Boundary Conditions For Mean-Reverting Square Root Process

by

Jonathan Aquan-Assee

A research paper
presented to the University of Waterloo
in fulfillment of the
requirement for the degree of
Master of Mathematics
in
Computational Mathematics

supervisors: Peter Forsyth and George Labahn

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this paper. This is a true copy of the paper, including any required final revisions, as accepted by my examiners.

I understand that my paper may be made electronically available to the public.

Abstract

The Cox-Ingersoll-Ross (CIR) interest rate model is used to model interest rates and interest rate derivatives. We study the term structure equation for single-factor models that predict non-negative interest rates. It is shown using finite difference techniques that if the boundary is attainable, then this boundary behaviour serves as a boundary condition and guarantees a uniqueness of solutions. However, if the boundary is non-attainable, then the boundary condition is not needed to guarantee uniqueness. The finite difference solution is verified by use of non-negative numerical approximations.

Acknowledgements

I would like to thank Peter Forsyth and George Labahn for helping make this paper possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The mean-reverting square-root process is a stochastic differential equation (SDE) that has found considerable use as a model for volatility, interest rates, and other financial quantities. The equation has no general, explicit solution, although its transition density can be characterized. Some standard expectations can be analytically calculated which can be useful for calibrating the parameters. These processes were initially introduced to model the short interest rate by Cox, Ingersoll, and Ross [1], and are now widely used in modeling because they present interesting features like non-negativity and mean reversion. The equation for this process is

$$dR(t) = \kappa(b - R(t))dt + \sigma\sqrt{R(t)}dW(t), \tag{1.1}$$

$R(t)$ will denote a Cox-Ingersoll-Ross (CIR) process. The parameters κ, b, and σ are all non-negative and $dW(t)$ denotes the increments of a Brownian motion. Notice that the drift term in equation (1.1) is positive if $R(t) < b$ and negative if $R(t) > b$. Thus $R(t)$ is pulled toward the level b, a property generally referred to as mean reversion. We may interpret b as the long-run interest rate level and κ as the speed at which $R(t)$ is pulled toward b. In this paper, we assume that κ, b, and σ do not vary with time. Under the above assumption on the parameters, that we will suppose valid through all the paper, it is well known that this SDE has a non-negative solution, and this solution is path-wise unique (see [2] and [3]). Also, under the conditions

$$2\kappa b > \sigma^2 \text{ and } R(0) > 0 \tag{1.2}$$

the process is always positive.

It is well known that the increments of the CIR process are non-central chi-squared random variables that can be simulated exactly. However, the exact simulation in general

requires more computational cost than a simulation with approximation schemes. It may also be restrictive if one wishes to study a 'generalized' version of the mean reverting process. For these reasons, studying approximation schemes are relevant.

The CIR process is frequently used for pricing of different interest rate derivatives such as bonds and bond options. For valuing interest rate derivatives, stochastic methods seem to be more commonly used than finite difference (FD) methods. When using FD methods it is necessary to use the hyperbolic partial differential equation (PDE) called the *term structure equation*, for valuing bond prices. The authors of [4] believe that stochastic methods are preferred over FD methods because the pricing equation with appropriate boundary conditions is not fully developed from a mathematical sense. Moreover, in [5], the authors claim that there are multiple solutions to the term structure equation that satisfy the same boundary conditions.

In this paper we compare stochastic and FD methods for pricing of bonds using the CIR process. We will show using simulation that the unique solution to the term structure equation is the one given using FD methods subject to the correct boundary conditions. We verify these results by performing a Monte Carlo (MC) simulation using a non-negative approximation of the CIR equation. We are interested in the solution of the term structure equation for vanishing interest rates (i.e as $r \to 0$) and the behavior of the solution at the boundary using different boundary conditions. Our hypothesis is that if the boundary is attainable, then this boundary behavior serves as a boundary condition and guarantees uniqueness of solutions. However, if the boundary is non-attainable, then the boundary condition is not needed to guarantee uniqueness, but it is nevertheless very useful from a numerical perspective.

The project is structured as follows: Chapter 2 develops the methods for simulating the stochastic interest rate model given by the CIR process. A discussion of boundary conditions for the model is presented. Chapter 3 details the numerical schemes for the CIR model. Discretization methods for stochastic integration schemes are discussed. A Crank-Nicolson method is also developed for the CIR model. Chapter 4 presents simulation results of the integration schemes developed previously. The theoretical results are verified using the numerical tests. Finally, Chapter 5 concludes by pointing out the main results of the paper.

Chapter 2

Problem Formulation

This chapter develops the methods for simulating the stochastic interest rate model given by the CIR process in equation (1.1). The CIR process models the dynamics of an instantaneous continuously compounded interest rate $R(t)$. As shown in [6], the interest rate for the CIR model is given by

$$R(t) = R(0) + \int_0^t \kappa(b - R(s))ds + \int_0^t \sigma\sqrt{R(s)}W(s)ds$$

which is just an integral form of the differential equation in (1.1). An investment in a money market account earning interest at a rate $R(t)$ at time t grows from a value of one at time zero to a value of

$$D(t) = e^{\int_0^t R(u)du}$$

at time t. The price at time zero of a derivative security that pays X at time T is the expectation of $X/D(T)$, or

$$E\left[e^{-\int_0^t R(u)du}X\right].$$

In particular, we define:

Definition 1 (Zero coupon bond) *A financial contract promising to pay a certain "face" amount, which we take to be one, at a fixed maturity date T is called a zero coupon bond. Prior to the expiration date the bond makes no payments.*

Given the stochastic process to the interest rate $R(t)$ as shown in equation (1.1), the value of the zero coupon bond $V(r,t)$ is given by

$$V(r,t) = E_Q\left[e^{-\int_t^T R(s)ds}|R(t) = r\right] \tag{2.1}$$

with E_Q denoting the expectation operator under the risk-neutral probability measure Q and r is the interest rate for the chosen dynamics $R(t)$ conditional on information available at time t. This pricing formula is a martingale under the risk-neutral measure.

As shown in [6], by applying Itô's lemma to the expectation in equation (2.1), it is possible to derive the *term structure* PDE satisfied by $V(r,t)$ as given by

$$V_t(r,t) + \kappa(b-r)V_r(r,t) + \frac{1}{2}\sigma^2 r V_{rr}(r,t) = rV(r,t). \tag{2.2}$$

2.1 Analytical positivity and boundedness

The SDE in equation (1.1) is defined on the domain $[0,\infty)$. An important concept in the study of SDEs is the concept of positivity or non-negativity. We must therefore define the following concept

Definition 2 *Assuming the process X_t is defined on the domain $[L,\infty)$ and $X_0 > L$, then*

- *the process X_t has a non-attainable boundary L, if*

$$P(X_t > L \text{ for all } t > 0) = 1.$$

- *The process X_t has an attainable boundary L, if*

$$P(X_t \geq L \text{ for all } t > 0) = 1 \quad \text{and} \quad P(\exists t^* > 0 : X_{t^*} = L) = 1.$$

For $L = 0$, attainability is equivalent to positivity and non-negativity.

Throughout the rest of the paper we will make use of Definition 2.

2.2 Boundary Conditions

The boundary behavior for the SDE shown in equation (1.1) has been studied at great detail by Feller in [7]. For the process in (1.1), Feller showed that the boundary at the origin is instantaneously reflecting, regular, and attainable if $\sigma^2 > 2\kappa b$, and unattainable, non-attracting, boundary otherwise. As can be seen in Appendix A, interesting properties related to the boundary behavior for the SDE (1.1) can be uncovered by looking at the probability distribution for the process.

The results of Feller can be used to derive appropriate boundary conditions for finite differencing schemes of the term structure equation given in (2.2). For example, it is possible that the case $\sigma^2 < 2\kappa b$ can be handled by simple Dirichlet boundary conditions at $r = 0$, while the same cannot be done if $\sigma^2 > 2\kappa b$. This can be generalized with the following definition.

Definition 3 (Feller Condition) *Given a process X defined by a SDE of the form $dX(t) = \beta(X(t), t)dt + c(X(t), t)dW$. If the condition*

$$\lim_{x \to 0} \left(\beta(x) - \frac{1}{2} \frac{\partial c^2}{\partial x}(x) \right) \geq 0 \tag{2.3}$$

is true, then the boundary at the origin is non-attainable for the process X.

According to [4], the appropriate boundary condition for the term structure equation in (2.2) can be obtained by plugging in $r = 0$ into the equation. The resulting condition is then

$$V_t(0, t) + \kappa b V_r(0, t) = 0. \tag{2.4}$$

This boundary condition will result in a unique solution to the term structure equation when $\sigma^2 > 2\kappa b$. However, when $\sigma^2 < 2\kappa b$, the boundary condition (2.4) is not needed to guarantee a unique solution and is thus redundant from a mathematical perspective [4].

At the upper boundary we must truncate the domain to $[0, r_{max}]$. Consequently, it is necessary to impose a boundary condition at $r = r_{max}$. The theoretical upper boundary condition $V = 0$ as $r \to \infty$ is of critical use if we retain r as the coordinate. In [8] the authors propose a boundary technique that does not require additional linearity assumptions. This FD method involves using one-sided derivatives for the spatial discretization of the PDE.

