CONVERGENCE OF NUMERICAL METHODS FOR VALUING PATH-DEPENDENT OPTIONS USING INTERPOLATION

P.A. Forsyth
Department of Computer Science
University of Waterloo
Waterloo, ON
Canada N2L 3G1
E-mail: paforsyt@elora.math.uwaterloo.ca

K.R. Vetzal
Centre for Advanced Studies in Finance
University of Waterloo
Waterloo, ON
Canada N2L 3G1
E-mail: kvetzal@watarts.uwaterloo.ca

and

R. Zvan
Financial Analytics and Structured Transactions
Bear Stearns
383 Madison Avenue
New York, NY
U.S.A. 10179
E-mail: rzvan@bear.com

First version: November 30, 1998
This revision: May 21, 2002

JEL Classification: G12, G13

Acknowledgement: This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Communications and Information Technology Office of Ontario. A previous version of this paper was presented at the 9th Annual Derivative Securities Symposium at Boston University in April 1999. We thank conference participants for their comments and suggestions.

Abstract

One method for valuing path-dependent options is the augmented state space approach described in Hull and White (1993) and Barraquand and Pudet (1996), among others. In certain cases, interpolation is required because the number of possible values of the additional state variable grows exponentially. We provide a detailed analysis of the convergence of these algorithms. We show that it is possible for the algorithm to be non-convergent, or to converge to an incorrect answer, if the interpolation scheme is selected inappropriately. We concentrate on Asian options, due to their popularity and because of some errors in the previous literature.

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 et al. (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 et al. (1993). An illustration of how this type of approach may be used to value lookback options is provided in Forsyth et al. (1999). 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 et al. (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 et al. (1995) and Ritchken and Chuang (1999) have used this general kind of approach to value interest rate contingent claims. Another application is provided by Ritchken and Trevor (1999) 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, but this does not prove convergence to the correct answer. 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 during the last timestep of the tree, ignoring the additional errors that occur at each preceding timestep.

After outlining the Asian option pricing problem in Section 2, and describing the forward shooting grid algorithm in Section 3, Section 4 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 5 demonstrates that the Hull and White (1993) method is convergent provided that the grid quantization parameter is chosen appropriately. Section 6 describes a partial differential equation (PDE) based method and shows that it is convergent as well. Section 7 presents some numerical examples. Section 8 provides a brief illustration of the use of interpolation for mortgage-backed securities. In particular, we use this example to illustrate how the frequency with which the interpolation must be applied affects the rate of convergence. Section 9 concludes.

2 Formulation: Asian Options

A standard approach for valuing Asian and other path-dependent options is to augment the state space. Let the discretely observed average be given by observing the asset price at discrete times $t_0, t_1, \ldots, t_n$, with corresponding asset prices $S^0, S^1, \ldots, S^n$. The average is then defined as

$$A^n = \frac{\sum_{i=0}^{n} S^i}{n+1} \tag{2.1}$$

with $A^0 = S^0$. A recursive expression for the average at observation time t_{n+1} is given by

$$A^{n+1} = A^n + \frac{(S^{n+1} - A^n)}{n+2} \ . \tag{2.2}$$

The value of an option whose payoff depends on the (discretely observed) average asset price is given by $V = V(S, A, t)$, where the average A can take on any value. We assume that the underlying asset price follows the process

$$dS = \mu S dt + \sigma S dZ$$

where μ is the drift rate, σ is the volatility, and dZ is the increment of a Wiener process. At times other than observation dates, standard arguments imply that the option satisfies the usual Black-Scholes equation

$$V_t + \frac{\sigma^2 S^2}{2} V_{SS} + rSV_S - rV = 0 \tag{2.3}$$

where r is the risk free interest rate. At observation dates no-arbitrage considerations imply that

$$V(S, A^{n+1}, t_n^+) = V(S, A^n, t_n^-) \tag{2.4}$$

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

$$A^{n+1} = A^n + \frac{(S - A^n)}{n+2} \tag{2.5}$$

with $A^0 = S^0$. Equations (2.3-2.5) represent an infinite set of one-dimensional PDEs, where the average value A appears as a parameter. These one-dimensional PDEs (for a given value of A) depend only on the asset price S. At each observation date t_n, these one-dimensional problems exchange information based on conditions (2.4-2.5).

To complete the specification of the problem, terminal and boundary conditions are required. Equations (2.3-2.5) are posed on the domain $0 \leq S \leq \infty$ and $0 \leq A \leq \infty$. At $t = T$, the terminal condition is given by the payoff function. For example, a fixed strike Asian call/put would result in the condition

$$\begin{aligned} V(S, A, t = T) &= \max(E - A, 0) \quad \text{for a put} \\ &= \max(A - E, 0) \quad \text{for a call} \ , \end{aligned} \tag{2.6}$$

where E is the exercise price. At $S = 0$, equation (2.3) reduces to the ordinary differential equation $V_t - rV = 0$, so no boundary condition needs to be explicitly enforced here. Since equation (2.3) contains no explicit dependence on A, we need only specify the behavior of V as $S \to \infty$. Initially, the asymptotic form of V as $S \to \infty$ is given by the payoff condition (2.6). At observation dates, the asymptotic form is updated using the jump conditions (2.5). Following Hull and White (1993) and Barraquand and Pudet (1996), we assume (for simplicity) that there is a constant interval between the discrete observations, i.e. $t_{n+1} - t_n = \delta t$. Note that the interval between observations is completely unrelated, in general, to a timestep Δt in a discretization of the PDE.

In some cases (e.g. the floating strike as described in Andreasen (1998)), the two state variable problem can be reduced to a single state variable problem. However, in general (e.g. an American fixed strike, discretely observed option), the problem cannot be reduced to one with a single factor. Since our focus in this article is on convergence issues related to interpolation of discrete quantities, we will not concentrate on options with early exercise. However, all of the algorithms discussed here can be trivially generalized to handle this, and we do provide some numerical results for this case.

Although most (if not all) Asian option contracts are based on discrete observation of the underlying asset, in some cases the assumption of continuous monitoring can provide a reasonable approximation. In this case, the average is defined as

$$\begin{aligned} A &= \frac{\int_0^t S(\tau)\, d\tau}{t} \\ &\Rightarrow \quad dA = \frac{(S-A)\, dt}{t} \ . \end{aligned} \tag{2.7}$$

Using standard arguments, the value of an option whose payoff is a function of S and A, where A is continuously observed, is given by $V^c = V^c(S,A,t)$, where V^c satisfies (see Barraquand and Pudet (1996))

$$V_t^c + rS\frac{\partial V^c}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V^c}{\partial S^2} + \frac{(S-A)}{t}\frac{\partial V^c}{\partial A} - rV^c = 0 \ . \tag{2.8}$$

Clearly, as the number of observations becomes infinite, and the time between observations tends to zero, then definition (2.1) $\to$ definition (2.7), and the solution of equations (2.3-2.5) tends to the solution of equation (2.8). Note however, that if constant observation intervals of size δt are used in equations (2.3-2.5), then the discretely observed model will converge only at a rate of $O(\delta t)$ to the continuously observed limit.

In the following, we will base our analysis on the discretely observed Asian options defined by equations (2.3-2.5), where we take the limit as the number of observations becomes large, and hence converge to the continuously observed limit. This approach is a convenient starting point for analyzing the algorithms in Barraquand and Pudet (1996) and Hull and White (1993), as well as PDE methods for discretely observed Asian options. In addition, it is straightforward to modify these methods so that they can be used to price options where δt represents a fixed observation interval (e.g. one day). Note that PDE methods for directly solving the continuous limit equation (2.8) are discussed in Zvan et al. (1998).

3 Analysis of the Forward Shooting Grid Method

In this section, we introduce a general framework which will be used to analyze the FSG method. We use the notation in Barraquand and Pudet (1996) to facilitate comparison with that work. Let

$$\begin{aligned} \Delta Z &= \sigma\sqrt{\Delta t} \\ \Delta Y &= \rho \Delta Z \end{aligned} \tag{3.1}$$

where σ is the volatility, Δt is the (discrete) 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. Also, note that in Barraquand and Pudet (1996) it is implicitly assumed that the discrete timestep Δt is equal to the observation frequency δt, so that convergence to the continuously observed limit is desired as $\Delta t \to 0$.

Let discrete values of the asset price S and average price A be given by

$$\begin{aligned} S_j^n &= S_0 e^{j\Delta Z} \\ A_k^n &= S_0 e^{k\Delta Y} \\ & n = 0, \ldots, N; \;\; j = -n, \ldots, +n; \;\; k = -k_m(n), \ldots, +k_m(n) \end{aligned} \tag{3.2}$$

where N is the number of timesteps and

$$k_m(n) = n/\rho \; . \tag{3.3}$$

Recall that $1/\rho$ is an integer which controls the fineness of the quantization in the average direction. To avoid unnecessary algebraic complication, and without any loss of generality, take $S_0 = 1$ in equation (3.2), which then becomes

$$\begin{aligned} S_j^n &= e^{j\Delta Z} \\ A_k^n &= e^{k\Delta Y} \; . \end{aligned} \tag{3.4}$$

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 $S_j^n \to S_{j+1}^{n+1}$ with (risk-neutral) probability p, and a downward transition $S_j^n \to S_{j-1}^{n+1}$ with (risk-neutral) 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:

$$\begin{aligned} A_{k^+(j,k)}^{n+1} &= A_k^n + \frac{\left(S_{j+1}^{n+1} - A_k^n\right)}{(n+2)} \\ A_{k^-(j,k)}^{n+1} &= A_k^n + \frac{\left(S_{j-1}^{n+1} - A_k^n\right)}{(n+2)} \end{aligned} \tag{3.5}$$

with $A_0^0 = S_0^0 = 1$. Each asset price node in the tree has associated with it a set of average values A_k^n and

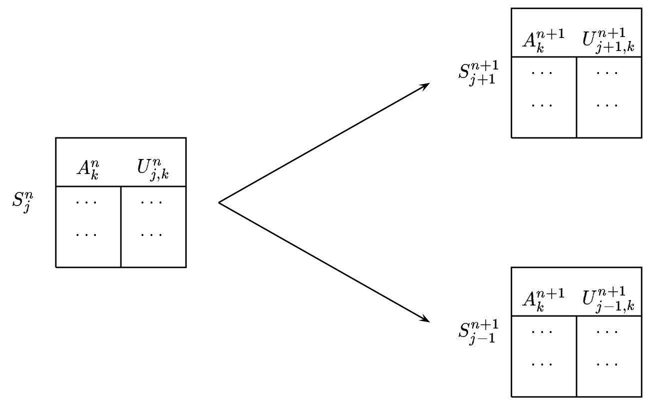

FIGURE 1: *Asset price tree indicating that a set of discrete averages (A_k^{n+1}) and approximate option values ($U_{j+1,k}^{n+1}$) exists at each node of the tree.*

approximate 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 (3.5) do not necessarily coincide with the lattice values in equation (3.4). This necessitates some form of interpolation (Hull and White (1993); Barraquand and Pudet (1996)). For future reference, define

$$\begin{aligned} k_{floor}^{\pm}(j,k) &= floor\left[\frac{\log\left(A_{k^{\pm}(j,k)}^{n+1}\right)}{\rho\Delta Z}\right] \\ k_{ceil}^{\pm}(j,k) &= k_{floor}^{\pm}(j,k)+1\ . \end{aligned} \tag{3.6}$$

These are simply the indices for the lattice average values in equation (3.4) which bracket the updated average values in equation (3.5).

