A POWER PENALTY METHOD FOR A CLASS OF LINEARLY CONSTRAINED VARIATIONAL INEQUALITY

MING CHEN*
School of Applied Mathematics, Xiamen University of Technology
Xiamen 361024, China

CHONGCHAO HUANG
School of Mathematics and Statistics, Wuhan University
Wuhan 430072, China

(Communicated by Song Wang)

Abstract. This paper establishes new convergence results for the power penalty method for a mixed complementarity problem (MiCP). The power penalty method approximates the MiCP by a nonlinear equation containing a power penalty term. The main merit of the method is that it has an exponential convergence rate with the penalty parameter when the involved function is continuous and $\xi$-monotone. Under the same assumptions, we establish a new upper bound for the approximation error of the solution to the nonlinear equation. We also prove that the penalty method can handle general monotone MiCPs. Then the method is used to solve a class of linearly constrained variational inequality (VI). Since the MiCP associated with a linearly constrained VI does not $\xi$-monotone even if the VI is $\xi$-monotone, we establish the new convergence result for this MiCP. We use the method to solve the asymmetric traffic assignment problem which can be reformulated as a class of linearly constrained VI. Numerical results are provided to demonstrate the efficiency of the method.

1. Introduction. The finite-dimensional variational inequality (VI), which is a generalization of the nonlinear complementarity problems (NCP), provides a general setting for the study of optimization and equilibrium problems arising from mechanics, engineering, economics, and transportation. Extensive studies have been done on the theoretical and computational aspects of the VI and NCP models. For details of these results, we refer the reader to the excellent monograph [2] and the references therein. Popular numerical methods for solving VIs and NCPs include semismooth Newton methods, smoothing methods, projection methods, and interior point methods [2, 11, 7, 3, 10].

Penalty methods have been widely used for solving constrained optimization problems for decades. In recent years, the power penalty methods have been developed for linear and nonlinear complementarity problems in both infinite and finite dimensions [17, 14, 16, 4, 1]. The penalty method approximates a complementarity problem by a nonlinear penalty equation containing a power penalty term. The

2010 Mathematics Subject Classification. Primary: 90C33, 90B20; Secondary: 90B10.
Key words and phrases. Variational inequality, mixed complementarity problem, power penalty method, traffic assignment problem, column generation.

* Corresponding author: M. Chen.
main merit of the method is its exponential convergence rate when the involved function is continuous and \(\xi\)-monotone, just as established in those studies. In [5], the method was extended to solve a mixed complementarity problem (MiCP) when the nonlinear function involved is \(\xi\)-monotone. However, the condition of \(\xi\)-monotonicity may be restrictive in some practical applications. In particular, the first order necessary conditions of constrained optimization problems and the KKT conditions of VIs can both be reformulated as MiCPs, but the functions involved are usually not \(\xi\)-monotone. In [15] the authors proposed a penalty method for the bounded NCPs, and in [12, 13] penalty methods for a discretized single and double obstacle problem were presented respectively. In these studies, the complementarity problems do not satisfy \(\xi\)-monotonicity condition, but the authors also established the similar convergence results under some assumptions. In this work, we first establish new convergence results for the power penalty method for the MiCPs. We give a new upper bound for the distance between the solutions to the MiCP and the penalty equation. Moreover, we prove that the method can handle the general monotone MiCP. Then the method is extended to solve a class of linearly constrained VI arising from the asymmetric traffic assignment problem. We first reformulate this VI as an MiCP based on the KKT conditions of the VI. However, the function defining this MiCP is not \(\xi\)-monotone even if the VI is \(\xi\)-monotone. Then we establish the similar convergence results with the help of results established under monotonicity assumption.

The VI model with efficient algorithms for the asymmetric traffic assignment problem based on Wardrop user equilibrium principle has been widely studied [6, 8]. However, efficient and accurate methods for the VI model are still scarce. In this paper, we focus on the route-based traffic assignment model which can be reformulated as a linearly constrained VI. An important algorithm principle for the solution of traffic assignment problem is to generate routes as needed rather than enumerate all the routes in the network. This approach is usually referred to as column generation. Once the new favorable routes are generated and added to the route set, we use the proposed power penalty method to solve the equilibrium problem defined by this route set, which is also a linearly constrained VI. Our numerical results confirm the theoretical findings.

The rest of this paper is organized as follows. Section 2 and Section 3 present the power penalty methods and their convergence results for the MiCP and a class of linearly constrained VI, respectively. In Section 4, we use our method to solve the route-based traffic assignment problem and we give a conclusion in the last section.

### 2. A power penalty method for the MiCP.

#### 2.1. The penalty method.

**Problem 1.** Given a set \(K \subseteq \mathbb{R}^n\) and a mapping \(F : \mathbb{R}^n \rightarrow \mathbb{R}^n\), the variational inequality, denoted as \(VI(K,F)\), is to find a vector \(x^* \in K\) such that

\[
(x - x^*)^T F(x^*) \geq 0, \quad \forall x \in K.
\]

**Definition 2.1.** A mapping \(F : K \rightarrow \mathbb{R}^n\) is said to be

1. monotone on \(K\) if

\[
(x - y)^T (F(x) - F(y)) \geq 0, \quad \forall x, y \in K;
\]
(2) $\xi$–monotone on $K$ for some $\xi \in (1, 2]$ if there exists a constant $\alpha > 0$ such that 
\[
(x - y)^T (F(x) - F(y)) \geq \alpha \|x - y\|_2^\xi, \quad \forall x, y \in K;
\]
(3) coercive with respect to $K$ if there exists a vector $x_0 \in K$ such tat
\[
\lim_{x \in K, \|x\|_2 \to \infty} \frac{(x - x_0)^T F(x)}{\|x\|_2} = \infty.
\]
Note that a $\xi$–monotone mapping must be monotone. Moreover, if $F$ is $\xi$–monotone on $K$, then for any $x, x_0 \in K$, we have
\[
(x - x_0)^T (F(x) - F(x_0)) \geq \alpha \|x - x_0\|_2^\xi.
\]
If $x \in K$ and $\|x\|_2$ sufficiently large, then it follows that
\[
\frac{(x - x_0)^T F(x)}{\|x\|_2} \geq \frac{(x - x_0)^T F(x_0) + \alpha \|x - x_0\|_2^\xi}{\|x\|_2} \\
\geq \frac{\alpha \|x - x_0\|_2^\xi - \|x - x_0\|_2 \|F(x_0)\|_2}{\|x\|_2} \\
\geq \frac{\alpha \|x - x_0\|_2^\xi - \|x - x_0\|_2 \|F(x_0)\|_2}{\|x - x_0\|_2 + \|x_0\|_2} \\
\geq \frac{\alpha \|x - x_0\|_2^{\xi-1} - \|F(x_0)\|_2}{1 + \|x_0\|_2/\|x - x_0\|_2}.
\]
Since $\xi > 1$, we have
\[
\lim_{x \in K, \|x\|_2 \to \infty} \frac{(x - x_0)^T F(x)}{\|x\|_2} = \infty.
\]
Hence, $F$ is coercive with respect to $K$. However, the coercivity of $F$ can not deduce the monotonicity of $F$ on $K$.

