An interior penalty method for a large-scale finite-dimensional nonlinear double obstacle problem

Song Wang

Department of Mathematics & Statistics, Curtin University, GPO Box U1987, Perth, Australia

Abstract

We propose and analyze an interior penalty method for a finite-dimensional large-scale bounded Nonlinear Complementarity Problem (NCP) arising from the discretization of a differential double obstacle problem in engineering. Our approach is to approximate the bounded NCP by a nonlinear algebraic equation containing a penalty function with a penalty parameter $\mu > 0$. The penalty equation is shown to be uniquely solvable. We also prove that the solution to the penalty equation converges to the exact one at the rate $\mathcal{O}(\mu^{1/2})$ as $\mu \to 0$. A smooth Newton method is proposed for solving the penalty equation and it is shown that the linearized system is reducible to two decoupled subsystems. Numerical experiments, performed on some non-trivial test examples, demonstrate the computed rate of convergence matches the theoretical one.

Keywords: Variational inequality, double obstacle problem, nonlinear complementarity problem, interior penalty method, convergence.

1. Introduction

Double obstacle problems appear in many areas such as physics, engineering, game theory, investment and financial engineering (cf., for example, [4, 7, 3, 10, 8]). A typical example of double obstacle problems is the following optimization problem in an infinite-dimensional setting

$$\min_{u \in H} \mathcal{F}(u) \quad \text{subject to} \quad u_* \leq u \leq u^*, \tag{1}$$

Email address: song.wang@curtin.edu.au (Song Wang)

where H is a partially ordered functional space, $\mathcal{F}$ is a functional on H involving usually differential operators, and u_* and u^* are known functions in a proper function space satisfying $u_* < u^*$. In general, an infinite-dimensional double obstacle problem can be stated as follows.

Problem 1.1. *Find a pair (u, v) in a proper function space such that*

$$\mathcal{L}(u) + v \leq 0, \quad u - u^* \leq 0, \quad (\mathcal{L}(u) + v)(u - u^*) = 0, \tag{2}$$

$$u_* - u \leq 0, \quad v \leq 0, \quad v(u_* - u) = 0, \tag{3}$$

where $\mathcal{L}$ denotes a nonlinear (differential) operator and v is a multiplier.

(A more rigorous statement of the above problem is beyond the scope of this work and we refer interesting readers to [44].) Mathematically, Problem 1.1 may be the optimality (KKT) conditions (or Euler-Lagrange equations for unconstrained problems) for (1) via Calculus of Variation in which v can be regarded as a multiplier. However, Problem 1.1 represents a larger class of problem than that of the form (1) as it includes double obstacle problems which may not be formulated as an optimization problem of the form (1). Problems of this type often appear in financial engineering such as pricing European and American options under proportional transaction costs [10, 8, 24, 25, 26].

Problem 1.1 is an infinite-dimensional problem and is not solvable analytically unless for some trivial cases. Thus, numerical approximations to it are usually sought in practice. To solve Problem 1.1 numerically, two different types of methods are needed – schemes for the discretization of (2)–(3) and methods for solving the discretized problem. Various methods can be found in the open literature for the discretization of differential LCPs, such as those in [6, 40, 32, 12, 22, 25, 39]. Application of a proper discretization scheme to (2)–(3) usually yields the following n-dimensional double obstacle problem.

Problem 1.2. *Find $x, y \in \mathbb{R}^n$ such that*

$$\begin{aligned} f(x) + y &\leq 0, && (4)\\ x &\leq c, && (5)\\ (x - c)^\top (f(x) + y) &= 0, && (6) \end{aligned}$$

and

$$\begin{aligned} b - x &\leq 0, && (7)\\ y &\leq 0, && (8)\\ y^\top (b - x) &= 0, && (9) \end{aligned}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ and b and c are known n-dimensional vectors satisfying $b < c$ which define the lower and upper bounds on x.

Without loss of generality, we assume that $c = 0$ and $b < 0$, as the general case can be transformed into this one by a simple translation. In the special case that there exists a $\phi(x) : \mathbb{R}^n \mapsto \mathbb{R}$ such that $f(x) = \nabla\phi(x)$ (or f is conservative), it is easy to show that (4)–(9) are the 1st-order optimality (KKT) conditions for the minimization problem $\min_{b \le x \le 0} \phi(x)$.

Problem 1.2, containing a set of two complementarity problems, is equivalent to the bounded Nonlinear Complementarity Problem (NCP) discussed in [16] in which a numerical procedure for the problem is also proposed. Various other numerical methods have also been developed for complementarity problems such as those in [14, 21, 27, 9, 30]. Obviously, existing numerical methods may be used or extended for solving Problem 1.2. However, most of the existing methods have been developed for generic NCPs and they are not normally designed for large-scale problems. On the other hand, (4)–(9) arises from the discretization of (2)–(3) by a discretization scheme which is usually designed in such a way so that the resulting nonlinear mapping f has some special properties such as sparsity and strong monotonicity, and its Jacobian matrix has positive diagonal and non-negative off-diagonal entries and is diagonally dominant. These properties can be exploited in the computation of approximate solutions to Problem 1.2, particularly when it is large-scale, as will be commented later in this work. Therefore, it is desirable to design algorithms for Problem 1.2 which allows full use of these special properties in computation.

Recently, penalty methods have been used extensively for solving both finite- and finite-dimensional NCPs arising in classic and financial engineering [2, 17, 28, 37, 18, 35, 19, 20, 43, 33, 26, 5, 23, 34]. More specifically, a power penalty method is proposed for Problem 1.2 in [36] in which a penalty term is used to penalize the infeasible components of an approximation to the relaxed unconstrained problem. This power penalty method has the merits that it has an exponential convergence rates, is simple to implement and does not introduce any new local minima. However, the power penalty methods are 'exterior' in the sense that some of the constraints, more specifically some of (5) and (8), can only be satisfied approximately, not strictly, by the solutions from these power penalty methods. An empirical interior penalty method in infinite-dimensions for a linear complementariy problem (LCP) arising in option pricing is proposed in [28] and a convergence analysis for a similar

method in an infinite-dimensional functional space setting is conducted in [42]. Motivated by the logarithmic barrier functions used extensively for general linear and nonlinear programming problems [15], an interior penalty method is proposed and analyzed in [38] for an LCP.

In this work, we propose an interior penalty approach to the numerical solution of Problem 1.2 in which (4)–(9) is approximated by a nonlinear algebraic equation with a penalty term linked to a barrier function often used for solving nonlinear constrained optimization problem. As demonstrated in our previous work [38], unlike most of the existing interior methods for NCPs (cf., for example, [30, 13]), the method we propose in this work does not require any auxiliary functions such as potential or merit functions. It is known that the introduction of auxiliary functions in solving NCPs often results in a global optimization problem even when the original problem is monotone [21, 41]. In fact, we are able to show that the penalty equation has a unique solution under some conditions which are usually satisfied by the discretized forms of (2)–(3) using one of the aforementioned discretization schemes. A convergence theory for the approximate solution will also be established. Moreover, we show that the penalty equation from this method inherits properties from the original problem such as sparsity and monotonicity. The rest of this paper is organised as follows.

