

A penalty approach to a discretized double obstacle problem with derivative constraints

Song Wang

Department of Mathematics & Statistics

Curtin University

GPO Box U1987, Perth, WA6845, Australia

song.wang@curtin.edu.au

Abstract

This work presents a penalty approach to a nonlinear optimization problem with linear box constraints arising from the discretization of an infinite-dimensional differential obstacle problem with bound constraints on derivatives. In this approach, we first propose a penalty equation approximating the mixed nonlinear complementarity representing the Karush-Kuhn-Tucker conditions of the optimization problem. We then show that the penalty solution converges to that of the complementarity problem with an exponential convergence rate depending on the parameters used in the penalty equation. Numerical experiments, carried out on non-trivial test problems to verify the theoretical finding, show that the computed rates of convergence match the theoretical ones well.

Keywords. Double obstacle problem, mixed nonlinear complementarity problem, variational inequalities, bounded linear constraints, global optimizer, penalty method, convergence rates.

1 Introduction

Many real-world problems in engineering, physics, finance and investment are governed by infinite-dimensional constrained optimization problems of the following form:

$$\min_{u \in \mathcal{H}} \mathcal{F}(u) \tag{1.1}$$

$$\text{subject to } g_1 \leq \mathcal{L}u \leq g_2, \tag{1.2}$$

where \mathcal{H} is a functional space, \mathcal{F} is a functional on \mathcal{H} of usually u and its derivatives, \mathcal{L} a given linear differential operators, and g_1 and g_2 satisfying $g_1 \leq g_2$ are known functions defining the lower and upper bounds on $\mathcal{L}u$. Examples of such type of problems are shape-preserving interpolation in which it is required that the interpolation satisfies prescribed slope or convexity conditions, various problems in engineering, physics with gradient or

curl constraints, see, for example, [6, 25, 17, 1, 2] just to name a few. Optimization problems with derivative constraints also arise in differential game theory and pricing financial options under transaction costs in financial engineering [3, 7, 9, 31, 20, 21, 22].

The above infinite-dimensional problem cannot be usually solved exactly except for some trivial cases. In practice, a discretization scheme is usually applied to (1.1)–(1.2) so that the resulting optimization problem is in finite dimensions. Various effective discretization schemes such as finite difference, finite volume and finite element methods are available for the discretization of (1.1)–(1.2) depending on the problem in question [26, 15, 32, 18, 22]. The discretized problem of (1.1)–(1.2) is usually of the following form:

$$\min_{x \in \mathbb{R}^m} F(x) \tag{1.3}$$

$$\text{subject to } -b \leq Ax \leq c, \tag{1.4}$$

where m is a positive integer, $F : \mathbb{R}^m \mapsto \mathbb{R}$ is a nonlinear differentiable function, $A : \mathbb{R}^m \mapsto \mathbb{R}^n$ is an $n \times m$ matrix with n an positive integer, and $b, c \in \mathbb{R}^n$ are given vectors satisfying $-b < c$. Clearly, $F, A, -b$ and c are discretized forms of respectively $\mathcal{F}, \mathcal{L}, g_1$ and g_2 , and the solution x to (1.3)–(1.4) provides an approximation to the solution u to (1.1)–(1.2). We assume that $n \leq m$ and both $Ax = c$ and $Ax = -b$ have at least one feasible solution. In this case, we may simply assume that $c = 0$ and $b > 0$, as the case that $c \neq 0$ can be transformed into this by a simple substitution $z = x - x_0$, where x_0 is a solution to $Ax = c$, and the lower bound for Ax then becomes $-b - c$ due to the transformation.

We comment that in [29] we propose a penalty method for a (single) obstacle problem, i.e., the minimization problem (1.3) subject to the constraint $Ax \leq c$ with $Ax = c$ having at least one solution. However, the double obstacle problem (1.3)–(1.4) is completely different from that in [29] because the former has both lower and upper bounds on Ax . Thus, the theory developed in [29] does not apply to (1.3)–(1.4). As can be seen below, both the formulation and analysis for (1.3)–(1.4) are different from and, in particular, the analysis substantially more difficult than that in [29].

The KKT conditions for (1.3)–(1.4) (with $c = 0$) and $b > 0$ are

$$f(x) + A^\top \mu_1 - A^\top \mu_2 = 0, \tag{1.5}$$

$$\mu_1 \geq 0, \quad \mu_2 \geq 0, \tag{1.6}$$

$$Ax \leq 0, \quad -Ax \leq b, \tag{1.7}$$

$$\mu_1^\top Ax = 0, \quad \mu_2^\top (Ax + b) = 0, \tag{1.8}$$

where $f(x) = \nabla F(x) : \mathbb{R}^m \mapsto \mathbb{R}^m$ and $\mu_1, \mu_2 \in \mathbb{R}^n$ are unknown multipliers. The system (1.5)–(1.8) is a mixed nonlinear complementarity problem (MCP) for the unknowns x, μ_1

and μ_2 containing a set of inequalities and equations. Extensive studies on theoretical aspects and computational algorithms of conventional nonlinear complementarity problems (NCP) have been done for the last two decades and many of these results can be found in a recent outstanding monograph [10] and the references therein. For mixed complementarity problems, some numerical methods have also been developed. These include the semi-smooth Newton method [16], the smooth methods [4, 19], the active-set Newton methods [8] and the interior point method [24], just to name a few.

An NCP or MCP is usually equivalent to a concave minimization problem even for the linear case, as pointed out in [12, pp.24-25]. Thus, how to find a global optimizer for the above MCP becomes an issue. Also, popular numerical methods for solving an MCP such as (1.5)–(1.8) are based on the minimization of a merit function for the MCP. However, the resulting problem is usually a global optimization problem even when the MCP is monotone or strictly monotone [16, 23, 30]. The main problem in the global optimization issue of the aforementioned methods is that either auxiliary cost functions or variables (multipliers), or both introduced in the methods may change the nature of the original problem. In this work, we propose a power penalty method for (1.5)–(1.8) and show that the nonlinear penalty equation from the method is uniquely solvable when the original mapping is ξ -monotone, and thus this formulation does not introduce any local optimal solutions. We also show that the solution to the penalty equation converges exponentially to that of (1.5)–(1.8), and so it provide a global optimal solution to the original problem when the penalty constants approach infinity.

Let $y = -\mu_1$ and $z = -\mu_2$, (1.5)–(1.8) can be written as the following problem:

Problem 1.1 Find $(x^\top \ y^\top \ z^\top)^\top \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$f(x) - A^\top y + A^\top z = 0, \quad (1.9)$$

$$y \leq 0, \quad z \leq 0, \quad (1.10)$$

$$Ax \leq 0, \quad -Ax \leq b, \quad (1.11)$$

$$y^\top Ax = 0, \quad z^\top (Ax + b) = 0. \quad (1.12)$$

Let $\mathcal{K} = \left\{ (x^\top \ y^\top \ z^\top)^\top \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n : y \leq 0, z \leq 0 \right\}$. It is easy to see that \mathcal{K} is a convex subset of $\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$. Using \mathcal{K} we define a variational inequality as follows.

