



## Efficient Calibration of Trinomial Trees for One-Factor Short Rate Models

MARKUS LEIPPOLD\*

*Swiss Banking Institute, University of Zurich, Switzerland*

leippold@isb.unizh.ch

ZVI WIENER

*School of Business Administration, The Hebrew University of Jerusalem, Israel and  
Marshall School of Business, University of Southern California, Los Angeles, CA, USA*

**Abstract.** In this paper we propose a computationally efficient implementation of general one factor short rate models with a trinomial tree. We improve the Hull–White’s procedure to calibrate the tree to bond prices by circumventing the forward rate induction and numerical root search algorithms. Our calibration procedure is based on forward measure changes and is as general as the Hull–White procedure, but it offers a more efficient and flexible method of constructing a trinomial term structure model. It can be easily implemented and calibrated to both prices and volatilities.

**Keywords:** short rate models, trinomial trees, forward measure.

**JEL classification:** G13, C6

In this paper, we elaborate on the implementation and calibration of one-factor short rate models. Taking the Hull–White model as a starting point, we span trinomial trees for the short rate using a set of general branching processes. Generalizing the branching processes not only allows to avoid negative interest rates (see Hull and White (1994)), but also allows to obtain a “slender” tree at the edges. This can substantially reduce the computational time. Moreover, the additional flexibility in defining branching processes becomes important when pricing certain types of exotic options. For example, for barrier options a finer grid around the barrier helps to increase the convergence of the numerical tree method. Recent contributions in this direction are, e.g., Tian (1999), Figlewski and Gao (1999), and Baule and Wilkens (2004).

Our main contribution is a computational improvement of the procedure of Hull and White (1994) to calibrate the tree to bond prices. In the Hull and White (1994) model, when pricing an interest rate derivative, one has first to perform a forward induction to match the tree, and secondly one has to do backward induction to price the instrument. The forward induction can become computationally intensive. Furthermore, for non-Gaussian models, forward induction has to be combined with numerical root search algorithms. Using the approach presented in this paper, we are only left with backward induction. No forward induction and no root-search algorithms are necessary.

\* Corresponding author. University of Zurich, ISB, Plattenstr. 14, 8032 Zurich, Switzerland.

Our matching procedure is based on the representation of the short rate under the forward measure. This allows us to substantially reduce the computational costs. Moreover, our approach is not restricted to Gaussian interest rate models, but it is applicable to a wide range of short rate models. It is most efficient for all models which allow a closed-form expectation of the interest rate either under the risk-neutral or the forward measure. However, even if such a closed form model does not exist, our calibration is applicable and still leads to efficiency gains. Therefore, we claim that our calibration procedure is as general as the forward induction method used by Hull and White.

Finally, we show how the tree can be adjusted to the volatility structure in such a manner that our approach to match the initial term structure is still applicable. This is done by exploiting the flexibility embedded in the trinomial tree model.

Our paper is structured as follows. Section 1 presents the basic setup and introduces the notation. Furthermore, this section gives a brief overview of term structure models, which can be handled within our approach. Section 2 elaborates on the modelling of the state variable process within a trinomial tree. Section 3 presents our calibration procedures. Section 4 provides some explicit examples and comparisons with the Hull–White model. Section 5 concludes.

## 1. Model Setup

We assume that the market is complete and arbitrage opportunities are absent. Then, there exists a unique risk-neutral pricing measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  such that every claim discounted by the money market account is a martingale under  $\mathbb{P}$ . The state space is described by the one-dimensional process  $X_t$  with  $X_0 = x$ . We assume that  $X_t$  follows a time-homogeneous Markov process. Interest rates and bond prices can be expressed as functions of  $X_t$ , i.e., we assume  $P(X_t, t, T)$  to be the time- $t$  pricing functional for a zero bond in state  $X_t$ , which pays \$1 at the maturity date  $T$ . Then,

$$P(X_t, t, T) = \mathbb{E} \left[ e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right],$$

where  $r_t$  is the short-term interest rate and  $\mathbb{E}[\cdot \mid \mathcal{F}_t]$  is the expectation operator under  $\mathbb{P}$ . We further require the zero bond to satisfy the conditions  $P(X_t, t, t) = 1$  and  $\lim_{T \rightarrow \infty} P(X_t, t, T) = 0$  for all  $X_t$  and  $t$ . To abbreviate notation, bond prices observed from the initial term structure,  $P(x, 0, T)$ , are denoted by  $P^*(T)$ . Given the above setting, specifying a short rate model can be done in two steps.

1. Defining a state variable process  $X_t$ . Here, we make the basic assumption that  $X$  follows an Itô diffusion,

$$dX_t = \mu_X(X_t) dt + \sigma_X(X_t) dW_t, \quad X_0 = x, \quad (1)$$

under the measure  $\mathbb{P}$ .

2. Defining the functional form  $g(X, t) = r_t$ , which translates the state variable to the interest rate. Under the appropriate technical conditions, we can then write the general dynamics for the short rate as

$$dr_t = \mathcal{L}g(X, t) dt + \frac{\partial g(X, t)}{\partial X} \sigma_X(X_t) dW_t,$$

where  $\mathcal{L}$  is the extended generator.<sup>1</sup>

Thus, depending on the choices  $\mu_X$ ,  $\sigma_X$  and  $g(X, t)$ , we can construct a wide class of short rate models. However, since the tree construction starts with spanning a trinomial tree for  $X_t$  by matching its moments, a simple choice for the process  $X_t$  is appropriate. Therefore, when calculating examples in Section 3, we will focus on a Gaussian specification for  $X_t$ .

### 1.1. Affine Function $g(X)$

To illustrate the generality of the above short rate specification, we briefly discuss a few possible models. We start with affine term structure models. These are obtained by setting  $r_t = a + bX_t$ . If  $X_t$  follows a mean-reverting Ornstein–Uhlenbeck process with  $\mu_X(X_t) = -\kappa X_t$  and  $\sigma_X(X_t) = \sigma$ , then

$$dr_t = (a\kappa - \kappa r_t) dt + b\sigma dW_t, \quad (2)$$

which corresponds to the short rate model of Vasicek (1977). With time-dependent parameters Equation (2) is often referred to as the extended Vasicek model of Hull and White (1994). When  $\mu_X(X_t) = \theta - \kappa X_t$  and  $\sigma_X(X_t) = \sigma\sqrt{X_t}$ , and  $b > 0$ , we get

$$dr_t = (\bar{\theta} - \kappa r_t) dt + \sigma\sqrt{b(r_t - a)} dW_t, \quad (3)$$

with  $\bar{\theta} = b\theta + a\kappa$ . Setting  $a = 0$  we obtain the classical Cox, Ingersoll, and Ross (1985) term structure model. A lognormal model would be obtained by setting  $\sigma_X(X_t) = \sigma X_t$ .

### 1.2. Quadratic Function $g(X)$

Suppose now that we can represent the short rate as a quadratic function of the state variable, i.e., we set

$$r_t = a + bX_t + cX_t^2. \quad (4)$$

Such a specification defines a short rate model which belongs to the class of quadratic models.<sup>2</sup> Recently, Filipović (2002) proves, under certain regularity conditions, that if one represents the forward rate as a polynomial function of the diffusion state vector, then the maximal order of the polynomial is two for the model to be consistent. Consistency in this context, as discussed in Björk and Christensen (1999) and Filipović (1999), means that the interest rate model will produce forward rate curves belonging to the parameterized family. These findings stress the potentially important role played by quadratic models.

The functional specification in Equation (4) not only extends the affine class, but also embeds some well-known nonlinear term structure models. To see this, we assume

$\mu_X(X) = -\kappa$ ,  $\sigma_X(X) = \sigma$ , and derive the interest rate dynamics implied by (4) in its general form as

$$dr_t = \left( c\sigma^2 - \frac{\kappa}{2c}(b^2 - 4ac + 4cr_t) + \sigma \frac{b\kappa}{2c} \sqrt{b^2 - 4ac + 4cr_t} \right) dt + \sigma \sqrt{b^2 - 4ac + 4cr_t} dW_t.$$

By setting  $a = b = 0$ , i.e.,  $r_t = cX_t^2$ , we obtain again a parameterized version of the Cox, Ingersoll, and Ross (1985) model,

$$dr_t = (c\sigma^2 - 2\kappa cr_t) dt + 2\sigma \sqrt{cr_t} dW_t. \quad (5)$$

Note the subtle difference between the specification in (3) (with  $a = 0$ ) and (5). In the latter parameterization, the interest rate can become zero. In (3), a suitable choice of parameters prohibits the interest rate from attaining zero.

As another special case of the dynamics in (4), we obtain the double square-root interest rate model of Longstaff (1989; 1992), and Beaglehole and Tenney (1992). These authors investigate a one-factor model with the short rate process given as

$$dr_t = \hat{\kappa} \left( \frac{\hat{\sigma}^2}{4\hat{\kappa}} - \sqrt{r_t} \right) dt + \hat{\sigma} \sqrt{r_t} dW_t, \quad (6)$$

for some parameters  $\hat{\kappa}$ ,  $\hat{\sigma}$ . The interest rate dynamics in (6) fits into our framework in (4) by setting  $a = b = 0$ ,  $\mu_X(X) = -\frac{1}{2}\sqrt{c}$ ,  $\sigma_X(X) = \sigma = \frac{1}{2}\hat{\sigma}/\sqrt{c}$ , and  $c = \hat{\kappa}$ .

### 1.3. Exponential Function $g(X)$

In lognormal term structure models, the short rate is of the form

$$r_t = \exp(a + bX_t). \quad (7)$$

A popular example is the Black and Karasinski (1991) model, which is a generalization of the continuous-time formulation of the Black, Derman, and Toy (1990) model. The Black and Karasinski (1991) model assumes the logarithm of the interest rate to evolve according to

$$d \ln r_t = (\theta_t - \kappa_t \ln r_t) dt + \sigma_t dW_t. \quad (8)$$

To obtain the process in (8) we have to set  $r_t = \exp(a_t + b_t X_t)$  with  $X_t$  following a normal distribution. As shown by Hogan and Weintraub (1993), such dynamics for the short interest rate would lead to infinite prices for Eurodollar futures. Nevertheless, the Black and Karasinski (1991) model is one of the most popular models used in practice.

