

Efficient option valuation using trees

DAVID C. HEATH¹ and STEFANO HERZEL²

¹ Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA

² University of Perugia, Perugia, Italy

Received 19 May 2000 and in revised form 22 February 2002

An algorithm is proposed for the discrete approximation of continuous market price processes that uses trees instead of lattices. It is shown that it is convergent when used for pricing both European and American options and that it is more efficient, for some models, than the usual recombining schemes.

Keywords: option pricing, discrete-time approximations, non-recombining trees

1. Introduction

Traditionally, numerical calculation of prices of options under the Black–Scholes model (1973) (henceforth BS model) is carried out by the method of Cox *et al.* (1979). This method uses a discrete approximation of the set of possible dates and stock prices in which each state can be specified by two numbers: the ‘date’, k , of the state, and the number of times the stock price has moved ‘up’, i , between date 0 and date k , where $0 \leq i \leq k \leq n$, n being the date corresponding to expiration of the option to be valued. As time marches forward, the state of the system can evolve from (i_1, k_1) to (i_2, k_2) provided $i_1 \leq i_2$, $k_1 \leq k_2$, and $i_2 - i_1 \leq k_2 - k_1$. Such a structure is called a ‘tree’ in the finance literature, and a ‘lattice’ by mathematicians. These structures are used because they have the property that, starting from a fixed state, an ‘up’ move followed by a ‘down’ move results in the same new state as a down move followed by an up move, and in the BS model the stock price follows a Markov process, so the future probability distribution of the process depends only on the current state.

For more general situations such as HJM term structure models (Heath *et al.*, 1992), however, the state of the system cannot be specified by just two numbers; a state is specified by a ‘curve’ (the ‘term structure’) and by a date. For most models, an up movement of the term structure followed by a down movement cannot lead to the same state as a down movement followed by an up movement. Thus to build a discrete approximation for the evolution of such a system one needs to allow each sequence of ups and downs to lead to a different state. Hence, starting from a fixed term structure at date 0, there will be 2^k possible term structures at date k , and to model the system’s evolution up to date n requires $2^{n+1} - 1$ states. Such a structure is called a ‘non-recombining tree’ in the finance literature and simply a ‘tree’ by mathematicians. In this paper we shall use the mathematical names for these structures.

Because the number of nodes in trees grows exponentially with their depth, and because the usual lattice methods are commonly used with depths of 40 or more, many people have rejected tree-based methods as impractical. In this paper we shall:

1. Give a method for the construction of 'good' trees.
2. Show that, for any model, the discrete processes and the approximate valuations converge to the correct values.
3. Analyse carefully the valuation errors for lattices and good trees in valuing a European option under the BS model and show that:
 - (a) For lattice models the error decreases no faster than the reciprocal of the depth, while the work grows as the square of the depth.
 - (b) For good trees, the error decreases at least as fast as the reciprocal of the number of terminal nodes, while the work increases as the number of terminal nodes.

Since the work needed for the valuation is proportional, for both cases, to the total number of nodes, this means that getting an extra decimal place of precision requires at least a one hundred-fold increase in work for lattices, but at most a ten-fold increase for good trees.
4. Show some applications of the method, both to the BS model and to the interest rate model of Vasicek (1977).

A careful analysis of the rate of convergence for European options for several binomial approximations to the BS model has been performed by Leisen and Reimer (1996). They show that, for any fixed strike price, the error produced by most of the recombining discretization schemes for the BS model goes to zero as fast as $1/N$, when N is the number of time steps. In our analysis we do not fix the strike price and find a lower bound on the worst case error. From a theoretical point of view, relevant contributions have also been given by Amin and Khanna (1994) who studied the problem of convergence of American options prices. Duffie and Protter (1992) developed a general toolbox to check the convergence of discrete models.

Many discretization schemes have been proposed for the numerical approximation of the continuous-time model, starting from the seminal paper by Cox *et al.* (1979). For a review of the most commonly used methods and a discussion about non-recombining schemes in the HJM framework, see Chapter 4 of James and Webber (2000). Hull and White (1994) proposed a trinomial model which is among the most widely used in practice. McCarthy and Webber (2001) addressed the problem of constructing a recombining lattice in several dimensions, to approximate models with more factors. Although designed to work with three factors, their approach can also be used for a single-factor model.

The structure of the paper is as follows: Section 2 shows how to construct a good tree. Section 3 determines the convergence properties of the tree process. In Section 4, lower and upper bounds on the evaluation error for tree- and lattice-based models are determined, to make a theoretical comparison between them. Section 5 shows some numerical experiments for the BS and Vasicek models.

2. Construction of efficient trees

This section shows how to construct T^N , a discrete-time stochastic process defined on a finite probability space $(\Omega^N, \{\mathcal{F}_k\}_{k=0}^N, P^N)$. Ω^N is the set of all the sequences of +1s and -1s with length equal to N . Let ω be any such sequence. $\{\mathcal{F}_k\}_{k=0}^N$ is an increasing family of σ -fields: specifically, \mathcal{F}_k is the σ -field generated by the first k coordinates of ω . The N components of ω are independent and identically distributed; they are equal to +1 or to -1 with the same probability $1/2$. Therefore, the probability measure P^N is uniform

over Ω^N , with $P^N(\omega) = 2^{-N}$. One can think of such a probability space as one generated by N independent tosses of a fair coin.

