# A model of the short rate with regime shifts and reflecting barriers

by

Matthew Christian Smith

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Applied Mathematics

Program of Study Committee:
Ananda Weerasinghe, Major Professor
Dermot Hayes
Hailing Liu

Iowa State University

Ames, Iowa

2006

Graduate College
Iowa State University

This is to certify that the Master's thesis of

Matthew Christian Smith

has met the thesis requirements of Iowa State University

Major Professor

For the Major Program

# TABLE OF CONTENTS

# 1 Introduction

During the last three decades a plethora of dynamic, stochastic models of the term structure of interest rates have been proposed. Early examples of these modeled the instantaneous spot interest rate (known as the "short rate") as a diffusion process. However, these models have their flaws. Those who use the models to price bonds and other interest rate derivatives find that they must frequently reestimate the models' parameters in order for them to perform well. If the models perfectly described the evolution of the short rate there would be no need to reestimate model parameters. Furthermore, some econometric studies (including [Ait-Sahalia, Y. (1996)]) find that these models are not good descriptions of the time series behavior of the short rate. Such defects have stimulated the development of other types of models for the short rate, like models employing Levy processes, random field models and regime shift models.

Regime shift (also known as Hidden Markov) models introduce an unobserved Markov point process that affects the drift and volatility of a diffusion process, in this case the short rate. Pricing European type claims in this framework can be done via a system of partial differential equations. However, proving that there is a unique solution to the system is difficult for "regime shift versions" of some popular interest rate models. To circumvent this difficulty one may place reflecting barriers on the short rate where both barriers are greater than zero. One can then apply results of V.A. Solonnikov (see [Eidelman, S. D. and Zhitarashu, N. V. (1998)], [Solonnikov, V. A. (1967)]) to show the existence and uniqueness of a solution to the resulting PDE system.

The remainder of this article is organized as follows. Section 1 outlines a general regime shift model, shows how to derive the parabolic PDE system that prices a European contingent claim and explains the difficulties one may encounter in proving the existence and uniqueness of its solution. Section 2 looks at how imposing reflecting boundaries on the short rate adds boundary conditions to the PDE system and gives sufficient conditions for the solution of the system. Section 3 shows the conditional convergence of an explicit finite difference numerical scheme to the actual solution of the system in the case of the Cox-Ingersoll-Ross regime shift model.

## 2 The regime shift model, the infinitesimal generator and derivation of the PDE system

A general version of a regime shift model for the short rate, $r(t)$, under the risk neutral measure is

$$dr(t) = a(t, r(t), u(t))dt + b(t, r(t), u(t))dW(t) \tag{2.1}$$

where $r(t)$ is the spot rate, $W(t)$ denotes a $d$-dimensional standard Brownian motion, $b(t, r(t), u(t))$ is a $d \times d$ matrix, $a(t, r(t), u(t))$ is a scalar and $u(t)$ denotes an $m$-dimensional unobservable Markov process with dynamics given by

$$du(t) = \int_E \delta(t, u_{t-}, z)\mu(dt, dz)$$

The process $u(t)$ represents the regime, and $\mu(dt, du)$ is marked point process on a Lusin mark space. I will assume that $u(t)$ is a Markov chain, $u(t)$ and $W(t)$ are independent and that $d = m = 1$. The states of the Markov chain will be denoted by the integers from 1 to $n$.

Assuming there is no arbitrage and the market is complete, to price a financial asset that pays $\psi(r(T), T)$ at some future time $T$ requires that one computes the following expectation under the risk neutral measure.

$$
\begin{aligned}
V(t,T) &= E\left[\exp\left(-\int_t^T r(s)\,ds\right)\psi(r(T),T)\,\Big|\, r(t)\right] \\
&= E\left[E\left[\exp\left(-\int_t^T r(s)\,ds\right)\psi(r(T),T)\,\Big|\, r(t),u(t)\right]\,\Big|\, r(t)\right] \\
&= \sum_{i=1}^{n} E\left[\exp\left(-\int_t^T r(s)\,ds\right)\psi(r(T),T)\,\Big|\, r(t),u(t)=i\right] \\
&\qquad \times P[u(t)=i \,|\, r(t)]
\end{aligned}
$$

Making the change of variable $y = T - s$ one can write

$$
\begin{aligned}
V(t,T) = \sum_{i=1}^{n} E&\left[\exp\left(-\int_0^{T-t} r(y)\,dy\right)\psi(r(0),0)\,\Big|\, r(0),u(0)=i\right] \\
&\times P[u(0)=i \,|\, r(0)]
\end{aligned}
$$

Thus, to compute the price of the claim, one needs to compute

$$
E\left[\exp\left(-\int_0^{T-t} r(y)\,dy\right)\psi(r(0),0)\,\Big|\, r(0),u(0)=i\right]
$$

the price of the claim if the Markov chain were, in fact, observed and in state $i$ for $i = 1,\ldots,n$. These expectations are probably impossible to compute directly. However, one can show they satisfy a system of PDEs. Assuming that the short rate $r$ is bounded from below, then the Feynman-Kac theorem says that $V(t,T)$ satisfies:

$$
\begin{aligned}
&\frac{\partial V}{\partial t} = AV - rV; \qquad T - t > 0 \\
&V(0,r) = \psi(r)
\end{aligned}
$$

Thus, to apply the theorem one must determine $A$, the infinitesimal generator of the process. By definition, the operator $A$ applied to a function $f$ is

$$
Af = \lim_{t\to 0} \frac{E[f(r(t),u(t)) - f(r(0),u(0))]}{t}
$$

if the limit exists. One can use the generalized Itôs formula presented in [Protter, P. (1990)] to compute $f(r(t),u(t)) - f(r(0),u(0))$.