In Section 2 we will formulate the the penalty equation and establish the unique solvability of the penalty equation. We prove in Section 3 both weak and strong convergence of the approximate solutions to the exact one. In Section 4, we will first design a Newton's method for the penalty equation. We then show that the linearized system can be decomposed into two decoupled subsystems by a block Gaussian elimination and prove their system matrices are M-matrices. In Section 5, we use some numerical experimental results to verify the theoretical rate of convergence and to demonstrate the usefulness of the methods for solving practical problems.

2. The interior penalty method

Let z and function w be defined respectively by

$$z = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad w(z) = \begin{pmatrix} f(x) + y \\ b - x \end{pmatrix}. \tag{10}$$

Problem 1.2 can then be written as the following unbounded NCP (recall $c = 0$):

Problem 2.1. *Find $z \in \mathbb{R}^{2n}$ such that*

$$\begin{aligned} w(z) &\leq 0, & (11)\\ z &\leq 0, & (12)\\ z^\top w(z) &= 0. & (13) \end{aligned}$$

Interior methods have been used very successfully for solving constrained optimization problems and an excellent overview of these methods can be found in [15]. These methods guarantee that a solution from such a method satisfies the constraints in question strictly. Motivated by the interior penalty formulation in [38] for an LCP, we propose to use the nonlinear equation defined below to approximate Problem 2.1.

$$W(z_\mu) := w(z_\mu) - \mu./z_\mu = 0, \tag{14}$$

where μ is a (small) positive constant. $z_\mu = (x_\mu^\top, y_\mu^\top)^\top \in \mathbb{R}^{2n}$ and $u./v$ denotes (Hadamard) element by element division of two matrices of the same size with the exception that $u./v = (u/v_1, ..., u/z_m)^\top$ if u is a scalar and $v \in \mathbb{R}^m$. Any negative solution z_μ to (14) defines an approximation to the solution z of Problem 2.1. We expect that $z_\mu \to z$ when $\mu \to +0$. Before further discussion, we first make some assumptions on the mapping f.

A discretization scheme such as any of [32, 12, 22, 25] is designed so that the nonlinear coefficient matrix of the resulting system is strongly monotone, but usually not symmetric. Therefore, it is reasonable to make the following assumptions on the mapping f:

A1. f is strongly monotone, i.e., there exists a constant $\alpha > 0$, independent of n, such that

$$(x_1 - x_2)^\top (f(x_1) - f(x_2)) \geq \alpha \|x_1 - x_2\|_2^2, \tag{15}$$

for any $x_1, x_2 \in \mathbb{R}^n$, where $||\cdot||_2$ denotes the l_2 (Euclidean) norm on $\mathbb{R}^m$ for any positive integer m.

A2. $f(x)$ is continuously differentiable on $\mathbb{R}^n$.

We comment that, in practice, Assumption A1 is usually satisfied by a discretized system if an appropriate discretization scheme is used, as we have exampled in [23] using an NCP arising from the discretization of a nonlinear Black–Scholes equation. Using the above assumptions we are able to show the following lemma.

Lemma 2.2. *For any given $0 < \mu \leq \mu_0$ with μ_0 a positive constant, if $z_\mu \in \mathbb{R}^{2n}$ is a solution to (14) such that $z_\mu < 0$, then z_μ satisfies*

$$\|z_\mu\|_2 \leq M, \tag{16}$$

where $M > 0$ is a constant, independent of μ.

Proof. Let $z_\mu = (x_\mu^\top, y_\mu^\top)^\top$, where $x_\mu, y_\mu \in \mathbb{R}^n$. Left-multiplying (14) by $(x_\mu^\top, 0^\top)$ gives

$$x_\mu^\top f(x_\mu) + x_\mu^\top y_\mu - \mu n = 0,$$

since $x_\mu^\top (1./x_\mu) = n$. From this equation we obtain

$$x_\mu^\top (f(x_\mu) - f(0)) + x_\mu^\top y_\mu = \mu n - x_\mu^\top f(0).$$

Because $x_\mu, y_\mu < 0$, we have $x_\mu^\top y_\mu > 0$. Thus, using this observation and (15) we have from the above equation

$$\alpha \|x_\mu\|_2^2 \leq \mu n + \|x_\mu\|_2 \|f(0)\|_2.$$

Using the ε-inequality with $\varepsilon = \alpha$, we have from the above estimate

$$\alpha \|x_\mu\|_2^2 \leq \mu n + \frac{\alpha}{2}\|x_\mu\|_2^2 + \frac{1}{2\alpha}\|f(0)\|_2^2$$

from which we thus have

$$\|x_\mu\|_2^2 \leq \frac{2}{\alpha}\left(\mu_0 n + \frac{1}{2\alpha}\|f(0)\|_2^2\right),$$

since $\mu \leq \mu_0$. Taking the square root on both sides of the above we have $\|x_\mu\|_2 \leq C_1$ for some constant $C_1 > 0$, independent of μ.

We now consider y_μ. Left-multiplying (14) by $(x_\mu^\top, y_\mu^\top)$ gives

$$x_\mu^\top f(x_\mu) + y_\mu^\top b - 2\mu n = 0.$$

Note that $y_\mu < 0$ and $b < 0$ is a constant. We have from the above equation

$$(\min_i |b_i|)\|y_\mu\|_1 \leq y_\mu^\top b = 2\mu n - x_\mu^\top f(x_\mu) \leq 2\mu_0 n + \|x_\mu\|_2 \|f(x_\mu)\|_2, \tag{17}$$

where $\|\cdot\|_1$ denotes the usual l_1-norm on $\mathbb{R}^n$. From the first part of this proof we have $\|x_\mu\|_2 \leq C_1$. Also, from Assumption A2 we see that f is continuous which, combined with the boundedness of x_μ, implies that $\|f(x_\mu)\|_2$ is bounded uniformly in μ. Thus, we see from (17) that $\|y_\mu\|_1$ is also bounded above by a positive constant. Since all norms on $\mathbb{R}^n$ are equivalent, we see that $\|y_\mu\|_2 \leq C_2$ for some positive constant C_2, independent of μ. Combining this with $\|x_\mu\|_2 \leq C_1$ obtained above we have (16). □

Using Lemma 2.2, we are able to show that (14) has a unique negative solution in the following theorem.

