CONVERGENCE OF LATTICE AND PDE METHODS FOR PRICING ASIAN OPTIONS

P.A. Forsyth
Department of Computer Science
University of Waterloo
Waterloo, ON
Canada N2L 3G1
Telephone: (519) 888-4567 ext. 4415
E-mail: paforsyth@yoho.uwaterloo.ca

K.R. Vetzal
Centre for Advanced Studies in Finance
University of Waterloo
Waterloo, ON
Canada N2L 3G1
Telephone: (519) 888-4567 ext. 6518
E-mail: kvetzal@watarts.uwaterloo.ca

and

R. Zvan
Department of Computer Science
University of Waterloo
Waterloo, ON
Canada N2L 3G1
Telephone: (519) 888-4567 ext. 3363
E-mail: rzvan@yoho.uwaterloo.ca

November 30, 1998

Abstract

In a recent article, Barraquand and Pudet (1996) state that the lattice based Forward Shooting Grid (FSG) method is convergent for Asian options if either nearest lattice point or linear interpolation is used. Moreover, this result is claimed to be independent of any relationship between the grid quantization parameter (for the spacing of the nodal averages) and the timestep size. However, a more detailed analysis of the propagation of interpolation error reveals a problem. A worst case error analysis shows that the error may be large as the number of timesteps becomes large if nearest lattice point interpolation is used. We demonstrate that the worst case error is indeed attained in some numerical examples. Moreover, if a linear interpolation scheme is employed the FSG algorithm does not converge to the correct price, being off by a constant error which does not vanish in the limit as $\Delta t \to 0$, unless the limit is carried out in a certain way. Similarly, the method proposed by Hull and White (1993) is convergent provided that the node spacing in the average direction is selected appropriately as $\Delta t \to 0$. It is also a straightforward matter to show that partial differential equation (PDE) based methods are convergent. Numerical examples comparing convergence for all three techniques are presented.

1 Introduction

The valuation of path-dependent contingent claims continues to be an active area of research in finance. With the general absence of analytic solutions, the development of effective numerical algorithms has taken on added importance. Broadly speaking, these fall into three categories. Monte Carlo methods are relatively straightforward to implement, though there are some significant issues with regard to variance reduction methods as well as monitoring frequencies. The general survey paper on Monte Carlo techniques by Boyle, Broadie, and Glasserman (1997) includes some discussion of these aspects of path-dependent option valuation and provides references to this literature. A general approach based on partial differential equations is described in Wilmott, Dewynne, and Howison (1993). An illustration of how this type of approach may be used to value lookback options is provided in Forsyth, Vetzal, and Zvan (1998). Finally, given their popularity and simplicity in the context of vanilla options, it is not surprising that much effort has been devoted to adapting lattice based methods (i.e. binomial and trinomial trees) to the context of path-dependent contracts. Although there are numerous examples of this type of approach in the literature, we wish to concentrate on a subset of these. In particular, certain authors have proposed a method in which the usual tree is augmented by a second state vector which is intended to capture the path-dependent aspects of the claim. The elements of this auxiliary vector may be, for example, possible values for the maximum or minimum stock price reached thus far in the case of a lookback or candidates for the average stock price in the case of an Asian option.

For present purposes, an important feature of the auxiliary state vector is whether it contains exact values of the path-dependent feature or whether it is a representative grid spanning the range of possible values. In the case of a lookback, the highest or lowest price is necessarily one in the stock price tree. Consequently it is easy to construct the second state vector so that each element corresponds to a possible value of the maximum or minimum price reached thus far. On the other hand, the number of possible values for the arithmetic average grows exponentially with the number of timesteps. It is not feasible to track every possible average in the auxiliary vector. Instead the vector contains a grid which covers the range of possible averages, and interpolation between the nodes of this grid is required when solving backwards through the tree to find the initial value of the claim.

The first authors to propose this type of method were Ritchken, Sankarasubramanian, and Vijh (1993) and Hull and White (1993). Ritchken et al. examined European and American style Asian options, whereas Hull and White considered a variety of path-dependent claims including American and European lookbacks and Asians. A similar set of contracts was studied by Barraquand and Pudet (1996) using a slightly different algorithm which they called the forward shooting grid (FSG) method. Li, Ritchken, and Sankarasubramanian (1995) and Ritchken and Chuang (1998) have used this general kind of approach to value interest rate contingent claims. Another application is provided by Ritchken and Trevor (1998) in the context of pricing options where the underlying stock price follows various kinds of GARCH processes.

Given the wide applicability of this methodology, it is clearly important to understand its convergence properties. Somewhat surprisingly, only Barraquand and Pudet (1996) have provided much analysis in this regard. Most authors have confined themselves to illustrating convergence through numerical examples. Unfortunately, although the convergence proof provided by Barraquand and Pudet is correct for situations which do not require interpolation, there is a problem with their proof for contracts where interpolation is needed. More precisely, Barraquand and Pudet claim that

- the FSG method is convergent if nearest lattice point interpolation is used;
- unconditional convergence is obtained provided that the timestep Δt and the spacing of the nodal averages tend to zero, regardless of any quantitative relationship between these two quantization parameters (Barraquand and Pudet (1996), p. 42).

Since an interpolation error is introduced at each timestep, it is clear that the cumulative effect of a finite error applied an infinite number of times (as the timestep tends to zero) must be carefully monitored. The basic problem with Barraquand and Pudet's analysis is that they consider the interpolation error only at the expiry date of the option, ignoring the additional errors that occur at each timestep in the tree. Section 2 of this paper presents a worst case error analysis for the propagation of the interpolation error which shows that:

- if nearest lattice point interpolation is used, then the FSG method may not be convergent;
- if linear interpolation is used, then the error is not reduced in the limit as $\Delta t \to 0$, unless the limit is carried out in a certain way. In particular, the grid spacing in the auxiliary vector must be an appropriate function of Δt.

This latter point illustrates the importance of a formal convergence analysis. Numerical examples intended to demonstrate convergence are not sufficient here because it is possible to converge to a value which differs from the correct price by a constant. Now in practice, it should be pointed out that this constant appears to be quite small, at least for the examples which we have examined. This means that there do not appear to be any significant problems from using a theoretically inappropriate grid spacing in the auxiliary vector. Nonetheless, such problems might occur and our recommended approach provides a simple means of ensuring that they do not.

Using a similar analysis, section 3 demonstrates that the Hull and White (1993) method is convergent provided that the grid quantization parameter is chosen appropriately. Section 4 describes a partial differential equation (PDE) based method and shows that it is convergent as well. Section 5 presents some numerical examples. Section 6 concludes. As we are interested only in cases where interpolation is required, we focus our discussion on (arithmetic average) Asian options. We emphasize, however, that there is nothing unique about this. Our results apply to any situation where interpolation is needed.

2 Analysis of the Forward Shooting Grid Method

We will use the notation in Barraquand and Pudet (1996) to facilitate comparison with that work. Let

$$\begin{aligned} \Delta Z &= \sigma\sqrt{\Delta t} \\ \Delta Y &= \rho \Delta Z \end{aligned} \tag{1}$$

where σ is the volatility, Δt is the timestep, and ρ is a quantization parameter for spacing in the Y (average) direction. In the following, we assume that $1/\rho$ is an integer. Let discrete values of the asset price S and average price A be given by

$$\begin{aligned} S_j^n &= S_0 e^{j\Delta Z} \\ A_k^n &= S_0 e^{k\Delta Y} \\ & n = 0, \ldots, N; \quad j = -n, \ldots, +n; \quad k = -k_m(n), \ldots, +k_m(n) \end{aligned} \tag{2}$$

where N is the number of timesteps and

$$k_m(n) = n/\rho \ . \tag{3}$$

To avoid unnecessary algebraic complication without any loss of generality, take $S_0 = 1$ in equation (2). It becomes:

$$\begin{aligned} S_j^n &= e^{j\Delta Z} \\ A_k^n &= e^{k\Delta Y} \ . \end{aligned} \tag{4}$$

It follows that all error estimates will be relative to S_0, consistent with Barraquand and Pudet (1996).

