Pricing American Stock Options by Linear Programming

M.A.H. Dempster & J.P. Hutton
Department of Mathematics
University of Essex
Wivenhoe Park, Colchester, England CO4 3SQ
mahd@essex.ac.uk & hutto@essex.ac.uk

October 24, 1996

Abstract

We investigate numerical solution of finite difference approximations to American option pricing problems, using a novel direct numerical method — simplex solution of a linear programming formulation. This approach is based on a new result extending to the parabolic case 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 put and lookback put. 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 are certainly possible over the results given here.

1 Introduction

The aim of this paper is to investigate numerical solution of finite difference approximations to partial differential equation (PDE) problems arising in pricing American options using a novel linear programming approach, and to test this empirically against other methods for such problems. We argue 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 §2 we outline the classical Black-Scholes model for pricing standard European and American options, paying particular attention to the related PDEs and boundary conditions. We present, for the case of the American vanilla put option, the well-known

equivalent formulations of an American option problem as a free boundary problem, a linear order complementarity problem (OCP) and a variational inequality (VI). Results from the literature on uniqueness of the variational inequality solution give us uniqueness of the order complementarity problem solution. 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 formulation 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 problems in §2. Again, results from the literature on convergence of the solution to the discretised variational inequality to the continuous American put value function give convergence for the equivalent discretised linear programme. Standard numerical algorithms for the solution of these problems are projected successive over-relaxation (PSOR) for the complementarity problem and simplex and interior point methods for the linear programme. We investigate these in §5.

In §5 we test the new linear programming approach empirically for the American vanilla put and lookback put — against the PSOR algorithm for the complementarity problem — using modern simplex and interior point algorithms from IBM's Optimization Subroutine Library (OSL) running on an IBM RS/6000 590. We reproduce known solution values from the literature and detail the behaviour of solution time and iteration counts with varying discretisation and market parameters for both algorithms, giving plots of the solution surfaces. Overall, we find similar performance for PSOR and simplex algorithms on average, with the interior point algorithm much worse than either. We also notice that the solution time of the simplex method is highly robust to changes in the risk-free rate and volatility parameters, as one expects from a direct method. We conclude that the linear programming approach has much potential, and should, with its speed and flexibility, prove superior to other methods for a wide range of option types — particularly for certain exotic options, where one may exploit the *parametric* simplex method. Furthermore, it is worth pointing out that any PDE-based numerical method for option valuation yields an approximation to the option value *surface*, and hence provides immediately numerical approximations of the various partial derivatives (the 'Greeks') fundamental to practical hedging schemes. (See Carr [5] for details.)

2 The Black-Scholes Model

To begin the classical Black-Scholes analysis we must make certain assumptions about the nature of our economy and the agents that operate within it. Uncertainty in our security market is modelled by a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q)$, supporting a Wiener process $\mathbf{W}$ with finite time horizon T. World events affect prices only in so far as they drive the Wiener process and so the simplest construction of Ω is as the set of all paths of the Wiener process over $t \in [0, T]$, the filtration as that naturally generated by the Wiener process over time and Q as the Wiener measure over paths. (To be technically correct we must apply the *usual conditions*, namely, that the filtration is increasing, right-continuous

and augmented with the null sets of Q.)

Furthermore, we make the following assumptions about the structure of our economy and the agents that operate within it: continuous trading; perfectly divisible assets; no transaction costs; no restriction on short sales; agents have symmetric information and are non-satiated; no arbitrage opportunities.

We allow only two financial instruments in our market. Firstly, we have a savings account, whose value is continuously compounded at the constant risk-free interest rate $r \geq 0$. Secondly, we model the price process $\mathbf{S}$ followed by the dividend-free stock by a *geometric Brownian motion* (GBM). With the assumptions given above and some technical restrictions, it can be shown that the absence of arbitrage is necessary and sufficient for the existence of a probability measure $\tilde{Q}$, equivalent to the objective measure Q, such that the discounted stock price process $e^{-rt}\mathbf{S}(t)$ is a martingale. (The equivalent martingale measure approach is due to Harrison and Kreps [14], to which the reader is referred for more details.) This implies that the stock price process has drift r under this measure; indeed, by the equating the result of applying Ito's lemma to the discounted stock price and the calculation of its total derivative we obtain the SDE

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

where $S(0) > 0$, $\sigma > 0$ is the constant *volatility* of the stock, a measure of the price variability and $\tilde{\mathbf{W}}$ is a Wiener process under the *equivalent martingale measure* $\tilde{Q}$ (see, for example, Karatzas and Shreve [19]).

2.1 Derivative securities

We define a *derivative security* as a risky security whose value is entirely determined by other risky assets. We consider here standard (*vanilla*) options, which form the basis for most other more exotic varieties. A *call* or *put option* confers the right to buy or sell respectively one share of stock for *strike price* K, only at a maturity date T for a *European* option, or at any stopping time in $[0,T]$ for the *American* equivalent. Under the model here, closed form formulae are known for all these options apart from the American put — in general an American option will have no known closed form formula. We study the vanilla put in more detail, in terms of its *payoff function* $\psi\colon \mathbb{R}^+ \to \mathbb{R}$ given by definition as $\psi(\mathbf{S}(\boldsymbol{\tau})) = (K - \mathbf{S}(\boldsymbol{\tau}))^+$, received by the holder on exercise at any *stopping time* $\boldsymbol{\tau} \in \mathcal{T}_{0,T}$, where $\mathcal{T}_{0,T}$ is the set of stopping times in $[0,T]$. 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 $t \in [0,T]$.

2.2 The equivalent martingale measure and Black-Scholes PDE for derivative securities

Under the equivalent martingale measure the discounted price process of any European-style derivative security is a martingale. If $u(x,t)$ is the value of any European-style

derivative when the price of the underlying stock is x at time $t \in [0,T]$, then this martingale property immediately gives us u as the conditional expectation of the discounted payoff. However, we may derive another characterisation of u: its discounted price process must have zero drift to be a martingale, so that applying Ito's lemma to give the process followed by $e^{-rt}u(\mathbf{S}(t),t)$ and equating the drift to zero gives the PDE derived by Black and Scholes [3], *viz.*

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

for $(x,t) \in \mathbb{R}^+ \times [0,T)$, where the differential operator $\mathcal{L}_{BS} := \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + rx\frac{\partial}{\partial x} - r$. We supply the terminal condition $u(.,T) = \psi$ to determine the value function u as the solution to this PDE.

American options 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

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

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

Now for a PDE characterisation of u, only the discounted *stopped* price process of such a derivative is a martingale, and only up to the stopping time, so that u satisfies the same PDE (2), but 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* $\mathcal{S}$ is given by

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

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

On the whole domain $\mathbb{R}^+ \times [0,T)$, 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 this boundary, and for the American put this is usually taken to be the *smooth fit* condition $\frac{\partial u}{\partial x} = -1$ on the optimal stopping boundary. For further discussion of these matters see Van Moerbecke [25], Jacka [17] and Myneni [24].

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.