**Proposition 1.** Let $K \subseteq \mathbb{R}^n$ be closed and convex and $F : K \to \mathbb{R}^n$ be continuous.

1. If $F$ is $\xi$-monotone on $K$, then $VI(K, F)$ has a unique solution. When $K = \mathbb{R}^n$, a vector $x^*$ solves $VI(K, F)$ if and only if $x^*$ is a zero of $F$ (i.e., $F(x^*) = 0$).
2. If $F$ is coercive with respect to $K$, then $VI(K, F)$ has a nonempty, compact solution set.

The proof of Proposition 1 can be found in [2, Theorem 2.3.3].

**Problem 2.** Let $G$ and $H$ be two mappings from $\mathbb{R}^m \times \mathbb{R}^n$ into $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. The mixed complementarity problem, denoted as $MiCP(G, H)$, is to find a pair of vectors $(x^*, y^*)$ belonging to $\mathbb{R}^m \times \mathbb{R}^n$ such that
\[
G(x^*, y^*) = 0, \quad x^* \text{ free} \\
0 \leq y^* \perp H(x^*, y^*) \geq 0.
\]

**Proposition 2.** A vector $(x^*, y^*) \in \mathbb{R}^m \times \mathbb{R}^n$ is a solution to $MiCP(G, H)$ if and only if it is a solution to $VI(\mathbb{R}^m \times \mathbb{R}^n_+, \Phi)$, where $\Phi(x, y) = \begin{pmatrix} G(x, y) \\ H(x, y) \end{pmatrix}$. It follows that for all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n_+$,
\[
\left( \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x^* \\ y^* \end{pmatrix} \right)^T \Phi(x^*, y^*) \geq 0. \quad (1)
\]
Proposition 3. If \( -\text{MiCP} \) \( \Phi \) \( \text{Problem 3.} \)

Find \( (x, y) \in \mathbb{R}^m \times \mathbb{R}^n \) such that

\[
\Psi(x, y) = \Phi(x, y) - \lambda \left( [y]_+^{1/k} \right) = 0,
\]

where \( k > 0 \) is a power parameter, \( \lambda > 1 \) is a penalty parameter and \( [v]_- = -\min\{v, 0\} \). (2) is the power penalty equation approximating \( \text{MiCP}(\Phi) \).

Proposition 3. If \( \Phi \) is continuous and \( \xi \)-monotone on \( \mathbb{R}^m \times \mathbb{R}^n \), then \( \text{MiCP}(\Phi) \) has a unique solution by Proposition 1 and Proposition 2. Moreover, \( \Psi(x, y) \) is also continuous and \( \xi \)-monotone on \( \mathbb{R}^m \times \mathbb{R}^n \) for any \( \lambda \geq 0 \). By Proposition 1, the power penalty equation (2) has a unique solution for any \( \lambda \geq 0 \).

2.2. Convergence analysis. In this subsection, we establish the convergence results for the penalty method for the MiCP. We start the convergence analysis with the following lemma.

Lemma 2.2. Assume that \( \Phi \) is monotone on \( \mathbb{R}^m \times \mathbb{R}^n \). If \( (x^*, y^*) \) solves \( \text{MiCP}(\Phi) \) and \( (x, y) \) solves the power penalty equation (2) for a given \( \lambda > 1 \), then

\[
\|\|y\|_{-}\|_{p} \leq \frac{\|H(x^*, y^*)\|^k_{q}}{\lambda^k},
\]

where \( p = 1 + 1/k \) and \( q = 1 + k \).

Proof. Since \((x^*, y^*)\) is the solution to \( \text{MiCP}(\Phi) \), it satisfies (1) by Proposition 2. Note that \( y_{\lambda} = [y]_{+} - [y]_{-} \), thus the vector \( \left( \frac{x^*}{y^*} \right) - \left( \frac{x}{y} \right) \) can be decomposed as

\[
\frac{x^*}{y^*} - \frac{x}{y} = \frac{x^*}{y^*} - \frac{x}{[y]_{+}} + \frac{0}{[y]_{-}} = r_{\lambda} + \frac{0}{[y]_{-}},
\]

where \( [v]_{+} = \max\{v, 0\} \), \( r_{\lambda} = \frac{x^*}{y^*} - \frac{x}{[y]_{+}} \).

Since \( \frac{x^*}{y^*} - r_{\lambda} = \left( \frac{x}{[y]_{+}} \right) \in \mathbb{R}^m \times \mathbb{R}^n \), by taking \( \left( \frac{x}{y} \right) = \left( \frac{x^*}{y^*} \right) - r_{\lambda} \) in (1), we obtain

\[
-r_{\lambda}^T \Phi(x^*, y^*) \geq 0.
\]

Left-multiplying both sides of (2) by \( r_{\lambda}^T \), it follows that

\[
r_{\lambda}^T \Phi(x, y) = 0.
\]

Adding up (4) and (5) yields

\[
r_{\lambda}^T (\Phi(x, y) - \Phi(x^*, y^*)) - \lambda r_{\lambda}^T \left( [y]_{-}^{1/k} \right) \geq 0.
\]

Furthermore, note that

\[
r_{\lambda}^T \left( [y]_{-}^{1/k} \right) = \left( \left( \frac{x^*}{y^*} - \frac{x}{[y]_{+}} \right) \right)^T \left( [y]_{-}^{1/k} \right) = (y^*)^T [y]_{-}^{1/k} \geq 0,
\]

because \( y^* \geq 0 \) and \( [y]_{-} \geq 0 \), we deduce

\[
r_{\lambda}^T (\Phi(x^*, y^*) - \Phi(x, y)) \leq 0.
\]
Since $r_{\lambda} = (x^*) - (x_\lambda) - (0) - (y_\lambda)$, from the above inequality and the assumption that $\Phi$ is monotone on $\mathbb{R}^m \times \mathbb{R}^n$ we have

\[
0 \leq \left( (x^*) - (x_\lambda) - (y_\lambda) \right)^T (\Phi(x^*, y^*) - \Phi(x_\lambda, y_\lambda)) \\
\leq \left( 0 - (y_\lambda) \right)^T (\Phi(x^*, y^*) - \Phi(x_\lambda, y_\lambda)) \\
= [y_\lambda]^T (H(x^*, y^*) - H(x_\lambda, y_\lambda)).
\]  