In this paper we take the approach of extending the computational domain in order to minimize any boundary errors at r_{max}. In this way the boundary condition at r_{max} can be handled with a simple Dirichlet boundary condition with

$$V(r_{max}, t) = 0. \tag{2.5}$$

Chapter 3

Numerical Methods

It is well known that an analytical formula for the term structure equation (2.2) exists (see Appendix B). However, recall from the introduction that it has been claimed by previous studies, that multiple solutions to the term structure equation exist that satisfy the same boundary conditions. All numerical methods seek to find a solution for the pricing relationship given in equation (2.1). In this chapter we will examine the numerical methods that will be used to price zero coupon bonds in the CIR model. Using these numerical methods we will be able to examine the effects of different boundary conditions for the term structure equation.

3.1 Finite Difference Discretization Methods

In this section we will derive the discretized equations using an approach to approximate the PDE (2.2) using finite differencing techniques. The pricing equation as shown in [9], is given by the following equation

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sigma^2 r \frac{\partial^2 V}{\partial r^2} + \kappa(b-r)\frac{\partial V}{\partial r} - rV, \qquad 0 < r < r_{max}, \qquad t > 0, \tag{3.1}$$

where $V(r,\tau)$ is the bond price, r is the spot interest rate, $\tau = T - t$, T is the expiry time of the bond, t is the time (in the forward direction).

This is a first-order hyperbolic equation and it must be augmented with an initial condition and boundary conditions in order to define a valid initial boundary value problem. In this case we define them as

$$V(r_{max}, t) = 0 \qquad \text{and} \qquad V(r,T) = 1. \tag{3.2}$$

If we take the limit as $r \to 0$, we get the following boundary condition at $r = 0$

$$\frac{\partial V}{\partial \tau} - \kappa b \frac{\partial V}{\partial r} = 0, \quad \text{if} \quad \sigma^2 > 2\kappa b. \tag{3.3}$$

We can then implement a number of finite differencing discretization schemes. In particular, we will use a forward-backward differencing scheme. Define a grid of points in the (r, τ) plane

$$\begin{aligned} r_0, r_1, ..., r_m \qquad & 0 \leq r_i \leq r_{max} \\ \tau^n = n\Delta\tau \qquad & 0 \leq \tau^n \leq N\Delta\tau = T \end{aligned}$$

and let

$$V(r_i, \tau^n) = V_i^n.$$

Let

$$\Delta r_{i+1/2} = r_{i+1} - r_i \qquad \text{and} \qquad \Delta r_{i-1/2} = r_i - r_{i-1}.$$

Equation (3.1) can be then approximated by

$$\left(\frac{\partial V}{\partial \tau}\right)_i^n = \left(\frac{\sigma^2 r^2}{2} V_{rr}\right)_i^n + (\kappa(b-r)V_r)_i^n - (rV)_i^n, \tag{3.4}$$

where the time derivative in equation (3.4) is approximated by

$$\left(\frac{\partial V}{\partial \tau}\right)_i^n \simeq \frac{V_i^{n+1} - V_i^n}{\Delta\tau}. \tag{3.5}$$

The second order derivative term is approximated by

$$(V_{rr})_i^n \simeq \left(\frac{\left(\frac{V_{i+1}^n - V_i^n}{\Delta r_{i+1/2}}\right) - \left(\frac{V_i^n - V_{i-1}^n}{\Delta r_{i-1/2}}\right)}{\frac{\Delta r_{i+1/2} + \Delta r_{i-1/2}}{2}} \right) \tag{3.6}$$

with the discounting term in equation (3.4) as

$$(V)_i^n \simeq V_i^n \tag{3.7}$$

while the derivative term can be approximated by either central, forward, or backward differencing.

A central difference is

$$(V_r)_i^n \simeq \left(\frac{V_{i+1}^n - V_{i-1}^n}{\Delta r_{i+1/2} + \Delta r_{i-1/2}} \right), \tag{3.8}$$

a forward difference is

$$(V_r)_i^n \simeq \left(\frac{V_{i+1}^n - V_i^n}{\Delta r_{i+1/2}} \right), \tag{3.9}$$

and a backward difference is

$$(V_r)_i^n \simeq \left(\frac{V_i^n - V_{i-1}^n}{\Delta r_{i-1/2}} \right). \tag{3.10}$$

It is well known that central differencing is preferred because it is second order, while forward and backwards differencing are first order. Unfortunately central differencing can result in problems if used in all cases.

Using any of (3.8), (3.9), or (3.10) gives

$$V_i^{n+1} = V_i^n(1 - (\alpha_i + \beta_i + r_i)\Delta\tau) + V_{i-1}^n \Delta\tau\alpha_i + V_{i+1}^n \Delta\tau\beta_i. \tag{3.11}$$

Then α_i and β_i are defined as

$$\alpha_i^{central} = \left[\frac{\sigma^2 r_i^2}{(r_i - r_{i-1})(r_{i+1} - r_{i-1})} - \frac{\kappa(b - r_i)}{r_{i+1} - r_{i-1})} \right]$$
$$\beta_i^{central} = \left[\frac{\sigma^2 r_i^2}{(r_{i+1} - r_i)(r_{i+1} - r_{i-1})} + \frac{\kappa(b - r_i)}{r_{i+1} - r_{i-1})} \right]$$

in the case of central differencing, or

$$\alpha_i^{forward} = \left[\frac{\sigma^2 r_i^2}{(r_i - r_{i-1})(r_{i+1} - r_{i-1})} \right]$$
$$\beta_i^{forward} = \left[\frac{\sigma^2 r_i^2}{(r_{i+1} - r_i)(r_{i+1} - r_{i-1})} + \frac{\kappa(b - r_i)}{r_{i+1} - r_i)} \right]$$

in the case of forward differencing, or

$$\alpha_i^{backward} = \left[\frac{\sigma^2 r_i^2}{(r_i - r_{i-1})(r_{i+1} - r_{i-1})} - \frac{\kappa(b - r_i)}{r_i - r_{i-1})} \right]$$
$$\beta_i^{backward} = \left[\frac{\sigma^2 r_i^2}{(r_{i+1} - r_i)(r_{i+1} - r_{i-1})} \right]$$

in the case of backwards differencing.

For stability reasons it is vital that all α_i and β_i values be always positive. We can then decide between the central or upstream (i.e. forward or backward) discretization by:

For $i = 0, ..., n-1$
  If $\alpha_i^{central} \geq 0$ and $\beta_i^{central} \geq 0$
    $\alpha_i = \alpha_i^{central}$
    $\beta_i = \beta_i^{central}$
  ElseIf $\beta_i^{forward} \geq 0$
    $\alpha_i = \alpha_i^{forward}$
    $\beta_i = \beta_i^{forward}$
  Else
    $\alpha_i = \alpha_i^{backward}$
    $\beta_i = \beta_i^{backward}$
  EndIf
EndFor

The above algorithm guarantees that α_i and β_i are non-negative. As shown by [10], requiring that all α_i and β_i be non-negative has important ramifications for the stability of scheme (3.11).

Remark 1 (Unconditional stability) *As shown by [10], the discretization method (3.11) is unconditionally stable provided that α_i and β_i are non-negative.*

Remark 2 (Convergence to the viscosity solution) *As shown by [10], requiring that all α_i and β_i are non-negative has important ramifications for the convergence of scheme (3.11). It is possible to show that the discretization (3.11) is monotone and consistent. Since it is also unconditionally stable, then the discretized solution converges to the viscosity solution.*

3.1.1 Crank-Nicolson Discretization

In this section we consider a Crank-Nicolson time-stepping strategy for the FD discretization in equation (3.11).

Ordinary explicit FD schemes are known to become unstable if certain time-step conditions are not satisfied. For these reasons we will study the Crank-Nicolson method which is known to be unconditionally stable. In [4] the authors proposed a backward differencing formula of order two (BDF2), however we found that this method produced oscillations in the solution when the boundary condition (2.4) was not explicitly defined.

The FD equation (3.11) can be written into the Crank-Nicolson time-stepping form as follows

$$V_i^{n+1}\left[1+(\alpha_i+\beta_i+r_i)\frac{\Delta\tau}{2}\right]-\frac{\Delta\tau}{2}\beta_i V_{i+1}^{n+1}-\frac{\Delta\tau}{2}\alpha_i V_{i-1}^{n+1} \\ =V_i^n\left[1-(\alpha_i+\beta_i+r_i)\frac{\Delta\tau}{2}\right]+\frac{\Delta\tau}{2}\beta_i V_{i+1}^n+\frac{\Delta\tau}{2}\alpha_i V_{i-1}^n \tag{3.12}$$

which is an implicit equation rather than explicit. Let $\hat{M}$ be a tridiagonal matrix with entries

$$[\hat{M}V^n]_i=-\frac{\Delta\tau\alpha_i}{2}V_{i-1}^n+\frac{\Delta\tau(\alpha_i+\beta_i+r_i)}{2}V_i^n-\frac{\Delta\tau\beta_i}{2}V_{i+1}^n \tag{3.13}$$

that is, $\hat{M}$ is defined by

$$\hat{M}=\frac{\Delta\tau}{2}\begin{bmatrix} \gamma_1 & \gamma_2 & 0 & & & \\ -\alpha_2 & r_2+\alpha_2+\beta_2 & -\beta_2 & & & \\ & & & \ddots & & \\ & & & -\alpha_{m-1} & r_{m-1}+\alpha_{m-1}+\beta_{m-1} & -\beta_{m-1} \\ & & & 0 & 0 & 0 \end{bmatrix} \tag{3.14}$$

where m is the number of grid points, γ_1 and γ_2 are the boundary points at $r = 0$. Note that the values γ_1 and γ_2 depend on which boundary conditions we impose at $r = 0$. In Section 3.1.2, we will show the discretization for γ_1 and γ_2 using two different boundary conditions. The last row of zeros in matrix $\hat{M}$ is due to the Dirichlet boundary condition being imposed at $r = r_{max}$, which is further discussed in section 3.1.2.

