A Power Penalty Approach to Numerical Solutions of Two-Asset American Options

K. Zhang[1,2,*], S. Wang[2], X. Q. Yang[3] and K. L. Teo[4]

[1] *Department of Finance, Business School, Shenzhen University, Shenzhen, China.*
[2] *School of Mathematics and Statistics, University of Western Australia, Australia.*
[3] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong.*
[4] *Department of Mathematics and Statistics, Curtin University of Technology, Australia.*

Abstract. This paper aims to develop a power penalty method for a linear parabolic variational inequality (VI) in two spatial dimensions governing the two-asset American option valuation. This method yields a two-dimensional nonlinear parabolic PDE containing a power penalty term with penalty constant $\lambda > 1$ and a power parameter $k > 0$. We show that the nonlinear PDE is uniquely solvable and the solution of the PDE converges to that of the VI at the rate of order $\mathcal{O}(\lambda^{-k/2})$. A fitted finite volume method is designed to solve the nonlinear PDE, and some numerical experiments are performed to illustrate the usefulness of this method.

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

Key words: Complementarity problem, option pricing, penalty method, finite volume method.

1. Introduction

An option is a contract tradable in a financial market which gives to its owner the right to buy (*call option*) or to sell (*put option*) a fixed quantity of a specified asset or stock at a fixed price (*exercise* or *strike price*) on (European option) or before (American option) a given date (*expiry date*). The market prices of the rights to buy and to sell are called *call prices* and *put prices*, respectively. Clearly, the price of an option is dependent on the market price(s) of its underlying stock(s). How to valuate an option has long been a hot topic for financial engineers, economists and mathematicians. In the case of a European type option on a single asset, it was shown by Black and Scholes (cf. [4]) that the price satisfies a second-order partial differential equation with respect to the time horizon t and the underlying asset price x, known as the Black-Scholes equation. The value of an

---

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

American option is determined by a linear complementarity problem involving the Black-Scholes operator [19]. Since this complementarity problem is, in general, not analytically solvable, numerical approximation to the solution is normally sought in practice. Various numerical techniques have been proposed for the numerical solution of the single-asset American option pricing problem. Among them, lattice method [6], explicit method [13], projected successive over relaxed method (PSOR) [14], linear programming method [7], Monte-Carlo method [5], and penalty method [2, 8, 9, 19–21] are the most popular ones in both practice and research.

A linear penalty approach to the linear complementarity problem was proposed and analyzed in [3] which has been used in [9, 21]. Compared with other methods mentioned above, the penalty method possesses several advantages. First, a desirable accuracy in the approximate solution can be achieved by a judicious choice of the penalty parameter. Second, the resulting penalized PDE is of a simple form that is easy to discretize in any dimensions on both structured and unstructured meshes. Finally, the penalty method can easily be extended to other option models such as those of American options with stochastic volatilities and/or transaction costs.

In the application of the penalty approach to American option pricing, a penalty term is added to the Black-Scholes equation. In [9], an l_1 penalty method is used, resulting in a convergence rate of order $\mathscr{O}(\lambda^{-1/2})$, where $\lambda > 1$ denotes the penalty parameter. A power penalty method is proposed and analyzed in [19, 20], of which the convergence rate is shown to be of order $\mathscr{O}(\lambda^{-k/2})$ for any power parameter $k > 0$. This contains the l_1 case as the special one when $k = 1$ and provides an exponential convergence rate when $k > 1$.

For a single asset American option, it has been shown in [1, 9, 19, 20] that the solution to the penalized equation converges to that of the original problem. To our best knowledge, there are no advances in the use of penalty methods for two dimensional problems in the open literature except for the quadratic and l_1 penalty methods for solving the American option pricing problem with stochastic volatility (cf. [21]) period. On the other hand, it has been shown in [20] that the penalty methods are superior to other methods such as PSOR mentioned above.

The main purpose of this paper is to develop a power penalty method for the linear complementarity problem arising from the two-dimensional American option valuation, which comes from many models, such as stochastic volatility model, interest rate model, basket options model, and so on (cf. [17]). Without loss of generality, in this paper we put our focus on the two-asset basket option model. We will approximate the linear complementarity problem by a nonlinear parabolic PDEs in two spatial dimensions with an l_k penalty term. We will then show that the solution to the nonlinear PDE converges to that of the original complementarity problem at the rate of order $\mathscr{O}(\lambda^{-k/2})$. To solve the penalized nonlinear equation, the fitted finite volume method is proposed, based on the results in [12, 18–20]. Numerical results will be presented to verify our theoretical findings.

The paper is organized as follows. In the next section, the pricing of two-asset American options will be formulated as a linear complementarity problem. This complementarity problem will, in Section 3, be reformulated as a variational inequality problem in a functional setting, and its unique solvability will be established as well. The power penalty

problem approximating the complementarity problem will be posed in Section 4. In Section 5, we will present a convergence analysis of the power penalty method. In Section 6, a fitted volume method for the discretization of the penalized PDE is proposed. Numerical experiments, performed to illustrate the usefulness of the power penalty method and confirm the theoretical results, will be presented in Section 7.

2. The mathematical model

Let x and y denote the market prices of two assets, respectively. They follow respectively the following geometric Brownian motion processes

$$dx = \mu_1 x \, dt + \sigma_1 x \, dW_1 \quad \text{and} \quad dy = \mu_2 y \, dt + \sigma_2 y \, dW_2,$$

where μ_1 and μ_2 are the drift rates, σ_1 and σ_2 are the deterministic local volatilities, and W_1 and W_2 are the Brownian motions correlated by ρ.

Let $V(x, y, t)$ represent the value of an American put option on the two assets with expiry date T. If we define

$$
\begin{aligned}
LV = &-\frac{\partial V}{\partial t} - \frac{1}{2}\left[\sigma_1^2 x^2 \frac{\partial^2 V}{\partial x^2} + 2\rho\sigma_1\sigma_2 xy \frac{\partial^2 V}{\partial x \partial y} + \sigma_2^2 y^2 \frac{\partial^2 V}{\partial y^2}\right] \\
&- r\left[x\frac{\partial V}{\partial x} + y\frac{\partial V}{\partial y}\right] + rV,
\end{aligned}
\tag{2.1}
$$

with r being the risk free interest rate, then it is well known that V satisfies the following partial differential complementarity problem:

$$
\begin{cases}
LV \geq 0, \\
V - V^* \geq 0, \\
LV \cdot (V - V^*) = 0,
\end{cases}
\tag{2.2}
$$

for $(x, y, t) \in (0, X) \times (0, Y) \times ([0, T)$ with the boundary conditions

$$
\begin{aligned}
V(0, y, t) &= g_1(y, t), \quad V(x, 0, t) = g_2(x, t), \\
V(X, y, t) &= 0, \qquad\quad V(x, Y, t) = 0,
\end{aligned}
\tag{2.3}
$$

and terminal condition

$$V(x, y, t = T) = V^*(x, y),\tag{2.4}$$

where

$$V^*(x, y) = \max(K - w_1 x - w_2 y, 0)$$

is the payoff function, $K > 0$ is the striking price, $w_1, w_2 \geq 0$ are the weights of the assets x and y, respectively, and X, Y and T are given positive constants. We assume that $X \gg K$ and $Y \gg K$. Here, g_1 and g_2 are given functions that provide suitable boundary conditions.

Typically, g_1 and g_2 are determined via solving the associated one-dimensional American put option problems (cf. [12, 17]).

For convenience of theoretical analysis, we rewrite (2.1) as the following conservative form

$$LV = -V_t - \nabla \cdot \left(A\nabla V + \underline{b}V \right) + \overline{c}V, \tag{2.5}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sigma_1^2 x^2 & \frac{1}{2}\rho\sigma_1\sigma_2 xy \\ \frac{1}{2}\rho\sigma_1\sigma_2 xy & \frac{1}{2}\sigma_2^2 y^2 \end{pmatrix}, \tag{2.6a}$$

$$\underline{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} rx - \sigma_1^2 x - \frac{1}{2}\rho\sigma_1\sigma_2 x \\ ry - \sigma_2^2 y - \frac{1}{2}\rho\sigma_1\sigma_2 y \end{pmatrix}, \tag{2.6b}$$

$$\overline{c} = 3r - \left(\sigma_1^2 + \sigma_2^2 + \rho\sigma_1\sigma_2 \right). \tag{2.6c}$$

For discussion convenience, we transform (2.2)-(2.4) into an equivalent form satisfying homogeneous Dirichlet boundary conditions. Note that this transformation is needed only for the theoretical discussion, but not necessary in computations.