$$f(r(t),u(t)) - f(r(0),u(0)) = \int_{0^+}^{t} f_r(r(s_-),u(s_-))\,dr(s)$$
$$+ \int_{0^+}^{t} f_u(r(s_-),u(s_-))\,du(s) + \frac{1}{2}\int_{0^+}^{t} f_{rr}(r(s_-),u(s_-))\,(dr(s))^2$$
$$+ \int_{0^+}^{t} f_{ru}(r(s_-),u(s_-))\,d[r(s),u(s)]^c + \frac{1}{2}\int_{0^+}^{t} f_{uu}(r(s_-),u(s_-))\,d[u(s),u(s)]^c$$
$$+ \sum_{0\le\tau_i\le t}\Big[f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))$$
$$-f_r(r(\tau_{i-}),u(\tau_{i-}))\Delta r(\tau_i) - f_u(r(\tau_{i-}),u(\tau_{i-}))\Delta u(\tau_i)\Big]$$

The notation $\int_{0^+}^{t}$ denotes the integral over the interval $(0,t]$. $u(s_-)$ denotes $\lim_{v\to s, v<s} u(s)$. The $\tau_i$ denote the times in which there is a jump in the Markov chain from 0 to $t$, so, consequently, $\Delta u(\tau_i)$ denotes the size of the jump occurring at time $\tau_i$ in the process $u$. Note that there are no jumps in $r(t)$. The quantity $[r(s),u(s)]^c$ is the path by path continuous part of the quadratic covariation of $r$ and $u$. By theorem 28, p.68 of [Protter, P. (1990)], $[r(s),u(s)]$, the quadratic variation of $r$ and $u$ is $r(0)u(0) + \sum_{0\le\tau_i\le t}\Delta r(\tau_i)\Delta u(\tau_i)$. The $r$ process has no jumps so the quadratic variation is simply $r(0)u(0)$. The continuous part of this is $r(0)u(0)$. Thus $d[r(s),u(s)]^c = 0$ since the quadratic variation is constant, which makes the integral

$$\int_{0^+}^{t} f_{ru}(r(s_-),u(s_-))\,d[r(s),u(s)]^c$$

zero. Furthermore, $[u(s),u(s)]^c$ is zero since it is a Markov chain, unless it stays in one state in which case $d[u(s),u(s)]^c$ is zero. Thus, in any case, integrating with respect to $[u(s),u(s)]$ yields zero. The integral $\int_{0^+}^{t} f_u(r(s_-),u(s_-))\,du(s)$ is by definition $\sum_{0\le\tau_i\le T} f_u(r(\tau_{i-}),u(\tau_{i-}))\Delta u(\tau_i)$ and the term in the summation multiplied by $\Delta r(s)$ is zero since the $r$ process has no jumps. Thus, one obtains:

$$\begin{aligned} f(r(t),u(t)) - f(r(0),u(0)) &= \int_{0^+}^{t} f_r(r(s_-),u(s_-))\,dr(s) \\ &+ \frac{1}{2}\int_{0^+}^{t} f_{rr}(r(s_-),u(s_-))\,(dr(s))^2 \\ &+ \sum_{0\le\tau_i\le t} [f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))] \end{aligned}$$

Recalling that the dynamics of the short rate are given by equation (1) we get:

$$\begin{aligned} f(r(t),u(t)) - f(r(0),u(0)) &= \int_{0^+}^{t} f_r(r(s_-),u(s_-))\mu(r(s),u(s))\,ds \\ &+ \int_{0^t}^{t} \frac{1}{2} f_{rr}(r(s_-),u(s_-))\sigma^2(r(s),u(s))\,ds \\ &+ \int_{0^+}^{t} f_r(r(s_-),u(s_-))\sigma(r(s),u(s))\,dW(s) \\ &+ \sum_{0\le\tau_i\le t} [f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))] \end{aligned}$$

Taking the expectation of this, dividing by $t$ and then taking the limit as $t$ goes to zero yields:

$$\begin{aligned} \lim_{t\to 0}\frac{E\left[f(r(t),u(t) - f(r(0),u(0))\right]}{t} &= \mu(r(0),u(0))f_r(r(0),u(0)) \\ &+\frac{1}{2}\sigma^2(r(0),u(0))f_{rr}(r(0),u(0)) \\ +\lim_{t\to 0}\frac{E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\right]}{t} \end{aligned}$$

To calculate $\lim_{t\to 0} E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\right]/t$ one can proceed as follows. Let $J_t$ denote the number of jumps in the Markov chain from 0 to time $t$ and

$P[A]$ the probability of event $A$. By the rule of iterated expectations

$$
\begin{aligned}
&E\left[\sum_{0\le\tau_i\le T=t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\right]\\
=&E\left(E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\middle|\, J_t\right]\right)\\
=&\sum_{j=0}^{\infty} E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\middle|\, J_t=j\right] P[J_t=j]\\
=&E\left[f(r(\tau_1),u(\tau_1)) - f(r(\tau_{1-}),u(\tau_{1-}))\middle| J_t=1\right] P[J_t=1] \qquad (2.2)\\
+&\sum_{j=2}^{\infty} E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i-}),u(\tau_{i-}))\middle|\, J_t=j\right] P[J_t=j] \qquad (2.3)
\end{aligned}
$$