Then we can write equation (3.12) as

$$[I + \hat{M}]V^{n+1} = [I - \hat{M}]V^n \tag{3.15}$$

where V^n is the value of the solution at the n-th time step.

Theorem 1 (Necessary conditions for stability) *Let the Crank-Nicolson time-stepping be defined as in equation (3.15), where $\hat{M}$ is of size $(m+1) \times (m+1)$, where there are $m+1$ nodes in the discretization, and $\hat{M}$ has the properties*

1. *The off-diagonal entries of $\hat{M}$ are all strictly negative,*
2. *The diagonal entries of $\hat{M}$ are non-negative,*
3. *$\hat{M}$is row diagonally dominant.*

Then Crank-Nicolson time-stepping satisfies the necessary conditions for unconditional stability.

Proof 1 *The proof follows similarly as in [11]. Assume that $\hat{M}$ has $m+1$ linearly independent eigenvectors. If X_k is the k^{th} eigenvector of $\hat{M}$, then*

$$\hat{M}X_k = \lambda_k X_k$$

where λ_k is the eigenvalue associated with X_k. Then, rewrite equation (3.15 as

$$V^{n+1} = BV^n, \qquad B = [I + \hat{M}]^{-1}[I - \hat{M}]. \tag{3.16}$$

If V^0 is the initial condition, then after n time-steps we have

$$V^n = B^n V^0. \tag{3.17}$$

Assuming $\hat{M}$ has a complete set of eigenvectors, then

$$V^0 = \sum_k C_k X_k \tag{3.18}$$

for some coefficients C_k*. Substituting equation (3.18) into equation (3.17) then gives*

$$V^n = \sum_k C_k X_k \left[\frac{1-\lambda_k}{1+\lambda_k}\right]^n X_k^0. \tag{3.19}$$

The requirement of equation (3.19) to remain bound as $n \to \infty$*, is*

$$\left[\frac{1-\lambda_k}{1+\lambda_k}\right] \leq 1 \quad \forall k, \tag{3.20}$$

which will be true as long as

$$Re(\lambda_k) \geq 0 \quad \forall k. \tag{3.21}$$

From the Gershgorin circle theorem, we know that every eigenvalue λ *of* $\hat{M}$ *satisfies at least one of the inequalities*

$$\begin{aligned} |\hat{M}_{ii} - \lambda| &\leq \sum_{j \neq i} |\hat{M}_{ij}| \\ &\leq |\hat{M}_{ii}| \end{aligned}$$

since $\hat{M}$ *is diagonally dominant. Therefore, in the complex plane* $\lambda = x + \sqrt{-1}y$*, every* λ_k *lies in some disk with center* $x = |\hat{M}_{ii}|$ *and radius strictly less then* $|\hat{M}_{ii}|$*, and therefore*

$$Re(\lambda_k) \geq 0 \quad \forall k.$$

□

Theorem 1 states that the CN method satisfies the necessary conditions for stability, however this does not prove that it satisfies the sufficient conditions for stability. The precise definition of stability is

$$||V^n|| = bounded. \tag{3.22}$$

However, it takes a bit of work to prove (3.22) for the CN method, so the idea of algebraic stability has been introduced [10].

Definition 4 (Stability of Crank-Nicolson) *If n is the number of time-steps, m is the number of grid nodes, and B is defined as in equation (3.16), then given an arbitrary norm $\|\cdot\|$, we say that B is algebraic stable if*

$$\|B^n\| \leq Cn^\alpha m^\beta, \forall n, m. \tag{3.23}$$

where C, α, β are independent of n, m. This is often referred to as bounding the power of a matrix. We remark that algebraic stability is a weaker condition than strong stability or strict stability.

It can be shown that for the CN method, algebraic stability is guaranteed and is summarized in the following theorem.

Theorem 2 *Algebraic Stability of Crank-Nicolson*

The Crank-Nicolson discretization is algebraically stable in the sense that, the CN method satisfies equation (3.23) with $\beta = 0, \alpha = 1/2$.

Proof 2 *The proof is the same as in [10]. All the Gershgorin disks of $\hat{M}$ are in the right half of the complex plane.*

□

3.1.2 Discretization of the Boundary condition

The spatial discretization given by matrix (3.14) depends on γ_1 and γ_2, the discretization of the boundary condition imposed at $r = 0$. In a financial context, we would like to insure convergence to the viscosity solution. Viscosity solutions have been discussed in [12]. It is therefore important that the discretization of the boundary conditions are monotone in order for the solution to converge to the viscosity solution.

Boundary condition at** $r = 0$**:

At the lower boundary point we solve the initial value problem using boundary condition from equation (2.4). In addition, we will alternatively use an arbitrary Neumann boundary condition

$$\frac{\partial V}{\partial \tau} = 0. \tag{3.24}$$

We will now refer to equation (2.4) and equation (3.24) as BC1 and BC2 respectively. BC1 is discretized using first order forward differencing as

$$(V_r)_0^n \simeq \left(\frac{V_1^n - V_0^n}{\Delta r_{0+1/2}} \right). \tag{3.25}$$

Therefore γ_1 and γ_2 of matrix (3.14) are

$$\gamma_1 = \kappa b \frac{1}{\Delta r_{0+1/2}},$$
$$\gamma_2 = -\kappa b \frac{1}{\Delta r_{0+1/2}},$$

and the discretization for BC2 is

$$\gamma_1 = 0,$$
$$\gamma_2 = 0.$$

Boundary condition at $r = r_{max}$***:***

We can implement the Dirichlet boundary condition $V(r_{max}, t) = 0$, by computing the solution on the grid $\hat{M}$ (3.14) with the last row and last column removed. Therefore, the matrix $\hat{M}$ becomes $(m-1) \times (m-1)$, where m is the number of grid points.

3.2 Discretization of stochastic numerical methods

In this section we discuss numerical schemes for the solution of the mean-reverting square root process given by equation (1.1). We are examining stochastic numerical methods

because these methods are able to 'naturally' implement boundary conditions that are imposed in stochastic differential equations. We therefore seek to compare the solution of our FD approach to that of a Monte-Carlo simulation with stochastic numerical methods.

In the numerical solution of SDEs, the convergence and numerical stability properties of the schemes play a fundamental role as well as in a deterministic framework. As discussed in [13], the region of absolute stability defines possible restrictions on the maximum allowed step size. However, based on the fact that the coefficients of (1.1) are non-linear and non-Lipschitzian it is not possible to appeal to standard convergence theory for numerical simulations to deduce the numerically computed paths are accurate for small step-sizes. These issues were discussed by [14], where the authors showed that a natural Euler-Maruyama discretization provides qualitatively correct approximations to the first and second moments of the SDE (1.1).

3.2.1 Numerical positivity

It is well known that the domain for the SDE in equation (1.1) is $[0, \infty)$. In order to simulate this process using numerical approximation we must be assured of not getting negative values. The concept of positivity or non-negativity for this process has been discussed in [15] and [16].

Definition 5 *Let $R(t)$ be a stochastic process defined on a domain of $[L, \infty)$ with*

$$P(R(t) > L \text{ for all } t > 0) = 1.$$

Then the stochastic integration scheme possesses an eternal life time if

$$P(R_{n+1} > L | R_n > L) = 1.$$

Otherwise it has a finite life time.

Based on the above definition we are therefore interested in numerical methods that have an eternal life time.

3.2.2 Numerical discretization

We will consider six numerical discretization schemes:

- Euler-Maruyama,

- Milstein,
- Second-order Milstein,
- Implicit Milstein,
- Balanced implicit method,
- Exact transition distribution.

A natural way to simulate the process in (1.1) is with the explicit Euler-Maruyama scheme

Definition 6 (Euler-Maruyama method)

$$R_{n+1} = R_n + \kappa(b - R_n)\Delta t + \sigma\sqrt{R_n \Delta t}(Z_{n+1}), \tag{3.26}$$

where the time-steps are discretized as $t_0 < t_1 < ... < t_n < T$, R_n is the interest rate value at the n^{th} time-step, $\Delta t = t_{n+1} - t_n$, and $Z_1, Z_2, ...$ are independent, m-dimensional standard normal random vectors. The Euler-Maruyama scheme can lead to negative values since the Gaussian increment is not bounded from below. Therefore, the standard Euler-Maruyanam scheme has a finite life time. A natural fix adopted in [14], is to replace the method with the computationally safer method

$$R_{n+1} = R_n + \kappa(b - R_n)\Delta t + \sigma\sqrt{|R_n|\Delta t}(Z_{n+1}). \tag{3.27}$$

Theorem 3 *For SDE (1.1) the first moment is*

$$\lim_{t\to\infty} E[R(t)] = b \tag{3.28}$$

and second moment is given by

$$\lim_{t\to\infty} E[R(t)^2] = b^2 + \frac{\sigma^2 b}{2\kappa}. \tag{3.29}$$

Proof 3 *As shown in [14], the first moment results by taking expectations of equation (1.1). The second moment result can be obtained by applying the Ito formula to $R(t)^2$ and taking expectations, using the result for $E[R(t)]$.*

□

Corollary 1 *We can use properties (3.28) and (3.29) to estimate the step-size needed to obtain qualitatively correct solutions. We can rewrite (3.27) as*

$$R_{n+1} = R_n(1 - \kappa\Delta t) + \kappa b\Delta t + \sigma\sqrt{|R_n|\Delta t}(Z_{n+1}). \tag{3.30}$$

Using induction, the exception of (3.30) can be shown to be

$$E(R_n) = (1 - \kappa\Delta t)^n(E(R_0) - b) + b$$

and hence

$$\begin{array}{llll} for & \Delta t < 2/\kappa, & E(R_n) \to b \ as \ n \to \infty, \\ for & \Delta t = 2/\kappa, & E(R_n) = (-1)^n E(R_0) + ((-1)^{n+1} + 1)b, \\ for & \Delta t > 2/\kappa, & |E(R_n)| \to \infty \ as \ n \to \infty. \end{array}$$

Proof 4 *As shown in [14], the proof follows after taking expected values in (3.30).*

□

Theorem 3 and Corollary 1 show that as $t_n \to \infty$ the correct mean is produced if and only if the step-size satisfies $\Delta t < 2/\kappa$. This result will be used in step-size selection for the following methods.

