A NUMERICAL SCHEME FOR PRICING AMERICAN OPTIONS WITH TRANSACTION COSTS UNDER A JUMP DIFFUSION PROCESS

DONNY CITRA LESMANA
Department of Mathematics, Bogor Agricultural University
Kampus IPB Darmaga, Bogor, Jawa Barat 16680, Indonesia

SONG WANG
Department of Mathematics & Statistics, Curtin University
GPO Box U1987, WA 6845, Australia

(Communicated by Chao Xu)

Abstract. In this paper we develop a numerical method for a nonlinear partial integro-differential complementarity problem arising from pricing American options with transaction costs when the underlying assets follow a jump diffusion process. We first approximate the complementarity problem by a nonlinear partial integro-differential equation (PIDE) using a penalty approach. The PIDE is then discretized by a combination of a spatial upwind finite differencing and a fully implicit time stepping scheme. We prove that the coefficient matrix of the system from this scheme is an $M$-matrix and that the approximate solution converges to the viscosity solution to the PIDE by showing that the scheme is consistent, monotone, and unconditionally stable. We also propose a Newton’s iterative method coupled with a Fast Fourier Transform for the computation of the discretized integral term for solving the fully discretized system. Numerical results will be presented to demonstrate the convergence rates and usefulness of this method.

1. Introduction. In a complete market without transaction costs, Black-Scholes [6] proposed a partial differential equation pricing model for European options, based on the assumptions that the underlying stock price follows a geometric Brownian motion with a log-normal diffusion and the volatility is constant. However, there is evidence that these assumptions are not consistent with the market price movements in practice, which is often called the volatility skew or smile [20]. A number of models have been proposed to remedy this problem. These improved models can be categorized into three classes: stochastic volatility model [16, 21], the deterministic volatility function model [12], and the jump diffusion model [10, 27, 42]. Among them, the jump diffusion model is very popular.

It has been shown that the price of a European option under a jump diffusion process without transaction costs satisfies a linear partial integro-differential equation (PIDE) and the price of its American counterpart satisfies a linear complementarity problem involving the partial integro-differential operator. Since exact solutions to

2010 Mathematics Subject Classification. Primary: 65K15, 65M06; Secondary: 91G60.
Key words and phrases. American option pricing, nonlinear partial integro-differential equation, nonlinear complementarity problem, upwind finite difference, penalty method, convergence.
such problems can hardly be found, numerical approximations are usually sought in practice in order to price such options. This is challenging as the PIDE and differential complementarity problems involve non-local integral terms. In [11, 42], the authors treat the integral term in the conventional jump-diffusion model explicitly and the remaining terms implicitly in their numerical approaches. However, the method is only conditionally stable. An operator splitting method coupled with a Fast Fourier Transform (FFT) for the evaluation of the integral term has been used in [2]. The method is unconditionally stable and 2nd-order accurate. There are various other methods for the linear PIDE such as [14, 39, 40].

In [28, 4] the authors showed that, in the presence of transaction costs, the price of a European option whose underlying asset price satisfies a jump diffusion process is governed by a nonlinear PIDE. Likewise, the price of an American option of this type is governed by a nonlinear complementarity problem. Although some theoretical properties of the non-linear model are established in [28, 4], to our best knowledge, no numerical methods for nonlinear partial integro-differential complementarity problem can be found in the open literature. In this work, we propose a numerical method for pricing American options under transaction costs and jump diffusion based on a penalty approach and some finite difference schemes for the time and spatial derivatives as well as the integral term. We show that the system matrix of the fully discretized equation is an M-matrix and establish a convergence theory for the numerical method. To evaluate the discretized integral term, we use the FFT technique. A Newton iterative method coupled with a regular operator splitting is used to avoid the inversion of the dense matrix arising from the non-local term. The rest of this work is organized as follows.

In the next section, we will give a brief account of the American option pricing model. In Section 3, we will present some discretization schemes for the PIDE. A convergence analysis is given in Section 4 in which we show that the numerical scheme is consistent, unconditionally stable and monotone. In Section 5, we propose a Newton iterative algorithm for solving the nonlinear algebraic system and show it is convergent. In Section 6, we present some numerical experimental results to demonstrate the rates of convergence and usefulness of the numerical method for solving a number of model problems.

2. The continuous model. Consider a European option with strike price $K$ and expiry time $T$ on an asset whose price $S$ follows the following stochastic differential equation

$$\frac{dS}{S} = (\nu - \lambda \kappa) \, dt + \sigma_0 \, dZ + (\eta - 1) \, dq,$$

(2.1)

where $dZ$ is an increment of the standard Gauss-Wiener process and $dq$ is the independent Poisson process with deterministic jump intensity $\lambda$. Here, $\sigma_0$ is the volatility, $\nu$ is the drift rate, $\eta - 1$ is an impulse function producing a jump from $S$ to $S\eta$, and $\kappa = E(\eta - 1)$ with $E$ the expectation operator. Following [37], it is easy to show using (2.1) that the value $V(S,t)$ of the European option at any $t \in [0,T)$ and $S > 0$ when transactions do not incur any costs satisfies the following linear PIDE:

$$V_{\tau}(S,\tau) = \frac{1}{2} \sigma_0^2 S^2 V_{SS}(S,\tau) + (r - \lambda \kappa)SV_S(S,\tau) - (r + \lambda) V + \lambda \int_0^\infty V(S\eta, \tau) g(\eta) \, d\eta$$

(2.2)
for \((S, \tau) \in [0, +\infty) \times [0, T)\), where \(r\) is the risk-free interest rate, \(\tau = T - t\), and \(g(\eta)\) is the probability density function of the jump amplitude \(\eta\), satisfying \(\int_0^\infty g(\eta) \, d\eta = 1\). Note that (2.2) is defined on an infinite domain. In practice, we are only interested in the computational domain \(0 < S < S_{\text{max}}\), where \(S_{\text{max}}\) is a positive number much greater than the strike price of the option. Also, various choices of \(g\) are available, but for brevity, we only consider, in this work, the following lognormal density function in Merton’s model:

\[
g(\eta) = \frac{1}{\sqrt{2\pi \sigma \rho}} \exp\left(-\frac{(\ln \eta - \mu)^2}{2\sigma^2}\right),
\]

(2.3)

where \(\mu\) and \(\sigma_j\) determine the mean and variance of the jumps. In this case, \(\kappa = E(\eta - 1) = \exp(\mu + \sigma_j^2/2) - 1\).

When transactions on selling and buying the asset involve costs, the value of a European option \(V(S, \tau)\) is governed by the following nonlinear PIDE:

\[
V_r(S, \tau) = \frac{1}{2} \sigma^2 (\tau, S, V_S, V_{SS}) S^2 V_{SS}(S, \tau) + (r - \lambda \kappa)SV_S(S, \tau) - (r + \lambda) V(S, \tau) + \lambda \int_0^\infty V(S\eta, \tau) g(\eta) \, d\eta,
\]

(2.4)

where \(\sigma\) is the modified volatility as a function of \(\tau, S, V_S\) and \(V_{SS}\). Assuming that the transaction cost is proportional to the value of the transaction and the portfolio is revised every \(\Delta t\) units of time, where \(\Delta t\) denotes a small, but non-infinitesimal, fixed time interval, Mocioalca [28] showed that the option price is the solution to (2.4) with the modified volatility models

\[
\sigma^2 = \sigma_0^2 + \frac{2}{\Delta t} \rho E \left( \left| \frac{\Delta S}{S} \right| \right) \quad \text{and} \quad \sigma^2 = \sigma_0^2 - \frac{2}{\Delta t} \rho E \left( \left| \frac{\Delta S}{S} \right| \right)
\]

for, respectively, long and short positions in the option, where \(\rho\) is the transaction cost parameter, measured as a fraction of the volume of transactions. It is shown in [22] that these can be written as

\[
\sigma^2(S, V_{SS}) = \sigma_0^2(1 + Le \, \text{sign}(V_{SS})),
\]

(2.5)

where \(Le\) is the Leland number given by

\[
Le = \sqrt{\frac{2}{\pi}} \left( \frac{\rho}{\sigma_0 \sqrt{\Delta t}} \right).
\]

Clearly, \(Le\) has to be such that \(\sigma^2(\text{V}_{SS}) > 0\), that is \(0 < Le < 1\).

In [4], the author assumes that the transaction costs behave like a non-increasing positive linear function, \(h(x) = a - bx\), where \(a, b \geq 0\) are two transaction cost parameters. Using this cost function, the author derived the following modified volatility

\[
\sigma^2(S, V_{SS}) = \sigma_0^2(1 - k_1 \text{sign}(V_{SS}) + k_2SV_{SS})
\]

(2.6)

with

\[
k_1 = \frac{a}{\sigma_0} \sqrt{\frac{2}{\pi \Delta t}} \geq 0, \quad k_2 = 2b \geq 0.
\]

We assume that \(k_1, k_2\) and \(V_{SS}\) satisfy the following inequalities

\[
1 - k_1 > 0 \quad \text{and} \quad 1 - k_1 \text{sign}(V_{SS}) + 2k_2SV_{SS} > 0 \quad \forall S \in (0, S_{\text{max}}).
\]

(2.7)

Under these conditions, One can easily see from (2.6) that \(\sigma^2(S, V_{SS}) > 0\). Note that \(V_{SS} \geq 0\) is usually satisfied by European and American vanilla options as used.
in [4] and thus the 2nd condition in (2.7) is usually not needed in practice unless $V_{SS}$ is negative.