Let $s_j = \tau_j - \tau_{j-1}$, the amount of time between jumps $j$ and $j-1$, where $s_0 = 0$. These are often called the "holding times". In a Markov chain they are independently and exponentially distributed with parameter $\lambda(u(\tau_{j-1}))$. This parameter is often referred to as the "jump rate" out of state $u(\tau_{j-1})$. Let $Y = \{s_1, s_2, \ldots, s_{j+1} : s_1 \le t, s_2 \le t-\tau_1, \ldots, s_j \le t-\tau_j, s_{j+1} > t-\tau_j\}$, $p_{kl}$ be the probability (assumed to be independent of time) that the chain jumps to state $l$ given it is in state $k$ at any jump time and $I = \{(i_1, i_2, \ldots, i_j) : u(\tau_1) = i_1, u(\tau_2) = i_2, \ldots, u(\tau_j) = i_j\}$. Then,

$$
\begin{aligned}
&P[J_t=j] = P[\tau_1\le t, \tau_2\le t, \ldots, \tau_j\le t, \tau_{j+1}>t]\\
&= P\left[s_1\le t, s_2\le t-\tau_1, \ldots, s_j\le t-\tau_{j-1}, s_{j+1}>t-\tau_j\right]\\
&= \sum_I P\left[s_1\le t, s_2\le t-\tau_1, \ldots, s_j\le t-\tau_{j-1}, s_{j+1}>t-\tau_j \middle|\, i_1,\ldots,i_n\right]\\
&\qquad \times P[u(\tau_1)=i_1, u(\tau_2)=i_2, \ldots, u(\tau_j)=i_j]\\
&= \sum_I \left(\int\underset{Y}{\cdots}\int \lambda(i_0)\ldots\lambda(i_j)\exp\left(-\sum_{k=1}^{j+1}\lambda(u(\tau_{k-1}))s_k\right) ds_{j+1}\ldots ds_1\right)\\
&\qquad \times p_{i_0 i_1}\cdots p_{i_{j-1} i_j} \qquad (2.4)
\end{aligned}
$$

Let $I_j = \{i_j : i_j \neq i_{j-1}\}$. Thus[1]

$$
\begin{aligned}
P[J_t = 1] &= \sum_{I_1} p_{i_0 i_1} \int_0^t \int_{t-s_1}^{\infty} \lambda(i_0)\lambda(i_1) e^{-\lambda(i_0)s_1 - \lambda(i_1)s_2} ds_2\, ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \int_0^t \lim_{x\to\infty} \int_{t-s_1}^{x} \lambda(i_0)\lambda(i_1) e^{-\lambda(i_0)s_1 - \lambda(i_1)s_2} ds_2\, ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \lambda(i_0)\lambda(i_1) \int_0^t e^{-\lambda(i_0)s_1} \lim_{x\to\infty} \int_{t-s_1}^{x} e^{-\lambda(i_1)s_2} ds_2 \; ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \lambda(i_0)\lambda(i_1) \int_0^t e^{-\lambda(i_0)s_1} \lim_{x\to\infty} -\frac{1}{\lambda(i_1)} e^{-\lambda(i_1)s_2} \Big|_{t-s_1}^{x} \; ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \lambda(i_0)\lambda(i_1) \int_0^t e^{-\lambda(i_0)s_1} \frac{1}{\lambda(i_1)} e^{-\lambda(i_1)(t-s_1)}\, ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \lambda(i_0) e^{-\lambda(i_1)t} \int_0^t e^{(\lambda(i_1)-\lambda(i_0))s_1}\, ds_1 \\
&= \sum_{I_1} p_{i_0 i_1} \lambda(i_0) e^{-\lambda(i_1)t} \frac{1}{(\lambda(i_1)-\lambda(i_0))} e^{(\lambda(i_1)-\lambda(i_0))s_1} \Big|_0^t \\
&= \sum_{I_1} p_{i_0 i_1} \frac{\lambda(i_0)}{\lambda(i_1)-\lambda(i_0)} \left(e^{-\lambda(i_0)t} - e^{-\lambda(i_1)t}\right)
\end{aligned}
$$

Now I will find an upper bound for $P[J_t = j]$ for $j \geq 2$ by obtaining a bound on the value of the multiple integral in (4). Since $\exp(-x)$ is a decreasing function of $x$

$$
\begin{aligned}
&\int \underset{Y}{\cdots} \int \lambda(u(0)) \ldots \lambda(u(\tau_j)) \exp\left(-\sum_{k=1}^{j+1} \lambda(u(\tau_{k-1}))s_k\right) ds_{j+1} \ldots ds_1 \\
&\leq \int \underset{Y}{\cdots} \int \lambda_{Max}^{j+1} \exp\left(-\sum_{k=1}^{j+1} \lambda_{min} s_k\right) ds_{j+1} \ldots ds_1 \qquad (*)
\end{aligned}
$$

where $\lambda_{min}$ and $\lambda_{Max}$ are the minimum and maximum jump rates respectively. Since the Markov chain has only finitely many states these quantities are well-defined.

The innermost integral is with respect to $s_{j+1}$ and its integration limits are $t - \tau_j$ and $\infty$. One can pull $\lambda_{Max}^{j+1}$ and $\exp(-\sum_{k=1}^{j} \lambda_{min} s_k)$ out of this integral leaving

$$
\int_{t-\tau_j}^{\infty} \exp(-\lambda_{min} s_{j+1})\, ds_{j+1}
$$

[1] I assume here that no two states of the Markov chain have the same jump rate; if, say, states $k$ and $l$ have the same jump rate then the integrand below becomes $\lambda^2 e^{-\lambda(s_1+s_2)}$, where $\lambda = \lambda(k) = \lambda(l)$. Computing the double integral, one finds it equals $\lambda t e^{-\lambda t}$.

inside. The computation of the integral is as follows.

$$\int_{t-\tau_j}^{\infty} \exp(-\lambda_{min}\tau_{j+1})\, d\tau_{j+1} = \lim_{x\to\infty} \int_{t-\tau_j}^{x} \exp(-\lambda_{min}s_{j+1})\, ds_{j+1}$$
$$= \lim_{x\to\infty} \left.\frac{-\exp(-\lambda_{min}s_{j+1})}{\lambda_{min}}\right|_{t-\tau_j}^{x} = \frac{\exp(-\lambda_{min}(t-\tau_j)}{\lambda_{min}}$$

