A NEWTON METHOD FOR AMERICAN OPTION PRICING

THOMAS F. COLEMAN∗, YUYING LI∗, AND ARUN VERMA∗

December 13, 1999

Abstract. The variational inequality formulation provides a mechanism to determine both the option value and the early exercise curve implicitly [17]. Standard finite difference approximation typically leads to linear complementarity problems with tridiagonal coefficient matrices. The second order upwind finite difference formulation gives rise to finite dimensional linear complementarity problems with nontridiagonal matrices, whereas the upstream weighting finite difference approach with the van Leer flux limiter for the convection term [19, 22] yields nonlinear complementarity problems.

We propose a Newton type interior-point method for solving discretized complementarity/variational inequality problems that arise in the American option valuation. We illustrate that the proposed method on average solves a discretized problem in $2 \sim 5$ iterations to an appropriate accuracy. More importantly, the average number of iterations required does not seem to depend on the number of discretization points in the spatial dimension; the average number of iterations actually decreases as the time discretization becomes finer.

The arbitrage condition for the fair value of the American option requires that the delta hedge factor be continuous. We investigate continuity of the delta factor approximation using the complementarity approach, the binomial method, and the explicit payoff method. We illustrate that, while the (implicit finite difference) complementarity approach yields continuous delta hedge factors, both the binomial method and the explicit payoff method (with the implicit finite difference) yield discontinuous delta approximations. Hence the early exercise curve computed from the binomial method and the explicit payoff method can be inaccurate. In addition, it is demonstrated that the delta factor computed using the Crank-Nicolson method with complementarity approach oscillates around the early exercise curve.

∗ Computer Science Department and Cornell Theory Center, Cornell University, Ithaca, NY 14850. Research partially supported by the Applied Mathematical Sciences Research Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under grant DE-FG02-90ER25013.A000, NSF through grant DMS-9805602 and ONR through grant N00014-96-1-0050, and the Finance Industry Solutions Center, an SGI/Cornell joint venture in New York.

1. Introduction. An American option gives the holder of the option the right, but not the obligation, to buy or sell the underlying for a fixed exercise price any time before expiry. Valuation of the American option is an optimal stopping (or a free boundary) problem; it is significantly more complex than the European option pricing. Computing the fair value of an American option requires determination of the early exercise curve. This early exercise curve divides the domain of the underlying asset price and time into a continuation region $\mathcal{C}$ and a stopping region $\mathcal{S}$; it is optimal to exercise in $\mathcal{S}$ while continually hold the contract in $\mathcal{C}$. This early exercise curve is characterized by two boundary conditions. First, the option value is equal to the payoff at the early exercise curve. Second, the delta hedge factor is continuous at the early exercise curve.

American option valuation has been an active research area; many methods have been proposed to approximate American option values. Reviews of these methods can be found in Broadie and Detemple [4, 5]. A method frequently used in practice is the binomial method proposed by Cox, Ross and Rubinstein [9]. Convergence of this method for pricing American options is established in Amin and Khanna [1]. Another popular method in the partial differential equation framework is to approximate the American option value by simply taking the maximum of the continuation value and the payoff; this method is referred to as the *explicit payoff* method in this paper. Convergence properties of this method are, however, unclear.

A partial differential equation framework for option pricing has the advantage of providing the entire option value surface as well as the hedge factors. This can be useful for risk management, e.g., calculating VaR. Brennan and Schwartz [3] introduce a simple procedure using the standard implicit finite difference method for the classical Black-Scholes partial differential operator. Convergence of the Brenan and Schwartz method is established by Jaillet, Lamberton and Lapeyre in [17].

It is also shown, in [17], that determination of the early exercise curve can be made implicit with a variational inequality formulation in a generalized Black-Scholes framework. The variational inequality formulation for the American option in a jump diffusion model is analyzed in [21]. Advantages of the variational inequality approach include established convergence of computational methods in both the option values as well as the hedge factors [17, 21].

Unfortunately, the Brenan and Schwartz [3] method can be applied only when the standard finite difference approximation, which uses central difference to approximate the first order derivative in the spatial dimension, is used. This finite difference approximation generates discretized linear complementarity problems with tridiagonal coefficient matrices and negative off diagonals; the method of Brenan and Schwartz [3] explicitly exploits this special structure.

Linear complementarity problems with pentadiagonal matrices arise when the second order upwind finite difference approximation is used [15]. In addition, linear complementarity problems with non-tridiagonal coefficient matrices can arise in different asset pricing models, e.g., a jump diffusion model [21]. Computational investigation of American option pricing using the discretized linear complementarity has been made in [11, 10, 12, 15]. In [20], the projected SOR approach has been considered. In [11], the discretized linear complementarity problems from the standard finite difference approximation are solved as linear programming problems by the simplex method. More sophisticated methods such as Lemke's algorithm and interior point method have been used for the discretized linear complementarity problems [15, 14].

Numerical consideration may require use of a more sophisticated finite difference method, e.g., when the Black-Scholes operator has little or no diffusion [22]. This occurs for some path-dependent exotic options, e.g., the Asian option. For the Asian option, the diffusion term in one of the spatial dimensions is absent in a Black-Scholes partial differential equation [22]. Standard finite difference using central weighting for the convection term can produce solutions with spurious oscillations [22]. To prevent these oscillations, an upstream weighting for the convection term can be used to introduce numerical diffusion. Unfortunately this can lead to excessive numerical diffusion. To overcome this, a nonlinear flux limiter with the upstream weighting can be used for the convection term [22]. In §2, we illustrate that the upstream weighting with the van Leer limiter leads to discretized nonlinear complementarity problems. The Brenan and Schwartz method [3] or a linear programming method cannot be applied to these discretized problems.

In the partial differential equation framework, hundreds or even thousands of discretized problems need to be solved sequentially backwards in time to value options. Thus efficient computational methods for solving discretized problems are crucial in pricing American options, particularly when high accuracy is needed or a multi-factor model is used. In this paper, we propose a Newton type interior-point method to solve the discretized complementarity problem for American option pricing. The proposed method is applicable to both discretized linear complementarity problems as well as nonlinear complementarity problems; hence the standard finite difference approximation as well as the more complex upstream weighting with the van Leer flux limiter can be used. Our proposed method is based on the observation that these discretized problems, linear or nonlinear, are closely related: the difference between the solutions of the consecutive problems decreases as the time step size decreases. The proposed method uses a local Newton process for fast convergence; this Newton process is then globalized using a quadratic penalty function.

In §3, we illustrate that the average number of iterations required to solve each discretized problem is typically $2 \sim 5$ using the proposed method (motivated and described in §4). More importantly, the number of iterations required seems to be insensitive to the number of discretization points in the spatial dimension; it actually decreases as the number of discretization points in time increases. This property becomes particularly attractive for a multi-factor model or exotic options.

The arbitrage condition for the fair value of the American option requires that the delta hedge factor be continuous. Satisfaction of this condition is important in computing accurate delta hedge factor, the early exercise curve, as well as option values. We demonstrate in §3 that the delta hedge factor computed using the complementarity approach is continuous, thus satisfying the continuity requirement of the early exercise curve for no arbitragy. On the other hand, the delta hedge factors computed using the popular explicit payoff method and the binomial method typically have jumps.

Stability of a numerical method for a partial differential complementarity problem is different from that of a partial differential equation. We illustrate, in §3, that the Crank-Nicolson finite difference method is typically unstable for the partial differential complementarity problem; the computed delta hedging factor exhibits oscillations.

2. Discretized Problems for American Options. To compute an option value numerically, the associated partial differential operator is approximated through discretization and finite difference approximation. We illustrate the discretized American option pricing problem is a nonlinear complementarity problem when using the upstream weighting finite difference with a nonlinear flux limiter.

For simplification, we consider the generalized Black-Scholes 1-factor model: assume that the underlying price follows,

$$\frac{dS_t}{S_t} = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \tag{1}$$

where dW_t is a standard Brownian motion, and $\sigma(S, t)$ denotes a deterministic local volatility function. Let $\Lambda(S)$ denote the payoff function of an American option when the underlying equals S at time t. Under the no arbitrage assumption, the early exercise curve is characterized by continuity of $\frac{\partial V}{\partial S}$, and the condition $V(S, t) = \Lambda(S)$, see, e.g., [13].

It is established in a generalized Black-Scholes model [17], and in a jump diffusion model [21], that necessary and sufficient conditions for the American option value $V(S, t)$ is for $V(S, t)$ to solve a partial differential complementarity/ variational inequality problem. Assume that r is the risk free interest and q is the continuous dividend yield. If the volatility σ is a function of time only, and the payoff function is a convex function of S satisfying some technical conditions, then the American option value is a solution to the following partial differential complementarity problem,

$$\begin{array}{l} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV \leq 0, \\ V(S,t) - \Lambda(S) \geq 0, \\ (V(S,t) - \Lambda(S))(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV) = 0, \end{array} \tag{2}$$

with the final condition $V(S, T) = \Lambda(S)$. Subsequently we omit notationally the dependence of σ and q on S and t for simplicity. Partial differential complementarity problems for many exotic options, e.g., Asian options, can be formulated similarly in the Black-Scholes framework, e.g., [20].

