Fields Institute Communications
Volume 00, 0000

Embedding and the convergence of the binomial and trinomial tree schemes

John B. Walsh
Department of Mathematics
University of British Columbia
Vancouver, B.C. Canada

Owen D. Walsh
FinancialCAD
207-7565 132 St.
Surrey, B.C. Canada

Abstract. This paper studies the regularities and irregularities of the convergence of the binomial and trinomial tree schemes. It is shown that these and other schemes can be embedded in the Black-Scholes model by Skorokhod embedding. This method is used to study their detailed behavior. Numerical examples illustrate many of the conclusions.

Draft, February 27, 2002

1 Introduction

The Black-Scholes model of a simple financial market [**1**] takes the stock price S_t to be a logarithmic Brownian motion. Under the martingale measure, the discounted stock price $e^{-rt}S_t$ is a martingale, so that one can write $S_t = e^{\sigma W_t + (r-\frac{1}{2}\sigma^2)t}$, where σ is the stock volatility, W_t is a standard Brownian motion, and r is the (constant) interest rate. Black and Scholes' hedging argument then shows that an option or other financial derivative which pays $f(S_T)$ at its maturity date T has the value $V(S_0, 0)$ at time zero, where

$$V(S_0, 0) \stackrel{\text{def}}{=} e^{-rT}E\{f(S_T)\} \tag{1}$$

and the expectation is taken under the martingale measure.

The binomial tree, or Cox-Ross-Rubinstein model [**2**], and the trinomial tree models are two—among many—natural discrete versions of the Black-Scholes model. In the binomial tree, the stock price is a Markov chain, representing the actual price at times $t = 0,\ \delta,\ 2\delta, \ldots$ for some $\delta > 0$, and, at each time, it has two possibilities: it can either go up by a factor, or it can go down by another factor. The trinomial tree offers a third choice at each stage: the stock can also stay at the same value.

The same hedging argument applies here: to compute option prices in these models, the discounted stock price should be a martingale. In both models, if one lets δ tend to zero and adjusts the parameters correctly, the models will converge to the Black-Scholes model, and the option prices they calculate will converge to the Black-Scholes prices.

This article has two aims: the first is to look closely at the convergence of these option prices. The second is to call attention to the fact that one can unify these, and many other models as well, by embedding them in the Black-Scholes model by a procedure called Skorokhod embedding [**11**].

It is possible to look at the relations between the above models either probabilistically: the Markov chain is an approximation to the logarithmic Brownian motion, or analytically: the discrete model leads to a difference scheme which is a numerical method for the solution of the Black-Scholes partial differential equation. The interesting aspect of the financial problems from the analytic viewpoint is that the data is usually not smooth. For a call option of strike price K, for instance, the function f above is $f(s) = (s - K)^+$, whose derivative is discontinuous at K. This affects the convergence rather dramatically. Our aim is to use purely probabilistic considerations to develop some results on the convergence of the numerical schemes. In the process, we will show how to unify a large number of different models under the Black-Scholes banner.

This hedging argument does not depend strongly on the fact that the stock price is a logarithmic Brownian motion, but only on the fact that the market is complete, so it applies quite generally. In fact, the discounted stock price in any other complete-market models should also be a martingale, at least for the purposes of pricing options.

This is true in particular for the binomial tree, since the binomial market is complete. It is not quite true for the trinomial tree, however: the market is not complete for any strictly positive δ, but it is complete in the limit as δ tends to zero. In any case, even in incomplete markets it is common practice to assume the discounted stock price is a martingale, and use this as a first approximation to price options.

It is known [**7**] [**8**] that for standard call and put options, the rate of convergence is of order one: that is, the error is asymptotically $O(1/n)$, where n is the number of time steps. We will look closely at this question, and see that for one version of the binomial tree at least, it is possible to calculate the exact constants of the first-order convergence for a large set of European-type derivatives. This includes European calls and puts, but we will also treat more general payoff functions, including those with discontinuities, such as in binary options.

The general flavor of the convergence of the binomial and trinomial schemes depends strongly on the smoothness properties of the payoff function f. If f is continuously differentiable, the error in an n-step scheme is asymptotically C/n. If f' has discontinuities in the derivative—as with a call or put option—there is still order one convergence, but it is convergence "with a wobble": there are constants $\alpha < \beta$ such that the error is asymptotically between α/n and β/n. More exactly, the error is a quasi-periodic function of the number of time-steps. It converges to zero at the overall rate of $1/n$, but it does so quasi-periodically. Roughly the same applies to the case where f itself has jump discontinuities, except there, the order of convergence is $1/2$ unless all the discontinuities are at lattice points and $f(x) = (f(x+) + f(x-))/2$ at each discontinuity. If the discontinuities are at the lattice points, the convergence is of order one. Moreover, for continuous f, such as

European options, the derivative, or delta, which is used in the hedging strategy, also converges with order one.

The detailed proofs for the binomial tree will appear elsewhere [**11**], so we will just outline them here. We will prove some of the results for the trinomial tree, and we will give a number of numerical examples and some tentative conclusions on the two schemes in section 4.

2 Embeddings

It is a general fact [**9**] that any right-continuous martingale can be embedded in a Brownian motion with the same initial value by what is called Skorokhod embedding. It follows that strictly positive martingales can be embeded in a logarithmic Brownian martingale. That means that one can embed the discounted stock price from other models in the discounted Black-Scholes stock price. Suppose for example, that $Y_k,\ k = 0, 1, 2, \dots$ is the stock price in a discrete model. Under the martingale measure the discounted stock price $\tilde{Y}_k \stackrel{\text{def}}{=} e^{-kr\delta} Y_n$ is a martingale. Then there are (possibly randomized) stopping times $0 = \tau_0 < \tau_1 < \dots$ for S_t such that the processes $\{\tilde{Y}_k,\ k = 0, 1, 2, \dots\}$ and $\{\tilde{S}_{\tau_k},\ k = 0, 1, 2, \dots\}$ have exactly the same distribution. Thus the process $(\tilde{Y}_k)$ is embedded in $\tilde{S}_t$. One can see that $\tilde{Y}_k$ is just the process $\tilde{S}_t$ sampled at discrete times. However, the sampling times are random, not fixed. This is what we mean by embedding.

Because it is the discounted prices that are martingales, we will usually work with discounted stock prices. If we need to emphasize that a price is the real, undiscounted price, we will call it the *raw* price. We will use tilde's to denote the discounted quantities. If S_t is the raw stock price, for instance, $\tilde{S}_t$ will denote the discounted price, and if K is the strike price of an option maturing at time T, then $\tilde{K} = Ke^{-rT}$ will denote the discounted strike.

Note that certain things, such the dividend yield of a stock, can be treated as a negative interest rate; that is, if the dividends are paid out at rate r_1, then it will be $e^{-(r-r_1)t} S_t$ rather than $e^{-rt} S_t$ which will be the martingale under the martingale measure, so that is the quantity which would be denoted $\tilde{S}_t$. This also happens with commodities and their holding cost, and with foreign exchange and the foreign currency interest rate, for example. This is well-understood and would just complicate the formulas without adding enlightenment, so we will not carry r_1 separately.

To see how to use this embedding, suppose our option or financial derivative pays an amount $f(S_T) = f(e^{rT}\tilde{S}_T)$ at its maturity time T in the Black-Scholes model. Its value at time zero is given by (1). On the other hand, in the discrete model, if $Y_0 = S_0$, and if $\delta = T/n$ the same derivative pays $f(Y_n)$ at maturity and has a value at time zero of

$$U(s_0, 0) \stackrel{\text{def}}{=} e^{-rT} E\{f(Y_n)\}\,. \tag{2}$$

Now $Y_n = e^{rT}\tilde{Y}_n$ has the same distribution as $e^{rT}\tilde{S}_{\tau_n}$. Thus the value at time zero is $e^{-rT}E\{f(e^{rT}\tilde{S}_{\tau_n})\}$ and the difference between the two values can be expressed entirely in terms of the discounted stock price $\tilde{S}$:

$$\mathcal{E}_{\text{tot}}(f) \stackrel{\text{def}}{=} U(s_0, 0) - V(s_0, 0) = e^{-rT} E\big\{f(e^{rT}\tilde{S}_{\tau_n}) - f(e^{rT}\tilde{S}_T)\big\}\,. \tag{3}$$

This involves the same process at two different times, the fixed time T and the random time τ_n. In cases such as the binomial tree, we have a good hold on the embedding times τ_n and can use this to get quite accurate estimates of the error, as we shall see below.

In fact, the embedding described above works for continuous time models as well: suppose $\{Z_t,\ t \geq 0\}$ is another stock model under its martingale measure, then $\tilde{Z}_t \equiv e^{-rt} Z_t$ is a strictly positive martingale. If $Z_0 = S_0 = s_0$, then $\tilde{Z}$ can be embedded: there exist stopping times $\tau_t,\ t \geq 0$ such that $\{\tilde{S}_{\tau_t},\ t \geq 0\}$ has the same distribution as $\{\tilde{Z}_t,\ t \geq 0\}$. This might be useful to analyze the case where the stock prices can have jumps.