In addition to the Euler method, we will consider five other numerical schemes for the solution of SDEs. First, we consider the Milstein method which offers an improvement over the Euler scheme. The Euler scheme approximation expands the drift to $O(\Delta t)$ but the diffusion term only to $O(\sqrt{\Delta t})$. In order to ease notation we define the following $a = \kappa(b - R_n)$ and $c = \sigma\sqrt{|R_n|}$, where a, c, and their derivatives are all evaluated at time t.

Definition 7 (Milstein method)

$$R_{n+1} = R_n + a\Delta t + c\sqrt{\Delta t}Z_{n+1} + \frac{1}{2}c'c\Delta t(Z_{n+1}^2 - 1). \tag{3.31}$$

The approximation method in (3.31) adds a term to the Euler scheme. It expands both the drift and diffusion terms to $O(\Delta t)$. The new term $\frac{1}{2}c'c\Delta t(Z_{n+1}^2 - 1)$, has mean zero and is uncorrelated with the Euler terms because $Z_{n+1}^2 - 1$ and Z_{n+1} are uncorrelated.

Next, we consider a higher order version of the Milstein method. We consider a simplified version of the scheme as shown in [13].

Definition 8 (Second-order Milstein method)

$$\begin{aligned} R_{n+1} = R_n + a\Delta t + c\sqrt{\Delta t}Z_{n+1} + \frac{1}{2}\left(a'c + ac' + \frac{1}{2}c^2c''\right)\Delta t\sqrt{\Delta t}Z_{n+1} \\ + \frac{1}{2}cc'[Z_{n+1}^2 - \Delta t] + (aa' + \frac{1}{2}c^2a'')\frac{1}{2}\Delta t^2. \end{aligned} \tag{3.32}$$

This method was introduced by Milstein and has a weak order convergence of 2 under the conditions that a and c be six times continuously differentiable with uniformly bounded derivatives. This scheme should produce more accurate results than the Euler method, but it is computationally more expensive.

Next, we consider an implicit Milstein method. As noted in [13], implicit Euler schemes can reveal better stability properties. This implicit scheme is obtained by making implicit only the purely deterministic term of the equation, while at each time step, the coefficients of the random part of the equation are retained from the previous step. Note for the implicit method we use $a(R_{n+1}) = \kappa(b - R_{n+1})$.

Definition 9 (Implicit Milstein method) *The implicit Milstein method is defined by*

$$R_{n+1} = R_n + a(R_{n+1})\Delta t + c\sqrt{\Delta t}Z_{n+1} + \frac{1}{2}c'c\Delta t(Z_{n+1}^2 - 1). \tag{3.33}$$

Removing the implicitness and expanding $a(R_{n+1})$ and c yields

$$R_{n+1} = \frac{R_n + \kappa b + \sigma\sqrt{R_n\Delta t}Z_{n+1} + \frac{1}{4}\sigma^2\Delta t(Z_{n+1}^2 - 1)}{1 + \kappa\Delta t}. \tag{3.34}$$

The Milstein methods discussed previously fall under the category of dominating methods. Another idea to guarantee non-negativity for an SDE is to use balancing methods. The balancing method that we will examine was discussed in [16].

Definition 10 (Balanced implicit method (BIM)) *The integration scheme for the BIM is given as follows*

$$R_{n+1} = R_n + a\Delta t + c\sqrt{\Delta t}Z_{i+1} + (R_n - R_{n+1})Q_n(R_n), \tag{3.35}$$

$$Q_n(R_n) = q_0(R_n)\Delta + q_1(R_n)|\sqrt{\Delta t}Z_{i+1}|. \tag{3.36}$$

Removing the implicitness yields

$$R_{n+1} = \frac{R_n + a\Delta t + c\sqrt{\Delta t}Z_{i+1} + R_n(q_0(R_n)\Delta t + q_1(R_n)|\sqrt{\Delta t}Z_{i+1}|}{1 + q_0(R_n)\Delta t + q_1(R_n)|\sqrt{\Delta t}Z_{i+1}|}. \tag{3.37}$$

In this method the functions q_0 and q_1 are called control functions. To guarantee convergence of the method, the control functions must be bounded and have to satisfy

$$1 + q_0(R_n)\Delta t + q_1(R_n)|\sqrt{\Delta t}Z_{i+1}| > 0.$$

Finally, we consider a scheme of sampling the exact transition distribution (ETD) of the CIR process as shown in [17]. The SDE (1.1) is not explicitly solvable, however the transition density for the process is known. Based on work by Feller in [7], the distribution of $R(t)$ given $R(u)$ for some $t > u$ is, up to a scale factor, a non-central chi-squared distribution. This property can then be used to simulate the CIR process. The approach to simulate the process is suggested in [17]. Further information on the CIR distribution can be seen in Appendix A.

Definition 11 (Exact Transition Distribution (ETD)) *Given $R(u)$ and $t > u$, $R(t)$ is distributed as $\sigma^2\left(1 - e^{-\kappa(t-u)}\right)/(4\kappa)$ times a non-central chi-square random variable $\chi'^2_d(\lambda)$ with d degrees of freedom and non-centrality parameter*

$$\lambda = \frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2\left(1 - e^{-\kappa(t-u)}\right)} R(u).$$

Therefore,

$$R(t) = \frac{\sigma^2\left(1 - e^{-\kappa(t-u)}\right)}{4\kappa} \chi'^2_d\left(\frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2\left(1 - e^{-\kappa(t-u)}\right)} R(u)\right), \quad t > u,$$

where

$$d = \frac{4\kappa b}{\sigma^2}.$$

The algorithm that performs the ETD method is shown in Appendix C.

Chapter 4

Experiments and Results

In this section, we will analyze numerically the convergence of the FD method and numerical discretization schemes for different parameters of the CIR process. The parameters used in the simulations are shown in Table 4.1. Note that we are simulating two cases: Case 1 when the boundary at the origin is attainable (i.e $2\kappa b < \sigma^2$) and Case 2 the boundary at the origin is non-attainable (i.e $2\kappa b > \sigma^2$).

Table 4.1: CIR model parameters for the Case 1: attainable boundary and Case 2: non-attainable boundary

	Case 1	Case 2
Initial interest rate	2%	2%
κ	0.55	1.8
b	0.035	0.035
σ	0.3	0.3
Time to expiry	4.0 years	4.0 years

The parameters in Table 4.1 model the behaviour of a zero coupon bond that expires after 4 years. The interest rate starts at 2% and has a long term value or mean-reverting value of 3.5%. The volatility for the interest rate process is 30%. In case 1 the interest rate approaches the mean-reverting value at a slower rate than when in case 2.

We first consider pricing a zero coupon bond using the FD method. First, however we mention a few notes regarding the Crank-Nicolson scheme. It was shown in Section 3.1.1 that the Crank-Nicolson scheme is unconditionally stable. However, this does not prevent oscillations from forming in the solution. Modifying the boundary behavior of the solution may create discontinuities in the solution. As noted in [18], one way to avoid

the oscillations is to use a fully implicit method for a small number of time-steps after any discontinuities arise and Crank-Nicolson thereafter. This method called "Rannacher smoothing" as described by Rannacher (1984) restores quadratic convergence. We will use this method by applying two fully implicit time-steps and the remaining time steps will be Crank-Nicolson.

We seek to carry out a convergence study of the FD method. Each grid has twice as many nodes as the previous grid (new nodes inserted halfway between the coarse grid nodes) and the timestep size is halved. The error for the FD method is then give by the following

$$Error = O(\max((\Delta t)^2, (\Delta r)^2));\ \Delta r = \max_i (r_{i+1} - r_i).$$

Let

$$h = \max(C_1 \Delta r, C_2 \Delta t)$$

where C_1 and C_2 are constants. Then the solution on each grid (at a given point) has the form

$$\begin{aligned} V(h) &= V_{exact} + Ah^2, \\ V(h/2) &= V_{exact} + A(h/2)^2, \\ V(h/4) &= V_{exact} + A(h/4)^2, \end{aligned} \tag{4.1}$$

where V_{exact} is the exact analytical solution and we have assumed that the mesh size and timestep are small enough that the coefficient A is approximately constant. This implies that

$$\frac{V(h) - V(h/2)}{V(h/2) - V(h/4)} \simeq 4. \tag{4.2}$$

We are therefore interested in a convergence rate of around four. This is called quadratic convergence.

As a comparison to the Crank-Nicolson scheme we compute the price of a bond using the numerical methods with a Monte Carlo (MC) approach. In the MC approach there are two sources of error: time-stepping error and sampling error. The error in MC is then given by

$$\begin{aligned} Error &= O\left(\max(\Delta t, \frac{1}{\sqrt{M}})\right), \\ \Delta t &= \text{timestep}, \\ M &= \text{number of Monte Carlo paths}. \end{aligned}$$

We seek to balance the time-stepping error and sampling error. In order to make these two errors the same order we choose $M = O\left(\frac{1}{(\Delta t)^2}\right)$. In this way we increase the number of paths by four and divide the timestep by two for each simulation.