Thus the integral labeled $(*)$ above is equal to

$$\int_0^t \int_0^{t-\tau_1} \cdots \int_0^{t-\tau_{j-1}} \frac{\lambda_{Max}^{j+1}}{\lambda_{min}} \exp\left(-\sum_{k=1}^{j} \lambda_{min}s_k - \lambda_{min}(t-\tau_j)\right) ds_j \ldots ds_1$$
$$= \int_0^t \int_0^{t-\tau_1} \cdots \int_0^{t-\tau_{j-1}} \frac{\lambda_{Max}^{j+1}}{\lambda_{min}} \exp\left(-\lambda_{min}\tau_j - \lambda_{min}(t-\tau_j)\right) ds_j \ldots ds_1$$
$$= \frac{\lambda_{Max}^{j+1}}{\lambda_{min}} \exp\left(-\lambda_{min}t\right) \int_0^t \int_0^{t-\tau_1} \cdots \int_0^{t-\tau_{j-1}} ds_j \ldots ds_1$$
$$\leq \frac{\lambda_{Max}^{j+1}}{\lambda_{min}} \exp\left(-\lambda_{min}t\right) \int_0^t \int_0^{t} \cdots \int_0^{t} ds_j \ldots ds_1 = \frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j \exp\left(-\lambda_{min}t\right)$$

Therefore, it has been shown that

$$P[J_t = j] \leq \sum_I p_{i_0 i_1} \cdots p_{i_{j-1} i_j} \frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j \exp\left(-\lambda_{min}t\right) = \frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j \exp\left(-\lambda_{min}t\right)$$

which implies that

$$\left|\sum_{j=2}^{\infty} E\left[\left.\sum_{0\leq\tau_i\leq t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))\right| J_t = j\right] P[J_t = j]\right|$$
$$\leq \sum_{j=2}^{\infty} E\left[\left.\sum_{0\leq\tau_i\leq t} |f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))|\right| J_t = j\right] P[J_t = j]$$
$$\leq \sum_{j=2}^{\infty} E\left[\left.\sum_{0\leq\tau_i\leq t} |f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))|\right| J_t = j\right] \frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j e^{-\lambda_{min}t}$$

If the function $f$ is a continuous function of $r$ then it is also a continuous function of $t$, since the process $r(t)$ has continuous sample paths. Thus, for a given state $l$ of the Markov chain, $f$ achieves a maximum on the interval $[0,t]$, so there exists a constant $M_i$ such that $|f(r(t^*),l)| \leq M_l$ for all $t^* \in [0,t]$. Since the Markov chain has only finitely many states, then $|f(r(t^*),u(t^*)| \leq M$ for all $t^* \in [0,t]$ where $M = \sup M_l$. Thus,

$$|f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))| \leq |f(r(\tau_i),u(\tau_i))| + |f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))| \leq 2M$$

at any jump time $\tau_i$. Consequently,

$$\sum_{j=2}^{\infty} E\left[\sum_{0\le\tau_i\le t} |f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))|\,\middle|\, J_t = j\right] \frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j e^{-\lambda_{min} t}$$
$$\le \sum_{j=2}^{\infty} 2Mj\frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^j e^{-\lambda_{min} t} = t^2 \sum_{j=2}^{\infty} 2Mj\frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^{j-2} e^{-\lambda_{min} t}$$

if the infinite series converges. One can use the ratio test to investigate the convergence of this series by computing

$$\lim_{j\to\infty} \frac{2M(j+1)\frac{\lambda_{Max}^{j+2}}{\lambda_{min}} t^{j-1} e^{-\lambda_{min} t}}{2M\frac{\lambda_{Max}^{j+1}}{\lambda_{min}} t^{j-2} e^{-\lambda_{min} t}} = \lim_{j\to\infty} \frac{j+1}{j}\lambda_{Max} t = \lambda_{Max} t$$

Thus, the series converges if $t < \lambda_{Max}^{-1}$. It is clear that each term in the series is an increasing function of $t$, so if we consider values of $t$ such that $t \le T < \lambda_{Max}^{-1}$, then, under this restriction, the series reaches its maximum value when $t = T$. Denote this maximum value by $S$. Thus, for sufficiently small $t$

$$t^2 \sum_{j=2}^{\infty} 2Mj\frac{\lambda_{Max}^{j+1}}{\lambda_{min}}\, t^{j-2} e^{-\lambda_{min} t} \le t^2 S$$

From the results derived thus far

$$\left|\sum_{j=2}^{\infty} E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))\,\middle|\, J_t = j\right] P[J_t = j]\right| \div t \le tS$$

The right-hand side of this inequality goes to 0 as $t$ goes to 0 so the limit of the expression inside the absolute value on the left-hand side is 0. Hence,

$$\lim_{t\to 0} \frac{E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))\right]}{t}$$
$$= \lim_{t\to 0} \frac{E\left[f(r(\tau_1),u(\tau_1)) - f(r(\tau_{1\text{-}}),u(\tau_{1\text{-}}))\,\middle|\, J_t = 1\right] P[J_t = 1]}{t}$$
$$+ \lim_{t\to 0} \frac{\sum_{j=2}^{\infty} E\left[\sum_{0\le\tau_i\le t} f(r(\tau_i),u(\tau_i)) - f(r(\tau_{i\text{-}}),u(\tau_{i\text{-}}))\,\middle|\, J_t = j\right] P[J_t = j]}{t}$$
$$= \lim_{t\to 0} E\left[f(r(\tau_1),u(\tau_1)) - f(r(\tau_{1\text{-}}),u(\tau_{1\text{-}}))\,\middle|\, J_t = 1\right]$$
$$\times \sum_{I_1} p_{i_0 i_1} \frac{\lambda(i_0)}{\lambda(i_1) - \lambda(i_0)}\, \frac{(e^{-\lambda(i_0)t} - e^{-\lambda(i_1)t}}{t}) \tag{2.5}$$

By L'Hopital's rule

$$\lim_{t\to 0}\frac{e^{-\lambda(i_0)t}-e^{-\lambda(i_1)t}}{t}=\lim_{t\to 0}\frac{-\lambda(i_0)e^{-\lambda(i_0)t}+\lambda(i_1)e^{-\lambda(i_1)t}}{1}=\lambda(i_1)-\lambda(i_0)$$