If we define

$$L(V)(S, \tau) := V_{\tau} - \frac{1}{2}\sigma^2(S, \tau, V_S(S, \tau), V_{SS}(S, \tau)) S^2 V_{SS}(S, \tau)$$

$$- (r - \lambda \kappa)SV_S(S, \tau) + (r + \lambda) V(S, \tau) - \lambda \int_0^\infty V(S\eta, \tau)g(\eta) \, d\eta,$$  \hspace{1cm} (2.8)

then the American option pricing problem can be stated as the following nonlinear partial integro-differential complementarity problem:

$$LV \geq 0, \quad V - V^* \geq 0, \quad LV \cdot (V - V^*) = 0,$$  \hspace{1cm} (2.9)

where $V^*$ is a given non-negative function (usually the payoff function) defining a lower bound for the value of the option. Note that (2.9) contains (2.4) as the special case when $V^* = 0$.

To solve (2.4) or (2.9), boundary and payoff conditions need to be defined. There are various types of boundary and initial conditions and payoff functions depending on the type of contingent claims. For European vanilla call and put options, they are given by

$$V(S, 0) = g_1(S) := \begin{cases} \max(S - K, 0) & \text{for a call} \\ \max(K - S, 0) & \text{for a put} \end{cases}, \quad S \in (0, S_{max}),$$  \hspace{1cm} (2.10)

$$V(0, \tau) = g_2(\tau) := \begin{cases} 0 & \text{for a call} \\ Ke^{-r\tau} & \text{for a put} \end{cases}, \quad \tau \in (0, T],$$  \hspace{1cm} (2.11)

$$V(S_{max}, \tau) = g_3(\tau) := \begin{cases} S_{max} - Ke^{-r\tau} & \text{for a call} \\ 0 & \text{for a put} \end{cases}, \quad \tau \in (0, T],$$  \hspace{1cm} (2.12)

where $K$ denotes the strike price of the option and $S_{max} \gg K$ is a positive constant defining a computational upper bound on $S$. For American vanilla put options, they are given by

$$V(S, 0) = V^*(S) = g_1(S) := \max(K - S, 0),$$  \hspace{1cm} (2.13)

$$V(0, \tau) = g_2(\tau) := K,$$  \hspace{1cm} (2.14)

$$V(S_{max}, \tau) = g_3(\tau) := 0.$$  \hspace{1cm} (2.15)

Penalty methods have long been used in solving constrained optimization problems. In recent years, penalty approaches have been developed and used for solving nonlinear complementarity problems arising in financial engineering [34, 41, 7, 8, 24]. We apply the linear penalty method to (2.9) in which the complementarity problem is approximated by the following nonlinear PIDE:

$$LV_\vartheta - \vartheta [V^* - V_\vartheta]_+ = 0$$  \hspace{1cm} (2.16)

satisfying the initial and boundary conditions (2.10)–(2.12), where $\vartheta > 1$ is the penalty constant and $[z]_+ = \max(z, 0)$ for any $z$ and $\mathcal{L}$ is defined in (2.8).

The unique solvability and convergence of the solution of (2.16) to that of (2.9) when $\vartheta \to +\infty$ in both infinite and finite dimensions have been discussed extensively in the open literature. In [34, 43], the authors show the solution to a problem similar to (2.16) is unique and $V_\vartheta$ converges to the solution $V$ of (2.9) at the rate of order $O(\vartheta^{1/2})$ in a Sobolev norm. In [32, 17, 18, 30, 33, 31], the authors have developed convergence theories of the penalty approach for various finite-dimensional complementarity problems similar to the discretized form of (2.16) and the rates of convergence established in these works are of the order $O(\vartheta)$. Extensive numerical
results have also been reported in the aforementioned works to demonstrate the theoretical convergence rates of $V_\theta$ to $V$ as $\theta \to +\infty$. For brevity, we will omit these discussion, but refer the reader to these papers.

We comment that this penalty method, which can be regarded as a special case of the power penalty methods proposed in the references mentioned in the above paragraph, is different from the regularization technique appearing in the variational adjoint method such as that in [36]. The former is a technique for solving complementarity problems while the latter is usually for inverse problems.

3. Discretization of the nonlinear PIDE. We divide this discussion into two parts: the discretization of the integral term and that of the other terms in (2.16). For notational simplicity, we omit the subscript $\theta$.

3.1. Discretization of the integral. We follow the idea in [39] to approximate the integral in (2.16) or (2.4) numerically. First, we use a logarithmic transformation to transform the integral into a correlation integral so that it can be computed by the FFT algorithm. For clarity, we will suppress the time variable $\tau$ for most of the discussion in this subsection.

Let $x = \ln S$ and $y = \ln \eta$. Then the integral in (2.4) can be written as

$$Q(S, \cdot) := \int_0^\infty V(S\eta, \cdot)g(\eta) \, d\eta = \int_{-\infty}^\infty \bar{V}(x + y, \cdot)f(y) \, dy =: \bar{Q}(x, \cdot),$$  

(3.1)

where $f(y) = g(e^y)e^y$ and $\bar{V}(z, \tau) = V(e^z, \tau)$. Let the real number set $\mathbb{R}$ be partitioned uniformly into $x_i = i\Delta x$ for $i = 0, \pm 1, \pm 2, \ldots$ where $\Delta x > 0$ is a constant step length. Let $y_i = x_i$ for all $i = 0, \pm 1, \pm 2, \ldots$ and $\Delta y = \Delta x$. Note that the density function defined in (2.3) satisfies $g(\eta) \to 0$ exponentially as $|\eta| \to \infty$. So does $f(y)$. Therefore, we approximate the correlation integral in (3.1) at $x_i$ by the following finite sum

$$\bar{Q}(x_i) \approx \bar{Q}_i := \sum_{j=\frac{-J}{2}+1}^J \bar{V}_{i+j}f_j\Delta y$$  

(3.2)

for all feasible $i$, where $J \gg 0$ is a positive even integer, $\bar{V}_k$ denotes a nodal approximation of $V(k\Delta x, \cdot)$ for any $k$ and

$$f_j = \frac{1}{\Delta y} \int_{(j-\frac{1}{2})\Delta y}^{(j+\frac{1}{2})\Delta y} f(y) \, dy.$$

Eq.(3.2) defines nodal approximations to $\bar{Q}(x, \cdot)$ at the mesh nodes of the transformed region using nodal values of $\bar{V}$. However, as can be seen in the next subsection, we will discretize (2.16) in the original solution domain $(0, S_{\text{max}})$. Therefore, it is necessary to define nodal approximations to $V(S)$ using those of $\bar{V}(x)$ defined in (3.2) which requires interpolating $\bar{V}$ using nodal values of $V$ and $Q$ using nodal values of $Q$. Given a set of nodes $\{S_1, S_2, \ldots, S_n, \ldots\}$ on the $S$-axis satisfying $0 = S_1 < S_2 < \ldots < S_n < \ldots$, let $\bar{V}_i$ be a nodal approximation to $V(S_i)$ for $i = 1, 2, \ldots$. We use the following steps to approximate $Q(S_i)$.

1. Approximation of $\{\bar{V}_j\}$ by the linear interpolation of $\{V_i\}$.

For any integer $j$, let $p(j)$ be the index such that $S_{p(j)} \leq e^{x_j} \leq S_{p(j)+1}$. Then, we define the following approximation to $\bar{V}_j$

$$\bar{V}_j = \phi_{p(j)} V_{p(j)} + (1 - \phi_{p(j)}) V_{p(j)+1},$$  

(3.3)
where
\[
\phi_{p(j)} = \frac{e^{x_j} - S_{p(j)}}{S_{p(j)+1} - S_{p(j)}}.
\]

2. Approximation of \(\{Q_i\}\) by the piecewise linear interpolation of \(\{\bar{Q}_j\}\).

For any \(i = 1, 2, \ldots\), let \(q(i)\) be the integer satisfying \(x_{q(i)} \leq \ln S_i \leq x_{q(i)+1}\). Then, \(Q_i\) is defined as
\[
Q_i = \psi_{q(i)} \bar{Q}_{q(i)} + (1 - \psi_{q(i)}) \bar{Q}_{q(i)+1},
\]
where
\[
\psi_{q(i)} = \frac{\ln S_i - x_{q(i)}}{x_{q(i)+1} - x_{q(i)}}.
\]

From the definition, it is clear that \(0 \leq \phi_{p(j)} \leq 1\) and \(0 \leq \psi_{q(i)} \leq 1\). Using (3.2), (3.3) and (3.4), it is easy to verify that \(Q_i\) is defined by
\[
Q_i = \sum_{j=-\frac{N}{2}+1}^{\frac{N}{2}} \Pi^j_i(V) f_j \Delta y,
\]
where
\[
\Pi^j_i(V) = \psi_{q(i)} [\phi_{p(q(i)+j)} V_{p(q(i)+j)} + (1 - \phi_{p(q(i)+j)}) V_{p(q(i)+j)+1}] + (1 - \psi_{q(i)}) \times [\phi_{p(q(i)+1+j)} V_{p(q(i)+1+j)} + (1 - \phi_{p(q(i)+1+j)}) V_{p(q(i)+1+j)+1}].
\]

The operator \(\Pi^j_i(V)\) is linear in \(V\). This approximation will be used in the spatial discretization of the problem as given below.

