Affine Lattice Models [1]

Claudio Albanese

Department of Mathematics, Imperial College of Science and Technology, University of London, SW7 2AZ, London, United Kingdom. mailto:claudio.albanese@imperial.ac.uk

Alexey Kuznetsov

Department of Mathematics, 100 St. George Street, University of Toronto, M5S 3G3, Toronto, Canada. mailto:kuznecov@math.toronto.edu

Abstract. We introduce a new class of lattice models based on a continuous time Markov chain approximation scheme for affine processes, whereby the approximating process itself is affine. A key property of this class of lattice models is that the location of the time nodes can be chosen in a payoff dependent way and one has the flexibility of setting them only at the relevant dates. Time stepping invariance relies on the ability of computing node-to-node discounted transition probabilities in analytically closed form. The method is quite general and far reaching and it is introduced in this article in the framework of the broadly used single-factor, affine short rate models such as the Vasiček and CIR models. To illustrate the use of affine lattice models in these cases, we analyze in detail the example of Bermuda swaptions.

Key words: Interest Rate Models, Affine Models, Birth and Death Processes, Discretization Scheme

JEL Classification: G13

Mathematics Subject Classification (2000): 35Q58, 39A12, 60H30

1 Introduction

Let $(r_t)_{t\geq 0}$ be a stationary Markov process. The (discounted) *generating function* of the process (r_t) is defined as

$$G(\tau, r_0, \alpha) = E\left(e^{\alpha r_\tau - \int_0^\tau r_s ds} \big| r_0\right), \tag{1.1}$$

[1] This work was written while the authors were visiting the Mathematics Department at the National University of Singapore, whose hospitality is gratefully acknowledged. The authors were supported in part by the National Science and Engineering Council of Canada under grant RGPIN-171149 and by the Ontario Council of Graduate Studies. This work was presented at ETH Zurich, USI Lugano, King's College London and at the National University of Singapore. We thank the participants for discussions. Special thanks go to Stephan Lawi and Ken Jackson. Finally, we are grateful to the referee for his many helpful comments. Remaining errors are our own.

where $\tau \geq 0$ and $\alpha \in \mathbb{C}$. The process (r_t) is called *affine* if there exist deterministic functions $m(t, \alpha)$ and $n(t, \alpha)$ such that for all r, α and $t \geq 0$

$$G(\tau, r, \alpha) = e^{m(\tau,\alpha)r + n(\tau,\alpha)}, \tag{1.2}$$

or, in other words, that the logarithm of generating functions is affine in r.

Affine processes play a central role in several classes of derivative models, ranging from interest rate to stochastic volatility and credit models. The success of affine processes is due to their analytical tractability and to the fact that they capture certain key statistical properties of processes such as those for short term interest rates and asset price volatility. The archetypical affine models are based on diffusion processes and are described by stochastic differential equations of the form

$$dr_t = (a - br_t)dt + \sigma r_t^{\beta} dW \tag{1.3}$$

where a, b, σ are constants and β equals either 0 or $\frac{1}{2}$. The case $\beta = 0$ corresponds to the Vasiček (Gaussian Ornstein-Uhlenbeck) process [6] and the case $\beta = \frac{1}{2}$ corresponds to the Cox-Ingersoll-Ross (CIR) process [7]. It has been shown in [8] and [9] that any affine process which is a time-homogenous, nonnegative diffusion is necessarily of the CIR type. However, there are also affine processes with jumps. As discussed in the Finance literature in [10], general non-negative affine processes correspond to the so called conservative CBI-processes (continuous state branching processes with immigration) and have been well studied, among others, by Kawazu and Watanabe in [11].

Empirical analysis based on affine interest-rate models include [12], [13], [14], [15], while the paper [16] extends the analysis to foreign exchange rates. Statistical methods developed specifically for the analysis of time-series data from affine models have been based on the approximation of the likelihood function [17] or on spectral properties, making use of the easily calculated complex moments of affine processes [18].

Affine diffusion models allow one to price in analytically closed form not only bonds, but also European style interest rate derivatives such as caplets and swaptions. However, to price path dependent claims including for instance barriers and optimal exercise features, one has to resort to numerical methods, either based on lattice discretizations or on Monte Carlo simulations. A class of binomial lattice models was proposed by Hull and White in [19], [20], see also [2]. These lattice models are characterized by a fixed set of time dates at which lattice nodes are placed and can be calibrated in such a way to achieve consistency with the initial term structure of interest rates at the nodes. These models admit a continuous limit corresponding to a version of the Vasiček with time dependent coefficients and are known as the Hull-White (HW) model. Monte Carlo methods are discussed by Longstaff and Schwarz in [4], and apply to multifactor models such as the market LIBOR models of Brace, Gatarek and Musiela [1] and Jamshidian [3].

