Pricing American Stock Options by Linear Programming

M.A.H. Dempster
Judge Institute of Management Studies
University of Cambridge
Trumpington Street, Cambridge, England CB2 1AG
and
J.P. Hutton
Asset Trading Group,
Nomura Research Institute (Europe) Ltd.,
1 St. Martins-le-Grand,
London EC1A 4NP

August 24, 1999

Abstract

We investigate numerical solution of finite difference approximations to American option pricing problems, using a new direct numerical method — simplex solution of a linear programming formulation. This approach is based on an extension to the parabolic case of the equivalence between linear order complementarity problems and abstract linear programmes known for certain elliptic operators. We test this method empirically, comparing simplex and interior point algorithms with the projected successive overrelaxation (PSOR) algorithm applied to the American vanilla and lookback puts. We conclude that simplex is roughly comparable with projected SOR on average (faster for fine discretisations, slower for coarse), but is more desirable for robustness of solution time under changes in parameters. Furthermore, significant speed-ups over the results given here have been achieved and will be published elsewhere.

1 Introduction

The aim of this paper is to investigate the numerical solution of finite difference approximations to partial differential equation (PDE) problems arising in pricing American options. We use a novel linear programming approach and test it empirically against other methods for such problems. Our results show that with the current state of solver and computer technology it is efficient to solve numerically for the value function of a wide range of American derivative securities by simplex solution of the linear programming formulation.

In §3, after a brief summary of well-known results for the American vanilla put option in §2, we study equivalent formulations of an American option problem as a free boundary problem, a linear complementarity problem and a variational inequality. Results from the literature on uniqueness of the variational inequality solution give us uniqueness of the order formulation of the complementarity problem. Our main theoretical result is an extension to the parabolic case of a known equivalence for coercive elliptic partial differential operators of type Z; namely that this order complementarity problem is equivalent to a least element problem and hence to an abstract linear programme.

In §4, we consider finite difference approximations to the various equivalent formulations of the American put problem in §3. Again, results from the literature on convergence of the solution of the discretised variational inequality to the continuous American put value function give convergence for the equivalent discretised linear programme. Standard numerical algorithms are described.

In §5, we test empirically – for the American vanilla and lookback puts – the new linear programming approach against the PSOR algorithm for the complementarity problem using modern simplex and interior point algorithms. We reproduce known solution values from the literature and investigate solver behaviour with respect to discretisation and market parameters.

2 The American Put Option

Consider the well-known problem of pricing an American stock option in the standard Black-Scholes economy. This problem has been extensively studied and we refer the reader to any standard text, for example, Duffie (1992); for a literature review see Myneni (1992). The *stock price* process is modelled under the *equivalent martingale (risk neutral) measure* as

$$\frac{d\mathbf{S}(t)}{\mathbf{S}(t)} = r dt + \sigma d\tilde{\mathbf{W}}(t) \qquad t \in [0,T], \tag{1}$$

where $S(0) > 0$, $\sigma > 0$ is the constant volatility of the stock, and $\tilde{\mathbf{W}}$ is a Wiener process under this measure.

We consider here a standard (vanilla) *American put option* on a stock, a security whose payoff to the holder on exercise at any *stopping time* $\boldsymbol{\tau} \in [0,T]$ is given by the *payoff function* $\psi(\mathbf{S}(\boldsymbol{\tau})) = (K - \mathbf{S}(\boldsymbol{\tau}))^+$, for a given *maturity date* $T > 0$ and *exercise price* K. The security would be *European* in case exercise were only possible at maturity. We wish to characterise, in a manner suitable for numerical solution, the *value function* $u \colon \mathbb{R}^+ \times [0,T] \to \mathbb{R}$, giving the option fair value $u(x,t)$ to the holder at stock price $x > 0$ and time $t \in [0,T]$.

In case the security were European, then u is simply the solution of the quasilinear parabolic partial differential equation (PDE) derived by Black & Scholes (1973), *viz.*

$$\mathcal{L}_{BS}u + \frac{\partial u}{\partial t} = 0 \tag{2}$$

for $(x,t) \in \mathrm{I\!R}^+ \times [0,T)$ and terminal condition $u(\cdot,T) = \psi$, where the differential operator $\mathcal{L}_{BS} := \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + rx\frac{\partial}{\partial x} - r$. In this case Mania (1997) has recently obtained results which show that u possesses two weak derivatives for $t \in [0,T)$, see § 3.2.

The case of American-style payoffs is more difficult. The value function is the solution of a classical *optimal stopping* problem, namely to choose the stopping time that maximises the conditional expectation of the discounted payoff — indeed the optimal stopping time $\boldsymbol{\rho}(t)$ may be shown to be given by

$$\boldsymbol{\rho}(t) = \inf\left\{s \in [t,T] : u(\mathbf{S}(s),s) = \psi(\mathbf{S}(s)\right\}, \tag{3}$$

i.e. the first time the option value falls to simply that of the payoff for immediate exercise.

The discounted *stopped* price process is a martingale, but only up to the stopping time, so that u satisfies the same PDE (2) on an implicitly defined region $\mathcal{C}$ where $u(x,t) > \psi(x)$, since (3) tells us that exercise occurs when $u(x,t)$ falls to $\psi(x)$. Thus the domain of the value function may be partitioned into a *continuation region* $\mathcal{C}$ and a *stopping region* $\mathcal{S}$ given by

$$\begin{aligned} \mathcal{C} &:= \{(x,t) \in \mathrm{I\!R}^+ \times [0,T) : u(x,t) > \psi(x)\} \\ \mathcal{S} &:= \{(x,t) \in \mathrm{I\!R}^+ \times [0,T) : u(x,t) = \psi(x)\}. \end{aligned} \tag{4}$$

Clearly this is a partition, because we have $u(x,t) \geq \psi(x)$ everywhere.

On the whole domain $\mathrm{I\!R}^+ \times [0,T)$, we have $\mathcal{L}_{BS}u + \frac{\partial u}{\partial t} \leq 0$, since, to preclude arbitrage opportunities, the drift of the (undiscounted) price process cannot be greater than the risk-free rate. However, as long as the current position of the stock price process $(t, \mathbf{S}(t))$ is in $\mathcal{C}$, it is optimal to continue, and hence the PDE (2) *is* satisfied on this region. As soon as the process crosses into $\mathcal{S}$, it is apparent from (3) it is optimal to stop, and on the stopping region $u(x,t) = K - x$, hence $\mathcal{L}_{BS}u + \frac{\partial u}{\partial t} < 0$. These features will be neatly encapsulated in the complementarity problem of §3.

Instead of a simple terminal condition for the PDE, however, we now have a *free boundary* condition: that $u(x,t) = \psi(x)$ for (x,t) on the *optimal stopping boundary* between $\mathcal{C}$ and $\mathcal{S}$. One more condition is necessary to define the optimal stopping boundary, and for the American put this is usually taken to be the *smooth fit* condition $\frac{\partial u}{\partial x} = -1$ on the boundary. For further discussion of these matters see van Moerbecke (1976), Jacka (1991) and Myneni (1992).

Figure 1 is a sketch of the American put value function. The *projections* of the continuation and stopping regions on to the value surface are labelled $\mathcal{C}^p$ and $\mathcal{S}^p$ respectively.

3 Equivalent Formulations of the American Put Problem

The characterisations of the American put value function as optimal stopping and free boundary problems are adequate mathematically, but are not sufficiently explicit to lead to simple numerical schemes. The following formulations of the American put problem as a *linear order complementarity problem* and a *variational inequality* allow us to treat the domain of the value function as an entire region, dispensing with the need to consider explicitly the optimal stopping boundary. For the remainder of this section, we make the usual change of variables to the log-stock price $\xi := \log x$, with respect to which the Black-Scholes PDE for the American put is given by $\mathcal{L}u + \frac{\partial}{\partial t}u = 0$, where $\mathcal{L}$ is the constant coefficient elliptic operator

$$\mathcal{L} := \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial \xi^2} + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial}{\partial \xi} - r, \tag{5}$$

and u now refers to the option value as a function of ξ. The various inequalities for the operator $\mathcal{L}_{BS}$ in the §2 carry over to the log-transformed version $\mathcal{L}$. Note that we now have a new payoff function given by $\tilde{\psi}(\xi) := (K - e^{\xi})^+$ and continuation and stopping regions $\hat{\mathcal{C}}$ and $\hat{\mathcal{S}}$ defined with respect to the new variable ξ.

3.1 The order complementarity problem

The American put value function u satisfies $\mathcal{L}u + \frac{\partial u}{\partial t} = 0$ and $u > \tilde{\psi}$ on $\hat{\mathcal{C}}$, so that $(\mathcal{L}u + \frac{\partial u}{\partial t}) \wedge (u - \tilde{\psi}) = 0$, where $\wedge$ denotes pointwise minimum of the two functions. On $\hat{\mathcal{S}}$, $\mathcal{L}u + \frac{\partial u}{\partial t} < 0$ and $u(\xi, t) = \tilde{\psi}(\xi)$, so that if we formally require $\mathcal{L}u + \frac{\partial u}{\partial t} \geq 0$ we again have $(\mathcal{L}u + \frac{\partial u}{\partial t}) \wedge (u - \tilde{\psi}) = 0$. We may express the free boundary problem for the American put option in a form that encapsulates these main complementary properties as the following *order complementarity problem* of Borwein & Dempster (1989).