(6)

Left-multiplying both sides of (2) by $\left( \begin{array}{c} 0 \\ [y_\lambda] \end{array} \right)^T$, we obtain

\[
[y_\lambda]^T H(x_\lambda, y_\lambda) - \lambda [y_\lambda]^T [y_\lambda]_1^{1/k} = 0,
\]

then

\[
[y_\lambda]^T H(x_\lambda, y_\lambda) = \lambda \|[y_\lambda]_1\|_p^p.
\]  

(7)

where $p = 1 + 1/k$ and $q = 1 + k$ so that $1/p + 1/q = 1$.

Combining (6) and (7) and using Hölder inequality we can establish that

\[
0 \leq [y_\lambda]^T (H(x^*, y^*) - H(x_\lambda, y_\lambda)) \\
= [y_\lambda]^T H(x^*, y^*) - \lambda \|[y_\lambda]_1\|_p^p \\
\leq \|[y_\lambda]_1\|_p (\|H(x^*, y^*)\|_q - \lambda \|[y_\lambda]_1\|_p^{1/k}),
\]  

(8)

hence, we obtain

\[
\|[y_\lambda]_1\|_p (\|H(x^*, y^*)\|_q - \lambda \|[y_\lambda]_1\|_p^{1/k}) \geq 0.
\]  

(9)

If $\|[y_\lambda]_1\|_p = 0$, then

\[
\|H(x^*, y^*)\|_q - \lambda \|[y_\lambda]_1\|_p^{1/k} = \|H(x^*, y^*)\|_q \geq 0;
\]

when $\|[y_\lambda]_1\|_p > 0$, by (9) we have

\[
\|H(x^*, y^*)\|_q - \lambda \|[y_\lambda]_1\|_p^{1/k} \geq 0.
\]

Thus, $\|H(x^*, y^*)\|_q \geq \lambda \|[y_\lambda]_1\|_p^{1/k}$ always holds. Consequently, we can deduce

\[
\|[y_\lambda]_1\|_p \leq \frac{\|H(x^*, y^*)\|_q^k}{\lambda^k}.
\]

Theorem 2.3. Suppose that $\Phi$ is continuous and $\xi$-monotone on $\mathbb{R}^m \times \mathbb{R}^n$. Let $(x^*, y^*)$ and $(x_\lambda, y_\lambda)$ be the solutions to MiCP($\Phi$) and the power penalty equation (2), respectively. Then

\[
\left\| \left( \begin{array}{c} x^* \\ y^* \end{array} \right) - \left( \begin{array}{c} x_\lambda \\ y_\lambda \end{array} \right) \right\|_2 \leq \frac{C}{\lambda^{k/\xi}},
\]

where $C = \frac{1}{\alpha^{1/\xi} \|H(x^*, y^*)\|_{1+k/\xi}}$. 


Proof. By the $\xi$–monotonicity of $\Phi$ on $\mathbb{R}^m \times \mathbb{R}^n$, we have
\[
\left( \begin{array}{c} x^* \\ y^* \end{array} \right) \left( \begin{array}{c} x^* \\ y^* \end{array} \right)^T (\Phi(x^*, y^*) - \Phi(x_\lambda, y_\lambda)) \geq \alpha \| \left( \begin{array}{c} x^* \\ y^* \end{array} \right) - \left( \begin{array}{c} x_\lambda \\ y_\lambda \end{array} \right) \|^\xi.
\]
Moreover, (6), (8) and (3) hold since a $\xi$–monotone function must be monotone. Thus, we can deduce
\[
\alpha \left\| \left( \begin{array}{c} x^* \\ y^* \end{array} \right) - \left( \begin{array}{c} x_\lambda \\ y_\lambda \end{array} \right) \right\|_2^\xi \leq \|[y_\lambda]_-\|_p \| H(x^*, y^*) \|_q \leq \frac{\| H(x^*, y^*) \|_q^{1+k}}{\lambda^k}.
\]
Let $C = \frac{1}{\alpha^2 \epsilon} \| H(x^*, y^*) \|_{1+k}^{(1+k)/\xi}$, then we can obtain from the above inequality
\[
\left\| \left( \begin{array}{c} x^* \\ y^* \end{array} \right) - \left( \begin{array}{c} x_\lambda \\ y_\lambda \end{array} \right) \right\|_2 \leq \frac{C}{\lambda^{k/\xi}}.
\]
\[
\square
\]

Theorem 2.4. Assume that $\Phi$ is continuous and monotone on $\mathbb{R}^m \times \mathbb{R}^n$. If $(x^*, y^*)$ solves $\text{MiCP}(\Phi)$ and $(x_\lambda, y_\lambda)$ solves the power penalty equation (2) for any $\lambda > 1$, and $\lim_{\lambda \to \infty} (x_\lambda, y_\lambda) = (x_\infty, y_\infty)$. Then $(x_\infty, y_\infty)$ is a solution to $\text{MiCP}(\Phi)$.