In this article we introduce a new class of lattice models that admit the CIR and HW model as continuous limits. The key idea is to approximate the corresponding affine diffusion process by means of a discrete birth and death process, which is also affine. Our lattice models have the advantage of being defined in continuous time and of being analytically tractable to the same extent as their continuous limits are. This implies that, in the calibration stage, one can match the initial term structure of interest rates or implied volatilities at all dates in the future. In the valuation stage instead, one has the flexibility of setting the lattice nodes at the time dates which are most appropriate to the given payoff at hand without changing the model. Figure 1 illustrates the concept. Of central importance is the ability to analytically compute the following node-to-node discounted transition probabilities in a discrete state Markov process:

$$q(\tau, r_0, r_1) \equiv E_0 \left[\delta(r_\tau - r_1) e^{-\int_0^\tau r_s ds} \big| r_0 \right]. \tag{1.4}$$

Since the underlying process is defined in continuous time, for all sequences of dates $t_0 < t_1 < \cdots <$

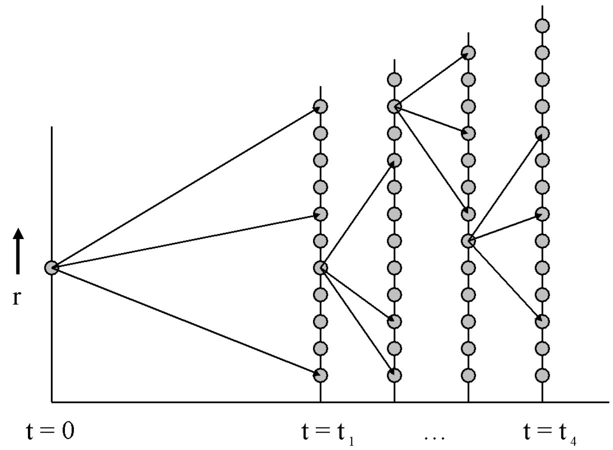

Figure 1: Lattice geometry

t_n we have that discounted transition probabilities satisfy the Chapman-Kolmogorov equation:

$$q(t_n - t_0, r_0, r_n) \equiv \sum_{r_1,\ldots,r_{n-1}} q(t_1 - t_0, r_0, r_1)\ldots q(t_n - t_{n-1}, r_{n-1}, r_n). \tag{1.5}$$

Notice that in our approach European style payouts can be priced with lattices involving only nodes at the maturity date. As an example, below we discuss in detail the case of Bermuda swaptions with three opportunities to exercise prior to maturity; this case corresponds to the lattice sketched in Figure 1; notice that nodes are required only at the exercise and maturity dates.

The paper is organized as follows. In Section 2, we define affine discrete birth-and death approximations for the CIR and HW processes. We also define and compute generating functions which, by inverting a Fourier series as discussed in Section 3, allows one to compute node-to-node transition probabilities. In Section 4, we discuss how to match the initial term structure of interest rates and, in Section 5, we give a step by step description of the construction of a lattice model for Bermuda swaptions.

2 Generating Functions

In this section we define the continuous time, discrete space birth-and-death approximations for the CIR and HW processes and compute the corresponding generating and characteristic functions. Here and throughout the article we assume that the process (r_t) which represents the instantaneous interest rate is a stationary Markov process and denote with $\mathcal{L}$ the corresponding generator.

We will remind that the (discounted) generating function of the process (r_t) was defined (see equation (1.1) as:

$$G(\tau, r, \alpha) = E\left[e^{\alpha r_\tau - \int_0^\tau r_s ds} \middle| r_0 = r\right], \tag{2.1}$$

where $t \geq 0$ and $\alpha \in \mathbb{C}$. Since the process (r_t) is stationary we could equivalently define the generating function as follows

$$G(\tau, r, \alpha) = E\left[e^{\alpha r_T - \int_t^T r_s ds} \middle| r_t = r\right], \tag{2.2}$$

where $\tau = T - t$. The *characteristic function* $F(t, r, \alpha)$ of the process (r_t) is defined as follows:

$$F(t, r, \alpha) = E\left[e^{i\alpha r_t} \middle| r_0\right]. \tag{2.3}$$

In the particular case $\alpha = 0$, the generating function reduces to the *discount function*:

$$B(t, T) = G(T - t, r_t, 0) = E\left[e^{-\int_t^T r_s ds} \middle| r_t\right].$$

Yields $y_t(T)$ are defined as follows:

$$y_t(T) = -\frac{\ln(B(t, T))}{T - t} \tag{2.4}$$

while the *continuously compounded instantaneous forward rate* is given by

$$f_t(T) = -\frac{\partial \ln(B(t, T))}{\partial T}. \tag{2.5}$$

The function $G(\tau, r, \alpha)$ can be evaluated by means of the following standard result:

Lemma 2.1. *Let $\mathcal{L}$ be the Markov generator of the process (r_t). The generating function $G(\tau, r, \alpha)$ solves the following evolution equation:*

$$-\frac{\partial G}{\partial \tau} + \mathcal{L}G = rG \tag{2.6}$$

with initial condition $G(0, r, \alpha) = e^{\alpha r}$.

Proof. Let $0 \leq t < T$ and $\tau = T - t$. Using equation (2.1), the generating function can be written as follows:

$$G(\tau, r_t, \alpha) = e^{\int_0^t r_s ds} E\left[e^{\alpha r_T - \int_0^T r_s ds} \middle| r_t\right]. \tag{2.7}$$

Let's define the function

$$g(t, T, r, \alpha) = E\left[e^{\alpha r_T - \int_0^T r_s ds} \middle| r_t = r\right]. \tag{2.8}$$

Since the process $X_t = g(t, T, r_t, \alpha)$ is a martingale for $t \leq T$, the function g solves the following equation:

$$\frac{\partial g}{\partial t} + \mathcal{L}g = 0. \tag{2.9}$$

Using equation (2.8) we can express the process $X_t = g(t, T, r_t, \alpha)$ as follows:

$$g(t, T, r_t, \alpha) = e^{-\int_0^t r_s ds} G(T - t, r_t, \alpha). \tag{2.10}$$

Inserting expression (2.10) into equation (2.21) leads to the equation (2.6) for G. □