Problem 1.2 Find $u = (x^\top \ y^\top \ z^\top)^\top \in \mathcal{K}$, such that for all $v \in \mathcal{K}$,

$$(v - u)^\top G(u) \geq 0, \quad (1.13)$$

where $G : \mathbb{R}^{m+2n} \mapsto \mathbb{R}^{m+2n}$ is given by

$$G(u) := \begin{pmatrix} f(x) - A^\top y + A^\top z \\ Ax \\ -Ax - b \end{pmatrix}. \quad (1.14)$$

Using a standard argument it is easy to show that Problems 1.1 and 1.2 are equivalent, as given in the following theorem.

Theorem 1.1 *A vector $u = (x^\top \ y^\top \ z^\top)^\top \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$ is a solution to Problem 1.1 if and only if it is a solution to Problem 1.2.*

The proof of this theorem can be found in [14].

In what follow we use $\|\cdot\|_p$ to denote the usual l_p -norm on \mathbb{R}^k for any $p \geq 1$ and positive integer k . When $p = 2$, it becomes the usual Euclidean norm. Before further discussion, we make the following assumptions:

A1. $f(x)$ is continuous on \mathbb{R}^m .

A2. $f(x)$ is ξ -monotone, i.e. there exist constants $\alpha > 0$ and $\xi \in (1, 2]$ such that for all $x_1, x_2 \in \mathbb{R}^m$,

$$(x_1 - x_2)^\top (f(x_1) - f(x_2)) \geq \alpha \|x_1 - x_2\|_2^\xi. \quad (1.15)$$

A3. Without loss of generality, we assume that the linear independent constraint qualification (LICQ) holds for (1.11). i.e., $\text{Rank}(A) = n$.

A4. The set of solutions to (1.3)–(1.4) is non-empty.

Note that Assumption A3 is realistic. When $\text{Rank}(A) < n$, some of the constraints in (1.11) are linearly dependent on others and thus they can be eliminated from the constraint set. In the rest of this paper we assume that A1, A2, A3 and A4 are fulfilled. The following theorem, based on Assumption A1, establishes the monotonicity of $G(u)$.

Lemma 1.1 *For any $u_1 = (x_1^\top \ y_1^\top \ z_1^\top)^\top, u_2 = (x_2^\top \ y_2^\top \ z_2^\top)^\top \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$, the mapping G defined in (1.14) satisfies*

$$(u_1 - u_2)^\top (G(u_1) - G(u_2)) \geq \alpha \|x_1 - x_2\|_2^\xi,$$

where α is the constant in (1.15).

PROOF. From the definition of G we have

$$\begin{aligned} & (u_1 - u_2)^\top (G(u_1) - G(u_2)) \\ &= \begin{pmatrix} f(x_1) - f(x_2) - A^\top(y_1 - y_2) + A^\top(z_1 - z_2) \\ A(x_1 - x_2) \\ -A(x_1 - x_2) \end{pmatrix}^\top \begin{pmatrix} x_1 - x_2 \\ y_1 - y_2 \\ z_1 - z_2 \end{pmatrix} \\ &= (f(x_1) - f(x_2))^\top (x_1 - x_2) - (y_1 - y_2)^\top A(x_1 - x_2) + (z_1 - z_2)^\top A(x_1 - x_2) \\ &\quad + (x_1 - x_2)^\top A^\top (y_1 - y_2) - (x_1 - x_2)^\top A^\top (z_1 - z_2) \\ &= (f(x_1) - f(x_2))^\top (x_1 - x_2) \geq \alpha \|x_1 - x_2\|_2^\xi \end{aligned}$$

by (1.15). Thus, the lemma is proved. \square

Using this lemma, we establish the following theorem.

Theorem 1.2 *Problem 1.2 has a unique solution.*

PROOF. From Assumption A4 we see that (1.3)–(1.4) has a solution. Therefore, its KKT condition Problem 1.1, or, equivalently, Problem 1.2, has at least one solution. In what follows we show that the solution is unique.

Let $u_i = (x_i^\top \ y_i^\top \ z_i^\top)^\top$, $i = 1, 2$, be two solutions to Problem 1.2. Then, for $i = 1, 2$, we have

$$(u - u_i)^\top G(u_i) \geq 0, \quad \forall u \in \mathcal{K}. \quad (1.16)$$

Replacing u in (1.16) with u_2 and u_1 respectively for $i = 1, 2$, we get

$$(u_2 - u_1)^\top G(u_1) \geq 0 \quad \text{and} \quad (u_1 - u_2)^\top G(u_2) \geq 0.$$

Therefore, combining these two inequalities gives

$$(u_2 - u_1)^\top (G(u_2) - G(u_1)) \leq 0.$$

Using the monotonicity of G established in Lemma 1.1, we see $\|x_1 - x_2\|_2 = 0$.

We now assume that the two solutions are $u_i = (x^\top \ y_i^\top \ z_i^\top)^\top$, $i = 1, 2$. Since u_1 and u_2 are also solutions to Problem 1.1, they satisfy (1.9), i.e.,

$$f(x) - A^\top y_1 + A^\top z_1 = 0 \quad \text{and} \quad f(x) - A^\top y_2 + A^\top z_2 = 0.$$

From these we have

$$A^\top (y_1 - z_1) = A^\top (y_2 - z_2).$$

Multiplying both sides of the above by A and noticing that AA^\top is invertible by Assumption A3, we have

$$y_1 - z_1 = y_2 - z_2. \quad (1.17)$$

From the KKT conditions (1.10)–(1.11) and the first complementarity in (1.12), it is easy to see that, for $i = 1$ or 2 , if the j th component of y_i is non-zero for some $j \in \{1, 2, \dots, n\}$, i.e., $y_i^j < 0$, then, $(Ax_i)^j = 0$. In this case, the 2nd complementarity condition in (1.12) gives $z_i^j = 0$, since $b^j > 0$. Conversely, if $z_i^j < 0$, then $y_i^j = 0$. Therefore, y_i and z_i satisfies the following complementarity conditions:

$$y_i \leq 0, \quad z_i \leq 0, \quad y_i^\top z_i = 0 \quad \text{for} \quad i = 1, 2. \quad (1.18)$$