Theorem 2.3. *For any given $\mu \in (0, \mu_0]$, where μ_0 is a given positive constant, there exists a unique solution z_μ to (14) satisfying $z_\mu < 0$.*

Proof. Let us first show if the set of solutions to (14) is non-empty, then it has only one element. For simplicity, we ignore the subscript μ in the rest of the proof. Suppose both $z_1 < 0$ and z_2 are solutions to (14). Then, they satisfy

$$w(z_1) - w(z_2) - \mu(1./z_1 - 1./z_2) = 0.$$

Left-multiplying the above equation by $(z_1 - z_2)^\top$ gives

$$(z_1 - z_2)^\top (w(z_1) - w(z_2)) - \mu(z_1 - z_2)^\top (1./z_1 - 1./z_2) = 0. \tag{18}$$

Let $z_k = (x_k^\top, y_k^\top)^\top$ for $k = 1, 2$. Using the definition of w in (10) we see that

$$\begin{aligned}(z_1 - z_2)^\top (w(z_1) - w(z_2)) &= (x_1 - x_2)^\top (f(x_1) - f(x_2)) + (x_1 - x_2)^\top (y_1 - y_2) \\ &\quad + (y_1 - y_2)^\top (x_1 - x_2) \\ &= (x_1 - x_2)^\top (f(x_1) - f(x_2)) \\ &\geq \alpha \|x_1 - x_2\|_2^2 \end{aligned} \tag{19}$$

by (15). Also,

$$\begin{aligned}(z_1 - z_2)^\top (1./z_1 - 1./z_2) &= (z_{1,1} - z_{2,1}, ..., z_{1,n} - z_{2,n}) \left(\frac{1}{z_{1,1}} - \frac{1}{z_{2,1}}, ..., \frac{1}{z_{1,n}} - \frac{1}{z_{2,n}} \right)^\top \\ &= - \sum_{i=1}^{n} \frac{(z_{1,i} - z_{2,i})^2}{z_{1,i} z_{2,i}}.\end{aligned}$$

Thus, using this equality and (19), we have from (18)

$$\alpha ||x_1 - x_2||_2^2 + \mu \sum_{i=1}^{n} \frac{(z_{1,i} - z_{2,i})^2}{z_{1,i} z_{2,i}} \leq 0.$$

Since $z_1 < 0$ and $z_2 < 0$, the above inequality implies that $z_1 = z_2$. Therefore, the solution to (14) is unique.

We now show that (14) has a solution. To prove this, let $S = \{z \in \mathbb{R}^{2n} : -\varepsilon^{-1} e < z < -\delta e\}$, where $e = (1, ..., 1)^\top \in \mathbb{R}^{2n}$ and ε and δ are (small)

positive constants. Clearly, $W : \bar{S} \subset \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is continuous, where W is defined in (14). We now show $0 \notin W(\partial S)$ when both $\varepsilon > 0$ and $\delta > 0$ are sufficiently small, where ∂S denotes the boundary of S. This is contained in the following two cases.

Case 1. We let $\partial S_1 \subset \partial S$ such that for any $z \in \partial S_1$, there is an $i \in \{1, ..., 2n\}$ such that $z_i = -\delta$. If $0 \in W(\partial S_1)$, then, there exists a $z \in \partial S_1$ satisfying

$$W_k(z) = w_k(z) + \frac{\mu}{\delta} = 0, \text{ or, } w_k(z) = -\mu/\delta \tag{20}$$

for a feasible k. Since w is continuous by Assumption A2 and z is bounded uniformly in μ by (16), we have

$$\|w(z)\|_2 \leq C,$$

where C denotes a generic positive constant, independent of μ. Clearly, from (20) we see that $w(z)$ violates the above boundedness of $w(z)$ when $\delta > 0$ is sufficiently small. Therefore, $0 \notin W(\partial S_1)$.

Case 2. Similarly, we let $\partial S_2 \subset \partial S$ such that for any $z \in \partial S_2$, there is an $i \in \{1, ..., 2n\}$ such that $z_i = -\varepsilon^{-1}$. If $0 \in F(\partial S_2)$, then, there is an $z \in S_2$ with $z_k = -\varepsilon^{-1}$ for a feasible k such that $W_k(z) == 0$. In fact this is not possible when ε is sufficiently small, as otherwise, z violates (16).

Combining the above two cases we see that when ε and δ are both sufficiently close to $+0$, $0 \notin W(\partial C)$. Furthermore, from (10) we have

$$\nabla W(z) = \begin{pmatrix} \nabla f(x) & I_n \\ -I_n & 0 \end{pmatrix} + \mu \text{diag}((1./z)./z), \tag{21}$$

where I_n denotes the $n \times n$ identity matrix and 0 is the $n \times n$ zero matrix. Since $f(x)$ is strongly monotone and continuous differentiable by Assumptions A1 and A2, from [29, p.142] we see that $\nabla f(x)$ is positive-definite. Thus, for any $u, v \in \mathbb{R}^n$ and fixed $z < 0$, it is easy to see that

$$(u^\top \ v^\top) \nabla W(z) \begin{pmatrix} u \\ v \end{pmatrix} = u^\top \nabla f(x) u + \mu \sum_{i=1}^{n} \left(\frac{u_i^2}{x_i^2} + \frac{v_i^2}{y_i^2} \right) \geq C(||u||_2^2 + ||v||_2^2)$$

for some positive constant C. Therefore, $\nabla W(z)$ is positive-definite. Thus, it is non-singular at any point and the degree of W, defined as $\text{sign}(\det(\nabla F))$, at any point is non-zero. Therefore, by the Kronecker Theorem, $W(z) = 0$ has a solution [29, p.161].

Thus we have proved the theorem. □

3. Convergence

While we have shown (14) has a unique negative solution, in this section we prove the solution converges to that of (11)–(13) when $\mu \to +0$. We first prove the weak convergence of the solution in the following theorem.

Theorem 3.1. *For any $\mu > 0$, the solution $z_\mu < 0$ to (14) satisfies the NCP (11)–(13) as $\mu \to +0$.*

Proof. Since $z_\mu < 0$ and $\mu > 0$, from (14) we have

$$w(z_\mu) = \mu./z_\mu < 0.$$

Left-multiplying both sides of (14) by $z_\mu^\top$ and rearranging the resulting equation give

$$z_\mu^\top w(z_\mu) = \mu z_\mu^\top (1./z_\mu) = 2\mu n \to 0 \tag{22}$$

as $\mu \to +0$. Therefore, z_μ satisfies (11)–(13) when $\mu \to +0$. □

Note that in the proof of Theorem 3.1 we have not explicitly used Assumptions A1 and A2. This theorem does not provide any convergence rate for z_μ. However, if we use Assumption A1, we are able to establish the rate of convergence for z_μ. We start this discussion with rewriting (11)–(13) as the following variational inequality problem.

Let $\mathcal{K} = \{x \in \mathbb{R}^n : x \leq 0\}$ and denote $\mathcal{K}^2 = \mathcal{K} \times \mathcal{K} \subset \mathbb{R}^{2n}$. We also use $\mathring{\mathcal{K}}$ and $\mathring{\mathcal{K}}^2$ to denote the sets of interior points of $\mathcal{K}$ and $\mathcal{K}^2$ respectively. It is known (see, for example, [13, 20, 36]) that Problem 1.2 is equivalent to the following variational inequality problem:

Problem 3.2. *Find $z^* \in \mathcal{K}^2$ such that, for all $z \in \mathcal{K}^2$,*

$$(z - z^*)^\top w(z^*) \geq 0. \tag{23}$$

Using this variational inequality, we establish the rate of convergence for z_μ in the following theorem.