We should note that it is the discounted stock prices, not the raw prices which are being embedded. In terms of the raw stock price, the error is:

$$\mathcal{E}_{\text{tot}}(f) = e^{-rT} E\big\{ f(e^{-r(\tau_n - T)} S_{\tau_n}) - f(S_T) \big\}\,. \tag{4}$$

The exponential term in the argument of f is random rather than deterministic, which makes this more awkward to handle than (3).

Rogers and Stapleton [**10**] have suggested modifying the binomial tree slightly in order to embed the stock prices directly, making (3) true for the raw stock prices (S_t) rather than $(\tilde{S}_t)$. They report that this increases its accuracy, and note that this embedding also makes it straightforward to handle things like knock-out options. They make an adjustment for end-effects at the same time, and it seems likely that the majority of the increase in accuracy is due to this adjustment.

The embedding scheme makes it possible to analyze the method probabilistically. The cause of the convergence wobble becomes obvious—it comes from the discontinuities in the payoff function and its derivative, and their relation to the lattice points. In fact it is not a random wobble: the error is simply a quasi-periodic function of the number of iterations[1]. It is even possible to get exact expressions for the first order error terms. It would be possible to get exact expressions for higher-order terms too, but given the complexity already present in the first-order terms, it would probably not be worth the effort.

The wobble is not present, for example, in the zero-interest rate at-the-money options, or, more generally, if the discounted price is at the money, i.e. if $Ke^{-rT} = s_0$, so that the strike price is always at a lattice point, but it is present whenever the strike price is not on a lattice point, and we can see from this that one can choose certain values of n—the number of time steps—for which the results are much better than average, and that if one chooses n carelessly, it is possible to, say, double the number of time steps and still increase the error. Moreover, once we understand the wobble, we can modify the Richardson extrapolation technique to bootstrap ourselves to a scheme of order three-halves.

3 The Binomial Tree Scheme

Let r be the interest rate. We will consider the Cox-Ross-Rubinstein binomial tree model for the *discounted* stock price. Let $\delta > 0$, and let the stock price at time $t = k\delta$ be Y_k; the discounted price is $\tilde{Y}_k = e^{-rk\delta} Y_k$. We will assume the probability measure is the martingale measure, so that $(\tilde{Y}_k)$ is a martingale, and we assume it takes values in the discrete set of values $a^j,\ j \in \mathbb{Z}$. At each step, $\tilde{Y}_k$ can jump to

[1] For the option case, this has also been shown recently and independently by F. and M. Diener [**3**], using quite different methods.

one of two possible values: either $\tilde{Y}_{k+1} = a\tilde{Y}_k$ or $\tilde{Y}_{k+1} = a^{-1}\tilde{Y}_k$, where $a > 1$ is a real number. The martingale property assures us that

$$P\{\tilde{Y}_{j+1} = a\tilde{Y}_j \mid \tilde{Y}_j\} = \frac{1}{a+1} \stackrel{\text{def}}{=} q\,, \quad P\{\tilde{Y}_{j+1} = a^{-1}\tilde{Y}_j \mid \tilde{Y}_j\} = 1 - q\,.$$

and we see that $(\tilde{Y}_k)$ is a Markov chain and these are its transition probabilities.

Let $\mathcal{K}$ be the class of real-valued functions f on $\mathbb{R}$ which are piecewise twice differentiable, polynomially bounded (i.e. there exist C and m such that $|f(x)| \leq C(1 + |x|^m)$ for all x) and which satisfy $f(x) = \frac{1}{2}(f(x+) + f(x-))$ at each discontinuity. We suppose that $f \in \mathcal{K}$ is the payoff of a European-type derivative with date of maturity T, which pays $f(s)$ at maturity if the stock price is s at time T. Fix an integer n and let $\delta = T/n$. Then the value of the derivative at time zero is $e^{-rT}E\{f(Y_n)\}$ and its value at some intermediate time $t = k\delta$ is

$$U(\tilde{Y}_k, k) \stackrel{\text{def}}{=} e^{-r(T-k\delta)}E\{f(Y_n) \mid \tilde{Y}_k\} = e^{-r(T-k\delta)}E\{f(e^{rT}\tilde{Y}_n) \mid \tilde{Y}_k\}\,. \tag{5}$$

Let $u(j,k) = U(s_0 a^j, k)$ Then $u = u(j,k)$ is a solution of the difference scheme

$$\begin{aligned} u(j,k-1) &= e^{-r\delta}\big(q\,u(j+1,k) + (1-q)\,u(j-1,k)\big),\ j \in \mathbb{Z},\ k = 1,\ldots,n\,, \\ u(j,n) &= f(e^{rT}a^j),\ j \in \mathbb{Z}\,. \end{aligned} \tag{6}$$

To approximate the Black-Scholes model with volatility σ, the parameters a, δ, n, σ and T are related:

$$\delta = \frac{T}{n}, \quad \log a = \sigma\sqrt{\frac{T}{n}}$$

and we define $h = \log a = \sigma\sqrt{T/n}$.

Remark 1 This is the binomial tree scheme for the *discounted* stock prices. It is more common to express it in terms of the raw stock price: i.e. to have the raw stock price increase or decrease by a factor, but there are some advantages of this scheme. In any case, the relations between the two schemes are rather simple, and they are identical if the interest rate is zero.

We will start with the exact results for the binomial scheme, and then compare it with the trinomial tree and some other modifications of the binomial scheme in the next section. A function $f \in \mathcal{K}$ can be written as a linear combination of options, smooth functions, and step functions. Since the scheme is linear, we can give the errors for each type and add them at the end. Let us start with the most interesting case, that of call and put options:

Theorem 3.1 *If f is the payoff of a European call or put option of strike price K, and if the initial stock price is s_0, there exist constants A and B such that*

$$\mathcal{E}_{\text{tot}}(f) = \frac{A + B\theta(1-\theta)}{n} + O(n^{-3/2}). \tag{7}$$

where $\theta \equiv frac\Big(\frac{\log(\tilde{K}/s_0)}{2h}\Big)$ is the fractional part of $\frac{\log(\tilde{K}/s_0)}{2h}$, $\tilde{K} \equiv Ke^{-rT}$ is the discounted strike price, and $h = \sigma\sqrt{\frac{T}{n}}$. The constants A and B can be given explicitly in terms of f.

The cases where f is smooth, or is discontinuous can be handled a bit more generally:

Theorem 3.2 *Suppose $f \in \mathcal{K}$. (a) If f is smooth, there is a constant A such that $\mathcal{E}_{\text{tot}}(f) = A/n$ + higher order terms.*

(b) If f has one or more jump discontinuities, and if at least one of its discontinuities is not at a lattice point, then $\mathcal{E}_{\text{tot}}(f) = O(n^{-1/2})$. If all its discontinuities are at lattice points, then $\mathcal{E}_{\text{tot}}(f) = O(n^{-1})$.

Remark 2 In fact, B has a rather simple formula: for a European call option of strike price K, $f(x) = (x-K)^+$, and

$$B = 2\sigma^2 T K \hat{p}(\log(\tilde{K}/s_0)),$$

where s_0 is the initial stock value and $\hat{p}$ is the P-density of $\sigma W_T - \sigma^2 T/2$ given below. The formula for A is not at all simple. It appears in all its glory in Theorem 3.4.

The hedging strategy depends on what is called delta, which is just the space derivative $\frac{\partial v}{\partial x}$, where $v(x,t)$ is the solution of the Black-Scholes problem. This can be estimated by a difference quotient in u. This estimate turns out to be first order as well, which may or may not be surprising. (In effect, estimating delta for a payoff function f is equivalent to estimating f' by the tree scheme. If f is smooth, Theorem 3.2 (a) applies to f', and implies first order convergence. But if f' is discontinuous, as it will be with a call or put option, the Theorem 3.2 (b) only guarantees order 1/2 convergence. So there is something to prove.)

It is important to use the symmetric estimate of the derivative: we estimate delta by

$$\frac{u(j+1,k) - u(j-1,k)}{s_0(e^{(j+1)h} - e^{(j-1)h})}. \tag{8}$$

For practical reasons, one would usually use an estimate gotten by looking one step ahead:

$$e^{-\delta}\frac{u(j+1,k+1) - u(j-1,k+1)}{s_0(e^{(j+1)h} - e^{(j-1)h})}.$$

The convergence rates of the two will be the same.