The first step in the definition of T^N consists of ordering the set Ω^N . Let $u_N(\omega)$ be the number of +1s in ω ; let B_h^N be the subset of Ω^N such that

$$B_h^N := \{\omega \in \Omega^N \mid u_N(\omega) = h\}$$

Note that the cardinality of B_h^N is equal to $\binom{N}{h}$, therefore for any h it is possible to define a one-to-one correspondence between B_h^N and the set

$$I_h^N := \left\{ i \in \mathbf{N} \mid \sum_{k=0}^{h-1} \binom{N}{k} < i \leq \sum_{k=0}^h \binom{N}{k} \right\}$$

where \mathbf{N} is the set of natural numbers. It is easy to see that any choice of such a correspondence gives a complete ordering on the set Ω^N . Now that we have defined an ordering we can call ω_k the k th element of Ω^N . Note that the first element, ω_1 , is the sequence of all -1s, while the last one, ω_{2^N} , is the sequence of all +1s.

The second step of the construction consists of associating to each ω the ending value, $T_N^N(\omega)$, of the tree process. To do that, we partition the real line into 2^N intervals A_i^N defined as

$$A_i^N := \left[\Phi^{-1}\left(\frac{i-1}{2^N}\right), \Phi^{-1}\left(\frac{i}{2^N}\right) \right], \quad 1 \leq i \leq 2^N$$

where $\Phi(\cdot)$ is the standard normal distribution function. For each integer number i , between 1 and 2^N , let χ_i^N be any point in the interval A_i^N . Set

$$T_N^N(\omega_i) := \chi_i^N, \quad 1 \leq i \leq 2^N$$

Note that for this step there is some freedom, too. A natural choice would be the expected value of a standard normal random variable conditioned on the subset A_i^N , that is

$$\begin{aligned} \chi_i^N &:= 2^N \int_{(i-1)/2^N}^{i/2^N} \Phi^{-1}(x) dx \\ &= 2^N [\phi(\Phi^{-1}((i-1)/2^N)) - \phi(\Phi^{-1}(i/2^N))] \end{aligned} \quad (2.1)$$

where $\phi(\cdot)$ is the density function of the standard normal distribution.

These results will hold for any particular construction of the final value of the tree process provided that $ET_N^N = 0$ and the fourth moment of T_N^N is uniformly bounded; henceforth it will be assumed that there is a positive real number c such that, for all N ,

$$E(T_N^N)^4 \leq c \quad (2.2)$$

When the 'natural choice' (2.1) is made, we have $c = 3$.

The third, and last, step consists in defining the random variables T_k^N , for $0 \leq k < N$. This is done by requiring the process T^N to be a martingale, that is

$$T_k^N := E[T_N^N | \mathcal{F}_k]$$

Note that $T_0^N = 0$ and that T^N associates to each ω a unique path along a binary tree of depth N .

3. Convergence

In this section it is proved that the tree processes converge to the Brownian motion. To do that it is shown that they are ‘close’ to the one-dimensional random walks.

Consider the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_{k=0}^\infty, P)$ of infinite sequences of -1 s and $+1$ s. Assume that $\mathcal{F} = \mathcal{F}_\infty$. For any $\omega \in \Omega$, let ω^N be the sequence of the first N components of ω . Then we can define (with a small abuse of notation) $T^N(\omega)$ as $T^N(\omega^N)$.

Let $S_n(\omega)$ be the sum of the first n components of ω . Indicate with L^N the discrete-time stochastic process such that

$$L_k^N := \frac{S_k}{\sqrt{N}}, \quad 0 \leq k \leq N$$

Observe that L^N is a symmetric, one-dimensional, random walk (stopped after N steps); it is called ‘the lattice process’.

This first result shows that the variance of the distance of the terminal values of the tree and the lattice goes to zero as the number of time periods goes to infinity.

Theorem 3.1. Let L_N^N and T_N^N be the terminal values for the N -step lattice and tree processes. Then

$$\lim_{N \rightarrow \infty} E(L_N^N - T_N^N)^2 = 0$$

Proof. In the appendix. ■

By using the fact that $L^N - T^N$ is a martingale one obtains the following result on the distance of the two processes in the intermediate steps.

Theorem 3.2. We have:

$$\lim_{N \rightarrow \infty} E\left(\max_{1 \leq k \leq N} |L_k^N - T_k^N|\right)^2 = 0 \quad (3.1)$$

Proof. Doob’s martingale maximal inequality (Durrett, 1991, p. 216) states that, for all N ,

$$E\left(\max_{1 \leq k \leq N} |L_k^N - T_k^N|\right)^2 \leq 4E(L_N^N - T_N^N)^2$$

and the result follows from Theorem 3.1. ■

To complete this analysis of the tree process, we want to show that a continuous-time version of it weakly converges to Brownian motion. From now on we denote by h the length of time intervals, that is

$$h := 1/N$$

We use a superscript N for discrete-time processes and a superscript h for their continuous-time counterparts.

Define the continuous-time processes T^h, L^h as follows:

$$T_t^h := T_{\lfloor Nt \rfloor}^N, \quad 0 \leq t \leq 1$$

and

$$L_t^h := L_{\lfloor Nt \rfloor}^N, \quad 0 \leq t \leq 1$$

Of course, we can think of T^h and L^h as random variables from Ω to the set of right continuous with left limits (RCLL) functions on the time interval $[0,1]$. This set is called ID and we consider it endowed with the Skorohod topology; it then follows that it is a complete and separable metric space (Billingsley, 1971, Chapter 3).

Now we can show

Theorem 3.3. T^h weakly converges to a one-dimensional Brownian motion W .