Theorem 1 *The American put value function is the unique solution to the linear order complementarity problem*

$$\text{(OCP)} \qquad \begin{cases} u(\cdot, T) = \tilde{\psi} \\ u \geq \tilde{\psi} \\ \mathcal{L}u + \frac{\partial u}{\partial t} \geq 0 \\ (\mathcal{L}u + \frac{\partial u}{\partial t}) \wedge (u - \tilde{\psi}) = 0 \qquad \text{a.e. } \mathbb{R} \times [0, T]. \end{cases}$$

■

For (OCP) to be well-posed, we must restrict it to a *vector lattice*, which is a vector space with a partial order defined by a positive cone P such that for any points x and y the maximum $x \vee y$ and the minimum $x \wedge y$ exist in the given order. See Borwein &

Dempster (1989) and Cryer & Dempster (1980) for further discussion. We give the precise setting in the sequel. To prove that the American put value function is the unique solution of (OCP), we express it in another equivalent form, namely as a parabolic *variational inequality*, in which form we may apply some standard results on uniqueness of solutions to such variational inequalities.

3.2 (OCP) as a variational inequality

Before we give the variational inequality formulation some definitions will be needed. Technically, we must specify a function space for the variational inequality solution, chosen ideally as a minimal set of restrictions so that it is well-posed. Define the *Sobolev* space $W^{m,p,\mu}(\mathbb{R}_2)$ as the space of functions $u \in L^p\left(\mathbb{R}_2, e^{-\mu|x|}\,dx\right)$ whose *weak derivatives* of order not exceeding $m \in \mathbb{N}$ exist and are also in $L^p\left(\mathbb{R}_2, e^{-\mu|x|}\,dx\right)$, for $p \in [0,\infty]$ and $\mu \in (0,\infty)$. (Here $|.|$ denotes the L^1 norm on $\mathbb{R}^2$ and dx denotes Lebesgue measure on $\mathbb{R}_2$, and it should be noted that the extension of the results in the sequel to $\mathbb{R}_{n+1}$, for arbitrary $n \in \mathbb{N}$, is completely straightforward.) We shall be interested in the *Hilbert* space $H^1(\mathbb{R}_2) := W^{1,2,\mu}(\mathbb{R}_2)$, for some fixed $\mu > 0$, of square integrable functions with square integrable derivatives defined on $\mathbb{R}_2$. The Hilbert space $H^1(\mathbb{R}_2)$ has as Banach dual the Sobolev space $H^{-1}(\mathbb{R}_2) := W^{-1,2,\mu}(\mathbb{R}_2)$, also a Hilbert space of Radon measures, with which it may be identified. Consider the pairing $\langle\cdot,\cdot\rangle : H^1 \times H^{-1} \to \mathbb{R}$ between dual spaces given by

$$\langle u, v\rangle := \int_{\mathbb{R}_2} u(\xi,t)v(\xi,t)e^{-\mu(|\xi|+|t|)}\,d\xi\,dt, \tag{6}$$

where we may interpret $v \in H^{-1}$ as the density function of the Radon measure element of the dual space H^{-1} of H^1 with respect to $e^{-\mu(|\xi|+|t|)}\,d\xi\,dt$. Alternatively, we may consider $\langle\cdot,\cdot\rangle$ given by (6) as an inner product on the Hilbert space $H^0(\mathbb{R}_2) := L^2(\mathbb{R}_2, e^{-\mu|x|}dx)$ by virtue of the canonical injections $H^1 \hookrightarrow H^0 \hookrightarrow H^{-1}$, see Baiocchi & Capelo (1984, p.79). In this setting the partial differential operator $\mathcal{L}$ may be interpreted either as a map $H^1 \to H^{-1}$ or as an operator on H^1. Consider also the bilinear form $a(\cdot,\cdot) : H^1 \times H^1 \to \mathbb{R}$ given by

$$\begin{aligned} a(u,v) \;:=\; & \int_{\mathbb{R}_2} \frac{\sigma^2}{2} u_\xi v_\xi e^{-\mu(|\xi|+|t|)}\,d\xi\,dt - \int_{\mathbb{R}_2}\left((r-\sigma^2/2) + \mu\frac{\sigma^2}{2}\frac{\xi}{|\xi|}\right) u_\xi v e^{-\mu(|\xi|+|t|)}\,d\xi\,dt \\ & + \int_{\mathbb{R}_2} ruve^{-\mu(|\xi|+|t|)}\,d\xi\,dt, \quad \forall u,v \in H^1. \end{aligned} \tag{7}$$

Finally, note that H^1 (and hence H^{-1}) is a vector lattice Hilbert space (but *not* a Hilbert lattice) with positive cone defined in terms of (Lebesgue) almost everywhere nonnegativity. See Baiocchi & Capelo (1984), Cryer & Dempster (1980) and Borwein & Dempster (1989, p.553–554), for more details on these ideas, which have been adapted here to match the more general setting of Jaillet, Lamberton & Lapeyre (1990). In particular, we shall assume all functions in H^1 ($\cong H^{-1}$) considered to be defined as $u(\cdot, |t|)$ on $\mathbb{R} \times (-\infty, 0)$ and as $u(., T)$ on $\mathbb{R} \times [T, \infty)$ (see Cryer & Dempster (1980, p.89 *et seq.*).

The following lemma relates the bilinear form $a(\cdot, \cdot)$ to the elliptic part of the partial differential operator $\mathcal{L}$, and we will use it to show variational inequality (VI) and (OCP) equivalence.

Lemma 1 *The bilinear form a satisfies*

$$a(u, v) = \langle v, \mathcal{L}u \rangle \quad u, v \in H^1. \tag{8}$$

Proof: See Cryer & Dempster (1980, p.80–81), and Jacka (1991, p.72 *et seq.*). However, the idea is simple enough — simply integrate the first term of (7) by parts. ■

With the preceding definitions and Lemma 1 we may now state the variational inequality formulation in the form of the following theorem.

Theorem 2 *The variational inequality (VI) given by*

$$\text{(VI)} \qquad \begin{cases} u(., T) = \tilde{\psi} \\ u \geq \tilde{\psi} \\ v \geq \tilde{\psi} \text{ a.e. } \Rightarrow a(u, v-u) + \langle v - u, \frac{\partial u}{\partial t} \rangle \geq 0 \quad \text{a.e. } [0, T] \end{cases}$$

is equivalent to the order complementarity problem (OCP).

Proof: This is again a well known result, due to Borwein & Dempster (1989). We can rewrite the third line of (VI) using Lemma 1 as

$$\left\langle v - u, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle \geq 0 \;\; \forall v \geq \tilde{\psi}. \tag{9}$$

Let u solve (VI): Choosing arbitrary $v \geq u$ in (9) gives $\mathcal{L}u + \frac{\partial u}{\partial t} \geq 0$, which is the second constraint in (OCP). This in turn implies, since $u - \tilde{\psi} \geq 0$, that $\left\langle u - \tilde{\psi}, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle \geq 0$. Now if we choose $v = \tilde{\psi}$, (9) becomes $\left\langle u - \tilde{\psi}, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle \leq 0$, which two inequalities together give the complementarity condition $\left\langle u - \tilde{\psi}, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle = 0$, equivalent to the third constraint of (OCP) (Borwein & Dempster (1989, p.549)).
Let u solve (OCP): Then

$$\begin{aligned} \left\langle v - u, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle &= \left\langle v - \tilde{\psi}, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle + \left\langle \tilde{\psi} - u, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle \\ &= \left\langle v - \tilde{\psi}, \mathcal{L}u + \frac{\partial u}{\partial t} \right\rangle \geq 0 \quad \forall v \geq \tilde{\psi}. \end{aligned} \tag{10}$$

So we see that (VI) and (OCP) are equivalent. ■

The key property that will determine uniqueness of the solution to (VI), and hence (OCP), is that of *coercivity* of the bilinear form or differential operator, defined as follows.

Definition A continuous bilinear form $a(\cdot,\cdot)$ defined on a Hilbert space H is *coercive* on H iff

$$\exists \alpha \in \mathbb{R}^+ \text{ s.t. } a(u,u) \geq \alpha \|u\|^2 \quad \forall u \in H.$$

Similarly, an operator $\mathcal{T}$ on H is coercive iff

$$\exists \alpha \in \mathbb{R}^+ \text{ s.t. } \langle u, \mathcal{T}u \rangle \geq \alpha \|u\|^2 \quad \forall u \in H.$$

■

It can be shown that a given by (7), and hence $\mathcal{L}$, is coercive (see Jaillet *et al.* (1990, p.267), whose spaces $L^2([0,T],V_\mu)$ and $L^2([0,T],H_\mu)$ may be considered restrictions respectively of our spaces H^1 and H^0.). Then the Lions-Stampacchia theorem implies that the solution to (VI) is unique (see, for example, Baiocchi & Capelo (1984, p. 24 *et seq.*)). This result completes the proof of Theorem 1, since (VI) has a unique solution and is equivalent to (OCP). The formulation (VI) is a type of classical physical problem, termed the *(Stefan) obstacle problem*, where the payoff function $\tilde{\psi}$ is the obstacle below which the solution cannot fall.

3.3 Abstract linear programme equivalent formulation

In the previous section we have established the uniqueness of the solution to (OCP) by considering its formulation as the variational inequality (VI). We now derive the key result which will eventually enable us to compute a numerical approximation to the value function of the American put – and indeed many other types of American derivative securities – as an ordinary linear programme. First we need some definitions.