Under the usual binomial approximation, we associate an upward transition $S_j^n \rightarrow S_{j+1}^{n+1}$ with probability p, and a downward transition $S_j^n \rightarrow S_{j-1}^{n+1}$ with probability $(1-p)$, during the time $t = n\Delta t$ to the time $t = (n+1)\Delta t$. The average is updated based on the transitions:

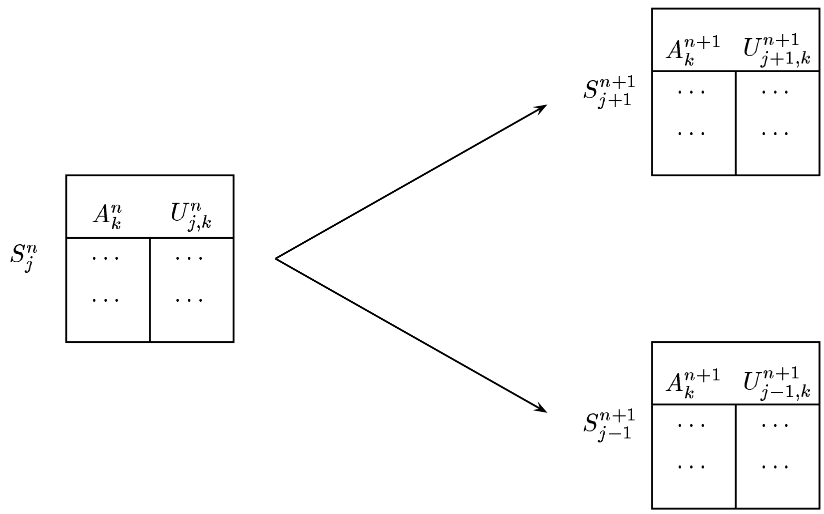

FIGURE 1: *Asset price tree indicating that a set of discrete averages and option values exists at each node of the tree.*

$$
\begin{aligned}
A_{k^+(j,k)}^{n+1} &= A_k^n + \frac{\left(S_{j+1}^{n+1} - A_k^n\right)}{(n+2)} \\
A_{k^-(j,k)}^{n+1} &= A_k^n + \frac{\left(S_{j-1}^{n+1} - A_k^n\right)}{(n+2)}
\end{aligned}
\qquad (5)
$$

with $A_0^0 = S_0 = 1$. Each asset price node in the tree has associated with it a set of average values A_k^n and option prices U_k^n. This is illustrated in Figure 1.

Note that $A_{k^+(j,k)}^{n+1}$ and $A_{k^-(j,k)}^{n+1}$ in equation (5) do not necessarily coincide with the lattice values in equation (4). This necessitates some form of interpolation (Hull and White (1993); Barraquand

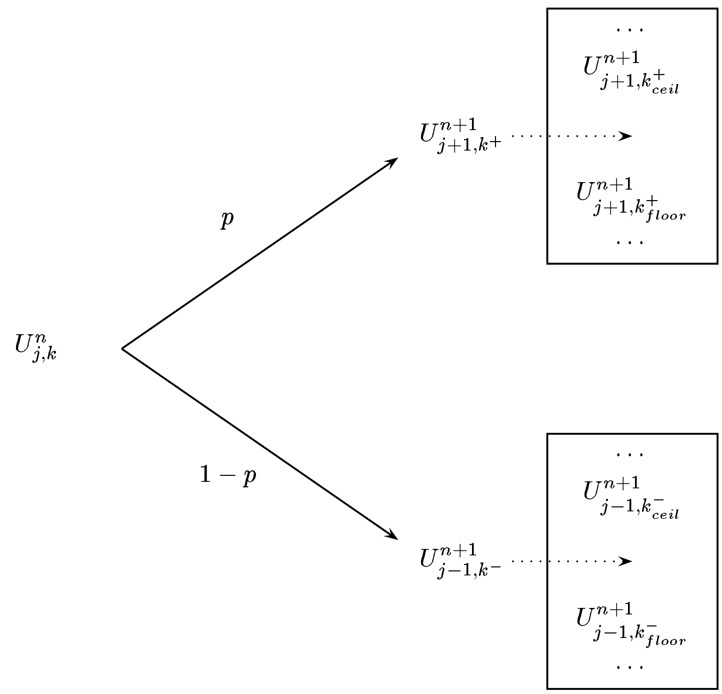

FIGURE 2: *The values of* $U^{n+1}_{j+1,k^+(j,k)}$ *and* $U^{n+1}_{j-1,k^-(j,k)}$ *must be interpolated from the known values at* $t = n + 1$.

and Pudet (1996)). For future reference, define

$$\begin{aligned} k^{\pm}_{floor}(j,k) &= floor\left[\frac{\log\left(A^{n+1}_{k^{\pm}(j,k)}\right)}{\rho \Delta Z}\right] \\ k^{\pm}_{ceil}(j,k) &= k^{\pm}_{floor}(j,k) + 1 \ . \end{aligned} \tag{6}$$

These are simply the indices for the lattice average values in equation (4) which bracket the updated average values in equation (5).

Let $U^n_{j,k} = U(S^n_j, A^n_k, n\Delta t)$ be the value of the option at $t = n\Delta t, A = A^n_k, S = S^n_j$. The value of the option given a suitable terminal payoff condition $U^N_{j,k}$ is given by the usual backward recursion, bearing in mind that the required values of the averages at $t = n + 1$ must be interpolated from the given lattice values at $t = n + 1$ (as shown in Figure 2):

$$
\begin{aligned}
U^n_{j,k} &= e^{-r\Delta t}\Big[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}U^{n+1}_{j+1,k^+_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)U^{n+1}_{j+1,k^+_{ceil}(j,k)}\right) \\
&+ (1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}U^{n+1}_{j-1,k^-_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)U^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)\Big] \\
&\quad n = N-1,\ldots,0; \quad j=-n,\ldots,+n; \quad k=-k_m(n),\ldots,+k_m(n)\ . \qquad (7)
\end{aligned}
$$

In equation (7), p is the risk neutral probability

$$
p = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{+\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}\ , \qquad (8)
$$

r is the risk-free interest rate, and the α's are determined by the type of interpolation used (nearest lattice point or linear). Note that

$$
\begin{aligned}
0 &\le \alpha^{n+1}_{k^+_{floor}(j,k)} \le 1 \\
0 &\le \alpha^{n+1}_{k^-_{floor}(j,k)} \le 1\ . \qquad (9)
\end{aligned}
$$

In the following, we will restrict attention to the propagation of the interpolation error. Let $(U^n_{j,k})^E$ denote the value of the option obtained using backward recursion, assuming that no interpolation is required, i.e.