3.2. Full discretization. To discretize (2.16) with the nonlinear volatility (2.5) or (2.6) and boundary and initial conditions (2.10)–(2.12), we first define a mesh for \((0, S_{\text{max}}) \times (0, T)\). For a given positive even integer \(M\), let \((0, S_{\text{max}})\) be divided uniformly into \(M\) sub-intervals with mesh nodes
\[
S_i = (i - 1)h, \quad i = 1, \ldots, M + 1,
\]
where \(h = S_{\text{max}}/M\).

Similarly, we divide \((0, T)\) into \(N\) sub-intervals with mesh nodes \(\{\tau_n\}_{n=1}^{N+1}\) satisfying \(0 = \tau_1 < \tau_2 < \ldots < \tau_{N+1} = T\). We put \(\Delta \tau_n = \tau_{n+1} - \tau_n\).

Given any matrices \(W^n = (W^n_1, W^n_2, \ldots, W^n_{M+1})^\top\) and \(W_i = (W_i^1, W_i^2, \ldots, W_i^{N+1})^\top\) for \(i = 1, \ldots, M + 1\) and \(n = 1, \ldots, N + 1\), we define the following finite difference operators on the mesh defined above:
\[
(\delta_t W_i)(n) = \frac{W_i^{n+1} - W_i^n}{\Delta \tau_n}, \quad (\delta_t^+ W^n)(i) = \frac{W_i^{n+1} - W_i^n}{h},
\]
\[
(\delta_S W^n)(i) = \frac{W_i^n - W_{i-1}^n}{h}, \quad (\delta_S S^n)(i) = \frac{W_i^n - 2W_{i-1}^n + W_{i+1}^n}{h^2}.
\]

Using these operators and \(Q_i\) defined in (3.5) for a sufficiently large even integer \(J\), we approximate (2.4) by the following finite difference equation:
\[
(\delta_t V_i)(n) - \frac{1}{2} \sigma^2 (\delta_S S^n V_i^{n+1})(i) S_i^2 (\delta_S S^n V_i^{n+1})(i) - \frac{1}{2} \sigma^2 (\delta_S S^n V_i^{n+1})(i) \right]
\[
- \left( \frac{1}{2} \sigma^2 \right) \bar{r} S_i (\delta_S S^n V_i^{n+1})(i) - \left( \frac{1}{2} \sigma^2 \right) S_i (\delta_S S^n V_i^{n+1})(i) - \left( \frac{1}{2} \sigma^2 \right) \bar{r} S_i (\delta_S S^n V_i^{n+1})(i) - \left( \frac{1}{2} \sigma^2 \right) S_i (\delta_S S^n V_i^{n+1})(i)
\]
\[
+ (r + \lambda) V_i^{n+1} - \lambda Q_i - \theta [V_i^n - V_i^{n+1}]_+ = 0
\]
(3.7)
for \(i = 2, \ldots, M\) and \(n = 1, \ldots, N\), where
\[
\bar{r} = r - \lambda k, \quad V^{n+1} = (V^{n+1}_1, V^{n+1}_2, \ldots, V^{n+1}_M)^\top, \quad V_i = (V^1_i, V^2_i, \ldots, V^{n+1}_i)^\top,
\]
and \(V^n_i\) denotes an approximation to \(V(S_i, \tau_n)\) to be determined for any feasible index pair \((i, n)\). The discretization of \(V_S\) is based on the so-called upwind finite difference scheme which has been used for solving various problems in classic and financial engineering [25]. For a detailed discussion of the above discretization scheme and its convergence for the nonlinear Black-Scholes equation without any jump, we refer the reader to [23]. We comment that there are various other spatial discretization schemes for solving Black-Scholes equations such as [3, 9, 13, 15, 19, 35] and any of these methods can be used for the discretization of the differential operators in (2.16). Using (2.10)–(2.12), we define the initial and boundary conditions for the above system as follows
\[
V^1_i = g_1(S_i), \quad i = 1, \ldots, M + 1, \quad (3.8)
\]
\[
V^n_1 = g_2(\tau_n), \quad n = 1, \ldots, N + 1, \quad (3.9)
\]
\[
V^n_{M+1} = g_3(\tau_n), \quad n = 1, \ldots, N + 1. \quad (3.10)
\]
The system (3.7) can be rewritten as
\[
\alpha_i^{n+1}(V^{n+1})V^n_{i-1} + \beta_i^{n+1}(V^{n+1})V^n_i + \gamma_i^{n+1}V^n_{i+1} - \lambda \Delta y \sum_{l=-J/2+1}^{J/2} \Pi^I_l(V^{n+1})f_l = \frac{1}{\Delta \tau_n} V^n_i + \frac{\vartheta}{2}(1 + \text{sign}(V^n_i - V^{n+1}_i))V^n_i \quad (3.11)
\]
where \(J\) is a large even positive number as used in (3.5),
\[
\alpha_i^{n+1}(V^{n+1}) = \frac{1}{2} \frac{\sigma^2((\delta SS V^{n+1})(i))S^2_i}{h^2} + \left(1 - \text{sign}(\bar{r})\right) \bar{r} S_i, \quad (3.12)
\]
\[
\beta_i^{n+1}(V^{n+1}) = \frac{1}{\Delta \tau_n} \frac{\sigma^2((\delta SS V^{n+1})(i))S^2_i}{h^2} + \frac{|\bar{r}| S_i}{h} + r + \lambda + \frac{\vartheta}{2}(1 + \text{sign}(V^n_i - V^{n+1}_i)), \quad (3.13)
\]
\[
\gamma_i^{n+1}(V^{n+1}) = -\frac{1}{2} \frac{\sigma^2((\delta SS V^{n+1})(i))S^2_i}{h^2} - \left(1 + \text{sign}(\bar{r})\right) \bar{r} S_i. \quad (3.14)
\]
The nonlinear finite-difference system (3.11) along with (3.8)–(3.10) form a recursive system that determines the unknown vector \(\hat{V}^{n+1} = (V^{n+1}_2, \ldots, V^{n+1}_M)^\top\) for \(n = 1, \ldots, N\). From (3.5) and (3.6) we see that the sum on the LHS of (3.11) is linear in \(V^{n+1}\). Let \(D = (d_{ij})\) be the \((M - 1) \times (M - 1)\) matrix such that
\[
[D \hat{V}^n]_i = \sum_{j=2}^{M} d_{ij} V^n_j = \sum_{l} \Pi^I_l(V^n) f_l \Delta y. \quad (3.15)
\]
Then, (3.11) can be written in a matrix form as follows
\[
[A^{n+1}(V^{n+1}) - \lambda D] \hat{V}^{n+1} = \frac{1}{\Delta \tau_n} \hat{V}^n + \frac{\vartheta}{2}(I + \text{diag}(\hat{V}^* - \hat{V}^{n+1}))\hat{V}^* + B^{n+1}, \quad (3.16)
\]
where $I$ denotes the identity matrix, $\hat{V}^* = (V^*_2, ..., V^*_M)^\top$ and $A^{n+1}$ and $B^{n+1}$ are given by

$$A^{n+1}(V^{n+1}) = \begin{pmatrix}
\beta_2^{n+1} & \gamma_2^{n+1} & 0 & \cdots & 0 & 0 & 0 \\
\alpha_3^{n+1} & \beta_3^{n+1} & \gamma_3^{n+1} & \cdots & 0 & 0 & 0 \\
0 & \alpha_4^{n+1} & \beta_4^{n+1} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \beta_{M-2}^{n+1} & \gamma_{M-2}^{n+1} & 0 \\
0 & 0 & 0 & \cdots & \alpha_{M-1}^{n+1} & \beta_{M-1}^{n+1} & \gamma_{M-1}^{n+1} \\
0 & 0 & 0 & \cdots & 0 & \alpha_M^{n+1} & \beta_M^{n+1}
\end{pmatrix},$$

$$B^{n+1} = \begin{pmatrix}
-\alpha_2^{n+1}(V^{n+1})V_1^{n+1} + \lambda\Delta y(f_{-1}V_1^{n+1} + f_{M-1}V_M^{n+1}) \\
\lambda\Delta y(f_{-2}V_1^{n+1} + f_{M-2}V_M^{n+1}) \\
\vdots \\
\lambda\Delta y(f_{M-1}V_1^{n+1} + f_{1}V_M^{n+1}) \\
-\gamma_{n+1}^{n+1}(V^{n+1})V_1^{n+1} + \lambda\Delta y(f_{-1}V_1^{n+1} + f_{1}V_M^{n+1})
\end{pmatrix}.$$  

Clearly, $B^{n+1}$ contains contributions from the boundary conditions in (3.9)–(3.10). For the system matrix of (3.16) we have the following results.

