(To appear in *Journal of Computational Finance*, vol 2 no 3, 1998)

Option Pricing and Linear Complementarity*

Jacqueline Huang and Jong-Shi Pang
Department of Mathematical Sciences
The Johns Hopkins University
Baltimore, Maryland 21218-2682, U.S.A.
Emails: jhuang@mts.jhu.edu and jsp@vicp1.mts.jhu.edu.

October 1997
(Revised May 1998)

Abstract

Many American option pricing models can be formulated as linear complementarity problems (LCPs) involving partial differential operators. While recent work with this approach has mainly addressed the model classes where the resulting LCPs are highly structured and can be solved fairly easily, this paper discusses a variety of option pricing models that are formulated as partial differential complementarity problems (PDCPs) of the convection-diffusion kind whose numerical solution depends on a better understanding of LCP methods. Specifically, we present second-order upwind finite difference schemes for the PDCPs and derive fundamental properties of the resulting discretized LCPs that are essential for the convergence and stability of the finite difference schemes and for the numerical solution of the LCPs by effective computational methods. Numerical results are reported to support the benefits of the proposed schemes. A main objective of this presentation is to elucidate the important role that the LCP has to play in the fast and effective numerical pricing of American options.

1 Introduction

The numerical valuation of options and derivative securities is of central importance to financial management. The paper by Broadie and Detemple [3] presents a survey of recent numerical methods for the pricing of these financial products. For a more theoretical treatment, see the paper by Myneni [23]. For many American and exotic options, the discretized linear complementarity approach described in Wilmott, Dewynne, and Howison [26] provides a very promising numerical procedure that has great potential for dealing with a wide range of options, ranging from a vanilla American option to a highly complex multi-asset option with such complicating factors as stochastic volatility, transaction costs, path dependency of payoffs, and so on. One of the earliest papers that used a linear complementarity method (although it was not labeled as such) for pricing an American put option is by Brennan and Schwartz [2]. In spite of the pioneering efforts of these authors

*This work was based on research supported by the National Science Research Foundation under grant CCR-9624018.

and others, including Jaillet, Lamberton, and Lapeyre [18] and Dempster and Hutton [10, 11], it is our belief that the linear complementarity approach to pricing American options is very much at its infancy and its full potential has yet to be realized.

This paper aims to further explore the linear complementarity problem (LCP) and its extensions as a powerful mathematical tool for the valuation of American options and interest rate sensitive financial products. The LCP was introduced within the mathematical programming community more than three decades ago and has now developed into a very fruitful discipline in applied mathematics. A comprehensive study of this problem, including a complete theory and extensive algorithms, is presented in the book by Cottle, Pang, and Stone [5]. Among the classic application areas of the LCP is that of optimal stopping (see Subsection 1.2 in the cited reference), the latter being at the heart of all option pricing models with a "free" expiry date. Recognizing this connection, it is then easy for one to accept the LCP as having an important role to play in the treatment of American options.

Today, there are numerous option and derivative valuation models. Typically, these models are formulated as second-order, time-dependent, parabolic partial differential equations (PDEs) for European options with fixed exercise dates and partial differential complementarity problems (PDCPs) for American options with free exercise dates. In this paper, we study the numerical solution of the latter option models by discretized LCP methods. We are particularly interested in those models where the partial differential operators are of the "convection-diffusion" kind [17, 19, 22] that can not be conveniently converted into pure diffusion operators. We discuss second-order upwind finite difference schemes [4, 22, 24, 27] for approximating the PDCPs and derive some basic properties of the discretized LCPs, stressing those that are essential for the stability of the finite difference approximation schemes and for the validity of the LCP methods. Obviously it is not possible for us to consider in one single paper all the option models that have appeared in the literature; hopefully, based on the treatment of the models presented herein, one can easily adopt the methodology to other models. Also, due to space limitation, we have omitted in this paper the error analysis of the discretization approximations.

Upwind finite difference schemes provide an effective numerical procedure to stabilize the potentially adverse sign oscillation in the convection term of the partial differential operator (that is, the first-order partial derivative with respect to the state variable). First-order schemes of this kind typically lead to discretized problems involving Z matrices (that is, square matrices whose off-diagonal entries are non-positive). For a recent study of these schemes for solving general algebraic obstacle problems, see the Ph.D. dissertation [25] and the paper [19]; an interesting application of a first-order finite difference scheme for solving an American option pricing problem in a jump-diffusion model (which leads to a PDCP with a convection term) is presented in [28]. In the present paper, inspired by the approach taken in [26, Section 22.4.1] and [13] for pricing the continuously sampled average strike Asian option, we present second-order upwind methods in order to obtain second-order approximations of the partial differential operators. As we shall see, the resulting discretized LCPs, although no longer of the Z-matrix type, can still be solved effectively by a host of iterative and/or pivotal methods (such as those presented in [5]).

2 One-Factor American Option Models

The options and derivative models treated in this paper are all of the American type; that is, the option owner can exercise the option at any time before its expiry. The basic framework is that of Black-Scholes [1] as expanded by Wilmott, Dewynne and Howison [26]. The vanilla American option with one underlying asset is the most basic of all American options; the discretized LCP has a defining matrix which is highly structured and has very desirable theoretical properties that render the problem solvable by most well-known methods, such as the those described in [5]. For the detailed numerical treatment of the vanilla American option by linear complementarity and linear programming methods, we refer to [26, 10, 11] and the references therein. We begin our discussion with a variant of the basic Black-Scholes model.

2.1 Cox's model of constant elasticity of variance

In 1975, John Cox proposed an option pricing model that generalized the Black-Scholes model, by postulating a certain non-stationarity of the variance of the stock returns. Although Cox's original note was not published until recently [6], his model and the related work [7] had been the basis of various alternative models, including those with stochastic volatilities; see Subsection 6.4.A of the book by Gibson [15] for more discussion on the issue of non-stationarity of stock returns' variance. In what follows, we discuss Cox's model of constant elasticity of variance in the context of an American option. A special case of this model yields the classic Black-Scholes vanilla American option.

Assuming that the asset pays out the dividend $D(S,t)dt$ in a time interval dt, the Cox model postulates that the asset price $S(t)$ satisfies the following stochastic differential equation:

$$dS \;=\; (\,\mu\,S - D(S,t)\,)\,dt + \sigma\,S^{\alpha/2}\,dW,$$

where, for simplicity, we assume that μ and σ are positive constants, α is an elasticity parameter in the interval $(0,2]$, and dW is a standard Wiener process with mean zero and variance dt. Let $r > 0$ denote the constant interest rate. Let $\Lambda(S,t)$ denote the given payoff function of an American option whose time of expiry is denoted T. Let $V(S,t)$ denote the value of this option as a function of the asset price S and time t. By a standard no-arbitrage argument, this value must satisfy the following complementarity conditions:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\,\sigma^2\,S^\alpha\,\frac{\partial^2 V}{\partial S^2} + (\,r\,S - D(S,t)\,)\,\frac{\partial V}{\partial S} - r\,V \;\leq\; 0 \tag{1}$$

$$V(S,t) \;\geq\; \Lambda(S,t) \tag{2}$$

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\,\sigma^2\,S^\alpha\,\frac{\partial^2 V}{\partial S^2} + (\,r\,S - D(S,t)\,)\,\frac{\partial V}{\partial S} - r\,V\right)\,(\,V(S,t) - \Lambda(S,t)\,) \;=\; 0, \tag{3}$$

and some appropriate boundary conditions. Unless $\alpha = 2$ and $D(S,t)$ is proportional to S (this case yielding the Black-Scholes model), it is not clear that there is a convenient transformation that will reduce the linear partial differential operator with variable coefficients in the above complementarity system to the constant coefficient case.

For the option model described by (1)–(3), the specific form of the payoff function $\Lambda(S,t)$ and the boundary conditions are not critical to the numerical solution of the model, provided that a proper discretization scheme (to be described shortly) is employed. Of course, this function and the boundary conditions are essential in practice; however, for reasons given below, the numerical treatment of the model does not rely on the explicit knowledge of the payoff function and boundary conditions. Indeed, when the partial differential system is discretized, the function $\Lambda(S,t)$ and the boundary conditions determine the constant vector, but not the matrix, in the resulting discretized LCP. From well-known LCP theory [5], such a constant vector is not expected to have a dramatic effect on the problem, provided that the matrix that defines the problem belongs to some broad class (such as the class of P-matrices; that is, those whose principal minors are all positive). Since this paper will provide sufficient conditions for the matrix in the discretized LCP to have desirable properties, the constant vector is of secondary importance, and thus will not be given too much emphasis in the analysis.

The finite difference discretization

We apply a finite difference approximation scheme to the partial differential complementarity system (1)–(3). Since there does not appear to be any advantage to make the usual logarithmic change of variable for the asset price S, we apply finite differences to this system directly. The partial differential operator

$$\mathcal{L}_{\text{Cox}} \equiv \frac{\partial}{\partial t} + \tfrac{1}{2}\,\sigma^2\,S^\alpha\,\frac{\partial^2}{\partial S^2} + (\,r\,S - D(S,t)\,)\,\frac{\partial}{\partial S} - r$$

is of the convection-diffusion type [17, 19, 22]; see also the technical discussion in [26, page 114]. The first and second partial derivatives, $\partial/\partial S$ and $\partial^2/\partial S^2$, are the convection and diffusion terms, respectively. For a partial differential operator of this kind, an upwind finite difference approximation scheme is frequently used in the PDE literature in order to prevent oscillations due to a potentially large convection term. In the case of $\mathcal{L}_{\text{Cox}}$ where the dividend yield function $D(S,t)$ is such that $rS - D(S,t)$ takes both positive and negative values, this oscillation in sign could have an adverse effect on the stability of the approximation scheme as well as on the numerical solution of the discretized LCP. In this case, the upwind scheme should be utilized. The main idea of such a scheme is that care should be taken in approximating the convection term $\partial/\partial S$ when its multiplicative factor has a dominant effect on the differential operator.

In what follows, we apply a second-order upwind finite difference approximation to the operator $\mathcal{L}_{\text{Cox}}$ in conjunction with the θ-method as discussed in [26, Chapter 22]. For the partial derivative with respect to time, we choose a positive but small time step δt and use the following approximation:

$$\frac{\partial V}{\partial t}(S,t) = \frac{V(S,t+\delta t) - V(S,t)}{\delta t} + O(\delta t),$$

where a function $f(x)$ is said to be $O(g(x))$ if

$$\limsup_{g(x)\to 0} \frac{|\,f(x)\,|}{|\,g(x)\,|} < \infty.$$

For the second partial derivative with respect to the variable S, we choose a positive but small scalar δS and use an θ-weighted central difference approximation:

$$\begin{aligned}\frac{\partial^2 V}{\partial S^2}(S,t) &= \theta\,\frac{V(S+\delta S,t) - 2\,V(S,t) + V(S-\delta S,t)}{(\delta S)^2} + \\ &\qquad (\,1-\theta\,)\,\frac{V(S+\delta S,t+\delta t) - 2\,V(S,t+\delta t) + V(S-\delta S,t+\delta t)}{(\delta S)^2} + O((\delta S)^2),\end{aligned}$$

where $\theta \in [0,1]$ is a given parameter whose specializations yield the explicit method ($\theta = 0$), the implicit method ($\theta = 1$), and the Crank-Nicolson method ($\theta = 1/2$).