In particular, if $y_1^j < 0$ for some $j \in \{1, 2, \dots, n\}$, from (1.17) and (1.18) we have

$$y_1^j = y_2^j - z_2^j.$$

Using (1.18) again we see that the only solution to the above equation is

$$y_1^j = y_2^j < 0 \quad \text{and} \quad z_1^j = z_2^j = 0.$$

By symmetry we also have that, if $z_1^k < 0$ for a $k \in \{1, 2, \dots, n\}$, then, $z_1^k = z_2^k < 0$ and $y_1^k = y_2^k = 0$. Combining these two cases we have $y_1 = y_2$ and $z_1 = z_2$. Therefore, $u_1 = u_2$, and thus this theorem is proved. \square

2 The penalty equation and its solvability

Penalty methods have been used successfully for solving constrained nonlinear optimization problems for decades. In recent years, various penalty methods have been developed for linear and nonlinear complementarity problems in both infinite and finite dimensions [27, 28, 13, 5, 33]. In [14] the authors proposed such a method for a mixed complementarity problem in the case that the nonlinear function involved is ξ -monotone. However, mixed complementarity problems arising from the discretization of continuous obstacle problems with gradient constraints usually do not satisfy the strong monotonicity condition used in [14] so that the results in [14] do not apply. In [29], we extend the penalty method to an obstacle problem with a one-sided derivative constraint. In what follows we shall extend the method further to Problem 1.1 which contains both lower and upper bounds on Ax .

Consider the following problem:

Problem 2.1 Find $u_\lambda = (x_\lambda^\top \quad y_\lambda^\top \quad z_\lambda^\top)^\top \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$, such that

$$G(u_\lambda) + \lambda \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} \equiv \begin{pmatrix} f(x_\lambda) - A^\top y_\lambda + A^\top z_\lambda \\ Ax_\lambda \\ -Ax_\lambda - b \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} = 0, \quad (2.1)$$

where G is the mapping defined in (1.14), $\lambda \geq 1$ and $k > 0$ are penalty parameters, $[v]_+ = \max\{v, 0\}$ and $w^\sigma = (w_1^\sigma, \dots, w_n^\sigma)^\top$ for any $w = (w_1, \dots, w_n)^\top \in \mathbb{R}^n$ and constant $\sigma > 0$.

Problem 2.1 is to find a solution to the nonlinear algebraic equation (2.1) which is a penalty equation approximating (1.9)–(1.12). In the penalty equation, the penalty terms $\lambda[y_\lambda]_+^{1/k}$ and $\lambda[z_\lambda]_+^{1/k}$ penalize the positive components of y_λ and z_λ respectively. In general, the constant k can be any positive number, but in this work we assume that $k \geq 1$. When $k \in (0, 1)$, the penalty terms in (2.1) are smooth and all of our analysis below hold true. However, the method with $k \in (0, 1)$ provides a slow convergence rate, as will be seen in the next section. The solvability of Problem 2.1 is given in the following theorem.

Theorem 2.1 *For any $\lambda \geq 1$, Problem 2.1 has a unique solution.*

PROOF. In Lemma 1.1 we showed that G is monotone. Since $[w]_+$ is monotone in w for any function w , $\lambda \left(0 \quad ([y_\lambda]_+^{1/k})^\top \quad ([z_\lambda]_+^{1/k})^\top \right)^\top$ is also monotone in u_λ . Also, the solution space $\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$ and the usual dot product form a Hilbert space. Therefore, from [11] we see that there exists a solution to Problem 2.1.

To show that the solution to Problem 2.1 is unique, we let $u_{\lambda,j} = (x_{\lambda,j}^\top \quad y_{\lambda,j}^\top \quad z_{\lambda,j}^\top)^\top$, $j = 1, 2$, be two solutions to Problem 2.1, and we show that $u_{\lambda,1} = u_{\lambda,2}$. In fact, from the monotonicity of the function on the left-hand side of (2.1) and Lemma 1.1 we have

$$\alpha \|x_{\lambda,1} - x_{\lambda,2}\|_2^\xi \leq (u_{\lambda,1} - u_{\lambda,2})^\top \left(G(u_{\lambda,1}) - G(u_{\lambda,2}) + \lambda \begin{pmatrix} 0 \\ [y_{\lambda,1}]_+^{1/k} - [y_{\lambda,2}]_+^{1/k} \\ [z_{\lambda,1}]_+^{1/k} - [z_{\lambda,2}]_+^{1/k} \end{pmatrix} \right) = 0.$$

This gives $x_{\lambda,1} = x_{\lambda,2}$.

Since $x_{\lambda,1} = x_{\lambda,2}$, from the 2nd and 3rd blocks of equations in (2.1) it is easy seen that

$$[y_{\lambda,1}]_+ = [y_{\lambda,2}]_+, \quad [z_{\lambda,1}]_+ = [z_{\lambda,2}]_+. \quad (2.2)$$

Therefore, the positive parts of both y_λ and z_λ are uniquely defined.

Now, using the first m scalar equations of (2.1) we have

$$A^\top (y_{\lambda,1} - z_{\lambda,1}) = A^\top (y_{\lambda,2} - z_{\lambda,2}).$$

Multiplying both sides of the above by A and noticing AA^\top is invertible we get

$$y_{\lambda,1} - z_{\lambda,1} = y_{\lambda,2} - z_{\lambda,2}. \quad (2.3)$$

Note, for any feasible j , if the j th component of $y_{\lambda,1}$ satisfies $y_{\lambda,1}^j \leq 0$, from the 2nd and 3rd blocks of equations in (2.1) we see that $z_{\lambda,i}^j > 0$ for $i = 1, 2$. In this case, from (2.3) we have

$$y_{\lambda,1}^j = y_{\lambda,2}^j,$$

since $z_{\lambda,1}^j = z_{\lambda,2}^j$ by (2.2). This implies that $[y_{\lambda,1}]_- = [y_{\lambda,2}]_-$, where $[w]_- := -\min\{w, 0\}$ for any w .

By symmetry, we also have $[z_{\lambda,1}]_- = [z_{\lambda,2}]_-$. Combining these with (2.2) we see that $y_{\lambda,1} = y_{\lambda,2}$ and $z_{\lambda,1} = z_{\lambda,2}$ and therefore, $u_{\lambda,1} = u_{\lambda,2}$. \square

Finally, we note that adding up the 2nd and 3rd blocks of equations in (2.1) gives

$$\lambda \left([y_\lambda]_+^{1/k} + [z_\lambda]_+^{1/k} \right) = b. \quad (2.4)$$

(Recall that the RHS of (2.4) should be $b+c$ when the constraints in (1.4) are transformed into the form $-b \leq Ax \leq 0$ as commented before.) We will use this relation to check the correctness of our numerical solutions late in this work.

