PRE-PRINT
UD 2015-01

Review of Lattice Construction Methods - Generalizing Hull & White -

Jorrit-Jaap de Jong* Michele Maio†

uGly Duckling B.V.
Burgemeester le Fevre de Montignylaan 30,
3055LG Rotterdam, the Netherlands

Stratos Nikolakakis‡

ABN AMRO
Gustav Mahlerlaan 10,
1000 EA Amsterdam, the Netherlands

December 29, 2015

Abstract

This paper extends the generalized procedure for building trees for short rates by Hull & White. A generalization for any mean and standard deviation of the underlying short rate model is presented. In addition we review the methodologies for constructing lattice models and give a step-by-step explanation on how to construct trinominal trees. We apply the formalism to some explicit examples of various complexity.

Keywords: Short rates, Hull-White, lattice construction.

*j.jong@uglyduckling.nl
†m.maio@uglyduckling.nl
‡stratos.nikolakakis@nl.abnamro.com - **Disclaimer**: The views and opinion expressed in this paper are those of the author and do not represent the views of ABN AMRO

Contents

1 Introduction

In a recent contribution Hull & White [1] generalize their previous work on short rate models using trinomial trees [2]-[3] to a wide class of models where the volatility is an arbitrary rate-dependent quantity and the drift is allowed to have a time dependency. However such time dependence is not completely arbitrary, since in their set-up the drift can be written as the sum of a time-dependent and a rate-dependent function.

In this paper we will present the most general framework for evaluating short rate models. Specifically we present a framework to deal with all short rate models of the form

$$dr = F(r,t)dt + G(r,t)dZ(t)\,,$$

for generic functions F and G, which can in principle contain both time and rate dependencies.

In most cases analytical formulas for discounts and (option) prices will not be available. For those cases one must resort to numerical methods. Here a lattice construction algorithm is presented that can be used for (arbitrage free) pricing for any specification of F and G.

The fundamental insight is that there exist a function $x = f(r)$ so that

$$dx = H(x,t)dt + dZ.$$

This implies that in general we can construct a lattice for x using a corresponding function H which depends on the given model specifications.

Once the lattice has been constructed, the parameters of F and G need to be taken in such a way that the model prices the market. To do this we present a pricing function that aims to match the spot yield curve and cap volatility in the market. The way this is achieved is illustrated using some explicit choices for F and G.

2 Generalized Continuous Short Rate Model

We consider the following set up for our short rate process:

$$dr = F(r,t)dt + G(r,t)dZ(t) \tag{1}$$

Equation 1 generalizes the set up used by Hull & White [1] (the latter will be reviewed in section 3.2.1). We define the Uhlenbeck-Ornstein process [4] for x as

$$\frac{dx}{dr} = \frac{1}{G(r,t)} \tag{2}$$

or equivalently

$$x = f(r) = \int_c^r \frac{1}{G(r',t)}dr'\,. \tag{3}$$

Using Ito for the process dx we find:

$$dx = \left[\int \frac{\partial}{\partial t}\left(\frac{1}{G(r',t)}\right)dr' + \frac{F(r,t)}{G(r,t)} - \frac{1}{2}\frac{\partial G(r,t)}{\partial r}\right]dt + dZ(t). \tag{4}$$

Note that we have moved the partial time-derivative inside the integral sign: for this to be allowed it is enough that the integrand, in this case the function G^{-1}, is continuous and differentiable and the integration set is finite[1].

To ease notation, we introduce

$$H(x,t) = -\int \frac{\dot{G}(r',t)}{G^2(r',t)}dr' + \frac{F(r,t)}{G(r,t)} - \frac{1}{2}G'(r,t)\,, \tag{8}$$

where

$$\dot{G}(r,t) \equiv \frac{\partial G}{\partial t} \qquad \text{and} \qquad G'(r,t) \equiv \frac{\partial G}{\partial r}\,. \tag{9}$$

So equation (4) reduces to

$$dx = H(x,t)dt + dZ(t). \tag{10}$$

Note that we can compute r from x in $H(x,t)$ as

$$r = f^{-1}(x) \tag{11}$$

which follows directly from (3).

Next we need the mean and variance, or first (m_1) and second (m_2) moments of the process of x. We recover the process for x by taking the integral of equation (10):

$$x(s) = x(s_0) + \int_{s_0}^{s} H(x(s'),s')ds' + \int_{s_0}^{s} dX(s) \tag{12}$$

We can now compute the first moment by taking the expectation

$$m_1 = \mathbb{E}\left[x(t)|x(s)\right]\,,$$

where $t > s$. We find:

$$m_1 = x(s) + \int_{s}^{t} H(x(s'),s')ds'\,. \tag{13}$$

The second moment is then given by

$$\mathbb{V}\left[x(t)|x(s)\right] = m_1^2 + (t-s)\,. \tag{14}$$

[1] More generally, using the definition of derivative, for any function $f(x,t)$ we can write:

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} f(x,t) = \lim_{h\to 0}\int_{x_1}^{x_2}\left(\frac{f(x,t+h)-f(x,t)}{h}\right)\,. \tag{5}$$

Recalling Lebesgue's Dominated Convergence Theorem, as long as for all h the integrand is bounded by an integrable function, we can move the limit sign inside the integral. This is always the case for continuous and differentiable functions on a finite set. In fact, it is easy to prove that

$$\left|\frac{f(x,t+h)-f(x,t)}{h}\right| \leq \max_{x,t}\left|\frac{\partial f}{\partial t}\right| = M < +\infty\,. \tag{6}$$

Proof: using the Lagrange form of the Taylor remainder at order one, there exists a $\theta \in [0,1]$ such that the following relations hold

$$\left|\frac{f(x,t+h)-f(x,t)}{h}\right| = \left|\frac{\partial f}{\partial t}(x,t+\theta h)\right| \leq \max_{t}\left|\frac{\partial f}{\partial t}\right| \leq \max_{x,t}\left|\frac{\partial f}{\partial t}\right| = M < +\infty\,. \tag{7}$$

3 Discrete Framework

3.1 Introduction

For most model specifications it will not be possible to compute analytic bonds and option prices, hence we must resort to numerical methods. Hull & White suggest a lattice building approach in [1, 2, 3]. We will do the same here only using a more generic specification of the short rate model. We will first derive the discretization for finite difference scheme (section (3.2)). The construction of the lattice for dx will be discussed in details later in section 3.3.

