A SEMI-LAGRANGIAN APPROACH FOR AMERICAN ASIAN OPTIONS UNDER JUMP DIFFUSION *

Y. D'HALLUIN[†], P.A. FORSYTH[‡], AND G. LABAHN[§]

Abstract. A semi-Lagrangian method is presented to price continuously observed £xed strike Asian options. At each timestep a set of one dimensional partial integro differential equations (PIDEs) is solved and the solution of each PIDE is updated using semi-Lagrangian timestepping. Crank-Nicolson and second order backward differencing timestepping schemes are studied. Monotonicity and stability results are derived. With low volatility values, it is observed that the non-smoothness at the strike in the payoff affects the convergence rate; sub-quadratic convergence rate is observed.

Key words. Continuously observed Asian option, semi-Lagrangian, American option, jump diffusion, implicit discretization.

AMS subject classi£cations. 65M12, 65M60, 91B28

1. Introduction. An Asian option gives the holder a payoff that depends on the average price of the underlying asset over a speci£ed period of time (12). Asian-style derivatives have a wide variety of applications in equity, energy, interest rate, and insurance markets. These options tend to be less expensive than vanilla options (since the volatility of the average price is less than the price itself) and average prices over a period of time are far more dif£cult to manipulate for illiquid commodities. For example, airline companies are certainly more interested in buying oil based on its average price instead of its spot price. For a historical review of Asian options we refer the reader to (11).

The price of an Asian option at any time is a function of both the underlying asset at that time and the average of the underlying prices up to that time. As such these options are considered path-dependent. In practice, Asian option contracts typically specify that the average is monitored discretely. A typical situation would be to base the average on the daily closing price. If daily averaging is used, then for typical market parameters, for options with expiry times more than three months, we can consider these options as being continuously monitored, for all practical purposes. In addition, if we need to price long term Asian options (for example greater than one year), then using timesteps of one day (which would be required in a discrete observation model (52)) would clearly be computationally wasteful. Consequently, in this paper we focus on continuously observed Asian options. For details on numerical methods for discretely observed Asian options, we refer the reader to (18; 17; 52; 25)

In this paper we are interested in the pricing of Asian options with various general payoff conditions (for example, American style, asset dependent barriers or volatility). When the underlying asset follows a standard brownian motion stochastic process then the price of an Asian option can be determined by solving a two dimensional PDE. In some special cases (e.g. constant volatility, no barrier features, and a ¤oating strike contract) this problem can be reduced to a one-dimensional PDE (4). In addition, for either ¤oating or £xed strike, but

---

*This work was supported by the Natural Sciences and Engineering Research Council of Canada, and by Bell University Laboratories at the University of Waterloo.

[†]School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 `ydhallui@elora.uwaterloo.ca`,

[‡]School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 `e-mail: paforsyt@elora.uwaterloo.ca`,

[§]School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 `glabahn@scg.uwaterloo.ca`

not American style or asset dependent features (e.g. volatility a function of asset price), a one dimensional PDE can also be derived (46). However, in the general case which we are interested in, the two dimensional PDE cannot be reduced to one dimension.

The two dimensional PDE that appears in our general case of the Asian option pricing problem has no diffusion in one of the coordinate directions and as such is well known to be dif£cult to solve numerically. In (51; 50), a ¤ux limiter was used to retain accuracy while preventing oscillations. In (32), the £rst order hyperbolic term was discretized using a £rst order upwind type method, resulting in at most £rst order accuracy. A related approach based on a combination of a WENO discretization and grid stretching was used for Asian options in (37). In (38), a semi-Lagrangian method was used to discretize the hyperbolic term in the average direction. Semi-Lagrangian schemes were £rst introduced by (21) and (40) for atmospheric and weather numerical predictions. These are time marching schemes that integrate convection-diffusion equations by tracing backward in time the position of the ¤ow. These schemes are used to reduce numerical problems raised by convection dominated equations. In principle, provided an appropriate time discretization is used, and a high enough order of interpolation is used to recover values at the feet of the characteristic curves (23; 1; 10), then this method is capable of greater than £rst order convergence as the grid and timestep size is reduced.

In this paper, we will explore the use of a semi-Lagrangian method for pricing Asian options. The semi-Lagrangian method has many advantages in our case. For example we are able to easily solve the pricing problem in the more general context of a jump diffusion stochastic processes, and contracts with American early exercise features. Jump diffusion models were introduced in the option valuation context in (33). They are important because of the increasing empirical evidence that the usual assumption of geometric Brownian motion should be augmented by discontinuous jump processes (22).

We make no assumptions regarding the contract or the form of the deterministic volatility function. We remark that although we focus exclusively on Asian options in this paper, similar PDEs (no diffusion in one of the space-like directions) occur in certain interest rate models (44) and employee reload options with various constraints (9). Hence the methods developed here will be applicable to these cases as well.

The main results in this paper are

- We demonstrate that a semi-Lagrangian method can be used to price continuously observed American Asian options under jump diffusion processes. The implementation suggested here reduces this problem to solving a decoupled set of one dimensional nonlinear discrete partial integro differential equations (PIDEs) at each timestep. This makes implementation of this method very straightforward in a software library which is capable of pricing discretely observed path dependent options (52). In the general case, a numerical method must be used to integrate the characteristic equation in a semi-Lagrangian method. However, in the case of Asian options, the characteristic equation can be solved analytically, which makes implementation very straightforward.
- We show that in the fully implicit case, the semi-Lagrangian method is algebraically identical to a standard numerical method for pricing discretely observed Asian options, if the observation interval is equal to the discrete timestep. Since lattice methods (47) can be regarded as explicit £nite difference methods it follows that the usual binomial forest method for Asian options (28) can also be regarded as an explicit semi-Lagrangian method.
- Since the discretized problem at each timestep reduces to a set of decoupled one dimensional PIDEs, we can make use of the techniques developed in (24; 19; 20) to

prove certain properties of the discrete scheme, including convergence of the iterative method used to solve the implicit discrete equations. In the fully implicit case, it is straightforward to prove l_∞ stability and monotonicity, which are important properties of discrete schemes for option pricing (5; 41; 15; 13).

In addition to the above contributions, we also include experimental computations which indicate that, even if second order timestepping methods are used, observed convergence as the mesh and timestep is re£ned occurs at a sub-second order rate. The problem can be traced to the non-smoothness of the payoff function.

2. Mathematical Model. In this section we give the mathematical model for options with jump diffusion processes. We do this for both European and American options. If the underlying asset follows a jump diffusion process, the usual portfolio hedging arguments cannot be used. As such, we will also present a brief discussion of various strategies for hedging jump risk.

Let S represent the underlying stock price. We restrict attention in the following to £nite activity processes, that is, we assume that the probability density of a jump occurring with a given jump size, in some interval $[t, t+dt]$ is always £nite (14). The potential stock paths followed by the stock can be modeled by a stochastic differential equation given by

$$\frac{dS}{S} = (\xi - \kappa\lambda)dt + \sigma dZ + (\eta - 1)dq\ , \tag{2.1}$$

where

ξ	is the drift rate,
dq	is the independent Poisson process $= \begin{cases} 0 & \text{with probability } 1-\lambda dt \\ 1 & \text{with probability } \lambda dt, \end{cases}$
λ	is the mean arrival time of the Poisson process,
$\eta - 1$	is an impulse function producing a jump from S to $S\eta$,
σ	is the volatility,
dZ	is an increment of the standard Gauss-Wiener process,
κ	is $E[\eta - 1]$, where $E[\cdot]$ is the expectation operator.

When the average is monitored continuously (7; 47; 25), the arithmetic average A is de£ned as

$$A = \frac{\int_0^t S(u)du}{t} \quad \text{with} \quad dA = \frac{(S-A)}{t}dt. \tag{2.2}$$

2.1. The PIDE for Asian Options. Using standard arguments (7), the value of an option depending on S (2.1) and A (2.2) and assuming no jumps (that is, $\lambda = 0$) is given by

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \frac{(S-A)}{t}V_A + rSV_S - rV = 0, \tag{2.3}$$

where r is the continuously compounded risk free interest rate. Since we are solving backward in time from the expiration time $t = T$ to the present time $t = 0$, equation (2.3) becomes

$$V_\tau = \frac{1}{2}\sigma^2 S^2 V_{SS} + \frac{(S-A)}{T-\tau}V_A + rSV_S - rV, \tag{2.4}$$

where $\tau = T - t$. It is important to note that equation (2.4) has no diffusion term in the A direction and this is the source of many numerical dif£culties (51).

Extending equation (2.4) to the case of jumps gives

$$V_\tau = \frac{(S-A)}{T-\tau}V_A + \frac{1}{2}\sigma^2 S^2 V_{SS} + (r-\lambda\kappa)SV_S - rV + \left(\lambda \int_0^\infty V(S\eta)g(\eta)d\eta - \lambda V\right), \quad (2.5)$$

where

$$g(\eta) \quad \text{is the probability density function of the jump amplitude } \eta,$$
$$\text{thus for all } \eta : g(\eta) \geq 0 \text{ and } \int_0^\infty g(\eta)d\eta = 1. \quad (2.6)$$

As a speci£c example, if

$$g(\eta) = \frac{e^{\left(-\frac{(\log(\eta)-\mu)^2}{2\gamma^2}\right)}}{\sqrt{2\pi}\gamma\eta}, \quad (2.7)$$

the probability density function suggested by (33; 45), then its expectation is given by $E[\eta] = \exp(\mu + \gamma^2/2)$. This means that the expected relative change in the stock price is given by $\kappa = E[\eta - 1] = \exp(\mu + \gamma^2/2) - 1$.

For brevity, the details of the derivation of equation (2.5) have been omitted (see (3; 33; 47)). In general, it is not possible to construct a hedging portfolio which eliminates jump risk. However, by adding options to the hedging portfolio, a hedging strategy can be constructed which minimizes jump risk (3). If equation (2.5) is calibrated to market prices, then the parameters so obtained should be regarded as risk-adjusted (31), not historical.

If we de£ne

$$\mathcal{H}V \equiv V_\tau - \left[\frac{(S-A)}{T-\tau}V_A + \frac{\sigma^2 S^2}{2}V_{SS} + (r-\lambda\kappa)SV_S - (r+\lambda)V + \lambda \int_0^\infty V(S\eta)g(\eta)d\eta\right] \quad (2.8)$$

and if $V^*(S,A)$ is the payoff, then the American option pricing problem can be stated as (39)

$$\min(\mathcal{H}V; V - V^*) = 0\ . \quad (2.9)$$

2.2. Boundary Conditions for our PIDE. In order to completely specify our problem we still need to give boundary conditions for our American Asian option pricing PIDE. For the terminal boundary conditions, a number of common payoffs for pricing different types of Asian securities can be used. Typical examples include

- £xed strike call: $V(S,A,\tau = 0) = \max(A - K, 0.0)$,
- £xed strike put: $V(S,A,\tau = 0) = \max(K - A, 0.0)$.

Although the original domain is $0 \leq S < \infty$, $0 \leq A < \infty$, for computational purposes we make the typical assumption of using a £nite computational domain $[0, S_{max}] \times [0, A_{max}]$, with $A_{max} = S_{max}$. A discussion of the errors involved in using a £nite computational domain (the localization error) for the speci£c case of option pricing PIDEs is discussed in (16). However, as pointed out in (16; 13), unlike the usual option pricing PDE case, it is necessary to approximate the solution outside the computational domain $S > S_{max}$ in order that the integral in equation (2.5) is well de£ned. The results in (16) suggest that the error in misspeci£cation of the boundary condition at $S = S_{max}$ (for values of S of practical interest) can be made arbitrarily small by making S_{max} (and hence $A_{max} = S_{max}$) suf£ciently large. This will be veri£ed in some numerical tests in Section 8.

In the case of Asian options without jumps (equation (2.4)), Meyer (34) provides an extensive discussion of the boundary conditions. Using the results in (36), Meyer concludes that no boundary conditions are required at $S = 0, A = 0, A = A_{max}$. Intuitively, this is because the PDE degenerates on the boundary in such a way that either the PDE becomes an ODE, or the resulting degenerate parabolic PDE has a normal hyperbolic term with outgoing characteristics. Note that rigorous arguments justifying this result can be deduced from (36).

We now turn our attention to the complete PIDE with a non-zero jump term (2.5). We note that the jump integral term contains no A dependence and so consequently, we can simply solve the PIDE along the boundaries $A = A_{max}, A = 0$ as in the no-jump case. In this case, the *boundary condition* for the PIDE is simply the continuous limit of $A \to 0, A \to A_{max}$, from points interior to the computational domain.