Proof. By Lemma 2.2, we know (3) holds, which implies $[y_\lambda]_- \to 0$ as $\lambda \to \infty$. Since $y_\lambda = [y_\lambda]_+ - [y_\lambda]_-$, it satisfies $y_\lambda \to [y_\lambda]_+$ as $\lambda \to \infty$. Thus, $y_\infty \geq 0$. It follows that
\[
(x_\infty, y_\infty) \in \mathbb{R}^m \times \mathbb{R}^n_+.
\]
By (2), we have for any $\lambda \geq 0$,
\[
\Phi(x_\lambda, y_\lambda) = \lambda \left( \begin{array}{c} 0 \\ [y_\lambda]_-^{1/k} \end{array} \right),
\]
thus $G(x_\lambda, y_\lambda) = 0$ and $H(x_\lambda, y_\lambda) \geq 0$. By the continuity of $\Phi$, it follows that
\[
G(x_\infty, y_\infty) = 0, \quad H(x_\infty, y_\infty) \geq 0.
\]
Left-multiplying both sides of (2) by $\left( \begin{array}{c} 0 \\ y_\lambda \end{array} \right)^T$, we obtain
\[
y_\lambda^T H(x_\lambda, y_\lambda) - \lambda y_\lambda^T [y_\lambda]_-^{1/k} = 0,
\]
which implies
\[
y_\lambda^T H(x_\lambda, y_\lambda) = \lambda y_\lambda^T [y_\lambda]_-^{1/k} = \lambda ([y_\lambda]_+ - [y_\lambda]_-) [y_\lambda]_-^{1/k} = -\lambda \|[y_\lambda]_-\|_p^p,
\]
by using (3) and $p = 1 + 1/k$, we have
\[
\|[y_\lambda]_- H(x_\lambda, y_\lambda)\| = \lambda \|[y_\lambda]_-\|_p^p \leq \lambda \frac{\| H(x^*, y^*) \|_{1+k}^{1+k}}{\lambda^k} = \frac{\| H(x^*, y^*) \|_{1+k}^{1+k}}{\lambda^k},
\]
so $y_\lambda^T H(x_\lambda, y_\lambda) \to 0$ as $\lambda \to \infty$, thus by the continuity of $\Phi$, we have
\[
y_\infty^T H(x_\infty, y_\infty) = 0.
\]
Consequently, by (10), (11) and (12), we deduce $(x_\infty, y_\infty)$ is a solution to $\text{MiCP}(\Phi)$.
\[
\square
Theorem 2.3 establishes an upper bound for the distance between the solution to $MiCP(\Phi)$ and the power penalty equation (2). Unlike the bound established in [5, Theorem 3.1], the new bound is closely related to the solutions to $MiCP(\Phi)$, but is independent of the penalty parameter $\lambda$. This result may provide certain information for choosing the parameters when we refer to solve an MiCP. Theorem 2.4 shows that the power penalty method can also handle a monotone MiCP. Although the same upper bound can not be obtained under monotonicity assumption, by the proof of Theorem 2.4, the task of solving the monotone MiCP reduces to choose a sufficiently large $\lambda$ such that $\|H(x^*,y^*)\|_{1+\frac{1}{k}}$ is sufficiently small.

The monotonicity of $\Phi$ on $\mathbb{R}^m \times \mathbb{R}^n$ is not sufficient for $MiCP(\Phi)$ to have a solution. Next we give a sufficient condition for $MiCP(\Phi)$ and the corresponding power penalty equation to have a solution. The condition is weaker than the continuity and $\xi$–monotonicity of $\Phi$ on $\mathbb{R}^m \times \mathbb{R}^n$.

Theorem 2.5. Assume that $\Phi : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ is continuous and there exists a vector $z_0 = (x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}_+^n$ such that

$$\lim_{\|z\|_2 \to \infty} \frac{(z - z_0)^T \Phi(z)}{\|z\|_2} = \infty,$$

where $z = (x, y) \in \mathbb{R}^m \times \mathbb{R}^n$. Then $MiCP(\Phi)$ has a nonempty, compact solution set, and for all $\lambda \geq 0$, the power penalty equation also has a nonempty, compact solution set.

Proof. Since $\mathbb{R}^m \times \mathbb{R}_+^n \subseteq \mathbb{R}^m \times \mathbb{R}^n$ and $z_0 = (x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}_+^n$, then by Proposition 1, we know $MiCP(\Phi)$ has a nonempty, compact solution set.

For a given $\lambda \geq 0$ and a vector $z = (x, y) \in \mathbb{R}^m \times \mathbb{R}^n$,

$$(z - z_0)^T \Psi(z) = (z - z_0)^T \Phi(z) - \lambda(y - y_0)^T |y|_{\frac{1}{k}}$$

$$= (z - z_0)^T \Phi(z) + \lambda \|[y]_+\|_{1+\frac{1}{k}} + y_0^T |y|_{\frac{1}{k}}$$

$$\geq (z - z_0)^T \Phi(z),$$

Thus we have from the assumption that

$$\lim_{\|z\|_2 \to \infty} \frac{(z - z_0)^T \Psi(z)}{\|z\|_2} = \infty.$$

By Proposition 1, we have $VI(\mathbb{R}^m \times \mathbb{R}^n, \Psi)$ has a nonempty, compact solution set. Thus for each $\lambda \geq 0$, the power penalty equation also has a nonempty, compact solution set.

In the implementation of the power penalty method, we need to choose the penalty parameter $\lambda$ and the power parameter $k$, and then solve the power penalty equation. Note that the penalty term $\lambda |y|_{\frac{1}{k}}$ is non-smooth, but it can be smoothed out by the following formula proposed in [17]:

$$s(y) = \begin{cases} \lambda (-y)^{1/k} & y \leq -\varepsilon \\ \lambda (3 - 1/k) \varepsilon^{1/k} - 2y^2 + (2 - 1/k) \varepsilon^{1/k} - 3y^3 & -\varepsilon < y < 0 \end{cases}$$
where $0 < \varepsilon << 1$, we set $\varepsilon = 10^{-10}$. Then non-smooth power penalty equation becomes smooth if $\Phi(x, y)$ is smooth, so we can use any efficient method to solve this smooth equation. In this paper, we use the damped Newton method.

3. A power penalty method for a class of linearly constrained VI. In this section, we first reformulate a class of linearly constrained VI as an MiCP. Since the MiCP satisfies a different property, we then present the penalty equation to approximate the MiCP and discuss the convergence results. Consider the following linearly constrained VI $(K, F)$, where $K = \{x : Ax = b, x \geq 0\}$, $A \in \mathbb{R}^{m \times n}$ is of full row rank and $b \in \mathbb{R}^m$.

**Proposition 4.** If $F$ is continuous and $\xi$-monotone on $K$, then Proposition 1 shows that VI $(K, F)$ has a unique solution.

The following theorem establishes the equivalence between the linearly constrained VI and an MiCP; therefore, we could solve the MiCP instead of solving the original problem.

**Theorem 3.1.** Let $K = \{x : Ax = b, x \geq 0\}$. A vector $x^*$ solves VI $(K, F)$ if and only if there exists a vector $u^* \in \mathbb{R}^m$ such that

$$Ax^* = b,$$

$$0 \leq x^* \perp F(x^*) - A^T u^* \geq 0. \quad (13)$$

The proof of Theorem 3.1 uses the simple linear programming duality theory and can be found in [2, Proposition 1.2.1].