3.2 Finite Difference set up

In this section we will review the set up of Hull and White [1]. First we will recover the result from [1] (section 3.2.1). Then we will derive the same result for the generalized case given by equation (1) (section 3.2.2). Following most of the literature, our tree will be defined in terms of nodes labelled by a time coordinate t and a space coordinate x, both of which assume discrete values. Starting at time $t_0 = 0$, all the subsequent time values will be denoted by $t_i = i\Delta t$, $i \in \mathbb{N}$, for some small -but finite- increment Δt. Similarly, the discrete x-values will be denoted by x_j, $j \in [j_d, j_u]$, where j_d and j_u are the lower and upper bound for the index j. For each given time t_i, many x_j values are possible. Hence, each node will be uniquely fixed in terms of its coordinate (t_i, x_j), or more compactly (i, j).

3.2.1 Hull-White set up

In the set up of Hull and White, the starting point is the process

$$dr = [\theta(t) + F(r)]\, dt + G(r) dZ_t \,, \tag{15}$$

for some functions θ, F and G. Then one changes to the new process

$$x = f(r) \equiv \int \frac{dr}{G(r)} \qquad \Longleftrightarrow \qquad \frac{df}{dr} = \frac{1}{G(r)} \tag{16}$$

which satisfies

$$dx = H(x,t)dt + dZ_t \,, \tag{17}$$

with

$$H(x,t) = \frac{\theta(t) + F(r)}{G(r)} - \frac{1}{2} G'(r) \tag{18}$$

(the prime denotes derivative with respect to r).

The first moment at time t_{i+1} can now be written as

$$m_1 = f(r_j + \delta r_j)\,. \tag{19}$$

One way of computing m_1 is to use the x-process in its finite version:

$$x_{j+1} = x_j + \int_{t_i}^{t_{i+1}} H(x_j, t_i) dt + (Z_{t_{i+1}} - Z_{t_i})\,. \tag{20}$$

Hence

$$m_1 \equiv \mathbb{E}\left(x_{j+1}|\mathcal{F}_t\right) = x_j + \int_{t_i}^{t_{i+1}} H(x_j, t_i)dt\,, \tag{21}$$

which is equal to

$$\begin{aligned} m_1 &= x_j + H(x_j, t_i)\Delta t \\ &= f(r_j) + \left.\frac{df}{dr}\right|_{r_j} G(r_j)H(x_j, t_i)\cdot\Delta t \\ &= f\left(r_j + G(r_j)H(x_j, t_i)\cdot\Delta t\right) \\ &= f\left(r_j + \left[\theta(t_i) + F(r_j) - \frac{1}{2}G(r_j)G'(r_j)\right]\cdot\Delta t\right)\,, \end{aligned} \tag{22}$$

for small Δt and the equality sign holds up to order $O\left((\Delta t)^2\right)$. Here we use the fact that the equation in the second step is of the same shape as the first order Taylor expansion if we consider $\frac{df}{dr}$ at the point r_j, since $f(r_j + \delta r_j) = f(r_j) + f'(r_j)\delta r_j + O(\delta r_j^2)$. By virtue of equation(14) for the variance we find

$$m_2 \equiv \mathbb{V}\left(x_{j+1}|\mathcal{F}_t\right) = m_1^2 + \Delta t. \tag{23}$$

Another way to compute m_1 is to start directly from δr_j and integrate the infinitesimal r-process:

$$\delta r_j = \int_{t_i}^{t_{i+1}} dr_t = \int_{t_i}^{t_{i+1}} \left[\theta(t) + F(r_t)\right]dt + \int_{t_i}^{t_{i+1}} G(r_t)dZ_t\,. \tag{24}$$

At order $O\left((\Delta t)^2\right)$, the first contribution is simply

$$\int_{t_i}^{t_{i+1}} \left[\theta(t) + F(r_t)\right]dt = \left[\theta(t_i) + F(r_{t_i})\right]\cdot\Delta t + O\left((\Delta t)^2\right)\,. \tag{25}$$

To compute the second contribution we have to do more work:

Lemma 3.1. *The following relation holds true:*

$$\int G(r_t)dZ_t = -\frac{1}{2}\int G(r_t)G'(r_t)dt\,. \tag{26}$$

Proof. Recall that

$$\frac{\partial r_t}{\partial Z_t} = G(r_t) \tag{27}$$

and hence

$$\frac{\partial^2 r_t}{\partial Z_t^2} = \frac{\partial G(r_t)}{\partial Z_t} = \frac{\partial G}{\partial r_t}\frac{\partial r_t}{\partial Z_t} = G'(r_t)G(r_t)\,. \tag{28}$$

Then, using Ito, we have:

$$\begin{aligned} \int G(r_t)dZ_t &= \int\left(\frac{\partial r_t}{\partial Z_t}\right)dZ_t \\ &= \int\left[dr_t - dt\frac{\partial r_t}{\partial t} - \frac{1}{2}\frac{\partial^2 r_t}{\partial Z_t^2}dZ_t^2\right] \\ &= \int dr_t - \int dt\frac{\partial r_t}{\partial t} - \frac{1}{2}\int\frac{\partial^2 r_t}{\partial Z_t^2}dt \\ &= -\frac{1}{2}\int G'(r_t)G(r_t)dt\,, \end{aligned} \tag{29}$$

which is what we wanted to show. □

Now, going back to eq. (24), we have

$$\delta r_j = \left[\theta(t_i) + F(r_j) - \frac{1}{2}G'(r_j)G(r_j)\right] \cdot \Delta t + O\left((\Delta t)^2\right) , \tag{30}$$

which is the same result that we have computed in (22).