Let $U_{j,k}^n = U(S_j^n, A_k^n, n\Delta t)$ be the approximate value of the option obtained using the FSG method at $t = n\Delta t$, $A = A_k^n$, $S = S_j^n$. The value of the option given a suitable terminal payoff condition $U_{j,k}^N$ 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):

$$\begin{aligned} U_{j,k}^n &= e^{-r\Delta t}\Big[p\Big(\alpha_{k_{floor}^+(j,k)}^{n+1} U_{j+1,k_{floor}^+(j,k)}^{n+1} + \Big(1-\alpha_{k_{floor}^+(j,k)}^{n+1}\Big) U_{j+1,k_{ceil}^+(j,k)}^{n+1}\Big) \\ &+ (1-p)\Big(\alpha_{k_{floor}^-(j,k)}^{n+1} U_{j-1,k_{floor}^-(j,k)}^{n+1} + \Big(1-\alpha_{k_{floor}^-(j,k)}^{n+1}\Big) U_{j-1,k_{ceil}^-(j,k)}^{n+1}\Big)\Big] \\ & n = N-1,\ldots,0;\quad j=-n,\ldots,+n;\quad k=-k_m(n),\ldots,+k_m(n)\ . \end{aligned} \tag{3.7}$$

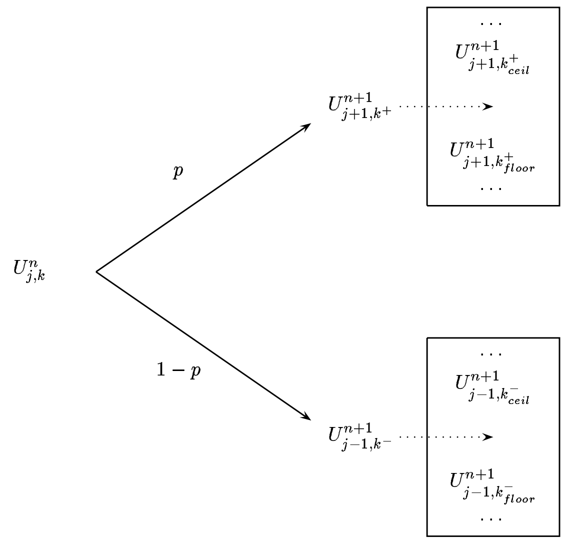

FIGURE 2: *The values of* $U^{n+1}_{j+1,k^+(j,k)}$ *and* $U^{n+1}_{j-1,k^-(j,k)}$ *must be interpolated from the known values at* $t = n+1$.

In equation (3.7), the risk-neutral probability p is

$$p = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{+\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \,, \tag{3.8}$$

and the α's are determined by the type of interpolation used, nearest lattice point (a.k.a. nearest neighbor) or linear. Note that in both of these cases

$$\begin{aligned} 0 &\leq \alpha^{n+1}_{k^+_{floor}(j,k)} \leq 1 \\ 0 &\leq \alpha^{n+1}_{k^-_{floor}(j,k)} \leq 1 \,. \end{aligned} \tag{3.9}$$

4 Error Analysis

In this section, we will analyze the FSG method as described above, and relate this method to the problem as posed in equations (2.3-2.5). For expository purposes, we shall make the following two simplifying assumptions:

1. The exact solution V to equations (2.3-2.5) for a given value of the observation interval δt has continuous bounded derivatives up to fourth order with respect to S, and up to second order with respect to A and t. The payoff condition has continuous bounded first and second derivatives with respect to A and S. (A fixed strike payoff is independent of S. In Appendix D, we show how to take into account the fact that the payoff derivatives w.r.t. A are not bounded everywhere).

2. We will assume that the effect of interpolation errors, introduced at large (but finite) values of $A > A^*$ have negligible effect on the approximate solution $U^0_{0,0}$. This is plausible, since we expect that states with very large values of A are exceedingly improbable. This assumption can be removed as shown in Appendix C.

The above assumptions allow us to focus on the effect of the interpolation error (as in equation (3.7)), without tedious algebraic complication. In fact, the derivatives of usual payoffs are not smooth at the strike. However, intuitively, we can expect that a diffusion type equation will rapidly smooth out any initial rough data. Similarly, we would also expect that the effect of states with large values of the average would have a vanishingly small effect ($A^* \to \infty$) at any finite initial value of S, since these states have a very low probability of being reached in a finite time. In Appendix D, we indicate in a heuristic way how these assumptions can be relaxed. However, we anticipate that a completely rigorous argument to account for lack of smoothness of the payoff would be quite long and involved. We leave this as a topic for future research.

Let $V^n_{j,k}$ denote the *exact* solution of equations (2.3-2.5), evaluated at $S = S^n_j$, $A = A^n_k$, $t = (n\delta t)^+ = t^+_n$, for a given discrete observation interval δt:

$$V^n_{j,k} = V(S^n_j, A^n_k, t^+_n) \ . \tag{4.1}$$

Note that this exact solution is independent of any approximation used to solve the PDEs, but does depend on the discrete observation interval δt. Also note that $V^n_{j,k}$ refers to values the instant after a discrete observation. We denote values of V at the instant before a discrete observation by $(V^n_{j,k})^-$, at $S = S^n_j$, $A = A^n_k$, $t = (n\delta t)^- = t^-_n$:

$$(V^n_{j,k})^- = V(S^n_j, A^n_k, t^-_n) \ . \tag{4.2}$$

This distinction is required in view of the jump conditions (2.4). Recall that, from equation (2.1), $A^0 = S^0$. Therefore no jump condition is required at $n = 0$, and so $(V^0_{0,0})^- = V^0_{0,0}$. In order to be consistent with the FSG algorithm as described above, we let the discrete observation period be equal to the discrete timestep, i.e. $\delta t = \Delta t$.

Observe that if we define $X = \log S$, then equation (2.3) becomes (for fixed A)

$$V_t + \frac{\sigma^2}{2} V_{XX} + (r - \frac{\sigma^2}{2}) V_X - rV = 0 \ . \tag{4.3}$$

In Appendix A, by means of Taylor series, we show the following result:

Proposition 1 (Recursion Satisfied by the Exact Solution to Equation (4.3)) *The exact solution to equation (4.3) at discrete points (j,k,n) satisfies*

$$\begin{aligned} V^n_{j,k} &= e^{-r\Delta t}\left[p\left(V^{n+1}_{j+1,k}\right)^- + (1-p)\left(V^{n+1}_{j-1,k}\right)^-\right] + \text{truncation error} \ , \\ &\quad \text{truncation error} = O\left[(\Delta t)^2\right] . \end{aligned} \tag{4.4}$$

In order to obtain a recursion in terms of $V^n_{j,k}$ in equation (4.4), we must eliminate the dependence on $\left(V^{n+1}_{j\pm1,k}\right)^-$. Let $V^{n+1}_{j+1,k^+(j,k)}$ denote the value of the exact solution to equations (2.3-2.5), evaluated at asset price S^{n+1}_{j+1}, and average value $A^{n+1}_{k^+(j,k)}$ (as defined in equation (3.5)) with discrete observation interval $\delta t = \Delta t$, and at t^+_{n+1}. Note that the jump conditions (2.4) can be written (at discrete points) as

$$\begin{aligned} \left(V^{n+1}_{j+1,k}\right)^- &= V^{n+1}_{j+1,k^+(j,k)} \\ \left(V^{n+1}_{j-1,k}\right)^- &= V^{n+1}_{j-1,k^-(j,k)} \; . \end{aligned} \tag{4.5}$$

Substituting equation (4.5) into equation (4.4), we obtain

$$\begin{aligned} V^n_{j,k} &= e^{-r\Delta t}\left[pV^{n+1}_{j+1,k^+(j,k)} + (1-p)V^{n+1}_{j-1,k^-(j,k)}\right] + \text{truncation error} \\ & \quad n = N-1,\ldots,0; \quad j = -n,\ldots,+n; \quad k = -k_m(n),\ldots,+k_m(n) \; . \end{aligned} \tag{4.6}$$

From the Taylor series expansion we have

$$\begin{aligned} V^{n+1}_{j+1,k^+(j,k)} &= \alpha^{n+1}_{k^+_{floor}(j,k)} V^{n+1}_{j+1,k^+_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right) V^{n+1}_{j+1,k^+_{ceil}(j,k)} \\ &+ \left(\beta^q_{k^+_{floor}(j,k)}\right)^{n+1} \\ V^{n+1}_{j-1,k^-(j,k)} &= \alpha^{n+1}_{k^-_{floor}(j,k)} V^{n+1}_{j-1,k^-_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right) V^{n+1}_{j-1,k^-_{ceil}(j,k)} \\ &+ \left(\beta^q_{k^-_{floor}(j,k)}\right)^{n+1} \end{aligned} \tag{4.7}$$

where for $q = 1$ (nearest lattice point interpolation) and $q = 2$ (linear interpolation)

$$\begin{aligned} \left(\beta^1_{k^+_{floor}(j,k)}\right)^{n+1} &= \min\left[\left(A^{n+1}_{k^+_{ceil}(j,k)} - A^{n+1}_{k^+(j,k)}\right), \left(A^{n+1}_{k^+(j,k)} - A^{n+1}_{k^+_{floor}(j,k)}\right)\right] \frac{\partial V^{n+1}_{j+1}(\eta)}{\partial A} \\ \left(\beta^2_{k^+_{floor}(j,k)}\right)^{n+1} &= -\frac{\left(A^{n+1}_{k^+(j,k)} - A^{n+1}_{k^+_{floor}(j,k)}\right)}{2} \left(A^{n+1}_{k^+_{ceil}(j,k)} - A^{n+1}_{k^+(j,k)}\right) \frac{\partial^2 V^{n+1}_{j+1}(\eta)}{\partial A^2} \end{aligned} \tag{4.8}$$

with $\eta \in \left[A^{n+1}_{k^+_{floor}(j,k)}, A^{n+1}_{k^+_{ceil}(j,k)}\right]$ in each case. Substituting equation (4.7) into equation (4.6) gives

$$\begin{aligned} V^n_{j,k} &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)} V^{n+1}_{j+1,k^+_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right) V^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)} V^{n+1}_{j-1,k^-_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right) V^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)\right] \\ &+ e^{-r\Delta t}\left[p\left(\beta^q_{k^+_{floor}(j,k)}\right)^{n+1} + (1-p)\left(\beta^q_{k^-_{floor}(j,k)}\right)^{n+1}\right] \\ &+ \text{truncation error} \; . \end{aligned} \tag{4.9}$$

Let the difference between the exact solution V (of equations (2.3-2.5) with observation interval $\delta t = \Delta t$), and the approximate solution U (from the FSG algorithm) be denoted by $E^n_{j,k}$ where

$$E^n_{j,k} = V^n_{j,k} - U^n_{j,k} \ . \tag{4.10}$$

Then an equation for the propagation of the error due to interpolation and truncation error can be deduced by subtracting equation (3.7) from equation (4.9) to obtain

$$\begin{aligned} E^n_{j,k} &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)} E^{n+1}_{j+1,k^+_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right) E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)\right. \\ &+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)} E^{n+1}_{j-1,k^-_{floor}(j,k)} + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right) E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)\right] \\ &+ \text{interpolation error} + \text{truncation error} \ , \end{aligned} \tag{4.11}$$

where

$$\text{interpolation error} = e^{-r\Delta t}\left[p\left(\beta^q_{k^+_{floor}(j,k)}\right)^{n+1} + (1-p)\left(\beta^q_{k^-_{floor}(j,k)}\right)^{n+1}\right] . \tag{4.12}$$

From the recursion (4.11), we can bound the cumulative effect of the interpolation error and the truncation error on the solution at $S^0_{0,0}$, which is denoted by $E^0_{0,0}$. Details of this are given in the Appendices. However, for expository purposes, we will use a heuristic argument to obtain the main result.