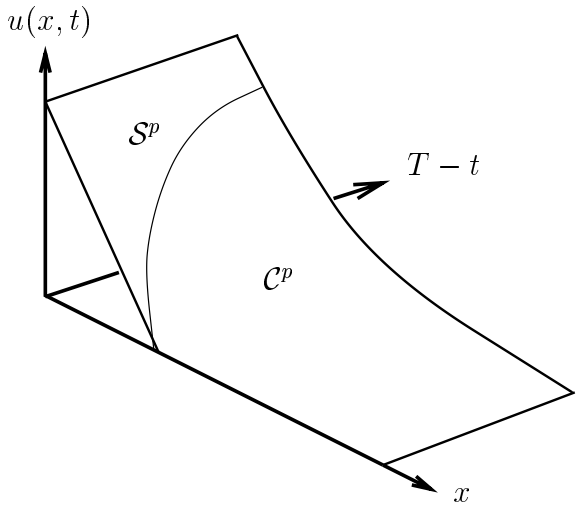

Figure 1: Sketch of the American put value function

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, but are not explicit enough 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 coefficients 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 previous section 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

As seen in §2.2, 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 $u(\xi, t) = \tilde{\psi}(\xi)$, and 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* [4].

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

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

■

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 and Dempster [4] and Cryer and Dempster [8] 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 $u \in L^p\left(\mathbb{R}_2, e^{-\mu|\xi|}\, d\xi\right)$ whose *weak derivatives* of order not exceeding $m \in \mathbb{N}$ exist and are also in $L^p\left(\mathbb{R}_2, e^{-\mu|\xi|}\, d\xi\right)$, for $p \in [0,\infty]$ and $\mu \in (0,\infty)$. (Here $|.|$ denotes the L^1 norm in $\mathbb{R}^2$ 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 $\langle .,. \rangle : H^1 \times H^1 \to \mathbb{R}$ 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 $\langle .,. \rangle$ 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 $H^1 \hookrightarrow H^0 \hookrightarrow H^{-1}$ [2] 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 $a(.,.) : H^1 \times H^1 \to \mathbb{R}$ given by

$$a(u,v) \;:=\; \int_{\mathbb{R}_2} \frac{\sigma^2}{2} u_\xi v_\xi e^{-\mu(|\xi|+|t|)}\, d\xi\, dt - \int_{\mathbb{R}_2} \left((r - \sigma^2/2) + \mu \frac{\sigma^2}{2} \frac{\xi}{|\xi|} \right) u_\xi v e^{-\mu(|\xi|+|t|)}\, d\xi\, dt$$

$$+\int_{\mathbb{R}_2} ruve^{-\mu(|\xi|+|t|)}\,d\xi\,dt, \quad \forall u,v \in H^1. \tag{7}$$

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 [4, 8]. See Baiocchi and Capelo [2] and Borwein and Dempster [4], p.553–554, for more details on these ideas, which have been adapted here to match the more general setting of Jaillet *et al* [18]. In particular, we shall assume all functions in H^1 ($\cong H^{-1}$) considered to be defined as $u(.,|t|)$ on $\mathbb{R}\times(-\infty,0)$ and as $u(.,T)$ on $\mathbb{R}\times[T,\infty)$ (see [2], p.89 *et seq.*).

The following lemma relates the bilinear form $a(.,.)$ 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*

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

Proof: See [8], p.80–81, and [17] 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*

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

is equivalent to the order complementarity problem (OCP).

Proof: This is again a well known result, due to Borwein and Dempster [4] . We can rewrite the third line of (VI) using Lemma 1 as

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

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 third constraint in (OCP). This in turn implies, since $u - \tilde{\psi} \geq 0$, that $\left\langle -\mathcal{L}u - \frac{\partial u}{\partial t}, u-\tilde{\psi}\right\rangle \geq 0$. Now if we choose $v = \tilde{\psi}$, (9) becomes $\left\langle -\mathcal{L}u - \frac{\partial u}{\partial t}, u-\tilde{\psi}\right\rangle \leq 0$, which two inequalities together give the complementarity condition $\left\langle -\mathcal{L}u - \frac{\partial u}{\partial t}, u-\tilde{\psi}\right\rangle = 0$, which is equivalent to the third constraint of (OCP) (Borwein and Dempster [4], p.549).

Let u solve (OCP): Then

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

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 $a(.,.)$ 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* [18], 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 and Capelo [2], 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, a *(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

We have established in the previous section 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 gives equivalence between (OCP), (LE) and (LP), and is an extension of a result of Cryer and Dempster [8] 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:*

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

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

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

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

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

Proof: We first prove the equivalence between (OCP) and (LE), after the trivial domain extensions of the problem functions to set them in H^1 given above. 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

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

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

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

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

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 $\hat{q}(.,\lambda) :\equiv -u(.,0)$ to agree with the notation of Borwein and Dempster [4], (OCP) is equivalent to the transformed order complementarity problem ($\widehat{\text{OCP}}$), also posed in H^1, given by

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

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, so is $\mathcal{T}+|\lambda|+\mu$, and hence $\hat{u}$ is the unique solution to ($\widehat{\text{OCP}}$). This assertion is apparent from: $\langle u,(\mathcal{T}+|\lambda|+\mu)u\rangle = \langle u,\mathcal{T}u\rangle + (|\lambda|+\mu)\langle u,u\rangle \geq (\alpha+|\lambda|+\mu)\|u\|^2$; and $\langle u,(\mathcal{T}+|\lambda|+\mu)v\rangle = \langle u,\mathcal{T}v\rangle + (|\lambda|+\mu)\langle u,v\rangle \leq 0$ for $\langle u,v\rangle = 0$, for all $\lambda \in \mathbb{R}$. We can now apply the order complementarity-least element equivalence result of Borwein and Dempster [4] for coercive type Z *elliptic* operators, so that $\hat{u}$ is also the solution (necessarily unique) to the least element problem ($\widehat{\text{LE}}$) defined by $\text{LE}(\hat{F})$, where $\hat{F}$ is defined by

$$\hat{F} := \{\hat{u} :\ \hat{u} \geq \hat{\psi},\ (\mathcal{T}+|\lambda|+\mu)\hat{u}+\hat{q} \geq 0\}. \qquad (15)$$

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

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

solves both (LE), given by $\text{LE}(F)$, 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. ■

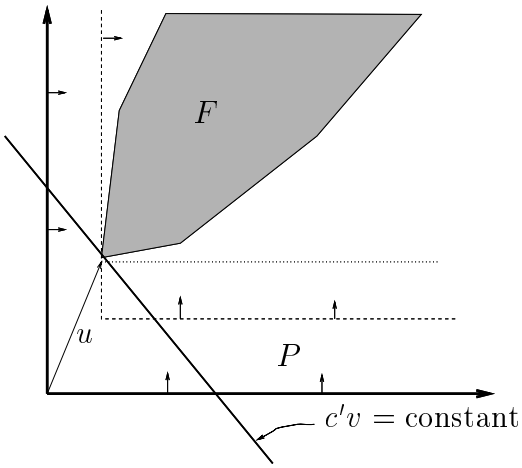

Figure 2: The least element problem as a linear programme

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 $u(.,0)$ 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 $\langle c, u\rangle$, where $c > 0$, over the set $u \in F$; 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* [18]). 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; 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 will be 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. We use the novel formulation of the value function in §2 as the solution to an abstract linear programme, which, when we discretise space and time by standard finite differences, becomes an ordinary linear programme which we may solve by well-known and highly developed algorithms.

