Numerical Performance of Penalty Method for American Option Pricing

K. Zhang [a] X.Q. Yang [b] S. Wang [c] K.L. Teo [d]

[a] *Department of Finance, Business School, Shenzhen University, Nanshan District, Shenzhen, P.R.China, 518060*

[b] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong*

[c] *School of Mathematics and Statistics, The University of Western Australia, Perth, Australia*

[d] *Department of Mathematics and Statistics, The Curtin University of Technology, Perth, Australia*

Abstract

This paper is devoted to studying the numerical performance of a power penalty method for a linear parabolic complementarity problem arising from American option valuation. The penalized problem is a nonlinear parabolic partial differential equation (PDE). A fitted finite volume method and an implicit time-stepping scheme are used for respectively the spatial and time discretizations of the PDE. The rate of convergence of the penalty methods with respect to the penalty parameters is investigated both theoretically and numerically. The numerical robustness and computational effectiveness of the penalty method with respect to the market parameters are also studied and compared with those from an existing popular method, PSOR.

Key words: Complementarity Problem, Option pricing, Penalty Method, Finite Volume Method.

AMS subject classification: 65M12, 65M60, 91B28

Email addresses: mazhangkai@gmail.com (K. Zhang), mayangxq@polyu.edu.hk (X.Q. Yang), swang@maths.uwa.edu.au (S. Wang), K.L.Teo@curtin.edu.au (K.L. Teo).

1 Introduction

The valuation and hedging of financial option contracts have attracted much attention from both mathematicians and financial engineers. European and American options are the two major types of options. The holders of European options have the right to exercise the contracts only at their maturity dates, while the American options can be exercised at any time up to their maturity dates. In ([1]) the authors formulated the pricing of an European option as a 2nd-order partial differential equation, known as the Black-Scholes equation. Though an analytical solution is available for the basic Black-Scholes equation, numerical methods still provide the most effective way for solving general Black-Scholes equations. American options are governed by a linear complementarity problem involving Black-Scholes operators ([10,12]) which is much hard to solve then a single Black-Scholes equation. Thus, efficient and robust numerical methods are crucial for valuating American options, as will be seen later in this paper.

There are several existing methods for solving the linear complementarity problems arising from American option valuation. These include the explicit lattice method ([4,2]), the projected successive over relaxation (PSOR) method ([12]) and the penalty methods ([14,11,3]). The explicit lattice method is widely used in practice because it is computationally inexpensive. But it has some disadvantages such as the lack of accuracy. The PSOR method is a commonly used method in the literature, see ([12]). In general, this method is easy to implement. However, its convergence rate depends crucially on the choice of the relaxation parameter and its computational cost increases exponentially as the number of spatial discretization points increases.

Recently, the power penalty method has shown its great potential for valuating American options accurately and efficiently ([14,15,11,9,3]). In this method a linear complementarity problem is approximated by a nonlinear equation (or equation system) by adding a power penalty term, called an l_k ($k > 0$) penalty. The advantages of this method are that it is simple to implement and independent of the discretization schemes and meshes used. It also works for multiple-connected problems and problems with nonlinearity, such as uncertain volatility models, drift-dominated problems, transaction cost models ([3]) and jump diffusion models ([13]). The convergence rate of the l_k penalty method for the valuation of American-style options has been theoretically studied in ([14,11,15]). Although some preliminary numerical results have been given in these papers, a thorough investigation of the numerical performance of the method is still missing from the open literature.

In this paper, we present some interesting results on the the numerical performance of the power penalty method. First, we will establish the convergence

theory of the l_k power penalty method applied to the discretized system of the complementarity problem. We will then study the numerical behaviors of method with the $l_{1/2}$, l_1 and l_2 penalty terms for solving the American option pricing problem. The numerical performance of the penalty method will be compared with that of the PSOR method to demonstrate the superiority of the penalty method to the PSOR in terms of computational costs with respect to the number of spatial mesh nodes and to the values of the system parameters. More specifically, we will show that the computational time of the penalty method increases linearly as the number of spatial steps increases, while the computational time of PSOR grows exponentially as the same parameter increases. The numerical results also show that the computational cost of the penalty method is almost constant for different values of the system (market) parameters (i.e. interest rate and volatility) while that of the PSOR method varies significantly with respect to different values of the parameters.

The organization of this paper is as follows. In Section 2, we introduce the American option pricing model in terms of a differential complementarity problem. We will also give a briefly account of the power penalty method. In Section 3, we will briefly discuss the discretization of the penalized equation by a fitted finite volume method in space and an implicit scheme in time. Section 4 studies the convergence property of the penalty method. Numerical performance of the $l_{1/2}$, l_1, l_2 penalty methods and a comparison of the performances of the penalty method and the PSOR method will be given in Section 5.

2   The model and the penalty approach

Consider an asset with price S which satisfies the following stochastic differential equation
$$dS = \mu S\, dt + \sigma S\, dW,$$
where W is a standard Brownian motion, μ is the drift rate and σ denotes a deterministic local volatility. Let $U(S,t)$ denote the value of a standard American put option on the underlying stock price S and time t, T the *expiry time*, and K the *striking price*. It is known that, under the non-arbitrage assumption, the American option pricing problem can be mathematically stated as a backward differential complementarity problem as follows ([6,12]):

$$
\begin{cases}
\mathcal{L}U \geq 0, \\
U - U^* \geq 0, \\
(\mathcal{L}U = 0) \wedge (U - U^* = 0), \\
\quad \text{a.e. in } \ \Omega = I \times (0, T),
\end{cases}
\tag{1}
$$

3

where

$$\mathcal{L}U(S,\tau) = \frac{\partial U}{\partial \tau} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} - rS\frac{\partial U}{\partial S} + rU$$

denotes the Black-Scholes differential operator, $I = (0, S_{\max}) \subset \mathbb{R}$ for a constant $S_{\max} > 0$, $\tau = T - t$, r is the risk-free interest rate and $U^*(S)$ is the *payoff function* defined by

$$U^*(S) = \max\{K - S, 0\}.$$

At $\tau = 0$, $U(S, \tau)$ is given by

$$U(S, \tau = 0) = U^*(S). \tag{2}$$

We assume $S_{\max} \gg K$ and the boundary conditions for the complementarity problem are:

$$U(0, \tau) = K, \qquad U(S_{\max}, \tau) = 0. \tag{3}$$

We comment that the real boundary condition should be $U(S, \tau) = 0$ when $S \to \infty$. However, for computational purpose, we truncate it.