3 Convergence

In this section we establish the convergence of the solution to Problem 2.1 to that of Problem 1.2. We first show that the x -component of the solution to Problem 2.1 is bounded for any $\lambda \geq 1$, as given in the following lemma.

Lemma 3.1 *For any $\lambda \geq 1$, let $u_\lambda = (x_\lambda^\top \ y_\lambda^\top \ z_\lambda^\top)^\top$ be a solution to Problem 2.1. Then, there exists a positive constant M , independent of λ and u_λ , such that*

$$\|x_\lambda\|_2 \leq M. \quad (3.1)$$

PROOF. Left-multiplying both sides of (2.1) by u_λ^\top gives

$$u_\lambda^\top G(u_\lambda) + \lambda u_\lambda^\top \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} = 0.$$

Taking $u_\lambda^\top G(0)$ away from both sides of the above equation and re-arranging the resulting equation, we have

$$u_\lambda^\top [G(u_\lambda) - G(0)] = -\lambda \left(y_\lambda^\top [y_\lambda]_+^{1/k} + z_\lambda^\top [z_\lambda]_+^{1/k} \right) - u_\lambda^\top G(0).$$

Using (1.14) and the monotonicity of G established in Lemma 1.1, we have from the above equality

$$\alpha \|x_\lambda\|_2^\xi \leq -x_\lambda^\top f(0) + z_\lambda^\top b - \lambda y_\lambda^\top [y_\lambda]_+^{1/k} - \lambda z_\lambda^\top [z_\lambda]_+^{1/k}. \quad (3.2)$$

But $y_\lambda = [y_\lambda]_+ - [y_\lambda]_-$, and thus

$$-y_\lambda^\top [y_\lambda]_+^{1/k} = -([y_\lambda]_+ - [y_\lambda]_-)^\top [y_\lambda]_+^{1/k} = -[y_\lambda]_+^\top [y_\lambda]_+^{1/k} \leq 0, \quad (3.3)$$

since $[y_\lambda]_-^\top [y_\lambda]_+^{1/k} = 0$. Similarly, we have

$$-z_\lambda^\top [z_\lambda]_+^{1/k} \leq 0. \quad (3.4)$$

Also, since $b > 0$, we have

$$z_\lambda^\top b = [z_\lambda]_+^\top b - [z_\lambda]_-^\top b \leq [z_\lambda]_+^\top b. \quad (3.5)$$

Now, left-multiplying the last block of n equations in (2.1) by $(b^{1/k})^\top$ and rearranging the result give

$$\lambda (b^{1/k})^\top [z_\lambda]_+^{1/k} = (b^{1/k})^\top A x_\lambda + (b^{1/k})^\top b \leq (b^{1/k})^\top b, \quad (3.6)$$

because $A x_\lambda \leq 0$ from the middle block of equations of (2.1) containing $[y_\lambda]_+$. Using the well-known inequality $(\sum_{i=1}^n w_i)^r \leq \sum_{i=1}^n w_i^r$ for any $w_i \geq 0$ and $r \in [0, 1]$, we have from (3.6)

$$\lambda (b^\top [z_\lambda]_+)^{1/k} \leq \lambda (b^{1/k})^\top [z_\lambda]_+^{1/k} \leq (b^{1/k})^\top b = \|b\|_p^p,$$

where $p = 1 + 1/k$. Thus, from the above inequality and (3.5) we have

$$z_\lambda^\top b \leq \frac{\|b\|_p^{1+k}}{\lambda^k},$$

Finally, combining the above estimate, (3.3) and (3.4), we have from (3.2)

$$\alpha \|x_\lambda\|_2^\xi \leq -x_\lambda^\top f(0) + \frac{\|b\|_p^{1+k}}{\lambda^k} \leq C \left(\|x_\lambda\|_2 + \frac{1}{\lambda^k} \right),$$

where C denotes a positive constant, independent of λ . Since $\xi \in (1, 2]$, the above estimate can be written as

$$\|x_\lambda\|_2 \left(\alpha \|x_\lambda\|_2^{\xi-1} - C \right) \leq \frac{C}{\lambda^k}. \quad (3.7)$$

We now consider the following two cases:

Case 1. If $\alpha \|x_\lambda\|_2^{\xi-1} - C \leq 1$, then $\|x_\lambda\|_2 \leq \left(\frac{1+C}{\alpha} \right)^{1/(\xi-1)}$.

Case 2. If $\alpha \|x_\lambda\|_2^{\xi-1} - C > 1$, from (3.7) we see that $\|x_\lambda\|_2 \leq \frac{C}{\lambda^k}$.

Combining these two cases, we have (3.1) for a positive constant M , independent of λ . \square

We now prove the following lemma.

Lemma 3.2 *For any $\lambda \geq 1$, let $u_\lambda = (x_\lambda^\top \ y_\lambda^\top \ z_\lambda^\top)^\top$ be the solution to (2.1). Then, there exists a positive constant C , independent of u_λ and λ , such that*

$$\|[y_\lambda]_+\|_2 + \|[z_\lambda]_+\|_2 \leq \frac{C}{\lambda_e^k} \quad (3.8)$$

where $\lambda_e = \lambda / (\|x\|_q^k + 1)^{1/k}$ with $q = 1 + k$.

PROOF. Left-multiplying (2.1) by $(0^\top \ [y_\lambda]_+^\top \ 0^\top)$ yields

$$[y_\lambda]_+^\top A x_\lambda + \lambda [y_\lambda]_+^\top [y_\lambda]_+^{1/k} = 0.$$

Let $p = 1 + 1/k$ and $q = 1 + k$ satisfying $1/p + 1/q = 1$. Then, using Holder's inequality we have from the above equality

$$\lambda [y_\lambda]_+^\top [y_\lambda]_+^{1/k} \equiv \lambda \|[y_\lambda]_+\|_p^p = -[y_\lambda]_+^\top A x_\lambda \leq C \|[y_\lambda]_+\|_p \|x_\lambda\|_q,$$

from which we have

$$\|[y_\lambda]_+\|_p^{p-1} \leq \frac{C \|x_\lambda\|_q}{\lambda},$$

where C is a generic positive constant, independent of λ and u_λ . Taking the $(p-1)$ -th root on both sides of the above gives

$$\|[y_\lambda]_+\|_p \leq \frac{C \|x_\lambda\|_q^{\frac{1}{p-1}}}{\lambda^{\frac{1}{p-1}}} = \frac{C \|x_\lambda\|_q^k}{\lambda^k}.$$