**Theorem 3.1.** For any $n = 2, ..., N + 1$, both $A^n(V^n)$ and $A^n(V^n) - \lambda D$ are $M$-matrices for any given $V^n$.

**Proof.** To prove $A^n$ defined above is an $M$-matrix, it suffices to show that

$$\begin{align*}
\alpha_i^n < 0, & \quad \beta_i^n > 0, \quad \gamma_i^n < 0, \quad (3.17) \\
\beta_i^n > |\alpha_i^n| + |\gamma_i^n| & \quad (3.18)
\end{align*}$$

for $i = 2, ..., M$.

From (3.12)–(3.14) we see that (3.17) is obviously true and that

$$\beta_i^n = |\alpha_i^n| + |\gamma_i^n| + \frac{1}{\Delta t^n} \left( r + \lambda + \frac{y}{2} (1 + \text{sign}(V_i^* - V_i^n)) \right) > |\alpha_i^n| + |\gamma_i^n|,$$

since $\frac{1}{\Delta t^n} > 0$ and the other terms are nonnegative. Therefore, (3.18) is also satisfied. Also, it is obvious that $A^n$ is irreducible and thus it is irreducibly diagonally dominant with positive diagonal and non-positive off-diagonal entries. By [29], $A^n$ is an $M$-matrix for any given $V^n$.

Now we investigate $d_{ij}$ defined in by (3.15). From (3.5) and (3.6) we see that when $V_i = 1$ for all $i = 1, 2, ..., M + 1$,

$$\sum_{j=2}^M d_{ij} = \sum_{l} f_l \Delta y. \quad (3.19)$$

Since $f$ is a probability density function, we have

$$\int_{-\infty}^{\infty} f(y) \, dy = \int_{0}^{\infty} g(\eta) \, d\eta = 1, \quad f(y) = g(e^y)e^y \geq 0. \quad (3.20)$$

Therefore, it follows from (3.20) that

$$\sum_i f_i \Delta y \leq 1 \quad \text{and} \quad f_i \geq 0.$$
From (3.19), we have

\[ 0 \leq d_{ij} \leq 1 \quad \text{and} \quad 0 \leq \sum_j d_{ij} \leq 1. \]

Using these inequalities we can conclude that all the off-diagonal entries of \( A^n(V^n) - \lambda D \) are non-positive and

\[
\alpha_i^n + \beta_i^n + \gamma_i^n - \lambda \sum_{j=2}^M d_{ij} \geq \frac{1}{\Delta r_{n-1}} + \lambda - \lambda \sum_{j=2}^M d_{ij} + \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n-1}))
\]

\[
= \frac{1}{\Delta r_{n-1}} + \lambda \left( 1 - \sum_{j=2}^M d_{ij} \right) + \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n-1}))
\]

\[
\geq \frac{1}{\Delta r_{n-1}} + \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n-1})) > 0.
\]  

This implies

\[
\beta_i^n - \lambda d_{ii} \geq \frac{1}{\Delta r_{n-1}} + \left( 1 - \sum_j d_{ij} \right) \lambda + |\alpha_i^n| + |\gamma_i^n| + \lambda \sum_{j \neq i} d_{ij}
\]

\[
= \frac{1}{\Delta r_{n-1}} + (1 - d_{ii}) \lambda + |\alpha_i^n| + |\gamma_i^n| > |\alpha_i^n| + |\gamma_i^n|. \tag{3.22}
\]

Therefore, all the diagonal entries of \( A^n(V^n) - \lambda D \) are positive, off-diagonal entries are non-positive, and the matrix is strictly diagonally dominant. Hence, it is an \( M \)-matrix by [29].

Using Theorem 3.1 and the fact that the initial and boundary conditions in (3.8)–(3.10) are all non-negative, it is trivial to show the following corollary.

**Corollary 3.1.** For any \( n = 1, 2, ..., N \), the solution \( V^{n+1} \) to (3.11) is non-negative.

4. **Convergence of the numerical scheme.** In [4], the author shows the existence and uniqueness of the viscosity solution to (2.4). In this section we will prove that the solution to (3.16) converges to the viscosity solution to (2.4) as the mesh sizes approach zero. It has been shown in [5] that the convergence of the fully discretized system (3.16) to the viscosity solution of a full nonlinear 2nd-order PDE is guaranteed if the discretization is consistent, stable and monotone. Thus, in the rest of this section we will prove the convergence of our numerical scheme by showing that it satisfies these properties. For \( i = 2, 3, ..., M \) and \( n = 1, ..., N \), introduce a functional \( H_i^{n+1} \) defined by

\[
H_i^{n+1}(V_i^{n+1}, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_j^n, V_i^n) := \eta_i V_i^{n+1} + \xi_i V_{i+1}^{n+1} + \zeta_i V_{i-1}^{n+1} - \lambda \sum_{j \neq i} d_{ij} V_j
\]

\[
- \frac{1}{\Delta r_n} V_i^n - \frac{1}{2} S_i^2 \sigma^2(S_i, (\delta_{SS} V^{n+1})(i))(\delta_{SS} V^{n+1})(i) - \vartheta [V_i^n - V_i^{n+1}]_+ \tag{4.1}
\]

where

\[
\eta_i = \frac{(1 - \text{sign}(\bar{r}))}{2} \bar{r} S_i, \quad \xi_i = \frac{\bar{r}}{\Delta r_n} + \frac{|\bar{r}| S_i}{\bar{h}} + r + (1 - d_{ii}) \lambda, \quad \zeta_i = - \frac{(1 + \text{sign}(\bar{r}))}{2} \bar{r} S_i.
\]
Then, it is easy to see that (3.11) (or (3.7)) becomes
\[
H_{i}^{n+1}(V_{i}^{n+1}, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_{j}^{n+1}, V_{i}^{n}) = 0
\]
for all feasible \( i \) and \( n \). For this scheme, we have the following lemma.

**Lemma 4.1 (Monotonicity).** The discretization (3.11) with \( \sigma \) defined in (2.5) is monotone. Furthermore, (3.11) with \( \sigma \) defined in (2.6) is also monotone for all \( V^{n} \) such that (2.7) is satisfied with \( V_{SS} \) replaced with \( (\delta_{SS}V^{n+1}). \)

**Proof.** To prove this lemma, we show that, for any \( \varepsilon > 0 \) and \( i = 2, 3, \ldots, M, \)
\[
H_{i}^{n+1}(V_{i}^{n+1}, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_{j}^{n+1}, V_{i}^{n}) \leq H_{i}^{n+1}(V_{i}^{n+1} + \varepsilon, V_{i+1}^{n+1} + \varepsilon, V_{i-1}^{n+1} + \varepsilon, V_{j}^{n+1} + \varepsilon)
\]
(4.2)
\[
H_{i}^{n+1}(V_{i}^{n+1} + \varepsilon, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_{j}^{n+1}, V_{i}^{n}) \geq H_{i}^{n+1}(V_{i}^{n+1}, V_{i+1}^{n+1}, V_{i-1}^{n+1}, V_{j}^{n+1}, V_{i}^{n}).
\]
(4.3)

Since \( \eta_{i} \leq 0, \xi_{i} > 0, \zeta_{i} \leq 0, \leq \sum_{j \neq i} d_{ij} \leq 1, \lambda \geq 0, \) and \( \frac{1}{\Delta n} > 0, \) the first five terms on the RHS of (4.1) are respectively non-increasing in \( V_{i}^{n+1}, \) increasing in \( V_{i}^{n+1}, \) non-increasing in \( V_{i}^{n+1}, \) non-increasing in \( V_{i}^{n+1}, \) and decreasing in \( V_{i}^{n}. \)

Let \( E_{k} = (0, 0, \ldots, 1, 0, 0, 0)^{T} \) be the \((M - 1) \times 1 \) column vector. From the definition of \( \delta_{SS} \) and (2.6) we have
\[
(\delta_{SS}(V^{n+1} + \varepsilon E_{n+1} + \varepsilon E_{i+1}))(i) = \frac{(V_{i}^{n+1} + \varepsilon - 2V_{i}^{n+1} + (V_{i}^{n+1} + \varepsilon)}{h^{2}}
\]
(4.4)
\[
(\delta_{SS}(V^{n+1} + \varepsilon E_{i}))(i) = \frac{V_{i}^{n+1} - 2(V_{i}^{n+1} + \varepsilon) + V_{i}^{n+1}}{h^{2}}
\]
(4.5)

Let us consider the first nonlinear term on the RHS of (4.1). We discuss this term in the following two cases corresponding to the choices of \( \sigma \) defined in (2.5) and (2.6).