Theorem 3.3 *Suppose f is continuous and both f and f' are in $\mathcal{K}$. Then the estimate (8) of delta is first-order.*

The estimate of the delta is at worst of order 1/2, being the difference of two quantities with errors on the order 1/n, divided by a quantity on the order of $1/\sqrt{n}$. However, the errors are highly correlated, and tend to cancel out. This can be seen by a coupling argument. Without loss of generality, it is enough to consider the case $k = 0$. We notice that if S_t starts from s_0, then $e^h S_t$ and $e^{-h} S_t$ start from $e^h s_0$ and $e^{-h} s_0$ respectively, so that

$$\frac{u(j+1,0) - u(j-1,0)}{s_0(e^{(j+1)h} - e^{(j-1)h})} = e^{-rT} E\{\hat{f}(e^{rT}\tilde{S}_{T_n}, h)\}$$

where $\hat{f}(s,h) \equiv (f(e^h s) - f(e^{-h}s))(s_0(e^h - e^{-h}))^{-1}$. The error can be written

$$e^{-rT} E\{\hat{f}(e^{rT} S_{\tau_n}, h) - \hat{f}(S_T, h)\} + e^{-rT} E\{\hat{f}(S_T, h) - \frac{S_T}{s_0} f'(S_T)\}.$$

The first term is the error using $\hat{f}(\cdot, h)$ instead of $f(\cdot)$, and one uses Theorem 3.4 to bound it *uniformly in h*. The second term is straightforward analysis. This estimate also exhibits a wobble in convergence.

For completeness, we will give the values of the constants for a general payoff function $f \in \mathcal{K}$. Let $\Delta f(s)$ and $\Delta f'(s)$ be the discontinuities at s of f and f' respectively. For any x, let $\theta(x) = \text{frac}\left(\frac{x}{2h}\right)$ be the fractional part of $\frac{x}{2h}$. Let $h\mathbb{Z}$ be the set of all multiples of h, $\mathbb{N}_e^h \stackrel{\text{def}}{=} 2h\mathbb{Z}$ the set of all *even* multiples of h, and $\mathbb{N}_o^h \stackrel{\text{def}}{=} h + \mathbb{N}_e^h$ the set of all *odd* multiples of h. The density of $\log \tilde{S}_T \equiv \sigma W_T - \sigma^2 T/2$ is

$$\hat{p}(x) \equiv \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{(x+\frac{1}{2}\sigma^2 T)^2}{2\sigma^2 T}} . \tag{9}$$

Theorem 3.4 *Suppose that $f \in \mathcal{K}$. Let $s_1, s_2, \ldots, s_k$ be the set of discontinuity points of f and f'. and let s_0 be the initial stock price. Denote their discounted values by $\tilde{s}_i = s_i e^{-rT}$. Then the error in the tree scheme at $(s_0, 0)$ is*

$$\begin{aligned}
\mathcal{E}_{\text{tot}}(f) = \frac{e^{-rT}}{6n} \Bigg[& \Big(\frac{5}{2} + \sigma^2 T + \frac{\sigma^4 T^2}{32}\Big) E\Big\{f(S_T)\Big\} - \frac{1}{\sigma^2 T} E\Big\{(\log(S_T/s_0))^2 f(S_T)\Big\} \\
& - \frac{1}{2\sigma^4 T^2} E\Big\{(\log(S_T/s_0))^4 f(S_T)\Big\} + 4\sigma^2 T\, E\Big\{S_T^2 f''(S_T)\Big\} \\
+ \sigma^2 T \sum_i & \Big(s_i\, \Delta f'(s_i) - \frac{1}{2}\Delta f(s_i)\Big)\Big(2 + 12\theta(\tilde{s}_i/s_0)(1 - \theta(\tilde{s}_i/s_0))\Big)\hat{p}(\log((\tilde{s}_i/s_0)) \\
& - 2 \sum_{i:\log(\tilde{s}_i/s_0)\in\mathbb{N}_e^h} \log(\tilde{s}_i/s_0)\, \Delta f(s_i)\, \hat{p}(\log(\tilde{s}_i/s_0)) \\
& + \sum_{i:\log(\tilde{s}_i/s_0)\in\mathbb{N}_o^h} \log(\tilde{s}_i/s_0)\, \Delta f(s_i)\, \hat{p}(\log(\tilde{s}_i/s_0)) \Bigg] \\
+ \frac{\sigma\sqrt{T} e^{-rT}}{\sqrt{n}} & \sum_{i:\log(\tilde{s}_i/s_0)\notin h\mathbb{Z}} (2\theta(\tilde{s}_i/s_0) - 1)\, \Delta f(s_i) \hat{p}(\log(\tilde{s}_i/s_0)) + O\Big(\frac{1}{n^{3/2}}\Big)
\end{aligned} \tag{10}$$

where the expectations above are taken with respect to the martingale measure given $S_0 = s_0$.

Remark 3 We have expressed the errors in terms of $E\{f(S_T)\}$. However, it might be better stated in terms of $E\{f(e^{rT}\tilde{S}_{\tau_n})\}$ since this is exactly what the binomial tree scheme computes. In fact we need more than that: we need the expectation of a linear combination of $f(e^{rT}\tilde{S}_{\tau_n})$, $f(e^{rT}\tilde{S}_{\tau_n})(\log(e^{rT}\tilde{S}_{\tau_n}/s_0))^2$ and $f(e^{rT}\tilde{S}_{\tau_n})(\log(e^{rT}\tilde{S}_{\tau_n})/s_0))^4$. The tree scheme will compute this exactly, too, but we have to run it again to do it.

In fact, the expectations of $f(e^{rT}\tilde{S}_{\tau_n})$ differ from those of $f(S_T)$ by $O(1/n)$. Since they occur as coefficients multiplying $1/n$, we can replace S_T by $e^{rT}\tilde{S}_{\tau_n}$ throughout (10). This will only change the terms by $O(n^{-2})$, so the formula remains correct. Thus S_T and $e^{rT}\tilde{S}_{\tau_n}$—and, for that matter, S_{τ_n}—are interchangeable as far as (10) goes.

Remark 4 The term involving θ is responsible for the wobble in convergence. Indeed, θ is the fractional part of $\frac{\sqrt{n}\log(\tilde{K}/s_0)}{2\sigma\sqrt{T}}$, so θ varies strongly with n. In fact, what happens is that the strike price K is fixed, but the lattice itself varies with n, so the distance from K to the nearest lattice point changes. Note that the term involving θ has a factor of $1/3 + 2\theta(1-\theta)$, which can vary by nearly a factor of three as theta goes from zero to one. It can be minimized by careful choice of n; one just has to choose n to make $\theta =$ frac $\frac{\log(\tilde{K})}{2h}$ close to either zero or one, or equivalently, make $\tilde{K}/s_0$ close to an even lattice point. Since the price can fluctuate by a large factor as n changes, this can be advantageous. (Note that this term is largest when $\theta = 1/2$, which happens when $\tilde{K}/s_0$ is at an odd lattice point. Thus it is not enough just to assure that the strike price is near a lattice point; it should be at an even lattice point.)

We hasten to add that this "good" choice of n may not always give the best possible values, for this is not the only error term, even though it is often the largest one. If the other error terms are negative, the error may change sign as n increases, and this "good" choice of n can possibly give a local maximum of the absolute error. See section 4.3 for some illustrations of this.

Remark 5 The above remarks hold for the binomial scheme for the discounted stock prices: it is the discounted stock prices, not the raw prices, which evolve on the lattice. The other binomial scheme, in which the raw stock prices are on the lattice, will behave in about the same way, except that it is K, not the discounted value $\tilde{K}$, which should be close to an even lattice point, and θ should be defined with K instead of $\tilde{K}$. We will see that the trinomial scheme has roughly similar behavior.

4 Examples

4.1 Close Relations. The trinomial tree is closely related to the binomial tree. It is also a first order scheme. In it, we allow three possibilities at each step instead of two: the stock price can go up or down by one step, as in the binomial scheme, but it can also stay the same. As with the binomial scheme, the transition probabilities are determined by the requirement that the stock price be a martingale, and that the expected time step be T/n.

Since there is one more free parameter—the probability of staying the same—there is some flexibility in choosing the size of the space steps. We analyze the errors in section 7. Its convergence behavior is broadly similar to that of the binomial tree, with one exception: the even-odd swings of the binomial tree are smoothed out. But the option-pricing error is again a quasi-periodic function of the number of time-steps, with a different quasi-frequency.

This quasi-periodic behavior comes from the relation of the strike price to the lattice points. It is possible to re-adjust the trinomial tree at each time step to put the strike at a node. We call this the "adjusted trinomial." This indeed smooths out the convergence: the error is now asymptotically A/n, but it doesn't seem to improve the accuracy; on the contrary, it can be uniformly worse than the binomial scheme, even though it involves more calculation. However, this does suggest Richardson extrapolation to get a scheme of order three-halves.

Another way of smoothing the convergence is either to move a single lattice point over to the strike price, or add a new lattice point at the strike price. This requires changing the transition probabilities of the new point and the points on

either side. In order to assure that the stock price remains a martingale and that the expected time of the time step remains T/n, one has to make these points into (at least) trinomial jump points.

We find that for the usual European options, at least in the cases we have checked, these modifications can smooth out the fluctuations in the error with the number of time steps, but give roughly comparable accuracy for a given number of iterations.