$$
\begin{aligned}
\left(U^n_{j,k}\right)^E &= e^{-r\Delta t}\left[p\left(U^{n+1}_{j+1,k^+(j,k)}\right)^E + (1-p)\left(U^{n+1}_{j-1,k^-(j,k)}\right)^E\right] \\
&\quad n = N-1,\ldots,0; \quad j=-n,\ldots,+n; \quad k=-k_m(n),\ldots,+k_m(n)\ . \qquad (10)
\end{aligned}
$$

Here $\left(U^{n+1}_{j+1,k^+(j,k)}\right)^E$ denotes the value of U for asset price S^{n+1}_{j+1}, and average value $A^{n+1}_{k^+(j,k)}$, assuming no interpolation is required. This assumes the existence of

$$
\left(U^{n+1}_{j+1,k^+(j,k)}\right)^E = U^{n+1}_{j+1}\left(A^{n+1}_{k^+(j,k)}\right) \qquad (11)
$$

where $U^{n+1}_{j+1}(A)$ is a continuous well-defined function of A in the range $\left[A^{n+1}_{-k_m(n)}, A^{n+1}_{+k_m(n)}\right]$. Note that

$$
\begin{aligned}
\left(U^{n+1}_{j+1,k^+(j,k)}\right)^E &= \alpha^{n+1}_{k^+_{floor}(j,k)}\left(U^{n+1}_{j+1,k^+_{floor}(j,k)}\right)^E + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left(U^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)^E \\
&+ \left(\beta^q_{k^+_{floor}(j,k)}\right)^{n+1} \\
\left(U^{n+1}_{j-1,k^-(j,k)}\right)^E &= \alpha^{n+1}_{k^-_{floor}(j,k)}\left(U^{n+1}_{j-1,k^-_{floor}(j,k)}\right)^E + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left(U^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)^E
\end{aligned}
$$

$$+ \quad \left(\beta^{q}_{k^-_{floor}(j,k)}\right)^{n+1} . \tag{12}$$

For $q = 1$ (nearest lattice point interpolation) and $q = 2$ (linear interpolation) we have

$$\begin{aligned} \left(\beta^{1}_{k^+_{floor}(j,k)}\right)^{n+1} &= \min\left[\left(A^{n+1}_{k^+_{ceil}(j,k)} - A^{n+1}_{k^+(j,k)}\right), \left(A^{n+1}_{k^+(j,k)} - A^{n+1}_{k^+_{floor}(j,k)}\right)\right] \frac{\partial U^{n+1}_{j+1}(\eta)}{\partial A} \\ \left(\beta^{2}_{k^+_{floor}(j,k)}\right)^{n+1} &= -\frac{\left(A^{n+1}_{k^+(j,k)} - A^{n+1}_{k^+_{floor}(j,k)}\right)}{2} \left(A^{n+1}_{k^+_{ceil}(j,k)} - A^{n+1}_{k^+(j,k)}\right) \frac{\partial^2 U^{n+1}_{j+1}(\eta)}{\partial A^2} \end{aligned} \tag{13}$$

where $\eta \in \left[A^{n+1}_{k^+_{floor}(j,k)}, A^{n+1}_{k^+_{ceil}(j,k)}\right]$ in each case. Substituting equation (12) into equation (10) gives

$$\begin{aligned} \left(U^n_{j,k}\right)^E &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}\left(U^{n+1}_{j+1,k^+_{floor}(j,k)}\right)^E + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left(U^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)^E\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}\left(U^{n+1}_{j-1,k^-_{floor}(j,k)}\right)^E + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left(U^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)^E\right)\right] \\ &+ e^{-r\Delta t}\left[p\left(\beta^{q}_{k^+_{floor}(j,k)}\right)^{n+1} + (1-p)\left(\beta^{q}_{k^-_{floor}(j,k)}\right)^{n+1}\right] . \end{aligned} \tag{14}$$

Let the error in the value of $U^n_{j,k}$ due to interpolation be denoted by $E^n_{j,k}$ where

$$E^n_{j,k} = \left(U^n_{j,k}\right)^E - U^n_{j,k} . \tag{15}$$

Then an equation for the propagation of the error due to interpolation can be deduced by subtracting equation (7) from equation (14) to obtain

$$\begin{aligned} E^n_{j,k} &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)} E^{n+1}_{j+1,k^+_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right) E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)} E^{n+1}_{j-1,k^-_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right) E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)\right] \\ &+ e^{-r\Delta t}\left[p\left(\beta^{q}_{k^+_{floor}(j,k)}\right)^{n+1} + (1-p)\left(\beta^{q}_{k^-_{floor}(j,k)}\right)^{n+1}\right] . \end{aligned} \tag{16}$$

In order to bound the error terms due to interpolation, let

$$\begin{aligned} M_1 &= \max_{n,j}\left|\frac{\partial U^{n+1}_{j+1}(A)}{\partial A}\right| \\ M_2 &= \max_{n,j}\left|\frac{\partial^2 U^{n+1}_{j+1}(A)}{\partial A^2}\right| \end{aligned} \tag{17}$$

where $A_{-k_m(n+1)} \le A \le A_{+k_m(n+1)}$ in each case. We can then write equation (16) as

$$\begin{aligned} \left|E^n_{j,k}\right| &\le e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}\left|E^{n+1}_{j+1,k^+_{floor}(j,k)}\right| + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left|E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right|\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}\left|E^{n+1}_{j-1,k^-_{floor}(j,k)}\right| + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left|E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right|\right)\right] \\ &+ e^{-r\Delta t}\left[M_q\left(A^{n+1}_{k^+_{ceil(j,k)}}\right)^q\left(1-e^{-\rho\Delta Z}\right)^q\right] \ . \end{aligned} \tag{18}$$

We can further bound the interpolation error term in equation (18) by noting that (from equations (5-6))

$$\begin{aligned} A^{n+1}_{k^+_{ceil(j,k)}} &\le \max\left(S^{n+1}_{j+1}, A^k_n\right) \\ A^{n+1}_{k^-_{ceil(j,k)}} &\le \max\left(S^{n+1}_{j-1}, A^k_n\right) \ . \end{aligned} \tag{19}$$

Combining equation (19) with equation (18) gives

$$\begin{aligned} \left|E^n_{j,k}\right| &\le e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}\left|E^{n+1}_{j+1,k^+_{floor}(j,k)}\right| + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left|E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right|\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}\left|E^{n+1}_{j-1,k^-_{floor}(j,k)}\right| + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left|E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right|\right)\right] \\ &+ e^{-r\Delta t}\left[M_q\left(\max\left(S^{n+1}_{j+1}, A^n_k\right)\right)^q\left(1-e^{-\rho\Delta Z}\right)^q\right] \ . \end{aligned} \tag{20}$$

Let

$$\|E^n\| = \max_{j,k}\left|E^n_{j,k}\right| \tag{21}$$

and note that (with $S_0 = 1$)

$$\max\left(S^{n+1}_{j+1}, A^n_k\right) \le e^{(n+1)\sigma\sqrt{\Delta t}} \ . \tag{22}$$

Then, since the interpolation coefficients α and the probability p are all in the range $[0,1]$, from equations (20-22)

$$\|E^n\| \le e^{-r\Delta t}\left(\left\|E^{n+1}\right\| + \left[2M_q\left(e^{(n+1)\sigma\sqrt{\Delta t}}\right)^q\left(1-e^{-\rho\Delta Z}\right)^q\right]\right) \ . \tag{23}$$

