TAMKANG JOURNAL OF MATHEMATICS
Volume 50, Number 1, 37-60, March 2019
doi:10.5556/j.tkjm.50.2019.2625

This paper is available online at http://journals.math.tku.edu.tw/index.php/TKJM/pages/view/onlinefirst

PERTURBED SMOOTHING APPROACH TO THE LOWER ORDER EXACT PENALTY FUNCTIONS FOR NONLINEAR INEQUALITY CONSTRAINED OPTIMIZATION

NGUYEN THANH BINH, YANQIN BAI, XIN YAN AND TOUNA YANG

Abstract. In this paper, we propose two new smoothing approximation to the lower order exact penalty functions for nonlinear optimization problems with inequality constraints. Error estimations between smoothed penalty function and nonsmooth penalty function are investigated. By using these new smooth penalty functions, a nonlinear optimization problem with inequality constraints is converted into a sequence of minimizations of continuously differentiable function. Then based on each of the smoothed penalty functions, we develop an algorithm respectively to finding an approximate optimal solution of the original constrained optimization problem and prove the convergence of the proposed algorithms. The effectiveness of the smoothed penalty functions is illustrated through three examples, which show that the algorithm seems efficient.

1. Introduction

Consider the following nonlinear constrained optimization problem:

$$
\begin{aligned}
(P):\qquad & \min\ f(x)\\
& \text{s.t. } g_i(x) \le 0, \quad i \in I = \{1,2,\dots,m\},\\
& \qquad x \in \mathbb{R}^n,
\end{aligned}
$$

where $f:\mathbb{R}^n \to \mathbb{R}$ and $g_i:\mathbb{R}^n \to \mathbb{R}$, $i \in I$, are twice continuously differentiable functions.

Let $X_0 = \{x \in \mathbb{R}^n \mid g_i(x) \le 0,\ i \in I\}$ be the feasible set of (P) and we assume that X_0 is not empty.

Penalty function method is a powerful method for solving general nonlinear constrained optimization problem. It converts the constrained optimization problem to a series of unconstrained problems, by adding a penalty term to the objective function. By adjusting the

Received October 16, 2017, accepted May 14, 2018.
2010 *Mathematics Subject Classification.* 90C30, 65K05, 57R12.
Key words and phrases. Nonlinear constrained optimization, exact penalty function, smoothing method, optimal solution.
Corresponding author: Nguyen Thanh Binh.

penalty parameter, the solutions of these unconstrained problems converge to the optimal solution of the original constrained optimization problem.

Conventional quadratic penalty function method usually requires that the penalty parameter tends to infinity, which is undesirable in practical computation. To tackle this issue, the l_1 exact penalty function is developed:

$$\psi_\rho^1(x) = f(x) + \rho \sum_{i=1}^{m} \max\{0, g_i(x)\}, \tag{1.1}$$

where $\rho > 0$ is a penalty parameter. It is proved that there exists a fixed constant $\rho_0 > 0$, for any $\rho > \rho_0$, any global solution of the exact penalty problem is also a global solution of the original problem. Therefore, the exact penalty function methods have been widely used for solving constrained optimization problems (see, e.g., [1, 6, 7, 8, 12, 19]).

Since the traditional l_1 exact penalty function is not a smooth function, which prevents the use of gradient-based method and causes some numerical instability problems in its implementation, when the value of the penalty parameter becomes large[6, 7, 8, 12]. In order to use existing gradient-based algorithms, such as Newton method, it is necessary to smooth the exact penalty function. Thus, the smoothing of the exact penalty function attracts much attention (see, e.g., [3, 5, 9, 11, 14, 20]).

In the literature of [16], a novel exact penalty method was proposed for solving semi-infinite programming problems, and later, by introducing a new variable, this exact penalty function method was extended to solve nonlinear mixed discrete programming problems [17]. Furthermore, exact penalty function method was proposed for solving a class of discrete-valued optimal control problems [18]. It is shown that if the value of the penalty parameter is sufficiently large, then any local minimizer of the corresponding unconstrained optimization problem is a local minimizer of the original problem.

Recently, a class of lower order penalty functions has been investigated in [13] as the following form

$$\psi_\rho^k(x) = f(x) + \rho \sum_{i=1}^{m} \max\{0, g_i(x)\}^k, \tag{1.2}$$

where $k \in (0,\ 1)$. Correspondingly, the lower order penalty problem and the original problem have the same set of global minima when the penalty parameter is sufficiently large. Obviously, if $k = 1$ the lower order penalty function $\psi_\rho^k(x)$ is reduced to the l_1 exact penalty function. However, both the penalty function $\psi_\rho^k(x)$ $(0 < k < 1)$ and the l_1 exact penalty function are not differentiable at x such that $g_i(x) = 0$ for some $i \in I$. When $k = \frac{1}{2}$, Meng et al. [9] discussed two smoothing approximations to the lower order penalty function for inequality constrained optimization. Binh [4] and Wu et al. [13] also proposed the ϵ-smoothing of (1.2), and obtained a modified exact penalty function under some mild conditions. Wu et al. [14]

proposed a quadratic smoothing approximation to the l_1 exact penalty function. It is shown that under certain conditions, any global minimizer of the smoothed penalty problem is a global minimizer of the original problem when the penalty parameter is sufficiently large.

In this study, motivated by the smoothing techniques in [13, 14], we introduce a new smoothing approximation to the lower order penalty functions. First, we define a new smoothing function $p^k_{\epsilon,\rho}(t) : R \to R$ as follows:

$$p^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \dfrac{m^2\rho^2}{6\epsilon^2} t^{3k} + \dfrac{m\rho}{4\epsilon} t^{2k} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ t^k - \dfrac{7\epsilon}{12m\rho} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

where $\frac{1}{2} < k < 1$, $\epsilon > 0$ and $\rho > 0$. By considering this smoothing function, a perturbed smooth exact penalty function $\psi^k_{\epsilon,\rho}(x)$ is obtained. Using the perturbed smoothing exact penalty function, we are able to convert a constrained optimization problem into the minimizations of a sequence of continuously differentiable functions. Then we propose an algorithm for solving the corresponding penalty problem and discuss its convergence property. We test problems to demonstrate the effectiveness of the proposed algorithm and compare the results obtained with other similar algorithms.

The rest of this paper is organized as follows. In Section 2, we propose a new smoothing method for smoothing the lower order penalty function (1.2) to obtain a first-order continuously differentiable penalty function, we prove some results for error estimates among the optimal objective function values of the nonsmooth penalty problem, smoothed penalty problem and original constrained optimization problem. In Section 3, we propose another method for smoothing the lower order penalty function (1.2) to obtain a second-order continuously differentiable penalty function, and some of its fundamental properties are proved. In Section 4, based on each of the smoothed penalty function, we construct the minimization algorithm respectively to finding an approximate optimal solution of the constrained optimization problems. In Section 5, some numerical examples are given. Finally, conclusions are discussed in Section 6.

2. A new first-order perturbed smoothing method

Consider the non-lipschitz function:

$$p^k(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ t^k & \text{if} \quad t \ge 0, \end{cases}$$

where $0 < k < 1$. Clearly, the function $p^k(t)$ is not differentiable for $0 < k < 1$. It was shown in [4, 11] that the $p^k(t)$ is used to define an exact penalty function for solving constrained optimization problems. We have the following low order penalty function:

$$\psi_\rho^k(x) = f(x) + \rho \sum_{i=1}^{m} p^k(g_i(x)), \tag{2.1}$$

and the corresponding penalty problem

$$(P_\rho): \quad \min\ \psi_\rho^k(x) \qquad \text{s.t. } x \in \mathbb{R}^n.$$

To proceed, we need the following assumption.

Assumption 2.1. *$f(x)$ satisfies the following coercive condition:*

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Under Assumption 2.1, there exists a box X such that $G(P) \subset \operatorname{int}(X)$, where $G(P)$ is the set of global solutions of (P). Let $\operatorname{int}(X)$ be the interior of the set X.

