J Glob Optim (2013) 56:1313–1323
DOI 10.1007/s10898-012-9843-1

Convergence analysis of power penalty method for American bond option pricing

K. Zhang · K. L. Teo

Received: 28 December 2011 / Accepted: 3 January 2012 / Published online: 11 January 2012

Abstract This paper is concerned with the convergence analysis of power penalty method to pricing American options on discount bond, where the single factor Cox–Ingrosll–Ross model is adopted for the short interest rate. The valuation of American bond option is usually formulated as a partial differential complementarity problem. We first develop a power penalty method to solve this partial differential complementarity problem, which produces a nonlinear degenerated parabolic PDE. Within the framework of variational inequalities, the solvability and convergence properties of this penalty approach are explored in a proper infinite dimensional space. Moreover, a sharp rate of convergence of the power penalty method is obtained. Finally, we show that the power penalty approach is monotonically convergent with the penalty parameter.

Keywords Complementarity problem · Variational inequalities · Option pricing · Penalty method

1 Introduction

The valuation and hedging of interest rate derivatives, like bond options, interest rate caps and swaps, have attracted a large number of attention from both mathematicians and financial engineers. Compared to stock derivatives, the pricing and hedging of interest rate derivatives pose greater challenges. For instance, for a simple bond option, unlike stock derivatives, its underlying asset is a bond whose price is dependent on interest rate and time [13]. It is thus necessary to develop dynamic models that describe the stochastic evolution of the whole yield curve, which makes pricing interest rate derivatives a complex task. A feature distinguishing interest models from equity models is the need for the interest rate models

K. Zhang (✉)
Business School, Shenzhen University, Shenzhen, Guandong Province, China
e-mail: mazhangkai@gmail.com

K. L. Teo
Department of Mathematics and Statistics, The Curtin University of Technology, Perth, WA, Australia

to exhibit mean reversion and for the volatility to be dependent on the interest rate. Many approaches to modeling interest rate derivatives have been established among academics and practitioners, such as Black model [2], Vasicek model, CIR model, HW model [4,11,15] and so on.

In this paper, we focus on the valuation method to American options on zero-coupon bond under the CIR model. This problem is formulated as a parabolic partial differential complementarity problem with suitable boundary and terminal conditions [7,9,12,14,20]. Due to the early exercise feature, this complementarity problem is, in general, not analytically solvable. Hence, numerical approximation methods are normally sought for pricing American bond options. Various approximation techniques have been developed for the solution of American bond option pricing problem. Among them, lattice method [2,17], explicit method [3,11], projected successive over relaxed method (PSOR) [19], semidefinite programming method [5], are the most popular ones in both practice and research.

It is well known that complementarity problems can be solved by penalty methods (cf. [1,6,10]). Recently, the penalty method for pricing American options on stocks was also presented in [8,18,21]. Compared with other methods mentioned above, the penalty method possesses several advantages [16]. 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. The power penalty method to American options on stocks has been well investigated in [18,21], where the convergence properties of the power penalty method were given. However, rare works are available for the study of the power penalty approach to pricing American options on bond. Hence, the main purpose of this paper is to develop a power penalty method for the complementarity problem arising from the valuation of American options on bond. Via the theory of variational inequalities, we first approximate the complementarity problem by a nonlinear parabolic PDEs with an l_k power penalty term. Then, strong convergence of the penalization is under investigation. At the same time, we prove that the solution to the nonlinear PDE converges to that of the original complementarity problem at the rate of order $\mathscr{O}(\lambda^{-k/2})$. Furthermore, the monotonicity of convergence of the power penalty method with the penalty parameter λ is established.

The organization of this paper is as follows. In Sect. 2, we introduce the mathematical model for pricing American options on bond and its equivalent formulations: complementarity problem and variational inequalities. Section 3 gives a power penalty approach to the complementarity problem. The strong convergence, rate of convergence and the monotonicity of convergence of the power penalty method is given in Sect. 5.

