UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

TRINOMIAL-TREE DISTRIBUTION OF A NONLINEAR POSITIVE STOCHASTIC INTEREST RATE MODEL WITH CONNECTIONS TO THE POTENTIAL APPROACH AND ITS APPLICATION TO COMPUTING CORPORATE DEFAULT RISK

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

WEIHUA LIN
Norman, Oklahoma
2013

TRINOMIAL-TREE DISTRIBUTION OF
A NONLINEAR POSITIVE STOCHASTIC INTEREST RATE MODEL
WITH CONNECTIONS TO THE POTENTIAL APPROACH AND
ITS APPLICATION TO COMPUTING CORPORATE DEFAULT RISK

A DISSERTATION APPROVED FOR THE
GRADUATE COLLEGE

BY

Dr. Kevin Grasse, Co-Chair

Dr. S. Lakshmivarahan, Co-Chair

Dr. Duane Stock

Dr. Louis Ederington

Dr. Luther White

Dr. Christian Remling

DEDICATION

to

My father Longkun Lin, and mother Qinying Lin

and

my brother, Weibin Lin

For

Encouraging me to follow my dreams

Acknowledgements

This dissertation owes many thanks to many people...

First, I wish to express my sincere gratitude to my research advisors, Professor **Kevin Grasse** and Professor **S. Lakshmivarahan**, for teaching me a great deal of mathematics, computer science and finance, for encouraging to explore and work in computational finance, for their guidance and unconditional support during all these years of graduate school, and for being a excellent mentors. Without their guidance and inspiring advices, this dissertation will never come to real.

Professor **Kevin Grasse** guided me every step through my Ph.D. research with his amazing patience, motivation, enthusiasm and immense knowledge. He not only helped me build up a solid mathematical foundation on stochastic analysis, but also taught me the first financial mathematics class and helped me tune my career path to a financial model builder. I greatly appreciate his kind understanding and help to me when I went through all kinds of tough troubles in my research and my life. I could not have imaged having a better advisor and mentor for my Ph.D research.

Professor **S. Lakshmivarahan** is the mentor who makes my dream comes true. His enthusiasm and immense knowledge open the door of scientific computing for me and guided my steps in exploring a career to computational finance. With his knowledge and expertise in financial modeling, he successfully migrated me from an mathematician to an analyst in finance. This is a milestone in my career life. I also have learnt tremendous knowledge from his algorithm analysis class, which definitely will play a big role on my future career.

I wish to thank the members of my dissertation committee, Professor **Duane Stock** and Professor **Louis Ederington**, from OU finance division. They showed me the way of finance research. They taught me the investment fund, financial market, term structures, and the risk management, which are all the keys of the background and the application of my work. With their knowledge and expertise in investments, they successfully lead me the finance world.

The members of my dissertation committee, Professor **Luther White** and Professor. **Christian Remling**, have generously given their time and expertise to better my work. The real analysis and functional analysis courses taught by Dr Remling, and the numerical analysis course taught by Dr. White helped me built up my solid mathematical fundamentals to my research. I thank them for their contribution and their support.

I also thank **Dan Balan**, **Dana King** and **Minnan Zhao**, who are my managers in MSCI Inc. They provided me the chance to work in this financial risk analysis firm, which is absolutely a unique treasure toward my Ph.D. research. I learnt a lot of real finance applications from MSCI before stepping out of school.

My parents, **Longkun Lin** and **Qinying Lin**, and my brother **Weibin Lin** deserve my every appreciation in heart. I dedicate any success to them. I thank for their encouragement, straightly trust and continuous support all along my goal pursuing.

Contents

List of Tables

List of Figures

Abstract

In this dissertation, I consider a new nonlinear stochastic interest rate model that is adapted from a stochastic population growth model and exhibits the desirable properties of positivity of interest rates and mean reversion. We show that in the constant parameter case this model falls within the paradigm of the Rogers approach for generating positive interest rate models.

Moreover, motivated by a procedure initiated by Hull and White, we also offer a variant of the model with a time-dependent parameter that allows calibration of the model to a specified initial term structure when a trinomial-tree method is implemented to obtain discrete approximations of the distributions of the interest-rate process. Although nonlinear, our model has a closed form solution, which facilitates the generation of sample paths by standard numerical methods. This allows us to carry out the trinomial-tree method to obtain approximate distribution of the interest-rate process and compare that result to the approximate distributions obtained by Monte Carlo simulation.

We incorporated the positive interest rate to derive the firm's default probability, which thereby extends Qian's work from a linear interest rate model to a non-linear interest rate model. In the research, first comparing to Qian's method, we used the first passage time method based on the Fortet integral equation to derive the firm's default probability as driven by the Vasicek inter-

est rate model.

As an alternative, we also proposed the coupled trinomial tree method to derive the default probability. With the comparison of the numerical results among the three methods, we successfully extended the coupled trinomial tree algorithm for default probability from the linear model to a nonlinear model and obtained reasonably consistent results.

Chapter 1

Introduction

Rogers in (1997) [35] developed a very elegant and a useful framework called the potential approach for systematically deriving many models for positive short-term interest rates.(Also see Cairns(2004) [6]). A thorough and comprehensive review of the potential approach is contained in the recent thesis by Parbhoo(2009) [29]; also refer to chapter 8 in Cairns(2004) [6].

In this dissertation, we investigate a nonlinear stochastic interest-rate model that has desirable attributes of positivity of the resulting (short term) interest rate and mean reversion. A version on this model has been used previously in a biological context to model the growth of a population in a crowded stochastic environment (see Chapter 5, Oksendal(2003) [28]). Typically, such models are obtained by adding a (possibly time and/or state dependent) stochastic forcing term. We will show that the stochastic differential equation (SDE) for the population-growth model can also be used as a model for the short rate, and that this model fits nicely into Rogers' potential framework, at least in the case where the parameters are constant. However, another desirable property of interest rate models is the ability to calibrate the models to a specified initial term structure. To incorporate this feature into our model, we will generalize the SDE borrowed from the population growth application by allowing one of the parameters to be non-random function of time. We show that this generalized model still admits a closed form solution (a feature already known for the constant-parameter version of the model), which makes apparent the positivity and mean reversion properties of the solutions. The closed-form solution for our generalized model facilitates simulation via sample-path generation, but it does not give direct information about the distribution of the short rate stochastic process governed by the model. We also apply the well known trinomial-tree discretization process to our model to obtain discrete approximations of its probability distributions. The trinomial-tree process is carried out in such a way as to allow calibration of the model to an initial term structure in a manner similar to that done by Hull and White(1994) [17].

Having obtained the distribution of the interest rate process, we will use it to derive the default probability of the firm value that follows a stochastic process (Acharya and Carpenter(2002) [1]) involving the interest rate process under a risk-neutral measure. The unleveraged firm value is the expected future cash flows discounted at an appropriate rate for an all-equity firm, while the levered firm value is the sum of the unlevered firm value and the gain from leverage due to a tax shield provided by the debt. In recent years, there have been several studies of models of the firm value. Improved models for firm value were developed by Leland and Toft(1996) [24] and Qi(2007)[33] that address optimal capital structure, debt maturity and credit spreads. The importance of the interest rate model studies has been enhanced with the growth of credit derivatives and the credit crisis of 2007 to 2009. Our research will extend the work of Lakshmivarahan et. al.(2012) [22] et al., which derives the probability of default of a firm following a common variety of narrow sense linear models(Arnold(1974) [2], such models being commonly used for credit risk. In our work, we consider both linear models and non-linear models. By transforming the coupled process into independent processes, we derive two methods to obtain the default probability of the firm value V_t: (1) derive the closed form expression of the firm value by Ito's formula and use the first passage time(Pierre Collin-Dufresne and Robert S. Goldstein(2001) [32])to derive the default probability; (2) implement the trinomial tree method for the coupled process, as motivated by Hull-White(1994b) [18]. Here we implement both the Vasicek model and population growth model as the interest rate model and compute the resulting default probabilities for each model. These probabilities are presented over a 20 year period for various values of the correlation ρ of the two underlying Brownian motion processes.

In Chapter 2, we will present the preliminaries. In Chapter 3 we provide a short summary of the potential approach. Chapter 4 presents the derivation of the positive interest model using the potential approach, and the utility of the model for use as a short term interest rate. The calibration of the interest rate model generated by our approach will be presented in Chapter 5. The application of the above models for computing the firm's default probability is explored in Chapter 6. Concluding observations are contained in Chapter 7.

Chapter 2

Preliminaries

This chapter presents several important mathematical theories and financial concepts which will be used in this dissertation. Some selected definitions and results, such as martingale, Ito's formula, and risk-neutral measure will be provided. In the last section, we will also introduce some basic financial theory for the interest rate model. For succinctness, we will omit all proofs, and provide references where appropriate. The more knowledgeable reader may omit this chapter.

2.1 Stochastic Processes and Martingale Theory

This section is mostly taken from Protter(2003) [31] and Oksendal(2003) [28]. We begin with some definitions from basic probability theory and stochastic processes. First, given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we have following definition.

Definition 2.1. Let p be a fixed real number such that $p \geq 1$. Let $f : \Omega \to \mathbb{R}$ be a measurable function. Then f is said to be $\mathbf{p-integrable}$ if

$$\int_\Omega |f(\omega)|^p d\mathbb{P}(w) < \infty. \tag{2.1.1}$$

We denote the set of p-integrable functions, the $L^p-\mathbf{space}$, by $L^p(\Omega, \mathcal{F}, \mathbb{P})$.

Now given an increasing filtration $\{\mathcal{F}_t\}_{t\geq 0}$ of $\mathcal{F}$, i.e. given a family of sub-σ-algebras $\mathcal{F}_t$ of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$, we define a desirable technical property if a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ as follows.

Definition 2.2. A filtered probability space is said to satisfy the "**usual conditions**" if

i $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets of $\{\mathcal{F}\}_{t\geq 0}$, i.e. sets $B \subseteq \Omega$ for which $\exists\ A \in \mathcal{F}$ such that $B \subseteq A$ and $\mathbb{P}(A) = 0$;

ii $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous, i.e. $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$ for all $t \geq 0$.

Definition 2.3. A **stochastic process** X on $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of $\mathbb{R}$-valued or $\mathbb{R}^d$-valued random variables (X_t) indexed by $t \geq 0$. The process is said to be **adapted** if X_t is $\mathcal{F}_t$-measurable for each $t \geq 0$.

Notation: For typographical considerations we may also denote stochastic process random variables by $X(t)$ and filtration σ-algebras by $\mathcal{F}(t)$.

Definition 2.4. Given two stochastic process X and Y, we say that X is a **modification** of Y if $X_t = Y_t$ almost surely (a.s.), for each t. Two processes X and Y are **indistinguishable** if for all t $X_t = Y_t$ a.s. .

Definition 2.5. A random variable $T : \Omega \to [0, \infty)$ is a **stopping time** if the event $\{T \leq t\} \in \mathcal{F}_t$, for every $t \geq 0$.

Henceforth we assume the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfies the usual conditions. Next we will present the definition of martingale.

Definition 2.6. A real-valued, adapted process X_t for $t \geq 0$ is a **martingale** (resp. **supermartingale**, **submartingale**) with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if

(i) $\mathbb{E}^{\mathbb{P}}[|X_t|] < \infty$ for $t \geq 0$,

(ii) if $0 \leq s \leq t$, then $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s] = X_s$, a.s. (resp. $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s] \leq X_s$, resp. $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s] \geq X_s$), where $\mathbb{E}^{\mathbb{P}}$ is the expectation under probability measure $\mathbb{P}$, and $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s]$ is the conditional expectation of X_t given $\mathcal{F}_s$.

As is well known, the expected value of a martingale is constant: $\mathbb{E}^{\mathbb{P}}[X_t] = \mathbb{E}^{\mathbb{P}}[X_0]$ for $t \geq 0$. Here we introduce the Martingale Convergence Theorem from Oksendal(2003) [28].

Theorem 2.7. *(***Doob′s Martingale Convergence Theorem I***) Let X_t be a right continuous supermartingale with the property that*

$$\sup_{t>0} E[X_t^-] < \infty, \tag{2.1.2}$$

where $X_t^- = \max(-X_t, 0)$. Then the positive limit

$$X(\omega) = \lim_{t\to\infty} X_t(\omega) \tag{2.1.3}$$

exists for almost all ω and $E[X^-] < \infty$.

However, note that the convergence need not be in $L^1(P)$. In order to obtain this we need uniform integrability:

Theorem 2.8. *(***Doob′s Martingale Convergence Theorem II***) Let X_t be* *a right continuous supermartingale. Then the following are equivalent:*

(i) $\{X_t\}_{t\geq 0}$ *is* **uniformly integrable**

(ii) There exists $X \in L^1(P)$ *such that* $X_t \to X$ *a.e.* (P) *and* $X_t \to X$ *in* $L^1(P)$*, i.e.* $\int |X_t - X| dP \to 0$ *as* $t \to \infty$*.*

Definition 2.9. An adapted process X is a **local martingale** if there exists a sequence of increasing stopping times, T_n, with $\lim_{n\to\infty} T_n = \infty$ a.s. such that $X_{T\wedge T_n}\mathbf{1}_{T_n>0}$ is a uniformly integrable martingale for each n, where $\mathbf{1}_{T_n>0}$ denotes the indicator function of $T_n > 0$, and $\wedge$ means minimum.

Definition 2.10. The **quadratic variation** of X, $\langle X \rangle$, is a process defined as follows: let X be a continuous **local martingale** with respect to $\{\mathcal{F}_t\}_{t\geq 0}$. We denote by $\langle X \rangle$ the unique continuous, increasing and adapted process with $\langle X \rangle_0 = 0$ such that the process $X^2 - \langle X \rangle$ is a continuous local martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$.

Definition 2.11. An adapted stochastic process X has **finite variation** if for each $t \geq 0$

$$\mathrm{Var}_{[0,t]}(X) := \sup_{\Delta} \sum_{i=1}^{n(\Delta)} |X_{t_i} - X_{t_{i-1}}| < \infty \quad \mathbb{P} - a.s. \tag{2.1.4}$$

where $\Delta = 0 = t_0 < t_1 < ... < t_{n(\Delta)} = t$ is a partition of the interval [0,t].

Definition 2.12. A stochastic process $X = (X_t)_{t\geq 0}$ is called **càdlàg** if it almost surely has sample paths which are right-continuous with left limits. A stochastic process $X = (X_t)_{t\geq 0}$ is called **càglàd** if it almost surely has sample paths which are left-continuous with right limits.

Definition 2.13. A real-valued, continuous, adapted process X is a **semi-**

martingale if it admits the decomposition[31]

$$X_t = X_0 + M_t + A_t \tag{2.1.5}$$

where X_0 is an $\mathcal{F}_0$-measurable random variable, M is a local martingale with $M_0 = 0$ and A is an adapted *càdlàg* process whose almost all sample paths are of finite variation, with $A_0 = 0$.

Definition 2.14. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed position number, and let $\{\mathcal{F}_t\}_{t \geq 0}$ be the filtration of $\mathcal{F}$. We call an adapted stochastic process $X(t), 0 \leq t \leq T$ a **Markov process** if for all $0 \leq s \leq t \leq T$ and for every nonnegative, Borel-measurable function f, there is another Borel-measurable function g such that

$$E[f(X_t)|\mathcal{F}_s] = g(X_s) \tag{2.1.6}$$

This concludes our basic definitions of martingale, supermartingale, submartingale, local martingale, semimartingale and Markov process. For the further properties of martingales, please refer to Protter(2003) [31].

2.2 Ito's Formula, Change of Measure and Some Basic Theory for Interest Rate Models

This section is mostly taken from Shreve(2004) [38] and Oksendal(2003) [28]. First, we introduce Brownian motion and Ito processes.

Definition 2.15. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there $W(t)$ is a stochastic process with continuous sample paths. Then $W(t), t \geq$

0, is a **Brownian motion** if $W(0) = 0$ and if for all $0 = t_0 < t_1 < ... < t_m$ the increments

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), ..., W(t_m) - W(t_{m-1}) \tag{2.2.1}$$

are independent and each of these increments is normally distributed with

$$E[W(t_{i+1} - W(t_i)] = 0, \tag{2.2.2}$$

$$\text{Var}[W(t_{i+1} - W(t_i)] = t_{i+1} - t_i \tag{2.2.3}$$

An **associated filtration** for W_t is a filtration $\{\mathcal{F}_t\}_{t\geq 0}$, which satisfies the following conditions:

(i) Each W_t is $\mathcal{F}_t$ measurable.

(ii) $s < t \Rightarrow W_t - W_s$ is independent of $\mathcal{F}_s$.

Proposition 2.16. *Given a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ *and a Brownian motion* $(W_t)_{t\geq 0}$ *w.r.t* $\{\mathcal{F}_\sqcup\}_{t\geq 0}$. *Assume the stochastic process* $\Delta(t), t \geq 0$ *is adapted to* $\{\mathcal{F}_\sqcup\}_{t\geq 0}$ *and satisfies the square-integrability condition*

$$\mathbb{E}^{\mathbb{P}}[\int_0^T \Delta(t)^2 dt] < \infty. \tag{2.2.4}$$

Then there is a stochastic process $I(t) = \int_0^t \Delta(u) dW(u)$ *called the* **Ito integral process** *of* Δ *that has the following properties:*

*(i) (***Continuity***) As a function of the upper limit of integration t, the paths of* $I(t)$ *are continuous.*

*(ii) (***Adaptivity***) For each t,* $I(t)$ *is* $\mathcal{F}(t)$*-measurable.*

*(iii) (***Linearity***) If* $I(t) = \int_0^t \Delta(u) dW(u)$ *and* $J(t) = \int_0^t \Gamma(u) dW(u)$*, then* $I(t)\pm$

$J(t) = \int_0^t(\Delta(u)+\Gamma(u))dW(u)$, and for every constant c, $cI(t) = \int_0^t c\Delta(u)dW(u)$.

*(iv)(**Martingale**) $I(t)$ is a martingale.*

*(v) (**Ito isometry**) $\mathbb{E}^{\mathbb{P}}[I^2(t)] = \mathbb{E}^{\mathbb{P}}[\int_0^t \Delta(u)^2du]$.*

*(vi) (**Quadratic variation**) $\langle I, I\rangle(t) = \int_0^t \Delta(u)^2du$.*

Definition 2.17. A d-dimensional Brownian motion is a process

$$W(t) = (W_1(t), W_2(t), ..., W_d(t))$$

with the following properties.

(i) Each $W_i(t)$ is a one-dimensional Brownian motion.

(ii) If $i \neq j$, then the processes $W_i(t)$ and $W_j(t)$ are independent.

Associated with a d-dimensional Brownian motion, we have a filtration $\{\mathcal{F}_{\sqcup}\}_{t\geq 0}$, such that the following holds.

(iii) (**Information accumulates**) For $0 \leq s < t$, every set in $\mathcal{F}_s$ is also in $\mathcal{F}_t$.

(iv)(**Adaptivity**) For each $t \geq 0$, the random vector $W(t)$ is $\mathcal{F}_t$-measurable.

(v) (**Independence of future increments**) For $0 \leq t < u$, the vector of increments $W(u) - W(t)$ is independent of $\mathcal{F}_t$.

Give these definitions and properties, we now introduce the notion of an Ito process.

Definition 2.18. Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ and a d-dimension Brownian motion $(W_t)_{t\geq 0}$ w.r.t $\mathcal{F}_t$, an adapted continuous process X is called an **Ito process** if it admits a representation

$$X_t = X_0 + \int_0^t \alpha_u du + \int_0^t \beta_u dW_u \tag{2.2.5}$$

for all $t \geq 0$, where α is a vector and β is a matrix, and both are adapted processes and satisfy suitable conditions. An Ito process is usually also represented as a stochastic differential of the form

$$dX_t = \alpha_t dt + \beta_t dW_t \tag{2.2.6}$$

Theorem 2.19. (**Ito's Formula**) *Let $X(t), t \geq 0$ be a d-dimensional Ito process described in equation 2.2.6, while X_i is the component processes of X for each i, and let $f(X_t)$ be a function: $\mathbb{R}^d \to \mathbb{R}$ for which the partial derivatives f_{x_i}, and $f_{x_i x_j}$ are defined and continuous. Then for every $T \geq 0$,*

$$df(X_t) = \sum_{i=1}^{k} f_{x_i} dX_i(t) + \frac{1}{2} \sum_{i,j=1}^{k} f_{x_i x_j} d\langle X_i, X_j \rangle(t) \tag{2.2.7}$$

More specifically, if for $i = 1, 2, ..., k$, each X_i is an Ito process, then

$$df(X_t) = \sum_{i=1}^{k} f_{x_i} \alpha_i(t) dt + \sum_{i=1}^{k} f_{x_i} \beta_i(t) dW_t + \frac{1}{2} \sum_{i,j=1}^{k} f_{x_i x_j} \beta_i(t) \beta_j(t) dt \tag{2.2.8}$$

Definition 2.20. A probability measure $\mathbb{Q}$ defined on $(\Omega, \mathcal{F})$ is **absolutely continuous** with respect to $\mathbb{P}$, and denoted by $\mathbb{Q} \ll \mathbb{P}$, if for all $A \in \mathcal{F}$,

$$\mathbb{P}(A) = 0 \Rightarrow \mathbb{Q} = 0. \tag{2.2.9}$$

If $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$, then we say that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures and denote this by $\mathbb{P} \sim \mathbb{Q}$.

Theorem 2.21. (**Radon – Nikodym Theorem**) *Let $\mathbb{Q}$ be a probability measure defined on $(\Omega, \mathcal{F})$. If $\mathbb{Q} \ll \mathbb{P}$, then there exists an a.s. unique random* *variable* $\rho \geq 0$ *which satisfies* $\mathbb{E}^{\mathbb{P}}[\rho] = 1$*, such that for* $A \in \mathcal{F}$*,*

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\rho \mathbf{1}_A] \tag{2.2.10}$$

where $\mathbf{1}_A$ *is the indicator of* A*. If* $\mathbb{P} \sim \mathbb{Q}$*, then* $\rho > 0$ *a.s. .*

Definition 2.22. Let $\mathbb{Q}$ be a probability measure defined on $(\Omega, \mathcal{F})$, such that $\mathbb{Q} \ll \mathbb{P}$. The random variable ρ from the above theorem is known as the **Radon-Nikodym derivative of** $\mathbb{Q}$ with respect to $\mathbb{P}$ and is denoted by $\rho = \frac{d\mathbb{Q}}{d\mathbb{P}}$.

Theorem 2.23. *(***Girsanov Theorem***) Let* T *be a fixed positive time and suppose that* $\Theta = (\Theta_1, ..., \Theta_d)$ *is an d-dimensional adapted process. Define*

$$Z(t) = e^{-\int_0^t \Theta(u)\cdot dW(u) - \frac{1}{2}\int_0^t ||\Theta(u)||^2 du} \tag{2.2.11}$$

$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u) du, \tag{2.2.12}$$

and assume that

$$\mathbb{E}\int_0^T ||\Theta(u)||^2 Z^2(u) du < \infty. \tag{2.2.13}$$

Set $Z = Z(T)$*. Then* $\mathbb{E}Z = 1$*, and under the probability measure* $\tilde{\mathbb{P}}$ *given by*

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \quad for \quad all \quad A \in \mathcal{F} \tag{2.2.14}$$

the process $\tilde{W}(t)$ *is a d-dimensional Brownian motion.*

Theorem 2.24. *(***Exponential martingale***) Let* $W(t), t \geq 0$*, be a Brownian motion with a filtration* $\mathcal{F}(t), t \geq 0$*, and let* σ *be a positive constant, The process* $Z_t, t \geq 0$*, given by*

$$Z(t) = e^{\sigma W(t) - 1/2\sigma^2 t}. \tag{2.2.15}$$

is a martingale.

Theorem 2.25. *(***Reflection principle***) For each Brownian motion path $W(t)$ that reaches a level m prior to time t but is at a level w below m at time t, there is a "reflected path" that is at level $2m-w$ at time t. This leads to the* **reflection equation**

$$P[\max_{0\leq s\leq t} W(s) \geq m, W(t) \leq w] = P[W(t) \geq 2m-w], \quad w \leq m, m > 0. \tag{2.2.16}$$

Moreover, the joint density of $(\max_{0\leq s\leq t} W(s), W(t))$ is

$$f(t,m,w) = \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}, \quad w \leq m, m > 0 \tag{2.2.17}$$

Let $W(t), 0 \leq t \leq T$, be a Brownian motion on a filtered probability space $(\Omega, \mathcal{F}), \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}$. Consider a stock price process whose differential is

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t), 0 \leq t \leq T \tag{2.2.18}$$

where the mean rate of return $\alpha(t)$ and the volatility $\sigma(t)$ are allowed to be adapted processes. Define the **discount process** (also known as the **money market account** process) by

$$D(t) = e^{-\int_0^t R(s)ds}, \quad i.e. \quad dD_t = R_t D_t dt \tag{2.2.19}$$

where $R(s)$ is an adapted interest rate process. Then, by using Ito's product rule, we can express the discounted stock price process in differential form:

$$d(D(t)S(t)) = \sigma(t)D(t)S(t)[\Theta(t)dt + dW(t)] \tag{2.2.20}$$

where $\Theta(t)$ is the **market price of risk** defined by

$$\Theta(t) = \frac{\alpha(t) - R(t)}{\sigma(t)} \tag{2.2.21}$$

Now introduce the probability measure $\tilde{\mathbb{P}}$ defined in Girsanov's Theorem, where $\Theta(t)$ is as defined in (2.2.21). We call this probability measure $\tilde{\mathbb{P}}$ the **risk neutral measure** because it is equivalent to the original measure $\mathbb{P}$ and it renders the discounted stock price $D(t)S(t)$ into a martingale. The risk-neutral measure is heavily used in the pricing of financial derivatives due to the fundamental theorem of asset pricing.

Definition 2.26. An **arbitrage** is a portfolio value process $X(t)$ satisfying $X(0) = 0$ and also satisfying for some time $T > 0$

$$\mathbb{P}\{X(T) \geq 0\} = 1, \quad \mathbb{P}\{X(T) > 0\} > 0.$$

An **arbitrage** is a way of trading so that one starts with zero capital and at some time point T is sure not to have lost money and furthermore has a positive probability to earn money.