Definition A linear operator $\mathcal{T}$ on a Hilbert space H is of *type Z* iff

$$\langle u, v \rangle = 0 \Rightarrow \langle u, \mathcal{T}v \rangle \leq 0 \quad \forall u, v \in H. \tag{11}$$

■

Definition Define, for a closed subset $F \subseteq P \subseteq H$ of a vector lattice Hilbert space H with positive cone $P := \{v \in H : v \geq 0\}$, the *least element problem*

$$\text{(LE)} \qquad \text{find } u \in F \text{ s.t. } u \leq v \quad \forall v \in F.$$

■

The least element is denoted by $u = \mathrm{LE}(F)$, and is illustrated in Figure 2. Note that if it exists, the least element is always unique since, if u_1 and u_2 are least elements of F, then $u_1 \leq u_2$ and $u_2 \leq u_1$, so from the vector lattice property $u_1 = u_2$.

We now define an associated problem, the abstract linear programme (LP).

Definition Define, for a subset $F \subseteq P \subseteq H$ of a vector lattice Hilbert (function) space H with positive cone P and constant vector $c > 0$ a.e. with respect to Lebesgue measure on its domain, the *abstract linear programme*

$$\text{(LP)} \qquad \inf_{v \in P} \langle c, v \rangle \text{ s.t. } v \in F.$$

■

The following theorem, giving equivalence between (OCP), (LE) and (LP), is an extension of a result of Cryer & Dempster (1980) for elliptic partial differential operators to the parabolic case.

Theorem 3 *In the setting described above, if $\mathcal{T}$ is a coercive type Z temporally homogeneous elliptic differential operator, then there exists a unique solution u to the following equivalent problems:*

$$\text{(OCP)} \qquad \begin{cases} u(\cdot, T) = \tilde{\psi} \\ u \geq \tilde{\psi} \\ \mathcal{T}u + \frac{\partial u}{\partial t} \geq 0 \\ (\mathcal{T}u + \frac{\partial u}{\partial t}) \wedge (u - \tilde{\psi}) = 0 \quad \text{a.e. } \mathbb{R} \times [0, T], \end{cases}$$

$$\text{(LE)} \qquad \text{find } u = \text{LE}(F),$$

$$\text{(LP)} \qquad \inf_{v} \langle v, c \rangle \text{ s.t. } v \in F,$$

for any $c > 0$ a.e. on $\mathbb{R} \times [0, T]$, where

$$F := \left\{ v : \ v(.\cdot, T) = \tilde{\psi}, \ v \geq \tilde{\psi}, \ \mathcal{T}v + \frac{\partial v}{\partial t} \geq 0 \right\}. \tag{12}$$

Proof: We first prove the equivalence between (OCP) and (LE), after making the trivial domain extensions of the problem functions given above to set them in H^1. It will be necessary to reverse time, so that in backwards time $\mathcal{T}u - \frac{\partial u}{\partial t} \geq 0$ and the terminal condition becomes the *initial* condition $u(., 0) = \tilde{\psi}$. Let L denote the Laplace transform operator with respect to the measure $e^{-\mu|t|}$, so that, for $(\xi, \lambda) \in \mathbb{R}_2$, the Laplace transform $\hat{u} \in H^1$ of a function $u \in H^1$ is defined by

$$\hat{u}(\xi, \lambda) := Lu(\xi, \cdot)(\lambda) := \int_0^\infty e^{-|\lambda| t} u(\xi, t) e^{-\mu t} dt. \tag{13}$$

As noted above, we have extended the temporal domain of our value functions u to $[0, \infty)$ as constant on (T, ∞), so that this generalised Laplace transform is well defined. L is a linear operator and $\mathcal{T}$ is temporally homogeneous, i.e. has time-independent coefficients, and therefore commutes with the Laplace operator, so that taking the Laplace transform of the operator $\mathcal{T} - \frac{\partial}{\partial t}$ gives $\mathcal{T}L - L\frac{\partial}{\partial t}$. The Laplace transform of the first order time derivative is given by

$$\begin{aligned}\left(L\frac{\partial u}{\partial t}\right)(\xi,\lambda) &:= \int_0^\infty e^{-|\lambda|t}\frac{\partial u}{\partial t}(\xi,t)e^{-\mu t}dt \\ &= \left[e^{-(|\lambda|+\mu)t}u(\xi,t)\right]_0^\infty + (|\lambda|+\mu)\int_0^\infty e^{-|\lambda|t}u(\xi,t)e^{-\mu t}dt \\ &= -u(\xi,0) + (|\lambda|+\mu)\hat{u}(\xi,\lambda) \end{aligned} \tag{14}$$

and $u(\xi,0)$ is given by the initial condition $u(\cdot,0) = \tilde{\psi}$.

Now, note that the Laplace transform is *positivity-preserving* in the sense that $u \geq 0 \Rightarrow \hat{u} \geq 0$ a.e. on $\mathbb{R}_2$. Then, writing the initial condition, constant in λ, as $\hat{q}(\cdot,\lambda) :\equiv u(\cdot,0) - |\lambda| - \mu$ to agree with the notation of Borwein & Dempster (1989), (OCP) is equivalent to the transformed order complementarity problem $(\widehat{\text{OCP}})$, also posed in H^1, given by

$$(\widehat{\text{OCP}}) \qquad \begin{cases} \hat{u} \geq \hat{\psi} \\ \mathcal{T}\hat{u} + \hat{q} \geq 0 \\ (\mathcal{T}\hat{u} + \hat{q}) \wedge (\hat{u} - \hat{\psi}) = 0 \quad \text{a.e. on } \mathbb{R}_2, \end{cases}$$

where $\hat{\psi}$ is the Laplace transform of the log-transformed payoff function $\tilde{\psi}$, given by $\hat{\psi}(\xi,\lambda) = \tilde{\psi}(\xi)/(|\lambda|+\mu)$. Since $\mathcal{T}$ is coercive, type Z and elliptic, $\hat{u}(\cdot,\lambda)$ is the unique solution to the projected (OCP) obtained from $(\widehat{\text{OCP}})$ by fixing $\lambda \in \mathbb{R}$. We can now apply the order complementarity-least element equivalence result of Borwein & Dempster (1989) for coercive type Z *elliptic* operators, so that for each $\lambda \in \mathbb{R}$, $\hat{u}(\cdot,\lambda)$, is the solution (necessarily unique) to the least element problem defined by LE$(\hat{F}_\lambda)$, where $\hat{F}_\lambda$ is defined by

$$\hat{F}_\lambda := \left\{\hat{u}(\cdot,\lambda) : \hat{u}(\cdot,\lambda) \geq \tilde{\psi}(\cdot,\lambda), \mathcal{T}\hat{u}(\cdot,\lambda) + \hat{q}(\cdot,\lambda) \geq 0\right\}. \tag{15}$$

It follows that $\hat{u}$ is the unique solution to the least element problem $(\widehat{\text{LE}})$ defined by LE$(\hat{F})$, where $\hat{F}$ is defined by

$$\hat{F} := \{\hat{u} : \ \hat{u} \geq \hat{\psi}, \ \mathcal{T}\hat{u} + \hat{q} \geq 0\}. \tag{16}$$

Applying the inverse Laplace transform L^{-1} to $\hat{u}$ shows that

$$u = L^{-1}\hat{u}$$

solves both (LE), given by LE(F), and (OCP), as required. Indeed suppose the contrary, i.e. that there exists $v \in F$ such that $v \leq u$, $v \neq u$. Then it follows since L is positivity preserving that $\hat{v} \in \hat{F}$ and $\hat{v} \leq \hat{u}$, $\hat{v} \neq \hat{u}$, a contradiction to $\hat{u} = \text{LE}(\hat{F})$.

With this least element result, the LP equivalence is immediate — u is the least element of $F \iff u \leq v$ for all $v \in F$, and so $\langle c,u\rangle \leq \langle c,v\rangle$ for all $v \in F$ and any vector $c > 0$.

Therefore u minimises $\langle c, v\rangle$ over all v in F and is thus the solution to the abstract linear programme (LP). Restricting to the original problem domain yields the result. ■

It should be noted that the above proof depends on time running 'backwards', that is, expressed in terms of time to maturity, otherwise we cannot substitute $\tilde{\psi}$ for $u(\cdot, 0)$ in (14). The finite dimensional least element-linear programme equivalence is illustrated for $\mathbb{R}^2$ in Figure 2. The least element result tells us that the linear constraint set lies within the positive cone translated so that its apex lies at u, since in finite dimensions the least vector is least in every element. We see immediately that we pick out the least element of the constraint set by minimising $\langle c, u\rangle$ over the set $u \in F$, where $c > 0$; specifically in $\mathbb{R}^2$ by minimising the intercept of negatively sloped lines defined by $c'u$ with normal $c > 0$ intersecting F.

This general result gives equivalence between (VI), (OCP), (LE) and (LP) for the American put, since $\mathcal{L}$ is coercive type Z (see Jaillet *et al.* (1990)). It should be stressed that Theorem 3 is very general, and applies to virtually any parabolic partial differential operator with a temporally homogeneous coercive type Z elliptic part, and virtually any payoff function. For example, it may be applied to the Black-Scholes operator $\mathcal{L}_{BS}$ directly.

The first part of the proof of Theorem 3 is easily generalised to the case of parabolic operators with time-dependent coefficient elliptic part $\mathcal{T}$ by considering the operator $L\mathcal{T}L^{-1}$ on functions defined on the price-frequency domain. However, the difficulty in extending the result to the time-dependent coefficient case by this method lies in verifying that the new operator inherits the coercive type Z properties from $\mathcal{T}$. Replacing L by the norm-preserving orthogonal Fourier transform verifies the required inheritance trivially, but introduces complex valued functions which cannot be naturally ordered. A more delicate argument involving step function coefficient approximation and a suitable passage to the limit can be however be used to establish the results of Theorem 3 for time-dependent coefficient operators; the details will appear elsewhere.