Consider the following problem:

$$\begin{aligned} (P'): \qquad & \min\ f(x) \\ & \text{s.t. } g_i(x) \le 0, \quad i \in I, \\ & \qquad x \in X \subset \mathbb{R}^n. \end{aligned}$$

Let $G(P')$ denote the set of global solutions of (P'). Then $G(P') = G(P)$.

In this paper, we say that the pair (x^*, λ^*) satisfies the second-order sufficiency condition [[2], page 169] if

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) &= 0, \\ g_i(x^*) &\le 0, \qquad i = 1, \dots, m, \\ \lambda^* &\ge 0, \qquad i = 1, \dots, m, \\ y^T \nabla^2 L(x^*, \lambda^*) y &> 0, \qquad \text{for any } y \in V(x^*), \end{aligned}$$

where $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$ and

$$\begin{aligned} V(x^*) &= \left\{ y \in \mathbb{R}^n \,\middle|\, \begin{array}{l} \nabla^T g_i(x^*) y = 0, \ \ i \in A(x^*) \\ \nabla^T g_i(x^*) y \le 0, \ \ i \in B(x^*) \end{array} \right\}, \\ A(x^*) &= \{ i \in I \,|\, g_i(x^*) = 0, \ \ \lambda_i^* > 0 \}, \\ B(x^*) &= \{ i \in I \,|\, g_i(x^*) = 0, \ \ \lambda_i^* = 0 \}. \end{aligned}$$

Now, we consider following penalty problem:

$$(P'_\rho): \quad \min\ \psi^k_\rho(x) \qquad \text{s.t. } x \in X.$$

By Corollary 2.3 in [13] and Theorem 2.1 in [14], we have the following lemma:

Lemma 2.1. *Under Assumption 2.1, and for any $x^* \in G(P)$, there exists a $\mu \in \mathbb{R}^m_+$ such that the pair (x^*, μ^*) satisfies the second-order sufficiency condition of problem (P). Suppose the set $G(P)$ is a finite set. Then there exists a $\rho_0 > 0$ such that when $\rho > \rho_0$, $G(P) = G(P'_\rho)$, where $G(P'_\rho)$ is the set of global solutions of (P'_ρ).*

Next, we consider the smoothing penalty function of the lower order penalty function (2.1). As previously mentioned, for $\frac{1}{2} < k < 1$, $\epsilon > 0$ and $\rho > 0$, the function $p^k_{\epsilon,\rho}(t)$ is defined as:

$$p^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \dfrac{m^2\rho^2}{6\epsilon^2} t^{3k} + \dfrac{m\rho}{4\epsilon} t^{2k} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ t^k - \dfrac{7\epsilon}{12 m\rho} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

Note that, the behavior of $p^k(t)$ and $p^k_{\epsilon,\rho}(t)$ is illustrated in Figure 1. In the following, we discuss the properties of $p^k_{\epsilon,\rho}(t)$.

Lemma 2.2. *For any $\epsilon > 0$, $\rho > 0$, we have*

(i) *$p^k_{\epsilon,\rho}(t)$ is continuously differentiable for $\frac{1}{2} < k < 1$ on $\mathbb{R}$, where*

$$[p^k_{\epsilon,\rho}(t)]' = \begin{cases} 0 & \textit{if} \quad t \le 0, \\ \dfrac{km^2\rho^2}{2\epsilon^2} t^{3k-1} + \dfrac{km\rho}{2\epsilon} t^{2k-1} & \textit{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ kt^{k-1} & \textit{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

(ii) $\lim\limits_{\epsilon \to 0} p^k_{\epsilon,\rho}(t) = p^k(t)$.

(iii) $p^k(t) \ge p^k_{\epsilon,\rho}(t), \ \forall t \in \mathbb{R}$.

Proof. (i) We prove that $p^k_{\epsilon,\rho}(t)$ is continuously differentiable, i.e. $[p^k_{\epsilon,\rho}(t)]'$ is continuous. Actually, we only need to prove that $[p^k_{\epsilon,\rho}(t)]'$ continuous at the separating points: 0 and $(\frac{\epsilon}{m\rho})^{\frac{1}{k}}$.
(1) For $t = 0$, we have

$$\lim_{t\to 0^-} [p^k_{\epsilon,\rho}(t)]' = \lim_{t\to 0^-} 0 = 0, \quad \lim_{t\to 0^+} [p^k_{\epsilon,\rho}(t)]' = \lim_{t\to 0^+} \left[\frac{km^2\rho^2}{2\epsilon^2} t^{3k-1} + \frac{km\rho}{2\epsilon} t^{2k-1} \right] = 0,$$

which implies that

$$\lim_{t\to 0^-}[p^k_{\epsilon,\rho}(t)]' = \lim_{t\to 0^+}[p^k_{\epsilon,\rho}(t)]' = 0 = [p^k_{\epsilon,\rho}(0)]'.$$

Thus, $[p^k_{\epsilon,\rho}(t)]'$ is continuous at $t=0$.

(2) For $t = (\frac{\epsilon}{m\rho})^{\frac{1}{k}}$, we have

$$\lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-}[p^k_{\epsilon,\rho}(t)]' = \lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-}\left[\frac{km^2\rho^2}{2\epsilon^2}t^{3k-1} + \frac{km\rho}{2\epsilon}t^{2k-1}\right] = k\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}},$$

$$\lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+}[p^k_{\epsilon,\rho}(t)]' = \lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+} kt^{k-1} = k\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}},$$

which implies that

$$\lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-}[p^k_{\epsilon,\rho}(t)]' = \lim_{t\to\left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+}[p^k_{\epsilon,\rho}(t)]' = k\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}} = \left[p^k_{\epsilon,\rho}\left((\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right)\right]'.$$

Thus, $[p^k_{\epsilon,\rho}(t)]'$ is continuous at $t = (\frac{\epsilon}{m\rho})^{\frac{1}{k}}$.

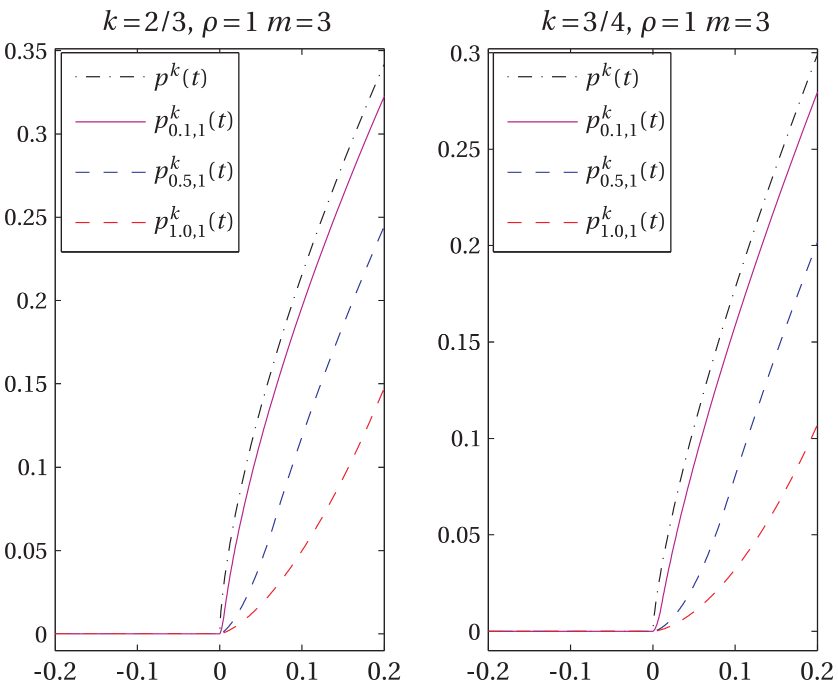

Figure 1: The behavior of $p^k(t)$ and $p^k_{\epsilon,\rho}(t)$.