Since all norms in a finite dimensional space are equivalent, we have

$$\|[y_\lambda]_+\|_2 \leq \frac{C\|x_\lambda\|_q^k}{\lambda^k}. \quad (3.9)$$

Now, left-multiplying (2.1) by $(0^\top \quad 0^\top \quad [z_\lambda]_+^\top)$ and re-arranging the resulting equations, we have

$$\|[z_\lambda]_+\|_p^p = \frac{1}{\lambda} ([z_\lambda]_+^\top Ax_\lambda + [z_\lambda]_+^\top b) \leq \|[z_\lambda]_+\|_p \frac{\|b\|_q}{\lambda}.$$

In the above we used the fact that $Ax_\lambda \leq 0$ and Holder's inequality. Using the same argument for (3.9) and noting all norms in a finite dimensional space are equivalent, we have from the above

$$\|[z_\lambda]_+\|_2 \leq \frac{C}{\lambda^k}.$$

Finally, adding up the above inequality and (3.9) we have

$$\|[y_\lambda]_+\|_2 + \|[z_\lambda]_+\|_2 \leq \frac{C}{\lambda^k} (\|x_\lambda\|_q^k + 1) \leq \frac{C}{\lambda_e^k}$$

with λ_e defined above. Thus, we have proved (3.8). \square

Remark 3.1 *Note that (3.8) implies that $\|[y_\lambda]_+\|_2 + \|[z_\lambda]_+\|_2$ converges to zero at the rate $\mathcal{O}(\lambda^{-k})$ uniformly in λ , since $\|x_\lambda\|_q$ is bounded above uniformly by a positive constant according to Lemma 3.1. However, we leave the upper bound in a form of λ_e which can be regarded as an **effective penalty constant**. Also, though the constant C is independent of u_λ and λ , it does depend on m and n because the equivalence of norms on a finite-dimension space depends on the dimensions of the space.*

We are now ready to prove the following main convergence result.

Theorem 3.1 *For any $\lambda \geq 1$, let $u = (x^\top \quad y^\top \quad z^\top)^\top$ and $u_\lambda = (x_\lambda^\top \quad y_\lambda^\top \quad z_\lambda^\top)^\top$ be the solutions to Problems 1.1 and 2.1 respectively. There exists a constant $C > 0$, independent of u_λ and λ , such that*

$$\|x - x_\lambda\|_2 \leq \frac{C}{\lambda_e^{k/(\xi-1)}}, \quad (3.10)$$

where λ_e is the effective penalty constant defined before. Furthermore, we have

$$\lim_{\lambda \rightarrow \infty} (y - y_\lambda) = 0 = \lim_{\lambda \rightarrow \infty} (z - z_\lambda). \quad (3.11)$$

PROOF. Let C be a generic positive constant, independent of u_λ and λ . We decompose $u - u_\lambda$ into

$$u - u_\lambda = u - \begin{pmatrix} 0 \\ [y_\lambda]_+ \\ [z_\lambda]_+ \end{pmatrix} + \begin{pmatrix} -x_\lambda \\ [y_\lambda]_- \\ [z_\lambda]_- \end{pmatrix} =: r_\lambda - \begin{pmatrix} 0 \\ [y_\lambda]_+ \\ [z_\lambda]_+ \end{pmatrix},$$

where

$$r_\lambda := u + \begin{pmatrix} -x_\lambda \\ [y_\lambda]_- \\ [z_\lambda]_- \end{pmatrix}. \quad (3.12)$$

Since $u - r_\lambda = (x_\lambda^\top \quad -[y_\lambda]_-^\top \quad -[y_\lambda]_-^\top)^\top \in \mathcal{K}$, replacing v in (1.13) with $u - r_\lambda$, we have

$$-r_\lambda^\top G(u) \geq 0. \quad (3.13)$$

Also, left-multiplying both sides of (2.1) by r_λ^\top yields

$$r_\lambda^\top G(u_\lambda) + \lambda r_\lambda^\top \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} = 0. \quad (3.14)$$

Adding up both sides of (3.13) and (3.14) and multiplying the resulting inequality by -1 , we get

$$r_\lambda^\top (G(u) - G(u_\lambda)) - \lambda r_\lambda^\top \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} \leq 0. \quad (3.15)$$

From (3.12) and recalling $w = [w]_+ - [w]_-$ with $[w]_+^\top [w]_- = 0$ for any column vector w , we have

$$\begin{aligned} r_\lambda^\top \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} &= \begin{pmatrix} x - x_\lambda \\ y + [y_\lambda]_- \\ z + [z_\lambda]_- \end{pmatrix}^\top \begin{pmatrix} 0 \\ [y_\lambda]_+^{1/k} \\ [z_\lambda]_+^{1/k} \end{pmatrix} \\ &= (y + [y_\lambda]_-)^\top [y_\lambda]_+^{1/k} + (z + [z_\lambda]_-)^\top [z_\lambda]_+^{1/k} \\ &= y^\top [y_\lambda]_+^{1/k} + z^\top [z_\lambda]_+^{1/k} \\ &\leq 0, \end{aligned}$$

since $y \leq 0$ and $z \leq 0$. Combining the above inequality with (3.15) leads to

$$r_\lambda^\top [G(u) - G(u_\lambda)] \leq 0.$$

Using (3.12) again we see that the above inequality becomes

$$\left[u + \begin{pmatrix} -x_\lambda \\ [y_\lambda]_- \\ [z_\lambda]_- \end{pmatrix} \right]^\top (G(u) - G(u_\lambda)) \leq 0.$$

By the definition of G in Problem 1.2, we have from the above inequality,

$$\begin{aligned} (u - u_\lambda)^\top [G(u) - G(u_\lambda)] &\leq - \begin{pmatrix} 0 \\ [y_\lambda]_+ \\ [z_\lambda]_+ \end{pmatrix}^\top [G(u) - G(u_\lambda)] \\ &= -[y_\lambda]_+^\top A(x - x_\lambda) + [z_\lambda]_+^\top A(x - x_\lambda). \end{aligned}$$

Using Lemma 1.1 and (3.8), we have from the above

$$\alpha \|x - x_\lambda\|_2^\xi \leq \|A\|_2 (\|[y_\lambda]_+\|_2 + \|[z_\lambda]_+\|_2) \|x - x_\lambda\|_2 \leq \frac{C}{\lambda_e^k} \|x - x_\lambda\|_2,$$

Therefore, (3.10) follows from the above inequality.

Let us now show (3.11). Since u and u_λ are solutions to, respectively, (1.9) and (2.1), taking the difference between the two equations gives

$$A^\top (y - z - y_\lambda + z_\lambda) = f(x) - f(x_\lambda),$$

from which we have

$$(y - y_\lambda) - (z - z_\lambda) = (AA^\top)^{-1} A(f(x) - f(x_\lambda)), \quad (3.16)$$

since A is of rank n by Assumption A3. From (1.18) we have that if the j th component of y satisfies $y^j < 0$, then $z^j = 0$. Therefore, when $y^j < 0$, the j th component of the above equality becomes

$$y^j - y_\lambda^j + z_\lambda^j = [(AA^\top)^{-1} A(f(x) - f(x_\lambda))]^j. \quad (3.17)$$

Now, from the proof of Theorem 2.1 we see that if $y_\lambda^j \leq 0$, then $z_\lambda^j > 0$, and if $z_\lambda^j \leq 0$, then $y_\lambda^j > 0$. Therefore, we have the following cases when λ is sufficiently large.