Definition 2.27. A **complete market** is one in which the complete set of possible gambles on future states-of-the-world can be constructed with existing assets without transaction costs.

In other words, for a complete market, all cash flows for a trading strategy can be replicated using a similar synthetic trading strategy.

Theorem 2.28. *If a market model has a risk-neutral probability measure, then an arbitrage does not exist. Consider a market model that has a risk-neutral* *probability measure. The model is complete if and only if the risk-neutral probability measure is unique.*

We end this section with some terminology from interest rate theory taken from Cairns(2004) [6].

Definition 2.29. A T-maturity **zero – coupon bond** is a contract that guarantees its holder the payment of one unit of currency at time T, with no intermediate payments. The contract value at time $t < T$, is denoted by $P(t, T)$. Clearly, $P(T, T) = 1$ for all T.

Definition 2.30. Let $0 < T < S$. Then the **forward rate** at time t with $t < T < S$(continuous compounding) which applies between times T and S is defined as

$$F(t, T, S) = \frac{1}{S - T} \log \frac{P(t, T)}{P(t, S)}. \tag{2.2.22}$$

The **instantaneous forward rate** at time t

$$f(t, T) = \lim_{S \to T^+} \frac{1}{S - T} \log \frac{P(t, T)}{P(t, S)} = -\frac{\partial \log P(t, T)}{\partial T}. \tag{2.2.23}$$

Equivalently, we have

$$P(t, T) = e^{-\int_t^T f(t,u)du}. \tag{2.2.24}$$

Definition 2.31. The **short rate** at time t is defined as the instantaneous forward rate at time t for the maturity t, i.e. $r_t := f(t, t)$

The short rate model can be classified as narrow sense linear, general linear, or nonlinear models. The dynamics of the short term interest rate models under the risk neutral measure, are modeled by a (scalar) SDE:

$$dr_t = \alpha(r_t, t)dt + \sigma(r_t, t)dW_t, \tag{2.2.25}$$

where the instantaneous drift, $\alpha(r_t, t)$, is a smooth function and the volatility,$\sigma(r_t, t)$, denotes the volatility term (Cairns(2004) [6]).

The interest rate models can be classified into two broad classes: single factor models and multi-factor models. In this dissertation we focus on the single factor models which are further subdivided into linear and nonlinear models. The model is called a nonlinear model if either $\alpha(r_t, t)$ or $\sigma(r_t, t)$ are nonlinear functions of the short rate r_t. Furthermore, following Arnold (1974), the linear models can be subdivided into two subclasses: narrow sense linear models if

$$\alpha(r_t, t) = a_1(t)r_t + a_2(t), \quad \sigma(r_t, t) = \sigma_r(t)$$

and general linear models if $\alpha(r_t, t)$ is of the form as above and

$$\sigma(r_t, t) = b_1(t)r_t + b_2(t),$$

where $a_i(t), b_i(t), i = 1, 2$ and $\sigma_r(t)$ are smooth functions of time t. Here are some examples of the above models in the table (2.1).

Table 2.1: Short rate model samples

	Narrow sense linear models
Merton(1973) [26]	$dr_t = \theta dt + \sigma_r dW_r$
Vasicek(1977) [39]	$dr_t = (\theta - cr_t)dt + \sigma_r dW_r$
Ho-Lee(1986) [15]	$dr_t = \theta(t)dt + \sigma_r dW_r$
Hull and White(1990) [16]	$dr_t = (\theta(t) - cr_t)dt + \sigma_r dW_r$
Generalized Hull and White [19]	$dr_t = (\theta(t) - c(t)r_t)dt + \sigma_r dW_r$
	General linear models
Brennan-Schwartz (1979) [5]	$dr_t = (\theta(t) - cr_t)dt + \sigma_r r_t dW_r$
	Non-linear models
Cox-Ingersoll-Ross(1985) [7]	$dr_t = \theta(\mu - r_t)dt + \sigma_r \sqrt{r_t} dW_r$

The narrow sense linear models define Gaussian processes which allow negative interest rates. The general linear model of Brennan and Schwartz (1979) gives rise to lognormal processes. The nonlinear models of Cox, Ingersoll and Ross (1985) define the so called square root processes. We will discuss these models in the following chapters.

2.3 Transition Function and Kolmogorov's Backward Equation

For this section, we inroduce to use the transition probability and Kolmogorov's backward equation(Arnold(1974) [2]).

First, let's review the definitions of transition probability and Chapman-Kolmogorov equation.

Consider a filtered probability space $(\Omega, \{\mathcal{F}\}_{t\geq 0}, \mathbb{P})$. Let X_t, for $t_0 \leq t \leq T$, denote a Markov process,

Definition 2.32. A function $P(s, x, t, B)$ is called a **transition probability** (**transition function**) of the markov process X_t if it has the following properties:

(a) For fixed $s \leq t$ and $B \in \mathcal{B}^d$, we have with probability 1 that $P(s, X_s, t, B) = P(X_t \in B|X_s)$, where $\mathcal{B}^d$ the σ-algebra generated by the Borel sets in R^d.

(b) $P(s, x, t, \cdot)$ is a probability measure on $\mathcal{B}^d$ for fixed $s \leq t$ and $x \in R^d$.

(c) $P(s, \cdot, t, B)$ is $\mathcal{B}^d$-measurable for fixed $s \leq t$ and $B \in \mathcal{B}^d$.

(d) For $t_0 \leq s \leq u \leq t \leq T$ and $B \in \mathcal{B}^d$ and for all $x \in R^d$ with the possible exception of a set $N \subset R^d$ such that $P[X_s \in N] = 0$, we have the **Chapman-Kolmogorov equation**.

$$P(s, x, t, B) = \int_{R^d} P(u, y, t, B) P(s, x, u, dy). \tag{2.3.1}$$

(e) For all $s \in [t_0, T]$ and $B \in \mathcal{B}^d$, we have

$$P(s, x, s, B) = I_B(x) = \begin{cases} 1 & for \quad x \in B, \\ 0 & for \quad x \notin B. \end{cases}$$

Now, we introduce a family of operators associated with the Markov processes in terms of which many important definitions and results follow. First, denote by $B(R^d)$ the space of bounded measurable scalar functions defined on R^d equipped with the norm $\|g\| = \sup_{x \in R^d} |g(x)|$.

Definition 2.33. Given a Markov process X_t, its family of **transition operators**, $(T_{s,t})_{0 \leq s \leq t}$ on the space $B(R^d)$, is defined by

$$(T_{s,t}f)(x) = \mathbb{E}^{\mathbb{P}}[f(X_t)|X_s = x] = \int_{R^d} f(y)P(s, x, t, dy) \tag{2.3.2}$$

for each $f \in B(R^d)$ and $x \in R^d$. For the **homogeneous** case, by setting $s = 0$, we denote the transition operator by T_t as follow

$$(T_t f)(x) = \mathbb{E}^{\mathbb{P}}[f(X_t)|X_0 = x] = \int_{R^d} f(y)P(0, x, t, dy). \tag{2.3.3}$$

Definition 2.34. The **infinitesimal operator(generator)** A of a homogeneous Markov process X_t, for $0 \leq t \leq T$, is defined by

$$Af(x) = \lim_{t \to 0} \frac{T_t f(x) - f(x)}{t}, \quad f \in B(R^d), \tag{2.3.4}$$

where the limit is uniform with respect to x.

Definition 2.35. A Markov process X_t, for $0 \leq t \leq T$, with values in R^d and almost surely continuous sample functions is called a **diffusion process** if its transition probability $P(s, x, t, B)$ satisfies the following three conditions for every $s \in [0, T)$, $x \in R^d$, and $\epsilon > 0$:

(a) $\lim_{t \to s} \frac{1}{t-s} \int_{|y-x|>\epsilon} P(s, x, t, dy) = 0$;

(b) there exists an R^d-valued function $f(s, x)$ such that

$$\lim_{t \to s} \frac{1}{t-s} \int_{|y-x|\le\epsilon} (y-x) P(s, x, t, dy) = f(s, x);$$

(c) there exists a $d \times d$ matrix-valued function $B(s, x)$ such that

$$\lim_{t \to s} \frac{1}{t-s} \int_{|y-x|\le\epsilon} (y-x)(y-x)' P(s, x, t, dy) = B(s, x).$$

The functions f and B are called the coefficients of the diffusion process. Also, f is called the **drift vector**, and B is called the **diffusion matrix**.

Using Kolmogorov's criterion and some elementary reasoning, we can safely choose R^d as the region of integration in conditions (b) and (c). So, we have the first moment of the increment $X_t - X_s$ under the condition $X_s = x$ as $t \to s$: $E_{s,x}(X_t - X_s) = f(s, x)(t-s) + o(t-s)$ and the covariance matrix of $X_t - X_s$ with respect to the probability $P(s, x, t, \cdot)$: $Cov_{s,x}(X_t - X_s) = B(s, x)(t-s) + o(t-s)$. Since we are only concerned with the distributions, we can resort to a first-order approximation (in $t - s$) and assert that X_t satisfies the following stochastic differential equation:

$$dX_t = f(t, X_t)dt + G(t, X_t)dW_t \tag{2.3.5}$$

where $GG' = B$ and W_t is the Brownian motion process.

Before we introduce the backward equation, we assign the following second order differential operator for each diffusion process with coefficients f and $B = (b_{i,j})_{i,j=1}^{n}$:

$$\mathcal{D} \equiv \sum_{i=1}^{d} f_i(s,x)\frac{\partial}{\partial x_i} + 1/2\sum_{i=1}^{d}\sum_{j=1}^{d} b_{i,j}(s,x)\frac{\partial^2}{\partial x_i \partial x_j}. \tag{2.3.6}$$

By means of Taylor expansion of $f(s+t,y)$ about (s,x) under the assumption that f is defined and bounded on $[t_0,T] \times R^d$ and is twice differentiable with respect to s, one can show that $A = \frac{\partial}{\partial s} + \mathcal{D}$. Therefore, the diffusion process is uniquely determined by f and B. Now let us look at Kolmogorov's backward equation.

Theorem 2.36. *Let X_t, for $t_0 \leq t \leq T$, denote a d-dimensional diffusion process with continuous coefficients $f(s,x)$ and $B(s,x)$. The limit relations in definition 2.35 hold uniformly in $s \in [t_0,T]$. Let f(x) denote a continuous bounded scalar value function, fix $t > t_0$ and for $s < t$ and $x \in R^d$ assume that the function*

$$u(s,x) = \mathbb{E}_{s,x} f(X_t) = \int_{R^d} f(y)P(s,x,t,dy)$$

*is continuous and bounded, as well as its derivatives $\frac{\partial u}{\partial x_i}$, and $\frac{\partial^2 u}{\partial x_i \partial x_j}$ for $1 \leq i,j \leq d$. Then $u(s,x)$ is differentiable with respect to s and satisfies **Kolmogorov's backward equation***

$$\frac{\partial u}{\partial s} + \mathcal{D}u = 0, \tag{2.3.7}$$

where $\mathcal{D}u$ is the operator defined as above. with the end condition $\lim_{s\to t} u(s,t) = f(x)$.

Theorem 2.37. *Suppose that the assumption of Theorem (2.36) regarding X_t holds. If $P(s,x,t,\cdot)$ has a density $p(s,x,t,y)$ that is continuous with respect* *to s and if the derivatives $\partial p/\partial x_i$ and $\partial^2 p/\partial x_i \partial x_j$ exist and are continuous with respect to s, then p is called the **fundamental solution** of the backward equation*

$$\frac{\partial p}{\partial s} + \mathcal{D}p = 0, \tag{2.3.8}$$

that is, it satisfies the end condition

$$\lim_{s \to t} p(s, x, t, y) = \delta(x - y), \tag{2.3.9}$$

where δ is Dirac's delta function.

We will need these results to derive the corporate default probability in Chapter 6. This concludes the preliminary results needed for this dissertation. More details and proofs can be found in the references.

Chapter 3

The Potential Approach

The earliest published paper using this approach appears to be Constantinides (1992) [8]. Subsequently, variations of this approach were developed by Rogers (1997) [35], Rutkowski(1997) [36], and Flesaker and Hughston(1996) [9].

3.1 Potentials

First, we discuss the pricing kernel and the construction of the positive interest rate term structure. Consider the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leqq t \leqq T}, \mathbb{P})$, where $\{\mathcal{F}_t\}_{0 \leqq t \leqq T}$ is the natural filtration generated by the Brownian motion B. Let $P(t,T)$ denote the value at time t of a zero-coupon bond with maturity T. We have discussed this in the previous chapter in equation (2.2.24). Here based on the bond pricing formula and arbitrage pricing paradigm, the term structure process is related to the short-rate process $(r_t)_{t \geq 0}$ by the formula:

$$P(t,T) = \mathbb{E}[e^{-\int_t^T r_s ds} | \mathcal{F}_t] \tag{3.1.1}$$

where the underlying probability measure is assumed to be risk neutral.

Definition 3.1. (Meyer(1966) [25]) A positive right continuous supermartingale tending to 0 in expectation as $t \to \infty$ is called a **potential**.

Let X_t, for $0 \leq t \leq T$, denote a d-dimensional Markov process and $P(s, x, t, B)$ be its **transition probability**(**transition function**), where $s \leq t$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}^d$, the σ-algebra generated by the Borel sets in $\mathbb{R}^d$.

Definition 3.2. (Arnold(1974) [2]) The family of **transition operators**, $(T_{s,t})_{0 \leq s \leq t}$ on the space $B(\mathbb{R}^d)$, associated to the Markov process X_t is given by

$$(T_{s,t}f)(x) = \mathbb{E}^P[f(X_t)|X_s = x] = \int_{R^d} f(y)P(s, x, t, dy) \tag{3.1.2}$$

for each $f \in B(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, where $B(\mathbb{R}^d)$ is the linear space of all bounded Borel measurable functions from $\mathbb{R}^d$ to $\mathbb{R}$. For the homogeneous case, where $T_{s,t}$ depends only on the difference $t - s$, we denote the transition operator by T_t. The **infinitesimal operator(generator)** $\mathfrak{g}$ of a homogeneous Markov process X_t, for $0 \leq t \leq T$, is defined by

$$\mathfrak{g}f(x) = \lim_{t \to 0^+} \frac{T_t f(x) - f(x)}{t}, \quad f \in B(\mathbb{R}^d), \tag{3.1.3}$$

where the limit is uniform with respect to x.

3.2 Pricing Kernels and the Rogers' Framework

Given a d-dimensional homogeneous Markov process X_t, given a positive Borel measurable function $g : \mathbb{R}^d \to [0, \infty)$, and given $\lambda > 0$, we define an increasing process A_t by $A_t = \int_0^t e^{-\lambda s} g(X_s)ds$ and set up a **pricing kernel**:

$$\varsigma_t = \mathbb{E}_t^P[A_\infty] - A_t \quad for \quad t \geq 0 \tag{3.2.1}$$

where $A_\infty = \lim_{t\to\infty} A_t$. For $\lambda > 0$, we define the **resolvent** operator as follows:

$$R_\lambda = (\lambda - \mathfrak{g})^{-1} = \int_0^\infty e^{-\lambda t} T_t dt$$

Then it can be proved that the pricing kernel is of the form (Rogers(1997) [35])

$$\varsigma_t = e^{-\lambda t} \frac{R_\lambda g(X_t)}{R_\lambda g(X_0)} \tag{3.2.2}$$

while it is also easy to prove that ς_t is a potential. Here $R_\lambda g(X_0)$ acts as a normalizing factor to ensure that $\varsigma_0 = 1$. Since this factor is trivial, we discount it hereafter for convenience. Now, using the method in Glasserman(2001) [13], the short rate can be derived by the dynamics of the pricing kernels as:

$$r_t = \frac{1}{\varsigma_t} \mathbb{E}_t^P \left[\frac{\partial A_T}{\partial T} \right] |_{T=t} = \frac{1}{\varsigma_t} \left[\frac{\partial A_t}{\partial t} \right] = \frac{e^{-\lambda t} g(X_t)}{\varsigma_t}. \tag{3.2.3}$$

From this, we obtain the short rate formula

$$r_t = \frac{g(X_t)}{R_\lambda g(X_t)}. \tag{3.2.4}$$

By substitution of ς_t into the bond pricing formula, we also have

$$P(t,T) = \frac{\mathbb{E}_t^P[e^{-\lambda(T-t)} R_\lambda g(X_T)]}{R_\lambda g(X_t)}. \tag{3.2.5}$$

Here is the algorithm to derive the interest rate models by the Rogers framework(1997) [35]:

Step 1: Choose a Markov process $(X_t)_{t\geq 0}$;

Step 2: Choose a positive twice differentiable function f defined in $B(\mathbb{R}^d)$, select $\lambda > 0$ and define g by

$$g = (\lambda - \mathfrak{g})f.$$

where $\mathfrak{g}$ is the infinitesimal generator of the process $(X_t)_{t\geq 0}$ Here we attempt to choose λ to make g positive. Then we get $f := R_\lambda g$;

Step 3: Compute the pricing kernel as follows:

$$\varsigma_t = e^{-\lambda t}\frac{f(X_t)}{f(X_0)};$$

Step 4: Substitution of the previous expression into the bond pricing formula yields

$$P(t,T) = e^{-\lambda(T-t)}\frac{\mathbb{E}_t^P[f(X_T)]}{f(X_t)} \tag{3.2.6}$$

and the short rate by

$$r_t = \frac{(\lambda - \mathfrak{g})f(X_t)}{f(X_t)} = \frac{g(X_t)}{f(X_t)}. \tag{3.2.7}$$

Rogers(1997) [35] applied this method to several classical Markov processes X_t, including the Ornstein-Uhlenbeck process, the mean-reverting Bessel process, and the univariate quadratic model. Also he examined various choices for the function f including affine functions, quadratic functions, exponential linear functions, exponential quadratic functions, and hyperbolic functions. In the next section we will use the Rogers framework to investigate another class of models with a nonlinear drift process that have mean reversion properties, which thereby have some intuitive appeal as interest rate models.

Chapter 4

Derivation of a Class of Nonlinear Interest-Rate Models

4.1 A Simple Example

4.1.1 Solution of the Process and the General Form of the Infinitesimal Generator

Consider the Markov process

$$dX_t = \mu dt + \alpha X_t dB_t \tag{4.1.1}$$

where α is a real constant, $\mu \in \mathbb{R}^d$, $X_t \in \mathbb{R}^d$ and $B_t \in \mathbb{R}$ is standard Brownian motion. (Note that this process is not among the examples presented by Rogers).

Using the integrating factor $F_t = e^{-\alpha B_t + \frac{1}{2}\alpha^2 t}$ and Ito's formula, we obtain the closed form solution

$$X_t = \int_0^t \mu e^{-\alpha(B_t - B_s) + \frac{1}{2}\alpha^2(t-s)} ds + \frac{X_0}{F_t} = \int_0^t \mu \frac{F_s}{F_t} ds + \frac{X_0}{F_t}. \tag{4.1.2}$$

We can see that in the scalar case ($d = 1$) this solution is already positive if $X_0 > 0$ and $\mu > 0$. Now, we will use Rogers' method which will not only give the expression of the pricing kernel, but also the bond price. To implement the algorithm we discussed in the previous section, we first compute the infinitesimal generator $\mathfrak{g}$ for this process. For a general Ito process $dX_t = b(X_t)dt + \sigma(X_t)dB_t$, where $b(X_t)$ is a d-vector function of X_t and $\sigma(X_t)$ is a $d \times m$ matrix, B_t is a m-dimensional Brownian motion, and $X_t \in \mathbb{R}^d$, the formula for $\mathfrak{g}$ is as follows:

$$(\mathfrak{g}f)(x) = \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} (\sigma(x)\sigma(x)^T)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}$$

for any bounded $\mathcal{C}^2$ function f.

So, for the simple process(4.1.1), we derive the infinitesimal generator $\mathfrak{g}$ as

$$(\mathfrak{g}f)(x) = \mu \cdot \nabla f + \frac{1}{2} tr[\alpha^2 X_t X_t^T \cdot Hf] \tag{4.1.3}$$

where ∇f is the gradient vector of f, Hf is the Hessian matrix of f and $tr(\cdot)$ denotes the trace of the matrix.

4.1.2 Implementation of Different Functions f of the Rogers' Method for the Process(4.1.1)

Supposing the dynamics of $(X_t)_{t \geq 0}$ are given by equation (4.1.1), we can implement the algorithm with different choices of the function f: exponential quadratic function, exponential linear function, hyperbolic cosine function, linear function, and quadratic function. The results provide a useful comparison with Rogers' work(1997) [35].

First consider an exponential quadratic function function f of the form

$$f(X_t) = e^{\frac{1}{2}(X_t - c)^T Q (X_t - c)} \tag{4.1.4}$$

where c is a $d \times 1$ vector of real numbers and Q is a $d \times d$ positive-definite, diagonal matrix. In order to derive the r_t, recalling that $r_t = \frac{(\lambda - \mathfrak{g}) f(X_t)}{f(X_t)} = \frac{g(X_t)}{f(X_t)}$ for some λ, we need to derive the infinitesimal generator $\mathfrak{g}$. So, first we compute ∇f and Hf.

$$\nabla f = f(X_t) Q (X_t - c)$$

$$Hf = f(X_t) \begin{bmatrix} q_{11} + q_{11}^2 (X_1 - c_1)^2 & \dots & q_{11} q_{dd} (X_1 - c_1)(X_d - c_d) \\ \vdots & \ddots & \vdots \\ q_{11} q_{dd} (X_1 - c_1)(X_d - c_d) & \dots & q_{dd} + q_{dd}^2 (X_d - c_d)^2 \end{bmatrix}.$$

This leads to the infinitesimal generator:

$$\begin{aligned} (\mathfrak{g} f)(X_t) &= \mu \cdot \nabla f + \frac{1}{2} tr[(\alpha^2 X_t X_t^T) \cdot Hf] \\ &= f(X_t)[\mu^T Q (X_t - c) + \frac{1}{2} \alpha^2 \cdot X_t^T (Q + Q(X_t - c)(X_t - c)^T Q) X_t] \end{aligned}$$

and the interest rate:

$$r_t = \lambda - \mu^T Q (X_t - c) - \frac{1}{2} \alpha^2 \cdot X_t^T (Q + Q(X_t - c)(X_t - c)^T Q) X_t. \tag{4.1.5}$$

Rogers's example with Ornstein-Uhlenbeck (OU) process $dX_t = k(\theta - X_t)dt +$

dW_t, and the same function (4.1.4) yields the interest rate

$$r_t = \lambda - \frac{1}{2}tr(Q) - (\theta - X_t)^T k^T Q(X_t - c) - \frac{1}{2}(X_t - c)^T Q^2 (X_t - c)$$

In our cases, we can see that r_t as given in (4.1.5) is a fourth order function of X_t, and we could try to choose the parameters such that r_t is positive. Since the OU process is a Gaussian process, the positivity of X_t is not assured, but it is useful to have the quadratic term to make r_t positive, like in this simple case. But due to the fourth order appearance of X_t in (4.1.5) and the complicated computation, we will not consider this case any further.

For the exponential linear function f of the form

$$f(X_t) = e^{\gamma^T X_t} \tag{4.1.6}$$

where γ is a $d \times 1$ vector, we have the infinitesimal generator:

$$(\mathfrak{g} f)(X_t) = f(X_t)[\mu^T \gamma + \frac{1}{2}(\alpha X_t^T \gamma)^2]$$

and the interest rate:

$$r_t = \lambda - \mu^T \gamma - \frac{1}{2}(\alpha X_t^T \gamma)^2. \tag{4.1.7}$$

By comparison with the special Bessel process $dX_t = (a + KX_t)dt + 2\sqrt{diag(X_t)}dW_t$ in Rogers(1997), his result is

$$r_t = \lambda - (a + KX_t)^T \gamma - 2\sum_{i=1}^{d} (\gamma_i)^2 X_t^{(i)}.$$

From here we can see that Roger's result is a linear function of X_t. Both

Rogers' Bessel process and our processes are positive processes, with simple representations of r_t, so it is convenient to run numerical experiments. For our case, the r_t derived from exponential linear function is a quadratic function of X_t, which is also a very good fit with the positive interest rate if the parameters are chosen wisely.

For the hyperbolic cosine function f of the form

$$f(X_t) = \cosh(\gamma(X_t + c) \tag{4.1.8}$$

where γ is a $1 \times d$ vector and c is a $d \times 1$ vector, a similar computation yields the infinitesimal generator:

$$(\mathfrak{g}f)(X_t) = (\gamma\mu)\sinh(\gamma(X_t + c)) + \frac{1}{2}\alpha^2(\gamma X_t)^2 \cosh(\gamma(X_t + c))$$

and the interest rate:

$$r_t = \lambda - \frac{1}{2}\alpha^2(\gamma X_t)^2 - (\gamma\mu)\tanh(\gamma(X_t + c)). \tag{4.1.9}$$

We compare this interest rate with the OU process in with Rogers(1997) and the hyperbolic cosine function

$$r_t = \lambda - \frac{1}{2}\gamma^2 - \gamma k X_t \tanh(\gamma(X_t + c)).$$

Consider the special case of $c = 0$ and $\lambda = \frac{1}{2}\gamma^2$ for Rogers' example: $r_t = \gamma k X_t \tanh(\gamma(X_t))$, which is a combination of a Gaussian model and a squared-Gaussian model with X_t taking the extreme value. For our case, setting $c = 0$ and noting that $\tanh(\gamma X_t)$ is bounded by -1 and 1, we see that r_t is a function combining X_t^2 and $\tanh(\gamma(X_t + a))$. When X_t is small, it is very complicated to express the trend of r_t; when γX_t is large, it acts like the quadratic function. We can simulate r_t numerically after we choose the parameters of the process and the function.