Proof. From Donkser's Theorem (Durrett, 1991, Theorem 7.6.6), we know that L^h converges to W weakly (as h goes to zero). Moreover, from the construction of the continuous processes and Doob's inequality,

$$\begin{aligned} E\left(\sup_{0 \leq t \leq T} |L_t^h - T_t^h|\right)^2 &= E\left(\max_{0 \leq k \leq N} |L_k^N - T_k^N|\right)^2 \\ &\leq 4E|L_N^N - T_N^N|^2 \end{aligned}$$

From Chebyshev's inequality it follows that, for any $\delta > 0$,

$$\begin{aligned} P\left(\sup_{0 \leq t \leq T} |L_t^h - T_t^h| > \delta\right) &= P\left(\max_{0 \leq k \leq N} |L_k^N - T_k^N| > \delta\right) \\ &\leq E\left[\max_{0 \leq k \leq N} |L_k^N - T_k^N|\right]^2 / \delta^2 \end{aligned}$$

Therefore, from Theorem 3.2, the distance (in the norm of the sup) of the two processes L^h and T^h converges to 0 in probability. Then (Billingsley, 1971, Theorem 3.1) the tree processes T^h also converge to W weakly, as h goes to zero. ■

Now we can show how to use the tree to compute an approximation of the prices of contingent claims for a given continuous-time model. We will show that the approximated prices converge to the exact ones for both European and American contingent claims.

Consider an underlying risky asset X whose dynamics (under an equivalent martingale measure) are given by

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad (3.2)$$

with initial value a given real number X_0 . Assume that the functions $\mu(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are regular enough for X to be the unique, strong solution of (3.2).

We approximate X from time 0 to time 1 in two different ways with the tree and the lattice processes. The usual Euler approximation is employed. Of course, other schemes, with higher order, can also be used. The discrete-time tree approximation $X^{T,N}$ is defined by

$$X_{n+1}^{T,N} = X_n^{T,N} + \mu(t_n, X_n^{T,N})1/N + \sigma(t_n, X_n^{T,N})\Delta T_n^N \quad (3.3)$$

$$X_0^{T,N} = X_0 \quad (3.4)$$

where $n = 1, \dots, N - 1$,

$$t_n := nh$$

and

$$\Delta T_n^N := T_{n+1}^N - T_n^N$$

Analogously, using the lattice process,

$$X_{n+1}^{L,N} = X_n^{L,N} + \mu(t_n, X_n^{L,N})1/N + \sigma(t_n, X_n^{L,N})\Delta L_n^N \quad (3.5)$$

$$X_0^{L,N} = X_0 \quad (3.6)$$

where

$$\Delta L_n^N := L_{n+1}^N - L_n^N$$

The continuous-time versions with RCLL paths of these discrete processes are then defined as

$$X_t^{T,h} := X_{\lfloor Nt \rfloor}^{T,N}, \quad 0 \leq t \leq 1$$

and

$$X_t^{L,h} := X_{\lfloor Nt \rfloor}^{L,N}, \quad 0 \leq t \leq 1$$

Theorem 3.4. Both $X^{L,h}$ and $X^{T,h}$ weakly converge to X , the strong solution of (3.2).

Proof. To prove this result we use Theorem 4.4 of Duffie and Protter (1992).

From the definitions it follows that $X^{L,h}$ and $X^{T,h}$ solve, respectively, the following equations

$$dX_t^{T,h} = \mu(t, X_{t-}^{T,h})df_t^h + \sigma(t, X_{t-}^{T,h})dT_t^h \quad (3.7)$$

and

$$dX_t^{L,h} = \mu(t, X_{t-}^{L,h})df_t^h + \sigma(t, X_{t-}^{L,h})dL_t^h \quad (3.8)$$

where

$$f_t^h := h \lfloor t/h \rfloor$$

The sequences of semimartingales

$$Y_t^{T,h} := f_t^h + T_t^h$$

$$Y_t^{L,h} := f_t^h + L_t^h$$

converge to $t + W_t$ weakly.

Since $E|L_1^h|$ and $E|T_1^h|$ are uniformly bounded (with respect to h), Condition A of Duffie and Protter (1992) is verified and the result then follows from their Theorem 4.4. ■

Since $X^{L,h}$ and $X^{T,h}$ weakly converge to X , we can apply the results of Amin and Khanna (1994) to prove convergence for European and American contingent claims.

4. Error analysis

This section compares upper and lower bounds on the worst case error estimates of the tree and the lattice methodologies. We start by estimating a lower bound on the approximating error for discrete methods

with N final states, then relate such a lower bound to the number of steps necessary to attain the given number of states. The idea is that with an efficient tree fewer steps are needed to get a set of final nodes with a satisfactory approximation. We compute an upper bound on the approximating error for the tree and show that the total number of nodes necessary to get a given accuracy grows much slower than the corresponding one for the lattice.

Let us first determine a lower bound on the approximating error for a discretization with N final nodes. We study the case of European Put options, although the same analysis can be applied to many other derivatives.

Let $P(K)$ be the exact price at time 0 of a European Put option written on an underlying S , with maturity T and strike K . Let us suppose that S_T has a (risk-neutral) density $v(\cdot)$ with respect to Lebesgue measure. Hence

$$P(K) = \int_{-\infty}^{+\infty} B_T(K-x)^+ v(x) dx$$

where B_T is the discount factor at time T . The function $P(K)$ is twice differentiable, with

$$\frac{d^2 P(K)}{dK^2} = v(K)$$

Let S^N be a discrete approximation of S_T with N terminal nodes. Denote S_i^N the i th terminal node, with $S_i^N > S_{i-1}^N$, for $i = 1, \dots, N$ and call I_i^N the interval (S_{i-1}^N, S_i^N) . Suppose that the approximation S_N has been obtained through a discretization of a continuous-time process, by using a recombining lattice. It is well known that, to get convergence, when the length of the time discretization interval is of the order of h , the length of the space discretization $|I_i^N| = S_i^N - S_{i-1}^N$ has to be of the order of \sqrt{h} . For a recombining lattice, the number of steps needed to get N final states is of the order of N , hence the time step h is of the order of $1/N$. Therefore, to get convergence, $|I_i^N|$ has to be of the order of $\sqrt{1/N}$.