The payoff constraint $V(S, t) \geq \Lambda(S)$ comes directly from the no arbitrage assumption. In contrast to the Black-Scholes equation for the European option, the inequality,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV \leq 0, \tag{3}$$

reflects the asymmetric relationship between the long and short positions of the American contract; only the holder of the option controls the early exercise feature.

If we let $t^* = T - t$, the partial differential inequality in (3) becomes $\mathcal{L}_{BS}[V] \geq 0$ where

$$\mathcal{L}_{BS}[V] \stackrel{\text{def}}{=} \frac{\partial V}{\partial t^*} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - (r-q)S\frac{\partial V}{\partial S} + rV.$$

Let $\{t_i^*\}_{i=0}^M$ denote the discretization in the time interval $[0, T]$ and $\{S_i\}_{i=0}^N$ denote the discretization in the spatial dimension, where S_N is sufficiently large. Let $\delta S_i = \frac{S_{i+1}-S_{i-1}}{2}$ and $\delta S_{i+\frac{1}{2}} = S_{i+1} - S_i$. Using the finite volume approach [19] and adopting the notation in [22], a finite difference approximation to the Black-Scholes partial differential operator can be described by

$$\begin{aligned} \mathcal{L}_{BS}[V] \approx \tfrac{V_i^{j+1}-V_i^j}{\delta t^*} \quad & -\theta(H_{i-\frac{1}{2}}^{j+1} - H_{i+\frac{1}{2}}^{j+1} + f_i^{j+1}) \\ & -(1-\theta)(H_{i-\frac{1}{2}}^{j} - H_{i+\frac{1}{2}}^{j} + f_i^{j}), \end{aligned} \tag{4}$$

where $i = 1, \cdots, N-1$, $j = 0, \cdots, M-1$, and $0 \leq \theta \leq 1$ denotes the temporal weighting. In addition,

$$\begin{aligned} H_{i-\frac{1}{2}}^{j+1} &= \frac{1}{\delta S_i}\left[(-\frac{1}{2}\sigma_i^{j\,2} S_i^2)\frac{V_i^{j+1} - V_{i-1}^{j+1}}{\delta S_{i-\frac{1}{2}}} - (r-q)S_i V_{i-\frac{1}{2}}^{j+1}\right], \\ H_{i+\frac{1}{2}}^{j+1} &= \frac{1}{\delta S_i}\left[(-\frac{1}{2}\sigma_i^{j\,2} S_i^2)\frac{V_{i+1}^{j+1} - V_{i}^{j+1}}{\delta S_{i-\frac{1}{2}}} - (r-q)S_i V_{i+\frac{1}{2}}^{j+1}\right], \end{aligned}$$

and

$$f_i^{j+1} = (-r)V_i^{j+1}.$$

The approximation scheme is fully implicit when $\theta = 1$, and explicit when $\theta = 0$. The Crank-Nicolson method is obtained when $\theta = \frac{1}{2}$. The central weighting scheme for the convection term corresponds to

$$V_{i+\frac{1}{2}}^{j+1} = \frac{V_{i+1}^{j+1} + V_i^{j+1}}{2}. \tag{5}$$

If the grid spacing is uniform, this approximation has second-order accuracy and the above finite difference approximates the Black-Scholes operator by a linear function. For notational simplicity, we sometimes suppress the dependence on the index j and denote $x = V - \Lambda$. Let $Bx + c$ be the linear function from the finite difference approximation (4) to the Black-Scholes operator. Introducing an auxiliary variable y: $y = Bx + c$, the discretized problem (2) can be written as

$$\begin{cases} G(x, y) = Bx + c - y = 0, \\ Xy = 0, \\ x \geq 0,\ y \geq 0, \end{cases} \tag{6}$$

where $X \stackrel{\text{def}}{=} \text{diag}(x)$, and $G(x, y) : \Re^{2(N-1)} \to \Re^{N-1}$ is a linear function. Problem (6) is a linear complementarity problem.

When the partial differential equation in (2) has little or no diffusion, the standard central weighting (5) for the convection term can produce solutions with oscillations. In addition, the standard finite difference methods are inaccurate due to introduction of excessive numerical diffusion [2, 22]. To overcome this, the following upstream weighting with the van Leer flux limiter has been suggested [22],

$$V_{i+\frac{1}{2}}^{j+1} = V_{\text{up}}^{j+1} + \frac{\phi(q_{i+\frac{1}{2}})}{2}(V_{\text{down}}^{j+1} - V_{\text{up}}^{j+1}), \tag{7}$$

where

$$\text{up} = \begin{cases} i & \text{if } \; -rS_i \geq 0 \\ i+1 & \text{if } \; -rS_i < 0, \end{cases} \qquad \text{down} = \begin{cases} i+1 & \text{if } \; -rS_i \geq 0 \\ i & \text{if } \; -rS_i < 0, \end{cases}$$

and $\phi \in [0, 2]$ is the weight which is a nonlinear function of the option values,

$$\begin{aligned} \phi(q_{i+\frac{1}{2}}^{j+1}) &= \frac{|q_{i+\frac{1}{2}}^{j+1}| + q_{i+\frac{1}{2}}^{j+1}}{1 + |q_{i+\frac{1}{2}}^{j+1}|}, \\ q_{i+\frac{1}{2}}^{j+1} &= \frac{V_{\text{up}}^{j+1} - V_{\text{2up}}^{j+1}}{S_{\text{2up}} - S_{\text{up}}} \Big/ \frac{V_{\text{down}}^{j+1} - V_{\text{up}}^{j+1}}{S_{\text{up}} - S_{\text{down}}}. \end{aligned} \tag{8}$$

When $V_{\text{down}}^{j+1} = V_{\text{up}}^{j+1}$, the finite difference approximation does not depend on the corresponding components of q; it can be eliminated. Note that when the weight $\phi \equiv 1$, the scheme becomes the central weighting scheme.

Using the van Leer flux limiter, the finite difference approximation to the partial differential operator in (2) gives rise to nonlinear functions. In addition, these nonlinear functions are piecewise smooth due to the presence of the absolute values in the weight. Standard mathematical programming methods do not apply directly to piecewise continuous functions. To deal with this nondifferentiability, let $\bar{q}_i^{j+1} \geq 0$ denote the negative part of $q_{i-\frac{1}{2}}^{j+1}$, and $\hat{q}_i^{j+1}$ denote the positive part of the $q_{i-\frac{1}{2}}^{j+1}$, $i = 1, \cdots, N$. It is clear that

$$|q_{i-\frac{1}{2}}^{j+1}| = \hat{q}_i^{j+1} + \bar{q}_i^{j+1} \quad \text{and} \quad q_{i-\frac{1}{2}}^{j+1} = \hat{q}_i^{j+1} - \bar{q}_i^{j+1}.$$

Hence the weight function becomes

$$\phi(q_{i+\frac{1}{2}}^{j+1}) = \frac{2\hat{q}_{i+1}^{(j+1)}}{1 + \hat{q}_{i+1}^{j+1} + \bar{q}_{i+1}^{j+1}},$$

and $\hat{G}$ denote the nonlinear function from the upstream weighting with the van Leer limiter for the convection term, i.e., for $i = 1, \cdots, N-1$,

$$\hat{G}_i = \frac{V_i^{j+1} - V_i^j}{\delta t^*} - \left[\theta(F_{i-\frac{1}{2}}^{j+1} - F_{i+\frac{1}{2}}^{j+1} + f_i^{j+1}) + (1-\theta)(F_{i-\frac{1}{2}}^j - F_{i+\frac{1}{2}}^j + f_i^j)\right]. \tag{9}$$

Let $\bar{G}$ be the following functions from the definition of the q,

$$\bar{G}_{i+1} = \frac{V_{\text{up}}^{j+1} - V_{\text{2up}}^{j+1}}{S_{\text{2up}} - S_{\text{up}}} - \frac{V_{\text{down}}^{j+1} - V_{\text{up}}^{j+1}}{S_{\text{up}} - S_{\text{down}}}(\hat{q}_{i+1}^{j+1} - \bar{q}_{i+1}^{j+1}), \; i = 0, \cdots, N-1. \tag{10}$$

Let $G : \Re^{4N-2} \rightarrow \Re^{2N-1}$ denote

$$G(V^{j+1}, \hat{q}^{j+1}, w^{j+1}, \bar{q}^{j+1}) \stackrel{\text{def}}{=} (\hat{G} - w^{j+1}, \bar{G}),$$

$x \stackrel{\text{def}}{=} (V - \Lambda, \hat{q}) \in \Re^{2N-1}$, and $y \stackrel{\text{def}}{=} (w, \bar{q}) \in \Re^{2N-1}$. Thus when the the upstream weighting with the van Leer flux limiter is used, the discretized American option problem can be formulated as the following nonlinear programming problem

$$\left\{ \begin{array}{l} G(x, y) = 0, \\ Xy = 0, \\ x \geq 0, \ y \geq 0, \end{array} \right. \tag{11}$$

where $G : \Re^{2n} \rightarrow \Re^n$, $x \in \Re^n$, and $y \in \Re^n$ with $n = 2N - 1$. Note that the formulation (6) is a special case of the formulation (11) where $G(x, y)$ is linear.