3.2.2 Generalized set up

For our convenience let us recall the main definitions:

$$dr_t = F(r_t,t)dt + G(r_t,t)dZ_t \tag{31}$$

$$x_t = f(r_t,t) = \int \frac{dr_t}{G(r_t,t)} \qquad \Longleftrightarrow \qquad \frac{df}{dr} = \frac{1}{G(r_t,t)} \tag{32}$$

$$dx_t = H(x_t,t)dt + dZ_t \tag{33}$$

where the function $H(x_t,t)$ is:

$$H(x_t,t) = -\int^{r_t} \frac{\dot{G}(r_t,t)}{G^2(r_t,t)}dr + \frac{F(r_t,t)}{G(r_t,t)} - \frac{1}{2}G'(r_t,t) \,. \tag{34}$$

Here the first term with the integral comes from the partial derivative of f with respect to t. Moreover we use the notation:

$$\dot{G}(r,t) \equiv \frac{\partial G}{\partial t} \qquad \text{and} \qquad G'(r,t) \equiv \frac{\partial G}{\partial r} \,. \tag{35}$$

The first approach to compute the moment m_1 is to start from the integrated x-process:

$$x_{j+1} = x_j + H(x_j,t_i)\Delta t + (Z_{t_{i+1}} - Z_{t_i}) \tag{36}$$

and hence

$$\begin{aligned} m_1 &\equiv \mathbb{E}\left(x_{j+1}|\mathcal{F}_{t_i}\right) \\ &= x_j + H(x_j,t_i)\Delta t \\ &= f(r_j,t_i) + \left.\frac{\partial f}{\partial r}\right|_{(r_j,t_i)} \left(F(r_j,t_i) - \frac{1}{2}G(r_j,t_i)G'(r_j,t_i)\right)\Delta t + \left.\frac{\partial f}{\partial t}\right|_{(r_j,t_i)} \Delta t \\ &= f\left(r_j + \left[F(r_j,t_i) - \frac{1}{2}G(r_j,t_i)G'(r_j,t_i)\right]\Delta t, t_i + \Delta t\right) \\ &\equiv f(r_j + \delta r_j, t_{i+1}) \,, \end{aligned} \tag{37}$$

where equalities hold up to order $O\left((\Delta t)^2\right)$. Now the mean can be used to compute the variance to be

$$m_2 \equiv \mathbb{V}\left(x_{j+1}|\mathcal{F}_t\right) = m_1^2 + \Delta t. \tag{38}$$

The alternative approach is to start from the integrated r-process:

$$\delta r_j = \int_{t_i}^{t_{i+1}} dr_t = \int_{t_i}^{t_{i+1}} \left[F(r_t,t)dt + G(r_t,t)dZ_t\right] . \tag{39}$$

The first term is straightforward while for the second we use again Lemma 3.1, which readily extends also to this case where the function G has an explicit time dependence. The final answer is

$$\delta r_j = \left[F(r_j, t_i) - \frac{1}{2}G(r_j, t_i)G'(r_j, t_i)\right]\Delta t\,, \tag{40}$$

which is again true up to order $O\left((\Delta t)^2\right)$. This is the same as (37).

3.3 Lattice Construction

It follows from equation (10) that given the process dx one can retrieve any short rate model that has the shape of (1) for given F and G via transformation functions. In this chapter we will show how the process for dx can be constructed using a lattice approach. There are a few of alternatives to produce the lattice, for example Daglish [5] or Vetzal [6]. However, we will follow the framework provided by Hull and White here ([2, 7]). The steps are reproduced to make this paper a complete review of the methodology. We will consider a fixed grid for x. We will match the market by adjusting the drift of r. The time step Δt, can vary at each step.

3.3.1 Constructing the Lattice

To construct the lattice we must define an algorithm to compute the x_j, j_d and j_u as in section 3.2. We begin with x by taking the step size:

$$\Delta x_i = \sqrt{3\Delta t_i} \tag{41}$$

The step size Δx_i can be chosen between $\frac{\sqrt{3\Delta t_i}}{2}$ and $\sqrt{3\Delta t_i}$ as pointed out in [8]. Consider $x_0 = f(r_0)$, with r_0 being the rate corresponding to the initial time step. At a given time $i\Delta t_i$ at level j in the lattice node (i, j) we have

$$x = x_0 + j\Delta x_i \tag{42}$$

Note that the value of x is not directly dependent on the time step, but only on the level. The time step element is introduced indirectly via $\hat{j}$, which indicates the level of the nodes at $i + 1$. We have

$$(i, j) \to \left\{(i+1, \hat{j}+1), (i+1, \hat{j}), (i+1, \hat{j}-1)\right\},$$

with

$$\hat{j} = \mathrm{int}\left(\frac{m_1 - x_0}{\Delta x_i} + \frac{1}{2}\right), \tag{43}$$

where m_1 is time dependent. Here $\hat{j}$ is used to indicate the nodes that can be reached in the next time step. Consider figure 1 for the most typical branching that can be encountered. A general overview of all branching options is given in [2].

