Simple, fast, and flexible pricing of Asian options

Timothy R. Klassen†

Department of Physics, Pupin Hall, Columbia University, New York, New York 10027, USA

The author describes a modified binomial method that provides a simple and unified framework for the valuation of various kinds of Asian options (American or European, arithmetic or geometric, fixed or floating strike, discrete or continuous sampling and dividends, and partial Asians). The greeks can also be calculated accurately and stably. The method is a refinement of that of Hull and White (1993), where at each node of a standard binomial tree one also considers a table of possible values of the average. To avoid the exponential explosion of the size of this table in the arithmetic average case, one considers a smaller set of representative values for the average, interpolates when necessary, and otherwise uses standard backward recursion to value the option. In this paper, an efficient implementation of this idea is presented. In particular, option values are insured to be smooth as a function of the number N of binomial time periods, so that Richardson extrapolation can be applied to eliminate $1/N$ (and sometimes higher-order) corrections, dramatically increasing the speed of the method. Detailed checks and illustrations are provided, showing that this approach can achieve any desired level of accuracy for convection- or diffusion-dominated regimes and for long or short maturities. It is typically much faster than standard PDE and Monte Carlo approaches.

1. INTRODUCTION

Asian options involve a payoff that depends on the average of the asset price over some prespecified period of time. They are an important class of path-dependent options. For the buyer, they provide a cost-efficient way of hedging cash or asset flows over extended periods, e.g., for foreign exchange, interest rates, or commodities like oil or gold. The averaging feature also makes the payoff less susceptible to large asset price movements, including attempts to manipulate the price just before maturity. For the writer of an Asian option, the advantages include more manageable hedge ratios and the ability to unwind his position more gracefully at the end.

These features make Asian options attractive conservative financial instruments. Whereas other exotic options have lost much of their luster in the last few years, especially in American markets, Asian options have become more and more commonplace. For commodities such as gold and oil the vast majority of options sold in the OTC market are Asian.

Much effort has recently been spent on the pricing of various Asian options. Valuing such options is nontrivial since the standard form of averaging in practice is *arithmetic* rather than *geometric*. With the canonical assumption of a

† Present address: Equity Derivatives, Goldman, Sachs & Co., 1 New York Plaza, New York, New York 10004, USA.

log-Gaussian distribution for stock prices, the geometric average has a simple analytic form (log-Gaussian again, with different mean and variance), whereas the arithmetic average does not. As usual, the early exercise opportunity allowed by American options[1] makes valuation more difficult, often impossible, in most approaches.

For the European case (especially continuous fixed strike options), a variety of methods have been suggested. A partial list includes:

- exact expressions involving Laplace transforms (Geman and Yor 1993) or an infinite sum over recursively defined integrals (Dufresne 1999);
- analytic approximations based on moment matching (Turnbull and Wakeman 1991; Levy and Turnbull 1992) or conditioning on some average (Curran 1994; Rogers and Shi 1995);
- convolution methods using the fast Fourier transform (Carverhill and Clewlow 1990; Reiner 1999);
- Monte Carlo or quasi–Monte Carlo methods (Kemna and Vorst 1990; Fu, Madan, and Wang 1998/9; Glasserman, Heidelberger, and Shahabuddin 1998);
- a number of PDE methods (Ingersoll 1987; Dewynne and Wilmott 1993; Wilmott, Dewynne, and Howison 1993; Rogers and Shi 1995; Zvan, Vetzal, and Forsyth 1998; Zvan, Forsyth, and Vetzal 1997/8; Forsyth, Vetzal, and Zvan 1999; Zvan, Forsyth, and Vetzal 1998; Dempster, Hutton, and Richards 1998; Andreasen 1998; Wilmott 1998) of varying degrees of generality—in some cases, such as with European floating strike options, the PDEs are one- (space) dimensional, but, in general, they are two-dimensional;
- modified (or "thick") tree methods (Hull and White 1993; Barraquand and Pudet 1996).

The exact expressions are difficult to evaluate (Geman and Eydeland 1995; Shaw 1999; Fu, Madan, and Wang 1998/9), especially for small volatilities or short maturities; they also do not allow for many of the features important in practice. Analytic approximations can take into account some of the latter, like discrete sampling for example, but are often not accurate enough. For European options, the other four methods can be applied in a wide range of situations, but can be rather slow. If early exercise is allowed, only the last two, PDE and modified tree methods, are currently practical.

Ultimately, improved finite-difference (FD) or finite-element methods for solving PDEs are presumably the most flexible and efficient way of incorporating many of the details that are required for valuing realistic contracts. For example, in addition to early exercise, the option might involve discrete (cash) dividends, knock-in or knock-out barriers, Parisian-type features, etc. We refer the reader to Wilmott (1998) for an introduction to PDE methods in finance and to Zvan,

[1] In the OTC market, a significant number of Asian options have early exercise features, especially in Europe (and Australia) and for commodities such as gold or oil (J. Andreasen, P. Forsyth, H. Geman, and A. Lipton, private communications)

Forsyth, and Vetzal (1997/8), Forsyth, Vetzal, and Zvan (1999), Zvan, Forsyth, and Vetzal (1998), Dempster, Hutton, and Richards (1998), and Andreasen (1998) for recent studies of some of these features.

Setting up improved FD schemes that can deal with the various sources of financial and numerical complications is, however, not a trivial task. In particular, for path-dependent options there is the problem of *degeneracy*, first pointed out by Barraquand and Pudet (1996). For small volatility, as well as other situations, the pricing PDE is *convection dominated*, leading to numerical problems in the form of spurious oscillations when improved and/or implicit FD schemes are used. The greeks especially can suffer from dramatic and disastrous oscillations as a function of stock price (or strike). For an introduction to these problems and possible remedies, we refer the reader to Zvan, Vetzal, and Forsyth (1998) and references therein.

Given these facts, it is of some interest to have a simple alternative framework. Here we present such a framework using a tree method. We concentrate on options without barriers, for which a binomial tree is sufficient. Barriers, at least not too complicated ones, can be handled by using a trinomial tree.

Note that tree methods *are* special FD schemes. More precisely, they correspond to special explicit FD schemes that are unconditionally stable (a generic explicit scheme is not stable) and will not exhibit the oscillations mentioned above. They also have the great advantage of simplicity. In particular, there are no complications—and no overhead—due to solving linear systems of equations, which is necessary for more sophisticated FD schemes (all of which are implicit schemes). On the other hand, this advantage is offset by the fact that tree methods necessitate a fixed relation between the time and space discretizations (for a trinomial tree there is some flexibility, but not nearly as much as for generic FD schemes). For a given spatial lattice spacing tree methods force one to use a smaller time step than possible for more sophisticated FD schemes.

However, perhaps the biggest (potential) problems with tree methods lie in their convergence properties. Typically, they only converge like $1/N$, where N is the number of time periods in the tree. Even worse, the convergence in the standard approach is oscillatory (not to be confused with the oscillatory behavior as a function of stock price or strike mentioned earlier). This makes it impossible to apply convergence acceleration techniques such as Richardson extrapolation to eliminate the leading $O(1/N)$ corrections. If these corrections could be eliminated, the tree methods would converge as fast as improved FD schemes.

The oscillation problem exists already for simple European vanilla options.[2] It arises because the one scale in the problem, the strike K (or rather K/S, the ratio

[2] As will become clear later, for the Asian case our valuation method takes care of these oscillations in a different manner, so they are not really a problem there. Nevertheless, to counter the widespread belief that the oscillations are inevitable already in the European vanilla case, and to lend credit to the idea that extrapolation should be an efficient tool to improve convergence, we feel it is useful to discuss the vanilla case first.

of the strike and the initial asset price) is, for standard binomial trees, not in a fixed position relative to the nodes of the tree as we vary N. A simple modification of the standard approach eliminates this problem. Deferring details to Section 2, the basic idea is to modify the tree so that the strike is exactly in the middle of the tree at maturity for any N.

For a European vanilla option, the oscillations are completely eliminated and after Richardson extrapolation one finds rapid $1/N^2$ convergence to the (in this case, analytically known) continuum answer. Further extrapolation can improve the convergence to $1/N^3$ or even higher powers. For American options, there is another scale—or rather an infinity of scales—in the problem: the early exercise boundary. It seems impossible to adjust the tree to lie in the same relative position to this boundary for all N (or even a suitable subsequence of N). Nevertheless, the dominant source of oscillations is inappropriate positioning of the strike (especially for the American vanilla options considered in practice, where the early exercise premium tends to be a small fraction of the European price). Hence, the oscillations are very much suppressed in the American case also, allowing extrapolation techniques to become effective.[3]

The complication when valuing arithmetic Asian options in a tree approach is that the number of possible values for the average grows exponentially with the number of time steps in the tree; no recombination takes place as for a geometric average. Hull and White (1993) suggested handling this problem by keeping track only of a smaller number of possible values for the average at each node, using interpolation when intermediate values are needed. This is the approach we will adopt. A crucial question, we will see, is how many values one has to keep track of at each node. For the method to be efficient, we would like to get away with the smallest possible number of values for the average (which will be a function of N, the position of the node, and other parameters) that guarantees convergence to the continuum in a smooth nonoscillatory manner.

In Section 2, we briefly recall the binomial tree method and expand on our discussion of the problem of oscillations for vanilla options. In Section 3, we describe our version of the Hull–White method of pricing Asian options. A number of questions have to be answered before we can provide an efficient implementation of their idea: how to choose the boundaries of the table of average values at each node, how to interpolate between neighboring values, and how the size of the tables should scale with N. After studying theses issues, we present an efficient valuation scheme, which involves the use of Richardson extrapolation to get rid of the (now smooth) leading $1/N$ corrections. For the European case, we can also eliminate the next correction. In Section 4, we perform a number of numerical experiments for European and American Asian

[3] A different but related idea was presented by Broadie and Detemple (1996), namely, to use the continuum Black–Scholes formula in the first step of the backward recursion, just before maturity. This very much reduces, but does not completely eliminate, the oscillations, even in the European case. The reduction is sufficient, however, to make the early exercise boundary the dominant source of oscillations in the American case. After Richardson extrapolation, this method is then virtually identical to ours for American vanilla options.

options with continuous, discrete, or delayed averaging, comparing the efficiency and convergence of our method with others from the literature. We also consider sensitivities (the "greeks"). For European Asian options, our method is much faster than virtually all other convergent methods, including standard PDE methods. For American options, where the convergence of our method is somewhat slower, it is still usually much faster than the only available alternative method, the PDE approach. We present our conclusions in Section 5.

2. BINOMIAL TREES AND OSCILLATIONS

2.1 *Basic Notation and Setup*

We will use language appropriate to a single stock, but it will be obvious how to translate this to cases where the underlying is a foreign exchange rate or commodity, for example.

We consider a Black–Scholes world, with the stock S evolving lognormally under the risk-neutral measure:

$$\mathrm{d}S_t = S_t[(r-q)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t]. \tag{1}$$

Here, r is the interest rate, q the continuous dividend yield, and σ the variance or volatility; W_t is a standard normalized Wiener process. The usual generalizations like a time-dependent r or a time- and asset-dependent σ are possible, but will not be considered here.

We discretize the risk-neutral evolution of S_t on a binomial tree, as shown in Figure 1. If T is the maturity of the option (in years), then the temporal lattice spacing of the tree is $\Delta t \equiv T/N$; where the time slices are labeled $i = 0, \ldots, N$. Given a stock price[4] S_i at time slice $i = t/\Delta t$, the stock at the next time slice $i+1$ can be either up, uS_i, or down, dS_i, with probabilities p_u and $p_d = 1 - p_u$. There are infinitely many choices for these probabilities and the factors u and d that will produce the correct evolution in the continuum limit. A standard choice for the probabilities is

$$p_u \equiv p = \frac{\mathrm{e}^{(r-q)\Delta t} - d}{u-d} \qquad p_d = 1 - p, \tag{2}$$

together with

$$u = \mathrm{e}^{\sigma\sqrt{\Delta t}}, \quad d = \mathrm{e}^{-\sigma\sqrt{\Delta t}} \qquad \text{(CRR)}. \tag{3}$$

As indicated, we will refer to this choice of values of u and d as the CRR values, short for Cox, Ross, and Rubinstein (1979).

Note that the binomial model is recombining for any (time- and stock-independent) choice of u and d, so we can label the nodes of the tree as (i, j), with $i = 0, \ldots, N$ and $0 \leqslant j \leqslant i$, which is short for $(t_i, S_{ij}) \equiv (i\,\Delta t, S_0 u^j d^{i-j})$.