When the upstream weighting with the van Leer flux limiter is used for the convection term, it is easy to verify that a European option value can be computed by solving,

$$\left\{ \begin{array}{l} G(x, y, z) = 0, \\ Xy = 0, \\ x \geq 0, \ y \geq 0, \end{array} \right. \tag{12}$$

with $G : \Re^{2n+m} \rightarrow \Re^m$, $x \in \Re^n$, $y \in \Re^n$, and $z \in \Re^m$ for some positive integers n and m.

Problem (11) can be written as a nonlinear complementarity problem when $V_{\text{down}} \neq V_{\text{up}}$ while problem (12) is not a standard nonlinear complementarity problem. Problem (11) has $2n$ nonlinear equality constraints. The simple bound constraints adds additional complexity. One may be tempted to compute the solution to (11) by solving the optimization problem,

$$\begin{array}{ll} & \min_{x,y} \|G(x, y)\|^2 + \|Xy\|^2 \\ \text{subject to} & x \geq 0, y \geq 0, \end{array} \tag{13}$$

using a standard optimization software. Since this optimization problem is degenerate at the solution of the discretized problem (11), even when problem (11) itself is not, this approach will encounter numerical difficulty. We propose a specialized Newton type interior-point algorithm for solving (11); this algorithm can be easily modified to solve (12).

Next, we examine the continuity of the delta hedge factors using the complementarity approach, the binomial method, and the explicit payoff method. In addition, we illustrate the efficiency of our proposed algorithm for solving (11) ; details of the proposed algorithm are delayed until §4 due to its technical nature.

3. Ensuring Continuity of Delta. The discretized American option problem (11), includes a system of n equations $G(x, y) = 0$, complementarity conditions $x_i y_i = 0, i = 1, \cdots, n$, and nonnegativity constraints on the variable x and y. Depending on the finite difference method, the discretized problem can be a linear complementarity problem with a tridiagonal structure, a linear complementarity problem with a more general band structure, or a genuine nonlinear complementarity problem.

An efficient computational method for solving the discretized problems (11) is crucial for American option pricing since a discretized problem needs to be solved at each time step. If a discretized linear complementarity problem has a special tridiagonal coefficient matrix, the Brennan-Schwartz method [3] can be used. In addition, two popular methods are the binomial method [9] and the explicit payoff method, e.g., [16, 13]; the latter approximates the American option value by taking the maximum of the continuation value and the payoff at each time step. For example, when the finite difference with central weighting (4,5) is used (which leads to linear complementarity problems), the explicit payoff method approximates the American option value as below,

$$\begin{array}{ll} \textbf{for} & j = 1, 2, \cdots, M \\ & B^j x + c^j = 0, \\ & V^j = \max(x(1:n), \Lambda), \\ \textbf{end} & \end{array} \tag{14}$$

where $Bx + c$ is the finite difference approximation (4,5) to the Black-Scholes operator.

When the upstream weighting with the van Leer flux limiter (4,7,8) is used, the continuation value at each time step can be computed by solving

$$\begin{cases} \hat{G}(V, \hat{q}, \bar{q}) = 0, \\ \bar{G}(V, \hat{q}, \bar{q}) = 0, \\ \hat{q}_i \bar{q}_i = 0, \ i = 1, \cdots, N, \\ \hat{q} \geq 0, \ \bar{q} \geq 0, \end{cases} \tag{15}$$

where $\hat{G}$ and $\bar{G}$ are defined by (9) and (10) respectively. The explicit payoff method is used in [22].

One difficulty with the explicit payoff method is that its convergence properties are unclear. Although the computed option value is known to converge using the binomial method, we do not know of any convergence result of the hedge factors computed from finite difference approximation. On the other hand, convergence in both the option values and the hedge factors computed using the complementarity approach has been established [17, 21].

Unlike the Brennan and Schwartz method [3], the proposed algorithm does not have restriction on the type of the finite difference method used; it is applicable to the discretized problem from the standard finite difference as well as more sophisticated approximation schemes, such as the upstream weighting with the van Leer flux limiter (4,7,8). Using the vanilla American option in a generalized Black-Scholes model as an example, we illustrate the following:

- The (implicit finite difference) complementarity approach for American option pricing produces continuous delta hedge factors, whereas the popular binomial method and the explicit payoff method yield discontinuous delta factors.

- The Crank-Nicolson method, while stable for solving partial differential equations, is unstable for the partial differential complementarity problem. The delta hedge factor computed using the Crank-Nicolson method with the complementarity approach is typically oscillatory.

- The proposed algorithm for solving the discretized problem (11) is computationally efficient; it typically solves the discretized problem in $2 \sim 5$ iterations to an appropriate accuracy.

For subsequent discussion, we consider two types of finite difference approximations. First, the central weighting (4,5) for the convection term is used; the discretized problem (11) is a linear complementarity problem. Second, the upstream weighting scheme with the van Leer nonlinear flux limiter (4,7,8) is used; the discretized problem (11) is a nonlinear complementarity problem. In our experiments, the finite region $[0,T] \times [0, 2S_0]$ is used to approximate the region $[0,T] \times [0, +\infty)$. The finite region $[0,T] \times [0, 2S_0]$ is then discretized by a uniform grid $\{(t_i^*, S_j)\}$, $i = 0, \cdots, M$ and $j = 0, \cdots, N$. Unless explicitly stated otherwise, we use the implicit finite difference scheme, i.e., $\theta = 1$ in (4). The number of steps in the state dimension, shown in the first column, is chosen so that $\delta S = \sqrt{\delta t^*}$; hence the discretization error is $O(\delta t^*)$.

The starting point is crucial for any interior point method. Although the solution for the discretized problem at t_i^* offers a good approximation to the solution for the discretized problem at t_{i+1}^*, this point is typically at the boundary of the feasible region $\mathcal{F} = \{(x,y) : x \geq 0,\ y \geq 0\}$. We introduce a small perturbation. When the discretized problem is a linear complementarity problem, we set the starting point $(x^{(0)}, y^{(0)}) = \max((x,y), \min(.001, \delta t^*))$ at t_{i+1}^* where (x,y) is the solution at t_i^*. When the upstream weighting with the van Leer flux limiter is used, we also would like to start from a point such that $x_i \neq x_{i+1},\ i = 1, \cdots, N$, since the Newton step is not defined at such a point; a small random perturbation is then used to avoid $x_i = x_{i+1}$. The stopping criterion for the optimization algorithm is

$$\max({x^{(k)}}^T y^{(k)}, \|G(x^{(k)}, y^{(k)})\|_2) < \epsilon_{\text{opt}},$$

where $\epsilon_{\text{opt}} > 0$ is an error tolerance. Since the computed $(x^{(k)}, y^{(k)})$ is strictly positive, the discretized option pricing problem (11) is solved with at least ϵ_{opt} accuracy.

3.1. Continuity of Delta Hedge Factors. Continuity of the delta hedge factor is an arbitrage condition for the American option value. We compare first the computed delta hedge factors using the complementarity approach, the explicit payoff method, and the binomial method. The delta hedge factors are computed using finite difference, e.g., at time $t = 0$, the delta factor is approximated,

$$\left[\frac{\partial V}{\partial S}\right]_{t=0, S=S_i} \approx \frac{V_{i+1}^M - V_i^M}{\delta S}.$$

FIG. 1 displays the computed delta factors using the complementarity method and the explicit payoff method (14); the central weighting (4,5) is used in these examples (hence the discretized problem is a linear complementarity problem). The top plot displays the delta factors at $t = 0$, in the neighborhood of the early exercise curve, computed using the complementarity approach and the explicit payoff method respectively. It is clear that the hedge factor from the complementarity approach is continuous while the hedge factor from the explicit payoff method has a jump near the the early exercise curve. The left plot on the bottom graphs the price differences using the two methods; a relatively large difference can be observed around the early exercise curve. Notice that, although the difference in delta is localized, the price difference is propagated through a wide range of underlying values. The right plot on the bottom graphs the early exercise curves computed using the complementarity approach and the explicit payoff method; the early exercise curves for the put option is computed by locating, at each time step, the smallest underlying value S_i such that $|V(S_i, t) - \Lambda(S_i)| > 10^{-8}$.