Consider the mean and variance of the lattice

$$m_1 = p_u\left[(\hat{j}+1)\Delta x_i + x_0\right] + p_m\left[\hat{j}\Delta x_i + x_0\right] + p_d\left[(\hat{j}-1)\Delta x_i + x_0\right] \tag{44}$$

$$m_2 = p_u\left[(\hat{j}+1)\Delta x_i + x_0\right]^2 + p_m\left[\hat{j}\Delta x_i + x_0\right]^2 + p_d\left[(\hat{j}-1)\Delta x_i + x_0\right]^2 - m_1^2 \tag{45}$$

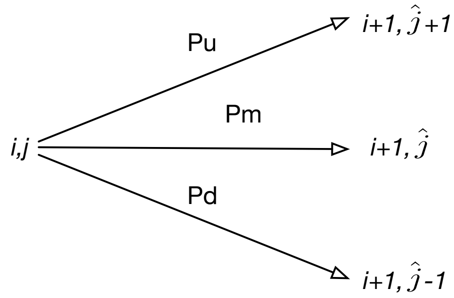

Figure 1: Branching process

and the fact that

$$p_m = 1 - p_u - p_d\,. \tag{46}$$

By inverting these relations we can derive the expressions for p_u,p_m and p_d. They are:

$$p_u = \frac{\Delta t + (m_1 - x_0 - \hat{j}\Delta x_i)^2}{2\Delta x^2} + \frac{m_1 - x_0 - \hat{j}\Delta x_i}{2\Delta x_i} \tag{47}$$

$$p_d = \frac{\Delta t + (m_1 - x_0 - \hat{j}\Delta x_i)^2}{2\Delta x_i^2} - \frac{m_1 - x_0 - \hat{j}\Delta x_i}{2\Delta x_i} \tag{48}$$

$$p_m = 1 - p_u - p_d \tag{49}$$

These probabilities look the same as in the Hull White framework. However, the formulas contain the mean and variance of the distribution (equation (44) and (45)) and those differ for the Hull White setup. In the Hull White setup the first moment m_1 is function of r_j, which is an explicit function of the level j of lattice and an implicit function of time t_i. For the general case we have $m_1(r_j, t_i)$, so that the i and j dependence is explicit. The fact that the probabilities look the same in both frameworks is logical as in both case they are taken to match mean and variance of process x using a lattice approach.

Using equations (47), (48) and (49), we can work out the the highest and lowest level that the tree reaches at a given time step i. These quantities are simply a matter of "book keeping". For the generalized model (1) it still holds that all probabilities are positive since $\Delta x = \sqrt{3\Delta t}$ and $m_1 - x_0 - \hat{j}\Delta x_i \leq \frac{1}{2}\Delta x_i$ as pointed out in [1].

For a given time step, the Arrow Debreu prices (as described in [9]) can be computed as

$$Q_{i,j} = \sum_{k=j_d(i-1)}^{j_u(i-1)} Q_{i-1,k}\, q(k,j) \exp\left(-r_k \Delta t_i\right) \tag{50}$$

where $Q_{0,0} = 1$ and the probability of moving from node $(i-1, k)$ to (i, j) is denoted by $q(k, j)$. This results in a discount at time $i + 1$ of

$$D_{i+1} = \sum_{k=j_d(i)}^{j_u(i)} Q_{i,j} \exp\left(-r_k \Delta t_i\right) \tag{51}$$

To price back the market we must find the Arrow Debreu prices so that the discounts (51) match those found in the market. This matching is done by adjusting the mean of short rate, respectively $\theta(t)$ in the Hull White set up (equation (15)) and $F(r_t, t)$ in the generalized set up (equation (31)). Calibration methods are described in general in [2, 3, 7, 10, 11]. More details on calibration of the Hull-White set up can be found in [1]. Here we will consider the same approach:

1. match discounts by adjusting drift $F(r_t, t)$
2. find $G(r_t, t)$ that minimizes differences in option prices quoted in the market.

4 Some Examples

In this chapter we will recover some well known models to demonstrate the methodology.

4.1 Black Karasinski

We consider the Black Karasinski short rate model as presented in their paper [12] formula (1)

$$d(\ln r) = [b(t) - a \ln r]\, dt + \sigma dZ(t)\,, \tag{52}$$

where we restrict ourselves to the case where a and σ are constant. To bring this into the form of eq. (1) we apply the coordinate transformation:

$$y = \ln r \quad \text{and thus} \quad \exp(y) = r\,.$$

Using Ito we find

$$dr = d\exp(y) = \left\{ [b(t) - a \ln r] \exp(y) + \frac{1}{2}\sigma^2 \exp(y) \right\} dt + \sigma \exp(y) dZ(t)\,. \tag{53}$$

This can by rewritten as

$$dr = r\left[b(t) - a \ln r + \frac{1}{2}\sigma^2 \right] dt + \sigma r dZ(t) \tag{54}$$

The transformed function does fit (1) with

$$F(r,t) = r\left[b(t) - a \ln r + \frac{1}{2}\sigma^2 \right]$$

$$G(r,t) = r\sigma$$

resulting in

$$H(x) = \frac{r\left[b(t) - a\ln r + \frac{1}{2}\sigma^2\right]}{r\sigma} - \frac{1}{2}\sigma\,.$$

Finally we find the transformation function to be

$$x = f(r) = \int_c^r \frac{1}{G(r',t)} dr' = \frac{\ln r}{\sigma}, \text{ where c=1}$$

and

$$r = f^{-1}(x) = \exp(x\sigma).$$

<u>Check A</u>

Based on the process of dx and transformation funtion f^{-1} we can recover the original process for dr. Using Ito we find

$$dr = \left[\left\{ \frac{r\left[b(t) - a\ln r + \frac{1}{2}\sigma^2\right]}{r\sigma} - \frac{1}{2}\sigma \right\} \sigma r + \frac{1}{2}\sigma r \right] dt + \sigma r dZ(t)\,. \tag{55}$$

By reorganizing this equation we find

$$dr = r\left[b(t) - a\ln r + \frac{1}{2}\sigma^2\right] dt + \sigma r dZ(t)\,,$$

which corresponds to the transformed Black Karasinski model that fitted our generalized formula (1).