Now consider a vanilla European call or put option with payoff $[\pm(S_N - K)]_+$,

[4] To keep the notation simple, we sometimes label the same quantity in terms of physical time, as in S_t, sometimes in units of lattice time, as in S_i. This should cause no confusion.

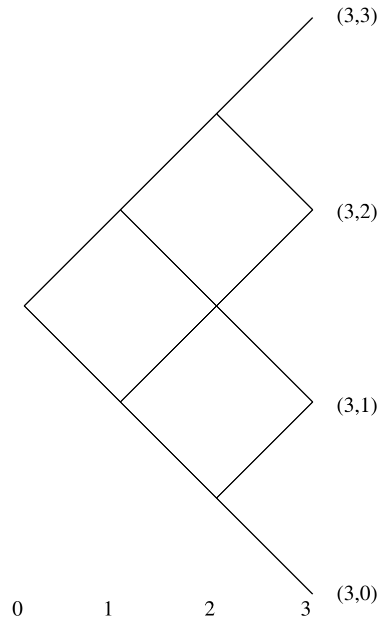

FIGURE 1. A binomial tree with $N = 3$. We show the labeling of the time slices and of the nodes at maturity.

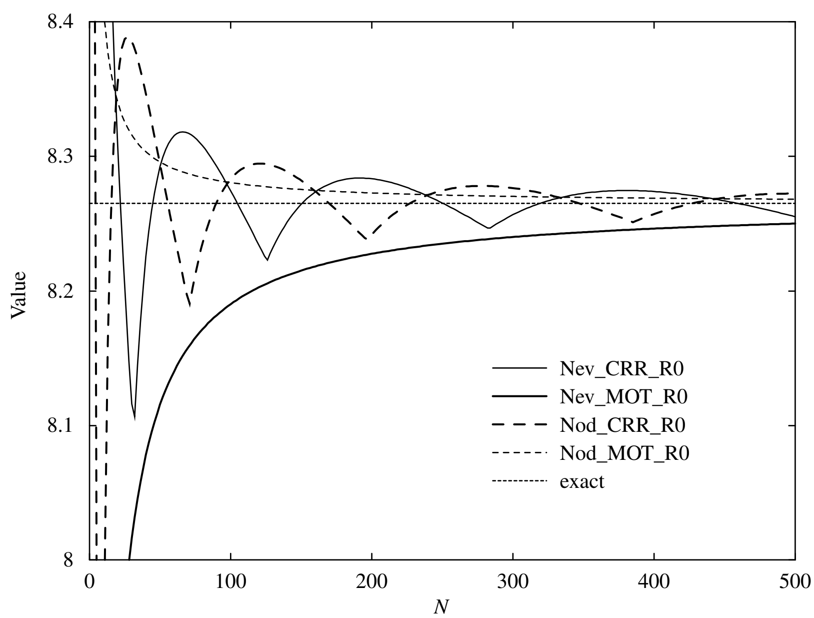

FIGURE 2. European vanilla put option value as a function of the number of binomial time periods ($\sigma = 0.5$, $r = 0.1$, $q = 0$, $T = 4$, $K = 70$, $S = 100$), illustrating the generic presence of oscillations when using the CRR values for u and d, and their absence when using the MOT choice of u and d, and considering even and odd N separately.

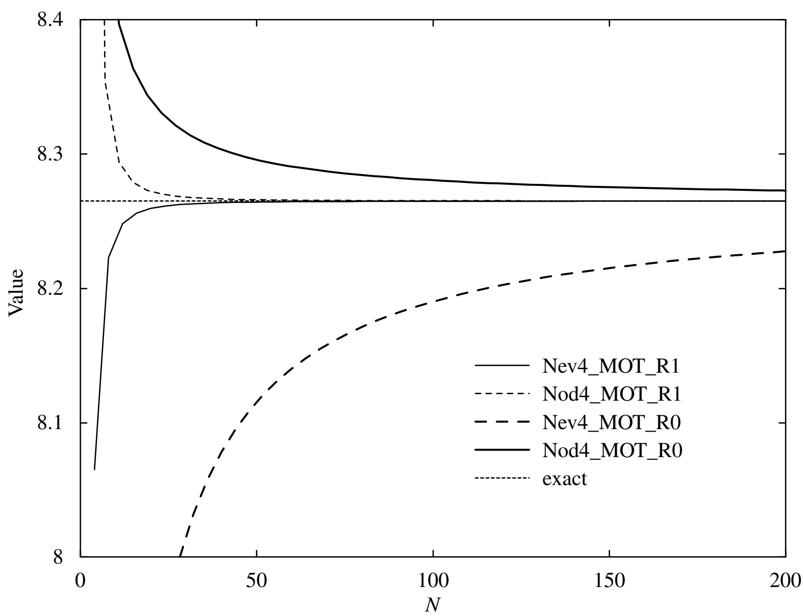

FIGURE 3. As in Figure 2 ($\sigma = 0.5$, $r = 0.1$, $q = 0$, $T = 4$, $K = 70$, $S = 100$), showing that with the MOT choice of u and d one can perform Richardson extrapolation (R1) on even and odd N separately to obtain very fast convergence to the continuum Black–Scholes value of the option.

where $[x]_+ \equiv \max(0, x)$. The notorious oscillatory convergence of tree methods is due to *mismatched scales*: the fixed strike K moves erratically, in general, relative to the nodes of the tree as we vary N. A simple solution to this problem is provided by Leisen (1996) and T. R. Klassen (unpublished) by arranging the up and down steps such that the strike always, for any N, ends up in the middle of the tree at maturity. In other words, choose

$$u = \exp\left(\sigma\sqrt{\Delta t} + \frac{\log(K/S)}{N}\right), \quad d = \exp\left(-\sigma\sqrt{\Delta t} + \frac{\log(K/S)}{N}\right) \qquad \text{(MOT)}, \tag{4}$$

where MOT stands for "middle of tree".

This implies $S(ud)^{N/2} = K$, so that for even N the strike K will lie exactly on the middle node of the tree at maturity, and for odd N it will be halfway between the two middle nodes. This completely eliminates oscillations for European vanilla options, as long as we consider even and odd N separately: the value of the option will lie on separate "smooth"[5] curves for even and odd N.

[5] The notion of smoothness is, of course, not well defined for a function of an integer N. It becomes well defined only for a function of $1/N$ as $N \to \infty$, where we could take it to mean that the function has a well-defined expansion in $1/N$. Actually, although this is the case for many options, in some cases, such as with barrier options (where a trinomial tree method should be used), there are "boundary effects" leading to first-order corrections in $\sigma\sqrt{\Delta t} \propto 1/\sqrt{N}$ (see Broadie, Glasserman, and Kou 1999). In this case, one should consider an expansion in $1/\sqrt{N}$. In general, smoothness could be taken to mean the existence of a convergent or asymptotic expansion in some power of $1/N$, perhaps modified by functions like $\log N$; see below.

Figures 2 and 3 illustrate the above remarks. As an aside, we should point out that the option value obtained using the MOT values of u and d with odd N typically converge much faster than the values for even N, as seen in Figure 2. However, this is a pure $1/N$ effect. After applying Richardson extrapolation[6] to eliminate the $1/N$ terms, the even N series converges just as fast as (and often even a bit faster than) the odd N series (see Figure 3).

In the American case, inappropriate positioning of the tree relative to the strike is usually also the main source of oscillations, but not the only one. The early exercise boundary provides an infinity of scales (as $N \to \infty$), which lead to additional oscillations that are not easy, presumably impossible, to eliminate in a standard tree approach. This effect is particularly pronounced when the early exercise premium is large. Figures 4 and 5 illustrate the oscillations that remain with the MOT choice of u and d. Richardson extrapolation is still effective in the sense that it brings, for most N, the value of the option closer to the continuum limit. But, at the same time, it amplifies the remaining oscillations, as can be seen quite dramatically in the "difficult" case of Figure 5 (difficult because the early exercise premium is very large). Note again the large difference in the even and odd N series before extrapolation, which almost completely disappears once the $1/N$ terms have been eliminated by extrapolation.

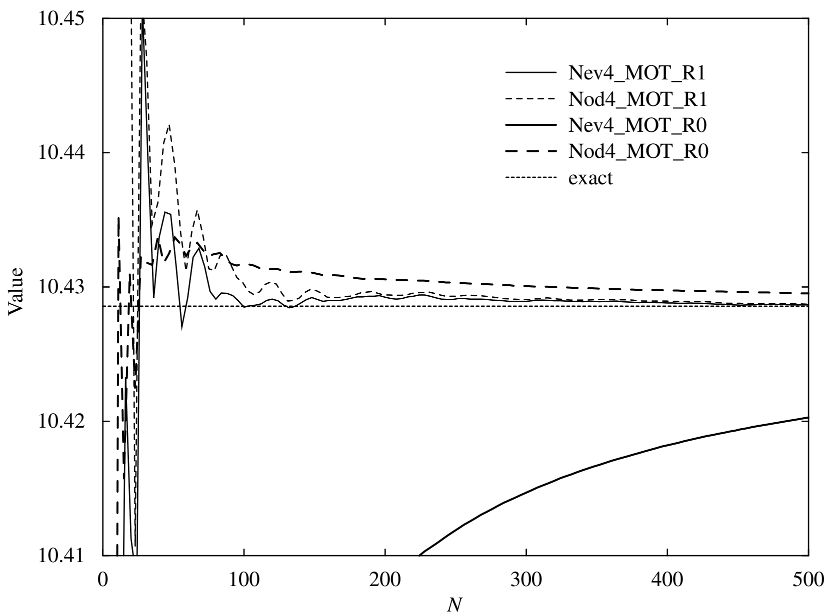

FIGURE 4. The American version of Figure 3 ($\sigma = 0.5$, $r = 0.1$, $q = 0$, $T = 4$, $K = 70$, $S = 100$). Some oscillations remain, even with the MOT choice of u and d, owing to the early exercise boundary. Richardson extrapolation still leads to much faster convergence.

[6] We employ the standard procedure, using the values at N and $\frac{1}{2}N$ (plotted at N) modified to make sure that we do not mix values from the even and odd N series: for even N, we therefore consider only multiplies of 4, $N = 4k$; for odd N, we use values $N = 4k + 3$ and $N = 2k + 1$ in the extrapolation. See Section 3.4 for more details about Richardson extrapolation.

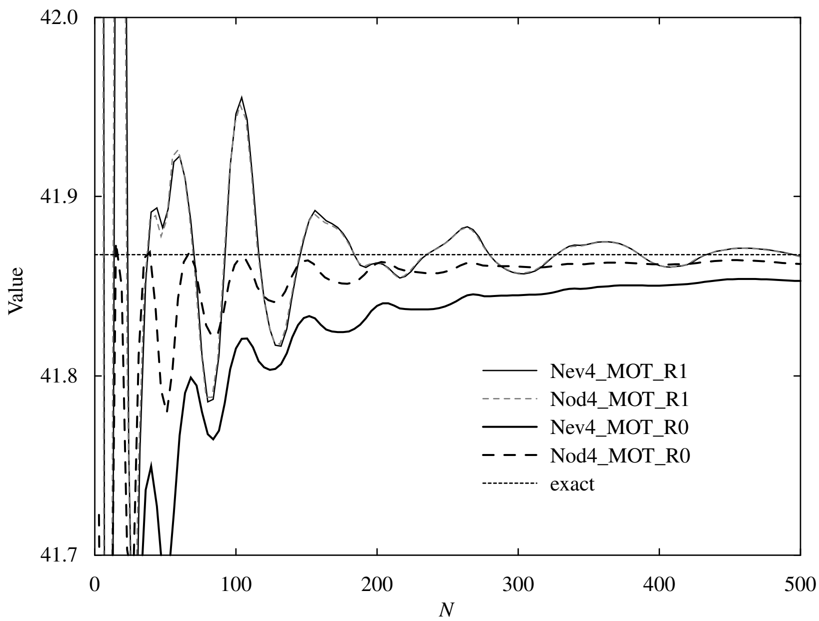

FIGURE 5. Another American vanilla put option ($\sigma = 0.5$, $r = 0.1$, $q = 0$, $T = 4$, $K = 130$, $S = 100$), with parameters chosen (rather unrealistically) to give a very large early exercise premium. This leads to relatively large oscillations that are amplified when performing Richardson extrapolation.

3. BINOMIAL ASIANS

3.1 *The Problem*

The value of a path-dependent option at some time t, depends, by definition, on other properties of the stock price path than just the current stock S_t. For options considered in practice, one often has to track only one more quantity in addition to the stock. Asian options fall into this category, where the quantity in question is some form of average A of the stock over some period in the past.