For the first partial derivative with respect to the variable S, we use a θ-weighted second-order upwind approximation consistent with the second-order approximation for $\partial^2 V/\partial S^2$. Specifically, the term $(rS - D(S,t))\partial V/\partial S$ is approximated as follows:

(a) if $rS - D(S,t) > 0$,

$$\begin{aligned}
&(\,rS - D(S,t))\,\frac{\partial V}{\partial S}(S,t) \\
&= |\,rS - D(S,t)\,|\,\Bigg[\,\theta\,\frac{4\,V(S+\delta S,t) - 3\,V(S,t) - V(S+2\,\delta S,t)}{2\,\delta S} + \\
&\quad (\,1-\theta\,)\,\frac{4\,V(S+\delta S,t+\delta t) - 3\,V(S,t+\delta t) - V(S+2\,\delta S,t+\delta t)}{2\,\delta S}\,\Bigg] + O((\delta S)^2).
\end{aligned}$$

(b) if $rS - D(S,t) < 0$,

$$\begin{aligned}
&(\,rS - D(S,t))\,\frac{\partial V}{\partial S}(S,t) \\
&= |\,rS - D(S,t)\,|\,\Bigg[\,\theta\,\frac{4\,V(S-\delta S,t) - 3\,V(S,t) - V(S-2\,\delta S,t)}{2\,\delta S} + \\
&\quad (\,1-\theta\,)\,\frac{4\,V(S-\delta S,t+\delta t) - 3\,V(S,t+\delta t) - V(S-2\,\delta S,t+\delta t)}{2\,\delta S}\,\Bigg] + O((\delta S)^2).
\end{aligned}$$

Notice that in case (a), a forward difference is used to approximate the first partial derivative $\partial H/\partial R$; whereas in case (b) a backward difference is used. The effect of this approximation is that the diagonal entry of the resulting matrix in the discretized system is always positive, for all positive values of δS and δt.

The PDCP (1)–(3) is approximated on a regular grid with step sizes δt and δS, and with the variable S truncated so that

$$0 \,\leq\, S \,\leq\, N\,\delta S,$$

where N is a (large) positive integer. Write

$$V_n^m \,\equiv V(n\,\delta S, m\,\delta t) \text{ and } \Lambda_n^m \,\equiv \Lambda(n\,\delta S, m\,\delta t), \quad \text{for } 0 \,\leq\, n \,\leq\, N \text{ and } 0 \,\leq\, m \,\leq\, M,$$

where $M \equiv T/\delta\, t$ is the total number of time steps in the discretization (the step size δt is chosen that M is an integer). The system (1)–(3), under the above finite difference approximations and with the boundary conditions suitably accounted for, leads to the following finite-dimensional LCP at time $t = m\,\delta t$ for $m = M-1,\, M-2,\, \ldots,\, 2,\, 1,\, 0,$

$$0 \leq (\,\boldsymbol{V}^m - \boldsymbol{\Lambda}^m\,) \,\perp\, (\,\boldsymbol{q}^{m+1} + \boldsymbol{M}_{\text{Cox}}^{\text{up}}\boldsymbol{V}^m\,) \,\geq\, 0, \tag{4}$$

where the perp notation $\perp$ denotes the orthogonality relation of two vectors; i.e., $a \perp b$ means $a^T b = 0$;

$$\boldsymbol{V}^m \,\equiv\, \begin{pmatrix} V_1^m \\ \vdots \\ V_{N-1}^m \end{pmatrix}$$

is the $(N-1)$-vector of variables of the LCP,

$$\mathbf{\Lambda}^m \equiv \begin{pmatrix} \Lambda_1^m \\ \vdots \\ \Lambda_{N-1}^m \end{pmatrix} \quad \text{and} \quad \boldsymbol{q}^{m+1} \equiv \begin{pmatrix} q_1^{m+1} \\ \vdots \\ q_{N-1}^{m+1} \end{pmatrix}$$

are two $(N-1)$-vectors of constants of the LCP, and

$$\boldsymbol{M}_{\text{Cox}}^{\text{up}} \equiv \begin{pmatrix} b_1 & c_1 & d_1 & 0 & 0 & 0 & \cdots & 0 \\ a_2 & b_2 & c_2 & d_2 & 0 & 0 & \cdots & 0 \\ e_3 & a_3 & b_3 & c_3 & d_3 & 0 & & 0 \\ 0 & e_4 & a_4 & b_4 & c_4 & d_4 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & e_{N-3} & a_{N-3} & b_{N-3} & c_{N-3} & d_{N-3} \\ 0 & 0 & \cdots & 0 & e_{N-2} & a_{N-2} & b_{N-2} & c_{N-2} \\ 0 & 0 & \cdots & 0 & 0 & e_{N-1} & a_{N-1} & b_{N-1} \end{pmatrix}$$

is the $(N-1)$ by $(N-1)$ matrix with entries given by:

$$\begin{aligned}
e_n &\equiv \frac{\theta\,(\,r\,n\,\delta S - D_n^m\,)^-}{2\,\delta S} \quad \text{for } n = 3, \ldots, N-1 \\
a_n &\equiv -\frac{\theta\,\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2}}{2} - \frac{2\,\theta\,(\,r\,n\,\delta S - D_n^m\,)^-}{\delta S} \quad \text{for } n = 2, \ldots, N-1 \\
b_n &\equiv \frac{1}{\delta t} + \theta\,\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2} + \frac{3\,\theta\,|\,r\,n\,\delta S - D_n^m)\,|}{2\,\delta S} + r \quad \text{for } n = 1, \ldots, N-1 \\
c_n &\equiv -\frac{\theta\,\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2}}{2} - \frac{2\,\theta\,(\,r\,n\,\delta S - D_n^m\,)^+}{\delta S} \quad \text{for } n = 1, \ldots, N-2 \\
d_n &\equiv \frac{\theta\,(\,r\,n\,\delta S - D_n^m\,)^+}{2\,\delta S} \quad \text{for } n = 1, \ldots, N-3,
\end{aligned}$$

where $D_n^m \equiv D(n\delta S, m\delta t)$ and for a real number x, $x^+ \equiv \max(x, 0)$ denotes its nonnegative part and $x^- \equiv \max(-x, 0)$ its non-positive part.

The structure of the matrix $\boldsymbol{M}_{\text{Cox}}^{\text{up}}$ is typical in a second-order upwind finite difference approximation of a convection-diffusion operator with one state variable. For one thing, the diagonal entries of $\boldsymbol{M}_{\text{Cox}}^{\text{up}}$ are all positive; but its off-diagonal entries are not all non-positive (cf. the entries e_n and d_n). Thus $\boldsymbol{M}_{\text{Cox}}^{\text{up}}$ is in general not a Z-matrix. More generally, only first-order finite difference schemes of a convection-diffusion operator will yield an LCP with a Z-matrix. Subsequently, we will establish some common properties of the matrices obtained by a second-order upwind scheme.

The constant vector $\boldsymbol{q}^{m+1}$ is closely tied to the matrix $\boldsymbol{M}^{\text{up}}_{\text{Cox}}$. Indeed, by defining the entries:

$$\begin{aligned}
e'_n &\equiv \frac{(1-\theta)\,(r\,n\,\delta S - D^m_n)^-}{2\,\delta S} \quad \text{for } n = 3,\ldots,N-1 \\
a'_n &\equiv -(1-\theta)\left(\frac{\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2}}{2} + \frac{2\,(r\,n\,\delta S - D^m_n)^-}{\delta S}\right) \quad \text{for } n = 2,\ldots,N-1 \\
b'_n &\equiv -\frac{1}{\delta t} + (1-\theta)\left(\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2} + \frac{3\,|\,r\,n\,\delta S - D^m_n\,|}{2\,\delta S}\right) \quad \text{for } n = 1,\ldots,N-1 \\
c'_n &\equiv -(1-\theta)\left(\frac{\sigma^2\,n^\alpha\,(\delta S)^{\alpha-2}}{2} + \frac{2\,(r\,n\,\delta S - D^m_n)^+}{\delta S}\right) \quad \text{for } n = 1,\ldots,N-2 \\
d'_n &\equiv \frac{(1-\theta)\,(r\,n\,\delta S - D^m_n)^+}{2\,\delta S} \quad \text{for } n = 1,\ldots,N-3
\end{aligned}$$

and forming the matrix $\boldsymbol{M}'_{\text{Cox}}$ in the same way as $\boldsymbol{M}^{\text{up}}_{\text{Cox}}$ (using the above primed entries instead), we see that

$$\boldsymbol{q}^{m+1} = \boldsymbol{M}'_{\text{Cox}}\,\boldsymbol{V}^{m+1},$$

where $\boldsymbol{V}^{m+1}$ is computed from the previous time step (or from the boundary conditions when $m = M-1$).

2.2 The continuously sampled average strike option

We choose to discuss the continuously sampled average strike Asian option with early exercise for two purposes. One is that we wish to clarify the convergence of the (point) projected successive over-relaxation (PSOR) method for solving the model as mentioned in the references [13, 26]. The other is that we will use this model to compare the upwind versus the no-upwind finite difference scheme in order to highlight the benefit of the former for approximating a convection-diffusion operator. (The no-upwind scheme was used in the references.) The matrix in the resulting discretized LCP from the upwind scheme can be easily shown to be positive definite (albeit asymmetric) under a mild condition on the time step δt. In the discussion that follows, we assume that the asset pays no dividend.

As shown in these references, after a suitable change of variables (similarity transformation) and reformulation, the continuously sampled average strike Asian option leads to the following PDCP with the function $H(R,t)$ to be computed:

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + (1 - r\,R)\frac{\partial H}{\partial R} \le 0 \tag{5}$$

$$H(R,t) \ge \Lambda(R,t) \tag{6}$$

$$\left(\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + (1 - r\,R)\frac{\partial H}{\partial R}\right)(H - \Lambda) = 0, \tag{7}$$

with $\Lambda(R,t) = \max(1 - R/t, 0)$. As derived in [26], the no-upwind finite difference approximation scheme of the partial differential operator leads to a discretized LCP parameterized by the time step. Specifically, at time $t = m\,\delta t$ for $m = M-1,\ M-2,\ \ldots,\ 2,\ 1,\ 0$, this LCP is

$$0 \le (\boldsymbol{H}^m - \boldsymbol{\Lambda}^m) \perp (\boldsymbol{b}^{m+1} + \boldsymbol{M}^{\text{nup}}_{\text{Asi}}\boldsymbol{H}^m) \ge 0, \tag{8}$$

where

$$
\boldsymbol{M}^{\rm nup}_{\rm Asi} \equiv \begin{pmatrix} b_1 & c_1 & d_1 & 0 & 0 & \cdots & 0 \\ a_2 & b_2 & c_2 & d_2 & 0 & \cdots & 0 \\ 0 & a_3 & b_3 & c_3 & d_3 & & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & & a_{N-3} & b_{N-3} & c_{N-3} & d_{N-3} \\ 0 & 0 & \cdots & 0 & a_{N-2} & b_{N-2} & c_{N-2} \\ 0 & 0 & \cdots & 0 & 0 & a_{N-1} & b_{N-1} \end{pmatrix}
$$

is the $(N-1)$ by $(N-1)$ matrix with entries given by:

$$
\begin{aligned}
a_n &\equiv -\frac{\theta}{2}\sigma^2 n^2 \quad \text{for } n = 2, \ldots, N-1 \\
b_n &\equiv \frac{1}{\delta t} + \frac{\theta}{2}\left(2\sigma^2 n^2 - 3\, r\, n + \frac{3}{\delta R}\right) \quad \text{for } n = 1, \ldots, N-1 \\
c_n &\equiv \frac{\theta}{2}\left(-\sigma^2 n^2 + 4\, r\, n - \frac{4}{\delta R}\right) \quad \text{for } n = 1, \ldots, N-2 \\
d_n &\equiv \frac{\theta}{2}\left(-r\, n + \frac{1}{\delta R}\right). \quad \text{for } n = 1, \ldots, N-3.
\end{aligned}
$$

Consistent with the matrix $\boldsymbol{M}^{\rm up}_{\rm Cox}$, we have left the time step δt in the denominator in the above quantity b_n. The vector $\boldsymbol{b}^{m+1}$ is also closely tied to the matrix $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ in the same way as $\boldsymbol{q}^{m+1}$ is to $\boldsymbol{M}^{\rm up}_{\rm Cox}$ in the Cox model. The details are not repeated.

It is noted in [26, page 355] that the "positive definiteness" of $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ is related to the stability of the finite difference scheme. It is further noted on the next page of the same reference that this property of $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ will ensure the convergence of the PSOR method for all relaxation parameters between 0 and 2. Before we formally establish some key properties of the matrix $\boldsymbol{M}^{\rm nup}_{\rm Asi}$, we caution the reader when interpreting the "positive definiteness" property of this matrix.