Check B

In the previous check A we recover the transformed Black Karasinski equation, but actual we would like to recover the original equation as published. To do this, define:

$$\hat{f}(x) \equiv \ln r = x\sigma\,. \tag{56}$$

Then

$$d\ln r = \left\{ \frac{r\left[b(t) - a\ln r + \frac{1}{2}\sigma^2\right]}{r\sigma} - \frac{1}{2}\sigma \right\} \sigma dt + \sigma dZ(t)\,, \tag{57}$$

which can be written as:

$$d\ln r = [b(t) - a\ln r]\, dt + \sigma dZ(t)\,.$$

We have now recovered the Black Karasinski equation as noted in (52).

4.2 Piece-Wise Model

The piece-wise model shows how the generalization can be used to create a regime shifting model. Here a combination of Hull White, Squared Gaussian and Black Karasinski are used. The Black Karasinski model is an elegant solution for dealing with high and negative rates. The Hull White model gives well understood distribution of rates for the middle of the lattice. To avoid too abrupt transitions (in terms of distribution) between Black Karasinski and Hull White the Gaussian model is used.

The model is defined as follows:

$$dr_t = \mu(r_t, t)dt + \sigma h(r_t)dZ_t\,, \tag{58}$$

where $\mu(r_t, t)$ is the drift and

$$h(r) = \begin{cases} c_0 r & \text{if } r < r_0 \\ c_1 1 & \text{if } r_0 \leq r < r_1 \\ c_2\sqrt{r} & \text{if } r_1 < r < r_2 \\ c_3 r & \text{if } r_2 \leq r \end{cases} \tag{59}$$

In principle the volatility can be time-dependent, but we will consider it constant for simplicity. To make $h(r)$ continuous we set $c_3 = 1$ and find $c_2 = \sqrt{r_2}$, $c_1 = \sqrt{r_1 r_2}$ and

$c_0 = \frac{\sqrt{r_1 r_2}}{r_0}$:

$$h(r) = \begin{cases} \frac{\sqrt{r_1 r_2}}{r_0} r & \text{if } r < r_0 \\ \sqrt{r_1 r_2} & \text{if } r_0 \le r < r_1 \\ \sqrt{r_2}\sqrt{r} & \text{if } r_1 \le r < r_2 \\ r & \text{if } r_2 \le r \end{cases} \tag{60}$$

The x-process is obtained as

$$x_t = f(r_t) = \int_c^{r_t} \frac{dr'}{G(r', t)} = \frac{1}{\sigma} \int_c^{r_t} \frac{dr'}{h(r')} \equiv \frac{I(r_t)}{\sigma} \qquad \text{(such that } f(c) = 0\text{)}\,. \tag{61}$$

We use $c = 1$ so that there are no negative rates. We solve the integral

$$I(r) \equiv \int_1^r \frac{dr'}{h(r')} \tag{62}$$

in 4 parts:

$$I_0(r) = \int_1^r \frac{r_0}{\sqrt{r_1 r_2}} \frac{1}{r'} dr' = \frac{r+0}{\sqrt{r_1 r_2}} \ln(r')\Big|_1^r = \frac{r_0}{\sqrt{r_1 r_2}} \ln(r) \tag{63}$$

$$\begin{aligned} I_1(r) &= \frac{r_0}{\sqrt{r_1 r_2}} \ln(r_0) + \int_{r_0}^r \frac{r_0}{\sqrt{r_1 r_2}} dr' \\ &= \frac{r_0}{\sqrt{r_1 r_2}} \ln(r_0) + \frac{r'}{\sqrt{r_1 r_2}}\Big|_{r_0}^r \\ &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) + \frac{r}{\sqrt{r_1 r_2}} \end{aligned} \tag{64}$$

$$\begin{aligned} I_2(r) &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) + \frac{r_1}{\sqrt{r_1 r_2}} + \int_{r_1}^r \frac{1}{\sqrt{r_2}\sqrt{r'}} dr' \\ &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) + \frac{r_1}{\sqrt{r_1 r_2}} + 2\frac{\sqrt{r'}}{\sqrt{r_2}}\Big|_{r_1}^r \\ &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2\frac{\sqrt{r}}{\sqrt{r_2}} \end{aligned} \tag{65}$$

$$\begin{aligned} I_3(r) &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2 + \int_{r_2}^r \frac{1}{r'} dr' \\ &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2 + \ln(r')|_{r_2}^r \\ &= \frac{r_0}{\sqrt{r_1 r_2}} (\ln(r_0) - 1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2 + \ln\left(\frac{r}{r_2}\right) \end{aligned} \tag{66}$$

so that

$$I(r) = \begin{cases} I_0(r) & \text{if } r < r_0 \\ I_0(r) + I_1(r) & \text{if } r_0 \le r < r_1 \\ I_0(r) + I_1(r) + I_2(r) & \text{if } r_1 \le r < r_2 \\ I_0(r) + I_1(r) + I_2(r) + I_3(r) & \text{if } r_2 \le r \end{cases} \tag{67}$$

At the boundaries r_1, r_2 and r_3 we find:

$$x_1 \equiv \frac{f(r_1)}{\sigma} = \frac{I_1(r_1)}{\sigma} = \frac{1}{\sigma}\left(\frac{r_0}{\sqrt{r_1 r_2}}\ln(r_0)\right)$$

$$x_2 \equiv \frac{f(r_2)}{\sigma} = \frac{I_2(r_2)}{\sigma} = \frac{1}{\sigma}\left(\frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) + \frac{\sqrt{r_1}}{\sqrt{r_2}}\right)$$

$$x_3 \equiv \frac{f(r_3)}{\sigma} = \frac{I_3(r_3)}{\sigma} = \frac{1}{\sigma}\left(\frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2\right)$$