**Remark 1.** Let $\Phi(u, x) = \begin{pmatrix} Ax - b \\ F(x) - A^T u \end{pmatrix}$, then (13) constitutes a special MiCP, denoted as MiCP($\Phi$). If $F$ is continuous and $\xi$-monotone on $K$, then the solution to MiCP($\Phi$) exists and $x^*$ is unique since VI $(K, F)$ has a unique solution and Theorem 3.1 holds. It is clear that $u^*$ is the vector of optimal Lagrange multiplier for the constraint $Ax = b$. Since $u^*$ may not be unique, let $U$ denote the set of optimal Lagrange multipliers.

**Theorem 3.2.** If $F(x)$ is $\xi$-monotone on $\mathbb{R}^n$, then $\Phi(u, x)$ satisfies the following partial $\xi$-monotone property: \( \forall \begin{pmatrix} u_1 \\ x_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ x_2 \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n, \)

\[
\begin{pmatrix} u_1 \\ x_1 \end{pmatrix} - \begin{pmatrix} u_2 \\ x_2 \end{pmatrix}^T \left( \Phi(u_1, x_1) - \Phi(u_2, x_2) \right) \geq \alpha \|x_1 - x_2\|_2^\xi.
\]

**Proof.** For any $\begin{pmatrix} u_1 \\ x_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ x_2 \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n$, we have

\[
\begin{pmatrix} u_1 \\ x_1 \end{pmatrix} - \begin{pmatrix} u_2 \\ x_2 \end{pmatrix}^T \left( \Phi(u_1, x_1) - \Phi(u_2, x_2) \right) \\
= \begin{pmatrix} u_1 - u_2 \\ x_1 - x_2 \end{pmatrix}^T \left( \Phi(u_1, x_1) - \Phi(u_2, x_2) \right) \\
= \begin{pmatrix} u_1 - u_2 \\ x_1 - x_2 \end{pmatrix}^T \left( \begin{pmatrix} Ax_1 - b - Ax_2 + b \\ F(x_1) - A^T u_1 - F(x_2) + A^T u_2 \end{pmatrix} \right) \\
= (u_1 - u_2)^T A(x_1 - x_2) + (x_1 - x_2)^T (F(x_1) - F(x_2))
\]
Assume that 

\[ \Phi(u, x) = \Phi(u, x) - \lambda \left( \frac{0}{[x_\lambda]_1^{1/k}} \right) = 0, \]  

(14)

where \( k > 0 \) and \( \lambda > 1 \) are parameters, \([v]_- = -\min\{v, 0\}\).

Remark 2. Although \( \Phi(u, x) \) is \( \xi \)-monotone in \( x \), it might not be \( \xi \)-monotone in \((u, x)\) when \( F(x) \) is \( \xi \)-monotone on \( \mathbb{R}^n \). In this case, the results established in [5] do not hold. However, \( \Phi(u, x) \) is monotone on \( \mathbb{R}^m \times \mathbb{R}^n \) if \( F(x) \) is monotone on \( \mathbb{R}^n \). Hence, Lemma 2.2 and Theorem 2.4 still hold.

Now, we present the power penalty equation for the MiCP and prove the existence of a unique solution to the MiCP.

Problem 4. Find \( u_\lambda \in \mathbb{R}^m \) and \( x_\lambda \in \mathbb{R}^n \) such that

\[ \Psi(u_\lambda, x_\lambda) = \Phi(u_\lambda, x_\lambda) - \lambda \left( \frac{0}{[x_\lambda]_1^{1/k}} \right) = 0, \]  

(14)

where \( k > 0 \) and \( \lambda > 1 \) are parameters, \([v]_- = -\min\{v, 0\}\).

Since \( \Psi(u_\lambda, x_\lambda) \) is still not \( \xi \)-monotone with respect to \((u, x)\), we cannot obtain the existence of a solution by Proposition 1.

Theorem 3.3. If \( F(x) \) is continuous and \( \xi \)-monotone on \( \mathbb{R}^n \), then for any \( \lambda \geq 0 \), the power penalty equation (14) has a unique solution.

Proof. Let \( \tilde{K} = \{ x : Ax = b \} \) and \( \tilde{F}(x) = F(x) - \lambda [x]_1^{1/k} \). It is not difficult to show that \( \tilde{F}(x) \) is continuous and \( \xi \)-monotone for any \( \lambda \geq 0 \). Since \( \tilde{K} \) is a closed convex set, then the existence of a unique solution \( x_\lambda \) to VI(\( \tilde{K}, \tilde{F} \)) follows from Proposition 1.

Note that if \( x_\lambda \) is the solution to VI(\( \tilde{K}, \tilde{F} \)), then \( x_\lambda \) solves the following linear program in the variable \( y \):

\[ \min \quad y^T \tilde{F}(x_\lambda) \]

\[ \text{s.t.} \quad y \in \tilde{K}. \]

Since (14) is precisely the KKT condition of this linear program, then there exists \( u_\lambda \) such that \((u_\lambda, x_\lambda)\) solves (14) and \( x_\lambda \) is unique.

Since \((u_\lambda, x_\lambda)\) is the solution to (14), we have

\[ A^T u_\lambda = F(x_\lambda) - \lambda [x_\lambda]_1^{1/k}, \]

it follows that

\[ AA^T u_\lambda = A(F(x_\lambda) - \lambda [x_\lambda]_1^{1/k}). \]

Since \( A \) is of full row rank, then \( AA^T \) is invertible. Thus,

\[ u_\lambda = (AA^T)^{-1}A(F(x_\lambda) - \lambda [x_\lambda]_1^{1/k}), \]

the uniqueness of \( u_\lambda \) follows readily.

Theorem 3.4. Assume that \( F(x) \) is continuous and \( \xi \)-monotone on \( \mathbb{R}^n \). Let \((u^*, x^*)\) and \((u_\lambda, x_\lambda)\) be the solutions to MiCP(\( \Phi \)) and the power penalty equation (14), respectively. Assume that \( U \) is bounded, then

\[ \|u^* - x_\lambda\|_2 \leq \frac{C}{\lambda^{k/\xi}}, \]

where \( C = \max_{u^* \in U} \|F(x^*) - A^T u^*\|_1^{(1+k)/\xi}. \)
Proof. Since $\Phi(u, x)$ is monotone on $\mathbb{R}^n \times \mathbb{R}^n$, Lemma 2.2 holds. It follows that
\[
\|\|x_\lambda\|-\|F(x^*) - AT^*u^*\|_q^k \leq \frac{\|F(x^*) - AT^*u^*\|_q^k}{\lambda^k},
\]
where $p = 1 + 1/k$ and $q = 1 + k$. Moreover, following the proof of Lemma 2.2, we can obtain
\[
\alpha \|x^* - x_\lambda\|_2^\xi \leq \left(\left(\frac{u^*}{x^*}\right) - \left(\frac{u_\lambda}{x_\lambda}\right)\right)^T (\Phi(u^*, x^*) - \Phi(u_\lambda, u_\lambda))
\leq \|\|x_\lambda\|-\|F(x^*) - AT^*u^*\|_q^1 \leq \frac{\|F(x^*) - AT^*u^*\|_q^1}{\lambda^k},
\]
by Theorem 3.2 and the inequality (15).