When the upstream weighting with the van Leer flux limiter is used, similar behavior is observed for the complementarity approach and the explicit payoff method (15). Fig. 2 compares the computed delta factors using the complementarity approach and the explicit payoff method in this case. We note that the proposed method is easily adapted to compute a solution of the discretized problem (15) for the explicit payoff method; hence it can be used to solve the discretized nonlinear programming problem (12) for European options.

The computed delta hedge factors using the explicit payoff method is discontinuous; this may be explained by the fact that the computed option values from the explicit payoff method do not solve the discretized complementarity problem (11). Indeed, the error measured as $\max(-\min(0, Bx + c)) + x^T|Bx + c|)$ of the solution obtained using the explicit payoff method is greater than 10 at each time step for the American put option with $S_0 = 100,\ T = 1,\ r = 0.1,\ \sigma = 0.1,\ q = 0.05,\ M = 2000$ and $N = 8943$.

We also investigate continuity of delta when computed using the binomial method. Although the option value computed using the binomial method is proven to converge, the convergence of the delta factor from the finite difference is unclear for the binomial method. FIG. 3 compares the delta factors at $t = 0$ computed using the complementarity method and binomial method with 2000 time steps. The discontinuity of the computed delta using the binomial method is clearly exhibited in the top plot. The middle plot exhibits that the size of the jump diminishes as S increases. The bottom plot illustrates the price difference.

Next we provide computational evidence illustrating that the Crank-Nicolson method is typically unstable for solving partial differential complementarity (American option pricing) problems. The Crank-Nicolson method has been a favorite finite difference approximation in both the European and American option pricing because its convergence rate is quadratic in time. For solving partial differential complementarity problems, however, it is noted in [17] that the unconditional stability seems to be established only for the fully implicit scheme.

The instability of the Crank-Nicolson scheme is illustrated in Fig. 4. We observe from the top-right plot that the delta hedge factor computed using the Crank-Nicolson method has oscillations around the early exercise curve; this is demonstrated more clearly in the middle plot. The bottom plot graphs the computed option prices using the two methods. Moreover, we note that instability of the Crank-Nicolson method worsens as the volatility parameter becomes larger.

3.2. Efficiency of the Proposed Algorithm. We have illustrated that the (implicit finite difference) complementarity approach produces continuous delta hedge factors while the

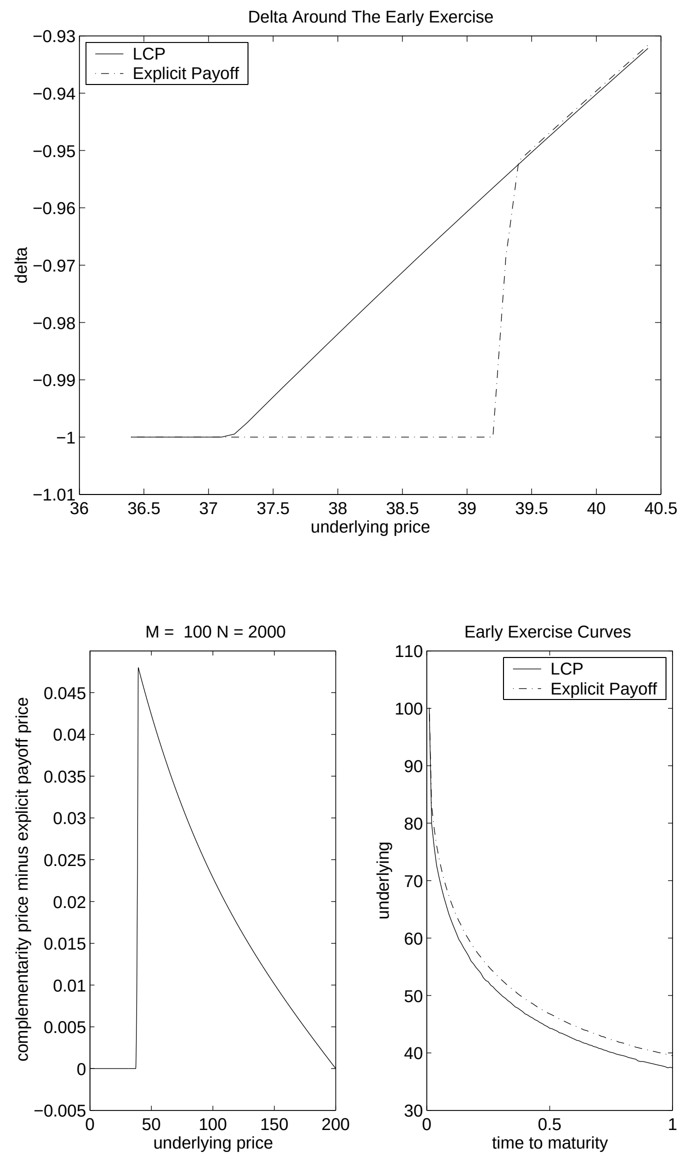

FIG. 1. *LCP Comparisons :* $S_0 = 100,\ K = 100,\ T = 1,\ \ r = 0.1,\ q = 0,\ \sigma = 0.8,\ \epsilon_{opt} = 10^{-4}$

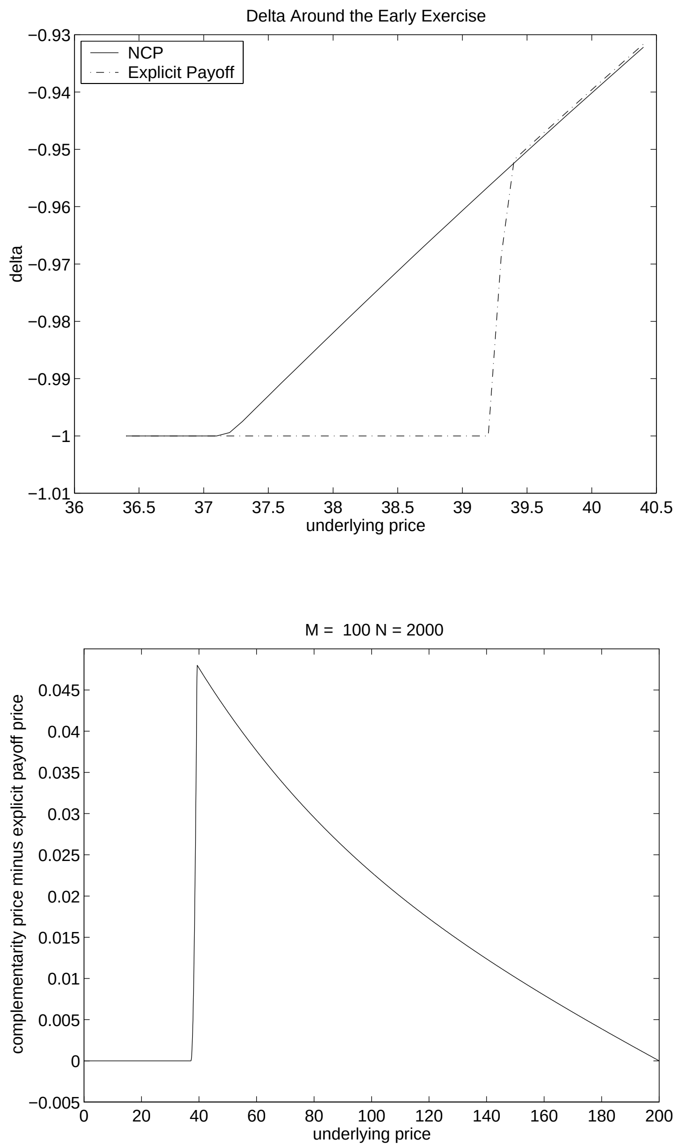

FIG. 2. *NCP Comparisons:* $S_0 = 100,\ K = 100,\ T = 1,\ \ r = 0.1,\ q = 0,\ \sigma = 0.8,\ \ \epsilon_{opt} = 10^{-4},\ \delta t^* = 0.01$