Taking the limit of the PIDE (2.5) as $S \to 0$, we obtain

$$\lim_{S\to 0} \mathcal{H}V = V_\tau - \left[\frac{-A}{T-\tau} V_A - rV \right] . \tag{2.10}$$

Note that the above limit is obtained by formally interchanging the limit and the integral in equation (2.5). Equation (2.10) is, of course, the same limit as in the PDE (no jump) case, since the integral terms disappear as $S \to 0$. Again, in this case, the boundary condition is simply the continuous limit of the PIDE as $S \to 0$. We will develop consistent, stable and monotone discrete schemes which do not require any data outside the computational domain at $S = 0$, and hence satisfaction of the boundary condition as $S = 0$ is ensured (see the discussion of viscosity solutions in subsection 2.3).

For the boundary condition as $S \to \infty$ we need to deal with two major issues. The £rst problem is that there is no obvious Dirichlet type condition which can be imposed for S large. The second issue concerns the fact that on any £nite domain $[0, S_{max}]$, the integral term appears to require information from outside the computational domain.

If we make the common assumption that $V_{SS} \to 0$ as $S \to \infty$ (27), then this implies that

$$V \simeq f(A,\tau)S + g(A,\tau) \tag{2.11}$$

as $S \to \infty$ which then means that equation (2.8) becomes

$$\mathcal{H}V \simeq V_\tau - \left(\frac{S-A}{T-\tau} V_A + rSV_S - rV \right) \ ; \ S \to \infty. \tag{2.12}$$

For $A \gg K$ and $S \to \infty$, we can approximate the solution to $\mathcal{H}V = 0$ by a linear function in A and S

$$V \simeq H_1(\tau)A + H_2(\tau)S + H_3(\tau) \tag{2.13}$$

so that

$$V \simeq \frac{D_1}{T} e^{-r\tau}(T-\tau)A + \left[\frac{D_1}{rT}(1-e^{-r\tau}) + D_2 \right] S + D_3 e^{-r\tau} \tag{2.14}$$

where D_1, D_2, D_3 are independent of (S, A, τ) and are determined by the payoff. For example, for a £xed strike call, $D_1 = 1, D_2 = 0, D_3 = -K$. Assuming (2.14) holds for $S \to \infty$, $\forall A$ (which is clearly an approximation for A small) we obtain

$$V_S \simeq \left[\frac{D_1}{rT}(1-e^{-r\tau}) + D_2 \right] \ ; \ S \to \infty \ . \tag{2.15}$$

Substituting equation (2.15) into equation (2.12) gives

$$\mathcal{H}V \equiv V_\tau - \left(\frac{(S-A)}{T-\tau} V_A + \chi(S,\tau) - rV \right) \ ; \ S \to \infty \ , \tag{2.16}$$

where

$$\chi(S,\tau) = \left[\frac{D_1}{rT}(1-e^{-r\tau}) + D_2\right] rS \ . \tag{2.17}$$

For the payoffs mentioned earlier, we have that $\chi(\tau,S) = \frac{(1-e^{-r\tau})}{T}S$ for a £xed strike call and $\chi(\tau,S) = 0$ for a £xed strike put.

The use of approximation (2.15) is discussed in (35), where it is mentioned that estimate (2.15) is in fact an upper bound for V_S. It must be admitted that use of equation (2.14) for all A as $S \to \infty$ is not rigorously justi£ed. However, we note that other authors (32) simply specify that the boundary condition at $S = S_{max}$ is set to the payoff. In (32), the size of the computational domain is increased as the grid size is reduced, so that the effect of poor speci£cation of the boundary condition becomes negligible. However, as discussed in (35), the boundary condition (2.16) is at least qualitatively correct.

In view of the above, we impose the boundary conditions as $S \to \infty$ in the following manner. We assume that S_{max} is suf£ciently large so that the solution can be well approximated by a linear function of S in the region $[S_{max} - \delta, S_{max}]$. Provided that we choose δ carefully (see (20)), the integral term in the PIDE (2.8) in $[0, S_{max} - \delta]$ has suf£cient data for accurate computation. In the region $[0, S_{max} - \delta] \times [0, A_{max}]$, the operator $\mathcal{H}$ is de£ned as in equation (2.8). In the region $[S_{max} - \delta, S_{max}] \times [0, A_{max}]$, we assume that V is linear in S, so that in this region the operator $\mathcal{H}V$ reduces to the PDE

$$\mathcal{H}V = V_\tau - \left[\frac{\sigma^2 S^2}{2} V_{SS} + \left(\frac{S-A}{T-\tau}\right) V_A + rSV_S - rV\right] \quad \text{for } S \in [S_{max} - \delta, S_{max}] \ . \tag{2.18}$$

Although we have assumed that $V_{SS} = 0$ in $[S_{max} - \delta, S_{max}]$, we leave the V_{SS} term in the operator (2.18) so that numerical solution is straightforward.

At $S = S_{max}$, based on equation (2.16), we de£ne $\mathcal{H}V$ as

$$\mathcal{H}V \equiv V_\tau - \left(\frac{S-A}{T-\tau} V_A + \chi(S,\tau) - rV\right) \ . \tag{2.19}$$

The above arguments can be used to justify formally solving the PIDE in the region $[0, S_{max}] \times [0, A_{max}]$, with the requirement that $\lambda = 0$ for $S \in [S_{max} - \delta, S_{max}]$. We then have a well de£ned problem in $[0, S_{max}] \times [0, A_{max}]$, which does not require data outside the computational domain. Note that in $[S_{max} - \delta, S_{max}]$, only a few nodes are required, since the solution is assumed to be linear. Hence this method of handling the dif£culties as $S \to \infty$ is not computationally expensive (at least as far as the PIDE solve is concerned).

2.3. Viscosity Solutions. In general, there may be no smooth solutions to equation (2.9). In what follows it will be understood that we are seeking weak viscosity solutions. A detailed discussion concerning existence and uniqueness of viscosity solutions to equation (2.9), can be found in (39; 2), for European and American options under jump diffusion. In addition, suf£cient conditions to ensure convergence of a discrete numerical scheme to the viscosity solution in the PDE case, is given in (6; 5). Finally, an extension of the results in (6; 5) to the case of option pricing problems with integro differential equations (jump diffusions) is given in (13) for nonlinear, multidimensional, degenerate problems and in (15) for linear problems. By exploiting the equivalence of Binomial Trees and explicit £nite difference methods, viscosity solution methods are used to prove converge of a lattice method for Asian options in (29).

3. Semi-Lagrangian Discretization. Typically when pricing continuously observed arithmetic average Asian option, the two dimensional problem (2.4) must be solved. In (51), the authors use a £nite volume approach combined with ¤ux limiters to solve equation (2.4). This requires solution of a set of nonlinear discretized algebraic equations at each timestep. When the convection terms become very large, (note that the convection term in the A direction in equation (2.5) becomes in£nite as $\tau \to T$), ¤ux-limiters revert to a £rst order upwind scheme which affects the accuracy of the solution (1). A related approach was used for Asian options in (37). In this paper we solve equation (2.4) using a semi-Lagrangian scheme. This idea was also suggested for American Asian options (without jumps but with stochastic volatility) in (38). In this section we explore different discretization methods for the partial differential equation using the semi-Lagrangian approach.

Before proceeding any further, let us introduce the following de£nitions. We use an unequally spaced grid in S coordinates for the PDE discretization $[S_0, \dots, S_M]$, and similarly use an unequally spaced grid in the A direction $[A_0, \dots A_M]$. Let

$$V_{i,j}^n = V(S_i, A_j, \tau^n) \tag{3.1}$$

denote the solution at asset price node S_i for the average A_j and time level n. Let $\mathcal{C}$ be the differential operator represented by

$$\mathcal{C}V = \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - (r - \lambda)V, \tag{3.2}$$

and

$$\mathcal{B}V = \lambda \int_0^\infty V(S\eta)g(\eta)d\eta \; . \tag{3.3}$$

Equation (2.5) can then be rewritten as

$$V_\tau + \frac{(A-S)}{T-\tau}V_A = \mathcal{C}V + \mathcal{B}V. \tag{3.4}$$

In order to impose the American early exercise constraint we will use a penalty method (19; 24; 49). Brie¤y, this replaces equation 2.9 by a non-linear PIDE

$$V_\tau + \frac{(A-S)}{T-\tau}V_A = \mathcal{C}V + \mathcal{B}V + q(V). \tag{3.5}$$

where the penalty term is

$$q(V) = \rho \max(V^* - V, 0). \tag{3.6}$$

The positive penalty parameter ρ is selected suf£ciently large so that either $V \geq V^*$ or

$$\frac{|V - V^*|}{V^*} < \varepsilon \;\; ; \;\; \varepsilon \ll 1 \;\; ; \quad \text{when } V < V^*, V^* > 0 \tag{3.7}$$

After discretizing the PDE with the non-linear penalty term, it is straightforward to select a suitably normalized discrete penalty parameter ρ', so that ε in equation (3.7) is less than a user speci£ed tolerance. We refer the reader to (24) for more details about this method.

We use standard £nite difference methods to discretize the operator $\mathcal{C}V$ (41) (see Appendix A). If we impose boundary condition (2.16), and use forward and backward differencing as appropriate, it is easy to see that the discrete form of $I - \mathcal{C}V$ is an M-matrix (see

Section 5.2). As discussed in (41), for typical values of σ, r, upwind differencing of the V_S term in equation (3.5) is required only rarely, and usually remote from regions of interest, so that in practice this does not impact solution quality. Requiring the discrete form of $\mathcal{C}V$ to be an M-matrix has interesting theoretical properties. In the following, we denote the discrete form of $\mathcal{C}V$ at $S = S_i, A = A_j, \tau = \tau^n$ by $(\mathcal{C}V)^n_{i,j}$. As described in (20; 19) the integral $(\mathcal{B}V)_{i,j}$ can be ef£ciently computed by transforming to equally spaced $\log S$ coordinates, approximating the integral using a Trapezoidal rule, using an FFT, and then transforming back to S coordinates. Special care is taken to avoid problems with wrap around (20). If linear interpolation is used to transform from equally spaced $\log S$ coordinates to unequally spaced S coordinates (and vice versa), this introduces a second order error consistent with the discretization of the PDE terms (20). Effectively, we are approximating $\mathcal{B}V$ by

$$\begin{aligned}(\mathcal{B}V)_{i,j} &\simeq \lambda \sum_k b_{ik} V_{kj} = \lambda B \cdot V_j \\ &\text{with } 0 \leq b_{ik} \leq 1 \quad \text{and } \sum_k b_{ik} \leq 1. \end{aligned} \tag{3.8}$$

Note that conditions (2.10) and (2.18) imply that we must also have

$$b_{ik} = 0 \quad \text{for} \quad i = 1 \quad \text{and} \quad i = M; \;\; \forall k \;\; . \tag{3.9}$$

The dense matrix multiply $B \cdot V_j$ can be evaluated ef£ciently using an FFT. For details, see (20; 19).

The Lagrangian derivative along a trajectory $A = A(S, \tau)$, for S £xed, is

$$\frac{DV}{D\tau} = \frac{\partial V}{\partial \tau} + \frac{\partial V}{\partial A} \frac{dA}{d\tau} \; . \tag{3.10}$$

Along the trajectory

$$\frac{dA}{d\tau} = \frac{A - S}{T - \tau} \tag{3.11}$$

equation (3.5) can be written as

$$\frac{DV}{D\tau} = \mathcal{C}V + \mathcal{B}V + q(V). \tag{3.12}$$

Let $A = A(S_i, A_j, \tau^{n+1}, \tau)$ along a trajectory satisfying equation (3.11), which passes through the discrete grid point (S_i, A_j) at $\tau = \tau^{n+1}$ for S_i being held constant. Let $A^n_{j(i,n+1)}$ be the departure point of this trajectory at $\tau = \tau^n$. Note that $A^n_{j(i,n+1)}$ will not necessarily coincide with a grid point A_j. Rather $A^n_{j(i,n+1)}$ is determined by solving

$$\frac{dA}{d\tau} = \frac{A - S_i}{T - \tau} \text{ where } A = A_j \;\text{ for }\; \tau = \tau^{n+1} \; , \tag{3.13}$$

from $\tau = \tau^{n+1}$ to $\tau = \tau^n$, that is,

$$A^n_{j(i,n+1)} = A_j + \lim_{\tau^* \to (\tau^{n+1})} \int_{\tau^*}^{\tau^n} \frac{A - S_i}{T - \tau} d\tau. \tag{3.14}$$

The limit in equation (3.14) is used to avoid problems at $\tau = T$, where ODE (3.13) becomes unde£ned.