So (5) equals

$$\begin{aligned}&\lambda(i_0)E\left[f(r(\tau_1),u(\tau_1))-f(r(\tau_{1\text{-}}),u(\tau_{1\text{-}}))\middle|J_t=1\right]\\ =&\lambda(i_0)\sum_{I_1}p_{i_0 i_1}[f(r(0),i_1)-f(r(0),i_0)]\end{aligned}$$

If the Markov chain is initially in state $i$ this can be written as

$$\lambda(i)\sum_{j\neq i}p_{ij}[f(r(0),j)-f(r(0),i)]$$

Now that we have derived the generator, we see that the Feynman-Kac formula gives the following PDE for the European contingent claim price given the initial interest rate and initial state of the Markov chain (suppose the Markov chain is initially in state i):

$$\begin{aligned}\frac{\partial V}{\partial t}(t,r,i) &= \frac{\sigma^2(r,u)}{2}\frac{\partial^2 V}{\partial r^2}(t,r,i)+\mu(r,u)\frac{\partial V}{\partial r}(t,r,i)\\ &+ \lambda(i)\sum_{j\neq i}p_{ij}[V(t,r,j)-V(t,r,i)]-rV(t,r,i)\end{aligned}$$

The initial condition is:

$$V(0,r,i)=\psi(r,i)$$

Therefore, to determine the value of a European contingent claim, given that the Markov chain is initially in state $i$ and the initial short rate, we must know what the value of the bond would be if the Markov chain were in any other initial state. Let $V_i(t,r)=V(t,r,i)$, $\mu_i(r)=\mu(r,i)$, $\sigma_i(r)=\sigma(r,i)$, $\lambda_i=\lambda(i)$ and $\psi_i(r)=\psi(r,i)$. Thus, the value of the claim satisfies the following system of PDEs if there are $n$ states in the Markov chain.

$$\begin{pmatrix} \frac{\partial V_1}{\partial t} \\ \frac{\partial V_2}{\partial t} \\ \vdots \\ \frac{\partial V_n}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_1^2(r)}{2} & & & \\ & \frac{\sigma_2^2(r)}{2} & & \\ & & \ddots & \\ & & & \frac{\sigma_n^2(r)}{2} \end{pmatrix} \begin{pmatrix} \frac{\partial^2 V_1}{\partial r^2} \\ \frac{\partial^2 V_2}{\partial r^2} \\ \vdots \\ \frac{\partial^2 V_n}{\partial r^2} \end{pmatrix}$$

$$+ \begin{pmatrix} \mu_1(r) & & & \\ & \mu_2(r) & & \\ & & \ddots & \\ & & & \mu_n(r) \end{pmatrix} \begin{pmatrix} \frac{\partial V_1}{\partial r} \\ \frac{\partial V_2}{\partial r} \\ \vdots \\ \frac{\partial V_n}{\partial r} \end{pmatrix}$$

$$+ \begin{pmatrix} -r-\lambda_1 & \lambda_1 p_{12} & \lambda_1 p_{13} & \dots & \lambda_1 p_{1n} \\ \lambda_2 p_{21} & -r-\lambda_2 & \lambda_2 p_{23} & \dots & \lambda_2 p_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_n p_{n1} & \lambda_n p_{n2} & \dots & \lambda_n p_{n(n-1)} & -r-\lambda_n \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{pmatrix}$$

The initial conditions are $V_i(0,r) = \psi(r)$, $i = 1, \dots, n$.

This PDE system is **parabolic** since at any point $(t,r)$, $\sigma_i^2(t,r) \geq 0$ for $i = 1, \dots, n$. It is **uniformly parabolic** if there is a number $\delta > 0$ such that for all $(t,r) \in [0,T] \times \mathbb{R}$

and for all $i = 1, \ldots, n$, $\sigma_i^2(t, r) \geq \delta$. To my knowledge, all general results on parabolic systems with initial conditions treat only uniformly parabolic systems.

Consider now the "regime shift version" of the popular Cox-Ingersoll-Ross short rate model. The dynamics of the short rate in this model are

$$dr(t) = \kappa_i(\theta_i - r(t))dt + \sigma_i\sqrt{r(t)}dW(t)$$

where $W(t)$ is a standard Brownian motion, the subscript $i$ denotes the state of the Markov chain, and $\kappa_i$, $\theta_i$ and $\sigma_i$ are constants. For this model $\sigma_i^2(t, x) = \sigma_i^2|r(t)|$. As long as the initial short rate is non-negative, then $r(t)$ is always non-negative (see [Cox, J., Ingersoll, J. and Ross, S. (1985)], p. 391) so the Feynman-Kac formula can be applied. However, it is not bounded away from zero, so the system is not uniformly parabolic and, therefore, the question of existence and uniqueness is not addressed by the existing literature on parabolic systems.

Thus, one is left with two options: one, prove existence and uniqueness for the resulting parabolic system or two, modify the short rate process somehow so that the resulting parabolic system has a unique solution. I have done the latter by imposing reflecting boundaries on the short rate that bound it away from zero and also give an upper bound. While this has been done for mathematical reasons, it is not without economic justification. If the central bank exercises control of the short rate to maximize some criteria (for example, the utility of a representative economic agent), placing such bounds on the short rate may be optimal. For example, consider the stochastic control problem in [Weerasinghe, A. (2005)]. In his article a bounded variation control may be applied to a diffusion process. It turns out that placing reflecting barriers on the diffusion is the optimal control when the cost function satisfies certain assumptions.

## 3 Reflecting boundaries and regime shifts

Consider the following model for the short rate

$$dr(t) = \mu_i(t,r) + \sigma_i(t,r)dW(t) + dL_\alpha(t) - dU_\beta(t)$$

where $\beta > \alpha > 0$ are the reflecting boundaries and $L_\alpha$ and $U_\beta$ are the local times at $\alpha$ and $\beta$ respectively. The subscript $i$ denotes the state of the regime at time $t$, while $W(t)$ is a standard Brownian motion.