Theorem 3.3. *Let Assumptions A1 and A2 be fulfilled. For any $\mu \in (0, \mu_0]$, where μ_0 is a sufficiently small positive constant, let $z^* \in \mathcal{K}^2$ and $z_\mu \in \mathring{\mathcal{K}}^2$ be the solutions to Problem 3.2 and (14) respectively. Then, there exists a positive constant C, independent of μ, such that z^* and z_μ satisfy*

$$||z^* - z_\mu||_2 \leq C\sqrt{\mu}. \tag{24}$$

Proof. Let $z^* = \begin{pmatrix} x^* \\ y^* \end{pmatrix}$ and $z_\mu = \begin{pmatrix} x_\mu \\ y_\mu \end{pmatrix}$, where $x^*, y^* \in \mathcal{K}$ and $x_\mu, y_\mu \in \mathring{\mathcal{K}}$. For simplicity, we assume $\mu_0 \ll 1$, as we are only interested in the behaviour of w_μ when $\mu \to +0$.

Since $z_\mu \in \mathring{\mathcal{K}}^2 \subset \mathcal{K}^2$, replacing z in (23) with z_μ gives

$$(z_\mu - z^*)^\top w(z^*) \geq 0. \tag{25}$$

Left-multiplying (14) by $(z_\mu - z^*)^\top$, we have

$$(z_\mu - z^*)^\top w(z_\mu) - (z_\mu - z^*)^\top (\mu./z_\mu) = 0. \tag{26}$$

Subtracting both sides of (26) from the corresponding sides of (25) yields

$$(z_\mu - z^*)^\top (w(z^*) - w(z_\mu)) + \mu(z_\mu - z^*)^\top (1./z_\mu) \geq 0.$$

From this inequality we have

$$\begin{aligned} (z_\mu - z^*)^\top (w(z^*) - w(z_\mu)) &\leq \mu(z_\mu - z^*)^\top (1./z_\mu) \\ &= \mu(2n - (z^*)^\top)(1./z_\mu)) \\ &\leq 2\mu n, \end{aligned} \tag{27}$$

since $-(z^*)^\top (1./z_\mu)) < 0$. Therefore, using (19) we have from the above estimate

$$\alpha ||x^* - x_\mu||_2^2 \leq (z_\mu - z^*)^\top (w(z^*) - w(z_\mu)) \leq 2\mu n. \tag{28}$$

We now establish the rate of convergence for y_μ. Without loss of generality, we assume that $x^* = ((x_N^*)^\top, x_O^\top)^\top$ where $x_N^* \in \mathbb{R}^m$ and $x_O^* \in \mathbb{R}^{n-m}$ for some non-negative integer m satisfying $x_N^* < 0$ and $x_O^* = 0$. Correspondingly, we also partition y^*, x_μ, y_μ and f in the same way. From the complementarity condition (6) (recall $c = 0$) and (14) we have

$$f_N(x^*) + y_N^* = 0, \tag{29}$$

$$f_N(x_\mu) + y_{\mu,N} - \mu./x_{\mu,N} = 0. \tag{30}$$

Taking both side of (29) from the corresponding sides of (30) gives

$$f_N(x_\mu) - f_N(x^*) + y_{\mu,N} - y_N^* - \mu./x_{\mu,N} = 0$$

from which we have

$$y_N^* - y_{\mu,N} = f_N(x_\mu) - f_N(x^*) - \mu./x_{\mu,N}. \tag{31}$$

Without loss of generality, we now further assume that there is a non-negative integer m_1 such that $|x_{\mu,N,i}| \geq \mu_0$ for $i = 1, 2, ..., m_1$ and $|x_{\mu,N,i}| < \mu_0$ for $i = m_1 + 1, ..., m$. We then consider these two cases separately below.

Case 1. $|x_{\mu,N,i}| \geq \mu_0$ for $i = 1, 2, ..., m_1$.

Let $x_{\mu,N_1} := (x_{\mu,N,1}, ..., x_{\mu,N,m_1})^\top$ and other variables can be decomposed in the same way. Left-multiplying both sides of the first m_1 equations of (31) by $(y^*_{N_1} - y_{\mu,N_1})^\top$ and noting f is continuously differentiable, we have

$$\begin{aligned} \|y^*_{N_1} - y_{\mu,N_1}\|_2^2 &= (y^*_{N_1} - y_{\mu,N_1})^\top (f_{N_1}(x_\mu) - f_{N_1}(x^*)) - \mu(y^*_{N_1} - y_{\mu,N_1})^\top (1./x_{\mu,N_1}) \\ &\leq C\|y^*_{N_1} - y_{\mu,N_1}\|_2 \left(\|x^* - x_\mu\|_2 + \mu\sqrt{n}\right) \\ &\leq C\|y^*_{N_1} - y_{\mu,N_1}\|_2 \sqrt{n\mu}, \end{aligned} \tag{32}$$

since $\|\mu./x_{\mu,N_1}\|_2 \leq \sqrt{n}\mu/\mu_0$, where C denotes a generic positive constant, independent of μ. From this inequality we have

$$\|y^*_{N_1} - y_{\mu,N_1}\| \leq Cn\sqrt{\mu}, \tag{33}$$

In the above deduction we used Cauchy-Schwarz inequality, the mean-value theorem in calculus and (28).

Case 2. $|x_{\mu,N,i}| < \mu_0$ for $i = m_1 + 1, ..., m$.

Let $x_{\mu,N_2} := (x_{\mu,N,m_1+1}, ..., x_{\mu,N,m})^\top$ and other variables and functions can be decomposed in the same way. In this case,

$$|b_i - x_{\mu,N,i}| \geq |b_i| - |x_{\mu,N,i}| \geq \min_i |b_i| - \mu_0, \;\; i = m_1 + 1, ..., m. \tag{34}$$

Left-multiplying (14) by $(0, ..., y^\top_{\mu,N_2}, 0, ..., 0)$ we have

$$y^\top_{\mu,N_2}(b_{N_2} - x_{\mu,N_2}) - (m - m_1)\mu = 0$$

from which, together with (34), we have

$$(m - m_1)\mu \geq (\min_i |b_i| - \mu_0)\|y_{\mu,N_2}\|_1, \tag{35}$$

where $\|\cdot\|_1$ denotes the usual l_1-norm. Since $x_\mu \to x^*$ as $\mu \to +0$ by (28), we also have $b_{N_2} - x^*_{N_2} < 0$ when μ_0 is sufficiently small. Thus, from (9) we see that $y^*_{N_2} = 0$. Combining this with (35) we have

$$\|y^*_{N_2} - y_{\mu,N_2}\|_1 = \|y_{\mu,N_2}\|_1 \leq \frac{(m - m_1)\mu}{\min_i |b_i| - \mu_0} \leq Cn\mu.$$

Since all norms on a finite-dimensional space are equivalent, we have $\|y^*_{N_2} - y_{\mu,N_2}\|_2 \le Cn\mu$. Combining this with (33) we have, when μ_0 is sufficiently small,

$$\|y^*_N - y_{\mu,N}\|_2 \le C\sqrt{n\mu}. \tag{36}$$

We now consider the convergence of $y_{\mu,O}$. Since $b_O - x^*_O = b_O < 0$, from (9) we see that $y^*_O = 0$. Furthermore, it is easy to see from (28) that, when μ_0 is sufficiently small, $|x_{\mu,O,i}| < \mu_0$ for $i = m+1, ..., n$. Therefore, following the proof of Case 2 above we have that $\|y^*_O - y_{\mu,O}\|_2 \le Cn\mu$. Finally, combining this with (28) and (36) we obtain (24). Thus, we have proved the theorem. □

4. Newton's method for (14)

For the solution of (14)in $\mathring{\mathcal{K}}^2$, we propose the following smooth Newton algorithm for (14).

