximation of the conditional expectation

$$\widehat{E}(Y|Z_1, \ldots, Z_m) = \sum_{j=1}^{M} \widehat{a}_j f_j(Z_1, \ldots, Z_m).$$

### 4.3.2 The Longstaff and Schwartz Algorithm

In this section, we will shortly present the adaptation of the Longstaff and Schwartz [47] algorithm for Monte Carlo simulation of American options to the case of stochastic interest rate of Cheyette type.

The general idea of the LS algorithm is first to simulate using MC the forward components of the FBSDE and then starting at maturity, to calculate stepwise backwards the optimal stopping rule using the reflection principle with continuation value that has been estimated via Least-squares minimization which on its turn takes into account the information of all simulated paths.

Assuming we perform $N \in \mathbb{N}$ simulations of the underlying process we have now two steps:

**Step 1.** At maturity $t_n = T$, the holder of an American option has the choice only to exercise the option if it is in the money or leave it if it is not. The situation is exactly the same as by the European type of options. Thus, we are going to construct an optimal stopping strategy, starting from the maturity so that at time $t_n$, we have $N$ decisions to make whether to exercise or not.

Let us keep in an $n \times N$ matrix $\Lambda = (\alpha)_{i,j}$ the optimal stopping strategy, so that at the end of the algorithm we will have for each time point $t_i$, $i = 1 \ldots, n$ and each path $\{j\}$, $j = 1, \ldots, N$

$$\alpha_{i,j} \quad := \quad \begin{cases} 1, & \text{if we stop} \\ 0 & \text{if we continue.} \end{cases}$$

Thus, at maturity $t_n$ we have for $j = 1, \ldots, N$

$$\alpha_{n,j} \quad := \quad \begin{cases} 1, & \text{if}\,(K - S_{t_n}) > 0 \\ 0 & \text{else.} \end{cases}$$

Further, for each time point until time $t_0$ we have to make the choice whether to continue keeping the option or exercise it. In order to estimate the continuation value, as shown in the motivation for the algorithm, we choose a quadratic function for the least squares regression. The algorithm is easily generalized for any polynomial.

**Step 2.** Assume that we are at time $t_{i-1}$ in the simulation, where we have already found the optimal stopping strategy from time $t_i$ to the maturity $t_n$. Thus, we have found all $\alpha_{i,j}, \ldots, \alpha_{n,j}$, where for each simulated path only one of them can be 1, since we can exercise the American option only once.

Next, looking for the optimal stopping rule at time $t_{i-1}$, we need to find the continuation value, which can be written for each path $j = 1, \ldots, N$ as

$$E^Q \left( \exp\left(-r_{t_{i-1}} \Delta_i\right) Y_{t_i} \middle| \mathfrak{F}^j_{t_{i-1}} \right). \tag{4.11}$$

Due to the fact that the stock price follows a Markov process and the Cheyette interest rate is an ingredient in a Markov process, (4.11) can be rewritten as

$$E^Q \left( \exp\left(-r_{t_{i-1}} \Delta_i\right) Y_{t_i} \middle| S^j_{t_{i-1}}, r^j_{t_{i-1}}, \phi^j_{t_{i-1}} \right). \tag{4.12}$$

Notice also that since by definition $\phi_{t_{i-1}}$ is uniquely determined by $r_{t_{i-2}}$ and $\phi_{t_{i-2}}$ and we condition only on the last state and not on the whole history of the process, we can thus omit it from the condition.

Next, for the following regression we need the discounted optimal payout of the American option for each path $\{j\}$ at time $t_i$, i.e. the value of $\exp\left(-r^j_{t_{i-1}} \Delta_i\right) Y^j_{t_i}$ and we notice that although we do not keep its running value, we have found the optimal strategy from time $t_i$ until the maturity $t_n$ for each simulated path. Hence, we can use the information hidden in it to find the possible payoff of the American option for each path. Therefore, let us denote the discounted up to time $t_{i-1}$ ex-post realized payoff for path $j$ from continuation from time $t_i$ to maturity $t_n$

by

$$\mathbf{Y} := \begin{pmatrix} Y_i^1 \\ \vdots \\ Y_i^N \end{pmatrix}$$

and for each path $j = 1, \ldots, N$ at time $t_i$ find it as

$$Y_i^j = \sum_{k=i}^{n} \alpha_{kj} \left(K - S_{t_k}^j\right)^+ \exp\left(-\sum_{p=i-1}^{k} r_{t_p}^j \Delta p\right).$$

Further, we denote with $\mathbf{X}$ the matrix of the current values of the underlying two processes needed for the regression such that

$$\mathbf{X} := \begin{pmatrix} \left(S_{t_{i-1}}^1\right)^2 & \left(r_{t_{i-1}}^1\right)^2 & S_{t_{i-1}}^1 & r_{t_{i-1}}^1 & S_{t_{i-1}}^1 r_{t_{i-1}}^1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \left(S_{t_{i-1}}^N\right)^2 & \left(r_{t_{i-1}}^N\right)^2 & S_{t_{i-1}}^N & r_{t_{i-1}}^N & S_{t_{i-1}}^N r_{t_{i-1}}^N & 1 \end{pmatrix}$$

If we denote $\mathbf{S^2} := \left(\left(S_{t_{i-1}}^1\right)^2, \ldots, \left(S_{t_{i-1}}^N\right)^2\right)'$, $\mathbf{r^2} := \left(\left(r_{t_{i-1}}^1\right)^2, \ldots, \left(r_{t_{i-1}}^N\right)^2\right)'$, $\mathbf{S} := \left(S_{t_{i-1}}^1, \ldots, S_{t_{i-1}}^N\right)'$, $\mathbf{r} := \left(r_{t_{i-1}}^1, \ldots, r_{t_{i-1}}^N\right)'$, $\mathbf{Sr} := \left(S_{t_{i-1}}^1 r_{t_{i-1}}^1, \ldots, S_{t_{i-1}}^N r_{t_{i-1}}^N\right)'$, $\mathbb{1} := (1, \ldots, 1)'$, $\Theta := (a_1, a_2, b_1, b_2, c, d)'$, then we can write the mean squared error as

$$\begin{aligned} S(\Theta) &= \sum_{j=1}^{N} \left(Y_i^j - a_1 \left(S_{t_{i-1}}^j\right)^2 - a_2 \left(r_{t_{i-1}}^j\right)^2 - b_1 S_{t_{i-1}}^j - b_2 r_{t_{i-1}}^j - c S_{t_{i-1}}^j r_{t_{i-1}}^j - d\right)^2 \\ &= \|\mathbf{Y} - a_1 \mathbf{S^2} - a_2 \mathbf{r^2} - b_1 \mathbf{S} - b_2 \mathbf{r} - c \mathbf{Sr} - d \mathbb{1}\|^2 \end{aligned}$$

and find by the Projection Theorem that there is a unique vector $\widehat{\Theta}$

$$\widehat{Y} = \mathbf{X}\widehat{\Theta} = \left(\widehat{a}_1 \mathbf{S^2} + \widehat{a}_2 \mathbf{r^2} + \widehat{b}_1 \mathbf{S} + \widehat{b}_2 \mathbf{r} + \widehat{c} \mathbf{Sr} + \widehat{d} \mathbb{1}\right)$$

which minimizes $S(\Theta)$, such that

$$\widehat{\Theta} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{Y}. \tag{4.13}$$

Notice that if $\mathbf{X}'\mathbf{X}$ is a singular matrix there are infinitely many solutions of $\widehat{Y} = \mathbf{X}\widehat{\Theta}$ but $\mathbf{X}\widehat{\Theta}$ is the same for all of them.

Having found an estimation for (4.12) for each path $j = 1, \ldots, N$, such that

$$\begin{aligned} \widehat{E}^Q \left(\exp\left(-r_{t_{i-1}} \Delta_i\right) Y_{t_i} \middle| S_{t_{i-1}}^j, r_{t_{i-1}}^j\right) &= \widehat{a}_1 \left(S_{t_{i-1}}^j\right)^2 + \widehat{a}_2 \left(r_{t_{i-1}}^j\right)^2 + \widehat{b}_1 S_{t_{i-1}}^j \\ &\quad + \widehat{b}_2 r_{t_{i-1}}^j + \widehat{c} S_{t_{i-1}}^j r_{t_{i-1}}^j + \widehat{d} \end{aligned}$$

we obtain directly the optimal stopping rule for the time $t_{i-1}$ as

$$\alpha_{i-1,j} \quad := \quad \begin{cases} 1, & \text{if} \quad \left(K - S^j_{t_{i-1}}\right)^+ > \widehat{E}^Q\left(\exp\left(-r_{t_{i-1}}\Delta_i\right) Y_{t_i} \middle| S^j_{t_{i-1}}, r^j_{t_{i-1}}\right) \\ 0 & \text{else.} \end{cases}$$

It is easy to notice that if $\left(K - S^j_{t_i}\right) \leq 0$ then the immediate exercise value of the American option is zero and in that case the option will **not** be exercised. Thus, for all out of the money payoffs we do not need to calculate the continuation value as for them we can directly set the respective $\alpha$ equal to zero. Longstaff and Schwartz [47] showed that this leads to a significant increase in the efficiency of the algorithm. Practically, instead of $N$ elements, the vector $Y$ will have for each time $t_i$ only $N^*_{j-1} \leq N$ elements such that

$$N^*_{j-1} = \sum_{j=1}^{N} \mathbb{1}_{\left\{\left(K - S^j_{t_{i-1}}\right) > 0\right\}}.$$

In addition, we should notice that as soon as it becomes optimal to exercise, the American option will be exercised and therefore stops existing. Thus, if $\alpha_{i-1,j} = 1$ then $\alpha_{i,j} = \alpha_{i+1,j} = \ldots = \alpha_{n,j} = 0$ for all $j = 1, \ldots, N$. The interpretation is that for each simulated path, the American option can be exercised at most once and if exercised, the payoff is paid and the option does not exist anymore.

**Remark 4.3.1.** *Notice that although the stock price and the interest rate are correlated, there will be no effect on the estimated expectation since the fitted value of the regression is unaffected by the degree of correlation between the explanatory variables.*

After we have found the optimal stopping strategy for all paths $j = 1, \ldots, N$ and all discrete time points $t_1, \ldots, t_n$, the approximated value of the American option at time $t_0$ is given by

$$Y_0 = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{N} \alpha_{ij} \exp\left(-\sum_{k=0}^{i} r^j_{t_k} \Delta_k\right) \left(K - S^j_{t_i}\right)^+.$$

## 4.4 The Lattice Approach

Using trees is the standard way to deal with reflected FBSDE or with linear programming problems. Both lattices from Chapter 3 are also applicable for pricing American options with an interest rate process of Cheyette type.

As in the previous section, we can calculate the value of an American type of option by the following backward routine

$$\begin{aligned} Y_T &= (K - S_T)^+, \quad \text{for all nodes of the tree at maturity} \\ Y^j_{t_{i-1}} &= \max\left(\exp\left(-r^j_{t_{i-1}}\Delta_i\right) E^Q\left(Y_{t_i}\middle|\mathfrak{F}^j_{t_{i-1}}\right), \left(K - S^j_{t_{i-1}}\right)^+\right), \quad \text{for node } \{j\} \text{ at time } t_{i-1}. \end{aligned}$$

The conditional expectation above is approximated as

$$E^Q\left(Y_{t_i}\middle|\mathfrak{F}^j_{t_{i-1}}\right) \approx p^j_{uu}Y^{uu}_{t_i,j} + p^j_{ud}Y^{ud}_{t_i,j} + p^j_{du}Y^{du}_{t_i,j} + p^j_{dd}Y^{dd}_{t_i,j}$$

where $p^j_{uu}, p^j_{ud}, p^j_{du}$ and $p^j_{dd}$ are the risk-neutral probabilities at node $\{j\}$ calculated by the lattice construction methods presented in Chapter 3.

## 4.5 Numerical Results

We have calculated as a first numerical example an American put option with maturity $T = 1$ and strike $K = 36$ and where the process parameters are $\sigma_r = 23\%, \gamma = 1, \kappa = 0.083, \rho = 0.8, S_0 = 40, \sigma_s = 30\%$ and $r_0 = 80\%$. Using the Longstaff and Schwartz Monte Carlo simulation algorithm we have obtained the following two tables:

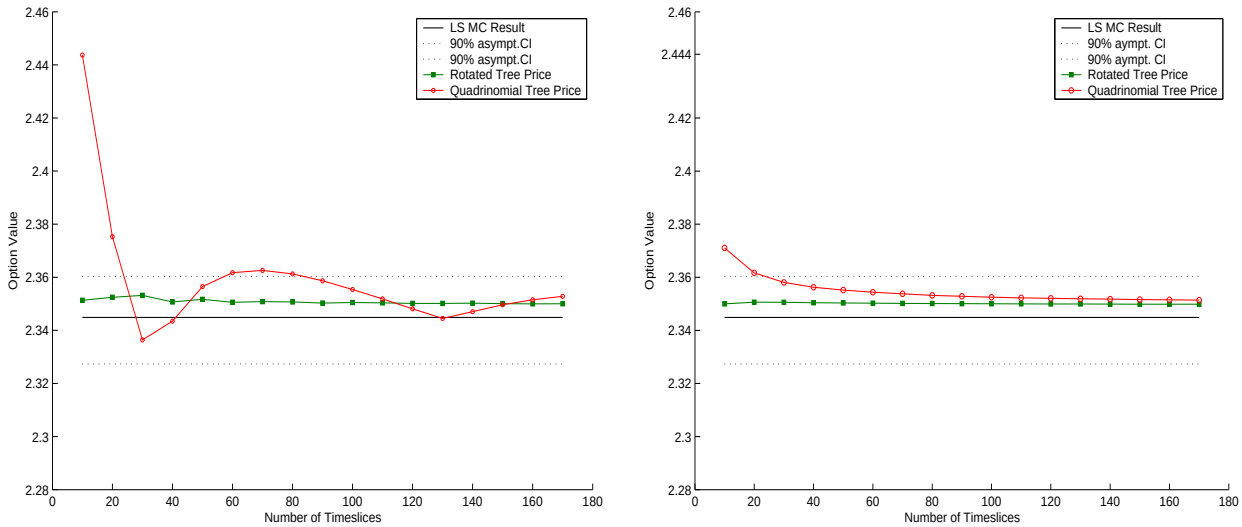
| **American Put Option, LS Monte Carlo Results** | | | |
|---|---|---|---|
| Number of Simulations | Discretization | MC | 90% asympt. CI |
| 1000 | 200 | 2,5150483 | [2.2755226,2.7545631] |
| 5000 | 200 | 2.4591275 | [2.3545425,2.5637043] |
| 10000 | 200 | 2.3621708 | [2.2912535,2.4330854] |
| 50000 | 200 | 2.3413370 | [2.3102514,2.3724324] |
| 100000 | 200 | 2.3419884 | [2.3201125,2.3638695] |
| 200000 | 200 | 2.3448583 | [2.3273747,2.3602747] |

| **American Put Option, LS Monte Carlo Results** | | | |
|---|---|---|---|
| Number of Simulations | Discretization | MC | 90% asympt. CI |
| 200000 | 50 | 2.323347 | [2.307922,2.338765] |
| 200000 | 100 | 2.341988 | [2.320113,2.363870] |
| 200000 | 150 | 2.332644 | [2.317256,2.348018] |
| 200000 | 200 | 2.344858 | [2.323347,2.360275] |

In the first table, we can observe the convergence of the result with the increase of the number of simulations and have cited in addition the 90% asymptotic confidence interval[1] (CI). We remark

[1] For an explanation of the MC estimation of the asymptotic confidence intervals, we refer to Chapter 6.

that even only with 200000 MC simulation and 200 time steps we are at the boundary of the current standard computer power (in the sense of computer time) but still we do not obtain sufficiently tight CI.




(a) Convergence of the trees with the increase of the number of time slices without BS smoothing.

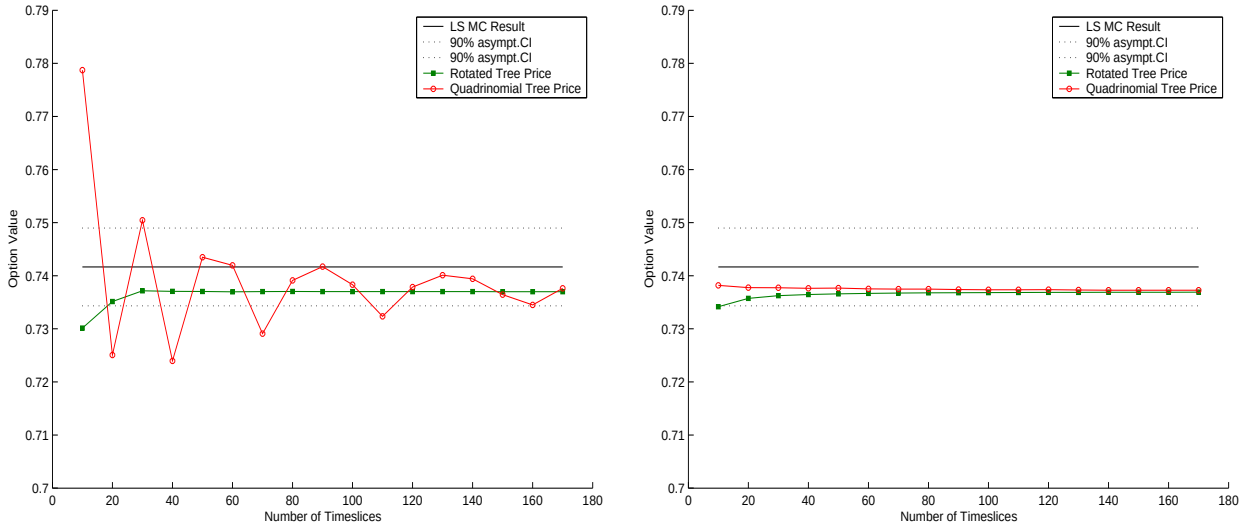
(b) Convergence of the trees with the increase of the number of time slices with BS smoothing.

Figure 4.1: Comparison of the convergence of the Rotated Tree and the Quadrinomial Tree in the calculation of an American Put Option with strike $K = 36$, maturity $T = 1$, $S_0 = 40$, $\rho = 0.8$. For the LS MC simulation we have used 200000 simulated paths and 200 timesteps.

The second table we have used to show that simply increasing the number of time steps cannot lead to improvement of the result if we keep the same number of MC simulations. In fact, it even leads to a deterioration of the confidence intervals since it increases the numerical error of the simulation (see e.g. Touzi and Bouchard [11] for an extensive analysis of the connection between the error of the regression and the error of the MC simulation to the total estimation error).

Further, we compare the performance of both trees for the same option and process parameters in Figure 4.1$a$). As an additional example we use $T = 1$, $K = 30$, $S_0 = 40$, $\rho = 0.6$, $\sigma_r = 23\%, \gamma = 1, \kappa = 0.083$ and $\sigma_s = 30\%$ and we plot the respective results in Figure 4.2$a$).

We notice that both trees deliver good results with insignificant additional time expense, compared to the European option pricing. We can also observe that the rotated tree delivers




(a) Convergence of the trees with the increase of the number of time slices without BS smoothing.

(b) Convergence of the trees with the increase of the number of time slices with BS smoothing.

Figure 4.2: Comparison of the convergence of the Rotated Tree and the Quadrinomial Tree in the calculation of an American Put Option with strike $K = 30$, maturity $T = 1$, $S_0 = 40$, $\rho = 0.6$. For the LS MC simulation we have used 200000 simulated paths and 200 timesteps.

smoother convergence, which was also the case for European options.

**Black-Scholes Smoothing:**

As already mentioned in the previous chapter, an improvement of the convergence the lattice constructions can be done using the method of Black-Scholes trees. This method, used for improving the convergence of the binomial tree for pricing of American options consists of substituting at the period exactly before the maturity of the option the continuation value with the Black-Scholes formula. The motivation is that at time $t_{n-1}$, the *continuation value* of the American claim is given by

$$\begin{aligned} E^Q\left(\exp\left(-r_{t_{n-1}}\Delta_{n-1}\right)Y_{t_n}\middle|\mathfrak{F}_{t_{n-1}}\right) &= E^Q\left(\exp\left(-r_{t_{n-1}}\Delta_{n-1}\right)(K-S_{t_n})^+\middle|\mathfrak{F}_{t_{n-1}}\right) \\ &= \mathrm{BS}\left(S_{t_{n-1}},\, r_{t_{n-1}},\, K,\, \Delta_{n-1}\right) \end{aligned}$$

where BS$(\cdot)$ denotes again the Black-Scholes price of an European option with time to maturity $\Delta_{n-1}$, spot price $S_{t_{n-1}}$, strike $K$ and constant interest rate $r_{t_{n-1}}$. Performing numerical simulation of the previous examples we obtain Figure 4.1*b*) and Figure 4.2*b*). Comparison to the results before the BS smoothing reveals again especially in the Quadrinomial tree construction

a significant improvement of the convergence.