We assume that there exists an A^* such that the effect of interpolation errors induced at $A > A^*$, is negligible at $S^0_{0,0}$ (this assumption is removed in Appendix C). We will also assume that

$$\begin{aligned} \left|\frac{\partial V^n_j(A)}{\partial A}\right| &\le M_1 \\ \left|\frac{\partial^2 V^n_j(A)}{\partial A^2}\right| &\le M_2 \end{aligned} \tag{4.13}$$

for any n, j, where M_1 and M_2 are constants independent of Δt. Consequently, the interpolation errors in equation (4.8) can be bounded by

$$\max\left[\left|\beta^q_{k^+_{floor}(j,k)}\right|^{n+1}, \left|\beta^q_{k^-_{floor}(j,k)}\right|^{n+1}\right] \le \left[M_q (A^*)^q \left(1-e^{-\rho\Delta Z}\right)^q\right] . \tag{4.14}$$

Equation (4.14) becomes, in the limit as $\Delta t \to 0$,

$$\text{interpolation error} = O\left((\rho\sigma\sqrt{\Delta t})^q\right) . \tag{4.15}$$

If we define the maximum error at step n as

$$\|E^n\| = \max_{j,k} \left| E^n_{j,k} \right| \tag{4.16}$$

then, since the interpolation coefficients α and the probabilities p are all in the range $[0,1]$, it follows from equation (4.11) that

$$\begin{aligned} \|E^n\| &\leq e^{-r\Delta t} \left[p \left(\alpha^{n+1}_{k^+_{floor}(j,k)} + 1 - \alpha^{n+1}_{k^+_{floor}(j,k)} \right) \right. \\ &+ \left. (1-p) \left(\alpha^{n+1}_{k^-_{floor}(j,k)} + 1 - \alpha^{n+1}_{k^-_{floor}(j,k)} \right) \right] \left\| E^{n+1} \right\| \\ &+ \text{interpolation error} + \text{truncation error} \\ &\leq \left\| E^{n+1} \right\| + \text{interpolation error} + \text{truncation error} \end{aligned} \tag{4.17}$$

Assuming that

$$\begin{aligned} \text{interpolation error} &= O((\sqrt{\Delta t})^q) \\ \text{truncation error} &= O((\Delta t)^2) \end{aligned} \tag{4.18}$$

then it follows from equations (4.17-4.18) that after $N = O(1/(\Delta t))$ steps, we have that the worst case error bound is

$$\begin{aligned} \left\| E^0 \right\| &\leq O(\Delta t) + O\left(\frac{(\sqrt{\Delta t})^q)}{\Delta t} \right) \\ &= O(\Delta t) + O\left((\Delta t)^{q/2-1} \right) \end{aligned} \tag{4.19}$$

so that if nearest neighbor $(q = 1)$ or linear interpolation $(q = 2)$ is used, there is no guarantee that the numerical result will converge to the correct solution as $\Delta t \to 0$. More precisely, in Appendix B, we show the following result:

Proposition 2 (Convergence of the Forward Shooting Grid Method) *Under the assumption that the derivatives w.r.t. A in equations (4.8) are bounded, then the cumulative error due to interpolation* $\left\| E^0_{0,0} \right\|^I$ *is bounded by*

$$\left\| E^0_{0,0} \right\|^I \leq \min \left[B, NC_q \left(1 - e^{-\rho \Delta Z} \right)^q \right] \tag{4.20}$$

where B is a constant which depends on the strike, but is independent of Δt*,* C_q *is a constant which is independent of* Δt*, but depends on the type of interpolation used, and* $N\Delta t = T$*. The cumulative error due to the truncation error is* $O(\Delta t)$*.*

Note that equation (4.20) can be approximated for small Δt as

$$\begin{aligned} \left\| E^0_{0,0} \right\|^I &\leq \min \left[B, NC_q \left(1 - e^{-\rho \Delta Z} \right)^q \right] \\ &\simeq \min \left[B, \frac{TC_q \left(\rho \sigma \sqrt{\Delta t} \right)^q}{\Delta t} \right] \quad ; \quad \Delta t \to 0 \end{aligned} \tag{4.21}$$

Observe that for $q = 1$ and $q = 2$ the bound does *not* tend to zero as $\Delta t \to 0$.

We have derived an upper bound for the error of the FSG method. If the upper bound tends to zero as $\Delta \to 0$, then convergence is ensured. However, the upper bound in equation (4.21) does not show this, so we are unable to make the claim that the FSG method is convergent. In fact, this suggests that we can expect problems with the FSG method. This is due to the fact that at each step, an interpolation error of size

$$(\sqrt{\Delta t})^q \tag{4.22}$$

is introduced. In the worst case, the cumulative error is

$$O\left(\frac{(\sqrt{\Delta t})^q}{\Delta t} \right) . \tag{4.23}$$

Of course, this analysis does not say give us any information about whether this worst case is actually attained. In a subsequent section, we will give numerical examples which show that these worst case errors are in fact attained, and that the cumulative error does not tend to zero for the FSG method.

It is clear, then, that in order to guarantee convergence as $\Delta t \to 0$, we must construct a method where the interpolation error at each step tends to zero faster than $O(\sqrt{\Delta t})$. In the context of a lattice method, it is seemingly natural to choose $\Delta Z = \sigma \sqrt{\Delta t}$. However, this spacing in the A direction is not fine enough to ensure convergence if linear or nearest neighbor interpolation is used, due to the cumulative effect of these local errors after $O(1/(\Delta t))$ timesteps.

For payoffs of type (2.6), V_{AA} becomes unbounded as $t \to T$ (the first derivative with respect to A is discontinuous at $A = E$). However, the exact payoff is available at $t = T$ $(n = N)$ so that no interpolation error is induced in making the transition from $t_N^+ \to t_N^-$. Also, note that even if V_{AA} does not exist, the interpolation error does not become unbounded, but simply reduces to a first order error in the spacing in the A direction. At $n = 0$ or $t = 0$, we have that $S_0^0 = A_0^0$, so that $V_A = V_{AA} = 0$. This can also be seen from the continuously observed model equation (2.8), where the coefficient of the V_A term becomes infinite for $S \neq A$, which means that the solution becomes independent of A as $t \to 0$. Consequently, although V_{AA} is large at $A = E$ as $t \to T$, $V_{AA} \to 0$ as $t \to 0$. The jump conditions (2.4) tend to smooth derivatives in the A direction, while the diffusion term tends to smooth in the S direction. In Appendix D, we discuss the form of V_{AA} as $t \to T$, and indicate how the error analysis would have to be modified in order to take this into account. However, as mentioned previously, a complete analysis is beyond the scope of this work.

In equation (4.20), 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

$$\rho = O\left[(\Delta t)^{(2-q/2)/q}\right] \ . \tag{4.24}$$

For the case of nearest neighbor interpolation $(q = 1)$, $\rho = O\left[(\Delta t)^{3/2}\right]$. This implies that at timestep n (from equation (3.3)),

$$k_m \quad = \quad O\left(\frac{n}{(\Delta t)^{3/2}}\right) \ . \tag{4.25}$$

This 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 (4.24), 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.

5 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

$$A_k^n = S_0 e^{kh} \tag{5.1}$$

where, for given h, the range in k values in equation (5.1) is selected to span the possible averages at timestep n. Recall that in equation (3.2) 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 (B.14), we obtain the estimate

$$\left|E_{0,0}^0\right|^I \leq \frac{TC_q\left(1-e^{-h}\right)^q}{\Delta t} \ . \tag{5.2}$$

Hull and White (1993) suggest either linear or quadratic interpolation. It is worth emphasizing that Hull and White specify h as a constant, but our analysis indicates that convergence of this method requires that h be specified as an appropriate function of Δt. If we take $h = C\Delta t$, for example, then

$$\left|E_{0,0}^0\right| \leq \frac{TC_q\left(1-e^{-h}\right)^q}{\Delta t} \simeq TC_q C^q (\Delta t)^{q-1} \ . \tag{5.3}$$

We will refer to this version of the method as the "modified Hull and White method". Equation (5.3) indicates that the modified 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 coefficients 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 (5.3) considers only the effect of the interpolation error. There will also be the usual truncation error of size $O(\Delta t)$, so that the global convergence rate of the modified Hull and White method should be of $O(\Delta t)$.

Following Chalasani et al. (1999), 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

$$\begin{aligned} A^n_{max} &= \frac{\sum_{k=0}^{n} e^{k\sigma\sqrt{\Delta t}}}{n+1} \\ &= \frac{1-e^{\sigma\sqrt{\Delta t}(n+1)}}{(n+1)\left(1-e^{\sigma\sqrt{\Delta t}}\right)} \\ &\simeq O\left[\frac{e^{\sigma(n+1)\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)}\right] \quad \text{as} \quad n \to \infty \,. \end{aligned} \tag{5.4}$$

The minimum possible value of the average after n steps is

$$\begin{aligned} A^n_{min} &= \frac{\sum_{k=0}^{n} e^{-k\sigma\sqrt{\Delta t}}}{n+1} \\ &= \frac{1-e^{-\sigma\sqrt{\Delta t}(n+1)}}{(n+1)\left(1-e^{-\sigma\sqrt{\Delta t}}\right)} \\ &\simeq O\left[\frac{e^{\sigma\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)}\right] \quad \text{as} \quad n \to \infty \,. \end{aligned} \tag{5.5}$$

Letting

$$\begin{aligned} e^{m_1 C\Delta t} &= \frac{e^{\sigma(n+1)\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)} \\ e^{m_2 C\Delta t} &= \frac{e^{\sigma\sqrt{\Delta t}}}{(n+1)\left(e^{\sigma\sqrt{\Delta t}}-1\right)} \,, \end{aligned} \tag{5.6}$$

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 $O(N^{7/2})$.

Note that there are other possibilities for choice of the node spacing in the average direction. The Hull and White method uses a fixed h, and the range of k (equation (5.1)) is adjusted at each node. Alternatively, one could fix the range of k, and adjust h at each node to span the maximum and minimum possible averages at each node. This latter approach would use smaller grid spacing at certain lattice nodes. However, the cumulative error will be proportional to the largest value of h at any timestep, so this use of *local refinement* in the average direction may improve the constant in the rate of convergence, but the order of convergence will remain the same.