resulting in

$$f(r) = \frac{1}{\sigma}\begin{cases} \frac{r_0}{\sqrt{r_1 r_2}}\ln(r) & \text{if } r < r_0 \\ \frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) + \frac{r}{\sqrt{r_1 r_2}} & \text{if } r_0 \leq r < r_1 \\ \frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2\frac{\sqrt{r}}{\sqrt{r_2}} & \text{if } r_1 \leq r < r_2 \\ \frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) - \frac{\sqrt{r_1}}{\sqrt{r_2}} + 2 + \ln(\frac{r}{r_2}) & \text{if } r_2 \leq r \end{cases} \tag{68}$$

and

$$r = f^{-1}(x) = \begin{cases} \exp\left(\frac{\sqrt{r_1 r_2}}{r_0}\sigma x\right) & \text{if } x < x_1 \\ \sqrt{r_1 r_2}\sigma x - r_0(\ln(r_0)-1)) & \text{else if } x < x_2 \\ \left(\frac{\sqrt{r_2}\sigma x + \sqrt{r_1} - \frac{r_0}{\sqrt{r_1}}(\ln(r_0)-1)}{2}\right)^2 & \text{else if } x < x_3 \\ r_2\exp\left(\sigma x - \frac{r_0}{\sqrt{r_1 r_2}}(\ln(r_0)-1) + \frac{\sqrt{r_1}}{\sqrt{r_2}} - 2\right) & \text{else} \end{cases} \tag{69}$$

Using eq. (34), the x-process satisfies

$$dx_t = H(x_t,t)dt + dZ_t \tag{70}$$

with

$$\begin{aligned} H(x_t,t) &= \frac{F(r_t,t)}{G(r_t,t)} - \frac{1}{2}G'(r_t,t) \\ &= \frac{\mu(r_t,t)}{\sigma h(r)} - \frac{1}{2}h'(r)\sigma\,. \end{aligned} \tag{71}$$

4.3 The Modified Squared Damped Harmonic Oscillator

In this subsection we consider an example where the volatility depends explicitly on the time. In order to find a proper candidate for the volatility, we assume the following assumptions:

- it may have an oscillating behaviour
- on average it decreases in time

Later we will also need to fact that the volatility is always positive. In the following, we will construct a volatility function that satisfies such assumptions.

One system that almost satisfies such assumptions is the damped harmonic oscillator. The damped harmonic oscillator belongs to a wider class of systems that can all in principle be used to pick different drift and diffusion terms in the stochastic differential equation. In Appendix A we review the main properties of one-dimensional non-linear dynamical systems and in particular in Theorem A.2 we recall an interesting result valid for monotonic maps.

The damped harmonic oscillator is solution to the motion equation

$$\ddot{x} + \nu \dot{x} + \omega_0^2 x = 0\,. \tag{72}$$

In principle, there might also be a driven force on the r..h.s., which produces very interesting effects, such as resonances etc., but we will not consider it here. The solution to such equation is of the form

$$x(t) \propto \exp(\lambda t)\,, \tag{73}$$

with λ satisfying the constraint $\lambda^2 + \nu\lambda + \omega_0^2 = 0$, which implies

$$\lambda = -\frac{\nu}{2} \pm \sqrt{\frac{\nu^2}{4} - \omega_0^2}\,. \tag{74}$$

In the case of complex-conjugate exponents

$$\lambda = -\frac{\nu}{2} \pm i\omega\,, \qquad \omega^2 \equiv \omega_0^2 - \frac{\nu^2}{4} > 0\,, \tag{75}$$

the solution can be written as

$$x(t) = a e^{-\frac{\nu}{2}t} \cos(\omega t + \phi)\,. \tag{76}$$

Here ϕ is the initial phase and a is the amplitude of the oscillator. This is the damped harmonic oscillator.

In order to construct a sensible volatility function out of the damped harmonic oscillator, we need to make a few adjustments. First,the positivity constraint suggests us to consider the squared function $y(t) \equiv x^2(t)$. Secondly, regularity of the H-function implies that the G-function must be non vanishing everywhere. This suggest to shift the cosine by a strictly-positive amount b^2 for some non-zero value of b, $z(t) = a^2 e^{-\nu t}\left[\cos^2(\omega t + \phi) + b^2\right]$.

In order to construct our volatility, we notice that the constant parameters can in principle be functions of the random variable r. Hence, we will consider the following choice for the G-function:

$$G(r,t) = a(r)^2 e^{-\nu t}\left[\cos^2(\omega t + \phi) + b^2(r)\right]\,, \tag{77}$$

where the functions $a(r)$ and $b(r)$ are required not to vanish for any value of r. The derivative w.r.t. r is given by

$$G'(r,t) = 2a(r)a'(r)e^{-\nu t}\left[\cos^2(\omega t + \phi) + b^2(r)\right] + 2b(r)b'(r)a^2(r)e^{-\nu t}\,, \tag{78}$$

while the derivative w.r.t. t is

$$\dot{G}(r,t) = -\nu G(r,t) - a^2(r)e^{-\nu t}\omega \sin(2(\omega t + \phi))\,. \tag{79}$$

These expressions can be used now to evaluate the H-function as

$$H(x,t) = -\int \frac{\dot{G}(r,t)}{G^2(r,t)} dr + \frac{F(r,t)}{G(r,t)} - \frac{1}{2}G'(r,t)\,. \tag{80}$$