Let $V^n_{i,j(i,n+1)} = V(S_i, A^n_{j(i,n+1)}, \tau^n)$ denote the value of the option price at the departure point of the trajectory. Then discretizing equation (3.12) along the characteristic trajectory for different timestepping schemes gives, in the case of fully implicit timestepping:

$$\frac{V^{n+1}_{i,j} - V^n_{i,j(i,n+1)}}{\Delta\tau} = (\mathcal{C}V)^{n+1}_{i,j} + (\mathcal{B}V)^{n+1}_{i,j} + q(V^{n+1}_{i,j}) \; , \tag{3.15}$$

and for Crank-Nicolson timestepping (CN),

$$\begin{aligned} \frac{V^{n+1}_{i,j} - V^n_{i,j(i,n+1)}}{\Delta\tau} &= \frac{1}{2}\left((\mathcal{C}V)^{n+1}_{i,j} + (\mathcal{B}V)^{n+1}_{i,j} \right) \\ &+ \frac{1}{2}\left((\mathcal{C}V_{i,j(i,n+1)})^n + (\mathcal{B}V)^n_{i,j(i,n+1)} \right) + q(V^{n+1}_{i,j}), \end{aligned} \tag{3.16}$$

and for second order backward differencing (BDF) (8)

$$\frac{\frac{1}{2}V^{n+1}_{i,j} - 2V^n_{i,j(i,n+1)} + \frac{1}{2}V^{n-1}_{i,j(i,n+1)}}{\Delta\tau} = (\mathcal{C}V)^{n+1}_{i,j} + (\mathcal{B}V)^{n+1}_{i,j} + q(V^{n+1}_{i,j}) \; . \tag{3.17}$$

For ease of exposition, we have written equation (3.17) for constant timesteps. This is trivially generalized to non-constant timesteps (8).

Unlike traditional applications of the semi-Lagrangian approach where the characteristic curve must be estimated numerically, for Asian options the solution along the characteristic curve can be determined exactly. Regarding S as a constant, and solving equation (3.11) gives

$$A = S_i + \frac{D}{T-\tau}, \tag{3.18}$$

where D is a constant independent of A (but a function of S_i). At $\tau = \tau^{n+1}$, $A = A_j$, so that

$$\begin{aligned} \text{At time } \tau^n : A^n_{j(i,n+1)} &= A_j + \frac{(S_i - A_j)(\tau^{n+1} - \tau^n)}{T - \tau^n}, \\ \text{At time } \tau^{n-1} : A^{n-1}_{j(i,n+1)} &= A_j + \frac{(S_i - A_j)(\tau^{n+1} - \tau^{n-1})}{T - \tau^{n-1}}, \end{aligned} \tag{3.19}$$

where $T \geq \tau^{n+1} > \tau^n > \tau^{n-1}$. Note that for the last step when $\tau^{n+1} = T$, equation (3.19) simpli£es to $A^n_{j(i,n+1)} = A^{n-1}_{j(i,n+1)} = S_i$, which is the correct limiting behaviour of equation (3.14). The various quantities $(\cdot)^n_{i,j(i,n+1)}$ in equations (3.15-3.17) are determined by interpolation along lines of constant $S = S_i$. Assuming that the S derivatives and the integral term are discretized using second order accurate methods, then it follows from (10; 23) that at least quadratic interpolation should be used for $(\cdot)^n_{i,j(i,n+1)}$ in order to retain global second order convergence.

4. Semi-Lagrangian Timestepping and Discrete Observations. Asian options which are continuously observed are typically considered to be the limit of discretely observed Asian options as the observation interval tends to zero (18; 45). A similar statement can also be made about the discretization of the continuous problem given in the previous section. In this section we compare these two discrete problems. We show that if the discrete sampling period is equal to the discrete PIDE timestep, then a fully implicit, discretely sampled model is algebraically identical to a fully implicit semi-Lagrangian discretization of a continuously observed model. In the following, we ignore the effect of the boundary condition (2.16).

Consider the discrete average computed at the discrete forward observation times $t^\ell = \ell \Delta t$

$$\hat{A}(t^\ell) = \frac{1}{\ell} \sum_{p=1}^{\ell} S(t^p)$$
$$= \hat{A}(t^{\ell-1}) + \frac{S(t^\ell) - \hat{A}(t^{\ell-1})}{\ell} \ . \tag{4.1}$$

Note that we are careful to distinguish the average computed at forward times as a function of t, from the average computed at backwards times as a function of τ, that is, $A(\tau) = \hat{A}(t)$ where $\tau = T - t$. Similarly, we can de£ne the value of an Asian option as a function of $(S, \hat{A}, t)$ as $\hat{V}(S, \hat{A}, t)$, where, in terms of $V(S, A, \tau)$ we have the identity $V(S, A(T-t), T-t) = \hat{V}(S, \hat{A}(t), t)$.

When using a PIDE method to price a discretely observed Asian option, we consider that $\hat{V} = \hat{V}(S, \hat{A}, t)$, and regard $(S, \hat{A})$ as independent variables. Suppose we have N observation dates, at the times $\Delta t, 2\Delta t, \ldots, N\Delta t$ with $N\Delta t = T$. Let $\hat{A}^\ell = \hat{A}(t^\ell)$. Then at the ℓ-th observation date we must have, by no arbitrage (47),

$$\hat{V}(S, \hat{A}^{\ell+1}, t^{(\ell+1)+}) = \hat{V}(S, \hat{A}^\ell, t^{(\ell+1)-}) \tag{4.2}$$

where $t^{(\ell+1)+}, t^{(\ell+1)-}$ are the instants just after and just before the observation date $t^{\ell+1}$. In equation (4.2) $\hat{A}^{\ell+1} = \hat{A}^\ell + \frac{S - \hat{A}^\ell}{\ell+1}$ and is regarded as constant for $t^{\ell+1} < t < t^{\ell+2}$.

Let $A^k = A(\tau^k)$ and set $k = N - \ell$, so that k counts backwards. Since $\tau^k = k\Delta\tau$ we have that $t^\ell = T - \tau^k$, $t^{\ell+1} = T - \tau^{k-1}$ and as well, $t^{(\ell+1)+} = T - \tau^{(k-1)-}$ and $t^{(\ell+1)-} = T - \tau^{(k-1)+}$. Writing the no-arbitrage condition (4.2) in terms of the variables $\tau, A(\tau)$ rather than t, $\hat{A}(t)$ then gives

$$V(S, A^{k+1}, \tau^{k+}) = V(S, A^k, \tau^{k-}) \text{ where } A^k = A^{k+1} + \frac{S - A^{k+1}}{N-k}. \tag{4.3}$$

As before we regard A^{k+1} as £xed during $\tau^k < \tau < \tau^{k+1}$.

Consider the case of a discretely observed European Asian option. In this case we solve

$$V_\tau = \mathcal{C}V + \mathcal{B}V + q(V) \tag{4.4}$$

on the domain $[0, S_{max}] \times [0, A_{max}]$, with the no-arbitrage conditions (4.3) imposed at observation times. Away from observation dates, if we discretize equation (4.4) in the A direction, then equation (4.4) represents a set of one dimensional PIDEs, which communicate only through no-arbitrage conditions (18).

From the no-arbitrage condition (4.3) we have that a fully implicit discretization of equation (4.4) gives

$$\frac{V(S_i, A_j^{k+1}, \tau^{(k+1)-}) - V(S_i, A_j^{k+1}, \tau^{k+})}{\Delta\tau} = (\mathcal{C}V)_{i,j}^{k+1} + (\mathcal{B}V)_{i,j}^{k+1} + q(V_{i,j}^{k+1}) \ . \tag{4.5}$$

Note that this is a set of independent one dimensional PDEs (there are no A derivatives in equation (4.5), A_j^{k+1} appears only as a parameter). Using the no-arbitrage condition (4.3) in equation (4.5) gives

$$\frac{V(S_i, A_j^{k+1}, \tau^{(k+1)-}) - V(S_i, A_{j(i,k+1)}^{k}, \tau^{k-})}{\Delta\tau} = (\mathcal{C}V)_{i,j}^{k+1} + (\mathcal{B}V)_{i,j}^{k+1} + q(V_{i,j}^{k+1}) \tag{4.6}$$

where

$$A^k_{j(i,k+1)} = A^{k+1}_j + \frac{(S - A^{k+1}_j)\Delta\tau}{T - \tau^k}, \tag{4.7}$$

which we recognize from equation (3.15) and equation (3.19) as being algebraically identical to a semi-Lagrangian, fully implicit discretization.

Note that in order for this result to hold, we must have discrete observations at $t = \Delta t, 2\Delta t, ..., N\Delta t$, that is, no observation at $t = 0$. Of course, in the limit as $\Delta t \to 0$, adding an extra observation at $t = 0$ will be the same to $O(\Delta t)$ as the semi-Lagrangian solution.

REMARK 4.1. *As discussed in (25), it is straightforward to show that the common lattice methods used to price Asian options (28) are simply explicit £nite difference methods for discretely observed models of Asian options. In many lattice applications, the observation interval is set to the lattice timestep, hence the continuously observed price is computed in the limit of vanishing timestep. A straightforward extension of the results above can be used to show that these lattice methods are simply explicit semi-Lagrangian methods. In this case, it is also easy to derive the conditions on the order of interpolation and the spacing on the lattice in the average direction to ensure optimal convergence. We note that, as discussed in (25), this is a point of confusion in the £nance literature, and has led to schemes which are not, in fact, convergent (25; 7). A convergence proof using viscosity solution ideas for lattice methods (for Asian options) is given in (29), but only for the case where the lattice contains all possible average values.*

5. Monotonicity and Stability of the Discrete Equations. As shown in (41), in the case of nonlinear option pricing problems, seemingly reasonable discretization schemes can converge to an incorrect (i.e. non-viscosity) solution. Convergence to the viscosity solution is guaranteed if the discretization is consistent, monotone and l_∞ stable (5; 13; 15). Usually, consistency follows if any reasonable discretization method is used, although in the case of jump-diffusion, the non-locality of the integral term requires care in showing consistency (13; 15). l_∞ stability is usually a consequence of monotonicity. Consequently, the most interesting requirement is monotonicity.

In the following, we will investigate the monotonicity and stability properties of the discrete equations. We will use a de£nition of monotonicity which is somewhat stricter than is usually the case in £nancial applications (5), but more in line with the de£nition used in computational ¤uid dynamics (CFD) (30). It appears to us that the CFD de£nition is a more useful aid to the design of suitable discrete schemes. In Appendix B, we will show that CFD de£nition of monotonicity can be deduced from the usual de£nition of monotonicity in the viscosity solution literature, as long as the discretization is consistent.

We remind the reader that use of semi-Lagrangian timestepping decouples the discrete equations at each timestep, resulting in a set of one dimensional discrete PIDEs. Hence we can use our techniques in (24; 41; 19; 20) to prove the desired properties of the discretized equations.

5.1. Preliminaries. De£ne the matrices B and C such that

$$\lambda \left[B \cdot V^n_j\right]_i = (\mathcal{B}V)^n_{i,j} + \text{ truncation error} \tag{5.1}$$

$$\left[C \cdot V^n_j\right]_i = (\mathcal{C}V)^n_{i,j} + \text{ truncation error} \tag{5.2}$$

where V^n_j is the vector of discrete solution values $\left[V^n_j\right]_i = V(A_j, S_i, \tau^n)$ for £xed A_j. A detailed description of B is given in (20; 19). For our purposes, we note that B has the properties given by equation (3.8). A detailed description of matrix C is given in Appendix A.

To avoid algebraic complication, we will describe the discrete equations and the method used to solve the algebraic equations, only for the fully implicit and Crank-Nicolson timestepping methods. The reader should have no dif£culty generalizing the results to the BDF case.