6 Analysis of PDE Methods

The discrete Asian option pricing problem can be solved using the system of one-dimensional PDEs (2.3-2.5) as described in Wilmott et al. (1993), Dempster et al. (1998) and Zvan et al. (1999). Convergence of the PDE method is easily demonstrated. Away from the observation dates, we simply solve a set of one-dimensional problems (equation (2.3)) for each discrete value of the average, using standard numerical methods. The PDEs are posed on a finite domain, $0 \le S \le S_{max}$ and $0 \le A \le A_{max}$. 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 (2.5). Since generally A^{n+1} will not coincide with a grid node, interpolation (linear or quadratic) is used to estimate the value of the approximate 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 ΔS_{max} be the maximum grid spacing in the S or A direction, then the interpolation error at each step is

$$\text{interpolation error at each observation} = O\left[(\Delta S_{max})^q\right] \tag{6.1}$$

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

$$\text{global interpolation error} = O\left[\frac{(\Delta S_{max})^q}{\Delta t}\right] . \tag{6.2}$$

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

$$\text{global discretization error} = O\left[\frac{(\Delta S_{max})^q}{\Delta t}\right] + O\left[(\Delta S_{max})^2\right] + O\left[(\Delta t)^2\right] . \tag{6.3}$$

If we use quadratic interpolation as in Zvan et al. (1999), and take the limit in such a way that $\Delta S_{max} = C\Delta t$ where C is a constant, then we obtain

$$\text{global discretization error} = O\left[(\Delta t)^2\right] . \tag{6.4}$$

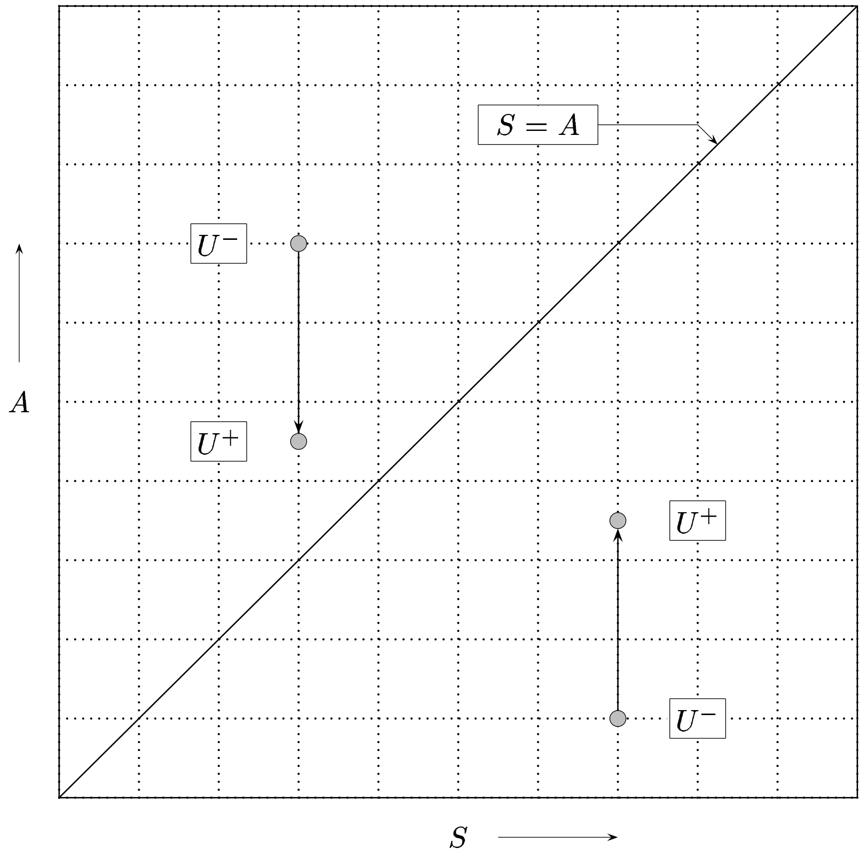

FIGURE 3: *Between each observation date, one-dimensional PDEs for each value of the average A are solved. The values of the approximate option price U^- before each observation date are interpolated from the values just after the observation date U^+.*

Further details concerning the PDE method can be found in Zvan et al. (1999).

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 PDEs 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 (3.1)). 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 (6.4) only takes into account the truncation error of the discretization of the PDE and the interpolation error. There is an additional error due to the fact that we are attempting to converge to a continuously observed Asian option using a discretely observed model. 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 asymptotic convergence of the PDE approach may be more useful.

7 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, modified 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, but modified in the manner described above in section 5. In particular, the average node spacing parameter h in equation (5.1) was specified as:

$$h = \alpha\sqrt{\frac{0.25}{T}\sigma^2\Delta t}. \tag{7.1}$$

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 (7.1) controls the fineness of the grid in the average direction. Three values of α were used for each test case. 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.

Preliminary evidence regarding the convergence properties of the FSG method is provided in Table 1, which shows results for two examples (one with low volatility and short time until maturity, one with high volatility and long time until maturity) of Asian call options where the exercise price of the option is set to zero. Although of little practical relevance, this is an interesting case because: i) as noted by Barraquand and Pudet (1996), there is an analytic solution; and ii) since the payoff function is linear, linear interpolation is exact. Our analysis above suggests that although the FSG method should perform poorly with nearest lattice point interpolation, it should converge to the analytic solution with linear interpolation. This behavior is clearly documented in the table. For Case 1, the analytic solution value is \$98.7604. When $\rho \neq 0.1$, the nearest lattice point scheme gives unsatisfactory answers. If $\rho = 0.1$, the calculated prices are reasonably accurate when the number of timesteps is small, but the performance of the method deteriorates markedly

TABLE 1: *Convergence of the forward shooting grid method for zero strike European-Asian call options.* ρ *is a parameter which specifies the grid spacing in the average direction (smaller* ρ *means a finer grid). CPU times are normalized to* 1 *second for the case where* $\rho = 1.0$ *and there are* 50 *timesteps. The exact solution for this problem is linear in A, so the interpolation error is identically zero for linear interpolation. Note the poor results for nearest neighbor interpolation.*

		Nearest Neighbor Interpolation		Linear Interpolation	
ρ	Timesteps	Option Value	CPU (sec)	Option Value	CPU (sec)
Case 1: $r = .10$, $\sigma = .10$, $T = 0.25$ years, $E = \$0$, $S = \$100$					
1.0	50	97.9020	1.0	98.7602	1.8
	100	97.7548	7.5	98.7603	13
	200	97.6709	59	98.7603	104
	400	97.6209	470	98.7603	828
0.5	50	98.2868	2.0	98.7602	3.3
	100	97.9820	15	98.7603	26
	200	97.7703	118	98.7603	207
	400	97.6550	943	98.7603	1655
0.1	50	98.7667	9.3	98.7602	16
	100	98.7636	14	98.7603	26
	200	98.7525	590	98.7603	1035
	400	98.6688	4705	98.7603	8273
Analytic Value: \$98.7604					
Case 2: $r = .10$, $\sigma = .50$, $T = 5.0$ years, $E = \$0$, $S = \$100$					
1.0	50	74.7619	1.0	78.6535	1.8
	100	69.7366	7.8	78.6736	13
	200	66.1876	62	78.6837	105
	400	63.6150	488	78.6888	835
0.5	50	81.0266	2.0	78.6535	3.3
	100	76.8897	16	78.6736	26
	200	72.4131	123	78.6837	209
	400	68.2997	978	78.6888	1670
0.1	50	78.5366	9.5	78.6535	17
	100	78.9473	77	78.6736	133
	200	80.7914	615	78.6837	1108
	400	82.5950	4890	78.6888	8413
Analytic Value: \$78.6939					

as the number of timesteps is increased to 400. By contrast, the linear interpolation results are virtually identical to the analytic value. Note that although the table provides results for the linear interpolation scheme for all three values of ρ, much of this information is redundant because the linear payoff structure implies that the calculated answers are independent of ρ (though they obviously do depend on the number of timesteps). This general pattern is repeated for Case 2. Once again the nearest lattice point technique fails to provide satisfactory answers, except possibly for $\rho = 0.1$ with a small number of timesteps. The results for the linear interpolation scheme are not as close to the analytic value as for Case 1, but it is worth noting that extrapolation to $\Delta t = 0$ of the calculated answers for each value of ρ gives a price estimate equal to the analytic solution value of \$78.6939.

For the remainder of this section, we concentrate on more realistic cases, using the same examples as above but changing the exercise price to \$100. 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 about \$ $1.8515 \pm .0001$. 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 \$1.8522, a little higher than the true price.

Turning to Case 2, we begin by noting that the correct price here is $\simeq$ \$$28.40525 \pm .00015$. 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 \$28.4147. 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 modified Hull and White algorithm results for both cases (with an exercise price of \$100) 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 (7.1), 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 $O(\Delta t)$ (consistent with our analysis). Extrapolation to $\Delta t = 0$ of the values when $\alpha = 1$ gives price estimates of \$1.8516 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 \$1.8514 for Case 1 and \$28.4054 for Case 2, in excellent agreement with the modified Hull and White extrapolated prices of \$1.8516 and \$28.4051. As both of these methods are convergent, this leads to the conclusion that the true prices are $\simeq$ \$$1.8515 \pm .0001$ and \$$28.40525 \pm .00015$. By contrast, recall that the FSG extrapolated prices were \$1.8522 and \$28.4147. This is clearly consistent with our analysis indicating that the FSG method

TABLE 2: *Convergence of the forward shooting grid method for fixed strike European-Asian call options. ρ is a parameter which specifies the grid spacing in the average direction (smaller ρ means a finer grid). CPU times are normalized to* 1 *second for the case where* $\rho = 1.0$ *and there are* 50 *timesteps. The analysis suggests that large errors will occur as* $\Delta t \rightarrow 0$ *for nearest lattice point interpolation. If linear interpolation is used, the FSG method should converge to the correct solution plus a small constant.*

		Nearest Neighbor Interpolation		Linear Interpolation	
ρ	Timesteps	Option Value	CPU (sec)	Option Value	CPU (sec)
Case 1: $r = .10$, $\sigma = .10$, $T = 0.25$ years, $E = \$100$, $S = \$100$					
1.0	50	0.5875	1.0	1.8738	1.8
	100	0.3892	8.3	1.8691	13.3
	200	0.2634	65	1.8649	106
	400	0.1806	520	1.8615	843
0.5	50	1.1974	2.0	1.8603	3.0
	100	0.8089	16	1.8592	23
	200	0.5208	130	1.8577	187
	400	0.3391	1035	1.8563	1480
0.1	50	1.8533	10	1.8492	16
	100	1.8524	82	1.8508	130
	200	1.8347	650	1.8516	1038
	400	1.7147	5175	1.8519	8305
Case 2: $r = .10$, $\sigma = .50$, $T = 5.0$ years, $E = \$100$, $S = \$100$					
1.0	50	16.2053	1.0	28.7217	1.5
	100	10.7957	7.5	28.6631	12
	200	6.9113	60	28.6052	92
	400	4.0803	478	28.5556	728
0.5	50	25.0508	2.0	28.5168	2.8
	100	19.9843	15	28.5107	23
	200	14.7166	120	28.4934	183
	400	9.9951	955	28.4745	1405
0.1	50	28.2968	9.5	28.3440	14
	100	28.6676	75	28.3816	114
	200	29.3198	603	28.3996	915
	400	29.1303	4765	28.4071	7270

TABLE 3: *Convergence of the modified Hull and White method for fixed strike European-Asian call options. The grid size is the number of nodes in the A direction at $t = T$. The parameter α controls the grid spacing in the A direction (smaller α means a finer grid). CPU times are normalized to* 1 *second for the forward shooting grid method with $\rho = 1.0$ and* 50 *timesteps. The analysis predicts that the modified H&W method is convergent as $\Delta t \to 0$ for any α.*

α	Timesteps	Grid Size	Option Value	CPU (sec)
Case 1: $r = .10$, $\sigma = .10$, $T = 0.25$ years, $E = \$100$, $S = \$100$				
40	50	200	1.8542	2.0
	100	577	1.8529	24
	200	1503	1.8523	247
	400	4279	1.8519	2760
20	50	391	1.8502	4
	100	1055	1.8509	42
	200	2969	1.8512	478
	400	8214	1.8514	5298
4	50	1794	1.8486	18
	100	5050	1.8501	204
	200	14247	1.8508	2293
	400	40198	1.8512	25918
Case 2: $r = .10$, $\sigma = .50$, $T = 5.0$ years, $E = \$100$, $S = \$100$				
10	50	163	28.5098	1.8
	100	450	28.4583	19
	200	1225	28.4319	207
	400	3394	28.4186	2283
5	50	308	28.4310	3.3
	100	844	28.4180	36
	200	2356	28.4115	400
	400	6579	28.4083	4455
1	50	1440	28.3899	15
	100	4051	28.3972	168
	200	11415	28.4011	1893
	400	32196	28.4031	21370

TABLE 4: *Convergence of the PDE method for fixed strike European-Asian call options. A Cartesian product grid is used. The grid size is given as number of nodes in the S and A directions. CPU times are normalized to* 1 *second for the forward shooting grid method with* $\rho = 1.0$ *and* 50 *timesteps. Convergence to the continuously observed limit should be at a first order rate as* $\Delta t \to 0$.