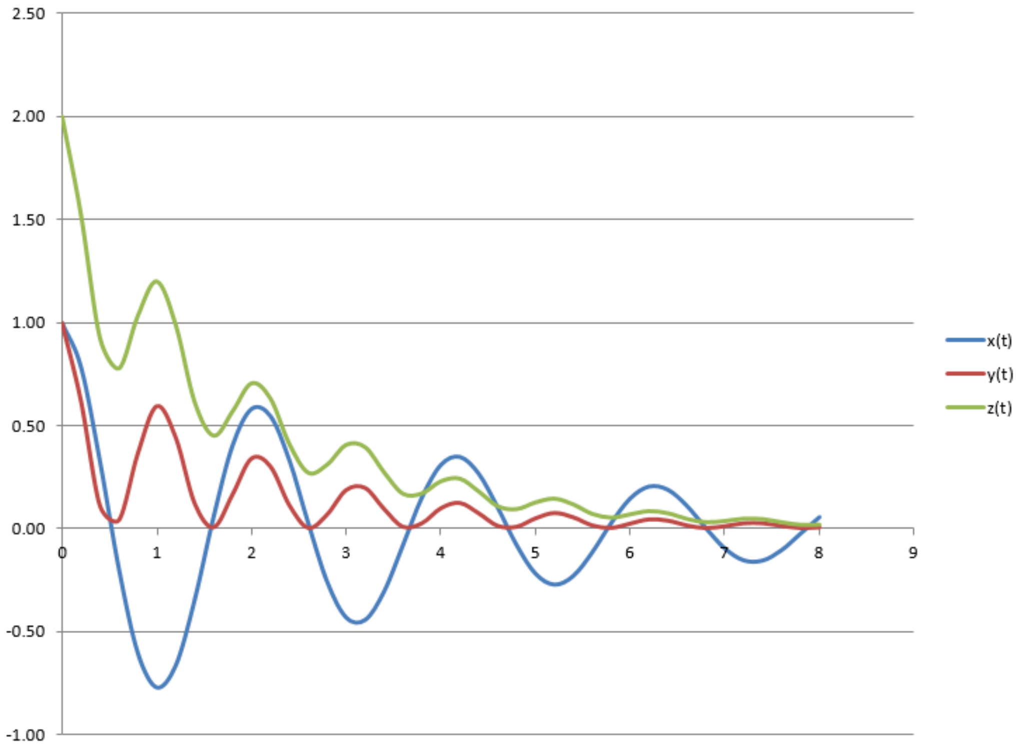

Figure 2: This plot shows the standard damped harmonic oscillator $x(t)$, the squared damped harmonic oscillator $y(t)$ and the modified squared damped harmonic oscillator $z(t)$. The parameters are: $a = 1$, $b = 1$, $\nu = 0.5$, $\omega = 3$, $\phi = 0$.

Observe that the convergence of the integral in the previous formula for $r \to +\infty$ depends on the convergence of the two integrals:

$$\int_{r_0}^{r} dr \frac{1}{a^2(r)\left(\cos^2(\omega t + \phi) + b^2(r)\right)} \qquad \text{and} \qquad \int_{r_0}^{r} dr \frac{1}{a^2(r)\left(\cos^2(\omega t + \phi) + b^2(r)\right)^2} . \tag{81}$$

We can distinguish two cases, depending on the function $b(r)$:

- $b(r) \overset{r \to +\infty}{\longrightarrow} +\infty$:
  in this case the relevant integrals are

$$\int_{r_0}^{r} dr \frac{1}{a^2(r) b^2(r)} \qquad \text{and} \qquad \int_{r_0}^{r} dr \frac{1}{a^2(r) b^4(r)} ; \tag{82}$$

- $b(r) \overset{r \to +\infty}{\longrightarrow}$ constant
  (the constant limit can also be arbitrarily small, as long as it is non-zero otherwise the cosine will periodically vanish): in this case the only relevant integral is

$$\int_{r_0}^{r} dr \frac{1}{a^2(r)} . \tag{83}$$

So, if convergence is an issue, these expressions fix constraints on the $a(r)$ and $b(r)$ functions. These constrains are the following for the two cases:

- $b(r) \overset{r\to+\infty}{\longrightarrow} +\infty$:
  when $r \to +\infty$, the functions a and b should behave as

$$a(r) \sim r^{\alpha-\frac{\beta}{2}} \qquad \text{and} \qquad b(r) \sim r^{\frac{\beta-\alpha}{2}} \tag{84}$$

  with both $\alpha, \beta > 1$.

- $b(r) \overset{r\to+\infty}{\longrightarrow}$ constant
  when $r \to +\infty$, the functions a should behave as

$$a(r) \sim r^{\frac{\alpha}{2}} \tag{85}$$

  with $\alpha > 1$, and no additional constraint on $b(r)$.

5 Summary and Conclusions

In this paper we have considered a generalization of the Hull-White approach to tree building in the case where both the drift and the diffusion terms in a stochastic differential equation are time dependent.

The framework presented here is general in the sense of the functional form of short rate models that can be used. However, it only allows for one factor models. In principal multi-factor short rates models seem a feasible extension. Such an extension would pose interesting further research.

One interesting application of short rate models is its embedded use in ALM[2] models. To integrate short rate models with the rest of the VAR model used for analysis one would ideally like to have both risk neutral and real world interest rate risk scenario's. The methodology to realize has been described for specific models using the market prices of risk (for example [13]).

Future research could deliver a method to acquire and estimate the market price of risk for the generalized model presented here.

In this paper we have looked at a very specific example of time-dependent volatility function, namely the MSDHO. We have also pointed out in the main body of the paper (see also section A) that the damped harmonic oscillator is one of the situations that can happen with non-linear dynamical systems. It would be interesting to address all the other types of possibility in the future and their numerical implementations.

Finally, it would be useful to perform a numerical calculation of the methods described above. When time-dependent drift and volatilities are considered, numerical calculation might turn out to be trick to perform due to initial conditions and root-finding issues when pricing back the market. We hope to report on this front soon in the future.

6 Acknowledgements

We could have not written this paper without all the wonderful support and feedback we received for our fellow quants at ABN AMRO.

[2]Assets and Liability Management

During our research we have had fruitful email exchanges with John Hull and Alan White on their recent paper [1]. Ideas sparked by this exchanged have been very valuable for this paper.

References

[1] John Hull and Alan White. A generalized procedure for building trees for the short rate and its application to determining market implied volatility functions. *Working Paper No. 2399615*, 2014.
Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2399615.

[2] John Hull and Alan White. Numerical procedures for implementing term structure models i - single-factor models. *The Journal of Derivatives*, 2(01):7–16, 1994.