Note that $P_N^L(K)$, the approximation to $P(K)$ produced by the lattice, is a piecewise-linear function, with breaks at S_i^N . Let us define the approximation error as

$$e_N(K) := P(K) - P_N(K)$$

A standard argument of real analysis (linear approximation of convex functions) gives a lower bound on $e_N(K)$:

$$\sup_{K \in I_i^N} |e_N(K)| \geq \frac{v_i^N |I_i^N|^2}{16}$$

where $v_i^N := \inf_{K \in I_i^N} v(K)$.

Now, let (K_1, K_2) be an interval where the density $v(\cdot)$ is strictly positive; since $|I_i^N|$ has to be of the order of $\sqrt{1/N}$,

$$\sup_{K \in I_i^N(K_1, K_2)} |e_N(K)| \geq \frac{C}{N} \tag{4.1}$$

where C is a constant that depends on the discretization scheme, on the density $v(\cdot)$ and on the interval (K_1, K_2) . In particular, for the classical binomial approximation of the BS model,

$$C = \frac{\bar{v}\sigma^2 K_1^2}{4}$$

where \bar{v} is the inf of $v(\cdot)$ on (K_1, K_2) .

Leisen and Reimer (1996) analyse three recombining schemes and show that the errors for fixed strikes converge as fast as $1/N$. Formula (4.1) states that, for any converging and recombining scheme, the worst case error for any $K \in (K_1, K_2)$ cannot decrease faster than $1/N$.

The work needed to compute the no arbitrage price for a derivative contract is proportional to $\mathcal{W}^L(N)$, the total number of nodes in the lattice, which is of the order of N^2 . Hence we can write

$$\sup_{K \in (K_1, K_2)} |e_N(K)| \geq \frac{\text{Const.}}{\sqrt{\mathcal{W}^L(N)}} \quad (4.2)$$

Equation (4.2) shows that, to get one more correct digit with a recombining scheme, one has to work one hundred times as much.

Lattice methods oversample the tails of the distribution. To avoid this and save computational work a common approach is 'pruning the tree', that is to cut off all the nodes displaced far away from the mean. Avellaneda and Laurence (1999) show that in this way the total number of nodes can be reduced to be of the order of $N^{3/2}$, without significantly increasing the error. In this case the lower bound on the worst case error would be of the order of $\mathcal{W}^{\bar{L}}(N)^{-3/2}$, where \bar{L} is the trimmed lattice.

For a non-recombining tree, only $\log_2 N$ steps are needed to get N terminal nodes, with a total number of nodes equal to $2N - 1$. Therefore, it might be expected that, if the approximation is good, a tree with fewer nodes can give better results than a lattice. The next result will be used to find an upper bound on the worst case error produced by efficient trees.

Proposition 4.1. Let $f(x)$ be a bounded function, with finite total variation and let Z be a standard normal random variable. Then

$$|Ef(Z) - Ef(T_N^N)| \leq C2^{-N}$$

where C is a positive constant which depends on f .

Proof. We have

$$\begin{aligned} |Ef(Z) - Ef(T_N^N)| &= \left| \int_0^1 f(\Phi^{-1}(x)) dx - \frac{1}{2^N} \sum_{k=1}^{2^N} f(x_k^N) \right| \\ &= \left| \frac{1}{2^N} \sum_{k=1}^{2^N} f(\Phi^{-1}(x_k)) - f(\Phi^{-1}(y_k)) \right| \end{aligned}$$

where x_k and y_k belong to the interval $(2^{-N}(k-1), 2^{-N}k)$. Since Φ^{-1} is continuous and monotone on $(0, 1)$, from the hypothesis on f it follows that $f(\Phi^{-1}(x))$ is a function of finite total variation on $(0, 1)$. The result then follows. ■

Let $\psi(S_T)$ be the discounted payoff function of a derivative security depending on the value at time T of an underlying $S_T = g(W_T)$, where W is Brownian motion. Using the notation of Proposition 4.1 take

$$f(Z) := \psi(S_T) = \psi(g(\sqrt{T}Z))$$

where the equality is in distribution. Assume that $g(\cdot)$ is regular enough so that $f(\cdot)$ has bounded variation. The exact value of the derivative security is

$$V := E\psi g(X_T) = Ef(Z)$$

Suppose that an N -step tree process is used to approximate the Brownian motion W . The approximated value of the derivative is

$$V^{T,N} := E\psi(g(S_T^{T,h})) = Ef(T_N^N)$$

Hence, from Proposition 4.1, it follows that

$$|V^{T,N} - V| \leq C2^{-N} \quad (4.3)$$

In particular, when the derivative security is a European Put with strike price K and the model is BS,

$$|V^{T,N} - V| \leq e^{-rT} K 2^{-N}$$

Denote $\mathcal{W}^T(N)$ the number of nodes for a tree of depth N , that is

$$\mathcal{W}^T(N) = 2^{N+1} - 1$$

In terms of the number of nodes, Equation (4.3) may be written as

$$|V^{T,N} - V| \leq \frac{\text{Const.}}{\mathcal{W}^T(N)} \quad (4.4)$$

Equation (4.4) shows that the worst case error for the tree approximation decreases at least with the same order as the inverse of the number of nodes. A comparison of this result to the analogue for the lattice algorithm (4.2) shows that, when worst case errors are compared, the tree method is more efficient than the lattice. In fact we have observed that while for the lattice one more digit costs at least one hundred times as much work, for the tree it costs at most ten times as much. A first illustration of the results of this section, applied to the BS model, is given in Figure 1. It shows that the tree needs fewer nodes to reach a given level of precision. More numerical examples are shown in the next section.