If the average were geometric, the option could be valued analytically in some situations. If that fails (e.g., in the American case), it would be relatively easy to value on a binomial tree, since a standard geometric average of the stock is *recombining*, i.e., the number of possible values the geometric average can take after N steps is much smaller than 2^N, namely, $1 + \frac{1}{2}N(N+1)$. This number is (just barely) small enough for all possible values of the geometric average to be kept track of up to moderate N. At least, when combined with extrapolation techniques, this might provide a viable strategy to price relatively general geometric Asian options.

For the arithmetic average, on the other hand, each path will generically (for generic values of u and d, that is) give rise to a different value of the average. After N steps, the numbers of possible values of the average when the stock is in the middle of the tree would be, for example,

$$\binom{N}{N/2} \approx \frac{2^N}{\sqrt{N}} \tag{5}$$

for even N. The total number of values of the average at maturity is 2^N. One cannot keep track of this exponentially exploding number of values.

Hull and White (1993) suggested a general strategy to solve this problem. Instead of keeping track of all possible values of the average, we consider only a representative set of values at each node. Otherwise we use standard backward recursion to value the option on a binomial tree; whenever a value of the average is needed that is not in the table of representative values, we use interpolation. The details will be explained below; here we would just like to stress the following.

- By finely discretizing the A direction, i.e., choosing a big table of representative values, any potential mismatch between the strike K and values of the average A is a small effect. The oscillation problem should therefore not be significant. For a floating strike, where the payoff involves $S_T - A_T$, the same applies.
- We can handle both continuously and discretely sampled averages.
- This "naive" way of solving the problem has the advantage that for American-style options it also suppresses discretization errors, and therefore oscillations, from the early exercise boundary.
- Since only one direction has to be finely discretized, the cost of this method should be acceptable.

Note that the mechanism by which oscillations are eliminated is different from the one discussed in the previous section for European vanilla options. Our remarks there were presented mainly for psychological reasons—if we claim to be able to solve the oscillation problem for Asian options, shouldn't we also be able to solve it for vanilla options? As empirical evidence that the mechanism is indeed different, we jump ahead slightly and point out that in this approach the precise choice of values of u and d is not important, i.e., there is hardly any difference between the CRR and MOT values (nor between even and odd N, for that matter).

3.2 *Details of the Binomial Asian Method*

We now describe our implementation of the idea sketched above in more detail.

- Choose a grid spacing h in the $\log A$ direction[7] and, for each node (i, j), consider the values of the option at

$$A_k \equiv A(i, j)_k = A_{\min}(i, j)\, \mathrm{e}^{hk}, \quad k = 0, \ldots, k_{\max}, \tag{6}$$

  where $k_{\max} = k_{\max}(i, j)$ is the smallest integer such that $A_{k_{\max}} \geqslant A_{\max}(i, j)$.

[7] This could be optimized further by choosing a node-dependent $h = h(i, j)$. We have not pursued this idea.

- The maximum $A_{\max}(i, j)$ and minimum $A_{\min}(i, j)$ possible values of the average at a given node are easily exactly determined. The paths giving rise to these extremal values are illustrated in Figure 6.
- Now determine the value of the option by backward recursion. We start at maturity, where the payoff is

  $[\pm(A_T - K)]_+$ for a fixed strike call/put,

  $[\pm(S_T - A_T)]_+$ for a floating strike call/put,

  and determine the value for all A_k for all nodes on the last time slice.
- The recursion step proceeds as follows. For given i, consider sequentially the nodes (i, j), with $j = 0, \ldots, i$. For given A_k, from our table at node (i, j) we know the two possible values A_{up} and A_{dw} of the average at the next time slice $i + 1$, corresponding to an up or down move of the stock. Explicitly, for the simplest case the average is of the form $(S_0 + \cdots + S_i)/(i + 1)$ (see below, however), so that

$$A_{\text{up}} = \frac{1}{i+2}[(i+1)A_i + uS_{ij}], \qquad A_{\text{dw}} = \frac{1}{i+2}[(i+1)A_i + dS_{ij}]. \tag{7}$$

- The values A_{up} andA_{dw} will generically not appear in the "A-tables" at the relevant nodes $(i + 1, j + 1)$ and $(i + 1, j)$. But we can *interpolate* the

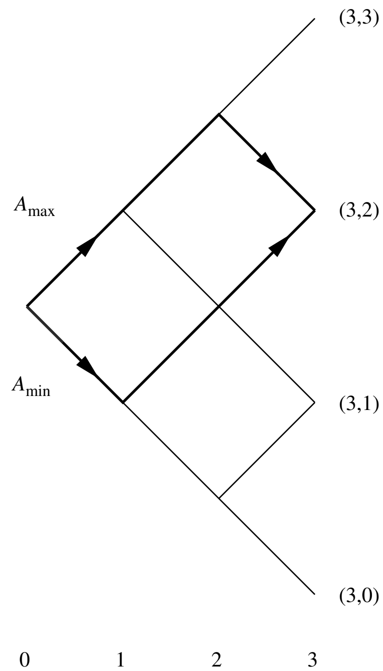

FIGURE 6. A binomial tree illustrating the paths leading to the minimal and maximal possible values of the average at node $(3, 2)$.

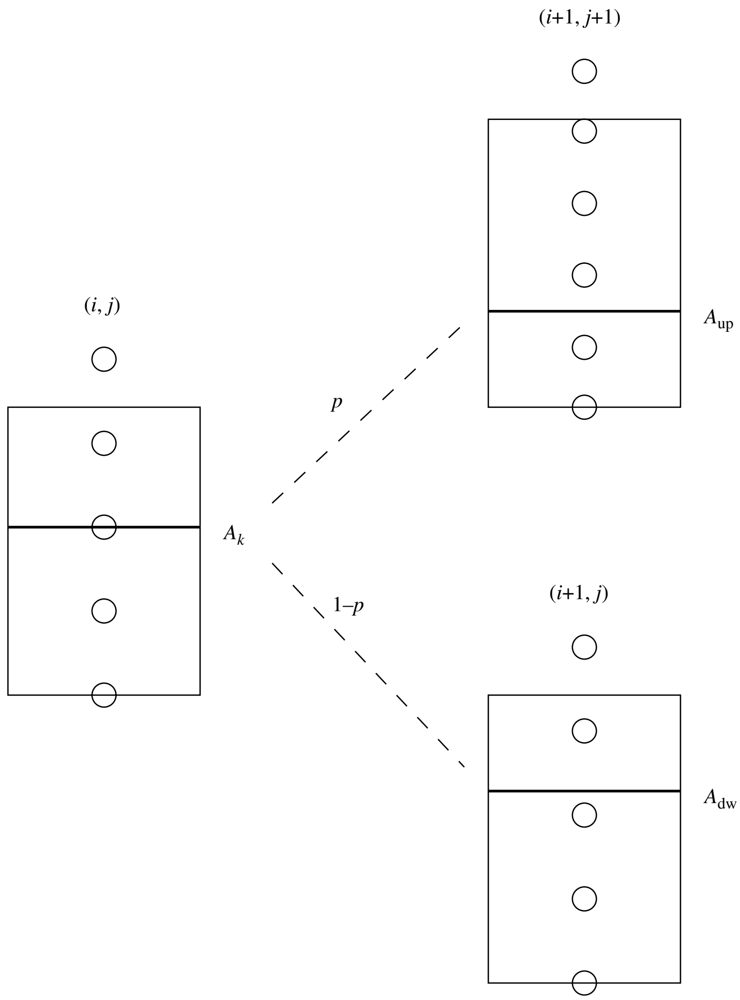

FIGURE 7. Schematic illustration of the backward propagation procedure for the value of an Asian option. The upper and lower edges of a box denote $A_{\max}$ and $A_{\min}$; the circles denote the values of the average considered at a given node.

previously determined values of the option at these nodes to obtain $V(i+1, j+1)$ at A_{up} and $V(i+1, j)$ at A_{dw}. These values will be denoted as $V(i+1, j+1, A_{\mathrm{up}})$ and $V(i+1, j, A_{\mathrm{dw}})$, respectively. We will return to details of the extrapolation below.

- Now we can use the standard backward propagation formula (see Figure 7) for the value of the option:

$$V(i, j, A_k) = \mathrm{e}^{-r\Delta t}[pV(i+1, j+1, A_{\mathrm{up}}) + (1-p)V(i+1, j, A_{\mathrm{dw}})]. \quad (8)$$

- The early exercise allowed by American-type options is also implemented in the standard manner, by replacing $V(i, j, A_k)$ with the maximum of the above expression and the immediate exercise value. For a fixed strike call, for example, the latter would be $[A_k - K]_+$.

- In this manner, we proceed backwards till we reach the root of the tree where only one value of the stock and average is possible, leading to a unique price $V(0, 0, A_0)$ for the option.

Some further remarks are in order. Note that the number of values in the "A-table" at a given node, $N_A(i, j)$, depends strongly on i and j. For given i, it is maximal close to the middle of the tree. Along the edges of the tree, where $j = 0$ or $j = i$, there is only one entry in the table. In the examples we will consider later, where option values are calculated with high or *very* high accuracy, typical values of $N_A(i, j)$ away from the edges are in the range of a few dozen to a few thousand.

As far as efficient memory use is concerned, note (i) that we do not actually need to store "A-tables" but only "V-tables", since we know from equation (6) what A the kth entry in the V-table at node (i, j) corresponds to, and (ii) that we only have to save the V-tables for one time slice at a time.[8]

We should briefly comment on the differences between various implementations of the general Hull–White idea. Besides the original work of Hull and White (1993), the "forward shooting grid" of Barraquand and Pudet (1996) also falls into this class of models. Neither of these groups uses an efficient choice of $A_{\min}(i, j)$ and $A_{\max}(i, j)$. Hull and White (1993) use the same $A_{\min}(i, j)$ and $A_{\max}(i, j)$ for all nodes at a given time slice, basically $A_{\min}(i) = \min_j A_{\min}(i, j)$ and $A_{\max}(i) = \max_j A_{\max}(i, j)$. Barraquand and Pudet (1996) choose the even less optimal $A_{\min}(i) = S_0 d^i$ and $A_{\max}(i) = S_0 u^i$. Our method gains an even larger speed-up from the use of extrapolation techniques. Such techniques work only with a suitable choice for the N dependence of the grid spacing h, an issue not discussed by Hull and White (1993) or by Barraquand and Pudet (1996). We will return to this important point below.

Concerning the interpolation required in the recursion formula, we should point out that, for the last average $A_{k_{\max}}$ at node (i, j), it might happen that A_{up} or A_{dw} are slightly larger than the last entries in the respective tables at time slice $i + 1$ (see Figure 7). In this case, we use an extrapolation rather than an interpolation.

There are some further questions to consider. First of all, what discretization of the average should we use? The answer depends on whether we are interested in a discretely or a continuously sampled average. (In practice, one is always interested in the former, of course, but for benchmarking and comparing different approaches it is useful to consider the continuously sampled average.)

For discrete sampling at times t_a $(a = 1, \ldots, n)$, corresponding to time slices $i_a = t_a/\Delta t$ the relevant average is clearly, by definition,

$$A_n = \frac{1}{n}\,(S_{i_1} + \cdots + S_{i_n}). \tag{9}$$

[8] One does, however, first have to store the values of the option at node (i, j) in a temporary array before one can overwrite the V-table at $(i + 1, j)$ with the new one from node (i, j).

On the other hand, if we are interested in continuous sampling, obtained as a limit $N \to \infty$, it would be better to use the trapezoidal rule and write[9]

$$A_i = \frac{1}{i}(\tfrac{1}{2}S_0 + S_1 + \cdots + S_{i-1} + \tfrac{1}{2}S_i). \tag{10}$$

(To avoid potential confusion, note that the averages in the above equations refer to actually realized values of the average, rather than an entry $A_k = A(i, j)_k$ of an A-table. Correspondingly, the subscripts n and i of A have a quite different meaning from k in A_k, referring to the nth sampling point and ith time slice, respectively.)

The difference between a naive discretization of the continuum average, as in equation (9), and the "improved" one in (10) is $O(1/N^2)$. Since we will use extrapolation methods to eliminate $1/N$ corrections, it is not obvious that there is much point in using an improved discretization of the continuum average. There is *a priori* no reason to believe that (10) has any advantage over a naive discretization as far as $O(1/N^2)$ errors are concerned. It turns out though, as we will see empirically, that the use of (10) is in general a very significant improvement even after first- (and also second-) order extrapolation has been applied.