Next we consider a quadratic function f of the form

$$f(X_t) = \frac{1}{2}(X_t - c)^T Q(X_t - c) + \gamma \tag{4.1.10}$$

where γ is a $1 \times d$ vector, Q is a $d \times d$ positive-definite, diagonal matrix and c is a $d \times 1$ vector. We compute $\nabla f = Q(X_t - c)$, $Hf = Q$, and we derive

$$(\mathfrak{g}f)(X_t) = \mu^T Q(X_t - c) + \frac{1}{2} tr(\alpha^2 X_t X_t^T Q) = \mu^T Q(X_t - c) + \frac{1}{2}\alpha^2 X_t^T Q X_t.$$

Choosing $\gamma = \frac{1}{2}\frac{v^T S v - \lambda c^T Q c - 2\mu^T Q c}{\lambda}$, we get the interest rate:

$$r_t = \frac{(X_t - v)^T S(X_t - v)}{(X_t - c)^T Q(X_t - c) + 2\gamma}. \tag{4.1.11}$$

For the quadratic function, the form of r_t is similar to that of Rogers' example but different from the choice of S and v. In Rogers' example, $S = \lambda Q + k^T Q + Qk$ and $v = S^{-1}(\lambda Q c + k^T Q c)$. From above we can see, with this different process, the function choice may be varied to yield the positive interest rate property.

Lastly we will focus on the linear function $f(X_t) = \gamma^T X_t$, where γ is a $d \times 1$ vector. We compute the interest rate:

$$r_t = \frac{\lambda \gamma^T X_t - \mu^T \gamma}{\gamma^T X_t} = \lambda - \frac{\mu^T \gamma}{\gamma^T X_t}. \tag{4.1.12}$$

Compare the interest rate of Rogers' example with Bessel process using linear function:

$$dX_t = (a + BX_t)dt + 2\sqrt{diag(X_t)}dW_t, \tag{4.1.13}$$

where $r_t = \frac{-\gamma^T a + X_t^T(\lambda - B^T)\gamma}{\gamma^T X_t}$. The r_t in our case is simpler than Rogers' example with the Bessel process, and thus may provide an appealing alternative.

4.2 Application to the General Markov Process $dX_t = h(t, X_t)dt + c(t)X_t dB_t$

4.2.1 Solution of the Process and the General Form of the Infinitesimal Generator

Expanding on the idea of the first simple Markov process(4.1.1), we consider the more general Markov process

$$dX_t = h(t, X_t)dt + c(t)X_t dB_t \tag{4.2.1}$$

where $X_t \in \mathbb{R}$ is a scalar process, $X_0 = x$, $h(t, X_t)$ is a function of t and X_t, $c(t)$ is a function of t, and $B_t \in \mathbb{R}$ is standard Brownian motion.

Using the integrator factor $F_t = e^{-\int_0^t c(s)dB_s + \frac{1}{2}\int_0^t c(s)^2 ds}$ and Ito's formula, we get

$$\begin{aligned}
dF_t &= c(t)^2 F_t dt - c(t)F(t)dB_t \\
d(F_t X_t) &= X_t dF_t + F_t dX_t + dF_t dX_t = h(t, X_t)F_t dt
\end{aligned}$$

Defining $Y_t(\omega) = F_t(\omega)X_t(\omega)$, i.e. $X_t = F_t^{-1}Y_t$ and integrating both sides of above equation, we obtain an integral representation the solution:

$$X_t = X_0 + F_t^{-1} \int_0^t F_s \cdot h(s, X_s) ds \tag{4.2.2}$$

The infinitesimal generator is seen to be

$$(\mathfrak{g}f)(x) = h(t, X_t) \cdot \nabla f + \frac{1}{2} tr[c(t)^2 X_t X_t^T \cdot Hf]. \tag{4.2.3}$$

Here are two examples with specific choices of the function h.

Example 1: $h(t, X_t) = \frac{1}{X_t}$ and $c(t) = \alpha$

Here we get the integrating factor $F_t = e^{-\alpha B_t + \frac{1}{2}\alpha^2 t}$ and define $Y_t(\omega) = F_t(\omega) X_t(\omega)$. So, $\frac{dY_t}{dt} = F_t \cdot (F_t^{-1} Y_t)^{-1}$, i.e. $Y_t dY_t = F_t^2 dt$. Integrating both sides of this equation, we get

$$\frac{1}{2}(Y_t^2 - Y_0^2) = \int_0^t e^{-2\alpha B_s + \alpha^2 s} ds$$

which yields the solution

$$Y_t = (x^2 + 2 \int_0^t e^{-2\alpha B_s + \alpha^2 s} ds)^{\frac{1}{2}} \tag{4.2.4}$$

i.e. $\|Y_t\|^2 = x^2 + 2 \int_0^t e^{-2\alpha B_s + \alpha^2 s} ds$.

Example 2: $h(t, X_t) = X_t^\gamma$ and $c(t) = \alpha$

Here we also get the integrating factor $F_t = e^{-\alpha B_t + \frac{1}{2}\alpha^2 t}$ and define $Y_t(\omega) = F_t(\omega) X_t(\omega)$. So, $\frac{dY_t}{dt} = F_t \cdot (F_t^{-1} Y_t)^\gamma$, i.e. $Y_t^{-\gamma} dY_t = F_t^{-\gamma+1} dt$. If $\gamma = 1$, we get the geometric Brownian motion. If $\gamma \neq 1$ we can integrate both sides to get

$$\frac{1}{-\gamma + 1}(Y_t^{-\gamma+1} - Y_0^{-\gamma+1}) = \int_0^t F_s^{-\gamma+1} ds$$

which yields the solution

$$Y_t = [x^{-\gamma+1} + \frac{1}{1-\gamma}\int_0^t (e^{-\alpha B_s + \frac{1}{2}\alpha^2 s})^{1-\gamma} ds]^{\frac{1}{1-\gamma}} \tag{4.2.5}$$

where $Y_0 = x$.

4.2.2 Derivation of the Interest Rate r_t

Here we will consider various choices of f for the infinitesimal generator (4.2.3) to derive the interest rate r_t. All of the results will be summarized in the table (4.1) at the end of this chapter.

Exponential quadratic function

Suppose the dynamics of $(X_t)_{t\geq 0}$ are given by equation (4.2.1), and f is an exponential quadratic function of the form

$$f(X_t) = e^{\frac{1}{2}(X_t - a)^T Q (X_t - a)} \tag{4.2.6}$$

where a is a $d \times 1$ vector of real numbers and Q is a $d \times d$ positive-definite, diagonal matrix.

In order to derive the r_t, recalling $r_t = \frac{(\lambda - \mathfrak{g})f(X_t)}{f(X_t)} = \frac{g(X_t)}{f(X_t)}$ for some λ, we need to derive the infinitesimal generator $\mathfrak{g}$. So, first we compute ∇f and Hf.

$$\nabla f = f(X_t)Q(X_t - a)$$

$$Hf = f(X_t)\begin{bmatrix} q_{11} + q_{11}^2(X_1 - a_1)^2 & \dots & q_{11}q_{dd}(X_1 - a_1)(X_d - a_d) \\ \vdots & \ddots & \vdots \\ q_{11}q_{dd}(X_1 - a_1)(X_d - a_d) & \dots & q_{dd} + q_{dd}^2(X_d - a_d)^2 \end{bmatrix}.$$

Similar to the previous section, we derive the infinitesimal generator:

$$(\mathfrak{g}f)(X_t) = f(X_t)[h(t,X_t)^T Q(X_t - a) + \frac{1}{2}c(t)^2 \cdot X_t^T (Q + Q(X_t - a)(X_t - a)^T Q)X_t)] \tag{4.2.7}$$

and the interest rate:

$$r_t = \lambda - h(t,X_t)^T Q(X_t - a) - \frac{1}{2}c(t)^2 \cdot X_t^T (Q + Q(X_t - a)(X_t - a)^T Q)X_t) \tag{4.2.8}$$

Exponential linear function

Suppose the dynamics of $(X_t)_{t\geq 0}$ are given by equation (4.2.1), and f is an exponential linear function of the form

$$f(X_t) = e^{\gamma^T X_t} \tag{4.2.9}$$

where γ is a $d \times 1$ vector.

By computing $\nabla f = f(X_t)\gamma$ and $Hf = f(X_t)\gamma\gamma^T$, we have the infinitesimal generator:

$$(\mathfrak{g}f)(X_t) = f(X_t)[h(t,X_t)^T\gamma + \frac{1}{2}(c(t)X_t^T\gamma)^2] \tag{4.2.10}$$

and the interest rate:

$$r_t = \lambda - h(t,X_t)^T\gamma - \frac{1}{2}(c(t)X_t^T\gamma)^2. \tag{4.2.11}$$

Hyperbolic cosine function

Suppose the dynamics of $(X_t)_{t\geq 0}$ are given by equation (4.2.1), and f is a hyperbolic cosine function of the form

$$f(X_t) = \cosh(\gamma(X_t + a) \tag{4.2.12}$$

where γ is a $1 \times d$ vector and a is a $d \times 1$ vector.

First, for simplicity, we consider the case $d = 1$. By computing $\nabla f = \gamma \sinh(\gamma(X_t + c))$ and $Hf = \gamma^2 \cosh(\gamma(X_t + c))$, we have the infinitesimal generator:

$$(\mathfrak{g}f)(X_t) = h(t, X_t)\gamma \sinh(\gamma(X_t + a)) + \frac{1}{2}c(t)^2\gamma^2 X_t^2 \cosh(\gamma(X_t + a))$$

and the interest rate:

$$r_t = \lambda - h(t, X_t)\gamma \tanh(\gamma(X_t + a)) - \frac{1}{2}c(t)^2\gamma^2 X_t^2. \tag{4.2.13}$$

For general case of d, we have $h(t, X_t) \in \mathbb{R}^d$ and $\gamma^T \in \mathbb{R}^d$. By computing $\nabla f = \gamma \sinh(\gamma(X_t + a))$ and $Hf = \gamma^T\gamma \cosh(\gamma(X_t + a))$, we derive the infinitesimal generator:

$$(\mathfrak{g}f)(X_t) = \gamma h(t, X_t) \sinh(\gamma(X_t + a)) + \frac{1}{2}c(t)^2(\gamma X_t)^2 \cosh(\gamma(X_t + a)) \tag{4.2.14}$$

and the interest rate:

$$r_t = \lambda - \gamma h(t, X_t) \tanh(\gamma(X_t + a)) - \frac{1}{2}c(t)^2(\gamma X_t)^2. \tag{4.2.15}$$

Linear function

Suppose the dynamics of $(X_t)_{t\geq 0}$ are given by equation (4.2.1), and f is a linear function of the form

$$f(X_t) = \gamma^T X_t \tag{4.2.16}$$

where γ is a $d \times 1$ vector.

By computing $\nabla f = \gamma$, $Hf = 0$, we derive

$$(\mathfrak{g}f)(X_t) = h(t, X_t)^T \gamma \tag{4.2.17}$$

and the interest rate:

$$r_t = \frac{\lambda \gamma^T X_t - h(t, X_t)^T \gamma}{\gamma^T X_t} = \lambda - \frac{h(t, X_t)^T \gamma}{\gamma^T X_t}. \tag{4.2.18}$$

Quadratic function

Suppose the dynamics of $(X_t)_{t \geq 0}$ are given by equation (4.2.1), and f is an quadratic function of the form

$$f(X_t) = \frac{1}{2}(X_t - a)^T Q (X_t - a) + \gamma \tag{4.2.19}$$

where γ is a $1 \times d$ vector, Q is a $d \times d$ positive-definite, diagonal matrix and a is a $d \times 1$ vector.

By computing $\nabla f = Q(X_t - a)$, $Hf = Q$, we derive

$$\begin{aligned}
(\mathfrak{g}f)(X_t) &= h(t, X_t)^T Q (X_t - a) + \frac{1}{2} tr(c(t)^2 X_t X_t^T Q) \\
&= h(t, X_t)^T Q (X_t - a) + \frac{1}{2} c(t)^2 X_t^T Q X_t.
\end{aligned}$$

The funcion $g(X_t)$ can be computed as follows:

$$\begin{aligned}
g(X_t) =& \lambda\gamma + \frac{1}{2}\lambda a^T Q a + h(t, X_t)^T Q a + \frac{1}{2}(\lambda - c(t)^2) X_t^T Q X_t \\
&- \frac{1}{2}\lambda X_t^T Q a - \frac{1}{2}(\lambda a^T Q + 2h(t, X_t)^T Q) X_t \\
=& \lambda\gamma + \frac{1}{2}\lambda a^T Q a + h(t, X_t)^T Q a + \frac{1}{2}(X_t - v)^T S (X_t - v) - \frac{1}{2} v^T S v
\end{aligned}$$

where $S = (\lambda - c(t)^2)Q$,$v = S^{-1}(\lambda Q a + Q h(t, X_t))$.

The precise form of r_t depends on $h(t, X_t)$, which can vary depending on the example. We will look at some specific examples in the following sections.

4.2.3 Results for Special Examples

With the above general form of infinitesimal generator 4.2.7,4.2.10,4.2.14, and 4.2.17, we apply Rogers' algorithm to the following nonlinear examples and derive the positive interest rate for each example. The results are summarized in Table 4.1.

1. The following stochastic differential equation provides a model for the growth of a population of size X_t in a stochastic crowded environment (Oksendal(2003) [28]):

$$dX_t = \alpha X_t(k - X_t)dt + \beta X_t dB_t \qquad X_0 = x > 0 \tag{4.2.20}$$

where the constant $k > 0$ is called the carrying capacity of the environment, the constant $\alpha \in \mathbb{R}$ is a measure of the quality of the environment and the constant $\beta \in \mathbb{R}$ is a measure of the size of noise in the system. To solve this system, we will bring in the geometric Brownian motion

$$dS_t = aS_t dB_t + bS_t dt, \quad S_0 = 1$$

and the linear S.D.E

$$dY_t = aY_t dB_t + (r + bY_t)dt, \quad Y_0 = y.$$

As we know, the solution of the G.B.M is $S_t = e^{(b-1/2a^2)t+aB_t}$. Now, use the 2-dim Ito's formula to $\frac{Y_t}{S_t}$, we get

$$d(\frac{Y_t}{S_t}) = -\frac{Y_t}{S_t^2}dS_t + \frac{1}{S_t}dY_t - \frac{1}{S_t^2}dS_t dY_t + \frac{Y_t}{S_t^3}dS_t dS_t = \frac{r}{S_t}dt$$

So, we get

$$Y_t = S_t(y + r\int_0^t \frac{1}{S_u}du)$$

Now if we take $Y(t) = \frac{1}{X_t}$ for the SDE (4.2.20), and use Ito's formula again, we get

$$\begin{aligned} dY_t =& d\frac{1}{X_t} = -\frac{1}{X_t^2}dX_t + \frac{1}{X_t^3}dX_t dX_t \\ =& [\alpha - (\alpha k - \beta^2)Y_t]dt - \beta Y_t dB_t. \end{aligned}$$

Using the Y_t solution with $a = -\beta, \quad r = \alpha$ and $b = \beta^2 - \alpha k$, we can directly derive the solution of the SDE (4.2.20):

$$Y_t = e^{-\beta B_t + (\frac{1}{2}\beta^2 - \alpha k)t}(x^{-1} + \alpha\int_0^t e^{\beta B_u + (\alpha k - \frac{1}{2}\beta^2)u}du)$$

i.e. we get the closed form expression for the solution of model(4.2.20):

$$X_t = \frac{e^{\beta B_t + (\alpha k - \frac{1}{2}\beta^2)t}}{x^{-1} + \alpha\int_0^t e^{\beta B_u + (\alpha k - \frac{1}{2}\beta^2)u}du}. \tag{4.2.21}$$

From the result table (4.1) of r_t for this model, due to the positivity of X_t, we find that the result for a linear function 4.2.16 is very attractive. Indeed using the formula 4.2.18 with $h(t, X_t) = \alpha X_t(k - X_t)$ and the formula 4.2.21 for X_t

we obtain the interest rate

$$r_t = \lambda - k\alpha + \frac{\alpha e^{\beta B_t + (\alpha k - \frac{1}{2}\beta^2)t}}{x^{-1} + \alpha \int_0^t e^{\beta B_u + (\alpha k - \frac{1}{2}\beta^2)u} du}. \tag{4.2.22}$$

2. The spot freight rate model in shipping (geometric mean reversion process): (Oksendal(2003) [28])

$$dX_t = k(\alpha - \log X_t)X_t dt + \sigma X_t dB_t \qquad X_0 = x > 0 \tag{4.2.23}$$

where k, α, σ and x are positive constants. Change variable by $Y_t = \log X_t$, so $Y_0 = \log x$, and use Ito's formula to obtain

$$\begin{aligned} dY_t =& \frac{1}{X_t} dX_t - \frac{1}{2}\frac{1}{X_t^2} dX_t dX_t \\ =& [(k\alpha - \frac{1}{2}\sigma^2) - kY_t]dt + \sigma dB_t. \end{aligned}$$

This is a narrow sense linear SDE, so by Arnold(1974) [2], we have the solution as follows:

$$Y_t = e^{-kt} logx + (\alpha - \frac{1}{2k}\sigma^2)(1 - e^{-kt}) + \sigma e^{-kt} \int_0^t e^{ks} dB_s$$

Thus, we get the closed form expression for the solution:

$$X_t = e^{e^{-kt} logx + (\alpha - \frac{1}{2k}\sigma^2)(1 - e^{-kt}) + \sigma e^{-kt} \int_0^t e^{ks} dB_s} n \tag{4.2.24}$$

Observe that X_t has a desirable geometric mean reversion property. Please refer to the appendix for the computation of the expectation of X_t . Now checking the result table (4.1), we find that r_t with the choice of a linear function of f has a nice form. We will use this linear function for the next step. First, $dr_t = kd(\log X_t)$. Then we use the same method that was used to solve the SDE (4.2.23). Take $Y_t = \log(X_t)$, so that $r_t = \lambda - k\alpha + kY_t$ and $dr_t = kdY_t$. Since Y_t satisfies $dY_t = [(k\alpha - \frac{1}{2}\sigma^2) - kY_t]dt + \sigma dB_t$, we get

$$dr_t = [(k\lambda - \frac{k}{2}\sigma^2) - kr_t)]dt + k\sigma dB_t \tag{4.2.25}$$

which is a narrow sense linear stochastic differential equation and has the solution

$$\begin{aligned} r_t &= \lambda - k\alpha + k\log X_t \\ &= \lambda - k\alpha + k[e^{-kt}\log x + (\alpha - \frac{1}{2k}\sigma^2)(1 - e^{-kt}) + \sigma e^{-kt}\int_0^t e^{ks}dB_s]. \end{aligned}$$

By way of comparison we recall that Black and Karasinski(1991) [4] used a general time-inhomogeneous SDE with $Y_t = \log X_t$:

$$dY_t = k(t)(\log\mu(t) - Y_t)dt + \sigma(t)dB_t$$

where $k(t)$, $\mu(t)$ and $\sigma(t)$ are deterministic functions of time. Applying Ito's formula we get:

$$dX_t = k(t)[\log\mu(t) + \frac{\sigma(t)^2}{2k(t)} - \log X_t]X_t dt + \sigma(t)X_t dB_t$$

So, the geometric mean reversion process is the time-homogeneous version of the Black-Karasinski model with $k(t) = k$, $\sigma(t) = \sigma$ and $\log\mu(t) + \frac{\sigma(t)^2}{2k(t)} = \alpha$. Please see Appendix A for more details of computing the expectation of geometric mean reversion process. This model fits the historical data significantly better with its lognormal property, so we believe that it will be a very good research model for the future work.

Finally we mention briefly two additional examples, but suppress any computational details.

3. Brennan and Schwartz(1979) [28] model:

$$dX_t = \alpha(\mu(t) - X_t)dt + \sigma(t)X_t dB_t \tag{4.2.26}$$

where α is positive constant, and $\mu(t)$ and $\sigma(t)$ are time-dependent functions.

4. Brownian Motion on an ellipse:

$$dX_t = -\frac{1}{2}X_t dt + MX_t dB_t \tag{4.2.27}$$

where $M = \begin{bmatrix} 0 & -\frac{a}{b} \\ \frac{b}{a} & 0 \end{bmatrix}$ with the solution: $X_t = \begin{bmatrix} acosB_t \\ bsinB_t \end{bmatrix}$.

We summarized all of the results in the following table(4.1) of r_t for various models and choices of the function f that we have considered. In our summary we label the models we have considered as follows:

Model 1: Mean-reverting Ornstein-Uhlenbeck(OU) process, $dX_t = k(\theta - X_t)dt + dB_t$;

Model 2: Mean-reverting Bessel process, $dX_t = k(\theta - X_t)dt + \Sigma\sqrt{diag(X_t)}dB_t$;

Model 3: Brownian bridge, $dX_t = -\frac{b-X_t}{1-t}dt + dB_t$;

Model 4: special Markov process, $dX_t = h(t, X_t)dt + c(t)X_t dB_t$:

Model 4.1: $dX_t = rdt + \alpha X_t dB_t$;

Model 4.2: population growth model, $dX_t = \alpha X_t(k - X_t)dt + \beta X_t dB_t$;

Model 4.3: geometric mean reversion process, $dX_t = k(\alpha - logX_t)X_t dt + \sigma X_t dB_t$;

Model 4.4: Brennan and Schwartz model, $dX_t = \alpha(\mu(t) - X_t)dt + \sigma(t)X_t dB_t$;

Model 4.5: Brownian Motion on ellipse, $dX_t = -\frac{1}{2}X_t dt + MX_t dB_t$.

Table 4.1: Table of Results(Forms of r_t)

	Exponential Quadratic	Exponential linear
Model 1	$\lambda - \frac{1}{2}tr(Q) - (\theta - X_t)^T k^T Q(X_t - c) - \frac{1}{2}(X_t - c)^T Q^2 (X_t - c)$	$\lambda - (\theta - x)^T k^T \gamma + \frac{1}{2}tr(\gamma\gamma^T)$
Model 2	no need(already positive)	$\lambda - \gamma^T(a + KX_t) - 2\sum_{i=1}^{d}(\gamma^{(i)})^2 X_t^{(i)}$
Model 3	$\lambda + \frac{b-X_t}{1-t}Q(X_t - a) - \frac{1}{2}Q - \frac{1}{2}Q^2(X_t - a)^2$	$\lambda + \frac{b-X_t}{1-t}\gamma - \frac{1}{2}\gamma^2$
Model 4.1	$\lambda - \mu^T Q(X_t - c) - \frac{1}{2}\alpha^2 \cdot X_t^T(Q + Q(X_t - c)(X_t - c)^T Q)X_t$	$\lambda - \mu^T\gamma - \frac{1}{2}(\alpha X_t^T \gamma)^2$
Model 4.2	$\lambda - \alpha Q X_t(k - X_t)(X_t - a) - \frac{1}{2}\beta^2 Q X_t^2 (1 + Q(X_t - a)^2)$	$\lambda - \alpha\gamma X_t(k - X_t) - \frac{1}{2}(\beta\gamma X_t)^2$
Model 4.3	$\lambda - Qk(\alpha - \log X_t)X_t(X_t - a) - \frac{1}{2}\sigma^2 Q X_t^2(1 + Q(X_t - a)^2)$	$\lambda - \gamma k(\alpha - \log X_t)X_t - \frac{1}{2}(\sigma\gamma X_t)^2$
Model 4.4	$\lambda - Q\alpha(\mu(t) - X_t)(X_t - a) - \frac{1}{2}\sigma(t)^2 Q X_t^2(1 + Q(X_t - a)^2)$	$\lambda - \gamma\alpha(\mu(t) - X_t) - \frac{1}{2}(\sigma(t)\gamma X_t)^2$
Model 4.5	$\lambda + \frac{1}{2}X_t^T Q(X_t - a) - \frac{1}{2}M^T M X_t^T(Q + Q(X_t - a)(X_t - a)^T Q)X_t$	$\lambda + \frac{1}{2}\gamma^T X_t - \frac{1}{2}(\gamma^T M X_t)^2$

	Hyperbolic cosine(1-dim)	Linear
Model 1:	$\lambda - \frac{1}{2}\gamma^2 + \gamma k X_t \tanh(\gamma(X_t + c))$	$\frac{-\theta k^T\gamma + (\lambda\gamma^T + \gamma^T k)X_t}{\gamma^T X_t}$
Model 2:	$\lambda - \frac{1}{2}\Sigma^2 X_t \gamma^2 + X_t^T k^T \gamma \tanh(\gamma(X_t + c))$	$\frac{-\gamma^T a + X_t^T(\lambda - K^T)\gamma}{\gamma^T X_t}$
Model 3	$\lambda + \frac{b-X_t}{1-t}\gamma \tanh(\gamma(X_t + a)) - \frac{1}{2}\gamma^2$	$\lambda - \frac{b-X_t}{(1-t)X_t}$
Model 4.1	$\lambda - \frac{1}{2}\alpha^2(\gamma X_t)^2 - (\gamma\mu)\tanh(\gamma(X_t + c))$	$\lambda - \frac{\mu^T\gamma}{\gamma^T X_t}$
Model 4.2	$\lambda - \gamma\alpha X_t(k - X_t)\tanh(\gamma(X_t + a)) - \frac{1}{2}\beta^2(\gamma X_t)^2$	$\lambda - k\alpha + \alpha X_t$
Model 4.3	$\lambda - \gamma k(\alpha - \log X_t)X_t \tanh(\gamma(X_t + a)) - \frac{1}{2}\sigma^2(\gamma X_t)^2$	$\lambda - k\alpha + k\log X_t$
Model 4.4	$\lambda - \gamma\alpha(\mu(t) - X_t)\tanh(\gamma(X_t + a)) - \frac{1}{2}\sigma(t)^2(\gamma X_t)^2$	$\lambda + \alpha - \frac{\alpha\mu(t)}{X_t}$
Model 4.5	$\lambda + \frac{1}{2}\gamma X_t \tanh(\gamma(X_t + a)) - \frac{1}{2}M^2(\gamma X_t)^2$	$\lambda + \frac{1}{2}$

	Quadratic: $r_t = \frac{(X_t - v)^T S(X_t - v)}{(X_t - a)^T Q(X_t - a) + 2\gamma}$
Model 1:	$S = \lambda Q + k^T Q + Qk, \quad v = S^{-1}(\alpha Q a + k^T Q a)$
Model 2:	$\lambda - \frac{2(\theta - X_t)^T k^T Q(X_t - c) + tr(\Sigma\Sigma^T diag X_t Q)}{(X_t - c)^T Q(X_t - c) + 2\gamma}$
Model 3	$S = -\frac{2Q}{1-t}$, $v = S^{-1}(\lambda a - \frac{a+b}{1-t})Q)$
Model 4.1	$S = (\lambda - \alpha^2)Q$, $v = S^{-1}(\lambda a Q + \mu^T Q)$
Model 4.2	$r_t = \frac{g(X_t)}{f(X_t)}$, complicated
Model 4.3	$r_t = \frac{g(X_t)}{f(X_t)}$, complicated
Model 4.4	$S = (\lambda - \beta^2 + \alpha)Q, \quad v = S^{-1}(\lambda a + \alpha a + \alpha\mu(t)))$
Model 4.5	$S = (\lambda - M^T M + 2)Q, \quad v = S^{-1}((\lambda + \frac{1}{2})a^T Q)$

4.3 Generating Positive Interest-Rate Models

4.3.1 Selection of Model Parameters

To ensure the positivity of the interest rate r_t, we need a judicious choice of the parameters in the model for r_t, such as $\lambda, Q, c, , a$ and so on. While this is not guaranteed to work in every case, here we consider some examples of one-dimensional models where the proper selection of parameters yields positive interest rates.