4.2 Richardson Extrapolation. Theorem 3.1 suggests a way to improve the estimate by extrapolation. Let the true value of the option be V, and let the value computed by the binomial scheme of n steps be $u(n)$. Define $\theta = \theta_n$ as in Theorem 3.1, and note that by (7) we have

$$V = u(n) - n^{-1}(A + B\theta_n(1-\theta_n)) + \epsilon(n) \tag{11}$$

where $\epsilon(n) = O(n^{-3/2})$. Choose three values of n, say $n_1 < n_2 < n_3$, which give at least two different values of θ_n. Then (11) gives a system of three linear equations in the unknowns V, A, and B in terms of the $u(n_i)$ and $\epsilon(n_i)$. Since we know nothing about the $\epsilon(n_i)$ except that they are relatively small, set them equal to zero and solve the resulting equations for v, A, and B. If $\hat{V}$ is the solution for V, it is easy to see that the error $u_n - \hat{V}$ will be on the order of the $\epsilon(n_i)$, which is $O(n_1^{-3/2})$.

This is the ordinary modification of Richardson extrapolation to our case. (Actually, we have an explicit formula for B, so we could get away with only two values of n in this.) Let us point out that there is a cheap way of improving the estimate. The estimates of A and B above have errors on the order of $1/\sqrt{n}$. If we wanted to improve the error for $n = n_0$, we could estimate A and B with much smaller values of n, and use these values to estimate the error to be subtracted off at n_0. This can give a greatly improved estimate, albeit not of order three-halves.

These remarks hold for the trinomial tree as well. While we have only analyzed discounted trees, they seem to hold quite well for the ordinary trees. This has implications for American options which we will explore elsewhere.

4.3 Pictures. We will give a number of numerical examples to illustrate some of our conclusions on the convergence of the binomial, trinomial, and related schemes. In all cases, the graph shows the error $\mathcal{E}_{\text{tot}}$ plotted against the number n of iterations. Unless stated otherwise, only even iterations are plotted. We will consider call/put options, where the payoff is $(x-K)^+$ or $(K-x)^+$, and binary options, where the payoff is 1 if $x \geq K$ and 0 if $x < K$.

One thing is clear from these graphs: the errors fluctuate enough with n so that it is misleading to compare different schemes for a single, or even for just a few values of n. One gets a much better picture by comparing them over a whole range.

So far, we have concentrated on the discounted binomial simply because it is easy to analyze. However, the usual binomial method should be quite similar, and all the numerical examples we have checked support this. All the methods: the binomial, discounted binomial, the trinomial and its discounted cousin, have roughly comparable accuracy. There doesn't seem to be a great reason to favor one over the other. Figure 1 shows the comparison of three methods on a typical European option. The (signed) error is plotted against the number n of time steps,

for all even values of n from $n = 50$ to $n = 350$, and the dotted lines $.75/n$ and $-.6/n$ are there for the sake of comparison.

Note that the error in all three is quasi-periodic, and changes sign several times. The trinomial is slightly better than the other two on the average: while the worst errors of the three are about equivalent, the trinomial has less fluctuation, so that it is somewhat better than the other two on the downside. But there are many values of n for which any given one is better than the other two, and it is hard to form a firm preference.

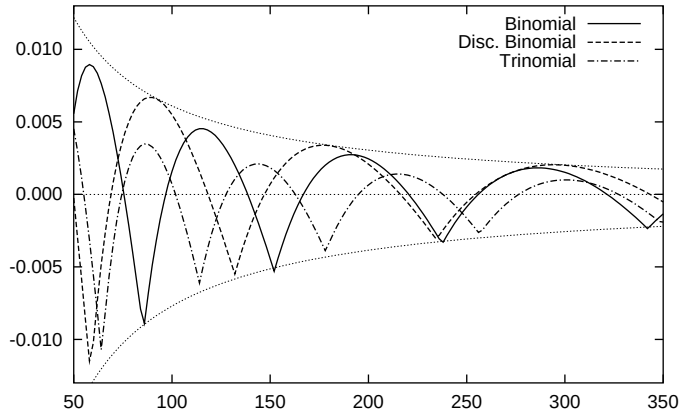
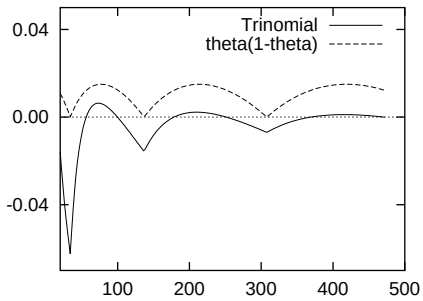


Fig. 1: Binomial, Discounted Binomial and Trinomial errors.
European call, $S_0 = 95$, $K = 100$, $T = 0.1$, vol $= 25\%$, $r = 10\%$

The following two figures show the signed error of the ordinary trinomial scheme, first superimposed on a graph of $\theta(1-\theta)$, and then compared with two multiples of $1/n$.



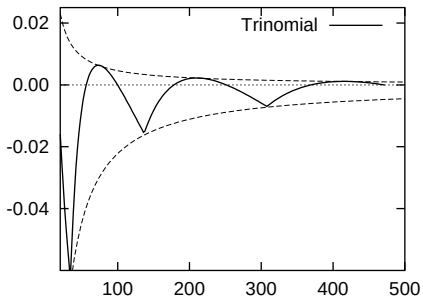


Fig. 2: Trinomial error with $C\theta(1-\theta)$. Fig. 3: Trinomial error with C/n.
European call, $S_0 = 95$, $K = 100$, $T = 1$, vol $= 15\%$, $r = 10\%$.

Note that the sign of the error changes as n increases. This is typical, though not universal. This means that there are certain values of n near where the error changes sign which give extremely close estimates. The problem is to find them. It is easy to find the local extreme values of the error, for those are where $\theta = 0$, 1 and $1/2$. On the other hand, if the error is known to always be of one sign, say positive, the good values of n are easily predicted: they will be the n for which θ is close to zero or one.

Here is an example in which the error does not change sign, and one can find "good" values of n by choosing θ close to one half. The upper envelope is $-.00036/n$,

and the lower envelope is $-.0045/n$, so the "good" values are more than ten times better than the "bad" ones.

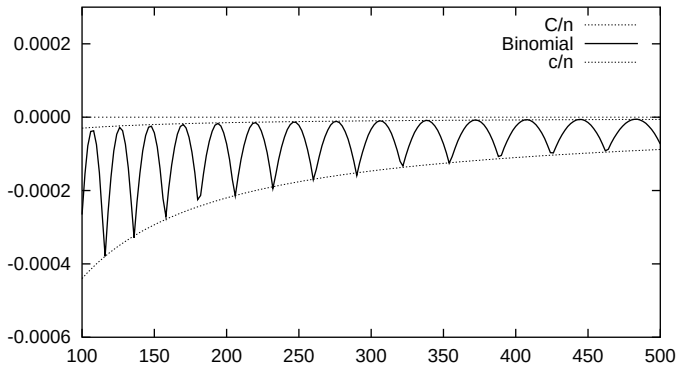


Fig. 4: European call, $S_0 = 100$, $K = 80$, $T = 1$, vol $= 10\%$, $r = 5\%$
Comparison of the error with C/n.

The binomial tree scheme is notoriously susceptible to odd-even fluctuations while the trinomial is not. The reason for the swing on the binomial is clear. The random walk is periodic, with period 2: it visits even nodes on even steps, odd nodes on odd steps. Going from an even to an odd step is roughly equivalent to interchanging even and odd nodes, which effectively replaces θ by $|\theta - 1/2|$. The random walk corresponding to the trinomial, however, is aperiodic, and its distribution varies little between one step and the next.

The quasi-periodicity comes from the term involving $\theta(1-\theta)$, where θ is the fractional part of $\log(\tilde{K}/s_0)\sqrt{n/4\sigma^2 t}$. This is a quasi-periodic function of n, and we could define its quasi-period at n iterations as $4\sigma\sqrt{nT}/\log(\tilde{K}/s_0)$. Note that the period increases as the volatility increases, and also as n and the term T increase. In fact, if the term is long and the volatility high, the fluctuation almost disappears. (This is for the discounted binomial. For the ordinary binomial, replace $\tilde{K}$, by K. For the trinomial, see (23).) Note, however, that this quasi-periodicity disappears if the option is at-the-money, and in fact, the convergence is quite smooth. To be more exact, the convergence is smooth along the even values of n, and it is also smooth along the odd values of n. However, it swings wildly between even and odd values. (Indeed, in this case, θ is effectively changing from 0 to 1/2 between the even and odd iterations.) Figure 5 below illustrates both the even-odd swings and the smoothness of convergence for at-the-money options. The error in the even iterations is asymptotic to C/n, and the error in the odd iterations is asymptotic to $-C'/n$, and, in this case at least, C and C' are nearly equal.

It is perhaps more interesting to see the odd-even swings when the option is not at-the-money, when the quasi-periodic behavior is also present. Both the binomial and trinomial error in the pricing of a call option are shown in Figure 6. (We added .02 to the trinomial error to get both on the same graph.) We see that the trinomial error does not have odd-even swings, while the binomial definitely does. Notice that both have similar quasi-periodic behavior, and that with the binomial, the odd-even swing is simply superimposed on the quasi-periodicity.