Another question concerns the interpolation in the V-tables. Should we use linear or quadratic interpolation? Naively, quadratic interpolation is clearly preferable, as it leads to $O(h^3)$ rather than $O(h^2)$ errors (for fixed N). However, the question arises of when these asymptotic laws start to hold. Perhaps there is a lot of "noise"—or oscillatory behavior—for moderate values of h, preventing quadratic interpolation from being effective in practice.

To answer these questions, we have studied the price of Asian options as a function of the grid spacing h for a fixed number of time periods N. An example is shown in Figure 8. This and other examples indicate that, even though there is some "noise" in the option value as a function of h, it is rather small, and for roughly $h < 0.02$ (in the case at hand) the asymptotic expectations hold extremely well. It is amusing to see that for large h (almost too large to be relevant in practice) the unextrapolated option value becomes roughly linear rather than quadratic in h, as illustrated in the right half of Figure 8. Nevertheless, the figure also shows that quadratic interpolation is still very effective in suppressing lattice artifacts due to the discreteness of the A direction.

We will henceforth always use quadratic interpolation between values in the V-tables; this incurs an overhead of only about 20% compared with linear interpolation. If a V-table has only two entries, we use linear interpolation. No interpolation is required if there is only one entry, as happens along the edges of the binomial tree. Recall also that in a few exceptional cases the interpolation is really an extrapolation.

[9] Note that with this choice of A one also has to modify the values of $A_{\min}(i, j)$ and $A_{\max}(i, j)$ in the appropriate manner. A similar remark applies to the calculation of A_{up} and A_{dw} from A_k.

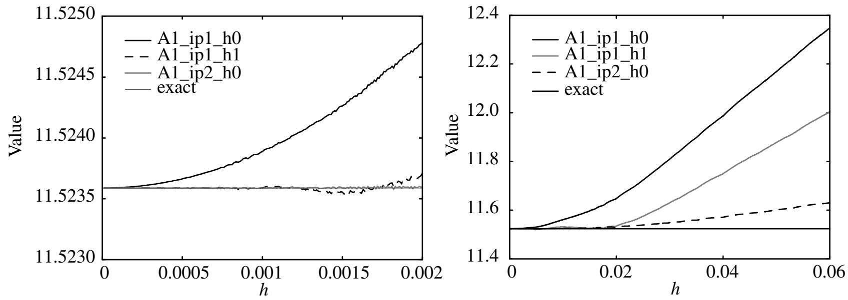

FIGURE 8. Illustration of the h dependence of Asian option values at fixed $N = 40$ for small h (left) and large h (right) ($\sigma = 0.2$, $r = 0.09$, $q = 0$, $T = 0.4$, $K = 90$, $S = 100$). Besides comparing linear and quadratic interpolation (ip1_h0 versus ip2_h0), we also check another method of improving convergence in h. We calculate the value of the option with grid spacings h and $2h$ in the $\log A$ direction and extrapolate to $h = 0$ assuming a quadratic dependence on h. This method (labeled ip1_h1) is much better than linear interpolation but not as good as a direct quadratic interpolation. In all cases, we used the improved discretization of the continuum arithmetic average (denoted A1).

3.3 *Continuum Limit and Convergence in* $1/N$

It was recently pointed out by Forsyth, Vetzal, and Zvan (1999) that one must be careful in how one takes the continuum limit; not any $N \to \infty$, $h \to 0$ trajectory will converge to the correct answer. They also remark that this is a problem for the approach of Barraquand and Pudet (1996), where the choice $h \propto \Delta S \propto \sqrt{\Delta t} \propto 1/\sqrt{N}$ is made. Forsyth, Vetzal, and Zvan (1999) provide numerical evidence that the method of Barraquand and Pudet (1996) converges to values that are not quite correct. Although no rigorous necessary conditions for convergence are spelled out by Forsyth, Vetzal, and Zvan (1999) for the most general case (e.g., the case of quadratic interpolation is not considered), their results suggest that a *sufficient* condition for convergence, also in our case, is to choose asymptotically $h = h(N) \propto \Delta t \propto 1/N \to 0$.

After some experimentation, we decided to use

$$h = h(N) = \frac{h_0}{1 + N/b}, \quad \text{with } b = 100,\ h_0 = a\sigma\sqrt{T}. \tag{11}$$

This is consistent with the desired asymptotic behavior and prevents h from getting too large for small N. (The precise choice is not crucial; only the asymptotic behavior is.) We have factored out the natural scale $\sigma\sqrt{T}$ of the $\log A$ direction by introducing the parameter a above. This has the practical advantage that, for given N and a, the cost of valuing an option will be roughly the same for all options.

Finally, we come to the crucial question of how to perform the extrapolation in $1/N$. This requires some knowledge of the form of the corrections to the $N = \infty$ limit. For a European vanilla option in a binomial (or, more generally, tree) framework, one can argue intuitively that these corrections should take the form of a power series in $1/N$: we know that the price is the solution of a standard PDE. The binomial method corresponds to a certain FD scheme of approximating the continuum PDE. Since diffusion equations smooth out nonsmooth initial conditions instantaneously, we can apply elementary calculus to relate lattice deriatives (a.k.a. finite differences) to continuum derivatives and conclude that the binomial value of the option differs from its continuum value by a power series in the lattice spacing $\Delta t \propto 1/N$.[10] Obviously it would require some effort to make these arguments rigorous.

There are two issues complicating the application of these arguments, even on the above intuitive level, to the Asian options we are interested in. First of all, we do not have a simple binomial model, since there is another variable, the average A, on which the option value depends. This should simply lead to additional errors of order h^2 (h^3) for linear (quadratic) interpolation in the V-tables (plus higher powers of h). With our choice of $h(N)$, these errors are of the same asymptotic form as the ones we have already considered. By comparing with cases where exact results are known, we can consider some of the numerical results presented in the next section for European Asian options as providing strong empirical evidence that the corrections are indeed a power series in $1/N$ (at least for the first few terms).

Secondly, and more critically, options can have various features which imply that their values are *not solutions of standard PDEs*. Barriers or early exercise features, for example, lead to additional inequalities and constraints (like the *smooth-pasting condition* for American options) that qualitatively change the nature of the problem. In the American case, one obtains a so-called *free-boundary problem*, for example. The simple arguments appealed to above for discretized PDEs cannot be applied anymore. As mentioned earlier, for barriers it is at least easy to guess the form of the new corrections, namely, $1/\sqrt{N}$ with a calculable coefficient (Broadie, Glasserman, and Kou 1999), but for American options this is much harder. Based on the representation of the early exercise premium of vanilla options in terms of the early exercise boundary (Carr, Jarrow, and Myneni 1992) it seems plausible that the leading corrections are still $1/N$,[11] but that subleading corrections are larger than $1/N^2$, perhaps involving $\log N$.

Empirically there is little question that the leading corrections are also $1/N$ for American vanilla and Asian options (without barriers), but the form of the subleading corrections is much harder to divine on purely empirical grounds. Note that these questions are relevant for all FD schemes, not just tree methods.

For a more careful discussion of at least the leading corrections in various approaches of pricing Asian options, we refer the reader to Zvan, Forsyth, and Vetzal (1998).

[10] Actually, this argument only shows that the *equations* satisfied by the binomial and continuum values differ by a power series in the lattice spacing. Although perhaps not obvious, it is generally true that then also the *solutions* differ by such a power series; see Thomas (1995), for example.

[11] In fact, this is claimed to have been proved by Leisen (1996), but D. Lamberton (private communication to M. Broadie) has criticized the supposed proof.

3.4 *Extrapolation in Practice*

As the final element of our "accelerated binomial Asian" (ABA) method, we now describe how we perform extrapolations in $1/N$. As discussed above, for both European- and American-style Asian options the leading corrections are of order $1/N$. They can therefore be eliminated by standard extrapolation techniques. Let us briefly recall how this works.

Given an increasing sequence $\{N_i\}$, assume we know that the function of interest has an expansion of the form

$$f(N_i) = f_0 + \frac{f_\alpha}{N_i^\alpha} + \frac{f_\beta}{N_i^\beta} + \cdots, \quad \text{with } \beta > \alpha > 0. \tag{12}$$

We can now form a new sequence $\{(Rf)(N_i)\}$ from $\{f(N_i)\}$ via

$$(Rf)(N_i) \equiv \frac{N_i^\alpha f(N_i) - N_{i-1}^\alpha f(N_{i-1})}{N_i^\alpha - N_{i-1}^\alpha} = f_0 + O(1/N_i^\beta). \tag{13}$$

This procedure can now be iterated to "peel off" the next layer of $1/N^\beta$ corrections. This idea can work with any sort of expansion. It does not have to be an expansion in powers only; it could also involve logarithms, for example. The only requirement is that we know the general form of the expansion. Then the above process can be repeated several times, the only limitation being rounding errors, which eventually become severe. It is therefore important to know the original sequence $\{f(N_i)\}$ quite accurately.

In practice, one typically uses sequences $\{N_i\}$ with $N_i = 2N_{i-1}$ for $i > 1$ and $N_0 = 2$, 3, or perhaps 5. The whole procedure is often referred to as *Richardson extrapolation*. It tends to work amazingly well in improving the convergence of a series with a known expansion.

Back to the problem at hand, we will always use first-order Richardson extrapolation to eliminate the $1/N$ corrections. The series so obtained will be labeled R1; the original, unextrapolated, series will be referred to as R0. For European options, we will also always perform second-order Richardson extrapolation to eliminate the $1/N^2$ corrections, giving a series R2.

For American Asian options, we do not know the form of the subleading corrections, so we will form two series going beyond R1. In one we assume the corrections are $1/N^2$ as in the European case. We know that at least in some cases this is a good assumption, since the value of an American can be arbitrarily close to, or even equal to, its European counterpart. Largely out of curiosity, we also try the assumption that the subleading corrections are $1/(N \log N)$; this is about the largest possible subleading correction one can imagine. We label the resulting sequence R1l.

Note that making the wrong assumption about the subleading corrections does not destroy the convergence of the series to the right answer. In the worst case, it might slow the order of convergence, but usually not even that; only the subleading rather than leading corrections of the series will be affected. Unless the original series is not very accurately known in a numerical sense, it is therefore quite harmless—and often helpful, as we will see—to play with various reasonable guesses about subleading corrections whose form is not precisely known.

3.5 *The Greeks*

For vanilla options, there are obvious ways of estimating the greeks Δ, Γ, Θ within a binomial tree at no extra cost (see, e.g., Hull 1997; Shaw 1999). This approach gives estimates of the three greeks at time $t = \Delta t$, $2\Delta t$, and Δt, respectively. One of the nice side-effects of our $1/N$ extrapolation is that it also automatically extrapolates the greeks to $t = 0$.

One must be more careful for path-dependent options, since now the option price depends not just on the stock S. For Asian options, the greeks Δ and Γ, for example, are defined as derivatives with respect to S at fixed A. Since S and A vary from node to node in a correlated manner, the calculation of the greeks within a binomial tree is now more subtle. There are various ways around this problem. One is the standard approach of extending the tree slightly to negative times, so that there are three nodes at time slice $t = 0$, which can all be taken to be at the same average A_0. Another is to set up a backward recursion scheme for the derivative of the option value with respect to some parameter in exactly the same way as for the option value itself (e.g., the value at maturity is obtained by taking the derivative of the payoff, etc.).

We will not pursue these matters here, but just point out the simplifications that arise for the case of continuous sampling. In this case, one can initialize the option with any value for the average. It will have no effect in the continuum limit; its contribution is drowned out by the infinitely many other values contributing to the running average. The effect of this is that we do not have to worry about taking the S derivatives in Δ and Γ at fixed A. For Δ, for example, we can use the standard formula to calculate it "from within" a binomial tree as

$$\Delta(t = T/N) = \frac{V\big(1, 1, A(1, 1)_0\big) - V\big(1, 0, A(1, 0)_0\big)}{(u - d)\, S_0}, \tag{14}$$

in terms of the option values at nodes $(1, 0)$ and $(1, 1)$. The desired greek, $\Delta(t = 0)$, differs from the above expression by an $O(T/N)$ term that can be eliminated by extrapolation.

Other sensitivities can always be calculated as a finite difference from neighboring values of the parameter in question. Since our method does not suffer from spurious oscillations or such problems, these finite differences will be quite accurate.