## 4.6 Suggestions for Further Research Topics

Since now we have all component to price reasonably a convertible bond, a topic for further research can be the practical implementation of its pricing algorithm by adding in addition a jump-diffusion dynamic of the underlying stock price (recall the previous chapter). One can also adapt and test the algorithm of Andersen [2] for the case of stochastic interest rate of Cheyette type and compare its behavior to the Longstaff and Schwartz [47] one.

# Chapter 5

# Pricing Cumulative Parisian Call/Put Options with Modified Binomial/Trinomial Trees

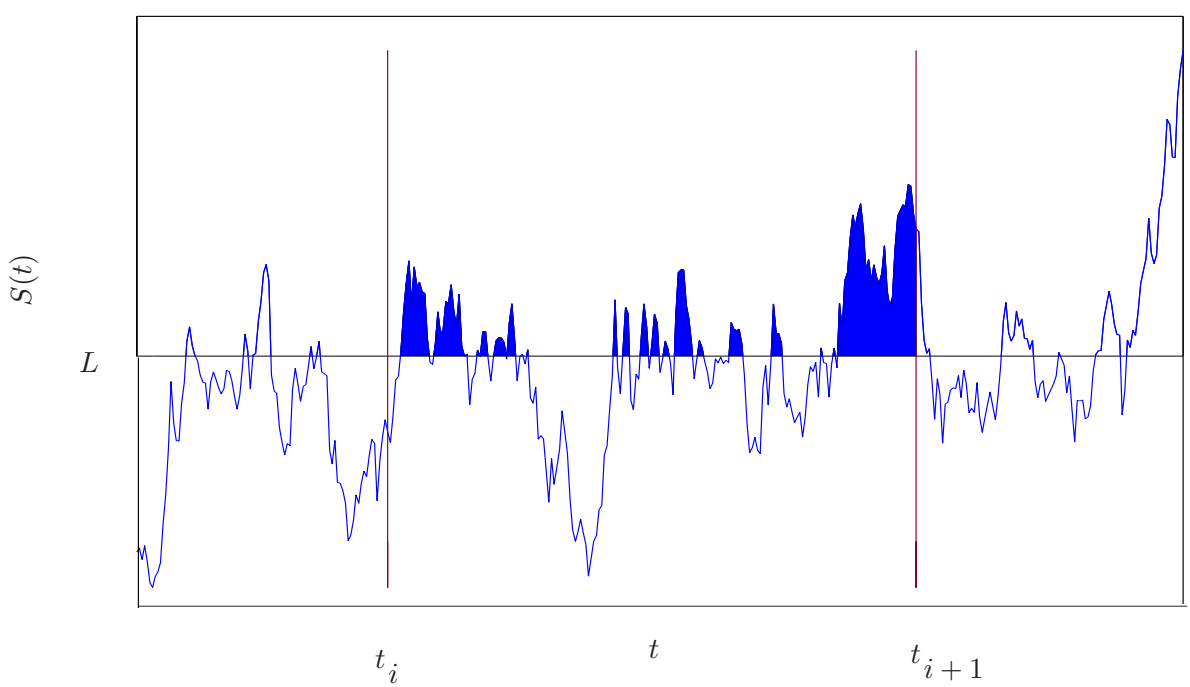
Both continuously and discretely monitored barrier options share the same drawback - they tempt either the option buyer or the option seller to try to influence the underlying stock price which leads to the so called "barrier wars". As a solution, the financial markets have offered several instruments that both preserve the barrier feature of the option and minimize the risk of manipulation. The oldest example is the so called "Asian barrier option" which is written on the weighted sum of the discretely observed underlying stock prices.

Another recent development of the financial markets is the "Parisian" option, which is knocked-in (or out) when the underlying stock price stays above (or below) the barrier a minimum period of time, referred to as a "window". In this case, we differentiate between continuous period of time (or consecutive days) and cumulative period of time (the sum of all days outside the barrier, not necessarily consecutive). For a quasi-analytical solution of the continuously monitored (also cumulative) Parisian barrier down-and-out option, we refer to Chesney, Jeanblanc-Picque and Yor [19] who define the option value in terms of an integral expressed as an inverse Laplace transform using the theory of Brownian excursions. Hugonnier [32] finds a quasi closed-form analytical solution for the continuously monitored cumulative Parisian barrier option and later Moraux [50] slightly corrects Hugonnier's quasi-analytical solution although he agrees with the correctness of his numerical implementation. Further, for a partial differential equation formula-

tion of the continuously monitored cumulative and consecutive Parisian barrier options we refer to Haber, Schönbucher and Wilmott [63] and Vetzal and Forsyth [62].

In this Chapter, we will introduce a new pricing method for the continuously monitored cumulative Parisian barrier option that combines a Cox, Ross and Rubinstein [26] binomial or trinomial lattice construction with a closed-form solution of the (conditional on the starting and ending stock value) knock-out probabilities using the theory of Brownian excursions. The presented method is both easy to implement and requires no extra interpolation. Further, it shows results very close to the analytical solution. For comparison with the existing lattice methods based on the Forward Shooting Grid algorithm for pricing continuously monitored Put or Call cumulative Parisian barrier options, we refer to Kwok and Lau [44].

**Definition 5.0.1.** *A continuously monitored "Cumulative Parisian Barrier-Up Call Option" (CPCall$^+$) with maturity $T$ and strike $K$ pays at maturity the amount $(S(T)-K)^+$ if in a pre-specified interval $[T_0, T], 0 \leq T_0 < T$ the stock price $S(t)$ remains "cumulatively" above a certain level $L$ time longer than a fraction $R \in \mathbb{R}$, $R \in (0,1)$ of the length of the interval.*




**Example 5.0.1.** *Usual case in the praxis is a "window" $[t_i, t_{i+1}]$ of 30 days and a fraction $R$ of "20 out of 30 days".*

## 5.1 Definition of the Problem

Let us assume we have a complete probability space $\left(\Omega, \{\mathfrak{F}\}_{t\in[0,T]}, \mathbb{P}\right)$ and an arbitrage-free market. Let us further assume that we are in the Black-Scholes world (see more about it in Appendix C) where the stock price follows a Geometric Brownian motion which is given directly under the unique risk-neutral measure $Q$ (equivalent to $\mathbb{P}$ with respect to $\{\mathfrak{F}_t\}$) as

$$\begin{aligned} dS_t &= S_t\left[(r-q)\,dt + \sigma dW_t^Q\right], \quad \sigma > 0 \\ S(0) &= S_0 \end{aligned}$$

with a constant riskless interest rate $r$ and a continuous dividend yield $q$.

The price $CPCall^+(0)$ of the cumulative Parisian barrier-up call option at time 0 for a window $[0,T]$ is given by

$$\begin{aligned} CPCall^+(0) &= E^Q\left(e^{-rT}\left[S_T - K\right]^+ \mathbb{1}_{\{A\}}\right) \\ A &:= \int_0^T \mathbb{1}_{\{S_t \geq L\}} dt \geq RT \end{aligned}$$

where $A$ denotes the the set of events on which the option is not knocked out.

If we approximate the distribution of the stock price via a discrete random walk and use in addition the formula of total probability we would obtain

$$\begin{aligned} CPCall^+(0) &\approx \sum_{x=0}^{\infty} e^{-rT} E^Q\left(\left[S_T - K\right]^+ \mathbb{1}_{\{A\}} \middle| S_T = x\right) Q\left(S_T = x\right) \\ &\approx \sum_{x=0}^{\infty} e^{-rT}\left[x - K\right]^+ Q\left(A \middle| S_T = x\right) Q\left(S_T = x\right). \end{aligned} \tag{5.1}$$

In this chapter, we choose to work with a binomial or trinomial lattice method for approximation of the distribution of the stock price, introduced by Cox, Ross and Rubinstein [26] (CRR).

Starting with the binomial tree, we denote the up-move by $u$ and the down-move by $d$ and choose $d = \frac{1}{u}$. Further, we denote the risk-neutral probabilities of an up-jump by $p_u$ and of a down-jump by $p_d$ and referring to CRR, we can write them as

$$\begin{aligned} p_u &= 1 - p_d = \frac{e^{(r-q)\Delta t} - d}{u - d} \\ u &= \frac{1}{d} = e^{\sigma \Delta t} \end{aligned}$$

where $\Delta t = \frac{T}{n}$ with $n$ denoting the number of timesteps in the binomial tree. With these choices, the binomial price process converges weekly to the geometric Brownian motion as $n \to \infty$.

Next, we notice that in a binomial lattice, the price of a plain vanilla call option can be approximated by

$$Call(0) \approx e^{-rT} \sum_{j=0}^{n} p_u^j p_d^{n-j} \begin{pmatrix} n \\ j \end{pmatrix} \left[ S_0 u^j d^{n-j} - K \right]^+ . \tag{5.2}$$

If we can now estimate the probability of the stock price starting in $S_0$ and ending at time $T$ in $S_0 u^j d^{n-j}$, $j = 0, \ldots, n$ to stay longer than a fraction $R$ of the time above the barrier $L$, we could approximate also the price of the cumulative Parisian call option in a way similar to (5.2) and corresponding to (5.1). Specifically, denoting the conditional probability of a survival, conditioned on the starting and ending stock value by $\tilde{p}\,(S_0, S_T)$, we can write it as

$$\tilde{p}\,(S_0, S_T) := Q\Big( \int_0^T \mathbb{1}_{\{S_t \geq L\}} dt \geq RT \,\Big|\, S_0, S_T \Big)$$

and the Parisian call option price can be approximated at time 0 as

$$CPCall^+(0) \approx e^{-rT} \sum_{j=0}^{n} p_u^j p_d^{n-j} \begin{pmatrix} n \\ j \end{pmatrix} \tilde{p}\left(S_0, S_0 u^j d^{n-j}\right) \left[ S_0 u^j d^{n-j} - K \right]^+ . \tag{5.3}$$

In the next section, we will deal with the calculation of $\tilde{p}\,(S_0, S_T)$, which is called "occupational time probability".

In the case of a trinomial tree construction, let us denote the risk-neutral probabilities for the stock price to keep its value, to jump down and to jump up by $p_m$, $p_d$ and $p_u$ respectively. Let us also denote the changes in the stock price in the middle-node, down-node and up-node by $m$, $d$ and $u$ respectively, where by following Hull [33] we assume that $m = 1$ and $u = \frac{1}{d}$. By matching the first and second moment of the stock price, we find

$$\begin{aligned}
p_u &= \sqrt{\frac{\Delta t}{12\sigma^2}} \left( r - q - \frac{\sigma^2}{2} \right) + \frac{1}{6} \\
p_m &= \frac{2}{3} \\
p_u &= -\sqrt{\frac{\Delta t}{12\sigma^2}} \left( r - q - \frac{\sigma^2}{2} \right) + \frac{1}{6} \\
u &= \frac{1}{d} = e^{\sigma\sqrt{3\Delta t}}
\end{aligned}$$

Using a trinomial lattice as a discrete approximation method, we can rewrite (5.1) for the cumulative Parisian call option price as

$$\begin{aligned} CPCall^{+}(0) \;\approx\; & e^{-rT}\sum_{j=0}^{n}\sum_{k=0}^{j} p_u^{j-k}p_m^{k}p_d^{n-j}\begin{pmatrix} n \\ j-k,k,n-j \end{pmatrix}\tilde{p}\left(S_0, S_0u^{j-k}m^kd^{n-j}\right) \\ & .\left[S_0u^{j-k}m^kd^{n-j}-K\right]^{+} \end{aligned} \tag{5.4}$$

where

$$\begin{pmatrix} n \\ j-k,k,n-j \end{pmatrix} = \frac{n!}{(j-k)!k!(n-j)!}$$

are the coefficients of a generalization of the Pascal's triangle to three dimensions called Pascal's pyramid or tetrahedron.

## 5.2 Calculation of the Occupational Time Probabilities

Recall that in order to incorporate the "cumulative Parisian" call feature in the tree, we need to find the following probability

$$Q\Big(\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt \geq RT \,\Big|\, S_0, S_T\Big) = E^Q\Big(\mathbb{1}_{\{\Gamma\geq RT\}} \,\Big|\, S_0, S_T\Big)$$

where we denote $\Gamma := \int_0^T \mathbb{1}_{\{S_t\geq L\}}dt$.

**Remark 5.2.1.** *Notice that the time the stock price process spends cumulatively above and below the barrier sums up to the total time $T$, i.e.*

$$\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt + \int_0^T \mathbb{1}_{\{S_t\leq L\}}dt = T.$$

*Therefore, we have the following relation between the probabilities of the occupational time of a Brownian motion below or above a barrier $L$*

$$\begin{aligned} Q\Big(\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt \geq RT \,\Big|\, S_0, S_T\Big) &= Q\Big(T - \int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \geq RT \,\Big|\, S_0, S_T\Big) \\ &= Q\Big(\int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \leq (1-R)\,T \,\Big|\, S_0, S_T\Big) \\ &= 1 - Q\Big(\int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \geq (1-R)\,T \,\Big|\, S_0, S_T\Big) \end{aligned}$$

*which allows us an easy transformation between the upper probabilities, once we have estimated one of them. Hence, with a little effort we can also calculate all types of down-and-in, up-and-in, down-and-out or up-and-out continuously monitored cumulative Parisian call/put options.*

Similar to the relationship between the upper probabilities, Chesney, Jeunblanc and Yor [19], Hugonnier [32]and Moraux [50] show that if we denote by $CPCall(\cdot)$ and $CPPut(\cdot)$ the prices of the continuously monitored cumulative Parisian barrier Call/Put option (with subindices " + " or " − " whether the occupational time is considered above or below the barrier) given as a function of the cumulative "window" time, then the following parity relationships hold:

$$CPCall^{\mp}(RT) + CPCall^{\pm}((1-R)T) = BSCall(T) \tag{5.5}$$

$$CPPut^{\mp}(RT) + CPPut^{\pm}((1-R)T) = BSPut(T) \tag{5.6}$$

with $BSCall(T)$ and $BSPut(T)$ denoting the Black-Scholes prices of a European Call or respectively Put option with the same stock price parameters as the cumulative Parisian barrier option. For the general proof of the these relationships, we refer again to Chesney, Jeanblanc and Yor [19], Hugonnier [32] and Moraux [50] and state here an easy proof of their validity in the case of the binomial/trinomial tree approximation we have offered.

**Lemma 5.2.1.** *If we denote the respective approximated (with the offered lattice method) prices of the different types of continuously monitored cumulative Parisian barrier Call/Put options with* $\widehat{CPCall}^{\mp}(\cdot)$, $\widehat{CPCall}^{\pm}(\cdot)$,$\widehat{CPPut}^{\mp}(\cdot)$ *and* $\widehat{CPPut}^{\pm}(\cdot)$ *and the approximated (with a CRR lattice) simple vanilla Call/Put option price by* $\widehat{Call}(\cdot)$ *and* $\widehat{Put}(\cdot)$*, then the following parity relationships hold*

$$\widehat{CPCall}^{\mp}(RT) + \widehat{CPCall}^{\pm}((1-R)T) = \widehat{Call}(T) \tag{5.7}$$

$$\widehat{CPPut}^{\mp}(RT) + \widehat{CPPut}^{\pm}((1-R)T) = \widehat{Put}(T) \tag{5.8}$$

**Proof:** Since the different cases are analogical, we will prove only

$$\widehat{CPCall}^{+}(RT) + \widehat{CPCall}^{-}((1-R)T) = \widehat{Call}(T)$$

for the binomial tree construction.

Recall that due to (5.3) we can write the price of the cumulative Parisian Call option with window time $(RT)$ above the barrier $L$ as

$$\widehat{CPCall}^{+}(RT) = e^{-rT}\sum_{j=0}^{n} p_u^j p_d^{n-j}\begin{pmatrix} n \\ j \end{pmatrix}\tilde{p}\left(S_0, S_0u^jd^{n-j}\right)\left[S_0u^jd^{n-j}-K\right]^{+}$$

with

$$\tilde{p}\left(S_0, S_T\right) := Q\Big(\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt \geq RT \,\Big|\, S_0, S_T\Big).$$

Then, using (recall Remark 5.2.1)

$$Q\Big(\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt \geq RT \,\Big|\, S_0, S_T\Big) = 1 - Q\Big(\int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \geq (1-R)\,T \,\Big|\, S_0, S_T\Big)$$

yields

$$\begin{aligned}
\widehat{CPCall}^{+}(RT) &= e^{-rT}\sum_{j=0}^{n} p_u^j p_d^{n-j}\begin{pmatrix} n \\ j \end{pmatrix}\Big(1 - Q\Big(\int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \geq (1-R)\,T \,\Big|\, S_0, S_T\Big)\Big) \\
&\quad \left[S_0u^jd^{n-j}-K\right]^{+} \\
&= e^{-rT}\sum_{j=0}^{n} p_u^j p_d^{n-j}\begin{pmatrix} n \\ j \end{pmatrix}\left[S_0u^jd^{n-j}-K\right]^{+} \\
&\quad -e^{-rT}\sum_{j=0}^{n} p_u^j p_d^{n-j}\begin{pmatrix} n \\ j \end{pmatrix} Q\Big(\int_0^T \mathbb{1}_{\{S_t\leq L\}}dt \geq (1-R)\,T \,\Big|\, S_0, S_T\Big)\left[S_0u^jd^{n-j}-K\right]^{+} \\
&= \widehat{Call}(T) - \widehat{CPCall}^{-}\left((1-R)T\right).
\end{aligned}$$

The rest of the cases follow by analogy due to Remark 5.2.1. The proof for the trinomial lattice approximation follows the same steps as the binomial one. ■

Now, let us denote $m := \frac{1}{\sigma}\left(r - q - \frac{\sigma^2}{2}\right)$ and notice that $S_t = S_0e^{\sigma(mt+W_t)} = S_0e^{\sigma Z_t}$, where $Z_t = W_t + mt$. Our next goal in calculating the occupational time probabilities is to move from the risk-neutral probability measure $Q$ under which $W_t$ is a driftless Brownian motion to an equivalent (with respect to $\{\mathfrak{F}\}_{t\in[0,T]}$ ) probability measure $Q^*$ under which $Z_t$ is a driftless Brownian motion.

The measure $Q^*$ is defined by the Radon-Nikodym density $D(T)$, given by Theorem B.0.2 as

$$D(T) = \frac{dQ}{dQ^*}\Big|_{F_T} = e^{mZ_T - \frac{m^2}{2}T} \text{ and } \frac{1}{D(T)} = \frac{dQ^*}{dQ}\Big|_{F_T} = \frac{1}{e^{-mW_T - \frac{1}{2}m^2T}}.$$

Denoting $L^* = \frac{1}{\sigma} log\left(\frac{L}{S_0}\right)$, $\Gamma^* = \int_0^T \mathbb{1}_{\{Z_t \geq L^*\}} dt$ and referring to the Cameron-Martin-Girsanov Theorem B.0.1, we obtain

$$E^Q\Big(\mathbb{1}_{\{\Gamma \leq RT\}} \Big| S_0, S_T\Big) = \frac{E^{Q^*}\Big(\mathbb{1}_{\{\Gamma^* \leq RT\}} D(T) \Big| S_0, S_T\Big)}{E^{Q^*}\Big(D(T) \Big| S_0, S_T\Big)}$$

where notice that $\mathcal{G} = \{S_0, S_T\}$ is a sub-sigma field of $\mathcal{F}$ and $\mathcal{G} = \{S_0, S_T\} = \{W_0, W_T\} = \{Z_0, Z_T\}$. Then, plugging in the Girsanov density $D(T)$ yields,

$$\begin{aligned}
&E^Q\Big(\mathbb{1}_{\{\Gamma \leq RT\}} \Big| S_0, S_T\Big) \\
&= \frac{E^{Q^*}\Big(\mathbb{1}_{\{\Gamma^* \leq RT\}} e^{mZ_T - \frac{m^2}{2}T} \Big| Z_0, Z_T\Big)}{E^{Q^*}\Big(e^{mZ_T - \frac{m^2}{2}T} \Big| Z_0, Z_T\Big)} = E^{Q^*}\Big(\mathbb{1}_{\{\Gamma^* \leq RT\}} \Big| Z_0, Z_T\Big) \\
&= Q^*\Big(\int_0^T \mathbb{1}_{\{Z_t \geq L^*\}} dt \leq RT \Big| Z_0, Z_T\Big) \qquad (5.9)
\end{aligned}$$

where $Z_0 = 0$ and $Z_T = \frac{1}{\sigma} log \frac{S_T}{S_0} =: y$.