Another possible specification within an exponential function  $g(X)$  would be to specify the state variable  $X$  as a normal process and define the interest rate as an exponential quadratic form of  $X$ , i.e.,

$$r_t = \exp(a + bX_t + cX_t^2).$$

Such a model specification is still tractable, since the expectation of  $r_t$  is available in closed form. However, the problem of infinite prices for Eurodollar futures remains.

#### 1.4. Other Specifications

There are many possible specifications for defining the function  $g(X)$  as long as the function is reasonably well-behaving. The following two examples are taken from Rogers (1996). The first one is a “rational-quadratic” model that specifies the short rate as

$$r_t = \frac{a + bX_t + cX_t^2}{\alpha + \beta X_t + \gamma X_t^2}, \quad (9)$$

where  $X$  is normally distributed. The specification in (9) allows highly nonlinear dynamics for the short rate. Zero bond prices are available in closed form (see Rogers (1996)). Another model with closed-form bond prices is

$$r_t = \beta\gamma X_t \tanh(\gamma X_t + \gamma c) + \alpha, \quad (10)$$

with  $X$  normally distributed. The interest rate dynamics in (10) share the feature that when interest rates are low they behave like in a quadratic Gaussian model, whereas for large  $X_t$  the interest rate behaves like in an affine Gaussian model. This feature makes (10) an attractive specification for the short rate from both a theoretical as well as from a practical perspective.

However, care has to be taken, since not all specification of the functional form for  $g$  result in arbitrage-free models. One possibility to check for admissability of a function  $g(X)$  is to follow the approach of Rogers (1996). There, the existence of a pricing kernel is linked to its representation as a potential. Nevertheless, the scope for generating interest rate models is immense.

## 2. State Variable Tree

The standard procedure to construct a trinomial tree approximation for the short rate process is to start spanning a tree for the state variable  $X_t$  on an equidistant tree. The process  $X_t$  is assumed to follow a time-homogeneous stochastic differential equation. Then, a tree representation can be constructed to provide a discrete-time and discrete-space Markov approximation for  $X_t$ . Usually, a trinomial tree is preferred to a binomial tree approximation, since the additional flexibility provided by the trinomial tree can be used to match not only the first, but also the second moment of the process  $X_t$ .

We make the following notational conventions. The nodes of the trinomial tree are denoted by  $(i, j)$ , where  $i$  is the vertical placement (the space axis) and  $j$  is the horizontal placement (the time axis). We define by  $\pi_{i,j}^{k,h}$  the probability by which the process moves from node  $(i, j)$  to node  $(i + k, j + h)$  within time interval  $h$ . For the standard branching process in the Hull and White (1994) model, it is assumed that (a) the time interval  $h$  is constant for the whole tree, and (b) the jump  $k$  is either 1, 0, or  $-1$ . Hence, the transition

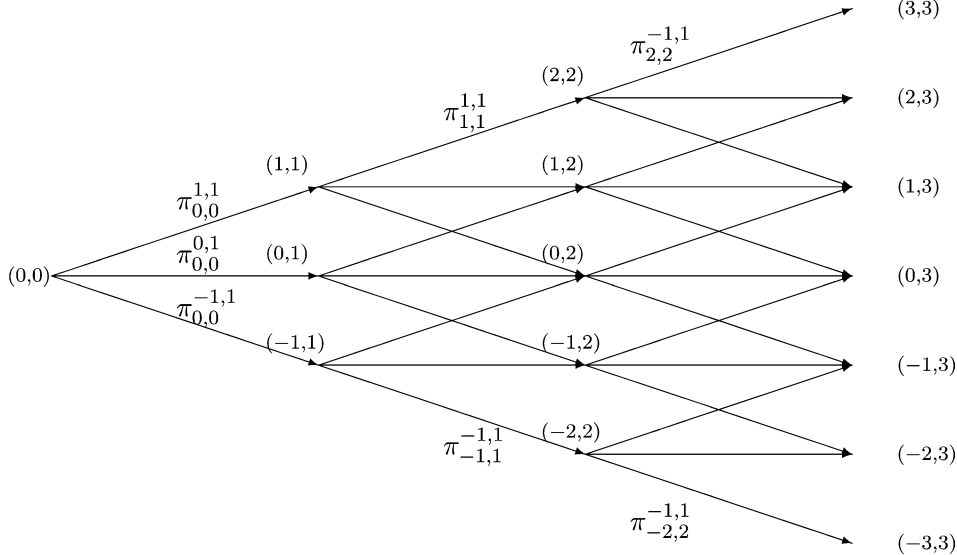


Figure 1. Trinomial tree. The tree starts at node  $(0, 0)$ . At each node there is a threesome  $\{\pi_{i,j}^{-1,1}, \pi_{i,j}^{0,1}, \pi_{i,j}^{1,1}\}$  evolving to the neighbor nodes  $(i - 1, j)$ ,  $(i, j)$ , and  $(i + 1, j)$ , respectively.

from a node  $(i, j)$  to nodes  $\{(i - 1, j + 1), (i, j + 1), (i + 1, j + 1)\}$  are equidistant both on the time as well as on the space axis. Furthermore, it is assumed that the nodes  $(i, j)$  and  $(i, j + 1)$  remain on the same vertical level. This branching process is illustrated in Figure 2(A) and the resulting trinomial tree structure is illustrated in Figure 1.

Having fixed the tree geometry, we have to determine the tree probabilities in such a way that the distributional properties of the state variable are mimicked at every node. However, instead of matching the variance as in Hull and White (1994), we match the second moment. This gives the same result, but simplifies the formulas somewhat. For the branching process given Figure 2(A), the system of equations is given by

$$\begin{aligned}
 1 &= \pi_{i,j}^{-1,1} + \pi_{i,j}^{0,1} + \pi_{i,j}^{1,1}, \\
 \mathbb{E}(X_{j+1}|X_{i,j}) &\equiv M_1 = \delta(\pi_{i,j}^{1,1} - \pi_{i,j}^{-1,1}) + X_{i,j}, \\
 \mathbb{E}(X_{j+1}^2|X_{i,j}) &\equiv M_2 = \pi_{i,j}^{-1,1}(X_{i,j} - \delta)^2 + \pi_{i,j}^{0,1}X_{i,j}^2 + \pi_{i,j}^{1,1}(X_{i,j} + \delta)^2,
 \end{aligned} \tag{11}$$

which is linear in the probabilities and hence straightforward to solve. In order to guarantee that the threesome  $\{\pi_{i,j}^{-1,1}, \pi_{i,j}^{0,1}, \pi_{i,j}^{1,1}\}$  can be interpreted as probabilities for all  $i$ 's, we have to guarantee  $\pi_{i,j}^{-1,1} + \pi_{i,j}^{0,1} + \pi_{i,j}^{1,1} = 1$  together with the three inequality constraints  $\{\pi_{i,j}^{-1,1} \geq 0, \pi_{i,j}^{0,1} \geq 0, \pi_{i,j}^{1,1} \geq 0\}$ . This can be done in several different ways. First, we can build some constraints on the number of time steps we are considering. However by doing so, we impose some severe restrictions on the depth of the tree. This method would probably fail to value either derivatives with complex payoff structures or long term instruments with intermediate payoffs, since such instruments require a reasonable depth

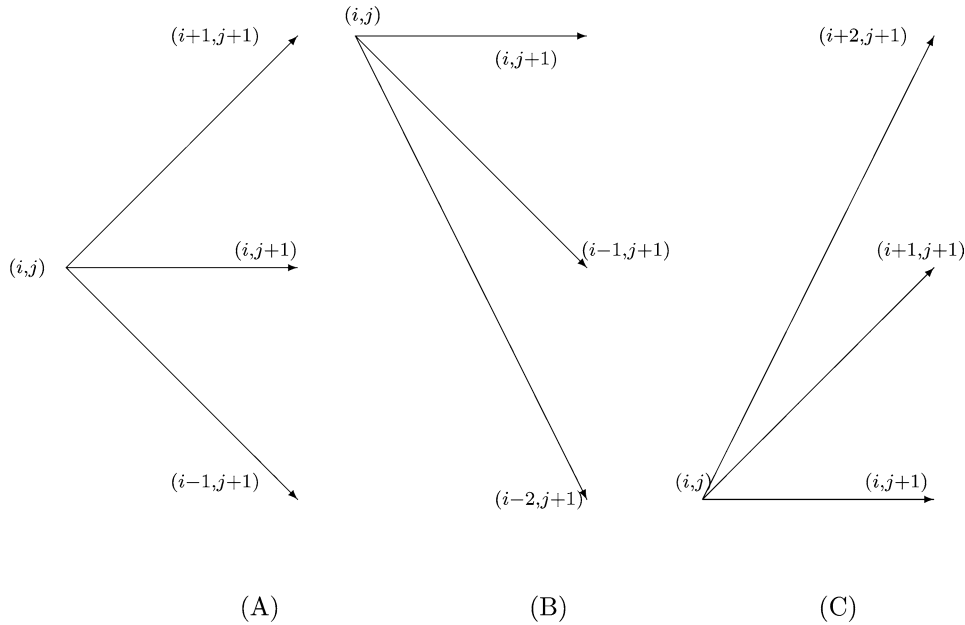


Figure 2. Branching processes. In order to control the state spanned by the tree, the common branching processes (A) is altered to either (B) for high interest rates, or (C) for low interest rates. With the latter branching processes, negative interest rates within the tree can be avoided.

for the tree. A more serious problem, in particular for Gaussian interest rate models, is the possibility of obtaining negative interest rates. As was done, e.g., in Hull and White (1994), this can be avoided by altering the geometry of the tree. Of course, altering the geometry is an arbitrary manipulation of the pricing problem and thus subject to some criticism. Nevertheless, it is widely used in practice.

An obvious way to avoid negative interest rates is to use the branching process illustrated in Figure 2(C). At the same time, the branching process in Figure 2(B) would avoid the possibility that the trinomial tree spans interest rates, which are unreasonably high. Depending on the branching process used, the system of equations for the tree probabilities in (11) has to be adjusted accordingly.