1. Exponential quadratic function with Model 1 from table(4.1)(Mean-reverting OU process):

$$r_t = \lambda - \frac{1}{2}Q - kQ(\theta - X_t)(X_t - c) - \frac{1}{2}Q^2(X_t - c)^2.$$

If we take $\widetilde{Q} = 2kQ - Q^2, \widetilde{c} = (kQc + kQ\theta - Q^2c)/\widetilde{Q}$, and $\lambda = \frac{1}{2}Q - \theta kQc + \frac{1}{2}c^2Q^2 + \frac{1}{2}\widetilde{c}^2\widetilde{Q}$, then we will have: $r_t = \frac{1}{2}\widetilde{Q}(X_t - \widetilde{c})^2$, which is a quadratic function of a normally distributed variable. Such models, which are known as squared-Gaussian models, guarantee positivity of interest rates if $\widetilde{Q} > 0$. These models were first studied by Beaglehole and Tenney(1991) and Jamshidian(1996).

2. Exponential quadratic function with Model 3 from table(4.1)(Brownian Bridge):

$$r_t = \lambda + \frac{b - X_t}{1 - t}Q(X_t - a) - \frac{1}{2}Q - \frac{1}{2}Q^2(X_t - a)^2.$$

If we take

$$\lambda = \frac{1}{2}Q + \frac{1}{2}Q^2a^2 + \frac{abQ}{1-t} - \frac{1}{2}Q\frac{(\frac{a+b}{1-t} + Qa)^2}{Q + \frac{1}{1-t}} \tag{4.3.1}$$

and $\frac{Q}{t-1} - Q^2 \geq 0$, then we will have:

$$r_t = \frac{1}{2}(\frac{Q}{t-1} - Q^2)(X_t - \frac{\frac{a+b}{1-t} + Qa}{Q + \frac{1}{1-t}})^2$$

which is quite similar to the model 1 case. However it is time-dependent, which makes it complicated.

3. Exponential linear function with Model 2 from table(4.1)(Bessel process):

$$r_t = \lambda - \gamma(a + KX_t) - 2\gamma^2 X_t$$

If we choose $\gamma K + 2\gamma^2 \leq 0$ and $\lambda - \gamma a \geq 0$, we will ensure the nonnegativity of interest rates.

4. Exponential linear function with model 4.2 from table (4.1)(population growth):

$$r_t = \lambda - \alpha\gamma X_t(k - X_t) - \frac{1}{2}(\beta\gamma X_t)^2$$

If we choose $\lambda = \frac{\alpha^2\gamma^2 k^2}{4(\alpha\gamma - \frac{1}{2}\beta^2\gamma^2)}$ and $\alpha\gamma - \frac{1}{2}\beta^2\gamma^2 \geq 0$, we have:

$$r_t = \alpha\gamma - \frac{1}{2}\beta^2\gamma^2[X_t - \frac{\alpha\gamma k}{2(\alpha\gamma - \frac{1}{2}\beta^2\gamma^2)}]^2. \tag{4.3.2}$$

This is a quadratic function, which might guarantee the positivity if we choose the right parameters. We can use this interest rate in future numerical experiments.

5. Hyperbolic consine function with model 1 from table (4.1)(OU process): $r_t = \lambda - \frac{1}{2}\gamma^2 + \gamma k X_t \tanh(\gamma(X_t + c))$. Taking $\lambda = \frac{1}{2}\gamma^2$ and $c = 0$, we have $r_t = \gamma k X_t \tanh(\gamma(X_t + c))$, which is a positive model. For large value of γX_t, as $\tanh(\gamma X_t) \to 1$, the models acts like a Gaussian model, but loses the positivity

property. For X_t near zero, $\tanh(\gamma X_t) \to \gamma X_t$, the model acts like a squared-Gaussian model.

6. Linear function:

Model 2: $r_t = \frac{-\gamma^T a + X_t(\lambda - K)\gamma}{\gamma X_t}$. We shall require of λ and γ that $-\gamma a \geq 0$ and $(\lambda - K)\gamma \geq 0$ for nonnegativity of the interest rate. This category is appealing for further use in numerical experiments.

Model 4.1: $r_t = \lambda - \frac{\mu}{X_t}$. The choices $\lambda \geq 0$ and $\mu \leq 0$ ensure nonnegativity of the interest rate.

Model 4.2: $r_t = \lambda - k\alpha + \alpha X_t$. The choices of $\lambda - k\alpha \geq 0$ and $\alpha \geq 0$ ensure nonnegativity of the interest rate.

Model 4.4: $r_t = \lambda + \alpha - \frac{\alpha\mu(t)}{X_t}$. The choices of $\lambda + \alpha \geq 0$ and $\alpha\mu_t \leq 0$ ensure nonnegativity of the interest rate.

7. Quadratic function: from the result table (4.1), we have specified S and v in the expression of r_t for different interest rate processes. This enables us to make a proper selection of γ in order to guarantee the positivity of $r_t = \frac{(X_t - v)^T S (X_t - v)}{(X_t - a)^T Q (X_t - a) + 2\gamma}$. Here we summarize some appropriate values of γ for the result table (4.1).

Model 1: $\gamma = \frac{1}{2}(\frac{Q - Sv^2}{\alpha} - c^2 Q)$,

Model 3: $\gamma = \frac{v^2 S - (\lambda a^2 - \frac{2ab}{1-t} - 1)}{2\lambda}$,

Model 4.1: $\gamma = \frac{v^2 S - \lambda c^2 Q - 2c\mu Q}{2\lambda}$,

Model 4.4: $\gamma = \frac{v^2 S - \lambda a^2 Q - 2a\alpha Q \mu_t}{2\lambda}$,

Model 4.5: $\gamma = \frac{v^2 S - \lambda a^2 Q}{2\lambda}$.

4.3.2 Deriving the Term Structure for the Positive Interest Models

In order to build the trinomial trees for our positive interest models, we can apply Ito's lemma to the expressions for our positive interest rates to get a pair of stochastic differential equations for the interest rate r_t and the underlying Markov process X_t.

1. Exponential quadratic function with OU process:

$$
\begin{aligned}
r_t =& \lambda - \frac{1}{2}tr(Q) - (\theta - X_t)^T k^T Q(X_t - c) - \frac{1}{2}(X_t - c)^T Q^2 (X_t - c) \\
dX_t =& k(\theta - X_t)dt + dB_t
\end{aligned}
$$

By proper selection of $\tilde{Q}$ and $\tilde{c}$, we can put r_t in the form $r_t = \frac{1}{2}\tilde{Q}(X_t - \tilde{c})^2$, and then by Ito's lemma, obtain

$$
dr_t = [\tilde{Q}k(\theta - \tilde{c})\sqrt{2r_t/\tilde{Q}} - 2kr_t + \tilde{Q}]dt + \sqrt{2\tilde{Q}r_t}dB_t. \tag{4.3.3}
$$

2. Exponential linear function with Bessel process:

$$
\begin{aligned}
r_t =& \lambda - \gamma a - (\gamma K + 2\gamma^2)X_t \\
dX_t =& (a + KX_t)dt + 2\sqrt{X_T}dW_t
\end{aligned}
$$

By Ito's lemma, we get

$$
dr_t = -(\gamma B + 2\gamma^2[(a + KX_t)dt + 2\sqrt{X_t}dW_t]. \tag{4.3.4}
$$

3. Exponential linear function with population growth model(model 4.2):

$$\begin{aligned} r_t =& (\alpha\gamma - 1/2\beta^2\gamma^2)[X_t - \frac{\alpha\gamma k}{(2\alpha\gamma - \beta^2\gamma^2)}]^2 \\ dX_t =& dX_t = \alpha X_t(k - X_t)dt + \beta X_t dB_t \end{aligned}$$

By setting $Q = \alpha\gamma - 1/2\beta^2\gamma^2$ and $c = \frac{\alpha\gamma k}{(2\alpha\gamma - \beta^2\gamma^2)}$ and using Ito's lemma, we get

$$dr_t = (Q + \alpha(2r_t + c\sqrt{2r_t Q})(k - \sqrt{2r_t/Q} - c))dt + \beta(2r_t + c\sqrt{2r_t Q})dB_t. \quad (4.3.5)$$

4. Linear function with Bessel process: (set $\lambda = B$)

$$\begin{aligned} r_t =& -a/X_t \\ dX_t =& (a + BX_t)dt + 2\sqrt{X_T}dW_t \end{aligned}$$

By Ito's lemma, we get

$$dr_t = (r_t^2 - r_t - 2r_t^3/a^2)dt + 2\sqrt{-a/r_t}dW_t. \quad (4.3.6)$$

5. Linear function with model 4.1: (set $\lambda = 0$)

$$\begin{aligned} r_t =& -\mu/X_t \\ dX_t =& \mu dt + \alpha X_t dB_t \end{aligned}$$

By Ito's lemma, we get

$$dr_t = (r_t^2 - 2\frac{r_t^3}{\mu^2})dt - \alpha\mu dB_t. \quad (4.3.7)$$

6. Linear function with geometric mean-reversion model: (set $\lambda = k\alpha$)

$$
\begin{aligned}
r_t =& \lambda - k\alpha + k \log X_t \\
dX_t =& k(\alpha - \log X_t) X_t dt + \sigma X_t dB_t
\end{aligned}
$$

By Ito's lemma, we get

$$
dr_t = (k\lambda - k/2\sigma^2 - kr_t)dt + k\sigma dB_t. \tag{4.3.8}
$$

7. Linear function with Brennan-Schwartz model: (set $\lambda = -\alpha$)

$$
\begin{aligned}
r_t =& -\frac{\alpha\mu_t}{X_t} \\
dX_t =& dX_t = \alpha(\mu(t) - X_t)dt + \sigma(t) X_t dB_t
\end{aligned}
$$

By Ito's lemma, we get

$$
dr_t = (r_t^2 + \alpha r_t - \frac{2r_t^3}{\alpha^2\mu_t^2})dt - \frac{\sigma_t\alpha\mu_t}{r_t} dB_t. \tag{4.3.9}
$$

In the next chapter, we will focus on the model generated by linear function for population growth model X_t in a stochastic crowded environment,which has very good mean reversion and positivity properties:

$$
dX_t = \alpha X_t(k - X_t)dt + \beta X_t dB_t \qquad X_0 = x > 0.
$$

Using a linear function with this model we have the interest rate $r_t = \lambda - k\alpha + \alpha X_t = \alpha X_t$ by choosing $\lambda = k\alpha$. For convenience, we take $\beta = \sigma$ and $k\alpha = \theta(t)$, which yields

$$dr_t = r_t(\theta(t) - r_t)dt + \sigma r_t dB_t. \tag{4.3.10}$$

4.3.3 A General Expression for dr_t

From above analysis of the range of λ and other parameters, here is the general form of r_t based on the Markov process X_t with $dX_t = b(X_t)dt + \sigma(X_t)dB_t$ in the one-dimensional case. First, as we know, $r_t = g(X_t)$ where

$$\begin{aligned} g(x) &= (\lambda - \mathfrak{g})f(x) \\ &= \frac{\lambda f(x) - b(x)f'(x) + \frac{1}{2}\sigma(x)^2 f''(x)}{f(x)} \\ &= \lambda - \frac{b(x)f'(x)}{f(x)} - \frac{\sigma(x)^2 f''(x)}{2f(x)}. \end{aligned}$$

So, we get

$$\begin{aligned} dr_t =& \frac{1}{f(X_t)^2}\Bigg\{ - [b(X_t)b'(X_t) + \frac{1}{2}b''(X_t)]f'(X_t) + [b(X_t)^2 + \frac{1}{2}b'(X_t)]f'(X_t)^2 \\ &- [b(X_t)^2 + b'(X_t) + b(X_t)\sigma(X_t)\sigma'(X_t) + \frac{1}{2}\sigma'(X_t) + \frac{1}{2}\sigma(X_t)\sigma''(X_t)]f''(X_t) \\ &+ \frac{1}{4}\sigma(X_t)^2 f''(X_t)^2 + [\frac{1}{2}b(X_t)\sigma(X_t)^2 + b(X_t) + \frac{1}{2}\sigma(X_t)\sigma'(X_t)]f'(X_t)f''(X_t) \\ &- [\frac{1}{2}b(X_t) + \frac{1}{2}b(X_t)\sigma(X_t)^2 + \sigma(X_t)\sigma'(X_t)]f'''(X_t) + \frac{1}{4}\sigma(X_t)^2 f'(X_t)f'''(X_t) \\ &- \frac{1}{4}\sigma(X_t)^2 f^{(4)}(X_t) + 2\frac{f'(X_t)}{2f(X_t)}[2b'(X_t)f'(X_t) - 2b(X_t)f''(X_t)^2 \\ &+ 2(b(X_t) + \sigma(X_t)\sigma'(X_t))f''(X_t) - \sigma(X_t)^2 f'(X_t)f''(X_t) \\ &+ \sigma(X_t)^2 f'(X_t)f''(X_t) + \sigma(X_t)^2 f'''(X_t)]\Bigg\}dt \\ &+ \frac{1}{f(X_t)^2}\Bigg\{ - b'(X_t)\sigma(X_t)f'(X_t) + b(X_t)\sigma(X_t)f'(X_t)^2 - [b(X_t)\sigma(X_t) \\ &+ \sigma(X_t)^2\sigma'(X_t)]f''(X_t) + \frac{1}{2}\sigma(X_t)^3 f'(X_t)f''(X_t) - \frac{1}{2}\sigma(X_t)^3 f'''(X_t)\Bigg\}dB_t. \end{aligned}$$

While this formula subsumes all of the cases we have previously considered, it is clearly very complicated. However, we intend to analyze this formula in more detail in future investigations to see if it might shed light on generating additional positive interest rate models.

Chapter 5

Calibration of the Interest Rate Model

Here we generalize slightly the short rate model generated by Rogers' method in the previous section to allow for calibration of the model to a specified initial term structure. Hull and White(1994) [17] have developed and implemented a series of short-rate models with the mean reversion property. We take the basic model(4.2.20) and set $\alpha = 1$, $\beta = \sigma$, and replace the constant k by a nonrandom function $\theta(t)$. The short rate $r_t = X_t$ is then governed by the SDE:

$$dr_t = r_t(\theta(t) - r_t)dt + \sigma r_t dB_t. \tag{5.0.1}$$

Using the substitution method $x_t = r_0 \ln r_t$ suggested by Hull and White, one can reduce this model to a model with constant volatility. We can apply Ito's lemma and obtain

$$dx_t = r_0(\theta(t) - e^{\frac{x_t}{r_0}} - \frac{1}{2}\sigma^2)dt + \sigma r_0 dB_t. \tag{5.0.2}$$

Then we can use the method for the constant volatility case to build the trinomial tree for the x_t process and then transform back to r_t. To build the tree for the x_t process, we choose a convenient Δt and set $\Delta x = \sigma r_0 \sqrt{3\Delta t}$(as suggested by Hull and White(1994) [17]).

5.1 Implementing Trinomial Trees

Now, let us recall the construction of the trinomial tree. For both of the binomial and trinomial frameworks, there are two ways to implement the model. One approach is to fix the time step Δt and branching probabilities, which leaves the freedom of adjusting the space step Δx. The other approach is to fix both Δt and Δx, while choosing the branching probabilities so that the change over each time interval Δt has the correct mean and standard deviation.

For the process (5.0.2), the drift function contains an unknown function of time $\theta(t)$ that permits calibration with the initial yield curve. Now introduce the following notation as in Hull and White(1994) [17]:

$\theta(i * \Delta t)$: time-dependent θ at the i-th time step.

$\mu_{i,j}$: the drift rate of r at node (i, j).

$p_{i,j,k}$: probability of moving from (i, j) to $(i + 1, k + \epsilon)$, where $\epsilon = -1, 0, 1$ and $k \in \mathbb{Z}$ is chosen such that x_k is the closest value to $x_j + \mu_{i,j}\Delta t$ (the expected value of x). To match the first and second moments in the change in x, the equations must be satisfied are:

$$p_{i,j,k-1}(k - j - 1)\Delta x + p_{i,j,k}(k - j)\Delta x + p_{i,j,k+1}(k - j + 1)\Delta x = \mu_{i,j}$$

$$p_{i,j,k-1}(k - j - 1)^2\Delta x^2 + p_{i,j,k}(k - j)^2\Delta x^2 + p_{i,j,k+1}(k - j + 1)^2\Delta x^2 = \mu_{i,j}^2 + \sigma^2\Delta t$$

$$p_{i,j,k-1} + p_{i,j,k} + p_{i,j,k+1} = 1$$

Straightforward computation yields the solution of these equations

$$p_{i,j,k-1} = \frac{1}{2}((k-j)^2 + (k-j) - (1+2(k-j))\frac{\mu_{i,j}}{\Delta x} + \frac{\mu_{i,j}^2}{\Delta x^2} + \frac{\sigma^2 \Delta t}{\Delta x^2}) \quad (5.1.1)$$

$$p_{i,j,k} = 1 - (k-j)^2 + 2(k-j)\frac{\mu_{i,j}}{\Delta x} - \frac{\mu_{i,j}^2}{\Delta x^2} - \frac{\sigma^2 \Delta t}{\Delta x^2} \quad (5.1.2)$$

$$p_{i,j,k+1} = \frac{1}{2}((k-j)^2 - (k-j) - (1-2(k-j))\frac{\mu_{i,j}}{\Delta x} + \frac{\mu_{i,j}^2}{\Delta x^2} + \frac{\sigma^2 \Delta t}{\Delta x^2}) \quad (5.1.3)$$

In our case, we have modeled the above trinomial tree equations for $\sigma * r_0$ instead of σ, which is the same process but with different symbols. From the model we derived at the beginning, the expression for $\mu_{i,j}$ is

$$\mu_{i,j} = r_0(\theta(i * \Delta t) - e^{\frac{x_{i,j}}{r_0}} - \frac{1}{2}\sigma^2)\Delta t \quad (5.1.4)$$

where $x_{i,j} = x_0 + j\Delta x$. To determine the values of $\theta(t)$ we use forward induction. Assume the tree has been constructed up to time $i\Delta t$. For $k \leq i$, define $Q(k,j)$ as the value at time 0 of a security that pays off 1 if node (k,j) is reached and zero otherwise. Then $Q(k,j)$ is calculated from the following relationship:

$$Q(k,j) = \sum_{j^*} Q(k-1, j^*) q(j^*, j) e^{-r_{k-1,j^*}\Delta t} \quad (5.1.5)$$

where $q(j^*, j)$ is the probability of moving from $(k-1, j^*)$ to node (k,j), and r_{k-1,j^*} is the short rate at node $(k-1, j^*)$. So once we get $Q(k,j)$ for all $k \leq i$, $\theta(i\Delta t)$ can be determined by the following scheme. The value at node (i,j) of a bond maturing at time $(i+2)\ \Delta t$ is

$$e^{-r_{i,j}\Delta t} E[e^{-r(i+1)\Delta t} | r(i) = r_{i,j}],$$

where E is the risk-neutral expectation operator and $r(i)$ is the value of r at time $i\Delta t$. So, the value at time zero of a discount bond maturing at time $(i+2)\Delta t$ is given by

$$e^{-(i+2)R(i+2)\Delta t} = \sum_j Q(i,j)e^{-r_{i,j}\Delta t}E[e^{-r(i+1)\Delta t}|r(i) = r_{i,j}]$$

where $R(i)$ is the term structure of the bond to calibrate. If we define

$$\epsilon(i,j) = E[r(i+1) - r(i)|r(i) = r_{i,j}],$$

then we get $E[e^{-r(i+1)\Delta t}|r(i) = r_{i,j}] = e^{-r_{i,j}\Delta t} * e^{\epsilon(i,j)\Delta t}$. By expanding $e^{-\epsilon(i,j)\Delta t}$ as a Taylor series, taking the expectation, and ignoring terms of higher order than Δt^2, we get the following expression:

$$E[e^{-r(i+1)\Delta t}|r(i) = r_{i,j}] = e^{-r_{i,j}\Delta t} * (1 - \mu_{r,i,j}\Delta t^2).$$

So, for our model with $\mu_{r,i,j} = (\theta(i * \Delta t) - r_{i,j})r_{i,j}$, which is the drift for the r_t, we can derive the expression of $\theta(i * \Delta t)$ as follows:

$$\theta(i * \Delta t) = \frac{\sum Q(i,j)e^{-2r_{i,j}\Delta t}(1 + r_{i,j}^2\Delta t^2) - e^{-(i+2)R(i+2)\Delta t}}{\sum Q(i,j)e^{-2r_{i,j}\Delta t}r_{i,j}\Delta t^2}. \tag{5.1.6}$$

At this point, we are all set to build the **trinomial-tree-algorithm**:

Step 1: From $\mu_{i-1,j}$ and the implied branching process create the short rate

$r_{i,j}$, $x_{i,j}$ and discount factors at time i. For any j at i,

$$x_{i,j} = x_{0,0} + j\Delta x$$
$$r_{i,j} = e^{x_{i,j}/r_0}$$
$$d_{i,j} = e^{-r_{i,j}\Delta t}$$

Step 2: Update the pure security prices for every node at time step i according to the expression of $Q(i,j)$ in equation(5.1.5).

Step 3: Determine $\theta(i * \Delta t)$ by equation(5.1.6).

Step 4: Using $\theta(i * \Delta t)$ and $x_{i,j}$ update $\mu_{i,j}$ for all j by equation(5.1.4).

Step 5: Decide the branching process to determine k.

Step 6: Calculate the branching probabilities using the solutions(5.1.1) to (5.1.3).

Here we will implement the algorithm and compare the distribution result with a Monte-Carlo simulation in the case where we have a constant θ in the equation (5.1.4) for μ. For constant θ, we omit step 2 and step 3 from the above algorithm. In the following figures, we only present results for the first 4 time steps due to the limitation of space, but we will present the distribution in the next subsection and compare it with the Monte Carlo method. Figure 5.1 is the simple case of the trinomial tree generated for the trend of the r_t with θ held constant with value 0.1 and R_0 held constant at 0.05. Figure 5.2 is the trinomial tree generated for the trend of the r_t with θ held constant with value 0.4 and real market term structure R_0. Figure 5.3 is the trinomial tree generated for the trend of the r_t with $\theta(t)$ time dependent, which we get by the above algorithm.

5.2 Comparing Simulation Results with the Closed Form Solution

Using a change of variable, we can find a closed form expression for the solution r_t of the modified population growth model(5.0.1):

$$dr_t = r_t(\theta(t) - r_t)dt + \sigma r_t dB_t.$$

This is a bit different from the model(4.2.20) due to the presence of the time-varying factor $\theta(t)$. Here we follow the method suggested in Chap.4 of Gard(1988) [11]. Let $Y_t = f(t, r_t) = e^{-\sigma \int_0^{r_t} \frac{1}{\sigma u} du} = \frac{1}{r_t}$, and use the Ito's formula.

$$\begin{aligned} dY_t &= f_t dt + f_r dr_t + 1/2 f_{rr}(dr_t)^2 \\ &= (1 - (\theta(t) - \sigma^2)Y_t)dt - \sigma Y_t dB_t. \end{aligned}$$

Now take $F_t = e^{\int_0^t (\theta(s) - 1/2\sigma^2)ds + \sigma dB_t}$, so we have

$$dF_t = F_t[(\theta(t) - 1/2\sigma^2)dt + \sigma dB_t].$$

Set $Z_t = F_t Y_t$ and use the Ito's formula again:

$$dZ_t = F_t dY_t + Y_t dF_t + 1/2 dF_t dY_t = F_t dt.$$

Integrating both sides and using $Z_t = F_t Y_t$, we obtain the solution:

$$r_t = \frac{e^{\sigma B_t + \int_0^t \theta(s)ds - \frac{1}{2}\sigma^2 t}}{r_0^{-1} + \int_0^t e^{\sigma B_u + \int_0^u \theta(s)ds - \frac{1}{2}\beta^2 u} du}. \tag{5.2.1}$$

With this closed form solution, we can numerically generate the standard Brownian motion B_t and thus simulate the expression of $e^{\sigma B_t}$. Also, we use a step function interpolation for $\theta(t)$. Our computations yield a pattern which is similar to the trinomial tree result. Figure 5.4 is the numerical result generated by this method for the trend of the r_t. Figure 5.5 is the distribution of the r_t at time $t = 3$. The computation of the probabilities in the previous section of each tranche at the final step yields:

$$P(r_{20} < 2.50\%) = 0.084, \quad P(2.50\% \leq r_{20} < 3.10\%) = 0.481,$$
$$P(3.10\% \leq r_{20} < 3.60\%) = 0.333, \quad P(r_{20} \geq 3.60\%) = 0.101.$$

Compare these probabilities to the distribution generated by the Monte Carlo method:

$$P(r_{20} < 2.477\%) = 0.086, \quad P(2.477\% < r_{20} < 3.13\%) = 0.482,$$
$$P(3.13\% \leq r_{20} < 3.658\%) = 0.3211, \quad P(r_{20} \geq 3.658\%) = 0.1105.$$

We can see that our trinomial tree distribution is very consistent with the Monte Carlo simulation obtained via the closed form solution. By way of illustration, we also present a comparison of numerical results in the case where θ is constant. Figure 5.6 and Figure 5.7 provide graphical summaries of the comparison simulations. The computation of the probabilities in the previous section of each tranche at the final step for the case of constant θ, yield:

$$P(r_{20} < 3.24\%) = 0.0332, \quad P(3.24\% \leq r_{20} < 3.51\%) = 0.3126,$$
$$P(3.51\% \leq r_{20} < 3.80\%) = 0.4722, \quad P(r_{20} \geq 3.80\%) = 0.181.$$

Compare these probabilities to the distribution generated by the Monte Carlo method:

$$P(r_{20} < 2.95\%) = 0.059, \quad P(2.50\% < r_{20} < 3.00\%) = 0.414,$$

$$P(3.00\% \leq r_{20} < 3.05\%) = 0.46, \quad P(r_{20} \geq 3.05\%) = 0.066.$$

We see that for constant θ, there is a small shift between trinomial tree method and Monte Carlo method, but the distributions are otherwise very close.