4. NUMERICAL STUDIES OF VARIOUS ASIAN OPTIONS

The first numerical studies we performed were, of course, directed towards verifying the correctness of our method and its implementation. The code was written in C++, using the Standard Template Library (STL) for convenient, safe, and efficient handling of numerical and character arrays (as represented by the `vector<double>` and `string` classes). No particular effort was made to optimize the code (besides avoiding obvious inefficiencies); we aimed more for flexibility to allow easy exploration of various algorithmic and financial parameters. Our numerical results quoted below were obtained on a 450 MHz Pentium II PC running RedHat Linux 5.2.[12]

In our numerical studies, we used three interleaving power-of-2 sequences:

$$N = 3, 4, 5,\ \ 6, 8, 10,\ \ 12, 16, 20,\ \ 24, 32, 40,\ \ 48, 64, 80,\ \ 96, 128, 160, \ldots \qquad (15)$$

(in some cases we start the series at $N = 6$). Richardson extrapolation is performed on each such subsequence separately. For example, the first value in the R2 series would be for $N = 12$ and use results from $N = 12, 6, 3$; the next for $N = 16$ would use $N = 16, 8, 4$, and so on.

In this first, exploratory, study of our method, it is useful to consider such a large number of N values, because plateaux (as a function of N) in the extrapolated series serve as a useful check of convergence and aid in providing error estimates for the final results. For real-world applications, one might decide that second-order Richardson extrapolation on the three points $N = 32, 16, 8$, or even first-order extrapolation on $N = 16, 8$, say, might provide sufficient accuracy.

We will use annualized units for option parameters. All options are initiated at time $t = 0$ and mature at $t = T$. The initial stock price will be denoted by S (sometimes S_0, for clarity). We will discuss representative examples of the various possible Asian options we can value with our approach: European- or American-style exercise, fixed or floating strike, discrete or continuous sampling, delayed averaging, and at- or off-the-money strikes. Obviously, not all combinations can be considered; we will also not discuss discrete dividends here, though there is no problem introducing discrete proportional dividends into our framework.

We will also not consider the case where the option was initialized at some time before the current $t = 0$. The payoff will then involve a weighted sum of the known average A_0 plus additional contributions from later times. It is easy to see that (for fixed strike options) this case can be reduced to an option initialized at $t = 0$ by modifying the strike and correcting the option price by a multiplicative factor.

4.1 *European Asian Options*

In this section and the following one, we will be interested in continuous sampling; discretely sampled Asian options are considered in Section 4.4.

The only exactly known results exist in the European case: geometric average Asian options and zero strike Asian options. It was checked that the easily

[12] CPU time comparisons with results from other methods obtained by other researchers on different hardware have to be taken with a grain of salt, of course. We quote all information concerning hardware provided by the authors when referring to other results. In any case, the quoted CPU times should provide a rough comparison on the order-of-magnitude level.

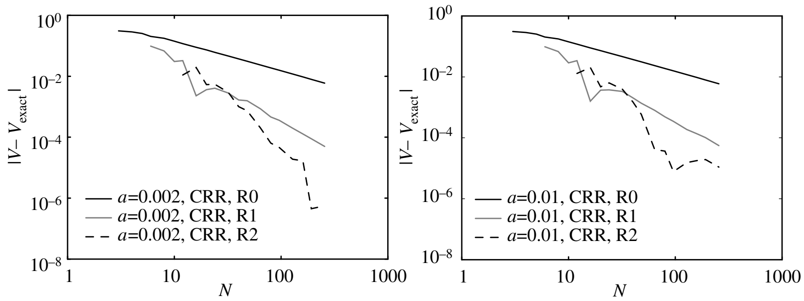

FIGURE 9. Convergence of geometric Asian option value ($\sigma = 0.3$, $r = 0.1$, $q = 0$, $T = 1$, $K = 110$, $S = 100$), calculated with the accelerated binomial method, towards the exact value with no (R0) and first-order (R1) or second-order (R2) Richardson extrapolation. The exact value is approximately 4.440 155 210 for the parameters indicated. The only difference between the left and right plots is the grid spacing in the $\log A$ direction; it is five times smaller on the left for a given N.

derived exact results for geometric Asians options (Kemna and Vorst 1990) could be reproduced to high precision by our method (with the obvious modifications in Section 3 for the case of geometric averaging). Note that in our approach the case of geometric averaging is in no way simpler than that of arithmetic averaging, so the geometric case provides a nontrivial check of the method and its convergence properties.

In Figure 9, we show the convergence of a typical case towards its exact value with no (R0) or first-order (R1) and second-order (R2) extrapolation. For given N, the grid spacing h differs by a factor of 5 between the two plots in this figure. One hardly sees a difference in R0 or R1 between the two curves, whereas for R2 the option value at $N = 256$ is 20 times closer to the exact value if one uses the smaller value of h. The results from the smaller values of h (left plot) are consistent with $1/N^{k+1}$ corrections after kth-order extrapolation is applied. (One might have hoped for a smoother power-law behavior in this figure. Perhaps this is expecting too much: for small N, the asymptotic behavior does not apply; for large N, and higher-order extrapolation rounding errors become relevant.)

Asymptotically, for very large N and in the absence of rounding errors, the curves in both plots will converge with the same order to the exact value. We see, however, that choosing too large a value of h results in erratic and slow convergence (for R2) for accessible values of N. Roughly speaking, choosing too large a value of h introduces "noise". In this sense, Figure 9 illustrates our earlier remark about the importance of accurate initial values when attempting to apply higher-order extrapolation methods. By the same token, however, this figure shows that if we are interested in moderate accuracy, on the 1 cent level, say, it might be more efficient to use a larger value of h (and perhaps just R1 instead of R2): an accuracy of \$0.01 is obtained about five times faster in the right than in the left panel of Figure 9.

Next we turn to the case of arithmetic Asian options with zero strike (and from now on "Asian" will always mean "arithmetic Asian", unless indicated

TABLE 1. Zero strike European Asian call values with $r = 0.1$, $q = 0$, $S = 100$. Besides the analytic result, we show our accelerated binomial Asian (ABA) values and results from Barraquand and Pudet (1996) (BP), who used their forward shooting grid method, Zvan, Forsyth, and Vetzal (1997/8) (ZFV), who applied a PDE approach (Crank–Nicolson with flux limiter), and Dempster, Hutton and Richards (1998) (DHR), who used a highly tuned LP approach to solve discretized PDEs. BP and ZFV used DEC Alphas in their calculations, and DHR an IBM RS6000/590 Power PC.

σ	T	Analytic	ABA	BP	ZFV	DHR
0.1	0.25	98.760351887	98.760351887(0)	98.780	98.770	98.761
	0.5	97.541150999	97.541150998(1)	97.579	97.558	97.541
	1.0	95.162581964	95.162581961(3)	95.232	95.186	97.163
0.2	0.25	98.760351887	98.760351886(1)	98.781	98.769	98.761
	0.5	97.541150999	97.541150997(2)	97.580	97.552	97.542
	1.0	95.162581964	95.162581960(4)	95.239	95.166	97.163
0.4	0.25	98.760351887	98.760351884(3)	98.785	98.760	98.747
	0.5	97.541150999	97.541150992(7)	97.587	97.541	97.514
	1.0	95.162581964	95.16258194(2)	95.277	95.186	95.099
CPU			0.9 s	15 s	40 s	4.3 s

otherwise). Here the exact value is trivially obtained as

$$V = \mathrm{e}^{-rT}\left\langle \frac{1}{T}\int_0^T S_t \,\mathrm{d}t \right\rangle = S_0\, \mathrm{e}^{-rT}\, \frac{\mathrm{e}^{(r-q)T} - 1}{(r-q)T}, \tag{16}$$

where the expectation value $\langle \cdots \rangle$ is taken in the risk-neutral measure. Note that the exact answer is independent of σ. In Table 1, we show the exact answer, results obtained with our accelerated binomial Asian method in 0.9 s of CPU time (using $a = 0.002$ in equation (11)), and results from the literature.[13] It is clear, and confirmed by other results not reported here, that for a given accuracy our method gives answers many orders of magnitude faster.

A graphical presentation of our results for zero strike Asians is given in Figure 10. We show results using the "naive" (A0) and the improved (A1) discretization of the continuous arithmetic average on the same scale. Amusingly, and at first sight confusingly, even though for the unextrapolated results the improved average is clearly converging much faster, this is not the case after first-order extrapolation has been applied. The resolution of this paradox is that the only $1/N$ errors in this case come from the unimproved average. In other words, the R0 results in the right-hand plot of Figure 10 only have $O(1/N^2)$

[13] We should point out that a set of typographical errors has propagated through the literature on Asian option pricing. Barraquand and Pudet (1996) give the correct formula for zero strike option values, but in their Table 7, where numerical results from their method are compared with the exact analytical values, the latter contain typographical errors in the second or third digit after the decimal point. Several sources quote these false values for the analytic result.

corrections to start with; when extrapolating assuming they are $O(1/N)$, one still has $O(1/N^2)$ corrections left (they happen to change sign, though). When second-order extrapolation is applied, both the improved and unimproved average lead to results that are virtually identical (up to the precision we output the results) to the exact answer. This holds as long as one uses sufficiently small grid spacings h in the $\log A$ direction, which our choice of $a = 0.002$ assures.

The absence of $1/N$ corrections when using the improved discretization of the average is clearly a nongeneric peculiarity of this example. The fact that the improved average leads to numerically smaller corrections *is* generic; we have observed it in all other cases of arithmetic Asian options we investigated. This property also seems to hold after first- or second-order extrapolation has been applied.

Let us now turn to generic European Asian options. A first check was provided by a comparison with a set of fixed strike Asians studied by Broadie and Kou (private communication) using the quasi–Monte Carlo method. The set covered a large range of strikes, long and short maturities (more precisely, large and small values of $\sigma^2 T$), as well as convection- and diffusion-dominated regimes ($r \gg \sigma^2$ and $r \ll \sigma^2$, respectively). With second-order Richardson extrapolation using $N = 256$, 128, 64 (but no variance-reduction techniques) the quasi–Monte Carlo method achieves a relative accuracy of about 10^{-7}. Our method gives perfectly consistent results about 500 times faster.

In Table 2, we consider two cases studied in some detail by the Waterloo group (Forsyth, Vetzal, and Zvan 1999). One case is a short-term low-volatility option; the other, a long-term high-volatility one. Forsyth, Vetzal, and Zvan (1999) used a PDE approach and compared it with the methods used by Barraquand and Pudet (1996) and Hull and White (1993).

For the first case, the PDE approach of Forsyth, Vetzal, and Zvan (1999) gives values of \$1.8478 in 1.9 s and \$1.8509 in 1016 s (all their calculations were performed on a Sun Ultraspace workstation). With first-order Richardson extrapolation, the results in their Table 4 lead to \$1.8506 in 24 s and \$1.8515 in 1141 s. Second-order extrapolation leads to \$1.8515 in 1163 s. For the second

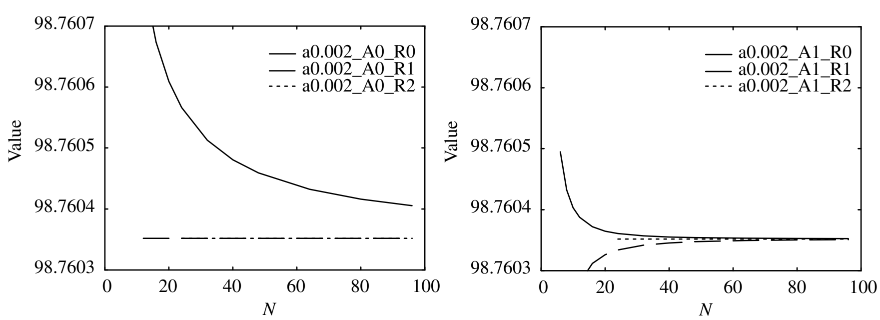

FIGURE 10. Convergence of ABA for a zero strike European Asian call option ($\sigma = 0.2$, $r = 0.1$, $q = 0$, $T = 0.25$, $S = 100$). The only difference between the two plots is the use of the unimproved (A0, left) versus improved (A1, right) discretization of the continuous arithmetic average, as discussed in Section 3. The CRR values of u and d were used.

TABLE 2. European Asian fixed strike calls with $r = 0.1$, $q = 0$, $K = S = 100$ (cases 1 and 2 of Forsyth, Vetzal, and Zvan 1999). We show the convergence of results from the ABA method with first- and second-order Richardson extrapolation. The CPU times quoted are the cumulative run times for the three N values needed for the R2 results; the times for R1 results are slightly lower. Our final result and error is 1.8515926(1) for the first and 28.405169(3) for the second case.