The main point is that $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ is <u>not</u> a symmetric matrix. Thus the positive definiteness of this matrix can only be taken to mean that its symmetric part (i.e., $\boldsymbol{M}^{\rm nup}_{\rm Asi} + (\boldsymbol{M}^{\rm nup}_{\rm Asi})^T$) is positive definite. Although this usage is very common in the LCP literature, it is not at all in the numerical analysis literature. Chapter 5 in [5] gives a very detailed convergence analysis of many iterative methods for solving an LCP, including the PSOR scheme. To the best of our knowledge, we are not aware of a result in the LCP arena which states that the PSOR method converges for all relaxation parameters between 0 and 2 when the defining matrix of the LCP is asymmetric positive definite (and without further property). A precise convergence result will be stated later for the LCP (8).

With an upwind finite difference scheme applied to the system (5)–(7), a backward difference is employed to approximate $\partial H/\partial R$ if the factor $1 - rR$ is negative. This is in contrast to the no-upwind scheme where a forward difference is used regardless of the sign of $1 - rR$. Specifically, in conjunction with the θ-method, the approximation is as follows:

(a) if $1 - rR > 0$,

$$(1 - rR)\,\frac{\partial H}{\partial R}(R,t) \;\approx\; |\,1 - rR\,| \left[\theta\, \frac{4\,H(R+\delta R,t) - 3\,H(R,t) - H(R+2\,\delta R,t)}{2\,\delta R} + \right.$$

$$\left. (\,1-\theta\,)\, \frac{4\,H(R+\delta R,t+\delta t) - 3\,H(R,t+\delta t) - H(R+2\,\delta R,t+\delta t)}{2\,\delta R} \right]$$

(b) if $1 - rR < 0$,

$$(1 - rR)\frac{\partial H}{\partial R}(R,t) \;\approx\; |\,1 - rR\,| \left[\theta\, \frac{4\,H(R-\delta R,t) - 3\,H(R,t) - H(R-2\,\delta R,t)}{2\,\delta R} + \right.$$

$$\left. (\,1-\theta\,)\, \frac{4\,H(R-\delta R,t+\delta t) - 3\,H(R,t+\delta t) - H(R-2\,\delta R,t+\delta t)}{2\,\delta R} \right].$$

The matrix that defines the discretized LCP is

$$\boldsymbol{M}^{\rm up}_{\rm Asi} \;\equiv\; \begin{pmatrix} b'_1 & c'_1 & d'_1 & 0 & 0 & 0 & \cdots & 0 \\ a'_2 & b'_2 & c'_2 & d'_2 & 0 & 0 & \cdots & 0 \\ e'_3 & a'_3 & b'_3 & c'_3 & d'_3 & 0 & & 0 \\ 0 & e'_4 & a'_4 & b'_4 & c'_4 & d'_4 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & e'_{N-3} & a'_{N-3} & b'_{N-3} & c'_{N-3} & d'_{N-3} \\ 0 & 0 & \cdots & 0 & e'_{N-2} & a'_{N-2} & b'_{N-2} & c'_{N-2} \\ 0 & 0 & \cdots & 0 & 0 & e'_{N-1} & a'_{N-1} & b'_{N-1} \end{pmatrix}$$

is the $(N-1)$ by $(N-1)$ matrix with entries given by:

$$e'_n \;\equiv\; \frac{\theta\,(\,1 - r\,n\,\delta R)^-}{2\,\delta R} \qquad \text{for } n \,=\, 3,\ldots,N-1$$

$$a'_n \;\equiv\; -\,\frac{\theta}{2}\,\sigma^2\,n^2 - \frac{2\,\theta\,(\,1 - r\,n\,\delta R)^-}{\delta R} \qquad \text{for } n \,=\, 2,\ldots,N-1$$

$$b'_n \;\equiv\; \frac{1}{\delta t} + \theta\,\sigma^2\,n^2 + \frac{3\,\theta}{2\,\delta R}\,|\,1 - r\,n\,\delta R\,| \qquad \text{for } n \,=\, 1,\ldots,N-1$$

$$c'_n \;\equiv\; -\frac{\theta}{2}\,\sigma^2\,n^2 - \frac{2\,\theta\,(\,1 - r\,n\,\delta R)^+}{\delta R} \qquad \text{for } n \,=\, 1,\ldots,N-2$$

$$d'_n \;\equiv\; \frac{\theta\,(\,1 - r\,n\,\delta R)^+}{2\,\delta R} \qquad \text{for } n \,=\, 1,\ldots,N-3.$$

The constant vector $\boldsymbol{b}^{m+1}$ of the LCP will also be affected by the modified finite difference scheme. The reader can easily write down the components of this modified vector that corresponds to a given set of boundary conditions (or inputs from previous time step).

One obvious difference between the entries b'_n in $\boldsymbol{M}^{\rm up}_{\rm Asi}$ and b_n in $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ is that the former entry is always positive for all positive values of δR and δt and all integers n whereas the latter entry is positive if δR and δt satisfy a certain condition. The positivity of such diagonal entries has an important effect on the stabilization of the numerical scheme.

2.3 Models with transaction costs

We consider a vanilla American option with proportional transaction costs cast in the periodic rebalancing framework of Leland [20]; see also [26, Chapter 13] and [16]. Let $\delta t > 0$ be a finite fixed time step so that the portfolio is revised every δt; the random walk of the asset price in discrete time is assumed to be:

$$\delta S \,=\, \mu\, S\, \delta t + \sigma\, S\, \phi \sqrt{\delta t},$$

where ϕ is a random variable with a standardized normal distribution. Let $k > 0$ denote the investor's constant of proportionality so that the transaction cost is equal to $kS|\nu|$ if ν shares of the asset are bought ($\nu > 0$) or sold ($\nu < 0$). From the derivation in the cited references, we obtain the following partial differential complementarity system for this model with transaction costs:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2} - \sqrt{\frac{2}{\pi}}\, \frac{k\,\sigma}{\sqrt{\delta t}} \left| S^2\, \frac{\partial^2 V}{\partial S^2} \right| + r\, S\, \frac{\partial V}{\partial S} - r\, V \,\leq\, 0 \tag{9}$$

$$V(S,t) \,\geq\, \Lambda(S,t) \tag{10}$$

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2} - \sqrt{\frac{2}{\pi}}\, \frac{k\,\sigma}{\sqrt{\delta t}} \left| S^2\, \frac{\partial^2 V}{\partial S^2} \right| + r\, S\, \frac{\partial V}{\partial S} - r\, V \right) \left(V(S,t) - \Lambda(S,t) \right) \,=\, 0, \tag{11}$$

along with some appropriate boundary conditions. As noted in [26, 16], the above involves a nonlinear parabolic partial differential operator, mainly because of the modulus term $|\cdot|$. Actually, such nonlinearity is not a handicap at all because the absolute value function is a simple piecewise linear function which is very easy to deal with. In what follows, we employ a well-known technique from linear programming to convert the system (in its discretized form) to a finite-dimensional LCP. Specifically, we recall the fact that every real number x is equal to the difference of its nonnegative part x^+ and non-positive part x^- (i.e., $x = x^+ - x^-$) and the absolute value of x is equal to the sum of x^+ and x^-. Moreover, x^+ and x^- are "complementary" to each other; that is, $x^+ x^- = 0$.

The discretized LCPs

We introduce the change of variables:

$$S \,\equiv\, \exp(X) \quad \text{and} \quad \tau \,\equiv\, T - t,$$

and define the model constant:

$$K \,\equiv\, \frac{k}{\sigma} \sqrt{\frac{8}{\pi\, \delta t}}.$$

Letting

$$\hat{V}(X,\tau) \,\equiv\, V(S,t), \quad \text{and} \quad \hat{\Lambda}(X,\tau) \,\equiv\, \Lambda(S,t),$$

we can easily show that the function $\hat{V}$ satisfies the following modified differential complementarity system:

$$\frac{\partial \hat{V}}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 \hat{V}}{\partial X^2} - K\left|\frac{\partial^2 \hat{V}}{\partial X^2}\right|\right] - r\,\frac{\partial \hat{V}}{\partial X} + r\,\hat{V} \geq 0 \tag{12}$$

$$\hat{V}(X,\tau) \geq \hat{\Lambda}(X,\tau) \tag{13}$$

$$\left(\frac{\partial \hat{V}}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 \hat{V}}{\partial X^2} - K\left|\frac{\partial^2 \hat{V}}{\partial X^2}\right|\right] - r\,\frac{\partial \hat{V}}{\partial X} + r\,\hat{V}\right)\left(\hat{V}(X,\tau) - \hat{\Lambda}(X,\tau)\right) = 0, \tag{14}$$

For notational simplification, we drop the hat in $\hat{V}$ and $\hat{\Lambda}$. We may write

$$\left|\frac{\partial^2 V}{\partial X^2}(X,\tau)\right| = u^+ + u^-,$$

where u^+ and u^- satisfy

$$\frac{\partial^2 V}{\partial X^2}(X,\tau) = u^+ - u^-,$$

and

$$(u^+, u^-) \geq 0, \quad \text{and} \quad u^+ u^- = 0. \tag{15}$$

Substituting these expressions into (12)–(14), we obtain the following equivalent system:

$$\frac{\partial V}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 V}{\partial X^2} - K\left(\frac{\partial^2 V}{\partial X^2} + 2\,u^-\right)\right] - r\,\frac{\partial V}{\partial X} + r\,V \geq 0 \tag{16}$$

$$V(X,\tau) \geq \Lambda(X,\tau) \tag{17}$$

$$\left(\frac{\partial V}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 V}{\partial X^2} - K\left(\frac{\partial^2 V}{\partial X^2} + 2\,u^-\right)\right] - r\,\frac{\partial V}{\partial X} + r\,V\right)\left(V(X,\tau) - \Lambda(X,\tau)\right) = 0, \tag{18}$$

$$u^+ = \frac{\partial^2 V}{\partial X^2}(X,\tau) + u^-, \tag{19}$$

along with (15). Following the same finite difference scheme as in Subsection 2.1 (in particular, using a weighted central difference for the second partial derivative $\partial^2 V/\partial X^2$), letting

$$V_n^m \equiv V(n\,\delta X, m\,\delta\tau), \quad u_n^m \equiv u^-(n\,\delta X, m\,\delta\tau), \quad \text{and} \quad \Lambda_n^m \equiv \Lambda(n\,\delta X, m\,\delta\tau),$$

for $-N^- \leq n \leq N^+$ and $0 \leq m \leq M$, and defining the vectors

$$\boldsymbol{V}^m \equiv \begin{pmatrix} V^m_{-N^-+1} \\ V^m_{-N^-+2} \\ \vdots \\ V^m_{N^+-2} \\ V^m_{N^+-1} \end{pmatrix}, \quad \boldsymbol{u}^m \equiv \begin{pmatrix} u^m_{-N^-+1} \\ u^m_{-N^-+2} \\ \vdots \\ u^m_{N^+-2} \\ u^m_{N^+-1} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\Lambda}^m \equiv \begin{pmatrix} \Lambda^m_{-N^-+1} \\ \Lambda^m_{-N^-+2} \\ \vdots \\ \Lambda^m_{N^+-2} \\ \Lambda^m_{N^+-1} \end{pmatrix},$$

we obtain the discretized LCP at time $t = m\delta\tau$ which is stated in the pair of variables $(\boldsymbol{V}^m, \boldsymbol{u}^m)$:

$$0 \leq \left[\begin{pmatrix} \boldsymbol{V}^m \\ \boldsymbol{u}^m \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Lambda}^m \\ 0 \end{pmatrix} \right] \perp \left[\begin{pmatrix} \boldsymbol{q}^{m-1} \\ \boldsymbol{r}^{m-1} \end{pmatrix} + \boldsymbol{M}^{\text{cen}}_{\text{trn}} \begin{pmatrix} \boldsymbol{V}^m \\ \boldsymbol{u}^m \end{pmatrix} \right] \geq 0,$$