4.1 Weak Convergence Criterion

Usually we do not know the solution of a stochastic differential equation explicitly, so we will use simulation to try to discover the solution. It was shown by [4] that the solution to the bond pricing problem is given exactly by the solution to the FD method with the appropriate boundary condition. To this end we shall repeat M different simulations of sample paths of the CIR process using the numerical approximations. We shall estimate the weak error as

$$\text{Weak error} = \left| \frac{1}{M} \sum_{m=1}^{M} [V^h(T)_m] - V(T) \right| \tag{4.3}$$

where step-size $h = T/N$, with N being the number of time-steps, and $V^h(T)_m$ is the solution obtained by the Monte Carlo time-stepping along the m^{th} Brownian path, and $V(T)$ is the solution from the FD method.

Table 4.2: Finite difference simulation of bond price using the CIR model where the boundary at the origin is **attainable**. Grid fixed at $r = 0.02$. Parameters are given in Table (4.1).

Method	**Nodes**	**Time-steps**	**Value**	**Change**	**Ratio**
BC1	102	5	0.89646	-	-
	203	10	0.896184	-2.75544×10^{-4}	-
	405	20	0.896116	-6.83711×10^{-5}	4.03012
	809	40	0.896099	-1.66775×10^{-5}	4.09959
	1617	80	0.896095	-4.08764×10^{-6}	4.08000
	3233	160	0.896094	-1.00939×10^{-6}	4.04960
	6465	320	0.896094	-2.50617×10^{-7}	4.02763
BC2	102	5	0.937245	-	-
	203	10	0.936078	-1.16676×10^{-3}	-
	405	20	0.935761	-3.16871×10^{-4}	3.68214
	809	40	0.93567	-9.08804×10^{-5}	3.48668
	1617	80	0.935641	-2.92177×10^{-5}	3.11046
	3233	160	0.935629	-1.14844×10^{-5}	2.54412
	6465	320	0.935624	-5.66127×10^{-6}	2.02860

Table 4.3: Finite difference simulation of bond price using the CIR model when the boundary at the origin is **non-attainable**. Grid fixed at $r = 0.02$. Parameters are given in Table (4.1).

Method	**Nodes**	**Time-steps**	**Value**	**Change**	**Ratio**
BC1	102	5	0.878503	-	-
	203	10	0.878015	-4.88367×10^{-4}	-
	405	20	0.877892	-1.22579×10^{-4}	3.98409
	809	40	0.877862	-3.07156×10^{-5}	3.99078
	1617	80	0.877854	-7.68650×10^{-6}	3.99605
	3233	160	0.877852	-1.92238×10^{-6}	3.99842
	6465	320	0.877852	-4.80678×10^{-7}	3.99931
BC2	102	5	0.881929	-	-
	203	10	0.880122	-1.80663×10^{-3}	-
	405	20	0.879356	-7.65886×10^{-4}	2.35888
	809	40	0.878937	-4.18938×10^{-4}	1.82816
	1617	80	0.878662	-2.75726×10^{-4}	1.51940
	3233	160	0.878462	-1.99039×10^{-4}	1.38529
	6465	320	0.878314	-1.48614×10^{-4}	1.33930

4.2 Convergence and Verification of methods

In order to test out the convergence of the FD and MC methods we developed simulations using MATLAB. All simulations were run in MATLAB 7.8.0 on a MacOSX with 4GB of memory.

We validate our FD method by demonstrating the convergence of bond prices using the CIR process. We consider the Crank-Nicolson discretization using boundary conditions BC1 and then BC2 with constant time-stepping. The number of points in the grid of interest rates is doubled for each simulation. For each test, as we double the number of points we divide the time-step in half. The range of the grid is from $[r_0 = 0, r_{max} = 10]$ with initially 102 points and eventually expanded to 6465 points. The price of the bond is computed using the different sets of parameters as shown in Table 4.1. The results are contained in Tables 4.2 and 4.3. In the Tables we fix a point on the grid which is close to the boundary ($r = 0.02$ in our case) and compare the convergence rate to the theory as given in equation (4.2).

Looking at Table 4.2 for the case where the boundary is attainable it can be seen that BC1 and BC2 do not converge to the same solution (the absolute difference between the two boundary conditions at the finest grid is 3.95×10^{-2}). In Table 4.3 are the results of the FD method when the boundary is non-attainable. In the case of Table 4.3, BC1 and BC2 are converging to the same solution (the absolute difference between the two boundary conditions at the finest grid is 4.62×10^{-4}), however with BC2 the convergence rate is much slower. Note that with BC2 the solution does not appear to be converging quadratically (the ratio of changes is about two (first order rate) for the attainable boundary case and about 1.3 for the non-attainable boundary case instead of four, which we would expect for quadratic convergence).[1]

The results demonstrate that where there is a non-attainable boundary, the boundary condition is not required as both BC1 and BC2 will eventually converge to the same solution. Thus, if the boundary is non-attainable, then the boundary behavior is not needed to guarantee uniqueness of solution. While convergence is guaranteed using both boundary conditions, from a numerical perspective using BC1 will provide faster convergence even if it is not required from a mathematical perspective.

The results are further emphasized by comparing the plots of the FD solution in Figures 4.3 and 4.2. In the case of the attainable boundary in Figure 4.1(a), as the number of grid points is increased the solution with BC2 will never converge to the solution with BC1.

[1] We remark that a variable time-step selector was tested in order to see if the convergence ratio could be improved. The time-step selector was based on one used in [19]. However, it was found that the variable time-step selector did not improve the convergence ratio when used for the FD method with either BC1 or BC2.

In the non-attainable case in Figure 4.1(b), as the number of grid points is increased the solution with BC2 converges to the solution in BC1. The difference in the two boundary conditions is also evident in Figure 4.2, where we can see plots of bond price in the attainable and non-attainable cases using the different boundary conditions for time values from 0 to T.

The Figures 4.3 and 4.2 also partly explain the slower convergence rate when using BC2. When using BC2, the FD solution at $r = 0$ is pinned down to the value one, so in order to converge to the true solution at $r = 0$ requires a much finer grid than in the case of BC2. However, it was observed that the convergence rate when using BC2 was improved with larger initial interest rates (i.e fixing the grid at larger interest values). This is to be expected since the solution values computed using larger initial interest rates do not depend as much on the solution values close to the boundary.

We now turn to the numerical solution of the SDE in equation (1.1) using the MC approach. We tested the six SDE discretization schemes - Euler, Milstein, 2nd Order Milstein, implicit Euler, BIM, and ETD - described Section 3.2. In order to obtain bond prices we simulate the discretization $R(t) = R_n,\ n = 0, 1, ..., N-1$ with the step-size $h = T/N$, the simplest estimate of the bond price is then

$$V(t) = \frac{1}{M} \sum_{m=0}^{M-1} \left[\exp\left(-h \sum_{n=0}^{N-1} R_n^m \right) \right]$$

with m being the m-th brownian path.

The convergence results are shown in Tables 4.4 and 4.5. Looking at Table 4.4 for the case where the boundary is attainable, we can see that all methods except the BIM and ETD method produce negative paths. For the case where the boundary is non-attainable, as shown in 4.5 the BIM and EDT method still out performs all other methods in terms of non-negative paths.

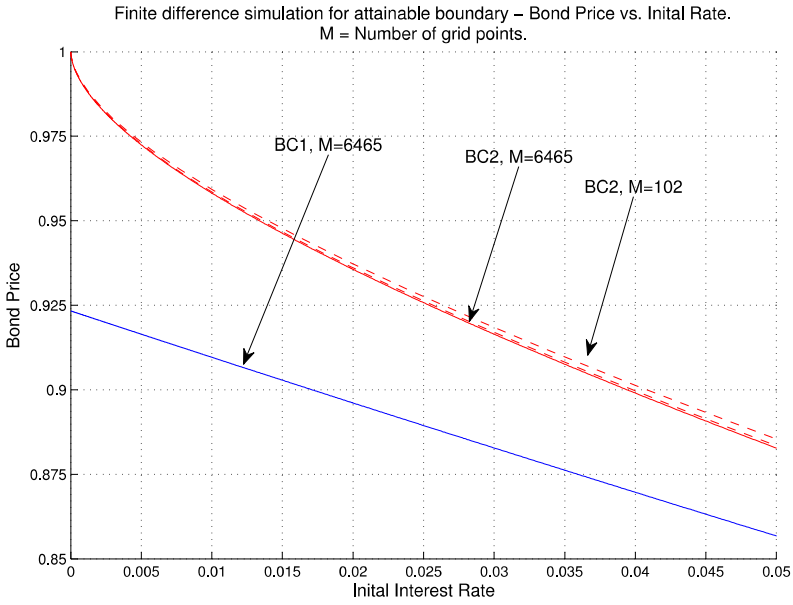
We wish to show that the numerical approximation methods converge to the solution given by the FD method. Comparing the value of the bond and change-in-value in Tables 4.4 and 4.5, it is evident that all numerical methods except the BIM method converge to the solution given by the FD method.[2] Note also that all methods except the BIM are converging to the solution given by BC1 when the boundary is attainable. Figures 4.3(a) and 4.3(b) confirm the convergence results by showing a comparison of the bond price for the different methods versus computation cost.