Before proceeding, some standard notation is to be used in the paper. For an open set $\mathbf{S} \subset \mathbb{R}$ and $1 \leq p \leq \infty$, let

$$L^p(\mathbf{S}) = \left\{ v : \left(\int_{\mathbf{S}} |v(x)|^p dx \right)^{1/p} < \infty \right\}$$

denote the space of all p-power integrable functions on $\mathbf{S}$. We use the $\|\cdot\|_{L^p(\mathbf{S})}$ to denote the norm on $L^p(\mathbf{S})$. The inner product on $L^2(\mathbf{S})$ is denoted by $(\cdot, \cdot)$. For $m = 1, 2, \ldots$ and $p = 2$, we let $H^m(\mathbf{S})$ denote the usual Sobolev space with the norm $\|\cdot\|_{H^m(\mathbf{S})}$. We put

$$H_0^m(\mathbf{S}) = \{v : v \in H^m(\mathbf{S}),\ v|_{\partial\mathbf{S}} = 0\}$$

where $\partial \mathbf{S}$ is the boundary of $\mathbf{S}$. Finally, for any Hilbert space $W(\mathbf{S})$, the norm of $L^p(0, T; W(\mathbf{S}))$ is denoted by

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

Obviously, $L^p(0, T; L^p(\mathbf{S})) = L^p(\mathbf{S} \times (0, T))$.

For clarity, we will often simply write $v(\cdot, t)$ as $v(t)$ when we regard $v(\cdot, t)$ as an element of $H_0^1(\mathbf{S})$. From time to time, we will also suppress the independent time variable t when it causes no confusion in doing so.

2 Mathematic model

In this paper, we assume the CIR model is applied for the interest rate term structure. That means the short-term interest rate r is governed by the following mean-reverting version of the square-root process

$$dr = \kappa (\theta - r) \, dt + \sigma \sqrt{r} dW,$$

where dW is the increment of a Wiener process, θ is the long-term level of the short rate, $\kappa > 0$ stands for the reversion speed, $\sigma^2 r$ $(\sigma > 0)$ is the variance. In practice the interest rate r is a positive defined quantity, which enforces the following constraint (Feller's condition [4])

$$0 < \sigma^2 < 2\kappa\theta. \tag{1}$$

In [4], it has been shown that the price $P(r, t, s)$ of a pure discount bond with face value \$1 at its maturity date s is given as follows

$$P(r, t, s) = A(t, s) \, e^{-B(t,s)r},$$

where

$$A(t, s) = \left[\frac{\phi_1 e^{\phi_2 (s-t)}}{\phi_2 \left[e^{\phi_1 (s-t)} - 1 \right] + \phi_1} \right]^{\phi_3}, \quad B(t, s) = \frac{e^{\phi_1 (s-t)} - 1}{\phi_2 \left[e^{\phi_1 (s-t)} - 1 \right] + \phi_1},$$
$$\phi_1 = \sqrt{\mu^2 + 2\sigma^2}, \quad \phi_2 = (\mu + \phi_1)/2, \quad \phi_3 = 2\overline{\theta}/\sigma^2,$$
$$\overline{\theta} = \kappa\theta, \quad \mu = \kappa + \lambda,$$

and λ is the market risk premium.

Now, let $V(r, t)$ be the value of an American option on a zero-coupon bond with striking price K, where the holder can receive the payoff $V^*(r, t)$ at expiry date T. Then, the option pricing problem can be formulated as the following parabolic partial differential complementarity problem (PDCP) [19].

Problem 1

$$\begin{cases} LV(r, t) \geq 0, \\ V(r, t) - V^*(r, t) \geq 0, \\ LV(r, t) \cdot (V(r, t) - V^*(r, t)) = 0, \end{cases} \tag{2}$$

a.e. in $(0, +\infty) \times (0, T)$, where

$$LV = -\frac{\partial V}{\partial t} - \left[\frac{1}{2}\sigma^2 r \frac{\partial^2 V}{\partial r^2} + \left(\overline{\theta} - \mu r\right)\frac{\partial V}{\partial r} - rV\right]$$

with the final condition

$$V(r, t = T) = V^*(r, T) = \begin{cases} \max\left[P(r, T, s) - K, 0\right], & \text{for a call,} \\ \max\left[K - P(r, T, s), 0\right], & \text{for a put.} \end{cases}$$

and the following boundary conditions

$$V(0, t) = \lim_{r \to +0} V^*(r, t), \tag{3}$$

$$\lim_{r \to +\infty} V(r, t) = \lim_{r \to +\infty} V^*(r, t). \tag{4}$$

For computational purpose, we restrict r in a region $[0, R]$, where R is sufficiently large to ensure the accuracy of the solution ([19]). Thus, (4) becomes

$$V(R, t) = V^*(R, t). \tag{5}$$

Remark 1 It is worth noting that $T < s$ and $K < P(0, T, s) = A(T, s)$ for a call option or $K > A(T, s)$ for a put option, since otherwise the option would never be exercised and would be worthless.