where $\boldsymbol{q}^{m-1}$ and $\boldsymbol{r}^{m-1}$ are appropriate constant vectors, and $\boldsymbol{M}^{\text{cen}}_{\text{trn}}$ is the partitioned matrix:

$$\boldsymbol{M}^{\text{cen}}_{\text{trn}} \equiv \begin{pmatrix} \boldsymbol{M}^{\text{cen}}_{11} & \sigma^2\, K\, \boldsymbol{I} \\ -\dfrac{\theta}{(\delta X)^2}\, \boldsymbol{M}^{\text{cen}}_{21} & \boldsymbol{I} \end{pmatrix}$$

with $\boldsymbol{I}$ being the identity matrix of order $(N^- + N^+ - 1)$, $\boldsymbol{M}^{\text{cen}}_{11}$ being the Toeplitz matrix of the same order:

$$\boldsymbol{M}^{\text{cen}}_{11} \equiv \begin{pmatrix} b & c & d & 0 & 0 & \cdots & 0 \\ a & b & c & d & 0 & \cdots & 0 \\ 0 & a & b & c & d & & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & & a & b & c & d \\ 0 & 0 & \cdots & 0 & a & b & c \\ 0 & 0 & \cdots & 0 & 0 & a & b \end{pmatrix}$$

whose entries are given by

$$\begin{aligned} a &\equiv -\frac{\theta\,\sigma^2\,(1-K)}{2\,(\delta X)^2} \\ b &\equiv \frac{1}{\delta\tau} + \frac{\theta\,\sigma^2\,(1-K)}{(\delta X)^2} + \frac{3\,\theta\,r}{2\,\delta X} + r \\ c &\equiv -\frac{\theta\,\sigma^2\,(1-K)}{2\,(\delta X)^2} - \frac{2\,\theta\,r}{\delta X} \\ d &\equiv \frac{\theta\,r}{2\,\delta X}, \end{aligned}$$

and $\boldsymbol{M}^{\text{cen}}_{21}$ being the tridiagonal matrix:

$$\boldsymbol{M}^{\text{cen}}_{21} \equiv \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & 0 & -1 & 2 \end{pmatrix}$$

For $K > 1$, the diagonal entry b could become negative for certain values of δX and $\delta\tau$. In order to ensure that this entry is always positive for all values of δX and $\delta\tau$, we can use a weighted backward difference approximation for the second partial derivative:

$$\begin{aligned}\frac{\partial^2 V}{\partial X^2}(X,\tau) \;&\approx\; \theta\,\frac{V(X,\tau) - 2\,V(X-\delta X,\tau) + V(X - 2\,\delta X,\tau)}{(\delta X)^2} + \\ &(1-\theta)\,\frac{V(X,\tau-\delta\tau) - 2\,V(X-\delta X,\tau-\delta\tau) + V(X-2\,\delta X,\tau-\delta\tau)}{(\delta X)^2};\end{aligned}$$

the matrix that defines the resulting LCP is:

$$M_{\text{trn}}^{\text{bac}} \equiv \begin{pmatrix} \boldsymbol{M}_{11}^{\text{bac}} & \sigma^2\,K\,\boldsymbol{I} \\ \dfrac{\theta}{(\delta X)^2}\,\boldsymbol{M}_{21}^{\text{bac}} & \boldsymbol{I} \end{pmatrix}$$

with $\boldsymbol{I}$ being the identity matrix as before, $\boldsymbol{M}_{11}^{\text{bac}}$ being the Toeplitz matrix of the same order:

$$\boldsymbol{M}_{11}^{\text{bac}} \equiv \begin{pmatrix} b & c & d & 0 & 0 & 0 & \cdots & 0 \\ a & b & c & d & 0 & 0 & \cdots & 0 \\ e & a & b & c & d & 0 & \cdots & 0 \\ 0 & e & a & b & c & d & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \\ 0 & & & e & a & b & c & d \\ 0 & 0 & \cdots & 0 & e & a & b & c \\ 0 & 0 & \cdots & 0 & 0 & e & a & b \end{pmatrix}$$

whose entries are given by

$$\begin{aligned} e &\equiv \frac{\theta\,\sigma^2\,(K-1)}{2\,(\delta X)^2} \\ a &\equiv -\frac{\theta\,\sigma^2\,(K-1)}{(\delta X)^2} \\ b &\equiv \frac{1}{\delta\tau} + \frac{\theta\,\sigma^2\,(K-1)}{2\,(\delta X)^2} + \frac{3\,\theta\,r}{2\,\delta X} + r \\ c &\equiv -\frac{2\,\theta\,r}{\delta X} \\ d &\equiv \frac{\theta\,r}{2\,\delta X}, \end{aligned}$$

and $\boldsymbol{M}_{21}^{\rm bac}$ being the lower triangular matrix:

$$\boldsymbol{M}_{21}^{\rm bac} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & & 0 \\ & \ddots & \ddots & \ddots & & \\ 0 & & 1 & -2 & 1 & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}$$

A unified approach

We propose a unified method of discretizing the PDCP defined by (9)–(11) that guarantees positive diagonal entries in the resulting LCP matrix $\boldsymbol{M}_{\rm trn}$, regardless of whether K is greater or smaller than unity. The idea is simply to interchange the roles of u^- and u^+. Specifically, we solve for u^- in terms of u^+ and the second partial derivative $\partial^2 V/\partial X^2$, and substitute u^- in (16) and (18). This results in the following equivalent system:

$$\frac{\partial V}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 V}{\partial X^2} - K\left(-\frac{\partial^2 V}{\partial X^2} + 2\,u^+\right)\right] - r\,\frac{\partial V}{\partial X} + r\,V \,\geq\, 0 \tag{20}$$

$$V(X,\tau) \,\geq\, \Lambda(X,\tau) \tag{21}$$

$$\left(\frac{\partial V}{\partial \tau} - \frac{\sigma^2}{2}\left[\frac{\partial^2 V}{\partial X^2} - K\left(-\frac{\partial^2 V}{\partial X^2} + 2\,u^+\right)\right] - r\,\frac{\partial V}{\partial X} + r\,V\right)(\,V(X,\tau) - \Lambda(X,\tau)\,) \,=\, 0, \tag{22}$$

$$u^- \,=\, -\frac{\partial^2 V}{\partial X^2}(X,\tau) + u^+ \tag{23}$$

$$(\,u^+,\,u^-\,) \,\geq\, 0, \quad \text{and} \quad u^+\,u^- \,=\, 0. \tag{24}$$

We use a weighted central difference to approximate the second partial derivative $\partial^2 V/\partial X^2$. The resulting LCP matrix is:

$$\boldsymbol{M}_{\rm trn}^{\rm uni} \equiv \begin{pmatrix} \boldsymbol{M}_{11}^{\rm uni} & \sigma^2\,K\,\boldsymbol{I} \\ \dfrac{\theta}{(\delta X)^2}\,\boldsymbol{M}_{21}^{\rm cen} & \boldsymbol{I} \end{pmatrix}$$

with $\boldsymbol{I}$ and $\boldsymbol{M}_{21}^{\rm cen}$ being the same matrices as before, and $\boldsymbol{M}_{11}^{\rm uni}$ being the following Toeplitz matrix of order $(N^+ + N^- - 1)$:

$$\boldsymbol{M}_{11}^{\rm uni} \equiv \begin{pmatrix} b & c & d & 0 & 0 & \cdots & 0 \\ a & b & c & d & 0 & \cdots & 0 \\ 0 & a & b & c & d & & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & & a & b & c & d \\ 0 & 0 & \cdots & 0 & a & b & c \\ 0 & 0 & \cdots & 0 & 0 & a & b \end{pmatrix}$$

whose entries are given by

$$\begin{aligned} a &\equiv -\frac{\theta\,\sigma^2\,(1+K)}{2\,(\delta X)^2} \\ b &\equiv \frac{1}{\delta\tau} + \frac{\theta\,\sigma^2\,(1+K)}{(\delta X)^2} + \frac{3\,\theta\,r}{2\,\delta X} + r \\ c &\equiv -\frac{\theta\,\sigma^2\,(1+K)}{2\,(\delta X)^2} - \frac{2\,\theta\,r}{\delta X} \\ d &\equiv \frac{\theta\,r}{2\,\delta X}. \end{aligned}$$

Note that the diagonal entry b is positive for all values of K ($K \geq 0$ by definition).

Other models of transaction costs

The absolute value function is piecewise linear in its argument. By exploiting this piecewise linearity property, we have derived a linear complementarity problem for an American option under Leland's local-time model of transaction costs. In [12], a general nonlinear function of the form $F(S, \partial^2 V/\partial S^2)$ was introduced as a unification of various transaction cost structures and hedging strategies. The resulting discretized problems become truly nonlinear complementarity problems which can be solved by a host of advanced numerical methods, such as those summarized in the survey article [14]. An alternative approach to include transaction costs in based on utility maximization [8, 9]; this leads to a "two-factor" complementarity problem of the "vertical" type [5]. The numerical solution of these other models with transaction costs is the subject of a separate study which is presently in progress.

2.4 A jump diffusion model

In a recent paper, Zhang [28] presented a detailed numerical analysis of American option pricing in the context of Merton's jump diffusion model [21]. Zhang used a first-order approximation of the convection term and obtained an LCP with a diagonally dominant Z-matrix under an appropriate choice of the discretization steps. We examine the same jump-diffusion model and discuss its numerical solution using a second-order finite difference scheme.

In Merton's jump diffusion model (where no transaction cost is assumed), the asset price satisfies the following stochastic process [21, 28]:

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW + d\left(\sum_{j=1}^{N_t} U_j\right)$$

where $(N_t)_{t\geq 0}$ is a Poisson process with parameter λ and $(U_j)_{j\geq 1}$ is a sequence of square integrable i.i.d. random variables, with values in the interval $(-1, \infty)$. For simplicity, we take the drift μ to be a constant; in particular, we assume that the asset pays no dividend. Making the logarithmic change of variable $x \equiv \log(S)$ and assuming that the random variable $Z_j \equiv \log(1 + U_j)$ has a probability density function $g(z)$, the value $u(t, x)$ of an American option with payoff $\psi(t, x)$ can be computed from the following complementarity system involving an integral-differential operator, where the integral is the result of the jump diffusion:

$$\frac{\partial u}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + \left(\mu - \frac{\sigma^2}{2}\right)\frac{\partial u}{\partial x} - r\,u + \lambda\left(\int_{-\infty}^{\infty} u(t, x+z)\, g(z)\, dz - u(t,x)\right) \leq 0$$

$$u(t,x) - \psi(t,x) \geq 0$$

$$\left[\frac{\partial u}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + \left(\mu - \frac{\sigma^2}{2}\right)\frac{\partial u}{\partial x} - r\,u + \lambda\left(\int_{-\infty}^{\infty} u(t, x+z)\, g(z)\, dz - u(t,x)\right)\right](u - \psi) = 0$$

along with suitable boundary conditions. In applying an upwind finite difference scheme, the sign of the constant $\mu - \sigma^2/2$ is the deciding factor of whether a forward (second order) or backward difference should be used to approximate the first partial derivative. In what follows, we consider only the case where this constant is positive and employ a forward difference. We approximate the infinite integral

$$\int_{-\infty}^{\infty} u(t, x+z)\, g(z)\, dz$$

by a finite sum, appended by a residual term which appears in the constant vector of the discretized LCP; see [28] for more details. The matrix that defines this LCP is:

$$\boldsymbol{M}_{\text{jmp}} \equiv \tilde{\boldsymbol{M}}_{\text{jmp}} + \boldsymbol{G}_{\text{jmp}}$$

where $\tilde{\boldsymbol{M}}_{\text{jmp}}$ is the matrix that is the result of discretizing the partial differential operator:

$$\mathcal{L}_{\text{jmp}} \equiv \frac{\partial}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2}{\partial S^2} + \left(\mu - \frac{\sigma^2}{2}\right)\frac{\partial}{\partial x} - r;$$

specifically,

$$\tilde{\boldsymbol{M}}_{\rm jmp} \equiv \begin{pmatrix} b & c & d & 0 & 0 & \cdots & 0 \\ a & b & c & d & 0 & \cdots & 0 \\ 0 & a & b & c & d & & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & & a & b & c & d \\ 0 & 0 & \cdots & 0 & a & b & c \\ 0 & 0 & \cdots & 0 & 0 & a & b \end{pmatrix}$$

where δt and δx are the discretization steps of time and state, respectively,

$$\begin{aligned} a &\equiv -\frac{\theta\,\sigma^2}{2\,(\delta x)^2} \\ b &\equiv \frac{1}{\delta t} + \frac{\theta\,\sigma^2}{(\delta x)^2} + \frac{3\,\theta\left(\mu - \frac{\sigma^2}{2}\right)}{2\,\delta x} + r \\ c &\equiv -\frac{\theta\,\sigma^2}{2\,(\delta x)^2} - \frac{2\,\theta\left(\mu - \frac{\sigma^2}{2}\right)}{\delta x} \\ d &\equiv \frac{\theta\left(\mu - \frac{\sigma^2}{2}\right)}{2\,\delta x}. \end{aligned}$$