σ	T	N	a	ABA R1	ABA R2	CPU
0.1	0.25	12	0.100	1.8510778	1.8478955	0.004 s
		12	0.010	1.8514232	1.8495053	0.03 s
		16	0.010	1.8519546	1.8458018	0.08 s
		32	0.010	1.8517046	1.8516213	0.53 s
		64	0.010	1.8516163	1.8515868	4.13 s
		128	0.010	1.8516005	1.8515953	29.8 s
		128	0.002	1.8515982	1.8515926	151.9 s
		256	0.010	1.8515930	1.8515905	242.1 s
		256	0.002	1.8515940	1.8515926	1267 s
0.5	5	12	0.100	28.361909	28.345165	0.004 s
		12	0.010	28.415007	28.413037	0.03 s
		16	0.010	28.392752	28.356624	0.07 s
		32	0.010	28.402836	28.406197	0.46 s
		64	0.010	28.404460	28.405002	3.21 s
		128	0.010	28.405011	28.405194	24.1 s
		128	0.002	28.405016	28.405175	125.5 s
		256	0.010	28.405126	28.405165	203.7 s
		256	0.002	28.405130	28.405168	1045 s

case, they quote values of \$28.3573 after 2.2 s and \$28.4003 after 911 s. With first-order Richardson extrapolation, their Table 4 gives \$28.4111 in 17 s and \$28.4054 in 1023 s. With second-order extrapolation, their table gives \$28.4051 in 1038 s.

In both cases, our method gives comparable results several hundred times faster. Table 2 also confirms our earlier remark that the most efficient way to obtain results of moderate precision (penny-level or somewhat better) is to use large $a \approx 0.01$–0.05 and perhaps only first-order rather than second-order Richardson extrapolation. If high precision with a reliable error bar is desired, the most efficient approach is to use small $a \leqslant 0.005$ and second-order Richardson extrapolation. Although the results will also converge with larger a to the continuum answer as $N \to \infty$, they tend to do so in an oscillatory manner. At $N = 256$, say, they *might* be correct to the desired level of accuracy as for smaller a, but because of the oscillations it is hard to be sure about this. In other words, a realistic and honest error bar would be significantly larger than for smaller a where the convergence of R2 results is monotonic, as long as N is not too small. This is illustrated in Figure 11.

Very accurate results for European fixed strike Asians were recently obtained by Zhang (1999) using an approach inspired by perturbation theory. To avoid

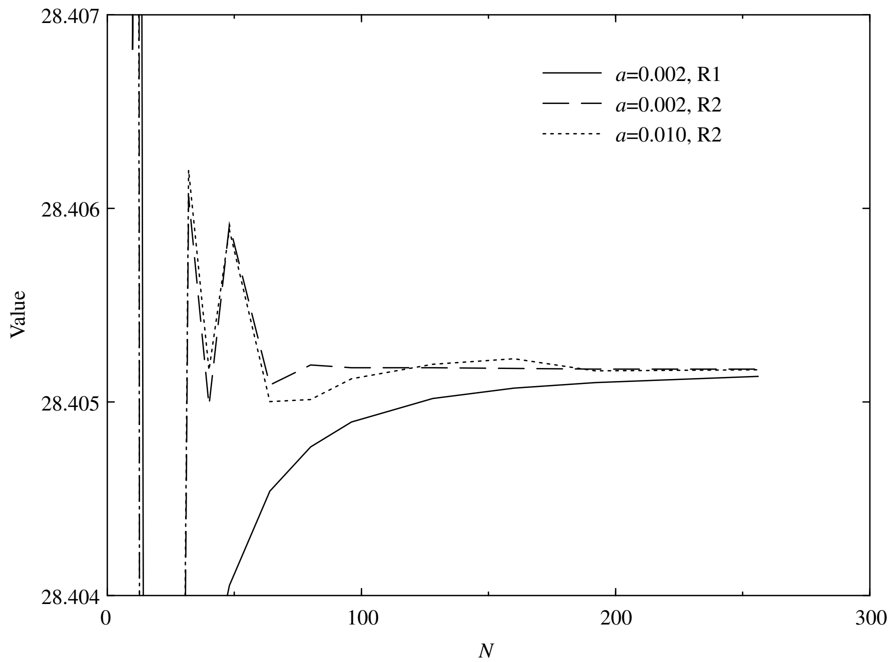

FIGURE 11. European Asian call option with fixed strike ($\sigma = 0.5$, $r = 0.1$, $q = 0$, $T = 5$, $K = 100$, $S = 100$), illustrating the (small) oscillations around the continuum value that occur for large values of a (compare the $a = 0.01$ R2 results with those for $a = 0.002$ R2).

the numerical problems associated with solving a PDE with nonsmooth payoff, he found a slightly different PDE with the same payoff which can be solved analytically. He shows that the remainder satisfies another PDE with smooth payoff (in fact, the payoff is 0), which is easily solved numerically to high precision. His results present, it seems, the current "world record" for European fixed strike Asians. As Table 3 shows, our much more general and flexible method is not far behind.

To conclude our discussion of European Asian options, we present some results for floating strike options. In Table 4, we show results for a case

TABLE 3. European fixed strike Asian calls with $T = 1$, $r = 0.09$, $q = 0$, $K = S = 100$. For comparison with our and Zhang's (1999) method, we also give results from the Turnbull and Wakeman (1991) approximation.

σ	ABA, $N = 256$	ABA, $N = 48$	Zhang	TW
0.05	4.30823352(2)	4.3082334	4.30823	4.3097
0.20	6.7773474(4)	6.777345	6.77735	6.8035
0.30	8.828758(1)	8.82872	8.82876	8.8858
0.50	13.028156(2)	13.0280	13.02817	13.2120
CPU	1100 s	2.8 s	1.5 s	

TABLE 4. European Asian floating strike put with $T = 0.25$, $\sigma = 0.4$, $r = 0.1$, $q = 0$, $S = 100$. Our final estimate of the price is 3.972036(1).

a	N	Value R1	Value R2	CPU
0.010	12	3.967525	3.957133	0.03 s
	16	3.973035	3.948535	0.07 s
	32	3.972428	3.972225	0.49 s
	64	3.972124	3.972022	3.46 s
	128	3.972049	3.972025	27.5 s
	256	3.972039	3.972036	240.1 s
0.005	12	3.967606	3.957232	0.06 s
	16	3.973272	3.948961	0.14 s
	32	3.972379	3.972081	1.03 s
	64	3.972123	3.972038	7.31 s
	128	3.972056	3.972034	56.5 s
	256	3.972041	3.972036	486.1 s

considered by Zvan, Forsyth, and Vetzal (1998), using a PDE method (Crank–Nicolson with van Leer flux limiter). On the finest grid considered, they obtain \$3.9653 in 722 s (on a Pentium-Pro 200 PC). Unfortunately, it is not possible to perform a proper continuum extrapolation with the results given in their Table 5, but rough estimates are certainly consistent with our result \$3.972036(1).

The problem of pricing a European Asian floating strike option can be reduced to a one-dimensional PDE (one space in addition to the time dimension; the results referred to in the previous paragraph used a two-dimensional PDE). Despite its low dimension, this PDE is difficult to solve, since it is convection dominated in most cases of interest in practice. Zvan, Forsyth, and Vetzal (1997/8) used flux-limiter techniques to alleviate this problem. Their results are compared with ours in Table 5. We can infer that the accelerated binomial method is at least 2–3 orders of magnitudes faster in reaching a given accuracy.

TABLE 5. European Asian floating strike put options with $T = 1$, $r = 0.15$, $q = 0$, $S = 100$. Besides our results with second-order Richardson extrapolation (we used $a = 0.005$ for $N = 256$, and $a = 0.01$ for $N = 48, 32$), we show those of Zvan, Forsyth, and Vetzal (1997/8) obtained by solving the one-dimensional PDE of Rogers and Shi (1995) using the second-order accurate Crank–Nicolson scheme with a flux limiter. Their calculations were performed on a DEC Alpha workstation.

σ	ABA, $N = 256$	ABA, $N = 48$	ABA, $N = 32$	PDE (ZFV)
0.1	0.2516754(3)	0.25170	0.2519	0.244
0.2	1.7101883(6)	1.71023	1.7105	1.708
0.3	3.609811(2)	3.60972	3.6098	3.609
CPU	503 s	1.71 s	0.51 s	27.3 s

4.2 *American Asian Options*

For the American case, there are no exact results for Asian options, so they cannot be used to check our method and implementation. However, the early exercise feature amounts to a simple one-line addition to the code in the binomial-type approach we are using. Given that our method has been extensively checked in the European case, we can therefore be fairly confident *a priori* about its correctness for American-style options. We can also compare the results with those in the literature, though there are far fewer results for this case: besides modified binomial methods, the only currently practical approach for American Asians is the PDE framework.

What complicates such comparisons is that often different types of discretization errors remain in varying degrees and are difficult to disentangle. We will discuss this issue in more detail in Section 4.4 on discretely sampled Asians, for which it is even more acute.

In Tables 6 and 7, we compare our method for fixed strike American Asian options with results from the PDE approach of Zvan, Forsyth, and Vetzal (1997/8) and the method of Barraquand and Pudet (1996). Our method is again orders of magnitude more efficient, especially for small $\sigma^2 T$. It seems that the PDE results do not actually converge to (what we consider) the correct answer; consider the $K = 100$ case in Table 6, for example. This is presumably due to the fact that the same time step was used when varying the spatial grid size in Table 7 of Zvan Forsyth, and Vetzal (1997/8). As mentioned earlier, Forsyth, Vetzal, and Zvan (1999) concluded that the method of Barraquand and Pudet (1996) does not converge to the correct continuum answer.

By comparing the results with those from the European case, it is immediately clear that the convergence is slower for American-style options. This is due to the fact that the subleading corrections are, in general, larger than $1/N^2$, as discussed in Section 3. To estimate the continuum limit and an error bar, we therefore consider both the R1l and R2 extrapolated series described in Section 3. We usually find that they converge from opposite directions, allowing us to bracket the limiting value. This is illustrated in Figure 12.

TABLE 6. American Asian fixed strike call options with $T = 0.25$, $\sigma = 0.2$, $r = 0.1$, $q = 0$, $S = 100$. Our ABA results use $a = 0.01$. For the accurate $N = 512$ results, we consider both the R1l and R2 extrapolations to estimate central value and error bar; the $N = 48$ results come from R2. For comparison, we show PDE results (Crank–Nicolson with flux limiter) of Zvan, Forsyth, and Vetzal (1997/8) at the indicated $A \times S$ grid sizes, as well as from the method of Barraquand and Pudet (1996). CPU times quoted for these two methods are from runs on DEC Alpha workstations.

K	ABA, $N = 512$	ABA, $N = 48$	PDE (161×177)	PDE (41×45)	BP
95	7.4660(5)	7.4573	7.465	7.521	7.371
100	3.21587(3)	3.2169	3.222	3.224	3.219
105	0.988153(2)	0.9885	0.992	1.009	1.001
CPU	2410 s	1.6 s	664 s	43 s	15 s

TABLE 7. American Asian fixed strike call options with $T = 1$, $\sigma = 0.4$, $r = 0.1$, $q = 0$, $S = 100$. Our ABA results use $a = 0.01$. For the accurate $N = 512$ results, we consider both the R1l and R2 extrapolations to estimate central value and error bar; the $N = 48$ results come from R2. For comparison, we show PDE results (Crank–Nicolson with flux limiter) of Zvan, Forsyth, and Vetzal (1997/8) and results from the method of Barraquand and Pudet (1996).

K	ABA, $N = 512$	ABA, $N = 48$	PDE (41×45)	BP
95	15.7747(4)	15.761	15.749	15.649
100	12.5094(2)	12.495	12.497	12.439
105	9.83047(5)	9.8275	9.825	9.790
CPU	2230 s	1.6 s	43 s	15 s

We will not consider further continuously sampled American Asians here. In all the cases we investigated, including floating strikes, we find that the accelerated binomial method is considerably faster than other methods. We will discuss the interesting case of discretely sampled American Asian options in Section 4.4.