Theorem 3 also suggests a simple way to solve the equivalent problems numerically — by a suitable discretisation, the infinite-dimensional abstract linear programme (LP) reduces to an ordinary linear programme with solutions in $\mathbb{R}^n$. This is a standard problem type with an extensive literature devoted to rapid solution, and efficient solution software is readily available. In the next section we discretise the problem and consider our suggested LP and alternative numerical solution methods.

4 Numerical Methods

In general, there is no known closed form solution to an American option problem, and we are unlikely to find one. In this section, we consider numerical solution of the American vanilla put problem using the novel formulation of the value function in §3.1 as the solution to an abstract linear programme. When we discretise space and time by standard finite differences this becomes an ordinary linear programme which we may solve by well-known algorithms.

4.1 Localisation and Discretisation of the value function

As a first approximation, we restrict the domain of the value function $\mathbb{R} \times [0, T]$ to a finite region $[L, U] \times [0, T]$, for any $L < \log K < U$, and set the value function on the boundaries as $u(L, \cdot) = \tilde{\psi}(L)$ and $u(U, \cdot) = \tilde{\psi}(U)$.

Defining the localised inner product as in (6) but integrated over $[L, U]$ in the first variable, gives a localised version of the linear programme (LP), for which there still exists a unique solution, since the operator is unchanged. As $L, U \to \infty$, this solution tends uniformly to the solution to (LP), i.e. the American put value function on the whole domain, a result demonstrated by Jaillet *et al.* (1990) for the equivalent localised variational inequality — naturally the equivalent localised (OCP) and (LP) inherit this same convergence property.

We discretise the localised LP by approximating the value function by a piecewise constant function, constant on rectangular intervals around points in a regular *lattice* or *mesh*, on the domain $[L, U] \times [0, T]$. (Note that everything that follows holds for irregular meshes with trivial modifications.) Write u_i^m as the value of the general function u at mesh points (i, m) defined by

$$u_i^m := u\left(L + i\Delta\xi, T - m\Delta t\right), \tag{17}$$

where $m \in \{0, 1, \ldots, M\} := \mathcal{M}$ and $i \in \{0, 1, \ldots, I\} := \mathcal{I}$. Writing $\tilde{\psi}_i := \tilde{\psi}(L + i\Delta\xi)$, we have the boundary values $u_0^m = \tilde{\psi}_0$, $u_I^m = \tilde{\psi}_I$ and, because m is a backwards time index, $u_i^0 = \tilde{\psi}_i$.

We now approximate the partial derivatives which appear in $\mathcal{L}$ by discrete analogues, using *finite difference* approximations. We approximate the partial derivatives of the value function at a point indexed by (i, m) in the interior of the index domain $\mathcal{I} \times \mathcal{M}$ by

$$\begin{aligned}
\frac{\partial u}{\partial \xi} &\approx \theta \frac{u_{i+1}^m - u_{i-1}^m}{2\Delta\xi} + (1-\theta)\frac{u_{i+1}^{m-1} - u_{i-1}^{m-1}}{2\Delta\xi} \\
\frac{\partial^2 u}{\partial \xi^2} &\approx \theta \frac{u_{i+1}^m - 2u_i^m + u_{i-1}^m}{(\Delta\xi)^2} + (1-\theta)\frac{u_{i+1}^{m-1} - 2u_i^{m-1} + u_{i-1}^{m-1}}{(\Delta\xi)^2} \\
\frac{\partial u}{\partial t} &\approx \frac{u_i^{m-1} - u_i^m}{\Delta t}
\end{aligned} \tag{18}$$

for $\theta \in [0, 1]$. The cases $\theta = 0$, $\theta = \frac{1}{2}$, $\theta = 1$ correspond to *explicit*, *Crank-Nicolson* and *implicit* discretisation schemes respectively, all of which are second order accurate in $\Delta\xi$ and first order accurate in Δt, except for $\theta = \frac{1}{2}$, which gives second order accuracy in Δt.

Substitution of these discrete forms for their counterparts in (LOCP) gives the discrete order complementarity problem (DOCP):

$$\begin{cases}
u_i^m \geq \tilde{\psi}_i, \ u_i^0 = \tilde{\psi}_i, \ u_I^m = 0, \ u_0^m = \tilde{\psi}_0 \\
a u_{i-1}^m + b u_i^m + c u_{i+1}^m + d u_{i-1}^{m-1} + e u_i^{m-1} + f u_{i+1}^{m-1} \geq 0 \\
\left(a u_{i-1}^m + b u_i^m + c u_{i+1}^m + d u_{i-1}^{m-1} + e u_i^{m-1} + f u_{i+1}^{m-1}\right) \wedge \left(u_i^m - \tilde{\psi}_i\right) = 0 \\
\qquad i \in \mathcal{I} \setminus \{0, I\}, \ m \in \mathcal{M} \setminus \{0\},
\end{cases} \tag{19}$$

where

$$
\begin{array}{ll}
a := -\theta\left[\frac{\sigma^2\Delta t}{2\Delta\xi^2} - \frac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] & b := 1 + r\Delta t + \theta\frac{\sigma^2\Delta t}{\Delta\xi^2} \\
c := -\theta\left[\frac{\sigma^2\Delta t}{2\Delta\xi^2} + \frac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] & d := -(1-\theta)\left[\frac{\sigma^2\Delta t}{2\Delta\xi^2} - \frac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] \\
e := (1-\theta)\frac{\sigma^2\Delta t}{\Delta\xi^2} - 1 & f := -(1-\theta)\left[\frac{\sigma^2\Delta t}{2\Delta\xi^2} + \frac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right].
\end{array}
\tag{20}
$$

We discuss well-posedness and convergence in the sequel — first of all we express the complementarity condition of (19) in matrix form by collapsing the space and time indices into vectors. Put

$$
u^m := \begin{pmatrix} u_1^m \\ \vdots \\ u_{I-1}^m \end{pmatrix} \quad \tilde{\psi} := \begin{pmatrix} \tilde{\psi}_1 \\ \vdots \\ \tilde{\psi}_{I-1} \end{pmatrix} \quad \phi := \begin{pmatrix} -(a+d)\tilde{\psi}_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{21}
$$

Then, substituting $u_0^m = \tilde{\psi}_0$ and $u_I^m = 0$ into (19), the complementarity condition becomes

$$
(Bu^{m-1} + Au^m - \phi) \wedge (u^m - \tilde{\psi}) = 0 \quad m \in M \setminus \{0\}, \tag{22}
$$

where, defining notation for a *tridiagonal* matrix as

$$
\mathrm{Td}_k(a_k, b_k, c_k) := \begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{K-2} & b_{K-2} & c_{K-2} \\ & & & a_{K-1} & b_{K-1} \end{pmatrix}, \tag{23}
$$

A and B are the $(I-1)$-square tridiagonal matrices

$$
A := \mathrm{Td}_i(a, b, c), \;\; B := \mathrm{Td}_i(d, e, f). \tag{24}
$$

Now we can collapse the time index m by putting

$$
u := \begin{pmatrix} u^1 \\ \vdots \\ u^M \end{pmatrix} \quad \tilde{\Psi} := \begin{pmatrix} \tilde{\psi} \\ \vdots \\ \tilde{\psi} \end{pmatrix} \quad \Phi := \begin{pmatrix} \phi - B\tilde{\psi} \\ \phi \\ \vdots \\ \phi \end{pmatrix} \tag{25}
$$

and may in turn express (22) as

$$
(Cu - \Phi) \wedge (u - \tilde{\Psi}) = 0, \tag{26}
$$

where C is the $M(I-1)$-square 'staircase' matrix given by

$$
C := \begin{pmatrix} A & & & \\ B & A & & \\ & \ddots & \ddots & \\ & & B & A \end{pmatrix}. \tag{27}
$$

So, the discretisation scheme we have described leads us to approximate $u(\xi, t)$ by a step function whose value on grid points $(u_i^m)_{(i,m)\in I\times M}$, in the vector form given by (21) and (25), is the solution $u \in \mathbb{R}^{M(I-1)}$ of the finite dimensional order complementarity problem

$$\text{(DOCP)} \qquad \begin{cases} u \geq \tilde{\Psi} \\ Cu \geq \Phi \\ (Cu - \Phi) \wedge (u - \tilde{\Psi}) = 0, \end{cases}$$

with the boundary values $u_I^{(\cdot)} = 0$, $u_0^{(\cdot)} = \tilde{\psi}_0$ and $u^0 = \tilde{\psi}$ — we give these separately because the boundary conditions have been substituted in to (DOCP) and do not appear in its solution as written.

However, before we can write down a well posed equivalent linear programme, we have to verify that the conditions of Theorem 3, namely the type Z property and coercivity of the operator, are satisfied in the matrix sense. Considering (22), since u^{m-1} is known at step m, the discretised operator $\mathcal{L}$ is represented in finite dimensions by the matrix A, so we require that A be type Z and coercive. It is simple to show that a matrix is type Z if an only if it has non-negative off-diagonal coefficients, the classical definition of a *Z matrix* Borwein & Dempster (1989). Clearly, A is type Z if and only if $a \leq 0$ and $c \leq 0$, which is the case if and only if

$$|r - \sigma^2/2| \leq \sigma^2/\Delta\xi. \tag{28}$$

This condition holds for all parameter values simply by taking I large enough, and indeed for realistic parameter values the critical value of I is very small. If condition (28) holds, it is a simple matter to show that A is then coercive — see Jaillet *et al.* (1990), or Hutton (1995).