However, we do not necessarily have to rely on the assumption that the tree evolves to the three neighboring states. We can be more general in two directions:

- The tree evolves from state  $(i, j)$  to the three states  $\{(i + k_1, j + 1), (i + k_2, j + 1), (i + k_3, j + 1)\}$ , requiring  $k_1 \neq k_2 \neq k_3$ .
- The tree directly evolves from state  $(i, j)$  to a threesome of states at time  $j + h, h \geq 1$ .

With these two generalizations, the equations to match the first two moments are given by

$$M_1^h = X_{i,j} + \delta(\pi_{i,j}^{k_1,h} k_1 + \pi_{i,j}^{k_2,h} k_2 + \pi_{i,j}^{k_3,h} k_3),$$

$$M_2^h = \pi_{i,j}^{k_1,h} (X_{i,j} + k_1 \delta)^2 + \pi_{i,j}^{k_2,h} (X_{i,j} + k_2 \delta)^2 + \pi_{i,j}^{k_3,h} (X_{i,j} + k_3 \delta)^2.$$

In this general setup, we then obtain

$$\begin{aligned}\pi_{i,j}^{k_1,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_2 + k_3))M_1^h + (X_{i,j} + \delta k_2)(X_{i,j} + \delta k_3)}{\delta^2(k_1 - k_2)(k_1 - k_3)}, \\ \pi_{i,j}^{k_2,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_1 + k_3))M_1^h + (X_{i,j} + \delta k_1)(X_{i,j} + \delta k_3)}{\delta^2(k_2 - k_1)(k_2 - k_3)}, \\ \pi_{i,j}^{k_3,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_1 + k_2))M_1^h + (X_{i,j} + \delta k_1)(X_{i,j} + \delta k_2)}{\delta^2(k_3 - k_1)(k_3 - k_2)}.\end{aligned}$$

With these formulas for the probabilities at hand, we can construct a large structure of possible tree geometries. Obviously, there are many degrees of freedom for building a trinomial tree. In general however, the more degrees of freedom, the less stable will the tree be when it comes to pricing derivatives. Therefore, one has to be careful, as in most cases flexibility comes at the price of less stability. Also, the choice of the tree geometry should depend on the properties of the instruments to be priced. For convergence purposes and also to minimize possible biases, a finer grid should be used in those regions of the tree where the value of the instrument is highly sensitive to changes in the interest rate (see, e.g., Figlewski and Gao (1999)).

### 3. Calibrating the Tree

In practice, term structure models are implemented by calibrating them to prices and volatilities of some subset of traded instruments. These instruments include, e.g., US T-bonds, interest rate swaps, and interest rate options like caps and swaptions. Typically, the drift of the short rate process is matched to the current term structure of bond prices. In Section 3.1 we present a novel method to achieve this goal. In a second step, the volatility function of the short rate may then be chosen to match the term structure of volatilities of the yield curve, or the term structure of implied volatilities of at-the-money interest rate options. The latter is of particular importance when it comes to pricing of exotic interest rate options. We elaborate on this method in Section 3.2.

#### 3.1. Matching Bond Prices Using Forward Measure

The common procedure for matching the tree to the term structure of bond prices is based on forward induction as used by Hull and White (1994) and first proposed by Jamshidian (1991). This procedure obviously becomes computationally demanding for more involved functions  $g(X, t)$ . For completeness, we review this method in the Appendix. In this section, we use an approach that builds on the use of the forward measure. To this end, we exploit two properties of the forward measure.

**Property 1.** *Under the forward measure the forward rate is an unbiased estimator of the future interest rate (see, e.g., Björk (1998, p. 288)).*



The first property allows us to match the tree in a straightforward manner and, moreover, the level shift is given in closed-form for a wide range of term structure models. Thus, compared to the analytical implementation of the Hull–White model recently proposed by Grant and Vora (2001), our procedure is not restricted to the extended Vasicek model. However, as we will see, even if there is no analytical expression available, we can easily fit the tree to the initial term structure with only minor additional computational costs. Neither forward induction nor any root search algorithm are required.

**Property 2.** *In discrete time the one-period forward measure equals the risk-neutral measure (see, e.g., Pliska (1997, p. 222)).*

From the second property it follows that the backward induction to determine derivative prices remains the same as in the standard trinomial tree. Therefore, no further adjustments have to be made and prices can be calculated the same way as with a tree spanned under the risk-neutral measure.

Changing the probability measure to the forward measure is a common tool in pricing derivative instruments and considerably facilitates the calculation of the corresponding expectation.<sup>3</sup> Suppose that  $\int_0^T g(X, s) ds < \infty$ . The  $T$ -forward measure  $\mathbb{P}^T$  is defined by

$$\frac{d\mathbb{P}^T}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \frac{\exp(-\int_0^T g(X, s) ds)}{\mathbb{E}[\exp(-\int_0^T g(X, s) ds)]} = P(t, T)^{-1} \exp\left(-\int_0^T g(X, s) ds\right).$$

Then, as stated in Property 1, we have

$$\mathbb{E}^T(r_T | \mathcal{F}_t) = f(t, T),$$

where  $\mathbb{E}^T$  is the expectation operator under the  $T$ -forward measure. Since it is more natural to work with simple compounded interest rates (e.g., LIBOR rates), we denote by  $f_\Delta^*(T) = (P(T)/P(T + \Delta) - 1) / \Delta$  the annualized, discrete time forward rate prevailing at time  $[T, T + \Delta]$  as observed from the initial term structure. In order to match the initial term structure, we have to insure that

$$\mathbb{E}^T(r_T | \mathcal{F}_t) = \mathbb{E}^T(g(X, T) | \mathcal{F}_t) = f_\Delta^*(T),$$

holds for each time step in the trinomial tree. How can this be achieved?

Up to now, we have constructed the tree for  $X$ , such that for every time step the conditional expectation of  $\Delta X$  is matched under the risk-neutral measure. Therefore, the probabilities  $\pi_{i,j}^{k,h}$  are  $\mathbb{P}$ -probabilities. Note that, in discrete time, the one-period forward measure equals the risk-neutral measure (see Property 2). So far we did not yet have specified the measures of path probabilities for more than one period. This additional degree of freedom will be used to efficiently match the tree to the initial term structure. How this can be achieved will be discussed next.

Given the appropriate technical conditions, the extended generator of  $X_t$  under  $\mathbb{P}^T$  following the SDE in (1) can be written as

$$\mathcal{L}^T g = \mathcal{L}g + \Gamma(\log P(t, T), g),$$

where  $\Gamma$  is the “carré du champ operator” corresponding to  $\mathcal{L}$  (see Davis (1998)), defined as

$$\Gamma(f, g) = \mathcal{L}(fg) - g\mathcal{L}f - f\mathcal{L}g, \quad f, g \in \mathcal{D}(\mathcal{L}).$$

This means that in continuous time, the drift of the process  $dX_t$  is changed from  $\mu(X_t)$  to  $\mu(X_t) - \sigma(X_t)\frac{\partial}{\partial X} \log P(t, T)$  under the forward measure  $\mathbb{P}^T$ :

$$\begin{aligned} dX_t &= \left( \mu(X_t) - \sigma(X_t)\frac{\partial}{\partial X} \log P(t, T) \right) dt + \sigma_X(X_t) dW_t^T \\ &= (\mu(X_t) - v(X, t, T)) dt + \sigma_X(X_t) dW_t^T, \end{aligned} \quad (12)$$

where  $v(X, t, T)$  is the instantaneous volatility of the bond price process  $dP(t, T)/P(t, T)$ , and  $W^T$  is the standard Brownian motion under the measure  $\mathbb{P}^T$ .

Now, the level shift of the original tree can be determined as follows. Using continuous time notation, we need to find a level shift such that the forward rate is an unbiased estimate of the future interest rate, i.e.,

$$\mathbb{E}^T(g(X, T)|\mathcal{F}_t) = f(t, T),$$

where

$$\begin{aligned} \frac{d\mathbb{E}^T[g(X, T)|\mathcal{F}_t]}{dt} &= \mathbb{E}^T[\mathcal{L}^T g(X, T)|\mathcal{F}_t] \\ &= \mathbb{E}[\mathcal{L}g(X, T) + \Gamma(\log P(t, T), g(X, T))|\mathcal{F}_t]. \end{aligned}$$

Thus,

$$\begin{aligned} f(t, T) &= f(t, t) + \int_t^T \mathbb{E}[\mathcal{L}g(X, s) + \Gamma(\log P(s, T), g(X, s))|\mathcal{F}_t] ds \\ &= \mathbb{E}[g(X, T)|\mathcal{F}_t] + \int_t^T \mathbb{E}[\Gamma(\log P(s, T), g(X, s))|\mathcal{F}_t] ds \\ &= \mathbb{E}^T[r_T|\mathcal{F}_t], \end{aligned} \quad (13)$$

subject to  $g(X, t) = f(t, t) = r_t$ . In order to match the tree, we alter the original tree for the Markov process  $X$  in the following way.

- At each time slice, we change the level of the tree for  $g(X, t)$  by a function  $\eta(t, T, X)$  defined by

$$\eta(t, T, X) = \int_t^T \mathbb{E}[\Gamma(\log P(s, T), g(X, s))|\mathcal{F}_t] ds. \quad (14)$$

To perfectly match the tree to the initial term structure at time  $t$ , the shift at each time slice  $j\Delta$  is determined by

$$\eta(t, j\Delta, X) = f_\Delta^*(j\Delta) - \mathbb{E}[g(x_{j\Delta}, j\Delta)|\mathcal{F}_t]. \quad (15)$$

- The tree for  $r_t$  is now spanned under forward probability measures. At time step  $j$ , we have  $\mathbb{E}^{j\Delta}[r_{j\Delta}|\mathcal{F}_0] = f(0, j\Delta)$  under measure  $\mathbb{P}^{j\Delta}$ . The one period forward measure equals the risk-neutral probability measure.