(ii) For any $t \in \mathbb{R}$, by the definition of $p^k(t)$ and $p^k_{\epsilon,\rho}(t)$ we have

$$p^k(t) - p^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ t^k - \dfrac{m\rho}{4\epsilon} t^{2k} - \dfrac{m^2\rho^2}{6\epsilon^2} t^{3k} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ \dfrac{7\epsilon}{12m\rho} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

If $0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$, let $u = t^k$. Then, we have $0 \le u \le \frac{\epsilon}{m\rho}$. Consider the function:

$$H(u) = u - \frac{m\rho}{4\epsilon} u^2 - \frac{m^2\rho^2}{6\epsilon^2} u^3, \quad 0 \le u \le \frac{\epsilon}{m\rho},$$

and we have

$$H'(u) = 1 - \frac{m\rho}{2\epsilon} u - \frac{m^2\rho^2}{2\epsilon^2} u^2, \quad 0 \le u \le \frac{\epsilon}{m\rho}.$$

Obviously, $H'(u) \ge 0$ for $0 \le u \le \frac{\epsilon}{m\rho}$. Moreover, $H(0) = 0$ and $H(\frac{\epsilon}{m\rho}) = \frac{7\epsilon}{12m\rho}$. Hence, we have

$$0 \le p^k(t) - p^k_{\epsilon,\rho}(t) \le \frac{7\epsilon}{12m\rho}.$$

That is,

$$\lim_{\epsilon \to 0} p^k_{\epsilon,\rho}(t) = p^k(t).$$

(iii) For any $t \in \mathbb{R}$, from (ii), we have

$$p^k(t) - p^k_{\epsilon,\rho}(t) \ge 0,$$

which is $p^k(t) \ge p^k_{\epsilon,\rho}(t)$.
This completes the proof. □

Consider the perturbed smooth exact penalty function as follows:

$$\psi^k_{\epsilon,\rho}(x) = f(x) + \rho \sum_{i=1}^{m} p^k_{\epsilon,\rho}\left(g_i(x)\right), \tag{2.2}$$

where $\epsilon > 0$, $\rho > 0$. Clearly, $\psi^k_{\epsilon,\rho}(x)$ is continuously differentiable at any $x \in \mathbb{R}^n$. The corresponding smoothed optimization problem is:

$$(SP_{\epsilon,\rho}): \quad \min\ \psi^k_{\epsilon,\rho}(x) \qquad \text{s.t. } x \in X.$$

Lemma 2.3. *We have that*

$$0 \le \psi^k_{\rho}(x) - \psi^k_{\epsilon,\rho}(x) \le \frac{7\epsilon}{12}, \tag{2.3}$$

for any $x \in X$, $\epsilon > 0$ and $\rho > 0$.

Proof. For any $x \in X$, we have

$$\psi^k_{\rho}(x) - \psi^k_{\epsilon,\rho}(x) = \rho \sum_{i=1}^{m} \left(p^k(g_i(x)) - p^k_{\epsilon,\rho}(g_i(x))\right).$$

Note that

$$p^k\left(g_i(x)\right) - p^k_{\epsilon,\rho}\left(g_i(x)\right) = \begin{cases} 0 & \text{if} \quad g_i(x) \le 0, \\ [g_i(x)]^k - \dfrac{m^2\rho^2}{6\epsilon^2}[g_i(x)]^{3k} - \dfrac{m\rho}{4\epsilon}[g_i(x)]^{2k} & \text{if} \quad 0 \le g_i(x) \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ \dfrac{7\epsilon}{12m\rho} & \text{if} \quad g_i(x) \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

for any $i \in I$.
From the proof of Lemma 2.2, we have

$$0 \le \sum_{i=1}^m \left(p^k(g_i(x)) - p^k_{\epsilon,\rho}(g_i(x))\right) \le \frac{7\epsilon}{12\rho},$$

which implies that

$$0 \le \rho\sum_{i=1}^m \left(p^k(g_i(x)) - p^k_{\epsilon,\rho}(g_i(x))\right) \le \frac{7\epsilon}{12},$$

and hence we have

$$0 \le \psi^k_\rho(x) - \psi^k_{\epsilon,\rho}(x) \le \frac{7\epsilon}{12}.$$

This completes the proof. □

Lemma 2.3 mean that the gap between $\psi^k_\rho(x)$ and $\psi^k_{\epsilon,\rho}(x)$ can be arbitrarily small if the smoothing parameter ϵ is sufficiently small.

Lemma 2.4. *Let x^* and $x^*_\rho \in X$ be the optimal solutions of (P') and (P'_ρ), respectively. If x^*_ρ is a feasible solution of (P'), x^*_ρ is an optimal solution of (P').*

Proof. Under the given conditions, we have

$$f(x^*_\rho) = \psi^k_\rho(x^*_\rho) \le \psi^k_\rho(x^*) = f(x^*),$$

which is

$$f(x^*_\rho) \le f(x^*).$$

Since x^* is an optimal solution and x^*_ρ is feasible of (P'), we have

$$f(x^*_\rho) \ge f(x^*).$$

Therefore, x^*_ρ is an optimal solution of (P'). □

Theorem 2.5. *Let x^*_ρ and $x^*_{\epsilon,\rho} \in X$ be the optimal solutions of (P'_ρ) and $(SP_{\epsilon,\rho})$, respectively, for some $\rho > 0$ and $\epsilon > 0$. Then, we have that*

$$0 \le \psi^k_\rho(x^*_\rho) - \psi^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}) \le \frac{7\epsilon}{12}. \tag{2.4}$$

Proof. From Lemma 2.3, we obtain

$$\begin{aligned}0 \le \psi_\rho^k(x_\rho^*) - \psi_{\epsilon,\rho}^k(x_\rho^*) &\le \psi_\rho^k(x_\rho^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*)\\ &\le \psi_\rho^k(x_{\epsilon,\rho}^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*)\\ &\le \frac{7\epsilon}{12}.\end{aligned}$$

Therefore,

$$0 \le \psi_\rho^k(x_\rho^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}.$$

This completes the proof. □

Lemma 2.1 and Theorem 2.5 yield the following theorem:

Theorem 2.6. *Suppose that x^* satisfies the assumptions of Lemma 2.1. Let x^* and $x_{\epsilon,\rho}^* \in X$ be the optimal solutions of (P) and $(SP_{\epsilon,\rho})$, respectively. Then there exists $\rho_0 > 0$ such that for any $\rho > \rho_0$, it holds that*

$$0 \le f(x^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}. \tag{2.5}$$

Proof. From Lemma 2.1, we have that x^* is an optimal solution of (P_ρ'). Then from Theorem 2.5, we have

$$0 \le \psi_\rho^k(x^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}.$$

Note that

$$\psi_\rho^k(x^*) = f(x^*) + \rho \sum_{i=1}^m p^k(g_i(x^*)).$$

Since $\sum_{i=1}^m p^k(g_i(x^*)) = 0$, we have $\psi_\rho^k(x^*) = f(x^*)$. Thus, we have that

$$0 \le f(x^*) - \psi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}.$$

This completes the proof. □

Theorem 2.7. *Suppose that x_ρ^* satisfies the conditions in Lemma 2.4 and $x_{\epsilon,\rho}^* \in X$ is an optimal solution of $(SP_{\epsilon,\rho})$ for some $\rho > 0$ and $\epsilon > 0$. If $x_{\epsilon,\rho}^*$ is feasible solution of (P), then we have that*

$$0 \le f(x_\rho^*) - f(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}. \tag{2.6}$$

i.e., $x_{\epsilon,\rho}^*$ is an approximate optimal solution of (P).