Recall that the process (r_t) is defined as being *affine* if the logarithm of its generating function is an affine function of the short rate r (see equation (1.2). Well known examples of affine short rate dynamics are given by the Ornstein-Uhlenbeck and the CIR processes. The corresponding Markov generators are of the form

$$\mathcal{L} = (a - br)\nabla + \frac{1}{2}\sigma^2 r^{2\beta}\Delta. \tag{2.11}$$

The case $\beta = 0$ corresponds to the Ornstein-Uhlenbeck process while $\beta = \frac{1}{2}$ corresponds to the CIR case. Here $\nabla = \frac{\partial}{\partial r}$ and $\Delta = \frac{\partial^2}{\partial r^2}$. The discrete versions of these processes are given by the Markov generator

$$\mathcal{L}^h = (a - br)\nabla_+^h + \frac{1}{2}\sigma^2 r^{2\beta}\Delta^h \tag{2.12}$$

where $\beta \in \{0, \frac{1}{2}\}$, $b, h > 0$. In this case, the short rate r is restricted to take values in the grid

$$\Lambda(h, D) = h\mathbb{Z} \cap D \tag{2.13}$$

where the domain D equals $\mathbb{R}$ in case $\beta = 0$, while $D = \mathbb{R}_+$ if $\beta = \frac{1}{2}$. The operators ∇_+^h and Δ^h are finite difference approximations to ∇ and Δ and are defined as follows:

$$\nabla_+^h f(x) = \frac{1}{h}(f(x+h) - f(x)); \qquad \Delta^h f(x) = \frac{1}{h^2}(f(x+h) + f(x-h) - 2f(x)). \tag{2.14}$$

We refer to the process with $\beta = 0$ as to the Charlier process and denote it by $CHA(a, b, \sigma; h)$, while the process for $\beta = \frac{1}{2}$ will be called the Meixner process and will be denoted by $MEI(a, b, \sigma; h)$. This terminology refers to the fact that the eigenfunctions of the corresponding Markov generators are given by the Charlier and Meixner polynomials, see [5]. (Notice though that, for the specific purposes of this article, knowledge of the eigenfunctions is not required.) Since in the limit as $h \to 0$ the one parameter family of Markov generators $\mathcal{L}^h$ converges to the generator $\mathcal{L}$, given by equation (2.11), the Charlier and Meixner processes can be regarded as the lattice approximations for the Ornstein-Uhlenbeck process $OU(a, b, \sigma)$ and of the CIR process $CIR(a, b, \sigma)$, respectively.

The Meixner and Charlier are sometimes called (discrete) birth and death processes. Birth and death processes behave very much like diffusion processes, the main similarity being that transitions from state i to state $j > i$ only occur if the process takes up all the intermediate values $i+1, i+2, \ldots, j-1$ between i and j. The following is the main theorem in this article and shows that both the Meixner and Charlier interest rate processes are affine and analytically solvable, by explicitly evaluating the functions $m(t, \alpha)$ and $n(t, \alpha)$:

Theorem 2.2. *In the two cases $\beta = 0$ and $\beta = \frac{1}{2}$, the short-rate model with generator (2.12) is affine for all values of the parameters.*

(i) If $\beta = 0$ (i.e. (r_t) is a Charlier process) the functions m and n are given by the following formulas:

$$\begin{cases} m(t,\alpha) = \alpha - \frac{1}{h}\log[1 + (1 - e^{-(b-h)t})(\frac{b}{b-h}e^{h\alpha} - 1)] \\ n(t,\alpha) = \frac{\sigma^2 + 2ah}{2h^2 b}\left(\log[1 + (1 - e^{-(b-h)t})(\frac{b}{b-h}e^{h\alpha} - 1)] - th\right) + \\ \qquad + \frac{\sigma^2}{2h^2(b-h)}\left((1 - e^{-(b-h)t})(e^{-h\alpha} - \frac{b}{b-h}) + th\right) \end{cases} \tag{2.15}$$

(ii) If $\beta = \frac{1}{2}$ (i.e (r_t) is a Meixner process) the functions m and n are given by the following formulas:

$$\begin{cases} m(t,\alpha) = \frac{1}{h}\log\left[1 + M_1 - \frac{\gamma}{\eta} + \frac{e^{h\alpha} - 1 - M_1 + \frac{\gamma}{\eta}}{1 + \frac{\eta}{\gamma}(1 - e^{\gamma t})(e^{h\alpha} - 1 - M_1)}\right] \\ n(t,\alpha) = \frac{a}{h}\left(M_1 t - \frac{1}{\eta}\log\left[1 + \frac{\eta}{\gamma}(1 - e^{\gamma t})(e^{h\alpha} - 1 - M_1)\right]\right) \end{cases} \tag{2.16}$$

where $\eta = \frac{\sigma^2}{2h} - b$, $\gamma = \sqrt{(b+h)^2 + 4h\eta}$ and $M_1 = \frac{b+h+\gamma}{2\eta}$.

Proof. By using (1.2) as an ansatz to solve the equation for the discount factor (2.6), we find

$$-\dot{m}r - \dot{n} + (a-br)\frac{1}{h}(e^{hm}-1) + \frac{1}{2}\sigma^2 r^{2\beta}\frac{1}{h^2}(e^{hm}+e^{-hm}-2) = r. \tag{2.17}$$

In case $\beta = 0$, this equation splits into the two differential equations

$$\begin{cases} \dot{m} = -b\frac{1}{h}(e^{hm}-1) - 1 \\ \dot{n} = a\frac{1}{h}(e^{hm}-1) + \frac{1}{2}\sigma^2\frac{1}{h^2}(e^{hm}+e^{-hm}-2), \end{cases} \tag{2.18}$$

while if $\beta = \frac{1}{2}$ we find the following system:

$$\begin{cases} \dot{m} = -b\frac{1}{h}(e^{hm}-1) - 1 + \frac{1}{2}\sigma^2\frac{1}{h^2}(e^{hm}+e^{-hm}-2) \\ \dot{n} = a\frac{1}{h}(e^{hm}-1). \end{cases} \tag{2.19}$$

In either case, the initial conditions are $m(0,\alpha) = \alpha$, $n(0,\alpha) = 0$.