Next, we denote the joint density of the occupational time $\Gamma^*$ and the location of $Z$ at time $T$ by $f_{\Gamma^*, Z_t}(x, y)$, the density of $Z$ at time $T$ by $f_{Z_T}(y)$ and the conditional on $Z_T = y$ density of $\Gamma^*$ by $f_{\Gamma^*, Z_T = y}(x)$. Then, we have that

$$Q^*\Big(\Gamma^* \leq RT \Big| Z_T = y\Big) = \int_0^{RT} f_{\Gamma^*, Z_T = y}(x) dx = \int_0^{RT} \frac{f_{\Gamma^*, Z_T}(x, y)}{f_{Z_T}(y)} dx.$$

The density of the Brownian motion at time $T$ is given by

$$f_{Z_T}(y) = \frac{1}{\sqrt{2\pi T}} \; e^{-y^2/2T}$$

and referring to Borodin and Salminen [9] p.158, we have the following four different cases for the joint density $f_{\Gamma^*, Z_t}(x, y)$:

1. $L \geq S_0, L \geq S_T \; (\Leftrightarrow \; L^* \geq 0, L^* \geq y)$

   $$f_{\Gamma^*, Z_t}(x, y) = \frac{a\sqrt{T - v}}{\pi T^2 \sqrt{v}} \, e^{-\frac{a^2}{2(T-v)}} dv dy + \frac{1}{\sqrt{2\pi} T^{\frac{3}{2}}} \, (1 - \frac{a^2}{T}) e^{-\frac{a^2}{2T}} \, \text{Erfc}(\frac{a\sqrt{v}}{\sqrt{2T(T-v)}}) dv dy$$

   where $a := 2L^* - y$.

2. $S_0 \leq L \leq S_T \, (\Leftrightarrow \, 0 \leq L^* \leq y)$

$$f_{\Gamma^*,Z_t}(x,y) = \Big(\frac{c\sqrt{v}}{\pi T^2\sqrt{T-v}} + \frac{L^*\sqrt{T-v}}{\pi T^2\sqrt{v}}\Big)e^{-\frac{c^2}{2v}-\frac{L^{*2}}{2(T-v)}}dvdy$$
$$+\frac{1}{\sqrt{2\pi}T^{\frac{3}{2}}}\,(1-\frac{a^2}{T})e^{-\frac{a^2}{2T}}\mathrm{Erfc}\Big(\frac{c\sqrt{T-v}}{\sqrt{2Tv}} + \frac{L^*\sqrt{v}}{\sqrt{2T(t-v)}}\Big)dvdy$$

where $c := y - L^*$, $a := 2L^* - y$.

3. $S_T \leq L \leq S_0 \, (\Leftrightarrow \, y \leq L^* \leq 0)$

$$f_{\Gamma^*,Z_t}(x,y) = \Big(\frac{-L^*\sqrt{v}}{\pi T^2\sqrt{T-v}} + \frac{d\sqrt{T-v}}{\pi T^2\sqrt{v}}\Big)e^{-\frac{L^{*2}}{2v}-\frac{d^2}{2(T-v)}}dvdy$$
$$+\frac{1}{\sqrt{2\pi}T^{\frac{3}{2}}}\,(1-\frac{a^2}{T})e^{-\frac{a^2}{2T}}\mathrm{Erfc}\Big(\frac{-L^*\sqrt{T-v}}{\sqrt{2Tv}} + \frac{d\sqrt{v}}{\sqrt{2T(t-v)}}\Big)dvdy$$

where $d := L^* - y$, $a := 2L^* - y$.

4. $L \leq S_0, L \leq S_T \, (\Leftrightarrow L^* \leq 0, L^* \leq y)$

$$f_{\Gamma^*,Z_t}(x,y) = \Big(\frac{-a\sqrt{v}}{\pi T^2\sqrt{T-v}}\Big)e^{-\frac{a^2}{2v}}dvdy$$
$$+\frac{1}{\sqrt{2\pi}T^{\frac{3}{2}}}\,(1-\frac{a^2}{T})e^{-\frac{a^2}{2T}}\mathrm{Erfc}\Big(\frac{-a\sqrt{T-v}}{\sqrt{2Tv}}\Big)dvdy$$

where $a := 2L^* - y$.

We have denoted with Erfc(.) the error function which is given by

$$\mathrm{Erfc}(x) := \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-v^2}dv$$

where substituting $v = \frac{u}{\sqrt{2}}$, $dv = \frac{1}{\sqrt{2}}du$ and noticing that we integrate $u$ in $[\sqrt{2}x, \infty]$, we obtain

$$\mathrm{Erfc}(x) \quad = \quad \frac{2}{\sqrt{2\pi}}\int_{\sqrt{2}x}^\infty e^{-\frac{u^2}{2}}du = 2\varphi(-\sqrt{2}x)$$

where $\varphi(\cdot)$ is the cumulative standard normal distribution function.

**Remark 5.2.2.** *For the special case of $S_0 = S_T = L \Leftrightarrow L^* = y = 0$, we have that*

$$Q^*\Big(\int_0^T \mathbb{1}_{\{Z_t \geq 0\}}dt \in dv \,\Big|\, Z_0 = Z_T = 0\Big) = \frac{1}{T}\mathbb{1}_{[0,T]}(v)dv$$

*and thus*

$$(5.9) = \frac{1}{T}\int_0^{RT}\mathbb{1}_{[0,T]}(v)dv = R.$$

Finally, analytically or numerically calculating (**??**) for the different four cases, we can obtain the required probability. Notice that the interest rate $r$ does not enter in the formula of the estimated probabilities, due to the conditioning on the starting and ending point of the stochastic process since the only uncertainty left is the one driven by the underlying Brownian motion.

**Remark 5.2.3.** *To speed-up the calculation of the upper probability, we can scale the time from* $[0,T]$ *to* $[0,1]$ *and thus be able to always integrate from* $0$ *to* $1$ *with a new variance* $\sigma\sqrt{T}$*. Therefore, notice that*

$$\begin{aligned} S_t &= S_0\, exp\left[\left(r-\frac{1}{2}\sigma^2\right)t+\sigma W^Q(t)\right] \\ &= S_0\, exp\left[\left(rT-\frac{(\sigma^2\sqrt{T})^2}{2}\right)\frac{t}{T}+\sigma\sqrt{T}W^Q\left(\frac{t}{T}\right)\right] \end{aligned}$$

*where* $\sqrt{T}W^Q(\frac{t}{T}) = W^Q(t)$ *both* $\sim \mathcal{N}(0,t)$

$$\int_0^T \mathbb{1}_{\{S_t\geq L\}}dt \leq RT \Leftrightarrow \int_0^T \mathbb{1}_{\left\{S_0\, exp\left[\left(rT-\frac{(\sigma\sqrt{T})^2}{2}\right)\frac{t}{T}+\sigma\sqrt{T}W^Q\left(\frac{t}{T}\right)\right]\right\}}dt \leq RT \tag{5.10}$$

*substitute* $t = uT$, $t\in[0,T] \Rightarrow dt = Tdu$ *and* $u\in[0,1]$ *and obtain,*

$$\int_0^1 \mathbb{1}_{\left\{S_0\, exp\left[\left(rT-\frac{(\sigma\sqrt{T})^2}{2}\right)u+\sigma\sqrt{T}W^Q(u)\right]\right\}}du\,T \leq RT$$

*Let us denote,*

$$S^*(u) = S_0\, exp\left[\left(rT-\frac{(\sigma\sqrt{T})^2}{2}\right)u+\sigma\sqrt{T}W(u)\right]$$

*Then,*

$$\begin{aligned} S^*(0) &= S_0 \\ S^*(1) &= S_0^*\, exp\left[\left(rT-\frac{(\sigma\sqrt{T})^2}{2}\right)+\sigma\sqrt{T}W(1)\right] \end{aligned}$$

*and*

$$Q\Big((5.10)\Big) = Q\left(\int_0^1 \mathbb{1}_{\{S^*(u)\}}du \leq R\,\Big|S^*(0),\, S^*(1)\right)$$

*Now, the same formulas apply as before, but for volatility* $\sigma\sqrt{T}$ *and time interval* $[0,1]$*. Note that although the interest rate is now* $rT$ *it causes no change in the calculations as it cancels out.*

If we want to price a cumulative Parisian barrier option with a forward starting window $[T_1, T_2]$, i.e. $T_1 > 0$, we can construct a standard binomial/trinomial tree until the beginning of the window and then integrate directly until the maturity of the option. Visually, this is shown in Figure 5.1 and assuming only for notational easiness that the number of time steps $n_1$ from $[0, T_1]$ and $n_2$ from $[T_1, T_2]$ are chosen so that $\Delta t = \frac{T_1}{n_1} = \frac{T_2 - T_1}{n_2}$ (and thus the jump heights are the same before or during the window) we can write the approximated via a binomial type of tree price of the option as

$$\begin{aligned} C(0) \quad \approx \quad & e^{-rT_2} \sum_{j=0}^{n_1} p_u^j p_d^{n_1-j} \begin{pmatrix} n_1 \\ j \end{pmatrix} \sum_{i=0}^{n_2} p_u^i p_d^{n_2-i} \begin{pmatrix} n_2 \\ i \end{pmatrix} \\ & .\tilde{p}\left(S_0 u^j d^{n_1-j}, S_0 u^j d^{n_1-j} u^i d^{n_2-i}\right) \left[S_0 u^j d^{n_1-j} u^i d^{n_2-i} - K\right]^+ . \end{aligned}$$

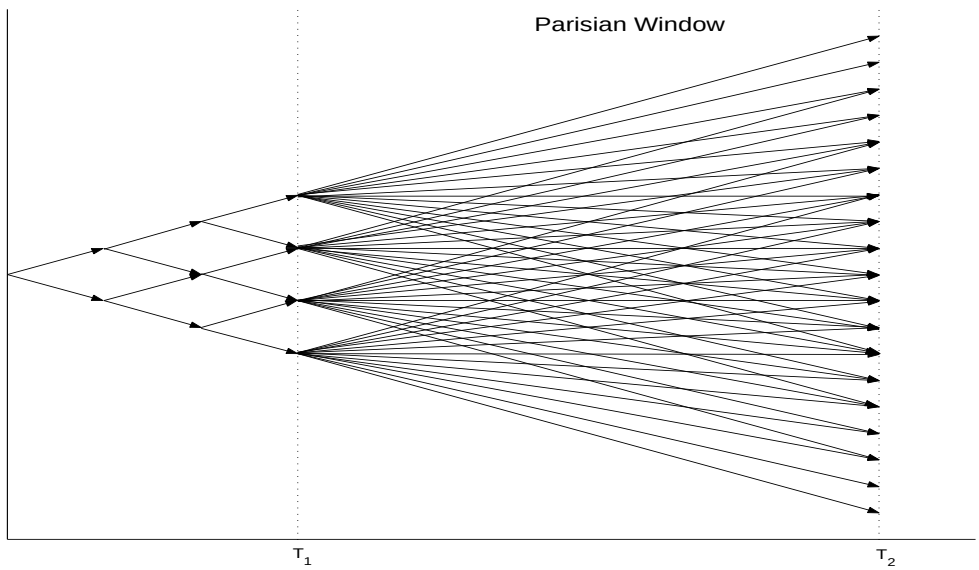



Figure 5.1: Adopting the method for the pricing of options with a cumulative Parisian window starting in the future.

One of the advantages of the offered method to the closed-form solution is exactly this flexibility in adapting different features of the Parisian option.

## 5.3 Results

In this section, we compare the prices of the continuously monitored cumulative Parisian call option, obtained from the offered modified lattice method to the quasi-analytical solution results

by Hugonnier [32] and the FSG lattice constructed by Kwok and Lau [44]. We set the starting stock price $S_0 = 95$, the strike $K = 100$, the barrier $L = 110$, the volatility $\sigma = 20\%$, the maturity $T = 1$, the interest rate $r = 5\%$ and the continuous dividend yield to 2%. We variate the fraction $R$ between the values 0.25, 0.5 and 0.75. For the results of the quasi-analytical solution by Hugonnier [32] and the FSG lattice by Kwok and Lau we refer to Kwok and Lau [44] (exhibit 3) and thus we obtain the following table:

| R | quasi-analytical solution | FSG n=500 | FSG n=1000 | FSG Extrapol. | **Direct Binom. Tree** $(n = 500)$ | **Direct Trinom. Tree** $(n = 500)$ |
|---|---|---|---|---|---|---|
| 0.25 | 4.88453 | 4.71818 | 4.76807 | 4.88851 | **4.88495** | **4.88459** |
| 0.5 | 3.08308 | 2.88602 | 2.94335 | 3.08175 | **3.08251** | **3.08212** |
| 0.75 | 0.98758 | 0.85923 | 0.89504 | 0.98149 | **0.98618** | **0.98620** |

where the results of the offered tree construction are given under the names "Direct Binomial Tree" and "Direct Trinomial Tree" and $n$ denotes the number of time steps in the tree constructions. Additionally, in Figure 5.2 we have plotted for $R = 0.25$ (keeping the other parameters the same as before) the direct binomial and trinomial lattice prices of the cumulative Parisian call option agains the number of time slices.

Notice that the results we obtain are better than the FSG lattice constructions of Kwok and Lau [44] already with only 200 time steps. The reason for this is that the "direct tree" construction uses analytically estimated conditional on the starting and ending point knock-out probabilities. Moreover, it has in addition no interpolation error as in the FSG algorithm. The offered tree construction is intuitive and in addition easy to construct for all cases of continuously monitored cumulative Parisian options. Furthermore, it allows for some flexibility in the features of the option such as a forward staring window.

However, its major drawback is that it is not flexible in the sense that for the pricing of any other type of Barrier options, the conditional on the starting and ending point knock-out probability has to be explicitly estimated while this is not needed in the FSG lattice pricing method.

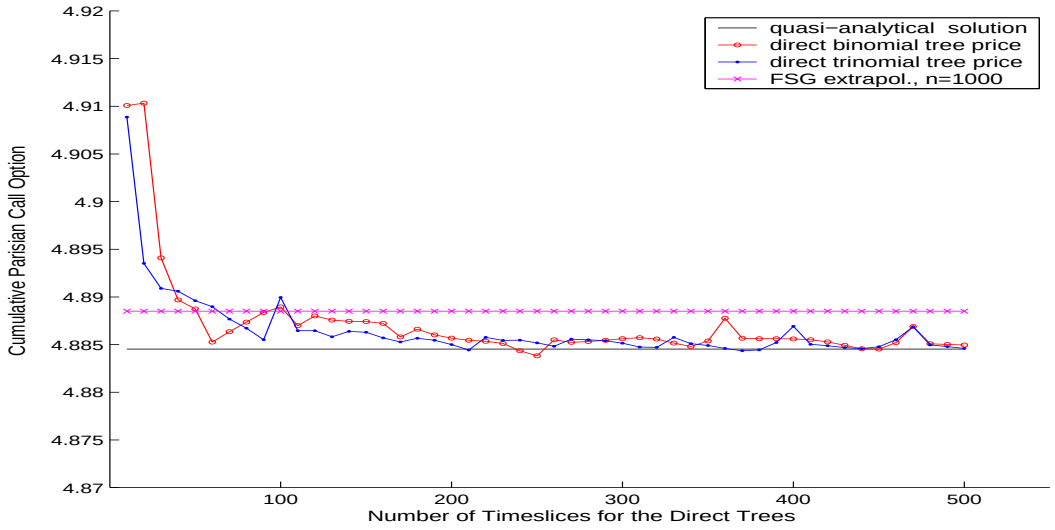



Figure 5.2: Convergence of the direct binomial and trinomial tree results with the number of time slices. The quasi closed-form solution result and the extrapolated FSG result are taken from Kwok and Lau [44].

## 5.4 Suggestions for Further Research Topics

A possible topic for further research is the pricing of the continuously monitored consequent Parisian options with the offered algorithm for which we have to estimate the respectice conditional knock-out probabilities. In addition, the algorithm can be applied for pricing convertible bonds with cumulative Parisian call feature. The method is also applicable when combining the stock price dynamic with a stochastic interest rate one but only as long as a direct integration until the maturity of the simple option (without the barrier feature) is possible. Additional difficulty in this respect is the usual mean-reversion feature of the stochastic short rate processes.

# Chapter 6

# Longevity Bonds - Pricing, Modeling and Application for German Data

This chapter is similar to the article of Korn, Natcheva and Zipperer [42] in which it additionally focuses on introducing new two-factor stochastic models for the mortality rate dynamics and consequently on delivering further numerical examples and analysis.

## 6.1 Introduction

Besides the interest rate risk, the longevity risk is the second major risk factor that has to be considered in the process of determining the premium of a life insurance. While it is possible to face the interest rate risk, due to the rich market of interest rate derivatives, it is generally impossible to hedge the longevity risk. In this thesis, with the usage of the term "longevity risk" it is meant the uncertain increase in life expectancy over time.

In the pricing of a life annuity, an insurer has to estimate the life expectancy for a long period of time. This estimation is typically performed in big time intervals and the life expectancy is assumed to be constant between two estimation dates. However, a research based on the data by the Federal Statistical Office of Germany reveals that in the last decade, mortality has been constantly decreasing and hence, the life annuities have been typically underpriced. As an extreme case, it may thus happen that the insurance premium does not suffice to cover the insured payments.

In this situation, a solution of the capital markets is to offer and trade specific insurance risks such as the so called "Longevity Bonds" with coupon payments depending on the development of the survival function of a given cohort[1], which can be used for hedging at least a big portions of the longevity risk (examples here are the SwissRe - Mortality Bond linked to short-term, catastrophic mortality risk and the EIB/BNP-Paribas Longevity Bond (see Cairns *et al.* [15] or Cowley and Cummins [24]). The estimate of the advantages of such financial products is strongly related to their price. In order to be able to assess them, we shall present in the this chapter a simplified (but flexible in respect to the applied model) approach for pricing longevity bonds which compounds ideas from Cairns, Blake and Dowd [15] and Milevsky and Promislow [49]. Further, we will offer two additional models - a stochastic two-factor variation of the well-known Gompertz model and a stochastic variation of the well-known Makeham model - which describe in a natural way the development of longevity and deliver very satisfactory results in their application on real data.

We stress here that in the pricing of the longevity bonds one has to consider the following difficulties:

- A reliable (implicit) calibration of the survival parameters, based on the prices of the respective financial instruments (e.g. the calibration of the short rate parameters in the financial market) is not possible due to the lack of corresponding market and the following non-tradability of the longevity risk.

- In the pricing of the bond, the trend of increasing life expectancy has to be explicitly calibrated.

- For the simultaneous pricing of interest rate risk and longevity risk, a combination between principals based on financial mathematics and actuarial science has to be found.

In this place, we have to point out that the pricing of longevity bonds makes sense only when we deal with mortality curves that change with time. If the mortality curve were to be exactly described by a mortality table that stays constant in (calender) time, then we could with the

[1] An example of a cohort is the population of 65-year-olds at the issue data of the bond.

means of that table directly estimate the price of the mortality risk (for a sufficiently big population on which the coupons of the longevity bonds are based). However, in order to avoid a systematic error in the estimation of the future payments of the longevity bond, a concept featuring the time dependence of the mortality curve has to be considered. Here, one can follow one of the (at least) two approaches

- Extrapolation of the past mortality curves, which are based on the realized number of deaths.
- Modeling the development of the mortality as a stochastic process and simulating the future mortality via Monte Carlo simulation.

In the following subsections, we will present and define the longevity bond we are going to deal with, consider some different mortality rate models and their time dependent variations. Finally, we propose methods for pricing the longevity bond.