For convenience of theoretical analysis, we introduce the following new variable

$$V(S, \tau) = e^{\xi\tau}U(S, \tau)$$

where ξ is a constant satisfying

$$\xi \geq \sigma^2. \tag{4}$$

Under this transformation, it is easy to show that, the complementarity problem in (1)–(3) becomes

$$\begin{cases} LV \geq 0, \\ V - V^* \geq 0, \\ (LV = 0) \wedge (V - V^* = 0) \end{cases} \tag{5}$$

with the initial condition

$$V(S, \tau = 0) = U(S, \tau = 0) = U^*(S) =: V^*(S) \tag{6}$$

and the boundary conditions

$$V(0, \tau) = Ke^{\xi\tau}, \qquad V(S_{\max}, \tau) = 0. \tag{7}$$

In (5),

$$LV(S, \tau) = \frac{\partial V}{\partial \tau} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rS\frac{\partial V}{\partial S} + (r + \xi)V.$$

Applying the penalty method in [15,11] to (5)–(7) yields

$$LW - \lambda \left[V^* - W\right]_+^{1/k} = 0, \tag{8}$$

with the initial and boundary conditions

$$W(S, \tau = 0) = V^*(S),$$
$$W(0, \tau) = Ke^{\xi\tau},$$
$$W(S = S_{\max}, \tau) = 0.$$

Here, W is the penalized solution, $k > 0$ is the power of the penalty term and $\lambda > 1$ and is the penalty parameter. The idea behind the penalty approach is simple. By adding a penalty term $-\lambda[V^* - W]_+^{1/k}$, we force the positive part $[V^* - W]_+^{1/k}$ to be close to zero as the penalty parameter λ gets sufficiently large. Hence, the complementarity conditions in (5) is approximately satisfied. The detailed study of the solvability and convergence property of the penalty approach can be found in ([11]).

In the next section, we will discuss the discretization of (5) and (8) based on the fitted finite volume discretization ([11,9]) in space and implicit scheme in time.

3 Discretization

In this section, we use the fitted finite volume discretization in space and the two level time stepping scheme in time with a splitting parameter θ for (5) and (8). For brevity, we omit the derivation of the fitted finite volume method which can be found in ([11,9]) and only list the final discretized forms.

First, we define two spatial partitions of $I = (0, S_{\max})$. The first is defined by dividing I into the following N sub-intervals

$$I_i = (S_i, S_{i+1}), \, i = 0, \ldots N - 1$$

with $0 = S_0 < S_1 < \cdots < S_N = S_{\max}$. For each $i = 0, \ldots N - 1$, let $h_i = S_{i+1} - S_i$. Also, we let $S_{i-1/2} = (S_{i-1} + S_i)/2$ and $S_{i+1/2} = (S_i + S_{i+1})/2$ for each $i = 0, \ldots N$ with $S_{-1/2} = S_0$ and $S_{N+1/2} = S_N$. These intervals $J_i = [S_{i-1/2}, S_{i+1/2}], i = 0, 1, ..., N$, form a second partition of $I = (0, S_{\max})$.

For a positive integer L, let the time interval $(0, T)$ be partitioned into a uniform mesh with mesh points $\tau_n = n\Delta\tau$ for $n = 0, 1, ..., L$, where $\Delta\tau = T/L$.

Let V_i^n denote the approximation of $V(S_i, \tau_n)$. The application of the finite volume method in ([11]) with on these partitions yields a fully discretized form of the complementarity problem (5) as follows:

$$\begin{cases} [I + (1-\theta)M]V^{n+1} \geq [I - \theta M]V^n, \\ V^{n+1} - V^* \geq 0, \\ ([I + (1-\theta)M]V^{n+1} = [I - \theta M]V^n) \wedge (V^{n+1} - V^* = 0), \end{cases} \quad (9)$$

where $V^n = (V_0^n, V_1^n, \ldots, V_N^n)^\top$ and M is a matrix satisfying

$$[MV^n]_i = -\Delta\tau\alpha_i V_{i-1}^n - \Delta\tau\gamma_i V_i^n - \Delta\tau\beta_i V_{i+1}^n, \quad i = 1, \ldots N - 1, \quad (10)$$

In (10),

$$\alpha_1 = \frac{S_1}{4l_1}(a - b),$$
$$\beta_1 = \frac{bS_{3/2}S_2^\eta}{(S_2^\eta - S_1^\eta)l_1},$$
$$\gamma_1 = -\frac{S_1}{4l_1}(a + b) - \frac{bS_{3/2}S_1^\eta}{(S_3^\alpha - S_2^\alpha)l_1} - c,$$

and

$$\alpha_i = \frac{bS_{i-1/2}S_{i-1}^\eta}{(S_i^\eta - S_{i-1}^\eta)l_i},$$
$$\beta_i = \frac{bS_{i+1/2}S_{i+1}^\eta}{(S_{i+1}^\eta - S_i^\eta)l_i},$$
$$\gamma_i = -\frac{bS_{i-1/2}S_i^\eta}{(S_i^\eta - S_{i-1}^\eta)l_i} - \frac{bS_{i+1/2}S_i^\eta}{(S_{i+1}^\eta - S_i^\eta)l_i} - c,$$

for $i = 2, \ldots N - 1$, where

$$a = \sigma^2/2, \ b = r - \sigma^2, \ c = r + \xi + b, \ \eta = b/a. \quad (11)$$

By defining a diagonal matrix $D(W^n)$ by

$$[D(W^n)]_i = d_i(W_i^n) = \lambda [V_i^* - W_i^n]_+^{1/k},$$

the fully discretized form of (8) by the finite volume method in [11] can be written as

$$[I + (1-\theta)M]W^{n+1} - D(W^{n+1}) = [I - \theta M]W^n, \quad (12)$$

where $W^n = (W_0^n, W_1^n, \ldots, W_N^n)^\top$ and W_i^n denote the approximation of $W(S_i, \tau_n)$. For (9) and (12), the fully implicit and Crank-Nicolson time stepping schemes correspond to $\theta = 0$ and $\theta = 1/2$, respectively.