Let Φ^{n+1} be the Lagrange interpolation operator such that

$$(\Phi^{n+1} \cdot V^n)_{i,j} = V(S_i, A^n_{j(i,n+1)}, \tau^n) + \text{ interpolation error} \tag{5.3}$$

where Φ^{n+1} is a linear operator for any order (linear, quadratic) of interpolation. We also let V^* be the vector of payoffs obtained upon exercise and P be the diagonal matrix given by

$$P(V_j^{n+1})_{ii} = \begin{cases} Large & \text{if } V_{i,j}^{n+1} < V_{i,j}^* \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

Then the matrix form of the discrete equations for the penalized method is given by

$$\begin{aligned}[] [I - (1-\theta)\Delta\tau C + P(V_j^{n+1})]V_j^{n+1} &= [\Phi^{n+1}[I + \theta\Delta\tau C]V^n]_j + (1-\theta)\lambda\Delta\tau BV_j^{n+1} \\ &\quad + [\Phi^{n+1}\theta\lambda\Delta\tau BV^n]_j + \left[P(V_j^{n+1})\right] V_j^* + \Delta\tau F_j^{n+1} \end{aligned} \tag{5.5}$$

for $j = 1,..,M$. Here $\theta = 0$ is fully implicit, and $\theta = 1/2$ is Crank-Nicolson timestepping. The term F_j^{n+1} is used to approximate the boundary condition at $S = S_{max}$, as discussed in subsection 2.2. Note that $B_{i,l} = 0$ for $S_i \in [S_{max} - \delta, S_{max}]$, as discussed in subsection 2.2. Condition (2.16) is enforced at $i = M$ by adjusting $C_{M,l}$ as discussed in Appendix A, and letting

$$\left[F_j^{n+1}\right]_i = \begin{cases} 0 & , \; i \neq M \\ (1-\theta)\chi(S_{max}, \tau^{n+1}) + \theta\chi(S_{max}, \tau^n) & , \; i = M \end{cases} \tag{5.6}$$

where $\chi(S,\tau)$ is discussed in subsection 2.2.

We note here that the penalty formulation of the American option pricing problem reduces problem (2.9) to the nonlinear PIDE (3.5), and hence the results in (13) apply. Brie¤y, if the numerical scheme is consistent, l_∞ stable, and monotone, then convergence to the viscosity solution is guaranteed. In order to obtain a monotone scheme, we can use at most linear interpolation in equation (5.3).

The choice of interpolation scheme is discussed in (25) and (23). Speci£cally, if the interpolation error does not get damped out, the global interpolation error after N timesteps is $O\left(\frac{((\Delta S)_{\max})^q}{\Delta\tau}\right)$, where $q = 2$ for linear interpolation, $q = 3$ for quadratic interpolation and $(\Delta S)_{\max} = \max_i (S_{i+1} - S_i)$. Assuming second order in space and time truncation errors, the global discretization error for a semi-Lagrangian method applied to a pure hyperbolic problem is (23; 10)

$$\text{global discretization error} = O\left[\frac{((\Delta S)_{\max})^q}{\Delta\tau} + ((\Delta S)_{\max})^2 + (\Delta\tau)^2\right] , \tag{5.7}$$

If we assume $(\Delta S)_{max} = const.h$ and $\Delta\tau = const.h$, then equation (5.7) reduces to

$$\text{global discretization error} = O\left[\min((h^{q-1}, h^2)\right] \tag{5.8}$$

As discussed in (10), estimate (5.8) is valid only for smooth solutions.

5.2. Monotonicity and Stability. The highest order interpolation method Φ which will result in a monotone scheme is linear interpolation. Equation (5.8) suggests that if linear interpolation is used ($q = 2$), we can obtain no more than £rst order convergence. With this in mind, in the following analysis, we will consider only a fully implicit timestepping, and a linear interpolant Φ (as in equation (5.3)). We will, however, carry out numerical experiments with Crank-Nicolson and BDF timestepping, and higher order interpolants. The fully implicit version of equation (5.5) is

$$[I - \Delta\tau C - \lambda\Delta\tau B + P(V_j^{n+1})]V_j^{n+1} = [\Phi^{n+1}V^n]_j + P(V_j^{n+1})V_j^* + \Delta\tau F_j^{n+1} \tag{5.9}$$

for $j = 1, ..., M$

As discussed in (5; 13), consistency, stability, and monotonicity are suf£cient conditions for a numerical scheme to ensure convergence to the viscosity solution. In view of the importance of discretizations which are stable and monotone, both from a theoretical and practical point of view, it is useful to gather together a set of results for the implicit discretization schemes.

LEMMA 5.1 (Properties of Matrix C). *The matrix C in equation (5.9) has the properties*

$$\begin{aligned} \sum_k C_{ik} &= -(r+\lambda) \ ; \ i = 2, ..., M-1 \\ &= -r \ ; \ i = 1 \ \ \textit{and} \ \ i = M \\ C_{ik} &\geq 0 \ ; \ i \neq k \ ; \ i = 1, ..., M \ . \end{aligned} \tag{5.10}$$

Proof. This follows directly from the discussion in Appendix A. □

Set

$$Q = I - \Delta\tau C - \lambda\Delta\tau B. \tag{5.11}$$

LEMMA 5.2 (M-matrix property of Q). *The matrix Q is an M matrix.*

Proof. From equations (3.8-3.9) we have that $-B$ has non-positive offdiagonal elements. From Lemma 5.1, we have that $-C - \lambda B$ has non-positive offdiagonal elements. From Lemma 5.1, and properties (3.8-3.9), we have that

$$\sum_k [-C - \lambda B]_{ik} \geq 0 \ ; \ i = 1, ..., M, \tag{5.12}$$

and hence Q is an M matrix. □

We can write the discrete equations at each node (S_i, A_j) as

$$\begin{aligned} g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) &= -[QV_j^{n+1}]_i + [\Phi^{n+1}V^n]_{i,j} + \Big[P(V_j^{n+1})\Big]_{ii}(V_{i,j}^* - V_{i,j}^{n+1}) + \Delta\tau F_{i,j}^{n+1} \\ &= 0 \end{aligned} \tag{5.13}$$

where $\{V_{k,j}^{n+1}\}_i$ is to be interpreted as the set of values $V_{k,j}^{n+1}, k \neq i$. $k = 1, ..., M$, and $\{V^n\}$ is the set $V_{k,\ell}^n, k = 1, .., M; \ell = 1, ..., M$.

DEFINITION 5.3 (Monotone Discretizations). *A discretization of the form* (5.13) *is monotone if*

$$\begin{aligned} g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1} + \rho_{k,j}^{n+1}\}_i, \{V_{k,\ell}^n + \rho_{k,\ell}^n\}) \geq g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) \quad \forall i, j; \quad \forall k \neq i \\ \forall \rho_{k,\ell}^n \geq 0, \ \forall \rho_{k,j}^{n+1} \geq 0, \end{aligned} \tag{5.14}$$

$$\begin{aligned} g_{i,j}(V_{i,j}^{n+1} + \rho_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) < g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) \quad \forall i, j; \quad \forall k \neq i \\ \forall \rho_{i,j}^{n+1} > 0 \end{aligned} \tag{5.15}$$

REMARK 5.1. *The above de£nition of monotonicity includes the condition (5.15). In the viscosity solution literature (5), only condition (5.14) is used to de£ne monotonicity. However, in the conservation law literature (30; 26) monotonicity is usually de£ned including condition (5.15). In Appendix B, we show that consistency (for Hamilton-Jacobi-Bellman PDEs) and condition (5.14) implies condition (5.15). However, we believe that it is more useful to include condition (5.15), in a de£nition of monotonicity. This allows for a simple interpretation of the meaning of monotonicity, without reference to any other conditions, which is a useful aid in designing discrete schemes.*

THEOREM 5.4 (Monotonicity of the Discretization). *The fully implicit discretization (5.13) is unconditionally monotone.*

Proof. We rewrite equation (5.13) as

$$g_{i,j} = -[QV_j^{n+1}]_i + [\Phi^{n+1}V^n]_{i,j} + \left[P(V_j^{n+1})\right]_{ii}(V_{i,j}^* - V_{i,j}^{n+1}) + \Delta\tau F_{i,j}^{n+1} \tag{5.16}$$

and examine each term in equation (5.16). From Lemma 5.2, matrix Q is an M matrix, hence $-[QV_j^{n+1}]_i$ is a strictly decreasing function of $V_{i,j}^{n+1}$, and a non-decreasing function of $\{V_{k,j}^{n+1}\}_i$. Since Φ^{n+1} is a linear interpolant operator, $[\Phi^{n+1}V^n]_{i,j}$ is a non-decreasing function of $\{V^n\}$. Finally we see that the term $\left[P(V_j^{n+1})\right]_{ii}(V_{i,j}^* - V_{i,j}^{n+1})$ is a non-increasing function of $V_{i,j}^{n+1}$. Hence the discretization is monotone from De£nition 5.3. $\square$

THEOREM 5.5 (Stability of the Fully Implicit Scheme). *The fully implicit method satis£es*

$$\|V^{n+1}\|_\infty \leq \max(\|V^n\|_\infty, \|V^*\|_\infty) + \Delta\tau\chi_{max} \tag{5.17}$$

where $\chi(S_{max},\tau)$ *is de£ned in subsection 2.2, and*

$$\chi_{max} = \max_{0\leq\tau\leq T} |\chi(S_{max},\tau)| \ . \tag{5.18}$$

In particular,

$$\|V^{n+1}\|_\infty \leq \|V^*\|_\infty + T\chi_{max} \ \ , \tag{5.19}$$

where $(n+1)\Delta\tau \leq T$.

Proof. Writing out equation (5.9) in component form gives (see Appendix A)

$$\begin{aligned} V_{i,j}^{n+1}(1+(\alpha_i+\beta_i+r+\lambda)\Delta\tau) - \alpha_i\Delta\tau V_{i-1,j}^{n+1} - \beta_i\Delta\tau V_{i+1,j}^{n+1} - \lambda\Delta\tau\sum_k b_{ik}V_{k,j}^{n+1} + P_{ii}(V_j^{n+1})V_{i,j}^{n+1} \\ = \sum_{k,l} w_{k,l}^{i,j}V_{k,l}^n + P_{ii}(V_j^{n+1})V_{i,j}^* + \Delta\tau F_{i,j}^{n+1} \ \ , \end{aligned} \tag{5.20}$$

where $w_{k,l}^{i,j}$ are linear interpolation weights satisfying

$$0 \leq w_{k,l}^{i,j} \leq 1 \text{ and } \sum_{k,l} w_{k,l}^{i,j} = 1 \ \ . \tag{5.21}$$

In addition, we recall that B has properties (3.8)

$$0 \leq b_{ik} \leq 1 \quad \text{and} \quad \sum_k b_{ik} \leq 1 \ \ , \tag{5.22}$$

while from Appendix A we have that

$$\alpha_i \geq 0 \quad \text{and} \quad \beta_i \geq 0 \ . \tag{5.23}$$

Let m be an index such that

$$|V_{m,j}^{n+1}| = \|V_j^{n+1}\|_\infty. \tag{5.24}$$

Then equations (5.20-5.23) imply that

$$\|V_j^{n+1}\|_\infty \left(1 + r\Delta\tau + P(V_j^{n+1})_{mm}\right) \leq \|V^n\|_\infty + P(V_j^{n+1})_{mm}\|V^*\|_\infty + \Delta\tau\chi_{max} \tag{5.25}$$

and so

$$\begin{aligned}\|V_j^{n+1}\|_\infty &\leq \max(\|V^n\|_\infty, \|V^*\|_\infty)\frac{1+P(V_j^{n+1})_{mm}}{1+r\Delta\tau+P(V_j^{n+1})_{mm}} + \frac{\Delta\tau\chi_{max}}{1+r\Delta\tau+P(V_j^{n+1})_{mm}} \\ &\leq \max(\|V^n\|_\infty, \|V^*\|_\infty) + \Delta\tau\chi_{max} \ .\end{aligned} \tag{5.26}$$

Hence

$$\|V^{n+1}\|_\infty \leq \max(\|V^n\|_\infty, \|V^*\|_\infty) + \Delta\tau\chi_{max}. \tag{5.27}$$

Therefore, by induction we have

$$\|V^{n+1}\|_\infty \leq \max(\|V^{n-i}\|_\infty, \|V^*\|_\infty) + (i+1)\Delta\tau\chi_{max} \tag{5.28}$$

for all i. Equation (5.19) follows from setting $i = n$ in equation (5.28). □

REMARK 5.2 (Extension to nonlinear models). *It is completely straightforward to include a transaction cost or uncertain volatility model in the basic option pricing PIDE (42), which makes the PIDE nonlinear (even in the European case). For example, using the methods in (42), it is a simple exercise to extend the above stability and monotonicity results to the case of an American Asian option, with jumps and transaction costs.*

5.3. Properties of a Semi-implicit Discretization. Suppose we alter the discretization (5.9) so that the jump integral term is evaluated explicitly. Then

$$[I - \Delta\tau C + P(V_j^{n+1})]V_j^{n+1} = \lambda\Delta\tau BV_j^n + [\Phi^{n+1}V^n]_j + P(V_j^{n+1})V_j^* + \Delta\tau F_j^{n+1} \tag{5.29}$$

for $j = 1,...,M$. The previous methods can also be applied to determine stability and monotonicity properties of this second discretization.