Let $V_0(x, y)$ be a twice differentiable function satisfying the boundary conditions in (2.3). We introduce a new function

$$u(x, y, t) = e^{\beta t}(V_0 - V), \tag{2.7}$$

where $\beta = \frac{1}{2}(\sigma_1^2 + \sigma_2^2 + \rho\sigma_1\sigma_2)$. Taking LV_0 away from both sides of the first inequality of (2.2) and transforming V in (2.2)-(2.4) into the new function u, we have

$$\begin{cases} \mathscr{L}u \leq f, \\ u - u^* \leq 0, \\ (\mathscr{L}u - f) \cdot (u - u^*) = 0, \end{cases} \tag{2.8}$$

where

$$\begin{aligned} \mathscr{L}u &= -u_t - \nabla \cdot (A\nabla u + \underline{b}u) + \underline{c}u, \\ \underline{c} &= \overline{c} + \beta, \quad u^* = e^{\beta t}(V_0 - V^*), \quad f(x, y, t) = e^{\beta t}LV_0. \end{aligned} \tag{2.9}$$

It is easy to see that under the transformation, the boundary and terminal conditions in (2.3)-(2.4) become, respectively,

$$u(0, y, t) = 0 = u(X, y, t), \quad \forall t \in [0, T] \text{ and } y \in [0, Y],$$
$$u(x, 0, t) = 0 = u(x, Y, t), \quad \forall t \in [0, T] \text{ and } x \in [0, X],$$

and $u(x, y, T) = u^*(x, y, T)$.

3. Reformulation of the problem

In this section, we will reformulate (2.8) as a variational inequality problem in an appropriate functional setting. Before proceeding, let us first introduce some standard notation to be used in the paper.

Let $\Omega = (0, X) \times (0, Y)$ and let Γ denote the boundaries of Ω. For $1 \le p \le \infty$, let $L^p(\Omega)$ denote the space of all p-integrable functions on Ω with the norm $\|\cdot\|_{L^p(\Omega)}$, and let $H^{m,p}(\Omega)$ denote the usual Sobolev space with the norm $\|\cdot\|_{(m,p,\Omega)}$. When $p = 2$, we simply use $H^m(\Omega)$ and $\|\cdot\|_{(m,\Omega)}$ to denote $H^{m,2}(\Omega)$ and $\|\cdot\|_{(m,2,\Omega)}$, respectively. We define the weighted Sobolev space $H^1_\varpi(\Omega)$ as follows:

$$H^1_\varpi(\Omega) = \{v : v, xv_x, yv_y \in L^2(\Omega)\}$$

with its norm denoted by $\|\cdot\|_{1,\varpi}$. We put

$$H^1_{0,\varpi}(\Omega) = \{v : v \in H^1_\varpi(\Omega),\ v|_\Gamma = 0\},$$
$$\mathcal{K}(t) = \left\{v(t) : v(t) \in H^1_{0,\varpi}(\Omega), v(t) \le u^*(t),\ \ a.e.\ \text{in } (0, T)\right\},$$

where $u^*(t)$ is defined by (2.9). It is easy to verify that $\mathcal{K}$ is a convex and closed subset of $H^1_{0,\varpi}(\Omega)$. Finally, for any Hilbert space $H(\Omega)$, the norm of $L^p(0, T; H(\Omega))$ is denoted by

$$\|v(\cdot, t)\|_{L^p(0,T;H(\Omega))} = \left(\int_0^T \|v(\cdot, \cdot, t)\|_H^p dt\right)^{1/p}.$$

Obviously, $L^p(0, T; L^p(\Omega)) = L^p(\Omega \times (0, T)) = L^p(\Theta)$. In what follows, we will simply write $v(t)$ when we regard $v(\cdot, \cdot, t)$ as an element of $H^1_{0,\varpi}(\Omega)$. We will also suppress the independent time variable t (or τ), when it causes no confusion in doing so.

Now, we define the following variational inequality problem.

Problem 3.1. *Find $u \in \mathcal{K}$ such that, for all $v \in \mathcal{K}$,*

$$\left(-\frac{\partial u}{\partial t}, v - u\right) + B(u, v - u; t) \ge (f, v - u),\ a.e.\ in\ (0, T), \tag{3.1}$$

where $B(u, v; t)$ is a bilinear form defined by

$$B(u, v; t) = (A\nabla u + \underline{b}u, \nabla v) + (\underline{c}u, v),\quad u, v \in H^1_{0,\varpi}(\Omega). \tag{3.2}$$

For this variational inequality problem, we have the following theorem.

Theorem 3.1. *Problem* 3.1 *is the variational form of the complementarity problem* (2.8).

The proof of Theorem 3.1 is standard (cf. [3, 10, 11]), we omit it here.

In order to establish the unique solvability of Problem 3.1, we study the properties of the operator $B(u, v; t)$. Let $|v|_{1,\varpi}$ be the functional on $H^1_{0,\varpi}(\Omega)$ define by

$$|v|^2_{1,\varpi} \triangleq \int_\Omega \left[x^2 v_x^2 + \rho\left(xv_x + yv_y\right)^2 + y^2 v_y^2\right] d\Omega, \tag{3.3}$$

for any $v \in H^1_{0,\varpi}(\Omega)$. It is easy to verify that $|\cdot|_{1,\varpi}$ is a weighted semi-norm on $H^1_{0,\varpi}(\Omega)$. Using this semi-norm, we define $\|v\|_{1,\varpi}$ by

$$\|v\|^2_{1,\varpi} \triangleq |v|^2_{1,\varpi} + \|v\|^2_{0,\varpi}.$$

It is easy to check $\|\cdot\|_{1,\varpi}$ is a weighted energy norm on $H^1_{0,\varpi}(\Omega)$. With these definitions and a stand argument (cf. [3]), we have the following lemma.

Lemma 3.1. *There exist positive constants C and M, independent of v and w, such that for any* $v, w \in H^1_{0,\varpi}(\Omega)$,

$$B(v, v; t) \geq C\|v\|^2_{1,\varpi},$$
$$|B(v, w; t)| \leq M\|v\|_{1,\varpi}\|w\|_{1,\varpi}.$$

Using Lemma 3.1 and the theory for abstract variational inequality problems in [3, 11], we establish the unique solvability of Problem 3.1 in the following theorem.

Theorem 3.2. *There exists a unique solution to Problem* 3.1.

4. The power penalty approach

To derive the power penalty approach, we first consider the following nonlinear variational inequality problem:

Find $u_\lambda \in H^1_{0,\varpi}(\Omega)$ such that, for all $v \in H^1_{0,\varpi}(\Omega)$,

$$\left(-\frac{\partial u_\lambda}{\partial t}, v - u_\lambda\right) + B(u_\lambda, v - u_\lambda; t) + j(v) - j(u_\lambda) \geq (f, v - u_\lambda), \quad \text{a.e. in } (0, T), \quad (4.1)$$

where

$$j(v) = \frac{\lambda k}{k+1}\left[v - u^*\right]^{\frac{k+1}{k}}_+, \quad k > 0, \quad \lambda > 1, \tag{4.2}$$

and $[z]_+ = \max\{0, z\}$ for any z.

The unique solvability of this problem is guaranteed by the coerciveness and continuity of the bilinear operator B and the lower semi-continuity of j, see [10]. It is easy to show by virtue of Lemma 3.1 and (4.2), that all the required conditions are satisfied by the bilinear form B. Hence, (4.1) is uniquely solvable.

From (4.2), we can see that $j(v)$ is differentiable. Thus, (4.1) is equivalent to the following problem.