6

Remark 1 *Note that there is no need to define (9) and (12) for $i = 0$ and N, because they correspond to the Dirichlet boundary points. Hence, M is an $(N-1) \times (N+1)$ matrix. However, by replacing V_0^{n+1} (respectively W_0^{n+1}) and V_N^{n+1} (respectively W_N^{n+1}) in (9) (respectively (13)) at $i = 1$ and $N + 1$ respectively with the the nodal values of the corresponding Dirichlet boundary conditions and moving the relevant terms to the right-hand side of (9) (respectively (13)), M can be made an $(N-1) \times (N-1)$ matrix and (9) (respectively (13)) is a system for $V^{n+1} = (V_1^{n+1}, ..., V_{N-1}^{n+1})^\top$ (respectively $W^{n+1} = (W_1^{n+1}, ..., W_{N-1}^{n+1})^\top$). In what follows, we will continue to use M to denote this $(N-1) \times (N-1)$ matrix.*

Lemma 2 *The matrix $I + (1 - \theta)M$ is an M-matrix.*

Proof. Let us first investigate α_i and β_i. From their respective definitions we have

$$-\alpha_i \leq 0, \qquad -\beta_i \leq 0. \tag{13}$$

This is because from the definitions of b and η in (11) we have

$$\frac{-b}{S_i^\eta - S_{i-1}^\eta} = \frac{-a\eta}{S_i^\eta - S_{i-1}^\eta} < 0$$

for all $i = 1, 2, \ldots N - 1$. On the other hand,

$$\begin{aligned}
&- \alpha_i - \beta_i - \gamma_i \\
&= -\frac{bS_{i-1/2}S_{i-1}^\eta}{(S_i^\eta - S_{i-1}^\eta)\, l_i} - \frac{bS_{i+1/2}S_{i+1}^\eta}{(S_{i+1}^\eta - S_i^\eta)\, l_i} + \frac{bS_{i-1/2}S_i^\eta}{(S_i^\eta - S_{i-1}^\eta)\, l_i} + \frac{bS_{i+1/2}S_i^\eta}{(S_{i+1}^\eta - S_i^\eta)\, l_i} + c \\
&= \frac{b}{l_i}\left[S_{i-1/2} - S_{i+1/2}\right] + c = c - b = r + \xi = r + \sigma^2 > 0. \tag{14}
\end{aligned}$$

Thus, from (13) and (14) we get

$$-\gamma_i > \alpha_i + \beta_i \geq 0. \tag{15}$$

Summarizing (13), (14) and (15), we conclude that $I + (1 - \theta)M$ has non-positive off-diagonals and positive diagonals, and is strictly diagonally dominant. Hence, $I + (1 - \theta)M$ is an M-matrix. ■

4 Convergence rate of penalty method

In this section, we investigate the convergence behavior of the penalty method (12) applied to the linear complementarity problem (9). To this end, we first prove that the penalty term is bounded independently of the penalty parameters.

Theorem 3 *For any $n = 0, 1, ..., L - 1$, there exists a constant $C > 0$, independent of λ and k such that*

$$\left\| [V^* - W^{n+1}]_+ \right\| \leq \frac{C}{\lambda^k},$$

where W^{n+1} denotes the solution to (12) and $\| \cdot \|$ denotes a norm on $\mathbb{R}^{N-1}$.

Proof. We first prove that the matrix M is semi-positive definite. For any $x = (x_1, x_2, \cdots, x_{N-1})^\top \in \mathbb{R}^{N-1}$, it is easy to show that

$$
\begin{aligned}
x^T M x &= (Mx, x) \\
&= \Delta\tau \left[\frac{S_1(a+b)}{4l_1} x_1^2 + b \sum_{i=1}^{N-1} S_{i+1/2} \frac{S_{i+1}^\eta x_{i+1} - S_i^\eta x_i}{(S_{i+1}^\eta - S_i^\eta) l_i} (x_{i+1} - x_i) + c \sum_{i=1}^{N-1} x_i^2 \right] \\
&= \Delta\tau \left[\frac{S_{1/2}(a-b)}{2l_1} x_1^2 + b \cdot I + c \sum_{i=1}^{N-1} x_i^2 \right], \tag{16}
\end{aligned}
$$

since $S_1 = 2S_{1/2}$. For the term I, we have

$$
\begin{aligned}
I &= \sum_{i=1}^{N-1} S_{i+1/2} \frac{S_{i+1}^\eta(x_{i+1} - x_i) + (S_{i+1}^\eta - S_i^\eta)x_i}{(S_{i+1}^\eta - S_i^\eta) l_i} (x_{i+1} - x_i) \\
&= \sum_{i=1}^{N-1} \frac{S_{i+1/2} S_{i+1}^\eta}{(S_{i+1}^\eta - S_i^\eta) l_i} (x_{i+1} - x_i)^2 + \sum_{i=1}^{N-1} \frac{S_{i+1/2}}{l_i} x_i (x_{i+1} - x_i) \\
&= \sum_{i=1}^{N-1} \frac{S_{i+1/2} S_{i+1}^\eta}{(S_{i+1}^\eta - S_i^\eta) l_i} (x_{i+1} - x_i)^2 + \sum_{i=1}^{N-1} \frac{S_{i+1/2}}{l_i} \left[\frac{1}{2}\left(x_{i+1}^2 - x_i^2\right) - \frac{1}{2}(x_{i+1} - x_i)^2 \right] \\
&= \frac{1}{2}\sum_{i=1}^{N-1} \frac{S_{i+1/2} S_{i+1}^\eta}{(S_{i+1}^\eta - S_i^\eta) l_i} (x_{i+1} - x_i)^2 + \frac{1}{2}\sum_{i=1}^{N-1} \frac{S_{i+1/2}}{l_i} \left(x_{i+1}^2 - x_i^2\right).
\end{aligned}
$$