THEOREM 5.6 (Stability and Monotonicity of Explicit Evaluation of the Jump Term). *The discretization (5.29) is unconditionally stable and monotone.*

Proof. Set $R = [I - \Delta\tau C]$. Then we can rewrite equation (5.29) as

$$g_{i,j} = -[RV_j^{n+1}]_i + [\Phi^{n+1}V^n]_{i,j} + \Big[P(V_j^{n+1})\Big]_{ii}(V_{i,j}^* - V_{i,j}^{n+1}) + \Delta\tau F_{i,j}^{n+1} + [\Delta\tau\lambda BV^n]_{i,j} \tag{5.30}$$

From Lemma 5.1 and the properties of matrix B (equation (3.8)), and following along the lines used to prove Theorem 5.4, it is straightforward to see that De£nition 5.3 holds unconditionally for equation (5.30). Using a similar maximum analysis as in the proof of Theorem 5.5, we obtain unconditional stability. □

REMARK 5.3. *Scheme (5.29) is very simple to implement, and retains unconditional monotonicity and stability. This method appears to have been completely overlooked. However, this scheme is only £rst order correct in time.*

6. Additional Properties of the Discrete Equations. In this section we investigate how well our discrete approximation (3.5) preserves important properties of our original problem (2.9). We focus on two important properties : how well does the discrete penalty method satisfy the inequality constraints in problem (2.9) and does the discretization preserve arbitrage inequalities (15).

6.1. Error in the Penalty Formulation. In our original problem (2.9) we need to solve $\min(\mathcal{H}V; V - V^*) = 0$. In particular, we require that

$$(V - V^*) \geq 0. \tag{6.1}$$

In discrete terms this becomes

$$(V_{i,j}^{n+1} - V_{i,j}^*) \geq 0 \ . \tag{6.2}$$

However, the penalty formulation (5.5) will result in $V_{i,j}^{n+1} < V_{i,j}^*$ at nodes in the exercise region. In this subsection we show that at these nodes we have $V_{i,j}^{n+1} = V_{i,j}^* - \varepsilon$, where $0 < \varepsilon \ll 1$. In particular, we have the following bound on the error in the penalty term.

LEMMA 6.1 (Error generated by the penalty formulation). *Assume that V^* satis£es a Lipschitz condition and suppose that*

$$\frac{\Delta\tau}{\Delta S_{min}} < const. \ \ as \quad \Delta\tau, \Delta S_{min} \rightarrow 0 \tag{6.3}$$

where $\Delta S_{min} = \min_i(S_{i+1} - S_i)$. Then

$$V_{i,j}^{n+1} - V_{i,j}^* \geq -\frac{C_1}{Large} \tag{6.4}$$

where C_1 is a positive constant independent of $\Delta S, \Delta\tau$.

Proof. Let k be an index such that

$$(V_{k,j}^* - V_{k,j}^{n+1}) = \max_i (V_{i,j}^* - V_{i,j}) \ . \tag{6.5}$$

Since the matrix Q de£ned by (5.11) is an M matrix from Lemma 5.2, it follows from equation (6.5) that

$$\left[Q(V_j^* - V_j^{n+1})\right]_k \geq 0 \ , \tag{6.6}$$

and hence

$$\left[QV_j^*\right]_k \geq \left[QV_j^{n+1}\right]_k \ . \tag{6.7}$$

From equation (5.16) we have that for all j

$$QV_j^{n+1} = [\Phi V^n]_j + P(V_j^{n+1})(V_j^* - V_j^{n+1}) + \Delta\tau F_j^{n+1} \ . \tag{6.8}$$

In particular, row k of equation (6.8) is

$$\left[QV_j^{n+1}\right]_k = \left[[\Phi V^n]_j\right]_k + \left[P(V_j^{n+1})(V_j^* - V_j^{n+1})\right]_k \ + \Delta\tau F_{k,j}^{n+1}. \tag{6.9}$$

Since

$$\left[P(V_j^{n+1})(V_j^* - V_j^{n+1})\right]_k = \|P(V_j^{n+1})(V_j^* - V_j^{n+1})\|_\infty \tag{6.10}$$

then equation (6.9) gives (using equation (6.7))

$$\|P(V_j^{n+1})(V_j^* - V_j^{n+1})\|_\infty \le \|V^n\|_\infty + \|QV^*\|_\infty + \Delta\tau\chi_{max} \ . \tag{6.11}$$

From Theorem 5.5, and equation (6.11) we have

$$\|P(V_j^{n+1})(V_j^* - V_j^{n+1})\|_\infty \le C_2 + \|QV^*\|_\infty, \tag{6.12}$$

where $C_2 = \|V^*\|_\infty + T\chi_{max}$, with $(n+1)\Delta\tau \le T$. Assuming that V^* satis£es a Lipschitz condition, then

$$\|QV^*\|_\infty \le C_3 \frac{\Delta\tau}{\Delta S_{min}} \tag{6.13}$$

which follows from Lemma 5.2 and Appendix A. Assuming $\Delta\tau/\Delta S_{min}$ is bounded, we have

$$(V_{i,j}^{n+1} - V_{i,j}^*) \ge -\frac{C_1}{Large} \tag{6.14}$$

with $C_1 = C_2 + \frac{C_3\Delta\tau}{\Delta S_{min}} > 0$. ∎

REMARK 6.1 (Signi£cance of Lemma 6.1). *Lemma 6.1 shows that the error induced by approximating problem (2.9) by the penalized system (3.6) can be made arbitrarily small by making the quantity Large (equation (5.4)) suf£ciently large, provided the grid size is reduced such that $\Delta\tau/\Delta S_{min}$ is bounded. In practice, this condition is not restrictive, since it does not make any sense to drive the spatial grid error to zero, leaving a £nite timestepping error. Of course, when using £nite precision arithmetic, our ability to distinguish (numerically) $(V_{i,j}^{n+1} - V_{i,j}^*)$ from zero is limited due to roundoff. As discussed in (24), this is not a problem of practical concern, since roundoff causes dif£culty only when seeking to enforce condition (6.2) to unrealistic levels of accuracy.*

THEOREM 6.2 (Discrete Comparison Principle). *The fully implicit discretization (5.9) satis£es a discrete comparison principle, that is, if $V^n > W^n$ and V^{n+1}, W^{n+1} satisfy equation (5.9), then $V^{n+1} > W^{n+1}$.*

Proof. Suppose $V^n > W^n$. Write equation (5.9) for V, W

$$\begin{aligned} QV_j^{n+1} &= [\Phi^{n+1}V^n]_j + \left[P(V_j^{n+1})\right](V_j^* - V_j^{n+1}) + \Delta\tau F_j^{n+1} \\ QW_j^{n+1} &= [\Phi^{n+1}W^n]_j + \left[P(W_j^{n+1})\right](V_j^* - W_j^{n+1}) + \Delta\tau F_j^{n+1} \end{aligned} \tag{6.15}$$

Some manipulation of equation (6.15) results in

$$\begin{aligned} Q(V_j - W_j)^{n+1} = &-P(W_j^{n+1})(V_j - W_j)^{n+1} + (P(V_j^{n+1}) - P(W_j^{n+1})(V_j^* - V_j^{n+1}) \\ &+ [\Phi^{n+1}(V^n - W^n)]_j \end{aligned} \tag{6.16}$$

or

$$\left[Q + P(W_j^{n+1})\right](V_j - W_j)^{n+1} = (P(V_j^{n+1}) - P(W_j^{n+1})(V_j^* - V_j^{n+1}) + [\Phi^{n+1}(V^n - W^n)]_j \tag{6.17}$$

Since Q is an M matrix we have that $[Q + P(W_j^{n+1})]$ is also an M matrix. From equation (5.4), we have that

$$(P(V_j^{n+1}) - P(W_j^{n+1})(V_j^* - V_j^{n+1}) \ge 0 \tag{6.18}$$

If linear interpolation is used, then $(V^n - W^n) > 0$ implies that $[\Phi^{n+1}(V^n - W^n)]_j > 0$. Finally, since $[Q + P(W_j^{n+1})]$ is an M matrix, its inverse satis£es $[Q+P(W_j^{n+1})]^{-1} \geq 0$, and $diag([Q + P(W_j^{n+1})]^{-1}) > 0$, and hence $(V_j - W_j)^{n+1} > 0$. ☐

REMARK 6.2. *As discussed in (15), Lemma 6.2 has the £nancial interpretation that the discrete option prices satisfy arbitrage inequalities, that is, the inequality of payoffs is preserved in the inequalities of option prices.*

7. Iterative Solution of the Discretized Equations. In order to solve equation (5.5), we use the following iteration scheme

Iteration

For $j = 1,2,\ldots$

Let $(V_j^{n+1})^0 = (V_j^n)$

Let $\hat{V}_j^{\,k} = (V_j^{n+1})^k$

Let $\hat{P}^k = P((V_j^{n+1})^k)$

For $k = 0,1,2,\ldots$ until convergence

Solve

$$\left[I - (1-\theta)C + \hat{P}^k\right]\hat{V}_j^{\,k+1} = \left[\Phi^{n+1}\left[I+\theta C\right]V^n\right]_j + \hat{P}^k V_j^* + \Delta\tau F_j^{n+1} + (1-\theta)\lambda\Delta\tau B\hat{V}_j^{\,k} + \theta\lambda\Delta\tau\left[\Phi^{n+1}BV^n\right]_j \tag{7.1}$$

If $\max_i \dfrac{|\hat{V}_{i,j}^{k+1} - \hat{V}_{i,j}^{k}|}{\max(1, |\hat{V}_{i,j}^{k+1}|)} < tol$ then break

EndFor

EndFor

For clarity, we have given algorithm 7.1 only for Crank-Nicolson and fully implicit timestepping. However, it is trivial to generalize this method to BDF timestepping. Note that each iteration of algorithm (7.1) requires a tridiagonal factor and solve, and a forward and back FFT (to evaluate $B \cdot \hat{V}_j^{\,k}$).

The following Theorem indicates that iteration scheme (7.1) is globally convergent.

THEOREM 7.1 (Convergence of Iteration). *Let matrices C, B and $\hat{P}$ be given by (5.2), (5.1) and (5.4), respectively. Assume that matrix $I - (1-\theta)C$ is an M-matrix (which follows from Lemma 5.1), and that B has properties (3.8). Then iteration (7.1) is globally convergent to the unique solution of equation (5.5) for any initial iterate $\hat{V}^0$.*

Proof. Note that the algebraic equations (5.5) are decoupled for each line of constant A_j. Hence the issue of convergence of scheme (7.1) reduces to the convergence of each set of equations for constant j. But for constant j, this iteration is equivalent to solution of the discrete penalized equations for one dimensional American options with jump diffusion as described in (19). Hence the result follows directly from Theorem 4.2 in (19). ☐

8. Computational Details and Numerical Results. This section presents numerical results for various options and payoffs, including vanilla European call/put and American options. We will use an unequally spaced grid in the A, S directions, on the domain $[0, S_{max}] \times$

$[0, A_{max}]$, with $A_{max} = S_{max}$. We remind the reader that δ is selected so that in $[S_{max} - \delta, S_{max}]$ we have that $\lambda = 0$, and we thus have suf£cient data for accurate computation of the jump integral term in $[0, S_{max}]$. The method used to determine δ is discussed in (20). Probabilistic arguments can be used to determine an appropriate value for S_{max} (45), so that a linear approximation to V is justi£ed in $[S_{max} - \delta, S_{max}]$. We use $S_{max} = 50K$, where K is the strike. We describe below some tests which were carried out to verify that the effect of imposing boundary conditions at $S = 50K$ results in insigni£cant error. The convergence tolerance in iteration (7.1) was $tol = 10^{-6}$. As suggested in (24), we choose $Large = 1/tol$. If Crank-Nicolson or BDF timestepping is used, then quadratic interpolation is used in equation (5.3). If fully implicit timestepping is employed, then linear interpolation is used in equation (5.3).