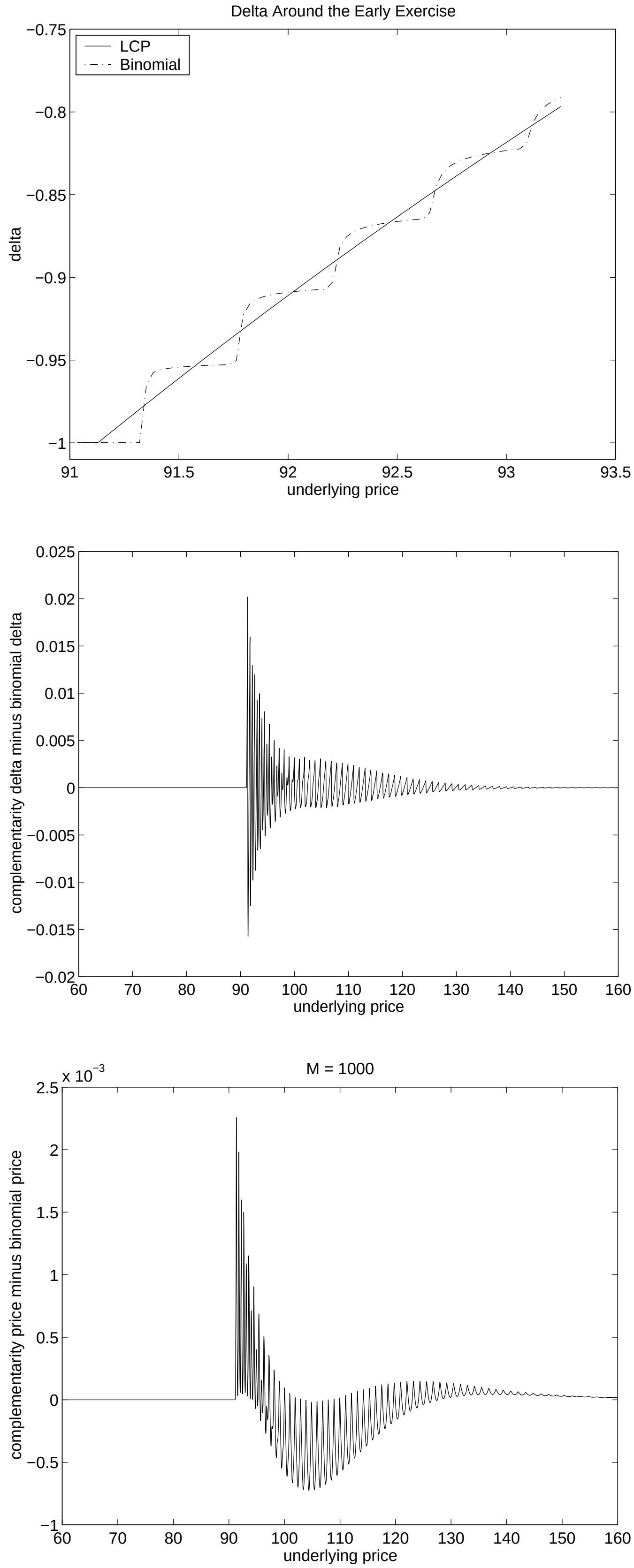

FIG. 3. *Comparison with the Binomial Method:* $S_0 = 100,\ K = 100,\ T = 1,\ ,\ r = 0.1,\ q = 0,\ \sigma = 0.15,\ \epsilon_{opt} = 10^{-4}$

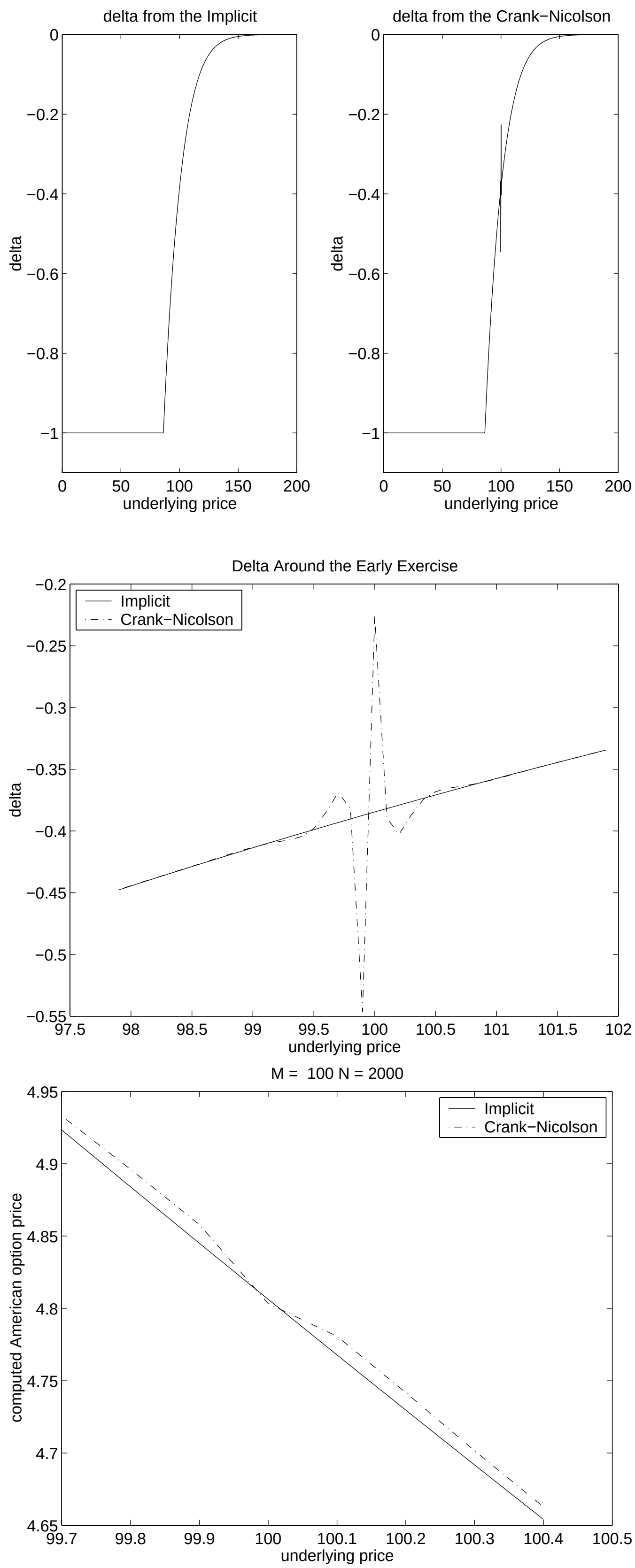

FIG. 4. *Implicit and Crank-Nicolson:* $S_0 = 100,\ K = 100,\ r = 0.1,\ q = 0.0,\ \sigma = 0.2,\ \epsilon_{opt} = 10^{-4}$

M	option value ($\sigma = 0.1$)	#its ($\sigma = 0.1$)	option value ($\sigma = 2$)	#its ($\sigma = 2$)
100	2.3804	4.00	42.8952	5.48
200	2.3829	3.20	42.9190	5.21
300	2.3837	3.06	42.9270	4.91
400	2.3841	2.61	42.9309	4.63
500	2.3844	2.42	42.9333	4.48
600	2.3845	2.33	42.9349	4.95
700	2.3846	2.27	42.9361	4.94
800	2.3847	2.22	42.9370	4.96
900	2.3848	2.19	42.9376	4.95
1000	2.3849	2.16	42.9382	4.88
1100	2.3849	2.14	42.9386	4.76
1200	2.3849	2.12	42.9390	4.68
1300	2.3850	2.11	42.9393	4.63
1400	2.3850	2.10	42.9395	4.58
1500	2.3850	2.09	42.9398	4.56
1600	2.3850	2.08	42.9400	4.44
1700	2.3851	2.07	42.9402	4.48
1800	2.3851	2.07	42.9403	4.44
1900	2.3851	2.06	42.9405	4.38
2000	2.3851	2.06	42.9406	4.31

Computed American Put Option Values and Average Numbers of Iterations
($S_0 = 100,\ K = 100,\ T = 1,\ \ r = .1,\ q = 0.05$)

TABLE 1

popular binomial method and the explicit payoff method yield discontinuous delta. Next we illustrate the discretized complementarity problem (11) can be solved efficiently using the proposed algorithm, making the complementarity approach affordable.

We note that it does not help in practice to solve the discretized option problem (11) to accuracy greatly surpassing the accuracy of the discretization. Setting the stopping tolerance $\epsilon_{\text{opt}} = 10^{-4}$, Table 1 display the computed option values using our proposed algorithm for discretized linear complementarity problems for two different volatility parameter settings, $\sigma = 10\%$ and $\sigma = 200\%$; the central weighting (5) is used for this example. The third and fifth columns list the average number of iterations required at each time step using the proposed algorithm to solve a discretized problem (11). The second and the fourth columns list the computed option values respectively.

Observing Table 1, the average numbers of iterations required by the proposed algorithm are $2 \sim 4$ in a more typical parameter setting with $\sigma = 10\%$. When the volatility is unusually high, $\sigma = 200\%$, the average numbers of iterations are slightly higher, $4 \sim 6$. The average number of iterations required, in both cases, decreases as the discretization becomes finer.