Case 1. $y_\lambda^j \leq 0$.

In this case, we have from (3.17)

$$\begin{aligned} |y^j - y_\lambda^j| &\leq \left| [(AA^\top)^{-1} A(f(x) - f(x_\lambda))]^j \right| + |[z_\lambda^j]_+| \\ &\leq \left| [(AA^\top)^{-1} A(f(x) - f(x_\lambda))]^j \right| + \frac{C}{\lambda_e^k} \end{aligned}$$

by (3.8). Since f is continuous by Assumption A1, from (3.10) we have

$$\lim_{\lambda \rightarrow \infty} |y^j - y_\lambda^j| = 0. \quad (3.18)$$

Case 2. $y_\lambda^j > 0$ and $z_\lambda^j \leq 0$.

In this case, from the 3rd block of equations in (2.1) we have $(Ax_\lambda)^j = -b^j < 0$ for any $\lambda \geq 1$. Since $y^j < 0$, from the first complementarity condition in (1.12) we have $(Ax)^j = 0$. Clearly, when $\lambda \rightarrow \infty$, these two equations violate the facts that $b^j > 0$ and $x_\lambda \rightarrow x$. Thus, $y_\lambda^j > 0$ will never happen when λ approaches ∞ .

Case 3. $y_\lambda^j > 0$ and $z_\lambda^j > 0$.

From (2.4) we have

$$\lim_{\lambda \rightarrow \infty} y_\lambda^j = 0 = \lim_{\lambda \rightarrow \infty} z_\lambda^j, \quad (3.19)$$

and this, along with (3.17) violates the assumption that $y^j < 0$.

Combining the above three cases we have that when $y^j < 0$, (3.18) holds.

By symmetry, we have that, when $z^j < 0$ for a feasible j , $\lim_{\lambda \rightarrow \infty} |z^j - z_\lambda^j| = 0$. Also, when $y^i = 0 = z^j$, from (3.17) we see that $-y_\lambda^j + z_\lambda^j \rightarrow 0$ as $\lambda \rightarrow \infty$. As commented above, from the proof of Theorem 2.1 we have if $y_\lambda^j \leq 0$, then $z_\lambda^j > 0$, and if $z_\lambda^j \leq 0$, then $y_\lambda^j > 0$. Therefore, if $y_\lambda^j \leq 0$ or $z_\lambda^j \leq 0$, y_λ^j or z_λ^j have the same sign, and thus from the above limit we see that (3.19) holds. Furthermore, if both $y_\lambda^j > 0$ and $z_\lambda^j \leq 0$, from (2.4) we also have (3.19). Therefore, we have proved (3.11). \square

In the case that f is Hölder continuous, we have the following convergence results for y_λ and z_λ :

Corollary 3.1 *For any $\lambda \geq 1$, let $u = (x^\top \ y^\top \ z^\top)^\top$ and $u_\lambda = (x_\lambda^\top \ y_\lambda^\top \ z_\lambda^\top)^\top$ be solutions to Problems 1.1 and 2.1, respectively. If f is Hölder continuous on \mathbb{R}^m , i.e., there exist constants $\beta > 0$ and $\gamma \in (0, 1]$ such that*

$$\|f(x_1) - f(x_2)\|_2 \leq \beta \|x_1 - x_2\|_2^\gamma, \quad \forall x_1, x_2 \in \mathbb{R}^m, \quad (3.20)$$

then, we have

$$\|y - y_\lambda\|_2 + \|z - z_\lambda\|_2 \leq \frac{C}{\lambda_e^{k\gamma/(\xi-1)}}, \quad (3.21)$$

where C is a positive constant, independent of λ and u_λ . Furthermore, when f is strongly monotone and Lipschitz continuous, i.e, $\xi = 2$ and $\gamma = 1$ in (1.15) and (3.20) respectively, we have

$$\|u - u_\lambda\|_2 \leq \frac{C}{\lambda_e^k}. \quad (3.22)$$

PROOF. Let C be a generic constant, independent of λ and u_λ . From (3.16) and (3.20) we have

$$\|y - y_\lambda - (z - z_\lambda)\|_2 = \|(AA^\top)^{-1}A\|_2 \|f(x_\lambda) - f(x)\|_2 \leq C\beta \|x_1 - x_2\|_2^\gamma.$$

Using an argument similar to the proof of (3.11) presented above and (3.10), we have (3.21).

When $\xi = 2$ and $\gamma = 1$, combining (3.10) and (3.21) we have (3.22). \square

4 Numerical experiments

In this section we present some numerical experimental results to demonstrate the theoretical findings in the previous sections. We also use these results to show the practicality

We also choose $N = 100$ ($h = 0.01$) and consider the solution of the penalty equation (2.1) corresponding to the above finite-dimensional problem. Note that the penalty equation is nonlinear, we use a damped Newton's method to solve it. Also, the penalty term is non-smooth, and thus in the computations, the function $[w]_+^{1/k}$ is smoothed out locally using the following formula in [13]:

$$\phi(w) = \begin{cases} w^{\frac{1}{k}}, & w \geq \varepsilon, \\ \varepsilon^{\frac{1}{k}-2}(3 - \frac{1}{k})w^2 + \varepsilon^{\frac{1}{k}-3}(\frac{1}{k} - 2)w^3, & w < \varepsilon \end{cases}$$

for any w , where ε is a small positive constant. In all the computational results given below, we choose $\varepsilon = 10^{-12}$. Replacing $[y_\lambda]_+^{1/k}$ and $[z_\lambda]_+^{1/k}$ in (2.1) by $\phi(y_\lambda)$ and $\phi(z_\lambda)$ respectively, the Jacobian matrix of the nonlinear function on the LHS of (2.1) is given by

$$J_G(u_\lambda) := \begin{pmatrix} B & -A & A \\ A^\top & 0 & 0 \\ -A^\top & 0 & 0 \end{pmatrix} + \text{diag}(3x_\lambda^\top, \lambda \nabla \phi(y_\lambda)^\top, \lambda \nabla \phi(z_\lambda)^\top),$$

where $\text{diag}(\dots)$ denotes the diagonal matrix and ∇ is the gradient operator with respect to the independent variable of ϕ . Note that the first part of J_G is a singular matrix. Thus, it is important that the initial guess in the Newton's method is chosen so that $[y_\lambda]_+ \neq 0$ and $[z_\lambda]_+ \neq 0$. (In what follows we will omit the subscript λ .)