From the above we see that

$$
\begin{aligned}
I &\geq \frac{1}{2}\sum_{i=1}^{N-1} \frac{S_{i+1/2}}{l_i} \left(x_{i+1}^2 - x_i^2\right) \\
&= \frac{1}{2}\left[\frac{S_{3/2}}{l_1}\left(x_2^2 - x_1^2\right) + \frac{S_{5/2}}{l_2}\left(x_3^2 - x_2^2\right) + \cdots + \frac{S_{N-1/2}}{l_{N-1}}\left(x_N^2 - x_{N-1}^2\right) \right] \\
&= \frac{1}{2}\sum_{i=1}^{N-1} \frac{(S_{i-1/2} - S_{i+1/2})}{l_i} x_i^2 - \frac{S_{1/2}}{2l_1} x_1^2 \\
&= -\frac{1}{2}\sum_{i=1}^{N-1} x_i^2 - \frac{1}{2} S_{1/2} \frac{x_1^2}{l_1}. \tag{17}
\end{aligned}
$$

8

Combining (16) and (17), it follows from (11) and (4) that

$$
\begin{aligned}
x^T M x &\geq \Delta\tau \left[\frac{S_{1/2}(a-b)}{2l_1} x_1^2 - \frac{b}{2} \sum_{i=1}^{N-1} x_i^2 - \frac{bS_{1/2}}{2l_1} x_1^2 + c \sum_{i=1}^{N-1} x_i^2 \right] \\
&= \Delta\tau \left[\frac{aS_{1/2}}{2l_1} x_1^2 + \left(c - \frac{b}{2} \right) \sum_{i=1}^{N-1} x_i^2 \right] \\
&= \Delta\tau \left[\frac{S_{1/2}}{4l_1} \sigma^2 x_1^2 + \frac{(3r+\sigma^2)}{2} \sum_{i=1}^{N-1} x_i^2 \right] \\
&\geq 0.
\end{aligned}
$$

Hence, M is semi-positive definite. Consequently, the matrix $I + (1-\theta)M$ is positive definite. Moreover, for the matrix $I + (1-\theta)M$, all the off-diagonal entries are negative and the diagonal entries are non-negative, by Lemma 2. Thus, under these two conditions, this theorem is just a consequence of Lemma 2.1 in ([10]), which gives a convergence result for a more general linear complementarity problem. ∎

The above theorem shows that at those nodes where $W^{n+1} - V^* < 0$,

$$
\lambda (V^* - W^{n+1})^{1/k} \leq C,
$$

or equivalently,

$$
(W^{n+1} - V^*) \geq -\frac{C}{\lambda^k}
$$

for a positive constant C, independent of λ and k. Thus, we see that the penalty method (12) solves the following complementarity problem.

$$
\begin{cases}
[I + (1-\theta)M]W^{n+1} \geq [I - \theta M]W^n, \\
W^{n+1} - V^* \geq -\frac{C}{\lambda^k}, \\
([I + (1-\theta)M]W^{n+1} = [I - \theta M]W^n) \wedge \left(-\frac{C}{\lambda^k} \leq W^{n+1} - V^* \leq 0 \right).
\end{cases} \tag{18}
$$

Obviously, (18) is an approximation of (9) with a relaxation term $-\frac{C}{\lambda^{1/k}}$ at the nodes at which $W^{n+1} - V^* < 0$. Also, (18) shows that we can select λ and k properly so that the solution to (18) approximates that of (9) with an arbitrary level of precision. This is given in the following theorem.

Theorem 4 *For any $n = 0, 1, ..., L-1$, let V^{n+1} and W^{n+1} be the solutions to (9) and (12) respectively. Then, there exists a constant $C > 0$, independent of λ and k, such that*

$$
\left\| W^{n+1} - V^{n+1} \right\| \leq \frac{C}{\lambda^k}
$$

for all $n = 0, 1, ..., L-1$.

9

Table 1
Data used to value American vanilla options

Parameter values	
r	0.1
σ	0.2 and 0.8
T	0.25
K	100

This theorem is just a consequence of Theorem 2.1 in ([10]) based on Theorem 3. We thus omit the proof of this theorem and refer the reader to ([10]).

Theorem 4 implies that the solution of (12) converges to that of (9) at the rate of $\mathcal{O}(1/\lambda^k)$ in the Euclidean norm. This rate is exponential in k for any $\lambda > 1$. In the next section, we will verify this theoretical result by some numerical experiments for some particular values of k and λ.

5 Numerical experiments

In this section, we investigate the numerical performance of the penalty approach discussed above using three special values of k, i.e., $k = 1/2, 1$ and 2. More specifically, we study the convergence behaviors of these penalty methods with respect to the penalty parameter λ. The test problem is a standard American vanilla put option with the parameters as given in Table 1.