4.1 Localisation of the value function

As a first simplifying 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 $L < \log K < U$. We then must specify explicitly the behaviour of the value function on the spatial boundaries — set $u(L,.) = \tilde{\psi}(L)$ and $u(U,.) = \tilde{\psi}(U)$, then we have a localised version (LOCP) on $[L,U] \times [0,T]$ of the order complementarity problem (OCP) given by

$$\text{(LOCP)} \qquad \begin{cases} u(L,.) = \tilde{\psi}(L), \;\; u(U,.) = \tilde{\psi}(U) \\ u(.,T) = \tilde{\psi} \\ u \geq \tilde{\psi} \\ -\mathcal{L}u - \frac{\partial u}{\partial t} \geq 0 \\ (-\mathcal{L}u - \frac{\partial u}{\partial t}) \wedge (u - \tilde{\psi}) = 0 \qquad \text{a.e. on } [L,U] \times [0,T]. \end{cases}$$

Furthermore, defining the localised inner product by $\langle u, v \rangle_\ell$ as in (6) but integrated over $[L,U]$ in the first variable, we have a localised version (LLP) of the linear programme (LP), identical in form except that the constraint set F is now given by

$$F := \left\{ v : \; v(L,.) = \tilde{\psi}(L), \;\; v(U,.) = \tilde{\psi}(U), \;\; v(.,T) = \tilde{\psi}, \;\; v \geq \tilde{\psi}, \;\; -\mathcal{L}v - \frac{\partial v}{\partial t} \geq 0 \right\}.$$

There still exists a unique solution to (LLP), 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* [18] for the equivalent localised variational inequality — naturally the equivalent problems (LOCP) and (LLP) inherit this same convergence property.

4.2 Discretisation of the value function

We discretise (LLP) 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

$$u_i^m := u\left(L + i\Delta\xi, T - m\Delta t\right), \tag{16}$$

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

$$\begin{aligned}
\frac{\partial u}{\partial \xi} &\approx \theta\frac{u_{i+1}^m - u_{i-1}^m}{2\Delta\xi} + (1-\theta)\frac{u_{i+1}^{m-1} - u_{i-1}^{m-1}}{2\Delta\xi} \\
\frac{\partial^2 u}{\partial \xi^2} &\approx \theta\frac{u_{i+1}^m - 2u_i^m + u_{i-1}^m}{(\Delta\xi)^2} + (1-\theta)\frac{u_{i+1}^{m-1} - 2u_i^{m-1} + u_{i-1}^{m-1}}{(\Delta\xi)^2} \\
\frac{\partial u}{\partial t} &\approx \frac{u_i^{m-1} - u_i^m}{\Delta t}
\end{aligned} \tag{17}$$

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.

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

$$\begin{cases}
u_i^m \geq \tilde{\psi}_i,\ u_i^0 = \tilde{\psi}_i,\ u_I^m = 0,\ u_0^m = \tilde{\psi}_0 \\
au_{i-1}^m + bu_i^m + cu_{i+1}^m + du_{i-1}^{m-1} + eu_i^{m-1} + fu_{i+1}^{m-1} \geq 0 \\
\left(au_{i-1}^m + bu_i^m + cu_{i+1}^m + du_{i-1}^{m-1} + eu_i^{m-1} + fu_{i+1}^{m-1}\right) \wedge (u_i^m - \tilde{\psi}_i) = 0 \\
\qquad\qquad i \in \mathcal{I} \setminus \{0,I\},\ m \in \mathcal{M} \setminus \{0\},
\end{cases} \tag{18}$$

where

$$\begin{aligned}
a &:= -\theta\left[\tfrac{\sigma^2\Delta t}{2\Delta\xi^2} - \tfrac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] & b &:= 1 + r\Delta t + \theta\tfrac{\sigma^2\Delta t}{\Delta\xi^2} \\
c &:= -\theta\left[\tfrac{\sigma^2\Delta t}{2\Delta\xi^2} + \tfrac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] & d &:= -(1-\theta)\left[\tfrac{\sigma^2\Delta t}{2\Delta\xi^2} - \tfrac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right] \\
e &:= (1-\theta)\tfrac{\sigma^2\Delta t}{\Delta\xi^2} - 1 & f &:= -(1-\theta)\left[\tfrac{\sigma^2\Delta t}{2\Delta\xi^2} + \tfrac{(r-\sigma^2/2)\Delta t}{2\Delta\xi}\right].
\end{aligned} \tag{19}$$

We discuss well-posedness and convergence in due course — first of all we express the complementarity condition of (18) in matrix form by collapsing the space and time indices into vectors. Put

$$u^m := \begin{pmatrix} u_1^m \\ \vdots \\ u_{I-1}^m \end{pmatrix} \quad \tilde{\psi} := \begin{pmatrix} \tilde{\psi}_1 \\ \vdots \\ \tilde{\psi}_{I-1} \end{pmatrix} \quad \phi := \begin{pmatrix} -(a+d)\tilde{\psi}_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{20}$$

Then, substituting $u_0^m = \tilde{\psi}_0$ and $u_I^m = 0$ into (18), the complementarity condition becomes

$$(Bu^{m-1} + Au^m - \phi) \wedge (u^m - \tilde{\psi}) = 0 \quad m \in M \setminus \{0\}, \tag{21}$$

where, defining notation for a tridiagonal matrix as

$$\mathrm{Td}_k(a_k, b_k, c_k) := \begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{K-2} & b_{K-2} & c_{K-2} \\ & & & a_{K-1} & b_{K-1} \end{pmatrix}, \tag{22}$$

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

$$A := \mathrm{Td}_i(a, b, c),\ B := \mathrm{Td}_i(d, e, f). \tag{23}$$

Now we can collapse the time index m by putting

$$u := \begin{pmatrix} u^1 \\ \vdots \\ u^M \end{pmatrix} \quad \tilde{\Psi} := \begin{pmatrix} \tilde{\psi} \\ \vdots \\ \tilde{\psi} \end{pmatrix} \quad \Phi := \begin{pmatrix} \phi - B\tilde{\psi} \\ \phi \\ \vdots \\ \phi \end{pmatrix} \tag{24}$$

and may in turn express (21) as

$$(Cu - \Phi) \wedge (u - \tilde{\Psi}) = 0, \tag{25}$$

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

$$C := \begin{pmatrix} A & & & \\ B & A & & \\ & \ddots & \ddots & \\ & & B & A \end{pmatrix}. \tag{26}$$

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 (20) and (24), 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 $u_I = 0$, $u_0 = \tilde{\psi}_0$ 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 (21), 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* [4]. Clearly, A is type Z if and only if $a \leq 0$ and $c \leq 0$, which is the case if and only if

$$|r - \sigma^2/2| \leq \sigma^2/\Delta\xi. \tag{27}$$

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 (27) holds, it is a simple matter to show that A is then coercive — see Jaillet *et al* [18], or Hutton [15] for equivalent results applied to the discretised operator expressed in terms of an operator acting on step functions — so that the original problem is well approximated .