Equation (23) states that the interpolation error generated during timestep $n+1 \to n$ does not become amplified, but propagates with non-increasing size throughout the remainder of the computation. However, the cumulative error grows linearly with each step, due to the fact that a new interpolation error occurs at each step. This results in an expression for the error after N

steps of the form (with $N\Delta t = T$; T being the expiry date of the contract)

$$\left\| E^0 \right\| = \left| E^0_{0,0} \right| \leq e^{-rN\Delta t} \left\| E^N \right\| + 2 \sum_{n=0}^{n=N-1} e^{-r\Delta t(n+1)} M_q \left(\Delta A^{n+1}_{max} \right)^q \tag{24}$$

with $\Delta A^{n+1}_{max} = e^{(n+1)\sigma\sqrt{\Delta t}} \left(1 - e^{-\rho \Delta Z}\right)$. This is of the same general form as that obtained by Barraquand and Pudet (1996, p. 40) except that the sum over all interpolation errors incurred at each step was omitted there — only the interpolation error at the last timestep was taken into account. Since $\left\| E^N \right\| = 0$ (the payoff is assumed exact) and since $T = N\Delta t$, taking the limit as $\Delta t \to 0$ in equation (24) gives

$$\begin{aligned} \left\| E^0 \right\| &\leq M_q e^{\sigma\sqrt{\Delta t} q - r\Delta t} \left(1 - e^{-\rho \Delta Z}\right)^q \sum_{n=0}^{n=N-1} e^{(\sigma\sqrt{\Delta t} q - r\Delta t)n} \\ &= M_q e^{\sigma\sqrt{\Delta t} q - r\Delta t} \left(1 - e^{-\rho \Delta Z}\right)^q \left(\frac{1 - e^{(\sigma\sqrt{\Delta t} q - r\Delta t)N}}{1 - e^{(\sigma\sqrt{\Delta t} q - r\Delta t)}} \right) \\ &\simeq \frac{M_q}{q} \rho^q \left(\sigma\sqrt{\Delta t}\right)^{q-1} e^{-rT} e^{(qT/\sqrt{\Delta t})\sigma} \quad \text{as} \quad \Delta t \to 0 \ . \end{aligned} \tag{25}$$

This clearly becomes unbounded as $\Delta t \to 0$.

The bound in equation (25) is too pessimistic. This is because the effect of interpolation errors for large values of A should be small at $t = 0$ since these states are very improbable. If we make the assumption that the interpolation errors are negligible for any value of $A > A_c$, then the estimate for the interpolation error bound becomes (replacing ΔA^{n+1}_{max} by $A_c \left(1 - e^{\rho \Delta Z}\right)$ in equation (24))

$$\begin{aligned} \left\| E^0 \right\| &\leq N M_q A_c^q \left(1 - e^{-\rho \Delta Z}\right)^q \\ &\simeq \frac{T M_q A_c^q \left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \end{aligned} \tag{26}$$

Note that this tends to zero only if $q \geq 3$. In particular, for $q = 1$ (nearest lattice point interpolation), equation (26) indicates that the scheme may be divergent. This will be demonstrated in some example computations below in Section 5.

More generally, a precise argument can be given to show that the effect of interpolation errors for large A becomes small at $t = 0$. In the Appendix, we show that the interpolation error is bounded by (as $\Delta t \to 0$ with $T = N\Delta t$)

$$\left| E^0_{0,0} \right| \leq N C_{max} M_q \left(1 - e^{-\rho \Delta Z}\right)^q \simeq \frac{T C_{max} M_q \left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \ . \tag{27}$$

where C_{max} is a constant independent of Δt. This bound has the same form as equation (26), which was obtained by a more intuitive argument.

Equation (27) suggests that the error can become unbounded if (for example) $q = 1$. However, this is an overestimate of the error. Note that equation (7) has all positive coefficients which sum to unity. Consequently

$$\min\left(U_{j,k}^{N}\right) = 0 \Rightarrow U_{0,0}^{0} \geq 0 \ . \tag{28}$$

In the case of a fixed strike put (with exercise price K), we have

$$\max\left(U_{j,k}^{N}\right) = K \Rightarrow U_{0,0}^{0} \leq K \ . \tag{29}$$

For a fixed strike call (with exercise price K), we note that the value of $U_{j,k}^{n}$ can be maximized at each step by choosing

$$\begin{aligned} A_{k^+(j,k)}^{n+1} = A_{k^+ceil(j,k)}^{n+1} = A_{k^+floor(j,k)}^{n+1} &= \max\left(S_{j+1}^{n+1}, A_n^k\right) \\ A_{k^-(j,k)}^{n+1} = A_{k^-ceil(j,k)}^{n+1} = A_{k^-floor(j,k)}^{n+1} &= \max\left(S_{j-1}^{n+1}, A_n^k\right) \\ & k^+(j,k) = \max(k, j+1) \\ & k^-(j,k) = \max(k, j-1) \ . \end{aligned} \tag{30}$$

Therefore $U_{j,k}^{n} \leq B_{j,k}^{n}$, where $B_{j,k}^{n}$ is given by

$$\left(B_{j,k}^{n}\right) = e^{-r\Delta t}\left[p\left(B_{j+1,k^+(j,k)}^{n+1}\right) + (1-p)\left(B_{j-1,k^-(j,k)}^{n+1}\right)\right] \tag{31}$$

for $n = N^* - 1, \ldots, 0$ where $B_{j,k}^{N} = \max\left(0, A_k^N - K\right)$ and $k^+(j,k), k^-(j,k)$ are given in equation (30). This is simply the binomial expression for a fixed strike lookback call. Consequently we have

$$0 \leq U_{0,0}^{0} \leq B(K) \tag{32}$$

where $B(K) = K$ for a fixed strike put, and is the value of a fixed strike lookback when bounding the computed price for a fixed strike Asian call. The above arguments can be repeated for the case of floating strike Asian options, with the upper bounds given in terms of the corresponding floating strike lookbacks. Thus, equation (60) is more precisely stated as

$$\left|E_{0,0}^{0}\right| \leq \min\left[B(K), N C_{\max} M_q \left(1 - e^{\rho \Delta Z}\right)^q\right] \tag{33}$$

where $B(K)$ is independent of Δt. This means that the error never becomes unbounded, but is of size $B(K)$ in the worst case (which may be very large, of course).

In equation (27), it is easy to see that convergence can be obtained if the grid quantization parameter ρ tends to zero as $\Delta t \to 0$ as a power of Δt. In particular, if we desire an overall convergence rate of at least Δt, then we must have

$$\rho = O\left[(\Delta t)^{(2-q/2)/q}\right] \tag{34}$$

For the case of nearest lattice node interpolation ($q = 1$) $\rho = O\left[(\Delta t)^{3/2}\right]$, which implies that at timestep n (using the notation of equation (2)),

$$\begin{aligned} e^{k_m \sigma \rho \sqrt{\Delta t}} &= e^{k_m \sigma (\Delta t)^2} \\ &= e^{n\sigma\sqrt{\Delta t}} \\ \Rightarrow \; k_m &= \frac{n}{(\Delta t)^{3/2}} \end{aligned} \tag{35}$$

which results in the total number of nodes at step n being $O\left[n^2 (\Delta t)^{-3/2}\right]$. The total computational complexity after N steps is then $O\left[N^3 (\Delta t)^{-3/2}\right] = O(N^{9/2})$. For linear interpolation, a similar calculation gives the total number of nodes at step n as $n^2(\Delta t)^{-1/2}$ with total complexity for N steps of $O(N^{7/2})$.

We emphasize here that the above complexities assume that ρ satisfies equation (34), but ρ is assumed to be a constant independent of Δt in Barraquand and Pudet (1996). For constant ρ, the complexity of the FSG method is $O(N^3)$, but convergence is problematic.