**3.1.1. The Extended Vasicek Model** To further clarify our point, we next discuss a few examples of term structure models that can be handled analytically within our approach. We start by discussing the extended Vasicek model. Assuming

$$dX_t = -\kappa X_t dt + \sigma dW_t, \quad X_0 = 0, \quad (16)$$

we fix the initial date to 0. We define the function  $g(X, t) = g(X_t) = X_t$ . Since in the extended Vasicek model, the volatility of the bond price process is a function of time only, we can write the forward rate in (13) as

$$f(t, T) = \mathbb{E}[g(X_T)|\mathcal{F}_t] + \eta(t, T).$$

Then, according to Equation (14), the level shift at time  $t = 0$  becomes

$$\eta(0, \Delta) = f_\Delta^*(\Delta) - \mathbb{E}(X_\Delta|\mathcal{F}_0) = f_\Delta^*(\Delta).$$

For the next time-step,  $[\Delta, 2\Delta]$ , we simply obtain

$$\eta(0, 2\Delta) = f_\Delta^*(2\Delta).$$

Obviously, we do not have to make any tedious calculations at all. The level shift at time step  $(j-1)\Delta$  is simply given by

$$\eta(0, j\Delta) = f_\Delta^*(j\Delta). \quad (17)$$

Finally, we can determine the measure change  $d\mathbb{P}^T/d\mathbb{P}$  as

$$\left. \frac{d\mathbb{P}^T}{d\mathbb{P}} \right|_{\mathcal{F}_T} = \exp\left(\int_0^T \frac{\partial \eta(u, T)}{\partial T} dW_u - \frac{1}{2} \int_0^T \left(\frac{\partial \eta(u, T)}{\partial T}\right)^2 du\right).$$

Since the derivation of the bond price is a straightforward task, the  $\mathbb{P}_T$ -dynamics of  $X_t$  in Equation (12) are explicitly obtained as

$$dX_t = \left(\frac{e^{-\kappa(T-t)} - 1}{\kappa} \sigma^2 - \kappa X_t\right) dt + \sigma dW_t^T, \quad X_0 = 0. \quad (18)$$

Indeed, whenever we have an affine term structure model, where bond prices allow the representation

$$P(t, T) = \exp(A(t, T) + B(t, T)r_t),$$

the change of measure is

$$\begin{aligned} & \left. \frac{d\mathbb{P}^T}{d\mathbb{P}} \right|_{\mathcal{F}_T} \\ &= \exp\left(\int_0^T B(u, T) \frac{\partial g(X, u)}{X} \sigma(X) dW_u - \frac{1}{2} \int_0^T \left(B(u, T) \frac{\partial g(X, u)}{X} \sigma(X)\right)^2 du\right). \end{aligned}$$

As was already pointed out in Kijima and Nagayama (1994) and Pelsser (1994), the level shift can be calculated analytically for the Vasicek model. They argue that the level shift equals the expected value of the future interest rate. Thus, there is no forward induction necessary. Hull and White (1996) object that this procedure does not provide an exact fit to the initial term structure, because the tree is a discrete time representation of an underlying continuous process. Hence, the tree is only fitted exactly using the forward induction procedure, which would justify the additional computational costs. Here, however, we calculate the expected value under the *forward measure* to determine the level shift for the one-period forward rate, which equals the discretized short rate. Hence, the tree is matched *exactly* to the initial term structure, while saving considerable amount of computation time.

**3.1.2. Lognormal Models** The extended Vasicek model has a significant drawback. The interest rate can become negative. This property is rather unsatisfactory from both a practical as well as a theoretical perspective. In lognormal models of the term structure the interest rate remains positive. Indeed, practitioners base their market formulas of caps and swaptions on the assumption of lognormal rates (Black (1976)).

We assume  $dX_t$  as given in Equation (16), but now with  $r_t = ae^{X_t}$ . The short rate becomes lognormally distributed with expectation

$$\mathbb{E}[r_T | \mathcal{F}_t] = a \exp(X_t e^{-\kappa(T-t)}) + a \exp\left(\frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(T-t)})\right).$$

Therefore, for the tree centered at  $X_0 = 0$ , and with fixed  $\kappa$  and  $\sigma$ , the relation

$$r_0 = f_{\Delta}^*(\Delta) = a + a \exp\left(\frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa\Delta})\right)$$

determines the constant  $a$ . Then, to match the whole term structure, we can explicitly derive the level shift in (13) as

$$\eta(0, j\Delta) = f_{\Delta}^*(j\Delta) - a - a \exp\left(\frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa j\Delta})\right)$$

to determine the level shift at the  $(j - 1)\Delta$ th time step.

We recall that with the standard Hull and White (1994) procedure one would have to perform a root search algorithm for each time step in order to determine the level shift for matching the tree with lognormal interest rates. Using, e.g., the Newton–Raphson procedure to find numerically the level shift may induce numerical instabilities, since such a procedure is sensitive to the starting values.

**3.1.3. General Quadratic Models** With the parameters  $a$ ,  $b$ , and  $c$  such that

$$r_t = a + bX_t + cX_t^2 \geq 0, \quad (19)$$

we can span an interest rate tree with strictly positive interest rates. For the strict inequality (19) to hold, we require

$$a > \frac{b^2}{4c}.$$

Again, we define  $X_t$  as an Ornstein–Uhlenbeck process given in Equation (16). Then,

$$\mathbb{E}[r_T | \mathcal{F}_t] = a + \frac{c\sigma^2}{2\kappa}(1 - e^{-2\kappa(T-t)}) + X_t e^{-\kappa(T-t)}(b + cX_t e^{-\kappa(T-t)})$$

and with  $X_0 = 0$ , we obtain for the short rate at time  $r_0$

$$r_0 = f_\Delta^*(\Delta) = a + \frac{c\sigma^2}{2\kappa}(1 - e^{-2\kappa\Delta}).$$

To match the entire initial term structure, the level shifts for  $j > 1$  can readily be calculated as

$$\eta(0, j\Delta) = f^*(j\Delta) - a - \frac{c\sigma^2}{2\kappa}(1 - e^{-2\kappa j\Delta}). \quad (20)$$

**3.1.4. The Cox–Ingersoll–Ross (CIR) Model** As noted in Section 1.2, depending on the state variable process, the CIR model can be either fitted into the affine class of term structure models or into the quadratic Gaussian class (see also Maghsoodi (1996)). As noted by Jamshidian (1996), there are some theoretical problems involved, when we specify the interest rate as  $r_t = X_t^2$ . First of all, when we would set  $\mu_X(X) = \theta - \kappa X_t$  with  $\theta \neq 0$ , the spot rate will not be a diffusion Markov process. But then, when we set  $r_t = X_t^2$  with  $\theta = 0$ , the interest rate, although never negative, can become zero.<sup>4</sup>

In practice, the CIR model is implemented by using some ad-hoc procedures. One example can be found, e.g., in Brigo and Mercurio (2001). They set

$$dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dW_t \quad (21)$$

and define the short rate as an affine function of the state variable, namely

$$r_t = X_t + \rho_t, \quad (22)$$

where  $\rho_t$  is the level shift that matches the model to the initial term structure. For convergence purposes, they span the tree for an auxiliary process  $Y_t$  with a constant instantaneous volatility, defined as

$$Y_t = \sqrt{X_t}. \quad (23)$$

The tree for  $Y$  still has the problem that some nodes may become negative. To avoid negative nodes, Brigo and Mercurio (2001) suggest to alter the tree geometry accordingly. Setting again  $X_t = Y_t^2$ , the final step is to find the value of  $\rho_t$  such that the tree is matched to the initial term structure.

We remark that our calibration procedure outlined in this section remains applicable. Indeed, our calibration procedure is as general as the forward induction of Hull–White.

If a closed-form expression for the correction term is not available, the calculation of this correction term can be approximated using the path-probabilities in the tree. For  $T = j\Delta$ , we obtain

$$\mathbb{E}^T [g(X_T) | \mathcal{F}_t] \approx \sum_i \Pi(i, j\Delta) g(X_{i,j}), \quad (24)$$

where  $\Pi(i, j\Delta)$  denotes the sum of all path probabilities leading to state  $(i, j)$ . Using (24), the tree is also matched exactly to the initial term structure. Since the path probabilities  $\Pi(i, j\Delta)$  have to be calculated only once, the additional computational costs of using Equation (24) instead of an analytical formula are low. Our numerical experiments for the Black–Karasinski model indicate that, compared to forward induction, the calculation of a one-year European option using a tree with 200 time steps is about five times faster when using our approach with (24). Using a tree with 3000 time steps, our method is still four times faster. If we do not recalculate the probabilities (24), then the efficiency gain is around a factor of 150 for a tree with 200 time steps and 140 for a tree with 3000 time steps.<sup>5</sup>

### 3.2. Calibrating to the Term Structure of Volatilities

In this section, we show how the trinomial tree can be fitted in an efficient way to the initial term structure of volatilities. Calibrating the volatility structure is often subject to some criticism. Hull and White (1996) suggest that in a Markov model there should be only one time-dependent parameter. Whenever the volatility is modeled as time-dependent, the resulting non-stationarity in the volatility curve may have many unexpected effects. In particular, any instrument whose price depends on future volatilities is liable to be mispriced.

Before we start discussing, how the trinomial tree can be manipulated to match the volatility structure, we explore the theoretical underpinning of calibrating the volatility. First, note that in the previous section we introduced a measure change to determine the level shift for the interest rate tree. An absolutely continuous change of measure only affects the drift of the process, but the quadratic variation will not be affected. Hence, for the tree construction, it is more appropriate to first match the volatility and then match the forward rate curve.

To alter the diffusion coefficient, we have to introduce a time change.<sup>6</sup> As we are only considering a deterministic time change, we can use the following result.