Grid Size	Timesteps	Option Value	CPU (sec)
Case 1: $r = .10$, $\sigma = .10$, $T = 0.25$ years, $E = \$100$, $S = \$100$			
50×50	50	1.8478	4.8
100×100	100	1.8492	55
200×200	200	1.8503	313
400×400	400	1.8509	2540
Case 2: $r = .10$, $\sigma = .50$, $T = 5.0$ years, $E = \$100$, $S = \$100$			
50×50	50	28.3573	5.5
100×100	100	28.3842	36
200×200	200	28.3952	280
400×400	400	28.4003	2278

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 modified Hull and White method, owing to an unnecessarily large number of nodes in the average direction.

Table 5 presents results for the FSG and modified Hull and White methods for American style fixed strike Asian options for Case 1 and Case2. As for the European case, the FSG method with nearest neighbor interpolation is divergent, whereas with linear interpolation the FSG method converges to a value which is slightly higher than the modified Hull and White method. In particular, for Case 1 with $\rho = 0.1$ the extrapolated FSG price is $\simeq 1.9605$ whereas the Hull and White extrapolated price with $\alpha = 4$ is $\simeq 1.9596$. Similarly, for Case 2 the extrapolated FSG price is $\simeq 34.3322$ while for the Hull and White scheme it is $\simeq 34.3065$.

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 modified 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 (1999). 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.

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

TABLE 5: *Convergence of the forward shooting grid and modified Hull and White methods for fixed strike American-Asian call options. The parameters* ρ *and* α *control the grid spacing in the average direction (smaller values indicate finer grids). The results are similar to those reported for the European cases in Tables 2 and 3. The forward shooting grid method is divergent if nearest neighbor interpolation is used, and it converges to a slightly higher value than the modified Hull and White method if linear interpolation is used.*

	Forward Shooting Grid			H & W	
Timesteps	ρ	Nearest Neighbor Interpolation	Linear Interpolation	α	
Case 1: $r = .10$, $\sigma = .10$, $T = 0.25$ years, $E = \$100$, $S = \$100$					
50	1.0	0.6423	1.9839	40	1.9460
100		0.4317	1.9827		1.9519
200		0.2950	1.9790		1.9555
400		0.2037	1.9750		1.9575
50	0.5	1.2363	1.9574	20	1.9397
100		0.8412	1.9634		1.9488
200		0.5498	1.9653		1.9540
400		0.3636	1.9653		1.9568
50	0.1	1.9422	1.9383	4	1.9374
100		1.9454	1.9487		1.9477
200		1.9200	1.9545		1.9534
400		1.7840	1.9575		1.9565
Case 2: $r = .10$, $\sigma = .50$, $T = 5.0$ years, $E = \$100$, $S = \$100$					
50	1.0	19.7377	34.8352	10	33.7946
100		13.5666	34.9001		34.0067
200		9.0702	34.8434		34.1450
400		5.7163	34.7529		34.2253
50	0.5	27.9919	33.9884	5	33.5438
100		22.9004	34.2788		33.8868
200		17.3613	34.4083		34.0866
400		12.2003	34.4454		34.1964
50	0.1	33.3721	33.4220	1	33.4484
100		33.7052	33.8425		33.8398
200		33.6270	34.0793		34.0636
400		32.7148	34.2057		34.1851

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. (1999). 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 somewhat more difficult.

8 An Additional Example: Mortgage-Backed Securities

A mortgage-backed security (MBS) is a fixed rate debt contract whose principal may be paid off prior to maturity. The debt in question is a pool of residential mortgages. The prepayment behavior of mortgage holders is usually analysed based on historical data. The results are encapsulated in a *prepayment function* which relates the amount of prepayment to various economic factors. We will assume that prepayment is a deterministic function of the spot interest rate. At any time, we assume that the principal outstanding is denoted by F. The prepayment function is then given by $\Pi = \Pi(F,r)$ where r is the spot risk free rate. We assume that prepayment can only occur at discrete time intervals (*prepayment times*).

If we assume a standard CIR square root model for the evolution of the spot risk free rate

$$dr = a(b-r)dt + \sigma_r \sqrt{r} dZ \ , \tag{8.1}$$

then the value of the MBS $V = V(r,F,t)$ is given by the solution to

$$V_t + \frac{\sigma_r^2 r}{2} V_{rr} + (a(b-r) - \lambda r)V_r - rV = 0 \ , \tag{8.2}$$

where λ is the market price of interest rate risk. At the expiry of the security, we have

$$V(r,F,t=T) = F \ . \tag{8.3}$$

As $r \to \infty$, then $V(r,F,t) = 0$. Depending on the values of a, b, and σ_r, a boundary condition may or may not need to be specified at $r = 0$ (see d'Halluin et al. (2001) for a detailed discussion).

We will assume that the mortgage can be prepaid at discrete times t_i, which coincide with the coupon payment dates. If t_i^-, t_i^+ are the times just before and after the prepayment dates, then no-arbitrage considerations yield the following jump condition

$$V(r,F,t_i^-) = V(r,F - \Pi(F,r), t_i^+) + \Pi(r,F) + K_i \ , \tag{8.4}$$

where K_i is the coupon at t_i. In the PDE context, we simply solve equation (8.2) for a set of discrete

TABLE 6: *Parameters used for the mortgage-backed security example.*

MBS parameters	
P_0	\$100
T	5 years
mortgage prepayments	every six months
coupon payments	\$25 every six months
α	.50
r_{min}	.02
r_{max}	.05
Interest rate model parameters	
σ	.03
a	.50
b	.05
λ	0.0

values of F_j, where $0 \leq F_j \leq P_0$, and where P_0 is the original principal. Along each line of constant F, we then discretize equation (8.2) in the r direction using a finite volume method with Crank Nicolson timestepping (see d'Halluin et al. (2001) for details). We then solve each one dimensional problem (for fixed F_j) backwards in time. At each prepayment/coupon payment date, we apply the jump condition (8.4). In general, interpolation is required to determine $V(r, F - \Pi(F,r), t_i^+)$. We will use either linear or quadratic interpolation (in the F direction). Note that, in contrast to the case of an Asian option in the limit of continuous observation, the interpolation error will be of the form

$$\text{MBS interpolation error} \quad < \quad \sum_{prepayment\ dates} O((\Delta F)^q) \ , \tag{8.5}$$

where $q = 2$ for linear interpolation and $q = 3$ for quadratic interpolation. In the MBS case, the sum in equation (8.5) is over the finite number of prepayment times. Consequently, the cumulative effect of the interpolation error is not as serious as for continuously observed Asian options.

In our numerical examples we use a prepayment function similar to that in Hull and White (1993),

$$\begin{aligned} \Pi' &= 0 \qquad ; \ r > r_{max} \\ &= P_0 \, \alpha \left(\frac{\frac{r_{max}}{r} - 1}{\frac{r_{max}}{r_{min}} - 1} \right) \ ; \ r_{min} < r < r_{max} \\ &= P_0 \, \alpha \qquad ; \ r < r_{min} \\ \Pi &= \min(\Pi', F) \end{aligned} \tag{8.6}$$

where r_{max}, r_{min}, and α are parameters, and P_0 is the original principal.

The parameters used in our numerical tests are shown in Table 6. We ran this problem on a sequence of grids. Each grid refinement doubled the number of nodes in the r and F directions and halved the timestep size. For the initial coarse grid, there were 10 timesteps per year. We used Crank-Nicolson timestepping

TABLE 7: *Results for the MBS example. Value of the security at $r = .05$, $t = 0$. Parameters are given in Table 6. Grid size is the number of nodes in the r and F directions. Change refers to the change in the solution from the previous (coarser) grid. Ratio is the ratio of successive changes.*

Grid Size	Timestep	Value	Change	Ratio
Linear Interpolation				
34x5	.1	99.644797		
67x9	.05	99.630542	.014255	
133x19	.025	99.626379	.004163	3.42
264x37	.00125	99.625338	.001041	4.00
529x63	.000625	99.625077	.000261	3.99
Quadratic Interpolation				
34x5	.1	99.643421		
67x9	.05	99.630441	.013040	
133x19	.025	99.626359	.004082	3.19
264x37	.00125	99.625333	.001026	3.98
529x63	.000625	99.625076	.000257	3.99

with the modification suggested in Rannacher (1984). The results for the value of the MBS at $r = .05, t = 0$ are shown in Table 7. There is clearly not much advantage here to using quadratic interpolation. This is in contrast to the results in Zvan et al. (1999) where quadratic interpolation noticeably improved the convergence of continuously observed Asian options. It is also in contrast to the results reported in Hull and White (1993), who report that quadratic interpolation produces significant improvement for coarse grids (in the F direction). This may be due to a variety of factors. Hull and White use a different term structure model, with a different prepayment function. As their reported values are lower than ours, the prepayment option is more valuable in their example. In any case, the relative performance on coarse grids of the linear and quadratic interpolation methods does not imply anything about the rate of convergence. Table 7 clearly shows that the convergence is quadratic as the grid/timestep is refined (the ratio of successive changes in the solution as the grid is refined is $\simeq 4$). This is to be expected for finite prepayment intervals.

9 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. We remark here that the analysis of this paper has been extended for the modified Hull and White method by Davidson (1999), who suggests a modification which results in convergence at a rate of $O(\Delta t)$ with complexity $O(N^{5/2})$ if linear interpolation is used.

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 Asian options, this method converges as $O\left[(\Delta t)^2\right]$, where Δt is the timestep size (Zvan et al. (1999)). 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 of the modified Hull and White or the PDE methods are effective. The FSG approach 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 solving a set of independent one-dimensional PDEs. 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 (Windcliff et al., 2000).

In more general terms, our results are indicative of when interpolation errors should be expected to cause difficulties. For cases where the interpolation must be performed relatively infrequently (such as our mortgage-backed security example above), there would not seem to be much to be gained from using higher order (e.g. quadratic) interpolation over linear interpolation, particularly in terms of the rate of convergence. Another context in which this would apply is pricing options with discrete dollar dividends: this can require interpolation at dividend payment dates, which obviously will be fairly infrequent. At the risk of stating the obvious, it is apparent that the effects of interpolation errors will be more serious when interpolation is applied more often. For the extreme case where we attempt to converge to a continous limit, we must ensure that the interpolation error at each timestep tends to zero faster than $O(\sqrt{\Delta t})$. In this situation, either the modified Hull and White algorithm or the PDE method will converge to the correct answer, and in some cases it may be advantageous to use quadratic (rather than linear) interpolation. As a final observation, note that our results here apply to cases where interpolation is applied to a variable which appears directly in the payoff function. In some other contexts, this is not the case. For instance, consider the GARCH option pricing algorithm proposed by Ritchken and Trevor (1999). For vanilla options, the interpolation used in this method applies to the conditional variance of the underlying asset, and so it does not directly affect the payoff function. However, for a contract such as a volatility or variance swap, then the payoff function is directly affected by the interpolation. We conjecture that in the GARCH framework there may be larger differences between various types of interpolation methods applied to this kind of pricing problem, as opposed to plain vanilla options. We leave the investigation of this as a topic for future research.