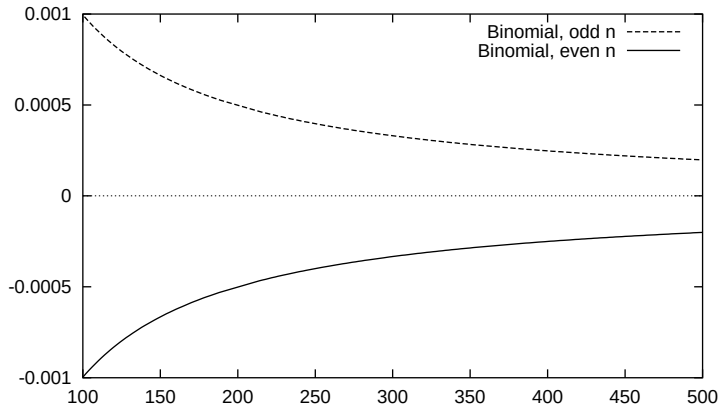


Fig. 5: Binomial errors, at the money, even and odd n plotted separately. European call, $S_0 = 10$, $K = 10$, $T = 1$, vol $= 10\%$, $r = 0$.

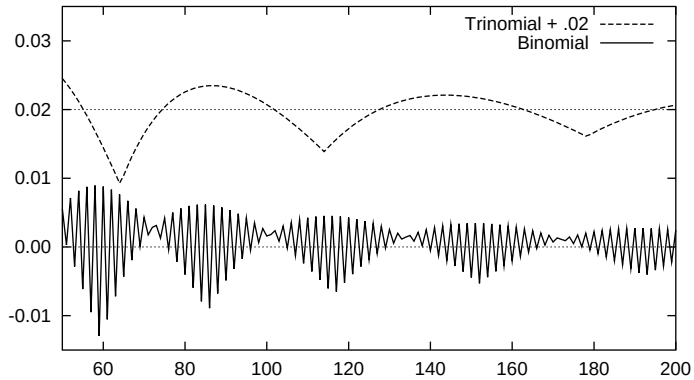


Fig. 6: Trinomial and binomial errors plotted for all n between 70 and 200. European call, $S_0 = 95$, $K = 100$, $T = .1$, vol $= 25\%$, $r = 10\%$.

Notice that while the convergence in the at-the-money case is smoother than it was in Figures 1—3, it is not substantially faster. The smooth convergence comes simply from the fact that $\theta(n) = 0$ for all n, so the fluctuation drops out.

There are various tricks used to smooth out the convergence, some having to do with arranging the grid so that the strike price is always on a node, as we mentioned in section 4.1. These methods can regularize the convergence; it is less clear that they actually improve it. Here is a typical problem comparing the usual trinomial with a trinomial which is adjusted to have the strike at a node for each n. While the adjusted trinomial error converges more smoothly to zero than the normal trinomial, it is uniformly larger. The apparent discontinuities in the adjusted trinomial error are an artifact of our method. The scale factor at each step is of the form $e^{\sigma\sqrt{\alpha T/n}}$; the parameter α is adjusted to put the strike K at a node. The optimal value of α is said to be 3, so we required $1 < \alpha < 4$ to keep the value close. We let α change smoothly with n until it reached one of the endpoints, and then let it jump back into the interval. This causes the discontinuity. The fact that we had to choose sub-optimal values of α to get the strike at a node is evidently enough to cause the larger error. Comparing this with Figure 1, which concerns the same problem, we see that the adjusted trinomial is also worse than the binomial.

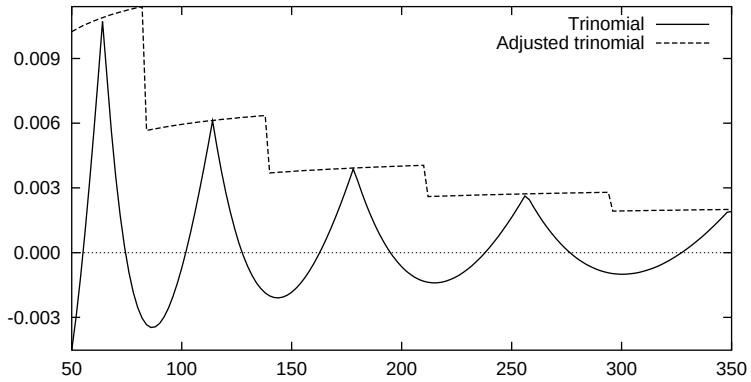


Fig. 7: Comparison of the errors in the trinomial and adjusted trinomial. European call, $S_0 = 95$, $K = 100$, $T = 0.1$, vol $= 25\%$, $r = 10\%$.

If it is not possible to divine the "good" values of n, one can try Richardson extrapolation. This will give an error on the order of $n^{-3/2}$, at the price of considerably more work. Here is a variant of that idea, a "cheap" Richardson. We ran the scheme for relatively small values of n—$n = 60$, 70 and 80—to estimate A and B in (11), and then used these estimates to subtract off the error for values of n up to 350. Because of the fixed error in the estimates of A and B, this only gives a scheme of order one, not three halves, but it does improve the constant of convergence.

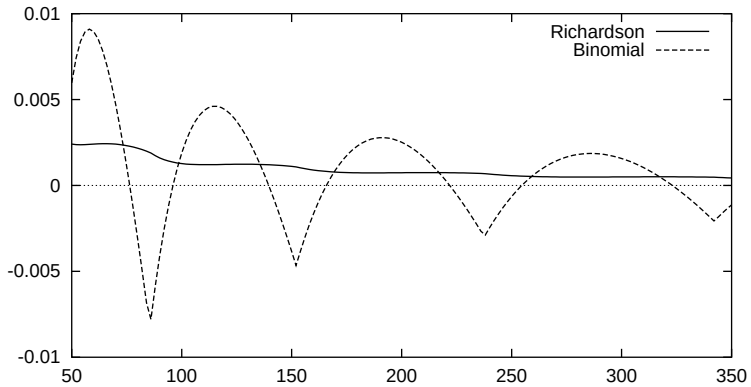


Fig. 8: Binomial with a variant of Richardson extrapolation. European call, $S_0 = 95$, $K = 100$, $T = .1$, vol $= 25\%$, $r = 0$.

Digital options, which have a genuine discontinuity in the payoff, only converge with order one half in general, but they converge with order one if the discontinuity sits exactly on a node. Here are two cases, the first a generic digital option, and the second an at-the-money digital option. In the at-the-money case, the discontinuity is always at a node, and the convergence is first order.

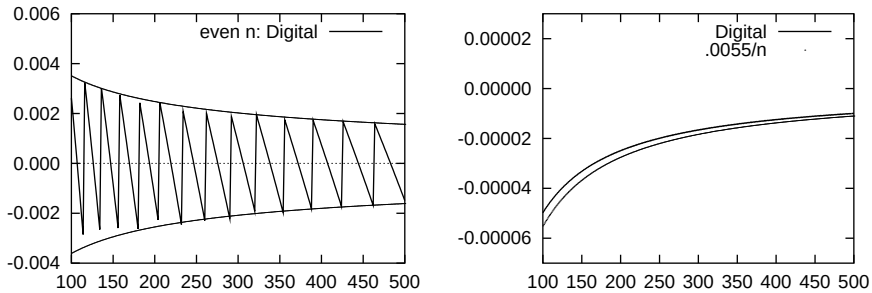


Fig. 9: Digital errors: in-the-money, with $\pm C/\sqrt{n}$; at-the-money, with C/n. Digital call, $S = 10$, $T = 1$, vol $= 10\%$, $r = 0$, with $K = 8$ and $K = 10$ resp.

In the first case the leading error term is

$$(A + B(2\theta - 1))\frac{1}{\sqrt{n}} \tag{12}$$

while in the second it is C/n. Note that the error is piecewise-linear in θ, rather than piecewise-quadratic, as with the options. This is evident in Figure 9.

Given the slow rate of convergence of digital options, Richardson extrapolation should be worthwhile. Indeed, it is a matter of estimating A and B again, which can be done just as with a call or put option, and the improvement essentially squares the error.

Finally, here is the graph of the binomial tree's estimate of the error in the delta. This is the same option considered in Figure 4.

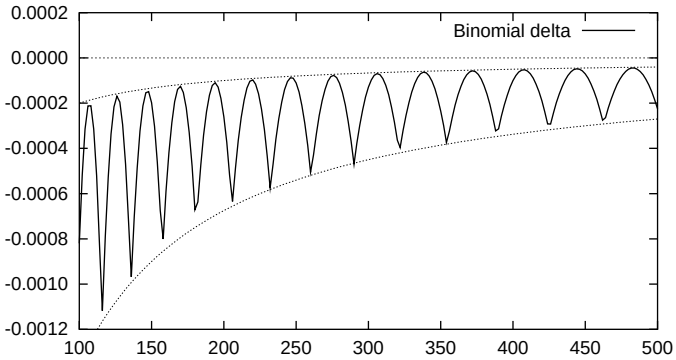


Fig. 10 Errors in the Delta with C/n
European call, $S_0 = 100$, $K = 80$, $T = 1$, vol $= 10\%$, $r = 5\%$.