Proof. By Theorem 2.5, we have

$$0 \le f(x_\rho^*) + \rho \sum_{i=1}^m p^k(g_i(x_\rho^*)) - \left(f(x_{\epsilon,\rho}^*) + \rho \sum_{i=1}^m p_{\epsilon,\rho}^k(g_i(x_{\epsilon,\rho}^*)) \right) \le \frac{7\epsilon}{12}.$$

Since x_ρ^* and $x_{\epsilon,\rho}^*$ are feasible solutions of (P), we have

$$\sum_{i=1}^m p^k(g_i(x_\rho^*)) = \sum_{i=1}^m p_{\epsilon,\rho}^k(g_i(x_{\epsilon,\rho}^*)) = 0.$$

Therefore,

$$0 \le f(x_\rho^*) - f(x_{\epsilon,\rho}^*) \le \frac{7\epsilon}{12}.$$

From Lemma 2.4, x_ρ^* is actually an optimal solution of (P). Thus, $x_{\epsilon,\rho}^*$ is an approximate optimal solution of (P). This completes the proof. □

By Theorem 2.7, we conclude that the difference between the objective function values on an optimal solution of $(SP_{\epsilon,\rho})$ and an optimal solution of (P) can be controlled through the smoothing parameter ϵ, and the optimal solution of $(SP_{\epsilon,\rho})$ is an approximate optimal solution of (P) if x_ρ^* and $x_{\epsilon,\rho}^*$ are feasible.

Now, we assume that the problem (P) is convex. By Proposition 2.1 in [14], the corresponding smooth penalty problem $(SP_{\epsilon,\rho})$ for (P) is a convex problem. The following theorem shows that under some mild conditions, an optimal solution of $(SP_{\epsilon,\rho})$ becomes an approximate optimal solution of (P). First, we recall the definition of KKT point.

Definition 2.8. A feasible solution x^* of (P) is called a KKT point, if there exists a $\mu^* \in \mathbb{R}^m$ such that the pair (x^*, μ^*) satisfies the following conditions

$$\nabla f(x^*) + \sum_{i=1}^m \mu_i^* \nabla g_i(x^*) = 0, \tag{2.7}$$

$$\mu_i^* g_i(x^*) = 0,\ g_i(x^*) \le 0,\ \mu_i^* \ge 0,\ \ i \in I. \tag{2.8}$$

Theorem 2.9. *Suppose the functions f, g_i $(i \in I)$ in problem (P) are convex. Let x^* and $x_{\epsilon,\rho}^* \in X$ be the optimal solutions of (P) and $(SP_{\epsilon,\rho})$, respectively. If $x_{\epsilon,\rho}^*$ is feasible of (P), and there exists a $\mu^* \in \mathbb{R}^m$ such that the pair $(x_{\epsilon,\rho}^*, \mu^*)$ satisfies the conditions in Equations (2.7) and (2.8), then we have that*

$$0 \le f(x_{\epsilon,\rho}^*) - f(x^*) \le \frac{7\epsilon}{12}. \tag{2.9}$$

Proof. Since f, g_i $(i \in I)$ are continuously differentiable and convex, we see that

$$f(x^*) \ge f(x_{\epsilon,\rho}^*) + \nabla f(x_{\epsilon,\rho}^*)^T (x^* - x_{\epsilon,\rho}^*), \tag{2.10}$$

$$g_i(x^*) \ge g_i(x_{\epsilon,\rho}^*) + \nabla g_i(x_{\epsilon,\rho}^*)^T (x^* - x_{\epsilon,\rho}^*), \quad i = 1,2,\dots,m. \tag{2.11}$$

By Equations (2.1), (2.7), (2.8), (2.10) and (2.11), we have that

$$\psi_\rho^k(x^*) = f(x^*) + \rho \sum_{i=1}^m p^k(g_i(x^*))$$

$$
\begin{aligned}
&\geq f(x^*_{\epsilon,\rho}) + \nabla f(x^*_{\epsilon,\rho})^T(x^* - x^*_{\epsilon,\rho}) \\
&= f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*_{\epsilon,\rho})^T(x^* - x^*_{\epsilon,\rho}) \\
&\geq f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* \left[g_i(x^*) - g_i(x^*_{\epsilon,\rho}) \right] \\
&= f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* g_i(x^*) \\
&\geq f(x^*_{\epsilon,\rho}).
\end{aligned}
$$

From Lemma 2.3, we obtain

$$\psi_\rho^k(x^*) \leq \psi_{\epsilon,\rho}^k(x^*) + \frac{7\epsilon}{12}.$$

It follows that

$$
\begin{aligned}
f(x^*_{\epsilon,\rho}) &\leq \psi_{\epsilon,\rho}^k(x^*) + \frac{7\epsilon}{12} \\
&= f(x^*) + \rho \sum_{i=1}^{m} p_{\epsilon,\rho}^k(g_i(x^*)) + \frac{7\epsilon}{12} \\
&= f(x^*) + \frac{7\epsilon}{12}.
\end{aligned} \tag{2.12}
$$

Since $x^*_{\epsilon,\rho}$ is feasible of (P), which is

$$f(x^*) \leq f(x^*_{\epsilon,\rho}). \tag{2.13}$$

Combining Equations (2.12) and (2.13), we have

$$f(x^*) \leq f(x^*_{\epsilon,\rho}) \leq f(x^*) + \frac{7\epsilon}{12},$$

which is

$$0 \leq f(x^*_{\epsilon,\rho}) - f(x^*) \leq \frac{7\epsilon}{12}.$$

This completes the proof. □

Note that, the smooth penalty function $\psi_{\epsilon,\rho}^k(x)$ is only first-order differentiable. If we want to use a Newton-type method, a smoothing penalty function must be second-order differentiable. In the next section, we present a second-order perturbed smooth penalty function method.

3. A new second-order perturbed smoothing method

This section, we propose a method for smoothing the lower order penalty function (1.2) to obtain a second-order continuously differentiable penalty function. We define the following smoothing function:

$$q^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \dfrac{m^2\rho^2}{2\epsilon^2} t^{3k} - \dfrac{m^3\rho^3}{5\epsilon^3} t^{4k} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ t^k + \dfrac{3\epsilon^2}{10m^2\rho^2} t^{-k} - \dfrac{\epsilon}{m\rho} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

where $\frac{1}{3} < k < 1$, $\epsilon > 0$ and $\rho > 0$.

Figure 2 shows the behavior of $p^k(t)$ and $q^k_{\epsilon,\rho}(t)$. Clearly, the function $q^k_{\epsilon,\rho}(t)$ has the following properties.

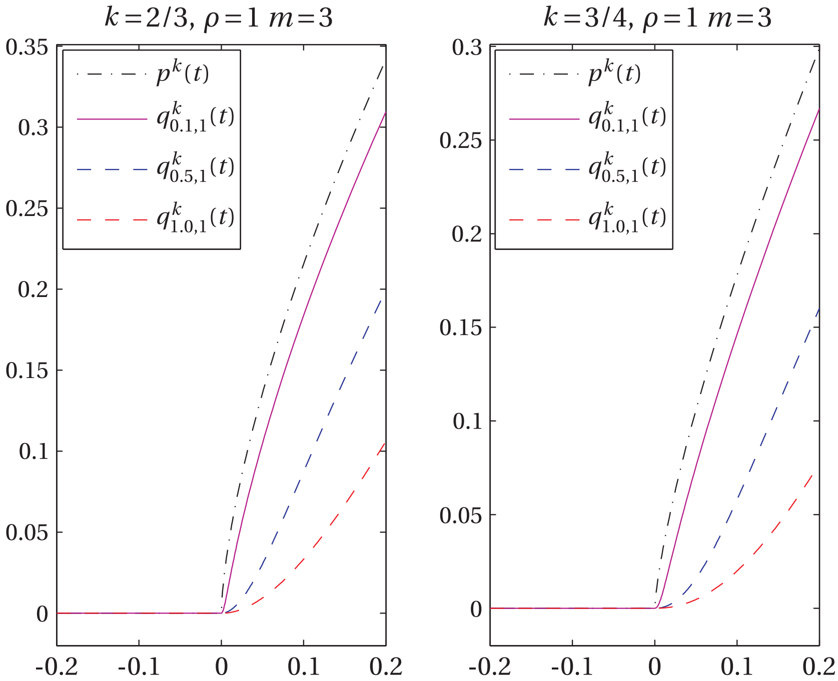

Figure 2: The behavior of $p^k(t)$ and $q^k_{\epsilon,\rho}(t)$.