**Lemma 1.** Consider the continuous time-dependent functions  $c(t) > 0$  and  $\tau(t) = \int_0^t c(s)^{-1} ds \geq 0$ . Define

$$\tilde{W}_t = \int_0^{\tau(t)} \sqrt{c(s)} dW_s.$$

Then  $\tilde{W}_t$  is a  $\mathcal{F}_{\tau(t)}$ -Brownian motion. Further,

$$\int_0^{\tau(t)} dW_s = \int_0^t \sqrt{\frac{\partial \tau(s)}{\partial s}} d\tilde{W}_s. \quad (25)$$

From Equation (25), the time-changed Brownian motion alters the volatility of the original Brownian motion. Hence, this technique offers a convenient tool to match the trinomial tree to the term structure of volatilities.

In the previous section, where we matched the tree to the forward rate curve only, we implicitly assumed  $\tau(t) = t$  and hence,  $c(t) = 1$ . By introducing the time change, we now construct a tree for the new process

$$dX_t = \mu_X(X) dt + \tilde{\sigma}_X(X, t) d\tilde{W}_t, \quad (26)$$

where  $\tilde{\sigma}_X(X, t) = \sigma_X(X_t) \sqrt{\partial_t \tau(t)}$ . Since we use now the process (26), care has to be taken when matching to the initial forward rate curve. The results in the previous section have to be adjusted accordingly. We will show below, how this can be achieved in an efficient manner. The introduction of a deterministic clock for volatility matching was already introduced in Schmidt (1997). Contrary to Schmidt (1997), we do not change the length of the interval between subsequent time-steps to fit this concept into our trinomial framework. Instead, we adjust the jump size in each time step in such a way that the tree probabilities remain unchanged. Hence, the jump size  $\delta$  is becoming a function of time and we denote it as  $\delta_j$ .

Assume now that we want to calibrate our interest rate tree to the term structure of forward rate volatilities.<sup>7</sup> We further assume that we are given a set of volatilities of one-period forward rates, denoted by  $V^*(j\Delta)$ . To construct a tree calibrated to the volatilities  $V^*(j\Delta)$ , we can follow a simple step-by-step procedure:

1. In the first step, we construct the tree as usual, i.e., we start with the process  $X$  and solve the system of equations in (11) to obtain the tree probabilities.
2. Next, we need to find the value  $\delta_j$  for each time step, such that the observed volatilities are calibrated. We obtain the values  $\delta_j$  by solving

$$V^*(j\Delta) = \sum_i \Pi(i, j\Delta) g(X_{i,j}, \delta_j)^2 - \left( \sum_i \Pi(i, j\Delta) g(X_{i,j}, \delta_j) \right)^2 \quad (27)$$

for  $\delta_j$ . Note that only  $g(X_{i,j}, \delta_j)$  depends on  $\delta_j$ , but not the path probabilities  $\Pi(i, j\Delta)$ . Solving (27), we end up with a tree for the process  $g(X_{i,j}) = g(X_{i,j}, \delta_j)$  that is calibrated to the observed volatilities.

3. Finally, we introduce the level shift as outlined in Section 3.1. Such a level shift does not alter the calibrated tree volatilities.

### 3.3. Convergence

Before we turn to numerical examples, we briefly address convergence issues. We do not provide theoretical results here. Instead, we refer to the work of Lesne, Prigent, and Scaillet (2000) that is directly applicable in our context. Based on the limit theorems of stochastic processes in Jacod and Shiryaev (1987), they extend results from Hubalek and Schachermayer (1998) and establish the convergence of trinomial trees using results for martingale difference arrays.

From Proposition 1 of Lesne, Prigent, and Scaillet (2000) and the continuous mapping theorem, the discretization of a continuous time financial market, including stocks and bonds, converges to its continuous-time limit, and so do the option prices. A requirement for the convergence are continuity and linear growth properties for the drift and diffusion terms of the processes involved.

These requirements on the functional forms of the drift and the diffusion terms put, in addition to the no-arbitrage condition, some restriction on the function  $g(X)$ . However, these requirements do not concern the way in which the trinomial tree is matched to the term structure. Therefore, whenever a term structure model converges within a Hull–White trinomial tree, it converges when using our calibration procedure. Since we can directly rely on the result in Lesne, Prigent, and Scaillet (2000), we restrict the discussion of convergence to a numerical exploration in Section 4.

Finally, we note that the prices obtained using the traditional forward induction and the prices obtained via the forward measure approach converge in the trinomial tree when  $\Delta \rightarrow 0$ . However, for a given  $\Delta$ , the prices must not be identical as we will see, e.g., in Figure 6. The reason for these difference lies in the use of different probability measures and different level shifts. Only in the limit, as  $\Delta \rightarrow 0$ , the two matching methods converge.

## 4. Numerical Examples

### 4.1. Illustrative Example

The best way to convey the intuition behind our calibration procedure is through a simple example. We start with an affine term structure model based on a Gaussian state variable  $X_t$ . We assume that  $X_t$  follows the process

$$dX_t = -\kappa X_t dt + \sigma dW_t.$$

The state spacing of the tree is given by

$$\delta = \sqrt{3}\sigma \sqrt{\frac{1 - e^{-2\kappa\Delta}}{2\kappa}},$$

for numerical reasons (see Hull and White (1994)). Setting

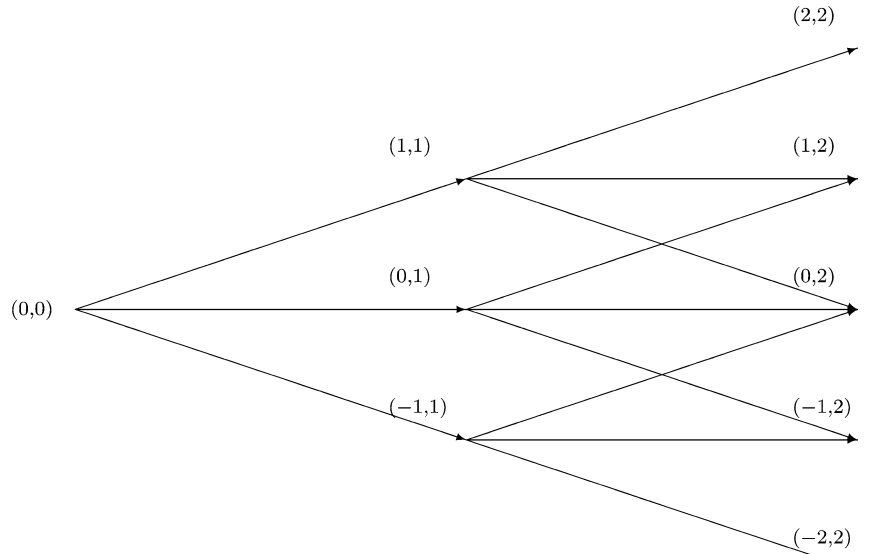
$$g(X) = X_t,$$

we obtain a Gaussian interest rate model. The branching probabilities are calculated in the same way as in the Hull–White trinomial tree. In the next step, instead of calculating the state prices as in Hull and White (1994), we match the tree to the initial term structure by using Equation (17), i.e., the interest rate  $r(i, j)$  at node  $(i, j)$  is simply given as

$$r(i, j) = x_{ij} + f_{\Delta}^*((j+1)\Delta),$$

where  $x_{ij}$  is value of the discretized version of the process  $X_t$  at node  $(i, j)$ . Hence, for each time slice, the tree is shifted upwards according to the prevailing forward rate. By doing so, the tree is matched to the initial bond prices.





Node	(0,0)	(1,1)	(0,1)	(-1,1)	(2,2)	(1,2)	(0,2)	(-1,2)	(-2,2)
$\pi_{i,j}^{1,1}$	0.1667	0.1569	0.1667	0.1768	0.1476	0.1569	0.1667	0.1768	0.1873
$\pi_{i,j}^{0,1}$	0.6666	0.6663	0.6666	0.6663	0.6651	0.6663	0.6666	0.6663	0.6651
$\pi_{i,j}^{-1,1}$	0.1667	0.1768	0.1667	0.1569	0.1873	0.1768	0.1667	0.1569	0.1476
$\Pi_{i,j}$	1	0.1667	0.6666	0.1667	0.0261	0.2222	0.5034	0.2222	0.0261
$r_{i,j}^{HW}$	4.9686	5.7381	4.9635	4.1889	6.6414	5.8668	5.0922	4.3176	3.5430
$r_{i,j}^{LW}$	4.9934	5.7549	4.9881	4.2211	6.6515	5.8846	5.1177	4.3508	3.5838

Figure 3. Construction of the trinomial tree for the extended Vasicek model. We assume an Ornstein–Uhlenbeck process for  $X$  with  $\kappa = 0.1, \sigma = 0.01, \Delta = 0.2$ , and base our calculations on the term structure given in Table 1. The interest rates  $r_{i,j}^{LW}$  are obtained by using our calibration approach, the interest rates  $r_{i,j}^{HW}$  are based on the Hull–White procedure.

The construction of the trinomial tree for  $g(X) = X_t, \kappa = 0.1, \sigma = 0.01, \Delta = 0.2$  is illustrated in Figure 3. We base our calculations on the term structure given in Table 1 taken from Hull (2003, p. 560). The interest rates  $r_{i,j}^{LW}$  are obtained by using our calibration approach, the interest rates  $r_{i,j}^{HW}$  are based on the Hull–White procedure.<sup>8</sup> We observe that in both cases the shifts of the trinomial tree is of a similar magnitude. As a consequence, there will only be a small difference between the interest rates, and particularly so for short time horizons.

The calculations of the tree probabilities for a single time step remain the same in both calibration procedures. However, we have to recall that the tree spans under the forward measure and not under the risk-neutral measure. This means that the path probabilities  $\Pi(i, j)$  in Figure 3 have to be interpreted as probabilities under the corresponding forward measures, i.e., under the  $j \Delta$ -forward measure. In the Hull and White model, these proba-

bilities would simply be risk-neutral probabilities for all  $j$ 's. These measure changes have to be kept in mind when pricing derivatives on the tree (see Section 4.2.1).

If we would expand the tree to more time steps, we will observe that interest rates soon turn negative in the extended Vasicek model (Figure 4(A)). A possibility to avoid negativity of interest rates is to consider a quadratic term structure model instead. The level shift is given by Equation (20). The upward shift by the forward rate has to be adjusted by a correction term. This correction term is available in closed form (see Equation (20)). The quadratic tree is plotted in Figure 4(B).