Assuming (27) 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 $u_I^{(.)} = 0$, $u_0^{(.)} = \tilde{\psi}_0$ and $u^0 = \tilde{\psi}$.

(OLP) is an *ordinary* linear programme which is easily solved numerically. Jaillet *et al* [18] show that as $M, I \to \infty$, 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).

4.3 Solution of the discrete problem

In the discrete problem given in §4.2, 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

$$\begin{aligned} &u^m \geq \tilde{\psi} \\ &Bu^{m-1} + Au^m - \phi \geq 0 \\ &(Bu^{m-1} + Au^m - \phi) \wedge (u^m - \tilde{\psi}) = 0 \qquad m = 1, \ldots, M, \end{aligned}$$

with the boundary conditions $u_I = 0$, $u_0 = \tilde{\psi}_0$ and $u^0 = \tilde{\psi}$. This decomposition just amounts to writing out rows of the global complementarity problem (DOCP) in blocks of $I - 1$ and noting that the resulting complementarity problem in each block 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 the decomposition (28) has an equivalent linear programme, so that we get the following decomposition of (OLP):

$$\min\ c'u^m$$

$$\begin{aligned} &\text{s.t. } u^m \geq \psi \\ &Au^m \geq \phi - Bu^{m-1} \qquad m = 1, \ldots, M. \end{aligned} \tag{28}$$

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 [23]. 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 [7], 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* [18] for further details on pivoting methods in the current setting.)

4.3.1 Solution of (OLP)

There are again two main algorithms for solving linear programmes such as (OLP), namely the (direct) *simplex* method, due to Dantzig [9], and the (iterative) *interior point* method, first applied to linear programmes by Koopmans in [21] and recently reintroduced by Karmarkar [20]. Our preliminary results from the interior point method were poor — see Hutton [15] for more details — so we concentrate here on the simplex method and outline the salient features that we propose to exploit in §5.

4.3.2 Simplex method

The dual of the mth sub-problem (28) of (OLP) is, for any $c > 0$,

$$\begin{aligned} &\max\ (\tilde{\psi}' | (\phi - Bu^{m-1})')y^m \\ &\text{s.t. } 0 \leq (I|A')y^m \leq c. \end{aligned} \tag{29}$$

In the primal sub-problem, given by (28), only the right-hand side changes from the preceding sub-problem. When translated into the dual this means that only the objective function of the dual sub-problem changes, as we see from (29), 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*.

4.3.3 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 (23) reduces to the $(I-1)$-square diagonal matrix $\mathrm{diag}(1 + r\Delta t)$, so that in fact the general mth sub-problem of (OLP) given by (28) reduces to

$$\begin{aligned} &\min\ c' u^m \\ &\text{s.t. } u^m \geq \tilde{\psi} \\ &(1 + r\Delta t) u^m \geq \phi - B u^{m-1}. \end{aligned} \tag{30}$$

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

$$u^m = \tilde{\psi} \vee \left(\frac{1}{1 + r\Delta t} \left(\phi - B u^{m-1} \right) \right). \tag{31}$$

This is clearly a very rapid calculation for each iteration, the only significant calculation being a single matrix multiplication. However, this method has a problem of stability, needing a large number of iterations to become stable. We discuss this further below.

The dual of the explicit method mth sub-problem, given in general in standard form by (29), is also very simple, and is given by, for any $c > 0$,

$$\begin{aligned} &\max\ (\tilde{\psi}' | (\phi - B u^{m-1})') y^m \\ &\text{s.t. } 0 \leq y^m \leq c. \end{aligned} \tag{32}$$

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

$$y^m = \mathrm{sgn}\left(\tilde{\psi}' | (\phi - B u^{m-1})' \right) c, \tag{33}$$

where $\mathrm{sgn}(x)$ 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, requiring only a matrix multiplication. The calculation of 'sgn' is perhaps quicker than computing a 'max', since a variable's sign is stored as a binary number in a computer, but this is not significant and overall the two methods can be made virtually identical. We will use the more standard primal method in our empirical tests in §5.

Accuracy and stability The finite difference approximations (17) to the first and second order spatial derivatives $\frac{\partial u}{\partial \xi}$ and $\frac{\partial^2 u}{\partial \xi^2}$ are both $O((\Delta\xi)^2)$, whilst the approximation to the time derivative $\frac{\partial u}{\partial t}$ is accurate to $O(\Delta t)$ — approximations accurate to a higher order would produce an unstable algorithm. Thus, in general, the solution of the discrete approximation will be accurate to $O(\Delta\xi^2 + \Delta t)$ — however, in the case of the Crank-Nicolson method ($\theta = 1/2$), accuracy is second order in time, i.e. $O(\Delta\xi^2 + \Delta t^2)$, due to fortuitous cancellation of remainders in the Taylor series.

However, these finite difference approximation choices do not guarantee stability. The general method is unconditionally stable for all $\theta \in [\frac{1}{2}, 1]$, but for $\theta \in [0, \frac{1}{2})$ — which includes the explicit method — we have stability of the scheme if and only if

$$0 \leq \rho \leq \frac{1}{\sigma^2(1 - 2\theta)}, \tag{34}$$

where the *mesh ratio* ρ is defined by $\rho := \frac{\Delta t}{(\Delta\xi)^2}$. The implications of this result for the explicit method ($\theta = 0$) are profound, because to satisfy (34) we must take a number of time steps of the order of the *square* of the number of space steps. This requirement can become computationally unmanageable, since we need to take a large number of space steps to get high accuracy. However, it is worth noting that, as a result, a stable explicit scheme has second order time accuracy.

Other constant values of θ in $[0, 1]$ are equally valid, but have no real advantages over the three just described. It may be advantageous to allow θ to be time and/or state dependent, so as to produce an *alternating direction implicit* (ADI) method suitable for higher space dimensions, but we discuss this idea further only briefly in the conclusions of §6. The above discussion of accuracy and stability is well-known — see any standard PDE text.