3 Analysis of the Hull and White Method

The method developed in Hull and White (1993) is actually a more efficient implementation of the method described in Barraquand and Pudet (1996). The node spacing in the A direction in Hull and White (1993) is

$$A_k^n = S_0 e^{kh} \tag{36}$$

where, for given h, the range in k values in equation (36) is selected to span the possible averages at timestep n. Recall that in equation (2) the range of A values at each timestep n is the same as the range of S values, which is clearly an overestimate. Consequently, the Hull and White method has a more efficient average node placement compared to the FSG method.

Using an argument similar to that used to derive equation (27), we obtain the estimate

$$\left|E_{0,0}^0\right| \le \frac{T C_{max} M_q \left(1 - e^{-h}\right)^q}{\Delta t} \; . \tag{37}$$

Hull and White (1993) suggest either linear or quadratic interpolation. If we take $h = C\Delta t$, for example, then

$$\left|E^0_{0,0}\right| \leq \frac{TC_{max}M_q\left(1-e^{-h}\right)^q}{\Delta t} \simeq TC_{max}M_qC^q(\Delta t)^{q-1} \ . \tag{38}$$

This implies that the Hull and White method is convergent as long as linear interpolation ($q = 2$) is used. The convergence arguments for lattice type methods used in this paper rely on the interpolation coefficiencts being in the range $[0,1]$. As such, they do not apply for the case of quadratic interpolation and so we do not consider such methods here. The expression in equation (38) considers only the effect of the interpolation error. There will also be the usual lattice error of size $O(\Delta t)$, so that the global convergence rate of the Hull and White method should be of $O(\Delta t)$.

Following Chalasani, Jha, Egriboyun, and Varikooty (1998), we can estimate the number of nodal averages at timestep n for large n. The maximum possible average value for a lattice after n steps is

$$\begin{aligned} A^n_{\max} &= \frac{\sum_{k=0}^{k=n} e^{k\sigma\sqrt{\Delta t}}}{n+1} \\ &= \frac{1-e^{\sigma\sqrt{\Delta t}(n+1)}}{(n+1)\left(1-e^{\sigma\sqrt{\Delta t}}\right)} \\ &\simeq O\left[\frac{e^{\sigma(n+1)\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)}\right] \quad \text{as} \quad n \to \infty \ . \end{aligned} \tag{39}$$

The minimum possible value of the average after n steps is

$$\begin{aligned} A^n_{\min} &= \frac{\sum_{k=0}^{k=n} e^{-k\sigma\sqrt{\Delta t}}}{n+1} \\ &= \frac{1-e^{-\sigma\sqrt{\Delta t}(n+1)}}{(n+1)\left(1-e^{-\sigma\sqrt{\Delta t}}\right)} \\ &\simeq O\left[\frac{e^{\sigma\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)}\right] \quad \text{as} \quad n \to \infty \ . \end{aligned} \tag{40}$$

Letting

$$\begin{aligned} e^{m_1 C\Delta t} &= \frac{e^{\sigma(n+1)\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)} \\ e^{m_2 C\Delta t} &= \frac{e^{\sigma\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)} \ , \end{aligned} \tag{41}$$

then the total number of average nodes $(m_1 - m_2)$ is $O(n/\sqrt{\Delta t})$. This gives the total number of nodes at each step as $O(n^2/\sqrt{\Delta t})$, with resulting complexity $O(N^{7/2})$.

4 Analysis of PDE Methods

The discrete Asian option pricing problem can be solved using a system of one-dimensional PDE's as described in Wilmott et al. (1993) and Zvan, Forsyth, and Vetzal (1998). We consider the option value U to be a function of two state variables, the asset price S and the average A. Away from each discrete observation date t_n, $U = U(S, A, t)$ satisfies the usual Black-Scholes equation

$$U_t + \frac{\sigma^2 S^2}{2} U_{SS} + rSU_S - rU = 0 \tag{42}$$

At observation dates no-arbitrage considerations imply that

$$U(S, A^{n+1}, t_n^+) = U(S, A^n, t_n^-) \tag{43}$$

where t_n^+ (t_n^-) is the time immediately after (before) the observation date t_n, and

$$A^{n+1} = A^n + \frac{(S - A^n)}{n+2} \tag{44}$$

with $A^0 = S^0$. Convergence of the above method is easily demonstrated. Away from the observation dates, we simply solve a set of one-dimensional problems (equation (42)) for each discrete value of the average, using standard numerical methods. For example, suppose that second order spatial discretization is used with Crank-Nicolson time weighting. Since this is a stable, consistent method, the solution converges at a rate $O\left[(\Delta S)^2, (\Delta t)^2\right]$. Note that this rate of convergence can be obtained even for *rough* initial data (Rannacher (1984)), which is characteristic of payoff functions. The only unusual feature in this problem is that at each observation date, a new initial condition is generated using the condition (44). Since generally A^{n+1} will not coincide with a grid node, interpolation (linear or quadratic) is used to estimate the value of the solution $U(S, A^{n+1}, t_n^+)$. The interpolation at each observation date is illustrated in Figure 3.

Since a stable method is being used, the interpolation errors do not become amplified by the difference scheme. In the worst case, the errors simply persist (i.e. do not get damped out). Consequently, if N interpolation errors are introduced at N observation times, then the worst case effect of these errors is simply N times the maximum interpolation error.

Assuming that the same grid spacing is used in the S and A direction, and letting $\Delta S_{\max}$ be

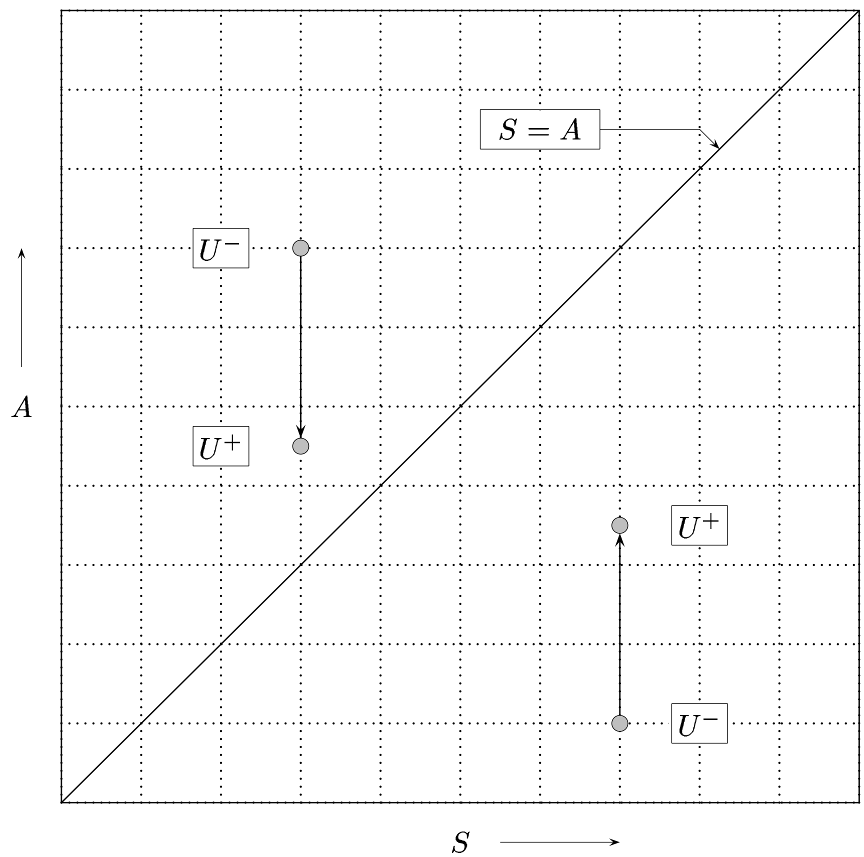

FIGURE 3: *Between each observation date, one-dimensional PDE's for each value of the average A are solved. The values of the option price U^- before each observation date are interpolated from the values just after the observation date U^+.*

the maximum grid spacing in the S or A direction, then the interpolation error at each step is

$$\text{interpolation error at each observation} = O\left[(\Delta S_{\max})^q\right] \tag{45}$$

where $q = 2$ for linear interpolation and $q = 3$ for quadratic interpolation. After $N = O(1/\Delta t)$ steps, we have

$$\text{global interpolation error} = O\left[\frac{(\Delta S_{\max})^q}{\Delta t}\right] . \tag{46}$$