When \( \sigma \) is defined in (2.5), we have, for any \( S, z_{1} \) and \( z_{2}, \)
\[
\frac{1}{\sigma_{0}}[\sigma^{2}(S, z_{1})z_{1} - \sigma^{2}(S, z_{2})z_{2}] = \left[(z_{1} - z_{2}) + Le(\text{sign}(z_{1})z_{1} - \text{sign}(z_{2})z_{2})\right]
\]
(4.6)
\[
= \left[1 + Le\text{sign}(z_{1})\right](z_{1} - z_{2}) + Le[\text{sign}(z_{1}) - \text{sign}(z_{2})]z_{2}.
\]
(4.7)
\[
= \begin{cases}
C_{1}(z_{1} - z_{2}) & z_{1} \cdot z_{2} > 0, \\
(1 + Le)z_{1} + (Le - 1)z_{2} & z_{1} > 0 > z_{2}, \\
(1 - Le)(z_{1} - z_{2}) - 2Lez_{2} & z_{1} < 0 < z_{2},
\end{cases}
\]
where \( C_{1} = (1 + Le \times \text{sign}(z_{1})) > 0, \) since \( 0 \leq Le < 1. \)

Therefore, \( \sigma^{2}(S, z)z \) defined in (2.5) is an increasing function of \( z. \)

When \( \sigma^{2}(S, z)z \) is by (2.6), we have
\[
\frac{d}{dz}[(1 - k_{1}\text{sign}(z) + k_{2}Sz)z] = (1 - k_{1}\text{sign}(z)) + 2k_{2}Sz > 0
\]
when \( z \) satisfies (2.7). Therefore, \( \sigma^{2}(S, z)z \) is an increasing function in \( z \) on the set defined by the 2nd inequality in (2.7).
Combining the monotonicity of \( \sigma^2(S, z)z \) and the properties of the linear terms on the RHS of (4.1) and (4.4) we have

\[
H_i^{n+1}(V_i^{n+1}, V_{i+1}^{n+1} + \varepsilon, V_{i-1}^{n+1} + \varepsilon, V_j^{n+1} + \varepsilon, V_i^n + \varepsilon)
\]

\[
= \eta_i(V_i^{n+1} + \varepsilon) + \xi_i V_i^{n+1} + \zeta_i (V_{i+1}^{n+1} + \varepsilon) - \lambda \sum_{j \neq i} d_{ij} (V_j^{n+1} + \varepsilon) - \frac{1}{\Delta \tau_n} (V_i^n + \varepsilon)
\]

\[
- \frac{1}{2} S_i^2 \sigma^2 \left( S_i, (\delta_{SS} V^{n+1})(i) + \frac{2\varepsilon}{\Delta^2} \right) \left( (\delta_{SS} V^{n+1})(i) + \frac{2\varepsilon}{\Delta^2} \right) - \vartheta |V_i^n - V_i^{n+1}|_+
\]

\[
\leq H_i^{n+1}(V_i^{n+1}, V_{i+1}^{n+1}, V_i^{n+1}, V_j^{n+1}, V_i^n).
\]

This is (4.2).

Similarly, using the monotonicity of \( \sigma^2(S, z)z \) and (4.5) and noticing that \([V_i^n - (V_i^{n+1} + \varepsilon)]_+ \leq [V_i^n - V_i^{n+1}]_+\) for any \( \varepsilon \geq 0 \), it is easy to show that (4.3) also holds true. Hence, the discretization scheme is monotone. \( \square \)

The stability of the method is given in the following lemma.

**Lemma 4.2 (Stability).** For \( n = 1, 2, \ldots, M \), let \( V^{n+1} = (V_i^{n+1}, \hat{V}^{n+1}, V_M^{n+1})^T \), where \( \hat{V}^{n+1} \) is the solution to (3.16). Then, \( V^{n+1} \) satisfies

\[
\|V^{n+1}\|_\infty \leq C \max \{ \|g_1\|_\infty, \|g_2\|_\infty, \|g_3\|_\infty \} + \vartheta \Delta \tau \|V^*\|_\infty,
\]

where \( \Delta \tau = \max_n \Delta \tau_n \), \( g_1, g_2 \) and \( g_3 \) are the initial and boundary conditions defined in (3.8)-(3.10) and \( \| \cdot \|_\infty \) denotes either the discrete or continuous \( L_\infty \)-norm.

**Proof.** For any \( n = 1, ..., N \), from (3.11) and (3.22) we have

\[
(\beta_i^{n+1} - \lambda d_{ii}) V_i^{n+1} = -\alpha_i^{n+1} V_i^{n+1} - \gamma_i^{n+1} V_{i+1}^{n+1} + \lambda \Sigma_{j \neq i} d_{ij} V_j^{n+1} + \frac{1}{\Delta \tau_n} V_i^n
\]

\[
+ \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n+1})) V_i^n
\]

for \( i = 2, ..., M \). Recall that \( \alpha_i^{n+1}, \gamma_i^{n+1} \leq 0, \beta_i^{n+1} > 0, 0 \leq d_{ij} \leq 1, \) and \( 0 \leq \Sigma_j d_{ij} \leq 1 \). From the above we get

\[
(\beta_i^{n+1} - \lambda d_{ii}) |V_i^{n+1}| \leq -\alpha_i^{n+1} |V_i^{n+1}| - \gamma_i^{n+1} |V_{i+1}^{n+1}| + \lambda \Sigma_{j \neq i} d_{ij} |V_j^{n+1}| + \frac{1}{\Delta \tau_n} |V_i^n|
\]

\[
+ \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n+1})) V_i^n
\]

\[
\leq -\alpha_i^{n+1} |V^{n+1}| - \gamma_i^{n+1} |V^{n+1}| + \lambda \Sigma_{j \neq i} d_{ij} |V^{n+1}| + \frac{1}{\Delta \tau_n} |V^n| + \frac{\vartheta}{2} (1 + \text{sign}(V_i^n - V_i^{n+1})) V_i^n
\]

for \( i = 2, ..., M \). We now consider the following two cases.

**Case I.** \( |V^{n+1}|_\infty = |V_k^{n+1}|_\infty \) for an index \( k \in \{2, 3, ..., M\} \).

In this case, the above estimate with \( i = k \) becomes

\[
(\alpha_k^{n+1} + \beta_k^{n+1} + \gamma_k^{n+1} - \lambda \Sigma_j d_{kj}) |V^{n+1}|_\infty \leq \frac{|V^n|_\infty + \frac{\vartheta}{2} (1 + \text{sign}(V_k^n - V_k^{n+1})) V_k^n}{\Delta \tau_n}
\]
Using (3.21) we have from this inequality
\[
\left( \frac{1}{\Delta \tau_n} + \frac{\vartheta}{2} \right) \left(1 + \text{sign}(V_k^* - V_k^{n+1})\right) \|V^{n+1}\|_\infty \leq \frac{\|V^n\|_\infty + \vartheta}{\Delta \tau_n} (1 + \text{sign}(V_k^* - V_k^{n+1})) V_k^*.
\]
It is easy to verify that the above system can be written as
\[
\|V^{n+1}\|_\infty \leq \frac{1 + p_n}{1 + p_n} \|V^n\|_\infty + \frac{p_n}{1 + p_n} \|V^*\|_\infty,
\]
where \( p_n = \vartheta \Delta \tau_n \) or 0. Repeatedly using this relationship, we obtain
\[
\|V^{n+1}\|_\infty = \frac{\|V^n\|_\infty}{(1 + p_n)(1 + p_{n-1})} + \left( \frac{p_n}{1 + p_n} + \frac{p_n p_{n-1}}{(1 + p_n)(1 + p_{n-1})} \right) \|V^*\|_\infty
\]
\[
\vdots
\]
\[
= \|V^1\|_\infty \prod_{i=1}^{n} \frac{1}{1 + p_i} + \|V^*\|_\infty \sum_{i=1}^{n} \left( \prod_{j=i}^{n} \frac{p_j}{1 + p_j} \right)
\]
\[
\leq \|V^1\|_\infty + \|V^*\|_\infty \left( \frac{p}{1 + p} + \frac{p^2}{(1 + p)^2} + \cdots + \frac{p^n}{(1 + p)^n} \right)
\]
\[
< \|V^1\|_\infty + \|V^*\|_\infty \left( \frac{p}{1 + p} + \frac{1}{1 - \frac{p}{1 + p}} \right)
\]
\[
= \|g_1\|_\infty + p \|V^*\|_\infty,
\]
where \( p = \vartheta \Delta \tau \). In the above we used the fact that \( z/(1 + z) \) is increasing for \( z \geq 0 \).