Problem 4.1. *Find* $u_\lambda \in H^1_{0,\varpi}(\Omega)$ *such that, for all* $v \in H^1_{0,\varpi}(\Omega)$,

$$\left(-\frac{\partial u_\lambda}{\partial t}, v\right) + B(u_\lambda, v; t) + (j'(u_\lambda), v) = (f, v), \quad \text{a.e. in } (0, T), \tag{4.3}$$

where

$$j'(v) = \lambda[v - u^*]^{1/k}_+. \tag{4.4}$$

We remark that (4.1)-(4.4) is a penalized variational equation corresponding to (3.1). The strong form of (4.1)-(4.4), which defines the penalized equation approximating (2.8), is given by

$$\mathscr{L}u_\lambda + \lambda[u_\lambda - u^*]_+^{1/k} = f, \quad (x, y, t) \in \Theta, \tag{4.5}$$

with the given boundary and final conditions

$$u_\lambda(x, y, t)|_\Gamma = 0 \quad \text{and} \quad u_\lambda(x, y, T) = u^*(x, y, T). \tag{4.6}$$

If $k = \frac{1}{2}$, this penalty approach corresponds to the quadratic penalty approach. While $k = 1$, the typical l_1 penalty approach is obtained. When $k > 1$, it is the so-called lower order penalty approach [16, 19]. In the next section, we will investigate the convergence rates of u_λ to u as $\lambda \to \infty$.

5. Convergence analysis

We now show that, as $\lambda \to \infty$, the solution to Problem 4.1 converges to that of Problem 3.1 at the rate of order $\mathcal{O}(\lambda^{-k/2})$ in a proper norm. We start this discussion by the following Lemma.

Lemma 5.1. *Let u_λ be the solution to Problem* 4.1. *If $u_\lambda \in L^p(\Theta)$, then there exists a positive constant C, independent of u_λ and λ, such that*

$$\|[u_\lambda - u^*]_+\|_{L^p(\Theta)} \le \frac{C}{\lambda^k}, \tag{5.1a}$$

$$\|[u_\lambda - u^*]_+\|_{L^\infty(0,T;L^2(\Omega))} + \|[u_\lambda - u^*]_+\|_{L^2(0,T;H_{0,\varpi}^1(\Omega))} \le \frac{C}{\lambda^{k/2}}, \tag{5.1b}$$

where k is the power of the power penalty function and $p = 1 + 1/k$.

Proof. Assume that C is a generic positive constant, independent of u_λ and λ. To simplify the notation, we let $\phi = [u_\lambda - u^*]_+$. Obviously, $\phi \in H_{0,\varpi}^1(\Omega)$ a.e. in $(0, T)$.

Setting $v = \phi$ in (4.3) and (4.4), we have

$$\left(-\frac{\partial u_\lambda}{\partial t}, \phi\right) + B(u_\lambda, \phi; t) + \lambda(\phi^{1/k}, \phi) = (f, \phi), \quad \text{a.e. in } (0, T), \tag{5.2a}$$

$$\left(-\frac{\partial(u_\lambda - u^*)}{\partial t}, \phi\right) + B((u_\lambda - u^*), \phi; t) + \lambda(\phi^{1/k}, \phi)$$
$$= (f, \phi) + \left(\frac{\partial u^*}{\partial t}, \phi\right) - B(u^*, \phi; t). \tag{5.2b}$$

Integrating both sides of (**??**) from t to T and using the coerciveness property of the

operator B and Holder's inequality, we get

$$\frac{1}{2}(\phi(t), \phi(t)) + \int_t^T \|\phi(\tau)\|_B^2 d\tau + \lambda \int_t^T (\phi^{1/k}, \phi) d\tau$$

$$\leq \int_t^T (f(\tau), \phi(\tau)) d\tau + \beta \int_t^T e^{\beta\tau}(V_0 - V^*, \phi(\tau)) d\tau - \int_t^T B(u^*, \phi(\tau); \tau) d\tau$$

$$\leq C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p} + \beta \int_t^T e^{\beta\tau}(V_0 - V^*, \phi(\tau)) d\tau - \int_t^T B(u^*, \phi(\tau); \tau) d\tau.$$

$$(5.3)$$

Noting that $|V_0 - V^*|$ is uniformly bounded and $\beta = \sigma_1^2 + \sigma_2^2 + \frac{1}{2}\rho\sigma_1\sigma_2$, we have

$$\frac{1}{2}(\phi(t), \phi(t)) + \int_t^T \|\phi(\tau)\|_B^2 d\tau + \lambda \int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau$$

$$\leq C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p} - \int_t^T B(u^*, \phi; \tau) d\tau. \tag{5.4}$$

Since $B(u, v; t) = (A\nabla u + \underline{b}u, \nabla v) + (\underline{c}u, v)$, it follows that

$$-\int_t^T B(u^*, \phi(\tau); \tau) d\tau = -\int_t^T (A\nabla u^* + \underline{b}u^*, \nabla\phi(\tau)) d\tau - \int_t^T (\underline{c}u^*, \phi(\tau)) d\tau. \tag{5.5}$$

Furthermore, by Green's theorem, we obtain

$$-\int_t^T \left(\underline{b}u^*, \nabla\phi(\tau) \right) d\tau = \int_t^T \int_\Omega \nabla \cdot \underline{b}u^*\phi(\tau) d\Omega d\tau - \int_t^T \int_\Gamma u^* \cdot n\phi(\tau) d\Gamma d\tau. \tag{5.6}$$

Let $\Omega_1 = \{0 < x < K/w_1, 0 < y < K/w_2, K - w_1 x - w_2 y > 0\}$ and $\Omega_2 = \Omega \setminus \overline{\Omega}_1$. We also let Γ_0 denote the interface of Ω_1 and Ω_2. Therefore, Γ_0 has two opposite orientations: Γ_0^+ when it is orientated in the same direction as $\partial\Omega_1$, and Γ_0^- when it is orientated in the same direction as $\partial\Omega_2$. Consider only the integrand $(A\nabla u^*, \nabla\phi)$ in (5.5). For $\phi \in H_{0,\varpi}^1(\Omega)$, note that $\phi = 0$ on Γ, we have

$$- (A\nabla u^*, \nabla\phi)$$

$$= -\int_\Omega (A\nabla u^*)^T \nabla\phi \, d\Omega = -\int_{\Omega_1} (A\nabla u^*)^T \nabla\phi \, d\Omega - \int_{\Omega_2} (A\nabla u^*)^T \nabla\phi \, d\Omega$$

$$= -\int_{\Gamma_0^+} A\nabla u^* \cdot n\phi \, ds + \int_{\Omega_1} \nabla \cdot (A\nabla u^*) \phi \, d\Omega - \int_{\Gamma_0^-} A\nabla u^* \cdot n\phi \, ds + \int_{\Omega_2} \nabla \cdot (A\nabla u^*) \phi \, d\Omega$$

$$= -\int_{\Gamma_0^+} \left(A\nabla u_-^* - A\nabla u_+^* \right) \cdot n\phi \, ds + \int_\Omega \nabla \cdot (A\nabla u^*) \phi \, d\Omega, \tag{5.7}$$

where n denotes the unit outward normal direction of the boundary segments and ∇u_-^* and ∇u_+^* denote, respectively, the values of ∇u^* evaluated on the left and right sides of Γ_0^+. From $u^* = e^{\beta t}(V_0 - V^*)$ and (2.6), it is easy to see that

$$\nabla u^* = e^{\beta t}(\nabla V_0 - \nabla V^*).$$

Since $V_0 \in H^2(\Omega)$, ∇V_0 is continuous on Ω, as mentioned before,

$$\nabla u_-^* - \nabla u_+^* = e^{\beta t}[(\nabla V_0 - \nabla V^*)_- - (\nabla V_0 - \nabla V^*)_+]$$
$$= e^{\beta t}(\nabla V_-^* - \nabla V_+^*) = e^{\beta t}(-w_1, -w_2)^T.$$

Furthermore, the unit outward-normal vector to Γ_0^+ is

$$n = \frac{\nabla(K - w_1 x - w_2 y)}{\|\nabla(K - w_1 x - w_2 y)\|} = \frac{(-w_1, -w_2)^T}{\sqrt{w_1^2 + w_2^2}}.$$

Therefore, estimate (5.7) becomes

$$-(A\nabla u^*, \nabla \phi) = -\int_{\Gamma_0^+} e^{\beta t} \frac{(w_1, w_2) A^T (w_1, w_2)^T}{(w_1^2 + w_2^2)^{1/2}} \phi \, ds + \int_\Omega \nabla \cdot (A\nabla u^*) \phi \, d\Omega \le C \int_\Omega \phi(\tau) d\Omega,$$

as A is positive definite, ϕ is non-negative and $\nabla \cdot (A\nabla u^*)$ is bounded above on Ω. Thus,

$$-\int_t^T (A\nabla u, \nabla \phi(\tau)) d\tau \le C \int_t^T \int_\Omega \phi(\tau) d\Omega d\tau \le C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p}. \quad (5.8)$$

Also, from (5.6), it follows that

$$-\int_t^T \left(\underline{b} u^*, \nabla \phi(\tau) \right) d\tau \le C \int_t^T \int_\Omega \phi(\tau) d\Omega d\tau \le C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p}, \quad (5.9)$$

because $\nabla \cdot \underline{b} u^*$ is bounded above on Ω.