Algorithm Newton

Step 1. Choose $\mu > 0$, $\varepsilon > 0$ and $\delta > 0$ sufficiently small and an initial guess $z^0 \in \mathbb{R}^n$ such that $z^0 < 0$. Let $k := 0$.

Step 2. Solve the following linear system for p^k:

$$\nabla W(z^k)p^k = -W(z^k), \tag{37}$$

where ∇W is given in (21).

Step 3. Set

$$z^{k+1} = \min\{z^k + p^k, -\delta e\},$$

where $e = (1, 1, ..., 1)^\top \in \mathbb{R}^{2n}$ as defined before in the proof of Theorem 2.3.

Step 4. If

$$\max\left\{\|z^{k+1} - z^k\|_2,\ \left|w^\top(z^{k+1})z^{k+1} - 2\mu n\right|\right\} \le \varepsilon,$$

then stop. Otherwise, let $k := k+1$ and go to Step 2.

Note in each Newton's iteration, the $2n \times 2n$ linear system (37) needs to be solved. However, the solution of this $2n \times 2n$ system reduce to the solution of an $n \times n$ linear system as given below.

To simplify our notation, we introduce, for any $u \in \mathbb{R}^n$, a diagonal matrix $D(u) := \mathrm{diag}((1./u)./u) = \mathrm{diag}(u_1^{-2}, u_2^{-2}, ..., u_n^{-2})$. Using (21), we write (37) as

$$\nabla W(z^k)p^k = \begin{pmatrix} \left[\nabla f(x^k) + \mu D(x^k)\right] p_x^k + p_y^k \\ -p_x^k + \mu D(y^k) p_y^k \end{pmatrix} = -W(z^k), \tag{38}$$

where $z^k = ((x^k)^\top, (y^k)^\top)^\top$ and $p^k = ((p_x^k)^\top, (p_y^k)^\top)^\top$. Decomposing $W(z^k)$ into $W(z^k) = (W_x^\top(z^k), W_y^\top(z^k))^\top$, we have from the 2nd block of equations in (38)

$$p_x^k = \mu D(y^k) p_y^k + W_x(z^k). \tag{39}$$

Substituting this into the 1st block of equations in (38) and rearranging the resulting equation we have

$$\left[I_n + \mu(\nabla f(x^k) + \mu D(x^k))D(y^k)\right] p_y^k = -\left[I_n + \nabla f(x^k) + \mu D(x^k)\right] W_x(z^k). \tag{40}$$

We can also rewrite (40) as

$$\left[D^{-1}(y^k) + \mu(\nabla f(x^k) + \mu D(x^k))\right] \bar{p}_y^k = -\left[I_n + \nabla f(x^k) + \mu D(x^k)\right] W_x(z^k), \tag{41}$$

$$p_y^k = D^{-1}(y^k)\bar{p}_y^k. \tag{42}$$

Therefore, we may replace Step 2 in Algorithm Newton with

Step 2* Solve (39) and (41)–(42) (or (40)) for p_y^k and p_x^k and set $p^k = ((p_x^k)^\top, (p_y^k)^\top)^\top$.

The advantage of splitting (40) into (41)–(42) is that the system matrix of (41) inherits any properties possessed by the Jacobian ∇f, as both D and D^{-1} are diagonal matrices with positive diagonal elements. More specifically, the system matrix of (41) has the following properties.

Theorem 4.1. *For any $\mu > 0$ and $z = (x^\top, y^\top)^\top \in \mathbb{R}^{2n}$ such that $z_i \neq 0$ for $i = 1, ..., 2n$, the system matrix on the RHS of (41), denoted as $A(z, \mu)$, satisfies the following properties.*

1. *$A(z, \mu)$ is positive definite.*
2. *If $\nabla f(x)$ is irreducibly diagonally dominant with positive diagonal and non-positive off-diagonal entries, then $A(z, \mu)$ is an M-matrix.*

Proof. Since $f(x)$ is strongly monotone and continuous differentiable by Assumptions A1 and A2, from [29, p.142] we see that $\nabla f(x)$ is positive-definite. Furthermore, from the definition of D we have that both $D(x)$ and $D^{-1}(y)$ are diagonal matrices with positive diagonal entries. Therefore, $A(z,\mu)$ is positive definite.

We now show A is an M-matrix. Note that both $D(x)$ and $D^{-1}(y)$ are diagonal matrices with positive diagonal elements. When $\nabla f(x)$ is irreducibly diagonally dominant, it is obvious that $A(z,\mu)$ is also irreducibly diagonally dominant. By, [31], $A(z,\mu)$ is an M-matrix. □

We comment that in practice ∇f is usually a sparse M-matrix as many discretization schemes such as those in [40, 32, 12, 1, 22, 26] are designed to guarantee this property is satisfied by f in (4). Theorem 4.1, particularly Item 2, has computational implication. Note that (41) is usually large-scale, particularly when the infinite-dimensional 2)–(3) is defined in multiple spatial or state dimensions, so that methods such as Gaussian Elimination or LU decomposition are computationally very expensive. Thus, Item 2 of Theorem 4.1 allows us to use iterative methods for large-scale sparse linear systems such as the preconditioned conjugate gradient type methods to solve (41) efficiently.

5. Numerical experiments

In this section we apply our numerical methods to two test problems with respectively a symmetric and an un-symmetric system matrices. The parameters in Step 1 of Algorithm Newton are chosen to be $\varepsilon = 10^{-10}$ and $\delta = 10^{-14}$. All experiments were carried out under Matlab environment.

In the following numerical experiments, we will estimate the rates of convergence of our method by solving each of the problems using a sequence of μ obtained by halving the previous value of μ. From Theorem 3.3 we see that, theoretically, the ratio of the errors corresponding to μ and $\mu/2$ is approximately $\|z^* - z_\mu\|_2/\|z^* - z_{\mu/2}\|_2 \approx \sqrt{2}$. However, from Algorithm Newton we see that, in computation, $|z_{\mu,i}| \geq \delta$ for all feasible i and thus from the first inequality in (27) and (19) we have

$$\alpha\|x^* - x_\mu\|_2^2 \leq \frac{C\mu}{\delta}(\|x^* - x_\mu\|_2^2 + \|y^* - y_\mu\|_2^2)^{1/2} \leq \frac{C\mu}{\delta}(\|x^* - x_\mu\|_2 + \|y^* - y_\mu\|_2) \tag{43}$$

where C denotes a *generic* positive constant, independent of μ and δ. Using (32) and following the proof for the error bound on $\|y^* - y_\mu\|_2$, we can easily show that

$$\|y^* - y_\mu\|_2 \le C(\|x^* - x_\mu\|_2 + \mu). \tag{44}$$

Combining this inequality and (43) and using the ε-inequality with $\varepsilon = \alpha$ we have

$$\alpha\|x^* - x_\mu\|_2^2 \le \frac{C\mu}{\delta}(\|x^* - x_\mu\|_2 + \mu) \le \frac{C\mu^2}{\alpha\delta^2} + \frac{\alpha}{2}\|x^* - x_\mu\|_2^2 + \frac{C\mu^2}{\delta}.$$

From the above inequality and after some manipulation, we obtain $\|x^* - x_\mu\|_2 \le \frac{C}{\delta}\mu$. Combining this and (44) we also have $\|y^* - y_\mu\|_2 \le \frac{C}{\delta}\mu$ for some positive constant C. Thus, we may expect that the ratio $\|z^* - z_\mu\|_2/\|z^* - z_{\mu/2}\|_2 \approx \mu/(\mu/2) = 2$. In computation, the ratio $\|z^* - z_\mu\|_2/\|z^* - z_{\mu/2}\|_2$ is bounded below by $\sqrt{2}$ and above by 2, as will be seen in the following tests.