Assuming second order space and time truncation errors, then the total error will be

$$\text{global discretization error} = O\left[\frac{(\Delta S_{\max})^q}{\Delta t}\right] + O\left[(\Delta S_{\max})^2\right] + O\left[(\Delta t)^2\right] . \tag{47}$$

If we use quadratic interpolation as in Zvan et al. (1998), and take the limit in such a way that

$\Delta S_{max} = C\Delta t$ where C is a constant, then we obtain

$$\text{global discretization error} = O\left[(\Delta t)^2\right] . \tag{48}$$

Further details concerning the PDE method can be found in Zvan et al. (1998).

Consider a limiting process whereby for a timestep of $\Delta t = T/N$ we take $O(N)$ nodes in both the average and asset grids. Since the cost of solving N implicit one-dimensional PDE's each consisting of N nodes is $O(N^2)$, we have a total complexity after N steps of $O(N^3)$. This complexity is smaller than that of the Hull and White (1993) method and is the same as that of the FSG method with constant ρ (equation (1)). The rate of convergence for the PDE method is $O\left[(\Delta t)^2\right]$, compared to at best $O(\Delta t)$ for the lattice methods. Therefore it would appear that the PDE method will be superior for sufficiently small convergence tolerances. However, equation (48) only takes into account the truncation error of the discretization of the PDE and the interpolation error. There is an additional error due the fact that we are attempting to converge to a continuously observed Asian option using a discretely observed approximation. This will introduce an $O(\Delta t)$ error which will eventually dominate the other errors. Note that the lattice methods suffer from this error as well, but these methods are only $O(\Delta t)$ to start with. Of course, in situations where we are attempting to price discretely monitored Asian options with a specified finite observation interval, then the faster convergence of the PDE approach will be more noticeable.

5 Numerical Examples

This section provides some numerical computations to support our analysis. We considered the example of a European fixed strike call option, and computed prices using the FSG, Hull and White, and PDE methods. The algorithms were coded in C++. Computations were performed on a Sun Ultrasparc workstation.

We begin by describing some further details about the algorithms. The FSG method was implemented as described in Barraquand and Pudet (1996). Both nearest lattice point and linear interpolation methods were examined. Barraquand and Pudet (p. 47) recommend values of $\rho = 0.5$ for linear interpolation and $\rho = 0.1$ for nearest lattice point interpolation. We computed values using both of these values of ρ for each interpolation scheme. In addition, we also used the value $\rho = 1.0$. In this case the number of nodes for the average was the same as that for the stock price. This particular scheme was not expected to perform very well, but it provided an interesting point of comparison.

The Hull and White (1993) method was implemented as described in that article, with the

TABLE 1: *Data for the test cases. European fixed strike call options. Limit as frequency of observation becomes continuous.*

	Case 1	Case 2
σ	0.40	0.50
r	0.10	0.10
Time to expiry	0.25 years	5.0 years
Exercise price	\$100	\$100
Asset price	\$100	\$100

additional feature that the average node spacing parameter h in equation (36) was specified as:

$$h = \alpha\sqrt{\frac{0.25}{T}\sigma^2 \Delta t}. \tag{49}$$

This choice of scaling factor for h was selected so as to give roughly the same number of average nodes at $t = T$ for the different maturities and volatilities we considered. The parameter α in equation (49) controls the fineness of the grid in the average direction. Three values of α were used: 10, 5, and 1. Linear interpolation was used.

The PDE method employed an irregularly spaced finite difference method with Crank-Nicolson timestepping. The finite difference method in one dimension is algebraically identical to a finite element discretization with linear basis functions and mass lumping. Constant timesteps were used to facilitate comparison with the lattice methods. The same grid spacing was used in both the A and S directions. On the initial coarse 50×50 grid, the spacing near the exercise price was selected to be similar to the spacing used in the lattice methods. Finer grids were constructed by halving the spacing of the coarse grids. The timestep size was also halved with each grid refinement. Quadratic interpolation was used.

We considered two cases of fixed strike Asian call options. The data for these cases is listed in Table 1. The results for the FSG method are given in Table 2. Consider Case 1 first. Clearly, the computations for nearest lattice point interpolation are in agreement with the convergence analysis presented above. As will be shown below, the correct price for this option is $\simeq$ \$5.1662. When $\rho = 1.0$ or $\rho = 0.5$, the computed values are nowhere near the true price. When $\rho = 0.1$, the results for a small number of timesteps (50-100) are reasonably close to the correct price. However, as Δt is decreased the solution begins to diverge.

When linear interpolation is used, our convergence analysis indicates that the FSG method will converge to the correct solution plus a constant error as $\Delta t \to 0$. Extrapolation of the prices in the table for $\rho = 0.1$ with linear interpolation to $\Delta t = 0$ gives a value of \$5.1688, a little higher than the true price.

TABLE 2: *Convergence of the Forward Shooting Grid method. CPU times are given for a Sun Ultrasparc.*

Case 1					
		Nearest Lattice Point Interpolation		Linear Interpolation	
ρ	Timesteps	Option Value	CPU (sec)	Option Value	CPU (sec)
1.0	50	2.0942	0.4	5.2672	0.6
	100	1.4311	2.9	5.2447	5.2
	200	0.9881	23.5	5.2254	41.5
	400	0.6873	188	5.2101	332
0.5	50	3.6866	0.7	5.2095	1.3
	100	2.6806	5.9	5.2019	10.3
	200	1.8415	46.9	5.1941	82.8
	400	1.2412	375	5.1874	663
0.1	50	5.1629	3.7	5.1613	6.5
	100	5.1688	29.5	5.1657	51.9
	200	5.1364	236	5.1678	415
	400	4.8737	1881	5.1685	3316
Case 2					
1.0	50	16.2053	0.4	28.7217	0.6
	100	10.7957	3.0	28.6631	4.6
	200	6.9113	24.1	28.6052	36.6
	400	4.0803	191	28.5556	291
0.5	50	25.0508	0.8	28.5168	1.1
	100	19.9843	6.0	28.5107	9.1
	200	14.7166	48.1	28.4934	73.0
	400	9.9951	382	28.4745	562
0.1	50	28.2968	3.8	28.3440	5.7
	100	28.6676	30.1	28.3816	45.7
	200	29.3198	241	28.3996	366
	400	29.1303	1906	28.4071	2908

Turning to Case 2, we begin by noting that the correct price here is $\simeq$ \$28.4052. Again, very poor results are obtained using nearest lattice point interpolation. The solution with linear interpolation is close to the true price with $\rho = 0.1$ and 400 timesteps. Extrapolation to $\Delta t = 0$ of the prices in the table for linear interpolation with $\rho = 0.1$ gives a value of \$28.4131. For both cases, the FSG method with linear interpolation converges to a number which although *close* to the correct price is *not* that price.