## 6.2 Longevity Bonds and Survival Probabilities

We want to present here as a base product the longevity bond we are going to further deal with and note that it borrows its form from the longevity bond, structured by BNP Paribas and issued by the European Investment Bank in November 2004. For this reason, we consider a reference cohort whose members are at age $x(x > 0)$ at the starting time $t = 0$. The coupon payments $S(t_i)$ of the longevity bond are payed at times $t_i$, $i = 1, \ldots, N$ and are defined by the fraction of survivors from the initial cohort at the respective time points. This allows us to write them in the simplified form

$$S(i) = \frac{\textit{number of survivors from the cohort at time } i}{\textit{number of individuals in the cohort at time } 0} \tag{6.1}$$

where we set $S(0) = 1$. Note that the height of the coupon payments as their time comes should be regulated in the contract and be possible to track (for example based on the data published by a National Statistical Office). For the further steps in the pricing of the longevity bonds, we make the following assumption:

**Assumption 6.2.1.** *The interest rate risk and the longevity risk are independent.*

This assumption directly implies that the value of the coupon payments $S(t)$ of the longevity bond can be written under an appropriate pricing measure (we do not discuss its form here but leave it for a later section) as

$$E\left(\exp\left(-\int_0^i r(s)\,ds\right)S(i)\right) = P(0,i)\,E(S(i)) \tag{6.2}$$

in which $P(0,i)$ denotes the market price at time 0 of a zero-coupon bond with maturity $i$. Further, using the notation[2]

$$p(t,T_0,T_1,x) := P(\textit{Individ. alive in } T_1 \,|\textit{Individ. alive in } T_0, \textit{Ind. has in } 0 \textit{ age } x, f_t) \tag{6.3}$$

for time points $t$, $T_1$, $T_0$ with $t \leq T_1$, $T_0 < T_1$ (and $f_t$ the $\sigma$-algebra that contains the complete information until time $t$ about the interest rate and mortality development) follows directly by the linearity of expectation and (6.1) that

$$E(S(i)) = p(0,0,i,x). \tag{6.4}$$

If we assume that there exists a current mortality table that models mortality rates sufficiently well, one can directly estimate from it the values of $p(0,0,i,x)$ - usually denoted by ${}_ip_x$ - and via (6.4) obtain as the fair price of the longevity bond

$$\sum_{i=1}^{n} P(0,i)\,{}_ip_x. \tag{6.5}$$

However, as a definite tendency towards longevity can be empirically observed (see Oliveri [51]) the usage of a current (and therefore underestimating the survival probabilities) mortality table would assign to the longevity bond a very low value. To avoid this, one has to model the survival probabilities as time dependent ones. One approach to do this consists of taking the already realizes survival probabilities (respectively their empirical analogies)

$${}_ip_x(t) = \frac{\textit{The number of } (x+i) \textit{ years olds at time } t+i}{\textit{The number of } x \textit{ years olds at time } t},\ t \in \{-(i+1), \cdots, -N_i\} \tag{6.6}$$

as values of a time dependent function, approximating it via appropriate interpolation (e.g. spline or polynomial interpolation) and afterwards from the interpolated function $f_x^{(i)}(.)$ obtain the required estimator through extrapolation

$${}_i\hat{p}_x(0) = f_x^{(i)}(0), \quad i = 1, ..., N. \tag{6.7}$$

[2]Cairns at al. [15] name this probability *"forward survival probability"*.

In this work, we are not going to deal with this method any further (see as a reference Oliveri and Pitacco [52] or Pitacco [54] for more details in that direction.) but will present several stochastic models, based on different variations of the traditional mortality laws.

## 6.3 Longevity Bonds and Dynamic Modeling of Mortality

Besides the extrapolation methods, there exists in the literature another stream of methods (applied by Cairns, Blake and Dowd [15] for the special case of the EIB/BNP longevity bond) which consists of modeling the survival probabilities via parameterization of the mortality rates. In this connection, let as first assume time-independent survival probabilities and define the mortality rates as

$$ {}_ip_x = \exp\left(-\int_0^i \mu_{x+s} ds\right) \tag{6.8} $$

which means again that the mortality rate $\mu_x$ depends only on the age $x$ of the single person. In the history of Insurance mathematics, there have been developed a variety of approaches which parametrise the development of the mortality rate with the change of the age of the single person (especially for non-integer ages $x$), where we refer to Pitacco [54] for historical overview. As an example of the most popular approaches for parametrical representation (called also "Mortality Laws"), we give

$$ \mu_x^{(G)} = \alpha e^{\beta x}, \ \ \alpha, \beta > 0 \qquad \text{(Gompertz model)}, \tag{6.9} $$

$$ \mu_x^{(M)} = \alpha e^{\beta x} + \gamma, \ \ \alpha, \beta > 0, \gamma \geq 0 \qquad \text{(Makeham model)}, \tag{6.10} $$

$$ \mu_x^{(T)} = \alpha_1 e^{-\beta_1 x} + \alpha_2 e^{-\beta_2 (x-\eta)^2} + \alpha_3 e^{\beta_3 x}, \ \ \alpha_i, \beta_i, \eta \geq 0 \quad \text{(Thiele model)} \tag{6.11} $$

where one can notice that the Thiele model contains the other two models as a special case since it is capable to model mortality satisfactory over the whole life span. In this respect, its first term is used to describe infant mortality, the second one to describe mortality at young adult ages (mainly accidental) and the last one to describe mortality at old ages. Since the longevity risk concerns only the people at old ages (higher than 50), for which the contribution of the first

two terms is negligible, we notice that for the purpose of pricing longevity bonds, the Gompertz and the Makeham model are more suitable.

Another form of generalization of the Makehams model that we will further use is the so called model of Perks (see Perks [53]) and it defines the mortality rate as

$$\mu_x^{(P)} = \frac{\alpha e^{\beta x} + \gamma}{1 + \delta e^{\beta x} + \varepsilon e^{-\beta x}}, \quad \alpha, \beta, \gamma, \delta, \varepsilon \geq 0 \tag{6.12}$$

Alternatively, it is also possible to model directly the survival probability $p_x$ as time independent e.g.

$$\frac{1 - p_x^H}{p_x^H} = A^{(x+B)^C} + De^{-E(\log x - \log F)^2} + GH^x, \qquad \text{(Heligan-Pollard model)} \tag{6.13}$$

where one can notice the similarity to the Thiele model of mortality rates and thus we have the same interpretation of the meaning of its three terms.

Since all these models are very well known, we are not going to deal further with their advantages and disadvantages, but will refer instead to the standard literature in which this is done (see e.g. Benjamin and Pollard [5]). Under the assumption that the mortality rates follow one of the above models, we can obtain via simple integration of (6.4), (6.8), (6.9), (6.10), (6.12) for example for the Gompertz, Makeham and Perks case the following closed form solutions for $S(t)$:

**Proposition 5.** *Under the assumption that the mortality rates follow one of the classical mortality rate models, the respective conditional survival probabilities until age $x + i$ are given by*

*a) in the Gompertz model:*

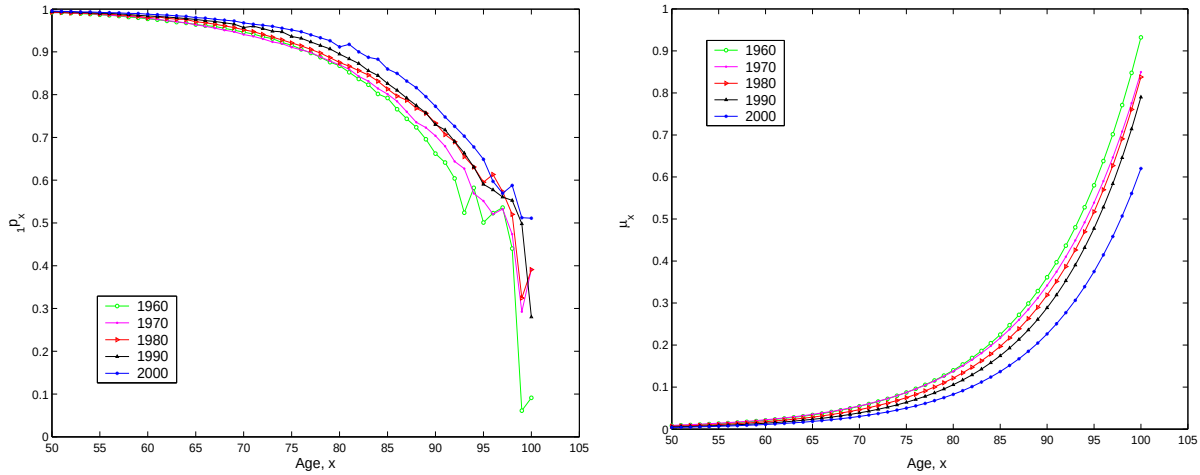
$$E\left(S\left(i\right)\right) = {}_ip_x = \exp\left(-\frac{\alpha}{\beta} e^{\beta x}\left(e^{\beta i} - 1\right)\right) \tag{6.14}$$

*b) in the Makeham model:*

$$E\left(S\left(i\right)\right) = {}_ip_x = \exp\left(-\left(\frac{\alpha}{\beta} e^{\beta x}\left(e^{\beta i} - 1\right) + \gamma\, i\right)\right) \tag{6.15}$$

*c) in the Perks modell for the special case $\gamma = \varepsilon = 0$*

$$E\left(S\left(i\right)\right) = {}_ip_x = e^{-\frac{\alpha i}{\delta}}\left(\frac{1 + \delta e^{\beta x}}{1 + \delta e^{\beta(x+i)}}\right)^{\frac{\alpha i}{\delta}}. \tag{6.16}$$




(a) Historical survival probabilities $p_x$.

(b) Calibrated with the classical Gompertz model mortality rates $\mu_x$.

Figure 6.1: Data for West German males by the SBA.

Proposition 5 points out that in order to be able to calculate the survival probabilities, one has to first estimate the parameters of the respective model using the already realized survival probabilities (or more precisely, using the observed relative frequencies).

However, if we plot the historical survival probabilities $p_x$ for some different past years or respectively the calibrated to them mortality rates using e.g. the calibrated parameters of the classical Gompertz model (see Figure 6.1), we notice a clear longevity trend. Hence, from now on we will concentrate on developing dynamic variations of the classical Mortality laws that are able to capture such a trend of longevity.

The most simple way to do this, is to choose one of the given Mortality laws and to substitute its parameters by time dependent functions. This approach corresponds to the already described interpolation (and extrapolation) of the survival probabilities approach.

However, we choose here to follow another approach, in which the parameters of the mortality law are substituted by stochastic processes which can be uniquely described by a parameter vector $\Theta$ (e.g. a Brownian motion with drift $\nu$ and volatility $\sigma$ so that $\Theta = (\nu, \sigma)$). The mortality rates can then be written as stochastic processes of the form

$$\mu_x(t) = \mu_x(t; \theta) \tag{6.17}$$

and the current survival probabilities of a $x$-years old individual until time $t$ are given by

$$ {}_tp_x(0) := p(0,0,t,x) = E(S(t)) = E\left(\exp\left(-\int_0^t \mu_{x+s}(s)\,ds\right)\right). \tag{6.18} $$

The clear advantage of this approach is that after estimating the parameter vector $\Theta$, one can simulate paths of the future mortality rates via Monte Carlo Simulation and additionally calculate confidence intervals for the survival probabilities. For the realization of this method, the following algorithm can be formulated (based on the ideas of Lee and Carter[45]):

**Algorithm:** "Stochastic Modeling of Mortality Rates "

0. Choose a parameterized stochastic form of the applied mortality rates law (respectively for the survival probabilities).

1. Using the observed death rates, estimate the realized mortality rates (respectively survival probabilities)

2. Calibrate the parameters of the chosen stochastic process to the realized time series of the mortality rates estimated in 1.

After the calibration of the stochastic process in step 2., we have now fully specified its parameters and we can proceed to its Monte Carlo Simulation.

In the next section, we are going to examine the application of two existing models of this approach, namely the Cairns *et al* and Milevski and Promislow MP models. We will also offer and examine two additional models which can be seen as stochastic variations of the Gompertz and the Makeham classical laws and we therefore name them respectively "stochastic Gompertz" and "stochastic Makeham" models.

### 6.3.1 Time dependent Perks Model

This model is offered by Cairns, Blake and Dowd [15], in which the authors offer a discrete in time modeling of the survival probabilities via a two-dimensional Brownian motion with drift. In specific, they define $p(t+1,t,t+1,x)$ to be the fraction of individuals with age $x$ at time 0

and still alive at time $t \geq 0$ that have survived until time $t+1$. With the assumption that

$$\hat{p}\left(t+1, t, t+1, x\right) = \frac{1}{1+e^{A_1(t+1)+A_2(t+1)(x+t)}} \tag{6.19}$$

for all historical time points $t$ and all ages $x$[3], one can estimate using historical data (with the help of least-squares minimization) the values of $(A_1(t), A_2(t))$.

In specific, Cairns, Blake and Dowd [15] assume that the $A$-process follows the dynamics

$$A\left(t+1\right) = A\left(t\right) + \nu + CZ\left(t+1\right), \;\; \nu \in \mathbb{R}^2, C \in \mathbb{R}^{2,2}, \quad t = 0, 1, 2, ... \tag{6.20}$$

where $C$ denotes an upper triangular matrix and the $Z$-process is the increment of a two-dimensional Brownian motion between time points $t$ and $t+1$. From 6.19 follows that the $A_1(t)$ process is responsible for time dependent changes of mortality that affect all ages, while the $A_2(t)$ process reflects time dependent changes of mortality that are different for the different ages.

**Remark 6.3.1.** *Cairns, Blake and Dowd's [15] model can also be seen as a stochastic variation of the classical Heligan-Pollard mortality law:*

$$\frac{1-p_x^H}{p_x^H} = A^{(x+B)^C} + De^{-E(\log x - \log F)^2} + GH^x.$$

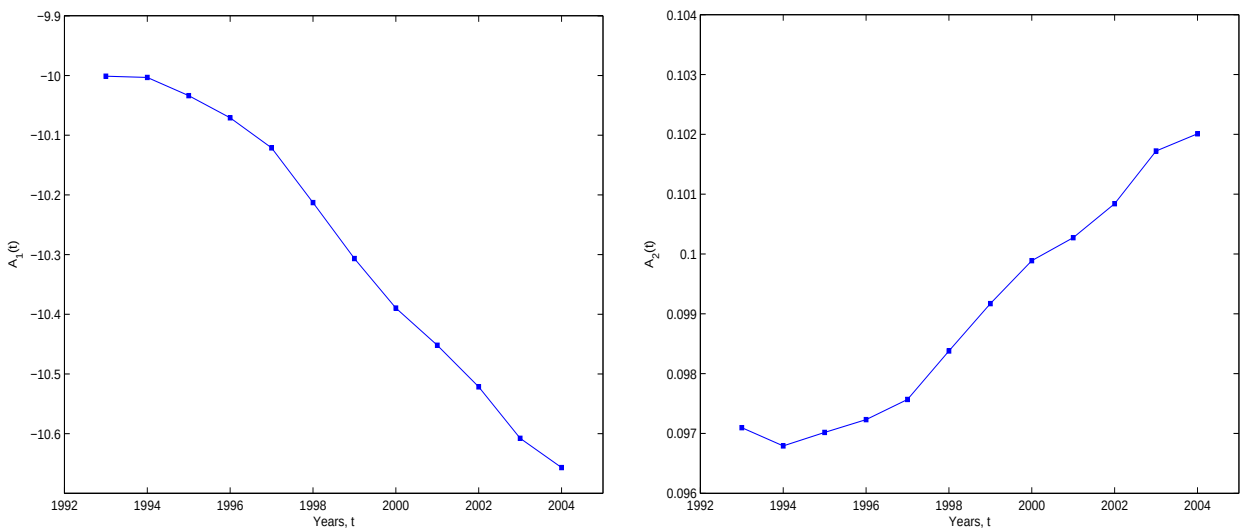
*by ignoring the irrelevant for modeling old-age mortality first two terms and setting* $G = e^{A_1}$ *and* $H = e^{A_2}$ *so that* $p_x = \frac{1}{1+e^{A_1+A_2 x}}$.

After estimating the past values of the A-process, one can easily find an estimate for its mean $\mu$ and covariance matrix $CC^\top$ using assumption (6.20) about the dynamic of $A$.

We want to demonstrate this method with the help of data for Germany. For that reason, we define as a reference population the German males aged 60-90 and use the official data published by the Federal Statistical Office from 1993-2004. For the past values of the $A$-vector, we obtain Figure 6.2. and for the needed estimators:

$$\hat{\nu} = \begin{pmatrix} -0.05959 \\ 0.0004465 \end{pmatrix}, \quad \widehat{\mathbf{CC}^\top} = \begin{pmatrix} 0.000842388 & -0.00001004 \\ -0.00001004 & 0.000000124 \end{pmatrix}, \quad \hat{A}(2004) = \begin{pmatrix} -10.65677 \\ 0.102009 \end{pmatrix}.$$

[3]For the needs of the pricing of the EIB/BNP longevity bond we limit the starting ages of the cohorts to 60 until 90 ones. The reason is that more than two time dependent factors are needed to match the survival probabilities at all ages and in addition, this is not needed for pricing longevity bonds.




(a) Estimated value of $A_1(t)$.

(b) Estimated value of $A_2(t)$.

Figure 6.2: Data for the German males (aged 60-89) from 1993-2002 from the Federal Statistical Office(SBA).

Although the covariance between the processes $A_1(t)$ and $A_2(t)$ is very small, their correlation is $-0.98238$. We can notice from the covariance matrix that both processes develop almost deterministically, which is additionally caused by the smoothing of the mortality data done by the Federal Statistical Office. To demonstrate this, we plot in Figure 6.3 the historical mortality rates cited by the Federal Statistical Office and their unsmoothed equivalents by the Human Mortality Database (http://www.mortality.org/ ) for the year 2002.

The respective estimated parameters using historical mortality data from 1991-2002 by the Human Mortality Database (HMD) are found as

$$\hat{\nu} = \begin{pmatrix} -0.03917 \\ 0.0001478 \end{pmatrix}, \quad \widehat{\mathbf{CC}^\top} = \begin{pmatrix} 0.002924269 & -0.00004349 \\ -0.00004349 & 0.000000690 \end{pmatrix}, \quad \hat{A}(2004) = \begin{pmatrix} -10.84976 \\ 0.105443 \end{pmatrix}$$

and the correlation between the processes is estimated to $-0.96818$.

Due to the big variation in the HMD data of the historical mortality rates at higher ages, we can perform weighted Least-squares method in order to estimate the needed parameters. We refer to Yue [64] for analysis on the usage of different weights in the estimation of the parameters from historical mortality data and chose as weights the number of individuals in each age group.

We obtain

$$\hat{\nu} = \begin{pmatrix} -0.05604 \\ 0.0003863 \end{pmatrix}, \quad \widehat{\mathbf{CC}^\top} = \begin{pmatrix} 0.000970319 & -0.00001166 \\ -0.00001166 & 0.000000145 \end{pmatrix}, \quad \hat{A}(2004) = \begin{pmatrix} -10.56728 \\ 0.1007317 \end{pmatrix}.$$

The correlation $\rho$ becomes now $-0.982504$, i.e. "smoothing" the data by using weights increased the correlation between the base processes. Due to its transparency, the "raw" data by the HMD can be used for a variety of analysis. For an example, one can analyze the effects of the different smoothing method on the correlation of the stochastic factors. In this respect, the SBA cites only the already smoothed data. However, since only the data by the SBA has an official character, in this thesis, we base our analysis on it.

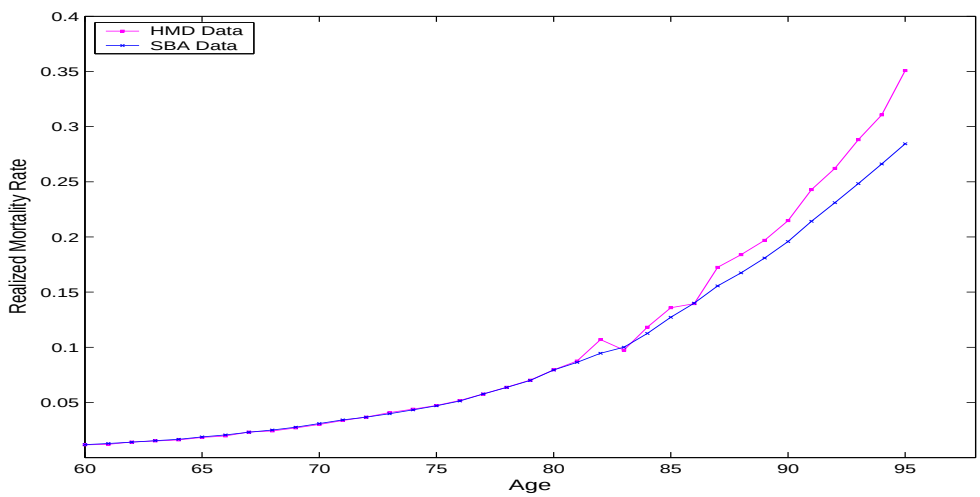



Figure 6.3: Mortality rates for the German males (aged 60-95) estimated by the HMD and the SBA for the year 2002.