We will illustrate how to solve these equation on the example of Meixner process (the case of Charlier process can be treated similarly). Let's introduce the new function $M(t,\alpha) = e^{hm(t,\alpha)} - 1$. By multiplying the first equation in the system (2.19) by he^{hm}, we arrive at

$$\begin{aligned} \dot{M} = he^{hm}\dot{m} &= -b(e^{2hm}-e^{hm}) - he^{hm} + \frac{1}{2}\frac{\sigma^2}{h}(e^{2hm}-2e^{hm}+1) \\ &= \eta M^2 - (b+h)M - h, \end{aligned} \tag{2.20}$$

where we denote $\eta = \frac{\sigma^2}{2h} - b$. The initial condition is $M(0,\alpha) = e^{h\alpha} - 1$. Equation (2.20) is a Riccati equation and can be solved explicitly:

$$M(t,\alpha) = \frac{M_1 - M_2 e^{\gamma t}\frac{M_1 - e^{h\alpha}-1}{M_2 - e^{h\alpha}-1}}{1 - e^{\gamma t}\frac{M_1 - e^{h\alpha}-1}{M_2 - e^{h\alpha}-1}}, \tag{2.21}$$

where $\gamma = \sqrt{(b+h)^2 + 4h\eta}$ and $M_{1,2} = \frac{b+h\pm\gamma}{2\eta}$ are the roots of the second order polynomial $\eta M^2 - (b+h)M - h$. By simplifying (2.21), we arrive at the expression for the function $m(t,\alpha,\lambda)$ in equation (2.16). Having found the function $M(t,\alpha)$, the function n can be computed using second equation in the system (2.19) as the following integral

$$n(t,\alpha) = \frac{a}{h}\int_0^t M(s,\alpha)ds.$$

□

Corollary 2.3. *In the two cases $\beta = 0$ and $\beta = \frac{1}{2}$, the characteristic function $F(t,r,\alpha)$ of the process can be computed analytically and is given by the following expression:*

$$F(t,r,\alpha) = e^{\bar{m}(t,\alpha)r + \bar{n}(t,\alpha)} \tag{2.22}$$

where the functions $\bar{m}(t,\alpha)$ and $\bar{n}(t,\alpha)$ are given as follows:

(i) if $\beta = 0$ (i.e (r_t) is a Charlier process), then

$$\begin{cases} \bar{m}(t,\alpha) = i\alpha - \frac{1}{h}\log[1 + (1-e^{-bt})(e^{ih\alpha}-1)] \\ \bar{n}(t,\alpha) = \frac{\sigma^2}{2bh^2}((1 + \frac{2ah}{\sigma^2})\log[1 + (1-e^{-bt})(e^{ih\alpha}-1)] + (1-e^{-bt})(e^{-ih\alpha}-1)) \end{cases} \tag{2.23}$$

(ii) if $\beta = \frac{1}{2}$ (i.e (r_t) is a Meixner process), then

$$\begin{cases} \bar{m}(t,\alpha) = -\frac{1}{h}\log\left[1 + \frac{\frac{2bh}{\sigma^2}(1-e^{-ih\alpha})e^{-bt}}{(1-e^{-ih\alpha})(1-e^{-bt}) - \frac{2bh}{\sigma^2}}\right] \\ \bar{n}(t,\alpha) = \frac{\frac{2a}{\sigma^2}}{\frac{2bh}{\sigma^2}-1}\log\left[1 + (e^{ih\alpha}-1)(1-e^{-bt})(1-\frac{\sigma^2}{2bh})\right] \end{cases} \tag{2.24}$$

3 Computing discounted transition probabilities by Fourier methods

Let (r_t) be a stationary Markov process on the lattice $\Lambda(h, \mathbb{R}_+)$ with generator $\mathcal{L}^h$ given by the equation in (2.12).

Definition 3.1. Let $0 \leq t_0 \leq t_1$ and $t = t_1 - t_0$. The *transition probabilities* for (r_t) are defined as follows:

$$p_{jk}(t) = \mathrm{P}\left[r_{t_1} = kh | r_{t_0} = jh\right] = E\left[\delta(r_{t_1} - kh) | r_{t_0} = jh\right], \tag{3.1}$$

Here $\delta(r)$ is the function such that $\delta(r) = 1$ if $r = 0$, while otherwise $\delta(r) = 0$. The *discounted transition probabilities* $q_{jk}(t)$ are defined as the following expectations:

$$q_{jk}(t) = E\left[\delta(r_{t_1} - kh) e^{-\int_{t_0}^{t_1} r_s ds} \big| r_{t_0} = jh\right] \tag{3.2}$$

By means of these functions, one can express discounted expectations as follows:

$$E\left[f(r_T) e^{-\int_{t_0}^{T} r_s ds} \big| r_{t_0} = jh\right] = E\left[\sum_k f(kh) \delta(r_T - kh) e^{-\int_{t_0}^{T} r_s ds} \big| r_{t_0} = jh\right] =$$
$$= \sum_k f(kh) E\left[\delta(r_T - kh) e^{-\int_{t_0}^{T} r_s ds} \big| r_{t_0} = jh\right] = \sum_k f(kh) q_{jk}(T - t_0). \tag{3.3}$$