5 Embedding the Binomial Tree

We normalize so that the initial value is $S_0 = 1$. Keeping the notation of section 3, let $h = \log a = \sigma\sqrt{\frac{T}{n}}$, and let $\tilde{Y}_k$ be the discounted stock price in the binomial model. Under the martingale measure P, the discounted stock price $\tilde{S}_t$ is also a martingale. Then we can embed $\tilde{Y}_k$ in $\tilde{S}_t$, and we can define the stopping times explicitly. Define $\tau_0, \tau_1, \tau_2, \ldots$ inductively by

$$\tau_0 = 0, \text{ and } \tau_{k+1} = \inf\{t > \tau_k : \tilde{S}_t = a\tilde{S}_{\tau_k} \text{ or } a^{-1}\tilde{S}_{\tau_k}\}.$$

From the martingale stopping theorem, we see that

$$P\{\tilde{S}_{\tau_{k+1}} = a\tilde{S}_{\tau_k} \mid \tilde{S}_{\tau_k}\} = \frac{1}{a+1} \quad \text{and} \quad P\{\tilde{S}_{\tau_{k+1}} = a^{-1}\tilde{S}_{\tau_k} \mid \tilde{S}_{\tau_k}\} = \frac{a}{a+1},$$

which are the same transition probabilities we found for $\tilde{Y}_k$. We conclude that the two processes $\{\tilde{S}_{\tau_k},\ k = 0,1,2,\dots\}$ and $\{\tilde{Y}_k,\ k = 0,1,2,\dots\}$ have exactly the same distributions, for they are both Markov chains and they have the same initial distribution and transition probabilities. For our purposes, they are the same process.

We will make two transformations to simplify the problem. First, we take logarithms. The logarithmic Brownian motion $\tilde{S}_t$ is a martingale, so it can be written $\tilde{S}_t = e^{\sigma W_t - \frac{1}{2}\sigma^2 t}$, where W_t is a standard Brownian motion and $\sigma > 0$. Let $X_t = \log S_t$, $\tilde{X}_t = \log \tilde{S}_t$. Then $\tilde{X}_t$ is a Brownian motion with drift, and $\{\tilde{X}_{\tau_k}, k = 0,1,2,\dots\}$ is a random walk on the lattice $h\mathbb{Z}$.

Next, we eliminate the drift in $\tilde{X}$ by a Girsanov transformation. Let ξ be a finite stopping time which dominates $\max(T, \tau_J, \tau_n)$, (where τ_J is defined in section 6.2) and let Q be the probability measure given by $dQ = e^{\frac{1}{2}\tilde{X}_\xi + \frac{1}{8}\sigma^2\xi} dP$. By Girsanov's Theorem [**5**], $\frac{1}{\sigma}\tilde{X}_t$ is a standard Brownian motion on $[0,\xi]$ under the probability Q. We call Q the *Brownian measure* to distinguish it from the martingale measure P, and we will write E^P and E^Q to indicate expectations under the measures P and Q respectively.

From (3), the error made by the tree scheme is

$$u(1,0) - v(1,0) = e^{-rT} E^P\{f\big(e^{\tilde{X}_{\tau_n} + rT}\big) - f\big(e^{\tilde{X}_T + rT}\big)\}.$$

Define

$$g(x) \stackrel{\text{def}}{=} f(e^{x+rT}) e^{-\frac{x}{2} - \frac{\sigma^2 T}{8}}. \tag{13}$$

Now $e^{-\frac{1}{2}\tilde{X}_t - \sigma^2 t/8}$ is a Q-martingale and $T < \xi$ so, using the martingale stopping theorem and writing E^P and E^Q for the expectations under the martingale and Brownian measures respectively, we see that

$$\begin{aligned} E^P\{f(S_T)\} &= E^Q\{f(e^{\tilde{X}_T + rT}) e^{-\frac{1}{2}\tilde{X}_\xi - \sigma^2\xi/8}\} \\ &= E^Q\{f(e^{\tilde{X}_T + rT}) e^{-\frac{1}{2}\tilde{X}_T - \sigma^2 T/8}\} = E^Q\{g(\tilde{X}_T)\}. \end{aligned} \tag{14}$$

and

$$E^P\{f(e^{rT}\tilde{S}_{\tau_n})\} = E^Q\{f e^{rT}\tilde{(S_{\tau_n})} e^{-\frac{1}{2}\tilde{X}_{\tau_n} - \sigma^2\tau_n/8}\} = E^Q\{g(\tilde{X}_{\tau_n}) e^{-\frac{\sigma^2(\tau_n - T)}{8}}\}$$

This gives us an expression for the total error in terms of the Brownian measure:

$$\mathcal{E}_{\text{tot}}(f) \stackrel{\text{def}}{=} e^{-rT} E^Q\{g(\tilde{X}_{\tau_n}) - g(\tilde{X}_T)\} + e^{-rT} E^Q\{g(\tilde{X}_{\tau_n})\big(e^{-\frac{\sigma^2}{8}(\tau_n - T)} - 1\big)\}. \tag{15}$$

The key to calculating these quantities is the following remark:

Remark 6 1. Under the Brownian measure, the stopping times $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2 \dots$ are i.i.d. and independent of the process $\tilde{X}_0, \tilde{X}_{\tau_1}, \tilde{X}_{\tau_2}, \dots$.

2. Under the Brownian measure, the τ_k are Brownian first-hitting times, whose distribution and moments can be calculated. In particular, the moment generating function of τ_1 can be written in terms of the complex cosine:

$$F(\lambda) \stackrel{\text{def}}{=} E\{e^{\lambda\tau_1}\} = \frac{1}{\cos\sqrt{2\lambda\frac{T}{n}}}, \quad -\infty < \lambda < \frac{n\pi^2}{8T}.$$

3. It follows that the expectation of τ_n is T, its variance is $\frac{2T^2}{3n}$, and that all moments of τ_n are finite.

6 Methods

6.1 A Heuristic Calculation. One can get some insight from the following calculation. Notice that τ_n and $\tilde{X}_{\tau_n}$ are independent. Let's go a little bit further and suppose that τ_n is independent of $\tilde{X}_T$, and even more, of the whole process $(\tilde{X}_t)$. Define $G(t) = E^Q\{g(\tilde{X}_t)\}$. Then

$$E^Q\{g(\tilde{X}_{\tau_n}) - g(\tilde{X}_T)\} = E^Q\{G(\tau_n) - G(T)\}$$

Expand G in a Taylor series:

$$= G'(T)E\{\tau_n - T\} + \frac{1}{2}G''(T)E\{(\tau_n - T)^2\} + \dots$$

But $E\{\tau_n\} = T$ and $\text{var}(\tau_n) = \frac{2T^2}{3n}$ so the first term drops out and this is

$$= \frac{T^2}{3n}G''(T) + \dots \tag{16}$$

The other terms are of higher order, so this has an error of order $1/n$, as claimed.

Of course this is not rigorous, since τ_n, far from being independent of the process $(\tilde{X}_t)$, is a function of it. However, it turns out that this dependence is rather local, and we can make the above heuristic argument rigorous by splitting off the local and the global parts. The complete details for the binomial tree are in [**11**]. We will just outline the essential parts of the argument here; we will adapt it to the trinomial tree in the next section.

6.2 Splitting the Error. Define

$$J = \inf\{2k : \tau_{2k} \geq T\}.$$

We concentrate on even n, because we want to avoid the even-odd fluctuations. Then τ_J is a stopping time for $(\tilde{X}_t)$, and J itself is a stopping time for $\tau_1, \tau_2, \dots$, and thus is independent of $\tilde{X}_{\tau_1}, \tilde{X}_{\tau_2}, \dots$

By independence, the second term in the right-hand side of (15) can be written

$$E^Q\{g(\tilde{X}_{\tau_n})\}E^Q\left\{e^{-\frac{\sigma^2}{8}(\tau_n - T)} - 1\right\}.$$

Expand the exponential and use the expectation and variance of τ_n from Remark 6 to see this is

$$\frac{\sigma^4 T^2}{192n}E^Q\{g(\tilde{X}_{\tau_n})\} + o\left(\frac{1}{n}\right). \tag{17}$$

The first term on the right hand side of (15) can be split into a global and a local part:

$$\begin{aligned} E^Q\{g(\tilde{X}_{\tau_n}) - g(\tilde{X}_T)\} &= E^Q\{g(\tilde{X}_{\tau_n}) - g(\tilde{X}_{\tau_J})\} + E^Q\{g(\tilde{X}_{\tau_J}) - g(\tilde{X}_T)\} \\ &\stackrel{\text{def}}{=} \mathcal{E}_{\text{glob}} + \mathcal{E}_{\text{loc}}. \end{aligned} \tag{18}$$