**Case II.** \( \|V^{n+1}\|_\infty = |V_1^{n+1}| \) or \( \|V^{n+1}\|_\infty = |V_{M+1}^{n+1}| \).

In this case, from (3.8), (3.9) and (3.10) it is easy to see that
\[
\|V^{n+1}\|_\infty \leq \max\{|V_1^{n+1}|, |V_{M+1}^{n+1}|\} \leq \max\{\|g_2\|_\infty, \|g_3\|_\infty\}.
\]
Combining the above two cases we have (4.6).

Note that for American Vanilla put options, \( V^* = g_1 \) as defined in (2.13). The consistency of the numerical scheme is given in the following lemma:

**Lemma 4.3 (Consistency).** The discretization scheme (3.11) is consistent.

The proof is standard since both of the time and spatial discretization schemes are standard and have been used extensively in the literature for 2nd-order partial differential equations. Therefore, we omit the proof of this lemma. Combining the above three lemmas we have the following convergence result.

**Theorem 4.1.** The solution to (3.16) converges to the viscosity solution to (2.4) as \((h, \Delta \tau) \to (0^+, 0^+)\), where \( \Delta \tau \) is defined in Lemma 4.2.

**Proof.** In [4] the authors show that if a discretization scheme for a fully nonlinear 2nd order PDE is monotone, stable and consistent, then the solution to the fully discretized system converges to the viscosity solution to the PDE. Therefore, this theorem is just a consequence of Lemmas 4.1, 4.2 and 4.3. \( \square \)
5. Solution of the nonlinear system.

5.1. The European case. In this case \( \vartheta = 0 \) in (3.11). Note that \( D \) arising from discretization of correlation product (integral) term is a dense matrix. Therefore, the solution of (3.16) is computationally expensive because of the inversion of \( D \).

To solve this, we use a Newton iterative method coupled with a regular splitting technique used in [38]. First, we write (3.16) in the following form

\[
H^{n+1}(V^{n+1}) := [A^{n+1}(V^{n+1}) - \lambda D]V^{n+1} - \frac{1}{\Delta t_n}V_n - B^{n+1} = 0.
\]

Let \( H^{n+1}(V^{n+1}) = (h_1^{n+1}(V^{n+1}), h_2^{n+1}(V^{n+1}), \ldots, h_M^{n+1}(V^{n+1}))^\top \). From (3.11) it is easy to see that the \( i \)th component of \( H^{n+1}(V^{n+1}) \) is

\[
h_i^{n+1}(V^{n+1}) = \alpha_i^{n+1}V_{i-1}^{n+1} + \beta_i^{n+1}V_i^{n+1} + \gamma_i^{n+1}V_{i+1}^{n+1} - \lambda \sum_{j=2}^{M} d_{ij}V_j^{n+1} - \frac{1}{\Delta t_n}V_i^n,
\]

where \( V_{i-1}^{n+1} \) and \( V_{M+1}^{n+1} \) are defined in (3.9) and (3.10). The Jacobian matrix of \( H^{n+1}(V^{n+1}) \) is given by \( E^{n+1}(V^{n+1}) - \lambda D \), where \( E^{n+1} \) is given by

\[
E^{n+1}(V^{n+1}) =
\begin{pmatrix}
E_{22}^{n+1} & E_{23}^{n+1} & 0 & \cdots & 0 & 0 & 0 \\
E_{32}^{n+1} & E_{33}^{n+1} & E_{34}^{n+1} & \cdots & 0 & 0 & 0 \\
0 & E_{43}^{n+1} & J_{44}^{n+1} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & E_{(M-2)(M-2)}^{n+1} & E_{(M-2)(M-1)}^{n+1} & 0 \\
0 & 0 & 0 & \cdots & E_{(M-1)(M-2)}^{n+1} & E_{(M-1)(M-1)}^{n+1} & E_{(M-1)(M)}^{n+1} \\
0 & 0 & 0 & \cdots & 0 & E_{(M)(M-1)}^{n+1} & E_{(M)(M)}^{n+1}
\end{pmatrix}
\]

with \( E_{ij}^{n+1} := \frac{\partial(A^{n+1}(V^{n+1}))_{ij}}{\partial V^{n+1}} \) for all feasible \( i \) and \( j \) and \( D \) is as defined before.

For the Jacobian \( E^{n+1}(V^{n+1}) - \lambda D \), we have the following results.

**Theorem 5.1.** For any given \( V^{n+1} \), both \( E^{n+1}(V^{n+1}) \) and \( E^{n+1}(V^{n+1}) - \lambda D \) are \( M \)-matrices.

**Proof.** For simplicity of notation, we omit the superscript \( n+1 \). To show that \( E \) is an \( M \)-matrix, from [29] we see that it suffices to prove that \( E_{ii} > 0 \), \( E_{ii-1}, E_{ii+1} \leq 0 \), \( E_{ii} \geq \vert E_{i,i-1} \vert + \vert E_{i,i+1} \vert \) and \( E_{ii} > \vert E_{i,i-1} \vert + \vert E_{i,i+1} \vert \) for at least one index \( i \). For volatility given by (2.5), it is easy to check that \( E^{n+1}(V^{n+1}) = A^{n+1}(V^{n+1}) \). Thus, from Theorem 3.1, both \( E \) and \( E - \lambda D \) are \( M \)-matrices.

Now we will prove that \( E \) and \( E - \lambda D \) are \( M \)-matrices for the volatility given in (2.6).

Let us first consider \( J_{i,i-1} \). Using the definition of \( \alpha_i^{n+1} \) in (3.12) and (2.6) we have

\[
E_{i,i-1} = \alpha_i^{n+1} + V_{n+1}^{i-1} \frac{\partial \alpha_i^{n+1}}{\partial V_{n+1}^{i-1}} + V_n^{i-1} \frac{\partial \alpha_i^{n+1}}{\partial V_n^{i-1}} + V_{n+1}^{i-1} \frac{\partial \alpha_i^{n+1}}{\partial V_{n+1}^{i-1}}
\]

\[
= \alpha_i^{n+1} - \frac{\sigma_i^2 k_v S^3}{2h^2} (d_{SS} V)(i)
\]
is strictly diagonally dominant. Hence, it is an M matrix. Therefore, all the diagonal entries of $E$ are positive and the matrix

$$
E = (1 - k_1 \text{sign}(\delta S_2 V)(i) + 2k_2 S_i(\delta S_2 V)(i)) S_i^2 \frac{1 - \text{sign}(\bar{r})}{2} \frac{\bar{r} S_i}{h} \leq 0.
$$

(by (2.7))

Similarly it can be shown that

$$
E_{i,i} = \frac{\sigma_i^2 (1 - k_1 \text{sign}(\delta S_2 V)(i) + 2k_2 S_i(\delta S_2 V)(i)) S_i^2}{h^2} + \frac{1}{\Delta r_n} + \frac{|\bar{r}| S_i}{h} + r + \lambda > 0,
$$

$$
E_{i,i+1} = -\frac{\sigma_i^2 (1 - k_1 \text{sign}(\delta S_2 V)(i) + 2k_2 S_i(\delta S_2 V)(i)) S_i^2}{2h^2} - \frac{1 + \text{sign}(\bar{r})}{2} \frac{\bar{r} S_i}{h} \leq 0.
$$

From these expressions we see that

$$
E_{i,i} = |E_{i,i-1}| + |E_{i,i+1}| + \frac{1}{\Delta r_n} + r + \lambda > |E_{i,i-1}| + |E_{i,i+1}|
$$

for any $i = 2, 3, ..., M$ with the convention that $E_{2,1} = 0 = J_{\lambda M + 1}$. Therefore, $E$ is an $M$-matrix by [29].