Consider now the price of a European contingent claim maturing at time $T$ under the assumption that the regime is observed, $V_i(T-t,r)$. Under the risk neutral measure:

$$V_i(r(t),T-t) = E\left[\exp\left(-\int_t^T r(s)\,ds\right)\middle|X(t)=i\right]$$

$X(t)$ denotes the state of the regime at time $t$. Applying Itôs lemma to $V_i(r(t),T-t)e^{\int_0^t r(s)\,ds}$ and taking expectations yields

$$\begin{aligned}
&E\left[V_i(r(t),T-t)e^{\int_0^t r(s)\,ds}\right] = V_i(x,0)\\
&+E\Bigg[\int_0^T\left(-\frac{\partial V_i}{\partial t}+\frac{\sigma_i^2 r}{2}\frac{\partial^2 V_i}{\partial r^2}+\kappa_i(\theta_i-r)\frac{\partial V_i}{\partial r}-\lambda_i r+\lambda_i\sum_{j\neq i}p_{i,j}[V_j-V_i]\right)\\
&\qquad\times e^{\int_0^t r(s)\,ds}\,ds\Bigg]\\
&+E\left[\int_0^T\left(\left.\frac{\partial V_i}{\partial r}\right|_{\alpha,T-t}dL+\left.\frac{\partial V_i}{\partial r}\right|_{\beta,T-t}dU\right)e^{\int_0^t r(s)\,ds}\,ds\right]
\end{aligned}$$

Since this must hold for any value of T and the quantity on the left-hand side of the equation is a martingale:

$$\frac{\partial V_i}{\partial t} = \frac{\sigma_i^2 r}{2}\frac{\partial^2 V_i}{\partial r^2} + \kappa_i(\theta_i - r)\frac{\partial V_i}{\partial r} - \lambda_i r + \lambda_i \sum_{j\neq i} p_{i,j}[V_j - V_i] \tag{3.1}$$

$$\left.\frac{\partial V_i}{\partial r}\right|_{\alpha,T-t} = \left.\frac{\partial V_i}{\partial r}\right|_{\beta,T-t} = 0 \tag{3.2}$$

Furthermore, we have know that the claim is worth $\psi(r)$ at its maturity date $T$. Thus for $i = 1, \ldots, n$ we also have the initial conditions:

$$V_i\Big|_{r,0} = \psi_i(r) \tag{3.3}$$

To guarantee the existence and uniqueness of a solution to the short rate stochastic differential equation it is sufficient to assume the following:

$$|\mu(t, r_1) - \mu(t, r_2)| \leq K_\mu |r_1 - r_2| \qquad |\sigma(t, r_1) - \sigma(t, r_2)| \leq K_\sigma |r_1 - r_2| \tag{3.4}$$

$$|\mu(t, r)| \leq K_1(1 + |r|) \qquad |\sigma(t, r)| \leq K_2(1 + |r|) \tag{3.5}$$

for any $r, r_1, r_2$ in the interval $[\alpha, \beta]$. The conditions on the first line above are known as **Lipschitz conditions** in r. Note that a function that satisfies a Lipschitz condition in a variable is continuous in that variable but the converse is not necessarily true.

To guarantee the existence and uniqueness of the solution to the PDE system (6) with boundary conditions (7) and initial conditions (8) one can invoke a theorem of V.A. Solonnikov (see [Eidelman, S. D. and Zhitarashu, N. V. (1998)], [Solonnikov, V. A. (1967)]). Let $\Omega = \{(t, r) : 0 < t \leq T,\ \alpha \leq r \leq \beta\}$. The theorem requires that the following conditions must be satisfied for $i = 1, \ldots, n$.

CONDITION 1: The functions $\mu_i(t, r)$ and $\sigma_i(t, r)$ are continuous in $t$ and $x$ on $\bar{\Omega}$

CONDITION 2: The functions $\mu_i(t, r)$ and $\sigma_i(t, r)$ are bounded on $\Omega$. This is guaranteed if these functions are continuous on $\Omega$.

CONDITION 3:

$$\sup_{(t,x),(t,y)\in\Omega} \frac{|\mu_i(t,x)-\mu_i(t,y)|}{|x-y|^a} \qquad \sup_{(t,x),(t,y)\in\Omega} \frac{|\sigma_i(t,x)-\sigma_i(t,y)|}{|x-y|^a}$$
$$\sup_{(t,r),(\tau,r)\in\Omega} \frac{|\mu_i(t,r)-\mu_i(\tau,r)|}{|t-\tau|^{a/2}} \qquad \sup_{(t,r),(\tau,r)\in\Omega} \frac{|\sigma_i(t,r)-\sigma_i(\tau,r)|}{|t-\tau|^{a/2}}$$

are all finite for some $0 < a < 1$.

CONDITION 4: $\psi_i'\big|_{r=\alpha} = \psi_i'\big|_{r=\beta} = 0$

CONDITION 5: $\psi_i(r) \in C^2$

CONDITION 6: $\sup_{(x,y)\in[\alpha,\beta]x\neq y} \frac{|\psi''(x)-\psi''(y)|}{|x-y|^a}$ is finite.

# 4 Numerical solution of the PDE system

Unfortunately, it is doubtful that the above boundary value problem can be solved analytically. However, one can solve it numerically using an explicit finite difference method. I will now show that this method is "good" in a certain sense for solving the parabolic boundary problem that arises from the Cox-Ingesoll-Ross model with regime shifts.