Assuming (28) holds then, we may now use an appropriate version of Theorem 3 and write down the equivalent discretised version of (LP) as, for any fixed $c > 0$ in $\mathbb{R}^{(I-1)M}$,

$$\text{(OLP)} \qquad \begin{cases} \min & c'u \\ \text{s.t.} & u \geq \tilde{\Psi} \\ & Cu \geq \Phi \end{cases}$$

with the boundary values $u_I^{(\cdot)} = 0$, $u_0^{(\cdot)} = \tilde{\psi}_0$ and $u^0 = \tilde{\psi}$.

(OLP) is an *ordinary* linear programme which is easily solved numerically. Jaillet *et al.* (1990) show that as $M, I \to \infty$, and in case $\theta < 1$, such that the *mesh ratio* $\rho := \frac{\Delta t}{(\Delta\xi)^2} \to 0$, the solution of the equivalent discretised localised variational inequality converges to the solution of the localised variational inequality, which as already mentioned, itself converges uniformly, as $L, U \to \infty$, to the American put value function on the whole domain. By virtue of Theorem 3, these same convergence properties are naturally inherited by (DOCP) and (OLP).

It seems that unconditional convergence and stability for the case $\frac{1}{2} < \theta < 1$ is not yet proven but is well known for the case of equations; for $0 < \theta < \frac{1}{2}$, we have convergence of the scheme if and only if

$$0 \le \rho \le \frac{1}{\sigma^2(1-2\theta)}. \tag{29}$$

4.2 Solution of the discrete problem

In the discrete problem given in §4.1, the equivalent (DOCP) and (OLP) are presented as global in time. We can, however, decompose completely the global problem suggested by (DOCP) and (OLP) by stepping through time — we can solve (DOCP) by solving the sequence of complementarity problems (22) for $m = 1 \dots M$. Noting that each complementarity problem is well-posed i.e. a unique solution exists, since A is a square coercive Z-matrix and u^{m-1} is known at time step m.

Each order complementarity problem in (22) has an equivalent linear programme, so that we get the following decomposition of (OLP):

$$\begin{aligned} &\min\ c' u^m \\ &\text{s.t. } u^m \ge \psi \\ &A u^m \ge \phi - B u^{m-1} \qquad m = 1, \dots, M. \end{aligned} \tag{30}$$

Note that this statement illustrates the fact that one can get an LP equivalence from the special case of the discrete complementarity problem - linear programme equivalence due to Mangasarian (1979). Solving either (DOCP) or (OLP) in this way is computationally far quicker and more memory-efficient than solving the global problem. We now consider suitable algorithms for solving the sub-problems.

The standard approach to solving the finite difference formulation for the American put is via the complementarity problem (DOCP), and there are two main approaches — one iterative, the other direct — to solving this problem. By far the most popular is the iterative method of *projected successive over-relaxation* (PSOR) due to Cryer (1971), and it is against this method that we test our proposed linear programming method. Pivoting methods (which are direct) may be used for the complementarity problem, however these tend to be more general and less well-developed than the simplex algorithm. (See Jaillet *et al.* (1990) for further details on pivoting methods in the current setting.)

4.2.1 Solution of (OLP)

There are again two main algorithms for solving linear programmes such as (OLP), namely the (direct) *simplex* method, due to Dantzig (1963), and the (iterative) *interior point* method, first applied to linear programmes by Koopmans (1951) and recently reintroduced by Karmarkar (1984). Our preliminary results from the interior point method were poor — see Hutton (1995) for more details — so we concentrate here on the simplex method and outline the salient features that we propose to exploit in §5.

4.2.2 Simplex method

The dual of the mth sub-problem (30) of (OLP) is, for any $c > 0$,

$$\begin{aligned} &\max\ (\tilde{\psi}' | (\phi - Bu^{m-1})')y^m \\ &\text{s.t. } 0 \leq (I|A')y^m \leq c. \end{aligned} \tag{31}$$

In the primal sub-problem, given by (30), only the right-hand side changes from the preceding sub-problem, which means that only the objective function of the dual sub-problem changes, as we see from (31), and so the optimal dual solution to the preceding sub-problem is still a basic feasible solution to the current dual sub-problem. This means at least that phase 1 — the feasibility search — is unnecessary. Since we assume Δt is small, the optimal solution of one sub-problem has only a few basic variables changed from that of the preceding problem; the preceding dual solution should not be too many pivots away from the current problem's dual solution. Not surprisingly, therefore, we see in §5 that for this problem once we 'hot-start' the solver from the previous time step's optimal basis, the dual simplex method is superior to the primal.

The simplex optimal basis has an immediate interpretation in terms of the problem: u_i is in the optimal basis if and only if $u_i^m > \tilde{\psi}_i$, i.e. the point indexed by (i, m) is in the discrete approximation of the continuation region $\mathcal{C}$. Similarly, u_i is non-basic if and only if $u_i^m = \tilde{\psi}_i$ and hence the point indexed by (i, m) is in the discrete approximation of the stopping region $\mathcal{S}$.

Finally, it is vital in any approach to PDE-type problems that the typically very large matrices in question, i.e. A and B, are stored in a way that exploits their *sparsity*. The OSL routines we use in §5 store only the non-zero elements and enough information to locate them, in so-called *storage-by-columns*.

4.2.3 The explicit method

The various algorithms for solving the discrete problem described above are in practice only applied to the class of *implicit* methods. (We distinguish between *the* implicit method, which has $\theta = 1$, and general implicit methods, which have $\theta > 0$.) For the *explicit* method ($\theta = 0$) we can write down the each time step's solution in a simple way. The constraint matrix A defined by (24) reduces to the $(I - 1)$-square diagonal matrix $\text{diag}(1 + r\Delta t)$, so that in fact the general mth sub-problem of (OLP) given by (30) reduces to

$$\begin{aligned} &\min\ c'u^m \\ &\text{s.t. } u^m \geq \tilde{\psi} \\ &(1 + r\Delta t)u^m \geq \phi - Bu^{m-1}. \end{aligned} \tag{32}$$

We can solve this by inspection — the solution u^m is explicitly determined from the previous time step's solution u^{m-1} by

$$u^m = \tilde{\psi} \vee \left(\frac{1}{1 + r\Delta t} \left(\phi - Bu^{m-1}\right) \right). \tag{33}$$

This is clearly a very rapid calculation for each iteration, the only significant calculation being a single matrix multiplication. However, the implications of (29) for the explicit method are profound, because we must take a number of time steps of the order of the *square* of the number of space steps, which can become computationally very demanding.

The dual of the explicit method mth sub-problem, given in general in standard form by (31), is also very simple, and is given by, for any $c > 0$,

$$\begin{aligned} &\max\ (\tilde{\psi}', (\phi - Bu^{m-1})')y^m \\ &\text{s.t. } 0 \leq y^m \leq c. \end{aligned} \tag{34}$$

Note that we have exploited the arbitrary value of c and the fact that A is diagonal to change, without loss of generality, the constraint $(I|A')y^m \leq c$ to $y^m \leq c$. Again this has a very simple solution, namely

$$y^m = \operatorname{sgn}\left(\tilde{\psi}', (\phi - Bu^{m-1})'\right) c, \tag{35}$$

where $\operatorname{sgn}(x)$ is the diagonal matrix with element i equal to 1 if $x_i > 0$, equal to 0 if $x_i < 0$ and arbitrary if $x_i = 0$. Note that the dual solution is not unique, since some elements of $\tilde{\psi}$ are zero and hence the corresponding variables may be set arbitrarily.

The dual solution is of the same order of computational complexity as the primal, but requires the previous time step's primal solution. We use the simpler primal method in our empirical tests in §5.

4.3 American lookback put

As a further test of the proposed linear programming method, we solve a variant of the vanilla put, namely the continuously-sampled lookback put, where the path dependent strike price is given by the maximum of the stock price process to date. In this case, the two state variables, the current stock price x and the current maximum y, may be encapsulated by a similarity transformation into the single variable $z := \log(y/x)$. It is straightforward to show that the normalised value function $V^* := \frac{1}{x}V$ solves the abstract linear programme (LP), but with a modified partial differential operator $\mathcal{L} := \frac{1}{2}\sigma^2\frac{\partial^2}{\partial z^2} - \left(r + \frac{1}{2}\sigma^2\right)\frac{\partial}{\partial z}$ and payoff function $\tilde{\psi}(z) := (e^z - 1)$. (To the authors' knowledge, this PDE was first derived in Babbs (1992).) Discretisation of (LP) is similar, except that we have spatial boundary condition $\lim_{z\to\infty} V^*(z,.) = \tilde{\psi}$ and the Neumann condition $\frac{\partial V^*}{\partial z}(0,.) \equiv 0$, which we approximate by a second order accurate estimate $u_{-1}^{(\cdot)} = u_1^{(\cdot)}$, after introducing an artificial node at $i = -1$. We give results for numerical solution of this problem along with those for the vanilla put in the next section.

5 Numerical Results

In this section, we give some results from empirical tests of the simplex and PSOR algorithms for solution of (OLP) or (DOCP) respectively. Accuracy of finite difference schemes is well-established (see Carr & Faguet (1994) in general or Dempster & Hutton (1997*b*) for the LP method investigated here), so here we only investigate timings and plot solution surfaces.

5.1 Computational Details

All results in the sequel were computed in double precision on an IBM RS/6000 590 serial computer with 128 megabytes (MB) of RAM, running under AIX 3.2.5. The simplex algorithm used was the routine `EKKSSLV` of IBM's Optimisation Subroutine Library (OSL) Release 2, described in IBM Corporation (1992), which consists of compiled FORTRAN subroutines. All routines were written in FORTRAN and compiled and optimised by IBM `xlf`. The code for PSOR applied to the American vanilla put was kindly supplied by Jeff Dewynne of Southampton University.