Now, let us consider $E - \lambda D$. From the definition of $D$, it is clear that all the off-diagonal entries of $E - \lambda D$ are non-positive and

$$
E_{i,i-1} + E_{i,i} + E_{i,i+1} - \lambda \sum_{j=2}^{M} d_{ij} = \frac{1}{\Delta r_n} + r + \lambda - \lambda \sum_{j=2}^{M} d_{ij}
$$

$$
= \frac{1}{\Delta r_n} + r + \lambda \left(1 - \sum_{j=2}^{M} d_{ij}\right)
$$

$$
\geq \frac{1}{\Delta r_n} + r > 0.
$$

This implies

$$
E_{i,i} - \lambda d_{ii} = \frac{1}{\Delta r_n} + r + \left(1 - \sum_{j} d_{ij}\right) \lambda + |E_{i,i-1}^n| + |E_{i,i+1}^n| + \lambda \sum_{j \neq i} d_{ij}
$$

$$
= \frac{1}{\Delta r_n} + r + (1 - d_{ii}) \lambda + |\alpha_i^n| + |\gamma_i^n| > |\alpha_i^n| + |\gamma_i^n|.
$$

Therefore, all the diagonal entries of $E^{n+1}(V^{n+1}) - \lambda D$ are positive and the matrix is strictly diagonally dominant. Hence, it is an $M$-matrix. $\Box$

Using the Jacobian of $H^{n+1}$, we propose the following Newton algorithm for (3.16):

**Algorithm N1**

1. Choose a tolerance $\varepsilon_1 > 0$. Let $n = 1$ and evaluate the discrete initial condition $V^1 = (V_2^1, ..., V_M^1)$ using (3.8).
2. Set $l = 1$ and $W^l = V^n$.
3. Solve

$$
[E^{n+1}(W^l) - \lambda D] \delta W = -H^{n+1}(W^l)
$$

for $\delta W$ and set $W^{l+1} = W^l + \delta W$.
4. If $||\delta W||_\infty \geq \varepsilon_1$, set $l := l + 1$ and go to Step 3. Otherwise, continue.
5. Set $V^{n+1} = W^{l+1}$. If $n < N - 1$, let $n := n + 1$ and go to Step 2. Otherwise, stop.
Note that $D$ is a dense matrix. Therefore, the solution of (5.1) is computationally expensive because of the inversion of $D$. To solve this, we use a regular splitting technique. Let

$$M = E^{n+1}(V^{n+1}) - \lambda D.$$ 

We split $M$ into

$$M = E^{n+1}(V^{n+1}) - (\lambda D) =: P - R. \quad (5.2)$$

Now, we introduce this following definition

**Definition 5.1.** A splitting $M = P - R$ is said to be a regular splitting if $P^{-1}$ is not expensive because of the inversion of $D$. If $R^{-1}$ is expensive because of the inversion of $D$, then we have a regular splitting. Using this splitting, we define an iterative technique. Let $\lambda D$ denote a regular splitting.

From Theorem 5.1 we have $E^{n+1}(V^{n+1})$ is an $M$-matrix for any given $V^{n+1}$ and hence $P^{-1}$ is not expensive. Also from (3.11) and (3.16) we have $D > 0$ and thus $R > 0$. Therefore, we have a regular splitting. Using this splitting, we define an iterative scheme for (5.1) as follows.

$$P(\delta W)^{k+1} = R(\delta W)^{k} - H. \quad (5.3)$$

The following lemma establishes the convergence of the iterative method (5.3).

**Lemma 5.1.** The iterative scheme (5.3) associated with the regular splitting (5.2) is convergent.

**Proof.** From Theorem 5.1, we know that $M = E^{n+1}(V^{n+1}) - \lambda D$ is an $M$-matrix. Hence we have $M^{-1} > 0$. By the result in [1], we have (5.3) is convergent. \hfill \square

Recall that matrix $D$ is a dense matrix, hence it is computationally expensive to evaluate the multiplication in (5.3) directly, because the computational cost is $O(N^2)$. Thus, we present a fast algorithm to evaluate this matrix. This algorithm, based on Fast Fourier Transform (FFT), has a computational cost of order $O(N \log N)$.

From (3.2) and (5.2) we know that $R$ is a Toeplitz matrix. Applying FFT to $R(\delta W)^k$ and $RV^{n+1}$ will produce wrap-round pollution. Hence, we embed the Toeplitz matrix $R$ into circulant matrix $C$ (cf. [26]) as follow

$$C = \begin{pmatrix}
    f_0 & f_1 & f_2 & \cdots & f_{M-3} & f_{M-2} & f_{M-3} & \cdots & f_{-2} & f_{-1} \\
    f_{-1} & f_0 & f_1 & \cdots & f_{M-4} & f_{M-3} & f_{M-2} & \cdots & f_{-3} & f_{-2} \\
    f_{-2} & f_{-1} & f_0 & \cdots & f_{M-5} & f_{M-4} & f_{M-3} & \cdots & f_{-4} & f_{-3} \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    f_{M-2} & f_{M-1} & f_{M-2} & \cdots & f_{1} & f_{0} & f_1 & \cdots & f_{M-3} & f_{M-2} \\
    f_{M-3} & f_{M-2} & f_{M-1} & \cdots & f_{2} & f_{1} & f_0 & \cdots & f_{M-4} & f_{M-3} \\
    f_{M-4} & f_{M-3} & f_{M-2} & \cdots & f_{3} & f_{2} & f_{1} & \cdots & f_{M-5} & f_{M-4} \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    f_1 & f_2 & f_3 & \cdots & f_{M-2} & f_{M-1} & f_{M-2} & \cdots & f_{-1} & f_0 
\end{pmatrix}.$$

Define

$$\delta U^k = [(\delta W)^k, \ldots, (\delta W)^k, 0, \ldots, 0]^T, \quad U^l = [V^l, \ldots, V^l, 0, \ldots, 0]^T.$$ 

Then, matrix-vector products $R(\delta W)^k$ and $RV^l$ are realized as the first $M - 1$ entries in $C(\delta U)^k$ and $CU^l$. 


The product $R(\delta W)^k$ is computed in the following two FFT operations. Let $FFT(u)$ denote the FFT of $u$ and define the vector

$$F = (f_0, f_1, \ldots, f_{N-2}, f_{2-N}, f_{3-N}, \ldots, f_{-1}),$$

which generates the row vector of $C$ by permutation. First, we compute $FFT(F)$ and $FFT((\delta U)^k)$. Then, we compute the inverse FFT of the product of $FFT(F)$ and $FFT((\delta U)^k)$.

The numerical implementation for the system (3.16) can be summarized in this following algorithm.

**Algorithm N2**

1. Choose tolerances $\varepsilon_1, \varepsilon_2 > 0$. Let $n = 1$ and evaluate the discrete initial condition $\nabla^1 = (V_1^1, \ldots, V_M^1)\top$ using (3.8).
2. Compute $FFT(F)$.
3. Set $l = 1$ and $V^l = V^n$.
4. Compute $FFT(U^l)$ and the inverse FFT on the product of $FFT(F)$ and $FFT(U^l)$ which gives $DV^l$ in $H$.
5. Set $k = 1$ and $(\delta W)^1 = [0, \ldots, 0] \top$.
6. Compute $FFT((\delta U)^k)$ and the inverse FFT on the product of $FFT(F)$ and $FFT((\delta U)^k)$ which gives $R(\delta W)^k$.
7. Solve

$$E^{n+1}(V^l)(\delta W)^{k+1} = R(\delta W)^k - H^{n+1}(V^l)$$

for $(\delta W)^{k+1}$.
8. If $\|((\delta W)^{k+1} - (\delta W)^k)\|_\infty < \varepsilon_1$ then stop. Otherwise, set $k := k + 1$, go to Step 6.
9. Set $(\delta W)^l = (\delta W)^{k+1}$. If $\|((\delta W)^l)\|_\infty < \varepsilon_2$ then stop. Otherwise, compute $V^{l+1} = V^l + (\delta W)^l$ and set $l = l + 1$, go to Step 4.
10. Set $\nabla^{n+1} = \nabla^{l+1}$. If $n < N$, let $n := n + 1$ and go to Step 3. Otherwise, stop.

5.2. The American case. From (3.7) we see that the discretized nonlinear system from American options differs from that from the European options by a penalty term $-\vartheta[V^* - V]_+$. Since this term is non-smooth in $V^{n+1}$ for each $n = 1, \ldots, N$, a smoothing technique such as those proposed in [34, 17] is usually used to locally smooth out the penalty term near zero. It has been shown in [34, 17] that the smoothed penalty term is an increasing function of $V^{n+1}$ and thus its derivative is non-negative. Therefore, all the analysis presented for European options above hold true in this case.

6. Numerical results. In this section we demonstrate the rates of convergence and numerical performance of our method using European call and put options as well as an American put option.

**Test Problem 1.** European vanilla call and put options with the system parameters: $r = 0.1$, $\sigma = 0.2$, $K = 40$, $T = 1$, $S_{\text{max}} = 80$, $\mu = 0.01$, $\sigma_J = 0.5$, and $\lambda = 0.1$. The payoff and boundary conditions are given in (2.10)–(2.12) for the call and put.

This test problem is solved using our method for the $\sigma$ defined in (2.6) with the transaction cost parameters $a = 0.01$ and $b = 0.07$. The computed European call and put option prices are plotted in Figure 6.1. From the figure we see that
Figure 6.1. Prices of the European call and put options with $a = 0.01$ and $b = 0.07$.

Figure 6.2. Prices of the European call and put option for different values of the transaction cost parameter.

Numerical solutions are stable. To see the effect of the transaction costs on the price of the call and put options, we plot the values of the options at $t = 0$ (or $\tau = T$) for $a = 0$ and three different values of $b$ on the interval $[30, 50]$ in Figure 6.2 for the European call option and put, respectively. From these plots we see that the prices of the options increase as the transaction cost parameter $b$ increases, as expected in practice.