In order to get a desirable accuracy, we choose the $S_{\max} = 10K$ in the original complementarity problem. The interval $I = (0, S_{\max})$ is partitioned uniformly into 2000 sub-intervals, yielding the spatial mesh size $h = 0.5$. The time interval $[0, T]$, where T is given in Table 1, is divided uniformly into 100 sub-intervals so that $\Delta t = 0.0025$. In all the numerical experiments, we use the Crank-Nicolson time stepping method, i.e. $\theta = 1/2$ in (12), so that the time discretization is of a 2nd-order accuracy. It should be noted that due to the non-smoothness of the payoff function, the usual second order convergence rate of Crank-Nicolson schemes may not be achieved. To remedy this, we use the Rannacher time stepping technique ([7]), that is, we use the fully implicit scheme ($\theta = 0$ in (12)) in the first two time steps and the Crank-Nicolson thereafter.

All the numerical results in this paper were computed in the double precision on a Pentium IV PC with 2.8 GHz and 512MB memory under the Visual C++.net environment. The values of the options at time $t = 0$ and $S = K = 100$ are computed and compared.

To determine numerically the numerical orders of convergence rate of the penalty method with respect to λ, we choose a sequence of λ defined by $\lambda_{n+1} = m\lambda_n$ for $n = 0, 1, \ldots$ for a given λ_0, where $m \geq 2$ is a constant. Then, we define

$$\text{Ratio} = \frac{W(\lambda_{n+2}) - W(\lambda_{n+1})}{W(\lambda_{n+1}) - W(\lambda_n)} = \frac{\Delta W_{n+2}}{\Delta W_{n+1}}, \tag{19}$$

where $W(\lambda)$ denotes the approximation of the exact option value at a given point (S, t) computed by (12) using the penalty parameter λ_n. The order of convergence is then defined by

$$\text{Rate} = \log_m \text{Ratio}. \tag{20}$$

5.1 $l_{1/2}$ penalty method

In the case of $k = 1/2$, the $l_{1/2}$ penalty method is essentially the quadratic penalty method. In this case, (12) becomes

$$\left[(I + (1-\theta)M)W^{n+1} \right]_i - \lambda \left[V_i^* - W_i^{n+1} \right]_+^2 = \left[(I - \theta M)W^n \right]_i. \tag{21}$$

It is clear that (21) is a smooth nonlinear system and thus we use the classical Newton method to linearize the system, yielding

$$\left[I + (1-\theta)M + J_D(\varpi^{l-1}) \right] \delta\varpi^l = [I - \theta M]W^n - (I + (1-\theta)M)\varpi^l - D(\varpi^{l-1}),$$
$$\varpi^l = \varpi^{l-1} + \gamma \cdot \delta\varpi^l \tag{22}$$

for $l = 1, 2, \ldots$ with ϖ^0 being a given initial guess, where $J_D(\varpi)$ denotes the Jacobian of the column vector $D(\varpi)$ and $\gamma \in (0, 1]$ denotes a damping parameter. From (21), it is easy to see that the Jacobian $J_D(\varpi)$ is the following diagonal matrix

$$J_D(\varpi) = \text{diag}\left(d_1'(\varpi_1), \cdots, d_{N+1}'(\varpi_{N+1}) \right),$$

where

$$d_i'(\varpi_i) = 2\lambda[V_i^* - W_i^{n+1}]_+.$$

We then choose

$$W^{n+1} = \lim_{l \to \infty} \varpi^l.$$

Table 2
Value of American put options at $S = K, t = 0$, computed by $l_{1/2}$ penalty method. $r = 0.10$, $\sigma = 0.2$ and 0.8, $T = 0.25$, $K = 100$. The fitted finite volume method is used. Ratio and ΔW are defined in (19). Rannacher technique is used. $tolerance = 10^{-6}$ for the Newton iteration

$\sigma = 0.2$				$\sigma = 0.8$			
λ_n	$W(\lambda_n)$	ΔW_n	Ratio	λ_n	$W(\lambda_n)$	ΔW_n	Ratio
250	3.02733			250	14.6310		
1000	3.04530	0.01797		1000	14.6515	0.0205	
4000	3.05481	0.00951	1.9	4000	14.6628	0.0113	1.8
16000	3.05969	0.00488	1.9	16000	14.6688	0.0060	1.9
64000	3.06215	0.00246	2.0	64000	14.6718	0.0030	2.0
256000	3.06337	0.00122	2.0	256000	14.6734	0.0016	1.9
1024000	3.06398	0.00061	2.0	1024000	14.6742	0.0008	2.0
4096000	3.06428	0.00030	2.0	4096000	14.6746	0.0004	2.0

Clearly, ϖ^l converges to W^{n+1} quadratically if ϖ^0 is sufficiently close to W^{n+1}.

Table 2 contains the results computed by the quadratic penalty method with the sequence of λ. Note that in this case $m = 4$ in (20). Thus, from the table and (20) we see that the computed rates of convergence for this case is about $1/2$ which well matches the theoretical rate in Theorem 4.

5.2 l_1 penalty method

In the case of $k = 1$, the l_1 penalty method is the linear penalty method. In this case, (12) becomes

$$\left[(I + (1 - \theta)M)W^{n+1}\right]_i - \lambda \left[V_i^* - W_i^{n+1}\right]_+ = \left[(I - \theta M)W^n\right]_i. \qquad (23)$$

It is clear that (23) is a semismooth nonlinear system. We solve this system using the algorithm (22) with the following constant Jacobian

$$J_D(\varpi) = \operatorname{diag}(\lambda, \cdots, \lambda).$$

Table 3 shows the results computed by the linear penalty method with a sequence of penalty parameters. From the table we see that the computed rates of convergence is about 1, coinciding with the theoretical result in Theorem 4.