5. Applications

This section shows some applications of the proposed methodology to the BS model and to the Vasicek model.

We start by comparing the performances of the tree and the lattice model when approximating the Brownian motion W for the BS price process $S = g(W)$. Figure 2 shows the error as a function of the number of nodes for a European Put option with strike price $K = 44$ and maturity $T = 1$. The remaining relevant parameters are $S_0 = 40$, $\sigma = 0.3$, $r = 0.03$. It confirms that the performances of the tree are generally better than those of the lattice, not only for the worst case error on a set of options, but also for a single option.

As a second example, the Euler scheme is applied to the BS model with the same parameters as above and compare the performances of the lattice and the tree for a fixed number of nodes when the strike price varies. More precisely, 12 steps of the tree process were performed, therefore generating $2^{13} - 1 = 8191$ nodes and 2^7 steps of the lattice, for a total number of 8256 nodes. The results are displayed in Figure 3. Note the characteristic parabolic behaviour of the lattice between two consecutive final nodes, while the tree is always very close to the exact price, consistently outperforming the lattice.

Let us move on to the Vasicek interest rate model. The stochastic differential equation for the short rate r_t is given by

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \quad (5.1)$$

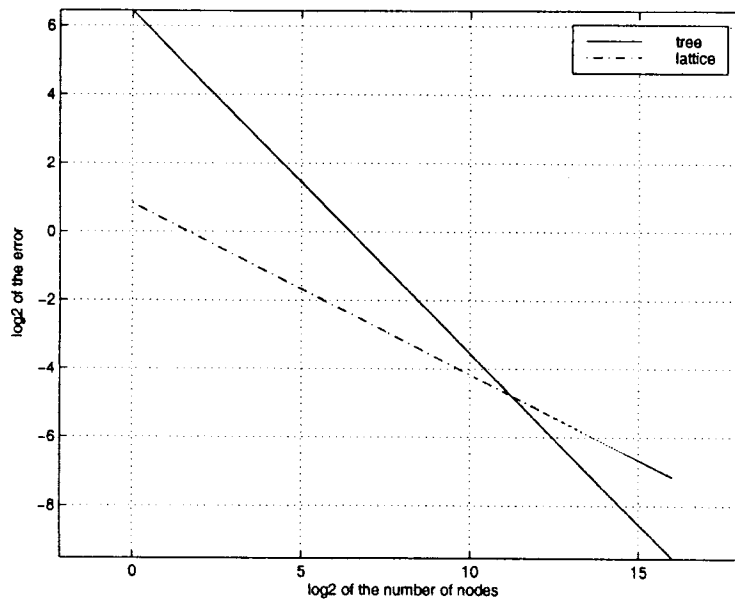


Fig. 1. The \log_2 of the bounds on the worst case errors as a function of the \log_2 of the number of nodes. The dotted line is a lower bound on the worst case error of the lattice as estimated by formula (4.2). The continuous line is an upper bound on the worst case error of the tree as estimated by formula (4.3). We consider European Put options with maturity $T = 1$ and strike prices ranging between $K_1 = 39$ and $K_2 = 48$. The remaining relevant parameters are $S_0 = 40$, $\sigma = 0.3$, $r = 0.03$.

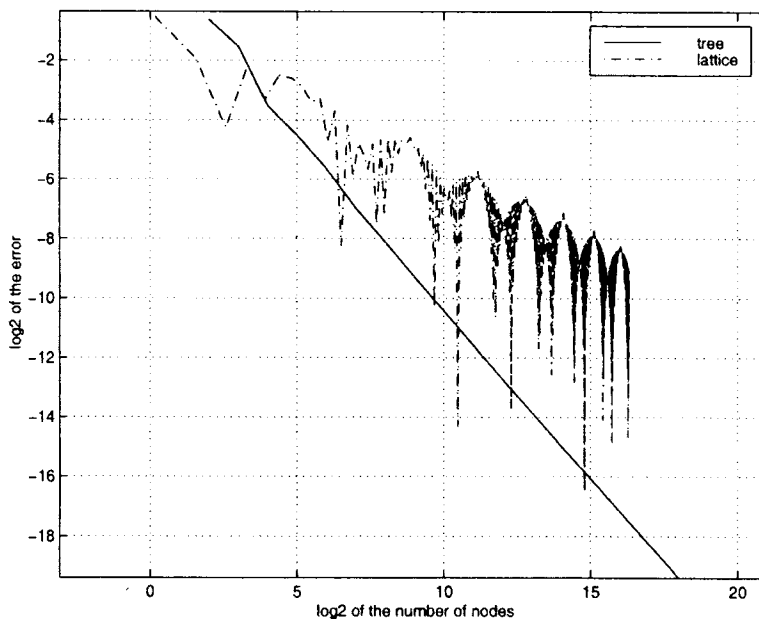


Fig. 2. The \log_2 of the absolute values of the difference between the approximation and the exact BS price as a function of the \log_2 of the number of nodes. The tree and the lattice have been used to get an approximation to the Brownian motion W in the BS price process $S = g(W)$. We consider a European Put option with strike price $K = 44$ and maturity $T = 1$. The remaining relevant parameters are $S_0 = 40$, $\sigma = 0.3$, $r = 0.03$.

where α , μ and σ are constant. The short rate r_T at time T is normally distributed, with mean and variance depending on the parameter of the model and on the initial rate r_0 . This model has closed-form expressions for bonds and options prices, therefore it is well suited for testing approximating algorithms. The Euler scheme is used to approximate Equation (5.1), that is

$$r_{n+1} = \alpha(\mu - r_n)\Delta t + \sigma\delta W_n \quad (5.2)$$

where $\Delta t := T/N$ and δW_n is the increment of the particular algorithm used. Note that in this case, even if the lattice is recombining, the discrete process determined by the Euler approximation is not, therefore the nodes of the tree approximation are as many as those produced by the binomial lattice.