Lemma 3.1. *For any* $\epsilon > 0,\ \rho > 0$, *we have*

(i) $q^k_{\epsilon,\rho}(t)$ *is twice continuously differentiable for* $\frac{2}{3} < k < 1$ *on* $\mathbb{R}$, *where*

$$[q^k_{\epsilon,\rho}(t)]' = \begin{cases} 0 & \textit{if} \quad t \le 0, \\ \dfrac{3km^2\rho^2}{2\epsilon^2} t^{3k-1} - \dfrac{4km^3\rho^3}{5\epsilon^3} t^{4k-1} & \textit{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ kt^{k-1} - \dfrac{3k\epsilon^2}{10m^2\epsilon^2} t^{-k-1} & \textit{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

and

$$[q^k_{\epsilon,\rho}(t)]'' = \begin{cases} 0 & \textit{if} \quad t \le 0, \\ \dfrac{3k(3k-1)m^2\rho^2}{2\epsilon^2} t^{3k-2} - \dfrac{4k(4k-1)m^3\rho^3}{5\epsilon^3} t^{4k-2} & \textit{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ k(k-1)t^{k-2} + \dfrac{3k(k+1)\epsilon^2}{10m^2\epsilon^2} t^{-k-2} & \textit{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

(ii) $\lim\limits_{\epsilon \to 0} q^k_{\epsilon,\rho}(t) = p^k(t)$.

(iii) $p^k(t) \ge q^k_{\epsilon,\rho}(t), \quad \forall t \in \mathbb{R}$.

Let

$$\varphi^k_{\epsilon,\rho}(x) = f(x) + \rho \sum_{i=1}^{m} q^k_{\epsilon,\rho}\left(g_i(x)\right), \tag{3.1}$$

where $\epsilon > 0,\ \rho > 0$. Then, $\varphi^k_{\epsilon,\rho}(x)$ is twice continuously differentiable at any $x \in \mathbb{R}^n$. We have the following smoothed optimization problem:

$$(PI_{\epsilon,\rho}): \quad \min\ \varphi^k_{\epsilon,\rho}(x) \qquad \text{s.t.}\ \ x \in X.$$

Lemma 3.2. *For any* $x \in X$, $\epsilon > 0,\ \rho > 0$, *we have*

$$0 \le \psi^k_\rho(x) - \varphi^k_{\epsilon,\rho}(x) \le \epsilon, \tag{3.2}$$

where $\psi^k_\rho(x)$ *and* $\varphi^k_{\epsilon,\rho}(x)$ *are given in* (2.1) *and* (3.1), *respectively.*

Proof. For any $x \in X$, we have

$$\psi^k_\rho(x) - \varphi^k_{\epsilon,\rho}(x) = \rho \sum_{i=1}^{m} \left(p^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x))\right).$$

Note that

$$\begin{aligned} & p^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x)) \\ &= \begin{cases} 0 & \text{if} \quad g_i(x) \le 0, \\ 0 \le [g_i(x)]^k - \left(\dfrac{m^2\rho^2}{2\epsilon^2}[g_i(x)]^{3k} - \dfrac{m^3\rho^3}{5\epsilon^3}[g_i(x)]^{4k}\right) < \dfrac{\epsilon}{m\rho} & \text{if} \quad 0 \le g_i(x) \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ 0 \le \dfrac{\epsilon}{m\rho} - \dfrac{3\epsilon^2}{10m^2\rho^2}[g_i(x)]^{-k} \le \dfrac{\epsilon}{m\rho} & \text{if} \quad g_i(x) \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases} \end{aligned}$$

for any $i = 1,2,\dots,m$. That is,

$$0 \le p^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x)) \le \frac{\epsilon}{m\rho}.$$

Thus,

$$0 \le \sum_{i=1}^{m} \left(p^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x))\right) \le \frac{\epsilon}{\rho},$$

which implies

$$0 \le \rho \sum_{i=1}^{m} \left(p^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x))\right) \le \epsilon.$$

Therefore,

$$0 \le \psi^k_\rho(x) - \varphi^k_{\epsilon,\rho}(x) \le \epsilon.$$

This completes the proof. □

Based on the Lemma 3.2, we have the following two theorems.

Theorem 3.3. *Let $\{\epsilon_j\} \to 0$, $\forall \epsilon_j > 0$, and x_j be a solution of $(PI_{\epsilon_j,\rho})$ for $\rho > 0$. Assume that x' is an accumulation point of $\{x_j\}$. Then, x' is an optimal solution of (P_ρ).*

Proof. Since x_j is a solution of $(PI_{\epsilon_j,\rho})$, we have

$$\varphi^k_{\epsilon_j,\rho}(x_j) \le \varphi^k_{\epsilon_j,\rho}(x).$$

By Lemma 3.2, we have

$$\psi^k_\rho(x_j) \le \varphi^k_{\epsilon_j,\rho}(x_j) + \epsilon_j,$$
$$\varphi^k_{\epsilon_j,\rho}(x) \le \psi^k_\rho(x).$$

It follows,

$$\begin{aligned}\psi^k_\rho(x_j) &\le \varphi^k_{\epsilon_j,\rho}(x) + \epsilon_j \\ &\le \psi^k_\rho(x) + \epsilon_j.\end{aligned}$$

Since $\{\epsilon_j\} \to 0$ and x' is an accumulation point of $\{x_j\}$, we obtain

$$\psi^k_\rho(x') \le \psi^k_\rho(x).$$

This completes the proof. □

Theorem 3.4. *For some $\rho > 0$ and $\epsilon > 0$, let x^*_ρ and $x^*_{\epsilon,\rho} \in \mathbb{R}^n$ be optimal solutions of (P_ρ) and $(PI_{\epsilon,\rho})$, respectively. Then, we have*

$$0 \le \psi^k_\rho(x^*_\rho) - \varphi^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}) \le \epsilon. \tag{3.3}$$

If both x^*_ρ and $x^*_{\epsilon,\rho}$ are feasible of (P), then

$$f(x^*_{\epsilon,\rho}) \le f(x^*_\rho) \le f(x^*_{\epsilon,\rho}) + \epsilon. \tag{3.4}$$

Proof. By Lemma 3.2, for $\rho > 0$ and $\epsilon > 0$, we have that

$$\begin{aligned} 0 &\le \psi_\rho^k(x_\rho^*) - \varphi_{\epsilon,\rho}^k(x_\rho^*) \\ &\le \psi_\rho^k(x_\rho^*) - \varphi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \\ &\le \psi_\rho^k(x_{\epsilon,\rho}^*) - \varphi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \\ &\le \epsilon. \end{aligned}$$

That is,

$$0 \le \psi_\rho^k(x_\rho^*) - \varphi_{\epsilon,\rho}^k(x_{\epsilon,\rho}^*) \le \epsilon$$

and

$$0 \le \left\{ f(x_\rho^*) + \rho \sum_{i=1}^m p^k(g_i(x_\rho^*)) \right\} - \left\{ f(x_{\epsilon,\rho}^*) + \rho \sum_{i=1}^m q_{\epsilon,\rho}^k(g_i(x_{\epsilon,\rho}^*)) \right\} \le \epsilon.$$

Furthermore, if x_ρ^* and $x_{\epsilon,\rho}^*$ are feasible of (P), then

$$\sum_{i=1}^m p^k(g_i(x_\rho^*)) = \sum_{i=1}^m q_{\epsilon,\rho}^k(g_i(x_{\epsilon,\rho}^*)) = 0.$$

Therefore,

$$0 \le f(x_\rho^*) - f(x_{\epsilon,\rho}^*) \le \epsilon.$$

That is,

$$f(x_{\epsilon,\rho}^*) \le f(x_\rho^*) \le f(x_{\epsilon,\rho}^*) + \epsilon.$$

This completes the proof. □

Theorem 3.5. *Suppose the functions $f(x)$, $g_i(x)$ $(i \in I)$ are convex. Let x^* and $x_{\epsilon,\rho}^* \in X$ be the optimal solutions of (P) and $(PI_{\epsilon,\rho})$, respectively. If $x_{\epsilon,\rho}^*$ is feasible of (P), and there exists a $\lambda^* \in \mathbb{R}^m$ such that the pair $(x_{\epsilon,\rho}^*, \lambda^*)$ satisfies the conditions in Equations (2.7) and (2.8), then we have*

$$f(x^*) \le f(x_{\epsilon,\rho}^*) \le f(x^*) + \epsilon. \tag{3.5}$$

Proof. The proof is similar to the proof of the Theorem 2.9. □

4. Algorithms for minimization procedure

In this section, by considering the above smoothed penalty functions, we propose algorithms to find an approximate optimal solution of (P), defined as Algorithm 4.2 and Algorithm 4.5.