Test 1. Problem 1.1 with the following nonlinear operator and functions:

$$\mathcal{L}(u) = -\nabla^2 u + u^3 - g,\ u_* = -s - t,\ u^* = 6[(s-0.5)^2 + (t-0.5)^2],$$
$$g = 4\pi\sin(2\pi s)(1 - 5\cos(4\pi t)) + \sin(2\pi s)^3(1 - \cos(4\pi t))^3$$

for $(s,t) \in \Omega = (0,1)\times(0,1)$ satisfying $u = 0$ on the boundary of Ω. The test problem, originally proposed in [11] with a different function g, is used in [36] and the unconstrained problem has the exact solution $u_{\text{unc}} = \sin(2\pi s)[1 - \cos(4\pi t)]$.

To solve this problem numerically, we first partition Ω into a uniform mesh with nodes $(s_i, t_j) = (ih, jh)$ for $i, j = 0, 1, ..., N$ for a positive integer N, where $h = 1/N$. Then, the central-finite difference scheme and the a one-point quadrature rule on the mesh are used to discretize the operator $\mathcal{L}$ and the given bounds. The resulting complementarity problem is of the form (4)–(9) with $f(x) = Ax + x.*x.*x - q$, where A is an $n \times n$ symmetric and positive-definite matrix and $.*$ denotes the element-by-element matrix multiplication (used in Matlab). The coefficient matrices A, b, c and q are the discretized forms of $-\nabla^2$, u_*, u^* and g using the schemes with the natural order of the nodes. Note that in this case $n = (N-1)\times(N-1)$ in Problem 1.2 which is the total number of interior nodes of the mesh. For this test, we choose $N = 50$. Since A is symmetric and positive-definite, Problem 1.2 is the KKT conditions of a constrained optimization problem as given in [36].

The finite-dimensional complementarity problem is first transformed into one with $c = 0$ in Problem 1.2 and then the corresponding penalty equation

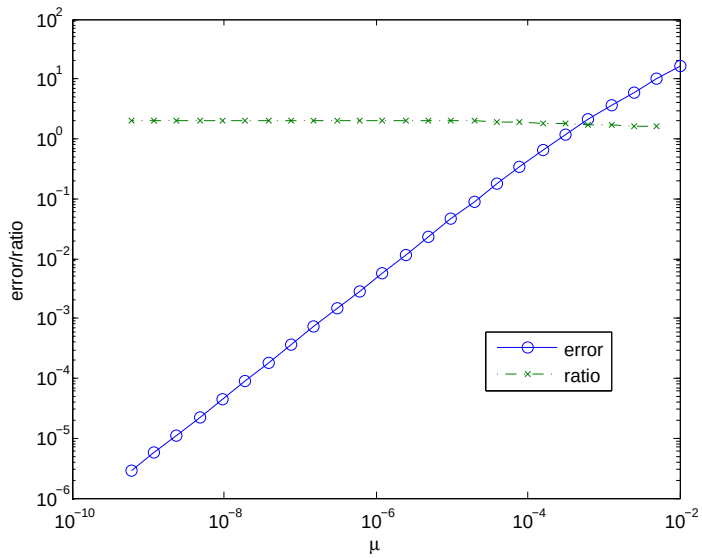


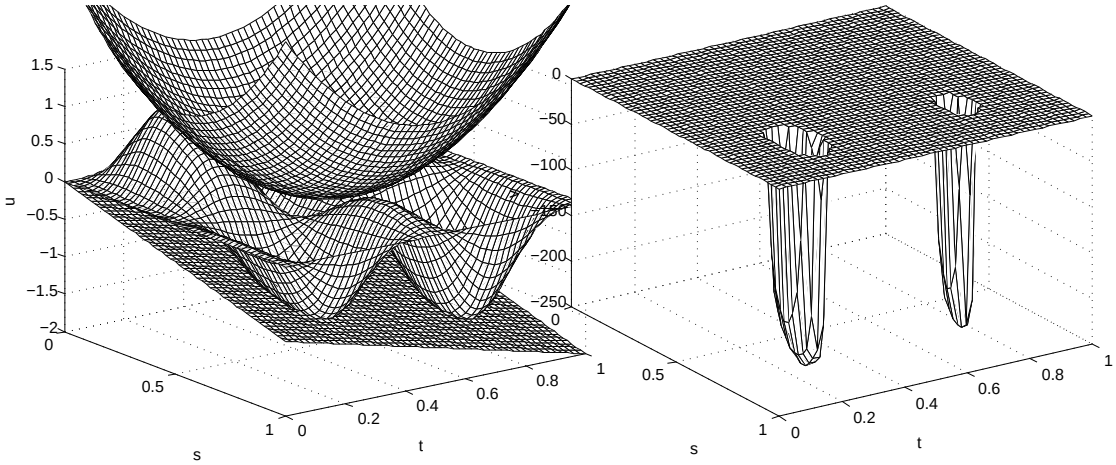
Figure 1: Computed errors and their ratios for Test 1.

(14) is solved. Since the exact solution is unknown, we use the numerical solution from this method with $\mu = 10^{-14}$ as the 'exact' or reference solution. Using this reference solution, we compute the errors in l^2-norm in the numerical solutions for $\mu = 10^{-2}/2^m, m = 0, 1, ..., 24$ and calculate the ratios of the errors from two consecutive values of μ. Our calculation shows that the ratio is initially 1.64. It increases to 2 as μ decreases and becomes stable around 2. This is in line our analysis earlier in this section. The computed errors and ratios are plotted in Figure 1 in log-log form from which we see that the computed rates of convergence are mostly of the order $\mathcal{O}(\mu)$. To further demonstrate the correctness of the numerical solution, we plot the numerical solution (u, v) to Problem 1.1, along with the lower and upper bounds u_* and u^* in Figure 2.

Test 2. Problem 1.1 with the following operator and functions:

$$\mathcal{L}(u) = -\nabla^2 u + \frac{\partial u}{\partial s} + \frac{\partial u}{\partial t} - g,$$
$$u_* = 0.7 - 10 * ((s - 0.7)^2 + (t - 0.3)^2), \quad u^* = 0.2 + |x - 0.25| + |y - 0.75|,$$
$$g = 5[6st(2 - s^2 - t^2) + (1 - 3s^2)(t - t^3) + (1 - 3t^2)(s - s^3)]$$

for $(s, t) \in \Omega := (0, 1)^2$ satisfying $u = 0$ on the boundary of Ω. The exact solution to the unconstrained problem $\mathcal{L}(u) = 0$ is $u_{\text{unc}} = 5(s - s^3)(t - t^3)$.