Given an A grid discretization, the discrete PIDEs (3.15-3.16) become decoupled. At each timestep, we have a set of independent one dimensional discrete PIDEs to solve. This property makes solution of the continuously observed Asian option straightforward to implement, given an existing library which supports pricing of path dependent options.

As pointed out in equation (5.8), it is necessary to use at least a quadratic Lagrange interpolation scheme to £nd the solution at the foot of the characteristic curve, if we hope to obtain quadratic convergence. This will, however, result in a scheme which is not monotone.

The convergence ratio R is de£ned in the following way. For each test, as we double the number of grid points in both S and A directions, we cut the timesteps $(\Delta\tau)$ in half. Let $\Delta\tau = \max_n(\tau^{n+1} - \tau^n)$, $(\Delta A)_{\max} = \max_j(A_{j+1} - A_j)$. Note that we are allowing here for the possibility of using variable timestep sizes (to be explained later), although most of our tests will simply use a constant timestep size. If we then carry out a convergence study, letting $h \to 0$, where $\Delta S_{\max} = \text{Const. } h$, $(\Delta A)_{\max} = \text{Const. } h$, and $\Delta\tau = \text{Const. } h$. then we can assume that the error in the solution (at a given node) is

$$V_{\text{approx}}(h) = V_{\text{exact}} + \text{Const. } h^{\xi}.$$

The convergence ratio is then de£ned as

$$R = \frac{V_{\text{approx}}(h/2) - V_{\text{approx}}(h)}{V_{\text{approx}}(h/4) - V_{\text{approx}}(h/2)}. \tag{8.1}$$

In the case of quadratic convergence ($\xi = 2$), then $R = 4$, while for linear convergence ($\xi = 1$), $R = 2$.

Table 8.1 Value of a continuously observed £xed strike European Asian call option (no jumps) with constant timesteps. The input parameters are $\sigma = .1$, $r = .1$, $T = .25$, $\lambda = 0$ and $K = 100$. We compare the results given using the Večeř (46) one dimensional model, and the semi-Lagrangian method presented here. Crank-Nicolson timestepping was used.

semi-Lagrangian				Večeř 1-D PDE (46)			
Size of	No. of	$S = 100$		Size of	No. of	$S = 100$	
S and A grids	timesteps	Value	R	S grids	timesteps	Value	R
51	25	1.857193	n.a.	51	25	1.839863	n.a.
101	50	1.853254	n.a.	101	50	1.848642	n.a
201	100	1.852120	3.475	201	100	1.850851	3.974
401	200	1.851781	3.338	401	200	1.851407	3.979
801	400	1.851660	2.815	801	400	1.851546	3.987

Table 8.2 Value of a continuously observed £xed strike Asian call option (no jumps) at $S = K = 100$, constant Crank-Nicolson timestepping. The input parameters are $\sigma = .1$, $\lambda = 0$, $r = .1$, $T = .25$, $K = 100$. Convergence ratios (8.1) are presented for different timestepping schemes. The right boundary of the space discretization $[0, S_{\max}]$ domain is truncated at different values.

	$S_{\max} = 5 \times K$		$S_{\max} = 50 \times K$	
Timesteps	S, A grid nodes	Value	S, A grid nodes	Value
25	51	1.857193	54	1.857193
50	101	1.853254	109	1.853254
100	201	1.852120	217	1.852120
200	401	1.851781	433	1.851781
400	801	1.851660	865	1.851660

Table 8.1 shows results for a low volatility case, European Asian option (no jumps), using the semi-Lagrangian approach. In this special case, the two dimensional PDE can be reduced to one dimension (46), which we will refer to as the Ve?ce? PDE (46) in the following. Results obtained by solving the Ve?ce? PDE numerically are also given in Table 8.1.

In Table 8.1, we can see that the convergence ratio R for the semi-Lagrangian method is not quadratic ($R \neq 4$), while for the Ve?ce? PDE (46) quadratic convergence is found. As discussed in (46), the Ve?ce? PDE is not convection dominated, hence it is straightforward to obtain accurate numerical solutions. We remind the reader that this clever reduction to one dimension cannot be used for American options. The discontinuity present in the payoff greatly affects the convergence of the semi-Lagrangian method, since there is very little diffusion in the A direction, and the non-smoothness in the payoff is not smoothed out during the solution phase. Since we need to use quadratic interpolation in the A direction in order to determine the values of the solution at the feet of the characteristic curves, the interpolation may be affected by the non-smooth payoff, and may lower the observed rate of convergence.

In order to test the effect of the boundary condition (2.16) at $S = S_{max}$, we show results using two different values of S_{max} in Table 8.2. This table would seem to indicate that there is a negligible error for options of this maturity incurred setting $S_{max} = 50K$, and all subsequent results will be reported imposing condition (2.16) at $S_{max} = 50K$.

Figures 8.1 and 8.2 graphically present the solution V and the £rst derivative of the solution with respect to the stock price V_S when Crank-Nicolson is used. The plots are all smooth and do not exhibit any oscillations. While not shown here, V_{SS} also did not show any oscillations.

We now explore numerical convergence for pricing Asian options for large values of volatility σ. Table 8.3 presents our results. As expected quadratic convergence is recovered. In this case, a suf£cient amount of diffusion in the S direction appears to compensate for zero diffusion in the A direction.

8.1. An In Depth Study of the Convergence Ratio. The results of the previous section indicated that the semi-Lagrangian approach, coupled with Crank-Nicolson timestepping, results in quadratic convergence, for large volatilities. However, for small volatility values, quadratic convergence was not recovered. The goal of this subsection is to explore in detail different numerical techniques that could improve the convergence rate.

Table 8.4 contains the convergence rate results for different timestepping schemes for small volatility ($\sigma = .1$ and $r = .1$). For implicit timestepping linear convergence is recovered

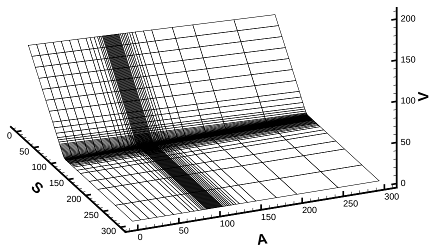

FIG. 8.1: *Value of a European £xed strike Asian put using Crank-Nicolson with constant timestepping* ($\Delta\tau = .01$)*.* 51 *grid points are used both in the A and S direction. The input parameters are* $\sigma = .1$, $r = .1$, $T = .25$, $K = 100$, *and* $\lambda = 0$.

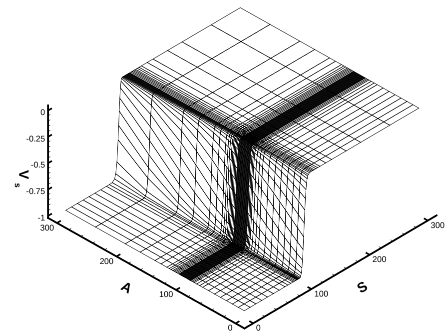

FIG. 8.2: *First derivative* (V_S) *value of a European £xed strike Asian put using Crank-Nicolson with constant timestepping* ($\Delta\tau = .01$)*.* 51 *grid points are used both in the A and S direction. The input parameters are* $\sigma = .1$, $r = .1$, $T = .25$, $K = 100$ *and* $\lambda = 0$.

Table 8.3 Value of a continuously observed £xed strike Asian call (no jumps) option with constant timesteps at $S = k$. The input parameters are $\sigma = .5$, $r = .05$, $T = .25$, $\lambda = 0$ and $K = 100$. We compare the results given using the Večeř 1-D PDE (46), and the semi-Lagrangian method presented here. Crank-Nicolson timestepping was used.

semi-Lagrangian				Večeř 1-D PDE (46)			
Size of S and A grids	No. of timesteps	$S = 100$ Value	R	Size of S grids	No. of timesteps	$S = 100$ Value	R
51	25	6.010203	n.a.	51	25	6.009821	n.a.
101	50	6.015092	n.a.	101	50	6.014848	n.a
201	100	6.016344	3.905	201	100	6.016251	3.582
401	200	6.016651	4.085	401	200	6.016619	3.816
801	400	6.016723	4.219	801	400	6.016713	3.915

($R = 2$), as expected. However for higher order timestepping schemes such as Crank-Nicolson and second order backward differencing, quadratic convergence is not found (see Table 8.4). These results are not surprising since the combination of small volatility with the non-smooth payoff, means that quadratic interpolation in the A direction is not $O((\Delta A)^3_{max})$, for small τ.

To try to remedy this problem, the initial payoff function is smoothed out. A classic method for handling discontinuities involves averaging the initial data. Speci£cally, values at each point are replaced with an average value over nearby space. Mathematically, we set

$$\text{PAYOFF}_{\text{smoothed}}(S_i, A_j) = \int_{K-\frac{\Delta A}{2}}^{K+\frac{\Delta A}{2}} \text{PAYOFF}(S_i, A)\, dA. \tag{8.2}$$

Table 8.4 Value of a continuously observed £xed strike Asian call option (no jumps) at the strike, constant timesteps. The input parameters are $\sigma = .1$, $r = .1$, $T = .25$, $\lambda = 0$, $K = 100$. Convergence ratios (8.1) are presented for different timestepping schemes: implicit, Crank-Nicolson and second order BDF.

		Implicit timestepping $S = 100$		CN timestepping $S = 100$		BDF timestepping $S = 100$	
Size of S and A grids	No. of timesteps	Value	R	Value	R	Value	R
51	25	1.911865	n.a.	1.857193	n.a	1.86096	n.a.
101	50	1.880801	n.a.	1.853254	n.a	1.854310	n.a.
201	100	1.865907	2.086	1.852120	3.475	1.852416	3.513
401	200	1.858681	2.061	1.851781	3.338	1.851868	3.453
801	400	1.855112	2.025	1.851660	2.815	1.851686	3.014

For a complete description of various smoothing methods the readers are referred to (42).

Table 8.5 Value of a continuously observed £xed strike call Asian call option (no jumps) at the strike with constant timesteps. The initial payoff is smoothed using the average scheme described by equation (8.2) The input parameters are $\sigma = .1$, $r = .1$, $T = .25$, $\lambda = 0$, and $K = 100$. Convergence ratios (8.1) are presented for different timestepping schemes: Crank-Nicolson and second order BDF.

		CN timestepping $S = 100$		BDF timestepping $S = 100$	
Size of S and A grids	No. of timesteps	Value	R	Value	R
51	25	1.870322	n.a.	1.874276	n.a.
101	50	1.856377	n.a.	1.857462	n.a.
201	100	1.852873	3.981	1.853179	3.925
401	200	1.851963	3.849	1.852053	3.803
801	400	1.851704	3.513	1.851731	3.497

Table 8.5 contains the convergence rate results. From a convergence point of view, the ratios have improved in comparison with the convergence ratio without smoothing (see Table 8.4). However, quadratic convergence is still not obtained. From a theoretical point of view, all the convergence analysis for semi-Lagrangian scheme is based on the smooth properties of the solution (10; 23). If the solution is smooth then quadratic convergence is recovered. However, if the solution is non-smooth, then we can expect some reduction in the convergence rate.

To con£rm our intuition that the non-smooth payoff is in fact the reason why quadratic convergence is not recovered, we create an arti£cial payoff that has the property of being quadratically smooth over the entire domain, e.g. $\text{PAYOFF}(A,K) = [\max(0,A-K)]^2$. In this case quadratic convergence is recovered for both Crank-Nicolson and second order backward differencing. Table 8.6 shows detailed convergence results for Crank-Nicolson timestepping.

Several other approaches were considered in an effort to improve convergence. We tried to use Rannacher timestepping (43); two or more implicit timesteps are taken before reverting to a higher order timestepping scheme such as Crank-Nicolson for example. Numerical experiments indicated that this did not improve the convergence rate. A convergence rate of

Table 8.6 Value of a continuously observed Asian call option (no jumps) at the strike with constant timesteps. The input parameters are $\Delta\tau = .01$, $\sigma = .1$, $r = .1$, $T = .25$, $\lambda = 0$ and $K = 1$. Convergence ratios (8.1) are presented for the Crank-Nicolson timestepping scheme.

		Call option (PAYOFF$(A,K) = [\max(0, A-K)]^2$)	
Size of	No. of	$S = 1$	
S and A grids	timesteps	Value	R
51	25	0.025749	n.a.
101	50	0.025622	n.a.
201	100	0.025591	4.131
401	200	0.025584	4.093
801	400	0.025582	4.054

approximately 3.5 is found in this case. Adaptive timestepping was also considered (24) but this technique did not improve the convergence rate.