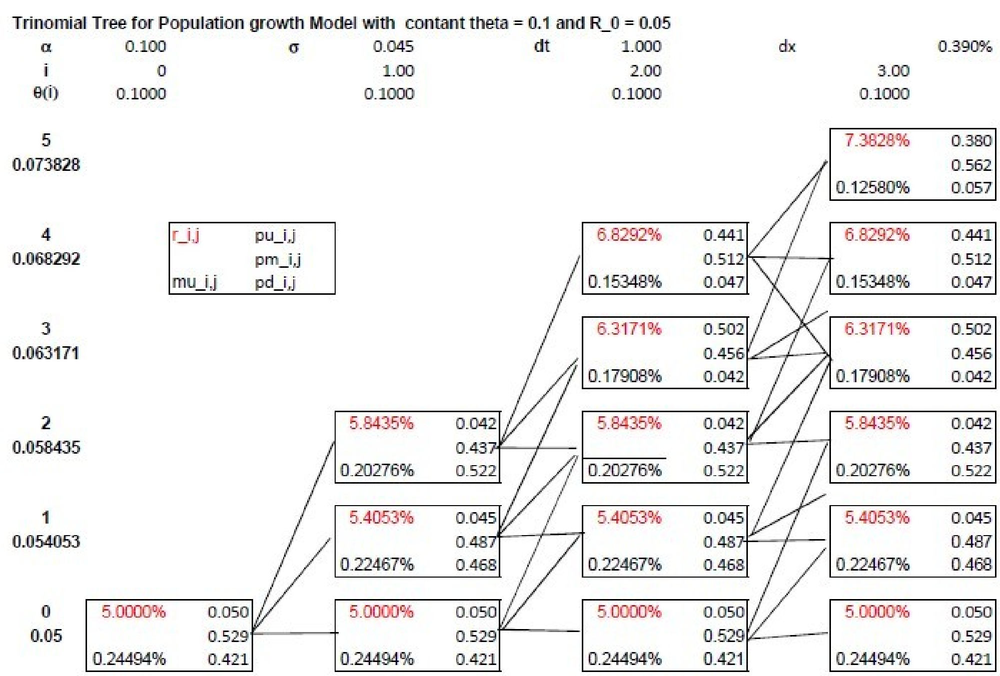

Figure 5.1: r_t generated by constant $\theta = 0.1$ and $R_0 = 0.05$ using trinomial tree

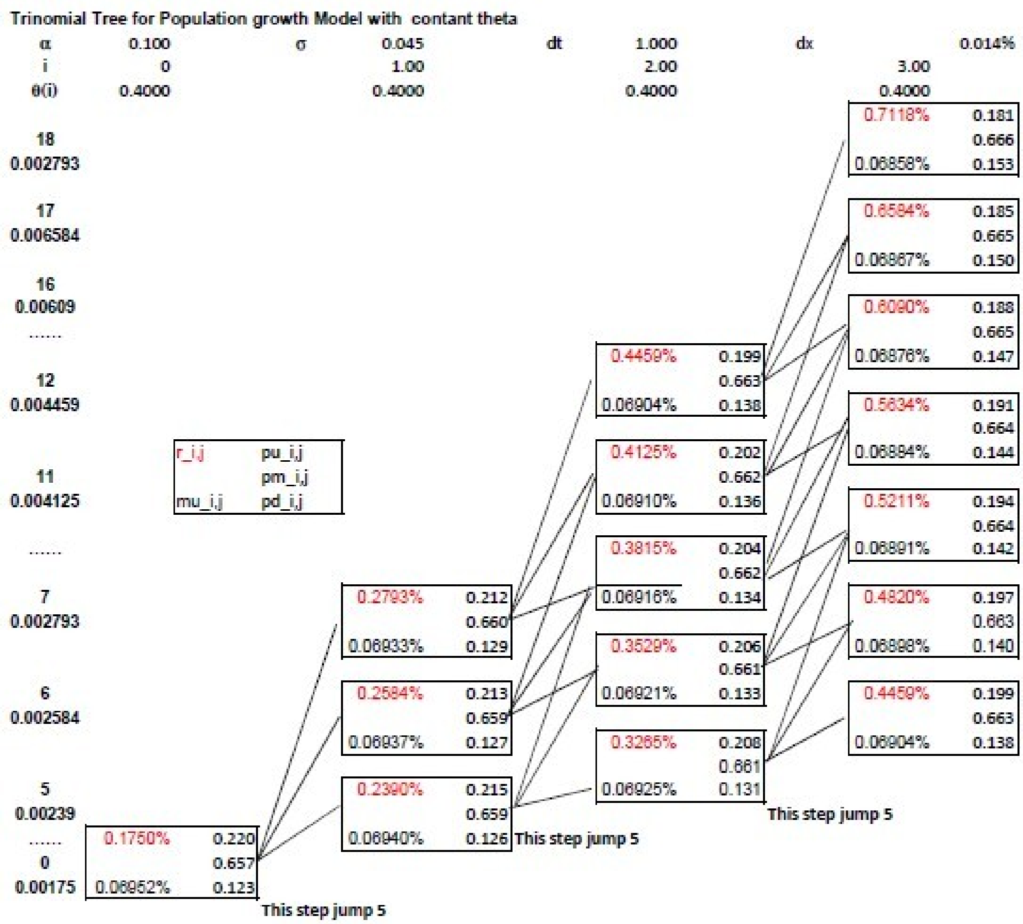

Figure 5.2: r_t generated by constant $\theta = 0.4$ and $R_0 = 0.00175$ using trinomial tree

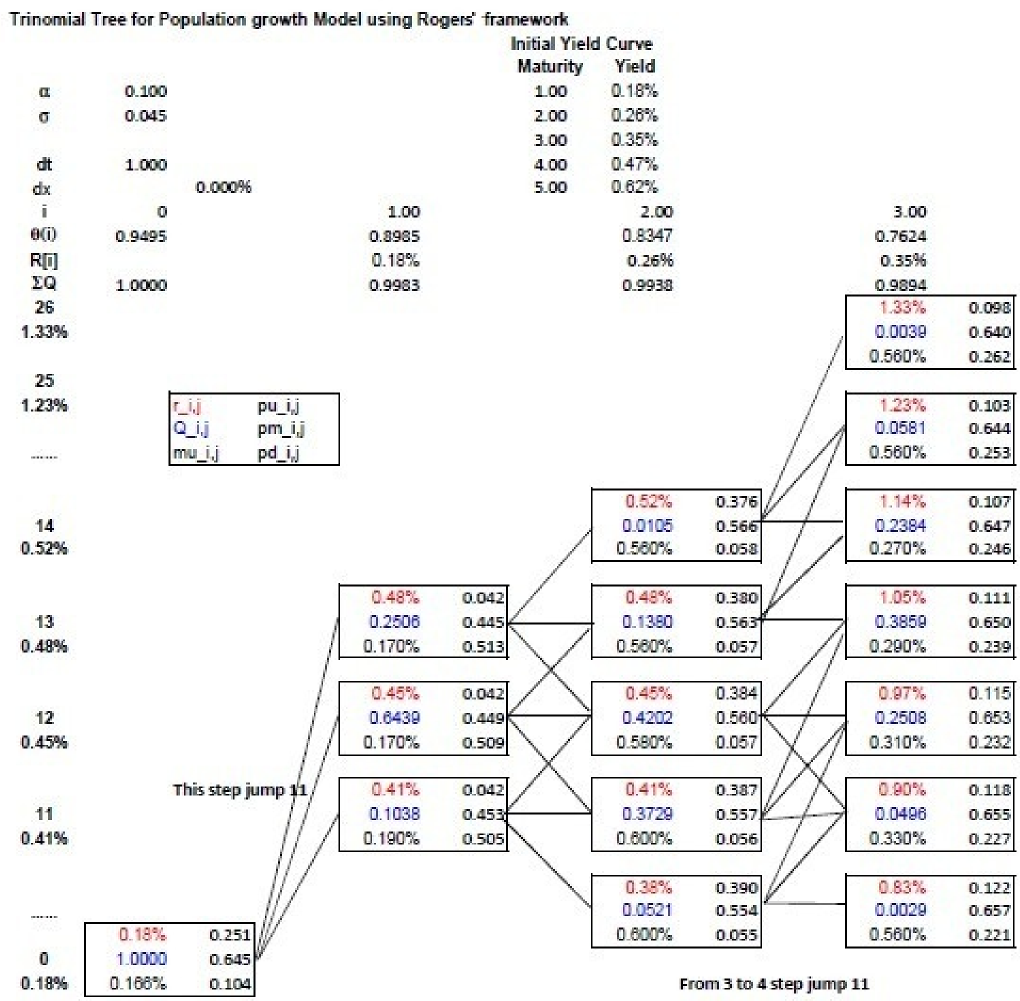

Figure 5.3: r_t generated by Rogers' method using trinomial tree with $\theta(t)$

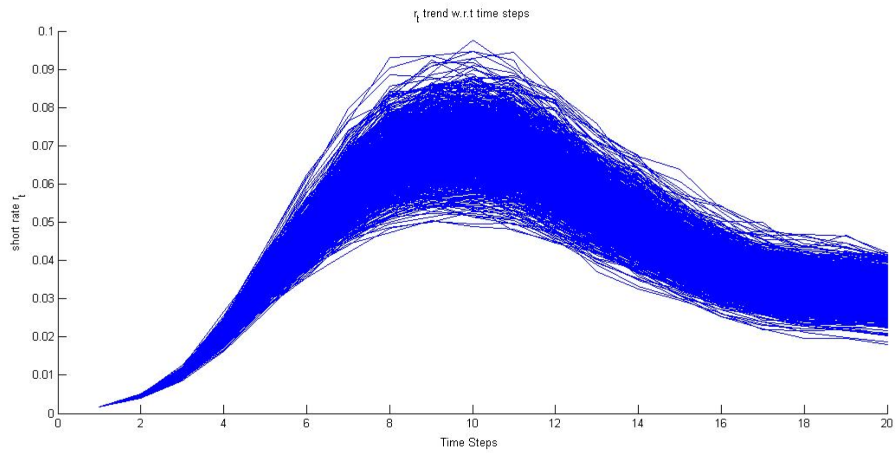

Figure 5.4: r_t generated by closed formula with $\theta(t)$ using Monte Carlo method

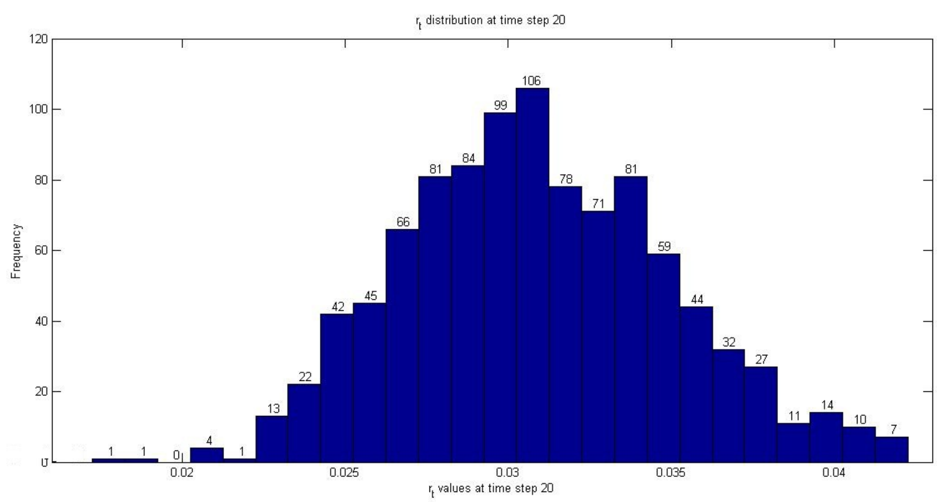

Figure 5.5: r_t Distribution at time 20 with $\theta(t)$

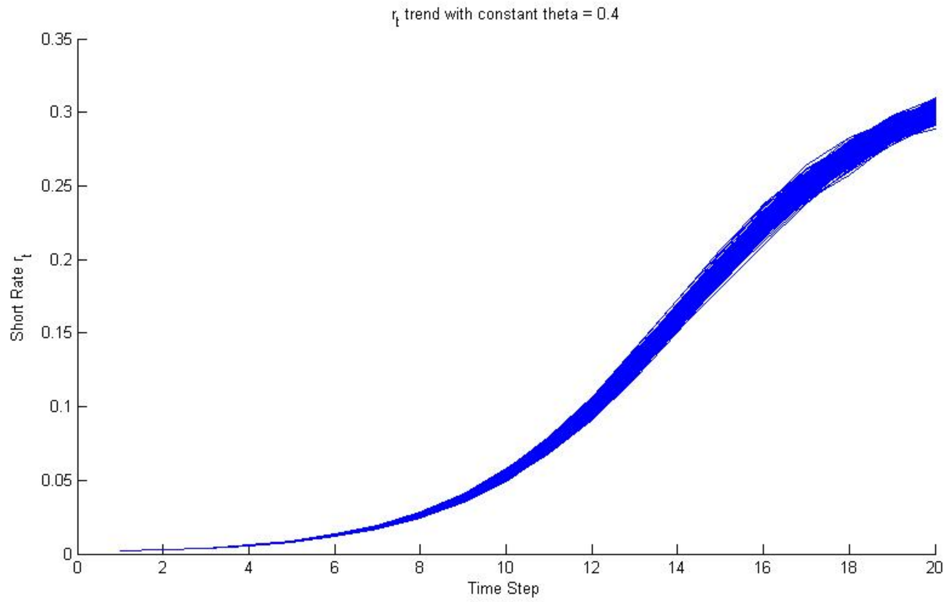

Figure 5.6: r_t generated by closed formula using Monte Carlo method with theta = 0.4

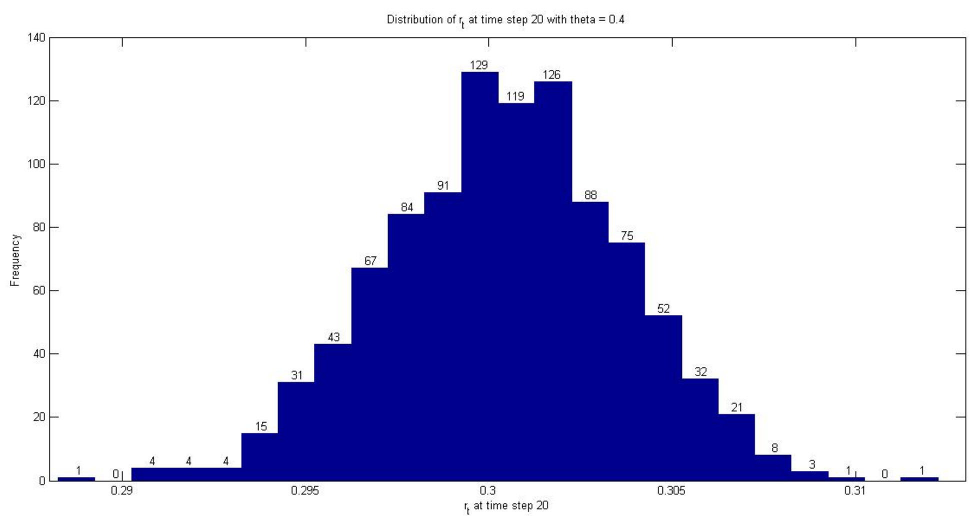

Figure 5.7: r_t Distribution at time 20 with theta = 0.4

Chapter 6

Application of the Interest Rate Model to the Firm Default Probability

6.1 Framework for Interest Rate and Firm Value

The simplest interest rate model for fixed income markets, under a risk-neutral measure, is given by the following (scalar) stochastic differential equation (we will use SDE for the abbreviation in this dissertation):

$$dr_t = \alpha(r_t, t)dt + \sigma(r_t, t)dW_r, \tag{6.1.1}$$

where the instantaneous drift, $\alpha(r_t, t)$ and the volatility $\sigma(r_t, t)$ are smooth functions with $\sigma > 0$(Cairns(2004) [6]). Sometimes (6.1.1) is called a short rate model because r_t is the interest rate for short-term borrowing.

Under a hypothetical risk-neutral measure, the dynamics of firm value, V_t, are routinely modeled by a linear, scalar SDE

$$\frac{dV_t}{V_t} = r_t dt + \sigma_v(t) dW_{v,t}, \tag{6.1.2}$$

where the instantaneous drift r_t denotes the stochastic interest rate process and σ_v(t) is the instantaneous volatility, which is a non-random function of time.(Acharya and Carpenter(2002) [1]).

Moreover, suppose that the shocks to the firm fundamentals and the default-free interest rates drive the variation of V_t. The shocks to the firm fundamentals under the risk-neutral measure $\widetilde{\mathbf{P}}$ are modeled by the Brownian motion processes dW_v, while the shocks to the default-free interest rates are modeled by the Brownian motion process dW_r. The instantaneous correlation, ρ_t, between dW_v and dW_r is assumed to be,

$$E[(dW_{v,t})(dW_{r,t})] = \rho_t dt \tag{6.1.3}$$

with $|\rho_t| < 1$.

Now, setting $g_t = \log V_t$, and applying Itô's formula, we get (Kloeden and Platen(1992) [21])

$$dg_t = [r_t - \frac{1}{2}\sigma_\nu^2(t)]dt + \sigma_\nu(t) dW_{\nu,t} \tag{6.1.4}$$

By the following manipulation, we can change the original correlated Brownian motion processes $W_{r,t}$ and $W_{\nu,t}$ to independent Brownian motion processes $W_{1,t}$ and $W_{2,t}$, respectively. From now on assume $\rho_t = \rho$ is a constant, set $dW_{r,t} = dW_{1,t}$ and $dW_{\nu,t} = \rho dW_{1,t} + \sqrt{1-\rho^2} dW_{2,t}$ (Shreve(2004) [38]). Levy's theorem implies $W_{2,t}$is a Brownian motion process and a simple computation shows that

$W_{1,t}$ and $W_{2,t}$ are uncorrelated. Then we can rewrite of the dynamics of interest rate and firm value as

$$\begin{aligned} dr_t &= \alpha(r_t, t)dt + \sigma_r(r_t, t)dW_{1,t}, \\ dg_t &= [r_t - \frac{1}{2}\sigma_\nu^2(t)]dt + \nu_1(t)dW_{1,t} + \nu_3(t)dW_{2,t} \end{aligned} \tag{6.1.5}$$

where $\nu_1(t) = \rho\sigma_v(t)$ and $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$. Integrating the above equation for g_t, we obtain

$$g_t = g_0 + \int_0^t r_s ds - 1/2 \int_0^t \sigma_\nu^2(s)ds + x_t + z_t \tag{6.1.6}$$

where $x_t = \int_0^t \nu_1(s)dW_{1,s}$, and $z_t = \int_0^t \nu_3(s)dW_{2,s}$.

We will use a narrow sense linear model for the interest rate dynamics. Setting

$$\alpha(r_t, t) = \theta(t) - c(t)r_t \quad and \quad \sigma_r(r_t, t) = \sigma_r(t), \tag{6.1.7}$$

in (6.1.1), we get the narrow sense linear model (generalized Hull-White model (2000) [19]) given by

$$dr_t = -c(t)r_t dt + \theta(t)dt + \sigma_r(t)dW_{1,t}. \tag{6.1.8}$$

Setting $\bar{c}(t) = \int_0^t c(s)ds$, the solution of (6.1.7) is given by (Arnold(1974) [2])

$$r_t = r_t(det) + r_t(ran), \tag{6.1.9}$$

where

$$
\begin{aligned}
r_t(det) &= e^{-\bar{c}(t)} r_0 + \int_0^t e^{-[\bar{c}(t)-\bar{c}(s)]} \theta(s) ds \\
r_t(ran) &= \int_0^t \nu_2(u) dW_{1,u},
\end{aligned}
\tag{6.1.10}
$$

and $\nu_2(u) = e^{-[\bar{c}(t)-\bar{c}(u)]}\sigma_r(u)$. So, it follows that

$$
\int_0^t r_s ds = \int_0^t r_s(det) ds + \int_0^t r_s(ran) ds.
$$

Here we observe that, by integration by parts, we have

$$
\int_0^t r_s(ran) ds = \int_0^t (t-s)\nu_2(s) dW_{1,s}.
$$

Substituting the above result into (6.1.6), we get

$$
g_t - g_0 = g_t(det) + g_t(ran), \tag{6.1.11}
$$

where

$$
\begin{aligned}
g_t(det) &= \int_0^t r_s(det) ds - 1/2 \int_0^t \sigma_\nu^2(s) ds \\
g_t(ran) &= [x_t + \int_0^t r_s(ran) ds] + z_t.
\end{aligned}
\tag{6.1.12}
$$

By straightforward computation we see that $g_t - g_0$ follows a normal distribution,

$$
g_t - g_0 \sim N(\mu(t), \sigma^2(t)), \tag{6.1.13}
$$

where $\mu(t) = g_t(det)$ and $\sigma^2(t) = \sigma_1^2(t) + \sigma_2^2(t)$, with

$$\begin{aligned} \sigma_1^2(t) &= \int_0^t [\nu_1(s) + (t-s)\nu_2(s)]^2 ds \\ \sigma_2^2(t) &= \int_0^t \nu_3^2(s) ds. \end{aligned} \tag{6.1.14}$$

Based on the time change properties of the Brownian motion process and its relation with Itô integrals ((Shiryaev(1999) [37]), we can get the formula for V_t:

$$V_t/V_0 = e^{g_t - g_0} = e^{\int_0^t r_s(det)ds - \frac{1}{2}\int_0^t \sigma_\nu(s)^2 ds + \sqrt{\frac{T_1(t)}{t}}W_{1,t} + \sqrt{\frac{T_2(t)}{t}}W_{2,t}} \tag{6.1.15}$$

where

$$\begin{aligned} T_1(t) &= \int_0^t [\rho\sigma^v(s) + (t-s)e^{-(\bar{c}(t)-\bar{c}(s))}]^2 \sigma_r(s)^2 ds \\ T_2(t) &= \int_0^t \sigma_\nu(s)^2(1-\rho^2) ds. \end{aligned}$$

From above, we can see the firm value follows the lognormal distribution. Recalling that every Ito integral is equivalent to a time scaled Brownian motion process(Shiryaev(1999) [37]), we can change the formula of $g_t(ran)$ in (6.1.11) as

$$g_t(ran) = B_1(\sigma_1(t)^2) + B_2(\sigma_2(t)^2) \tag{6.1.16}$$

where B_1 and B_2 are two independent Brownian motion processes with

$$B_1(\sigma_1(t)^2) = \int_0^t [\nu_1(s) + (t-s)\nu_2(s)] dW_{1,s}, \quad B_2(\sigma_2(t)^2) = \int_0^t \nu_3(s) dW_{2,s}$$

and $\sigma_1(t)^2$ and $\sigma_2(t)^2$ are given in (6.1.14). Since $B_1(t)$ and $B_2(t)$ are independent, there exists a Brownian motion process $B(t)$ such that

$$g_t(ran) = B(\sigma(t)^2) \tag{6.1.17}$$

where $\sigma^2(t) = \sigma_1^2(t) + \sigma_2^2(t)$.

Combining above results, we obtain that

$$g_t - g_0 = \mu(t) + B(\sigma(t)), \tag{6.1.18}$$

where $\mu(t) = g_t(det)$ and $\sigma(t)$ defined as above.

6.2 The Application of First Passage Time to Firm Default Probability

6.2.1 Firm Default Probability

Definition 6.1. Default risk refers to the possibility that the issuer fails to honor the agreement and obligation with respect to the timely payment of interest and principal for a given period.

For our purpose a firm defaults when its value falls under a prescribed barrier. There are two approaches to structurally modeling the occurrence of default: the classical approach by Merton(1974) [27] and the first-passage-time approach by Black and Cox(1976) [3]. Consider a firm with market value $V_t = E + K$ at time t, where E is the equity value and K is the debt value. Suppose the debt matures at time T, at which time the firm is obligated to repay the amount K to bond holders. In the classical approach, default happens

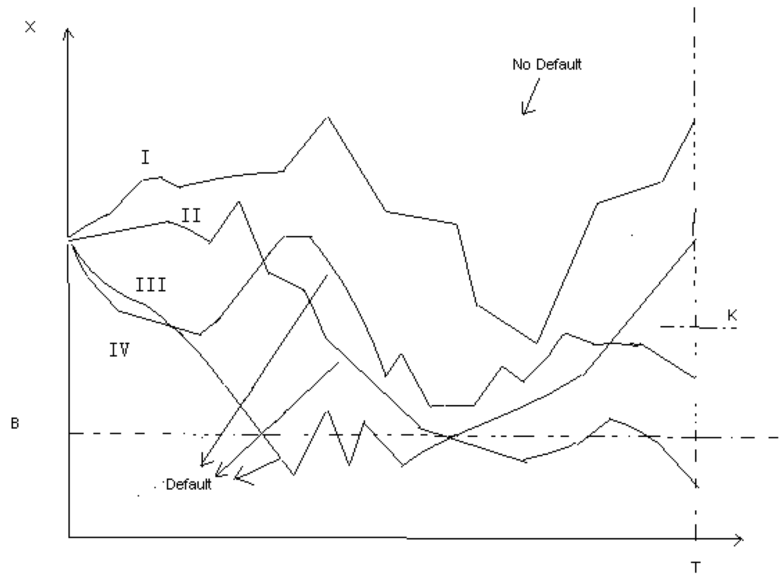

Figure 6.1: Default in the first-passage approach

when $V_T < K$. The previous section's discussion presents a model for which V_t follows the geometric Brownian motion:

$$\frac{dV_t}{V_t} = \mu dt + \sigma_v dW_{v,t}, \quad V_0 > 0, \tag{6.2.1}$$

where $\mu \in \mathbb{R}$ is a drift term, $\sigma_v > 0$ is the volatility term and $W_{v,t}$ is a standard Brownian motion. By Ito's lemma, we can obtain the following default probability result $P_d(K,T)$ due to Merton(1974) [27]:

$$\begin{aligned} P_d(K,T) &= P[V_T < K] = P[\sigma_v W_{v,T} < \log\frac{K}{V_0} - (\mu - 1/2\sigma_v^2)T] \\ &= \Phi(\frac{\log\frac{K}{V_0} - (\mu - 1/2\sigma_v^2)T}{\sigma_v\sqrt{T}}) \end{aligned} \tag{6.2.2}$$

where Φ is the standard normal cumulative distribution function.