The PDE system that prices a European contingent claim when the short rate follows a Cox-Ingersoll-Ross reflected diffusion with regimes shifts is

$$\frac{\partial V_i}{\partial t} = \frac{\sigma_i^2 r}{2}\frac{\partial^2 V_i}{\partial r^2} + \kappa_i(\theta_i - r)\frac{\partial V_i}{\partial r} - \lambda_i r + \lambda_i \sum_{j\neq i} p_{i,j}[V_j - V_i] \tag{4.1}$$
$$\left.\frac{\partial V_i}{\partial r}\right|_{\alpha,T-t} = \left.\frac{\partial V_i}{\partial r}\right|_{\beta,T-t} = 0$$
$$V_i\big|_{t=0} = \psi_i(r) \qquad i = 1,\ldots,n$$

Consider the grid obtained by partitioning the interval $[0,T]$ into $J$ equal subintervals and the interval $[\alpha,\beta]$ into $K$ equal subintervals. Let $\Delta t = T/J$ and $\Delta r = (\beta-\alpha)/K$. Approximate the derivatives of $V_i$ by finite differences as follows:

$$\frac{\partial V_i}{\partial t}(j\Delta t, \alpha + k\Delta r) \approx \frac{U^{(i)}_{j,k} - U^{(i)}_{j-1,k}}{\Delta t}$$
$$\frac{\partial^2 V_i}{\partial r^2}(j\Delta t, \alpha + k\Delta r) \approx \frac{U^{(i)}_{j-1,k+1} - 2U^{(i)}_{j-1,k} + U^{(i)}_{j-1,k-1}}{(\Delta r)^2}$$
$$\frac{\partial V_i}{\partial r}(j\Delta t, \alpha + k\Delta r) \approx \frac{U^{(i)}_{j-1,k+1} - U^{(i)}_{j-1,k-1}}{2\Delta r}$$

The notation $U^{(i)}_{j,k}$ denotes the value of the function $U^{(i)}$ at $(j\Delta t, \alpha + k\Delta r)$ where $U^{(i)}$ is the finite difference approximation of $V_i$. The subscript $j$ is an integer value from 1 to $T$ while $k$ is an integer from 1 to $K$. Thus, one might view $U^{(i)}_{j,k}$ as the value of $U^{(i)}$ at the "$j, kth$ node" of the grid. Suppose now that the Markov chain has only two states. Inserting the finite difference approximations into (11) we obtain the following finite difference equation for $U^{(i)}$:

$$
\begin{aligned}
U^{(i)}_{j,k} &= U^{(i)}_{j-1,k} + \frac{\sigma_i^2(\alpha + (k-1)\Delta r)\Delta t}{2} \left( \frac{U^{(i)}_{j-1,k+1} - 2U^{(i)}_{j-1,k} + U^{(i)}_{j-1,k-1}}{(\Delta r)^2} \right) \\
& \kappa_i(\theta_i - x)\Delta t \left( \frac{U^{(i)}_{j-1,k+1} - U^{(i)}_{j-1,k-1}}{2\Delta r} \right) \\
& + \lambda_i \Delta t U^{(l)}_{j-1,k} - (\lambda_i + r)\Delta t U^{(i)}_{j-1,k} \\
& i = 1,2 \quad l = 1,2 \quad l \neq i
\end{aligned}
$$

If we further assume that the derivatives of $V_i(t, r)$ of order two and greater are bounded and denote $V_i(j\Delta, k\Delta r)$ by $V^{(i)}_{j,k}$ then by Taylor's theorem :

$$
\begin{aligned}
\frac{\partial V_i}{\partial t}(j\Delta t, \alpha + k\Delta r) &= \frac{V^{(i)}_{j,k} - V^{(i)}_{j-1,k}}{\Delta t} + O(\Delta t) \\
\frac{\partial^2 V_i}{\partial r^2}(j\Delta t, \alpha + k\Delta r) &= \frac{V^{(i)}_{j-1,k+1} - 2V^{(i)}_{j-1,k} + V^{(i)}_{j-1,k-1}}{(\Delta r)^2} + O((\Delta r)^2) \\
\frac{\partial V_i}{\partial r}(j\Delta t, \alpha + k\Delta r) &= \frac{V^{(i)}_{j-1,k+1} - V^{(i)}_{j-1,k-1}}{2\Delta r} + O((\Delta r)^2)
\end{aligned}
$$

Thus, $V_j^{(i)}, k$ satisfies the following finite difference equation as well as (11):

$$
\begin{aligned}
V^{(i)}_{j,k} &= V^{(i)}_{j-1,k} + \frac{\sigma_i^2(\alpha + (k-1)\Delta r)\Delta t}{2} \left( \frac{V^{(i)}_{j-1,k+1} - 2V^{(i)}_{j-1,k} + V^{(i)}_{j-1,k-1}}{(\Delta r)^2} \right) \\
& \kappa_i(\theta_i - x)\Delta t \left( \frac{V^{(i)}_{j-1,k+1} - V^{(i)}_{j-1,k-1}}{2\Delta r} \right) \\
& + \lambda_i \Delta t V^{(l)}_{j-1,k} - (\lambda_i + r)\Delta t V^{(i)}_{j-1,k} \\
& + O((\Delta t)^2) + O((\Delta r)^2 \Delta t)
\end{aligned}
$$

Let $Z^{(i)}_{j,k}$ equal $V^{(i)}_{j,k} - U^{(i)}_{j,k}$ the difference between the solution of the $ith$ equation of the PDE system (11) and the finite difference approximation. Then:

$$\begin{aligned}
|Z^{(i)}_{j,k}| &\leq \left(1 - \frac{\sigma_i^2(\alpha + (k-1))\Delta r}{(\Delta r)^2} - \Delta t(\lambda_i + \alpha + (k-1))\Delta r\right)|Z^{(i)}_{j-1,k}| \\
&+ \left(\frac{\sigma_i^2(\alpha + (k-1))\Delta r\Delta t}{2(\Delta r)^2)} + \frac{\kappa_i(\theta_i - \alpha + (k-1))\Delta r}{2\Delta r}\right)|Z^{(i)}_{j-1,k+1}| \\
&+ \left(\frac{\sigma_i^2(\alpha + (k-1))\Delta r\Delta t}{2(\Delta r)^2)} - \frac{\kappa_i(\theta_i - \alpha + (k-1))\Delta r}{2\Delta r}\right)|Z^{(i)}_{j-1,k-1}| \\
&+ \Delta t\lambda_i|Z^{(l)}_{j-1,k}| + O((\Delta t)^2) + O((\Delta x)^2\Delta t) \qquad (4.2) \\
& \quad l \neq i
\end{aligned}$$