The matrix $\boldsymbol{G}_{\rm jmp}$ is a Toeplitz matrix given by

$$\boldsymbol{G}_{\rm jmp} \equiv \lambda\,\boldsymbol{I} - \lambda\,\delta x \begin{pmatrix} g_0 & g_1 & g_2 & g_3 & \cdots & g_{N-1} \\ g_{-1} & g_0 & g_1 & g_2 & \cdots & g_{N-2} \\ g_{-2} & g_{-1} & g_0 & g_1 & \cdots & g_{N-3} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ g_{N-2} & \cdots & g_{-2} & g_{-1} & g_0 & g_1 \\ g_{N-1} & g_{N-2} & \cdots & g_{-2} & g_{-1} & g_0 \end{pmatrix}$$

where the constants g_i are positive weights derived from the probability density function $g(z)$ evaluated at the grid points. With δx sufficiently small, the matrix $\boldsymbol{G}_{\rm jmp}$ is both row and column diagonally dominant. In [28] the matrix $\boldsymbol{G}_{\rm jmp}$ is multiplied by a factor $\bar{\theta} \in (0,1]$ similar to the θ-factor in $\tilde{\boldsymbol{M}}_{\rm jmp}$. For simplicity, we have taken $\bar{\theta} = 1$.

3 Matrix Classes and the LCP

In this section, we present some important properties of the matrices

$$\boldsymbol{M}^{\rm up}_{\rm Cox}, \quad \boldsymbol{M}^{\rm up}_{\rm Asi}, \quad \boldsymbol{M}^{\rm cen}_{\rm trn}, \quad \boldsymbol{M}^{\rm uni}_{\rm trn}, \quad \text{and} \quad \boldsymbol{M}_{\rm jmp} \tag{25}$$

and summarize the roles of these properties in the solution of the LCP. We omit the discussion of the matrix $\boldsymbol{M}^{\rm nup}_{\rm Asi}$ because, among other reasons, it is necessary to impose some restrictive conditions on δt, δR, and N in order for the diagonal entries of this matrix to be positive, whereas no such conditions are needed for the matrix $\boldsymbol{M}^{\rm up}_{\rm Asi}$. Also we do not discuss $\boldsymbol{M}^{\rm bac}_{\rm trn}$ because it is similar to $\boldsymbol{M}^{\rm uni}_{\rm trn}$. We note that the diagonal entries of the matrices listed in (25), with the possible exception of $\boldsymbol{M}_{\rm jmp}$ (which is due to the discrete jump matrix $\boldsymbol{G}_{\rm jmp}$), are all positive for all values of the time and state discretization steps; this is the consequence of the upwind finite difference scheme. (The matrix $\boldsymbol{M}_{\rm jmp}$ will have the same property provided that δx is sufficiently small.) All of the matrices in (25) are not symmetric.

We formally introduce several matrix classes that are relevant to the option models discussed in this paper. Fundamental properties of these matrices can be found in [5]. Throughout the discussion, matrices are not assumed to be symmetric.

An $n \times n$ real matrix $\boldsymbol{M}$ is said to be *positive definite* if $\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x} > 0$ for all nonzero n-vectors $\boldsymbol{x}$. Let PD denote the class of positive definite matrices. The matrix $\boldsymbol{M}$ is said to be a P-matrix if all its principal minors are positive. Given the matrix $\boldsymbol{M}$, its *comparison matrix* is $\overline{\boldsymbol{M}}$ whose entries are defined by

$$\overline{\boldsymbol{M}}_{ij} \equiv \begin{cases} \;\;|\,\boldsymbol{M}_{ij}\,| & \text{if } i = j \\ -|\,\boldsymbol{M}_{ij}\,| & \text{if } i \neq j. \end{cases}$$

The matrix $\boldsymbol{M}$ is said to be an H-matrix if $\overline{\boldsymbol{M}}$ is a P-matrix. An H-matrix with positive diagonals is said to be an $\rm H_+$-matrix. The matrix $\boldsymbol{M}$ is said to an S-matrix if there exists a vector $\boldsymbol{x} > 0$ such that $\boldsymbol{M}\boldsymbol{x} > 0$. There are many characterizations of a P-matrix; among these is the following:

(a) A matrix $\boldsymbol{M}$ is a P-matrix if and only if it reverses the sign of no nonzero vectors; that is, if and only if the following implication holds:

$$[\,0 \neq \boldsymbol{x}\,] \;\Rightarrow\; [\,\exists\, k \text{ such that } \boldsymbol{x}_k(\boldsymbol{M}\boldsymbol{x})_k > 0\,].$$

A class of matrices broader than the class of P-matrices turns out to contain the two matrices $\boldsymbol{M}^{\rm bac}_{\rm trn}$ and $\boldsymbol{M}^{\rm uni}_{\rm trn}$. This is the class of strictly semi-monotone matrices, denoted E (for B.C. Eaves who is among the pioneers in the LCP arena). Specifically, a matrix $\boldsymbol{M}$ is an E-matrix if it satisfies the following implication:

$$[\,0 \neq \boldsymbol{x} \geq 0\,] \;\Rightarrow\; [\,\exists\, k \text{ such that } \boldsymbol{x}_k(\boldsymbol{M}\boldsymbol{x})_k > 0\,].$$

We summarize the key facts of the above matrix classes and their connection to the LCP:

$$0 \leq \boldsymbol{x} \perp (\,\boldsymbol{q} + \boldsymbol{M}\boldsymbol{x}\,) \geq 0; \tag{26}$$

these results can all be found in the book [5]. We use the same letter to denote a matrix property and the corresponding matrix class.

(b) A matrix $\boldsymbol{M}$ is an H-matrix if and only if $\overline{\boldsymbol{M}}$ is an S-matrix. (Because of this property, H-matrices coincide with the so-called "quasi-diagonally dominant" matrices.)

(c) The following inclusions of matrix classes are valid:

$$\rm PD \,\cup\, H_+ \,\subset\, P \,\subset\, E \,\subset\, S.$$

(d) A matrix $\boldsymbol{M}$ is a P-matrix if and only if the LCP (26) has a unique solution for all constant vectors $\boldsymbol{q}$; moreover, this solution is a Lipschitz continuous function of q (with $\boldsymbol{M}$ fixed).

(e) If $\boldsymbol{M}$ is an H_+-matrix, then the PSOR method for solving the LCP (26) converges for all values of $\omega \in (0, \bar{\omega})$, where

$$1 < \bar{\omega} \equiv 2 \min_i \frac{\boldsymbol{M}_{ii}\boldsymbol{d}_i}{\sum_j |\boldsymbol{M}_{ij}|\boldsymbol{d}_j} \leq 2$$

and $\boldsymbol{d}$ is a vector such that $\overline{\boldsymbol{M}}\boldsymbol{d} > 0$ (such a vector $\boldsymbol{d}$ must exist, by (b)); the above estimate of the upper bound $\bar{\omega}$ tends to be conservative in practice.

(f) If $\boldsymbol{M}$ is an $n \times n$ H_+-matrix, then for any vector $\boldsymbol{d}$ such that $\overline{\boldsymbol{M}}\boldsymbol{d} > 0$, the vector

$$\boldsymbol{p} \equiv \left(\frac{\boldsymbol{M} + \overline{\boldsymbol{M}}}{2} \right)^T \boldsymbol{d}$$

is an "n-step" vector of the LCP (26) for all vectors $\boldsymbol{q}$; that is, with the above vector $\boldsymbol{p}$, Algorithm 4.8.2 (the revised parametric principal pivoting (PPP) algorithm) in [5] computes the unique solution of the LCP in at most n pivots.

(g) If $\boldsymbol{M}$ is an E-matrix, then Lemke's algorithm will compute a solution of the LCP (26) for all vectors q, under the usual non-degeneracy assumption as in the simplex method of linear programming.

Contrary to the case of an H-matrix with positive diagonals, the convergence of a PSOR method for solving the LCP (26) with an asymmetric positive definite matrix $\boldsymbol{M}$ is not very well understood; at best, the convergence requires highly restrictive conditions. Indeed, it is not common in the LCP community to apply this method for solving such an LCP; instead, alternative methods such as Lemke's method or other recent algorithms can be used.

The implication of (f) is that an LCP with an H_+-matrix can be solved by a pivotal algorithm whose computational complexity is a low order polynomial of the dimension of the problem. Furthermore, any special structure of the defining matrix (such as the penta-diagonality in the matrices $\boldsymbol{M}^{\mathrm{up}}_{\mathrm{Cox}}$ and $\boldsymbol{M}^{\mathrm{up}}_{\mathrm{Asi}}$ or the special block structure in the matrix $\boldsymbol{M}^{\mathrm{cen}}_{\mathrm{trn}}$) can be profitably exploited in the solution process; often, this will result in substantial reduction in the computational efforts. Finally, unlike the case of a P-matrix, an LCP with an E-matrix can have multiple solutions.

3.1 Diagonal dominance, P- and E-property

Besides having positive diagonals, the matrices

$$\boldsymbol{M}^{\mathrm{up}}_{\mathrm{Cox}}, \quad \boldsymbol{M}^{\mathrm{up}}_{\mathrm{Asi}}, \quad \boldsymbol{M}^{\mathrm{cen}}_{11}, \quad \boldsymbol{M}^{\mathrm{uni}}_{11}, \quad \text{and} \quad \boldsymbol{M}_{\mathrm{jmp}}$$

are all (row) strictly diagonally dominant under mild conditions on the discretization steps. Thus all these matrices are H_+-matrices and thus P-matrices. Statements (a)–(f) are therefore applicable to the LCPs arising from the Cox model, the Asian model, and the jump diffusion model. Additional treatment is needed for the model with transaction costs because the matrices $\boldsymbol{M}^{\mathrm{cen}}_{11}$ and $\boldsymbol{M}^{\mathrm{uni}}_{11}$

are only the leading principal submatrices of the matrices $\boldsymbol{M}^{\text{cen}}_{\text{trn}}$ and $\boldsymbol{M}^{\text{uni}}_{\text{trn}}$ respectively, the latter matrices being the ones that define the corresponding LCPs.

We give in the proposition below the precise condition for the matrix $\boldsymbol{M}^{\text{up}}_{\text{Asi}}$ to have the desired diagonal dominance property; similar conditions for the other matrices can be obtained easily.

Proposition 1 *With $\bar{R} \equiv N\delta R$, if*

$$\theta \frac{\delta t}{\delta R} < \min\left(1, \ \frac{1}{|\, r\,\bar{R} - 1\,|}\right)$$

then $\boldsymbol{M}^{\text{up}}_{\text{Asi}}$ is row strictly diagonally dominant.

Proof. We need to show that for each $n = 0, 1, \ldots, N$, the quantity $b'_n - (|e'_n| + |a'_n| + |c'_n| + |d'_n|)$ is positive (although the entries e'_0, e'_1, d'_{N-1}, d'_N, a'_0, and c'_N are not present in $\boldsymbol{M}^{\text{up}}_{\text{Asi}}$, we can include them in the following verification). We have for each n,

$$|\, b'_n \,| - (\,|\, e'_n \,| + |\, a'_n \,| + |\, c'_n \,| + |\, d'_n \,|\,) \geq \frac{1}{\delta t} - \frac{\theta}{\delta R}\,|\, 1 - r\, n\, \delta R \,|.$$

It is now easy to verify that under the assumption the right-hand quantity is indeed positive. This completes the proof. Q.E.D.