Theorem 3.2. *Discounted transition probabilities $q_{jk}(t)$ can be can be evaluated as the following inverse Fourier transform:*

$$q_{jk}(t) = \int_0^1 e^{-2\pi i \omega k} G\left(t, jh, 2\pi i \frac{\omega}{h}\right) d\omega. \tag{3.4}$$

Proof. The definition of generating function can be recast as follows:

$$G(t, jh, \alpha) = \sum_{l=0}^{\infty} q_{jl}(t) e^{\alpha l h}. \tag{3.5}$$

The integral on the right side of equation (3.4) equals:

$$\int_0^1 e^{-2\pi i \omega k} \sum_{l=0}^{\infty} q_{jl}(t) e^{2\pi i \omega l} d\omega = \sum_{l=0}^{\infty} q_{jl}(t) \int_0^1 e^{-2\pi i \omega k} e^{2\pi i \omega l} d\omega = q_{jk}(t) \tag{3.6}$$

since

$$\int_0^1 e^{-2\pi i \omega k} e^{2\pi i \omega l} d\omega = \delta(l - k.) \tag{3.7}$$

□

Corollary 3.3. *The transition probabilities $p_{jk}(t)$ can be evaluated as the following inverse Fourier transform:*

$$p_{jk}(t) = \int_0^1 e^{-2\pi i \omega k} F\left(t, jh, 2\pi \frac{\omega}{h}\right) d\omega. \tag{3.8}$$

4 Matching the term structure of interest rates

4.1 The Charlier case

The Hull-White dynamics for interest rates can be described as follows:

$$r_t = \phi(t) + x_t \tag{4.1}$$

where x_t is $OU(0, b, \sigma)$. We consider here the discrete version of the Hull-White model, where the process x_t is replaced by Charlier process $Cha(0, b, \sigma; h)$.

Definition 4.1. Bond prices are defined as follows:

$$B(t,T) = E\left[e^{-\int_t^T r_s ds} \big| r_t\right] \tag{4.2}$$

where $r_t = \phi(t) + x_t$ and x_t is $CHA(0, b, \sigma; h)$.

Notice that thanks to theorem 2.2, the process $B(t,T)$ is given by

$$\begin{aligned} B(t,T) &= e^{-\int_t^T \phi(s)ds} E\left[e^{-\int_t^T x_s ds} \big| r_t\right] = e^{-\int_t^T \phi(s)ds} G(T-t, x_t, 0) = \\ &= e^{-\int_t^T \phi(s)ds + m(T-t,0)x_t + n(T-t,0)} \end{aligned} \tag{4.3}$$

where the functions m and n are given by equation (2.15). The continuously compounded instantaneous forward rate is given by

$$f_t(T) = -\frac{\partial \ln(B(t,T))}{\partial T} = \phi(T) - \dot{m}(T-t,0)x_t - \dot{n}(T-t,0) \tag{4.4}$$

for all $t < T$, and yields $y_t(T)$ are

$$y_t(T) = -\frac{\ln(B(t,T))}{T-t} = \frac{1}{T-t}\left(\int_t^T \phi(s)ds - m(T-t,0)x_t - n(T-t,0)\right). \tag{4.5}$$

Notice that the initial yield curve $\hat{y}(T)$ at time $t = 0$ uniquely identifies the function $\phi(s)$ such that $y_0(T) = \hat{y}(T)$.

4.2 The Meixner case

Let X_t be the $MEI(a, b, \sigma, h)$. To match the term structure of interest rates in the Meixner case we make use of a deterministic time change of the form $s = s(t)$ where $s(t)$ is a monotonously increasing function with $s(0) = 0$ and inverse function $t(s)$. Let $S > 0$ be a fixed maturity and let $T = t(S)$. The interest rate process is defined as follows:

$$r_s = x_{t(s)} \frac{dt}{ds} \tag{4.6}$$

and the discount function under the new time coordinate s is given by

$$B(s,S) = E\left[e^{-\int_s^S r_s ds} \big| r_s\right] = E\left[e^{-\int_{s(t)}^{s(T)} x_{t(s)} \frac{dt}{ds} ds} \big| x_t\right] = E_t\left[e^{-\int_t^T x_u du} \big| x_t\right]. \tag{4.7}$$

Since x_t is a Meixner process, the discount function $B(s,S)$ can be computed explicitly as described in theorem (2.2). The yield curve is given by

$$y_s(S) = -\frac{m(t(S)-t(s),0)X_{t(s)} + n(t(S)-t(s),0)}{S-s}. \tag{4.8}$$

If $y_0(S)$ is the initial yield curve, the condition specifying the time-change function is

$$-m(t(S),0)X_0 - n(t(S),0) = S\hat{y}_0(S). \tag{4.9}$$

Notice that, in the Meixner case, rates stay positive, i.e. $X_0 \geq 0$, and the functions $m(t,0)$ and $n(t,0)$ are decreasing in t for all choices of the parameters. Since $-m(T,0)X_0 - n(T,0)$ is also increasing in T, for each maturity S there exists one and only one value $T = t(S)$ satisfying equation (4.9) and the time change function $t(s)$ is also monotonously increasing.