[2] As shown in [15], the BIM method is not an appropriate choice to integrate this SDE. For the SDE (1.1), it is necessary to choose the control functions as $q_0 = \kappa$ and $q_1(R(t)) = \sigma R(t)^{-\frac{1}{2}}$. However, in the limit as $R(t) \to 0$, q_1 is an unbounded function and therefore does not satisfy the condition in Definition 10.

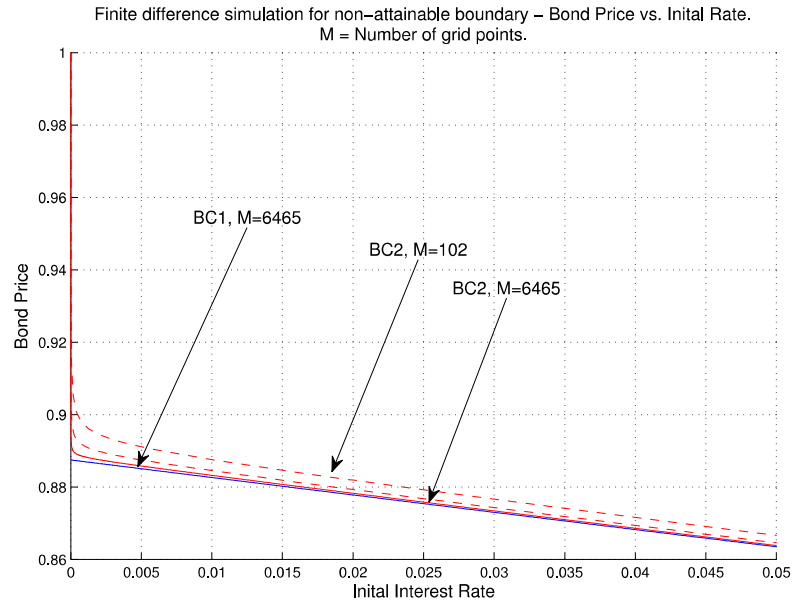
Next we compare the weak error of the MC methods relative to the FD method as defined in equation (4.3). Analyzing the error of the six different SDE methods in Table 4.6, we note that the absolute error of the methods lies in the interval $[10^{-4}, 10^{-3}]$ for the attainable boundary and non-attainable boundary case. In the attainable boundary case, the Euler method has the largest error, while for the non-attainable case the BIM method performs the worst. In both cases the ETD, 2nd Order Milstein, and ETD methods have the smallest error.

Finally, we examine the distribution produced by the numerical methods. Figure 4.4 compares the exact distribution of $R(t)$ as given by the equation (A.3) with the distribution produced by five numerical approximations. First note that for the case of an attainable boundary, the exact distribution approaches an exponentially increasing value in the limit as $R(t) \to 0$. The Milstein and 2nd order Milstein methods produce the closest match to the exact distribution in the attainable case.

Examining the case where the boundary is non-attainable in Figure 4.4(b), the exact distribution approaches zero in the limit as $R(t) \to 0$. Note that for this case, the distribution for the BIM method is skewed to the right of the true mode. The Milstein methods closely match the exact distribution, while the Euler method produces too many values close to or below zero.



(a)



(b)

Figure 4.1: Finite difference results for a zero coupon bond using the CIR model. Figure 4.1(a) is the bond price computed using different boundary conditions when boundary at the origin is attainable, Figure 4.1(b) is the bond price computed using two different boundary conditions when boundary at the origin is non-attainable. Parameters are given in Table 4.1.

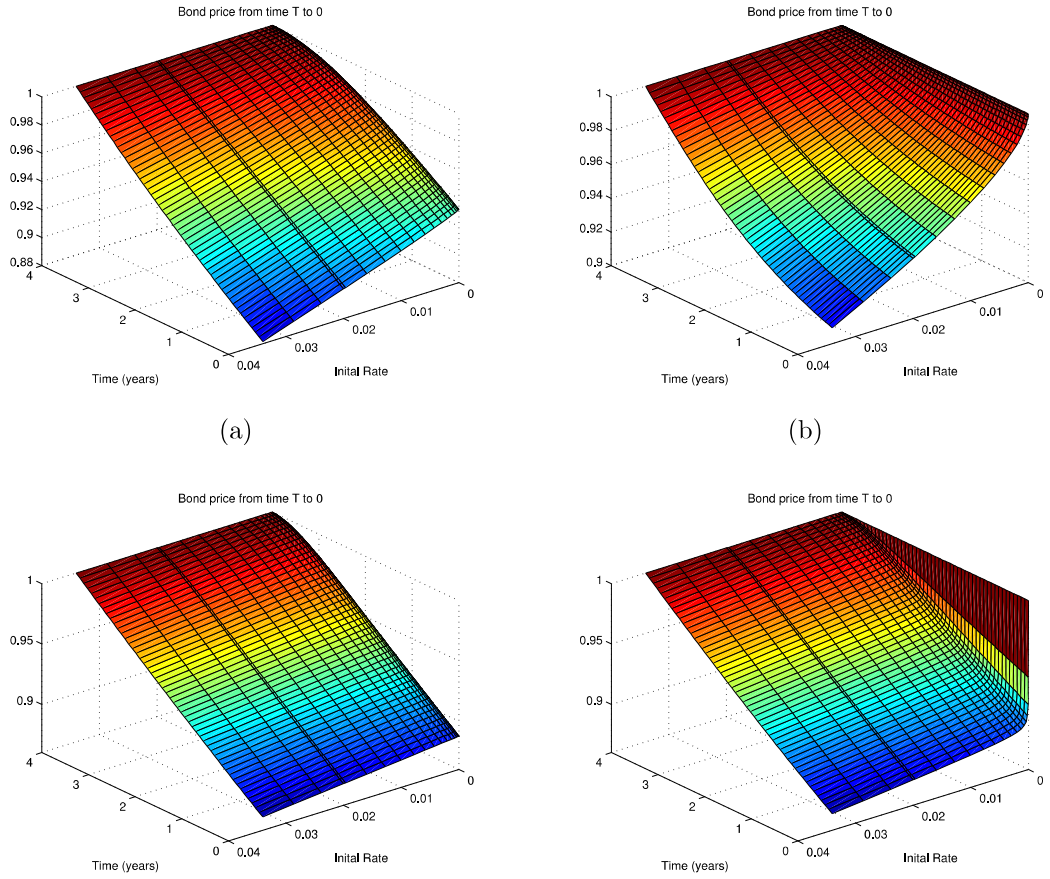


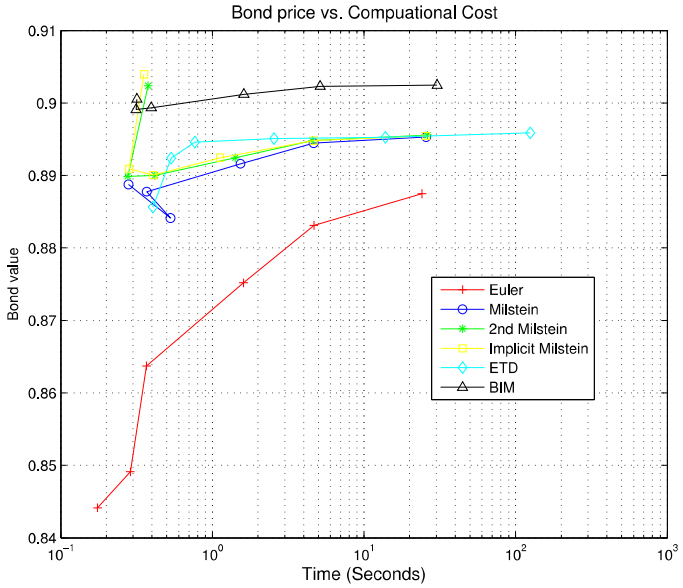
Figure 4.2: Figures 4.2(a) and 4.2(b), attainable boundary with BC1 and BC2 respectively. Figures 4.2(c) and 4.2(d), non-attainable boundary with BC1 and BC2 respectively. Parameters are given in Table 4.1.

Table 4.4: Monte-Carlo simulation of pricing a zero-coupon bond using the CIR model. Boundary at the origin is **attainable**. Parameters are given in Table 4.1.