For simplex solution of general implicit methods ($\theta > 0$), each time step was hot-started from the previous time step. The dual simplex was used (with solution accuracy set to 1×10^{-8}), since it was found to be roughly twice as fast as the primal method — this is as one expects from the discussion of dual simplex in §4.2.1. The projected successive over-relaxation (PSOR) algorithm is used with relaxation parameter $\omega := 1.5$, convergence tolerance $\epsilon := 1 \times 10^{-8}$ and starting vector given by the previous time step's solution vector. The optimal value of ω is not known analytically, so $\omega = 1.5$ was empirically found optimal for a range of problems.

For the explicit method a stable method must be used, for which (29) gives the restriction that $M \geq M_{\text{min}}$, where

$$M_{\text{min}} := \frac{\sigma^2 T I^2}{U - L}. \tag{36}$$

In the case of the lookback put we do not know M_{min} analytically, but it is very simple to determine the critical mesh ratio — and hence M_{min} — experimentally, since instability is very apparent in any numerical solution.

The American lookback put

We are rather limited in our ability to properly evaluate the accuracy of finite difference schemes applied to the American lookback put problem, since we have only one value against which to compare. Babbs (1992) computes, by a modified binomial method after 500 time steps, the solution at $t = 0$ as 10.17 with maturity $T = .5$, risk-free rate $r = .1$, volatility $\sigma = .2$, dividend rate $q = 0$ and current stock price 100. In Table 4 we give results at this same point computed by our LP method for the Crank-Nicolson and explicit scheme with spatial domain $z \in [0, 1]$. The critical mesh ratio for the explicit method was found by trial and error to be about 25.

5.2 Timings of Numerical Algorithms

Table 1 gives times for the vanilla put for the three main algorithms under consideration, and corresponding plots of each method's time as a function of space steps I are given in Figure 3. All times are CPU times in seconds, and are for Crank-Nicolson or explicit methods. We exclude the implicit method ($\theta = 1$), since times are virtually identical to the Crank-Nicolson method, except slightly less time is spent computing the matrix multiplication Bu_i^m since B is diagonal. Unless otherwise stated, all timed problems are solved on the truncated log-stock axis $[-1, 2]$ with maturity date $T = 1$, exercise price $K = 1$, riskless rate $r = .1$ and volatility $\sigma = .2$.

The number of time steps for the explicit was taken as the maximum of the number taken for the implicit methods (1000) and the minimum number for stability, so as to compare like with like, i.e. stable algorithms with at least 1000 time steps. The simplex solution gives near-linear solution time as a function of space steps I. PSOR is faster for smaller I, and explicit is faster still. PSOR and explicit methods exhibit rather similar behaviour as I increases, both increasing like I^2, so that for larger I, simplex is superior. This is a well known theoretical property of SOR methods.

Columns 3 and 5 of Table 1 give indicative first and last time step iteration counts for PSOR and simplex. From observation the simplex iteration count increases with the time step, and is near linear as a function of I, which in keeping with empirical evidence, according to Luenberger (1984), that the simplex method converges in a number of iterations proportional to the number of rows of the constraint matrix. We show the same information for PSOR — in this case, the iterations decrease as we step through time, declining rapidly for the first few steps due to the decreasing distance between successive time step solution vectors.

Similarly, Table 2 shows results and graphs of solution time for PSOR and simplex as the number of time steps M increases, for two cases $I = 600$ and $I = 2400$. Dependence of explicit method time on M is clear enough so we do not include it here. Again, simplex shows linear dependence on the number of time steps, so that each time step takes about the same amount of time. PSOR has the interesting property of being very flat as a function of M, particularly for $I = 2400$, for which case the solution time is virtually constant in M. This is probably because the previous time step's solution, used as the starting point for the iteration, is closer to the current time step's solution for smaller Δt. As in Table 1, we see PSOR is faster for smaller I, or equivalently, simplex is faster for smaller M.

Finally, we see in Table 3 the variation of PSOR and simplex with the financial parameters r, the risk-free rate, and σ, the volatility. Again we exclude explicit method times, since it is clear that they are constant with respect to r and are proportional to σ^2. In Table 3, an asterisk (*) represents failure to converge after several hours. Again, we see that the faster method is determined by the values of these parameters. PSOR is faster than simplex for low r and σ, slower for high r and σ. However, the most immediate feature of these results is that simplex time is virtually *constant* with respect to r and σ, whilst PSOR solution time explodes for high volatility. Note that stock volatilities greater than 80% are often observed in the financial markets.

We do not attempt to give such detailed solution timings for the American lookback put, since, as we can see from Table 4, the algorithms behave in a very similar manner on this problem. That table shows the same near-linear increase in solution time with the simplex method, and PSOR again does better for a smaller number of space steps I, but, as expected, as we increase I solution times increase as I^2. However, comparing with Table 1, note that both explicit (chosen with $M = \max(1000, M_{\min})$ again) and PSOR are slower on this problem than on the vanilla put, but simplex is *faster*.

5.3 American Option Solution Surfaces Graphically

Finally we give plots of the solution surfaces of the vanilla and lookback options solved for in this Section. Figure 6 shows the American vanilla put value function plotted with respect to the true stock price, and we may recognise in it all the theoretical features of Figure 1. Figure 7 shows the space-time domain, shaded according to whether the value function is equal to the obstacle or greater than it. In this figure, region A is exactly the stopping region $\mathcal{S}$, $B \cup C$ is the continuation region $\mathcal{C}$ truncated at $x = e^U$ and C represents where the solution is machine zero. The convex boundary between A and B is the optimal stopping boundary S^*.

Figure 8 shows the lookback put surface for a current stock price of 1 and spatial domain $z \in [0, 1]$, with the Neumann condition at $z = 0$ is clearly visible. Figure 9 shows the space-time domain with shaded continuation and stopping regions.

6 Conclusions

We conclude that the new linear programming algorithm presented is a very effective solution technique for finite difference approximations to American option free boundary problems like those considered here. It is efficient, especially for fine discretisations, and simple to implement when combined with modern commercially available simplex solvers. It is a *direct* method, and as such has the inherent advantages that solution times are predictable and robust with respect to changes in the parameters (whereas PSOR fails to converge for commonly observed levels of volatility), with the additional empirical feature of the simplex method of being near-linear in the spatial discretisation. We cannot claim on the basis of the results given in this paper that simplex solution is always superior, indeed PSOR is certainly faster for coarser space or finer time discretisations. As they stand, the implementations here are probably equally efficient. It may be that more recent simplex code would perform better (e.g. CPLEX or XPRESS), but more importantly the PSOR algorithm and code used here was optimised for this problem, whilst the simplex solver utilized was a general purpose algorithm. It is possible to write a simplex solver for tridiagonal constraint matrices, exploiting the rapid LU decomposition algorithms for such matrices, which produces dramatic speed-ups – this has been accomplished and will be reported elsewhere, see Dempster, Hutton & Richards (1998).

We have demonstrated that the LP solution works at least as well for the example of the continuously sampled similarity-transformed American lookback put. The simplex solution times are very similar to the vanilla put, even slightly better; whereas alternative methods appear to fare worse.

A further feature of the simplex method is that, once an optimal basic solution has been found, this solution may be rapidly recomputed after small changes in the variable bounds, the right hand side or the objective function. This *parametric simplex* method (for example the OSL routine `EKKSPAR` described in IBM Corporation (1992)) could be exploited for many path-dependent options, such as continuous and discretely sampled lookback and Asian options. For such options, the path-dependent variable does not appear in the PDE, and therefore the constraint matrix, but is simply a parameter of the payoff function and boundary conditions, i.e. the variable bounds and right-hand side — see Dempster & Hutton (1997*a*) and Dempster *et al.* (1998) for more details.

More effort might also be directed towards an efficient interior point solution of this problem, a method which is very popular in the optimisation community — largely because of its effectiveness in solving very large problems, and the possibility of efficient parallelisation.

Probably the most challenging extension of the linear programming method is to higher spatial dimensions, for example, to solve for the value function of American-style cross-currency interest rate derivatives, in which case we have a banded (nested tridiagonal) constraint matrix. Thought should be invested into exploiting such a structure. (See Hutton (1995) or Dempster & Hutton (1997*b*) for an application of finite difference methods to complex European-style cross-currency derivatives, driven by three stochastic variables.) However, the conclusions about the superiority of explicit schemes to standard implicit ones for a three state variable complex European option case in Hutton (1995) will apply equally to simplex solution of implicit schemes for American options in higher dimensions. One way to exploit the LP method in higher dimensions is to solve one-dimensional implicit steps as part of an ADI method. It may also be that a multi-grid method may be needed to obtain reasonable solution times, see Clarke (1998), and it would be very interesting to see how LP solution could fit into such a scheme.

Acknowledgements

Much of the work upon which this paper is based was undertaken while the authors were members of the Department of Mathematics in the University of Essex.

References

BABBS, S. H., (1992), Binomial Valuation of Lookback Options, Working Paper, Midland Montagu Capital Markets.

BAIOCCHI, C. & CAPELO, A. (1984), *Variational and Quasivariational Inequalities*, Wiley, New York.

Black, F. & Scholes, M. (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* **81**, pp. 637–659.

Borwein, J. M. & Dempster, M. A. H. (1989), The Linear Order Complementarity Problem, *Mathematics of Operations Research* **14**(3), pp. 534–558.