We now investigate the rate of convergence numerically, which requires an exact or reference solution. Since the exact solution to this problem is unknown, we use the numerical solution to (3.16) on the uniform mesh with $M = 5121$ ($h = \frac{1}{5120}$) and $N = 2560$ ($\Delta \tau = \frac{1}{2560}$) as an “exact” or reference solution. To determine these rates, we choose a sequence of meshes generated by successively halving the mesh sizes of the previous ones, starting from a given coarse mesh. Using the reference solution we then calculate the following ratios of the numerical solutions from two consecutive meshes:

$$\text{Ratio}(\| \cdot \|_{h,2}) = \frac{\| V_h^{\Delta \tau} - V_{\text{exact}} \|_{h,2}}{\| V_{h/2}^{\Delta \tau/2} - V_{\text{exact}} \|_{h,2}},$$
The computed errors in $\| \cdot \|_{h,2}$ and the ratios two consecutive errors are listed in Tables 6.1 and 6.2 for the European call and put respectively. From the tables we see that the rates of convergence of our method is about 2, showing that our numerical method is 2nd order accurate in the $L^2$-norm.

**Test Problem 2.** American put option with the system parameters: $r = 0.1$, $\sigma_0 = 0.2$, $K = 40$, $T = 1$, $S_{\text{max}} = 80$, $\mu = 0.01$, $\sigma_f = 0.5$, $\lambda = 0.1$, and $\vartheta = 1000$. The payoff and boundary conditions are given in (2.13)–(2.15).

To see the difference between the European and American options, we plot the computed values of both of the options at $t = 0$ (or $\tau = T$) for $S \in [20, 80]$ in Figure 6.3. From the figure it is clear that the American put option is more valuable than the European put option as expected.

Furthermore, we investigate the effect of the transaction cost parameters on the value of the option. To see this, we plot the values of the option at $t = 0$ (or $\tau = T$) for three different values of $b$ on the interval $[30, 50]$ in Figure 6.4 in which the curve for $a = 0$ and $b = 0$ is the price of the American put option without transaction cost. From this figure we see that the price of the option increases as the transaction parameter $b$ increases as expected in practice.
7. Conclusion. In this paper we have proposed and analyzed an upwind finite difference scheme for a nonlinear Partial Integro-Differential Equation (PIDE) and a nonlinear partial integro-differential complementarity problem arising from European and American option pricing with transaction costs under a jump diffusion process. The complementarity problem is first approximated by a nonlinear PIDE with a penalty term which contains the nonlinear PIDE for pricing European options as a special case. A finite difference scheme has then been developed for the penalized PIDE and the convergence of the solution to discretized system to the viscosity solution of the penalty equation has been proven. To solve the discretized system with a dense coefficient matrix, we have developed a fast iterative method.
coupled with a regular splitting technique. To speed up the computation of the discretized integral term, we have proposed to use the FFT algorithm. Numerical experiments have been performed to confirm the theoretical results and demonstrate the usefulness of the method. The numerical results show that the prices of both European and American options are increasing functions of the transaction cost parameters.

REFERENCES

[1] A. Almendral and C. W. Oosterlee, Numerical valuation of options with jumps in the underlying, Appl. Math. Comput., 53 (2005), 1–18.
[2] A. Anderson and J. Andresen, Jump diffusion process: volatility smile fitting and numerical methods for option pricing, Rev. Derivat. Res., 4 (2000), 231–262.
[3] J. Ankudinova and M. Ehrhardt, On the numerical solution of nonlinear Black-Scholes equations, Computers and Mathematics with Applications, 56 (2008), 799–812.
[4] C. G. Averbuj, Nonlinear differential evolution equation arising in option pricing when including transaction costs: a viscosity solution approach, R. Bras. Eco. de Emp., 12 (2012), 81–90.
[5] G. Barles, Convergence of numerical schemes for degenerate parabolic equations arising in finance theory, in: L.C.G. Rogers, D. Talay (Eds), Numerical Methods in Finance, Cambridge University Press, Cambridge, (1997), 1–21.
[6] F. Black and M. Scholes, The pricing of options and corporate liabilities, The Journal of Political Economy, 81 (1973), 637–654.
[7] W. Chen and S. Wang, A penalty method for a fractional order parabolic variational inequality governing American put option valuation, Computers & Mathematics with Applications, 67 (2014), 77–90.
[8] W. Chen and S. Wang, A finite difference method for pricing European and American options under a geometric Levy process, Journal of Industrial and Management Optimization, 11 (2015), 241–264.
[9] R. Company, L. Jodar and J. R. Pintos, A numerical method for European option pricing with transaction costs nonlinear equation, Mathematics and Computer Modelling, 50 (2009), 910–920.
[10] R. Cont and P. Tankov, Financial Modelling with Jump Processes, Chapman and Hall/CRC, Boca Raton, FL, 2004.
[11] R. Cont and E. Voltchkova, A finite difference scheme for option pricing in jump diffusion and exponential Levy models, SIAM J. Numer. Anal., 43 (2005), 1596–1626.
[12] B. Dupire, Pricing with a smile, Risk, 7 (1994), 18–20.
[13] B. D. üring, M. Fournier and A. J. üngel, High order compact finite difference schemes for a nonlinear Black-Scholes equation, International Journal of Theoretical and Applied Finance, 6 (2003), 767–789.
[14] Y. d’Halluin, P. A. Forsyth and K. R. Vetzal, Robust numerical methods for contingent claims under jump diffusion processes, IMA J. Numer. Anal., 25 (2005), 87–112.
[15] P. Heider, Numerical methods for nonlinear Black-Scholes equations, Applied Mathematical Finance, 17 (2010), 59–81.
[16] S. L. Heston, A closed-form solution for options with stochastic volatility with application to bond and currency options, Rev. Financial Stud., 6 (1993), 327–343.
[17] C. Huang and S. Wang, A power penalty approach to a nonlinear complementarity problem, Operations Research Letters, 38 (2010), 72–76.
[18] C. Huang and S. Wang, A penalty method for a mixed nonlinear complementarity problem, Nonlinear Analysis TMA, 75 (2012), 588–597.
[19] C. S. Huang, C. H. Hung and S. Wang, A fitted finite volume method for the valuation of options on assets with stochastic volatilities, Computing, 77 (2006), 297–320.
[20] J. Hull, Options, Futures, and Other Derivatives, Prentice-Hall, Englewood Cliffs, 2005.
[21] J. Hull and A. White, The pricing of options on assets with stochastic volatilities, J. Finance, 42 (1987), 281–300.
[22] H. E. Leland, Option pricing and replication with transaction costs, Journal of Finance, 40 (1985), 1283–1301.
[23] D.C. Lesmana and S. Wang, An upwind finite difference method for a nonlinear Black-Scholes equation governing European option valuation, Applied Mathematics and Computation, 219 (2013), 8818–8828.
[24] D.C. Lesmana and S. Wang, Penalty approach to a nonlinear obstacle problem governing American put option valuation under transaction costs, Applied Mathematics and Computation, 251 (2015), 318–330.
[25] W. Li and S. Wang, Pricing American options under proportional transaction costs using a penalty approach and a finite difference scheme, Journal of Industrial and Management Optimization, 9 (2013), 365–389.
[26] C. Van Loan, Computational Frameworks for the Fast Fourier Transform, Frontier in applied mathematics, Vol.10. SIAM, Philadelphia, PA, 1992.
[27] R. C. Merton, Option pricing when underlying stock returns are discontinuous, J. Financial Econ., 3 (1976), 125–144.
[28] A. Mocioalca, Jump diffusion options with transaction costs, Rev. Roumaine Math. Pures Appl., 52 (2007), 349–366.
[29] S. Wang, A penalty method for a finite-dimensional obstacle problem with derivative constraints, Optimization Letters, 8 (2014), 1799–1811.
[30] S. Wang, A penalty approach to a discretized double obstacle problem with derivative constraints, Journal of Global Optimization, 62 (2015), 775–790.
[31] S. Wang and X. Q. Yang, A power penalty method for linear complementarity problems, Operations Research Letters, 36 (2008), 211–214.
[32] S. Wang and X. Q. Yang, A power penalty method for a bounded nonlinear complementarity problem, Optimization, 64 (2015), 2377–2394.
[33] S. Wang, X. Q. Yang and K. L. Teo, Power penalty method for a linear complementarity problem arising from American option valuation, Journal of Optimization Theory & Applications, 129 (2006), 227–254.
[34] S. Wang, S. Zhang and Z. Fang, A superconvergent fitted finite volume method for Black-Scholes equations governing European and American option valuation, Numerical Methods for Partial Differential Equations, 31 (2015), 1190–1208.
[35] Y.-P. Wang and S.-L. Tao, Application of regularization technique to variational adjoint method: A case for nonlinear convection-diffusion problem, Applied Mathematics and Computation, 218 (2011), 4475–4482.
[36] K. Zhang and S. Wang, A penalty approximation method for a semilinear parabolic double obstacle problem, Journal of Global Optimization, 60 (2014), 531–550.

Received April 2016; revised July 2016.
E-mail address: donny.lesmana@graduate.uwa.edu.au
E-mail address: song.wang@curtin.edu.au