Let $C = \frac{1}{\alpha \xi} \max_{u \in U} \|F(x^*) - AT^*u^*\|_q^1 + \xi$, then $\|x^* - x_\lambda\|_2 \leq \frac{C}{\alpha \xi} \leq \frac{C}{\lambda^k}$ holds. \qed

Remark 3. When $F(x)$ is continuous and $\xi$-monotone on $\mathbb{R}^n$, we have $\|x^* - x_\lambda\|_2 \leq \frac{C}{\lambda^k}$, where $x^*$ is the unique solution to $VI(K, F)$ and $x_\lambda$ is the unique solution to the power penalty equation (14).

Similar to Theorem 2.5, we give a sufficient condition for $VI(K, F)$, MiCP($\Phi$) and the corresponding power penalty equation to have a solution.

Theorem 3.5. Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and there exists a vector $x_0 \in K$ such that
\[
\lim_{\|x\|_2 \to \infty} \frac{(x - x_0)^T F(x)}{\|x\|_2} = \infty,
\]
where $x \in \mathbb{R}^n$. Then $VI(K, F)$ has a nonempty, compact solution set. Moreover, there exists a solution to MiCP($\Phi$) and for all $\lambda \geq 0$, the power penalty equation has a solution.

Proof. Since $x_0 \in K$ and $K \subseteq \mathbb{R}^n$, then by Proposition 1, it follows that $VI(K, F)$ has a nonempty, compact solution set. By Theorem 3.1, the solution to MiCP($\Phi$) exists.

Let $\tilde{K} = \{x : Ax = b\}$, $\tilde{F}(x) = F(x) - \lambda [x]_{1/k}$. Since $x_0 \in K$, we have $x_0 \geq 0$. For a given $\lambda \geq 0$, we have
\[
(x - x_0)^T \tilde{F}(x) = (x - x_0)^T (F(x) - \lambda [x]_{1/k})
= (x - x_0)^T F(x) - \lambda x_0^T [x]_{1/k} + \lambda x_0^T [x]_{1/k}
= (x - x_0)^T F(x) - \lambda (x[+ - [x]_+]^T [x]_{1/k} + \lambda x_0^T [x]_{1/k}
= (x - x_0)^T F(x) + \lambda \|x[-]_{1+1/k} + \lambda x_0^T [x]_{1/k}
\geq (x - x_0)^T F(x),
\]
thus, $\lim_{\|x\|_2 \to \infty} \frac{(x - x_0)^T F(x)}{\|x\|_2} = \infty$. Since $\tilde{K} \subseteq \mathbb{R}^n$ and $x_0 \in \tilde{K}$, then by Proposition 1, we deduce that $VI(\tilde{K}, \tilde{F})$ has a nonempty, compact solution set.

Note that if $x_\lambda$ is the solution to $VI(\tilde{K}, \tilde{F})$, then $x_\lambda$ solves the following linear program in the variable $y$:
\[
\min \ y^T \tilde{F}(x_\lambda)
\text{ s.t. } y \in \tilde{K}.
\]
Since (14) is precisely the KKT condition of this linear program, then there exists $u_\lambda$ such that $(u_\lambda, x_\lambda)$ solves (14). Therefore, for any $\lambda \geq 0$, the power penalty equation (14) has a solution.

4. Application in asymmetric traffic assignment problem. In this section, we use the method developed in this work to solve the asymmetric traffic assignment problem. The traffic assignment problem is to allocate the origin-destination trips to routes in the transportation network, in order to estimate the traffic volumes and travel costs on the roads. The most widely used route choice model is the user-equilibrium (UE) principle, which stipulates that users of the traffic network will choose the minimum cost route between each origin-destination pair, and through this process the routes that are used (i.e., have positive flows) will have the same cost; moreover, routes with cost higher than the minimum will have no flow. The UE assignment finds the user-optimized patterns characterized by the UE principle. In the case of asymmetric problem, the travel cost on a link is allowed to depend upon the entire flow pattern. For more details about traffic assignment problem, we refer to [9].

4.1. The model and algorithm for route-based traffic assignment problem. Let us consider a road network represented by the graph $G = (N, A)$ where $N$ denote the set of nodes and $A$ denote the set of links. There are two distinguished subsets of $N$ that represent the set of origin nodes $O$ and destination nodes $D$. The set of origin-destination (O-D) pairs is a given subset $W$ of $O \times D$. For each $w \in W$, let $d_w$ denote the travel demand between the O-D pair $w$ and $d = (d_w)_{w \in W}$. For each $w \in W$, let $P_w$ denote the set of routes connecting the O-D pair $w$ and let $P = \bigcup_{w \in W} P_w$. Let $h_p$ be the flow on route $p \in P$ and $C_p(h)$ be the travel cost on this route that is a function of the entire vector $h = (h_p)_{p \in P}$ of route flows. Then for each $w \in W$, the travel demand must be satisfied: $\sum_{p \in P_w} h_p = d_w$. Let $\Omega$ be O-D pair-route incidence matrix whose entries are

$$\omega_{wp} = \begin{cases} 1 & \text{if route } p \in P_w \\ 0 & \text{otherwise.} \end{cases}$$

In vector notation, the flow conservation constraint implies $\Omega h = d$. Thus the set of feasible route flow vectors can be denoted as

$$H = \{h \geq 0 : \Omega h = d\}.$$

The flow $f_a$ on each link $a \in A$ is the sum of all flows on routes to which the link belongs, then the entire flow vector $f = (f_a)_{a \in A}$ is given by $f = \Delta h$, where $\Delta$ is the link-route incidence matrix with entries

$$\delta_{ap} = \begin{cases} 1 & \text{if route } p \in P \text{ traverses link } a \in A \\ 0 & \text{otherwise.} \end{cases}$$