M	NCP	#its	LCP	#its
100	2.3805	4.11	2.3804	4.00
200	2.3829	3.25	2.3829	3.20
300	2.3837	2.93	2.3837	3.06
400	2.3841	2.48	2.3841	2.61
500	2.3844	2.33	2.3844	2.42
600	2.3845	2.25	2.3845	2.33
700	2.3846	2.20	2.3846	2.27
800	2.3847	2.16	2.3847	2.22
900	2.3848	2.14	2.3848	2.19
1000	2.3849	2.12	2.3849	2.16
1100	2.3849	2.11	2.3849	2.14
1200	2.3850	2.09	2.3849	2.12
1300	2.3850	2.08	2.3850	2.11
1400	2.3850	2.07	2.3850	2.10
1500	2.3850	2.07	2.3850	2.09
1600	2.3851	2.06	2.3850	2.08
1700	2.3851	2.05	2.3851	2.07
1800	2.3851	2.05	2.3851	2.07
1900	2.3851	2.05	2.3851	2.06
2000	2.3851	2.04	2.3851	2.06

NCP: Computed Option Values and Average Numbers of Iterations
($S_0 = 100,\ K = 100,\ T = 1,\ \sigma = 0.1,\ r = .1,\ q = 0.05$)

TABLE 2

The cost of the American pricing using the proposed interior-point method is roughly $2 \sim 4$ times that of the European option pricing in the typical parameters setting. The performance of the proposed algorithm is similar when $G(x, y)$ is nonlinear. Table 2 displays the computed American put option values and the average numbers of iterations required when the upstream weighting with the van Leer flux limiter (4,7,8) is used ; the results from the central weighting are included for comparison.

Table 3 compares the computed American option values using the complementarity approach and the implicit finite difference ($\theta = 1$) with the complementarity approach and the Crank-Nicolson method ($\theta = \frac{1}{2}$). We note that the proposed interior-point method for the discretized American option value problem performs similarly when the Crank-Nicolson scheme is used.

In addition, we note that, although the delta computed using the explicit payoff has a large jump, the computed option values are surprisingly close to those computed from the complementarity approach. This, however, does not necessarily suggest that the computed option values from the explicit payoff method converge to the true values. FIG. 5 plots the

M	Implicit Complementarity	#it	Crank-Nicolson Complementarity	#it
100	2.3804	4.00	2.3836	3.61
200	2.3829	3.20	2.3852	3.27
300	2.3837	3.06	2.3853	3.03
400	2.3841	2.61	2.3853	2.77
500	2.3844	2.42	2.3853	2.59
600	2.3845	2.33	2.3853	2.49
700	2.3846	2.27	2.3853	2.41
800	2.3847	2.22	2.3853	2.37
900	2.3848	2.19	2.3853	2.32
1000	2.3849	2.16	2.3853	2.28
1100	2.3849	2.14	2.3853	2.25
1200	2.3849	2.12	2.3854	2.22
1300	2.3850	2.11	2.3853	2.21
1400	2.3850	2.10	2.3854	2.19
1500	2.3850	2.09	2.3854	2.17
1600	2.3850	2.08	2.3854	2.16
1700	2.3851	2.07	2.3854	2.14
1800	2.3851	2.07	2.3854	2.13
1900	2.3851	2.06	2.3854	2.12
2000	2.3851	2.06	2.3854	2.11

Comparisons of Implicit and Crank-Nicolson Scheme in American Option Pricing
($S_0 = 100,\ K = 100,\ T = 1,\ \sigma = 0.1,\ r = .1,\ q = 0.05$)

TABLE 3

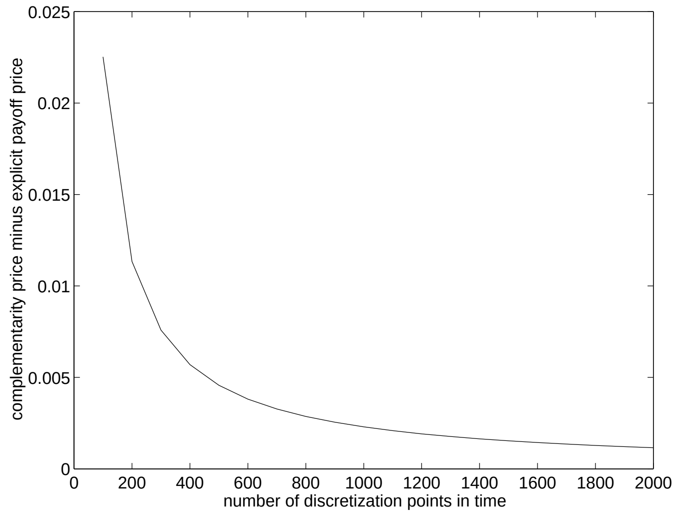

FIG. 5. *Price Difference:* $\sigma = 2,\ r = 0.1,\ q = 0.05,\ T = 1,\ K = 100,\ S_0 = 100,\ \epsilon_{opt} = 10^{-4}$

differences between the option values using these two methods against the accuracy of discretization (the number of time steps). Note that the convergence of the (implicit) complementarity method is known [17].

Finally, we demonstrate that the proposed algorithm can solve the discretized problem to high accuracy, if necessary. Table 4 illustrates the asymptotic behavior of the proposed algorithm when the central weighting (4,5) is used for the convection term; Table 5 demonstrates the convergence behavior when the discretized problem (11) is a nonlinear complementarity problem. In both cases, we display the optimality value $\frac{1}{2}\|G(x,y)\|^2 + x^T y$ at each iteration for the first three time steps $t^* = \delta t^*,\ 2\delta t^*$, and $3\delta t^*$. In our experience, this asymptotic behavior is typical. The stopping tolerance ϵ_{opt} is set to 10^{-7} in these examples.

The discretized linear complementarity problems are typically nearly degenerate since the two equations,

$$V(S,t) - \Lambda(S) = 0, \quad \text{and} \quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV = 0$$

are simultaneously approximately satisfied around the early exercise curve. The results in Table 4 show that this near degeneracy does not cause visible computational difficulty; superlinear convergence is observed for a couple of iterations to fairly high accuracy. After that convergence becomes fast linear.

Discretized nonlinear complementarity problems generated from the upstream weighting with the Van Leer nonlinear flux limiter have additional degeneracy: the Jacobian of the constraints $\bar{G}(x,y) = 0$ can be singular when $V_{\text{down}} = V_{\text{up}}$, as can be seen from the definition of $\bar{G}$ given in (10). We observe that, though the proposed method takes slightly more iterations when $G(x,y)$ is nonlinear, convergence is still rapid.

4. A Newton Type Interior-Point Method. In this section, we motivate and describe the proposed interior-point method for solving the discretized option pricing problem (11);

k	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_1^*$)	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_2^*$)	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_3^*$)
0	7.947819e+02	5.004075e+04	4.667865e+04
1	3.564160e+00	1.318422e-01	6.598802e-02
2	2.343340e-01	2.679068e-03	2.423917e-03
3	3.252325e-02	1.148443e-04	1.730893e-04
4	4.531208e-03	2.647152e-06	1.139226e-05
5	6.322716e-04	3.392189e-07	5.378067e-07
6	8.820751e-05	4.947867e-08	4.395356e-08
7	1.232954e-05		
8	1.727045e-06		
9	2.455198e-07		
10	3.851848e-08		

Asymptotic Behavior for an American Put Option: $G(x,y)$ *Is Linear*
($S_0 = 100,\ K = 100,\ T = 1,\ \sigma = 0.1,\ r = .05,\ q = 0,\ \epsilon_{\text{opt}} = 10^{-7},\ \delta t^* = 0.01,\ \delta S = 0.1$)

TABLE 4

k	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_1^*$)	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_2^*$)	$\frac{1}{2}\|G(x^{(k)},y^{(k)})\|^2+{x^{(k)}}^T y^{(k)}$ ($t^*=t_3^*$)
0	4.468042e+08	4.069523e+08	3.759767e+08
1	4.481993e+05	1.124695e+00	5.422102e-01
2	3.893335e+03	1.573953e-02	1.437938e-02
3	3.963988e+02	2.778598e-03	2.524083e-03
4	3.741587e-05	3.597233e-04	4.196310e-04
5	9.382774e-06	2.765571e-05	6.150641e-05
6	2.880286e-06	2.958813e-06	1.023294e-05
7	1.373253e-06	1.257544e-06	1.928311e-06
8	9.978982e-07	8.415812e-07	8.534015e-07

Asymptotic Behavior for an American Put Option: $G(x,y)$ *Is Nonlinear*
($S_0 = 100,\ K = 100,\ T = 1,\ \sigma = 0.1,\ r = .05,\ q = 0,\ \epsilon_{\text{opt}} = 10^{-7},\ \delta t^* = 0.01,\ \delta S = 0.1$)

TABLE 5

this method is used in §3 to obtain the computational results. If not interested in the details of the computational method, a reader can go directly to §5 for concluding remarks.