The Hull and White algorithm results for both cases are presented in Table 3. This method is well-behaved for all values of α and numbers of timesteps. This is consistent with our analysis because the grid spacing in the average direction is selected as in equation (49), providing a convergent method. The complexity estimate of $O(N^{7/2})$ is clearly confirmed in the table, both in terms of CPU time and the number of grid nodes at $t = T$. The rate of convergence implied by the numbers in the table is $O(\Delta t)$. Extrapolation to $\Delta t = 0$ of the values when $\alpha = 1$ gives price estimates of \$5.1663 for Case 1 and \$28.4051 for Case 2.

Table 4 contains the results for the PDE method for both cases. As expected, this method is also convergent and shows an $O(N^3)$ complexity. The rate of convergence is $O(\Delta t)$. As noted above, this is slower than the $O\left[(\Delta t)^2\right]$ convergence rate that one might expect due to the fact that we are taking the continuous limit of a discrete observation model. Extrapolating the results to $\Delta t = 0$ gives prices of \$5.1662 for Case 1 and \$29.4052 for Case 2, in excellent agreement with the Hull and White extrapolated prices of \$5.1663 and \$28.4051. As both of these methods are convergent, this leads to the conclusion that the true prices are $\simeq$ \$5.1662 and \$28.4052. By contrast, recall that the FSG extrapolated prices were \$5.1688 and \$28.4131. This is clearly consistent with our analysis indicating that the FSG method converges to a price with a constant error if linear interpolation is used. Of course, our analysis suggests that the FSG method could be modified so that it is convergent. This could be done, for example, by making ρ depend on $\sqrt{\Delta t}$. However, this would result in what amounts to an inefficient implementation of the Hull and White method, owing to an unnecessarily large number of nodes in the average direction.

Although our main emphasis is on convergence, it might be worth concluding this section by making some observations on the relative merits of the PDE and Hull and White methods. For this particular case, where we are attempting to converge to the continuous observation limit, the two approaches are quite comparable. It might be possible to employ quadratic interpolation to improve the efficiency of the Hull and White method. This has been suggested by both Hull and White and Ritchken and Chuang (1998). The tradeoff here would be between fewer nodes in the average direction (observe that our PDE approach using quadratic interpolation requires far fewer grid points than the Hull and White method to achieve comparable accuracy) versus more floating point operations being required for the interpolation. However, we stress that the convergence of such an approach has not been formally demonstrated.

TABLE 3: *Convergence of the Hull and White method for the test cases. CPU times are given for a Sun Ultrasparc. The grid size is the number of nodes in the A direction at $t = T$.*

Case 1				
α	Timesteps	Grid Size	Option Value	CPU (sec)
10	50	204	5.1824	0.8
	100	528	5.1745	8.6
	200	1535	5.1706	97.8
	400	4209	5.1686	1084
5	50	371	5.1651	1.5
	100	1051	5.1658	16.9
	200	2941	5.1661	189
	400	8192	5.1663	2120
1	50	1794	5.1580	7.2
	100	5053	5.1622	81.4
	200	14243	5.1644	917
	400	40205	5.1654	10296
Case 2				
10	50	163	28.5098	0.7
	100	450	28.4583	7.5
	200	1225	28.4319	82.7
	400	3394	28.4186	913
5	50	308	28.4310	1.3
	100	844	28.4180	14.4
	200	2356	28.4115	160
	400	6579	28.4083	1782
1	50	1440	28.3899	6.0
	100	4051	28.3972	67.2
	200	11415	28.4011	757
	400	32196	28.4031	8548

TABLE 4: *Convergence of the PDE method for the test cases. CPU times are given for a Sun Ultrasparc. A Cartesian product grid is used. The grid size is given as number of nodes in the S and A directions.*

	Case 1		
Grid Size	Timesteps	Option Value	CPU (sec)
50×50	50	5.1478	1.6
100×100	100	5.1559	13.8
200×200	200	5.1610	112
400×400	400	5.1636	908
	Case 2		
50×50	50	28.3573	2.2
100×100	100	28.3842	14.5
200×200	200	28.3952	112
400×400	400	28.4003	911

In practice, a typical contract would feature discrete monitoring. In such cases the PDE method can be expected to be superior. Both the Hull and White and FSG methods at best would converge at a rate of $O(\Delta t)$, and at best have $O(N^3)$ complexity. The PDE method also has complexity of order N^3, but its convergence rate is $O\left[(\Delta t)^2\right]$. This means that in order to obtain a given error, lattice based methods require work of order N^3 whereas the PDE method requires work of order $N^{3/2}$ due to the implicit discretization.

Even so, in our view the real strength of the PDE method lies in its flexibility in terms of handling more complex path-dependent features such as barrier provisions. Various types of Asian options with assorted barriers have been examined using the PDE method by Zvan et al. (1998). Examples include Parisian style cases where the barrier provisions depend on the length of time for which the underlying asset lies outside a pre-specified range, as well as situations where the barrier is in terms of the average rather than the price. It is also easy to adapt the PDE method to alternative stochastic processes for the underlying asset such as a CEV model. By comparison, we suspect that the incorporation of such characteristics into a lattice based approach would be relatively difficult.

6 Conclusion

The convergence analysis presented in this paper suggests that in the worst case, the forward shooting grid method proposed by Barraquand and Pudet (1996) with nearest lattice point interpolation will exhibit large errors as the number of timesteps becomes large. This analysis is confirmed by some numerical experiments. If linear interpolation is used, then the FSG method should converge to the correct solution plus a constant error term which is not reduced by decreasing the timestep. The constant appears to be fairly small if a large number of nodes is used in the average direction, but this method should be used with caution.

As long as the average node spacing parameter is selected appropriately, then the Hull and White (1993) method is convergent. If linear interpolation is used, then the complexity of this algorithm is $O(N^{7/2})$, where N is the number of timesteps. If quadratic interpolation is used, then it may be possible to reduce this complexity, although not to $O(N^3)$. Moreover, it has not been formally shown that such an approach would be convergent.

The PDE method is also convergent in the continuous limit for Asian options. The PDE method has complexity $O(N^3)$, where N is the number of timesteps. In the case of discretely observed barriers, this method converges as $O\left[(\Delta t)^2\right]$, where Δt is the timestep size (Zvan et al. (1998)). When using this method to converge to the continuously observed limit, the rate of convergence is reduced to $O(\Delta t)$.

Generally, when dealing with straightforward Asian options, either the Hull and White or the PDE methods are effective. The FSG method is somewhat problematic. The PDE method shows promise as being a flexible, general technique which can be used to price a wide variety of more complex path-dependent options. Between observation dates, the PDE algorithm consists of a set of independent one-dimensional PDE's. These one dimensional problems only exchange information at observation dates. This would seem to be ideally suited to a parallel implementation, if speed of computation is of paramount concern.

Appendix

In this Appendix, we make a more precise argument that the effect of interpolation errors for large A becomes small at $t = 0$. This means that a bound of the form (26) is actually a correct estimate, compared to the overly pessimistic (25).