In the classical approach, there is nothing to trigger the default event during the time period prior to maturity. Black and Cox(1976) [3] and Giesecke(2004) [12] propose that the default is allowed to occur before maturity when the value of the firm falls below some specific level B prior to maturity (See figure 6.1). Thus, $P_d(B,T)$ is the probability that the firm's value is below B at some time $t < T$, i.e.,

$$P_d(B,T) = P[\min_{\{t<T\}} V_t < B]. \tag{6.2.3}$$

If we assume the firm value follows the same dynamics (6.2.1), then by using the first passage time approach, the default probabilities can be determined as in Giesecke(2004) [12]:

$$\begin{aligned} P_d(K,T) &= P[\min_{t<T} V_t < B] = P[\min_{t<T}((\mu - 1/2\sigma_v^2)t + \sigma_v W_{v,s}) < \log \frac{B}{V_0}] \\ &= \Phi(\frac{\log \frac{B}{V_0} - \bar{\mu}T}{\sigma_v \sqrt{T}}) + (\frac{B}{V_0})^{\frac{2\bar{\mu}}{\sigma_v^2}} \Phi(\frac{\log \frac{B}{V_0} + \bar{\mu}T}{\sigma_v \sqrt{T}}) \end{aligned} \tag{6.2.4}$$

where $\bar{\mu} = (\mu - \frac{1}{2}\sigma_v^2)$. In the next section, we will also propose an new alternative algorithm to derive this default probability(6.2.4).

Now, suppose we allow either of the above default trigger events to occur and define the probability, $P_d(K,B,T)$, that the firm's value is either below B at some time $t < T$ or below K at time T, i.e.,

$$P_d(K,B,T) = P[\min_{\{t<T\}} V_t < B \quad or \quad V_T < K]. \tag{6.2.5}$$

Consequently, the default probability of the firm is essentially an instance of the first passage time problem. With the same firm value dynamics (6.2.1), to derive the default probabilities(6.2.5), we need to use the reflection principle method (refer to Shreve(2004) [38]). Giesecke(2004) [12] has given the default probability result for this dynamics:

$$P_d(K,B,T) = P[\min_{\{t<T\}} V_t < B \quad or \quad V_T < K]$$
$$= \Phi(\frac{\log\frac{B}{V_0} - \bar{\mu}T}{\sigma_v\sqrt{T}}) + (\frac{B}{V_0})^{\frac{2\bar{\mu}}{\sigma_v^2}}\Phi(\frac{\log\frac{B^2}{KV_0} + \bar{\mu}T}{\sigma_v\sqrt{T}}). \tag{6.2.6}$$

In the work of Qian(2008) [34], the computation of this default probability was generalized to allow a stochastic interest rate governed by the Vasicek model. This computation relies on the solution of a first-passage-time problem for a Brownian motion process with a linear (in time) drift term. Please see the figure (6.3) for the default probability for condition (6.2.4) and condition (6.2.5). Our next object is to indicate an approach to the first-passage-time problem for Brownian motion process with nonlinear drift terms. Such a generalization would allow the use of nonlinear interest rate models in the default probability computation.

6.2.2 Discussion on the First Passage Time

Following the result of section 6.1 we saw that $g_t - g_0 = \mu(t) + B(\sigma(t)^2)$. Let $\tau = \sigma(t)^2$ for $0 \leq t \leq T$. Note that $\sigma(t)^2$ is generally a strictly increasing function of t, so $t \mapsto \sigma(t)^2$ has a well defined inverse ξ, i.e. $\tau = \sigma(t)^2 \Leftrightarrow t = \xi(\tau)$ with $0 \leq \tau \leq \sigma(T)^2$, Then we obtain the following:

$$g(\xi(\tau)) - g_0 = \mu(\xi(\tau)) + B(\tau). \tag{6.2.7}$$

Defining $X(\tau) = g(\xi(\tau)) - g_0$ and $\lambda(\tau) = \mu(\xi(\tau))$, we obtain a diffusion process:

$$X(\tau) = \lambda(\tau) + B(\tau).$$

Now let us look at our default probability (6.2.5). As $X(\tau) = g_t - g_0 = \log \frac{V_t}{V_0}$, to avoid default we need $X(\tau) \geq \log B - \log V_0$. We set $\overline{B} = \log B - \log V_0$, $\overline{K} = \log K - \log V_0$ and $\bar{T} = \sigma(T)^2$. This results in equation (6.2.5) taking the form:

$$P_d(K, B, T) = P[\min_{\{0 \leq \tau \leq \bar{T}\}} X(\tau) \leq \overline{B} \quad or \quad X(\bar{T}) \leq \overline{K}] \tag{6.2.8}$$

Without loss of generality, we can change the time variable τ to t from now on, and change it back when we get the first passage probability, so the diffusion process takes the form:

$$X(t) = \lambda(t) + B(t) \quad i.e. \quad dX(t) = \dot{\lambda}(t)dt + dB(t). \tag{6.2.9}$$

Clearly, we can view this as a first passage time problem for a one-dimensional Markov process. In Qian's work(2008) [34], he derived the first passage time probability for the special case where $\lambda(t)$ is linear or constant. In that case, the Girsanov's theorem can be used to transfer the diffusion process (6.2.9) to a standard Brownian motion, and then the reflection principle can be used to compute the first passage time (default probability). Now let us consider the general case where $\lambda(t)$ is a smooth function. Define the exponential martingale process

$$\Lambda(t) = e^{-\int_0^t \dot{\lambda}(s)dB_s - 1/2\int_0^t \dot{\lambda}^2(s)ds}. \tag{6.2.10}$$

By Radon-Nikodym theorem and Girsanov's theorem, it defines a new equivalent measure P^Λ

$$dP^\Lambda = \Lambda(\bar{T})dP \tag{6.2.11}$$

so that $X(t)$ is a Brownian motion process under the new measure P^{Λ}. By defining $\tilde{X}(t) = -X(t), \tilde{K} = -\bar{K}$, and $\tilde{B} = -\bar{B}$, we have following relations:

$$\begin{aligned} X(\bar{T}) > \bar{K} &\Leftrightarrow \tilde{X}(\bar{T}) < \tilde{K} \\ \min_{\{0\le t\le \bar{T}\}} X(t) = -\max_{\{0\le t\le \bar{T}\}} (-X(t)) > \overline{B} &\Leftrightarrow \max_{\{0\le t\le \bar{T}\}} \tilde{X}(t) < \tilde{B}. \end{aligned} \tag{6.2.12}$$

So, the default probability (6.2.8) can be transformed to the following form:

$$\begin{aligned} P_d(K,B,T) &= 1 - P[\min_{\{0\le t\le \bar{T}\}} X(t) > \overline{B}, \quad X(\bar{T}) > \overline{K}] \\ &= 1 - P[\max_{\{0\le t\le \bar{T}\}} \tilde{X}(t) < \tilde{B}, \quad \tilde{X}(\bar{T}) < \tilde{K}]. \end{aligned} \tag{6.2.13}$$

The joint density function of $(\max_{0\le t\le \bar{T}})$ under P^{Λ} can be obtained by the joint density function (2.2.17) given in chapter 2. This allows us to compute the probability (6.2.13):

$$\begin{aligned} &P[\min_{\{0\le \tau\le \bar{T}\}} X(\tau) \le \overline{B} \quad or \quad X(\bar{T}) \le \overline{K}] \\ &= 1 - E[I(\max_{\{0\le t\le \bar{T}\}} \tilde{X}(t) < \tilde{B}, \quad \tilde{X}(\bar{T}) < \tilde{K})] \\ &= 1 - E^{\Lambda}[\Lambda(T)^{-1} I(\max_{\{0\le t\le \bar{T}\}} \tilde{X}(t) < \tilde{B}, \quad \tilde{X}(\bar{T}) < \tilde{K})] \\ &= 1 - \int_0^{\tilde{B}} \int_{\infty}^{\tilde{K}} e^{\int_0^{\bar{T}} \dot{\lambda}(s)dB_s + 1/2\int_0^{\bar{T}} \dot{\lambda}^2(s)ds} f(\bar{T},z,y)dzdy \end{aligned} \tag{6.2.14}$$

where

$$f(\bar{T},z,y) = \frac{2(2z-y)}{\bar{T}\sqrt{2\pi\bar{T}}} e^{-\frac{(2z-y)^2}{2\bar{T}}}$$

is the joint density function. Unfortunately the term $\int_0^{\bar{T}} \dot{\lambda}(s)dB_s$ prevents the derivation of a closed form solution for the above probability. It is still feasible to explore the use of numerical methods to approach the solution, which we leave for future work.

In light of these difficulties, another method is needed for the default probability computation involves a diffusion process with a nonlinear drift term. To solve this nonlinear problem, we will implement the method of Fortet(1943) [10]. He proposed an implicit formula for the probability density that the first passage time through a constant boundary occurs at time s. Here we will first consider the case of a one-factor continuous Markov process l_t where in terms of above notation $l_t = \overline{B} - X_t$.

Define $p_l(s, l_s, T, l_T)$ as the free transition density for l_t and $k(0, l_0, s, l_s = \underline{l})$ as the probability density that the first passage time through a constant boundary $\underline{l}$ occurs at time s. Then the implicit formula proposed by Fortet is expressed as:

$$p_l(0, l_0, T, l_T) = \int_0^T k(0, l_0, s, l_s = \underline{l}) p_l(s, l_s = \underline{l}, T, l_T) ds \qquad \forall (l_T > \underline{l} > l_0). \tag{6.2.15}$$

The transition density $p_l(s, l_s, T, l_T)$ can be derived from the transition density $p(s, x, t, y)$ for X_t by using the change of variable $l_t = \overline{B} - X_t$. With the default boundary at $\underline{l} = 0$ and initial position $l_0 < 0$, we obtain the desired default probability by applying $\int_0^\infty dl_T$ to both sides of equation (6.2.15). Before doing this, we need to use Kolmogorov's backward equation to derive the transition probability (Arnold(1974) [2]).

Now let us return to our first passage time problem and consider again the diffusion process (6.2.9):

$$dX(s) = \dot{\lambda}(s) ds + dB(s).$$

In our example, we consider the one-dimensional case and set $f(s) = \dot{\lambda}(s)$, which we assume is a smooth function with respect to t. According the definition of diffusion process, the drift vector is $f(s)$, and the one-by-one diffusion matrix is 1. Then using the Kolomogorov's backward equation (2.3.7), and setting $\tau = t - s$, we get the following heat equation with time-dependent drift:

$$\frac{\partial p}{\partial \tau} = f(t-\tau)\frac{\partial p}{\partial x} + 1/2\frac{\partial^2 p}{\partial x^2} \tag{6.2.16}$$

with initial condition $p(0, x) = \delta(x - y)$.

To get the solution, we use the change of variables: $\tilde{x} = x + \int_0^{\tilde{\tau}} f(t-u)du$ and $\tilde{\tau} = \tau$, which allows us to transform the PDE into a standard heat equation:

$$\frac{\partial p}{\partial \tilde{\tau}} = 1/2\frac{\partial^2 p}{\partial \tilde{x}^2}.$$

Using the well known solution of standard heat equation, changing variables $\tilde{x} \to x$, $\tilde{\tau} \to \tau \to t - s$, and noting $\int_0^{t-s} f(t-u)du = \lambda(t) - \lambda(s)$, for any fixed t and y, we get the transition probability density function of (6.2.9):

$$p(s, x, t, y) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(x+\lambda(t)-\lambda(s)-y)^2}{2(t-s)}}. \tag{6.2.17}$$

Recalling the default probability (6.2.2), to get the probability $P[X_T < \overline{K}]$, we just need to take the integration of $p(0, X_0, T, y)$ by $\int_{-\infty}^{\overline{K}} dy$. For the probability (6.2.4): $P[\min_{\{t<T\}} X_t < \overline{B}]$, we get the transition probability for l_t as follows:

$$p_l(s, x, t, y) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(x-\lambda(t)+\lambda(s)-y)^2}{2(t-s)}}. \tag{6.2.18}$$

This transition probability can also be alternatively derived by the Independence

Lemma, please refer to Shreve(2004) [38].

For the default probability (6.2.8), due to lack of independence, we are not able to derive the joint distribution for (6.2.8). So, to approximate this joint distribution, we will use a dynamic bounds condition to mimic the joint conditions:

$$k_t = v + \phi e^t \tag{6.2.19}$$

where v is the lower bound B as defined in 6.2.4, ϕ is a constant such that $k(T) = K$. Here is the graph for this dynamic bound. We will implement the following algorithm using this dynamic bound to compare the result with the work of Qian(2008) [34].

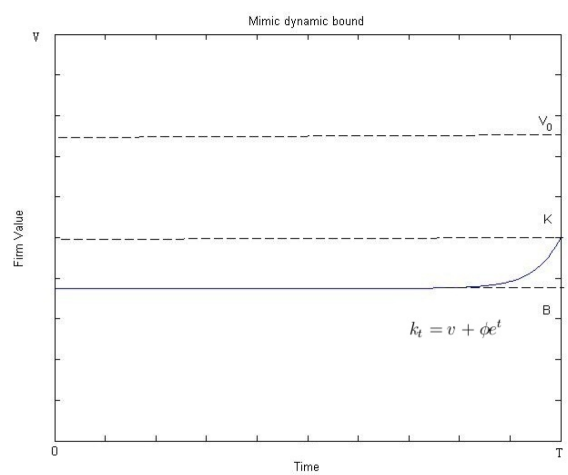

Figure 6.2: Mimic Dynamic bound

First, let us look at how to implement Fortet's method for $P[\min_{\{t<T\}} X_t < \overline{B}]$. With the default boundary at $\underline{l} = 0$ and initial position $l_0 < 0$, integrate both sides of equation (6.2.15) with $\int_0^\infty dl_T$ as follows:

$$\int_0^\infty p_l(0, l_0, T, l_T) dl_T = \int_0^\infty \int_0^T k(0, l_0, s, 0) p_l(s, 0, T, l_T) ds dl_T \qquad \forall (l_T > \underline{l} > l_0). \tag{6.2.20}$$

By discretizing the above equation according to the method used in Collin-Dufresne and Goldstein(2001) [32], we can obtain an approximate solution for the first passage time density. Define $Q(l_0, T)$ to be the risk-neutral probability of the event $[\min_{\{0 \le t \le T\}} : X(t) \le \overline{B}]$, i.e. the probability that default occurs before time T. First, we discretize the time into n equal intervals with $\Delta t = T/n$ and approximate the right hand side of equation (6.2.20) by estimating the value of an integral over an interval using the value of the integrand at the midpoint of the interval. More specifically, define $M(t) = \int_0^\infty p_l(0, l_0, t, l_t) dl_t$ and $N(s,t) = \int_0^\infty p_l(s, l_s = \underline{l}, t, l_t) dl_t$. Then by taking $a_i = i\Delta t$, we can approximate the first two terms of the discretizing approximation of the equation (6.2.20) as:

$$\begin{aligned}
M(a_1) =& \int_0^\infty \int_0^{\Delta t} k(0, l_0, s, 0) p_l(s, 0, t, l_t) ds dl_t \\
=& \Delta t k(0, l_0, \Delta t/2, 0) \int_0^\infty p_l(\Delta t/2, 0, \Delta t, l_{\Delta t}) dl_t \\
\approx& \Delta t k(0, l_0, \Delta t/2, 0) N(a_{1/2}, a_1) \\
M(a_2) =& (\int_0^\infty \int_0^{\Delta t} + \int_0^\infty \int_{\Delta t}^{2\Delta t}) k(0, l_0, s, 0) p_l(s, 0, t, l_t) ds dl_t \\
\approx& \Delta t k(0, l_0, \Delta t/2, 0) N(a_{1/2}, a_2) + \Delta t k(0, l_0, 3\Delta t/2, 0) N(a_{3/2}, a_2).
\end{aligned} \tag{6.2.21}$$

Continuing this pattern, we can obtain n equations for the n unknowns:

$$k(0, l_0, (i - 1/2)\Delta t, l_{(i-1/2))\Delta t} = \underline{l}), \tag{6.2.22}$$

where $i \in 1, ..., n$. If we define

$$q_i = \Delta t k(0, l_0, (i-1/2)\Delta t, l_{(i-1/2))\Delta t} = \underline{l}). \tag{6.2.23}$$

Then we can approximate the default probability as:

$$\tilde{\mathbb{P}}[\min_{\{0 \leq t \leq T\}} : X(t) \leq \overline{B}] \approx Q(l_0, T) = \sum_{i=1}^{n} q_i, \tag{6.2.24}$$

where

$$q_1 = M(a_1)/N(a_{1/2}, a_1), \tag{6.2.25}$$

$$q_i = (1/N(a_{i-1/2}, a_i)[M(a_i) - \sum_{j=1}^{i-1} q_j N(a_{j-1/2}, a_i)], \quad for \quad i = 2, 3, ..., n. \tag{6.2.26}$$

6.2.3 Application of the First Passage Time to the Vasicek Model

By introducing the log-firm value $g_t = \log V_t$ and supposing the interest rate dynamics follow a narrow sense linear model, we get the Longstaff-Schwartz(1992) [23] (LS) model specified under the risk-neutral measure. This model can be characterized by the following Markov system:

$$\begin{aligned} dg_t &= [r_t - \frac{1}{2}\sigma_\nu^2(t)]dt + \sigma_\nu(t) dW_{\nu,t}, \\ dr_t &= -c(t) r_t dt + \theta(t) dt + \sigma_r(t) dW_{r,t}. \end{aligned} \tag{6.2.27}$$

Here, we use a simple special case for r_t: the Vasicek model,

$$dr_t = k(\theta - r_t)dt + \sigma_r(t)dW_{r,t}. \tag{6.2.28}$$

So, $c(t) = k$ and $\theta(t) = k\theta$. In this section, we will consider the dynamic bound following the work of Pierre Collin Dufresne and Robert S. Goldstein(2001)[32] and the dynamic bound (6.2.19).

First, consider that the log-default lower bound follows a dynamical system of the form (Pierre Collin Dufresne and Robert S. Goldstein (2001)):

$$dk_t = \lambda(g_t - v - k_t - \phi(r_t - \theta))dt, \tag{6.2.29}$$

where v is the target log-leverage-ratio and λ is constant factor of the firm debt. This model can be interpreted that when k_t is less than $g_t - v$, the firm tends to increase k_t, and vice-versa. Generally, firms value tend to issue debt when their log-leverage-ratio falls below v, and tend to replace maturing debt when their log-leverage-ratio is above v. Here we also assume the threshold is a decreasing function of the interest rate following the work of Pierre Collin Dufresne and Robert S. Goldstein (2001).

Letting $l_t = k_t - g_t$ and applying Ito's lemma, we get

$$dl_t = \lambda(\bar{l}(r_t) - l_t)dt - \sigma_\nu dW_{\nu,t} \tag{6.2.30}$$

where $\bar{l}(r_t) = \frac{\sigma_\nu^2}{2\lambda} - v + \phi\theta - r_t(\frac{1}{\lambda} + \phi)$, a decreasing function of r_t. Now, set $dW_{r,t} = dW_{1,t}$ and $W_{\nu,t} = \rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}$ (Shreve(2004) [38]), where $dW_{1,t}$ and $dW_{2,t}$ are uncorrelated processes. Then we can rewrite the coupled dynamics of interest rate and firm value as

$$
\begin{aligned}
dr_t &= k(\theta - r_t)dt + \sigma_r(t)dW_{1,t}, \\
dl_t &= (-(\lambda\phi + 1)r_t - \lambda l_t + \lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(t)^2)dt - \nu_1(t)dW_{1,t} - \nu_3(t)dW_{2,t}
\end{aligned} \tag{6.2.31}
$$

where $\nu_1(t) = \rho\sigma_\nu(t)$ and $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$.

Thus, we can write the coupled dynamics of r_t, l_t in matrix form as

$$
\begin{aligned}
\begin{bmatrix} dr_t \\ dl_t \end{bmatrix} &= (\begin{bmatrix} -k & 0 \\ -(\lambda\phi+1) & -\lambda \end{bmatrix} \begin{bmatrix} r_t \\ l_t \end{bmatrix} + \begin{bmatrix} k\theta \\ \lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(t)^2 \end{bmatrix})dt \\
&\quad + \begin{bmatrix} \sigma_r(t) & 0 \\ -\nu_1(t) & -\nu_3(t) \end{bmatrix} \begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix}.
\end{aligned} \tag{6.2.32}
$$

We will use the notation

$$
\begin{aligned}
A &= \begin{bmatrix} -k & 0 \\ -(\lambda\phi+1) & -\lambda \end{bmatrix}, \quad Y_t = \begin{bmatrix} dr_t \\ dl_t \end{bmatrix}, \quad a(t) = \begin{bmatrix} k\theta \\ \lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(t)^2 \end{bmatrix}, \\
B(t) &= \begin{bmatrix} \sigma_r(t) & 0 \\ -\nu_1(t) & -\nu_3(t) \end{bmatrix}, \quad dZ_t^T = \begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix}, \quad Y_0 = c = \begin{bmatrix} r_0 \\ l_0 \end{bmatrix}.
\end{aligned}
$$

Then we see that the the vector process X_t follows the narrow sense linear stochastic differential equation:

$$
dY_t = (AY_t + a(t))dt + BdZ^T, \quad Y_{t_0} = c.
$$

For $t_0 = u$, we see by Arnold(1974) [2], the solution of the above SDE is

$$Y_t = e^{A(t-u)}c + \int_u^t e^{A(t-s)}(a(s)ds + B(s)dZ^T) \tag{6.2.33}$$

where $e^{A(t-u)} = \sum_{i=0}^{\infty} A^i \frac{(t-u)^i}{i!}$.

Here

$$A = \begin{bmatrix} -k & 0 \\ -(\lambda\phi + 1) & -\lambda \end{bmatrix}$$

is a lower triangular matrix, so it is not hard to compute e^A and $e^{A(t-u)}$. We need to consider two case: I) $k \neq \lambda$; II) $k = \lambda$.

Case I $k \neq \lambda$: In this case, matrix A has two distinct eigenvectors, which means A can always be diagonalized. Let

$$Q = \begin{bmatrix} 0 & \frac{k-\lambda}{1+\lambda\phi} \\ 1 & 1 \end{bmatrix},$$

the eigenvector matrix of A. Then the diagonalized matrix of A is

$$D = Q^{-1}AQ = \begin{bmatrix} -\lambda & 0 \\ 0 & -k \end{bmatrix}.$$

Because $QQ^{-1} = I$, $QD^2Q^{-1} = QD(Q^{-1}Q)DQ^{-1} = (QDQ^{-1})^2 = A^2$, we get $QD^iQ^{-1} = (QDQ^{-1})^i = A^i$. Consequently we have

$$\begin{aligned} e^A = e^{QDQ^{-1}} &= \sum_{i=0}^{\infty} \frac{(QDQ^{-1})^i}{i!} = \sum_{i=0}^{\infty} \frac{QD^iQ^{-1}}{i!} \\ &= Q\sum_{i=0}^{\infty} \frac{D^i}{i!}Q^{-1} = Qe^DQ^{-1} \end{aligned}$$

Similarly, we get $e^{A(t-u)} = Qe^{D(t-u)}Q^{-1}$. For D, as

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}^i = \begin{bmatrix} a^i & 0 \\ 0 & b^i \end{bmatrix},$$

we get

$$e^{D(t-u)} = exp \begin{bmatrix} -\lambda(t-u) & 0 \\ 0 & -k(t-u) \end{bmatrix} = \begin{bmatrix} e^{-\lambda(t-u)} & 0 \\ 0 & e^{-k(t-u)} \end{bmatrix}$$

which yields

$$e^{A(t-u)} = Qe^{D(t-u)}Q^{-1} = \begin{bmatrix} e^{-k(t-u)} & 0 \\ \frac{(1+\lambda\phi)(e^{-k(t-u)}-e^{-\lambda(t-u)})}{k-\lambda} & e^{-\lambda(t-u)}. \end{bmatrix}$$

Case II $k = \lambda$: For this case, the matrix A only has one independent eigenvector, which means it cannot be diagonalized. But we can rewrite A as:

$$\begin{bmatrix} -k & 0 \\ -(1+k\phi) & -k \end{bmatrix} = \begin{bmatrix} -k & 0 \\ 0 & -k \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -(1+k\phi) & 0 \end{bmatrix}.$$

Due to the fact of

$$\begin{bmatrix} 0 & 0 \\ -(1+k\phi)(t-u) & 0 \end{bmatrix}^i = 0, \quad for \quad i \geq 2$$

we see that

$$exp\begin{bmatrix} 0 & 0 \\ -(1+k\phi)(t-u) & 0 \end{bmatrix} = I + \begin{bmatrix} 0 & 0 \\ -(1+k\phi)(t-u) & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ -(1+k\phi)(t-u) & 1 \end{bmatrix}.$$

Thus we obtain

$$\begin{aligned} e^{A(t-u)} &= exp\begin{bmatrix} -k(t-u) & 0 \\ 0 & -k(t-u) \end{bmatrix} exp\begin{bmatrix} 0 & 0 \\ -(1+k\phi)(t-u) & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^{-k(t-u)} & 0 \\ 0 & e^{-k(t-u)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -(1+k\phi)(t-u) & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^{-k(t-u)} & 0 \\ -(1+k\phi)(t-u)e^{k(t-u)} & e^{-k(t-u)} \end{bmatrix}. \end{aligned}$$

Since generically, we have $k \neq \lambda$, we first consider case I. Substitute $e^{A(t-u)}$ into the solution of Y(t) (6.2.33) and $t_0 = 0$, to get the solution of (6.2.31) as given by:

$$\begin{aligned} r_t &= e^{-kt}r_0 + \int_0^t k\theta e^{-k(t-s)}ds + \int_0^t \sigma_r(t)e^{-k(t-s)}dW_{1,t} \\ l_t &= \bar{l}(0)r_0 + e^{-\lambda t}l_0 + \int_0^t [k\theta\bar{l}(s) + e^{-\lambda(t-s)}(\lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(s)^2)]ds \\ &\quad + \int_0^t [\bar{l}(s)\sigma_r(s) - e^{-\lambda(t-s)}\nu_1(s)]dW_{1,s} - \int_0^t e^{-\lambda(t-s)}\nu_3(s)dW_{2,s} \end{aligned}$$

where $\bar{l}(0) = \frac{1+\lambda\phi}{k-\lambda}(e^{-kt} - e^{-\lambda t})$.