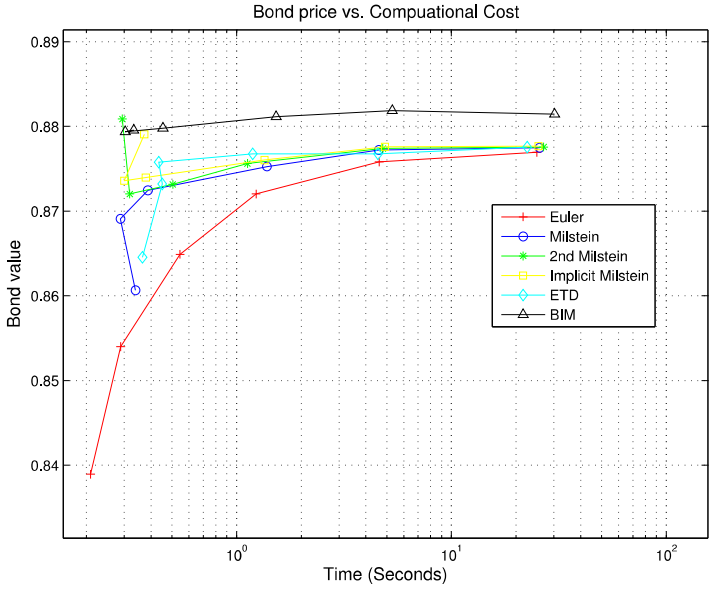
Method	**# Paths**	**#Time Steps**	**Neg. Paths**	**Value**	**Change**
Euler	100	9	9.00%	0.844147	-
	400	17	8.05%	0.849106	4.9586×10^{-3}
	1600	33	8.34%	0.863727	1.46211×10^{-2}
	6400	65	8.40%	0.875178	1.14512×10^{-2}
	25600	129	8.47%	0.883108	7.9302×10^{-3}
	102400	257	8.46%	0.88747	4.36155×10^{-3}
Milstein	100	9	9.20%	0.888752	-
	400	17	8.20%	0.884117	-4.63501×10^{-3}
	1600	33	8.20%	0.887759	3.64191×10^{-3}
	6400	65	8.20%	0.891614	3.85466×10^{-3}
	25600	129	8.30%	0.894456	2.84233×10^{-3}
	102400	257	8.30%	0.895302	8.45813×10^{-4}
2nd Order Milstein	100	9	6.30%	0.902384	-
	400	17	6.10%	0.889893	-1.24915×10^{-2}
	1600	33	6.90%	0.889999	-1.05806×10^{-4}
	6400	65	7.60%	0.892434	2.43495×10^{-3}
	25600	129	7.90%	0.89484	2.40593×10^{-3}
	102400	257	8.10%	0.895553	7.13607×10^{-4}
Implicit Milstein	100	9	7.70%	0.903941	-
	400	17	6.7%	0.890924	-1.30172×10^{-2}
	1600	33	7.4%	0.890057	-8.66521×10^{-4}
	6400	65	7.9%	0.892454	2.39712×10^{-3}
	25600	129	8.1%	0.894806	2.35181×10^{-3}
	102400	257	8.2%	0.895442	6.3583×10^{-4}
BIM	100	9	0%	0.900543	-
	400	17	0	0.899096	-1.44756×10^{-3}
	1600	33	0	0.899333	2.37411×10^{-4}
	6400	65	0	0.901201	1.86852×10^{-3}
	25600	129	0	0.902291	1.08979×10^{-3}
	102400	257	0	0.902452	1.60503×10^{-4}
ETD	100	9	0%	0.88611	-
	400	17	0	0.894203	8.09309×10^{-3}
	1600	33	0	0.893621	-5.82248×10^{-4}
	6400	65	0	0.895367	1.74598×10^{-3}
	25600	129	0	0.895894	5.27318×10^{-4}
	102400	257	0	0.895605	-2.89517×10^{-4}

Table 4.5: Monte-Carlo simulation of pricing a zero-coupon bond using the CIR model. Boundary at the origin is **non-attainable**. Parameters are given in Table 4.1.

Method	**# Paths**	**#Time Steps**	**Neg. Paths**	**Value**	**Change**
Euler	100	9	10.0%	0.838945	-
	400	17	9.70%	0.854002	1.50569×10^{-2}
	1600	33	3.70%	0.864918	1.09161×10^{-2}
	6400	65	0.22%	0.872035	7.11642×10^{-3}
	25600	129	2.73e-03%	0.875821	3.78631×10^{-3}
	102400	257	0%	0.87696	1.13861×10^{-3}
Milstein	100	9	10.0%	0.860662	-
	400	17	9.80%	0.869099	6.71405×10^{-1}
	1600	33	1.30%	0.872459	2.86081×10^{-2}
	6400	65	0%	0.875264	3.9011×10^{-3}
	25600	129	0	0.87722	1.80347×10^{-3}
	102400	257	0	0.877487	2.67708×10^{-4}
2nd Order Milstein	100	9	10.0%	0.880883	-
	400	17	3.85%	0.872025	-8.85794×10^{-3}
	1600	33	0%	0.873157	1.13106×10^{-3}
	6400	65	0	0.875632	2.47533×10^{-3}
	25600	129	0	0.877408	1.77643×10^{-3}
	102400	257	0	0.877578	1.69779×10^{-4}
Implicit Milstein	100	9	0%	0.87906	-
	400	17	0	0.873596	-5.46375×10^{-3}
	1600	33	0	0.873978	3.82041×10^{-4}
	6400	65	0	0.876024	2.04589×10^{-3}
	25600	129	0	0.877598	1.57397×10^{-3}
	102400	257	0	0.877678	7.96427×10^{-5}
BIM	100	9	0%	0.879563	-
	400	17	0	0.879377	-1.85705×10^{-4}
	1600	33	0	0.879793	4.16128×10^{-4}
	6400	65	0	0.881177	1.38406×10^{-3}
	25600	129	0	0.881864	6.86789×10^{-4}
	102400	257	0	0.881455	-4.09203×10^{-4}
ETD	100	9	0%	0.861204	-
	400	17	0	0.874199	3.9295×10^{-3}
	1600	33	0	0.874868	2.9482×10^{-3}
	6400	65	0	0.875975	2.72441×10^{-3}
	25600	129	0	0.876965	8.6216×10^{-4}
	102400	257	0	0.87752	3.5479×10^{-4}



(a)

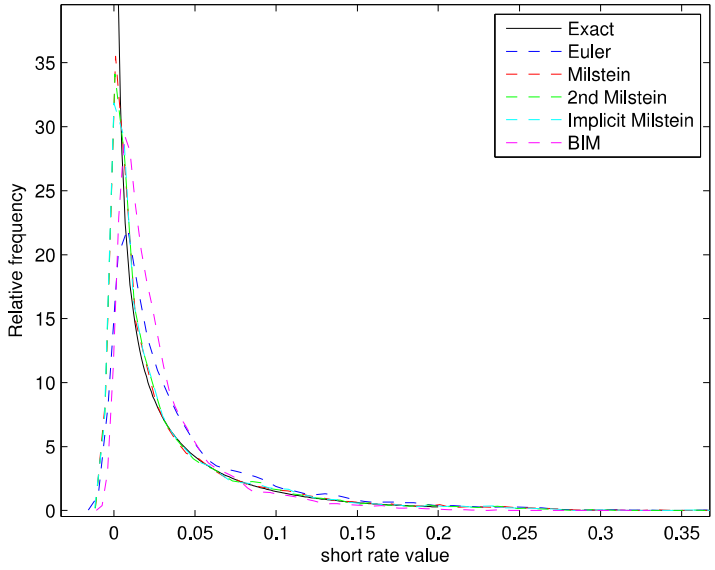


(b)

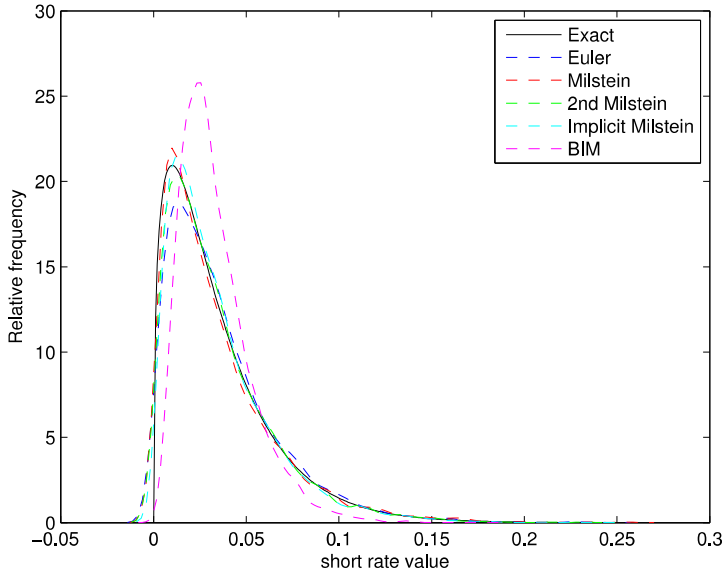
Figure 4.3: Comparison of numerical approximation bond value versus computational cost in a Monte-Carlo simulation. Figure 4.3(a), boundary is attainable. Figure 4.3(b), boundary is non-attainable. Parameters are given in Table 4.1.

Table 4.6: Weak error of numerical approximation schemes. Case 1: attainable boundary and Case 2: non-attainable boundary. FD method computed using BC1, 6465 grid points, and 320 time-steps. Numerical discretization computed at 102400 paths and 257 time-steps.

	Error for Case 1	Error for Case 2
Euler	8.624×10^{-3}	8.920×10^{-4}
Milstein	7.920×10^{-4}	3.650×10^{-4}
2nd Order Milstein	5.410×10^{-4}	2.740×10^{-4}
Implicit	6.520×10^{-4}	1.740×10^{-4}
BIM	6.358×10^{-3}	3.603×10^{-3}
ETD	4.890×10^{-4}	3.320×10^{-4}



(a)



(b)

Figure 4.4: Comparison of exact distribution and the distribution of the numerical approximation schemes for the square-root diffusion. Figure 4.4(a), boundary is attainable. Figure 4.4(b), boundary is non-attainable. Monte Carlo simulation with $M = 5000, \Delta t = 0.1$, and CIR parameters given in Table 4.1.

Chapter 5

Conclusions

In this paper we have provided an analysis of different boundary conditions for the CIR equation or square-root process. Boundary conditions are needed for finite difference methods, however in the CIR process, boundary conditions are only needed when the boundary is attainable. When the boundary of the CIR process is attainable there is only one true boundary condition that will satisfy the original CIR stochastic differential equation, and this is given by equation (2.4). While imposing another boundary condition is valid from a finite difference sense, it will artificially modify the density of the CIR stochastic differential equation. When the boundary of the CIR process is not attainable it is possible to use any arbitrary values at the boundary and still have convergence to a valid solution. However, based on our results we have shown that when using finite difference methods for the term structure equation, the convergence rate is far greater with the boundary condition given by equation (2.4).