Definition 4.1. A point $x_\epsilon^* \in X$ is called ϵ-feasible solution of (P), if it satisfies $g_i(x_\epsilon^*) \leq \epsilon,\ \forall i \in I$.

Algorithm 4.2. *Step* 1*:* Given the initial point $x_1^0 \in X$ and a stoping tolerance $\epsilon > 0$. Chooose $\epsilon_1 > \epsilon,\ \rho_1 > 0,\ 0 < \gamma < 1,\ N > 1$, and let $j = 1$.

Step 2*:* Star from the point x_j^0 and solve the following problem:

$$(SP_{\epsilon_j,\rho_j}): \quad \min_{x \in \mathbb{R}^n} \psi_{\epsilon_j,\rho_j}^k(x) = f(x) + \rho_j \sum_{i=1}^m p_{\epsilon_j,\rho_j}^k\left(g_i(x)\right).$$

Let x_{ϵ_j,ρ_j}^* be an optimal solution of (SP_{ϵ_j,ρ_j}). Here x_{ϵ_j,ρ_j}^* is obtained by the BFGS method given in [10].

Step 3*:* If x_{ϵ_j,ρ_j}^* is an ϵ-feasible of (P), then the algorithm stop. Otherwise, let $\rho_{j+1} = N\rho_j$, $\epsilon_{j+1} = \gamma\epsilon_j$, $x_{j+1}^0 = x_{\epsilon_j,\rho_j}^*$ and $j = j+1$. Then, go to Step 2.

Remark 4.3. In the Algorithm 4.2 shows that the sequence $\{\epsilon_j\}$ converges to 0 and the sequence $\{\rho_j\}$ converges to $+\infty$, as $j \to +\infty$.

Theorem 4.4. *For $\frac{1}{2} < k < 1$, suppose that for any $\epsilon \in (0, \epsilon_1]$, $\rho \in [\rho_1, +\infty)$, the set*

$$\arg\min_{x \in \mathbb{R}^n} \psi_{\epsilon,\rho}^k(x) \neq \emptyset. \tag{4.1}$$

Let $\{x_{\epsilon_j,\rho_j}^*\}$ be the sequence generated by Algorithm 4.2. If the sequence $\{\psi_{\epsilon_j,\rho_j}^k(x_{\epsilon_j,\rho_j}^*)\}$ is bounded, and Assumption 2.1 holds, then $\{x_{\epsilon_j,\rho_j}^*\}$ is bounded and the limit point of $\{x_{\epsilon_j,\rho_j}^*\}$ is a solution of (P).

Proof. First, we prove that $\{x_{\epsilon_j,\rho_j}^*\}$ is bounded. Note that

$$\psi_{\epsilon_j,\rho_j}^k(x_{\epsilon_j,\rho_j}^*) = f(x_{\epsilon_j,\rho_j}^*) + \rho_j \sum_{i=1}^m p_{\epsilon_j,\rho_j}^k(g_i(x_{\epsilon_j,\rho_j}^*)), \quad j = 0, 1, \ldots, \tag{4.2}$$

and by the definition of $p_{\epsilon,\rho}^k(t)$, we have

$$\rho_j \sum_{i=1}^m p_{\epsilon_j,\rho_j}^k(g_i(x_{\epsilon_j,\rho_j}^*)) \geq 0. \tag{4.3}$$

Suppose on the contrary that the sequence $\{x_{\epsilon_j,\rho_j}^*\}$ is unbounded. Without any loss of generality $\|x_{\epsilon_j,\rho_j}^*\| \to +\infty$ as $j \to +\infty$. Then, $\lim_{j \to +\infty} f(x_{\epsilon_j,\rho_j}^*) = +\infty$ by Assumption 2.1, and from Equations (4.2) and (4.3), we have

$$\psi_{\epsilon_j,\rho_j}^k(x_{\epsilon_j,\rho_j}^*) \geq f(x_{\epsilon_j,\rho_j}^*) \to +\infty, \quad j = 0, 1, \ldots,$$

which contradicts with the sequence $\{\psi_{\epsilon_j,\rho_j}^k(x_{\epsilon_j,\rho_j}^*)\}$ being bounded. Thus, $\{x_{\epsilon_j,\rho_j}^*\}$ is bounded.

Next, we prove that the limit point of $\{x^*_{\epsilon_j,\rho_j}\}$ is the solution of (P). Let x^* be a limit point of $\{x^*_{\epsilon_j,\rho_j}\}$. Then, there exists the subset $J \subset \mathbb{N}$ such that $x^*_{\epsilon_j,\rho_j} \to x^*,\ j \in J$, where $\mathbb{N}$ is the set of natural numbers. We have to show that x^* is an optimal solution of (P). Thus, it is sufficient to show $x^* \in X_0$, and $f(x^*) \le \inf_{x\in X_0} f(x)$.

(i) Suppose $x^* \notin X_0$. Then, there exist $\theta_0 > 0$ and the subset $J' \subset J$, such that $g_{i'}(x^*_{\epsilon_j,\rho_j}) \ge \theta_0$ for any $j \in J'$ and some $i' \in I$.

If $\left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}} > g_{i'}(x^*_{\epsilon_j,\rho_j}) \ge \theta_0$, from the definition of $p^k_{\epsilon,\rho}(t)$ and $x^*_{\epsilon_j,\rho_j}$ is the optimal solution according j-th values of the parameters ϵ_j, ρ_j for any $x \in X_0$, we have

$$\begin{aligned} f(x^*_{\epsilon_j,\rho_j}) + \frac{m^2\rho_j^3\theta_0^{3k}}{6\epsilon_j^2} + \frac{m\rho_j^2\theta_0^{2k}}{4\epsilon_j} &\le \psi^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}) \\ &\le \psi^k_{\epsilon_j,\rho_j}(x) = f(x), \end{aligned}$$

which contradicts with $\rho_j \to +\infty$ and $\epsilon_j \to 0$.

If $g_{i'}(x^*_{\epsilon_j,\rho_j}) \ge \theta_0 \ge \left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}}$ or $g_{i'}(x^*_{\epsilon_j,\rho_j}) \ge \left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}} \ge \theta_0$, from the definition of $p^k_{\epsilon,\rho}(t)$ and $x^*_{\epsilon_j,\rho_j}$ is the optimal solution according j-th values of the parameters ϵ_j, ρ_j for any $x \in X_0$, we have

$$\begin{aligned} f(x^*_{\epsilon_j,\rho_j}) + \rho_j\theta_0^k - \frac{7\epsilon_j}{12m} &\le \psi^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}) \\ &\le \psi^k_{\epsilon_j,\rho_j}(x) = f(x), \end{aligned}$$

which contradicts with $\rho_j \to +\infty$ and $\epsilon_j \to 0$.

Thus, $x^* \in X_0$.

(ii) For any $x \in X_0$, it holds that

$$f(x^*_{\epsilon_j,\rho_j}) \le \psi^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}) \le \psi^k_{\epsilon_j,\rho_j}(x) = f(x),$$

thus $f(x^*) \le \inf_{x\in X_0} f(x)$ holds.
This completes the proof. □

Algorithm 4.5. *Step* 1*:* Given the initial point $x_1^0 \in X$ and $\epsilon > 0$. Choose $\epsilon_1 > \epsilon,\ \rho_1 > 0,\ 0 < \gamma < 1,\ N > 1$, and let $j = 1$.