5 Example: Pricing Bermuda swaptions

In this section, we describe in detail how to construct lattice approximations for the HW and CIR models in the particular case of Bermuda swaptions. As an example, we consider a swaption struck at the rate $K = 7\%$, maturing in 5 years and early exercise possible on the following dates: $t_1 = 3.5y, t_2 = 4y, t_3 = 4.5y, t_4 = 5y$. If exercised at time t, the payoff function for the swaption is the present value of the underlying swap of rate K and unit nominal, i.e.

$$PV_t = PV(r_t,t) = (Z_t(T_m) - Z_t(T_0)) + \frac{1}{2}\sum_{l=1}^{m} KZ_t(T_l) \tag{5.1}$$

where $T_0, T_1, ...T_m$ are the cash-flow dates for the underlying swap. In our example, we chose $m = 10$ and $T_0, T_1, \ldots, T_m = (5.5, 6, \ldots, 10.5)$.

We make use of the following parameters for the short rate process:

(i) in the case of the Charlier process, we set $a = 0.005,\ b = 0.1,\ \sigma = 0.013,\ h = \frac{a}{50b}$.

(ii) in the case of the Meixner process, we set $a = 0.05,\ b = 1,\ \sigma = 0.1,\ h = \frac{a}{50b}$.

The risk neutral measure Q corresponds to choosing the money-market process $B_t = e^{\int_0^t r_s ds}$ as numeraire. The present value function $PV(r,t)$ satisfies the following martingale condition under Q:

$$PV(jh,t) = E\left[PV(r_{t+\Delta t}, t+\Delta t)\frac{B_t}{B_{t+\Delta t}}\ \middle|\ r_t = jh\right] = \sum_k q_{jk}(\Delta t)PV(kh, t+\Delta t) \tag{5.2}$$

where the q_{jk} are the discounted transition probabilities in (3.2). The error introduced in this relation by cutting off the lattice is a useful measure to assess the degree of accuracy of the lattice model.

A convenient method to numerically evaluate the matrix of discounted probabilities $q_{jk}(t)$ is based on the Fourier integral in equation (3.4). To assess the impact of the errors arising from the numerical quadrature, it is useful to measure the discrepancies in the martingale condition in equation (5.2). Notice that the probabilities $q_{jk}(t)$ decrease very rapidly as $|k-j|$ increases. Hence, for each fixed pair j and t, we can find an interval $[k_1, k_2]$ such that if $k \notin [k_1, k_2]$ then $q_{jk}(t)$ is negligibly small and thus truncate the support of this function. This enables us to pass to a discrete Fourier transform. Due to the steep decay properties of $q_{jk}(t)$ as a function of k, it turns out that the number of points required to achieve a precision of order 10^{-9} is of the order of a few hundreds and is thus within easy reach. Below we give a detailed, step-by-step description of this construction.

To avoid arbitrage, we need to ensure that the interest rate process r_t stays nonnegative. This is obvious by construction in the case of the Meixner process, while in the Charlier case (as in the corresponding Ornstein-Uhlenbeck case) the interest rate process r_t can possibly attain negative values with positive probability. We thus have to ensure that this probability is negligibly small. This can be accomplished by restricting the parameters $a,\ b$ and c as indicated above. Whether this restriction is too stringent depends obviously on how low the short rate is as compared to implied volatilities; in

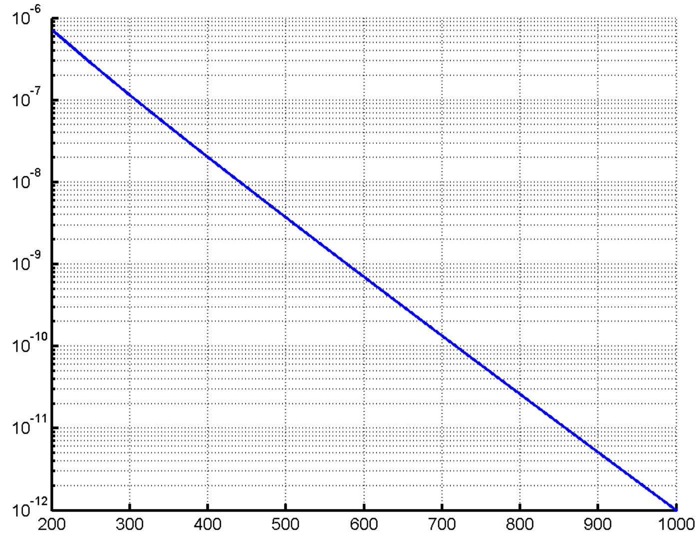

Figure 2: The dependence of the relative error in computing the discounted transition probabilities on the upper boundary k_2.

a low interest rate environment, the Charlier-Hull-White model may well not be applicable. Assuming that the probability that the short rate takes up negative values is either negligible or rigorously zero, the discounted probabilities $q_{jk}(t)$ are bounded from above by the (undiscounted) transition probabilities, i.e. $p_{jk}(t)$, that is

$$q_{jk}(t) \leq p_{jk}(t). \tag{5.3}$$

Hence, for each pair j, t, transition probabilities $q_{jk}(t)$ can be computed by the following procedure:

- Fix a small ϵ; in our experiment we set $\epsilon = 10^{-9}$.
- Fix the interval $[k_1, k_2]$. In our experiment we use the Meixner process to model the interest rate, thus we can choose $k_1 = 0$. In order to compute k_2 we use the closed form expressions for the transition probabilities (given by formula (A.2)) to approximate the corresponding probabilities on the lattice. In our experiment we found that to achieve the desired accuracy we can take $k_2 = 600$. Next, we truncate $q_{jk}(t)$ to the interval $[0, k_2]$.
- Compute the discounted transition probabilities using the following formula:

$$q_{jk}(t) = \frac{1}{k_2 + 1} \sum_{n=0}^{k_2} G(t, jh, \alpha_n) e^{-2\pi i \frac{nk}{k_2+1}} \tag{5.4}$$

  where $\alpha_n = 2\pi i \frac{n}{(k_2+1)h}$, $n = \{0, 1, .., k_2\}$. This formula can be justified as follows: Since $q_{jk}(t)$ is supported in $k \in [0, k_2]$ (neglecting the terms of order ϵ), the definition of the generating function can be recast as follows:

$$G(t, jh, \alpha_n) = \sum_{k=0}^{k_2} q_{jk}(t) e^{\alpha_n kh} = \sum_{k=0}^{k_2} q_{jk}(t) e^{2\pi i \frac{kn}{k_2+1}} \tag{5.5}$$

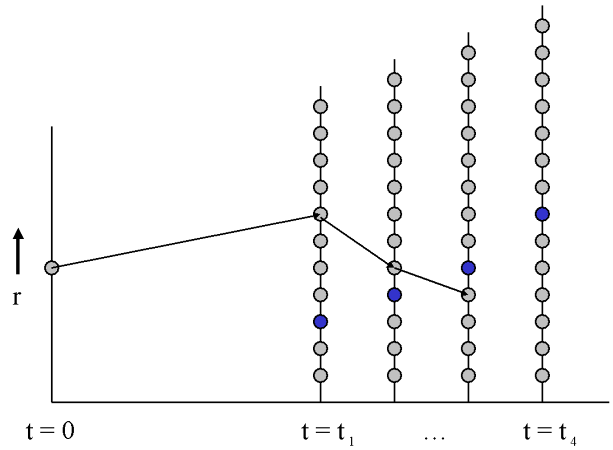

Figure 3: Continuation region and the optimal exercise boundary.

and we see that $G(t, jh, \alpha_n)$ as a function of n is the discrete Fourier transform of $q_{jk}(t)$. Thus $q_{jk}(t)$ can be computed as the inverse discrete Fourier transform of $G(t, jh, \alpha_n)$.

- After the matrix of discounted transition probabilities $q_{jk}(t)$ is computed, we verify the martingale condition (5.2): if all the estimates (for k_1, k_2) are correct, then the error will turn out to be of order ϵ.

After one computes with high precision the discounted transition probabilities, one can proceed to price the Bermuda swaption $PB_t = PB(r_t, t)$ by backward induction. Following the standard principles of stochastic optimization theory, we start with the time of maturity $t = t_4$ and make our way backwards up until current time. At maturity, the swaption price is given by $PB(r) = (PV_t(r, t_4) - K)^+$. To accomplish the first iteration step from time t_4 to time t_3, for each node j, the vector of discounted probabilities $q_{jk}(t_4 - t_3)$ is computed. Let $PB(r, t_3 + 0)$ be the discounted expectation of the option payoff conditioned to no exercise at time $t = t_3$. Our choice of notation indicates that this price coincides with the price of $PB(r_{t_3+\delta t}, t_3 + \delta t)$ for $\delta t \downarrow 0$, i.e.

$$PB(jh, t_3 + 0) = \sum_k q_{jk}(t_4 - t_3) PB(kh, t_4) \tag{5.6}$$

By comparison with the payoff in case on chooses to exercise at time $t = t_3$ and choosing the larger among the two one determines the exercise strategy and the pricing function

$$PB(jh, t_3) = \max(PB(jh, t_3 + 0), PV(jh, t_3)). \tag{5.7}$$

This procedure is then iterated backwards up until current time $t = 0$. The optimal exercise time for a Bermuda swaption is the first time $t \in \{t_1, ..t_4\}$ such that the value of Bermuda swaption PB_t equals the value of the swaption PV_t. Since for each fixed time t the price of the Bermuda swaption $PB(r, t)$ is monotone in the short rate r, the *optimal exercise boundary* r^*_t is defined as follows:

$$r^*_{t_k} = \sup\{r | PB(r, t_k) < PV(r, t_k)\} \tag{5.8}$$

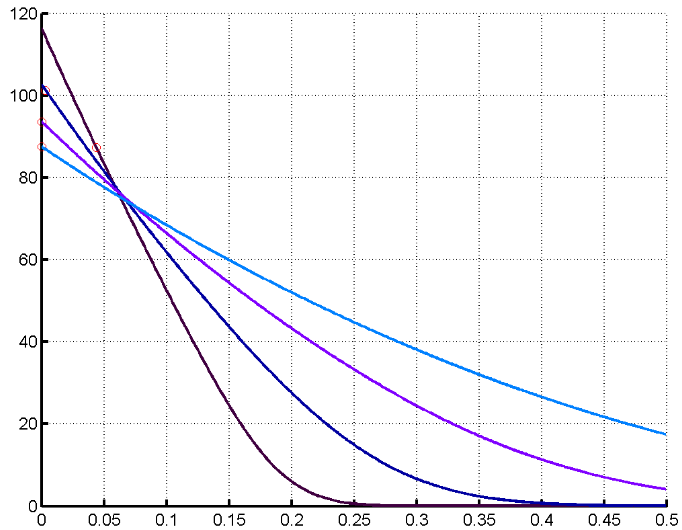

Figure 4: Prices of bermuda swaptions and the optimal exercise boundary.