### 6.3.2 Milevsky and Promislow's model

Cairns *et al.* offer as a modification of the Milevsky and Promislow's model for the development of the mortality rates the following stochastic model of the Gompertz type[4]

$$\mu_x^{(G)}(t) = \alpha e^{\beta x + \sigma Y(t)}, \quad \alpha, \beta, \sigma > 0 \tag{6.21}$$

[4]Originally, Milevsky and Promislow offer this model for the "instantaneous" mortality rate $h(t)$, which they call "hazard rate" such that

$$p_t(T) = e^{-\int_t^T \mu_s ds} = E\left(e^{-\int_t^T h_s ds}\right)$$

where $Y(t)$ follows an Ornstein-Uhlenbeck process given as

$$dY(t) = -\nu Y(t)\,dt + dW(t), \quad Y(0) = 0, \quad \nu \geq 0. \tag{6.22}$$

As we want to calibrate the model to the historical German data, we need first to estimate the values of $\mu_x^G(t)$. For that reason, using the relation

$$p(t,x) := p(t+1,t,t+1,x) = \exp\left(-\int_t^{t+1} \mu_{x+s}^{(G)}(s)\,ds\right) \tag{6.23}$$

we obtain for past, integer time points $t$

$$\frac{d}{dt}\ln(p(t,x)) = \mu_{x+t}^{(G)}(t) - \mu_{x+t+1}^{(G)}(t+1). \tag{6.24}$$

While we can now approximate the left side of (6.24) through a suitable difference such as for example $\ln(p(t,x)) - \ln(p(t-1,x)))$, for the proper representation of the right side, we are searching for an appropriate starting value $\mu_x^G(0)$ that can be used to approximate the rest of the realized mortality rates. One way to do this is to use the "earlier" realized survival probabilities and to calibrate a classical Gompertz model to them. For that purpose, for the starting year $t_0$ we assume that

$$p(t_0,x) = p_x = \exp\left(-\int_x^{x+1} \mu_s^{(G)} ds\right) = \exp\left(-\int_x^{x+1} \alpha e^{\beta s} ds\right) = \exp\left(-\frac{\alpha}{\beta}\left(e^{\beta(x+1)} - e^{\beta x}\right)\right)$$

and in order to estimate the parameter $\alpha$ and $\beta$ we use

$$\ln(-\ln(p_x)) = \left(\ln(\alpha) - \ln(\beta) + \ln\left(e^{\beta} - 1\right)\right) + \beta x.$$

Substituting

$$\begin{aligned} A: &= \ln(\alpha) - \ln(\beta) + \ln\left(e^{\beta} - 1\right) \\ B: &= \beta \end{aligned}$$

where

$$\begin{aligned} dh(t) &= \alpha e^{\beta t + \sigma Y(t)}, \quad \alpha, \beta, \sigma > 0 \\ dY(t) &= -\nu Y(t)\,dt + dW(t), \quad Y(0) = 0, \quad \nu \geq 0. \end{aligned}$$

Notice that in this form the "hazard rate" depends only on the time but not on the cohort age.

we obtain via (weighted) least squares the estimators $\hat{A}$ and $\hat{B}$ and therefore also

$$\begin{aligned}\hat{\beta} &= \hat{B} \\ \hat{\alpha} &= \exp\left(\hat{A} + \ln(\hat{\beta}) - \ln\left(e^{\hat{\beta}} - 1\right)\right).\end{aligned}$$

The parameters $\nu$ and $\sigma$ of the Ornstein-Uhlenbeck process can now be estimated via the Maximum-Likelihood method using that the process $Y(t)$ possesses a transition density of a normal distribution or more precisely, using that

$$\sigma\left(Y\left(t+1\right) - Y\left(t\right)\right) \sim N\left(\sigma Y\left(t\right) e^{-\nu}, \sigma^2 \frac{1-e^{-2\nu}}{2\nu}\right). \tag{6.25}$$

For the observations $\ln\left(\mu_x^{(G)}\left(t_i+1\right)/\mu_x^{(G)}\left(t_i\right)\right)$, $i = 1, ..., n$ one can obtain the Log-Likelihood function (see Ait-Sahalia [1]):

$$\begin{aligned}l_n\left(\nu,\sigma\right) &= -n\left(\ln\left(\sqrt{2\pi}\right) + \ln\left(\sigma\right) + \ln\left(\sqrt{\frac{1-e^{-2\nu}}{2\nu}}\right)\right) \\ &\quad - \sum_{i=1}^{n} \frac{\left(\ln\left(\frac{\mu_x^{(G)}(t_i+1)}{\mu_x^{(G)}(0)}\right) - \ln\left(\frac{\mu_x^{(G)}(t_i)}{\mu_x^{(G)}(0)}\right) e^{-\nu}\right)^2}{2\sigma^2 \frac{1-e^{-2\nu}}{2\nu}}.\end{aligned} \tag{6.26}$$

Finally, numerical maximization of the Log-likelihood function delivers the needed estimates of the parameters $\nu$ and $\sigma$ of the Ornstein-Uhlenbeck process.

If we recall that in the case of Germany we dispose only with data from 1993-2004, we chose 1993 as a starting year and the German males aged $60-89$ as a reference population. Thus, for the calibration of $\mu_x^{(G)}(0)$, we obtain

$$\begin{aligned}\hat{\alpha} &= 0.00005759 \\ \hat{\beta} &= 0.09263573.\end{aligned}$$

which we use in the Maximum-Likelihood Method to obtain

$$\begin{aligned}\hat{\sigma} &= 0.02751 \\ \hat{\nu} &= 0.\end{aligned}$$

### 6.3.3 Idea for a Model: Stochastic Gompertz

If we assume for a moment that the mortality rate follows the standard Gompertz model

$$\mu_x = \alpha e^{\beta x}, \alpha, \beta > 0$$

and take the case of Germany with observations for the last twelve years of the survival probabilities $p(t, x)$, then calibrating the parameters $\alpha$ and $\beta$ for each year separately using weighted LS (see Yue [64]) and plotting them over time yields Figure 6.4. It is easy to notice now that both parameters exhibit strong time dependence, which is a reflection of the improvement of longevity over time. In specific, the $\alpha$ parameter is responsible for the general improvement at all ages and the $\beta$ for the improvement at different ages.

Choosing a model for $\alpha(t)$, we need to ensure that it stays positive (required by the Gompertz construction) and also decreases with time (to reflect the improving mortality).

For that reason, we choose a model of the following form

$$\begin{aligned} d\alpha(t) &= -\kappa\alpha(t)dt + \sigma_1\alpha(t)dW_1(t), \quad \kappa > 0, \sigma > 0 \\ \alpha(0) &= \alpha_0. \end{aligned}$$

where $W_1(t)$ denotes a Brownian motion under the physical measure. The solution of the above SDE is given by

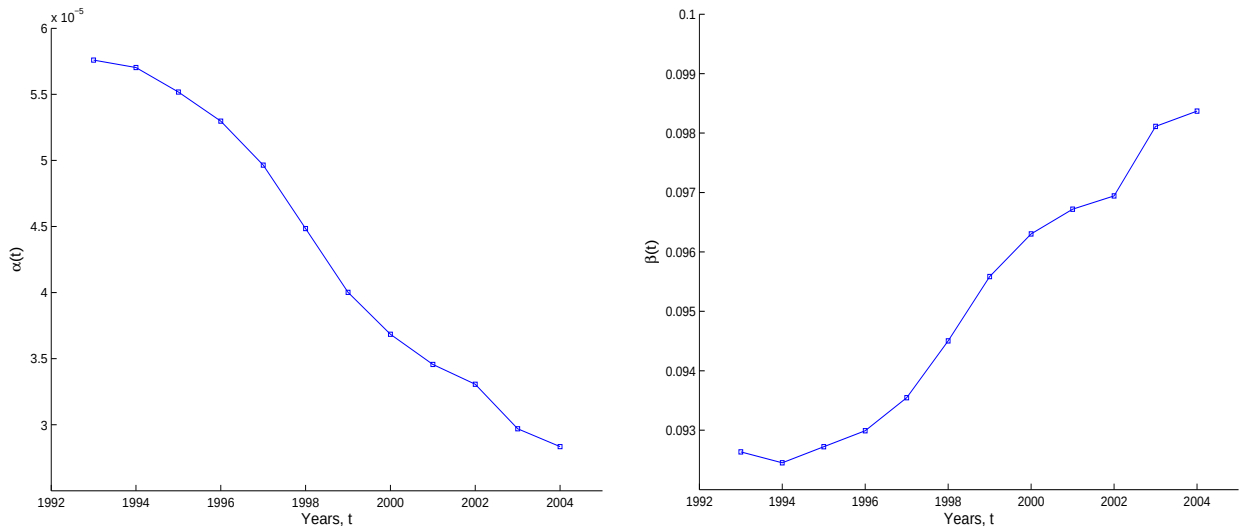
$$\alpha(t_i) = \alpha(t_{i-1})\left(-\kappa\left(t_i - t_{i-1}\right) - \frac{1}{2}\sigma_1^2\left(t_i - t_{i-1}\right) + \sigma_1\left(W_1(t_i) - W_1(t_{i-1})\right)\right), \quad \text{for} \quad t_{i-1} \leq t_i.$$

Thus, notice that since we have $\hat{\alpha}(2004) = 0.000028343 > 0$, the requirement of positivity of the $\alpha(t)$ process is fulfilled. We can easily calibrate the parameters $\kappa$ and $\sigma_1$ using the estimated time series $\hat{\alpha}(1993), \ldots, \hat{\alpha}(2004)$ and the natural logarithm of the $\alpha$-process:

$$d\ln\left(\alpha(t)\right) = -\left(\kappa + \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 dW_1(t), \quad \ln\left(\alpha(0)\right) = \ln\left(\alpha_0\right).$$

Denoting $\delta := -\kappa - \frac{1}{2}\sigma_1^2$ and using the estimators

$$\begin{aligned} \hat{\delta} &= \frac{1}{11}\sum_{i=1993}^{2003} \ln\left(\frac{\alpha(t_{i+1})}{\alpha(t_i)}\right) \\ \hat{\sigma}_1^2 &= \frac{1}{10}\sum_{i=1993}^{2003}\left(\ln\left(\frac{\alpha(t_{i+1})}{\alpha(t_i)}\right) - \hat{\delta}\right)^2, \end{aligned}$$




(a) Estimated value of $\hat{\alpha}(t)$.

(b) Estimated value of $\hat{\beta}(t)$.

Figure 6.4: Data for the German males (aged 60-89) from 1993-2002 from the Federal Statistical Office(SBA).

we find for the data of the SBA

$$\begin{aligned} \hat{\sigma}_1^2 &= 0.0009681 \\ \hat{\kappa} &= -\hat{\delta} - \frac{1}{2}\hat{\sigma}_1^2 = 0.0632616 \end{aligned}$$

At the same time, we have also estimated $\hat{\beta}(1993), \ldots, \hat{\beta}(2004)$ and assuming that $\beta$ follows a measurable stochastic process of the form

$$\begin{aligned} d\beta(t) &= \mu dt + \sigma_2 dW_2(t), \quad \sigma_2 > 0 \\ \beta(0) &= \beta_0, \end{aligned}$$

where $W_2(t)$ denotes a Brownian motion under the physical measure, we can calculate estimates of the mean $\mu$ and the variance $\sigma_2^2$ using

$$\begin{aligned} \hat{\mu} &= \frac{1}{11} \sum_{i=1993}^{2003} \left( \hat{\beta}(t_{i+1}) - \hat{\beta}(t_i) \right) \\ \hat{\sigma}_2^2 &= \frac{1}{10} \sum_{i=1993}^{2003} \left( \hat{\beta}(t_{i+1}) - \hat{\beta}(t_i) - \hat{\mu} \right)^2 . \end{aligned}$$

and obtain for the SBA data:

$$\begin{aligned}\hat{\mu} &= 0.000521236\\ \hat{\sigma}_2^2 &= 0.000000141394.\end{aligned}$$

Assuming a correlation between the Brownian motions $W_1$ and $W_2$ and denoting it by $\rho$, we obtain as an estimate of it $\hat{\rho} = -0.98279$. We need also for our simulation purposes $\hat{\beta}^{SBA}(2004) = 0.098369335$. We notice that although the $\beta$-process can become negative with a positive probability, this case is highly unlikely to happen due to the positive mean, the positive starting value (e.g. for a simulation starting in 2004 we use $\beta(0) = \hat{\beta}(2004) = 0.098369335$) and the small variance.

To sum up, we offer a model we name "Stochastic Gompertz" which models the mortality rate as

$$\mu_x^{SG}(t) = \alpha(t)e^{\beta(t)x}$$

for a stochastic, measurable process $\alpha(t)$

$$d\alpha(t) = -\kappa\alpha(t)dt + \sigma_1\alpha(t)dW_1(t), \quad \alpha(0) = \alpha_0, \quad \kappa, \sigma_1 > 0$$

and a second stochastic, measurable process $\beta(t)$ with the dynamics:

$$\begin{aligned}d\beta(t) &= \mu dt + \sigma_2 dW(t), \quad \sigma_2 > 0\\ \beta(0) &= \beta_0\end{aligned}$$

where both Brownian motions are given under the physical measure and their correlation is denoted by $\rho$.

**Remark 6.3.2.** *Similar to the model of Cairns et al., the stochastic Gompertz model is based on two stochastic factors. The first one takes into account the general improvement (or deterioration) of mortality that influences all ages and the second one takes into account the improvement (or deterioration) of mortality that influences higher ages more than lower ones. It is an interesting remark, that both Cairns et al. model and the stochastic Gompertz model show increasing trend of improving mortality at all ages (notice the decreasing trend in $A_1(t)$ and $\alpha(t)$) and deterioration of life expectancy at higher ages (notice the increasing trend of $A_2(t)$ and $\beta(t)$). Both*

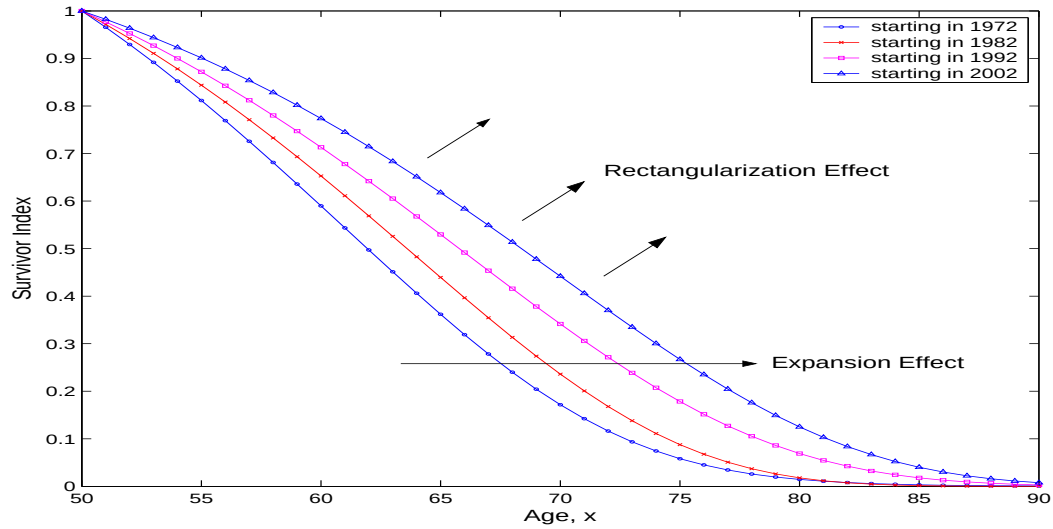



Figure 6.5: Simulation of the Survivor Index for the 65-years old males in 1972, 1982, 1992 and 2002 using data for West Germany from the HMD.

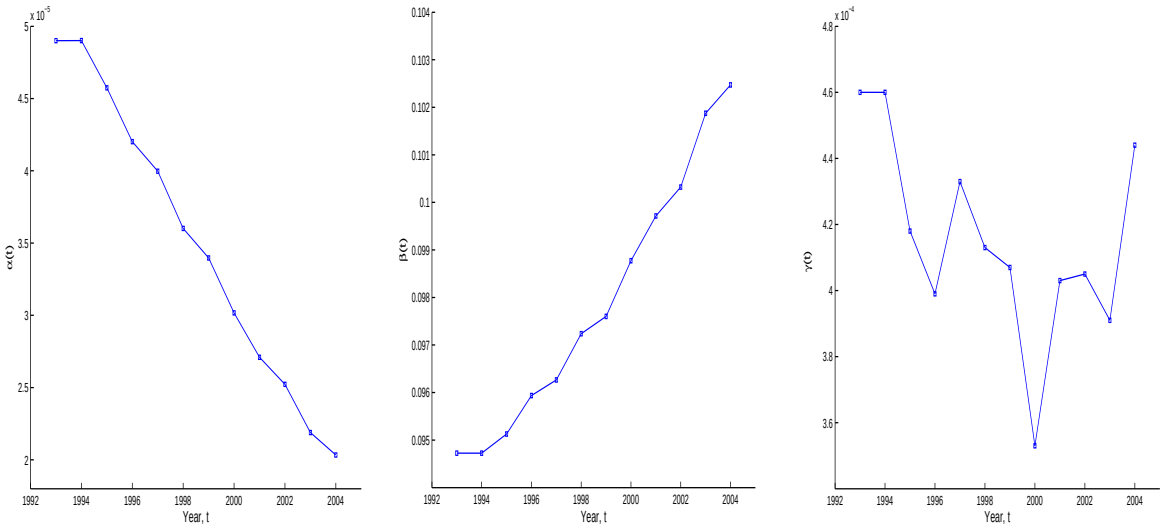
*effects lead to the so called* ***"rectangularization"*** *of the survival probabilities curve which has the interpretation that due to medical improvements or better life standards, the amount of people at middle ages was (and is) increasing with the years, but the amount of people at extremely high ages, was (is) increasing at much lower tempo, or was (is) even decreasing. The translation of the survival probabilities curve to right is called "expansion" effect. As an example of the both effects, we have plotted in Figure 6.5 the simulated expected value of the survivor index for the ages* $65 - 105$*, conditioned on survival until age* $65$ *and starting the simulation in years 1972, 1982, 1992 and 2002.*

### 6.3.4 Idea for a Model: Stochastic Makeham

If we consider the classical Makeham model, instead of the Gompertz one (recall (6.10)), we would have

$$p\left(t_0, x\right) = p_x = \exp\left(-\int_x^{x+1} \mu_s^{(G)} ds\right) = \exp\left(-\frac{\alpha}{\beta}\left(e^{\beta(x+1)} - e^{\beta x}\right) - \gamma\right).$$

In order to estimate the parameter $\alpha$, $\beta$ and $\gamma$ we have to use now a non-linear least squares minimization where a special care has to be taken to find the global solution due to the variety




(a) Estimated value of $\hat{\alpha}(t)$. (b) Estimated value of $\hat{\beta}(t)$. (c) Estimated value of $\hat{\gamma}(t)$.

Figure 6.6: Data for the German males (aged 60-89) from 1993-2002 from the Federal Statistical Office(SBA).

of very similar local minima in the objective function. We refer here to methods for global optimization such as the "Simulated annealing" and others.

In addition, since we have in the stochastic Makeham model two factors that influence the general mortality ($\alpha$ and $\gamma$) and one modeling the age-specific improvements in the mortality ($\beta$), the intuitive interpretation of the Cairns *et al.* [15] case or the stochastic Gompertz one is now lost. We choose to model only the $\gamma$-process as a stochastic one, such that

$$\begin{aligned} d\alpha(t) &= -\kappa\alpha(t)dt, \quad \alpha(0) = \alpha_0 \\ d\beta(t) &= \mu dt, \quad \beta(0) = \beta_0 \\ d\gamma(t) &= \nu dt + \sigma dW(t), \quad \gamma(0) = \gamma_0,\ \sigma > 0. \end{aligned}$$

For completeness, we present here also the results for the Makeham case, where the historical series are given in Figure 6.6 and the estimated parameters for the data by the SBA from 1993

until 2004 are found as

$$\begin{aligned}
\hat{\kappa} &= 0.0799 \\
\hat{\mu} &= 0.000704 \\
\hat{\nu} &= -0.00000144 \\
\hat{\sigma} &= 0.0007408,
\end{aligned}$$

where $\hat{\alpha}(2004) = 0.0000203$, $\hat{\beta}(2004) = 0.1025$ and $\hat{\gamma}(2004) = 0.000444$.

### 6.3.5 Mortality and Short Rate Modeling

On the basis of the formulas (6.4) and (6.8), one can notice that for the time dependent development of the stochastic mortality rates it is possible to consider a variety of well known models from the short rate world (as long as these models ensure the positivity of the short rates). However, some feature of these models, such as the mean reversion, will make no sense in the modeling of mortality rates so that we can already reject the Vasicek, CIR and Hull-White models from the list of possible models. In order to model the increase of the mortality rates, with the increase of the age of the individual, log-normal models such as the Dothan or the Black-Karasinski seem to be appropriate.

In this respect, we are not going to present here any specific model but we will remark that the stochastic mortality rates, modeled by Milevsky and Promislow, resemble the short rate Black-Derman-Toy model and the model by Cairns *et al.* [15] can be seen as a discrete in time variation of the Dothan short rate model.

In all the presented examples, we have shown how to calibrate the parameters of the chosen stochastic processes to the observed death frequency. As next, we are going to use these estimated parameters in a Monte Carlo simulation, in order to estimate the future values of the coupon payments $S(t)$, needed for the pricing of the longevity bond.