Thus, from (5.3) to (5.9), it follows that

$$\frac{1}{2}(\phi(t), \phi(t)) + \int_t^T \|\phi(\tau)\|_B^2 d\tau + \lambda \int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau$$
$$\le C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p}, \quad (5.10)$$

which implies that

$$\lambda \int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \le C \left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau \right)^{1/p}, \quad \text{a.e. in } (0, T).$$

From this, it follows that

$$\left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau\right)^{1/p} \le \frac{C}{\lambda^{1/(p-1)}} = \frac{C}{\lambda^k}, \quad \text{where } p = 1 + 1/k. \tag{5.11}$$

Now, from (5.10) and (5.11), we have

$$\frac{1}{2}\left(\phi(t),\phi(t)\right) + \int_t^T \|\phi(\tau)\|_B^2 d\tau \le C\left(\int_t^T \|\phi(\tau)\|_{L^p(\Omega)}^p d\tau\right)^{1/p} \le \frac{C}{\lambda^k},$$

from which, it follows that

$$\left(\phi(t),\phi(t)\right)^{\frac{1}{2}} + \left(\int_t^T \|\phi(\tau)\|_B^2 d\tau\right)^{\frac{1}{2}} \le \frac{C}{\lambda^{k/2}}, \quad \text{a.e. in } (0,T).$$

Clearly, by replacing ϕ with $[u_\lambda - u^*]_+$, we obtain readily (5.1). $\qquad\square$

Using Lemma 5.1, we are able to show that the solution to Problem 4.1 converges to that of Problem 3.1 at the rate of order $\lambda^{-k/2}$, as stated in the next theorem.

Theorem 5.1. *Let u and u_λ be the solutions to Problem 3.1 and Problem 4.1, respectively. If $u_\lambda \in L^p(\Theta)$ and $\frac{\partial u}{\partial t} \in L^{k+1}(\Theta)$, then there exists a positive constant C, independent of u_λ and λ, such that*

$$\left\|u - u_\lambda\right\|_{L^\infty(0,T;L^2(\Omega))} + \left\|u - u_\lambda\right\|_{L^2(0,T;H_{0,\varpi}^1(\Omega))} \le \frac{C}{\lambda^{k/2}}, \tag{5.12}$$

where k is the power of the power penalty function.

Proof. We still use the notation of Lemma 5.1. Setting $v_- = -\min(v,0)$ and $R_\lambda = u - u^* + [u_\lambda - u^*]_-$, it follows that

$$\begin{aligned} &u - u_\lambda = R_\lambda - \varphi, \\ &(\varphi^\alpha, [u_\lambda - u^*]_-) = [u_\lambda - u^*]_+^\alpha [u_\lambda - u^*]_- \equiv 0, \quad \alpha > 0. \end{aligned} \tag{5.13}$$

Set $v = u - R_\lambda$ in (3.1) and $v = R_\lambda$ in (4.3) respectively, we obtain

$$\left(-\frac{\partial u}{\partial t}, -R_\lambda\right) + B\left(u, -R_\lambda; t\right) \ge \left(f(t), -R_\lambda\right), \tag{5.14}$$

$$\left(-\frac{\partial u_\lambda}{\partial t}, R_\lambda\right) + B\left(u_\lambda, R_\lambda; t\right) + \lambda(\phi^{1/k}, R_\lambda) = (f(t), R_\lambda). \tag{5.15}$$

Combining (5.14) and (5.15) gives

$$\left(-\frac{\partial(u_\lambda - u)}{\partial t}, R_\lambda\right) + B\left(u_\lambda - u, R_\lambda; t\right) + \lambda(\phi^{1/k}, R_\lambda) \ge 0.$$

It follows from $u \leq u^*$ and $\phi \geq 0$ that

$$(\phi^{1/k}, R_\lambda) = (\phi^{1/k}, u - u^*) + (\phi^{1/k}, [u_\lambda - u^*]_-) = (\phi^{1/k}, u - u^*) \leq 0.$$

Therefore,

$$\left(-\frac{\partial(u - u_\lambda)}{\partial t}, R_\lambda \right) + B(u - u_\lambda, R_\lambda; t) \leq 0.$$

From (5.13), it follows that

$$\left(-\frac{\partial R_\lambda}{\partial t}, R_\lambda \right) + B(R_\lambda, R_\lambda; t) \leq \left(-\frac{\partial \phi(t)}{\partial t}, R_\lambda \right) + B\left(\phi, R_\lambda; t \right).$$

Integrating both sides of the above from $\tau = t$ to $\tau = T$ and then using Cauchy-Schwarz inequality and $(\varphi, [u_\lambda - u^*]_-) = 0$, we obtain

$$\frac{1}{2} \left(R_\lambda(t), R_\lambda(t) \right) + \int_t^T B\left(R_\lambda, R_\lambda; \tau \right) d\tau$$

$$\leq \int_t^T \left(-\frac{\partial \phi}{\partial \tau}, R_\lambda \right) d\tau + \int_t^T B\left(\phi, R_\lambda; \tau \right) d\tau$$

$$\leq (\phi, R_\lambda) + \int_t^T \left(\phi, \frac{\partial R_\lambda}{\partial \tau} \right) d\tau + \int_t^T B(\phi, R_\lambda; \tau) d\tau$$

$$\leq (\phi, R_\lambda) + \int_t^T B\left(\phi, R_\lambda; \tau \right) d\tau + \int_t^T \left(\phi, \frac{\partial u}{\partial \tau} \right) d\tau$$

$$\leq ||\phi||_{L^\infty(0,T;L^2(\Omega))} ||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + C||\phi||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))}$$

$$+ C||\phi||_{L^p(\Theta)} \left(\left\| \frac{\partial u}{\partial t} \right\|_{L^q(\Theta)} + ||V_0 - V^*||_{L^q(\Theta)} \right) \tag{5.16}$$

where $p = 1 + 1/k$, and $1/p + 1/q = 1$.

Since $u_\lambda \in L^p(\Theta)$, and $\frac{\partial u}{\partial t} \in L^{k+1}(\Theta)$, it follows from (5.1) that

$$||\phi||_{L^p(\Theta)} \left(\left\| \frac{\partial u}{\partial t} \right\|_{L^q(\Theta)} + ||V_0 - V^*||_{L^q(\Theta)} \right) \leq \frac{C}{\lambda^k}. \tag{5.17}$$

Using the coerciveness property of the operator B, we have

$$\frac{1}{2} \left(R_\lambda(t), R_\lambda(t) \right) + \int_t^T B\left(R_\lambda, R_\lambda; \tau \right) d\tau$$

$$\geq \frac{1}{2} ||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + C||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))}. \tag{5.18}$$

Therefore, from (5.16)-(5.18), we have

$$
\left(||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right)^2
$$

$$
\leq C \left(\tfrac{1}{2} ||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right)
$$

$$
\leq C \left[\left(||\phi||_{L^\infty(0,T;L^2(\Omega))} + ||\phi||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right) \left(||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} \right. \right.
$$

$$
\left. + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right) + \lambda^{\frac{1}{k}} \Big]
$$

$$
\leq C \left[\lambda^{-k/2} \left(||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right) + \lambda^{-k} \right].
$$

Clearly, the above inequalities imply that

$$
||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \leq \frac{C}{\lambda^{k/2}}.
$$

Using the triangle inequality and (5.1), also noting that $u - u_\lambda = R_\lambda - \varphi$, we finally have

$$
||u - u_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||u - u_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))}
$$

$$
\leq \left(||R_\lambda||_{L^\infty(0,T;L^2(\Omega))} + ||R_\lambda||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right) + \left(||\phi||_{L^\infty(0,T;L^2(\Omega))} + ||\phi||_{L^2(0,T;H^1_{0,\varpi}(\Omega))} \right)
$$

$$
\leq \frac{C}{\lambda^{k/2}}.
$$

This completes the proof of the theorem. $\square$

Remark 5.1. It is worth noting that the analysis of Sections 2-5 is not restricted to the two-asset option pricing problem. Essentially, there is no intrinsic difficulty in the analysis for treating higher dimensional problems.