For this experiment we compare the tree, the binomial lattice, the trinomial lattice, and the icosahedral lattice. The trinomial lattice is a popular method for approximating interest rate models proposed by Hull and White (1994). The icosahedral lattice was developed by McCarthy and Webber (2001) to approximate three factors models, however, we use its one-dimensional projection because it has been proved to perform well for this case.

For the trinomial lattice we have

$$\delta W_n := \begin{cases} C_1(\Delta t) & p = (1 - p_0)/2 \\ 0 & p = p_0 \\ -C_1(\Delta t) & p = (1 - p_0)/2 \end{cases} \quad (5.3)$$

where $p_0 = 2/3$ and $C_1(\Delta t) = \sqrt{\Delta t/(1 - p_0)}$.

The icosahedral lattice is defined as

$$\delta W_n := \begin{cases} \tau C_2(\Delta t) & p = (1 - p_0)/4 \\ C_2(\Delta t) & p = (1 - p_0)/4 \\ 0 & p = p_0 \\ -C_2(\Delta t) & p = (1 - p_0)/4 \\ -\tau C_2(\Delta t) & p = (1 - p_0)/4 \end{cases} \quad (5.4)$$

where $p_0 = 1/13$, $\tau = 2/(\sqrt{5} - 1)$, $C_2(\Delta t) = \sqrt{3\Delta t/[(1 - p_0)(1 + \tau^2)]}$.

We set $r_0 = 0.1$, $\alpha = 0.05$, $\mu = 0.1$, $\sigma = 0.01$, and performed $N = 10$ steps for each scheme, obtaining four approximations for the distribution of the short rate at time $T = 1$. A comparison of such distributions to the exact one through the quantile–quantile plots is displayed in Figure 4. Note the clustering effect of the binomial and trinomial lattices. As already observed by McCarthy and Webber (2001), this is due to the fact that these two methods produce a process that is not exactly, but ‘almost’ recombining. Also note that the final numbers of nodes are $2N$ for the tree and the binomial lattice, $3N$ for the trinomial and $5N$ for the icosahedral: this is the reason why the last two plots are more spread. It appears that the tree method produces the best fit of the short-rate distribution at time T . Other numerical experiments, not shown, confirm such behaviour.

In a second experiment, the approximations produced by the different methods were compared when used to price a zero coupon bond with maturity of 3 years. Figure 5 shows the errors of the approximations for different numbers of steps. The parameters used to produce the picture are the same as above. In this case also, efficient trees produce better results than the other methods.

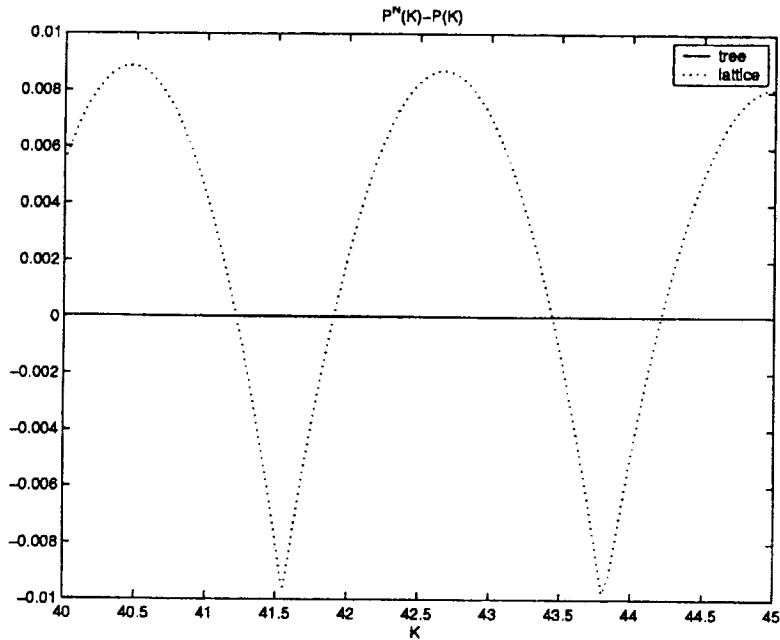


Fig. 3. Difference between the approximation and the exact (BS) price as a function of the strike price. The tree and the lattice are used in the Euler discretization of the Black–Scholes stochastic differential equation. A tree with 12 steps (i.e. 8191 nodes) and a lattice with 2^7 steps (i.e. 8256 nodes) were used. We consider European Put options with maturity $T = 1$. The remaining relevant parameters are $S_0 = 40$, $\sigma = 0.3$, $r = 0.03$