Table 3
Value of American put options at $S = K, t = 0$, computed by l_1 penalty method. $r = 0.10$, $\sigma = 0.2$ and 0.8, $T = 0.25$, $K = 100$. The fitted finite volume method is used. Ratio and ΔW are defined in (19). Rannacher technique is used. $tolerance = 10^{-6}$ is chosen for the Newton iteration

	$\sigma = 0.2$				$\sigma = 0.8$		
λ_n	$W(\lambda_n)$	ΔW_n	Ratio	λ_n	$W(\lambda_n)$	ΔW_n	Ratio
125	3.04804			125	14.6541		
250	3.05622	0.01797		250	14.6641	0.0100	
500	3.06039	0.00951	1.9	500	14.6695	0.0054	1.9
1000	3.06250	0.00488	1.9	1000	14.6722	0.0027	2.0
2000	3.06354	0.00246	2.0	2000	14.6736	0.0014	2.0
4000	3.06406	0.00122	2.0	4000	14.6743	0.0007	2.0
8000	3.06432	0.00061	2.0	8000	14.6746	0.0003	2.3

5.3 l_2 penalty method

In the case of $k = 2$, this is the l_2 penalty method which is also a so-called lower order penalty method ([8]). In this case, (12) becomes

$$\left[(I + (1 - \theta)M)W^{n+1}\right]_i - \lambda \left[V_i^* - W_i^{n+1}\right]_+^{1/2} = \left[(I - \theta M)W^n\right]_i. \qquad (24)$$

Clearly, (24) is neither a non-smooth, nonlinear system with an unbounded derivative when $V_i^* - W_i^{n+1} \to 0^+$. To over the nonsmooth difficulty, we first smooth the term $[V_i^* - W_i^{n+1}]_+^{1/2}$ in (24) by

$$[V_i^* - W_i^{n+1}]_+^{\frac{1}{2}} = \begin{cases} [V_i^* - W_i^n]_+^{\frac{1}{2}}, & V_i^* - W_i^n \geq \varepsilon, \\ \frac{5}{2}\varepsilon^{-\frac{3}{2}}[V_i^* - W_i^n]_+^2 - \frac{3}{2}\varepsilon^{-\frac{5}{2}}[W_i^* - W_i^n]_+^3, & V_i^* - W_i^m < \varepsilon, \end{cases} \qquad (25)$$

where $1 >> \varepsilon > 0$ is a regularization parameter. It has been shown in Corollary 5.1 of ([11]) that (25) is smooth and increasing on $(-\infty, +\infty)$. We then apply the classical Newton method to (24), yielding the same system as (22) with the Jacobian obtained by differentiating (25) with respect to W^{n+1}.

Table 4 shows the results computed by the l_2 penalty method with a sequence of λ for two different values of σ. From the table it is seen that for both cases, the computed ratios are close to 4, implying that the computed rates of convergence at the point are close to 2. This coincides with our theoretical results in the previous section.

Table 4

Value of American put options at $S = K, t = 0$, computed by l_2 penalty method with the smoothing interval $(0, 10^{-3})$. $r = 0.10$, $\sigma = 0.2$ and 0.8, $T = 0.25$, $K = 100$. The fitted finite volume method is used. Ratio and ΔW are defined in (19). Rannacher technique is used. $tolerance = 10^{-6}$ is chosen for the Newton iteration

	$\sigma = 0.2$				$\sigma = 0.8$		
λ_n	$\Delta W(\lambda_n)$	W_n	Ratio	λ_n	$\Delta W(\lambda_n)$	W_n	Ratio
10	2.98285			10	14.5838		
20	3.03279	0.04994		20	14.6352	0.0792	
40	3.05569	0.02290	2.2	40	14.6630	0.0278	2.8
80	3.06229	0.00660	3.5	80	14.6719	0.0089	3.1
160	3.06400	0.00171	3.8	160	14.6742	0.0023	3.9
320	3.06443	0.00043	4.0	320	14.6747	0.0005	4.6

To conclude this section, we comment that from Tables 2, 3 and 4, it is clear that the l_2 penalty method needs a much smaller penalty parameter than those used in $l_{1/2}$ and l_1 penalty methods in order to archive a comparable accuracy to those from the latter two methods. This confirms the theoretical results in the precious section as well in [8]. Furthermore, the computed convergence rates for the l_2 penalty method are respectively one and two order higher than those of of the l_1 and $l_{1/2}$ penalty methods, as predicted by the theoretical result in Theorem 4.

6 Comparison with PSOR method

In this section we will present a comparison of the l_2 penalty method and PSOR. As pointed out in the introduction, PSOR method is commonly used in American option pricing because of its simplicity. However, it turns out that its performance strongly depends on the number of partition points and the market parameters of a problem, as will be seen later in this section. In contrast to this, the l_2 penalty method has a much better performance than PSOR in terms of sensitivity to the choice of mesh and market parameters. This is given below.

Define

$$A = I + (1 - \theta)M,$$
$$b = [I - \theta M]W^n.$$

The PSOR method for American option pricing based on the discretization scheme developed in Section 3 can be stated as follows.

Algorithm 1

1: *Let* $(W^{n+1})^0 = W^n$;
2: *Let* $\varpi^l = (W^{n+1})^l$;
3: *For* $l = 1, 2, \ldots$ *until convergence*
   Solve

$$\varpi_i^{l+1} = \varpi_i^l + \frac{\rho}{A_{i,i}} \left[b_i - A_{i,i-1}\varpi_{i-1}^{l+1} - A_{i,i}\varpi_i^l - A_{i,i+1}^l \varpi_{i+1}^l \right],$$
$$\varpi_i^{l+1} = \max\{\varpi_i^{l+1}, W_i^*\}, \qquad (0 < \rho < 2)$$

If $\max_i \frac{\left|\varpi_i^{l+1} - \varpi_i^l\right|}{\max\left(1, \left|\varpi_i^{l+1}\right|\right)} < tolerance$ *then quit*
   EndFor

In Algorithm 1, ρ is the relaxation parameter. In what follows we will present some comparative results for the l_2 penalty method and PSOR. The test problem is chosen to be the one with the market parameters being listed in Table 1.