6. Discretization

The power penalty approach to the complementarity problem (2.8) yields a nonlinear parabolic partial differential equation (4.5). In this section, we will present the fitted finite volume method for (4.5). This method was first proposed in [18] and has been used for solving single-asset and stochastic volatility option pricing problem [12, 19]. The idea of this method is based on a finite volume formulation coupled with a fitted approximation technique. This fitting technique is to approximate the flux of a given function locally by a constant, yielding a locally nonlinear approximation to the function. Some results on error estimation for this method can be found in [1]. In what follows, we will give a brief account for the method applied to (4.5). More details can be found in [12] for a different, but related problem.

It is easy to show that, under the inverse of the transformation (2.7), (4.5)–(4.6) can be rewritten as equation

$$
-V_t - \nabla \cdot \left(A\nabla V + \underline{b}V \right) + \overline{c}V - \lambda \left[V^* - V \right]_+^{1/k} = 0, \quad (x,y,t) \in \Theta, \qquad (6.1)
$$

with the boundary and final conditions

$$V(x, y, t)|_\Gamma = 0 \quad \text{and} \quad V(x, y, T) = V^*(x, y).$$

To discretize this equation, we first partition $I_x := (0, X)$ into N_x sub-intervals $I_{x_i} := (x_i, x_{i+1})$, $i = 0, \cdots, N_x - 1$, with $0 = x_0 < x_1 < \cdots < x_{N_x} = X$. We also let

$$x_{i-1/2} = \frac{1}{2}(x_{i-1} + x_i) \quad \text{and} \quad x_{i+1/2} = \frac{1}{2}(x_i + x_{i+1}),$$

for each $i = 1, 2, \cdots, N_x - 1$. These mid-points form a second partition of $(0, X)$ if we define $x_{-1/2} = x_0$ and $x_{N_x+1/2} = x_{N_x}$. For each $i = 0, 1, \cdots, N_x - 1$, we put $h_{x_i} = x_{i+1/2} - x_{i-1/2}$.

Similar to the above, we define a partition on $I_y = (0, Y)$. Combining these partitions we obtain a mesh for $\Omega := I_x \times I_y$. Also, the mid-points form a second partition of Ω, which is called *boxes* which is a dual mesh to the original partition.

6.1. Boundary conditions

Before deriving the fitted finite volume method for (6.1), we first show how to determine the boundary condition functions $g_1(y, t)$ and $g_2(x, t)$ for a put option, as given below.

1. On the boundary $x = 0$, the boundary condition $g_1(y, t) = V(y, t)/w_2$ is determined by solving

$$\begin{cases} -\frac{\partial V(y,t)}{\partial t} - \frac{1}{2}\sigma_2^2 y^2 \frac{\partial^2 V(y,t)}{\partial y^2} - ry\frac{\partial V(y,t)}{\partial y} + rV(y,t) - \lambda[V^* - V]_+^{1/k} = 0, \\ V(0,t) = K/w_2, \quad V(Y,t) = 0, \\ V(y,T) = V^*(0,y) = \max(K/w_2 - y, 0). \end{cases} \tag{6.2}$$

2. On the boundary $y = 0$, the boundary condition $g_2(x, t) = V(x, t)/w_1$ is given by the solution of

$$\begin{cases} -\frac{\partial V(x,t)}{\partial t} - \frac{1}{2}\sigma_1^2 x^2 \frac{\partial^2 V(x,t)}{\partial x^2} - rx\frac{\partial V(x,t)}{\partial x} + rV(x,t) - \lambda[V^* - V]_+^{1/k} = 0, \\ V(0,t) = K/w_1, \quad V(X,t) = 0, \\ V(x,T) = V^*(x,0) = \max(K/w_1 - x, 0). \end{cases} \tag{6.3}$$

Both of the above cases fall into the framework of a single-asset American option problem which can be discretized by the scheme in [18].

6.2. The fitted finite volume method

As mentioned before, a two-dimensional version of the fitted finite volume method is developed in [12] for a linear equation. The difference between (4.5) and the equation in [12] is the nonlinear term on the left-hand size of (4.5). Since the discretization of this

nonlinear term is simple, we omit the lengthy discussion of the scheme and only present the resulting discretized form. Further details of the discretization can be found in [12].

Let $V_{i,j} = V(x_i, y_j, t)$, $V_{i,j}^* = V^*(x_i, y_j)$ and $R_{i,j} = (x_{i+1/2} - x_{i-1/2}) \times (y_{j+1/2} - y_{j-1/2})$, for all admissible (i, j). The application of the fitted finite volume method (6.1) yields

$$-\frac{\partial V_{i,j}}{\partial t} R_{i,j} + e_{i-1,j}^{i,j} V_{i-1,j} + e_{i,j-1}^{i,j} V_{i,j-1} + e_{i,j}^{i,j} V_{i,j} + e_{i,j+1}^{i,j} V_{i,j+1}$$

$$+ e_{i+1,j}^{i,j} V_{i+1,j} - \lambda [V_{i,j}^* - V_{i,j}]_+^{\frac{1}{k}} R_{ij} = 0, \tag{6.4}$$

for $i = 1, \cdots, N_x - 1$ and $j = 1, \cdots, N_y - 1$. Here

$$e_{i-1,j}^{i,j} = -h_{y_j} \frac{b_{i-1/2,j} x_{i-1/2} x_i^{\alpha_{i-1,j}}}{x_i^{\alpha_{i-1,j}} - x_{i-1}^{\alpha_{i-1,j}}}, \quad e_{i,j-1}^{i,j} = -h_{x_i} \frac{\overline{b}_{i,j-1/2} y_{j-1/2} y_{j-1}^{\overline{\alpha}_{i,j-1}}}{y_j^{\overline{\alpha}_{i,j-1}} - y_{j-1}^{\overline{\alpha}_{i,j-1}}}, \tag{6.5a}$$

$$e_{i,j}^{i,j} = h_{y_j} \left(\frac{b_{i+1/2,j} x_{i+1/2} x_i^{\alpha_{i,j}}}{x_{i+1}^{\alpha_{i,j}} - x_i^{\alpha_{i,j}}} + \frac{b_{i-1/2,j} x_{i-1/2} x_i^{\alpha_{i-1,j}}}{x_i^{\alpha_{i-1,j}} - x_{i-1}^{\alpha_{i-1,j}}} + \overline{d}_{i,j} \right)$$

$$+ h_{x_i} \left(\frac{\overline{b}_{i,j+1/2} y_{j+1/2} y_j^{\overline{\alpha}_{i,j}}}{y_{j+1}^{\overline{\alpha}_{i,j}} - y_j^{\overline{\alpha}_{i,j}}} + \frac{\overline{b}_{i,j-1/2} y_{j-1/2} y_j^{\overline{\alpha}_{i,j-1}}}{y_j^{\overline{\alpha}_{i,j-1}} - y_{j-1}^{\overline{\alpha}_{i,j-1}}} + d_{i,j} \right) + \overline{c}_{i,j} R_{i,j}, \tag{6.5b}$$

$$e_{i,j+1}^{i,j} = -h_{x_i} \left(\frac{\overline{b}_{i,j+1/2} y_{j+1/2} y_{j+1}^{\overline{\alpha}_{i,j}}}{y_{j+1}^{\overline{\alpha}_{i,j}} - y_j^{\overline{\alpha}_{i,j}}} + d_{i,j} \right), \quad e_{i+1,j}^{i,j} = -h_{y_j} \left(\frac{b_{i+1/2,j} x_{i+1/2} x_{i+1}^{\alpha_{i,j}}}{x_{i+1}^{\alpha_{i,j}} - x_i^{\alpha_{i,j}}} + \overline{d}_{i,j} \right). \tag{6.5c}$$

for $i = 1 \cdots N_x - 1$, $j = 1 \cdots N_y - 1$ and $e_{m,n}^{i,j} = 0$ if $m \neq i - 1, i, i + 1$ and $n \neq j - 1, j, j + 1$, where

$$b = r - \sigma_1^2 - \tfrac{1}{2}\rho\sigma_1\sigma_2, \quad a = \tfrac{1}{2}\sigma_1^2, \qquad d = \tfrac{1}{2}\rho\sigma_1\sigma_2 y,$$
$$b_{i,j+1/2} = b(x_i, y_{j+1/2}), \quad \alpha_{i,j} = b_{i,j+1/2}/a, \quad d_{i,j} = d(x_i, y_j),$$

and

$$\overline{b} = r - \sigma_2^2 - \tfrac{1}{2}\rho\sigma_1\sigma_2, \quad \overline{a} = \tfrac{1}{2}\sigma_2^2, \qquad \overline{d} = \tfrac{1}{2}\rho\sigma_1\sigma_2 x,$$
$$\overline{b}_{i,j+1/2} = \overline{b}(x_i, y_{j+1/2}), \quad \overline{\alpha}_{i,j} = \overline{b}_{i,j+1/2}/\overline{a}, \quad \overline{d}_{i,j} = \overline{d}(x_i, y_j).$$