Appendices

A Proof of Proposition 1

Under the assumption that V has the appropriate number of bounded derivatives with respect to X, a Taylor series expansion of equation (4.3) gives

$$
\begin{aligned}
V_{j,k}^{n} &= e^{-r\Delta t}\left[p^*\left(V_{j+1,k}^{n+1}\right)^- + (1-p^*)\left(V_{j-1,k}^{n+1}\right)^-\right] + (E_{dis})_{j,k}^{n+1} \\
& p^* = \frac{1}{2}\left[1+\sqrt{\Delta t}\left(\frac{r}{\sigma}-\frac{\sigma}{2}\right)\right] \\
& (E_{dis})_{j,k}^{n+1} = O\left[(\Delta t)^2\right](C^*)_{j,k}^{n+1} \\
& (C^*)_{j,k}^{n} = \max(L_1, L_2, L_3) \\
& L_1 = \left|[(V_{tt})_{j,k}^{n}]^-\right| \\
& L_2 = \left|[(V_{XXX})_{j,k}^{n}]^-\right| \\
& L_3 = \left|[(V_{XXXX})_{j,k}^{n})]^-\right|
\end{aligned}
\tag{A.1}
$$

which is simply an explicit finite difference approximation to equation (4.3) with a constant grid spacing in the X direction with $\Delta X = \sigma\sqrt{\Delta t}$.

Assume that

$$
\frac{\left(V_{j+1,k}^{n+1}\right)^- - \left(V_{j-1,k}^{n+1}\right)^-}{\sigma\sqrt{\Delta t}} = O\left(\left[(V_X)_{j,k}^{n+1}\right]^-\right) \text{ as } \Delta t \to 0 \tag{A.2}
$$

and define

$$
\begin{aligned}
(C)_{j,k}^{n} &= \max((C^*)_{j,k}^{n}, L_4) \\
& L_4 = \left|[(V_X)_{j,k}^{n}]^-\right| .
\end{aligned}
\tag{A.3}
$$

Then expanding p (equation (3.8)) in a Taylor series and comparing with p^* in equation (A.1) gives

$$
\begin{aligned}
V_{j,k}^{n} &= e^{-r\Delta t}\left[p\left(V_{j+1,k}^{n+1}\right)^- + (1-p)\left(V_{j-1,k}^{n+1}\right)^-\right] + \text{truncation error}, \\
& \text{truncation error} = O\left[(\Delta t)^2\right](C)_{j,k}^{n+1} .
\end{aligned}
\tag{A.4}
$$

B Proof of Proposition 2

In this Appendix, we provide a proof of Proposition 2. We start by considering equations (4.11-4.12). Since equation (4.11) is linear, we can decompose the error $E_{j,k}^{n}$ into contributions due to the interpolation error

$\left(E^n_{j,k}\right)^I$ and the truncation error $\left(E^n_{j,k}\right)^T$:

$$\begin{aligned}
\left(E^n_{j,k}\right)^I &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}\left(E^{n+1}_{j+1,k^+_{floor}(j,k)}\right)^I + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left(E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)^I\right)\right. \\
&+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}\left(E^{n+1}_{j-1,k^-_{floor}(j,k)}\right)^I + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left(E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)^I\right)\right] \\
&+ e^{-r\Delta t}\left[p\left(\beta^q_{k^+_{floor}(j,k)}\right)^{n+1} + (1-p)\left(\beta^q_{k^-_{floor}(j,k)}\right)^{n+1}\right] \qquad \text{(B.1)} \\
\left(E^n_{j,k}\right)^T &= e^{-r\Delta t}\left[p\left(\alpha^{n+1}_{k^+_{floor}(j,k)}\left(E^{n+1}_{j+1,k^+_{floor}(j,k)}\right)^T + \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left(E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right)^T\right)\right. \\
&+ \left.(1-p)\left(\alpha^{n+1}_{k^-_{floor}(j,k)}\left(E^{n+1}_{j-1,k^-_{floor}(j,k)}\right)^T + \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left(E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right)^T\right)\right] \\
&+ \text{truncation error} \qquad \text{(B.2)} \\
E^n_{j,k} &= \left(E^n_{j,k}\right)^T + \left(E^n_{j,k}\right)^I \ . \qquad \text{(B.3)}
\end{aligned}$$

A typical asymptotic form for V as $S \to \infty$ is $V \to S^\gamma = e^{\gamma X}$ (where $\gamma \geq 0$), which means that $C^n_{j,k}$ in equation (A.3) is bounded by $O\left(e^{\gamma X}\right) = O\left(e^{\gamma j \sigma \sqrt{\Delta t}}\right)$, $j \leq n$. Consider the truncation error $\left(E^0_{0,0}\right)^T$ as the sum of errors incurred for $X < X_c$ and $X \geq X_c$, where X_c is constant and independent of Δt. For discretization errors encountered for $X < X_c$, the last term of equation (B.2) is $O\left[(\Delta t)^2\right]$. Since $0 \leq p \leq 1$ and $0 \leq \alpha \leq 1$, we have immediately (from equation (B.2)) that ($N =$ number of timesteps)

$$\begin{aligned}
\left(\left|E^0_{0,0}\right|^T\right)_{X<X_c} &\leq N \times (\text{truncation error}) \\
&= \frac{T}{\Delta t} O\left[(\Delta t)^2\right] \\
&= O(\Delta t) \ . \qquad \text{(B.4)}
\end{aligned}$$

Now, if the last term of (B.2) is $O\left[e^{\gamma j \sigma \sqrt{\Delta t}} (\Delta t)^2\right]$ for $j > j_c = ceil\left[\log X_c/(\sigma\sqrt{\Delta t})\right]$, then this bound depends only on j (and not on k), so that

$$\begin{aligned}
\left(E^n_{j,k}\right)^T_{X \geq X_c} &\leq e^{-r\Delta t}\left[p\left(E^{n+1}_{j+1,k}\right)^T + (1-p)\left(E^{n+1}_{j-1,k}\right)^T\right] \\
&+ O\left[(\Delta t)^2\right] e^{\gamma j \sigma \sqrt{\Delta t}} \ . \qquad \text{(B.5)}
\end{aligned}$$

We choose $X_c = X_c(T, \sigma)$ sufficiently large so that the discrete Green's function (corresponding to the effect at $S^0_{0,0}$ of a perturbation at $S^n_j = e^{j\sigma\sqrt{\Delta t}}$, $j_c \leq j \leq n$) for equation (B.5) is bounded by

$$\begin{aligned}
G(j,n) &= O\left(\frac{e^{-\beta j^2/n}}{\sqrt{n}}\right) \ ; \ j \geq j_c \\
&\beta = \beta(T, \sigma, X_c) > 0 \ . \qquad \text{(B.6)}
\end{aligned}$$

This can be deduced from the results in Cox et al. (1979), and from the continuous Green's function described in Wilmott (1998). Intuitively, equation (B.6) states that the effect of a perturbation in a payoff at a large value of S is given by the tail of the lognormal distribution. From equation (B.5) and (B.6) we can deduce that ($N = T/\Delta t$)

$$\begin{aligned}
\left(\left|E^0_{0,0}\right|^T\right)_{X \geq X_c} &\leq \sum_{n=1}^{N} O\left[(\Delta t)^2\right] \sum_{j=j_c}^{n} e^{\gamma j \sigma \sqrt{T/N}} \left(\frac{e^{-\beta j^2/n}}{\sqrt{n}}\right) \\
&= \sum_{n=1}^{N} O\left[(\Delta t)^2\right] \\
&= O(\Delta t) \; . \qquad \text{(B.7)}
\end{aligned}$$

Thus we have

$$\left|E^0_{0,0}\right|^T = \left(\left|E^0_{0,0}\right|^T\right)_{X<X_c} + \left(\left|E^0_{0,0}\right|^T\right)_{X \geq X_c} \quad = \quad O(\Delta t) \; , \qquad \text{(B.8)}$$

so that the cumulative effect of the truncation error is $O(\Delta t)$. Another difficulty occurs since the payoff functions are not particularly smooth (as in equation(2.6)). For example, if the payoff condition does not have the necessary smoothness in derivatives with respect to S, it can still be shown that the exact solution to equation (2.3) is smooth for any time $t < T$. However, there are some technical difficulties in ensuring that the local truncation error will be $O\left[(\Delta t)^2\right]$. Intuitively, this is because the payoff conditions introduce high grid frequency errors which are not damped by the difference scheme. Provided that the initial data is smoothed appropriately, these high frequency errors can be damped sufficiently so as not to reduce the order of convergence. Consequently, the estimate (B.8) is valid even for *rough* initial data. For detailed discussions of this, see Kreiss et al. (1970), Rannacher (1984), and Wahlbin (1980). The problem near $t \to T$ has also been noted (in the binomial model context) by Leisen and Reimer (1996) and Heston and Zhou (2000). We shall not dwell on this topic further in this work. We assume in the following that an appropriate method has been used near the payoff so that estimate (B.8) holds.

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

$$\begin{aligned}
\left|\frac{\partial V^n_j(A)}{\partial A}\right| &\leq M_1(T,\sigma) \;\; ; \;\; \forall n, j \quad n, j \to \infty \\
\left|\frac{\partial^2 V^n_j(A)}{\partial A^2}\right| &\leq M_2(T,\sigma) \;\; ; \;\; \forall n, j \quad n, j \to \infty \qquad \text{(B.9)}
\end{aligned}$$

where M_1 and M_2 are bounded independent of Δt, and where $A_{-k_m(n+1)} \leq A \leq A_{+k_m(n+1)}$ in each case. We can then write equation (B.1) as

$$
\begin{aligned}
\left|E_{j,k}^{n}\right|^{I} &\leq e^{-r\Delta t}\left[p\left(\alpha_{k_{floor}^{+}(j,k)}^{n+1}\left|E_{j+1,k_{floor}^{+}(j,k)}^{n+1}\right|^{I}+\left(1-\alpha_{k_{floor}^{+}(j,k)}^{n+1}\right)\left|E_{j+1,k_{ceil}^{+}(j,k)}^{n+1}\right|^{I}\right)\right. \\
&+ \left.(1-p)\left(\alpha_{k_{floor}^{-}(j,k)}^{n+1}\left|E_{j-1,k_{floor}^{-}(j,k)}^{n+1}\right|^{I}+\left(1-\alpha_{k_{floor}^{-}(j,k)}^{n+1}\right)\left|E_{j-1,k_{ceil}^{-}(j,k)}^{n+1}\right|^{I}\right)\right] \\
&+ e^{-r\Delta t}\left[M_q\left(A_{k_{ceil}^{+}(j,k)}^{n+1}\right)^{q}\left(1-e^{-\rho\Delta Z}\right)^{q}\right] . \qquad \text{(B.10)}
\end{aligned}
$$

Assume that interpolation errors introduced for $A_{k_{ceil(j,k)}^{+}}^{n+1} > A^*$ can be ignored (this will be justified in Appendix C). Since the interpolation coefficients α and the probability p are all in the range $[0,1]$, if

$$
\|E^n\|^I = \max_{j,k}\left|E_{j,k}^{n}\right|^{I} \qquad \text{(B.11)}
$$

then from equations (B.10-B.11)

$$
\|E^n\|^I \leq e^{-r\Delta t}\left(\left\|E^{n+1}\right\|^{I}+\left[M_q\left(A^*\right)^q)\left(1-e^{-\rho\Delta Z}\right)^{q}\right]\right) . \qquad \text{(B.12)}
$$

Equation (B.12) states that the interpolation error generated during timestep $n+1 \rightarrow n$ 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. In the worst case, we obtain the bound after N steps from equation (B.12) (with $N\Delta t = T$) of

$$
\begin{aligned}
\left\|E^0\right\|^I &\leq NM_q(A^*)^q\left(1-e^{-\rho\Delta Z}\right)^q \\
&\simeq \frac{TM_q(A^*)^q\left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \\
&= \frac{TC_q\left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \qquad \text{(B.13)}
\end{aligned}
$$

where $C_q = M_q(A^*)^q$. Note that this tends to zero only if $q \geq 3$. In particular, for $q = 1$ (nearest lattice point interpolation), equation (B.13) indicates that the scheme may be divergent. This is illustrated in some example computations Section 7.