Solving a general linear complementarity problem is a NP-hard problem [17]. However, we observe that in the option pricing setting each discretized problem at t^*_{i+1} is endowed with a good approximation to its solution, i.e., the solution to the problem at t^*_i. Moreover, this approximation becomes more accurate as the time discretization parameter δt^* is decreased. This is the situation where a Newton process brings fast convergence. In addition, the nonlinear programming problem (11) is highly structured. We derive a Newton method which utilizes this initial point as well as the special structure of the discretized problem (11).

First, when $x \geq 0$ and $y \geq 0$, the inner product $x^T y \geq 0$ measures satisfaction of the constraints $Xy = 0$. Hence we consider the following auxiliary optimization problem,

$$\begin{aligned} &\min_{x,y} \quad \frac{1}{2}\|G(x,y)\|^2 + x^T y \\ &\text{subject to} \quad x \geq 0, y \geq 0. \end{aligned} \tag{16}$$

The objective function measures satisfaction of optimality of the orginal problem (11); it is simpler than that of (13) because it is less nonlinear; the term $x^T y$ is quadratic rather than quartic. The solution to the complementarity problem (11) is clearly always a global minimizer of (16). However, a local minimizer of (16) may not be a global minimizer and thus not a solution to (11). The possibility of computing a local minimizer which is not a solution can be alleviated through the use of an additional penalty parameter: consider

$$\begin{aligned} &\min_{x,y} \quad f(x,y) \stackrel{\text{def}}{=} \frac{\rho}{2}\|G(x,y)\|^2 + x^T y \\ &\text{subject to} \quad x \geq 0, y \geq 0, \end{aligned} \tag{17}$$

where $\rho > 0$ is a penalty parameter. The use of this penalty parameter is motivated from the following analysis.

The gradient of $f(x,y)$ is

$$\nabla f(x,y) = \begin{bmatrix} \rho \nabla G_x^T G + y \\ \rho \nabla G_y^T G + x \end{bmatrix}.$$

If the Kuhn Tucker condition of the minimization problem (17) is satisfied at a feasible point, then

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} \rho \nabla G_x^T G + y \\ \rho \nabla G_y^T G + x \end{pmatrix} = 0, \tag{18}$$

where $Y \stackrel{\text{def}}{=}$ diag(Y). Hence, a local minimizer of (17) satisfying $G(x,y) = 0$ is a global minimizer of (17) since (18) implies that $Xy = 0$. Therefore a large penalty parameter $\rho > 0$ makes it more likely for a local minimizer of (17) to satisfy $G(x,y) = 0$.

Using an algorithm which monotonically decreases the objective function measuring optimality of the original problem (11) further alleviates the potential global versus local minimizer problem. Recall that the discretized problem (11) is always endowed with a good starting point. A monotonically decreasing optimization algorithm restricts the possible local minimizer it computes; starting from a neighborhood of the global minimizer and insisting the objective function values decrease at each iteration further increases the likelihood of the iterates converging to the solution. Indeed, if the solution (x^*, y^*) of the discretized problem (11) is the only point in the intersection of the local minimizers of (17) and the level set $\mathcal{L} = \{x : f(x, y) \leq f(x^{(0)}, y^{(0)})\}$, then solving the minimization problem produces a solution to the discretized problem (11). The proposed algorithm below monotonically decreases the objective function $f(x, y)$; the problematic issue of a local minimizer of (17) which is not a global minimizer is not a practical concern in our experience.

We now consider a direct local Newton process for the discretized problem (11). Let

$$F(x, y) \stackrel{\text{def}}{=} [G(x, y); Xy],$$

$X \stackrel{\text{def}}{=} \text{diag}(x)$ and $Y \stackrel{\text{def}}{=} \text{diag}(y)$. The solution (x^*, y^*) to the discretized complementarity problem (11) can be considered as a feasibility problem: the solution satisfies the bound constraints $(x^*, y^*) \in \mathcal{F} = \{(x, y) : x \geq 0, y \geq 0\}$ and the system of nonlinear equations $F(x, y) = 0$.

A Newton step $d = [d_x; d_y]$ for $F(x, y) = 0$ satisfies

$$\nabla F^T d = -[G; Xy], \tag{19}$$

where ∇F is the Jacobian matrix of F, i.e.,

$$\nabla F^T = \begin{bmatrix} \nabla G_x^T & \nabla G_y^T \\ Y & X \end{bmatrix}.$$

The gradient of the objective function $f(x, y)$ can be written as,

$$\nabla f = \begin{bmatrix} \rho \nabla G_x^T G + y \\ \rho \nabla G_y^T G + x \end{bmatrix} = \nabla F^T [\rho G, e_n].$$

If (x, y) is not the solution of the discretized problem (11), the Newton step d defined by (19) is clearly a descent direction for $f(x, y)$ since

$$\nabla f^T d = -[\rho G, e_n]^T \nabla F^T \nabla F^{-T} [G; Xy] = -(\rho \|G\|^2 + x^T y) < 0.$$

Hence the Newton step (19) is compatible with the minimization problem (17). If the functions $F(x, y)$ were a general nonlinear function, the Newton step from $F(x, y) = 0$ could conflict with the inequality constraints $x \geq 0$ and $y \geq 0$; a truncated Newton step due to the nonnegativity constraints may not be able to produce sufficient progress. Fortunately, since the second half of the nonlinear constraints $Xy = 0$ is bilinear and intimately related to the nonnegativity constraints, the nonnegativity constraints $x \geq 0$ and $y \geq 0$ do not prevent a sufficiently large portion of the Newton step being taken as long as the components of x and y are not zero simultaneously. This can be seen from the equations

$$Y d_x = -X(y + d_y), \quad \text{and} \quad X d_y = -Y(x + d_x).$$

If x_i and y_i are not zero simultaneously, then the stepsize along d_x eventually converge to unity, and fast convergence of the Newton process occurs.

The Newton process described thus far is local: the iterates may not converge to a local minimizer unless the initial point is sufficiently close. To resolve the potential difficulty of the components of x and y being (near) zero simultaneously, and globalize the local Newton process, we combine the Newton step with the use of the affine scaling steepest descent direction used in the trust region and affine scaling method [7]. The scaled steepest descent direction is used in [7] to obtain global convergence to a minimizer of (17); it is simply $-D^2\nabla f$ with $D = \text{diag}(h)$ and h is defined in DEFINITION 4.1.

DEFINITION 4.1. *The vector $h \in \Re^{2n}$ is defined: for each component $1 \leq i \leq 2n$,*

(i) *If $(\nabla f_x)_i \geq 0$ then $h_i \stackrel{\text{def}}{=} x_i$,*

(ii) *If $(\nabla f_y)_i \geq 0$ then $h_{i+n} \stackrel{\text{def}}{=} y_i$,*

(ii) *If $(\nabla f_x)_i < 0$ then $h_i \stackrel{\text{def}}{=} -1$,*

(iii) *If $(\nabla f_y)_i < 0$ then $h_{i+n} \stackrel{\text{def}}{=} -1$.*

To guarantee convergence, each new iterate needs to yield a sufficiently large decrease for the objective function $f(x, y)$. For a discretized option pricing problem (11), a Newton step typically results in a sufficient decrease because it has a good starting point. Since computational efficiency is crucial in pricing the American option, we want to fully exploit this fact. Although sophisticated globalization techniques such as the trust region method [7] can be used, we choose here a simple line search technique to avoid as much unnecessary computation as possible. A full description of the line search method is tedious and unenlightening; we only point out that the projection below is used to ensure the strict feasibility when taking a Newton step:

$$(x^{(k+1)}, y^{(k+1)}) = \max((x^{(k)}, y^{(k)}) + d^{(k)}, \epsilon),$$

where $\epsilon > 0$ is a small parameter about machine precision. The iteration proceeds to the next if it yields a sufficient reduction of the objective function value; for example, if the new iterate $(x^{(k+1)}, y^{(k+1)})$ along the Newton direction satisfies

$$f(x^{(k+1)}, y^{(k+1)}) < f(x^{(k)}, y^{(k)}) - \gamma_0 f(x^{(k)}, y^{(k)})$$

for some $0 < \gamma_0 < 1$, then decrease is sufficient. If the projected Newton step yields a sufficient decrease, the algorithm proceeds to the next iteration. Otherwise, a subsequent line search is performed. A scaled steepest descent direction defined in DEFINITION 4.1 for (17) is computed if sufficient decrease cannot be obtained along the Newton direction. Strict feasibility $(x, y) > 0$ is maintained at each iteration. For more details of the use of the scaled steepest descent and the line search, see [7, 8].

The proposed algorithm is now summarized in FIG. 4. Typically the Newton step leads to sufficient decrease and the cost per iteration is roughly the cost of computing one Newton step, assuming a line search is implemented efficiently. We discuss below how a Newton step can be computed to exploit the sparsity structures of the discretized problems (11).