For each link $a \in A$, let $c_a(f)$ be the travel cost function and $c(f) = (c_a(f))_{a \in A}$. Assuming the route cost function $C(h) = (C_p(h))_{p \in P}$ is additive; that is, for each $p \in P$, $C_p(h)$ is the sum of the link costs $c_a(f)$ on all the links $a$ traversed by the route $p \in P$, thus $C_p(h) = \sum_{a \in A} c_a(\Delta h) \delta_{ap}$. In vector notation, this assumption says
\[ C(h) = \Delta^T c(f) = \Delta^T c(\Delta h). \]

The route-based asymmetric traffic equilibrium problem can be formulated as the following linearly constrained variational inequality: \( VI(H, C(h)) \). That is, find a route flow vector \( h^* \in H \) such that

\[ (h - h^*)^T C(h^*) \geq 0, \; \forall h \in H. \]

As mentioned in the introduction, we use a column generation procedure to generate new favorable routes as needed. This avoids enumerating all the routes in the road network. Additionally, we use a column dropping scheme, which means dropping the unused routes, to keep the route set compact and improve the computational efficiency of the algorithm. The detailed algorithmic steps are described as follows:

**Step 1. Initialization.**

**Step 1.1.** Let \( f^0 = 0, P^0_w = \emptyset \) for each O-D pair \( w \in W, \varepsilon_1 > 0 \).

**Step 1.2.** Solve the shortest route problem based on link travel cost \( c(f^0) \), obtain a shortest route \( p^1_w \) for each O-D pair \( w \in W \);

**Step 1.3.** Update the set of generated routes: \( P^1_w = \{p^1_w\} \cup P^0_w, \forall w \in W \);

**Step 1.4.** Perform an all-or-nothing (AON) assignment: \( h^1_w = d_w, \forall w \in W \);

**Step 1.5.** Compute link flows: \( f^1_w = \Delta^1 h^1 \), where \( \Delta^1 \) is the restricted link-route incidence matrix.

**Step 1.6.** Set \( m = 1 \);

**Step 2. Column generation.**

**Step 2.1.** Update link travel cost: \( c(f^m) \);

**Step 2.2.** Solve the shortest route problem based on link travel cost \( c(f^m) \), obtain a shortest route \( p^{m+1}_w \) and its travel cost \( u^{m+1}_w \) for each O-D pair \( w \in W \);

**Step 2.3.** If \( p^{m+1}_w \notin P^m_w \), then let \( P^{m+1}_w = \{p^{m+1}_w\} \cup P^m_w \), otherwise, let \( P^{m+1}_w = P^m_w \).

**Step 3. Convergence test.**

For each O-D pair, if the travel cost for each used routes (the routes which have positive flows) sufficiently approaches to the minimum travel cost, that is, if

\[ E = \max_{w \in W} \sum_{p \in P^m_w} \frac{h^m_w}{d_w} \left( \frac{C^m_p(h^m) - u^{m+1}_w}{C^m_p(h^m)} \right) \leq \varepsilon_1, \]

then stop; otherwise, set \( m = m + 1 \);

**Step 4. Equilibrium problem.**

**Step 4.1.** Use the power penalty method to solve the \( VI(H^m, C^m(h)) \), where \( H^m = \{h \in R_{+}^{|P^m|} : \Omega^m h = d\} \), \( P^m = \bigcup_{w \in W} P^m_w, |P^m| \) is the number of routes in \( P^m \), \( \Omega^m \) is the restricted O-D pair-route incidence matrix and \( C^m(h) = (C^m_p(h))_{p \in P^m} \), then obtain the route flow solution \( h^m \) and update link flow \( f^m = \Delta^m h^m \).

**Step 4.2.** Drop unused routes. For each \( w \in W \) and each \( p \in P^m_w \), if \( h^m_p \leq 0 \), then let \( P^m_w = P^m_w - \{p\} \).

**Step 4.3.** Go to Step 2.

Now we are ready to examine the performance of the power penalty method when it is applied to the equilibrium problems. It is obvious that \( \Omega^m \) is of full row rank. If we reformulate \( VI(H^m, C^m(h)) \) as an MiCP, then the multiplier \( u \) has a special meaning, that is, it represents the vector of minimum travel costs between all O-D pairs, thus \( u \) must be unique and bounded. Moreover, \( c(f) \) can be assumed to be continuous and monotone because of the congestion effect. Under this assumption, \( C^m(h) \) is a continuous and monotone mapping. Let \( (h^*, u^*) \) be
the solution to the MiCP associated with $VI(H^m, C^m(h))$. The column generation procedure generates the favorable routes that probably carry the positive flows, which means $C_p(h^*) - [(\Omega^m)^T u^*]_p = C_p(h^*) - u^*_w = 0$, and the column dropping scheme drops the unused routes, which means the route with $C_p(h^*) - u^*_w > 0$ will be dropped in the next iteration. Hence, in each iteration, the route set only contains the most likely used route. Consequently, $\|C^m(h^*) - (\Omega^m)^T u^*\|_{1+k}$ will be small. Therefore, the penalty parameter $\lambda$ is no need to be very large.

4.2. Numerical results. The traffic network used for the numerical experiment is depicted in Fig.1. The network consists of 4 O-D pairs 1-12, 9-4, 12-1 and 4-9, with demands 300, 400, 350 and 350 respectively. The network includes 34 links, the numbers by the link are the link number and the free-flow travel cost on that link. For example, 11(4.5) means that the free-flow travel cost on the 11th link is 4.5 min. We consider the following travel cost function:

$$c_a(f) = t^0_a \left(1 + 0.5 \left(\frac{f_a + 0.5f_{\bar{a}}}{2K_a}\right)^4\right),$$

where $t^0_a$ and $K_a$ are the free-flow travel cost and the capacity of the link $a$ respectively, $\bar{a}$ denotes the opposite link of $a$. The capacity on every link is set at 300. The cost function takes into account the traffic flow in the opposite directions, which leads to an asymmetric traffic assignment problem. The similar structure of the cost function is also considered for the networks used in [8].

In the implementation of the column generation algorithm, we choose the stopping criterion $\varepsilon_1 = 10^{-8}$. The power penalty method has two parameters, the penalty parameter is set at $\lambda = 20$ and the power parameter is set at $k = 2$. After 5 iterations, the column generation algorithm stops at $E = 3.83 \times 10^{-11}$. The detailed numerical results are listed in Table 1.