4.3 *Delayed Asian Options*

In practice, the averaging of the asset is often only applied during part of the lifetime of the option. Consider, for example, an investor with the view that the underlying will rise in the future. He might want to buy a call option to profit from the anticipated rise of the stock. If some averaging is applied towards the end, the effective volatility, and therefore the price of the option, will be lowered (with a relatively smaller decrease in the effective rise of the partially averaged asset). For the writer of the option, averaging at the end leads to more manageable hedge ratios. He can also unwind his position gracefully as expiry approaches. The other benefit of Asian options, suppressing the impact of

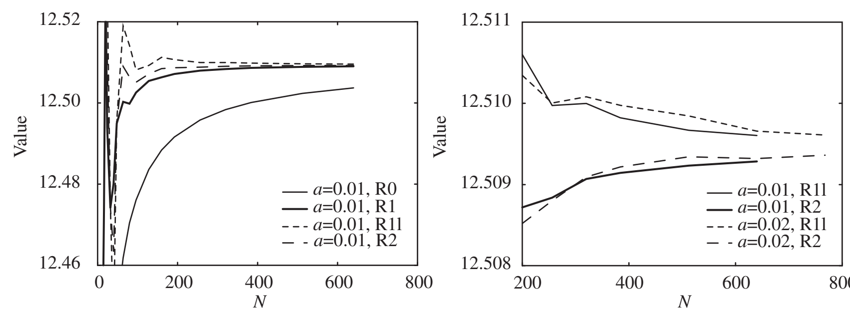

FIGURE 12. Convergence of American Asian fixed strike call option ($\sigma = 0.4$, $r = 0.1$, $q = 0$, $T = 1$, $K = 100$, $S = 100$) calculated with our accelerated binomial method. The right plot, which shows part of the left one on a finer scale, illustrates that the R1l and R2 extrapolated series converge from opposite directions.

attempts to manipulate the asset price just before expiration, also exists with delayed averaging, of course.

Assume that the asset is to be averaged over a period at the end of the option's lifetime, $t_s \leqslant t \leqslant T$. In our approach, we can implement delayed averaging simply by setting, for $t \leqslant t_s$, the average A_t for a node at time t equal to the stock S_t at that node. In other words, we choose $A_{\min}(i, j) = A_{\max}(i, j) = S_{ij}$ for $i \leqslant Nt_s/T$. Related to this, some changes should be implemented in the calculation of $A_{\min}(i, j)$ and $A_{\max}(i, j)$ for $i > Nt_s/T$.

The usual principles apply when it comes to speeding up the convergence of a delayed Asian option: convergence is fastest if $\{Nt_s/T\}$ forms a "nice" sequence of integers. For example, for $t_s/T = 0.9$, convergence will be fastest if we use $N = 10, 20, 40, 80, \ldots$ in our extrapolations. If one is interested in the value of an option where t_s/T is not a simple fraction, the quickest way to achieve high accuracy is to interpolate between two neighboring simple fractions for t_s/T.

Note that, as $t_s \to T$, the Asian option becomes a vanilla option. Correspondingly, for t_s/T close to 1, only a small fraction of nodes has a nontrivial V-table, so that the option can be valued very quickly. This will often offset the cost incurred by requiring Nt_s/T to be a sequence of integers (which might force one to use larger N than for Asians with $t_s = 0$).

A sample of results is presented in Table 8, where they are compared with the Turnbull and Wakeman (1991) approximation with a delayed continuous arithmetic average; for explicit formulas, see, e.g., Haug (1997). Note that the

TABLE 8. European Asian fixed strike call options with delayed averaging and $T = 1$, $r = 0.1$, $q = 0$, $K = S = 100$. Various volatilities and starting times for the averaging are shown. The first ABA column contains accurate results from second-order Richardson extrapolation requiring on average about 2000 s (except for the vanilla case $t_s = T$, where the calculations are almost instantaneous). The next column shows results that can be obtained with ABA in about 0.5 s. The last column gives estimates from the Turnbull and Wakeman (1991) approximation.

σ	t_s/T	ABA	ABA (0.5 s)	TW
0.05	0	4.72429825(3)	4.72429	4.72552852
	0.5	7.1072269(1)	7.10717	7.10736511
	0.9	9.0604403(1)	9.06072	9.06044363
	1	9.55663131	9.55663131	9.55663131
0.2	0	7.0410758(2)	7.04116	7.06857541
	0.5	10.286356(2)	10.2864	10.2903682
	0.9	12.682979(2)	12.6843	12.6831029
	1	13.2696766	13.2696766	13.2696766
0.5	0	13.206076(1)	13.2064	13.3933198
	0.5	18.999180(2)	19.0009	19.0300189
	0.9	22.981215(2)	22.9955	22.9822101
	1	23.9267448	23.9267448	23.9267448

Turnbull–Wakeman approximation becomes exact in two limits: $\sigma^2 T \to 0$ and $t_s \to T$. The table shows that this approximation is indeed quite accurate when $t_s/T \approx 1$ and/or $\sigma^2 T \ll 1$. It rapidly degrades if these conditions are not satisfied. Of course, it can also only handle the European case, whereas our method can easily incorporate early exercise.

4.4 *Discretely Sampled Asian Options*

Naively, one might think that discrete sampling of the average would allow faster evaluation of Asian option prices. However, this is not the case, at least not in our approach. The reason is that discrete sampling introduces *another* scale into the problem: $N/N_{\rm obs}$, where $N_{\rm obs}$ is the number of sampling (or "observation") periods.[14] Here we are assuming that the sampling points are regularly spaced. With irregular spacing, further scales would be introduced, creating further (although usually not fatal) complications.

For future reference, let us introduce the notation $\Delta t_{\rm obs} \equiv T/N_{\rm obs}$ for the sampling interval, by analogy with the time step $\Delta t = T/N$. Even with finite $N_{\rm obs}$, there are in general two reasons why we need further time slices in our tree between sampling points (i.e., $\Delta t < \Delta t_{\rm obs}$). One is to eliminate the *distributional error* from using the binomial rather than the continuum evolution of the asset between observations. The other, in the American case, is to allow for early exercise opportunities (assuming that early exericise is still allowed at all times, even when the asset is observed discretely; one could of course also allow early exercise only at discrete times, leading to Bermudan Asian options). Note that some methods can avoid the first type of error without having to take the limit $\Delta t \to 0$, e.g., the Monte Carlo or the convolution method (Carverhill and Clewlow 1990; Reiner 1999), but all methods will need a finer discretization than the sampling interval to adequately deal with American early exercise.

We stress that these various sources of discretization error all have to be dealt with in a way appropriate for the method in question. Some studies of Asian options use the same time discretization as the sampling interval, $\Delta t = \Delta t_{\rm obs}$. This can lead to significant errors, especially for American floating strike options, as we will see.

There are two strategies to take the continuum limit for discretely sampled Asians in our binomial approach:

- For small or moderate $N_{\rm obs}$, calculate the option value at $N/N_{\rm obs} = 1, 2, 4, \ldots$ with the desired number of sampling points and extrapolate to $N/N_{\rm obs} = \infty$.
- For large $N_{\rm obs}$, first extrapolate to the desired value of $N_{\rm obs}$ at fixed $N/N_{\rm obs}$, and then extrapolate to $N/N_{\rm obs} = \infty$. (Of course, for sufficiently large $N_{\rm obs}$, it might be accurate enough to use the approximation of continuous sampling,

[14] When $t = 0$ is one of the sampling points, it is natural to count the sampling points as starting at 0, which means we really have $N_{\rm obs} + 1$ sampling *points*, as compared to $N_{\rm obs}$ sampling *periods*. The potential confusion here is the same as that between the "number of binomial time steps" $N + 1$ and the "number of binomial time periods" N. We will be careful to always state our sampling scheme precisely.

where we only need a single extrapolation along $N = N_{\text{obs}}$ rather than a double extrapolation.)

We now turn to some examples. Levy and Turnbull (1992) compared a number of approximations for European fixed strike options. As a proxy for the exact answer, the authors used Monte Carlo results. Unfortunately, in the interesting case, large $\sigma^2 T$, where the accuracy of the approximations becomes questionable, the errors of Monte Carlo results were not very much smaller than the difference between (some of) the approximations and the Monte Carlo central values. We can easily give a more definite picture of the accuracy of various approximations using our approach, as seen in Table 9 for a 1-year Asian option with weekly sampling. We used the first strategy outlined above (which is always preferable if it is practical, since it avoids a double extrapolation): for fixed $N_{\text{obs}} = 52$, we calculate the option value at $N = 52, 104, \ldots, N_{\max}$ and extrapolate to $N/N_{\text{obs}} = \infty$. For Table 9, we used $N_{\max} = 416$ with second-order extrapolation in N_{obs}/N, and $N_{\max} = 104$ with first-order extrapolation.

We conclude that the Turnbull–Wakeman and Levy–Turnbull approximations are inadequate on the penny level. On the other hand, at least in the cases considered, the geometric conditioning approach of Curran (1994) is always accurate to within a penny.

We now turn to discretely sampled floating strike options. In the American case especially, they are difficult to value, since (a) the early exercise premium is large, and (b) there is a large gap between the price for $N/N_{\text{obs}} = 1$ and $N/N_{\text{obs}} = \infty$. Consider, for example, a discretely sampled American floating strike put. It might happen that initially the asset fluctuates up, and the running discrete average "gets stuck" at a large value. If the asset now moves down, it will make a big difference whether the buyer is allowed to exercise early or not. If he is, a large profit might be made by exercising at this point. This example

TABLE 9. European Asian fixed strike call options with weekly sampling, $N_{\text{obs}} = 52$, $T = 1$, $r = 0.09$, $q = 0$, $S = 100$. The asset is sampled every week for a total of 53 equally spaced samples (including the first and last day of the year, $t = 0, 1$). The accurate ABA results use second-order Richardson extrapolation and $a = 0.01$; the quicker ones use first-order extrapolation and $a = 0.02$. The last three columns contain results from various approximations: Turnbull and Wakeman (1991) (TW), Levy and Turnbull (1996) (LT), and the geometric conditioning approach of Curran (1994).

σ	K	ABA, $N = 416$	ABA, $N = 104$	TW	LT	Curran
0.3	90	14.963840(5)	14.96382	14.91	15.00	14.9627
	100	8.80152(1)	8.801454	8.78	8.84	8.8003
	110	4.671187(1)	4.671065	4.69	4.69	4.6694
0.5	90	18.14458(1)	18.14430	17.66	18.13	18.1386
	100	12.980970(3)	12.98088	12.86	13.00	12.9753
	110	9.080144(2)	9.079880	9.22	9.12	9.0737
CPU		1040 s	7.2 s			

TABLE 10. Price of European (EU) and American (US) Asian floating strike put option with discrete or continuous sampling, $T = 0.25$, $\sigma = 0.4$, $r = 0.1$, $q = 0$, $S = 100$, evaluated using the ABA approach. The asset is sampled at $t = 0$ and N_{obs} more sampling points separated by $\Delta t_{obs} = T/N_{obs}$. The third and fifth columns demonstrate that using a time step equal to the sampling interval (and no extrapolation in N) can significantly misprice the option. In the American case especially, the option would be very much undervalued. For details and results from the literature, refer to the main text.

N_{obs}	EU	EU ($N = N_{obs}$)	US	US ($N = N_{obs}$)
1	3.34695115	4.3046	6.922987	4.3046
3	3.6734(1)	3.7286	7.178(2)	4.4651
13	3.88890(1)	3.9394	6.690(2)	5.4178
45	3.94691(1)	3.9616	6.342(1)	5.8662
90	3.95934(1)	3.9667	6.258(1)	6.0024
250	3.9674(1)	3.9701	6.206(1)	6.1090
∞	3.972036(1)	3.9720	6.1821(5)	6.1821

illustrates why the American value of such an option will be much higher than its European counterpart, and also why the discretely sampled option can be significantly more valuable than the continuously sampled one if early exercise is allowed.

An example of this sort, $T = 0.25$, $\sigma = 0.4$, $r = 0.1$, has received some attention in the literature (Barraquand and Pudet 1996; Zvan, Forsyth, and Vetzal 1997/8; Dempster, Hutton, and Richards 1998; Zvan, Forsyth, and Vetzal 1998); in particular, it was pointed out (Zvan, Forsyth, and Vetzal 1998) that such options are difficult to value. Estimates of the continuum value ($\Delta t, \Delta t_{obs} \to 0$) of this option range from \$5.996 in Barraquand and Pudet (1996) to \$6.111 in Zvan, Forsyth, and Vetzal (1997/8). More recent work quotes \$6.181 for daily sampling (Dempster, Hutton, and Richards 1998) of a total of 90 samples, although we are not sure about the precise sampling scheme (i.e., whether $t = 0$ is a sampling date or not). Zvan, Forsyth, and Vetzal (1998) give a converged value ($\Delta t \to 0$) of \$6.206 for a sampling frequency of $\Delta t_{obs} = 0.001$.[15]

Our results for the European and American cases are shown in Table 10. For the American case, it typically took several thousand seconds of CPU time to obtain the estimates of the $N \to \infty$ limit, less in the European case, and *much* less for $N_{obs} = 1$. Note that there is difference of more than 2 cents between a continuously sampled American option, whose value we find to be \$6.1821(5), and one with 250 samples of the average, which corresponds to several samplings per day! These results are consistent with those of Zvan, Forsyth, and Vetzal (1998). All other results from the literature are inconsistent with ours.