Defining

$$E_{i,j} = (0, \cdots, 0, e_{i-1,j}^{i,j}, 0, \cdots, 0, e_{i,j-1}^{i,j}, e_{i,j}^{i,j}, e_{i,j+1}^{i,j}, 0, \cdots, 0, e_{i+1,j}^{i,j}, 0, \cdots, 0)$$

for $i = 2, \cdots, N_x - 1$, $j = 2, \cdots, N_y - 1$, and

$$\mathbf{V} = (V_{1,1}, \cdots, V_{1,N_y-1}, V_{2,1}, \cdots, V_{2,N_y-1}, \cdots, V_{N_x-1,1}, \cdots, V_{N_x-1,N_y-1})^T$$

with $V_{i,0}$, $i = 1, \cdots, N_x$ and $V_{0,j}$, $j = 1, \cdots, N_y$ in (6.4) being equal to the given boundary conditions, we can rewrite (6.4) as:

$$-\frac{\partial V_{i,j}}{\partial t} R_{i,j} + E_{i,j}\mathbf{V} + p\left(V_{i,j}\right) = 0, \tag{6.6}$$

where

$$p\left(V_{i,j}\right) = -\lambda R_{i,j}[V_{i,j}^* - V_{i,j}]_+^{1/k}. \tag{6.7}$$

This is a system of $(N_x - 1)^2 \times (N_y - 1)^2$ linear ordinary differential equations for $(N_x - 1) \times (N_y - 1)$ unknown values.

Now, we discretize the time by choosing a set of points t_i $(i = 0, 1, \cdots M)$ be a set of partition points on $[0, T]$ satisfying $T = t_0 > t_1, \cdots, > t_M = 0$. Apply the two-level implicit time-stepping method with a splitting parameter $\theta \in [0, 1/2]$ to (6.6) on this mesh, we get the following full discrete system

$$\left(\theta E^{m+1} + G^m\right)\mathbf{V}^{m+1} + \theta D(\mathbf{V}^{m+1}) = (G^m - (1-\theta)E^m)\mathbf{V}^m - (1-\theta)D(\mathbf{V}^m), \tag{6.8}$$

where

$$\mathbf{V}^m = (V_{1,1}^m, \cdots, V_{1,N_y-1}^m, V_{2,1}^m, \cdots, V_{2,N_y-1}^m, \cdots, V_{N_x-1,1}^m, \cdots, V_{N_x-1,N_y-1}^m)^T, \tag{6.9a}$$

$$E^m = (E_{1,1}^m, \cdots, E_{1,N_y-1}^m, E_{2,1}^m, \cdots, E_{2,N_y-1}^m, \cdots, E_{N_x-1,1}^m, \cdots, E_{N_x-1,N_y-1}^m)^T, \tag{6.9b}$$

$$G^m = diag(-R_{1,1}/(\triangle t_m), \cdots, -R_{N_x-1,N_y-1}/(\triangle t_m)), \tag{6.9c}$$

$$D(\mathbf{V}^m) = (p(V_{1,1}^m), \cdots, p(V_{N_x-1,N_y-1}^m))^T, \tag{6.9d}$$

for $m = 0, 1, \cdots, m-1$, where $\triangle t_m = t_{m+1} - t_m < 0$, where $\mathbf{V}^m$ denotes the approximation of $\mathbf{V}$ at $t = t_m$ and $E_{i,j}^m = E_{i,j}(t_m)$.

6.3. Solution of the discrete system

To solve the nonlinear discrete system (6.8), we use the standard Newton method. Note that when $k > 1$, from (6.7) it is easy to see that $p'(V_{i,j}^m) \to \infty$ as $V_{i,j}^* - V_{i,j} \to 0^+$. To remedy this, we use the technique proposed in [19] to smooth (6.7), yielding the following approximation to $p(V_{i,j}^m)$:

$$\frac{p(V_{i,j}^m)}{-\lambda R_{i,j}} \tag{6.10}$$

$$= \begin{cases} [V_{i,j}^* - V_{i,j}^m]_+^{1/k}, & V_{i,j}^* - V_{i,j}^m \geq \varepsilon, \\ \varepsilon^{\frac{1}{k}-n+1}(n - \frac{1}{k})[V_{i,j}^* - V_{i,j}^m]_+^{n-1} + \varepsilon^{\frac{1}{k}-n}(\frac{1}{k} - n + 1)[V_{i,j}^* - V_{i,j}^m]_+^n, & V_{i,j}^* - V_{i,j}^m < \varepsilon, \end{cases}$$

for $k > 0$ and positive integer n, where $1 \gg \varepsilon > 0$ is a transition parameter. It has been shown in Corollary 5.1 of [19] that when $n \geq 3$ and $k \geq 1/n$, (6.10) is smooth and increasing on $(-\infty, +\infty)$.

Applying Newton's method to (6.8) gives

$$\left[\theta E^{m+1} + G^m + \theta J_D(\varpi^{l-1})\right]\delta\varpi^l$$
$$= [G^m - (1-\theta)E^m]\mathbf{V}^m - (1-\theta)D(\mathbf{V}^m) - (\theta E^{m+1} + G^m)\varpi^l - \theta D(\varpi^{l-1}), \tag{6.11a}$$
$$\varpi^l = \varpi^{l-1} + \gamma \cdot \delta\varpi^l \tag{6.11b}$$

for $l = 1, 2, \cdots$, 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. We then choose

$$\mathbf{V}^{m+1} = \lim_{l \to \infty} \varpi^l.$$

It is easy to show that the system matrix of (6.11) is an M-matrix, as given in the following theorem.

Theorem 6.1. *For any given $m = 1, 2, \cdots, M - 1$, if $|\triangle t_m|$ is sufficiently small and $\overline{c} \geq 0$, then the system matrix of (6.11) is an M-matrix.*

Proof. From the definition of $D(\mathbf{V})$ in (6.9), it is easy to see that its Jacobian is the following diagonal matrix

$$J_D(\varpi^l) = \text{diag}\left(p'(V_{1,1}^m), \cdots, p'(V_{N_x-1,N_y-1}^m)\right).$$

From (6.10), we have $p'(V_{i,j}^m) \geq 0$ for all $i = 1, \cdots, N_x - 1$ and $j = 1, \cdots, N_y - 1$. Thus, to show that the system matrix of (6.11) is an M-matrix, it suffices to show that $\theta E^{m+1} + G^m$ is an M-matrix.

First, we note that $e_{m,n}^{i,j} \leq 0$ for all $m \neq i$, $n \neq j$ since

$$\frac{b_{i+1/2,j}}{x_{i+1}^{\alpha_{i,j}} - x_i^{\alpha_{i,j}}} = \frac{a\,\alpha_{i,j}}{x_{i+1}^{\alpha_{i,j}} - x_i^{\alpha_{i,j}}} > 0, \qquad \frac{\overline{b}_{i,j+1/2}}{y_{j+1}^{\overline{\alpha}_{i,j}} - y_j^{\overline{\alpha}_{i,j}}} = \frac{\overline{a}\,\overline{\alpha}_{i,j}}{y_{j+1}^{\overline{\alpha}_{i,j}} - y_j^{\overline{\alpha}_{i,j}}} > 0, \qquad (6.12)$$

for all $i = 1, \cdots, N_x - 1$, $j = 1, \cdots, N_y - 1$ and all $b_{i+1/2,j} \neq 0$, $\overline{b}_{i,j+1/2} \neq 0$. (6.12) also holds when $b_{i+1/2,j} \to 0$, $\overline{b}_{i,j+1/2} \to 0$. Furthermore, from (6.5), it follows that, when $c_{i,j} \geq 0$, for all $i = 1, \cdots, N_x - 1$, $j = 1, \cdots, N_y - 1$,

$$\left(e_{i,j}^{i,j}\right)^{m+1} \geq \left|(e_{i-1,j}^{i,j})^{m+1}\right| + \left|(e_{i,j-1}^{i,j})^{m+1}\right| + \left|(e_{i,j+1}^{i,j})^{m+1}\right| + \left|(e_{i+1,j}^{i,j})^{m+1}\right| + c_{i,j}^{m+1}R_{i,j}$$

$$= \sum_{p=1}^{N_x-1}\sum_{q=1}^{N_y-1} \left|(e_{p,q}^{i,j})^{m+1}\right| + \underline{c}_{i,j}^{m+1}R_{i,j},$$

since d and $\overline{d}$ are all non-negative. Therefore, E^{m+1} is diagonally dominant with respect to its columns. Hence, form the above analysis, we see that for all admissible i, j, E^{m+1} is a diagonally dominant matrix with positive diagonal elements and non-positive off-diagonal elements. This implies that E^{m+1} is an M-matrix.