4.4 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 to 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 $\tilde{\psi}(z) := (1 - e^z)$. (To the authors' knowledge, the PDE was first derived in Babbs [1].) Discretisation of (LP) is similar, except that we have spatial boundary conditions $\frac{\partial V^*}{\partial z}(0,.) \equiv 0$ and $\lim_{z\to\infty} V^*(z,.) = \tilde{\psi}$, and we approximate the Neumann condition $\frac{\partial V^*}{\partial z}(0,.) \equiv 0$ by the simple, albeit crude, first order accurate estimate $u_0 = u_1$. 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 results from empirical tests of the simplex and PSOR algorithms for solution of (OLP) or (DOCP) respectively. We demonstrate the accuracy of the solution of (OLP) by simplex, give timings for the different numerical algorithms and then give some plots of the 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 [16], 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 ($\theta > 0$), the first block of (OLP) was solved by a call to the OSL basis crash routine `EKKCRSH` at level 4, followed by a call to the dual simplex method of `EKKSSLV`. Level 4 crash tries to produce a triangular starting basis, maintaining dual feasibility, if possible, by not pivoting in variables that are in the objective function, and ensuring that the sum of infeasibilities never increases. Results were virtually identical to level 3 crash, which does not try to maintain dual feasibility. (See [16] for more details.) Since the constraint matrix A defined by (23) is the same in each block, successive blocks were 'hot started' from the previous block, i.e. the simplex solver started from the complete basis of the previously solved block, which should not be too far from the current sub-problem's optimal basis, since Δt is small. In practice, using the crash routine `EKKCRSH` to generate the starting basis at each time step was not much slower — perhaps 25%. No presolving routine was used, since the OSL presolver `EKKPRSSL` writes presolve information to a file, which slows down the procedure considerably. Initially, `EKKSSLV` was allowed to choose heuristically between primal and dual algorithms and always chose the dual algorithm. This proved to be the right choice, since *setting* primal simplex produced about half the speed of the dual method, and so for all experiments `EKKSSLV` is set to dual simplex, with solution accuracy set to 1×10^{-8}. This is as one expects from the discussion

of dual simplex in §4.3.1. 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.

For the explicit method for the vanilla put, according to (34), the number of time steps M must be chosen for stability so that $\rho \leq 1/\sigma^2$. This gives the restriction that $M \geq M_{\min}$, where

$$M_{\min} := \frac{\sigma^2 T I^2}{U - L}, \tag{35}$$

and for our explicit method experiments, we often choose $M = M_{\min}$. For 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 matrix multiplications Bu^{m-1} in the mth (OLP) sub-problem (28) and those in each PSOR iteration were computed in a sparse manner, skipping known zero elements, and so take only $O(I)$ operations, since the matrices involved are tridiagonal at most.

5.2 Accuracy of Finite Differences

The accuracy of finite difference schemes applied to the American put is already well-established — see, for example, Geske and Shastri [13] or more recently Carr and Faguet [6]. For the purposes of validation of the LP method, we present results for the different finite difference schemes solved by LP, specifically the simplex method, compared to option values found in the literature: the series solution of Geske and Johnson [12] for the vanilla put and the reflected binomial method of Babbs [1] for the lookback put.

5.2.1 The American vanilla put

Table 1 gives vanilla put option values for all three schemes, with strike price $K = 1$, maturity date $T = 1$ and varying risk-free rate and volatility. The current stock price is $x = 1$, i.e. the option is 'at the money', and calendar time $t = 0$, so that the value quoted is $P(1, 0)$. The discretisation in log price space is given by $L = -.58$, $U = 2.02$ and $I = 130$, chosen so that the current stock price coincides with a grid point, eliminating the need for interpolation. For implicit methods, we take $M = 70$, and for the explicit method, M was chosen as the minimum $M_{\min}$ according to (35). The third column, labelled $P_{\text{an}}(1, 0)$, gives values of the 'analytic' solution of Geske and Johnson [12] quoted in their paper, computed as the first three terms of their series expansion — it is hard to know whether the values quoted really are accurate, given that only a small number of terms are taken.

We see from this table that all schemes give at least 3 significant figures of accuracy with this grid size. From the root-sum-of-squared errors $(\sum e_i^2)^{\frac{1}{2}}$ quoted in the last row, with e_i defined as $P_{LP}(1, 0) - P_{\text{an}}(1, 0)$, and by inspection, Crank-Nicolson and the *stable* explicit method are more accurate than implicit, as is to be expected from their second order time accuracy. The Crank-Nicolson solution appears always to be above the implicit

one — this is because the solution surface is convex with respect to time, and the Crank-Nicolson solution better captures the curvature because of its higher order time accuracy, and so lies above the implicit solution. The explicit scheme produces very similar accuracy to Crank-Nicolson, which is to be expected, since with the number of time steps chosen according to (35), the explicit scheme also has second-order time accuracy.

5.2.2 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 [1] 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 5 we give results at this same point computed by our LP method for the three finite difference schemes, 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. For all three methods, we see that any reasonable agreement with the known value only happens for large numbers of space steps I. In general, accuracy is a whole significant figure less than for the vanilla put for a similar mesh size. This is because we are evaluating the function exactly on the boundary $z = 0$, where we have only a crude first order approximation to the Neumann condition $\frac{\partial V^*}{\partial z}(0, t) = 0$. However, this inaccuracy is an artifact of the discretisation and not the solution algorithm, and could be easily eliminated by more sophisticated approximation to the Neumann boundary condition, e.g. a small space step at the boundary, or a second order approximation.

5.3 Timings of Numerical Algorithms

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 $\sigma = .2$.

Table 2 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. 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 again. 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 2 give iteration numbers for PSOR and simplex. We cannot give iteration numbers for each time step, so we give the first and last time step iteration counts. In the case of simplex, from observation the iteration count always increases in a near linear manner from the first step count to the last, and the iteration count in each block is near linear as a function of I — as we expect from the solution time results. This is also in keeping with empirical evidence, according to Luenberger [22], that the simplex method converges in between m and $3m/2$ pivots, where m is the number of rows of the constraint matrix, in this case $I-1$. This clearly is an over-estimate for our results, perhaps due to the sophisticated solver we are using, but the same linearity is observed. 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.

Similarly, Table 3 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 2, we see PSOR is faster for smaller I, or equivalently, simplex is faster for smaller M.

Finally, we see in Table 4, 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 that table, 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.

We do not attempt to give such detailed solution timings for the American lookback put, since, as we can see from Table 5, 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 2, 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*.

5.4 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 surface plotted with respect to the true stock price, computed with a discretisation of 50×50 with log-transformed spatial domain $\xi \in [-1.5, 1.5]$, so that the true stock price domain is $x \in [e^{-1.5}, e^{1.5}]$. We recognise all the theoretical features of Figure 1 — we see that the option value at maturity

is simply the payoff or obstacle function $(K - x)^+$, the value function is convex in x, non-increasing with time, always greater than or equal to the obstacle and smoothly fitted onto the obstacle at the free boundary. Figure 7 shows the space-time domain, shaded according to whether the value function is equal to the obstacle or greater than it. (Note that Figure 7 was computed with the finer grid 200×200, so that S^* is better defined in the figure.) In that figure, there are three regions shown: region A is exactly the stopping region $\mathcal{S}$; $B \cup C$ is the continuation region $\mathcal{C}$ truncated at $x = e^U$; C represents where the solution is machine zero — theoretically, it should not be zero anywhere except for $t = T$ and $x \geq K$, but this region just represents the limits of machine accuracy. 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, a discretisation of 50×50 and spatial domain $z \in [0, 1]$. We see essentially the same theoretical features as those of the vanilla put regarding the obstacle $e^z - 1$, and the Neumann condition at $z = 0$ is clearly visible. Figure 9 shows the space-time domain with shaded continuation and stopping regions labelled $\mathcal{C}$ and $\mathcal{S}$. The optimal stopping boundary in this figure is more jagged than that in Figure 9 — this is simply because this figure was computed with the same coarse grid as the surface itself.