[15] Zvan, Forsyth, and Vetzal (1998) refer to this as a sampling frequency of four times a day, $1/(250 \times 0.001) = 4$; in another convention, it would correspond to $1/(360 \times 0.001) \approx 2.8$ daily samples.

Note that different valuation frameworks will give different results for a given sampling frequency when $\Delta t > 0$; they should only agree in the limit $\Delta t \to 0$. Table 10 demonstrates that, especially for American options valued in a tree approach, there can be large differences between the $\Delta t = \Delta t_{\text{obs}}$ and $\Delta \to 0$ values. For example, even with daily sampling this difference is almost 26 cents (out of \$6.26). This confirms our earlier qualitative remarks. According to Forsyth, Vetzal, and Zvan (1999), the difference is much smaller in a Crank–Nicolson PDE framework. This is, of course, not surprising, since a Crank–Nicolson scheme is second-order accurate in time.

We should comment briefly on the special case $N_{\text{obs}} = 1$. In this case, the average at maturity is $A_T = \frac{1}{2}(S_0 + S_T)$. For a European floating strike put, this means that the payoff at maturity is $[A_T - S_T]_+ = \frac{1}{2}[S_0 - S_T]_+$. In other words, the Asian option reduces to half a European vanilla put option struck at $K = S_0$. For the American case, note that $A_t = S_0$ for $t < T$, leading to an immediate exercise value of $[A_t - S_t]_+ = [S_0 - S_t]_+$. This is true for times arbitrarily close to maturity; *at* maturity, the payoff is as for the European case, i.e., half as big. The option would therefore never be exercised *at* maturity. At the latest, it would be exercised just before maturity. The payoff in the zero-measure set consisting of $t = T$ can be ignored for valuation, leading us to conclude that the value of the American Asian option reduces to an American vanilla put struck at $K = S_0$. Our method reproduces the accurately known prices for these two cases.

4.5 *The Greeks*

As our method is stable (as an FD scheme) and free of spurious oscillations, we can calculate any greek as a finite difference without too much loss of precision. As discussed in Section 3.5, at least for continuously sampled Asians we can also obtain Δ and Γ directly from the binomial tree at no extra cost. An example in the latter category is shown in Figure 13.

In Table 11, we compare greeks from ABA with values from the Curran (1994) approximation for European fixed strike options. We find that this

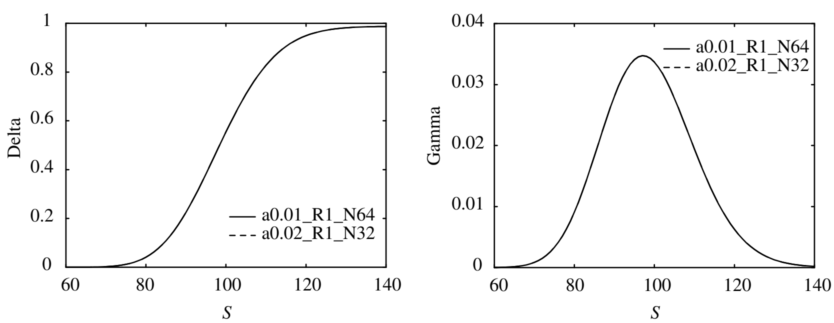

FIGURE 13. The greeks Δ and Γ for a European Asian fixed strike option ($\sigma = 0.4$, $r = 0.1$, $q = 0$, $T = 0.25$, $K = 100$) as a function of initial stock price, using first-order Richardson extrapolation. On the scale of the figures, no difference between the $N = 64, a = 0.01$ and $N = 32, a = 0.02$ results (which take about 3.5 s and 0.3 s, respectively, per S value) is detectable.

TABLE 11. Price and greeks for European fixed strike call options with continuous sampling, $\sigma = 0.5$, $r = 0.09$, $q = 0$, $S = 100$. We compare accurate values from ABA with the Curran (1994) approximation; the last column gives the relative error of the latter in percent.

T	K	ABA	Curran	Error (%)
		Price		
1	90	18.188843(1)	18.1829	−0.03
	100	13.028156(1)	13.0225	−0.04
	110	9.124312(1)	9.1179	−0.07
		Delta		
1	90	0.70652574(2)	0.706406	−0.02
	100	0.58093080(2)	0.580898	−0.01
	110	0.45809860(2)	0.458166	0.01
		Gamma		
1	90	0.01094624(1)	0.010942	−0.04
	100	0.01266518(1)	0.012654	−0.09
	110	0.01288438(1)	0.012876	−0.07
		Price		
5	90	31.42696(1)	31.3845	−0.14
	100	28.12934(1)	28.0856	−0.16
	110	25.23125(1)	25.1847	−0.18
		Delta		
5	90	0.6305474(1)	0.630172	−0.06
	100	0.5902421(1)	0.590011	−0.04
	110	0.5509437(1)	0.550880	−0.01
		Gamma		
5	90	0.00364498(1)	0.0036341	−0.30
	100	0.00399390(1)	0.0039777	−0.40
	110	0.00424059(1)	0.0042221	−0.43

approximation (in contrast to others we have investigated) is quite accurate also for Δ and Γ, even for options with large $\sigma^2 T$.

5. CONCLUSIONS

We have presented an efficient implementation of the idea of Hull and White (1993) for valuing Asian options in a modified binomial framework. The method is simple to code and use. It is unconditionally stable and does not suffer from the degeneracy and related oscillation problem that afflicts PDE methods and necessitates the use of fancy PDE solving technology such as flux limiters (though it should be remarked that these problems are most severe for continuously sampled Asians; less so for discrete sampling). When extrapolation techniques can be applied to eliminate at least the leading $1/N$ corrections, our method converges as fast as improved finite-difference schemes, but without the overhead of implicit schemes.

We have seen that our "accelerated binomial" approach works well for almost any kind of Asian option, with or without early exercise, in both convection- and diffusion-dominated regimes. It is often several orders of magnitude faster than standard PDE or Monte Carlo methods. In the European case especially, it is straightforward to obtain almost any (reasonable) desired accuracy. Greeks can also be obtained accurately (Δ and Γ come for free in the continuum case). At the very least, our method can therefore be used as a benchmarking tool for other methods, including very fast but potentially inaccurate analytic approximations. We have, in fact, used our approach to check various approximations in the European case. We find that only the method of Curran (1994), which is essentially identical (Thompson 1999) to the lower bound of Rogers and Shi (1995), is accurate enough for larger volatilities or maturities.

We have concentrated on the binomial case, but if one uses a trinomial tree then one can also value options with barrier features. The only situations where our approach might not be competitive with PDE methods are those involving complicated and irregular sampling schemes and barriers. In such cases, the multiple scales in the problem would prevent the effective use of extrapolation techniques. PDE methods can more easily be adapted to such situations by using suitably varying grid spacings in temporal and spatial directions.

As pointed out already by Hull and White (1993), the modified binomial approach can be used to value other path-dependent options. Examples include mortgage-backed securities and lockbacks; for the latter, our method can be taken over pretty much verbatim by replacing the average with the maximum (or minimum).

Acknowledgements

The author would like to thank Michael Bos, Kresimir Demeterfi, Emanuel Derman, Peter Forsyth, Michael Kamal, Joseph Zou, and especially Mark Broadie for useful discussions, suggestions, and comments. This work was supported in part by the US Department of Energy.

REFERENCES

Andreasen, J. (1998). The pricing of discretely sampled Asian and lookback options: A change of numeraire approach. *Journal Computational Finance*, **2**(1), 5–30.

Barraquand, J., and Pudet, T. (1996). Pricing of American path-dependent contingent claims. *Mathematical Finance*, **6**, 17–51.

Broadie, M., and Detemple, J. (1996). American option valuation: New bounds, approximations, and a comparison of existing methods. *Review of Financial Studies*, **9**, 1211–1250.

Broadie, M., Glasserman, P., and Kou, S. G. (1999). Connecting discrete and continuous path-dependent options. *Finance and Stochastics*, **3**, 55–82.

Carr, P., Jarrow, R., and Myneni, R. (1992). Alternative characterizations of American put options. *Mathematical Finance*, **2**, 87–106.

Carverhill, A. P., and Clewlow, L. J. (1990). Flexible convolution, *Risk*, **3**(4), April, 25–29.

Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, **7**, 229–264.

Curran, M. (1994). Valuing Asian and portfolio options by conditioning on the geometric mean price. *Management Science*, **40**, 1705–1711.

Dempster, M. A. H., Hutton, J. P., and Richards, D. G. (1998). LP valuation of exotic American options exploiting structure. *Journal Computational Finance*, **2**(1), 61–84.

Dewynne, J., and Wilmott, P. (1993). *Risk*, **6**(3), March, 38–46.

Dufresne, D. (1999). Laguerre series for Asian and other options. Working Paper, University of Melbourne.

Forsyth, P., Vetzal, K., and Zvan, R. (1999). Convergence of lattice and PDE methods for pricing Asian options. Working Paper, University of Waterloo, Ontario.

Fu, M., Madan, D. P., and Wang, T. (1998/9). Pricing continuous Asian options: A comparison of Monte Carlo and Laplace transform methods. *Journal of Computational Finance*, **2**(2), Winter, 49–74.

Geman, H., and Eydeland, A. (1995). Domino effect. *Risk*, **8**(4), April, 65–67.

Geman, H., and Yor, M. (1993). Bessel processes, Asian options and perpetuities. *Mathematical Finance*, **3**, 349–375.

Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1998). Asymptotically optimal importance sampling and stratification for pricing path-dependent options. Working Paper, Columbia University, New York.

Haug, E. G. (1997). *The Complete Guide to Option Pricing Formulas*. McGraw-Hill.

Hull, J. (1997). *Options, Futures, and Other Derivatives*, 3rd edn. Prentice-Hall.

Hull, J., and White, A. (1993). Efficient procedures for valuing European and American path-dependent options. *Journal of Derivatives*, **1**, Fall, 21–31.

Ingersoll, J. (1987). *Theory of Financial Decision Making*. Rowman and Littlefield.

Kemna, A. G. Z., and Vorst, E. C. F. (1990). A pricing method for options based on average values. *Journal of Banking and Finance*, **14**, 113–129.

Leisen, D. (1996). Pricing the American put option: A detailed convergence analysis for binomial models. Discussion Paper No. B-366, University of Bonn.

Levy, E., and Turnbull, S. M. (1992). Average intelligence. *Risk*, **5**(2), February, 53–59.

Reiner, E. (1999). Convolution methods for exotic options. Paper presented at 3rd Columbia–JAFFE Conference, Columbia University, New York.

Rogers, L., and Shi, Z. (1995). The value of an Asian option. *Journal of Applied Probability*, **32**, 1077–1088.

Shaw, W. T. (1999). *Modelling Financial Derivatives with Mathematica*. Cambridge University Press.

Thomas, J. W. (1995). *Numerical Partial Differential Equations: Finite Difference Methods*. Springer.

Thompson, G. (1999). Topics in mathematical finance. Ph.D. Thesis, University of Cambridge, UK.

Turnbull, S. M., and Wakeman, L. M. (1991). A quick algorithm for pricing European average options. *Journal of Financial and Quantitative Analysis*, **26**, 377–389.

Wilmott, P. (1998). *Derivatives*. Wiley.

Wilmott, P., Dewynne, J., and Howison, S. (1993). *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.

Zhang, J. E. (1999). Arithmetic Asian options with continuous sampling. Working Paper, City University of Hong Kong.

Zvan, R., Forsyth, P. A., and Vetzal, K. R. (1997/8). Robust numerical methods for PDE models of Asian options. *Journal of Computational Finance*, **1**(2), Winter, 39–78.

Zvan, R., Forsyth, P. A., and Vetzal, K. R. (1998). Discrete Asian barrier options. Working Paper, University of Waterloo, Ontario.

Zvan, R., Vetzal, K. R., and Forsyth, P. A. (1998). Swing low, Swing high. *Risk*, **11**(3), March, 71–75.