Let us first investigate the computed convergence rates of the method in λ for a fixed k . Since the exact solution to (4.1) is unknown, we use the numerical solution with $k = 2$ and $\lambda = 10^{10}$ as the 'exact' or reference solution u^* . Table 4.1 is a list of the computed errors in the l_2 -norm $\|u - u^*\|_2$ for different values of λ and k and the ratios between two consecutive values of λ . From (3.22) it is easy to see that the theoretical ratio for two consecutive values of λ for a fixed k is equal to $\lambda_{i+1}^k / \lambda_i^k = 2^k$. From Table 4.1, we see that our computed ratios match this theoretical one well for all $k = 1, 2, 3$ and 4, except when the numerical solution is too close to the reference solution.

To investigate numerical rates of convergence of the numerical solutions in k , we choose the sequence $k = 1, 2, \dots, 6$ for a given λ and evaluate the ratios of the errors corresponding to k and $k + 1$ for $k = 1, 2, \dots, 5$. These ratios are listed in Table 4.2 for different values of λ . From (3.10) and (3.21) we see that theoretically the ratio of the errors for k and $k + 1$ is a constant $\lambda_e^{k+1} / \lambda_e^k = \lambda_e$. From Table 4.2 we see that the ratios are almost constants, coinciding with the theoretical result. The only exception is when $\lambda = 800$ and $k = 6$ in which the numerical solution is too close to the reference solution.

To further demonstrate the performance of the penalty method, we plot the computed u and u_{UNC} in Figure 4.1(a). We also plot u' , u'_{UNC} and the constraints g_1 and g_2 in Figure 4.1(b) from which we see that u is bounded below by g_1 and above by g_2 . Figure 4.1(c) contains the computed multipliers y and z , along with the sum of the two penalty terms

	$\lambda = \frac{5^{2-k} \times 2^i}{h^2}$	$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
$k = 1$	Errors	1.03e-03	5.15e-04	2.58e-04	1.29e-04	6.44e-05	3.22e-05
	Ratios	–	2.00	2.00	2.00	2.00	2.00
$k = 2$	Errors	2.67e-06	6.67e-07	1.67e-07	4.17e-08	1.04e-08	2.60e-09
	Ratios	–	4.00	4.00	4.00	4.00	4.00
$k = 3$	Errors	1.88e-07	2.35e-08	2.93e-09	3.67e-10	4.60e-11	7.99e-12
	Ratios	–	8.00	8.00	8.00	7.97	5.76
$k = 4$	Errors	3.44e-07	2.15e-08	1.34e-09	8.41e-11	7.93e-12	4.66e-12
	Ratios	–	16.0	16.0	16.0	10.6	1.70

Table 4.1: Computed rates of convergence in λ for different values of k .

	$k =$	1	2	3	4	5	6
$\lambda = 100$	Errors	4.96e-1	2.56e-2	1.41e-3	8.13e-5	4.79e-6	2.88e-7
	Ratios	–	19.4	18.1	17.4	17.0	16.7
$\lambda = 200$	Errors	2.49e-1	6.40e-3	1.77e-4	5.08e-6	1.50e-7	4.50e-9
	Ratios	–	38.9	36.2	34.8	33.9	33.3
$\lambda = 400$	Errors	1.26e-1	1.60e-3	2.21e-5	3.18e-7	4.68e-9	7.05e-11
	Ratios	–	78.5	72.5	69.6	67.8	66.4
$\lambda = 800$	Errors	6.31e-2	4.00e-4	2.76e-6	1.98e-8	1.46e-10	6.25e-12
	Ratios	–	158	145	139	136	23.4

Table 4.2: Computed rates of convergence in k for different values of λ .

in (2.1). From Figure 4.1(c) we see that $y = 0 = z$ when the constraints are inactive, and either y or $z < 0$ when Ax touches either the upper or lower bound, i.e., when either of the two constraints is active. Figure 4.1(c) also shows that (2.4) is satisfied by y and z .

5 Conclusion

In this work we proposed a penalty approach to the discretized form of a double obstacle problem with derivative constraints. In this approach, we first write down the KKT conditions of the discrete problem which form a mixed nonlinear complementarity problem. The complementarity problem is then approximated by a nonlinear penalty equation with a term penalizing the parts of the solution violating the constraints. We have proved that the penalty equation is uniquely solvable and the solution to the penalty equation converges exponentially to that of the mixed complementarity of variational (or the inequality) problem. A non-trivial numerical example was solved using this method to demonstrate the rates of convergence of the method.

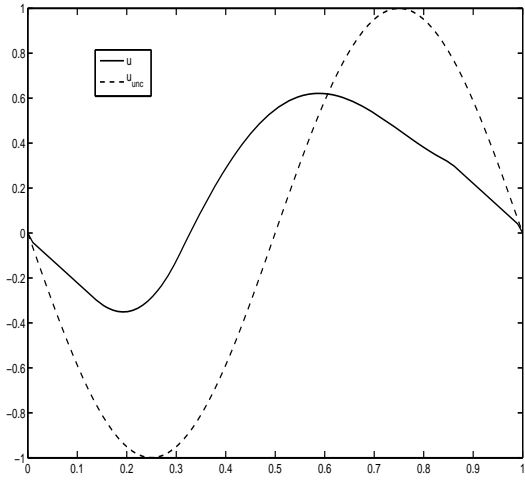
References

- [1] A. Azevedo, F. Miranda, L. Santos, Variational and quasivariational inequalities with first order constraints, *Journal of Mathematical Analysis & Applications*, **397**, 738–756 (2013).
- [2] J.W. Barrett and L. Prigozhin, A quasi-variational inequality problem in superconductivity, *Mathematical Models & Methods in Applied Sciences*, **20**, 679–706 (2010).
- [3] P. Cardaliaguet, A double obstacle problem arising in differential game theory, *Journal of Mathematical Analysis & Applications*, **360**, 95–107 (2009).
- [4] C.H. Chen and O. L. Mangasarian, A Class of Smoothing Functions for Nonlinear and Mixed Complementarity Problems, *Computational Optimization and Application*, **5**, 97-138 (1996).
- [5] M. Chen and C. Huang, A power penalty method for the general traffic assignment problem with elastic demand, *Journal of Industrial and Management Optimization*, **10**, 1019–1030 (2014).
- [6] H.J. Choe and Y. Shim, Degenerate variational inequalities with gradient constraints, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.(4)*, **22**, 25–53 (1995).
- [7] A. Damgaard, Computation of Reservation Prices of Options with Proportional Transaction Costs, *Journal of Economic Dynamics and Control*, **30**, 415–444 (2006).