We also provide a validation of the FD results by showing that numerical approximations of the CIR process and Monte Carlo simulations are able to provide natural convergence. We test six numerical approximations to the CIR process. It was shown that the EDT and 2nd Order Milstein, and Implicit methods converge to the FD solution the fastest, while the Euler method provides slower convergence. Further, it was shown that the BIM method is not appropriate for the CIR process because the control functions used in the BIM method become unbounded for vanishing interest rates.

Finally, we repeat the main result of the paper. If the boundary is attainable, then this boundary behavior servers as a boundary condition and guarantees uniqueness of the solutions. However, if the boundary is non-attainable, then the boundary behavior is not needed to guarantee uniqueness.

APPENDICES

Appendix A

Probability distribution of the CIR process

This appendix provides an overview of the probability distribution for the process given in equation (1.1). Some useful properties about the boundary behavior can be understood by examining the stationary probability distribution.

Let $P(R(t), t)$ be the probability density function of interest rate process $R(t)$ under the risk-neutral measure at time t, given an initial Dirac delta function distribution at time $t = t_0$

$$P(R(t), t = t_0) = \delta(R_t - R_0) \tag{A.1}$$

This function is referred to as a pricing kernel. The resulting equation for a pricing kernel is called the *Fokker-Planck* equation.

Theorem 4 (Fokker-Planck Equation) *The probability density function $P(r, t)$ under the risk-neutral measure for the interest rate values R at time t satisfying initial conditions (A.1), obeys the following equation:*

$$\frac{\partial P(r,t)}{\partial t} = \kappa P(r,t) - \kappa(b - r)\frac{\partial P(r,t)}{\partial r} + \frac{1}{2}\frac{\partial^2}{\partial r^2}(\sigma^2 r P(r,t)) \tag{A.2}$$

where $r = R(t)$.

Proof 5 *The proof follows exactly as in [6].*

An exact analytical solution of the time-dependent Fokker-Planck equation for the distribution function $P(r,t) = P(r, r_0; t)$, subject to the initial-time condition $P(r, t = 0) = \delta(r - r_0)$, can be shown as in [6], to take the form

$$P(r, r_0; t) = c_t \left(\frac{r e^{\kappa t}}{r_0} \right)^{q/2} \exp(-c_t(r_0 e^{-bt} + r)) I_q(2c_t(r_0 r e^{-bt})^{1/2}) \tag{A.3}$$

where $c_t = 2\kappa/(\sigma^2(1 - e^{-\kappa t}))$, $q = (2\kappa b/\sigma^2) - 1$, and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. In the long time limit $t \to \infty$ the distribution approaches a steady state with $\partial p/\partial t \to 0$. The stationary probability distribution is given by

$$P_\infty(r) = \frac{(2\kappa/\sigma^2)^{2\kappa b/\sigma^2}}{\Gamma(2\kappa b/\sigma^2)} r^{(2\kappa b/\sigma^2)-1} e^{-(2\kappa b/\sigma^2)r} \tag{A.4}$$

where $\Gamma(\cdot)$ is the gamma function. It is possible to deduce some interesting properties of the density near the boundary $r = 0$. The interested reader can refer to [20] and [7] for more details.

$$\lim_{r \to 0+} P_\infty(r) = \begin{cases} 0, & 2\kappa b > \sigma^2 \\ \text{const}, & 2\kappa b = \sigma^2 \\ \infty & 2\kappa b < \sigma^2 \end{cases} \tag{A.5}$$

The above equation says that when $2\kappa b > \sigma^2$, there is zero probability of attaining zero interest rates. When $2\kappa b < \sigma^2$, there is positive probability of attaining zero interest rates. Equation (A.5) allows us to determine when a boundary condition will be necessary for the process given in equation (1.1).

Appendix B

Analytical solution to the term structure equation

As shown by [21], a closed form solution for the time zero value of a bond under the CIR model is given by

$$B(t,T) = a(t,T)e^{-b(t,T)r(t)} \tag{B.1}$$

where

$$a(t,T) = \left[\frac{2h\exp\{(\alpha+h)(T-t)/2\})}{2h+(\alpha+h)(\exp\{(T-t)h\}-1)}\right]^{2\alpha b/\sigma^2},$$

$$b(t,T) = \frac{2(\exp\{(T-t)h\}-1)}{2h+(\alpha+h)(\exp\{(T-t)h\}-1)},$$

$$h = \sqrt{\alpha^2+2\sigma^2}.$$

Appendix C

Algorithm for simulation of square-root diffusion

Simulation of $dr(t) = \kappa(b - r(t))dt + \sigma\sqrt{r(t)}dW(t)$ on the time grid $0 = t_0 < t_1 < \ldots < t_n$ with $d = 4b\kappa/\sigma^2$

Case 1: $d > 1$
for $i = 0, ..., n-1$
    $c \leftarrow \sigma^2(1 - e^{-\kappa(t_{i+1}-t_i)}/(4\kappa)$
    $\lambda \leftarrow r(t_i)(e^{-\kappa(t_{i+1}-t_i)}/c$
    generate $Z \sim N(0,1)$
    generate $X \sim \chi^2_{d-1}$
    $r(t_{i+1}) \leftarrow c[(Z + \sqrt{(}\lambda)^2 + X]$
end
Case 2: $d \leq 1$
for $i = 0, ..., n-1$
    $c \leftarrow \sigma^2(1 - e^{-\kappa(t_{i+1}-t_i)}/(4\kappa)$
    $\lambda \leftarrow r(t_i)(e^{-\kappa(t_{i+1}-t_i)}/c$
    generate $N \sim$ Poisson$(\lambda/2)$
    generate $X \sim \chi^2_{d+2N}$
    $r(t_{i+1}) \leftarrow cX$
end

Simulation of square-root diffusion by sampling from the transition density as shown in [17].

Bibliography

[1] J. C. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, no. 2, pp. 385–407, 1985. 1

[2] L. Rogers and D. Williams, *Diffusions, Markov Processes and Martingale*. Cambridge: Cambridge University Press, second ed., 2000. 1

[3] S. Shreve, *Stochastic Calculus Models for Finance: Continuous Time Models*. New York: Springer Finance, 2004. 1

[4] E. Eskstrom, P. Lotstedt, and J. Tysk, "Boundary values and finite difference methods for the single factor term structure equation," *Applied Mathematical Finance,*, vol. 16, no. 3, pp. 253–259, 2009. 2, 5, 11, 23

[5] S. Heston, M. Loewenstein, and G. A. Willard, "Options and bubbles," *The Review of Financial Studies*, vol. 20, no. 2, pp. 359–390, 2007. 2

[6] C. Albanese and G. Campolieti, *Advanced Derivatives Pricing and Risk Management: Theory, Tools, and Hands-On Programming Applications*. Academic Press Advanced Finance Series, 2005. 3, 4, 37, 38

[7] W. Feller, "Two singular diffusion problems," *Annals of Mathematics*, vol. 54, no. 1, pp. 173–182, 1951. 5, 20, 38

[8] D. Tavella and C. Randall, *Pricing Financial Instruments, The Finite Difference Method*. New York: John Wiley & Sons, 2000. 5

[9] D. J. Duffy, *Finite Difference Methods in Financial Engineering - A partial Differential Equation Approach*. England: John Wiley & Sons Ltd, second ed., 2006. 7

[10] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal, "Robust numerical methods for contingent claims under jump diffusion processes," *IMA Journal of Numerical Analysis*, vol. 25, pp. 87–112, 2003. 10, 11, 14

[11] P. Forsyth and K. Vetzal, "Course notes for numerical computation for financial modelling," 2008. Instructor Yuying Li, School of Computer Science, University of Waterloo, Winter 2009. 12

[12] D. M. Pooley, P. A. Forsyth, and K. R. Vetzal, "Numerical convergence properties of option pricing pdes with uncertain volatility," *IMA Journal of Numerical Analysis*, vol. 23, p. 2003, 2003. 14

[13] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Verlag Berlin Heidelberg New York: Springer, 1992. 16, 18, 19

[14] D. Higham and X. Mao, "Convergence of monte carlo simulations involving the mean-reverting square root process," *Journal of Computational Finance*, vol. 8, no. 3, pp. 35–62, 2005. 16, 17, 18

[15] C. Kahl, "Positive Numerical Integration of Stochastic Differential Equations," 2004. Diploma thesis, University of Wuppertal. 16, 26

[16] C. Kahl, M. Gnther, and T. Rossberg, "Structure preserving stochastic integration schemes in interest rate derivative modeling," *Applied Numerical Mathematics*, vol. 58, no. 3, pp. 284 – 295, 2008. 16, 19

[17] P. Glasserman, *Monte Carlo Methods in Financial Engineering*. New York: Springer, second ed., 2004. 20, 40

[18] Y. d'Halluin, P. A. Forsyth, K. R. Vetzal, and G. Labahn, "A numerical pde approach for pricing callable bonds," *Applied Mathematical Finance*, vol. 8, pp. 49–77, 2001. 21

[19] P. Forsyth and K. Vetzal, "Quadratic convergence of a penalty method for valuing american options," *SIAM Journal on Scientific Computation*, vol. 23, p. 2002, 2002. 25

[20] V. Lucic, "Boundary Conditions for Computing Densities in Hybrid Models via PDE Methods." Available at SSRN: http://ssrn.com/abstract=1191962, 2008. 38

[21] D. Brigo and F. Mercurio, *Interest Rate Models - Theory and Practice*. Berlin Heidelberg New York: Springer, third ed., 2006. 39