Step 2*:* Star from the point x_j^0 and solve the following problem:

$$(PI_{\epsilon_j,\rho_j}):\quad \min_{x\in\mathbb{R}^n} \varphi^k_{\epsilon_j,\rho_j}(x) = f(x) + \rho_j \sum_{i=1}^m q^k_{\epsilon_j,\rho_j}\left(g_i(x)\right).$$

Let $x^*_{\epsilon_j,\rho_j}$ be the optimal solution obtained.

Step 3*:* If $x^*_{\epsilon_j,\rho_j}$ is an ϵ-feasible of (P), then stops and $x^*_{\epsilon_j,\rho_j}$ is an approximate optimal solution of (P). Otherwise, let $\rho_{j+1} = N\rho_j,\ \epsilon_{j+1} = \gamma\epsilon_j,\ x^0_{j+1} = x^*_{\epsilon_j,\rho_j}$ and $j = j+1$. Then, go to Step 2.

Theorem 4.6. *For $\frac{1}{3} < k < 1$, suppose that the set*

$$\arg\min_{x\in\mathbb{R}^n} \varphi^k_{\epsilon,\rho}(x) \neq \emptyset \tag{4.4}$$

for any $\epsilon \in (0, \epsilon_1]$ and $\rho \in [\rho_1, +\infty)$. Let $\{x'_j\}$ be the sequence generated by Algorithm 4.5. If the sequence $\{\varphi^k_{\epsilon_j,\rho_j}(x'_j)\}$ is bounded, and Assumption 2.1 holds, then $\{x'_j\}$ is bounded and the limit point of $\{x'_j\}$ is a solution of (P).

Proof. The proof is similar to the proof of the Theorem 4.4. □

5. Numerical examples

In this section, we apply our algorithms to test problems. The proposed algorithm is implemented in MATLAB R2011A (version, Manufacturer, City, US State if applicable, Country). In each example, we let $\epsilon = 10^{-6}$ is expected to get an ϵ-solution of (P) with both Algorithm 4.2 and Algorithm 4.5, j be the number of iterations, $x^*_{\epsilon_j,\rho_j}$ be the optimal solution of the j-th iteration, $f(x^*_{\epsilon_j,\rho_j})$ be the objective value at $x^*_{\epsilon_j,\rho_j}$, $g_i(x^*_{\epsilon_j,\rho_j})$, $i \in I$ is a constrain value at $x^*_{\epsilon_j,\rho_j}$, and the numerical results are presented in the tables following.

Example 5.1. Consider the following problem ([9], Example 4.5)

$$\begin{aligned}
(P5.1):\quad \min\ & f(x) = 10x_2 + 2x_3 + x_4 + 3x_5 + 4x_6\\
\text{s.t.}\ & g_1(x) = x_1 + x_2 - 10 = 0,\\
& g_2(x) = -x_1 + x_3 + x_4 + x_5 = 0,\\
& g_3(x) = -x_2 - x_3 + x_5 + x_6 = 0,\\
& g_4(x) = 10x_1 - 2x_3 + 3x_4 - 2x_5 - 16 \leq 0,\\
& g_5(x) = x_1 + 4x_3 + x_5 - 10 \leq 0,\\
& 0 \leq x_1 \leq 12,\\
& 0 \leq x_2 \leq 18,\\
& 0 \leq x_3 \leq 5,\\
& 0 \leq x_4 \leq 12,\\
& 0 \leq x_5 \leq 1,\\
& 0 \leq x_6 \leq 16.
\end{aligned}$$

For $k = \frac{3}{4}$, let $x^0_1 = (2,\ 1,\ 2,\ 2,\ 1,\ 2)$, $\rho_1 = 100$, $N = 3$, $\epsilon_1 = 0.2$, $\gamma = 0.1$. The results of Algorithm 4.2 for solving $(P5.1)$ are shown in Table 1 and Table 2.

For $k = \frac{2}{3}$, let $x^0_1 = (2,\ 1,\ 2,\ 1,\ 1,\ 2)$, $\rho_1 = 100$, $N = 3$, $\epsilon_1 = 0.4$, $\gamma = 0.1$. The results of Algorithm 4.2 for solving $(P5.1)$ are shown in Table 3 and Table 4.

Table 1: Results of Algorithm 4.2 with $x_1^0 = (2, 1, 2, 2, 1, 2)$ for $(P5.1)$.

j	ρ_j	$f(x^*_{\epsilon_j,\rho_j})$	$x^*_{\epsilon_j,\rho_j}$
1	100	117.029887	(1.697078, 8.302923, 0.217670, 0.481162, 0.998241, 7.522358)
2	300	117.000000	(1.715453, 8.284547, 0.260178, 0.455275, 1.000000, 7.544725)

Table 2: Results of Algorithm 4.2 with $x_1^0 = (2, 1, 2, 2, 1, 2)$ for $(P5.1)$.

j	ϵ_j	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$	$g_4(x^*_{\epsilon_j,\rho_j})$	$g_5(x^*_{\epsilon_j,\rho_j})$
1	0.2	-0.000000	-0.000004	-0.000006	-0.017559	-6.433999
2	0.02	-0.000000	-0.000000	-0.000000	-0.000000	-6.243834

Table 3: Results of Algorithm 4.2 with $x_1^0 = (2, 1, 2, 1, 1, 2)$ for $(P5.1)$.

j	ρ_j	$f(x^*_{\epsilon_j,\rho_j})$	$x^*_{\epsilon_j,\rho_j}$
1	100	117.145419	(1.618015, 8.381985, 0.036170, 0.581211, 1.000640, 7.417524)
2	300	117.000523	(1.691308, 8.308692, 0.197503, 0.493808, 0.999996, 7.506199)

Table 4: Results of Algorithm 4.2 with $x_1^0 = (2, 1, 2, 1, 1, 2)$ for $(P5.1)$.

j	ϵ_j	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$	$g_4(x^*_{\epsilon_j,\rho_j})$	$g_5(x^*_{\epsilon_j,\rho_j})$
1	0.4	-0.000001	-0.000006	-0.000008	-0.149835	-7.236664
2	0.04	-0.000000	-0.000000	-0.000000	-0.000495	-6.518683

Table 5: Results of Algorithm 4.5 with $x_1^0 = (1, 2, 1, 0, 1, 0)$ for $(P5.1)$.

j	ρ_j	$f(x^*_{\epsilon_j,\rho_j})$	$x^*_{\epsilon_j,\rho_j}$
1	1000	116.885478	(1.863857, 8.129143, 0.644751, 0.217036, 1.000069, 7.771825)
2	3000	116.961945	(1.627743, 8.369924, 0.031724, 0.595345, 1.000007, 7.400974)
3	9000	117.004035	(1.834835, 8.164391, 0.573505, 0.261764, 0.999341, 7.738332)

For $k = \frac{2}{3}$, let $x_1^0 = (1, 2, 1, 0, 1, 0)$, $\rho_1 = 1000$, $N = 3$, $\epsilon_1 = 0.01$, $\gamma = 0.1$. The results of Algorithm 4.5 for solving $(P5.1)$ are shown in Table 5 and Table 6.

The results in Tables 1-6 show that, the convergence of both Algorithm 4.2 and Algorithm 4.5, and the objective function values are almost the same. By Tables 1 and 2, we obtain an

Table 6: Results of Algorithm 4.5 with $x_1^0 = (1,\ 2,\ 1,\ 0,\ 1,\ 0)$ for $(P5.1)$.

j	ϵ_j	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$	$g_4(x^*_{\epsilon_j,\rho_j})$	$g_5(x^*_{\epsilon_j,\rho_j})$
1	0.01	-0.007000	-0.002000	-0.002000	0.000036	-4.557069
2	0.001	-0.002333	-0.000667	-0.000667	0.000004	-7.245353
3	0.0001	-0.000774	-0.000224	-0.000223	-0.012054	-4.871804

approximate optimal solution is

$$x^* = (1.715453,\ 8.284547,\ 0.260178,\ 0.455275,\ 1.000000,\ 7.544725)$$

after 2 iterations with objective function value $f(x^*) = 117.000000$. In [9], the obtained approximate optimal solution is

$$x^* = (1.805996,\ 8.194004,\ 0.497669,\ 0.308327,\ 1.000000,\ 7.691673)$$

with objective function value $f(x^*) = 117.010399$. Numerical results obtained by both of our algorithms are slightly better than the results in [9].