Since equation (20) is linear, we can consider $\left|E_{0,0}^0\right|$ to be

$$\left|E_{0,0}^0\right| \leq \sum_{N^*=0}^{N^*=N-1} \left|E_{0,0}^0(N^*)\right| \tag{50}$$

where $E_{j,k}^n(N^*)$ is the error propagated to node (j,k) at timestep n due to an interpolation error occurring during the transition from $N^*+1 \rightarrow N^*$, *assuming no other interpolation errors occur* during transitions from $N^* \rightarrow N^*-1, N^*-1 \rightarrow N^*-2, \ldots, 1 \rightarrow 0$. Clearly $E_{j,k}^n(N^*) = 0$ for $n > N^*$.

Consequently, we have (from equation (20))

$$\begin{aligned}\left|E_{j,k}^{N^*}(N^*)\right| &\leq e^{-r\Delta t}\left[M_q\left(\max\left(S_{j+1}^{N^*+1}, A_k^{N^*}\right)\right)^q\left(1-e^{-\rho\Delta Z}\right)^q\right] \\ &\leq M_q\left(1-e^{-\rho\Delta Z}\right)^q\left[\left(S_{j+1}^{N^*+1}\right)^q + \left(A_k^{N^*})\right)^q\right] .\end{aligned} \tag{51}$$

We can rewrite equation (51) as

$$\left|E_{j,k}^{N^*}(N^*)\right| \leq M_q\left(1-e^{-\rho\Delta Z}\right)^q\left[\left(E_{j,k}^{N^*}(N^*)\right)_A + \left(E_{j,k}^{N^*}(N^*)\right)_S\right] \tag{52}$$

with

$$\begin{aligned}\left(E_{j,k}^{N^*}(N^*)\right)_A &= \left(A_k^{N^*})\right)^q \\ \left(E_{j,k}^{N^*}(N^*)\right)_S &= \left(S_{j+1}^{N^*+1}\right)^q .\end{aligned} \tag{53}$$

Since the error equation (20) is linear, we can consider the error $E_{j,k}^n(N^*)$ to be bounded by the sum of errors due to $\left(E_{j,k}^{N^*}(N^*)\right)_S$ and $\left(E_{j,k}^{N^*}(N^*)\right)_A$

$$E_{j,k}^n(N^*) \leq M_q\left(1-e^{-\rho\Delta Z}\right)^q\left[\left(E_{j,k}^n(N^*)\right)_S + \left(E_{j,k}^n(N^*)\right)_A\right] \tag{54}$$

where $\left(E_{j,k}^n(N^*)\right)_S$ and $\left(E_{j,k}^n(N^*)\right)_A$ satisfy the error equation (20) with terminal conditions (53) and no interpolation error term.

Since $\left(E_{j,k}^n(N^*)\right)_S$ is independent of k, it follows from equation (20) (with zero interpolation error) that

$$\left(E^n_{j,k}(N^*)\right)_S = e^{-r\Delta t}\left[p\left(E^{n+1}_{j+1,k}\right)_S + (1-p)\left(E^{n+1}_{j-1,k}\right)_S\right] \tag{55}$$

for $n = N^* - 1, \ldots, 0$. This is precisely the binomial tree expression for the European call option with payoff (at $T^* = N^*\Delta t$)

$$\left(Se^{\sigma\sqrt{\Delta t}}\right)^q \simeq S^q \quad \text{as} \quad \Delta t \to 0 \ . \tag{56}$$

Let the value of this option $\left(E^0_{0,0}(N^*)\right)_S$ be bounded by $C(N^*)_S$. $\left(E^0_{0,0}(N^*)\right)_A$ can be bounded by noting that the payoff is A^q, so that $\left(E^n_{j,k}(N^*)\right)_A$ ($n < N^*$) is maximized at each timestep by selecting

$$\begin{aligned}
A^{n+1}_{k^+(j,k)} = A^{n+1}_{k^+ceil(j,k)} = A^{n+1}_{k^+floor(j,k)} &= \max(S^{n+1}_{j+1}, A^k_n) \\
A^{n+1}_{k^-(j,k)} = A^{n+1}_{k^-ceil(j,k)} = A^{n+1}_{k^-floor(j,k)} &= \max(S^{n+1}_{j-1}, A^k_n) \\
& k^+(j,k) = \max(k, j+1) \\
& k^-(j,k) = \max(k, j-1) \ .
\end{aligned} \tag{57}$$

This is simply an algebraic statement of the fact that the price of a fixed strike lookback call is always greater than the price of a fixed strike Asian call (with the same strike). With definition (57) in equation (20) we obtain

$$\left(E^n_{j,k}(N^*)\right)_A \le e^{-r\Delta t}\left[p\left(E^{n+1}_{j+1,k^+(j,k)}\right)_A + (1-p)\left(E^{n+1}_{j-1,k^-(j,k)}\right)_A\right] \tag{58}$$

for $n = N^* - 1, \ldots, 0$. The right hand side of inequality (58) is precisely the binomial tree expression for a lookback call with payoff A^q at $T^* = N^*\Delta t$, where A is maximum value attained by the asset (as defined in equation (57)). Note that there are no interpolation errors generated at any timestep of the lookback (58). Let $\left(E^0_{0,0}(N^*)\right)_A$ be bounded by some constant $C(N^*)_A$. Let

$$\begin{aligned}
\max_{N^*}\left(|C(N^*)_A| + |C(N^*)_S|\right) &\le C_{max} \\
N^* &\le N \ ; \ N \to \infty \\
N &= T/\Delta t \ .
\end{aligned} \tag{59}$$

Then, from equations (50) and (52) we have

$$\left|E^0_{0,0}\right| \le N C_{max} M_q \left(1 - e^{-\rho\Delta Z}\right)^q \ . \tag{60}$$

Taking the limit in equation (60) as $\Delta t \to 0$ with $T = N\Delta t$ we obtain

$$\left|E^0_{0,0}\right| \leq N C_{max} M_q \left(1 - e^{-\rho \Delta Z}\right)^q \simeq \frac{T C_{max} M_q \left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \ . \tag{61}$$

This estimate has the same form as that in equation (26), which was obtained by a more intuitive argument.

References

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

Boyle, P., M. Broadie, and P. Glasserman (1997). Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control* 21, 1267–1321.

Chalasani, P., S. Jha, F. Egriboyun, and A. Varikooty (1998). A refined binomial lattice for pricing American Asian options. 8th Annual Derivative Securities Conference, Boston.

Forsyth, P. A., K. R. Vetzal, and R. Zvan (1998). A finite element approach to the pricing of discrete lookbacks with stochastic volatility. *Applied Mathematical Finance*, forthcoming.

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

Li, A., P. Ritchken, and L. Sankarasubramanian (1995). Lattice models for pricing American interest rate claims. *Journal of Finance* 50, 719–737.

Rannacher, R. (1984). Finite element solution of diffusion problems with irregular data. *Numerische Mathematik* 43, 309–327.

Ritchken, P. and I. Chuang (1998). Interest rate option pricing with volatility humps. Working paper, Case Western Reserve University.

Ritchken, P., L. Sankarasubramanian, and A. M. Vijh (1993). The valuation of path dependent contracts on the average. *Management Science* 39, 1202–1213.

Ritchken, P. and R. Trevor (1998). Pricing options under generalized GARCH and stochastic volatility processes. *Journal of Finance*, forthcoming.

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

Zvan, R., P. A. Forsyth, and K. R. Vetzal (1998). Discrete Asian barrier options. Working paper, University of Waterloo, `http://yoho.uwaterloo.ca/~paforsyt/asian.ps`.