## 6.4 Pricing of Longevity Bonds

In Assumption 6.2.1, we require the independence of the interest rate and longevity risk, which directly implies that we can separate the pricing of both risks in a multiplicative way. Here, we should mention that the short rate risk is tradable (i.e. linear short rate products such as bonds are tradable) and therefore it should be priced under an appropriate martingale measure $Q$ (which means that the discounted with the money market account longevity bond price process should be a martingale under the measure $Q$). Therefore, if we further assume that the market price $P(t,T)$ at time $t$ of a zero coupon bond with maturity $T$ is given by

$$P\left(t,T\right)=E_{Q}\left(\exp\left(-\int_{t}^{T}r\left(s\right)ds\right)\middle|f_{t}\right)$$

(i.e. the interest market uses for pricing exactly measure $Q$, or an equivalent to it), then we can write the price of the longevity bond as

$$P^{(L)}\left(t\right)=E_{Q}\left(\sum_{i=1}^{N}\exp\left(-\int_{t}^{i}r\left(s\right)ds\right)S\left(i\right)1_{\{t\leq i\}}\middle|f_{t}\right)=\sum_{i=1}^{N}P\left(t,i\right)E_{Q}\left(S\left(i\right)|f_{t}\right)1_{\{t\leq i\}},$$

given that we have assumed that the interest rate risk and the mortality rate risk continue to be independent under the pricing measure $Q$.

Due to the non tradability of the mortality risk, we can not apply here any of the standard pricing principles of the financial mathematics (such as no arbitrage or replication principle) as they are here either not valid or not helpful. We offer to use instead one of the classical pricing principles, coming from the insurance mathematics (such as the expectation or the variance one). As a result, we offer a pricing concept that combines the no-arbitrage principle, on the interest rate side and a security margin in the spirit of the insurance mathematics as an insurance against the (non tradability of the) mortality risk, on the mortality side.

In this connection, we assume that the price of a zero coupon bond in the interest rate world is given as a conditional expectation under measure $Q$, which we use for the pricing of the longevity bond and which uniquely determines (only!) the interest rate component in the upper equation for the price of the longevity bond. Notice that the only probability measure that describes the time dependent development of mortality is the physical $P$-measure, which is defined by

the calibration of the stochastic processes to the historical mortality data. Independent of the underlying mortality model (be it a classical, time dependent deterministic or stochastic one), the measure $P$ is objectively verifiable and in addition it is compatible with the risk-neural pricing formula when

$$P^{(L)}(t) = E_Q\left(\sum_{i=1}^{N} \exp\left(-\int_t^i r(s)\,ds\right) S(i)\,1_{\{t\leq i\}}\right) = \sum_{i=1}^{N} P(t,i)\,E_P\left(S(i)\,|f_t\right)1_{\{t\leq i\}} \tag{6.27}$$

is valid, i.e. when $P$ is responsible for the "mortality component of Q".

In practice, the prices of the traded longevity bonds will however be greater that the ones calculated using the upper formula and obtained following the principles of financial mathematics. The reason for this, can be explained in two ways:

1.*Application of the Expectation Principle*

We can write the relationship between the market price of the longevity bond $P_M^{(L)}(0)$ and $P^{(L)}(0)$ as

$$P_M^{(L)}(0) = (1+\delta)\,P^{(L)}(0)$$

for a constant $\delta > 0$ and explain the higher market price $P_M^{(L)}(0)$, offered by the issuer, with the application of the actuarial "expectation principle" with a premium parameter $\delta$. It is also possible to explain the difference between the "fair" price of the longevity bond and its market price by means of other actuarial principles, such as the "(semi)-variance principle", the "exponential principle", the "quantile principle", the "Esscher principle" or others. However, the more complicated relationship using higher moments will also require a more complicated calculation of the part of the payments that is influenced by the mortality risk.

2.*Risk Neutral Pricing and Change of Measure*

In this section, keeping the risk-neutral pricing approach, we want to explain the difference between the "fair" price and the market price of the longevity bond. This is done with the help of a new martingale measure $Q(\lambda)$, which describes the mortality component with a measure $P_\lambda$, instead with the physical measure $P$. Cairns, Blake and Dowd [15] apply the measure $P_\lambda$ by simply substituting the drift vector $\nu$ of the underlying process, under measure $P$ by $\nu - C\lambda$,

under measure $P_\lambda$. Notice that in this sense, $\lambda$ plays the role of the "market price of mortality risk". Using numerical methods, one can try to estimate such a vector $\lambda$ by matching

$$P_M^{(L)}(0) = P_{Q(\lambda)}^{(L)}(0) = \sum_{i=1}^{N} P(0,i)\, E_{P_\lambda}(S(i)). \tag{6.28}$$

Notice that a solution vector $\lambda$ to (6.28) must not! necessarily exist.

We are now going to demonstrate the application of the change of measure method with the help again of the German historical data. For that purpose, let us consider the Cairns, Blake and Dowd [15] model and a reference population of German males aged 60-89. Recall we have already estimated for them the required parameters $\hat{\nu}, \hat{C}$ and $\hat{A}(2004)$. Setting $N = 25$, the initial cohort age to 65, $\delta = 0.0183$, the interest rate to 4% and solving numerically (6.28) yields the solution set

$$I := \left\{ (\lambda_1^*, \lambda_2^*) \quad s.t. \quad \left( P_M^{(L)}(0) - P_{Q(\lambda)}^{(L)}(0) \right)^2 \leq \varepsilon \right\}$$

for some small tolerance parameter $\varepsilon$. Plotting the surface of the squared error in Figure 6.7 (as a function of $\lambda_1$ and $\lambda_2$) we can notice that the solution set of the minimization problem converges to (an almost straight) line. Choosing points on the solution line[5], changes the curvature of the survivor index $S(t)$ (see Figure 6.8), and can be interpreted as a choice of a measure that puts more weight on earlier or later coupons. The motivation which measure to choose can depend for example on the term structure of interest rates or other internal bank principles.

Applying the same method on the stochastic Gompertz model such that the drift of the $\beta(\cdot)$ process is substituted by $\mu - \sigma\lambda$ for $\lambda \in \mathbb{R}$ and solving numerically (6.28), we obtain Figure 6.9. Note that as in the Cairns *at el.* [15] case, there is a variety of solutions solving (6.28).

**Monte Carlo Simulation of the Longevity Bond:**

In order to be able to estimate the price of the longevity bond at time 0, we need to simulate

$$E(S(i)|f_0), \quad \text{for} \quad i = 1, \ldots, N$$

where $N$ is the number of coupons of the longevity bond.

[5]Cairns *et al.* consider only measures of the form $Q_{(\lambda_1,0)}$, $Q_{(0,\lambda_2)}$ and $Q_{(\lambda,\lambda)}$ which seem to be close to the solution line.

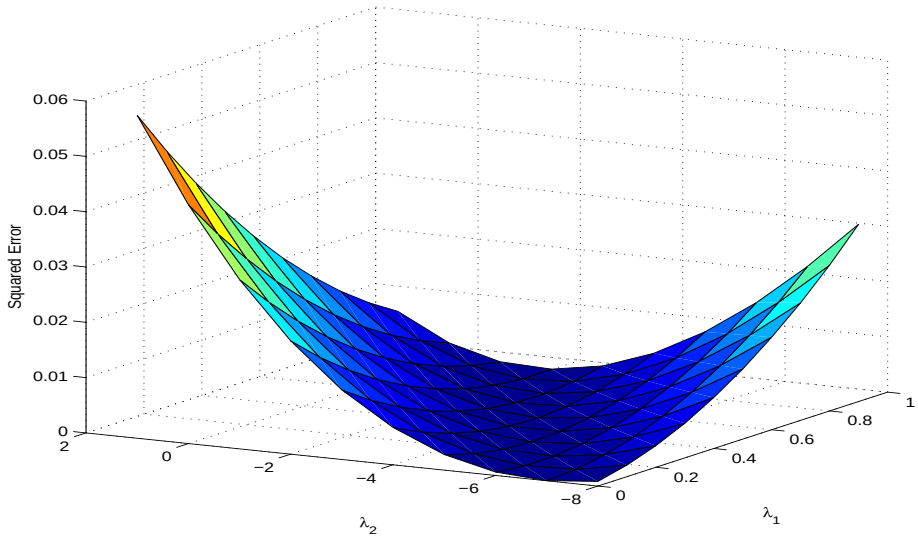



Figure 6.7: Surface of the squared differences between $P_M^{(L)}(0)$ and $P_{Q(\lambda)}^{L}(0)$.

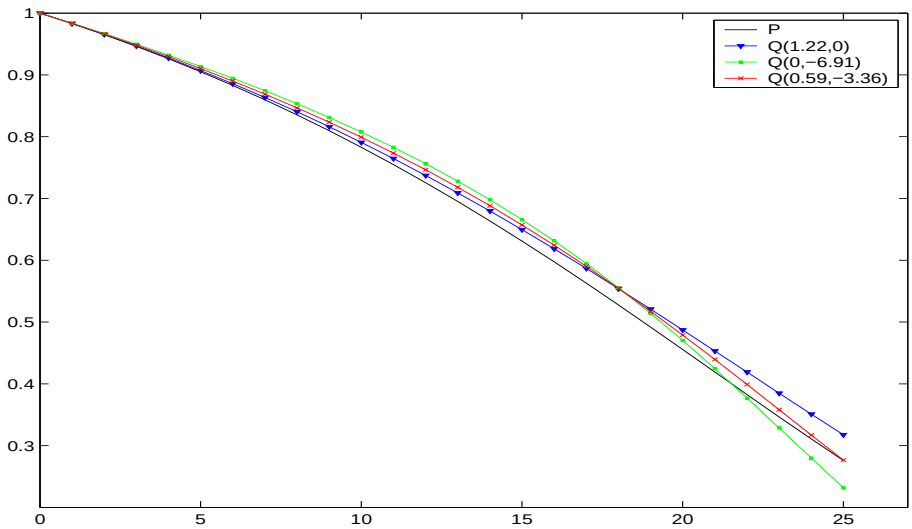



Figure 6.8: Estimated value of $E[S(t)]$ for the German males aged 60-89, using the Cairns *et al.* model, under the measures $P, Q(1.22, 0), Q(0, -6.91), Q(0.59, -3.36)$. Data from 1993-2004. Number of MC simulations 5000.

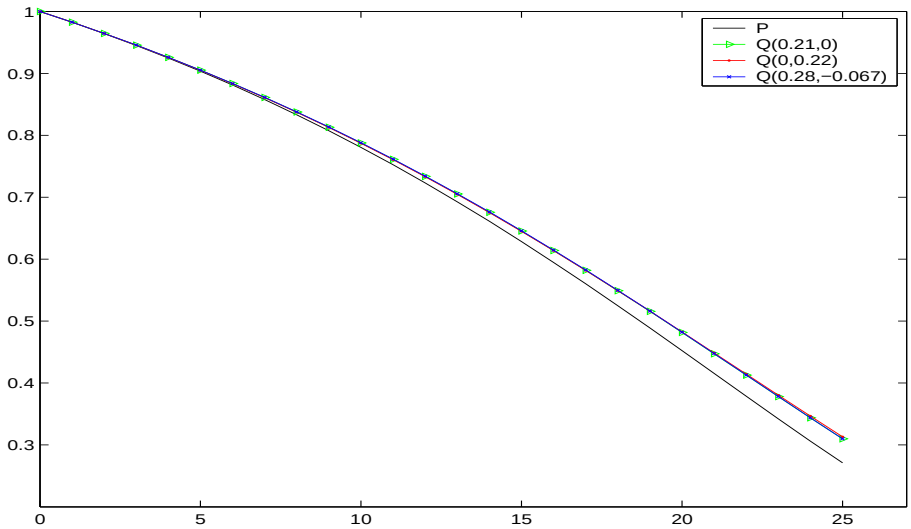



Figure 6.9: Estimated value of $E\left[S(t)\right]$ for the German males aged 60-89, using the stochastic Gompertz model, under the measures $P, Q(0.21, 0), Q(0, 0.22), Q(0.28, -0.067)$. Data from 1993-2004. Number of MC simulations 5000.

Then, in the Cairns *et al.* [15] case, using

$$S(i) = \hat{p}(1, 0, 1, x), \ldots, \hat{p}(i-1, i, i-1, x)$$

we approximate

$$E\left(S(i)|f_0\right) \approx \frac{1}{M}\sum_{j=1}^{M} \hat{p}_j(1, 0, 1, x), \ldots, \hat{p}_j(i-1, i, i-1, x), \quad \text{for} \quad i = 1, \ldots, N$$

where $M$ is the number of Monte Carlo simulations, $\hat{p}_j(1, 0, 1, x), \ldots, \hat{p}_j(i-1, i, i-1, x)$ is the $j$-th path of simulated realized survival probabilities and the convergence follows by the Strong Law of Large Numbers.

As we have already mentioned, using Monte Carlo simulation allows us to calculate confidence intervals for the survival probabilities. For that reason, let us denote each path $Y_{ij} := \hat{p}_j(1, 0, 1, x), \ldots, \hat{p}_j(i-1, i, i-1, x)$, $j = 1, \ldots, M$, $i = 1, \ldots, N$. Thus, since we simulate independent paths of the underlying process, we have that $Y_{i1}, \ldots, Y_{iM}$ are *i.i.d.* with mean $\mu_i = E(S(i)|f_0)$ and variance $\sigma_i$ and according to the Central Limit Theorem we have that

$$\frac{1}{M}\sum_{j=1}^{M} Y_{ij} \xrightarrow{a.s.} Z \sim \mathcal{N}\left(\mu_i, \frac{\sigma_i^2}{M}\right).$$

Using that $\frac{\bar{Y}_i-\mu_i}{\hat{S}_{i,M}/\sqrt{M}} \sim t_{M-1}$ where $\bar{Y}_i := \frac{1}{M}\sum_{j=1}^{M} Y_{ij}$, we obtain for the mean value $\mu_i = E\left(S(i)|f_0\right)$ and for $\gamma = 90\%$, the following asymptotic confidence interval

$$\left[\bar{Y}_i - \frac{\hat{S}_{i,M}}{\sqrt{M}} d_{\frac{1+\gamma}{2}}, \bar{Y}_i + \frac{\hat{S}_{i,M}}{\sqrt{M}} d_{\frac{1+\gamma}{2}}\right]$$

where $d_{\frac{1+\gamma}{2}}$ is the $\frac{1+\gamma}{2}$ quantile of the student $t_{M-1}$ distribution and $\hat{S}^2_{i,M} := \frac{1}{M-1}\sum_{j=1}^{M}\left(\bar{Y}_i - Y_{ij}\right)^2$.

In the stochastic Gompertz model, using

$$E\left(S(i)|f_0\right) = E\left(e^{-\int_0^i \mu_{x+s}(s)ds}\right), \quad \text{for} \quad i = 1, \ldots, N$$

we approximate

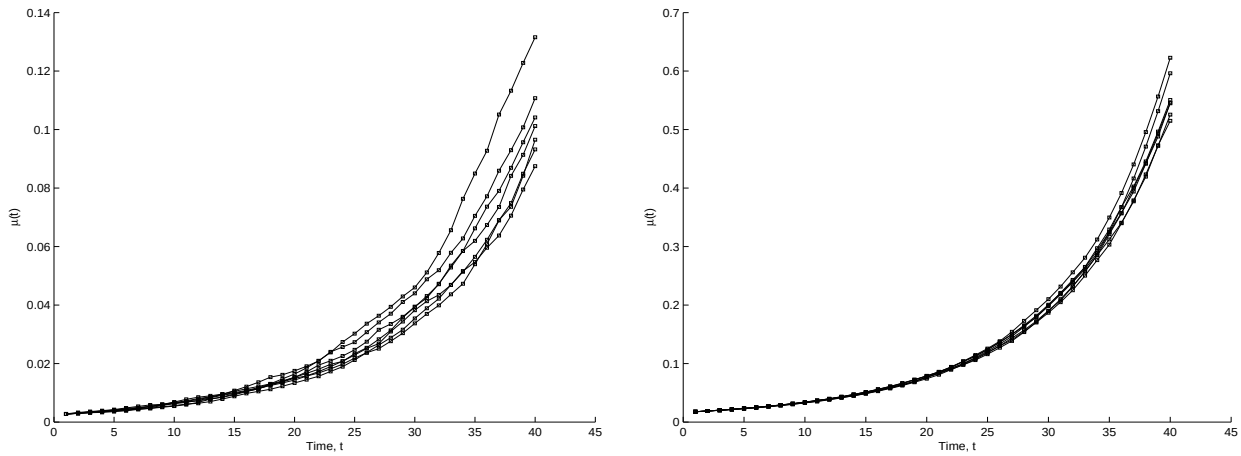
$$\begin{aligned} E\left(S(i)|f_0\right) &\approx \frac{1}{M}\sum_{j=1}^{M} \exp\left(-\sum_{t=1}^{i}\sum_{k=1}^{m} \mu^j_{x+(t-1)+\frac{k}{m}}\left(t-1+\frac{k}{m}\right)\frac{i}{m}\right), \quad \text{for} \quad i = 1, \ldots, N \\ &\approx \frac{1}{M}\sum_{j=1}^{M} \exp\left(-\sum_{t=1}^{i}\sum_{k=1}^{m} \alpha\left(t-1+\frac{k}{m}\right) e^{\beta\left(t-1+\frac{k}{m}\right)\left(x+t-1+\frac{k}{m}\right)\frac{k}{m}}\frac{i}{m}\right) \end{aligned}$$

where $m$ denotes the discretization of the integral between two coupon dates, needed for its numerical simulation. Note that if we set $m = 1$, the complexity of the problem will be the same as in the Cairns *et al.* [15] model. The derivation of the confidence intervals is the same as above, since it is based on the properties of the Monte Carlo simulation and not of the specific model.

## 6.5 Comparison Between the Models:

Let us now make a comparison between the four presented models, using the estimated parameters for the reference population of German males (aged 60-89) by the historical data of the SBA 1993-2004 For that purpose, setting the cohort age to 65 and $N = 40$, we simulate using MC the value of $E_P\left[S(i)\,|f_0\right]$ for $i = 1, \ldots, 40$ and plot the results in Figure 6.11. We can observe that the Cairns *et al.* curve is very close to the stochastic Gompertz and stochastic Makeham ones.

Further, notice that the Milevsky and Promislow curve predicts very high mortality rates, which is caused by the attempt to capture an almost deterministic time-dependent effect via parameters




(a) Milevski and Promislow's Model (b) Stochastic Gompertz Model

Figure 6.10: Simulated with Monte Carlo paths of the mortality rate $\mu_{x+t}(t)$ of the 65 (in 2004) year old German males.

responsible mostly for the volatility. Based on this results, one can conclude that the model of Milevsky and Promislow can not be used as a mortality rate dynamics (which was suggested by Cairns *et al.*) but should be mainly used to model the "instantaneous" mortality rate, as it was originally intended to.

To support additionally our statement, we compare in Figure 6.10 several Monte Carlo paths of the mortality rate $\mu(t)$ of the 65-year old (in 2004) German males, via the Milewski and Promislow's model and the stochastic Gompertz model . We observe strong fluctuations of the mortality rate in the Milewski and Promislow's model, and almost deterministic ones in the stochastic Gompertz model.

In this place, one can argue the need of a stochastic model for the development of the mortality rate and in fact, for short maturities, its stochastic modeling might be really not needed. However, since typically a longevity bond has a maturity of 25 years and longer, we can notice that for such long maturities, we do observe fluctuations in the mortality rate development, which motivates the effort of its modeling as a stochastic process in order to obtain a correct price for the longevity bond.

Finally, since we are interested in pricing a longevity bond, we want to compare the results using the different models. For that reason, we specify first the features of the bond to be as in

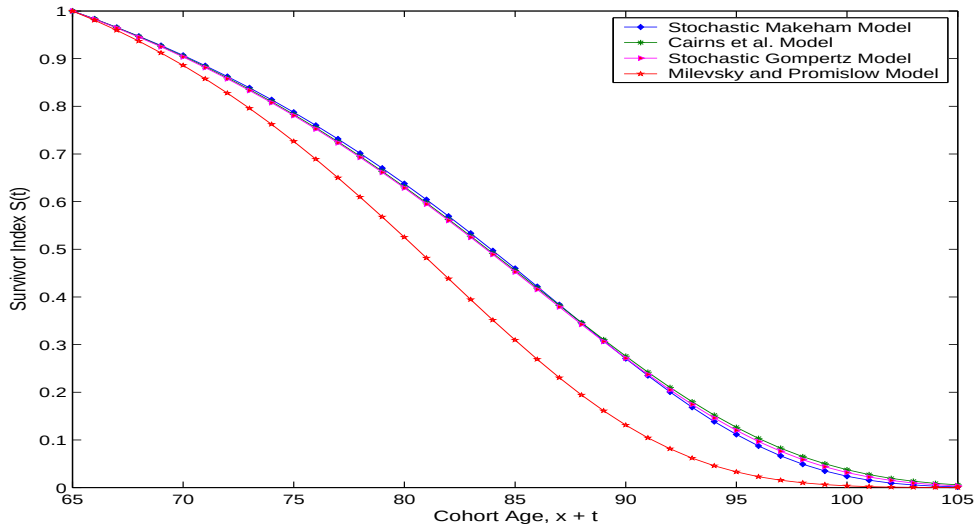



Figure 6.11: Estimated value of $E_P[S(t)]$ for the German males, using the models of Cairns *et al.*, Milevsky and Promislow and the stochastic Gompertz. Data from the SBA from 1993-2004.

the EIB/BNP longevity bond, i.e. maturity 25 years, cohort age 65 at 2004 and we also set for simplicity the interest rate to 4% and use 5000 MC simulations. Using (6.27), we obtain for the data published by the SBA:

| Modellart: | $P^L(0)$ |
|---|---|
| Cairns *et al.* | 11.388 |
| Milevsky and Promislow | 10.317 |
| stochastic Gompertz | 11.360 |
| stochastic Makeham | 11.416 |

Again, notice that the Cairns *et al.*, the stochastic Gompertz and the stochastic Makeham models deliver very similar results. Whereas, the Milevsky and Promoslow one strongly underprices the longevity bond (recall Figure 6.11).