In order to remove the degenerate factor r in the second order derivative term of L, by introducing

$$x = \sqrt{r},\ V(r, t) = x^{-\alpha} e^{\gamma t} U(x, t), \quad U^*(x, t) = x^{\alpha} e^{-\gamma t} V^*\left(x^2, t\right), \quad u = U^* - V$$

we transform (2)–(4) into the following equivalent form satisfying a homogeneous Dirichlet boundary condition, where α and γ are properly selected positive constants to have a coercive bilinear form.

Problem 2

$$\begin{cases} u(x, t) \le 0, \\ \mathscr{L}u(x, t) \le f(x, t), \\ (\mathscr{L}u(x, t) - f(x, t)) \cdot u(x, t) = 0, \end{cases} \tag{6}$$

in $\Omega = I \times (0, T),\ I = (0, X),\ X = \sqrt{R}$, with

$$u(x, T) = 0 \tag{7}$$

and

$$u(0, t) = u(X, t) = 0, \tag{8}$$

where

$$\mathscr{L}u = -\frac{\partial u}{\partial t} - \frac{1}{8}\sigma^2 \frac{\partial^2 u}{\partial x^2} + c_1(x)\frac{\partial u}{\partial x} + c_2(x) u,$$
$$f(x, t) = -\mathscr{L}U^*(x, t)$$

with

$$c_1(x) = \frac{\sigma^2}{8}\left(1 + 2\alpha - \frac{4a}{\sigma^2}\right) x^{-1} + \frac{b}{2} x,$$
$$c_2(x) = \frac{\sigma^2 \alpha}{8}\left(\frac{4a}{\sigma^2} - \alpha - 2\right) x^{-2} + x^2 + \gamma - \frac{\alpha b}{2}.$$

Remark 2 The choice of positive constants α and γ can be found in [20].

It is a standard result [10] that the linear complementarity problem (6)–(8) can be reformulated as the following equivalent variational inequalities.

Problem 3 Find $u \in \mathscr{K}$, such that, for all $v \in \mathscr{K}$,

$$\left(-\frac{\partial u}{\partial t}, v - u\right) + A(u, v - u; t) \geq (f, v - u) \tag{9}$$

a.e. in $(0, T)$, where $A(\cdot, \cdot; t)$ is in a bilinear form defined by

$$A(u, v; t) := \frac{\sigma^2}{8}\left(\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}\right) + \left(c_1\frac{\partial u}{\partial x} + c_2 u, v\right), \quad u, v \in H_0^1(I),$$

and $\mathscr{K} = \{v \in H_0^1(I) : v \leq 0\}$ is a convex and closed subset of $H_0^1(I)$.

For Problem 3, we establish the following unique solvability result.

Lemma 1 *Variational inequality* (9) *has a unique solution.*

Proof In fact, it has been shown in [20] that the operator $A(\cdot, \cdot; t)$ is coercive and continuous, i.e.

$$A(u, u; t) \geq \gamma_1 \|u\|^2_{H_0^1(I)}, \tag{10}$$

$$A(u, v; t) \leq \gamma_2 \|u\|_{H_0^1(I)} \|v\|_{H_0^1(I)}, \tag{11}$$

where γ_1 and γ_2 are two positive constants. Hence, by virtue of Theorem 2.3 in [1], the unique solvability for the parabolic variational inequalities (i.e., Problem 3) is established. □

Remark 3 It is pointed out in [20] that when $\frac{\kappa\theta}{\sigma^2} = \frac{1}{2}$, the bilinear form $A(\cdot, \cdot; t)$ is not coercive, hence the existence of the solution to variational inequality (9) cannot be guaranteed. However, since the Feller's condition (1) is applied, this case is ruled out in this work. Hence, with the choice of positive constants α and γ suggested in [20], the coerciveness of $A(\cdot, \cdot; t)$ is guaranteed.