The optimal stopping time τ^* is then

$$\tau^* = \inf\{t_n : n = 1, \dots 4; r_{t_n} \leq r^*_{t_n}\}. \tag{5.9}$$

The lattice is thus divided into two non-intersecting parts - the *continuation region*, $\{(t,r) : r > r^*_t\}$ and the *stopping region* $\{(t,r) : r \leq r^*_t\}$. The optimal exercise rule is that exercise occurs at the first point in time when we reach the stopping region.

Prices of bermuda swaptions for $t = 3.5,\ 4,\ 4.5,\ 5$ *and* $r = 1\% \dots 10\%$

	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
$t = 3.5$	110.29	103.60	97.00	90.49	84.05	77.71	71.44	65.26	59.16	53.15
$t = 4$	98.71	94.40	90.16	85.97	81.85	77.80	73.80	69.86	65.98	62.16
$t = 4.5$	91.01	88.13	85.30	82.51	79.77	77.07	74.40	71.78	69.20	66.66
$t = 5$	85.65	83.63	81.65	79.69	77.77	75.87	74.00	72.16	70.35	68.56

Remark 5.1. An alternative to using Fourier transforms is to evaluate the generating function at a sequence of real values α_n of the parameter α and invert a discrete Laplace transform. This leads to the problem of inverting the Vandermonde matrix $v_{jn} = (x_n)^j, n, j = 0, 1, ...N$ where $x_n = e^{h\alpha_n}$. The reason why we recommend using Fourier series instead, is that our numerical experiments indicate that if N is chosen sufficiently large, the Vandermonde matrix we obtained was ill-conditioned and impossible to invert numerically.

6 Concluding Remarks

In this article we introduce a new discretization scheme for a broadly used class of short rate models based on affine processes. Lattice models constructed in this framework have a direct financial interpretation and are thus fully consistent. Moreover, they are analytically tractable and lend themselves to efficient calibration schemes. In our framework, time nodes are defined only at those dates which are relevant to a specific payoff, while node-to-node transition probabilities are evaluated analytically. To illustrate the implementation of this class of lattice models, we discuss in full detail the example of Bermuda swaptions.

References

[1] A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7, 1997.

[2] D. Backus, L. Wu, and S. Ziu. Markov chain approximation for term structure models. *working paper*, 1999.

[3] F. Jamshidian. LIBOR and swap market models and measures. *Finance and Stochastics*, 1, 1997.

[4] Longstaff, F. A., and Scwarz, E. S. Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies*, 14, pages 113–147, 2001.

[5] G. Szego. Orthogonal Polynomials *American Mathematical Society*, 1959.

[6] O. Vasicek. An equilibrium characterization of the term structure. *J. Finan. Econom.* ,5, pages 177–188, 1977.

[7] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica 53*,pages 385–408, 1985.

[8] R. G. Brown and S. M. Schaefer. Interest rate volatility and the shape of the term structure. *Phil. Trans. R. Soc. Lond.*, A 347, pages 563–576, 1994.

[9] D. Duffie and N. Garleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, Forthcoming .

[10] D. Filipović. A general characterization of one factor affine term structure models. *Finance Stoch.*,5,no. 3, pages 389–412, 2001.

[11] K. Kawazu and S. Watanabe. Branching processes with immigration and related limit theorems. *Theory Probab. Appl.*,16, pages 36–54, 1971.

[12] R. Brown and S. Schaefer The term structure of real interest rates and the Cox, Ingersoll, and Ross model. *Journal of Financial Economics*, 35, pages 3–42, 1994.

[13] Q. Dai and K. Singleton. Specification analysis of affine term structure models. *Journal of Finance*,55, 1943–1978, 2000.

[14] M. Gibbons and K. Ramaswamy. A test of the Cox-Ingersoll-Ross model of the term structure of interest rates. *Review of Financial Studies*, 6, pages 619–658, 1993.

[15] N. Pearson and T.-S. Sun. An empirical examination of the Cox, Ingersoll, and Ross model of the term structure of interest rates using the method of maximum likelihood. *Journal of Finance*,54, pages 929–959, 1994.

[16] D. Backus, S. Foresi, and C. Telmer. Affine term structure models and the forward premium anomaly. *Journal of Finance*,56, pages 279–304, 2001.

[17] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68* ,no. 6, 1343-1376, 2000.

[18] K. Singleton. Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102, pages 111–141, 2001.

[19] Hull, J., and A. White. Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *Journal of Derivatives*,2, pages 7–16, 1994.

[20] Hull, J., and A. White. Using Hull-White Interest Rate Trees. *Journal of Derivatives*,3, pages 26–36, 1996.

Appendix

Here we present the formulas for the probability densities of the Ornstein-Uhlenbeck and CIR processes, which are used in section 5 for estimating the cut-off boundaries k_i.

(i) In the case of the Ornstein-Uhlenbeck process, the transition probability density is given by

$$p_t(x,y) = \sqrt{\frac{b}{2\pi \sinh(bt)\sigma^2}} \exp\left(-\frac{(b(x-y)e^{-bt} - a(1-e^{-bt}))^2}{2b\sigma^2 \sinh(bt)}\right) \tag{A.1}$$

(ii) In the case of the CIR process, the transition probability density is

$$p_t(x,y) = c_t \left(\frac{ye^{bt}}{x}\right)^{\frac{1}{2}(\frac{2a}{\sigma^2}-1)} \exp\left[-c_t(xe^{-bt}+y)\right] I_{\frac{2a}{\sigma^2}-1}\left(2c_t\sqrt{xye^{-bt}}\right), \tag{A.2}$$

where $c_t \equiv -2b/(\sigma^2(e^{-bt}-1))$.