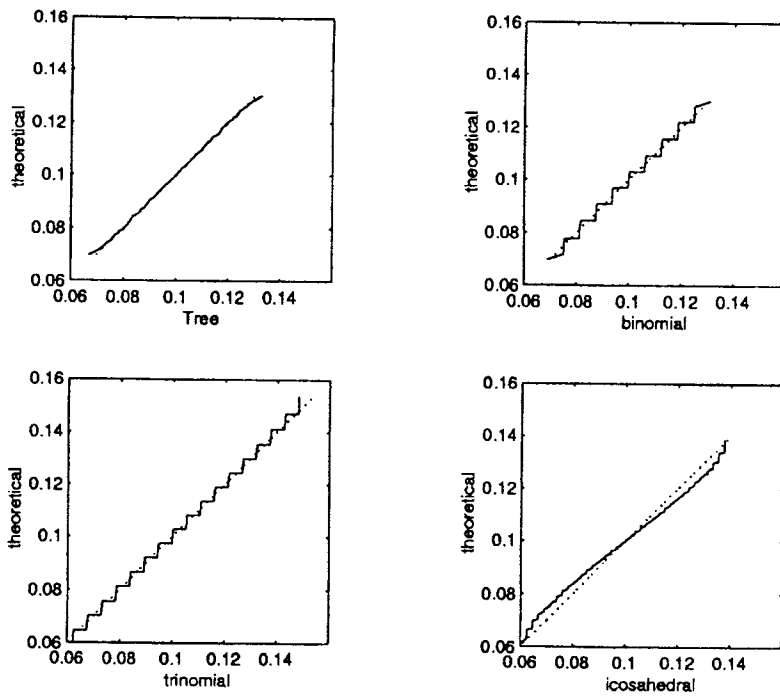


Fig. 4. Quantile–quantile plots comparing the approximated distributions for r_T with the exact one. The approximations were obtained by applying the Euler method to the Vasicek model with parameters $r_0 = 0.1$, $\alpha = 0.05$, $\mu = 0.1$, $\sigma = 0.01$, and taking $N = 10$ steps.

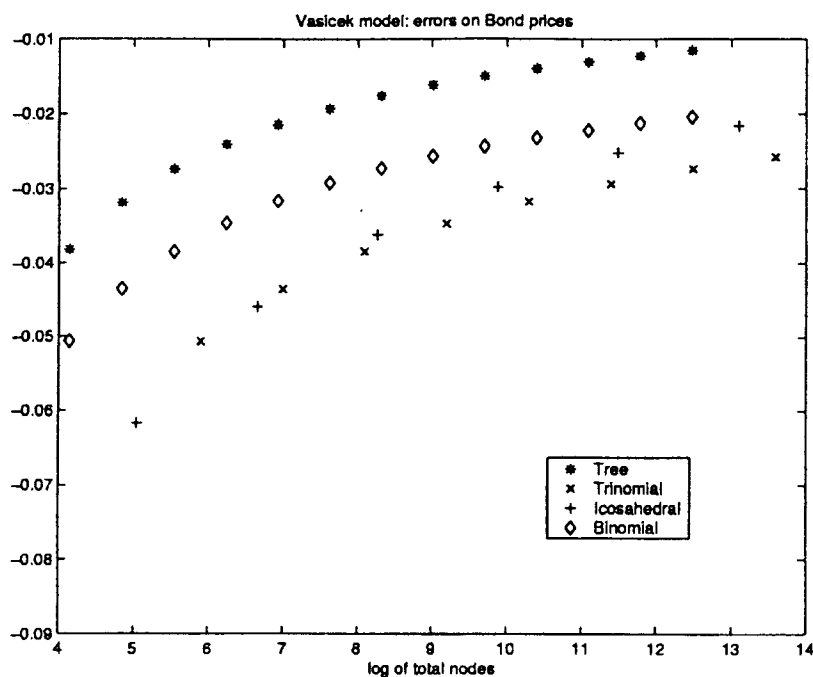


Fig. 5. Pricing errors on a zero-coupon bond of maturity $T = 3$ for different models and increasing number of nodes.

References

- Amin, K. and Khanna, A. (1994) Convergence of American option values from discrete- to continuous-time financial models, *Mathematical Finance*, **4**, 289–304.
- Avellaneda, M. and Laurence, P. (1999) *Quantitative Modelling of Derivatives Securities: From Theory to Practice*, CRC Press.
- Billingsley, P. (1968) *Convergence of Probability Measures*, John Wiley.
- Billingsley, P. (1971) *Weak Convergence of Measures: Applications in Probability*, SIAM.
- Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**, 637–54.
- Cox, J.C., Ross, S.A. and Rubinstein, M. (1979) Option pricing: a simplified approach, *Journal of Financial Economics*, **7**, 229–63.
- Durrett, R. (1991) *Probability: Theory and Examples*, Wadsworth.
- Duffie, D. and Protter, P. (1992) From discrete to continuous time finance: weak convergence of the financial gains processes, *Mathematical Finance*, **2**, 1–15.
- Heath, D.C., Jarrow, R. and Morton, A. (1992) Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation, *Econometrica*, **60**, 77–105.
- Hull, J.C. and White, A.D. (1994) Numerical procedures for implementing term structure models I: Single-factors models, *Journal of Derivatives*, Fall, 7–16.
- James, J. and Webber, N. (2000) *Interest Rate Modelling*, John Wiley.
- Leisen, D. and Reimer, M. (1996) Binomial models for option valuation - examining and improving convergence, *Applied Mathematical Finance*, **3**, 319–46.

McCarthy, L.A. and Webber, N.J. (2001) Pricing in three-factor models using icosahedral lattices, *Journal of Computational Finance*, **5**, 1–36.

Vasicek, O. (1977) An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5**, 177–88.

Appendix: Proof of Theorem 3.1

Before proceeding to the proof let us recall some useful results.

Let $F_N(x)$ be the cumulative distribution function of S_N . Note that, with probability one,

$$\Phi^{-1}(F_N(S_N - 2)) \leq T_N^N \leq \Phi^{-1}(F_N(S_N)) \quad (\text{A.1})$$

To prove the last inequality let us observe that, for any $j = 1, \dots, 2^N$,

$$\Phi(T_N^N(\omega_j)) = \Phi(\chi_j^N)$$

and, by construction,

$$2^{-N} \sum_{k=0}^{u_N(\omega_j)-1} \binom{N}{k} \leq \Phi(\chi_j^N) \leq 2^{-N} \sum_{k=0}^{u_N(\omega_j)} \binom{N}{k}$$

where $u_N(\omega_j)$ denotes the number of +1s in ω_j . Relation (A.1) then follows.