The heuristic argument above applies to $\mathcal{E}_{\text{glob}}$, and basically comes down to a Taylor expansion of binomial coefficients. Indeed, as $(\tilde{X}_{\tau_j})$ is a simple symmetric random walk, its distribution involves binomial probabilities, and, remembering that it is independent of J, we can write

$$\mathcal{E}_{\text{glob}} = \sum_{k \text{ even}} P^Q\{J = n+k\} \sum_{j \text{ even}} \big(P_n(j) - P_{n+k}(j)\big) g(jh)$$

where

$$P_k(j) = \frac{k!\,2^{-k}}{(\frac{k+j}{2})!(\frac{k-j}{2})!}\,.$$

One can show [**11**] that for the relevant values of j and k,

$$P_{n+k}(j) = \Big(1 - \frac{k}{2n} + \frac{3k^2}{8n^2} + \dots\Big) P_n(j)\,,$$

so that the above is

$$\begin{aligned} &= \sum_k P^Q\{J = n+k\}\Big(\frac{k}{2n} - \frac{3k^2}{8n^2} + \dots\Big) \sum_j g(jh) P_n(j) \\ &= \Big(\frac{1}{2n} E^Q\{J-n\} - \frac{3}{8n^2} E^Q\{(J-n)^2\} + \dots\Big) E^Q\{g(\tilde{X}_{\tau_n})\}\,. \end{aligned}$$

One can then show [**11**] that $E^Q\{J-n\} = 4/3 + o(1)$ and $E^Q\{(J-n)^2\} = 2n/3 + O(1)$ so that

$$\mathcal{E}_{\text{glob}} = \big(\frac{5}{12n} + \dots\big) E^Q\{g(\tilde{X}_{\tau_n})\} \tag{19}$$

The three dots in (19) hide some non-negligeable terms. We omitted them here, but included them in Theorem 3.4.

The next term, $\mathcal{E}_{\text{loc}}$, is where the local properties such as the continuity and differentiability of f and g enter. This is the term that causes the irregularity of convergence. We will see how this happens.

The calculation depends on whether T falls into an odd interval (τ_{2j-1}, τ_{2j}) or an even interval (τ_{2j}, τ_{2j+1}). The details can be found in [**11**], so we will just do the odd case, where $\tau_{2j-1} < T < \tau_{2j}$. This is enough to show where the convergence wobble comes from. In that case, evidently $J = 2j$, so $\tau_{J-1} < T < \tau_J$, and J is the first time after T that $\tilde{X}_t$ hits an even multiple of h.

If $\tilde{X}_T = x \in (2kh, (2k+2)h)$ for some k, then τ_J is the first time after T that $\tilde{X}_t$ leaves that interval, so from the well-known Brownian hitting probabilities,

$$P\{\tilde{X}_{\tau_J} = (2k+2)h \mid \tilde{X}_T = x,\ T \in (\tau_{J-1}, \tau_J)\} = \frac{x - 2kh}{2h}$$

$$P\{\tilde{X}_{\tau_J} = 2kh \mid \tilde{X}_T = x,\ T \in (\tau_{J-1}, \tau_J)\} = \frac{(2k+2)h - x}{2h}$$

These probabilities are linear in x on this interval. Consequently, if $\Pi g(x)$ is the continuous function which equals $g(x)$ if $x = 2kh$ for some k, and which is linear in between, then

$$E^Q\big\{g(\tilde{X}_{\tau_J}) \mid \tilde{X}_T = x,\ T \in (\tau_{J-1}, \tau_J)\big\} = \Pi g(x)\,.$$

Thus, if $p(x)$ is the conditional Q-density of $\tilde{X}_T$ given $T \in (\tau_{J-1}, \tau_J)$,

$$E^Q\{g(\tilde{X}_{\tau_J}) - g(\tilde{X}_T) \mid T \in (\tau_{J-1}, \tau_J)\} = \int_{-\infty}^{\infty} (\Pi g(x) - g(x))\, p(x)\, dx\,. \tag{20}$$

To calculate $p(x)$, we need the conditional distribution of $\tilde{X}_{\tau_{J-1}}$ given $\tilde{X}_T$. This can be found by a time-reversal argument. See (28) below for the details. For the moment, just suppose that the two conditional probabilities are the same—this is not quite correct, but is good enough for an initial estimate.

This is true for any g. Now suppose $g = (x-L)^+$ for some L. If $L \in (2kh, (2k+2)h]$ for some k, then $\Pi g \equiv g$ outside of the interval $(2kh, (2k+2)h)$, for g is already linear there. Thus the above is

$$= \int_{2kh}^{(2k+2)h} (\Pi g(x) - g(x))\, p(x)\, dx\,.$$

Assume $p(x)$ is Lipschitz and expand it about L: $p(x) = p(L) + O(h)$ in the interval of integration, so that this is

$$= p(L) \int_{2kh}^{(2k+2)h} (\Pi g(x) - g(x))\, dx + O(h^3)\,.$$

We know g and Πg explicitly, so we can do the integral. If $\theta = (L - 2kh)/2h$ this is

$$= 2p(L)\theta(1-\theta)h^2 + O(h^3)\,.$$

There is a similar contribution from the set $T \in (\tau_{J-2}, \tau_{J-1})$. Adding the two, we see that

$$\mathcal{E}_{\text{loc}} = \frac{\sigma^2 T}{n} p(L)\Big(\frac{1}{3} + 2\theta(1-\theta) + O(1/n^{3/2})\Big)\,. \tag{21}$$

A similar calculation in the case g is discontinuous, say $g(x) = I_{(L,\infty)}(x)$ shows why the error is $O(1/\sqrt{n})$) rather than $O(1/n)$ unless $L = kh$ for some integer k.

Since we can write a piecewise $C^{(2)}$ function as a linear combination of a $C^{(1)}$ function which is piecewise $C^{(2)}$, functions of the form $(x-K)^+$, and indicator functions, we can add to get the error for piecewise $C^{(2)}$ functions g. If g is defined by (13), we can rewrite the error in terms of f to get (10).

7 The Trinomial Tree

Let us look at the trinomial tree scheme for the *discounted* stock. We will keep the same notation: the time step is $\delta > 0$, the interest rate is $r \geq 0$, the stock price at time $t = k\delta$ is Y_k, and the discounted stock price is $\tilde{Y}_k = e^{-rk\delta} Y_k$. Set $a > 1$ and $0 < p \leq 1$. Then at each k, $\tilde{Y}_{k+1}$ equals either $\tilde{Y}_k$, $a\tilde{Y}_k$, or $a^{-1}\tilde{Y}_k$, with probabilities $1-p$, $\frac{p}{a+1}$ and $\frac{ap}{a+1}$ respectively. The constants a, p, and δ are related by $p(\log a)^2 = \sigma^2\delta$.

We can embed this in the discounted continuous model $(\tilde{S}_t)$ as follows. Define the embedding times $\tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots$, by induction: Set $\tau_0 = 0$. If $\tau_0, \ldots, \tau_k$ have been defined, toss a biased coin whose probability of heads is p, independent of (S_t) and of the previous τ_j. If tails, let $\tau_{k+1} = \tau_k$. If heads, let

$$\tau_{k+1} = \inf\{t > \tau_k : \tilde{S}_t = a\tilde{S}_{\tau_k} \text{ or } a^{-1}\tilde{S}_{\tau_k}\}\,.$$

Thus, τ_{k+1} is either equal to τ_k, or it is the first time after τ_k that the process goes up or down by a factor of a. Now $(\tilde{S}_t)$ is a martingale, hence $E\{\tilde{S}_{\tau_{k+1}} \mid \tilde{S}_{\tau_k}\} =$

$\tilde{S}_{\tau_k}$, and it is easily seen from this that $(\tilde{S}_{\tau_k})$ has the same transition probabilities as $(\tilde{Y}_k)$, hence the two processes have the same distribution.

Consider the derivative which pays $f(S_T)$ at time T. To price it with the trinomial scheme, choose an integer n, set $\delta = T/n$, and choose p and a so that $p(\log a)^2 = \sigma^2 T/n$, where σ is the volatility of the stock. Then let $u(j,k)$, $j \in \mathbb{Z}$, $k = 0, 1, \ldots, n$ be the solution of the difference scheme

$$\begin{aligned} u(j,k-1) &= e^{-rT/n}\big((1-p)u(j,k) + \frac{p}{a+1}u(j+1,k) + \frac{ap}{a+1}u(j-1,k)\big)\,, \\ u(j,n) &= f(e^{rT}a^j)\,. \end{aligned} \tag{22}$$

Then $u(j,k) = e^{-r(1-\frac{k}{n})T}E\{f(e^{rT}\tilde{Y}_n) \mid \tilde{Y}_k = a^j\}$. The error in this scheme is $\mathcal{E}_{\text{tot}}(f) = e^{-rT}E\{f(e^{rT}\tilde{S}_{\tau_n}) - f(S_T)\}$. Let $h = \log a$ be the space step and let $\hat{\theta}(x) = \text{frac}\,(\frac{x}{h})$ be the fractional part of x/h. Then we have the following. (We have not written out a complete proof of this, so we will call it a remark rather than a theorem.)