Obviously, we can not expect the matrices $\boldsymbol{M}^{\text{cen}}_{\text{trn}}$ and $\boldsymbol{M}^{\text{uni}}_{\text{trn}}$ to be diagonally dominant. Instead, we show below that $\boldsymbol{M}^{\text{cen}}_{\text{trn}}$ is a P-matrix if ($K < 1$ and) $\boldsymbol{M}^{\text{cen}}_{11}$ is row strictly diagonally dominant; furthermore $\boldsymbol{M}^{\text{uni}}_{\text{trn}}$ (with no restriction on K) is an E-matrix as long as its leading principal block, $\boldsymbol{M}^{\text{uni}}_{11}$, is a P-matrix. We also give examples to show that $\boldsymbol{M}^{\text{uni}}_{\text{trn}}$ is not necessarily P-matrices even if it leading principal block is so. Thus, the LCP arising from the option model with transaction coefficient $K < 1$ has a unique solution that can be computed by a host of efficient algorithms (including one that has a "strongly polynomial" complexity); in contrast, the LCP arising from the option model with transaction coefficient $K > 1$ has a (possibly non-unique) solution that can be computed by Lemke's algorithm; for the latter model, a central difference scheme applied to the formulation (20)–(24) is recommended.

The proof of the asserted P-property of the matrix $\boldsymbol{M}^{\text{cen}}_{\text{trn}}$ requires us to review several basic facts about a general matrix in partitioned form:

$$\boldsymbol{M} \equiv \begin{bmatrix} \boldsymbol{M}_{11} & \boldsymbol{M}_{12} \\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22} \end{bmatrix} \tag{27}$$

where $\boldsymbol{M}_{11}$ and $\boldsymbol{M}_{22}$ are square matrices with $\boldsymbol{M}_{22}$ having a positive determinant. The matrix

$$\boldsymbol{M}_{11} - \boldsymbol{M}_{12}(\,\boldsymbol{M}_{22}\,)^{-1}\boldsymbol{M}_{21}$$

is called the Schur complement of $\boldsymbol{M}_{22}$ in $\boldsymbol{M}$ and denoted $(\boldsymbol{M}/\boldsymbol{M}_{22})$. From the Schur determinantal formula [5, Chapter 2]:

$$\det \boldsymbol{M} \;=\; \det \boldsymbol{M}_{22}\, \det(\boldsymbol{M}/\boldsymbol{M}_{22}),$$

it follows that the sign of $\det \boldsymbol{M}$ coincides with that of the determinant of the Schur complement of $\boldsymbol{M}_{22}$ in $\boldsymbol{M}$.

Proposition 2 *If*

$$\frac{1}{\delta\tau} + r - \frac{\theta\, r}{\delta X} > 0, \tag{28}$$

then $\boldsymbol{M}^{\rm cen}_{\rm trn}$ *is a P-matrix.*

Proof. To simplify the notation, we drop the subscript "trn" and superscript "cen" throughout the proof; in particular, we simply write $\boldsymbol{M}$ for $\boldsymbol{M}^{\rm cen}_{\rm trn}$. Note that $\boldsymbol{M}$ is of the partitioned form (27) with $\boldsymbol{M}_{11}$ being strictly row diagonally dominant (implied by assumption) with positive diagonals and $\boldsymbol{M}_{22}$ being an identity matrix. To show that $\boldsymbol{M}$ is a P-matrix, let $\boldsymbol{M}'$ denote an arbitrary principal submatrix of $\boldsymbol{M}$; we need to show that $\boldsymbol{M}'$ has positive determinant. We may write

$$\boldsymbol{M}' \equiv \begin{bmatrix} \boldsymbol{M}'_{11} & \boldsymbol{M}'_{12} \\ \boldsymbol{M}'_{21} & \boldsymbol{M}'_{22} \end{bmatrix}$$

where $\boldsymbol{M}'_{ii}$ is a principal submatrix of $\boldsymbol{M}_{ii}$ (for $i = 1, 2$) and $\boldsymbol{M}'_{12}$ ($\boldsymbol{M}'_{21}$) is a submatrix (not necessarily principal) of $\boldsymbol{M}_{12}$ ($\boldsymbol{M}_{21}$). If $\boldsymbol{M}'_{22}$ is vacuous (thus so are the two off diagonal blocks), then $\boldsymbol{M}'$ is itself a principal submatrix of $\boldsymbol{M}_{11}$. Since $\boldsymbol{M}_{11}$ is itself a P-matrix by the matrix inclusions mentioned in statement (c) above, it follows that $\det \boldsymbol{M}'$ is positive. Similarly, if $\boldsymbol{M}'_{11}$ is vacuous, the same conclusion trivially holds too.

We now consider the case where both diagonal blocks in $\boldsymbol{M}'$ are not vacuous. Since $\boldsymbol{M}'_{22}$ is itself an identity matrix, from the Schur determinantal formula, it follows that

$$\text{sign } \det \boldsymbol{M}' = \text{sign } \det(\boldsymbol{M}'_{11} - \boldsymbol{M}'_{12}\boldsymbol{M}'_{21}).$$

By the structure of the off diagonal blocks in $\boldsymbol{M}$, it is not difficult to see that $-\boldsymbol{M}'_{12}\boldsymbol{M}'_{21}$ is a (possibly not strictly) row diagonally dominant matrix whose nonzero diagonal entries are all positive. Consequently, the matrix

$$\boldsymbol{M}'_{11} - \boldsymbol{M}'_{12}\boldsymbol{M}'_{21}$$

remains strictly row diagonally dominant with positive diagonals; hence its determinant is positive. Consequently $\det \boldsymbol{M}'$ is positive. Q.E.D.

The next result concerns the matrix $\boldsymbol{M}^{\rm uni}_{\rm trn}$. It is easy to show that under the same condition (28), $\boldsymbol{M}^{\rm uni}_{11}$ is row strictly diagonally dominant with positive diagonals.

Proposition 3 *If (28) holds, then* $\boldsymbol{M}^{\rm uni}_{\rm trn}$ *is an E-matrix.*

Proof. With $\boldsymbol{M}^{\rm uni}_{11}$ being a P-matrix (because it is row strictly diagonally dominant with positive diagonals) $\boldsymbol{M}^{\rm uni}_{\rm trn}$ is a partition matrix of the form (27) where $\boldsymbol{M}_{11}$ is a P-matrix, $\boldsymbol{M}_{12}$ is a positive multiple of the identity matrix, and $\boldsymbol{M}_{22}$ is the identity matrix. It remains to show that any partitioned matrix having these properties must be an E-matrix. For this purpose, let $(\boldsymbol{x}, \boldsymbol{y})$ be a pair of nonzero, nonnegative vectors. On the one hand, if $\boldsymbol{x}$ is nonzero, then the P-property of $\boldsymbol{M}_{11}$ implies (cf. statement (a) above) the existence of an index k such that

$$\boldsymbol{x}_k > 0 \quad \text{and} \quad (\boldsymbol{M}_{11}\boldsymbol{x})_k > 0.$$

Since $\boldsymbol{y}$ is nonnegative, we clearly have

$$(\boldsymbol{M}_{11}\boldsymbol{x} + \boldsymbol{M}_{12}\boldsymbol{y})_k > 0.$$

On the other hand, if $\boldsymbol{x} = 0$, then $\boldsymbol{y}$ must be nonzero; so there exists an index ℓ such that

$$\boldsymbol{y}_\ell > 0 \quad \text{and} \quad (\boldsymbol{M}_{21}\boldsymbol{x} + \boldsymbol{M}_{22}\boldsymbol{y})_\ell > 0.$$

Consequently, the E-property of $\boldsymbol{M}$ follows. Q.E.D.

Incidentally, the condition (28) does not imply that $\boldsymbol{M}_{\text{trn}}^{\text{bac}}$ is diagonally dominant; this is illustrated in the example below. The same example also shows that $\boldsymbol{M}_{\text{trn}}^{\text{bac}}$ and $\boldsymbol{M}_{\text{trn}}^{\text{uni}}$ are in general not P-matrices, although the submatrices $\boldsymbol{M}_{11}^{\text{bac}}$ and $\boldsymbol{M}_{11}^{\text{uni}}$ are both P-matrices.

Example. Let

$$T = 1, \qquad r = 0.05, \qquad \sigma = 0.4, \quad K = 1.3$$
$$\delta\tau = 0.25, \quad \delta X = 0.01, \quad N^- = 2, \quad N^+ = 2.$$

Condition (28) holds with these values. We have

$$\boldsymbol{M}_{\text{trn}}^{\text{bac}} = \begin{bmatrix} 127.8 & -5 & 1.25 & 0.208 & 0 & 0 \\ -240 & 127.8 & -5 & 0 & 0.208 & 0 \\ 120 & -240 & 127.8 & 0 & 0 & 0.208 \\ 50 & 0 & 0 & 1 & 0 & 0 \\ -100 & 50 & 0 & 0 & 1 & 0 \\ 50 & -100 & 50 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\boldsymbol{M}_{\text{trn}}^{\text{uni}} = \begin{bmatrix} 1847.8 & -925 & 1.25 & 0.208 & 0 & 0 \\ -920 & 1847.8 & -925 & 0 & 0.208 & 0 \\ 0 & -920 & 1847.8 & 0 & 0 & 0.208 \\ 100 & -50 & 0 & 1 & 0 & 0 \\ -50 & -100 & 50 & 0 & 1 & 0 \\ 0 & -50 & 100 & 0 & 0 & 1 \end{bmatrix};$$

both matrices have positive diagonals and negative determinants. The leading principal submatrix of order 3 in $\boldsymbol{M}_{\text{trn}}^{\text{bac}}$ is not diagonally dominant, whereas that in $\boldsymbol{M}_{\text{trn}}^{\text{uni}}$ is row strictly diagonally dominant. Nevertheless, both of these principal submatrices are P-matrices; the matrices $\boldsymbol{M}_{\text{trn}}^{\text{bac}}$ and $\boldsymbol{M}_{\text{trn}}^{\text{uni}}$ are both E-matrices.

4 Numerical Results

In this section, we report computational results pertaining to the finite difference schemes and the numerical solutions of the discretized LCPs. Details of the algorithms employed herein can be found in reference [5]. Three sets of results are presented, each in a separate subsection. The first set of results aims to demonstrate the benefit of the upwind finite difference scheme (proposed herein) versus the no-upwind scheme (used in [13]) for solving the Cox model and the continuously-sampled average strike option model using the PSOR method. The second set of results aims to demonstrate the effectiveness of Lemke's algorithm for solving options with transaction cost. A related aim is to confirm the benefit of the unified formulation for numerical purposes. The third set of results aims to demonstrate the effectiveness of the PPP algorithm for solving the jump diffusion model using the second-order finite difference scheme. Together, these results show that all the (discretized) option models discussed in the previous section can be solved efficiently by sound LCP methods.

4.1 The Cox model and average strike option

We explore the advantages of using the upwind finite difference scheme versus the no-upwind scheme for solving the Cox model and the continuously sampled average strike option model. As we have mentioned, the former scheme guarantees the positivity of the diagonals of the matrices $\boldsymbol{M}^{\text{up}}_{\text{Asi}}$ and $\boldsymbol{M}^{\text{up}}_{\text{Cox}}$; whereas the latter scheme could lead to matrices with negative diagonals. In addition, as can be seen from Tables 1–3, the upwind scheme also renders the matrices $\boldsymbol{M}^{\text{up}}_{\text{Asi}}$ and $\boldsymbol{M}^{\text{up}}_{\text{Cox}}$ positive definite; whereas this is not the case with the no-upwind scheme.