For the proof of the theorem, we also make use of the following theorems:

- The Berry–Esseen Theorem (Durrett, 1991, p. 108)

$$|F_N(x) - \Phi(x/\sqrt{N})| \leq 3/\sqrt{N} \quad (\text{A.2})$$

- The Local Central Limit Theorem (Durrett, 1991, p. 113)

$$p_N(S_N(\omega)) = \frac{2}{\sqrt{2N\pi}} \exp\left(-\frac{S_N(\omega)^2}{2N}\right) + o\left(\frac{1}{\sqrt{N}}\right) \quad N \rightarrow \infty \quad (\text{A.3})$$

uniformly with respect to ω , when $p_N(x) := P(S_N = x)$.

Now we can proceed to the proof.

Let M be a positive real number, then

$$\begin{aligned} E(L_N^N - T_N^N)^2 &= E(L_N^N - T_N^N)^2 \mathbf{1}_{\{|S_N/\sqrt{N}| < M\}} \\ &\quad + E(L_N^N - T_N^N)^2 \mathbf{1}_{\{|S_N/\sqrt{N}| \geq M\}} \end{aligned}$$

where $\mathbf{1}_A$ is the indicator function of the event A .

We show that, for any $M > 0$,

$$\lim_{N \rightarrow \infty} E(L_N^N - T_N^N)^2 \mathbf{1}_{\{|S_N/\sqrt{N}| < M\}} = 0 \quad (\text{A.4})$$

and that

$$E(L_N^N - T_N^N)^2 \mathbf{1}_{\{|S_N/\sqrt{N}| \geq M\}} \leq \frac{K}{M} \quad (\text{A.5})$$

where K is a constant independent of N .

Let us begin with (A.4). Observe that

$$\begin{aligned} (L_N^N - T_N^N)^2 &\leq 2(L_N^N - \Phi^{-1}(F_N(S_N)))^2 + 2(T_N^N - \Phi^{-1}(F_N(S_N)))^2 \\ &\leq 2(L_N^N - \Phi^{-1}(F_N(S_N)))^2 + 2(\Phi^{-1}(F_N(S_N - 2)) - \Phi^{-1}(F_N(S_N)))^2 \end{aligned}$$

We now determine an upper bound on the two terms in the last inequality, starting with the first one:

$$\begin{aligned} [L_N^N - \Phi^{-1}(F_N(S_N))] &= [\Phi^{-1}(\Phi(S_N/\sqrt{N})) - \Phi^{-1}(F_N(S_N))] \\ &= [\Phi(S_N/\sqrt{N}) - F_N(S_N)]\sqrt{2\pi} \exp(\Phi^{-1}(\xi_1)^2/2) \end{aligned}$$

where

$$\min(\Phi(S_N/\sqrt{N}), F_N(S_N)) \leq \xi_1 \leq \max(\Phi(S_N/\sqrt{N}), F_N(S_N))$$

From the Berry–Esseen Theorem (A.2) we get (a.s. and for all N)

$$\Phi(S_N/\sqrt{N}) - 3/\sqrt{N} \leq \xi_1 \leq \Phi(S_N/\sqrt{N}) + 3/\sqrt{N}$$

Therefore, for those ω s such that $|S_N/\sqrt{N}| \leq M$,

$$\Phi(-M) - 3/\sqrt{N} \leq \xi_1 \leq \Phi(M) + 3/\sqrt{N}$$

and, for any $N > \bar{N}$,

$$\Phi(-2M) \leq \xi_1 \leq \Phi(2M)$$

Hence,

$$\begin{aligned} [L_N^N - \Phi^{-1}(F_N(S_N))]^2 \mathbf{1}_{\{|S_N/\sqrt{N}| < M\}} &\leq [\Phi(S_N/\sqrt{N}) - F_N(S_N)]^2 2\pi \exp(4M^2) \\ &\leq 18\pi \exp(4M^2)/N \end{aligned}$$

therefore this part goes to zero a.s. as N goes to infinity, for any M . Since it is also uniformly bounded, its mean also converges to zero.

As for the second term, we have

$$[\Phi^{-1}(F_N(S_N - 2)) - \Phi^{-1}(F_N(S_N))] = p_N(S_N)\sqrt{2\pi} \exp(\Phi^{-1}(\xi_2)^2/2)$$

where

$$F_N(S_N - 2) \leq \xi_2 \leq F_N(S_N)$$

and $p_N(x) := P(S_N = x)$. From the same argument as above we have, for N ‘large enough’,

$$\Phi(-2M) \leq \xi_2 \leq \Phi(2M)$$

Hence, from the Local Central Limit Theorem (A.3), we get

$$2p_N(S_N)^2 2\pi \exp(\Phi^{-1}(\xi_2)^2) \mathbf{1}_{\{|S_N/\sqrt{N}| \geq M\}} \leq 4\pi \exp(4M^2) 3/(\pi N)$$

whose mean also goes to zero from the Bounded Convergence Theorem.

From the Cauchy–Schwartz inequality,

$$E(L_N^N - T_N^N)^2 \mathbf{1}_{\{|S_N/\sqrt{N}| \geq M\}} \leq [E(L_N^N - T_N^N)^4]^{1/2} P(|S_N/\sqrt{N}| \geq M)^{1/2}$$

The fourth moments of L_N^N and T_N^N (by assumption (2.2)) are uniformly bounded, then the same is true for that of their difference.

Moreover, from Chebyshev inequality:

$$P(|S_N/\sqrt{N}| \geq M) \leq \frac{E(S_N/\sqrt{N})^2}{M^2} = \frac{1}{M^2}$$

and (A.5) then follows, concluding the argument.