From above expression for l_t we can write

$$l_t = l_t(det) + l_t(ran) \tag{6.2.34}$$

where

$$\begin{aligned} l_t(det) =& \bar{l}(0)r_0 + e^{-\lambda t}l_0 + \int_0^t [k\theta\bar{l}(s) + e^{-\lambda(t-s)}(\lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(s)^2)]ds \\ l_t(ran) =& \int_0^t [\bar{l}(s)\sigma_r(s) - e^{-\lambda(t-s)}\nu_1(s)]dW_{1,s} - \int_0^t e^{-\lambda(t-s)}\nu_3(s)dW_{2,s}. \end{aligned}$$

From this expression we see that l_t follows a normal distribution,

$$l_t \sim N(\mu(t), \sigma(t)^2), \tag{6.2.35}$$

where $\mu(t) = l_t(det)$, and $\sigma(t)^2 = \sigma_1(t)^2 + \sigma_2(t)^2$, with $\sigma_1(t)^2 = \int_0^t [\bar{l}(s)\sigma_r(s) - e^{-\lambda(t-s)}\nu_1(s)]^2 ds$, and $\sigma_2(t)^2 = \int_0^t [e^{-\lambda(t-s)}\nu_3(s)]^2 ds$, and where $\nu_1(t) = \rho\sigma_\nu(t)$, $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$ and $\bar{l}(s) = \frac{1+\lambda\phi}{k-\lambda}(e^{-k(t-s)} - e^{-\lambda(t-s)})$.

Now, to compare the numerical result of Fortet's method to Qian's method, let us look at the dynamic bound (6.2.19) that mimics the joint default condition of Giesecke [12]. Letting $l_t = k_t - g_t = v + \phi e^t - g_t$ and applying Ito's lemma, we get

$$dl_t = (\phi e^t + \frac{1}{2}\sigma_v(t)^2 - r_t)dt - \sigma_\nu dW_{\nu,t}. \tag{6.2.36}$$

Now, set $dW_{r,t} = dW_{1,t}$ and $W_{\nu,t} = \rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}$ (Shreve(2004) [38]), so that $dW_{1,t}$ and $dW_{2,t}$ are uncorrelated processes. Then we can rewrite the coupled dynamics of interest rate and firm value as

$$\begin{aligned} dr_t &= k(\theta - r_t)dt + \sigma_r(t)dW_{1,t}, \\ dl_t &= (\phi e^t + \frac{1}{2}\sigma_v(t)^2 - r_t)dt - \nu_1(t)dW_{1,t} - \nu_3(t)dW_{2,t} \end{aligned} \tag{6.2.37}$$

where $\nu_1(t) = \rho\sigma_\nu(t)$ and $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$. Thus, we can write the coupled dynamics of r_t, l_t in matrix form as

$$\begin{bmatrix} dr_t \\ dl_t \end{bmatrix} = (\begin{bmatrix} -k & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} r_t \\ l_t \end{bmatrix} + \begin{bmatrix} k\theta \\ \phi e^t + \frac{1}{2}\sigma_\nu(t)^2 \end{bmatrix})dt + \begin{bmatrix} \sigma_r(t) & 0 \\ -\nu_1(t) & -\nu_3(t) \end{bmatrix} \begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix}. \tag{6.2.38}$$

We will use the notation

$$\begin{aligned} A &= \begin{bmatrix} -k & 0 \\ -1 & 0 \end{bmatrix}, \quad Y_t = \begin{bmatrix} dr_t \\ dl_t \end{bmatrix}, \quad a(t) = \begin{bmatrix} k\theta \\ \phi e^t + \frac{1}{2}\sigma_\nu(t)^2 \end{bmatrix}, \\ B(t) &= \begin{bmatrix} \sigma_r(t) & 0 \\ -\nu_1(t) & -\nu_3(t) \end{bmatrix}, \quad dZ_t^T = \begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix}, \quad Y_0 = c = \begin{bmatrix} r_0 \\ l_0 \end{bmatrix}. \end{aligned}$$

Then we see that the the vector process X_t follows the narrow sense linear stochastic differential equation:

$$dY_t = (AY_t + a(t))dt + BdZ^T, \quad Y_{t_0} = c.$$

For $t_0 = u$, we see by Arnold(1974) [2], the solution of the above SDE is

$$Y_t = e^{A(t-u)}c + \int_u^t e^{A(t-s)}(a(s)ds + B(s)dZ^T). \tag{6.2.39}$$

Following the similar computation, we obtain

$$
\begin{aligned}
e^{A(t-u)} &= exp\begin{bmatrix} -k(t-u) & 0 \\ 0 & 0 \end{bmatrix} exp\begin{bmatrix} 0 & 0 \\ -(t-u) & 0 \end{bmatrix} \\
&= \begin{bmatrix} e^{-k(t-u)} & 0 \\ -(t-u) & 1 \end{bmatrix}.
\end{aligned}
$$

Substitute $e^{A(t-u)}$ into the solution of Y(t) (6.2.39) and $t_0 = 0$, to get the solution of (6.2.37) as given by:

$$
\begin{aligned}
r_t &= e^{-kt}r_0 + \int_0^t k\theta e^{-k(t-s)}ds + \int_0^t \sigma_r(t)e^{-k(t-s)}dW_{1,t} \\
l_t &= -r_0 t + l_0 + \int_0^t [-(t-s)k\theta + \phi e^s + \frac{1}{2}\sigma_\nu(s)^2)]ds \\
&\quad - \int_0^t [(t-s)\sigma_r(s) + \nu_1(s)]dW_{1,s} - \int_0^t \nu_3(s)dW_{2,s}.
\end{aligned}
$$

From this expression we see that l_t follows a normal distribution,

$$
l_t \sim N(\mu(t), \sigma^2(t)), \tag{6.2.40}
$$

where

$$
\begin{aligned}
\mu(t) &= -r_0 t + l_0 + \int_0^t [-(t-s)k\theta + \phi e^s + \frac{1}{2}\sigma_\nu(s)^2)]ds \\
\sigma(t)^2 &= \int_0^t [(t-s)\sigma_r(s) + \nu_1(s)]^2 ds + \int_0^t [\nu_3(s)]^2 ds
\end{aligned}
$$

and $\nu_1(t) = \rho\sigma_\nu(t)$, $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$.

With above results (6.2.35) and (6.2.40), by using the algorithm introduced by previous section from (6.2.21) to (6.2.24), and setting all the parameters to be the same as shown the figure (6.3), we get we get the default probability figure (6.4). We also implemented the algorithm for the default probability with the constant bound (6.2.4) $P[\min_{\{t<T\}} V_t < B]$ and put these two results together in the graph to compare them with the numerical result figure (6.3) computed by Qian(2008) [34]. From the graph, we can see that the results obtained by Fortet's method and Qian's method have similar patterns, while the probabilities generated by Fortet's method are a bit higher. In figure (6.7) and figure (6.8), we will combine these two graphs with the graph generated by trinomial tree method in the next section.

Also, using the dynamic bound from Pierre Collin Dufresne and Robert S. Goldstein (2001), we get the figure (6.5). We can see that at the first stage, the default probability increases rapidly. As the maturity becomes longer, the default probability increases at a slower rate. Graphs of the default probability as a function of the maturity for various values of ρ are presented in figure (6.5).

6.3 Firm Default Probability Using a Nonlinear Interest Rate Model

In this section, we again consider the computation of the firm default probability, but here we replace the (linear) Vasicek model for r_t by the previously considered nonlinear model in Chap. 5. Following the same steps that led to equation(6.2.31), we see that the dynamics of interest rates and firm values are as follows:

$$dr_t = (\theta(t) - r_t)r_t dt + \sigma_r(t) r_t dW_{1,t},$$
$$dl_t = (-(\lambda\phi + 1)r_t - \lambda l_t + \lambda(\phi\theta - v) + \frac{1}{2}\sigma_\nu(t)^2)dt - \nu_1(t)dW_{1,t} - \nu_3(t)dW_{2,t} \tag{6.3.1}$$

where $\nu_1(t) = \rho\sigma_\nu(t)$ and $\nu_3(t) = \sqrt{1-\rho^2}\sigma_\nu(t)$.

Because of the nonlinear drift in the dynamics of r_t, we are not able to derive a closed form expression for the distribution of l_t using the previous subsection's method. As a result we are also not able to derive the transition probability for the l_t process, which means we cannot apply the Fortet integral equation. Consequently we resort to another approximation scheme and propose a new algorithm for implementing a trinomial tree for the coupled process (r_t, l_t). We describe this algorithm in terms of the general coupled process:

$$dr_t = \alpha(r_t, t)dt + \sigma_1 dW_{1,t},$$
$$dl_t = \xi(r_t, l_t)dt + \sigma_2 d\tilde{W}_{2,t}$$

where $W_{1,t}$ and $\tilde{W}_{2,t}$ are correlated Brownian motion processes with $dW_{1,t}d\tilde{W}_{2,t} = \rho dt (0 < \rho < 1)$ and the functions α, ξ, and the constants σ_1 and σ_2 are all appropriate to our model. We can use the standard decomposition trick to rewrite this system as

$$\begin{aligned} dr_t &= \alpha(r_t, t)dt + \sigma_1 dW_{1,t} \\ dl_t &= \xi(r_t, l_t)dt + \sigma_2\rho dW_{1,t} + \sigma_2\sqrt{1-\rho^2}dW_{1,t} \end{aligned} \tag{6.3.2}$$

where $W_{1,t}$ and $W_{2,t}$ are independent Brownian motion processes. A discrete approximation to the system (6.3.2) can be expressed as

$$
\begin{aligned}
&r(t+\Delta t)-r(t) \approx \alpha(r_t,t)\Delta t+\sigma_1 Z_1\sqrt{\Delta t}\\
&l(t+\Delta t)-l(t) \approx \xi(r(t),l(t))\Delta t+\sigma_2\rho Z_1\sqrt{\Delta t}+\sigma_2\sqrt{1-\rho^2}Z_2\sqrt{\Delta t}
\end{aligned}
\tag{6.3.3}
$$

where Z_1 and Z_2 are independent standard normal random variables. Suppose system (6.3.2) is initialized at (x_0, l_0) at time zero, choose a time-step $\Delta t > 0$, and set(following the Hull-White[18] suggestion for example)

$$\Delta r = \sigma_1\sqrt{3\Delta t}, \quad \Delta l = \sigma_2\sqrt{1-\rho^2}\sqrt{3\Delta t}.$$

At each time step $i = 1, 2, 3, ...$ the nodes of the tree are

$$(r_j, l_k), \quad r_j = r_0 + j\Delta r, \quad l_k = l_0 + k\Delta l, \quad j, k \in integers.$$

Our goal is to build a tree to approximate the transition probabilities

$$p(r(i+1) = r_{j^*}, l(i+1) = l_{k^*} | r(i) = r_j, l(i) = l_k)$$

In words, this is the probability that, conditioned on being at node (r_j, l_k) at time i, we make a transition to node (r_{j^*}, l_{k^*}) at time $i+1$. For convenience, we write $p_i(r_{j^*}, l_{k^*}|r_j, l_k)$ instead of $p(r(i+1) = r_{j^*}, l(i+1) = l_{k^*}|r(i) = r_j, l(i) = l_k)$. Using the definition of conditional probability we can write

$$p_i(r_{j^*}, l_{k^*}|r_j, l_k) = p_i(r_{j^*}, l_{k^*}|r_{j^*}, r_j, l_k)p_i(r_{j^*}, l_k|r_j, l_k). \tag{6.3.4}$$

Let us first focus on the second factor on the right-hand side of (6.3.4). Since $l_k$is unchanged in this transition and since the SDE for r_k is decoupled from the

SDE for l_t, we have

$$p_i(r_{j^*}, l_k | r_j, l_k) = p_i(r_{j^*} | r_j).$$

This allows us to approximate the second factor on the right-hand side of (6.3.4) by the trinomial-tree process we already used for the interest rate process by itself. Having done this, the other transition probability $p_i(r_{j^*}, l_{k^*} | r_{j^*}, r_j, l_k)$, can be approximated by a trinomial-tree that is constructed by matching the first and second conditional moments of the random variable $l(t+\Delta t) - l(t)$. Specially, the branching probabilities for the l-process tree(conditioned on the values of the interest rate obtained from the r-process tree) are obtained by matching the conditional moments

$$E[l(t+\Delta t) - l(t) | r(t+\Delta t), r(t), l(t)] \quad and \quad E[(l(t+\Delta t) - l(t))^2 | r(t+\Delta t), r(t), l(t)].$$

We can compute the conditional moments by taking the conditional expectation on both sides of the second equation in (6.3.3) conditioned of the random variables $r(t+\Delta t), r(t)$ and $l(t)$ and using the standard properties of conditional expectation(independence and "taking out what is known"). To this end note that: (1) $\xi(r(t), l(t))$ is "known" with respect to $r(t+\Delta t), r(t)$; (2) Z_1 is also "known" with respect to $r(t+\Delta t), r(t)$, and $l(t)$ because of the first equation in (6.3.3); (3) Z_2 is also "known" with respect to $r(t+\Delta t), r(t)$, and $l(t)$ because it is independent from $r(t)$ and $l(t)$ and Z_1, and it is also independent from $r(t+\Delta t)$, since $r(t+\Delta t)$ is a function of the random variables $r(t)$ and Z_1. If we write E_C as shorthand for the conditional expectation with respect to

$r(t+\Delta t), r(t)$, $l(t)$, we apply E_C to both sides of the second equation in (6.3.3):

$$E_C[l(t+\Delta t) - l(t)] = E_C[\xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t} + \sigma_2 \sqrt{1-\rho^2} Z_2 \sqrt{\Delta t}]$$

$$= \xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t} + E_C[\sigma_2 \sqrt{1-\rho^2} Z_2 \sqrt{\Delta t}]$$

$$(taking\ out\ what\ is\ known)$$

$$= \xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t} + \sigma_2 \sqrt{1-\rho^2} \sqrt{\Delta t} E[Z_2]$$

$$(E_C[Z_2] = E[Z_2]\ by\ independence\ of\ Z_2\ from\ x(t+\Delta t), x_t, l(t))$$

$$= \xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t}\ (since\ E[Z_2] = 0).$$

A similar computation for the second conditional moment gives

$$E_C[(l(t+\Delta t) - l(t))^2] = (\xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t})^2 + \sigma_2^2 (1-\rho^2)^2 \Delta t.$$

To approximate the first factor on the right-hand side of (6.3.4), we carry out the following steps:

(1) Assume at time step i we are at node (r_j, l_k). Use the trinomial-tree algorithm already devised to approximate the transition probabilities to approximate the transition probabilities of the r-process by itself. This means that we will use a "branching process" to find the node r_{j^*} so that at time step $i+1$. The r-process will be assumed to be at one of the three nodes $r_{j^*-1}, r_{j^*}, r_{j^*+1}$ and the algorithm will also give us approximations to the three transition probabilities

$$p_i(r_{j^*-1}|r_j), p_i(r_{j^*}|r_j), p_i(r_{j^*+1}|r_j). \tag{6.3.5}$$

This determines the second factor on the right-hand side of (6.3.4).

(2)Suppose for example we make a transition from r_j to r_{j^*+1} with probability

$p_i(r_{j^*+1}|r_j)$. Conditioned on this information, the first equation in (6.3.3) can be written as

$$r_{j^*+1} - r_j \approx \alpha(r_j, i)\Delta t + \sigma_1 Z_1 \sqrt{\Delta t} \tag{6.3.6}$$

so this conditioning effectively determines the value of the random variable Z_1 and we can solve for this value via equation (6.3.6).

(3) Now set $t = i$ and $r(t) = r(i) = r_j$ in the second equation in (6.3.3) and view the value of the random variable Z_1 as being determined (and so a constant) according to the previous step because we are conditioning on the three values r_j, r_{j^*+1} and l_k. We get

$$l(i+1) - l_k \approx \xi(r_j, l_k)\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t} + \sigma_2 \sqrt{1-\rho^2} Z_2 \sqrt{\Delta t}. \tag{6.3.7}$$

It is important to note that although r_{j^*+1} does not seem to appear explicitly in (6.3.7), this equation does in fact depend on r_{j^*+1} because we used that value to determine Z_1. Then we can use another one-dimensional trinomial tree process to approximate the transition probabilities $p_i(r_{j^*}, l_{k^*}|r_{j^*}, r_j, l_k)$. Specifically, as we discussed above, use the branching process to determine k^* so that (conditioned on r_j, r_{j^*+1}, x_j, and l_k) we transition to one of the three nodes $l_{k^*-1}, l_{k^*}, l_{k^*+1}$, with the transition probabilities

$$p_i(r_{j^*+1}, l_{k^*-1}|r_{j^*}, r_j, l_k), p_i(r_{j^*+1}, l_{k^*}|r_{j^*}, r_j, l_k), p_i(r_{j^*+1}, l_{k^*+1}|r_{j^*}, r_j, l_k). \tag{6.3.8}$$

These probabilities can be computed as in the one-dimensional trinomial tree process so as to match the first moment of (6.3.7)

$$\xi(r_j, l_k)\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t}$$

and the second moment

$$(\xi(r(t), l(t))\Delta t + \sigma_2 \rho Z_1 \sqrt{\Delta t})^2 + \sigma_2^2(1-\rho^2)^2 \Delta t$$

(note that in (6.3.7) the only random quantity is Z_2 due to the condition that determines Z_1).

So, by using (6.3.5) and (6.3.8) we can build a coupled tree that from the node (r_j, l_k) at time step i allows transitions to $3 \times 3 = 9$ possible nodes at time step $i+1$. Once we have the distribution of l_t, we can compute the default probability of the firm value at mature time T, denoting as P_T. We refer to this as the **coupled-tree-algorithm**.

Step 1: Following the trinomial-tree-algorithm generated by Chapter 5.1, derive the trinomial tree for r_t to get the probability $D_r(i, j_r)$ for each trinomial tree node j_r at time step i, for $i \leq T$:

$$D_r(i, j_r) = \sum_j D_r(i-1, j) p_r(i-1, j) \tag{6.3.9}$$

where j is determined by the paths leading to node (i, j_r), and $p_r(i-1, j)$ is the probability from $r(i-1, j)$ leading to (i, j_r).

Step 2: Similarly, by the trinomial-tree-algorithm as step 1, derive the trinomial tree for l_t to get the distribution $D_l(i, j_l)$ for each trinomial tree node j_l at time step i, for $i \leq T$:

$$D_l(i, j_l) = \sum_j D_l(i-1, j) p_l(i-1, j) \tag{6.3.10}$$

where j is determined by the paths leading to node (i, j_l), and $p_l(i-1, j)$ is the probability from $l(i-1,j)$ leading to (i,j_l).

Step 3: At time step i, we denote the coupled tree as (l,r), and compute to get the distribution $D_{lr}(i,j_{lr})$ for each tree node j_{lr} at time step i, for $i \leq T$:

$$D_{lr}(i,j_{lr}) = \sum_{j^*,j^{**}} D_r(i-1,j^*)p_r(i-1,j^*)D_l(i-1,j^{**})p_l(i-1,j^{**}) \quad (6.3.11)$$

where j^* is the r_t tree tranche path, and j^{**} is the l_t tree tranche path at time step $i-1$ determined by the coupled-paths leading the (l,r) tree to node (i,j_{lr}).

Step 4: At time step k, where $0 < k \leq i$, we compute the probability

$$P_{lr_k} = (1-P_{k-1})\sum_j D_{lr}(k,j) \quad (6.3.12)$$

where j is summed over all the nodes such that $lr_{k,j} > 0$, i.e. passing the bound.

Step 5: Finally, we have the recursive algorithm as follows to get P_T.

$$P_T = 1-(1-P_{T-1})(1-P_{lr_T}). \quad (6.3.13)$$

First, let us use the above coupled tree algorithm to compute the default probability for constant bounds as previous sections. We implemented this algorithm with Vasicek model using the same parameters as in the previous two sections for Qian's method and Fortet's method. For the default conditions (6.2.4) and (6.2.5), we have both of the results in figure (6.6). Then we compare the three methods for the default conditions (6.2.4) in figure (6.7) and the default conditions (6.2.5) in figure (6.8). We can see that results are quit consistent. This provides some justification to apply it to the population growth model.

Also, with the dynamic bounds proposed in Fortet's method and for different values of ρ, we implemented this coupled trinomial tree with the Vasicek model. Compared the result figure(6.5) and the figure (6.9) and we got from previous section, we see the results are roughly consistent. But the computation time of the coupled trinomial tree is much faster. Also, we derive the result for the joint default probability (6.2.5) using coupled trinomial tree method in figure (6.10).

Applying this approach to the population growth model which we proposed in the previous chapter, we can see the default trend with Dufresne and Goldstein's dynamic bound (6.2.29) and $V_T < K$ using our model. Similarly, we also implement the algorithm for different values of ρ for comparison. To better present our method, we use our model to calibrate the U.S. Treasury bond's term structure. The source of this term structure time series is MSCI's database and Yahoo Finance, and we linearly interpolate the curve into 20 one-year time steps as follow:

Maturity	1	2	3	4	5	6	7	8	9	10
R (%)	0.175	0.258	0.354	0.471	0.616	0.783	0.968	1.16	1.356	1.554
Maturity	11	12	13	14	15	16	17	18	19	20
R (%)	1.748	1.914	2.053	2.164	2.252	2.319	2.365	2.404	2.441	2.478

Now, using the the r_t distribution result computed by the algorithm from Chapter 5, we get two comparison default probability time series. One is the figure (6.11) with constant $\theta = 0.2$, and the second one is the figure (6.12) with $\theta(t)$ which calibrates to the U.S. Treasury bond's term structure. We can see both of these figures have the similar pattern as the Vasicek figure (6.9), with a very low default probability in the beginning stage, which then increases rapidly in the short term, and then increases somewhat more slowly in the long term. The default probabilities for the nonlinear interest rate model are consistently higher over the entire time range than for the Vasicek model. This is reasonable since the population growth model is a nonlinear model (which has larger variance of the short rate when compared to the Vasicek model; refer to Chapter 5, figure 5.3). The higher variance of the interest rate model will cause the higher variance of the coupled process, which causes higher risk. Also, we can see that the default probability pattern also reflects that the positive or negative correlation coefficients will have similar effect on the result, which we believe is reasonable for the risk factor.

Based on the above analysis, the method proposed in this chapter appears to provide a good alternative to approximate the default probability if the first passage time approach is infeasible.

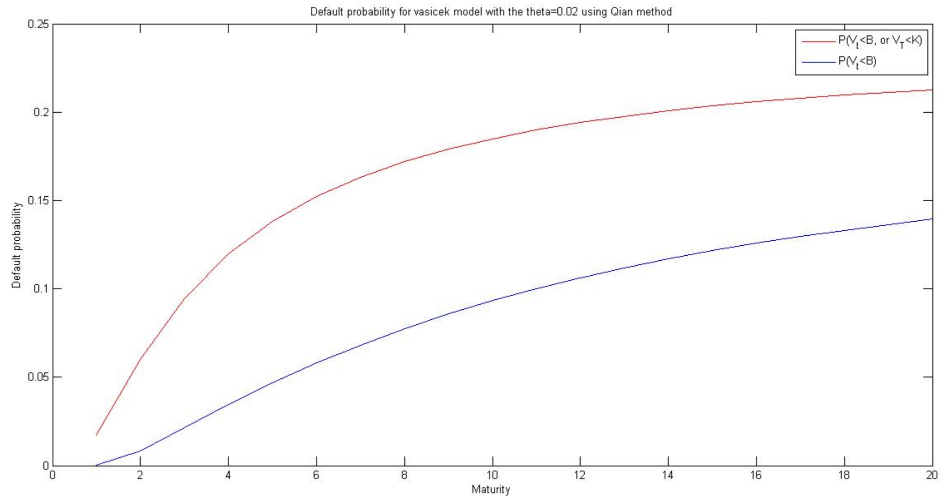

Figure 6.3: Default Probability of the firm using the Vasicek model and Qian's method, with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $K = 100$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\rho = 0$, $r_0 = 0.04$.