8.2. Exotic Asian Options. It is not generally possible to achieve second order convergence for American options using constant timesteps. In (24) it was demonstrated that in order to achieve second order convergence, it is necessary to use variable timestepping for American options. However, some initial tests showed that due to the large convective term in the A direction, near $\tau = T$, the timestep selector suggested in (24) required very small timesteps near $\tau = T$. Consequently, we will show results in the following using constant timesteps.

Table 8.7 presents the input parameters. The mean of the jump distribution is denoted by μ and the jump distribution standard deviation is denoted by γ (see equation (2.7)). These parameters are roughly the same as those estimated by (3) using European call options on the S&P 500 stock index in April of 1999.

Table 8.7 Input data used to value American £xed strike Asian options under the lognormal jump diffusion process (2.7). These parameters are approximately the same as those reported in (3) using European call options on the S&P 500 stock index in April of 1999.

Parameter values			
σ	0.15	λ	0.10
r	0.05	T	0.25
γ	0.45	K	100.00
μ	-0.90	σ_{implied}	0.1886

To ensure consistent comparison between American Asian options with jumps and American Asian options without jumps, we proceed as follows:

1. Given the parameters in Table 8.7, we compute the analytical solution V_{jump} at the strike K of a vanilla put option, under jump diffusion.
2. Use a constant volatility Black-Scholes model with no jump to determine the implied volatility σ_{implied} which matches the jump diffusion value V_{jump} at the strike K.
3. Price the American Asian option with jumps using the parameters in Table 8.7.
4. Price the American Asian option with no jumps but with the implied volatility σ_{implied} estimated in Step 2.

Table 8.8 Value of a continuously observed £xed strike put American Asian option (under jump diffusion) with constant timestepping. Crank-Nicolson timestepping is used. The input parameters are de£ned in Table 8.7. This table presents convergence rates with and without jumps. Iterations refers to the total (over all timesteps) of the maximum number of iterations required for any value of j (see algorithm 7.1) at each timestep.

		No jump			Jumps		
Size of	No. of	No. of	($S = 100$)		No. of	($S = 100$)	
S and A grids	timesteps	iterations	Value	R	iterations	Value	R
51	25	77	2.220443	n.a.	99	2.044636	n.a.
101	50	160	2.195726	n.a.	167	2.018530	n.a
201	100	319	2.188555	3.447	340	2.012220	4.138
401	200	692	2.186717	3.903	716	2.010691	4.126
801	400	1397	2.186243	3.874	1609	2.010281	3.728

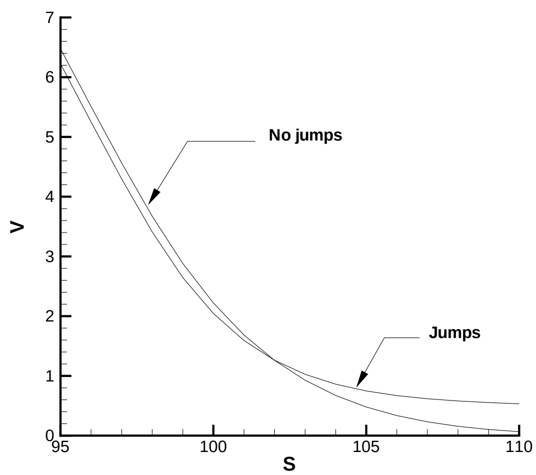

FIG. 8.3: *Comparison between the value of an American Asian £xed strike put option and the value of an American Asian £xed strike put option when the underlying stock follows the jump diffusion process described by (33). The input parameters are de£ned in Table 8.7.*

Table 8.8 compares the value of an American Asian £xed strike put option with the value of an American Asian £xed strike put option when the underlying stock follows the jump diffusion process described by (33). Second order backward timestepping is used and the initial payoff is smoothed out using equation (8.2). We observe that quadratic convergence is not recovered, the convergence ratios are ≈ 3.5. It is interesting to note that, at the strike, the price of an American Asian £xed strike put option with jumps is 9% cheaper than the price of the same option without jumps, while at $S = 1.05K$, the jump diffusion price is considerably higher than the no-jump price, as can be seen in Figure 8.3.

REMARK 8.1 (Alternative Boundary Condition). *A simpler method of imposing boundary condition (2.16), is to simply set $V_{SS} = 0$, and then discretize the V_S term using one sided*

£nite differences. As discussed in (48), this destroys the M matrix property of the discretized equations $I - \Delta\tau C - \lambda\Delta\tau B$. Most of the theoretical results in this paper require that this M matrix property hold, hence if we impose the boundary condition in this manner, these results cannot be proven to hold in this case. Nevertheless, we repeated all the computations reported above using this method of enforcing the boundary condition as $S \to \infty$. There was no change in the computed solution (at the strike) to eight digits.

9. Conclusion. In this paper we have put forward four primary contributions. First we have demonstrated that a semi-Lagrangian method can be used to price continuously observed American Asian options under jump diffusion processes. The implementation suggested here reduces this problem to solving a decoupled set of one dimensional discrete partial integro differential equations (PIDEs) at each timestep.

A second contribution is that since the discretized problem at each timestep reduces to a set of decoupled one dimensional PIDEs, we can make use of previous techniques developed by the authors to prove certain important properties of the discrete scheme, including convergence of the iterative method used to solve the implicit discrete equations.

In addition, we have included experimental computations which indicate that, even if second order timestepping methods are used, observed convergence as the mesh and timestep is re£ned occurs at a sub-second order rate. The problem can be traced to the non-smoothness of the payoff function.

Finally, we have also shown that in the fully implicit case, the semi-Lagrangian method for continuously observed Asian options is algebraically identical to a standard numerical method for pricing discretely observed Asian options, when the observation interval is equal to the discrete timestep.

Appendices

Appendix A. Discretization of the PDE.

In this appendix, we give the details of the discretization of the term $(\mathcal{C}V)_{i,j}$ in equation (3.15).

Using £nite differences, the matrix (5.2) is

$$(C \cdot V)_i = -V_i^{n+1}(\alpha_i + \beta_i + r + \lambda) + \beta_i V_{i+1}^{n+1} + \alpha_i V_{i-1}^{n+1} \ ; \ i = 2, \ldots, M-1 \tag{A.1}$$

where α_i, β_i depend on the type of approximations used for the derivatives and second derivatives. For $i = 1$, we impose condition (2.10) by setting $\alpha_1 = \beta_1 = \lambda = 0$, and for the row $i = M$, condition (2.16) is imposed by setting $\alpha_M = \beta_M = \lambda = 0$. There are a number of different discretizations of the derivative terms leading to various choices for α_i and β_i.

Discretizing the £rst derivative term of equation (3.2) with central differences leads to

$$\begin{aligned}
\alpha_{i,central} &= \frac{\sigma_i^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{(r - \lambda\kappa)S_i}{S_{i+1} - S_{i-1}} \\
\beta_{i,central} &= \frac{\sigma_i^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{(r - \lambda\kappa)S_i}{S_{i+1} - S_{i-1}}.
\end{aligned} \tag{A.2}$$

However if $\alpha_{i,central}$ or $\beta_{i,central}$ is negative, oscillations may appear in the solution. The oscillations can be avoided by using forward or backward differences at the problem nodes, leading to (forward difference)

$$\begin{aligned}\alpha_{i,forward} &= \frac{\sigma_i^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} \\ \beta_{i,forward} &= \frac{\sigma_i^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{(r - \lambda\kappa)S_i}{S_{i+1} - S_i},\end{aligned} \tag{A.3}$$

or, (backward difference)

$$\begin{aligned}\alpha_{i,backward} &= \frac{\sigma_i^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{(r - \lambda\kappa)S_i}{S_{i+1} - S_i} \\ \beta_{i,backward} &= \frac{\sigma_i^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})}.\end{aligned} \tag{A.4}$$

Algorithmically, we decide between a central or forward discretization at each node for equation (A.1) as follows:

Discretization

If $[\alpha_{i,central} \geq 0 \text{ and } \beta_{i,central} \geq 0]$ then
  $\alpha_i = \alpha_{i,central}$
  $\beta_i = \beta_{i,central}$
ElseIf $[\beta_{i,forward} \geq 0]$ then
  $\alpha_i = \alpha_{i,forward}$
  $\beta_i = \beta_{i,forward}$ (A.5)
Else
  $\alpha_i = \alpha_{i,backward}$
  $\beta_i = \beta_{i,backward}$
EndIf

Note that the test condition (A.5) guarantees that α_i and β_i are non-negative. For typical values of σ, r and grid spacing, forward differencing is rarely required for single factor options. In practice, since this occurs at only a small number of nodes remote from the region of interest, the limited use of a low order scheme does not result in poor convergence as the mesh is re£ned. For situations where the low order method causes excessive numerical diffusion, a ¤ux limiter can be used (50). As we have seen, requiring that all α_i and β_i are non-negative has important theoretical rami£cations.

Appendix B. Practical Monotonicity.

In this appendix, we will give a rationale for de£ning monotonicity using both conditions (5.14) and (5.15).

Writing out equation (3.12) gives

$$-\frac{DV}{D\tau} + \frac{\sigma^2 S^2}{2} V_{SS} + (r - \lambda\kappa) S V_S - rV + \rho \max(V^* - V, 0) + \mathcal{J}V = 0 \tag{B.1}$$

with $\mathcal{J}V = \lambda \int_0^\infty \{V(S\eta) - V(S)\} g(\eta)\, d\eta$, or

$$G(V, \frac{DV}{D\tau}, V_S, V_{SS}, \mathcal{J}V) = -\frac{DV}{D\tau} + \frac{\sigma^2 S^2}{2} V_{SS} + (r - \lambda\kappa) S V_S - rV + \rho \max(V^* - V, 0) + \mathcal{J}V = 0 \ . \tag{B.2}$$

Let $\zeta > 0$ be a constant, independent of (S, A, τ), then from equation (B.2) we have

$$\begin{aligned} G(V+\zeta, &\frac{D(V+\zeta)}{D\tau}, (V+\zeta)_S, (V+\zeta)_{SS}, \mathcal{J}(V+\zeta)) \\ &= G(V, \frac{DV}{D\tau}, V_S, V_{SS}, \mathcal{J}V) - r\zeta - \rho\left[\max(V^* - V, 0) - \max(V^* - V - \zeta, 0)\right] \\ &< G(V, \frac{DV}{D\tau}, V_S, V_{SS}, \mathcal{J}V) \quad \text{if} \quad \zeta, r > 0 \ . \end{aligned} \tag{B.3}$$

Recall our notation for the discretized form of equation (5.13). At each node $(S_i, A_j), \tau = \tau^{n+1}$, the discrete form of G is

$$\left[G(V, \frac{DV}{D\tau}, V_S, V_{SS}, \mathcal{J}V) \right]_{i,j}^{n+1} = g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) = 0 \ . \tag{B.4}$$

If $g_{i,j}$ is a consistent discretization of equation (B.2) then we must have $(\zeta = constant > 0)$

$$g_{i,j}(V_{i,j}^{n+1} + \zeta, \{V_{k,j}^{n+1} + \zeta\}_i, \{V^n + \zeta\}) < g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) \ . \tag{B.5}$$

Assuming that condition (5.14) holds (which is the usual monotonicity condition in the viscosity solution literature), then it follows that

$$g_{i,j}(V_{i,j}^{n+1} + \zeta, \{V_{k,j}^{n+1}\}_i, \{V^n\}) \le g_{i,j}(V_{i,j}^{n+1} + \zeta, \{V_{k,j}^{n+1} + \zeta\}_i, \{V^n + \zeta\}) \ . \tag{B.6}$$

Hence, equations (B.5-B.6) imply that

$$g_{i,j}(V_{i,j}^{n+1} + \zeta, \{V_{k,j}^{n+1}\}_i, \{V^n\}) < g_{i,j}(V_{i,j}^{n+1}, \{V_{k,j}^{n+1}\}_i, \{V^n\}) \ , \tag{B.7}$$

which is condition (5.15).

In summary, condition (5.14) and consistency of the discretization implies condition (5.15). However, as a stand-alone condition, equations (5.14-5.15) are more appealing, in that these conditions have a simple physical interpretation. This de£nition of monotonicity is also consistent with the de£nition of monotonicity in the CFD (Computational Fluid Dynamics) literature.