3 Power penalty approach

In this section, we propose a power penalty approach to the complementarity problem (6)–(8). To derive the power penalty approach, we first consider the following nonlinear variational inequalities problem:

Problem 4 Find $u_\lambda \in H_0^1(I)$ such that, for all $v \in H_0^1(I)$,

$$\left(-\frac{\partial u_\lambda}{\partial t}, v - u_\lambda\right) + A(u_\lambda, v - u_\lambda; t) + j(v) - j(u_\lambda) \geq (f, v - u_\lambda), \tag{12}$$

a.e. in $(0, T)$, where

$$j(v) = \frac{\lambda k}{k+1}[v]_+^{\frac{k+1}{k}}, \quad k > 0, \ \lambda > 1, \tag{13}$$

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

Since the bilinear operator A is coercive and continuous, and the operator j is lower semi-continuous, the unique solvability of Problem 4 is easily obtained (cf. [10]).

From (13), we can see that $j(v)$ is differentiable. Thus, Problem 4 is equivalent to the following problem.

Problem 5 Find $u_\lambda \in H_0^1(\Omega)$ such that, for all $v \in H_0^1(\Omega)$,

$$\left(-\frac{\partial u_\lambda}{\partial t}, v\right) + A(u_\lambda, v; t) + (j'(u_\lambda), v) = (f, v), \tag{14}$$

a.e. in $(0, T)$, where

$$j'(v) = \lambda [v]_+^{1/k}. \tag{15}$$

We remark that (5)–(15) is a penalized variational equation corresponding to (3). The strong form of (5)–(15), which defines the penalized equation approximating (6), is given by

$$\mathscr{L} u_\lambda + \lambda [u_\lambda]_+^{1/k} = f, \tag{16}$$

with the given boundary and final conditions

$$\begin{aligned} u_\lambda(0, t) &= 0, \\ u_\lambda(X, t) &= 0, \\ u_\lambda(x, T) &= 0. \end{aligned} \tag{17}$$

Remark 4 If $k = \frac{1}{2}$, this penalty approach corresponds to the quadratic penalty approach. For $k = 1$, the typical l_1 penalty approach is obtained. When $k > 1$, it is the so-called lower order penalty approach [18]. In the next section, we will investigate the convergence properties of u_λ to u as $\lambda \to \infty$.

Remark 5 The regularity results of the solution to the penalized problems have been studied extensively in several monographs such as [6] and [1]. In brief, under the assumption that $u_\lambda(x, t)$ and $f(x, t)$ are sufficiently smooth, we have the following regularity results.

$$\frac{\partial u_\lambda(x, t)}{\partial t},\ u_\lambda(x, t) \in L^2(0, T; H_0^1(I)) \cap L^\infty(0, T; L^2(I)).$$

4 Convergence analysis

4.1 Rate of convergence of the power penalization

We now show that, as $\lambda \to \infty$, the solution to Problem 5 converges to that of Problem 3 in a proper norm. We start this discussion by the following Lemma.

Lemma 2 *Let u_λ be the solution to Problem 5. If $u_\lambda \in L^p(\Omega)$, then there exists a positive constant C, independent of u_λ and λ, such that*

$$\begin{aligned} \|[u_\lambda]_+\|_{L^p(\Omega)} &\le \frac{C}{\lambda^k}, \\ \|[u_\lambda]_+\|_{L^\infty(0,T;L^2(I))} + \|[u_\lambda]_+\|_{L^2(0,T;H_0^1(I))} &\le \frac{C}{\lambda^{k/2}}, \end{aligned} \tag{18}$$

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

Proof Assume that C is a generic positive constant, independent of u_λ and λ. Clearly, $[u_\lambda]_+ \in H_0^1(I)$ for almost all $t \in (0, T)$. Now, setting $v = [u_\lambda]_+$ in (14), we have

$$\left(-\frac{\partial u_\lambda}{\partial t}, [u_\lambda]_+\right) + A(u_\lambda, [u_\lambda]_+; t) + \lambda([u_\lambda]_+^{1/k}, [u_\lambda]_+) = (f, [u_\lambda]_+). \tag{19}$$