For these two models, the matrices defining the LCPs are very sparse; thus a natural candidate for solve these LCPs is the PSOR method. The positive definiteness of the matrices is actually important for the convergence of PSOR. This consideration provides a sound reason for favoring the upwind scheme.

The columns in Tables 1–3 are self-explanatory. They give the values of the problem parameters and the relaxation parameter (ω) in the PSOR method we have used in the experiment. The four columns labeled "upwind", "dd", "pd", and "diag" refer, respectively, to whether upwind or no-upwind is used, whether the corresponding matrix is diagonally dominant, positive definite, and whether the diagonals are positive or not. The last column labeled "convergence" refers to the "average" number of iterations required by the PSOR method to converge under the following termination rule:

$$\text{residual} \ \equiv \| \min(\boldsymbol{x}, \boldsymbol{q} + \boldsymbol{M}\boldsymbol{x}) \|_\infty \leq 10^{-8}, \quad \text{or} \quad \text{max iteration} \ > 500,$$

where "min" is the component-wise minimum operator of two vectors. Each row in the tables corresponds to the solution of $T/\delta t$ (number of time steps) LCPs each of size N; the entries in the "convergence" column are the average PSOR iterations (over time steps) in solving these LCPs in each row, rounded to the nearest tenth (thus the "+" symbol). The symbol "∞" in the convergence column means that the method fails to terminate under the above rule; "erratic behavior" refers to runs where the method behaves unpredictably; more specifically, throughout the solution procedure, the residual was increasing and reached very high values; but all of a sudden (after many iterations), it dropped below the prescribed tolerance and the method terminated. This erratic behavior can be attributed to the fact that the diagonal entries of the matrix are not positive; thus there is no basis to expect reasonable performance of the PSOR method for solving the LCPs. (Similar results are also obtained with other values of ω.) For the Cox model, we use the dividend function $D(S,t) = D_0 S$, where D_0 is a constant.

Note that we have only tested cases where the volatility parameter σ is relatively small. These are the cases where the no-upwind matrices will easily have negative diagonals. Without the positivity of the diagonals, PSOR fails to converge in many cases. From these tables, we conclude that the upwind finite difference scheme is generally superior to the no-upwind scheme for solving the continuously-sampled average strike option and the Cox model.

T	r	σ	δt	δS	D_0	α	upwind	dd	pd	diag	convergence
1	0.04	0.03	0.25	0.25	0.07	1.5	no	-	-	+	20+
							yes	+	+	+	10+
1	0.04	0.03	0.5	1.0	0.07	1.5	no	-	-	-	∞
							yes	-	+	+	10+
1	0.04	0.01	0.25	0.25	0.07	1.5	no	-	-	-	20+
							yes	-	+	+	10+
2	0.04	0.04	0.5	1.0	0.07	1.4	no	-	-	-	∞
							yes	-	+	+	10+
2	0.09	0.04	0.5	1.0	0.07	1.4	no	+	+	+	10+
							yes	+	+	+	10+

Table 1: The Cox model: $N = 400$ and $\omega = 1.2$.

T	r	σ	δt	δR	N	ω	upwind	dd	pd	diag	convergence
1	0.04	0.1	0.25	0.25	400	1.2	no	+	+	+	10+
							yes	-	+	+	10+
1	0.04	0.001	0.25	0.25	400	1.2	no	-	-	-	10+
							yes	-	+	+	10+
1	0.04	0.001	0.25	0.25	400	1.9	no	-	-	-	160+
							yes	-	+	+	160+
1	0.04	0.01	0.25	0.25	400	1.2	no	-	-	+	10+
							yes	-	+	+	10+
1	0.04	0.01	0.25	0.25	400	1.9	no	-	-	-	140+
							yes	-	+	+	140+
1	0.05	0.01	0.25	0.25	400	1.2	no	-	-	-	∞
							yes	-	+	+	10+
1	0.05	0.01	0.25	0.25	400	1.9	no	-	-	-	erratic behavior
							yes	-	+	+	140+
1	0.09	0.02	0.25	0.25	800	1.2	no	-	-	+	∞
							yes	-	+	+	10+

Table 2: Continuously-sampled average strike option.

For the Cox model, we also compute the value of put options for various values of α. In all cases, upwind finite difference is used. The parameters are given as follows: the time to expiration $T = 1$, the strike price $E = 10$, the interest rate $r = 0.04$, the volatility $\sigma = 0.4$, the time step $\delta t = 0.25$, the state step $\delta S = 0.4$. The results are plotted in Figure 1.

T	r	σ	δt	δR	N	ω	upwind	dd	pd	diag	Convergence
1.5	0.09	0.03	0.5	1.0	400	1.2	no	-	-	+	∞
							yes	-	+	+	10+
1.5	0.09	0.03	0.5	0.75	400	1.2	no	-	-	+	∞
							yes	-	+	+	10+
2	0.09	0.02	0.25	0.25	400	1.2	no	-	-	+	∞
							yes	-	+	+	10+
2	0.09	0.02	0.25	0.25	1000	1.2	no	-	-	+	∞
							yes	-	+	+	10+
2	0.09	0.02	0.5	0.25	400	1.2	no	-	-	-	∞
							yes	-	+	+	10+
2	0.09	0.02	0.5	0.25	1000	1.2	no	-	-	-	∞
							yes	-	+	+	10+
5	0.05	0.01	0.25	0.25	400	1.2	no	-	-	-	erratic behavior
							yes	-	+	+	10+
5	0.05	0.01	0.25	0.25	1000	1.2	no	-	-	-	erratic behavior
							yes	-	+	+	16+

Table 3: Continuously-sampled average strike option.

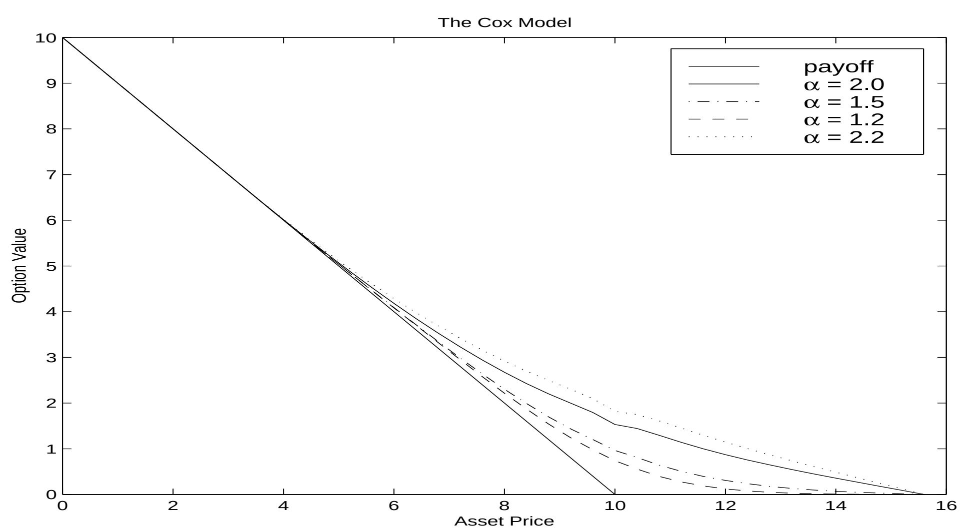

Figure 1: American put option values given by the Cox model

4.2 Options with transaction cost

Lemke's algorithm (Algorithm 4.4.5 in [5]) has been applied to solve the various formulations of the option models with transaction cost. The problem parameters are given as follows: the time to expiration $T = 0.5$, the interest rate $r = 0.10$, the strike price $E = 45$, the volatility $\sigma = 0.40$, the time step $\delta\tau = 0.1$, the state step $\delta x = 0.1$. For $K = 0.25$, we plot the value of the American put option obtained using the centered-difference and using the unified approach in Figures 2 and 3. In each figure, the solid lines denote the payoff at expiration and the vanilla American put option value. The dashed line is the put option value in the presence of transaction cost. If, instead, we apply the backward-difference to this problem, we do not obtain any meaningful result as expected. For $K > 1$, we also test Lemke's algorithm on each of the three approaches: unified, centered-difference, and backward-difference. In at least one case, though rather artificial, when $K = 15$, the centered-difference formulation causes Lemke's algorithm to fail while the unified formulation successfully yield a solution to the problem.

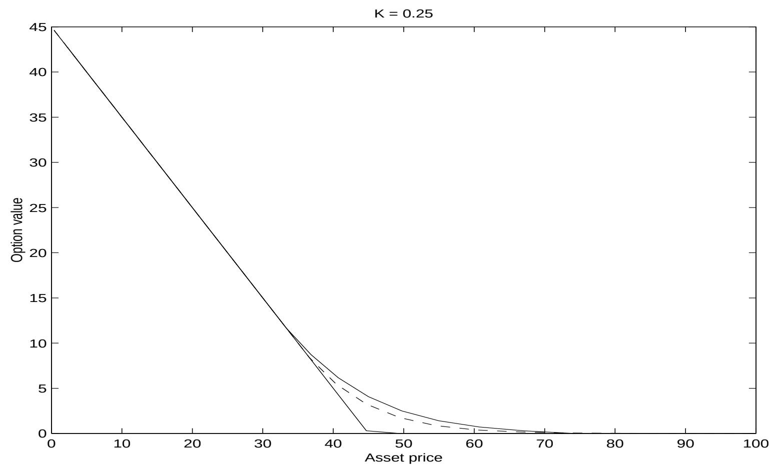

Figure 2: American put option with transaction cost: centered-difference, $K = 0.25$

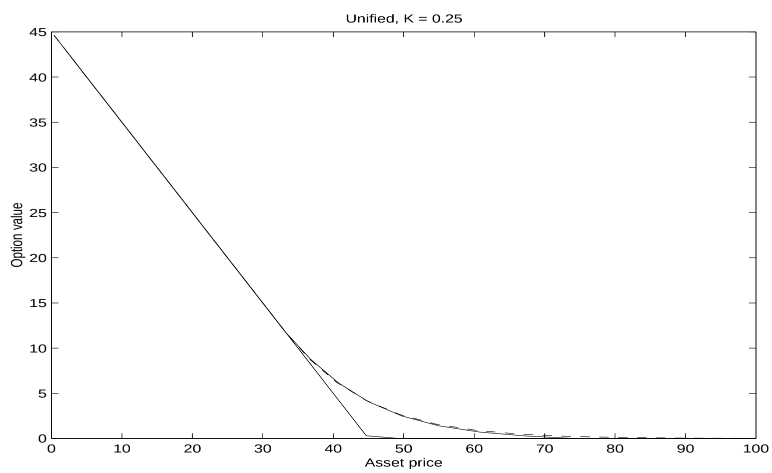

Figure 3: American put option with transaction cost: unified formulation, $K = 0.25$

4.3 Jump diffusion model

We have applied the revised PPP algorithm (Algorithm 4.8.2 in [5]) to solve the jump diffusion model. As in Zhang [28], we assume the random variable $1 + U_j$ follows a lognormal distribution. The problem parameters are as follows: the time to expiration $T = 0.2$, the interest rate $r = 0.04$, the strike price $E = 2$, the volatility $\sigma = 0.3$, the time step $\delta t = 0.05$, $\mathrm{E}[U_1] = 0$, $\mathrm{Var}(U_1) = 0.04$, and the Poisson parameter $\lambda = 1.5$. We test the PPP algorithm with matrices of difference sizes: while fixing $N^- \delta x = N^+ \delta x = 2$, we solve four problems for $\delta x = 0.1$, 0.05, 0.02, and 0.01. Hence, the order of the matrix $\boldsymbol{M}_{\mathrm{jmp}}$ is 41, 81, 201, and 401, respectively. For each problem, the solution obtained at each time step is plotted in Figures 4–7, with the lowest curve representing the payoff at expiration and the highest curve representing the value of option at time zero. Note that the reported number of iterations in each case is the average over the 4 ($= T/\delta t$) time steps.