6 Conclusions

We conclude that the new linear programming algorithm presented here 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, with the additional feature of the simplex method of being near-linear in the number of constraints — which is directly related to the space 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. However, the PSOR algorithm and code used here was optimised for this problem, whilst the simplex solver was a general purpose algorithm. There is thus enormous scope for speeding up the simplex solution, and one need only look at the historical improvements in linear programming to date to see that further speed-ups in general purpose algorithms are to be expected. Preliminary results on other problems suggest that the latest simplex routine of CPLEX gives a significant speed improvement over that of OSL.

It should also be possible to write a simplex solver for tridiagonal constraint matrices, exploiting the rapid LU decomposition algorithms for such matrices. This comment also illustrates one possible drawback of this approach — a modern simplex solver is a complex piece of code, and one cannot usually obtain source code for commercially-available solvers, so that alteration and fine-tuning may be impossible. However there are other simplex

solution algorithms in the public domain, and it might be profitable to adapt some of these to properly exploit the matrix structure — particularly with a view to solving the banded constraint matrices of higher dimensional problems.

We have demonstrated that the LP solution works at least as well for the continuously sampled similarity-transformed American lookback put — the poor accuracy of the solution at the boundary is due to the discretisation scheme, not the solution algorithm, and could easily be improved. The simplex solution times are very similar to the vanilla put, even slightly better; whereas alternative methods appear to fare much worse.

The explicit method is certainly very fast for smaller numbers of space steps. However, the stability problem eventually causes solution times to blow up for fine spatial discretisations, and is a constant headache when solving different problems, especially when the precise stability criterion is not known — for example, one small space step on the boundary could cause severe stability problems.

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 [16]) could be exploited for many path-dependent options, such as continuous and discretely sampled lookback and Asian options. For such options, the path-dependent parameter 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. This approach might effectively reduce the two dimensional nature of these problems by almost one dimension as far as solution times are concerned (see Dempster and Hutton [10] 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 interesting 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 [15] or Dempster and Hutton [10] 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 [15] will apply equally to simplex solution of implicit schemes for American options in higher dimensions. It appears, therefore, that the 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, and it would be very interesting to see how LP solution could fit into such a scheme.

References

[1] Babbs, S.H. (1992). Binomial Valuation of Lookback Options. Working Paper, Midland Global Markets, London, April 1992.

[2] Baiocchi, C. and Capelo, A. (1984). *Variational and Quasivariational Inequalities.* Wiley, New York.

[3] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *J. Political Econ.* **81** 637–659.

[4] Borwein, J.M. and Dempster, M.A.H. (1989). The Linear Order Complementarity Problem. *Maths of OR* **14** 534–558.

[5] Carr, P. (1993). Deriving Derivatives of Derivative securities. Working paper, Johnson Graduate School of Management, Cornell University, February 1993.

[6] Carr, P. and Faguet, D. (1994). Fast Accurate Valuation of American Options. Working paper, Johnson Graduate School of Management, Cornell University, June 1994.

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

[8] Cryer, C.W. and Dempster, M.A.H. (1980). Equivalence of Linear Complementarity Problems and Linear Programs in Vector Lattice Hilbert spaces. *SIAM J. Control Optim.* **18** (1) 76–90.

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

[10] Dempster, M.A.H. and Hutton, J.P. (1995). Fast Numerical Valuation of American, Exotic and Complex Options. Submitted to *Applied Mathematical Finance.*

[11] Friedman, A. (1988) *Variational Principles and Free-Boundary Problems.* Krieger, Malabar, Florida.

[12] Geske, R. and Johnson, H. (1984). The American Option Valued Analytically. *J. Finance* **39** 1511–1524.

[13] Geske, R. and Shastri, K. (1985). Valuation by Approximation: A Comparison of Option Valuation Techniques. *J. Financial Quantitative Anal.* **20** (1) 45–71.

[14] Harrison, J. M. and Kreps, D. (1979). Martingales and Arbitrage in Multi-Period Security Markets. *J. Economic Theory* **20** 381–408.

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

[16] IBM Corporation. (1992). *Optimization Subroutines Library Guide and Reference Release 2.* 4th Edition.

[17] Jacka, S.D. (1991). Optimal Stopping and the American Put. *Mathematical Finance* **1** 1–14.

[18] Jaillet, P., Lamberton, D. and Lapeyre, B. (1990). Variational Inequalities and the Pricing of American Options. *Acta Appl. Math.* **21** 263–289.

[19] Karatzas, I. and Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus.* 2nd edition, Graduate Texts in Mathematics, Springer-Verlag, Berlin.

[20] Karmarkar, N. (1984). A New Polynomial Time Algorithm for Linear Programming. *Combinatorica* **4** 373–394.

[21] Koopmans, T.C. (1951). *Activity Analysis of Production and Allocation.* Yale University Press.

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

[23] Mangasarian, O.L. (1979). Simplified Characterisations of Linear Complementarity Problems Solvable as Linear Programs. *Mathematics of Operations Research* **4** 3.

[24] Myneni, R. (1992). The Pricing of the American Option. *The Annals of Applied Probability* **2** (1) 1–23.

[25] van Moerbecke, P.L.J. (1976). On Optimal Stopping and Free Boundary Problems. *Arch. Rational Mech. Anal.* **60** 101–148.

Risk-free rate r	Volatility σ	Geske & Johnson $P_{\text{an}}(1,0)$	Implicit $P_{LP}(1,0)$	Crank-Nicolson $P_{LP}(1,0)$	Explicit time steps $M_{\min}$	Explicit $P_{LP}(1,0)$
.125	.5	.1476	.1475	.1479	1625	.1479
.080	.4	.1258	.1255	.1256	1040	.1260
.045	.3	.1005	.1001	.1004	585	.1004
.020	.2	.0712	.0708	.0710	260	.0712
.005	.1	.0377	.0374	.0375	65	.0376
.090	.3	.0859	.0858	.0861	1585	.0861
.040	.2	.0640	.0637	.0639	260	.0640
.010	.1	.0357	.0354	.0355	65	.0356
.080	.2	.0525	.0525	.0526	260	.0527
.020	.1	.0322	.0319	.0320	65	.0321
.120	.2	.0439	.0439	.0440	260	.0440
.030	.1	.0292	.0289	.0290	65	.0291
			$(\sum e_i^2)^{\frac{1}{2}} = 9$	$(\sum e_i^2)^{\frac{1}{2}} = 6$		$(\sum e_i^2)^{\frac{1}{2}} = 5$