Assuming that $\alpha < \theta_i < \beta$, i.e., that the short rate in state $i$ tends to a level between its boundaries, one can make choices of $\Delta t$ and $\Delta r$ to ensure that the coefficients of $Z^{(i)}_{j-1,k}$, $Z^{(i)}_{j-1,k+1}$ and $Z^{(i)}_{j-1,k-1}$ are all non-negative. To make the coefficient of $Z^{(i)}_{j-1,k}$ non-negative one requires that

$$1 - \frac{\sigma_i^2(\alpha + (k-1)\Delta r)\Delta t}{(\Delta r)^2} - \Delta t(\lambda_i + \alpha + (k-1)\Delta r) \geq 0$$

Since $\beta \geq \alpha + (k-1)\Delta r$, it suffices to choose $\Delta r$ and $\Delta t$ such that $1 - \sigma_i^2\beta\Delta t/(\Delta r)^2 - \Delta t(\lambda_i + \beta) \geq 0$ which is true if and only if

$$\frac{(\Delta r)^2}{\sigma_i\beta + (\Delta r)^2(\lambda_i + \beta)} \geq \Delta t$$

If $\kappa_i \geq 0$ it suffices to choose $\Delta r$ such that $\Delta r \leq -\alpha\sigma_i^2/\kappa_i(\theta_i - \beta)$ to make the coefficient of $Z^{(i)}_{j-1,k+1}$ non-negative and $\Delta r \leq \alpha\sigma_i^2/\kappa_i(\theta_i - \alpha)$ to make the coefficient of $Z^{(i)}_{j-1,k-1}$ non-negative. If $\kappa_i$ is negative then it suffices to choose $\Delta r \leq -\alpha\sigma_i^2/\kappa_i(\theta_i - \alpha)$ to make the coefficient of $Z^{(i)}_{j-1,k+1}$ and $\Delta r \leq \alpha\sigma_i^2/\kappa_i(\theta_i - \beta)$ to make the coefficient of $Z^{(i)}_{j-1,k-1}$ non-negative. Now define

$$\|Z_j\|_m = \sup_k\{|Z^{(1)}_{j,k}|, |Z^{(2)}_{j,k}|\}$$

This is the largest value of the European contingent claim under either state of the Markov Chain computed using the finite difference scheme at time $j\Delta t$. The finite difference scheme is said to be convergent if for any sequence of partitions of the state space $\{\Delta r_m\}$ such that $\Delta r_m$ converges to zero as $m$ goes to infinity, $(j+1)\Delta t$ converges to $t$ and $\Delta t$ converges to zero, $\|Z_j\|_m$ also converges to zero. Using the definition of $\|Z_j\|_m$, (12) and the non- negativity of the coefficients in (12) one obtains

$$\begin{aligned}\|Z_j\|_m &\leq (1-\Delta t(\alpha+(k-1)\Delta r))\|Z_{j-1}\|_m + O((\Delta t)^2) + O((\Delta r)^2\Delta t) \\ &\leq (1-\alpha\Delta t)\|Z_{j-1}\|_m + O((\Delta t)^2) + O((\Delta r)^2\Delta t)\end{aligned}$$

Recursing this inequality backwards yields

$$\begin{aligned}\|Z_j\|_m &\leq (1-\alpha\Delta t)^2\|Z_{j-2}\|_m + (1-\alpha\Delta t)[O((\Delta t)^2) + O((\Delta r)^2\Delta t)] \\ &+ O((\Delta t)^2) + O((\Delta r)^2\Delta t) \\ &\leq (1-\alpha\Delta t)^2\|Z_{j-2}\|_m + 2[O((\Delta t)^2) + O((\Delta r)^2\Delta t)] \\ &\leq \ldots \leq (1-\alpha\Delta t)^j\|Z_0\|_m + [O(\Delta t)^2 + O(\Delta t(\Delta r)^2)] \qquad (4.3)\end{aligned}$$

Since the value of the zero coupon bond at the initial time is known to be one, $\|Z_0\|_m$, the initial error in the finite difference approximation of the PDE system is zero. Furthermore, $j[O(\Delta t)^2 + O(\Delta t(\Delta r)^2)] = j\Delta t[O(\Delta t) + O((\Delta r)^2)]$, so since $j\Delta t$ converges to $t$ and $\Delta t$ and $\Delta x$ both converge to zero, the right hand side of (13) goes to zero. Thus, the explicit finite difference approximation converges to the true solution of the PDE system.

# BIBLIOGRAPHY

[Ait-Sahalia, Y. (1996)] Aït-Sahalia, Yacine. (1996). Testing continuous-time models of the spot interest rate. *R*eview of Financial Studies, *9*(2), 385–426.

[Cox, J., Ingersoll, J. and Ross, S. (1985)] Cox, J., J. Ingersoll and S. Ross. A theory of the term structure of interest rates. *E*conometrica, *5*3(2), 385–408.

[Eidelman, S. D. and Zhitarashu, N. V. (1998)] Eidelman, S.D. and N.V. Zhitarashu. *P*arabolic Boundary Value Problems. Basel: Birkhäuser Verlag.

[Protter, P. (1990)] Protter, P. *S*tochastic Integration and Differential Equations. A New Approach. New York: Springer-Verlag.

[Solonnikov, V. A. (1967)] Solonnikov, V.A. (1967). Boundary problems of mathematical physics III. Proceedings of the Steklov Institute of Mathematics *8*3, 1-184.

[Weerasinghe, A. (2005)] Weersinghe, A. (2005). A bounded variation control problem for diffusion processes. *S*IAM Journal of Control and Optimization, *4*4(2), 389–417.