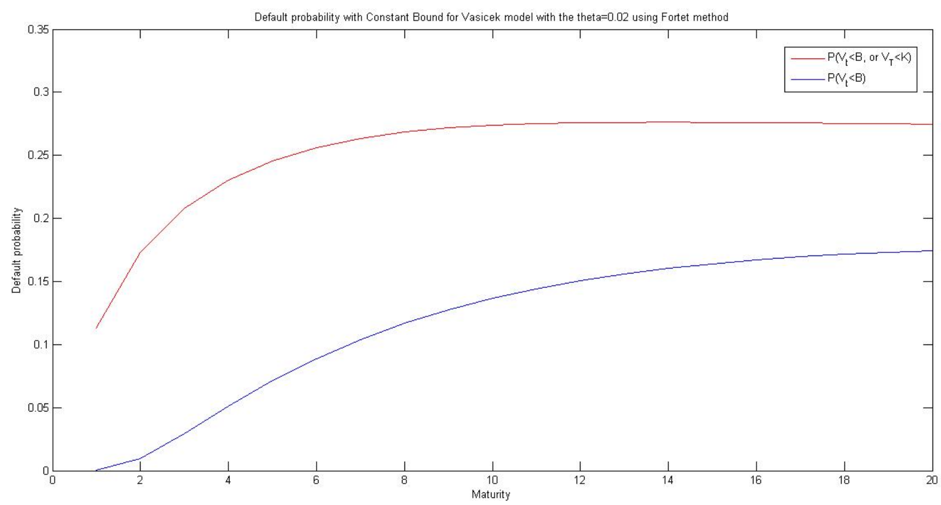

Figure 6.4: Default Probability of the firm using the Vasicek model and Fortet's method, with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $K = 100$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\rho = 0$, $r_0 = 0.04$, $v = 75$, $\phi = 5.1529 * 10^{-8}$.

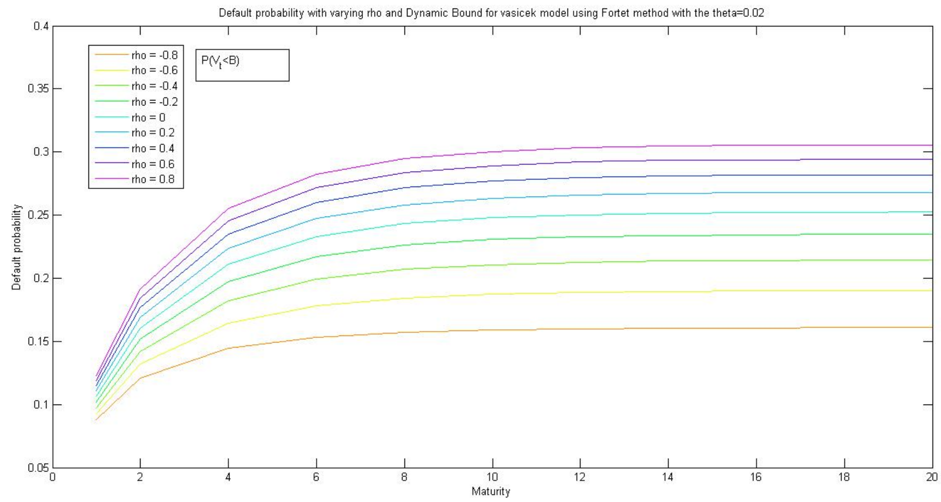

Figure 6.5: Default Probability of the firm using the Vasicek model and Fortet's method using Pierre Collin Dufresne and Robert S. Goldstein's dynamic bound, with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $r_0 = 0.04$, $v = 0.6$, $\lambda = 0.18$, $\phi = 2.8$.

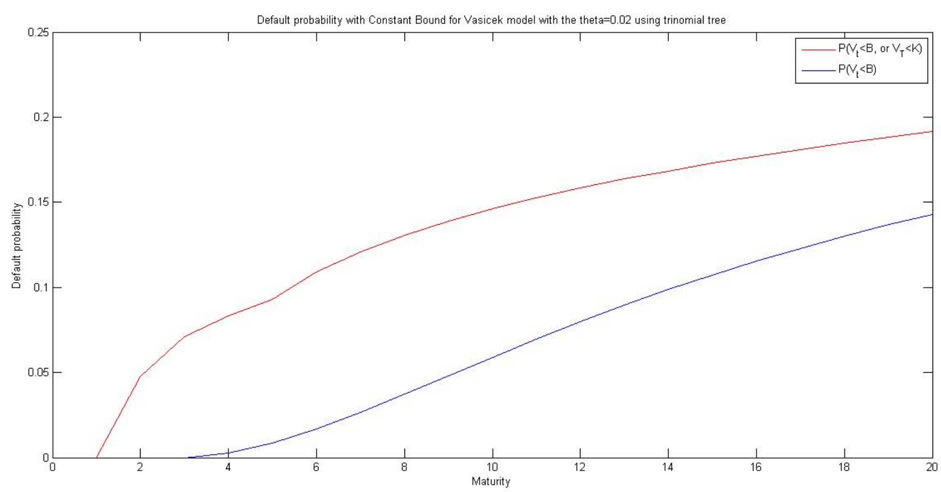

Figure 6.6: Default Probability of the firm using the Vasicek model and trinomial tree method, with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $K = 100$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\rho = 0$, $r_0 = 0.04$.

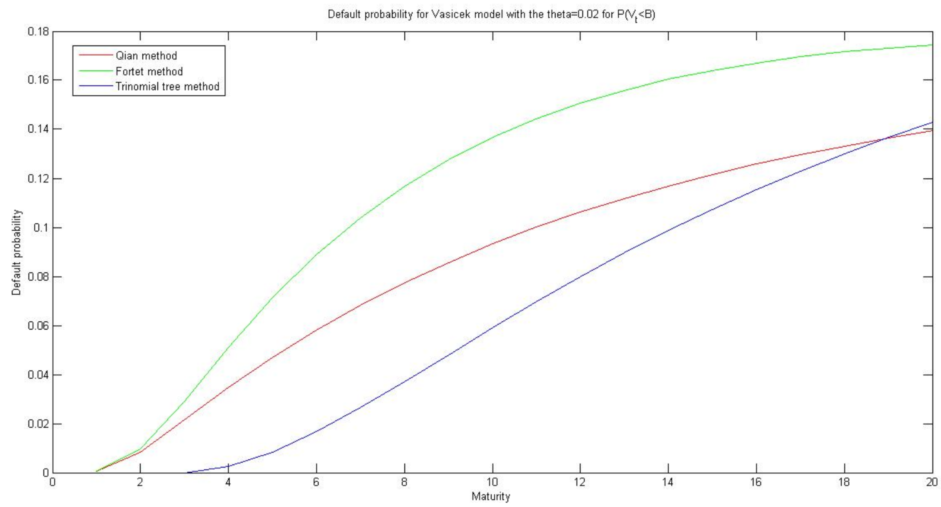

Figure 6.7: Default Probability Result Comparison using the Vasicek model with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\rho = 0$, $r_0 = 0.04$.

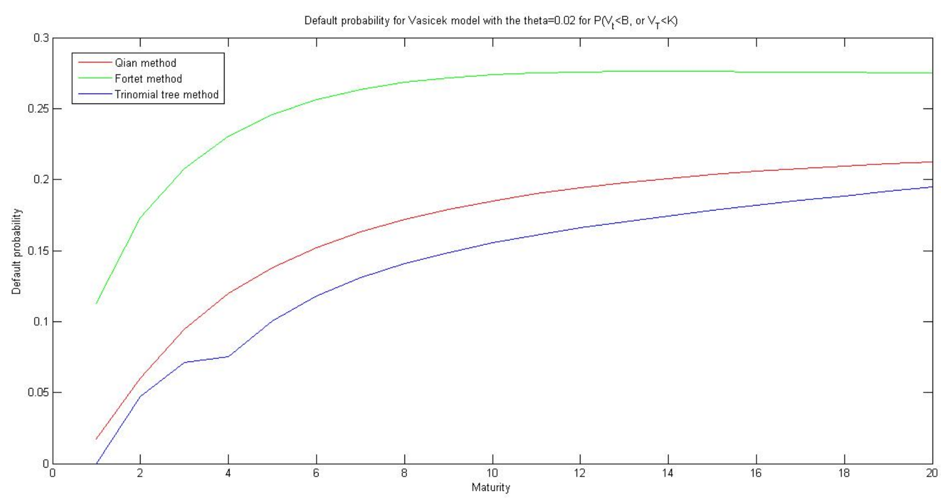

Figure 6.8: Default Probability Result Comparison using the Vasicek model with all the parameters in the algorithm as follows: $B = 75$, $V_0 = 150$, $K = 100$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\rho = 0$, $r_0 = 0.04$.

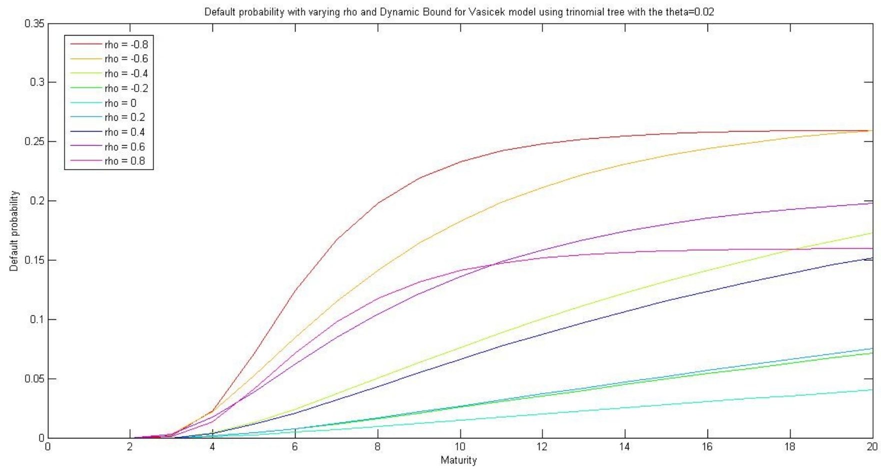

Figure 6.9: Default probability with varying rho from -0.8 to 0.8 by 0.2 with Pierre Collin Dufresne and Robert S. Goldstein's dynamic bound for Vasicek model using trinomial tree method, the parameters in the algorithm as follows $B = 75$, $V_0 = 150$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $r_0 = 0.04$, $\lambda = 0.18$, $\phi = 2.8$, $v = 0.6$, $r_0 = 0.06$.

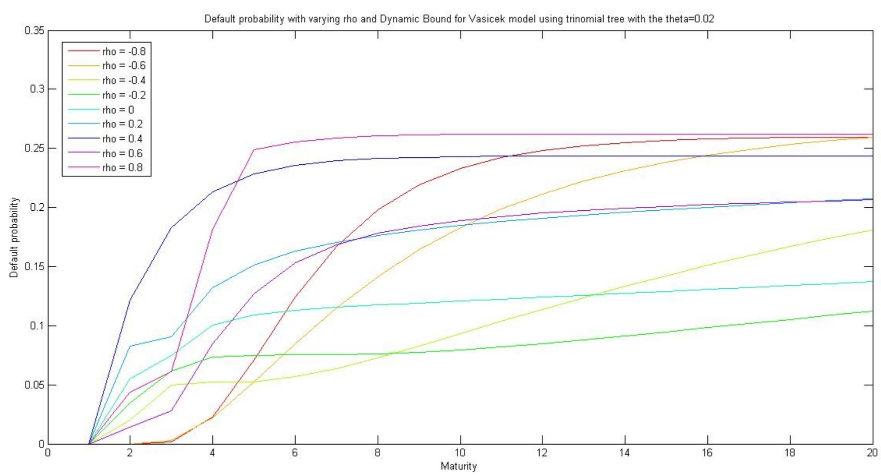

Figure 6.10: Default probability with varying rho from -0.8 to 0.8 by 0.2 with Pierre Collin Dufresne and Robert S. Goldstein's dynamic bound and $V_T < K$ for Vasicek model using trinomial tree method, the parameters in the algorithm as follows $B = 75$, $V_0 = 150$, $K = 100$, $k = 0.5$, $\theta = 0.04$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $r_0 = 0.04$, $\lambda = 0.18, \phi = 2.8, v = 0.6, r_0 = 0.06$.

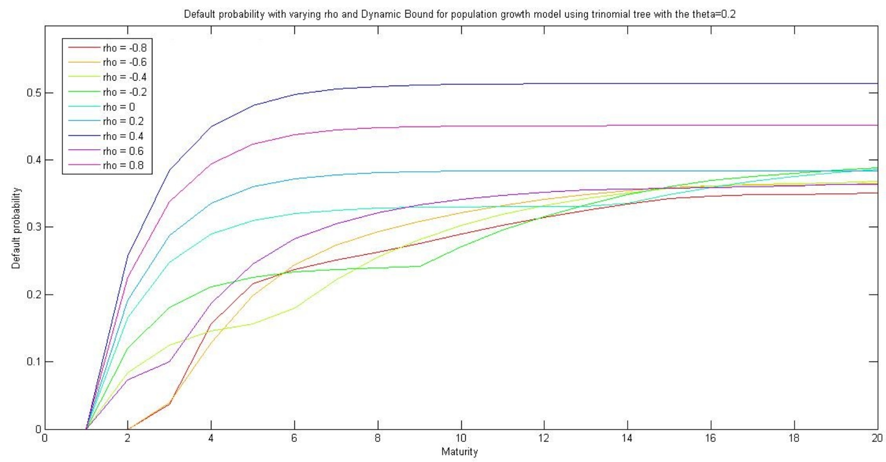

Figure 6.11: Default probability with varying rho from -0.8 to 0.8 by 0.2 with Pierre Collin Dufresne and Robert S. Goldstein's dynamic bound and $V_T < K$ for population growth model with the $\theta = 0.2$ using trinomial tree method, the parameters in the algorithm as follows $B = 75$, $V_0 = 150$, $K = 100$, $\theta = 0.2$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\lambda = 0.18, \phi = 2.8, v = 0.6, R = 0.00175$.

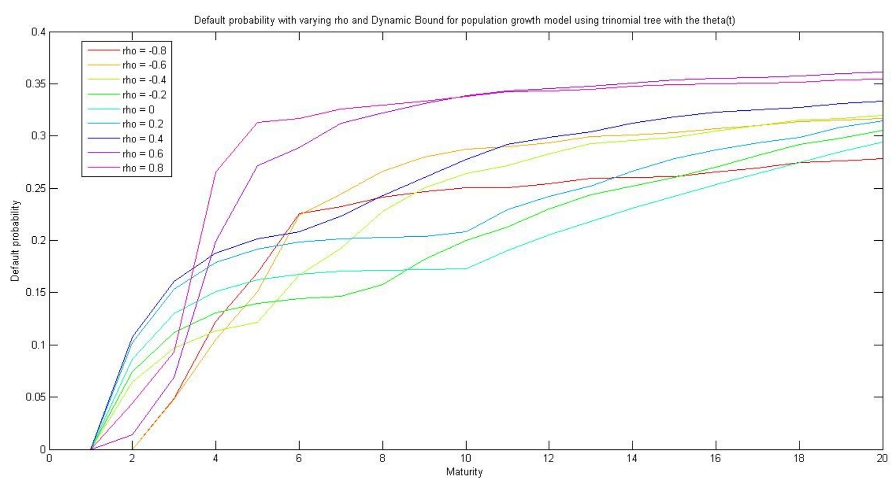

Figure 6.12: Default probability with varying rho from -0.8 to 0.8 by 0.2 with Pierre Collin Dufresne and Robert S. Goldstein's dynamic bound and $V_T < K$ for population growth model with the $\theta(t)$ using trinomial tree method, the parameters in the algorithm as follows $B = 75$, $V_0 = 150$, $K = 100$, $\sigma_r = 0.04$, $\sigma_v = 0.2$, $\lambda = 0.18, \phi = 2.8, v = 0.6$, the term structure is as table(2.1).

Chapter 7

Conclusion

In this dissertation I have investigated the use of a nonlinear stochastic population growth model as a stochastic model for the evolution of interest rates. We find the model to be attractive because it yields positive interest rates with mean reversion properties over time. In the constant parameter case we showed that the model fits the Rogers' scheme for the generation of positive interest rate models. Moreover, we also considered a variant of the model with a non-random time-dependent parameter, which allowed calibration of the model with respect to a specified initial term structure through an adaptation of a method introduced by Hull and White. Although the model is nonlinear, we were able to derive closed form solutions for the model in both the constant and time-varying parameter cases. For both of these cases we used a trinomial-tree discretization process to obtain discrete approximations of the distributions of the interest rates and we compared these results with approximate distributions obtained via Monte-Carlo simulations. The favorable comparisons obtained provide evidence of the validity of the application of the trinomial tree method to nonlinear interest-rate processes.

Moreover, we incorporated the positive interest rate to derive the firm's default probability, which thereby extends Qian's work from a linear interest rate model to a non-linear interest rate model. In order to get the distribution of the firm value under a risk-neutral measure, we first used the first passage time method based on the Fortet integral equation to derive the firm's default probability as driven by the Vasicek interest rate model. However, since the Vasicek model yields an interest rate with a Gaussian distribution and can thus assume negative values, we then proposed the population growth model. As we discussed this model satisfies both the positivity and mean reversion properties. Due to the nonlinearity, it was infeasible to use the first passage time approach in its solution. As an alternative we proposed a new coupled trinomial tree algorithm to derive the distribution of the firm value. To validate this algorithm, we applied this revised trinomial tree method to the Vasicek model and obtained reasonably consistent results. Given the success with this example, we should be able to extend this approach to the other nonlinear positive interest rate models.

In addition to the population growth model, Chapter 4 of this dissertation examined a variety of other interest rate models that were derived in the context of the Rogers' framework and can yield positive interest rates with the appropriate selection of certain parameter values. For example, we noted that when we applied the Rogers' method to the geometric mean reversion process (4.2.23) with a linear function f, the generated interest rate showed the potential for giving a superior fit to the historical data due to its lognormal property. We hope to analyze this model in more detail in future work. There are, of course, any number of other Markov processes and functions f (such as the sigmoid function) that would form the basis for future study and might offer a good fit to historical market data.

In the application of our nonlinear interest rate model to the computation of default probabilities we tried to extend the method (based on the reflection principle) of Giesecke (2004) [12] and Qian (2008) [34], but we were unsuccessful in deriving a closed form solution for the distribution of corresponding first passage time probability (6.2.14) due to the presence of an Ito integral that made closed form evaluation of a certain integral infeasible. In future work we hope to explore the use of numerical methods to obtain discrete approximations of this probability distribution.

In our application of the Fortet integral equation approach we only derived the default probability distribution for interest rate models with linear drift. This is because the Fortet method requires explicit knowledge of the transition function for the underlying Markov process. For nonlinear drift terms the determination of the required transition functions appears to be quite challenging.

Another avenue for future work involves the application of our various models to other topics of importance in fixed income research such as credit spreads, credit grades, and bond pricing.

Bibliography

[1] Acharya, V.V. and J.N. Carpenter(2002), "Corporate Bond Valuation and Hedging with Stochastic Interest Rate and Endogenous Bankruptcy", **The Review of Financial Studies**, Vol.15, pp 1355-1383.

[2] Arnold, L.(1974) **Stochastic Differential Equations: Theory and Application**,John Wiley & Sons, New York.

[3] Black, F. and J. Cox(1976), "Valuing Corporate Securities: Some effect of Bond Independence Provisions", **Journal of Finance**, Vol.31, pp 351-367.

[4] Black, F., and P. Karasinski. "Bond and Option Pricing when Short Rates are Lognormal", **Financial Analysts Journal**, Vol. 47, pp 52-59.

[5] Brennan, M. and E. Schwartz(1979),"A Continuous-Time Approach to the Pricing of Bonds", **Journal of Banking and Finance**, Vol. 3, pp 133-155.

[6] Cairns, A.J.G.(2004), **Interest Rate Models: An Inroduction**, Princeton Univerity Press, Princeton, N.J., 274 pages.

[7] Cox, J., J.Ingersoll and S.Ross(1985), "A Theory of the Term Structure of Interest Rates", **Econometrica**, Vol 53, pp 385-407.

[8] Constantinides, G. M. (1992): A Theory of the Nominal Term Structure of Interest Rates, **Rev.Financial Stud.**, Vol. 5, 531C552.

[9] Flesaker, B. and Hughston, L. P. (1996a), "Positive Interest", **Risk**, Vol. 9, No. 1, pp 46-49.

[10] Fortet, Robert(1943), Les fonctions alatoires du type de Markoff associes a certaines quations linaires aux drives partielles du type parabolique, **Journal de Mathematiques Pures et Appliquees** Vol. 22, pp 177-243.

[11] Gard, Thomas C.(1988), **Introduction to Stochastic Differential Equations**, New York, Marcel Dekker.

[12] Giesecke, K(2004), "Credit Risk Modeling and Valuation: An Introduction", Vol.2, edited by David Shimko, Risk Books, London, 638 pages.

[13] Glasserman, P. and Jin, Y. (2001), "Equilibrium Positive Interest Rates: A United View", **Review of Financial Studies**, Vol. 14, pp 187-214.

[14] Hackbarth, D., C.A. Hennessy and H.E. Lenland (2007), "Can the Trade-off Theory Explain Debt Structure?", **Review of Financial Studies**, Vol. 20, pp 1389-1428.

[15] Ho, S.Y. and S.B. Lee (1986), "Term Structure Movements and Pricing Interest Rate Contingent Claims", **Journal of Finance**, Vol.41, pp 1011-1029.

[16] Hull, J.C. and A.D. White(1990), "Pricing interest rate derivative securities", **Review of Financial Studies**, Vol.3, pp 573-592.

[17] Hull, J.C. and A.D. White(1994), "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models", **Journal of Derivatives**, Vol.2 pp 7-16.

[18] Hull, J.C. and A.D. White(1994), "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models", **Journal of Derivatives**, Vol.2, No.2: pp 37-48.

[19] Hull, J.C. (2000) "Options, Futures and other Derivatives, Prentice Hall", Pretice Hall, (4^{th} Edition), Englewood Cliff, N.J.

[20] James, J. and Webber, N. (2000), **Interest Rate Modelling**, John Wiley - Sons.

[21] Kloeden, P.E. and E. Platen(1992), **Numerical Solution of Stochastic Differential Equations**, Springer Verlag, 636 pages.

[22] Lakshmivarahan, S., Qian, Shengguang and Stock, Duane(2012), "The Distribution of the Value of the Firm and Stochastic Interest Rates", **Journal of Mathematical Finance**, Vol. 2, No. 1, pp 75-82.

[23] Longstaff, F. A. and Schwartz, E. S. (1992), "A Two-factor Interest Rate Model and Contingent Claims Valuation", **The Journal of Fixed Income** Vol. 2, No. 3, pp 16-23.

[24] Leland, H.E. and K.B. Toft(1996), "Optimal Captital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads", **Journal of Finance**, Vol.51, pp 987-1019.

[25] Meyer, P. A. (1966), **Probability and Potentials**, Blaisdell Publishing Company.

[26] Merton, R.C.(1973), "Theory of Rational Option Pricing", **Bell Journal of Economics and Management Science**, Vol.4, pp 141-183.

[27] Merton, R.C.(1974), "On the Pricing of Corporate Debt:The Risk Structure of Interest Rate", **Journal of Finance**, Vol.29, No.2, pp 449-470.

[28] Oksendal, B.(2003), **Stochastic Differential Equations: An Introduction with Applications**, Springer Verlag, New York(sixth Edition), 360 pages.

[29] Parbhoo, Priyanka Anjali (2009), "Positive Interest, Potentials and the Pricing Kernel", Master Thesis, School of Computational and Applied Mathematics, Univ. Witwatersrand.

[30] Pennacchi, G. (2007), **Theory of Asset Pricing**, Addison-Wesley.

[31] Protter, P. E. (2003), **Stochastic Integration and Differential Equations** (2nd edition), Springer-Verlag.

[32] Pierre Collin Dufresne and Robert S. Goldstein (2001), "Do Credit Spreads Reflect Stationary Leverage Ratios? **Journal of Finance**, Vol. 56, No. 5, pp 1929-1958.

[33] Qi,H.(2007), "Credit Spread by a modified Leland-Toft Model" presented at FMA meeting, Orlando.

[34] Qian, Shengguang(2008),"On Pricing Corporate Bonds with Default Risk and Embedded Call Options when the Interest Rate is Stochastic", Ph.D. dissertation, Dept. Comput. Science, Univ. Oklahoma.

[35] Rogers, L. C. G. (1997), "The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates", **Mathematical Finance**, Vol. 7, No. 2, pp 157C164.

[36] Rutkowski, M. (1997), "A Note on the Flesaker-Hughston Model of the Term Structure of Interest Rates", **Applied Mathematical Finance** Vol. 4, No. 3, pp 151-163.

[37] Shiryaev, A.N.(1999), **Essentials of Stochatic Finance: Facts, Models, Theory**, World Scientific, New York, 834 pages.

[38] Steven E.Shreve(2004), **Stochastic Calculus for Finance II: Continuous-Time Models**, Springer Verlag, New York, 550 pages.

[39] Vasicek, O.(1977), "An equilibrium characterization of the term structure", **Journal of Financial Economics**, Vol. 37, pp 339-348.

Appendix A

Expectation of the Geometric Mean Reversion Process

This process is of interest because of its geometric mean reversion property. As we know, most of the market assumptions for bond prices have this property. Now, let us look at the expectation of X_t. To get $E[X_t]$, the key step is to compute $E[e^{\sigma e^{-kt}\int_0^t e^{ks}dB_s}]$. From the form of X_t, it is a stochastic exponential. We will use Novikov's theorem to get the expectation. Set

$$M_t := \sigma e^{-kt}\int_0^t e^{ks}dB_s,$$

and

$$Z_t = e^{\int_0^t \sigma e^{k(s-t)}dBs - \frac{1}{2}\int_0^t [f(s)]^2 ds},$$

where $f(s)$ is a measurable function on $[0, \infty)$ with $\int_0^t [f(s)]^2 ds < \infty$ for each $t > 0$.

Theorem A.1. *Novikov Theorem: If*

$$E[e^{\frac{1}{2}<M>_T}] < \infty,$$

then $E[Z_t] = 1$*, in which case* $(Z_s)_{0\leq s\leq t}$ *is a martingale.*

Then, in our example, set $f(s) = \sigma^2 e^{2k(s-t)}$. We get the following:

$$\begin{aligned} E[e^{\sigma e^{-kt}\int_0^t e^{ks}dB_s}] =& E[Z_t]E[e^{-\frac{1}{2}\int_0^t [f(s)]^2 ds}] \\ =& \frac{\sigma^2}{4k}(1 - e^{-2kt}). \end{aligned}$$

Thus the expectation of X_t is:

$$E[X_t] = e^{e^{-kt}\log x + (\alpha - \frac{\sigma^2}{2k})(1-e^{-kt}) + \frac{\sigma^2(1-e^{-2kt})}{2k}}. \tag{A.0.1}$$