Table 1: Accuracy of three American vanilla put finite difference schemes

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

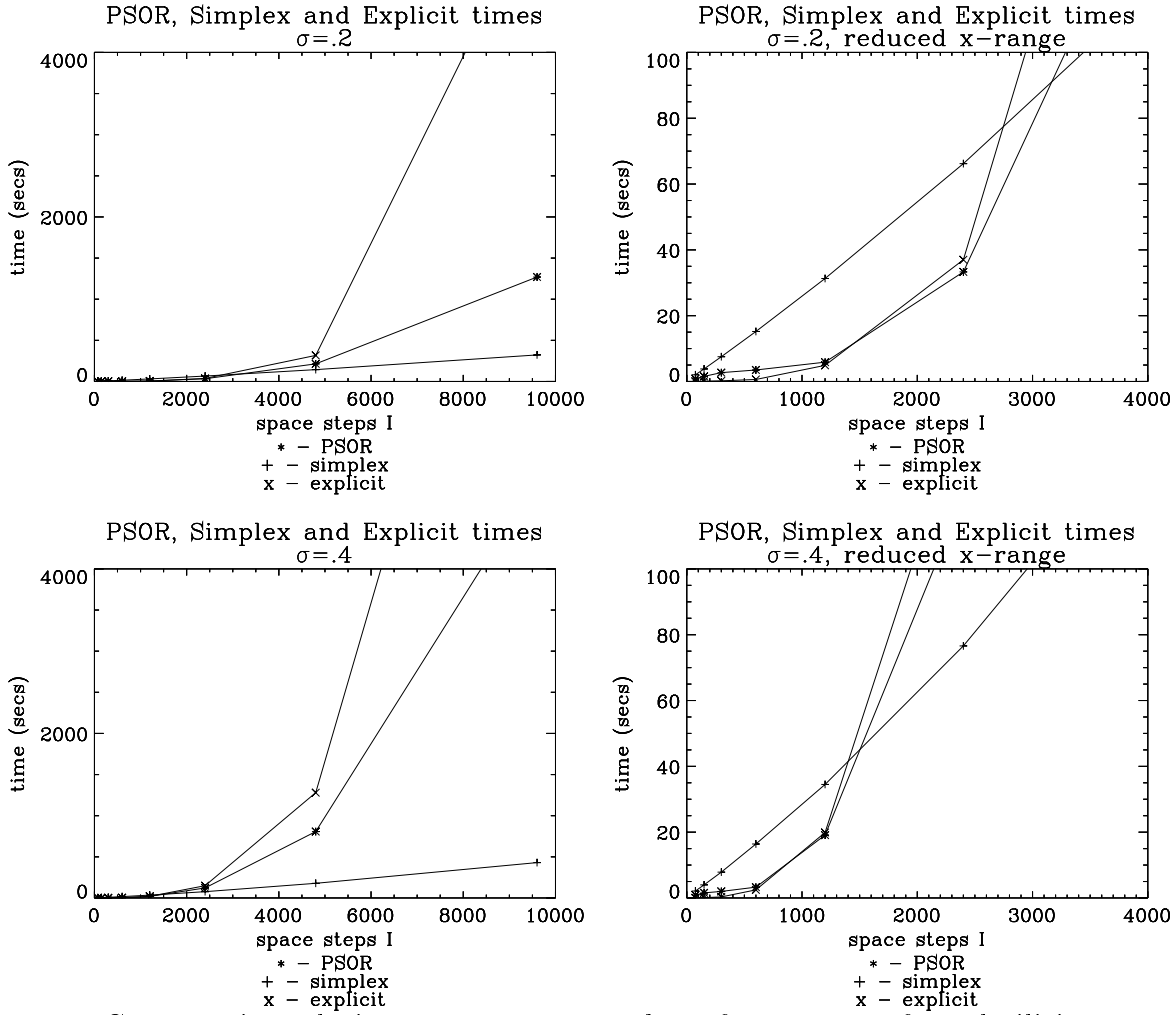

Figure 3: Comparative solution times versus number of space steps for volatilities $\sigma = .2$ and .4

time steps M	Space steps $I = 600$ PSOR	Simplex
10	.77	.31
20	.80	.53
40	.76	.85
80	.97	1.42
160	1.14	2.66
320	1.33	5.09
640	2.44	10.02
1280	5.97	19.4
2560	14.6	38.2
5120	35.0	76.7
10240	61.8	153

time steps M	Space steps $I = 2400$ 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 3: PSOR and Simplex times for varying time steps

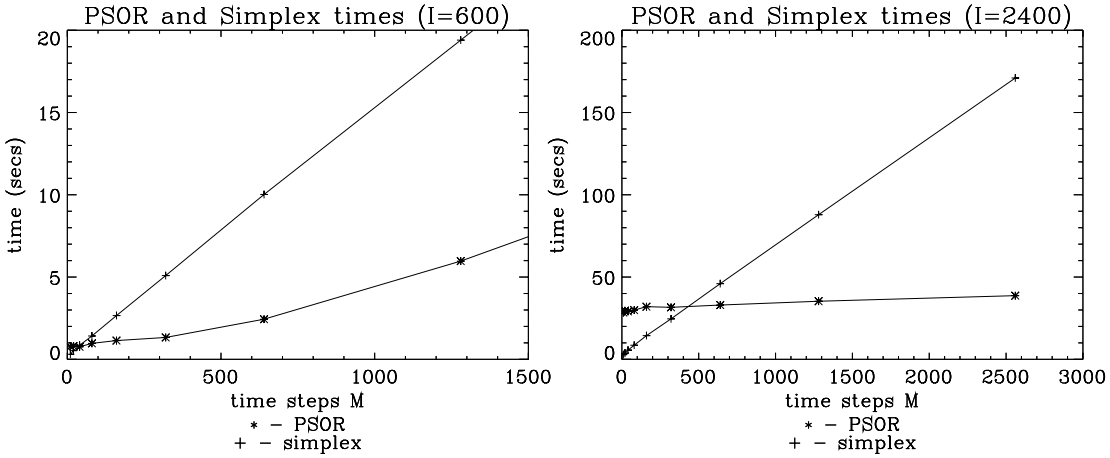

Figure 4: PSOR and Simplex times for varying time steps

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

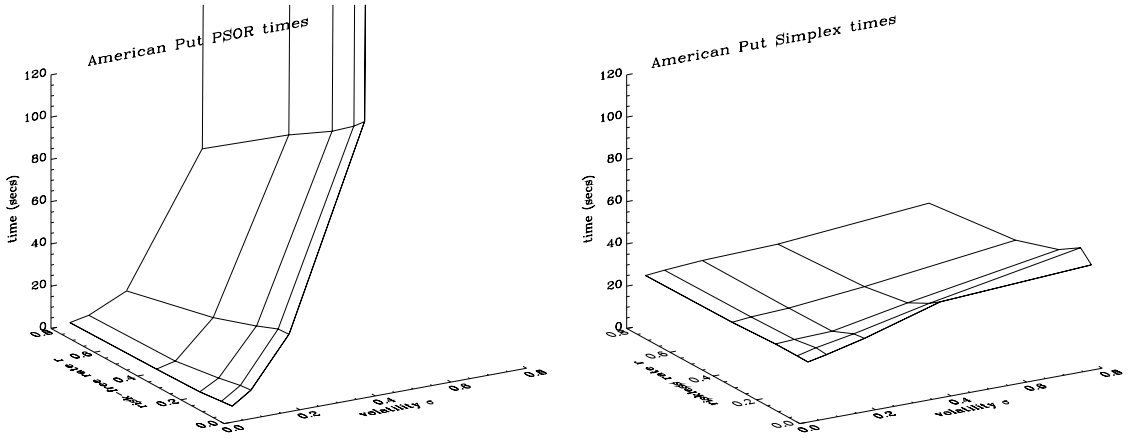

Figure 5: PSOR and simplex times for varying r and σ

space steps I	PSOR time (s)	Simplex time (s)	Explicit time (s)	Crank-Nicolson $P_{LP}(0,.5)$	Implicit $P_{LP}(0,.5)$	Explicit $P_{exp}(0,.5)$
75	.76	1.60	.08	.1091	.1091	.1091
150	1.36	2.85	.16	.1054	.1054	.1054
300	2.11	5.52	.66	.1036	.1035	.1036
600	3.63	11.4	5.27	.1026	.1026	.1026
1200	17.0	24.4	38.2	.1022	.1021	.1022
2400	102	54.9	315	.1020	.1019	.1020
4800	632	131	2580	.1018	.1018	.1018
9600	3330	324	21100	.1018	.1017	.1018
			Binomial value	.1017	.1017	.1017