6.1   Computational costs versus mesh parameters

The computational costs in CPU time in seconds for the l_2 penalty method and PSOR using various numbers of mesh points in S and t directions are listed respectively in Table 5 and Table 6. To look into these results, we plot them in Figures 1 and 2. From these figures we see that the CPU time for the penalty method increases linearly with respect to the increase in N_s, while the CPU time for PSOR method increases exponentially as N_s increases. From the figures it also can be seen that the computational costs for both the penalty method and PSOR are linearly dependent on N_t.

Table 5
Solution times of the penalty method and PSOR method for a sequence of space steps. N_s and N_t are the numbers of space and time steps, respectively. l_2 penalty method with the smoothing interval $(0, 10^{-3})$ is used

		$\sigma = 0.2$		$\sigma = 0.8$	
N_s	N_t	l_2	PSOR	l_2	PSOR
200	200	0.07005	0.01043	0.08750	0.26300
400	200	0.13047	0.01978	0.14765	0.13152
800	200	0.23255	0.05573	0.27995	0.86223
1600	200	0.46933	0.55182	0.56301	6.01483

15

Table 6
Solution times of the penalty method and PSOR method with a sequence of space steps. N_t is the time steps. l_2 penalty method with the smoothing interval $(0, 10^{-3})$ is used

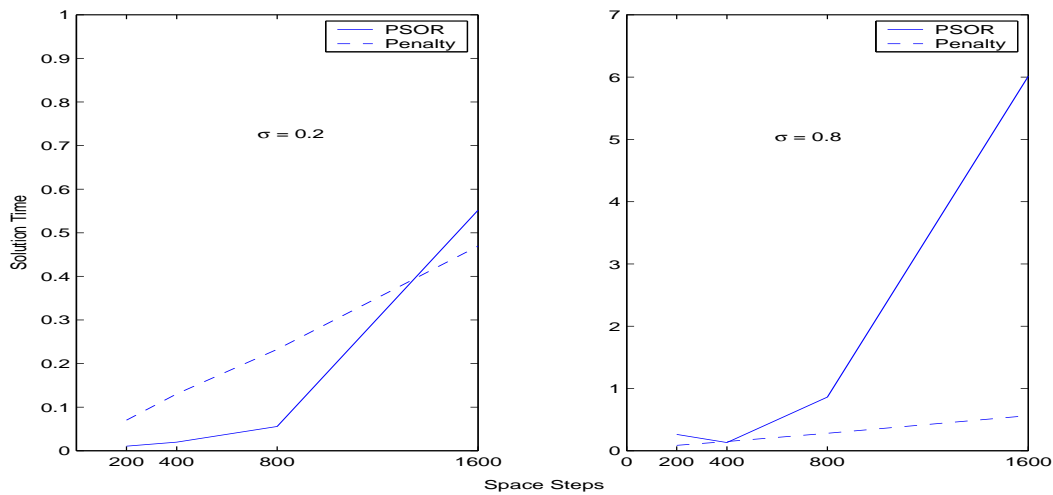| N_s | N_t | $\sigma = 0.2$ | | $\sigma = 0.8$ | |
		l_2	$PSOR$	l_2	$PSOR$
800	100	0.16380	0.04167	0.15858	0.83438
800	200	0.26198	0.05598	0.27995	0.86233
800	400	0.42343	0.08412	0.51983	0.90027
800	800	0.72187	0.13802	0.94115	0.96355

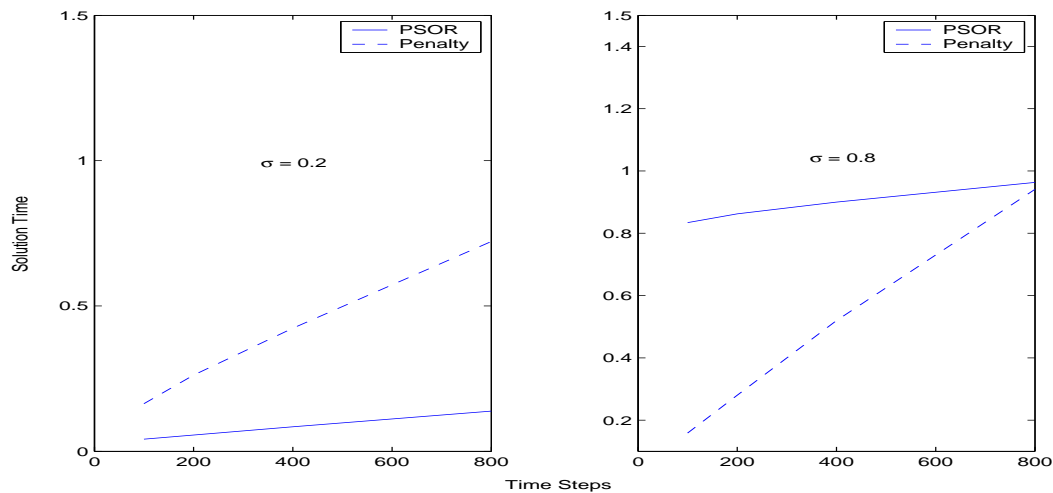Fig. 1. Solution times of the PSOR and penalty method for various space steps

Fig. 2. Solution times of the PSOR and penalty method for various time steps

16

Table 7
Solution times of the PSOR method with different market parameters. Space steps $N_S = 1600$, time steps $N_t = 200$

	r			
σ	0.10	0.20	0.40	0.80
0.10	0.10052	0.09895	0.93483	0.97642
0.20	0.25157	0.24662	0.23438	0.90339
0.40	1.12005	1.07238	1.03125	0.92473
0.80	6.05442	6.02578	5.91875	5.65860

Figures 1 and 2 show that the penalty method has computational cost almost as a linear function of number of space mesh points N_s, while PSOR method exhibits an exponential solution-time behavior as N_2 increases. For a smaller number, N_s, of the space steps, the PSOR method performs faster than the penalty method. However, as the number of spatial mesh points increases, the computational cost required by the penalty method is significantly less that that required by PSOR. From Figures 1 and 2, it also can be seen that both the penalty method and PSOR method have comparable computational costs which vary linearly as N_t, of the time steps, varies.