Integrating both sides of (19) from t to T and using the coerciveness property of the operator A and Hölder's inequality, it follows that

$$\begin{aligned}&\frac{1}{2}\,\|[u_\lambda\,(t)]_+\|_{L^2(I)}^2 + \gamma_1 \int_t^T ||[u_\lambda]_+||_{H_0^1(I)}^2 d\tau + \lambda \int_t^T [u_\lambda]_+^{1+1/k} d\tau \\ &\leq \int_t^T (f, [u_\lambda]_+) d\tau \leq C \left(\int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau\right)^{1/p},\end{aligned} \tag{20}$$

from which, we infer

$$\begin{aligned}&\frac{1}{2}\,\|[u_\lambda\,(t)]_+\|_{L^2(I)}^2 + \gamma_1 \int_t^T ||[u_\lambda]_+||_{H_0^1(I)}^2 d\tau + \lambda \int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau \\ &\leq C \left(\int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau\right)^{1/p}.\end{aligned} \tag{21}$$

This implies that

$$\lambda \int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau \leq C \left(\int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau\right)^{1/p}.$$

From this, it follows that

$$\left(\int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau\right)^{1/p} \leq \frac{C}{\lambda^{1/(p-1)}}. \tag{22}$$

Then, from (21) and (22), we have

$$\frac{1}{2}\,([u_\lambda]_+, [u_\lambda]_+) + \int_t^T ||[u_\lambda]_+||_{H_0^1(I)}^2 d\tau \leq C \left(\int_t^T ||[u_\lambda]_+||_{L^p(I)}^p d\tau\right)^{1/p} \leq \frac{C}{\lambda^{1/(p-1)}},$$

and hence

$$([u_\lambda]_+, [u_\lambda]_+)^{\frac{1}{2}} + \left(\int_t^T ||[u_\lambda]_+||_{H_0^1(I)}^2 d\tau\right)^{\frac{1}{2}} \leq \frac{C}{\lambda^{1/(2p-2)}},$$

i.e.

$$\|[u_\lambda]_+\|_{L^\infty(0,T;L^2(I))} + \|[u_\lambda]_+\|_{L^2(0,T;H_0^1(I))} \leq \frac{C}{\lambda^{k/2}}.$$

□

On the basis of Lemma 2, we obtain the following convergence rate.

Theorem 1 *Assume that the assumptions of Lemma 2 are satisfied. If* $u_\lambda \in L^p(\Omega)$ *and* $\frac{\partial u}{\partial t} \in L^q(\Omega)$*, then there exists a positive constant* C*, independent of* u_λ *and* λ*, such that*

$$\|u_\lambda - u\|_{L^\infty(0,T;L^2(I))} + \|u_\lambda - u\|_{L^2(0,T;H_0^1(I))} \le \frac{C}{\lambda^{k/2}}, \tag{23}$$

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

Proof First ,we note that $u - u_\lambda$ can be decomposed as

$$u - u_\lambda = R_\lambda - [u_\lambda]_+ \,,$$

where

$$R_\lambda = u + [u_\lambda]_- \,, [u_\lambda]_- = -\min(u_\lambda, 0)\,.$$

Then, it follows from (18) that, in order to prove (23), it is sufficient to show that

$$||R_\lambda||_{L^\infty(0,T;L^2(I))\cap L^2(0,T;H_0^1(I))} \le \frac{C}{\lambda^{k/2}}. \tag{24}$$

Set $v = u - R_\lambda$ in (9) and $v = R_\lambda$ in (14). Then, we have

$$\left(-\frac{\partial u}{\partial t}, -R_\lambda\right) + A(u, -R_\lambda; t) \ge (f, -R_\lambda), \tag{25}$$

$$\left(-\frac{\partial u_\lambda}{\partial t}, R_\lambda\right) + A(u_\lambda, R_\lambda; t) + \lambda\left([u_\lambda]_+^{1/k}, R_\lambda\right) = (f, R_\lambda). \tag{26}$$

Combining (25) and (26) gives

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

But, it follows from $u \le 0$ that.