[3] John Hull and Alan White. Using hull-white interest-rate trees. *The Journal of Derivatives*, 3:26–36, 1996.

[4] George E. Uhlenbeck and Leonard Salomon Ornstein. On the theory of the brownian motion. *Physical review*, page 823, 1930.

[5] Toby Daglish. Lattice methods for no-arbitrage pricing of interest rate securities. *Journal of Derivatives*, 18(2):7, 2010.

[6] Kenneth R Vetzal. An improved finite difference approach to fitting the initial term structure. *The Journal of Fixed Income*, 7(4):62–81, 1998.

[7] John Hull and Alan White. Numerical procedures for implementing term structure models ii - two-factor models. *The Journal of Derivatives*, 2(2):37–48, 1994.

[8] John Hull and Alan White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis*, 25(01):87–100, 1990.

[9] Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954. http://web.stanford.edu/class/msande311/arrow-debreu.pdf/.

[10] John Hull and Alan White. The general hull-white model and super calibration. *Working Paper*, FIN-00(024), 2000. http://ssrn.com/abstract=1295228.

[11] Sebastian Gurrieri, Masaki Nakabayashi, and Tony Wong. Calibration methods of hull-white model. *ssrn.com*, 2009. http://ssrn.com/abstract=1514192.

[12] Fischer Black and Piotr Karasinski. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, pages 52–59, 1991.

[13] GA Medvedev and SH Cox. The market price of risk for affine interest rate term structures. *Procedings 6th - Intern. AFIR Symposium Nuremberg*, 1996.
Available at http://www.elib.bsu.by/bitstream/123456789/8800/1/1996_Medvedev_Cox_Nuremberg.pdf.

[14] Hommes. Nonlinear economic dynamics. Technical report, University of Amsterdam, 2003.

A One dimensional maps

In this section we will put the damped harmonic oscillator into the larger context of non-linear dynamical systems. We will restrict ourselves to the one-dimensional case. The main reference is [14].

A one-dimensional non-linear dynamical system is defined by a map f_λ

$$x_{t+1} = f_\lambda(x_t)\,, \tag{86}$$

where λ is a parameter. To ease the notation, if not necessary, we will not write the parameter subscript explicitly. Given an initial value x_0, the orbits of the system are given by

$$\{x_0, x_1, x_2, \dots\} = \{x_0, f(x_0), f^2(x_0), \dots\}\,. \tag{87}$$

One can define a *fixed point* or *steady state* or *equilibrium state* x^{eq} as the point that satisfies

$$f(x^{eq}) = x^{eq}\,. \tag{88}$$

A fixed point is *locally* stable if changing by a little bit the initial condition x_0 in a small interval of x^{eq} the resulting orbit converges to $f(x^{eq})$:

$$\lim_{n\to+\infty} f^n(x_0) = x^{eq} \qquad \text{with} \qquad x_0 \in (x^{eq} - \varepsilon, x^{eq} + \varepsilon)\,. \tag{89}$$

A fixed point is *globally* stable if the limit holds for any x_0.

A fixed point is unstable if changing by a little bit the initial condition x_0 in a small interval of x^{eq} the resulting orbit diverges away from x^{eq}:

$$|f^n(x_0) - x^{eq}| > \varepsilon \qquad \text{for some } n > 0\,. \tag{90}$$

The following theorem is true:

Theorem A.1. *If x^{eq} is a fixed point of the dynamical system $x_{t+1} = f(x_t)$, then*

- *if $|f'(x^{eq})| < 1$, then x^{eq} is locally stable*
- *if $|f'(x^{eq})| > 1$, then x^{eq} is locally unstable*

Here the unit value on the r.h.s. is the slope of the function $f(x) = x$. The four typical cases are the following:

- $0 < f'(x^{eq}) < 1$: stable and monotonic convergence
- $f'(x^{eq}) > 1$: unstable and monotonic divergence
- $-1 < f'(x^{eq}) < 0$: stable and damped oscillations
- $-1 < f'(x^{eq}) < 0$: unstable and undamped oscillations

One can also define a *periodic* point x with period k as a fixed point of the map f^k. The orbits of a periodic point have exactly k elements, $\{x, f(x), f^2(x), \dots, f^{k-1}(x)\}$, since $f^k(x) = x$. This is called a *periodic orbit* or *k-cycle*.

For *monotonic* maps, the following theorem is true:

Theorem A.2. *If f is monotonic, then the only possibilities for x_t are:*

- *x_t converges to a fixed point*
- *x_t converges to a 2-cycle*
- *x_t diverges to $\pm\infty$ or exhibits unbounded oscillations.*

Hence, for monotonic maps, the dynamics is simple.

For non-monotonic maps, the dynamics can be very complicated. As an example, let us consider the quadratic difference equation

$$x_{t+1} = f_\lambda(x_t) = \lambda x_t(1 - x_t)\,, \tag{91}$$

The two steady states are $x = 0$ and $x^* = 1 - 1/\lambda$. Stability depends on whether the derivative of f_λ

$$f'_\lambda = \lambda(1 - 2x) \tag{92}$$

evaluated at zero and x^* are between ± 1. In particular, one can show that

Theorem A.3. *For the quadratic difference equation (91) with initial condition $x_0 \in [0, 1]$, we have:*

1. *for $0 \leq \lambda \leq 1$, $x = 0$ is the unique steady state and is globally stable*
2. *for $\lambda > 1$, the two states $x = 0$ and $x^* = 1 - 1/\lambda$ are both fixed points. Moreover, the state $x = 0$ is unstable. If $\lambda \leq 3$, the state x^* is stable and attracts all time paths.*

Another typical effect that can be seen from the quadratic difference map is the *sensitivity dependence to the initial conditions.* Slightly different initial conditions can have dramatic effects on the dynamics of the system. We will not go into the details here.