We make several computational remarks about this set of experiments. The parameters of the problems are such that the matrix $\boldsymbol{M}_{\mathrm{jmp}}$ is row strictly diagonally dominant; thus statement (f) in Section 3 is applicable and the indicated vector $\boldsymbol{p}$ is used in the PPP algorithm. Thus the algorithm is expected to terminate in no more than $N^- + N^+ - 1$ iterations; indeed it does so in less than half this number in all runs. The Toeplitz structure of the matrix $\boldsymbol{M}_{\mathrm{jmp}}$ can be profitably used in the implementation to improve the computational efficiency (e.g., a Toeplitz linear solver can be used for solving the linear equations in each pivot step). Nevertheless, our present implementation has not taken advantage of such structure.

One referee has pointed out certain non-convexity in Figures 6 and 7. Although Brennan and Schwartz [2] have provided some theoretical results on the convexity of the value of a vanilla American put option, we are not aware of a theoretical justification that the value of an American put in the jump diffusion model must be convex in the asset price. The non-convexity of the curves in these two figures (where δx is somewhat small) may also be related to the smallness of the discretization step; the curves appear to be convex in Figures 4 and 5 where δx is larger.

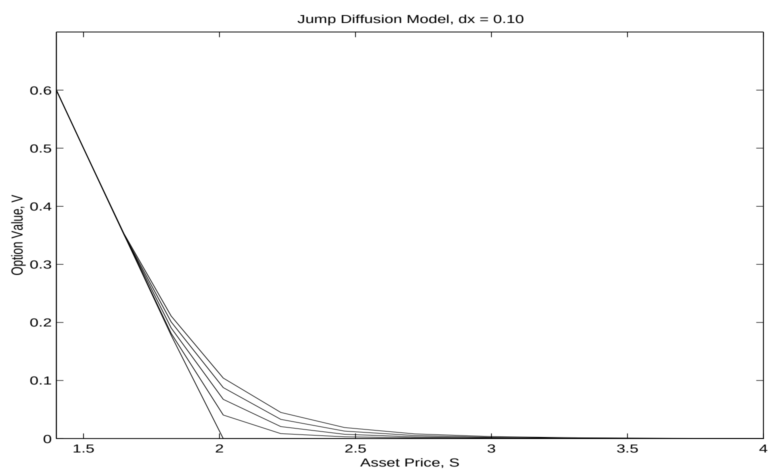

Figure 4: Jump diffusion model: $\delta x = 0.1$, number of pivots by PPP = 15

5 Conclusion

In this paper, we have examined in detail American option models of several types: Cox's model, the continuously sampled average strike option, options with transaction cost, and finally a jump diffusion model. Using the linear complementarity approach, we have obtained numerical solutions to these models via discretized LCPs through a time-stepping scheme. For Cox's model and the average strike option, we have introduced a second-order upwind finite difference scheme to deal with the convection term in the PDCP which may vary in sign. Numerical experiments show

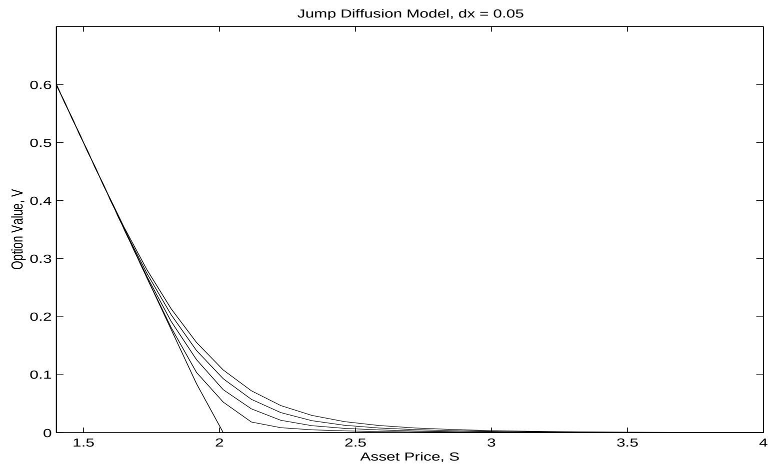

Figure 5: Jump diffusion model: $\delta x = 0.05$, number of pivots by PPP = 30

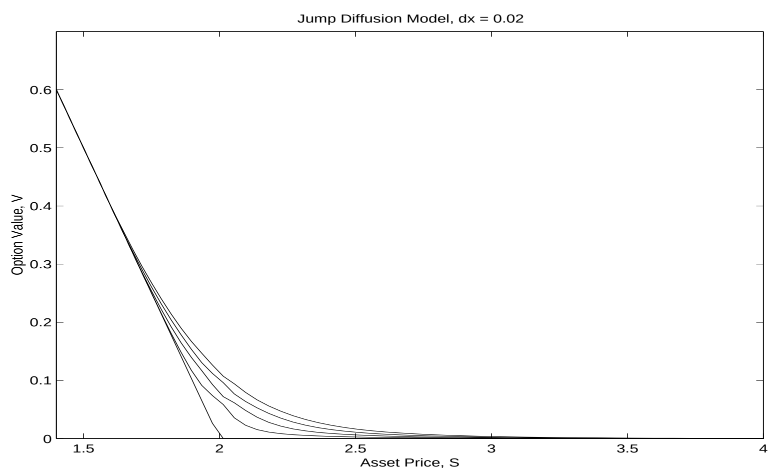

Figure 6: Jump diffusion model: $\delta x = 0.02$, number of pivots by PPP = 76

that the upwind finite difference scheme is preferred to to that without upwind, especially in the case of low volatilities. For American options with transaction costs, where the partial differential operator is initially nonlinear, we have applied a simple technique from linear programming to convert it into an LCP. The resulting LCP has a highly structured matrix, and we have established its properties and the solvability of the LCP by Lemke's algorithm. Finally, we have looked at an American option model that incorporates a jump diffusion process. Using a second-order finite difference approximation for the convection term, we have solved the model numerically via the PPP algorithm, which is guaranteed to terminate in n steps, where n is the order of the defining matrix of the discretized LCP.

In conclusion, our study has demonstrated the validity and the versatility of LCP methods for solving American option models of different types. We are confident that these methods can be similarly applied to such extensions as two-factor models, models with stochastic volatilities, and bond option models. Further investigation of the LCP approach applied to these and other

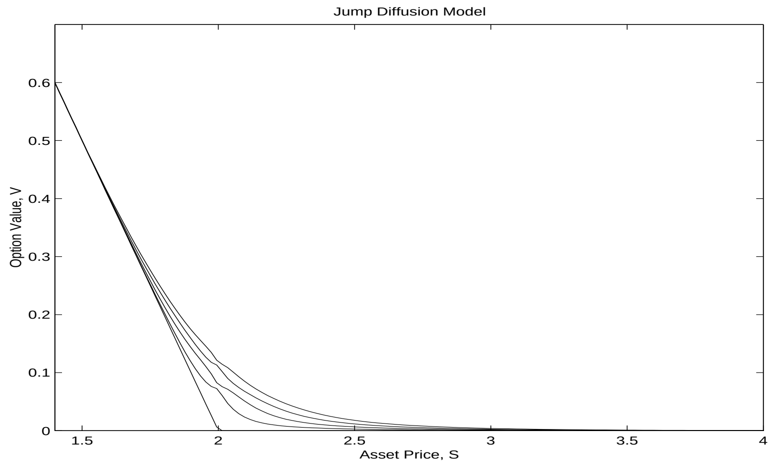

Figure 7: Jump diffusion model: $\delta x = 0.01$, number of pivots by PPP = 155

extended models is presently in progress and the findings will be reported in due time.

References

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities", *Journal of Political Economy* 81 (1973) 637–654.

[2] M.J. Brennan and E.S. Schwartz, "The valuation of American put option", *The Journal of Finance* 32 (1977) 449–462.

[3] M. Broadie and J. Detemple, "Recent advances in American option pricing", *Numerical Methods in Finance*, eds: L.C.G. Rogers and D. Talay, Cambridge University Press, Cambridge (1997) 43-66.

[4] P.G. Ciarlet, "Discrete maximum principle for finite-difference operators", *Aequationes Mathematicae* 4 (1970) 338-352.

[5] R.W. Cottle, J.S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston (1992).

[6] J.C. Cox, "The constant elasticity of variance option pricing model", *Journal of Portfolio Management*, Special Issue (1996) 15–17.

[7] J.C. Cox and S.A. Ross, "The valuation of options for alternative stochastic processes", *Journal of Financial Economics* 3 (1976) 145–166.

[8] M.H.A. Davis and J.M.C. Clark, "A note on super-replicating strategies", *Philosophical Transactions of the Royal Society of London, A* 347 (1994) 485–494.

[9] M.H.A. Davis and T. Zariphopoulou, "American options and transaction fees", in M.H.A. Davis, D. Duffie, W.H. Fleming, and S.E. Shreve, editors, *Mathematical Finance*, IMA Volumes in Mathematics and its Applications, # 165, Springer Verlag (1995) 47–61.

[10] M.A.H. Dempster and J.P. Hutton, "Fast numerical valuation of American, exotic and complex options", *Applied Mathematical Finance* 4 (1997) 1–20.

[11] M.A.H. Dempster and J.P. Hutton, "Pricing American stock options by linear programming", *Mathematical Finance*, forthcoming.

[12] J.N. Dewynne, A.E. Whalley, and P. Wilmott, "Path-dependent options and transaction costs", *Philosophical Transactions of the Royal Society of London, A* 347 (1994) 517–529.

[13] J.N. Dewynne and P. Wilmott, "Asian options as linear complementarity problems: analysis and finite-difference solutions", *Advances in Futures and Options Research* 8 (1995) 145–173.

[14] M.C. Ferris and J.S. Pang, "Engineering and economic applications of complementarity problems", *SIAM Review* 39 (1997) 669–713

[15] R. Gibson, *Option Valuation: Analyzing and Pricing Standardized Option Contracts*, McGraw-Hill, Inc. New York (1991).

[16] T. Hoggard, A.E. Whalley, and P. Wilmott, "Hedging option portfolios in the presence of transaction costs", *Advances in Futures and Options Research* 7 (1994) 21–35.

[17] T. Ikeda, *Maximum Principle in Finite Element Models for Convection-Diffusion Phenomena*, Lecture Notes in Numerical and Applied Analysis Vol. 4, North-Holland, Amsterdam (1983).

[18] P. Jaillet, D. Lamberton, and B. Lapeyre, "Variational inequalities and the pricing of American options", *Acta Applicandae Mathematicae* 21 (1990) 263–289.

[19] Y.A. Kuznetsov, P. Neittaanmäki, and P. Tarvainen, "Overlapping block methods for obstacle problems with convection-diffusion operators", in M.C. Ferris and J.S. Pang, editors, *Complementarity and Variational Problems: State of the Art*, SIAM Publications, Philadelphia (1997) pp. 165–180.

[20] H.E. Leland, "Option pricing and replication with transaction costs", *Journal of Finance* 40 (1985) 1283–1301.

[21] R.C. Merton, "Option pricing when underlying stock returns are discontinuous", *Journal of Financial Economics* 3 (1976) 125–144.

[22] K.W. Morton, *Numerical Solution of Convection-Diffusion Problems*, Chapman & Hall, London (1996).

[23] R. Myneni, "The pricing of the American option", *The Annals of Applied Probability* 2 (1992) 1–23.

[24] D.B. Spalding, "A novel finite difference formulation for differential expressions involving both first and second derivatives", *International Journal for Numerical Methods in Engineering* 4 (1972) 551–559.

[25] P. Tarvainen, "Block relaxation methods for algebraic obstacle problems with M-matrices: theory and applications", Ph.D. thesis, Department of Mathematics, University of Jyväskylä, Finland (1994).

[26] P. Wilmott, J.N. Dewynne and S. Howison, *Option Pricing: Mathematical models and computation*, Oxford Financial Press, Oxford, United Kingdom (1993).

[27] L.B. Zhang, "A second-order upwind finite difference scheme for the steady Navier-Stokes equations in primitive variables in a driven cavity with a multigrid solver", *RAIRO Mathematical Modelling and Numerical Analysis* 24 (1990) 133–150.

[28] X.L. Zhang, "Numerical analysis of American option pricing in a jump-diffusion model", *Mathematics of Operations Research* 22 (1997) 668–690.