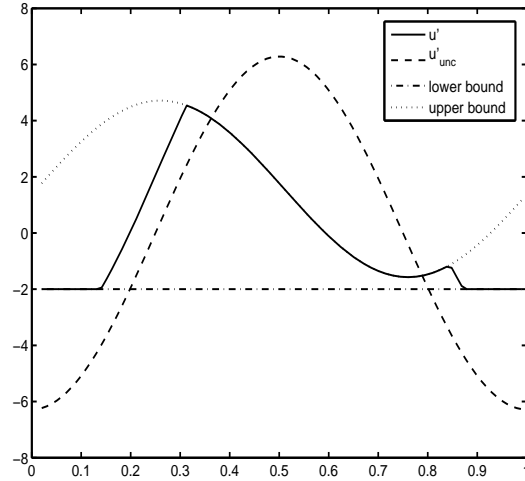
- [8] A.N. Daryina, A. F. Izmailov and M. V. Solodov, A Class of Active-Set Newton Methods for Mixed Complementarity Problems, *SIAM Journal on Optimization*, **36**, 409-429 (2004).
- [9] M.H.A. Davis and T. Zariphopoulou, American Options and Transaction Fees in *Mathematical Finance* (eds. M.H.A. Davis et al.), Springer-Verlag (1995).
- [10] F. Facchinei and J.S. Pang, *Finite-dimensional Variational Inequalities and Complementarity Problems, Vol. I & II*, Springer Series in Operations Research, Springer-Verlag, New York (2003).
- [11] N.S. Hoang and A.G. Ramm, An iterative scheme for solving nonlinear equations with monotone operators, *BIT*, **48**, 725–741 (2008).
- [12] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer-Verlag, Berlin, Heidelberg, 1996.
- [13] C.C. Huang and S. Wang, A Power Penalty Approach to a Nonlinear complementarity problem, *Operations Research Letters*, **38**, 72–76 (2010).
- [14] C.C. Huang and S. Wang, A penalty method for a mixed nonlinear complementarity problem, *Nonlinear Analysis*, **75**, 588–597 (2012).
- [15] C.S. Huang, S. Wang, K.L. Teo, On application of an alternating direction method to Hamilton-Jacobi-Bellman equations, *Journal of Computational & Applied Mathematics*, **166**, 153–166 (2004).
- [16] C. Kanzow, Global Optimization Techniques for Mixed Complementarity Problems, *J. Glob. Optim.*, **16**, 1-21 (2000).
- [17] M. Kunze and J.F. Rodrigues, An elliptic quasi-variational inequality with gradient constraints and some of its applications, *Mathematical Methods in the Applied Sciences*, **23**, 897–908 (2000).
- [18] D.C. Lesmana and S. Wang, An upwind finite difference method for a nonlinear BlackScholes equation governing European option valuation under transaction costs, *Applied Mathematics & Computation*, **219**, 8811–8828 (2013).
- [19] D. Li and M. Fukushima, Smoothing Newton and Quasi-Newton Methods for Mixed Complementarity Problems, *Computational Optimization and Applications*, **17**, 203-230 (2000).

- [20] W. Li and S. Wang, Penalty approach to the HJB equation arising in European stock option pricing with proportional transaction Costs, *Journal of Optimization Theory and Applications*, **143** 279–293 (2009).
- [21] W. Li and S. Wang, Pricing American options under proportional transaction costs using a penalty approach and a finite difference scheme, *Journal of Industrial and Management Optimization*, **9**, 365–398 (2013).
- [22] W. Li and S. Wang, A numerical method for pricing European options with proportional transaction costs, *J. Glob. Optim.*, **60**, 59–78 (2014).
- [23] O.L. Mangasarian and M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, *Mathematical Programming* **62**, 277–297 (1993).
- [24] R.D.C. Monteiro and Jong-Shi Pang, Properties of an Interior-Point Mapping for Mixed Complementarity Problems, *Mathematics of Operations Research*, **21**, 629–654 (1996).
- [25] L. Santos, Variational problems with non-constant gradient constraints, *Port. Math. (N.S.)*, **59**, 205–248 (2002).
- [26] S. Wang, A novel fitted finite volume method for the Black-Scholes equation governing option pricing, *IMA J. Numer. Anal.*, **24**, 699–720 (2004).
- [27] S. Wang and X.Q. Yang, A power penalty method for linear complementarity problems, *Operations Research Letters*, **36**, 211–214 (2008).
- [28] S. Wang, X.Q. Yang and K.L. Teo, Power penalty method for a linear complementarity problem arising from American option valuation, *Journal of Optimization Theory & Applications*, **129**, 227–254 (2006).
- [29] S. Wang, A power penalty method for a finite-dimensional obstacle problem with derivative constraints, *Optim. Lett.*, **8**, 1799–1811 (2014).
- [30] N. Yamashita and M. Fukushima, On stationary points of the implicit Lagrangian for nonlinear complementarity problems. *Journal of Optimization Theory and Applications*, **84**, 653–663 (1995).
- [31] V. Zakamouline, European option pricing and hedging with both fixed and proportional transaction costs, *J. Econ. Dyn. Control*, **30**, 1–25 (2006).
- [32] K. Zhang and S. Wang, Pricing American bond options using a penalty method, *Automatica*, **48**, 472–479 (2012).

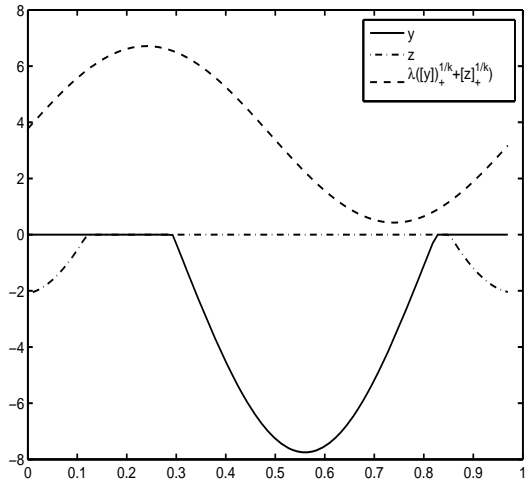
- [33] Y.Y. Zhou, S. Wang and X.Q. Yang, A penalty approximation method for a semilinear parabolic double obstacle problem, *J. Glob. Optim.*, in press, DOI 10.1007/s10898-013-0122-6.



(a) u and u_{unc}



(b) u' , u'_{unc} , g_1 and g_2



(c) multipliers y , z and $\lambda([y]_+^{1/k} + [z]_+^{1/k})$

Figure 4.1: Computed solutions using $\lambda = 10^5$ and $k = 2$.