As another example, consider the CIR model in Equations (21)–(22). To construct the tree, we follow the methodology of Brigo and Mercurio (2001). However, to match the tree to the initial term structure, we do not apply a forward induction, but we use our methodology based on the forward measure. To avoid negative interest rates, the tree geometry has to be adjusted. Therefore, we introduce an alternative branching process whenever the lower nodes would become negative. As we observe in Figure 5, such a procedure kinks the lower boundary of the tree. In order to avoid too high interest rates, we also kink the tree at the upper boundaries. Again, there are many different alternatives to manipulate the tree's geometry.

#### 4.2. Comparison with the Hull–White Trinomial Tree

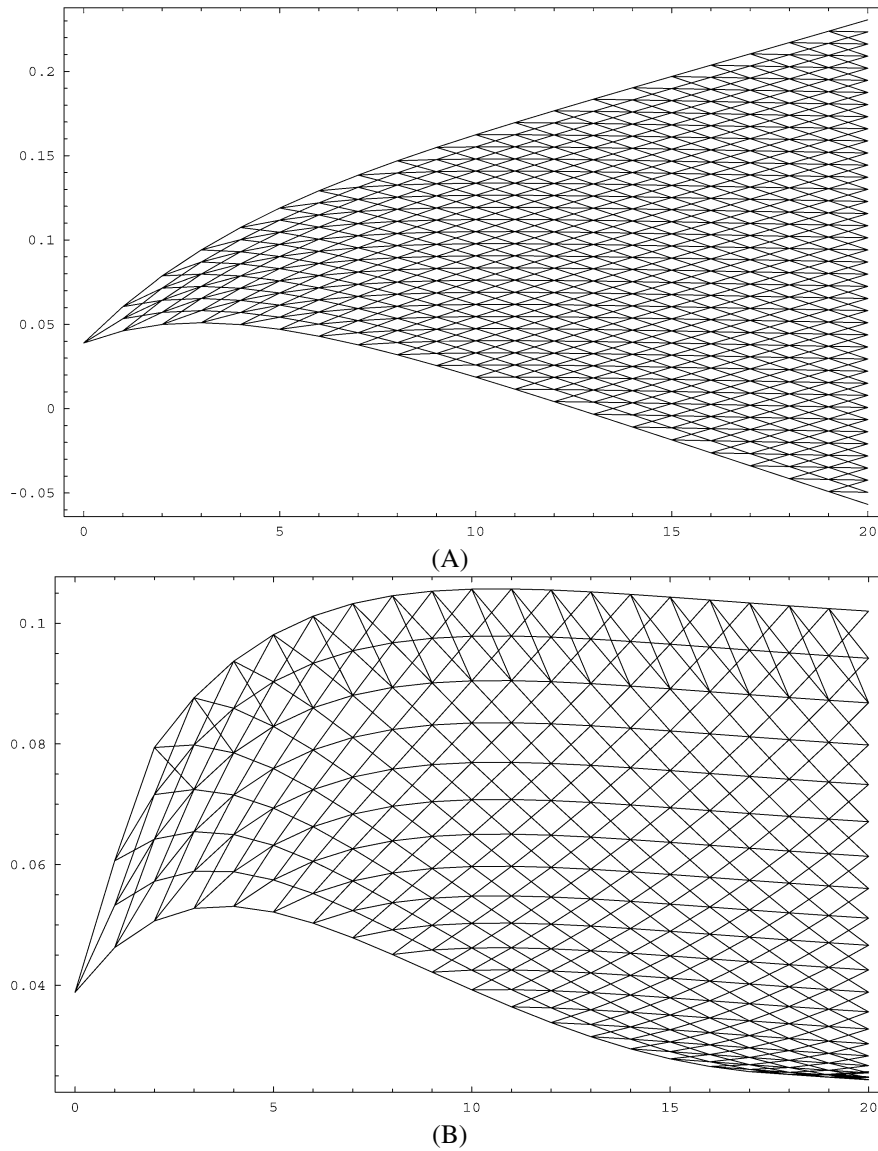
In this section, we compare the convergence of option prices and sensitivities using our calibration procedure based on the forward measure and the traditional Hull–White forward induction.<sup>9</sup>

**4.2.1. Option Prices** Under the forward measure, the time- $t$  price  $V_t$  of a European claim with payoff  $V_T$  at maturity  $T$ , we have

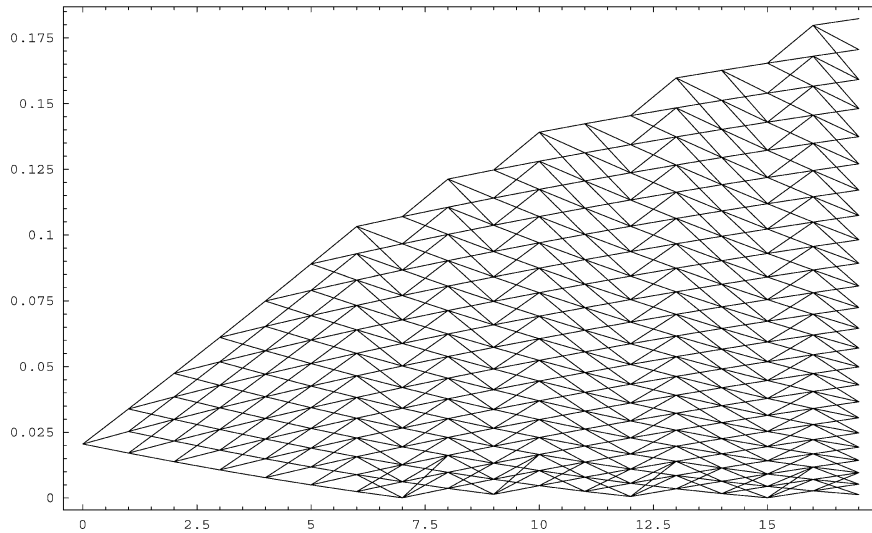
$$V_t = P^*(T) \mathbb{E}^T [V_T | \mathcal{F}_t].$$

Thus, when using the tree for valuation, we only need to discount the payoffs at  $T$  with the appropriate discount factor, i.e., the appropriate (and observable) bond prices. The same is true for all other path-independent options, which can be decomposed into a portfolio of European options. Note that, since in the current discrete time setting, the one-period forward measure equals the risk-neutral probability measure, we can determine American option prices by recursively working through the tree.

For comparison, we only calculate European option prices and base the subsequent calculations on the term structure given in Table 1. Based on this term structure, Hull (2003) calculates the value of a put option on a 9 year zero bond with face value 100. The maturity of the put option is 3 years and the strike price is set equal to 63. Furthermore, Hull assumes the interest rate to follow a mean-reverting Ornstein–Uhlenbeck process with parameters  $\kappa = 0.1$  and  $\sigma = 0.01$ . For a model with normally distributed interest rates no root search is necessary. Furthermore, the value of the put can be derived analytically, which makes this choice particularly suited to explore the convergence properties.



*Figure 4.* Trinomial trees for an affine (A) and a quadratic (B) Gaussian term structure model. The affine model leads to negative interest rates, when no alternative branching processes are introduced. The Gaussian model has only positive interest rates. To avoid to high interest rates for this model, we introduced the alternative branching process (B) from Figure 2.

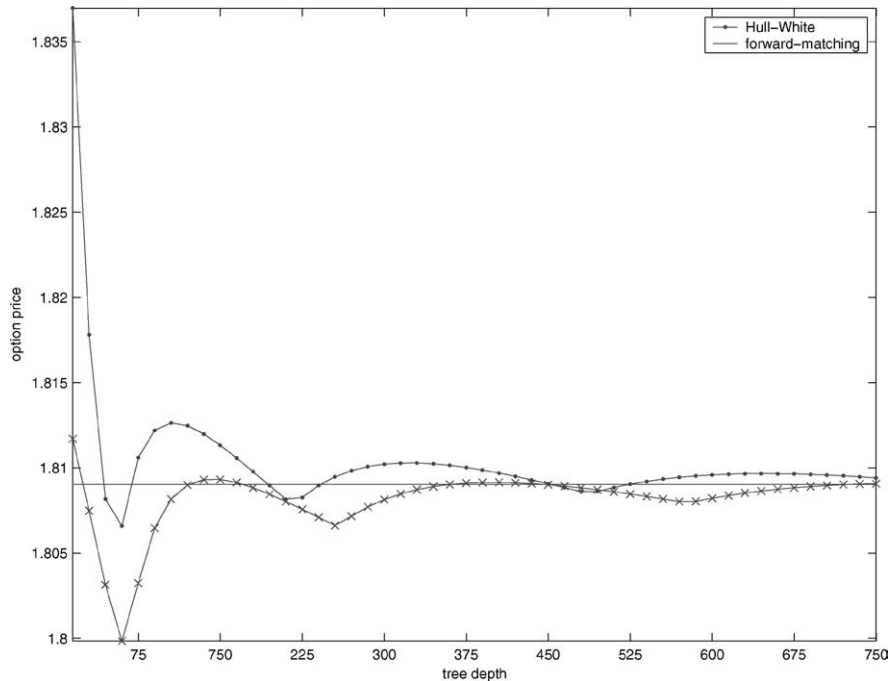


*Figure 5.* Trinomial tree for the CIR model. We match the CIR interest rate model to the initial term structure using the forward measure approach. Whenever the interest rate would fall below zero, we introduce an alternative branching process. The resulting tree stays positive, but has some kinks at its lower edge. Additionally, we introduce kinks at the upper edge.

*Table 1.* Zero Curve with Continuously Compounded Interest Rates

Maturity (Days)	Rate (%)
3	5.01772
31	4.98284
62	4.97234
94	4.96157
185	4.99058
367	5.09389
731	5.79733
1096	6.30595
1461	6.73464
1826	6.94816
2194	7.08807
2558	7.27527
2922	7.30852
3287	7.39790
3653	7.49015

Using Hull's (2003) example, we compare the two different calibration methods. To this end, we calculate the put option value for different tree depths up to a depth of 750. The results are plotted in Figure 6. Both methods converge to the true analytical value following a similar convergence pattern. Therefore, we conclude that the convergence properties do not favor one method over the other.