$$\left([u_\lambda]_+^{1/k}, R_\lambda\right) = \left([u_\lambda]_+^{1/k}, u\right) + \left([u_\lambda]_+^{1/k}, [u_\lambda]_-\right) = \left([u_\lambda]_+^{1/k}, u\right) \le 0.$$

Thus,

$$\left(-\frac{\partial(u - u_\lambda)}{\partial t}, R_\lambda\right) + A(u - u_\lambda, R_\lambda; t) \le 0, \tag{27}$$

and hence

$$\left(-\frac{\partial R_\lambda}{\partial t}, R_\lambda\right) + A(R_\lambda, R_\lambda; t) \le \left(-\frac{\partial [u_\lambda]_+}{\partial t}, R_\lambda\right) + A([u_\lambda]_+, R_\lambda; t).$$

Integrating both sides of the above from t to T and then using Cauchy-Schwartz inequality, we obtain

$$\frac{1}{2}\,||R_\lambda\,(t)||^2_{L^2(I)} + \gamma_1 \int_t^T ||R_\lambda||^2_{H_0^1(I)} d\tau$$

$$\le ([u_\lambda(t)]_+, R_\lambda(t)) + \int_t^T \left([u_\lambda]_+, \frac{\partial R_\lambda}{\partial \tau}\right) d\tau + \int_t^T A([u_\lambda]_+, R_\lambda; \tau) d\tau. \tag{28}$$

But

$$\left| \int_t^T \left([u_\lambda]_+, \frac{\partial R_\lambda}{\partial \tau} \right) d\tau \right| = \left| \int_t^T \left([u_\lambda]_+, \frac{\partial u}{\partial \tau} \right) d\tau \right| \leq C \| [u_\lambda]_+ \|_{L^p(\Omega)} \left\| \frac{\partial u}{\partial t} \right\|_{L^q(\Omega)} \leq \frac{C}{\lambda^k},$$

and hence (28) gives

$$\frac{1}{2} \| R_\lambda (t) \|^2_{L^2(I)} + \gamma_1 \int_t^T \| R_\lambda \|^2_{H_0^1(I)} d\tau \leq \frac{C}{\lambda^{k/2}} \left(|R_\lambda (t)| + \int_t^T \| R_\lambda \|^2_{H_0^1(I)} d\tau \right) + \frac{C}{\lambda^k},$$

so that (24) follows. □

4.2 The monotonic convergence of the penalization

In this subsection, we will show that when the penalty parameter λ increases, the solution to the penalized problem decreases. Moreover, the solution of the penalized problem is bounded below by that of the original complementarity problem. The monotonic convergence of the power penalization is stated in the following theorem.

Theorem 2 *Let* $0 < \lambda_1 \leq \lambda_2$ *be two different penalty parameters. Then,*

$$u \leq u_{\lambda_2} \leq u_{\lambda_1},$$

where u is the solution to Problem 3, u_{λ_1} *and* u_{λ_2} *are the solutions to Problem 5 for* $\lambda = \lambda_1$ *and* λ_2*, respectively.*

Proof In (14), we set $v = [u_{\lambda_2} - u_{\lambda_1}]_+$ and $\lambda = \lambda_1$ and λ_2, respectively. Then, it follows that

$$\left(-\frac{\partial u_{\lambda_1}}{\partial t}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) + A\left(u_{\lambda_1}, [u_{\lambda_2} - u_{\lambda_1}]_+ ; t\right) + \lambda_1 \left([u_{\lambda_1}]_+^{\frac{1}{k}}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) = \left(f, [u_{\lambda_2} - u_{\lambda_1}]_+\right),$$

$$\left(-\frac{\partial u_{\lambda_2}}{\partial t}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) + A\left(u_{\lambda_2}, [u_{\lambda_2} - u_{\lambda_1}]_+ ; t\right) + \lambda_2 \left([u_{\lambda_2}]_+^{\frac{1}{k}}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) = \left(f, [u_{\lambda_2} - u_{\lambda_1}]_+\right),$$

and hence

$$\begin{aligned} &\left(-\frac{\partial (u_{\lambda_2} - u_{\lambda_1})}{\partial t}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) + A(u_{\lambda_2} - u_{\lambda_1}, [u_{\lambda_2} - u_{\lambda_1}]_+; t) \\ &\quad = \lambda_1 \left([u_{\lambda_1}]_+^{1/k}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) - \lambda_2 \left([u_{\lambda_2}]_+^{1/k}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) \\ &\quad = (\lambda_1 - \lambda_2) \left([u_{\lambda_1}]_+^{1/k}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) + \lambda_2 \left([u_{\lambda_1}]_+^{1/k} - [u_{\lambda_2}]_+^{1/k}, [u_{\lambda_2} - u_{\lambda_1}]_+\right). \end{aligned} \tag{29}$$