As we can see from Table 1, for each O-D pair, all of the routes which have positive flows have the same cost and the cost equal to the minimum cost; on the other hand, with the benefit of column generation and column drop schemes, all the generated routes carry the positive flows.
Table 1. Computational Results

| O-D pair | Minimum cost | Route flow | Route cost |
|----------|--------------|------------|------------|
| 1-12     | 22.540478    | 51.1269    | 22.540478  |
|          |              | 170.1978   | 22.540478  |
|          |              | 21.8757    | 22.540478  |
|          |              | 21.5513    | 22.540478  |
|          |              | 35.2483    | 22.540478  |
| 9-4      | 22.006421    | 222.7626   | 22.006421  |
|          |              | 121.4219   | 22.006421  |
|          |              | 55.8155    | 22.006421  |
| 12-1     | 22.591574    | 88.8082    | 22.591574  |
|          |              | 189.4446   | 22.591574  |
|          |              | 18.6029    | 22.591574  |
|          |              | 24.3365    | 22.591574  |
|          |              | 28.8078    | 22.591574  |
| 4-9      | 21.965890    | 195.8051   | 21.965890  |
|          |              | 99.6358    | 21.965890  |
|          |              | 54.5591    | 21.965890  |

To show the effect of the penalty parameter $\lambda$ on the performance of the method, we set $k = 2$, and compute the numerical solutions for $\lambda = 5 \times i, i = 1, \ldots, 10$. The computational results are reported in Table 2. The first column gives the value of parameter $\lambda$, the second one reports the number of iterations, the third one gives the number of generated routes, the fourth one shows the number of routes which carry the positive flows, the fifth one provides the computational error and the last one gives the total travel costs of the network.

Table 2. Computational results when $k = 2$

| $\lambda$ | m | h | posi-h | $E$     | total costs  |
|------------|---|---|--------|--------|--------------|
| 5          | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 10         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 15         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 20         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 25         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 30         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 35         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 40         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 45         | 5 | 16| 16     | 3.83E-11| 31159.8244  |
| 50         | 5 | 16| 16     | 3.83E-11| 31159.8244  |

Table 2 shows that the performance of method is insensitive to the penalty parameter $\lambda$. Moreover, from the third and fourth columns of the table, we can see that all the generated routes carry the positive flows.

Now we investigate the performance of the algorithm under different power parameters $k$. We choose $\lambda = 5$, and then compute the numerical solutions for $k = 1, 2, 3, 4$. The computational results are listed in Table 3. The first column gives the value of parameter $k$, the meaning of the other columns are the same as...
Table 2. From Table 3 we see that the method obtains high accuracy solutions regardless of the power parameter values.

| k | m | h | posi-h | E     | total costs          |
|---|---|---|--------|-------|----------------------|
| 1 | 5 | 16| 16     | 3.83E-11 | 31159.8244          |
| 2 | 5 | 16| 16     | 3.83E-11 | 31159.8244          |
| 3 | 5 | 18| 16     | 1.31E-10  | 31159.8244          |
| 4 | 6 | 17| 16     | 2.43E-12  | 31159.8244          |

5. Conclusions. In this paper, we establish new convergence results for the power penalty method for solving MiCPs. Under $\xi$–monotone assumption, we prove that the solution to the penalty equation converges to that of the MiCP with an exponential convergence rate depending on the penalty parameter $\lambda$. In particular, we also prove that the algorithm has the ability to handle a general monotone problem. Then we extend the penalty method to solve a class of linearly constrained VI. Since the MiCP associated with a linearly constrained VI satisfies a new property, we establish the new convergence result for this MiCP. The convergence proof directly depends on the result established under monotonicity assumption. Using this method, we develop a column generation scheme for solving the asymmetric traffic assignment problem. The numerical results show that the new method can obtain the solution with high accuracy on a small network.

Acknowledgments. The authors would like to thank the editors and the anonymous referees for their valuable comments and suggestions. This work is supported by the science and technology project for high-level talents of the Xiamen University of Technology(Project No.YKJ14017R)and the Key Program of the National Natural Science Foundation of China(Grant No.71231007).

REFERENCES

[1] M. Chen and C. C. Huang, A power penalty method for the general traffic assignment problem with elastic demand, Journal of Industrial and Management Optimization, 10 (2014), 1019–1030.
[2] F. Facchinei and J. S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer-Verlag, New York, 2003.
[3] B. S. He and L. Z. Liao, Improvements of some projection methods for monotone nonlinear variational inequalities, Journal of Optimization Theory & Applications, 112 (2002), 111–128.
[4] C. C. Huang and S. Wang, A power penalty approach to a nonlinear complementarity problem, Operations Research Letters, 38 (2010), 72–76.
[5] C. C. Huang and S. Wang, A penalty method for a mixed nonlinear complementarity problem, Nonlinear Analysis: Theory, Methods & Applications, 75 (2012), 588–597.
[6] S. Lawphongpanich and D. Hearn, Simplicial decomposition of the asymmetric traffic assignment problem, Transportation Research Part B: Methodological, 18 (1984), 123–133.
[7] T. D. Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, Mathematical Programming, 75 (1996), 407–439.
[8] B. Panicucci, M. Pappalardo and M. Passacantando, A path-based double projection method for solving the asymmetric traffic network equilibrium problem, Optimization Letters, 1 (2007), 171–185.
[9] P. Patriksson, The Traffic Assignment Problem: Models and Methods, VSP, Utrecht, 1994.
[10] M. V. Solodov and B. F. Svaiter, A new projection method for variational inequality problems, *SIAM Journal on Control & Optimization*, 37 (1999), 765–776.

[11] K. Taji, M. Fukushima and T. Ibaraki, A globally convergent newton method for solving strongly monotone variational inequalities, *Mathematical Programming*, 58 (1993), 369–383.

[12] S. Wang, A penalty method for a finite-dimensional obstacle problem with derivative constraints, *Optimization Letters*, 8 (2014), 1799–1811.

[13] S. Wang, A penalty approach to a discretized double obstacle problem with derivative constraints, *Journal of Global Optimization*, 62 (2015), 775–790.

[14] S. Wang and C. S. Huang, A power penalty method for solving a nonlinear parabolic complementarity problem, *Nonlinear Analysis: Theory, Methods & Applications*, 69 (2008), 1125–1137.

[15] S. Wang and X. Q. Yang, A power penalty method for a bounded nonlinear complementarity problem, *Optimization: A Journal of Mathematical Programming and Operations Research*, 64 (2015), 2377–2394.

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

[17] 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.

Received October 2016; revised August 2017.

E-mail address: 201411009@xmu.edu.cn
E-mail address: cchuang@whu.edu.cn