*Figure 6.* Convergence of the trinomial trees. We compare option prices for the extended Vasicek model for the calibration procedure based on the Hull–White forward induction and based on the forward measure approach. The analytical option pricing formula serves as benchmark. We value a put option on a 9 year zero bond with face value 100, maturity of 3 years, and a strike equal to 63. The interest rate follows a mean-reverting Ornstein–Uhlenbeck process with parameters  $\kappa = 0.1$  and  $\sigma = 0.01$ . The tree is matched to the term structure given in Table 1.

However, from a computational viewpoint there is, even for the simple extended Vasicek model, a clear advantage from using our calibration method. Recall that the Hull–White method is still based on forward induction. This forward induction has to be performed every time when the tree has to be adjusted to the new term structure. Using our method, the most costly part is the calculation of the path probabilities. However, these probabilities do not depend on market values, and therefore only have to be calculated once. When the term structure changes from one day to the next, the only thing to do is to shift the tree at every time slice according to the new forward rates. In the extended Vasicek model, this shift involves just a simple addition. No recalculation of probabilities is needed.

As another example, we calculate the prices of a one year interest rate call options for the Black–Karasinski model. In Figure 7, we compare the option prices for different tree depths. The option price calculated with tree depth of 1000 time steps serves as benchmark. To calculate prices, we assume  $g(X) = e^X$ ,  $\kappa = 0.8$ ,  $\sigma = 0.075$ , and the caprate is fixed at 0.05.

In Figure 7, we observe that both calibration methods converge. For the calculated example, we see that the method based on the forward measure does a better job. The

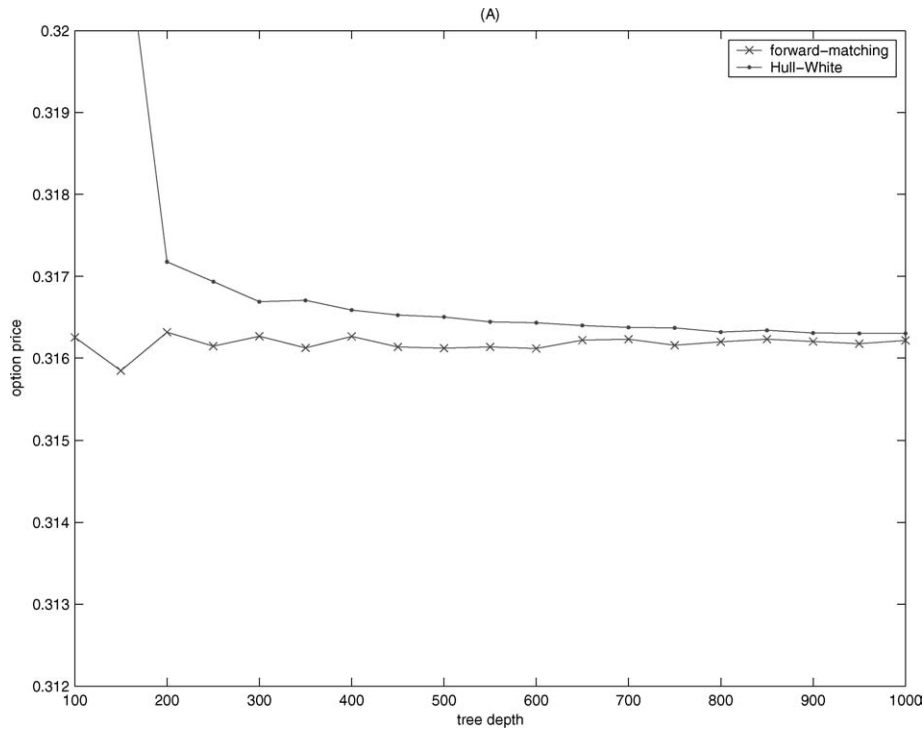


Figure 7. Convergence of the trinomial trees. We calculate the price of a one-year interest rate call option for the Black–Karasinski model for different tree depths. The calibration procedures are based on the Hull–White forward induction and on the forward measure approach. The option price calculated with tree depth 1000 serves as benchmark. To calculate prices, we assume  $g(X) = e^X$ ,  $\kappa = 0.8$ ,  $\sigma = 0.075$ , and the caprate is fixed at 0.05. The tree is matched to the term structure given in Table 1.

numerical values for small tree depths already start close to the benchmark value calculated with a tree depth of 1000 steps. However, at this stage, we are not able to come up with a theoretical justification of this point. We leave this issue open for future research.

**4.2.2. Sensitivities** For hedging and risk management purposes, sensitivities to changes in different model inputs can be calculated in the same manner as with trinomial trees calibrated by forward induction. As an example, we can calculate the sensitivity of a put option on a bond with respect to a one basis point upward-shift of the term structure. We assume an extended Vasicek model and consider a put option with maturity of one year on a three year bond. The strike is fixed at 89%. The remaining parameters are set equal to those we used in Figure 6.

From Figure 8, we see that both methods converge in a similar manner. This was to be expected, since prices also converge similarly (see Figure 6). Note, however, that our calibration procedure is again much more efficient. Since for the sensitivity analysis we do not have to recalculate the path probabilities, we can shift the tree according to the shifted

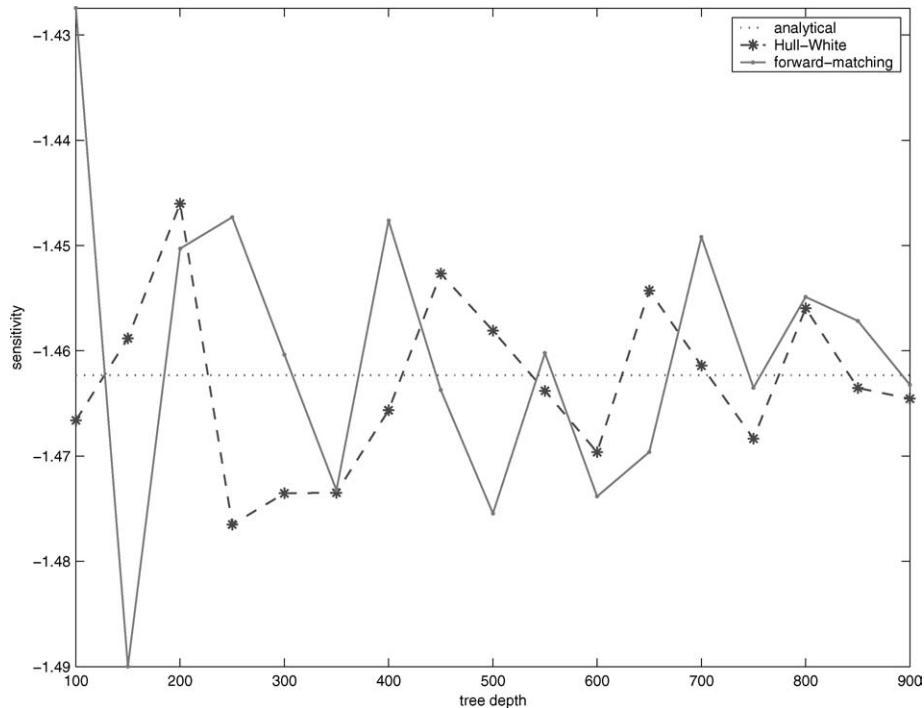


Figure 8. Convergence of the trinomial trees. We compare the sensitivity of a call option to an upward parallel shift in the yield curve. We calculate price changes for the extended Vasicek model for the calibration procedure based on the Hull–White forward induction and based on the forward measure approach. The analytical option pricing formula serves as benchmark. The call option is on a three year zero bond and has maturity of one year. The tree is matched to the term structure given in Table 1. The initial term structure is shifted by one basis point.

forward rate curve and use the same probabilities to calculate prices. In contrast, when using forward induction, we would have to recalibrate the tree, and when no closed form solution is available for the shift (such as, e.g., in the Black–Karasinski model), we have to restart the entire numerical root search algorithm. This may become time-consuming.

## 5. Conclusion

In this paper, we elaborated on some extensions and generalizations of the traditional trinomial tree models for interest rates. We furthermore showed how the tree matching procedure can be reformulated in a much more efficient way. Our approach is based on the forward measure methodology. It is as generic as the Hull–White methodology based on forward induction and also avoids numerical root-search algorithms necessary to determine the state prices in the Hull–White model, when the interest rate is no longer normally distributed. Therefore, our approach can simplify and considerably improve current prac-

tice. Furthermore, it is robust in the sense that it can still be applied when the tree is also required to match the term structure of volatilities.

### Acknowledgements

The authors thank seminar participants at the Z-Quants Workshop of the Zürcher Kantonalbank. Furthermore, we would like to thank Menachem Brenner, Paolo Vanini, Ton Vorst, and Simon Benninga for their comments. Markus Leippold acknowledges the financial support of the Swiss National Science Foundation (NCCR FINRISK). Zvi Wiener acknowledges the financial support of the Krueger and Rosenberg funds at the Hebrew University of Jerusalem. We welcome comments, including references to related papers we inadvertently overlooked.

### Appendix. Matching by Forward Induction

Consider a trinomial tree with starting point  $(0, 0)$ . All quantities are expressed as annualized quantities. One year is partitioned into subperiods of length  $\Delta$ . To simplify the subsequent analysis, we will consider standard branching processes only and set  $h = 1$  and  $k \in \{-1, 0, 1\}$ .