For the first term in (29), it is obvious that

$$(\lambda_1 - \lambda_2) \left([u_{\lambda_2}]_+^{1/k}, [u_{\lambda_2} - u_{\lambda_1}]_+\right) = (\lambda_1 - \lambda_2) \int_0^X [u_{\lambda_2}]_+^{1/k} [u_{\lambda_2} - u_{\lambda_1}]_+ dx \leq 0, \tag{30}$$

since $\lambda_1 < \lambda_2$.

For the second term in (29), we shall also show that

$$\lambda_2\left([u_{\lambda_1}]_+^{1/k}-[u_{\lambda_2}]_+^{1/k},\left[u_{\lambda_2}-u_{\lambda_1}\right]_+\right)=\lambda_2\int_0^X\left([u_{\lambda_1}]_+^{1/k}-[u_{\lambda_2}]_+^{1/k}\right)[u_{\lambda_2}-u_{\lambda_1}]_+dx\le 0.$$

In fact, since for the set where $u_{\lambda_2}\le u_{\lambda_1}$, $[u_{\lambda_2}-u_{\lambda_1}]_+=\max\{u_{\lambda_2}-u_{\lambda_1},0\}=0$, to calculate $\int_0^X([u_{\lambda_1}]_+^{1/k}-[u_{\lambda_2}]_+^{1/k})[u_{\lambda_2}-u_{\lambda_1}]_+dx$ we only need to integrate $u_{\lambda_2}-u_{\lambda_1}$ over the set for which $u_{\lambda_2}>u_{\lambda_1}$. On this set, by virtue of the monotonicity of the operator $[\cdot]_+^{1/k}=(\max\{\cdot,0\})^{1/k}$ we can infer that

$$\lambda_2\left([u_{\lambda_1}]_+^{1/k}-[u_{\lambda_2}]_+^{1/k},\left[u_{\lambda_2}-u_{\lambda_1}\right]_+\right)\le 0.$$

Consequently, on the whole set $I=(0,X)$, we have

$$\lambda_2\left([u_{\lambda_1}]_+^{1/k}-[u_{\lambda_2}]_+^{1/k},\left[u_{\lambda_2}-u_{\lambda_1}\right]_+\right)\le 0. \tag{31}$$

It then follows form (29), (30) and (31) that

$$\left(-\frac{\partial(u_{\lambda_2}-u_{\lambda_1})}{\partial t},[u_{\lambda_2}-u_{\lambda_1}]_+\right)+A(u_{\lambda_2}-u_{\lambda_1},[u_{\lambda_2}-u_{\lambda_1}]_+;t)\le 0,$$

so that

$$\frac{1}{2}||[u_{\lambda_2}(t)-u_{\lambda_1}(t)]_+||^2_{L^2(I)}+\gamma_1\int_t^T||[u_{\lambda_2}-u_{\lambda_1}]_+||^2_{H_0^1(I)}d\tau\le 0,$$

and hence

$$[u_{\lambda_2}-u_{\lambda_1}]_+=0\quad\text{and}\quad u_{\lambda_1}\ge u_{\lambda_2}.$$

Finally, passing to the limit as $\lambda_2\to\infty$ (for a converging subsequence), we deduce that

$$u_{\lambda_1}\ge u_{\lambda_2}\ge u.$$

□

Remark 6 By virtue of Lemma 2 and Theorem 2, we can see that the power penalized problem (16) solves the following complementarity problem

$$\begin{cases} u_\lambda\le \frac{C}{\lambda^k},\\ \mathscr{L}u_\lambda\le f,\\ (\mathscr{L}u_\lambda-f)\cdot\left(u_\lambda-\frac{C}{\lambda^k}\right)=0,\end{cases}$$

where C is independent of u_λ and λ. Intuitively, this complementarity problem is an approximation of the original complementarity problem (6).