Table 5: PSOR, simplex and explicit results for the American lookback put with varying space steps

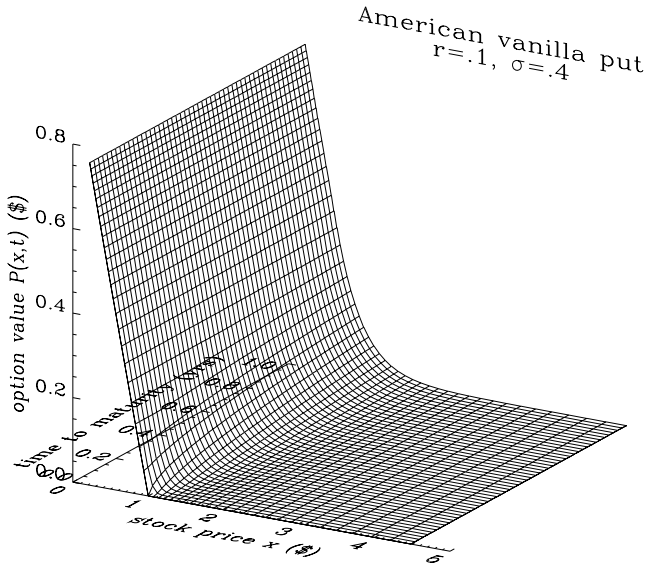

Figure 6: (LP′) solution surface with true stock price axis

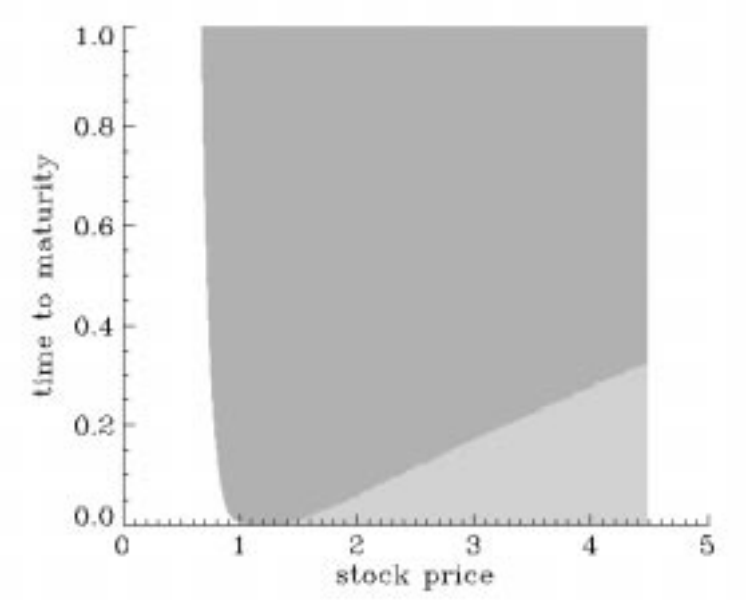

Figure 7: The computed optimal stopping boundary

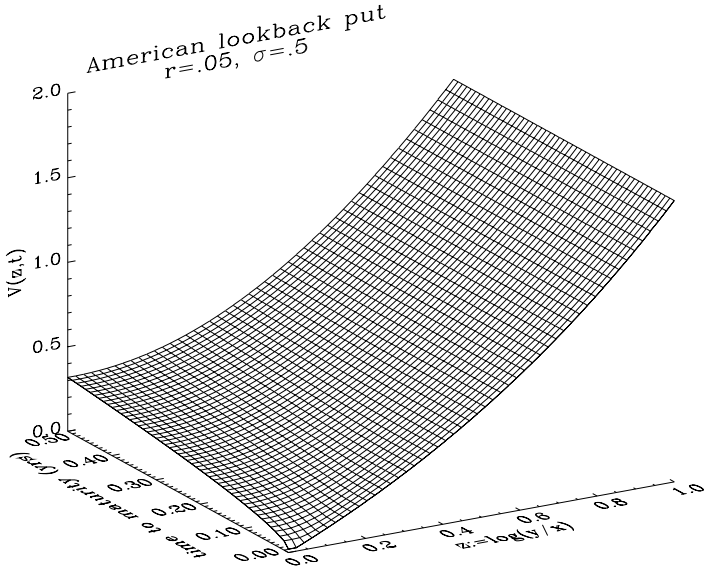

Figure 8: American lookback put value surface with exercise boundary

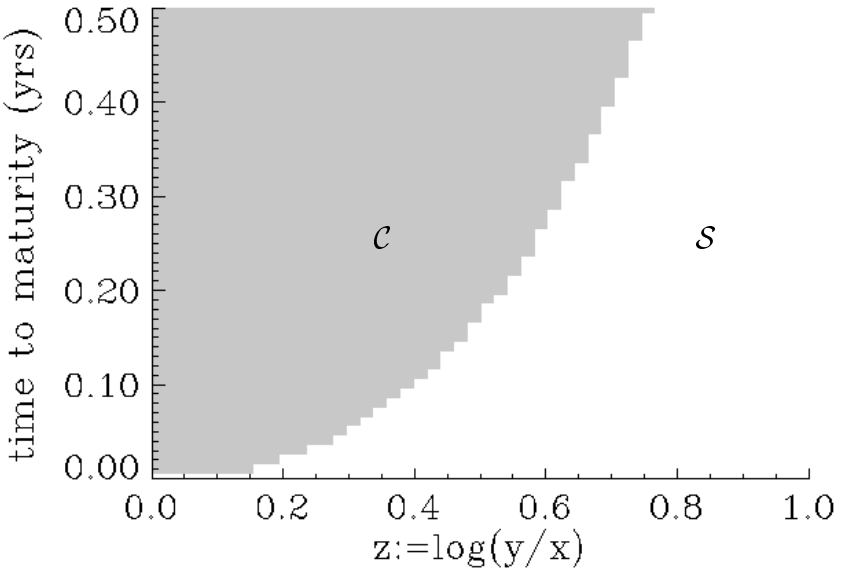

Figure 9: The optimal stopping boundary of the American lookback put