For the second part, we first note that G^m of the system matrix (6.11) is a diagonal matrix with positive diagonal entries. When $|\triangle t_m|$ is sufficiently small, we also have

$$\theta \underline{c}_{i,j} + \frac{R_{i,j}}{-\triangle t_m} > 0.$$

So, $\theta E^{m+1} + G^m$ is an M-matrix. $\qquad\square$

Theorem 6.1 implies that the fully discrete system (6.8) satisfies the discrete maximum principle and the discretization is monotone.

7. Numerical experiments

In this section, we demonstrate the efficiency and usefulness of the above penalty and numerical methods by solving the following test model problem with different values of k and λ. Throughout this section, the model parameters in Table 1 are used.

Table 1: Data used for two-asset model.

Parameter values			
r	0.1	ω_1	1.0
σ_1	0.2	ω_2	1.0
σ_2	0.2	K	1.0
ρ	0	T	0.25

In order to perform simulations, we choose an upper limit for the solution domain, that is a domain of which option values outside are regarded worthless. For our model, we choose $X = Y = 4$.

To compare the numerical performance of different power penalty methods, we consider three power penalty approaches: $k = 1/2, 1$ and 2. When $k = 1/2$, the nonlinear system (6.8) is smooth, for which the classical Newton's method is used. When $k = 1$, the nonlinear system (6.8) is semismooth. In this case we use the semismooth Newton's method [15]. However, for $k = 2$, this nonlinear system becomes nonsmooth. Hence, the smoothing technique (6.10) is adopted and the Newton's method is used for the resulting system.

For time discretization, we choose the time step splitting parameter $\theta = 1$, i.e. the Backward Euler time-stepping method. All the implementations of our numerical method are done under Matlab 7.0 environment.

First, we investigate the order of the convergence of the fitted finite volume method by numerical experiments. To achieve this, we use a sequence of uniform meshes as listed in the column 'Grid($N_x \times N_y$)' in Table 2, where the l_1 penalty method with $\lambda = 500$ is used and the number of time steps is fixed to 200. The results are presented in Table 2. From Table 2 it is clear that the computed rates of convergence for the method are very close to $\mathcal{O}(h)$ in the discrete maximum norm.

Table 2: Values of a put option at $x = 1, y = 0, t = T$ and $x = 1, y = 0, t = 0$ under grid refinement, using the fitted finite volume method combined with the implicit scheme, data as in Table 1. l_1 penalty method with $\lambda = 500$ is used. "Ratio" is the ratio of changes on successive grids.

Grid($N_x \times N_y$)	$x = 1.0, y = 0, t = 0$		$x = 0.5, y = 0.5, t = 0$	
	V	Ratio	V	Ratio
40×40	0.0287		0.0180	
80×80	0.0298		0.0189	
160×160	0.0303	1.8	0.0193	1.5
320×320	0.0305	2.5	0.0195	2.0
640×640	0.0306	2.0	0.0196	2.0

Table 3: Results computed by $l_{1/2}$ penalty method with a sequential penalty parameter λ_i, data as in Table 1. "Rate" is defined by (7.1). Grid:640×640, time steps$= 200$.

λ_i	$\|V_\lambda - V\|$	Rate	Average Iteration
$\lambda_1 = 16$	0.00045		1.8
$\lambda_2 = 256$	0.00022		1.9
$\lambda_3 = 4096$	0.00012	0.23	2.0
$\lambda_4 = 65536$	0.00008	0.25	2.0
$\lambda_5 = 1048576$	0.00006	0.25	2.3

Table 4: Results computed by l_1 penalty method with a sequential penalty parameter λ_i, data as in Table 1. "Rate" is defined by (7.1). Grid:640×640, time steps$= 200$.

λ_i	$\|V_\lambda - V\|$	Rate	Average Iteration
$\lambda_1 = 16$	0.00034		1.9
$\lambda_2 = 64$	0.00019		2.0
$\lambda_3 = 256$	0.00010	0.47	2.0
$\lambda_4 = 1024$	0.00006	0.49	2.0
$\lambda_5 = 4096$	0.00004	0.50	2.3

Table 5: Results computed by l_2 penalty method with a sequential penalty parameter λ_i, data as in Table 1. "Rate" is defined by (7.1). Grid:640×640, timesteps$= 200$.

λ_i	$\|V_\lambda - V\|$	Rate	Average Iteration
$\lambda_1 = 2$	0.00022		3.5
$\lambda_2 = 4$	0.00013		4.2
$\lambda_3 = 8$	0.00008	0.98	5.5
$\lambda_4 = 16$	0.00004	0.98	6.7
$\lambda_5 = 32$	0.00002	0.99	7.4

Second, we verify the rate of convergence of the penalty approach with respect to the penalty parameter λ. To do so, we choose a sequence of penalty parameters as listed in the column 'λ_i' in Tables 3, 4 and 5, where the number of time steps is fixed to 200 and the $N_x \times N_y$ grid is chosen to be 640×640. We use the solution with the greatest penalty parameter as the 'exact solution (V)'. Then, we compute the following ratios of the numerical solutions of the consecutive penalty parameters λ_i:

$$\text{Ratio}(\|\cdot\|) = \frac{\|V_{\lambda_{i+1}} - V\|}{\|V_{\lambda_i} - V\|}$$

in the solution domain, where V_{λ_i} denotes the computed solution with the ith penalty parameter λ_i, $\|\cdot\|$ denotes the corresponding discrete form of $\|\cdot\|_{L^\infty(0,T;L^2(\Omega))} + \|\cdot\|_{L^2(0,T;H^1_{0,\varpi}(\Omega))}$. The numerical order of convergence is then defined by

$$\text{Rate} = \log_{\lambda_{i+1}/\lambda_i} \text{Ratio}. \tag{7.1}$$

It is easy to see that for l_k penalty method 'Rate$= k/2$' is consistent with the theoretical convergence result (5.12). As is observed, the columns 'Rate' in Tables 3, 4 and 5 clearly

show the rates of convergence for $l_{1/2}$ ($k = 1/2$), l_1 ($k = 1$) and l_2 ($k = 2$) penalty methods are $\mathcal{O}(\lambda^{-1/4})$, $\mathcal{O}(\lambda^{-1/2})$ and $\mathcal{O}(\lambda^{-1})$ respectively, which confirms the theoretical results as stated in Section 5.

As different Newton nonlinear iterations are employed for $l_{1/2}$, l_1 and l_2 penalty methods, we also investigate the numerical performance of these different Newton's methods. In doing so, we calculate the average number of Newton iteration for each time step with each penalty parameter. The columns 'Average Iterations' in Tables 3, 4 and 5 show the computational costs of different Newton's methods. It can be seen that with the same level of accuracy the $l_{1/2}$ and l_1 penalty methods need the least computational cost while the l_2 penalty method needs the most expensive computational cost. This is due to the nonsmoothness (hence singularity) of the l_2 penalty function, which causes much more iterations for Newton's method with smoothing technique. At the same time, the $l_{1/2}$ and l_1 penalty functions are smooth and semismooth respectively, both of which possesses quadratic convergence rate if classical Newton's method and semismooth Newton's method are used respectively without any smoothing technique (cf. [9]). Moreover, Tables 3, 4 and 5 show that there is a trade-off between the penalty methods, penalty parameters and the computational costs: for the same level of accuracy the higher the penalty power (k) becomes, the less penalty parameter and the more expensive computational cost are required. In our example, the $l_1(k = 1)$ penalty method is a moderate choice, given the balance of computational efficiency and accuracy.