In Appendix D, we discuss the form of V_{AA} as $t \rightarrow T$, and indicate how the error analysis would have to be modified in order to take this into account. However, as mentioned previously, a complete analysis is beyond the scope of this work.

In the following, we assume that the interpolation error is given by equation (B.13), (as $\Delta t \rightarrow 0$ with $T = N\Delta t$)

$$
\left\|E^0\right\|^I \leq NC_q\left(1-e^{-\rho\Delta Z}\right)^q \simeq \frac{TC_q\left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \qquad \text{(B.14)}
$$

where C_q is a constant independent of Δt.

Equation (B.14) suggests that the error can become unbounded if (for example) $q = 1$. However, this is an overestimate of the error. Note that equation (3.7) has all positive coefficients which sum to $e^{-r\Delta t}$. Consequently

$$\min\left(U^N_{j,k}\right) = 0 \Rightarrow U^0_{0,0} \geq 0 \; . \tag{B.15}$$

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

$$\max\left(U^N_{j,k}\right) = E \Rightarrow U^0_{0,0} \leq E \; . \tag{B.16}$$

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

$$\begin{aligned} A^{n+1}_{k^+(j,k)} &= \max\left(S^{n+1}_{j+1}, A^k_n\right) \\ A^{n+1}_{k^-(j,k)} &= \max\left(S^{n+1}_{j-1}, A^k_n\right) \\ & k^+(j,k) = \max(k, j+1) \\ & k^-(j,k) = \max(k, j-1) \; . \end{aligned} \tag{B.17}$$

Therefore $U^n_{j,k} \leq B^n_{j,k}$, where $B^n_{j,k}$ is given by

$$\left(B^n_{j,k}\right) = e^{-r\Delta t}\left[p\left(B^{n+1}_{j+1,\max(k,j+1)}\right) + (1-p)\left(B^{n+1}_{j-1,\max(k,j-1)}\right)\right] \tag{B.18}$$

for $n = N^* - 1, \ldots, 0$ with $B^N_{j,k} = \max\left(0, A^N_k - E\right)$. This is simply the binomial expression for a fixed strike lookback call. Consequently we have

$$0 \leq U^0_{0,0} \leq B(E) \tag{B.19}$$

where $B(E) = E$ for a fixed strike put, and is the value of a fixed strike lookback call 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 (B.14) is more precisely stated as

$$\left\|E^0\right\|^I \leq \min\left[B(E), NC_q\left(1 - e^{-\rho\Delta Z}\right)^q\right] \tag{B.20}$$

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

C Effect of Interpolation Errors as $A \to \infty$.

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 (B.13) is correct, if we allow $A^* \to \infty$ as $n \to \infty$ in equation (B.12).

We will also assume that

$$\begin{aligned} \left| \frac{\partial V_j^n(A)}{\partial A} \right| &\leq M_1 \\ \left| \frac{\partial^2 V_j^n(A)}{\partial A^2} \right| &\leq M_2 \end{aligned} \tag{C.1}$$

for any n, j, where M_1 and M_2 are constants independent of Δt. Note that in some cases, the formal derivative may not exist at certain points, in which case we take $|\frac{\partial V_{j+1}^{n+1}(A)}{\partial A}|$ and $|\frac{\partial^2 V_{j+1}^{n+1}(A)}{\partial A^2}|$ to be defined as the maximum of left and right limits near these points.

Taking into account equations (C.1) instead of equations (B.9), then our starting point is a modified form of equation (B.10)

$$\begin{aligned} \left|E_{j,k}^n\right|^I &\leq e^{-r\Delta t} \Bigg[p \left(\alpha^{n+1}_{k^+_{floor}(j,k)} \left|E^{n+1}_{j+1,k^+_{floor}(j,k)}\right|^I + \left(1 - \alpha^{n+1}_{k^+_{floor}(j,k)}\right) \left|E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right|^I \right) \\ &+ (1-p) \left(\alpha^{n+1}_{k^-_{floor}(j,k)} \left|E^{n+1}_{j-1,k^-_{floor}(j,k)}\right|^I + \left(1 - \alpha^{n+1}_{k^-_{floor}(j,k)}\right) \left|E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right|^I \right) \Bigg] \\ &+ e^{-r\Delta t} \left[M_q \left(A^{n+1}_{k^+_{ceil}(j,k)} \right)^q \left(1 - e^{-\rho \Delta Z}\right)^q \right]. \end{aligned} \tag{C.2}$$

We can bound the interpolation error term in equation (C.2) by noting that (from equations (3.5-3.6))

$$\begin{aligned} A^{n+1}_{k^+_{ceil}(j,k)} &\leq \max\left(S^{n+1}_{j+1}, A^n_k\right) \\ A^{n+1}_{k^-_{ceil}(j,k)} &\leq \max\left(S^{n+1}_{j-1}, A^n_k\right) . \end{aligned} \tag{C.3}$$

Equations (C.2) and (C.3) then give

$$\begin{aligned} \left|E_{j,k}^n\right|^I &\leq e^{-r\Delta t} \Bigg[p \left(\alpha^{n+1}_{k^+_{floor}(j,k)} \left|E^{n+1}_{j+1,k^+_{floor}(j,k)}\right|^I + \left(1 - \alpha^{n+1}_{k^+_{floor}(j,k)}\right) \left|E^{n+1}_{j+1,k^+_{ceil}(j,k)}\right|^I \right) \\ &+ (1-p) \left(\alpha^{n+1}_{k^-_{floor}(j,k)} \left|E^{n+1}_{j-1,k^-_{floor}(j,k)}\right|^I + \left(1 - \alpha^{n+1}_{k^-_{floor}(j,k)}\right) \left|E^{n+1}_{j-1,k^-_{ceil}(j,k)}\right|^I \right) \Bigg] \\ &+ e^{-r\Delta t} \left[M_q \left(\max\left(S^{n+1}_{j+1}, A^n_k\right) \right)^q \left(1 - e^{-\rho \Delta Z}\right)^q \right] . \end{aligned} \tag{C.4}$$

Since equation (C.4) is linear, we can consider $\left|E^0_{0,0}\right|^I$ to be

$$\left|E^0_{0,0}\right| \leq \sum_{N^*=0}^{N-2} \left|E^0_{0,0}(N^*)\right|^I \tag{C.5}$$

where $\left(E^n_{j,k}(N^*)\right)^I$ 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 $\left(E^n_{j,k}(N^*)\right)^I = 0$ for $n > N^*$. Note that the upper bound in the sum in equation (C.5) is $N-2$ since there is no interpolation error at $n = N$.

Consequently, we have (from equation (C.4))

$$\begin{aligned} \left|E^{N^*}_{j,k}(N^*)\right|^I &\leq e^{-r\Delta t}\left[M_q\left(\max\left(S^{N^*+1}_{j+1}, A^{N^*}_k\right)\right)^q \left(1-e^{-\rho\Delta Z}\right)^q\right] \\ &\leq \left(1-e^{-\rho\Delta Z}\right)^q M_q\left[\left(S^{N^*+1}_{j+1}\right)^q + \left(A^{N^*}_k\right)\right)^q\right] . \end{aligned} \tag{C.6}$$

We can rewrite equation (C.6) as

$$\left|E^{N^*}_{j,k}(N^*)\right|^I \leq M_q\left(1-e^{-\rho\Delta Z}\right)^q\left[\left(E^{N^*}_{j,k}(N^*)\right)_A + \left(E^{N^*}_{j,k}(N^*)\right)_S\right] \tag{C.7}$$

with

$$\begin{aligned} \left(E^{N^*}_{j,k}(N^*)\right)_A &= \left(A^{N^*}_k\right)^q \\ \left(E^{N^*}_{j,k}(N^*)\right)_S &= \left(S^{N^*+1}_{j+1}\right)^q . \end{aligned} \tag{C.8}$$

For $n < N^*$ we can define

$$\left|E^n_{j,k}(N^*)\right|^I = M_q\left(1-e^{-\rho\Delta Z}\right)^q\left[\left(E^n_{j,k}(N^*)\right)_A + \left(E^n_{j,k}(N^*)\right)_S\right] \tag{C.9}$$

where $\left(E^n_{j,k}(N^*)\right)_\kappa$, $\kappa = S, A$, satisfy the recursions

$$\begin{aligned} \left(E^n_{j,k}(N^*)\right)_\kappa \leq\ & e^{-r\Delta t}\Bigg\{p\Bigg[\alpha^{n+1}_{k^+_{floor}(j,k)}\left(E^{n+1}_{j+1,k^+_{floor}(j,k)}(N^*)\right)_\kappa + \\ & \qquad \left(1-\alpha^{n+1}_{k^+_{floor}(j,k)}\right)\left(E^{n+1}_{j+1,k^+_{ceil}(j,k)}(N^*)\right)_\kappa\Bigg] \\ & + (1-p)\Bigg[\alpha^{n+1}_{k^-_{floor}(j,k)}\left(E^{n+1}_{j-1,k^-_{floor}(j,k)}(N^*)\right)_\kappa + \\ & \qquad \left(1-\alpha^{n+1}_{k^-_{floor}(j,k)}\right)\left(E^{n+1}_{j-1,k^-_{ceil}(j,k)}(N^*)\right)_\kappa\Bigg]\Bigg\} . \end{aligned} \tag{C.10}$$

Since $\left(E^n_{j,k}(N^*)\right)_S$ is independent of k, it follows from equation (C.10) that

$$\left(E^n_{j,k}(N^*)\right)_S = e^{-r\Delta t}\left[p\left(E^{n+1}_{j+1,k}(N^*)\right)_S + (1-p)\left(E^{n+1}_{j-1,k}(N^*)\right)_S\right] \tag{C.11}$$

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

$$\left(Se^{\sigma\sqrt{\Delta t}}\right)^q \simeq S^q \quad \text{as } \Delta t \to 0 \ . \tag{C.12}$$