Initialization. Let $\rho > 0$ and $(x^{(0)}, y^{(0)}) > 0$.

Step 1. Compute a Newton step $d^{(k)}$ satisfying

$$\nabla F^{(k)^T} d = -[G^{(k)}; X^{(k)} y^{(k)}].$$

Step 2. Perform the line search along the Newton direction to compute $(x^{(k+1)}, y^{(k+1)}) > 0$. If

$$f(x^{(k)}, y^{(k)}) - f(x^{(k+1)}, y^{(k+1)}) \le \gamma_1 f(x^{(k)}, y^{(k)}),$$

then go to **Step 1**. Otherwise, continue to **Step 3**.

Step 3. Compute the scaled steepest descent direction; Perform the line search along the scaled steepest descent direction to compute $(x^{(k+1)}, y^{(k+1)}) > 0$ which satisfies sufficient decrease conditions.

First, we examine the case when the discretized problem (11) is a linear complementarity problem. In this case, $\nabla G_x = B$ and $\nabla G_y = -I$. Instead of solving the $2n$-by-$2n$ system (19), the Newton step can be computed by solving a n-by-n linear system:

$$(XB + Y) d_x = -X(G + y), \tag{20}$$

and $d_y = G + B d_x$. Notice that B is typically asymmetric. If finite differencing with central weighting is used, system (19) is tridiagonal and this band sparsity structure can be exploited for efficiency. Moreover $(XB + Y)$ has the same sparsity structure as B and thus can be solved as efficiently as solving $Bx = c$.

In addition, the equation $Xy = 0$ is bilinear. The Newton step can be further improved by computing a correction step $(\hat{d}_x, \hat{d}_y)$ satisfying

$$(x + \hat{d}_x)(y + \hat{d}_y) = 0.$$

This can be approximately achieved by

$$(XB + Y)\hat{d}_x = -X(G + y) - \text{diag}(d_x) d_y. \tag{21}$$

and $\hat{d}_y = G + B\hat{d}_x$. Notice that the coefficient matrices of (19) and (21) are exactly the same.

For the Newton step $\hat{d}$ with correction, we have

$$\nabla f^T \hat{d} = -[\rho G, e_n]^T \nabla F^T \nabla F^{-T} [G; Xy + \text{diag}(d_x) d_y] = -(\rho \|G\|^2 + x^T y) - d_x^T d_y.$$

Hence $(\hat{d}_x, \hat{d}_y)$ is a descent direction for $f(x, y)$ if

$$d_x^T d_y > -(\rho \|G\|^2 + x^T y).$$

In addition, when $G(x, y) = Bx + c - y$, $f(x, y)$ is a quadratic function and the first exact local minimizer of f along a descent direction $d^{(k)}$ in the feasible region $\mathcal{F} = \{x : x \geq 0\}$ can be computed easily when performing a line search. Therefore, the corrected Newton step and the exact line search can be used when the finite difference leads to a linear complementarity problem.

Now we consider the case when $G(x, y)$ is nonlinear in the discretized problem (11); this happens when the upstream weighting with the nonlinear van Leer flux limiter is used. Here $-\nabla G_y$ does not equal to the identity matrix. In this case, the coefficient matrix ∇F for the Newton equation (19),

$$\nabla F^T = \begin{bmatrix} \nabla G_x^T & \nabla G_y^T \\ Y & X \end{bmatrix},$$

is still typically an asymmetric sparse matrix. The matrices ∇G_x and ∇G_y typically have band structures. The Newton step, in this case, can be computed using a sparse LU factorization. To exploit sparsity, it is important to use a column ordering to reduce fill-in. Moreover, for computational efficiency, it is essential that the column ordering is only computed once at the first time step $t^* = \delta t^*$; the computed ordering can be used for subsequent time steps since the sparsity structure remains the same. When $G(x, y)$ is nonlinear, we do not compute a correction step.

Under the nondegenerate assumptions, similar analysis to that in [6] can be used to show that the Newton process is locally quadratically convergent. A discretized problem (11) for American option pricing is typically near degenerate in the sense that there are components with x_i and y_i simultaneously near zero. We illustrate in §3 computationally that this does not prohibit fast local convergence of the Newton process which is derived directly from the discretized problem (11).

The Newton process proposed here is closely related to the one used in the primal and dual interior point methods for linear programming [18]. However, there is no barrier parameter embedded in our proposed method. This allows us to fully exploit the given initial approximation and avoid computational effort needed to find a starting point close to the central trajectory. The potential difficult of approaching the boundary prematurely is dealt with using the scaled steepest descent direction for globalization.

5. Concluding Remarks. The discretized American option problem using finite difference approximation can be formulated as a finite dimensional complementarity problem. Depending on the finite difference approximation used, this complementarity problem will be a linear complementarity with a tridiagonal coefficient matrix, a general linear complementarity problem, or a nonlinear complementarity problem. As can be seen from (12), even the discretized problem for the European option becomes a nonlinear programming problem with some complementarity constraints when the upstream weighting with the van Leer flux limiter is used for the convection term.

We propose a Newton type interior-point for solving these discretized option pricing problems; the close relationship between the option values of the consecutive time steps is exploited by the Newton process. We demonstrate that a small number of iterations, typically $2 \sim 5$, is required to compute a sufficiently accurate price. More importantly, the number of iterations does not increase with the number of discretization points in the spatial dimension; it actually decreases as the time discretization becomes finer. This makes the proposed method particularly suitable for higher dimension problems.

We investigate the computed (finite difference) delta hedge factors using the complementarity method, the binomial method, and the explicit payoff method. We illustrate that, while the (implicit finite difference) complementarity approach produces continuous delta hedge factors, the binomial method and the explicit payoff method yield discontinuous delta. Therefore, the early exercise curve from the binomial method and the explicit method can be inaccurate. Finally, we demonstrate computationally that the Crank-Nicolson method, which is stable for the European option valuation, can lead to oscillations in the delta hedge factor.

6. Acknowledgement. The authors would like to thank Jacqueline Huang for her comments which have improved the presentation of this paper.

REFERENCES

[1] K. Amin and A. Khanna, *Convergence of American option values from discrete- to continuous-time financial models*, Mathematical Finance, 4 (1994), pp. 289–304.

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

[3] M. Brennan and E. Schwartz, *The valuation of American put options*, Journal of Finance, 32 (1977), pp. 449–462.

[4] M. Broadie and J. Detemple, *American option valuation: New bounds, approximations, and a comparison of existing methods*, The Review of Financial Studies, 9 (1996), pp. 1211–1250.

[5] ———, *Recent advances in numerical methods for pricing derivative securities*, in Numerical Methods in Finance, L. C. G. Rogers and D. Talay, eds., Publication of the Newton Institute, 1997, pp. 43–66.

[6] T. F. Coleman and Y. Li, *On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds*, Mathematical Programming, 67 (1994), pp. 189–224.

[7] ———, *An interior, trust region approach for nonlinear minimization subject to bounds*, SIAM Journal on Optimization, 6 (1996), pp. 418–445.

[8] ———, *A reflective Newton method for minimizing a quadratic function subject to bounds on the variables*, SIAM Journal on Optimization, 6 (1996), pp. 1040–1058.

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

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

[11] ———, *Pricing American stock option by linear programming*, Mathematical Finance, 9 (1999), pp. 229–254.

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

[13] D. Duffie, *Dynamic Asset Pricing Theory*, Princeton, 1996.

[14] J. Huang, *American Options and Complementarity Problems*, Johns Hopkins University Ph.D thesis, 1999.

[15] J. Huang and J. Pang, *Option pricing and linear complementarity*, The Journal of Computational Finance, 2 (1998), pp. 31–60.

[16] J. Hull, *Options, Futures, and Other Derivatives*, Prentice Hull, 1997.

[17] P. Jaillet, D. Lamberton, and B. Lapeyre, *Variational inequalities and pricing of American options*, Acta Applicanda Mathematicae, 21 (1990), pp. 263–289.

[18] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point method for linear programming*, in Progress in Mathematical Programming Interior-point and Related Methods, N. Megiddo, ed., Springer-Verlag, 1989, pp. 29–47.

[19] P. Roache, *Computational Fluid Dynamics*, Hermosa, Albuquerque, New Mexico, 1972.

[20] P. Wilmott, J. Dewynne, and S. Howison, *Option Pricing: Mathematical models and computation*, Oxford Financial Press, 1993.

[21] X. L. Zhang, *Valuation of American options in a jump-diffusion model*, in Numerical Methods in Finance, L. C. G. Rogers and D. Talay, eds., Publication of the Newton Institute, 1997, pp. 93–114.

[22] R. Zvan, K. Vetzal, and P. Forsyth, *Robust numerical methods for pde models of Asian options*, Journal of Computational Finance, (1997), pp. 39–78.