6.2 Computational costs versus market parameters

We now investigate the computational performances of the l_2 penalty method and PSOR under different market parameters. More specifically, we will look into the computational costs for the two methods in term of CPU time in seconds with respect to two market parameters, the risk-free interest rate r and volatility σ, in the American put option. The computed results using a fixed mesh with 1600×200 mesh points and a range of values of r and σ are listed in Tables 7 and 8 respectively for PSOR and the l_2 penalty method. Plots representing the results in the two tables are given in Figure 3. As in the previous subsection, from the figure we see that the CPU time for the penalty method is rather independent o the values of the parameters, while the computational cost for PSOR grows exponentially as σ increases. This shows that our penalty method is much less sensitive to the change in the parameter values as PSOR. The robustness of the penalty method with respect to the market parameters is a significant advantage in option pricing theory and practice.

Table 8
Solution times of the penalty method with different market parameters. Space steps $N_S = 1600$, time steps $N_t = 200$

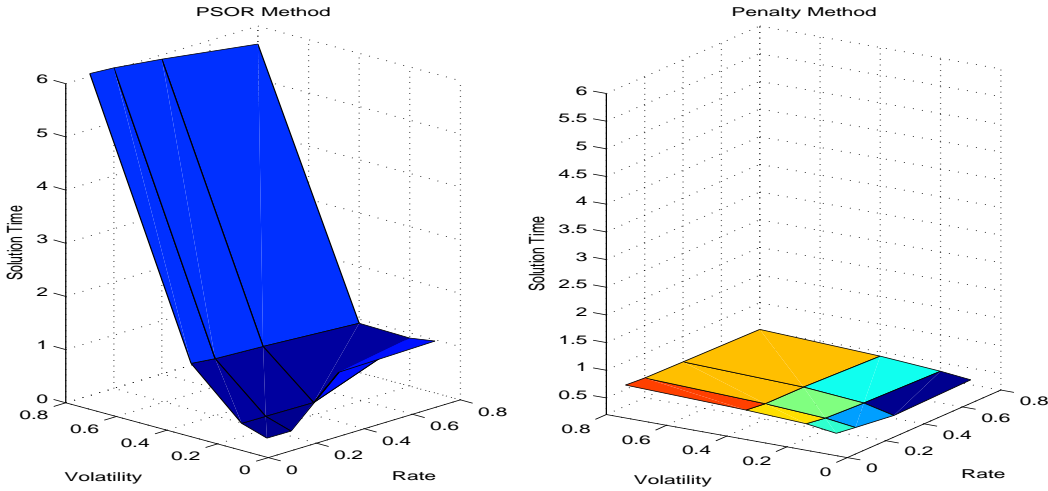|	r			
σ	0.10	0.20	0.40	0.80
0.10	0.38750	0.34113	0.25417	0.28658
0.20	0.46442	0.41407	0.38098	0.34505
0.40	0.51772	0.47890	0.47448	0.42473
0.80	0.58152	0.55572	0.54272	0.51900

Fig. 3. Solution times of the PSOR and penalty methods for different market parameters. Space steps $N_S = 1600$, time steps $N_t = 200$

7   Conclusions

We have studied the numerical performance of penalty method for valuating American options. The continuous model and its penalization were discussed. The fitted finite volume method was used to discretize the continuous nonlinear penalized problem. Convergence properties of the penalty method was given. Via detailed numerical tests, the convergence properties with the penalty parameter of the quadratic, linear and lower order penalty methods were investigated. Numerical experiments clearly confirmed the theoretical results. By comparing with the PSOR method, we also explored the stability of the penalty method under various space and time discretization schemes. Moreover, the robustness with respect to the marker parameters (r and σ) were revealed by numerical experiments.

References

[1] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Political Economy*, 81:637–659, 1973.

[2] J. C. Cox, S.A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *J. Financial Econ.*, 7:229-263, 1979.

[3] P.A. Forsyth and K.R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM J. on Sci. Comput.*, 23:2095-2122, 2002.

[4] J. Hull. *Options, Futures, and Other Derivatives.* Prentice-Hall, Englewood Cliffs, 2005.

[5] C.-S. Huang, C.-H. Hung, and S. Wang. A fitted finite volume method for the valuation of options on assets with stochastic volatilities. *Computing*, 77:297-320, 2006.

[6] J. Huang and J.-S. Pang. Option pricing and linear complementarity. *J. Comp. Fin.*, 2:31-60, 1998.

[7] R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numer. Math.*, 43:309–327, 1984.

[8] A.M. Rubinov, B.M. Glover, and X. Q. Yang. Decreasing functions with application to penalization. *SIAM J. Optim.*, 10:289–313, 1999.

[9] S. Wang. A novel fitted finite volume method for the Black-Scholes equation governing option pricing. *IMA J. Numer. Anal.*, 24:699-720, 2004.

[10] S. Wang, and X.Q. Yang. A power penalty method for linear complementarity problems, *Oper. Res. Lett.*, 36:211-217, 2008.

[11] S. Wang, X.Q. Yang, and K.L. Teo. Power penalty method for a linear complementarity problem arising from American option valuation. *J. Optim. Theory Appl.*, 129(2):227-254, 2006.

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

[13] K. Zhang and S. Wang. Pricing options under jump diffusion processes with fitted finite volume method. *Appl. Math. Comput.*, 201:398-413, 2008.

[14] K. Zhang, S. Wang, X.Q. Yang, and K.L. Teo. A power penalty approach to numerical solution of two-factor American option pricing. *Numer. Math: TMA* , 2:202-223, 2009.

[15] K. Zhang, X.Q. Yang, and K.L. Teo. Power penalty method for American option pricing with jump-diffusion process. *J. Ind. Manag. Optim.* , 4(4):983-799, 2008.