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 (in equation (C.10))

$$\begin{aligned}
A^{n+1}_{k^+(j,k)} = A^{n+1}_{k^+_{ceil}(j,k)} = A^{n+1}_{k^+_{floor}(j,k)} &= \max(S^{n+1}_{j+1}, A^k_n) \\
A^{n+1}_{k^-(j,k)} = A^{n+1}_{k^-_{ceil}(j,k)} = A^{n+1}_{k^-_{floor}(j,k)} &= \max(S^{n+1}_{j-1}, A^k_n) \\
&k^+(j,k) = \max(k, j+1) \\
&k^-(j,k) = \max(k, j-1) \ .
\end{aligned} \tag{C.13}$$

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 (C.13) in equation (C.10) we obtain

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

for $n = N^* - 1, \ldots, 0$. The right hand side of inequality (C) 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 (C.13)). Let $\left(E^0_{0,0}(N^*)\right)_A$ be bounded by some constant $C(N^*)_A$. Let

$$\begin{aligned}
\max_{N^*}\left(|C(N^*)_A| + |C(N^*)_S|\right) &\leq K \\
N^* &< N - 1 \ ; \ N \to \infty \\
N &= T/\Delta t
\end{aligned} \tag{C.14}$$

where K is independent of Δt. Then, from equations (C.5) and and (C.9) we have

$$\left|E^0_{0,0}\right|^I \leq NKD_q\left(1 - e^{-\rho\Delta Z}\right)^q \tag{C.15}$$

or

$$|E^0_{0,0}|^I = O(N(1 - e^{-\rho\Delta Z})^q) \ . \tag{C.16}$$

In equation (C.16) we replace $N = T/\Delta t$ to obtain

$$\left\|E^0\right\|^I \quad \leq \quad NC_q\left(1-e^{-\rho\Delta Z}\right)^q \simeq \frac{TC_q\left(\rho\sigma\sqrt{\Delta t}\right)^q}{\Delta t} \tag{C.17}$$

where C_q is a constant independent of Δt. This estimate has the same form as that in equation (B.13), which was obtained by a more intuitive argument.

D Effect of Non-Smooth Payoff

In this appendix, we consider the effect of a non-smooth payoff on the propagation of the interpolation error. The numerical evidence in Tables 3 and 4 indicate that this does not cause any significant problems. This appendix provides a heuristic argument as to why this is the case.

For brevity, we shall consider only a single representative example: the modified Hull and White method with linear interpolation. We do not consider the FSG method as this was shown to have a non-convergent error bound in Section 4 under the more generous assumption of a smooth payoff function. Arguments similar to that used for analysis of the FSG method result in the following analogue of equation (B.12) for the propagation of the interpolation error:

$$\|E^n\|^I \leq e^{-r\Delta t}\left(\left\|E^{n+1}\right\|^I + \left[M_2\left(A^*\right)^2\right)\left(1-e^{-C\Delta t}\right)^2\right]\right) \tag{D.1}$$

where

$$\left|\frac{\partial^2 V_j^n(A)}{\partial A^2}\right| \leq M_2 \; . \tag{D.2}$$

Now, consider a fixed strike Asian call option, with payoff

$$\text{payoff} = \max(A-E,0) \; . \tag{D.3}$$

Note that no interpolation error is incurred in making the transition $t_N^+ \to t_N^-$, since the exact payoff is available (equation (D.3)). Now, during the time interval $t \in [t_N^-, t_{N-1}^+]$ the exact solution to the discretely observed Asian option problem satisfies

$$V_t + \frac{\sigma^2 S^2}{2} V_{SS} + rSV_S - rV = 0 \; . \tag{D.4}$$

Note that the jump conditions (2.4) imply that

$$V(S,A,t_N^-) = V\left(S, A + \frac{S-A}{N+1}, t_N^+\right) \; . \tag{D.5}$$

Let $(V^-)_{SS}$ be the diffusion term in equation (D.4) evaluated before the discrete observation at $t = t_N$. If we write this in terms of V^+ (the value after the observation at $t = t_N$), we obtain

$$\begin{aligned} (V^-)_{SS} &= \left(V^+(S, A + \frac{S-A}{N+1}, t_N^+) \right)_{SS} \\ &= V_{SS}^+ + \frac{2}{N+1} V_{AS}^+ + \frac{1}{(N+1)^2} V_{AA}^+ \ . \end{aligned} \tag{D.6}$$

This means that there is a diffusion in the A direction during the step $t_N^+ \rightarrow t_{N-1}^+$ with effective volatility (noting that $N = O((\Delta t)^{-1})$)

$$\begin{aligned} \sigma_{eff} &= \frac{\sigma}{N+1} \\ &\simeq \sigma \Delta t \ . \end{aligned} \tag{D.7}$$

Now, from equation (D.3), we have that at $t = t_N^+$ the value of the option is independent of S. Therefore, during the first interval Δt, we can regard the behavior of V_{AA} as given approximately by the value of gamma for a vanilla option which is a function of A only, with payoff (D.3). From the usual expression for gamma (Wilmott, 1998), we then have

$$|V_{AA}| \leq \frac{D}{\sigma_{eff}\sqrt{T-t}} \tag{D.8}$$

where D is a constant independent of Δt. Note that this unbounded behavior occurs only at $A = E$. At the end of the first timestep $(T - t) = \Delta t$, we then have (from equations (D.7-D.8))

$$|V_{AA}| \leq \frac{D}{\sigma(\Delta t)^{3/2}} \ . \tag{D.9}$$

Continuing this heuristic argument for subsequent discrete observations, we obtain the bound

$$|V_{AA}| \leq \frac{D}{\sigma(T-t)^{3/2}} \ . \tag{D.10}$$

We emphasize at this point that equation (D.10) is probably exceedingly pessimistic.

Before proceeding further, we will simplify equation (D.1) (assuming $\Delta t \rightarrow 0$) to

$$\|E^n\|^I \leq \left\|E^{n+1}\right\|^I + M_2(A^*)^2 C^2 (\Delta t)^2 \ . \tag{D.11}$$

Now, replacing M_2 in equation (D.11) by the expression (D.10) gives

$$\begin{aligned} \|E^n\|^I &\leq \left\|E^{n+1}\right\|^I + (A^*)^2 C^2 (\Delta t)^2 \frac{D}{\sigma((N-(n+1))\Delta t)^{3/2}} \\ &= \left\|E^{n+1}\right\|^I + \frac{D_2\sqrt{\Delta t}}{((N-(n+1)))^{3/2}} \end{aligned} \tag{D.12}$$

where D_2 is a constant independent of Δt. However, this is overly pessimistic: since the error using linear interpolation can be no worse than the error for nearest neighbor interpolation, equation (D.12) must be modified to give

$$\|E^n\|^I \leq \left\|E^{n+1}\right\|^I + \min\left(\frac{D_2\sqrt{\Delta t}}{((N-(n+1)))^{3/2}}, M_1\Delta t\right) \ . \tag{D.13}$$

It then follows from equation (D.13) that

$$\begin{aligned} \left\|E^0\right\|^I &\leq \sum_{n=0}^{n=N-1} \min\left(\frac{D_2\sqrt{\Delta t}}{((N-n))^{3/2}}, M_1\Delta t\right) \\ &= \sum_{i=1}^{i=N} \min\left(\frac{D_2\sqrt{\Delta t}}{(i)^{3/2}}, M_1\Delta t\right) \ . \end{aligned} \tag{D.14}$$

We can break up the sum on the right hand side of equation (D.14) from $i = 0,\ldots,N^p$ and from $i = N^p+1,\ldots,N-1$, where N^p is selected so that the size of the two error terms (nearest neighbor and linear interpolation) are of the same order. This yields

$$\begin{aligned} \sum_{i=1}^{n=N} \min\left(\frac{D_2\sqrt{\Delta t}}{(i)^{3/2}}, M_1\Delta t\right) &= \sum_{i=1}^{i=N^p} M_1\Delta t + \sum_{i=N^p+1}^{i=N} \sqrt{\Delta t}\frac{D_2}{i^{3/2}} \\ &\simeq O\left((\Delta t)^{1-p}\right) + O\left((\Delta t)^{p/2+1/2}\right) \ , \end{aligned} \tag{D.15}$$

where we have used the fact that $\Delta t = O(1/N)$. Now, the two terms on the right hand side of equation (D.15) will be of the same order if $p = 1/3$, which (from equations (D.14-D.15)) means that

$$\left\|E^0\right\|^I \leq O(\Delta t)^{2/3} \ . \tag{D.16}$$

Note that the analysis in the main body of the paper, which assumes that V_{AA} is always bounded, results in the expression (5.3). In the case of linear interpolation this gives

$$\left\|E^0\right\|^I = O(\Delta t) \ . \tag{D.17}$$

The numerical experiments indicate that equation (D.17) appears to be the correct asymptotic form for the error for the modified Hull and White method with linear interpolation (see Table 3). This is not surprising, since the argument used to derive equation (D.16) assumes worst case scenarios, and hence is not very sharp. For example, V_{AA} is bounded except at the single point $A = E$. We have also ignored any additional smoothing due to diffusion in the S direction. Nevertheless, it is interesting to note that the assumption of unbounded behavior for V_{AA}, as in equation (D.10), as $t \to T$, still results in convergence for the modified Hull and White method using linear interpolation, albeit at a reduced rate. It is an interesting topic of further research to obtain a sharp bound for the interpolation error, taking into account the non-smooth payoff (D.3). We expect this sharp bound will be closer to equation (D.17) than equation (D.16).

References

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

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

Boyle, P. 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 (1999). A refined binomial lattice for pricing American Asian options. *Review of Derivatives Research* 3, 85–105.

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

Davidson, M. (1999). A modification of the FSG method for pricing average options. Working paper, Moscow State University, Faculty of Computational Mathematics.

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

d'Halluin, Y., P. A. Forsyth, K. R. Vetzal, and G. Labahn (2001). A numerical PDE approach for pricing callable bonds. *Applied Mathematical Finance* 8, 49–77.

Forsyth, P. A., K. R. Vetzal, and R. Zvan (1999). A finite element approach to the pricing of discrete lookbacks with stochastic volatility. *Applied Mathematical Finance* 6, 87–106.

Heston, S. and G. Zhou (2000). On the rate of convergence of discrete-time contingent claims. *Mathematical Finance* 10, 53–75.

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

Kreiss, H. O., V. Thomée, and O. Widlund (1970). Smoothing of initial data and rates of convergence for parabolic difference equations. *Communications on Pure and Applied Mathematics* 23, 242–259.

Leisen, D. and M. Reimer (1996). Binomial models for option valuation: Examining and improving convergence. *Applied Mathematical Finance* 3, 319–356.

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 (1999). Interest rate option pricing with volatility humps. *Review of Derivatives Research* 3, 237–262.

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 (1999). Pricing options under generalized GARCH and stochastic volatility processes. *Journal of Finance* 54, 377–402.

Wahlbin, L. B. (1980). A remark on parabolic smoothing and the finite element method. *SIAM Journal on Numerical Analysis* 17, 33–38.

Wilmott, P. (1998). *Derivatives*. Wiley, West Sussex, England.

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

Windcliff, H., K. R. Vetzal, P. A. Forsyth, A. Verma, and T. F. Coleman (2000). An object-oriented framework for valuing shout options on high-performance computer architectures. *Journal of Economic Dynamics and Control*, forthcoming.

Zvan, R., P. A. Forsyth, and K. R. Vetzal (1998). Robust numerical methods for PDE models of Asian options. *Journal of Computational Finance* 1 (Winter), 39–78.

Zvan, R., P. A. Forsyth, and K. R. Vetzal (1999). Discrete Asian barrier options. *Journal of Computational Finance* 3 (Fall), 41–67.