5 Conclusion

We have studied the power penalty method for pricing American options on pure discount bond under the CIR model. By using the equivalence of LCP and variational inequalities, the power penalty method was developed. Its solvability and convergence properties of the monotonic penalty method were established as well. We have shown that the solution to the power penalized nonlinear equation converges to that of the original LCP. Furthermore, a sharp convergence rate of the power penalty method was achieved. Finally, the monotonicity of convergence of the penalization was demonstrated.

Acknowledgments The authors would like to thank two anonymous referees for their helpful comments and suggestions toward the improvement of this paper. Project 11001178 supported by National Natural Science Foundation of China. This work was also partially supported by Foundation for Distinguished Young Talents in Higher Education of Guangdong, China (Grant No. WYM10099).

References

1. Bensoussan, A., Lions, J.L.: Applications of Variational Inequalities in Stochastic Control. North-Holland, Amsterdam (1982)
2. Black, F., Karasinski, P.: Bond and option pricing when short rates are lognormal. Financ. Anal. J. **47**, 52–59 (1991)
3. Brennan, M., Schwartz, E.: The valuation of American put options. J. Financ. **32**, 449–462 (1977)
4. Cox, J.C., Ingersoll, J.E., Ross, S.A.: A theory of the term structure of interest rates. Econometrica **53**, 385–407 (1985)
5. Dalakouras, G., Kwon, R., Pardalos, P.M.: Semidefinite programming approaches for bounding Asian option prices. In: Konthoghiorges, E.J., Rustem, B., Winker, P. (eds.) Computational Methods in Financial Engineering, pp. 103–116. Springer, New York (2008)
6. Dauvaut, G., Lions, J.L.: Inequalities in Mechanics and Physics. Springer, Berlin (1976)
7. Elliot, E.M., Ockendon, J.R.: Weak and Variational Methods for Moving Boundary Problems. Pitman, London (1982)
8. Forsyth, P.A., Vetzal, K.R.: Quadratic convergence for valuing American options using a penalty method. SIAM J. Sci. Comput. **23**, 2095–2122 (2002)
9. Giannessi, F., Maugeri, A., Pardalos, P.M.: Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models. Kluwer, Dordrecht (2002)
10. Glowinski, R.: Numerical Methods for Nonlinear Variational Problems. Springer, New York (1984)
11. Hull, J.: Options, Futures, and Other Derivatives. Prentice-Hall, Englewood Cliffs (2005)
12. Kwok, Y.K.: Mathematical Models of Financial Derivatives. Springer, Berlin (1998)
13. Kosmidou, K., Zopounidis, C.: Generating interest rate scenarios for bank asset liability management. Optim. Lett **2**(2), 157–169 (2008)
14. Pardalos, P.M., Tsitsiringos, V.: Financial Engineering, Supply Chain and E-commerce. Kluwer, Dordrecht (2002)
15. Rebonato, R.: Interest-rate Option Models: Understanding, Analyzing and Using Models for Exotic Interest-rate Options. Wiley, Chicherster (1996)
16. Rubinov, A., Yang, X.Q.: Lagrange-type functions in constrained non-convex optimization. Applied Optimization, 85. Kluwer, Boston (2003)
17. Tian, Y.: A simplified binomial approach to the pricing of interest-rate contingent claims. J. Financ. Eng. **1**, 14–37 (1992)
18. Wang, S., Yang, X.Q., Teo, K.L.: Power penalty method for a linear complementarity problem arising from American option valuation. J. Optim. Theory Appl. **129**, 227–254 (2006)
19. Wilmott, P.: Paul Wimott on Quantitative Finance. Wiley, New York (2000)
20. Yang, H.: American put options on zero-coupon bonds and a parabolic free boundary problem. IJNUM **1**(2), 203–215 (2004)
21. Zhang, K., Yang, X.Q., Teo, K.L.: A power penalty approach to American option pricing with jump-diffusion processes. J. Ind. Manag. Optim **4**, 767–782 (2008)