Remark 7 *(Trinomial Errors) If $f \in \mathcal{K}$ and $S_0 = s_0$, then*
(i) if f is smooth, there exists a constant A such that

$$\mathcal{E}_{\text{tot}}(f) = An^{-1} + O(n^{-\frac{3}{2}})\,. \tag{23}$$

(ii) If f is a European call or put of strike price K, there exist constants A and B depending on K such that if $\hat{\theta} = \hat{\theta}(\log \tilde{K}) = frac\Big(\frac{\log(\tilde{K}/s_0)}{h}\Big)$,

$$\mathcal{E}_{\text{tot}}(f) = \big(A + B\hat{\theta}(1-\hat{\theta})\big)n^{-1} + O(n^{-\frac{3}{2}})\,; \tag{24}$$

and

$$B = \frac{1}{2}\sigma^2 T K \hat{p}(\log \tilde{K}/s_0),$$

where $\hat{p}$ is given by (9). (iii) If $f(x) = I_{(K,\infty)}(x) + \frac{1}{2}I_{\{K\}}(x)$, then

$$\mathcal{E}_{\text{tot}}(f) = \begin{cases} O(n^{-\frac{1}{2}}) & \text{if } \tilde{K} \notin h\mathbb{Z} \\ O(n^{-1}) & \text{if } \tilde{K} \notin h\mathbb{Z}\,. \end{cases} \tag{25}$$

To see why this holds, let

$$J = \inf\{k : \tau_k > T\}\,.$$

Note that both J and θ are different from those in the binomial scheme. The trinomial scheme doesn't have the odd-even fluctuations that plague the binomial, so we look at all k here, not just even k.

As with the binomial, we take logs and switch to a Brownian measure, which redefines the problem in terms of the Brownian motion $\tilde{X}_t \equiv \log \tilde{S}_t$ and the function g defined in (13). The error term (17) is handled exactly as before, leaving the terms $\mathcal{E}_{\text{glob}}$ and $\mathcal{E}_{\text{loc}}$ defined in (18). The one new technical difficulty is with the term $\mathcal{E}_{\text{glob}}$. In the binomial case, we used the fact that the embedding times were independent of the embedded process. While this is no longer true with the trinomial scheme, the two processes are in fact uncorrelated, and there will still be a constant A such that

$$\mathcal{E}_{\text{glob}} = An^{-1} + O(n^{-\frac{3}{2}})\,. \tag{26}$$

To handle $\mathcal{E}_{\text{loc}}$, note that τ_J is necessarily strictly greater than T, so that $\tau_{J-1} \le T < \tau_J$. If $\tilde{X}_{\tau_{J-1}} = kh$, then τ_J is the first time after T that $\tilde{X}_t$ hits $(k \pm 1)h$. To calculate the distribution of $\tilde{X}_{\tau_J}$ given $\tilde{X}_T = x$, we use the Markov property and Brownian hitting probabilities.

First, the processes $\{\tilde{X}_t,\ 0 \le t \le T\}$ and $\{\tilde{X}_{T+t},\ t \ge 0\}$ are conditionally independent given $\tilde{X}_T$, so that

$$\begin{aligned} P\{\tilde{X}_{\tau_{J-1}} = jh,\ \tilde{X}_{\tau_J} = kh \mid \tilde{X}_T = x\} \\ = P\{\tilde{X}_{\tau_{J-1}} = jh \mid \tilde{X}_T = x\} P\{\tilde{X}_{\tau_J} = kh \mid \tilde{X}_T = x,\ \tilde{X}_{\tau_{J-1}} = jh\}\,. \end{aligned} \tag{27}$$

If $\tilde{X}_T = x \in (kh, (k+1)h)$, then $\tilde{X}_{\tau_{J-1}}$ is either at kh or $(k+1)h$, and it will be at kh if $\tilde{X}_t$ has hit kh more recently than $(k+1)h$. To calculate that probability, we reverse $\tilde{X}_t$ from T: let $Z_t = \tilde{X}_{T-t}$ and note that $\tilde{X}_{\tau_{J-1}} = kh$ if and only if Z_t hits kh before it hits $(k+1)h$. Now (Z_t) is a Brownian bridge, which is essentially a Brownian motion plus drift. If h is small, we can ignore this drift in calculating the hitting probabilities to see that

$$\begin{aligned} P\{\tilde{X}_{\tau_{J-1}} = kh \mid \tilde{X}_T = x\} &\sim \frac{(k+1)h - x}{h}, \\ P\{\tilde{X}_{\tau_{J-1}} = kh \mid \tilde{X}_T = x\} &\sim \frac{x - (k-1)h}{h}. \end{aligned} \tag{28}$$

But now, if $\tilde{X}_{\tau_{J-1}} = kh$, then τ_J is the first time after T that $\tilde{X}_t$ hits either $(k-1)h$ or $(k+1)h$. Since $\tilde{X}_T = x$, $(\tilde{X}_{T+t})$ is a Brownian motion from x, and this happens with probabilities $((k+1)h - x)/2h$ and $(x - (k-1)h)/2h$ respectively. Similarly, if $\tilde{X}_{\tau_{J-1}} = (k+1)h$, $\tilde{X}_{\tau_J}$ equals either kh or $(k+2)h$, with probabilities $((k+2)h - x)/2h$ and $(x - kh)/2h$ respectively. Putting this together with (27) and (28), we see that for any function $g(x)$, if $x \in (kh, (k+1)h)$ and $y = (x - kh)/h$,

$$\begin{aligned} \Pi\, g(x) \stackrel{\text{def}}{=} E\{g(\tilde{X}_{\tau_J}) \mid \tilde{X}_T = x\} \sim \frac{1}{2h^2}\big((h-y)^2 g((k-1)h) \\ + y(2h-y)g(kh) + (h^2 - y^2)g((k+1)h) + y^2 g((k+2)h)\big) \end{aligned} \tag{29}$$

Thus if $p(x)$ is the density of $\tilde{X}_T$, then

$$\begin{aligned} E\{g(\tilde{X}_{\tau_J}) - g(\tilde{X}_T)\} &= \int_{-\infty}^{\infty} (\Pi\, g(x) - g(x))\, p(x)\, dx \\ &= \sum_{k=-\infty}^{\infty} \int_{(k-1)h}^{(k+1)h} (\Pi\, g(x) - g(x))\, p(x)\, dx\,. \end{aligned} \tag{30}$$

It is now easy to compute the local error. If $g \in C^{(1)}$ is piecewise-$C^{(2)}$, it is straightforward to show that each term in (30) is $O(h^2) = O(n^{-1})$, and, as p is integrable, the sum is also $O(n^{-1})$, giving (23).

Now consider the case where $g(x) = (x - \tilde{K})^+$, choose k such that $kh \le \tilde{K} < (k+1)h$, and define θ by $\tilde{K} = (k+\theta)h$. Then $0 \le \theta < 1$. As g is linear outside of $(kh, (k+1)h)$, all but three of the terms in the sum (30) vanish, leaving

$$= p(\tilde{K}) \int_{(k-1)h}^{(k+2)h} \big(\Pi\, g(x) - (x - \tilde{K})^+\big)\, dx + O(n^{-\frac{3}{2}})\,.$$

We have Πg explicitly in (29), so we can do this integral: it is

$$p(\tilde{K})\Big(\frac{1}{3}+\frac{1}{2}\theta(1-\theta)\Big)\,.$$

We can go back to the original function f as we did with the binomial scheme, and combine this with the estimate (26) on $\mathcal{E}_{\text{glob}}$ to get (24). We leave the remaining case, where f is the indicator function of (K,∞), to the reader.

References

[1] Black and Scholes, The pricing of Options and Corporate Liabilities, J. Political Economy 81, 637–659

[2] Cox, J.C., S.A. Ross, and M. Rubinstein, Option pricing, a simplified approach, Journal of Financial Economics 7, (1979) 229–263.

[3] Diener, F. and M. Diener, Asymptotics of the price oscillations of a vanilla option in a tree model, (preprint)

[4] Diener, F. and M. Diener, Asymptotics of the price of a barrier option in a tree model (preprint)

[5] Karatzas, I. and S. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, 1988.

[6] Karatzas, I. and S. Shreve, Methods of Mathematical Finance, Springer-Verlag, 1998.

[7] Leisen, D.P. and M. Reimer, Binomial models for option valuation-examining and improving convergence, Applied Math. Finance 3, (1996) 319–346

[8] Leisen, D.P., Pricing the American put option, a detailed convergence analysis for binomial models, Journal of economic dynamics and control 22, (1998) 1419–1444.

[9] Monroe, Itrel, On embedding right-continuous martingales in Brownian motion, Ann. Math. Stat 43, 1293–1316.

[10] Rogers, L.C.G. and E.J. Stapleton, Fast accurate binomial pricing, Finance and Stochastics 2 (1998), 3–17.

[11] Walsh, J.B., The Rate of Convergence of the Binomial Tree Scheme (preprint)