We now investigate numerically the influence of the penalty parameter λ on the accuracy of the option value and its first and second derivatives with respect to x, denoted respectively by Delta and Gamma. To observe the influence, we use the l_2 penalty method with the sequence $\lambda = 4, 16$ and 64. The number of time steps is 200 and the $N_x \times N_y = 640 \times 640$. For brevity, we compute the option value V, Delta $\partial V / \partial x$ and Gamma $\partial^2 V / \partial x^2$ at $t = 0$ on the cross-section $y = 0$. The results are depicted in Figs. 1, 2 and 3 from which it can be seen that while the computed option prices are close to each other for the different values of λ, the differences in Delta and Gamma are significant. It can be also seen that the larger the value of λ is, the more accurate the option value and its derivatives are.

Finally, we depicts in Figs. 4 and 5 the option value V and $V - V^*$ on a cross-section of a fixed t with the parameters given in Table 1. From the figure we see that the constraint is always (up to a tolerance) satisfied. To finish this section, we have the following remark on the use of the smoothing parameter ε.

Remark 7.1. As mentioned before, when $k > 1$ in the penalty method, the smoothing technique (6.10) needs to be used in order that Newton's method applies. In this case, the convergence rate of the method is of order $\mathcal{O}(\lambda^{-k/2})$ when $V_{i,j}^* - V_{i,j}^m \geq \varepsilon$, but it reduces to a much lower order than this when $0 < V_{i,j}^* - V_{i,j}^m < \varepsilon$. On the other hand, the Jacobian of the penalty function is close to singular when $V_{i,j}^* - V_{i,j}^m = \varepsilon$. Therefore, when choosing ε, we need to balance the tolerances of singularity and convergence rate. For the above numerical experiments, ε was chosen to be 10^{-3}.
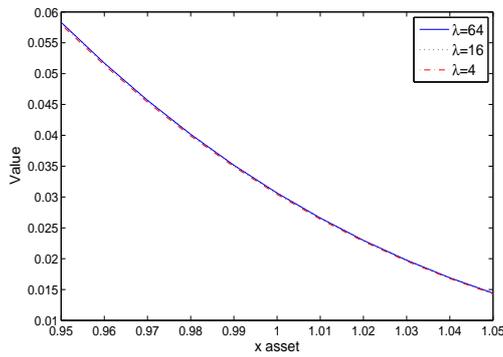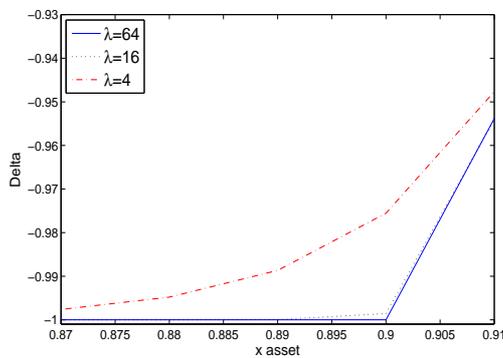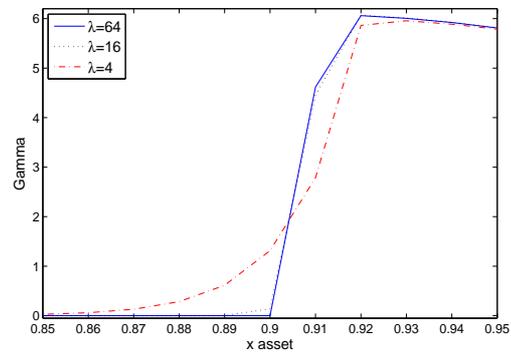
Figure 1: American put option value at $t = 0$, computed by l_2 penalty method with three penalty parameter $\lambda = 4, 16, 64$, data as in Table 1. Grid: 640×640, time steps $= 200$.

Figure 2: American put option Delta at $t = 0$, computed by l_2 penalty method with three penalty parameter $\lambda = 4, 16, 64$, data as in Table 1. Grid: 640×640, time steps $= 200$.
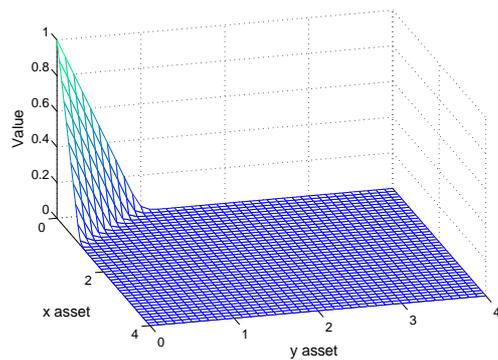

Figure 3: American put option Gamma at $t = 0$, computed by l_2 penalty method with three penalty parameter $\lambda = 4, 16, 64$, data as in Table 1. Grid: 640×640, time steps $= 200$.

Figure 4: V computed at $t = 0$ by l_2 penalty method with $\lambda = 32$. Grid: 640×640, time steps $= 200$.
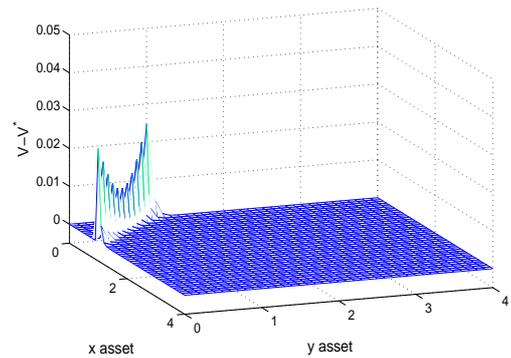

Figure 5: $V - V^*$ computed at $t = 0$ by l_2 penalty method with $\lambda = 32$. Grid: 640×640, time steps $= 200$.

8. Conclusions

In this paper, we presented the power penalty method for the two-asset American option pricing problem and applied the fitted volume method to the resulting nonlinear penalized parabolic PDE. The rate of convergence of the power penalty method was obtained in an infinite dimensional space. We showed some numerical results to confirm the theoretical findings.

References

[1] L. Angermann and S. Wang. Convergence of a fitted finite volume method for the penalized Black-Scholes equation governing European and American Option pricing. *Numer. Math.*, 106:1-40, 2007.

[2] A.Q.M. Khaliq, D.A. Voss and S.H.K. Kazmi. A linearly implicit predictor-corrector scheme for pricing American options using a penalty method approach. *J. Bank. Finance*, 30:489-502, 2006.

[3] A. Bensoussan and J.L. Lions. *Applications of Variational Inequalities in Stochastic Control*. North-Holland, Amsterdam-New York-Oxford, 1982.

[4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81:637-659, 1973.

[5] P.P. Boyle, M. Broadie and P. Glasserman. Monte Carlo methods for security pricing. *J. Econ. Dynam. Control*, 21:8-9, 1997.

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

[7] M.A.H. Dempster, J.P. Hutton and D.G. Richards. LP valuation of exotic American options exploiting structure. *J. Comp. Fin.*, 2:61-84, 1998.

[8] G. Fasshauer, A.Q.M. Khaliq and D.A. Voss. Using mesh free approximation for multi-asset American options, (special issue on mesh free methods, guest editor, C.S. Chen, university of Nevada, Los Vegas). *J. Chin. Inst. Eng.*, 27(4):563-571, 2004.

[9] 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.

[10] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer-Verlag, New York, 1984.

[11] J. Haslinger, M. Miettinen and D.P. Panagiotopoulos. *Finite Element Method for Hemivariational Inequalities: Theory, Methods and Applications*. Kluwer Academic Publishers, Dordrecht, 1999.

[12] 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.

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

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

[15] J.-S. Pang and L. Qi. Nonsmooth equations: motivation and algorithms. *SIAM J. Optim.*, 3:443-465, 1993.

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

[17] J. Topper. *Financial Engineering with Finite Elements*. John Wiley and Sons, NJ, 2005.

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

[19] 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.

[20] K. Zhang, X.Q. Yang and K.L. Teo. Augmented Lagrangian method applied to American option pricing. *Automatica*, 42:1407-1416, 2006.

[21] R. Zvan, P.A. Forsyth and K.R. Vetzal. Penalty methods for American options with stochastic volatility. *J. Comput. Appl. Math.*, 91:199-218, 1998.