## 6.6 Conclusion and Remarks

In this chapter, based on the article of Korn, Natcheva and Zipperer (2006), a general framework for pricing longevity bonds was presented. First the existing in the literature approaches for stochastic modeling of the survival probabilities (by Cairns *et al.*) and the mortality rates (by Milewski and Promislow) have been examined, with the help of the historical mortality data for

Germany from 1993-2004. And afterwards, two natural approaches for stochastic modeling of the mortality rates, expanding the well-known classical mortality laws of Gompertz and Makeham have been offered and examined.

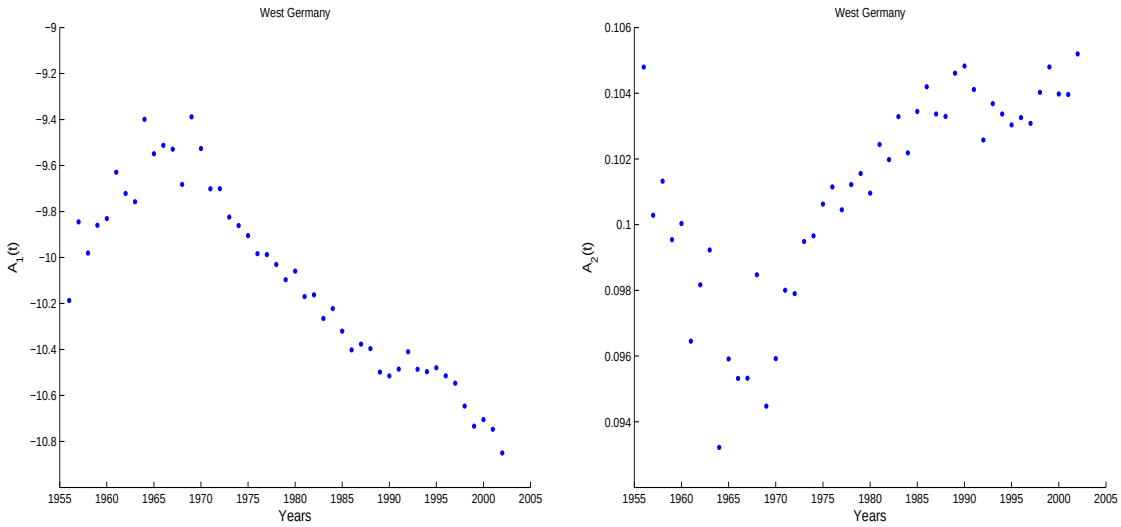
The comparison of the examined models has revealed, that besides the Milevsky and Promislow model, all models are capable of calibrating well the historical trend of increasing longevity and thus are appropriate to be used for its future forecasting. In addition, their stochastic components add fluctuation around that trend in a similar fashion and lead to very similar, though different results. The advantage of the offered models is based on fact that it is more natural to deal with the mortality rate stochastic dynamics (than the survival probability one) due to its strong similarity to the well-developed short rate modeling.

However, which of these models is more appropriate to use for pricing longevity bonds is a natural question to come, but not the goal of this Chapter. Our aim is to offer, additionally to the existing one, new possible ways to model stochastically mortality rates (and thus longevity) and leave it open for further discussion how the goodness of these models can be compared.

In this chapter, using the data for the German males, we have been limited to the historical data of only twelve years, in which well behaved, almost deterministic trend in longevity was to be observed. To show that this is generally not the case, we have plotted in Figure 6.12 the estimated values of $A_1(t)$ and $A_2(t)$ from the Cairns *et al.* model for the West German males, using data by the HMD from $1956 - 2002$. Additionally, in Figure 6.13 we have plotted their values for the East German males, from $1956 - 2002$, again using data by the HMD. We notice that political regime changes have a very strong influence on the trend of the mortality, where in specific, they cause jumps in the trend, change of its direction, etc.

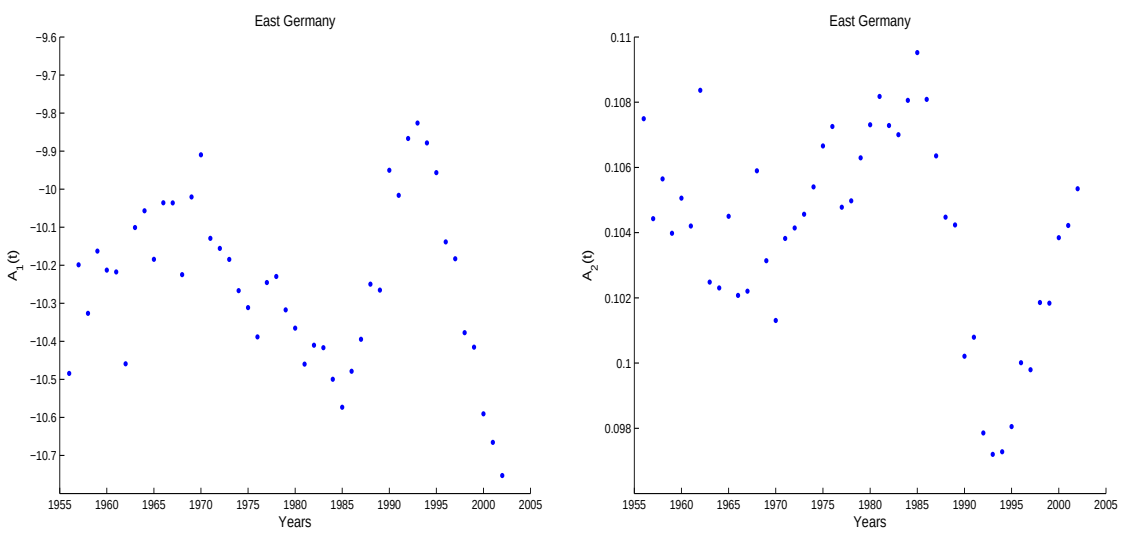
We shall now analyze the two possible jumps [6] in mortality, from the side of an insurer or a pension plan who buys a longevity bond:

[6] By the usage of "jump in mortality", it is meant a change in mortality, stronger that the usual fluctuation, e.g. Figure 6.12 and Figure 6.13. The small fluctuations have been modelled by adding Brownian motions to the mortality rate dynamics. The strong changes have not been explicitely modelled, due to the insufficient data for their calibration.




(a) Estimated value of $A_1(t)$. (b) Estimated value of $A_2(t)$.

Figure 6.12: Data for the West German males (aged 60-89) from 1956-2002 from the Human Mortality Database.




(a) Estimated value of $A_1(t)$. (b) Estimated value of $A_2(t)$.

Figure 6.13: Data for the East German males (aged 60-89) from 1956-2002 from the Human Mortality Database.

- **A regime change that causes longevity to strongly deteriorate**
  In this case, the buyer of a longevity bond will receive very small coupons, compared to the high longevity bond price he has payed[7]. At the same time, he will also have very small annuity payments. In total, he is still worse off and therefore he carries the risk of a possible regime change that causes longevity to strongly deteriorate.

- **A regime change or medical improvement that strongly improves longevity**
  In this case, the buyer of a longevity bond will receive very high coupon payment relative to the lower longevity bond price he has payed. At the same time, he will also have high annuity payments. In total, he profits from this situation and thus the risk lays in the longevity bond issuer.

To summarize, in the two possible mortality jump events the risk lays either in the longevity bond buyer or in he longevity bond issuer. Notice that in this respect, the higher offered price of the longevity bond, issued by the EIB/Paribas (than the one estimated by the historical data) has the explanation that the EIB/Paribas put a higher weight on the second event and this finds its application (as commented) in the so-called actuarial "expectation principle".

## 6.7 Suggestions for Further Research Topics

In this place, we mention that further research and a possibly more sophisticated approach might be needed in order to be able to incorporate a regime change in the underlying mortality model. However, such an approach will inevitably have to face the difficulty of insufficient historical data. It is also a very interesting idea to try to explain the trend of longevity (captured e.g. by the historical series of $A_1(t)$ and $A_2(t)$ in the Cairns *et al.* model) by regressing it on several macroeconomical quantities. This can then be used to predict the mentioned strong changes in the longevity trend, since a delay effect is to be expected.

[7]Recall that especially in the case of German data we have at our disposal only historical data of constantly improving mortality and thus this "optimism" is translated to the calibrated parameters and consequently leads to a high longevity bond price.

# Appendix A

# Proof of Theorem 2.2.1

**Proof:**

Adopting all notation from Chapter 1, we will consider here only the European call bond option and the proof for the put option will follow by analogy. For that reason, notice that the price of the European call bond option, can generally be written under a martingale measure $Q$ as

$$CB(t, T_1, T_2, K) = E^Q \left( e^{-\int_t^{T_1} r(s)ds} \left[ P(T_1, T_2) - K \right]^+ \middle| \mathfrak{F}_t \right).$$

In order to remove the stochastic discounting in the expectation, we can change to the equivalent $T_1$-forward risk-adjusted measure $Q^{T_1}$(associated with the bond maturing at time $T_1$), defined as

$$\frac{dQ^{T_1}}{dQ} = \frac{B(0)P(T_1, T_1)}{B(T_1)P(0, T_1)},$$

where $B(t)$ denotes the money-market account, driven by

$$dB(t) \quad = \quad B(t)r(t)dt, \quad B(0) = 1.$$

Thus, we obtain

$$CB(t, T_1, T_2, K) = P(t, T_1) E^{Q^{T_1}} \left( \left[ P(T_1, T_2) - K \right]^+ \middle| \mathfrak{F}_t \right).$$

Our next step will be to find the dynamic of the $P(T_1, T_2)$ bond, under the $Q^{T_1}$ measure. Notice that

$$\begin{aligned} \frac{dQ^{T_1}}{dQ} &= \frac{\exp\left( -\int_0^{T_1} r(u)du \right)}{P(0, T_1)} = \frac{\exp\left( -\int_0^{T_1} \left[ f(0, u) + X_1(u) + X_2(u) \right] du \right)}{P(0, T_1)} \\ &= \exp\left( -\int_0^{T_1} \left[ X_1(u) + X_2(u) \right] du \right). \end{aligned}$$

Now, by integration by parts we have

$$\int_0^{T_1} X_1(u)du = T_1X_1(T_1) - \int_0^{T_1} udX_1(u) = \int_0^{T_1}(T_1-u)dX_1(u) + T_1X_1(0).$$

As next, using that $X_1(0)=0$ and by the definition of the dynamic of the $X_1$ process, we obtain

$$\int_0^{T_1} X_1(u)du = \int_0^{T_1}(T_1-u)\left[-\kappa_1X_1(u) + \sum_{k=1}^{2}\varphi_{1k}(0,u)\right]du + \int_0^{T_1}(T_1-u)\sigma_1 dW_1^Q(u).$$

Adding the solution of the SDE of the $X_1$ process, which conditioned on $\mathfrak{F}_0$ can be written as

$$X_1(u) = \frac{\sigma_1^2}{2}\beta_1(0,u)^2 + \frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[\kappa_2 - (\kappa_1+\kappa_2)e^{-\kappa_1 u} + \kappa_1 e^{-(\kappa_1+\kappa_2)u}\right] + \int_0^u \sigma_1 e^{-\kappa_1(u-s)}dW_1^Q(s),$$

we have that

$$\begin{aligned}
& -\kappa_1\int_0^{T_1}(T_1-u)X_1(u)du \\
= \; & -\kappa_1\int_0^{T_1}(T_1-u)\left[\frac{\sigma_1^2}{2}\beta_1(0,u)^2 + \frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[\kappa_2-(\kappa_1+\kappa_2)e^{-\kappa_1 u}+\kappa_1 e^{-(\kappa_1+\kappa_2)u}\right]\right]du \\
& -\kappa_1\sigma_1\int_0^{T_1}(T_1-u)\int_0^u e^{-\kappa_1(u-s)}dW_1^Q(s)du.
\end{aligned}$$

Using again integration by parts yields

$$-\kappa_1\sigma_1\int_0^{T_1}(T_1-u)\int_0^u e^{-\kappa_1(u-s)}dW_1^Q(s)du = -\sigma_1\int_0^{T_1}\left((T_1-u) - \frac{1-e^{-\kappa_1(T_1-u)}}{\kappa_1}\right)dW_1^Q(u).$$

Summing up the upper calculations, we obtain

$$\begin{aligned}
\int_0^{T_1} X_1(u)du \;=\; & -\kappa_1\int_0^{T_1}(T_1-u)\left[\frac{\sigma_1^2}{2}\beta_1(0,u)^2\right]du \\
& -\kappa_1\int_0^{T_1}(T_1-u)\left[\frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[\kappa_2-(\kappa_1+\kappa_2)e^{-\kappa_1 u}+\kappa_1 e^{-(\kappa_1+\kappa_2)u}\right]\right]du \\
& -\sigma_1\int_0^{T_1}(T_1-u)dW_1^Q(u) + \frac{\sigma_1}{\kappa_1}\int_0^{T_1}\left(1-e^{-\kappa_1(T_1-u)}\right)dW_1^Q(u) \\
& +\int_0^{T_1}(T_1-u)\sum_{k=1}^{2}\varphi_{1k}(0,u)du + \sigma_1\int_0^{T_1}(T_1-u)dW_1^Q(u)
\end{aligned}$$

in which we notice that some of the terms cancel out. Next, by direct integration we estimate

$$\begin{aligned}
\int_0^{T_1} & (T_1-u)\left[-\frac{\kappa_1\sigma_1^2}{2}\beta_1(0,u)^2 + \varphi_{11}(0,u)\right]du \\
= & \int_0^{T_1}(T_1-u)\left[-\frac{\kappa_1\sigma_1^2}{2}\frac{\left(1-e^{-\kappa_1 u}\right)^2}{\kappa_1^2} + \frac{\sigma_1^2\left(1-e^{-2\kappa_1 u}\right)}{2\kappa_1}\right]du \\
= & \frac{\sigma_1^2}{2\kappa_1^2}\left[T_1 + \frac{2e^{-\kappa_1 T_1}}{\kappa_1} - \frac{3}{2\kappa_1} - \frac{e^{-2\kappa_1 T_1}}{2\kappa_1}\right]
\end{aligned}$$

in which we have used that

$$\int_0^{T_1}(T_1-u)e^{-\kappa_1 u}du = \frac{T_1 e^{-\kappa_1 T_1}}{\kappa_1}+\frac{e^{-\kappa_1 T_1}-1}{\kappa_1^2}.$$

and

$$\begin{aligned}&\int_0^{T_1}(T_1-u)\left[\frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[\kappa_2-(\kappa_1+\kappa_2)e^{-\kappa_1 u}+\kappa_1 e^{-(\kappa_1+\kappa_2)u}\right]\right]du\\ =\ &\frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[\left(\kappa_2^2+\kappa_1\kappa_2\right)T_1+(\kappa_1+\kappa_2)\frac{e^{-\kappa_1 T_1}-1}{\kappa_1}-\left(\kappa_2^2+\kappa_1\kappa_2\right)\frac{e^{-(\kappa_1+\kappa_2)T_1}-1}{(\kappa_1+\kappa_2)^2}\right].\end{aligned}$$

The same steps are applied for the integration of the $X_2$-process and therefore, we obtain finally

$$\frac{dQ^{T_1}}{dQ} = \exp\left(-\frac{1}{2}\Upsilon(T_1)-\sum_{i=1}^{2}\frac{\sigma_i}{\kappa_i}\int_0^{T_1}\left(1-e^{-\kappa_i(T_1-u)}\right)dW_i^Q(u)\right)$$

with

$$\begin{aligned}\Upsilon(T_1) &= \frac{\sigma_1^2}{\kappa_1^2}\left[T_1+\frac{2e^{-\kappa_1 T_1}}{\kappa_1}-\frac{3}{2\kappa_1}-\frac{e^{-2\kappa_1 T_1}}{2\kappa_1}\right]+\frac{\sigma_2^2}{\kappa_2^2}\left[T_1+\frac{2e^{-\kappa_2 T_1}}{\kappa_2}-\frac{3}{2\kappa_2}-\frac{e^{-2\kappa_2 T_1}}{2\kappa_2}\right]\\ &\quad+\frac{2\rho\sigma_1\sigma_2}{\kappa_1\kappa_2}\left[T_1+\frac{e^{-\kappa_1 T_1}-1}{\kappa_1}+\frac{e^{-\kappa_2 T_1}-1}{\kappa_2}-\frac{e^{-(\kappa_1+\kappa_2)T_1}-1}{\kappa_1+\kappa_2}\right].\end{aligned}$$

Further, we decompose the Brownian motions $W_1^Q(t)$ and $W_2^Q(t)$ into a sum of two independent Brownian motions $\tilde{W}_1^Q(t)$ and $\tilde{W}_2^Q(t)$ such that

$$\begin{aligned}dW_1^Q(t) &= d\tilde{W}_1^Q(t)\\ dW_2^Q(t) &= \rho d\tilde{W}_1^Q(t)+\sqrt{1-\rho^2}d\tilde{W}_2^Q(t)\end{aligned}$$

and then we can rewrite

$$\begin{aligned}\frac{dQ^{T_1}}{dQ} &= \exp\Bigg(-\frac{1}{2}\Upsilon(T_1)-\int_0^{T_1}\left(\frac{\sigma_1}{\kappa_1}\left(1-e^{-\kappa_1(T_1-u)}\right)+\rho\frac{\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2(T_1-u)}\right)\right)d\tilde{W}_1^Q(t)\\ &\quad-\frac{\sigma_2}{\kappa_2}\int_0^{T_1}\sqrt{1-\rho^2}\left(1-e^{-\kappa_2(T_1-u)}\right)d\tilde{W}_2^Q(t)\Bigg).\end{aligned}$$

Simple integration shows that

$$\begin{aligned}\Upsilon(T_1) &= \int_0^{T_1}\left(\frac{\sigma_1}{\kappa_1}\left(1-e^{-\kappa_1(T_1-u)}\right)+\rho\frac{\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2(T_1-u)}\right)\right)^2du\\ &\quad+\frac{\sigma_2^2}{\kappa_2^2}(1-\rho^2)\int_0^{T_1}\left(1-e^{-\kappa_2(T_1-u)}\right)^2du\end{aligned}$$

which, due to the Girsanov Theorem for the change of measure, implies (referring to Brigo [12]) that the Brownian motions $\tilde{W}_1^{Q^{T_1}}(t)$ and $\tilde{W}_2^{Q^{T_1}}(t)$, under the forward measure $Q^{T_1}$, are given as

$$\begin{aligned} d\tilde{W}_1^{Q^{T_1}}(t) &= d\tilde{W}_1^{Q}(t) + \left(\frac{\sigma_1}{\kappa_1}\left(1-e^{-\kappa_1(T_1-t)}\right) + \rho\frac{\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2(T_1-t)}\right)\right)dt \\ d\tilde{W}_2^{Q^{T_1}}(t) &= d\tilde{W}_2^{Q}(t) + \frac{\sigma_2}{\kappa_2}\sqrt{1-\rho^2}\left(1-e^{-\kappa_2(T_1-t)}\right)dt. \end{aligned}$$

Since $\tilde{W}_1^{Q}(t)$ and $\tilde{W}_2^{Q}(t)$ are independent Brownian motions, so are also $\tilde{W}_1^{Q^{T_1}}(t)$ and $\tilde{W}_2^{Q^{T_1}}(t)$. Then, the processes $X_1(t)$ and $X_2(t)$ are given under the forward $Q_{T_1}$ measure as

$$\begin{aligned} dX_1(t) &= \left[-\kappa_1 X_1(t) + \sum_{k=1}^{2}\varphi_{1k}(0,t) - \frac{\sigma_1^2}{\kappa_1}\left(1-e^{-\kappa_1(T_1-u)}\right) - \rho\frac{\sigma_1\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2(T_1-u)}\right)\right]dt \\ &\quad +\sigma_1 dW_1^{Q^{T_1}}(t), \quad X_1(0)=0 \\ dX_2(t) &= \left[-\kappa_2 X_2(t) + \sum_{k=1}^{2}\varphi_{2k}(0,t) - \frac{\sigma_2^2}{\kappa_2}\left(1-e^{-\kappa_2(T_1-u)}\right) - \rho\frac{\sigma_1\sigma_2}{\kappa_1}\left(1-e^{-\kappa_1(T_1-u)}\right)\right]dt \\ &\quad +\sigma_2 dW_2^{Q^{T_1}}(t), \quad X_2(0)=0 \end{aligned}$$

where we have defined

$$\begin{aligned} dW_1^{Q^{T_1}}(t) &= d\tilde{W}_1^{Q^{T_1}}(t) \\ dW_2^{Q^{T_1}}(t) &= \rho d\tilde{W}_1^{Q^{T_1}}(t) + \sqrt{1-\rho^2}d\tilde{W}_2^{Q^{T_1}}(t) \end{aligned}$$

in order to preserve the correlation between $X_1(t)$ and $X_2(t)$ under the forward $Q^{T_1}$ measure.