To see this, consider the case where g_{ij} in equation (5.13) is differentiable. Then we can restate De£nition 5.3 as

$$\begin{aligned} \frac{\partial g_{i,j}}{\partial V_{i,j}^{n+1}} &< 0 \\ \frac{\partial g_{i,j}}{\partial y} &\ge 0 \ ; \ y \in \{V_{k,j}^{n+1}\}_i \\ \frac{\partial g_{i,j}}{\partial z} &\ge 0 \ ; \ z \in \{V^n\} \end{aligned} \tag{B.8}$$

If g_{ij} satis£es conditions (B.8), then we have immediately that

$$\begin{aligned} \frac{\partial V_{i,j}^{n+1}}{\partial y} &\ge 0 \ ; \ y \in \{V_{k,j}^{n+1}\} \\ \frac{\partial V_{i,j}^{n+1}}{\partial z} &\ge 0 \ ; \ z \in \{V^n\}. \end{aligned} \tag{B.9}$$

In other words. a positive perturbation of any of $\{V_{k,j}^{n+1}\}_i$, $\{V_{k,l}^{n}\}$ results in a non-negative perturbation of $V_{i,j}^{n+1}$. This has the intuitive interpretation in terms of ¤uid or heat ¤ows, that is, discrete heat diffusion should always ¤ow from high temperature nodes to low temperature nodes. Hence De£nition (5.3) is commonly used (30; 26).

Another advantage of the de£nitions (5.14-5.15) which may not be obvious from condition (5.14) and consistency, is that a consequence of conditions (5.14-5.15) is a simple suf£cient test for l_∞ stability. Let

$$(V_{min}^n)_{i,j} \leq V_{i,j}^n \leq (V_{max}^n)_{i,j} \tag{B.10}$$

(these bounds need not be tight) and set

$$\max_{i,j}(V_{max}^n)_{i,j} = V_{max}^n \quad \text{and} \quad \min_{i,j}(V_{min}^n)_{i,j} = V_{min}^n \ . \tag{B.11}$$

Then if $g_{i,j}$ satis£es conditions (5.14-5.15) we have that

$$g_{i,j}((V_{max}^n)_{i,j}, V_{max}^{n+1}, V_{max}^n) = 0 \quad \text{and} \quad g_{i,j}((V_{min}^n)_{i,j}, V_{min}^{n+1}, V_{min}^n) = 0 \ . \tag{B.12}$$

Equation (B.12) holds for all (i,j), and so in particular, equation (B.12) holds for (p,q) where $V_{max}^{n+1} = (V_{max}^{n+1})_{p,q}$. Hence V_{max}^{n+1} satis£es

$$g_{p,q}(V_{max}^{n+1}, V_{max}^{n+1}, V_{max}^n) = 0 \ . \tag{B.13}$$

Similarly, at node (r,s), where $V_{min}^{n+1} = (V_{min}^{n+1})_{r,s}$, V_{min}^{n+1} satis£es

$$g_{r,s}(V_{min}^{n+1}, V_{min}^{n+1}, V_{min}^n) = 0 \ . \tag{B.14}$$

Hence by examining equations (B.13-B.14) for all p,q and for all r,s, we can obtain worst case error bounds for $V_{max}^{n+1}, V_{min}^{n+1}$ in terms of V_{max}^n, V_{min}^n. Equations (B.13-B.14) are usually easy to solve by inspection, since most terms will be zero for a consistent discretization.

Appendix C.

[1] A. ALLIEVI AND R. BERMEJO, *Finite element modi£ed method of characteristics for the Navier-Stokes equation*, International Journal for Numerical Methods in Fluids, 32 (2000), pp. 439–464.

[2] A. L. AMADORI, *The obstacle problem for nonlinear integro-differential equations arising in option pricing*. Working paper, Istituto pre le Applicazione del Calcolo "M. Picone", Rome, www.iac.rm.cnr.it/~amadori.

[3] L. ANDERSEN AND J. ANDREASEN, *Jump-diffusion processes: Volatility smile £tting and numerical methods for option pricing*, Review of Derivatives Research, 4 (2000), pp. 231–262.

[4] J. ANDREASEN, *The pricing of discretely sampled Asian and lookback options*, Journal of Computational Finance, 2 (Fall) (1998), pp. 5–30.

[5] G. BARLES, *Convergence of numerical schemes for degenerate parabolic equations arising in £nance*, in Numerical Methods in Finance, L. C. G. Rogers and D. Talay, eds., Cambridge University Press, Cambridge, 1997, pp. 1–21.

[6] G. BARLES AND P. SOUGANIDIS, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Analysis, 4 (1991), pp. 271–283.

[7] J. BARRAQUAND AND T. PUDET, *Pricing of American path-dependent contingent claims*, Mathematical Finance, 6 (1996), pp. 17–51.

[8] J. BECKER, *A second order backward difference method with variable timesteps for a parabolic problem*, BIT - Numerical Mathematics, 38 (1998), pp. 644–662.

[9] A. BELANGER, P. FORSYTH, AND K. VETZAL, *Valuing employee stock options with turnover*, (2004). Proceedings of the Northern Finance Association, St. Johns.
[10] R. BERMEJO, *Analysis of a class of quasi-monotone and conservative semi-Lagrangian advection schemes*, Numerische Mathematik, 87 (2001), pp. 597–623.
[11] P. P. BOYLE, *New life forms on the option landscape*, Journal of Financial Engineering, 2 (1996), pp. 217–252.
[12] P. P. BOYLE AND D. EMANUEL, *Options on the general mean*. Working paper, University of British Columbia, 1980.
[13] M. BRIANI, C. L. CHIOMA, AND R. NATALINI, *Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in £nancial theory*. Working paper, Istituto pre le Applicazione del Calcolo "M. Picone", Rome, www.iac.rm.cnr.it/˜natalini, To appear in Numerische Mathematik.
[14] R. CONT AND P. TANKOV, *Financial Modelling with Jump Processes*, Chapman & Hall, London, 2004.
[15] R. CONT AND E. VOLTCHKOVA, *A £nite difference scheme for option pricing in jump diffusion and exponential Levy models*. Internal Report 513, CMAP, Ecole Polytechnique, 2003.
[16] R. CONT AND E. VOLTCHKOVA, *A £nite difference scheme for option pricing in jump diffusion and exponential Levy models*, (2003). working paper, Ecole Polytechnique.
[17] M. DEMPSTER, J. HUTTON, AND D. RICHARDS, *LP valuation of exotic American options exploiting structure*, Journal of Computational Finance, 2 (Fall) (1998), pp. 61–84.
[18] J. DEWYNNE AND P. WILMOTT, *A note on average rate options with discrete sampling*, SIAM Journal on Applied Mathematics, 55 (1995), pp. 267–277.
[19] Y. D'HALLUIN, P. A. FORSYTH, AND G. LABAHN, *A Penalty method for American options with jump diffusion processes*, Numerische Mathematik, 97 (2004), pp. 321–352.
[20] Y. D'HALLUIN, P. A. FORSYTH, AND K. R. VETZAL, *Robust numerical methods for contingent claims under jump diffusion processes*. to appear in IMA Journal of Numerical Analysis, http://www.scicom.uwaterloo.ca/˜paforsyt/jump.pdf, 2004.
[21] J. DOUGLAS AND T. F. RUSSEL, *Numerical methods for convection-dominated diffusion problems based on combining the method of characteritics with £nite element or £nite difference procedures.*, SIAM Journal on Numerical Analysis, 19 (1982), pp. 871–885.
[22] B. ERAKER, M. JOHANNES, AND N. POLSON, *The impact of jumps in volatility and returns*, Journal of Finance, 58 (2003), pp. 1269–1300.
[23] M. FALCONE AND R. FERRETTI, *Convergence analysis for a class of high-order semi-Lagrangian advection schemes*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 909–940.
[24] P. A. FORSYTH AND K. R. VETZAL, *Quadratic convergence of a penalty method for valuing American options*, SIAM Journal on Scienti£c Computation, 23 (2002), pp. 2096–2123.
[25] P. A. FORSYTH, K. R. VETZVAL, AND R. ZVAN, *Convergence of numerical methods for valuing path-dependent option using interpolation*, Review of Derivatives Research, 5 (2002), pp. 273–314.
[26] J. FUHRMANN AND H. LANGMACH, *Stability and existence of solutions of time-implicit £nite volume schemes for viscous nonlinear conservation laws*, Applied Numerical Mathematics, 37 (2001), pp. 201–230.
[27] J. HUGGER, *Wellposedness of the boundary value formulation of a £xed strike Asian*

option. to appear in the Journal of Computational Methods in Sciences and Engineering, also, working paper, University of Copenhagen.

[28] J. HULL AND A. WHITE, *Ef£cient procedures for pricing European and American path-dependent options*, Journal of Derivatives, 1 (Fall) (1993), pp. 21–31.

[29] L. JIANG AND M. DAI, *Convergence of binomial tree methods for European/American path dependent options*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 1094–1109.

[30] R. LEVEQUE, *Numerical Methods for Conservation Laws*, Birkhauser, 1992. Lectures in Mathematics.

[31] A. L. LEWIS, *Fear of jumps*, Wilmott Magazine, (December) (2002), pp. 60–67.

[32] M. MARCOZZI, *On the valuation of Asian options by variational methods*, SIAM Journal on Scienti£c Computing, 24 (2003), pp. 1124–1140.

[33] R. C. MERTON, *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Engineering, 3 (1976), pp. 125–144.

[34] G. H. MEYER, *The numerical valuation of options with underlying jumps*, Acta Mathematica Universitatis Comenianae, 67 (1998), pp. 69–82.

[35] ———, *On pricing American and Asian options with PDE methods*, Acta Mathematica Universitatis Comenianae, 70 (2001), pp. 153–165.

[36] O. A. OLEINIK AND E. V. RADKEVIC, *Second Order Equations with Nonnegative Characteristic Form*, American Mathematical Society, Providence, R.I., 1973.

[37] C. OOSTERLEE, F. GASPAR, AND J. FRISCH, *Weno and blended BDF discretizations for Asian options*, Computing and Visualization in Science, 6 (2004), pp. 131–138.

[38] K. PARROTT AND N. CLARKE, *A parallel solution of early exercise Asian options with stochastic volatility*, (1998). Proceedings of the 11th Domain Decomposition Conference, Greenwich.

[39] H. PHAM, *Optimal stopping of controlled jump diffusion processes: a viscosity solution approach*, Journal of Mathematical Systems, Estimation and Control, 8 (1998), pp. 1–27.

[40] O. PIRONNEAU, *On the transport diffusion algorithm and its applications to the Navier-Stokes equations*, Numerische Mathematik, 38 (1982), pp. 309–332.

[41] D. POOLEY, P. FORSYTH, AND K. VETZAL, *Numerical convergence properties of option pricing PDEs with uncertain volatility*, IMA Journal of Numerical Analysis, 23 (2003), pp. 241–267.

[42] D. M. POOLEY, P. A. FORSYTH, AND K. R. VETZAL, *Convergence remedies for non-smooth payoffs in option pricing*, Journal of Computational Finance, 6 (Summer) (2003), pp. 25–40.

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

[44] P. RITCHKEN AND R. TREVOR, *Interest rate options with a volatility hump*, Review of Derivatives Research, 3 (1999), pp. 237–262.

[45] D. TAVELLA AND C. RANDALL, *Pricing Financial Instruments: The Finite Difference Method*, John Wiley & Sons, Inc, 2000.

[46] J. VEČEŘ, *A new PDE approach for pricing arithmetic average Asian options*, Journal of Computational Finance, 4 (Summer) (2001), pp. 105–113.

[47] P. WILMOTT, *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley & Sons Ltd., West Sussex, England, 1998.

[48] H. WINDCLIFF, P. FORSYTH, AND K. VETZAL, *Analysis of the stability of the linear boundary condition for the Black-Scholes equation*, (2003). To appear in Journal of Computational Finance.

[49] R. ZVAN, P. FORSYTH, AND K. VETZAL, *Penalty methods for American options with stochastic volatility*, Journal of Computational and Applied Mathematics, 91 (1998), pp. 199–218.

[50] ———, *A £nite volume approach for contingent claims valuation*, IMA Journal of Numerical Analysis, 21 (2001), pp. 703–731.

[51] R. ZVAN, P. A. FORSYTH, AND K. R. VETZAL, *Robust numerical methods for PDE models of Asian options*, Journal of Computational Finance, 1 (Winter) (1998), pp. 39–78.

[52] ———, *Discrete Asian barrier options*, Journal of Computational Finance, 3 (Fall) (1999), pp. 41–67.