Carr, P. & Faguet, D. (1994), Fast Accurate Valuation of American Options, Working Paper, Johnson Graduate School of Management, Cornell University.

Clarke, N. (1998), *Numerical Solution of Financial Derivatives*, Ph.D. Thesis, Linacre College, Oxford.

Cryer, C. W. (1971), The Solution of a Quadratic Programme using Systematic Overrelaxation, *SIAM Journal of Control and Optimization* **9**, pp. 385–392.

Cryer, C. W. & Dempster, M. A. H. (1980), Equivalence of Linear Complementarity Problems and Linear Programs in Vector Lattice Hilbert Spaces, *SIAM Journal of Control and Optimization* **18**(1), pp. 76–90.

Dantzig, G. B. (1963), *Linear Programming and Extensions*, Princeton University Press.

Dempster, M. A. H. & Hutton, J. P. (1997*a*), Fast Numerical Valuation of American, Exotic and Complex Options, *Applied Mathematical Finance* **4**(1), pp. 1–20.

Dempster, M. A. H. & Hutton, J. P. (1997*b*), Numerical Valuation of Cross-Currency Swaps and Swaptions, in *Mathematics of Derivative Securities*, M. A. Dempster & S. R. Pliska, eds., Cambridge University Press.

Dempster, M. A. H., Hutton, J. P. & Richards, D. G. (1998), LP Valuation of Exotic American Options Exploiting Structure, Working Paper, Judge Institute of Management Studies, University of Cambridge.

Duffie, D. (1992), *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, New Jersey.

Hutton, J. P. (1995), *Fast Pricing of Derivative Securities*, Ph.D. Thesis, Department of Mathematics, University of Essex.

IBM Corporation (1992), *Optimization Subroutine Library - Guide and Reference Release 2*, 4th edition.

Jacka, S. (1991), Optimal Stopping and the American Put, *Mathematical Finance* **1**, pp. 1–14.

Jaillet, P., Lamberton, D. & Lapeyre, B. (1990), Variational Inequalities and the Pricing of American Options, *Acta Applicandae Mathematicae* **21**, pp. 263–289.

KARMARKAR, N. (1984), A New Polynomial Time Algorithm for Linear Programming, *Combinatorica* **4**, pp. 373–394.

KOOPMANS, T. C. (1951), *Activity Analysis of Production and Allocation*, Yale University Press.

LUENBERGER, D. G. (1984), *Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 2nd edition.

MANGASARIAN, O. (1979), Simplified Characterisations of Linear Complementarity Problems Solvable as Linear Programs, *Mathematics of Operations Research* **4**(3).

MANIA, M., (1997) Semimartingale Characterization of Generalized Derivatives, *Stochastics and Stochastics Reports* **61**, pp. 35–66.

MYNENI, R. (1992), The Pricing of the American Option, *The Annals of Applied Probability* **2**(1), pp. 1–23.

VAN MOERBECKE, P. (1976), On Optimal Stopping and Free Boundary Problems, *Arch. Rational Mech. Anal.* **60**, pp. 101–148.

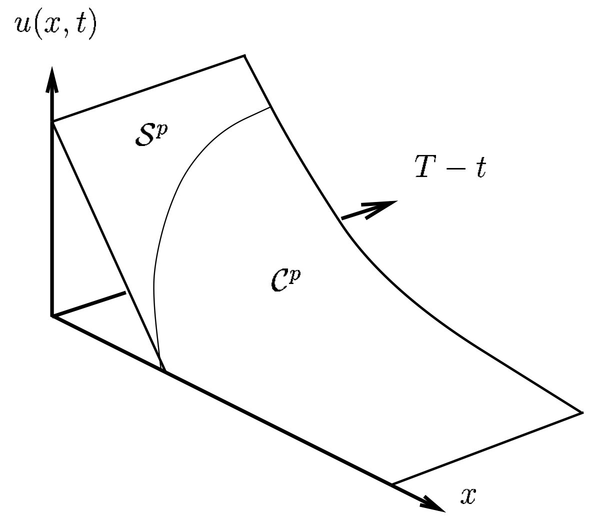

Figure 1: Sketch of the American put value function

Volatility $\sigma = .2$						
	time steps $M = 1000$				Explicit	
space	PSOR		Simplex		$M = \max\{1000, M_{\min}\}$	
steps I	time (s)	iterations	time (s)	iterations	M	time (s)
75	.83	17, 13	2.04	0, 3	1000	.05
150	1.56	17, 12	3.81	0, 6	1000	.1
300	2.69	17, 11	7.53	0, 13	1200	.2
600	3.50	16, 7	15.2	0, 27	4800	.61
1200	5.87	15, 6	31.3	1, 55	19200	4.9
2400	33.3	17, 16	66.2	7, 114	76800	37.0
4800	214	62, 47	144	17, 232	307200	317.0
9600	1270	214, 134	323	36, 468	1228800	5770
Volatility $\sigma = .4$						
	time steps $M = 1000$				Explicit	
space	PSOR		Simplex		$M = \max\{1000, M_{\min}\}$	
steps I	time (s)	iterations	time (s)	iterations	M	time (s)
75	.9	18,14	2.11	0, 9	1000	.05
150	1.55	18, 13	3.98	0, 18	1000	.1
300	1.99	18, 8	7.85	0, 38	1600	.32
600	3.29	18, 6	16.4	2, 78	6400	2.46
1200	19.1	20, 20	34.5	8, 159	25600	19.9
2400	122	72, 60	76.6	21, 323	102400	149
4800	807	250, 188	178	45, 650	409600	1280
9600	5080	831, 559	430	94, 1304	1638400	10500

Table 1: Comparative solution times for PSOR, simplex and explicit finite difference algorithms for varying space steps

time	Space steps $I = 2400$	
steps M	PSOR	Simplex
10	28.4	2.86
20	29.6	3.65
40	29.1	5.61
80	29.9	8.57
160	32.0	14.4
320	31.6	24.6
640	33.0	46.0
1280	35.3	87.9
2560	38.7	171

Table 2: PSOR and Simplex times for varying time steps

Risk-free rate r	Volatility σ				
	.05	.1	.2	.4	.8
.05	3.82	9.81	32.9	127	*
.1	3.26	9.15	32.6	122	*
.2	2.13	7.04	28.4	114	*
.4	1.64	3.80	21.1	101	*
.8	1.12	2.96	11.2	71.9	*

Risk-free rate r	Volatility σ				
	.05	.1	.2	.4	.8
.05	24.7	26.6	31.0	41.7	46.1
.1	24.8	27.0	31.3	38.2	51.4
.2	24.8	25.3	25.9	32.8	44.9
.4	23.8	24.7	25.6	29.2	38.1
.8	23.4	24.3	25.6	26.8	33.1

Table 3: PSOR and Simplex times for varying riskless rate r and volatility σ (* $\Rightarrow$ failure to converge in 2000s)

space steps I	PSOR time (s)	Simplex time (s)	Explicit time (s)	Crank-Nicolson $P_{LP}(0,.5)$	Explicit $P_{exp}(0,.5)$
75	.76	1.60	.08	.101661	.101671
150	1.36	2.85	.16	.101706	.101717
300	2.11	5.52	.66	.101718	.101730
600	3.63	11.4	5.27	.101721	.101725
1200	17.0	24.4	38.2	.101722	.101723
2400	102	54.9	315	.101722	.101723
4800	632	131	2580	.101722	.101723
9600	3330	324	21100	.101722	.101723
Binomial value				.1017	.1017

Table 4: PSOR, simplex and explicit results for the American lookback put with varying space steps

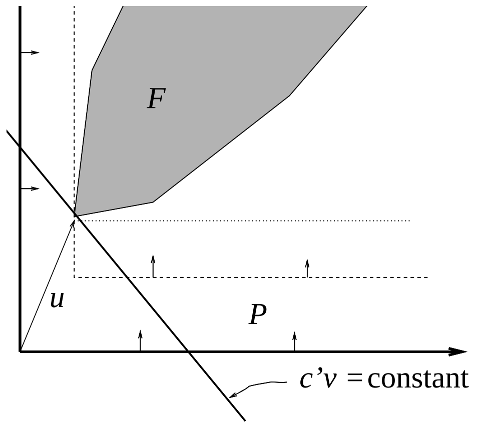

Figure 2: The least element problem as a linear programme

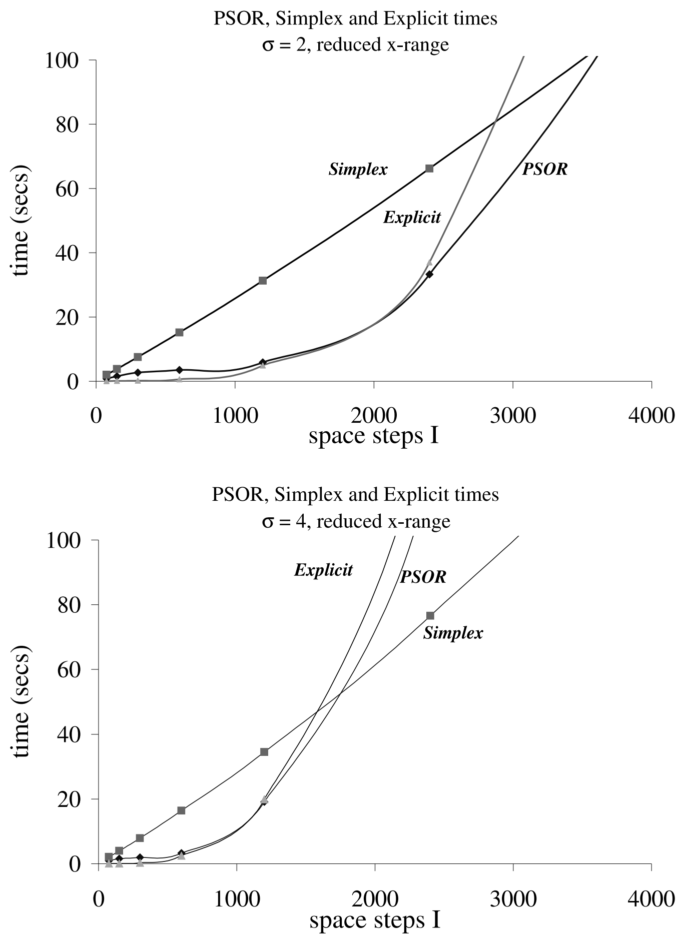

Figure 3: Comparative solution times versus number of space steps for volatilities $\sigma = .2$ and .4