(a) u, u_* and u^* (b) v

Figure 2: Computed constrained solution u and multiplier v for Test 1.

This is an infinite-dimensional LCP in which $\mathcal{L}$ consists of a 2nd order (diffusion) operator and a first-order (convection) one. To solve this infinite-dimensional problem numerically, we first choose the same uniform mesh with 51×51 nodes as used for Test 1 above. The conventional central difference scheme and upwind difference scheme are used to approximate the diffusion and convection terms in $\mathcal{L}(u)$ respectively [25, 22, 26]. All other terms are approximated by their respective values at a mesh node. This discretization method results in a bounded LCP of the form (4)–(9) with $f(x) = Ax - q$, where A is an $n \times n$ unsymmetric matrix with $n = 49^2$. As for Test 1, the matrices A, b, c and q in Problem 1.2 are discretized forms of $-\nabla^2$, u_*, u^* and g. It is easy to verify numerically that A is a positive-definite matrix. Unlike the case in Test 1, A is unsymmetric and may not be the KKT condition of a constrained optimization problem.

Again, since $c \neq 0$, we first transform Problem 1.2 into one with $c = 0$ and then consider the solution of (14). To compute the error in the solution to (14), we need the exact solution which is unknown. As in Test 1, we use the numerical solution with $\mu = 10^{-14}$ as the 'exact' or reference solution and solve (14) using $\mu = 10^{-2}/2^m$ for $m = 0, 1, ..., 24$. The computed errors for the sequence of μ and the ratios of the errors from two consecutive values of μ are plotted in Figure 3 in log-log form from which we see that the ratio

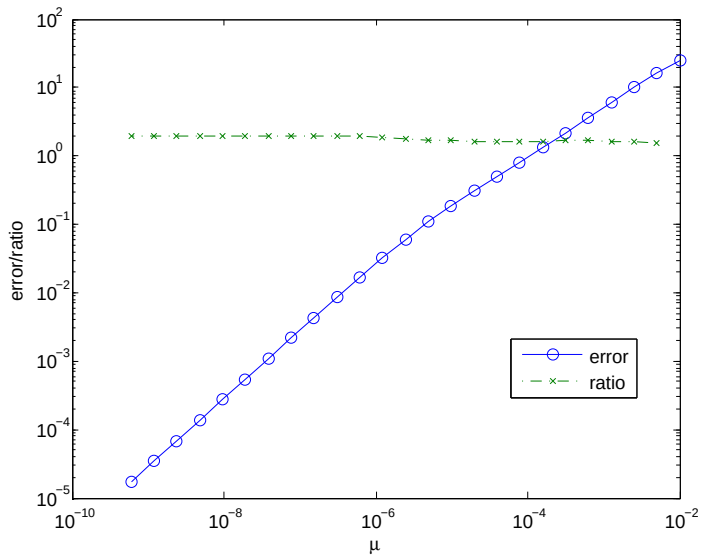


Figure 3: Computed errors and their ratios for Test 2.

is in between $\sqrt{2}$ and 2. In fact, the computed results show that the ratio is initially 1.54 when $\mu \approx 10^{-2}/2$ and increases to 2 as μ deceases from 10^{-2} to $10^{-2}/2^{19}$ and becomes stable when μ further decreases.

To gauge the influence of the size n of the problem on the rate of convergence, we have discretized the continuous problem using the uniform meshes with $N = 25, 50, 100$ and 150. Correspondingly, $n = 576, 2401, 9801$ and 22201. We also use the numerical solutions of the discretized problems with $\mu = 10^{-14}$ as the reference solutions and compute errors of the numerical solutions of each of the discretized systems using $\mu = 10^{-3}/2^m$ for $m = 0, 1, ..., 24$. The ratios of the errors in Matlab log-log form are plotted in Figure 4. From these results we see that n does affect the computed ratios, particularly when μ is away from 10^{-14}. This shows that n does affect the rate of convergence. However, when μ approaches 10^{-14}, the ratio converges to 2. Nevertheless, all the computed ratios are above the theoretical one $\sqrt{2}$.

To demonstrate the effectiveness of the method further, we plot the computed solution u and v, along with u_* and u^* in Figure 5. Figure 5(a) shows clearly that both of the constraints are satisfied by the computed u. Figure 5(b) demonstrates that multiplier v is non-zero when the lower bound constraint becomes active and zero when it is inactive.

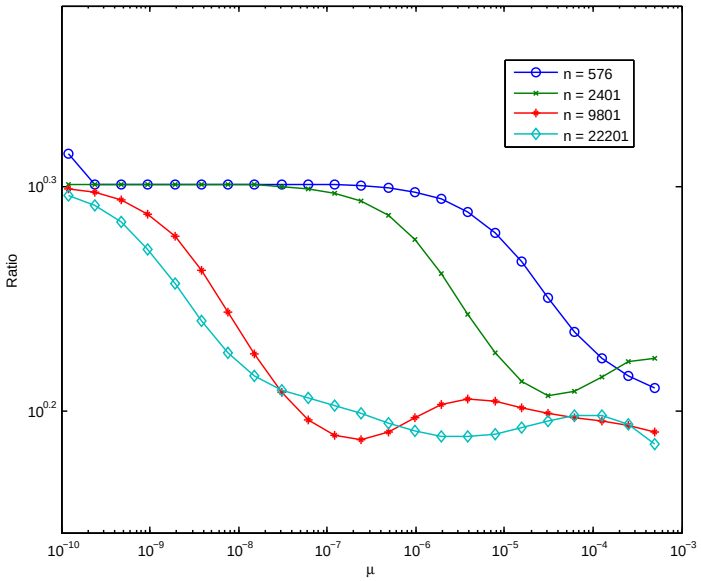
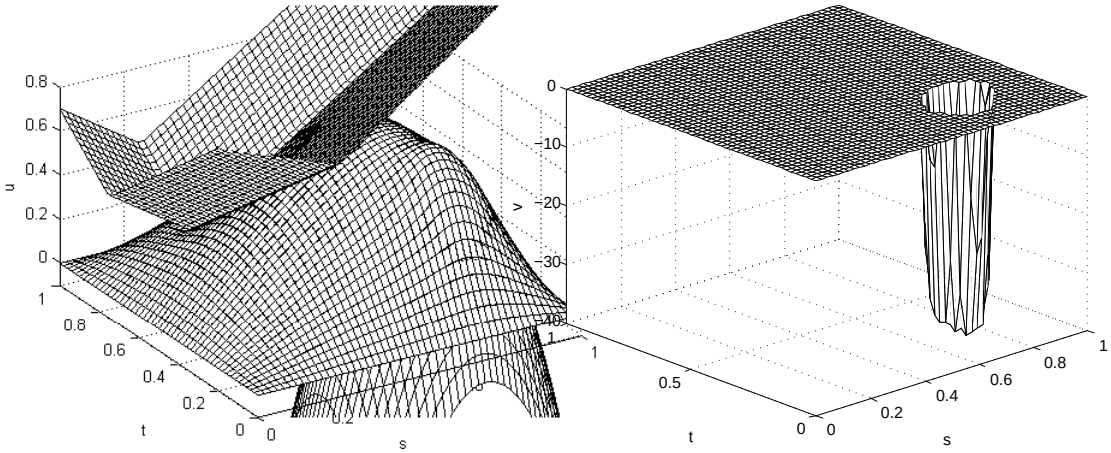


Figure 4: Computed ratios for different values of n for Test 2.