Today's one-period bond price  $P^*(\Delta)$  is assumed to be known, i.e., extracted from the market quotes by some estimation procedure. Then,

$$P^*(\Delta) = e^{-r_{0,0}\Delta},$$

with  $r_{i,j}$  the annualized, continuously compounded short rate in state  $(i, j)$  prevailing over the time period  $[j\Delta, (j+1)\Delta]$ . Before considering the second time step, we introduce the concept of a state-price. The state-price is denoted by  $Q_{i,j}$ . In the following, the state-price  $Q_{i,j}$  can be thought of today's price of a security that pays exactly \$1 if state  $(i, j)$  occurs, and \$0 in every other state. Then,  $Q_{0,0} = 1$  and in the standard trinomial tree we have

$$P^*(j\Delta) = \sum_{i=-j}^j Q_{i,j}. \quad (\text{A.1})$$

Now, moving from time  $\Delta$  to  $2\Delta$ , we observe the following:<sup>10</sup>

$$P^*(2\Delta) = e^{-r_{0,0}\Delta} (\pi_1 e^{-r_{1,1}\Delta} + \pi_0 e^{-r_{0,1}\Delta} + \pi_{-1} e^{-r_{-1,1}\Delta}).$$

For large  $j$ , it would be rather cumbersome to write this up. Using the state-price formulation, we can considerably simplify the above procedure by writing the bond price as

$$P^*(2\Delta) = \sum_{i=-2}^2 Q_{i,2} = \sum_{i=-1}^1 Q_{i,1} e^{-r_{i,1}\Delta}.$$



The  $Q_{i,\Delta}$  have still to be determined. This is achieved by forward induction. We know  $Q_{0,0} = 1$  and  $r_{0,0}$ . So, for the next time-step

$$Q_{i,1} = \pi_i e^{-r_{0,0}\Delta}.$$

Generalizing the above procedure, the bond price  $P^*((j+1)\Delta)$  can be written as

$$P^*((j+1)\Delta) = \sum_{i=-j}^j Q_{i,j} e^{-r_{i,j}\Delta}.$$

This form is much more amenable for determining the level shift needed to match the term structure. Once the interest rate at time-slice  $j$  is determined by matching the tree, the state-prices for the subsequent time step can be calculated as

$$Q_{i,j+1} = \sum_m Q_{m,j} \pi_m e^{-r_{m,j}},$$

where  $m$  is determined by the paths leading to node  $(i, j+1)$ . For now, we are still lacking a piece. Recall  $r_t = g(X_t, t)$  and consider now, e.g., the function  $r_t = g(\alpha_t + X_t)$  with  $\alpha_t$  a deterministic function of time. Then,

$$P^*((j+1)\Delta) = \sum_{i=-j}^j Q_{i,j} e^{-g(\alpha_j + x_{i,j})\Delta}.$$

To determine the level shift  $\alpha_j$ , we have to invert the above relation. For illustration, consider the affine function  $r_t = \alpha_t + \beta X_t$ . Hence, we obtain

$$\alpha_j = \frac{1}{\Delta} \log \left( \sum_{i=-j}^j Q_{i,j} e^{-\beta x_{i,j}\Delta} \right) - \frac{\log P^*((j+1)\Delta)}{\Delta}. \quad (\text{A.2})$$

Note, the above procedure implicitly assumes the  $\pi$ 's to be probabilities under the risk-neutral measure. Therefore, when pricing claims using the matched tree, one should recall that the tree is spanned under the risk-neutral measure.

## Notes

1. The extended generator of the process  $X$  is defined as  $\mathcal{L} = \frac{\partial}{\partial t} + \frac{\partial}{\partial X} \mu_X(X) + \frac{1}{2} \frac{\partial^2}{\partial X^2} \sigma_X^2(X)$ , see, e.g., Arnold (1974), p. 180.
2. General quadratic term structure models have been discussed by Ahn, Dittmar, and Gallant (2002), Leippold and Wu (2002).
3. See, e.g., Jamshidian (1989), Geman, El Karoui, and Rochet (1995), Benninga, Björk, and Wiener (2002).
4. However, the parameters in the CIR square-root model can be chosen in such a way that the spot rate never reaches zero. In particular, the condition for the parameters directly derives from the parameterized version of the CIR model in Equation (5). If we consider the general CIR model with

$$dr_t = (\theta - \kappa r_t) dt + \sigma \sqrt{r_t} dW_t,$$

the condition  $2\theta > \sigma^2$  guarantees a positive interest rate that never reaches zero.

5. These calculations have been done on a standard personal computer with a Matlab implementation of the two methodologies.
6. See, e.g., Durrett (1996).
7. Typically, they are either estimated historically or determined by at-the-money interest rate derivatives such as caps or swaptions.
8. Note that the difference in the interest rate in node (0, 0) arises due to the fact that we interpret the interest rate as a simple rate, whereas Hull and White define this rate as a continuously compounded short rate.
9. The computations are based on a Mathematica implementation of the trinomial trees.
10. Since in this simple setup probabilities are not time dependent, we just simplify the notation to  $\pi_j$ .

## References

- Ahn, D.-H., R.F. Dittmar, and A.R. Gallant. (2002). "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies* 15, 243–288.
- Arnold, L. (1974). *Stochastic Differential Equations: Theory and Application*. Melbourne: Krieger.
- Baule, R. and M. Wilkens. (2004). "Lean Trees—A General Approach for Improving Performance of Lattice Models for Option Pricing," *Review of Derivative Research* 7, 53–72.
- Beaglehole, D.R. and M. Tenney. (1992). "A Nonlinear Equilibrium Model of Term Structures of Interest Rates: Corrections and Additions," *Journal of Financial Economics* 32, 345–454.
- Benninga, S., T. Björk, and Z. Wiener. (2002). "On the Use of Numeraires in Option Pricing," *Journal of Derivatives* 10, 43–58.
- Björk, T. (1998). *Arbitrage Theory in Continuous Time*. Oxford: Oxford University Press.
- Björk, T. and B.J. Christensen. (1999). "Interest Rate Dynamics and Consistent Forward Rate Curves," *Mathematical Finance* 9, 323–348.
- Black, F. (1976). "The Pricing of Commodity Contracts," *Journal of Financial Economics* 3, 167–179.
- Black, F., E. Derman, and W. Toy. (1990). "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal* 46, 33–39.
- Black, F. and P. Karasinski. (1991). "Bond and Option Pricing when Short Rates Are Lognormal," *Financial Analysts Journal* 47, 52–59.
- Brigo, D. and F. Mercurio. (2001). *Interest Rate Models—Theory and Practice*. Springer Finance, Vol. 37. Berlin: Springer.
- Cox, J.C., J.E. Ingersoll, and S.R. Ross. (1985). "A Theory of the Term Structure of Interest Rates," *Econometrica* 53, 385–408.
- Davis, M. (1998). "A Note on the Forward Measure," *Finance and Stochastics* 2, 19–28.
- Durrett, R. (1996). *Stochastic Calculus*. Probability and Stochastics Series. Boca Raton, FL: CRC Press.
- Figlewski, S. and B. Gao. (1999). "The Adaptive Mesh Model: A New Approach to Efficient Option Pricing," *Journal of Financial Economics* 53, 313–351.
- Filipović, D. (1999). "A Note on the Nelson–Siegel Family," *Mathematical Finance* 9, 349–359.
- Filipović, D. (2002). "Separable Term Structures and the Maximal Degree Problem," *Mathematical Finance* 12, 341–349.
- Geman, H., N. El Karoui and J.-C. Rochet. (1995). "Changes of Numéraire, Changes of Probability Measure and Option Pricing," *Journal of Applied Probability* 32, 77–105.
- Grant, D. and G. Vora. (2001). "An Analytical Implementation of the Hull and White Model," *Journal of Derivatives* 4, 54–60.
- Hogan, M. and K. Weintraub. (1993). "The Log-Normal Interest Rate Model and Eurodollar Futures," Working Paper, Citybank, New York.
- Hubalek, F. and W. Schachermayer. (1998). "When Does Convergence of Asset Price Processes Imply Convergence of Option Prices?" *Mathematical Finance* 8, 385–403.
- Hull, J. (2003). *Options, Futures and other Derivatives*, 5th ed. Finance Series. Upper Saddle River, NJ: Prentice Hall.

- Hull, J. and A. White. (1994). "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives* 2, 7–16.
- Hull, J. and A. White. (1996). "Using Hull–White Interest Rate Trees," *Journal of Derivatives* 3, 26–36.
- Jacod, J. and A. Shiryaev. (1987). *Limit Theorems for Stochastic Processes*. Grundlehren der Mathematischen Wissenschaften, Vol. 228. Berlin: Springer.
- Jamshidian, F. (1989). "An Exact Bond Option Formula," *Journal of Finance* 44, 205–209.
- Jamshidian, F. (1991). "Forward Induction and the Construction of Yield Curve Diffusion Models," *Journal of Fixed Income* 1, 62–74.
- Jamshidian, F. (1996). "Bond, Futures and Option Valuation in the Quadratic Interest Rate Model," *Applied Mathematical Finance* 3, 93–115.
- Kijima, M. and I. Nagayama. (1994). "Efficient Numerical Procedures for the Hull–White Extended Vasicek Model," *Journal of Financial Engineering* 5, 275–292.
- Leippold, M. and L. Wu. (2002). "Asset Pricing under the Quadratic Class," *Journal of Financial and Quantitative Analysis* 37, 271–295.
- Lesne, J.-P., J.-L. Prigent, and O. Scaillet. (2000). "Convergence of Discrete Time Option Pricing Models under Stochastic Interest Rates," *Finance and Stochastics* 4, 81–93.
- Longstaff, F.A. (1989). "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics* 23, 195–224.
- Longstaff, F.A. (1992). "Multiple Equilibria and Term Structure Models," *Journal of Financial Economics* 32, 333–344.
- Maghsoodi, Y. (1996). "Solution to the Extended CIR Term Structure and Bond Option Valuation," *Mathematical Finance* 6, 89–109.
- Pelsser, A. (1994). "An Efficient Algorithm for Calculating Prices in the Hull–White Model," Working Paper, Derivative Product Research and Development, ABN Amro Bank.
- Pliska, S.R. (1997). *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell Publishers.
- Rogers, L.C.G. (1996). "Gaussian Errors," *RISK* 9, 42–45.
- Schmidt, W. (1997). "On a General Class of One-Factor Models for the Term Structure of Interest Rates," *Finance and Stochastics* 1, 3–24.
- Tian, Y. (1999). "Pricing Complex Barrier Options under General Diffusion Processes," *Journal of Derivatives* 7, 11–30.
- Vasicek, O. A. (1977). "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* 5, 177–188.