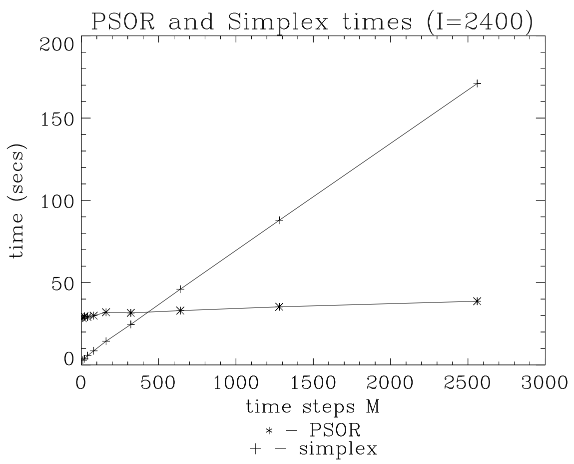

Figure 4: PSOR and Simplex times for varying time steps

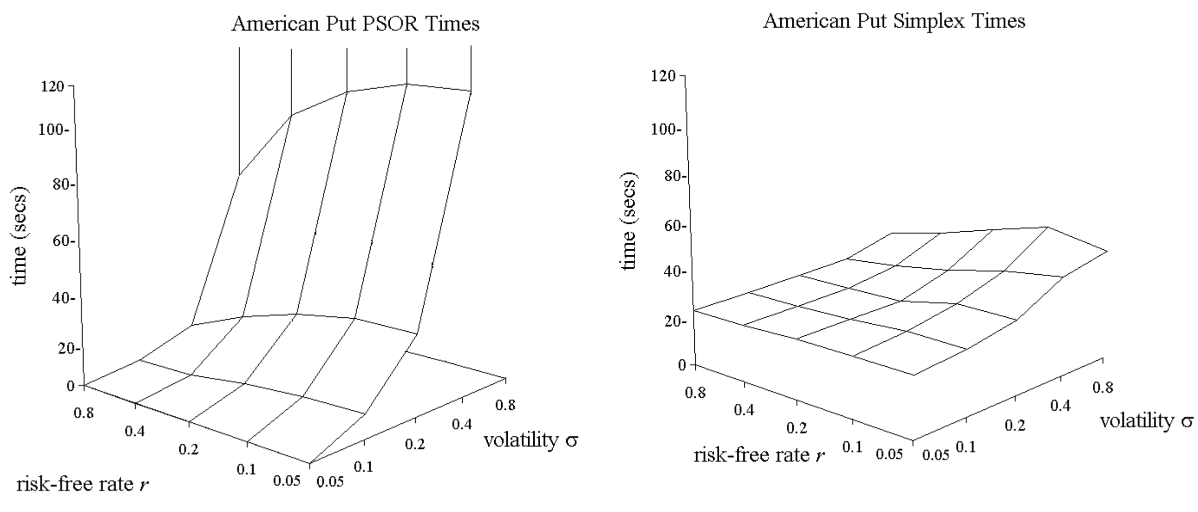

Figure 5: PSOR and simplex times for varying r and σ

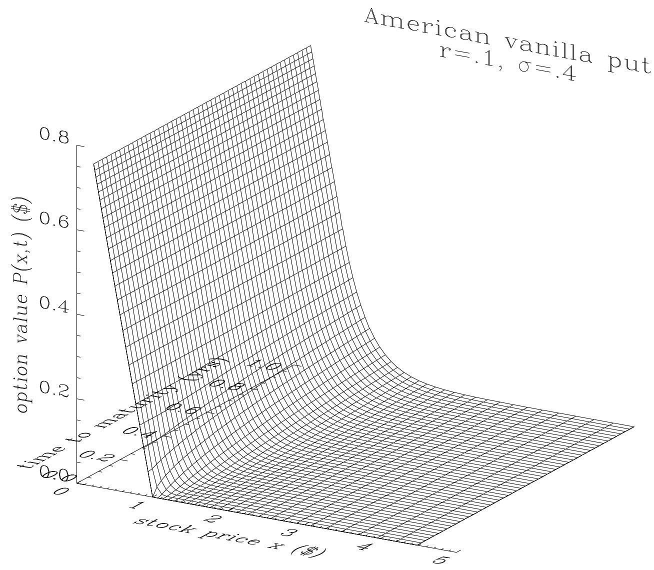

Figure 6: (LP′) solution surface with true stock price axis

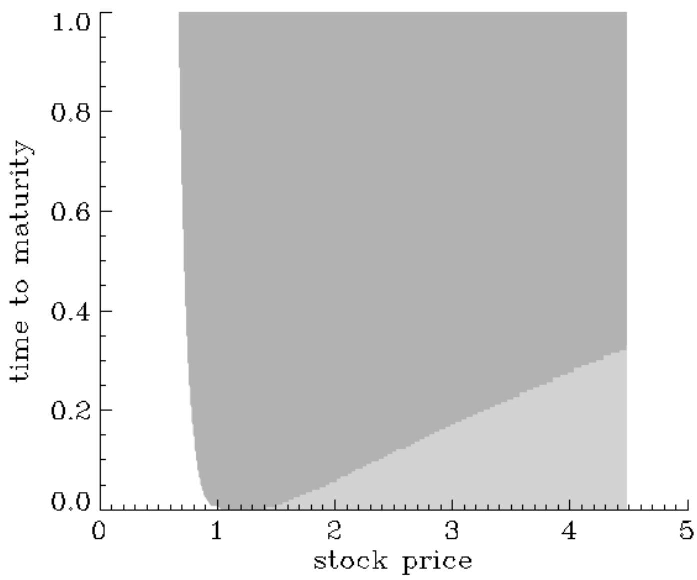

Figure 7: The computed optimal stopping boundary

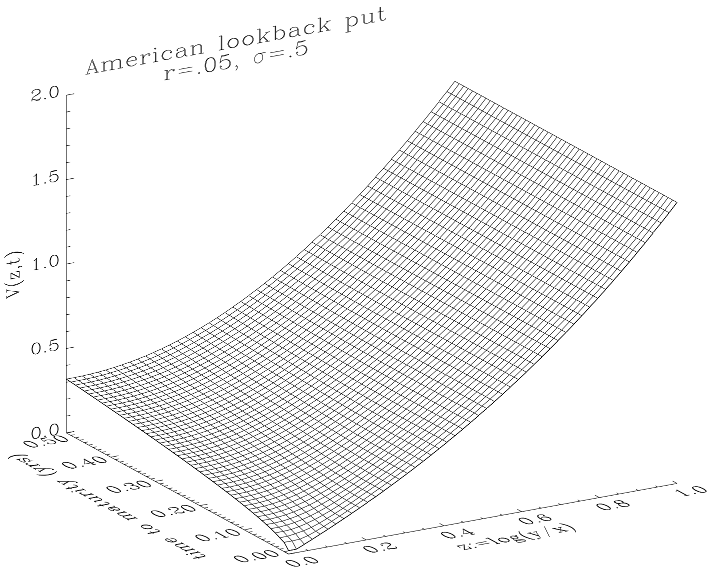

Figure 8: American lookback put value surface

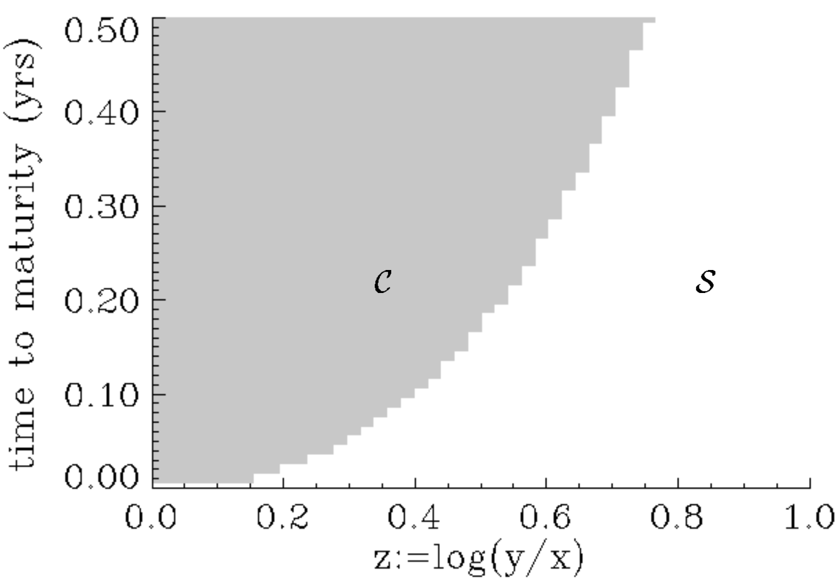

Figure 9: The optimal stopping boundary of the American lookback put