(a) u, u_* and u^* (b) v

Figure 5: Computed constrained solution u and multiplier v for Test 2.

6. Conclusions

In this work we have developed am interior penalty method for a large-scale, bounded NCP arising from the discretization of a double obstacle problem in engineering. The unique solvability of the penalty equation in the interior of the feasible region has been proved and a convergence theory for the approximate solution has been established. We have proposed a smooth Newton method for the nonlinear penalty equation and showed that the linearized system can be decomposed into two decoupled subsystems of which the coefficient matrices are both M-matrices. Numerical experiments using non-trivial differential double obstacle problems have been carried out. Numerical results presented have demonstrated that the computed rates of convergence are consistent with the theoretical one and that the method provide a efficient and accurate computational for double obstacle problems in both classic and financial engineering.

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. Bensoussan and J.L. Lions, *Applications of Variational Inequalities in Stochastic Control* North-Holland, Amsterdam-New York-Oxford (1978).

[3] P. Cardaliaguet, A double obstacle problem arising in differential game theory, *J. Math. Anal. Appl.*, **360** (2009) 95–107.

[4] M.M. Cecchi, R. Russo, Numerical approximation of free boundary problem by variational inequalities. Application to semiconductor devices, in *Variational Analysis & Applications, eds. F. Giannessi and A. Maugeri*, Springer, New York (2005).

[5] W. Chen, S. Wang, A penalty method for a fractional order parabolic variational inequality governing American put option valuation, *Computers and Mathematics with Applications*, 67 (2014) 77–90.

[6] G. Courtadon, A more accurate finite difference approximation for the valuation of options, *J. Financial Economics Quant. Anal*, 17 (1982) 697-703.

[7] M. Dai, F. Yi, Finite-horizon optimal investment with transaction costs: A parabolic double obstacle problem, *Journal of Differential Equations*, 246 (2009) 1445–1469 .

[8] A. Damgaard, Computation of Reservation Prices of Options with Proportional Transaction Costs, *Journal of Economic Dynamics and Control*, 30 (2006) 415–444 .

[9] 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** (2004) 409-429 .

[10] 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).

[11] S.P. Dirkse and M.C. Ferris, Mcplib: a collection of nonlinear mixed complementarity problems, *Optimization Methods and Software*, 5 (1995) 319–345 .

[12] D. Duffy, *Finite Difference Methods in Financial Engineering – A Partial Differential Equation Approach*, John Wiley & Sons Ltd (2006).

[13] 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).

[14] M.C. Ferris, J.S. Pang, Engineering and economic applications of complementarity problems, *SIAM Rev.*,**39** (1997) 669–713 .

[15] A. Forsgren, P.E. Gill and M.H. Wright, Interior methods for nonlinear optimization, *SIAM Rev.*,**44** (2002) 525–597 .

[16] S.A. Gabriel, An NE/SQP method for the bounded nonlinear complementarity problem, *Journal Optimization Theory & Applications*, 97 (1998) 493–506 .

[17] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems* Springer-Verlag, New York-Berlin-Heidelberg-Tokyo (1984).

[18] 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. Banking Finance*, 30 (2006) 489–502 .

[19] C.C. Huang and S. Wang, A Power Penalty Approach to a Nonlinear complementarity problem, *Operations Research Letters*, 38 (2010)72–76.

[20] C.C. Huang and S. Wang, A penalty method for a mixed nonlinear complementarity problem, *Nonlinear Analysis*, 75 (2012) 588–597.

[21] C. Kanzow, Global optimization techniques for mixed complementarity problems, *J. Glob. Optim.* (2000) 16, 1-21.

[22] D.C. Lesmana and S. Wang, An upwind finite difference method for a nonlinear Black–Scholes equation governing European option valuation under transaction costs, *Applied Mathematics & Computation*, 219 (2013) 8811–8828.

[23] D.C. Lesmana and S. Wang, Penalty approach to a nonlinear obstacle problem governing American put option valuation under transaction costs, *Applied Mathematics & Computation*, 251 (2015) 318–330.

[24] 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 (2009) 279-293.

[25] 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 (2013) 365–398.

[26] W. Li and S. Wang, A numerical method for pricing European options with proportional transaction costs, *J. Glob. Optim.*, 60 (2014) 59–78.

[27] D. Li and M. Fukushima, Smoothing Newton and Quasi-Newton Methods for Mixed Complementarity Problems, *Computational Optimization and Applications*, 17 (2000) 203-230.

[28] B.F. Nielsen, O. Skavhaug and A. Tveito. Penalty and front-fixing methods for the numerical solution of American option problems. *J. Comp. Fin.*, 5 (2001) 69–97.

[29] J.M. Ortega, W.C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press (1970).

[30] F.A. Potra and Y. Ye, Interior-point methods for nonlinear complementarity problems. *Journal of Optimization Theory & Applications*, 88 (1996) 617–642.

[31] R. S. Varga. *Matrix Iterative Analysis*. Prentice-Hall, Engelwood Cliffs, NJ (1962).

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

[33] S. Wang, A power penalty method for a finite-dimensional obstacle problem with derivative constraints, *Optim. Lett.*, 8 (2014) 1799–1811.

[34] S. Wang, A penalty approach to a discretized double obstacle problem with derivative constraints, *J. Glob. Optim.*, 62 (2015) 775–790.

[35] S. Wang and X.Q. Yang, A power penalty method for linear complementarity problems, *Operations Research Letters*, 36 (2008) 211-214.

[36] S. Wang and X.Q. Yang, A power penalty method for a bounded nonlinear complementarity problem, *Optimization*, 64 (2015) 2377–2394.

[37] 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 (2006) 227-254.

[38] S. Wang, K. Zhang, An interior penalty method for a finite-dimensional linear complementarity problem in financial engineering, *Optim. Lett.*, DOI 10.1007/s11590-016-1050-4.

[39] S. Wang, S. Zhang and Z. Fang, A superconvergent fitted finite volume method for Black-Scholes equations governing European and American option valuation, *Numerical Partial Differential Equations*, 31 (2015) 1190–1208.

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

[41] N. Yamashita anf M. Fukushima, M., On stationary points of the implicit Lagrangian for nonlinear complementarity problems, Journal of Optimization Theory and Applications, 84 (1995) 653–663.

[42] K. Zhang and S. Wang, Convergence property of an interior penalty approach to pricing American option, *J. Ind. Manag. Optim.*, 7 (2011) 435–447.

[43] K. Zhang and S. Wang, Pricing American bond options using a penalty method. *Automatica*, 48 (2012) 472–479.

[44] Y.Y. Zhou, S. Wang and X.Q. Yang, A penalty approximation method for a semilinear parabolic double obstacle problem, *J. Glob. Optim.*, 60 (2014) 531–550.