Having found the dynamic of the underlying processes under the forward $Q^{T_1}$ measure, we need as a next step the variance of the natural logarithm of the bond price $P(T_1,T_2)$, under that measure. Thus, due to (2.13) and since the change of measure changes only the drift of the process and not its variance, we can write

$$\begin{aligned} \mathrm{var}^{Q^{T_1}}\left(\ln\left(P(T_1,T_2)\right)|\mathfrak{F}_t\right) &= \beta_1(T_1,T_2)^2\mathrm{var}^{Q^{T_1}}\left(X_1(T_1)|\mathfrak{F}_t\right) + \beta_2(T_1,T_2)^2\mathrm{var}^{Q^{T_1}}\left(X_2(T_1)|\mathfrak{F}_t\right) \\ &\quad +2\beta_1(T_1,T_2)\beta_2(T_1,T_2)\mathrm{cov}^{Q^{T_1}}\left(X(T_1),X_2(T_2)|\mathfrak{F}_t\right) \\ &= \beta_1(T_1,T_2)^2\mathrm{var}^{Q}\left(X_1(T_1)|\mathfrak{F}_t\right) + \beta_2(T_1,T_2)^2\mathrm{var}^{Q}\left(X_2(T_1)|\mathfrak{F}_t\right) \\ &\quad +2\beta_1(T_1,T_2)\beta_2(T_1,T_2)\mathrm{cov}^{Q}\left(X(T_1),X_2(T_2)|\mathfrak{F}_t\right) \\ &= \beta_1(T_1,T_2)^2\varphi_{11}(t,T_1) + \beta_2(T_1,T_2)^2\varphi_{22}(t,T_1) \\ &\quad +2\beta_1(T_1,T_2)\beta_2(T_1,T_2)\varphi_{12}(t,T_1). \end{aligned}$$

where the subindex reflects under which measure the variance or covariance is being calculated. We denote

$$\Sigma_p^2(t,T_1,T_2) := \beta_1(T_1,T_2)^2\varphi_{11}(t,T_1) + \beta_2(T_1,T_2)^2\varphi_{22}(t,T_1) + 2\beta_1(T_1,T_2)\beta_2(T_1,T_2)\varphi_{12}(t,T_1).$$

Since any tradable asset discounted with a bond maturing at time $T_1$ is a martingale under the $Q^{T_1}$-forward measure, then denoting in addition the mean of the $\ln(P(T_1,T_2))$ under the $Q^{T_1}$-forward measure as $\mu_p(t,T_1,T_2)$ we have that

$$\begin{aligned}
\frac{P(t,T_2)}{P(t,T_1)} &= E^{Q^{T_1}}\left(P(T_1,T_2)|\mathfrak{F}_t\right) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\Sigma_p(t,T_1,T_2)^2} e^x e^{-\frac{(x-\mu_p(t,T_1,T_2))^2}{2\Sigma_p(t,T_1,T_2)^2}} dx \\
&= e^{\mu_p(t,T_1,T_2)+\frac{1}{2}\Sigma_p(t,T_1,T_2)^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\Sigma_p(t,T_1,T_2)^2} e^{-\frac{\left(x-\left(\Sigma_p(t,T_1,T_2)^2+\mu_p(t,T_1,T_2)\right)\right)^2}{2\Sigma_p(t,T_1,T_2)^2}} dx \\
&= e^{\mu_p(t,T_1,T_2)+\frac{1}{2}\Sigma_p(t,T_1,T_2)^2}
\end{aligned}$$

due to the normality of $\ln(P(T_1,T_2))$ as a sum of two normally distributed random variables. Thus, we do not have to additionally estimate the mean of $\ln(P(T_1,T_2))$, since we can present it as

$$\mu_p(t,T_1,T_2) = \ln\frac{P(t,T_2)}{P(t,T_1)} - \frac{1}{2}\Sigma_p(t,T_1,T_2)^2. \tag{A.1}$$

Finally, following the same steps as in the calculation of the Black-Scholes formula (see Appendix C), we obtain for the price of the European call bond option

$$\begin{aligned}
CB(t,T_1,T_2,K) &= P(t,T_1)\left[e^{\mu_p(t,T_1,T_2)+\frac{1}{2}\Sigma_p(t,T_1,T_2)^2}\Phi\left(\frac{\mu_p(t,T_1,T_2)-\ln K+\Sigma_p(t,T_1,T_2)^2}{\Sigma_p(t,T_1,T_2)}\right)\right. \\
&\quad \left. -K\Phi\left(\frac{\mu_p(t,T_1,T_2)-\ln K}{\Sigma_p(t,T_1,T_2)}\right)\right] \\
&= P(t,T_2)\Phi\left(\frac{\ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right)+\frac{1}{2}\Sigma_p(t,T_1,T_2)}{\Sigma_p(t,T_1,T_2)}\right) \\
&\quad -P(t,T_1)K\Phi\left(\frac{\ln\left(\frac{P(t,T_2)}{KP(t,T_1)}\right)-\frac{1}{2}\Sigma_p(t,T_1,T_2)}{\Sigma_p(t,T_1,T_2)}\right)
\end{aligned}$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution function. ■

# Appendix B

# Girsanov Theorems

**Theorem B.0.1 (Conditional Girsanov Theorem).** *Let $\mathcal{G}$ be a sub-$\sigma$-field of $\mathcal{F}$ on which two probability measures $Q$ and $P$ are given. If $Q$ is absolutely continuous with respect to $P$ with $dQ = \Lambda dP$ and $X$ is $Q$-integrable, then $\Lambda X$ is $P$-integrable and*

$$E^Q(X|\mathcal{G}) = \frac{E^P(X\Lambda|\mathcal{G})}{E^P(\Lambda|\mathcal{G})} \qquad Q-(a.s.)$$

**Proof:** *refer to Klebaner [40]* ■

**Theorem B.0.2 (Cameron-Martin-Girsanov Theorem).** *Let $B(t)$, $0 \leq t \leq T$, be a Brownian motion under a probability measure $Q$, and $\mu \neq 0$. Consider the process $W(t) = B(t) + \mu t$. There exists a measure $P$ equivalent to $Q$, such that $W(t)$ is $P$-Brownian motion. The Girsanov density is given by,*

$$\frac{dQ}{dP}(W) = \Lambda = e^{\mu W(T) - \frac{1}{2}\mu^2 T} \; and \frac{dP}{dQ}(B) = \frac{1}{\Lambda} = e^{-\mu B(T) - \frac{1}{2}\mu^2 T}$$

**Proof:** *refer to Klebaner [40]* ■

# Appendix C

# Black-Scholes Formula

Let $(\Omega, \mathfrak{F}, \mathbf{P})$ be a complete probability space endowed with a filtration $(\mathfrak{F}_t)_{t\geq 0}$. Further, let us assume that we have a **complete**, **arbitrage-free** market where the stock price $S(t)$ follows a lognormal walk under the unique (due to the complete market), equivalent (to the physical measure $\mathbf{P}$) martingale measure Q, such that

$$dS(t) = S(t)\left[r(t)dt + \sigma dW^Q(t)\right], \quad S(0) = S_0$$

where $W^Q(t)$ denotes a Brownian motion under the measure $Q$, the riskless interest rate $r(t)$ is a deterministic function of time and the volatility $\sigma > 0$ is a constant. Further, assuming **no transaction cost**, **no dividents** on the underlying and **continuous delta hedging**, we can derive (see Black and Scholes [7]) by replication that the price V(t) of a contingent (on the stock price) claim with payoff at maturity $T > t$ of $F(T, S(T))$ is a solution to the following **pricing linear parabolic PDE**:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rV\frac{\partial V}{\partial S} - rV = 0 \tag{C.1}$$

with the final condition

$$V(T, S(T)) = F(T, S(T)).$$

This can be written, due to the Feynman-Kac representation Theorem as

$$V(t, S(t)) = E^Q\left(e^{-\int_t^T r(s)ds} F(T, S(T)) \,\middle|\, \mathfrak{F}_t\right) \tag{C.2}$$

where $E^Q(\cdot)$ denotes an expectation under the equivalent martingale measure $Q$.

**Theorem C.0.3.** *For an European call option with payoff $F(T, S(T)) = (S(T) - K)^+$ at maturity $T$ and a strike $K > 0$, we have that*

$$V(t, S(t)) = S(t)\Phi(d_1) - Ke^{-\int_t^T r(s)ds}\Phi(d_2)$$

*where*

$$\begin{aligned} d_1 &= \frac{log\left(\frac{S(t)}{K}\right) + \left(\int_t^T r(s)ds + \frac{1}{2}\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned}$$

*and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.*

*The price of an European put option with payoff at maturity $F(T, S(T)) = (K - S(T))^+$ for a strike $K > 0$ is given by*

$$V(t, S(t)) = -S(t)\Phi(-d_1) + Ke^{-\int_t^T r(s)ds}\Phi(-d_2)$$

*for the same notations of $d_1$ and $d_2$.*

**Proof:** We will give the proof of this well-known theorem for the only reason that we refer to it in another place of this thesis. We show it only for an European call option.

From the equivalence of the pricing PDE (C.1) to the expectation in (C.2) follows that we are free to choose to solve any of them. Choosing to calculate (C.2) yields

$$\begin{aligned} V(t, S(t)) &= E^Q\left(e^{-\int_t^T r(s)ds}F(T, S(T))\,\middle|\,\mathfrak{F}_t\right) = E^Q\left(e^{-\int_t^T r(s)ds}(S(T) - K)^+\,\middle|\,\mathfrak{F}_t\right) \\ &= e^{-\int_t^T r(s)ds}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}(T-t)}e^{-\frac{x^2}{2(T-t)}}\left(S(t)e^{\int_t^T r(s)ds - \frac{1}{2}\sigma^2(T-t) + \sigma x} - K\right)^+ dx \\ &= e^{-\int_t^T r(s)ds}\int_{K^*}^{\infty}\frac{1}{\sqrt{2\pi}(T-t)}S(t)e^{\int_t^T r(s)ds - \frac{1}{2}\sigma^2(T-t) + \sigma x - \frac{x^2}{2(T-t)}}dx \\ &\quad -Ke^{-\int_t^T r(s)ds}\int_{K^*}^{\infty}\frac{1}{\sqrt{2\pi}(T-t)}e^{-\frac{x^2}{2(T-t)}}dx \\ &= S(t)\int_{K^*}^{\infty}\frac{1}{\sqrt{2\pi}(T-t)}e^{-\frac{(x-(T-t)\sigma)^2}{2(T-t)}}dx - Ke^{-\int_t^T r(s)ds}\int_{K^*}^{\infty}\frac{1}{\sqrt{2\pi}(T-t)}e^{-\frac{x^2}{2(T-t)}}dx \\ &= S(t)\Phi\left(\frac{\sigma(T-t) - K^*}{\sqrt{T-t}}\right) - Ke^{-\int_t^T r(s)ds}\Phi\left(\frac{-K^*}{\sqrt{T-t}}\right) \end{aligned}$$

with $K^* = \frac{1}{\sigma}\left[\log\left(\frac{K}{S} + \frac{1}{2}\sigma^2(T-t) - \int_t^T r(s)ds\right)\right]$. ■

# Bibliography


[1] J. Ait-Sahalia. Maximum likelihood estimation of discretely sampled diffusion: A closed-form approximation approach. *Econometrica*, 70:223–262, 2002.

[2] L. Andersen. A simple approach to the pricing of bermudan swaptions in the multi-factor libor market model. March 1999.

[3] M. Avellaneda and L. Wu. Pricing parisian-style options with a lattice method. *International Journal of Theoretical and Applied Finance*, 2:1–16, 1999.

[4] J. Barraquand and T. Pudet. Pricing of american path-dependent contingent claims. *Mathematical Finance*, 6:17–51, 1996.

[5] B. Benjamin and J. Pollard. The analysis of mortality and other actuarial statistics. *The Institute of Actuaries, Oxford.*, 1993.

[6] A. Bensoussan. On the theory of option pricing. *Acta. Appl. Math.*, 2:139–158, 1984.

[7] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[8] T. Börk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 1998.

[9] A. Borodin and P. Salminen. *Handbook of Brownian Motion-Facts and Formulae.* Birkheuser Verlag., 2002.

[10] B. Bouchard, I. Ekeland, and N. Touzi. On malliavin approach to monte carlo approximation of conditional expectations. Finance and Stochastics, 2002.

[11] B. Bouchard and N. Touzi. Discrete time approximation and monte-carlo simulation of backward stochastic differential equations. 2003.

[12] D. Brigo and F. Mercurio. *Interest Rate Models. Theory and Practice.* Springer-Verlag Berlin Heidelberg, 2001.

[13] M. Broadie and J. Detemple. American option valuation: New bounds, approximation and a comparison of existing methods. *Review of Financial Studies*, 9:1211–1250, 1996.

[14] P. Brockwell and R. Davis. *Time Series: Theory and Methods.* Springer Series in Statistics, 1990.

[15] A. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty. Heriot-Watt-University Edinburgh, 2005.

[16] J. Carrière. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, 19:19–30, 1996.

[17] A. Carverhill. When is the short rate markovian? *Mathematical Finance*, 2:305–312, 1994.

[18] K. Chan, G. Karolyi, F. Longstaff, and A. Sanders. An empirical comparison of alternative models of the term structure of interest rates. *Journal of Finance*, 47:1209–1228, 1992.

[19] M. Chesney, M. Jeanblanc-Picque, and M. Yor. Brownian excursions and parisian barrier options. *Adv.Appl.Proba.*, March 1997.

[20] O. Cheyette. Term structure dynamics and mortgage valuation. *The Journal of Fixed Income*, March 1992.

[21] O. Cheyette. Markov representation of the heath-jarrow-morton model. 1994.

[22] C. Chu and Y. Kwock. Target redemption notes. 2006.

[23] E. Clément, D. Lambertin, and P. Protter. An analysis of a least-squares regression method for american option pricing. *Finance and Stochastics*, 6:449–472, 2002.

[24] A. Cowley and J. Cummins. Securitization of life insurance assets and liabilities. *Journal of Risk and Insurance*, 72:193:226, 2005.

[25] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[26] J. Cox, S. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[27] I. Craig and A. Sneyd. An alternating-direction implicit scheme for parabolic equations with mixed derivatives. *Comput.Math.Appl.*, 16(4):341–350, 1988.

[28] P. Forsyth, K. Vetzal, and R. Zvan. Convergence of numerical methods for valuing path-dependent options using interpolation. 1998.

[29] P. Glynn and W. White. The asymptotic efficiency of simulation estimators. *Operations Research*, 40:505–520.

[30] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology. *Econometrica*, 60:77–105, 1992.

[31] S. Heston and G. Zhou. Mathematical finance. 10:53–75, 2000.

[32] J. Hugonnier. The feynman-kac formula and pricing occupational time derivatives. *International Journal of Theoretical and Applied Finance*, 2:153–178, 1999.

[33] J. Hull. *Options, Futures & Other Derivatives*. Fourth edition, 1999.

[34] J. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3:4, 1990.

[35] J. Hull and A. White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis*, 25:87–100, 1990.

[36] J. Hull and A. White. Efficient procedure for valuing european and american path-dependent options. *The Journal of Derivatives*, 1:21–31, 1993.

[37] J. Hull and A. White. Numerical procedure for implementing term structure models ii: Two-factor models. *Journal of Derivatives*, 2:37–48, 1994.

[38] F. Jamshidian and Y. Zhu. Scenario simulation: Theory and methodology. *Finance and Stochastics*, 1:43–67.

[39] I. Karatzas. On the pricing of american options. *Appl. Math. Optim.*, 17:37–60, 1998.

[40] F. C. Klebaner. *Introduction to Stochastic Calculus with Applications.* Imperial College Press, 2001.

[41] R. Korn and E. Korn. *Option Pricing and Portfolio Optimization.* American Mathematical Society, 2001.

[42] R. Korn, K. Natcheva, and J. Zipperer. Longevity bonds: Pricing, modelling and application for german data. *Deutsche Gesellschaft für Versicherung-und Finanzmathematik e.V.*, Band XXVII, Heft 3, 2006.

[43] M. Krekel and K. Natcheva. About the numerical solution of the cheyette model. ITWM, Kaiserslautern, 2005.

[44] Y. Kwok and K. Lau. Pricing algorithms for options with exotic path-dependence. *The Journal of Derivatives*, Fall:28–38, 2001.

[45] R. Lee and L. Carter. Modeling and forecasting u.s. mortality. *Journal of the American Statistical Association*, 87:659–675, 1992.

[46] A. Li, P. Ritchken, and L. Sankarasubramanian. Lattice models for pricing american interest rate claims. *The Journal of Finance*, 50:719–737, June 1995.

[47] F. Longstaff and E. Schwartz. Pricing american options by simulation: A simple least-squares approach. forthcoming in The Review of Financial Studies, 2000.

[48] H. McKean. Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics. *Indust.Manage.Rev.*, 6:32–39, 1965.

[49] M. Milevsky and S. Promislow. Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29:299–318, 2001.

[50] F. Moraux. On cumulative parisian options. to apper in Special Issue on Mathematical Finance, 2002.

[51] A. Olivieri. Uncertainty in mortality projections: an actuarial perspective. *Insurance. Mathematics and Economics*, 29(2):231–245, 2001.

[52] A. Olivieri and E. Pitacco. Solvency requirements for pension annuities. *Journal of Pension Economics and Finance.*, 2003.

[53] W. Perks. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63:12–57, 1932.

[54] E. Pitacco. Survival models in actuarial mathematics: From halley to longevity risk. University of Trieste, 2005.

[55] R. Rebonato. *Interest Rate Option Models.* Wiley, Chichester, 1998.

[56] R. Reider. An efficient monte carlo technique for pricing options. Wharton School, University of Pennsylvania, 1993.

[57] P. Ritchken, L. Snakarasubramanian, and A. Vijh. The valuation of path dependent contracts on the average. *Management Science*, 39:1202–1213, 1993.

[58] R. Ritchken and L. Sankarasubramanian. The importance of forward-rate volatility structure in pricing interest rate-sensitive claims. *Journal of Derivatives*, Fall:25–41, 1995.

[59] L. Rogers. Monte carlo valuation of american options. 2001.

[60] L. Rogers and Z. Shi. The value of an asian option. *Journal of Applied Probability*, 32:1077–1088, 1995.

[61] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[62] K. Vetzal and P. Forsyth. Discrete parisian and delayed barrier options: A general numerical approach. *Advances in Futures and Options Research*, 10:1–15, 1999.

[63] P. Wilmott. *Derivatives. The Theory and Practice of Financial Engineering.* John Wiley & Sons, 1998.

[64] J. Yue. Oldest-old mortality rates and the gompertz law: A theoretical and empirical study based on four countries. National Chengchi University, 1991.

[65] J. Zhang. A numerical scheme for backward stochastic differential equations: Approximation by step processes. 2001.

[66] Z. Zhu and N. Stokes. A finite element platform for pricing path-dependent exotic options. CSIRO Methematical and Information Sciences working paper, 1999.

[67] F. P. Zvan, R. and K. Vetzal. Negative coefficients in two factor option pricing models. 2001.

# Scientific Career

| | |
|---|---|
| 1982 – 1989 | Primary School "St. Cyril and Methodi", Bourgas |
| 1989 – 1995 | Mathematical High School "Acad.Nikola Obreshkov", Bourgas |
| 1990 – 1995 | English Language School "Geo Milev", Bourgas (as a private student) |
| 1995 – 2000 | University of Sofia "St.Kliment Ohridski", Faculty of Economics and Business Administration. Graduation: Bachelor in Business Administration |
| 2000 – 2002 | Technical University of Kaiserslautern, Department of Mathematics, Graduation: Master of Science in Financial Mathematics |
| since October 2002 | Scientific researcher at the Fraunhofer ITWM, Kaiserslautern, Department of Financial Mathematics |