Example 5.2. Consider the following problem ([9], Example 4.2)

$$\begin{aligned}
(P5.2):\quad \min\ & f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4\\
\text{s.t.}\ & g_1(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1 + x_2 + x_4 - 5 \le 0,\\
& g_2(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 \le 0,\\
& g_3(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \le 0.
\end{aligned}$$

For $k = \frac{2}{3}$, let $x_1^0 = (4,\ 4,\ 4,\ 4)$, $\rho_1 = 8$, $N = 6$, $\epsilon_1 = 0.1$, $\gamma = 0.05$. The results of Algorithm 4.2 for solving $(P5.2)$ are shown in Table 7.

Table 7: Results of Algorithm 4.2 with $x_1^0 = (4,\ 4,\ 4,\ 4)$ for $(P5.2)$

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$
1	8	0.1	(0.057466, 1.068184, 1.934446, -1.062742)	-43.932900	0.010075	0.002288	-0.708469
2	48	0.005	(0.169403, 0.836072, 2.008446, -0.965146)	-44.233835	-0.000000	-0.000000	-1.880657

For $k = \frac{3}{4}$, let $x_1^0 = (5,\ 5,\ 5,\ 5)$, $\rho_1 = 8$, $N = 6$, $\epsilon_1 = 0.1$, $\gamma = 0.01$. The results of Algorithm 4.2 for solving $(P5.2)$ are shown in Table 8.

Table 8: Results of Algorithm 4.2 with $x_1^0 = (5,\ 5,\ 5,\ 5)$ for $(P5.2)$.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$
1	8	0.1	(0.128519, 0.787242, 2.058425, -0.960793)	-44.497657	-0.026616	0.156996	-1.828351
2	48	0.001	(0.168480, 0.843691, 2.005069, -0.969559)	-44.233372	-0.000021	-0.000035	-1.846516

Table 9: Results of Algorithm 4.5 with $x_1^0 = (0,\ 0,\ 0,\ 0)$ for $(P5.2)$.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$
1	10	0.04	(0.135083, 0.866135, 2.016392, -0.975808)	-44.328665	0.013015	0.047622	-1.670410
2	30	0.004	(0.169560, 0.835531, 2.008636, -0.964877)	-44.233861	0.000006	0.000010	-1.883115
3	90	0.0004	(0.169560, 0.835531, 2.008634, -0.964876)	-44.233837	0.000000	0.000000	-1.883126

For $k = \frac{1}{2}$, let $x_1^0 = (0,\ 0,\ 0,\ 0)$, $\rho_1 = 10$, $N = 3$, $\epsilon_1 = 0.04$, $\gamma = 0.1$. The results of Algorithm 4.5 for solving $(P5.2)$ are shown in Table 9.

The results in Tables 7-9 show that, the convergence of both Algorithm 4.2 and Algorithm 4.5, and the objective function values are almost the same. By Table 7, we obtain an approximate optimal solution $x^* = (0.169403,\ 0.836072,\ 2.008446,\ -0.965146)$ after 2 iterations with objective function value $f(x^*) = -44.233835$. In [4], the obtained approximate optimal solution is $x^* = (0.170446,\ 0.834248,\ 2.008753,\ -0.964559)$ with function value $f(x^*) = -44.233627$. In the paper [9], the obtained approximate optimal solution is $x^* = (0.169234,\ 0.835656,\ 2.008690,\ -0.964901)$ with function value $f(x^*) = -44.233582$. Numerical results obtained by both of our algorithms are slightly better than the results in [4, 9]. Moreover, in [9] the approximate solution is found with 4 and 13 iterations in the Algorithms I and II, respectively. So, can be seen that both of our algorithms find the solutions with the lower iteration numbers than in [9].

Example 5.3. Consider the following problem ([14], Example 3.2)

$$
\begin{aligned}
(P5.3): \quad & \min \quad f(x) = -x_1 - x_2 \\
& \text{s.t.} \quad g_1(x) = -2x_1^4 + 8x_1^3 - 8x_1^2 + x_1 - 2 \le 0, \\
& \qquad\;\; g_2(x) = -4x_1^4 + 32x_1^3 - 88x_1^2 + 96x_1 + x_2 - 36 \le 0,
\end{aligned}
$$

$$0 \le x_1 \le 3,$$
$$0 \le x_2 \le 4.$$

For $k = \frac{3}{4}$, let $x_1^0 = (1,\ 3),\ \rho_1 = 6,\ N = 8,\ \epsilon_1 = 0.2,\ \gamma = 0.1$. The results of Algorithm 4.2 for solving $(P5.3)$ are shown in Table 10.

Table 10: Results of Algorithm 4.2 with $x_1^0 = (1,\ 3)$ for $(P5.3)$.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	6	0.2	(2.112082, 3.900138)	-6.012220	0.000003	0.000006
2	48	0.02	(2.112086, 3.900125)	-6.012210	-0.000001	-0.000001

For $k = \frac{2}{3}$, let $x_1^0 = (0,\ 2),\ \rho_1 = 6,\ N = 4,\ \epsilon_1 = 0.01,\ \gamma = 0.1$. The results of Algorithm 4.2 for solving $(P5.3)$ are shown in Table 11.

Table 11: Results of Algorithm 4.2 with $x_1^0 = (0,\ 2)$ for $(P5.3)$.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	6	0.01	(2.112036, 3.900287)	-6.012323	0.000054	0.000073
2	24	0.001	(2.112126, 3.899981)	-6.012108	-0.000046	-0.000072

For $k = \frac{3}{4}$, let $x_1^0 = (1,\ 1),\ \rho_1 = 10,\ N = 5,\ \epsilon_1 = 0.05,\ \gamma = 0.1$. The results of Algorithm 4.5 for solving $(P5.3)$ are shown in Table 12.

Table 12: Results of Algorithm 4.5 with $x_1^0 = (1,\ 1)$ for $(P5.3)$.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	10	0.05	(2.112084, 3.900131)	-6.012215	0.000001	0.000002
2	50	0.005	(2.112163, 3.899988)	-6.012152	-0.000087	-0.000000

The results in Tables 10-12 show that, the convergence of both Algorithm 4.2 and Algorithm 4.5, and the objective function values are almost the same. By Table 10, we obtain an approximate optimal solution $x^* = (2.112086,\ 3.900125)$ after 2 iterations with objective function value $f(x^*) = -6.012210$. In [14], the obtained global solution is $x^* = (2.3295,\ 3.1784)$ with objective function value $f(x^*) = -5.5080$. In the paper [15], the obtained approximate optimal solution is $x^* = (2.112103,\ 3.900086)$ with objective function value $f(x^*) = -6.012190$. Numerical results obtained by both of our algorithms are much better than the results in [14] and find the correct solutions as in [15]. Further, it can be seen that the approximate solutions obtained by our algorithms with the lower iteration numbers in comparison with the [14].

6. Conclusions

In this paper, two new smoothing approaches for the lower order penalty functions are proposed. Both of the perturbed smoothing approaches present lower errors among smoothed penalty problems, nonsmooth nonlinear penalty problems and original constrained optimization problems. By using these perturbed smooth penalty functions, we developed algorithms to solve nonlinear constrained optimization problems and obtained satisfactory results.

Our perturbed smoothing techniques provide good approximations to the nonsmooth function. In fact, both of perturbed smoothing techniques can be used for non-lipschitz $\max\{x, 0\}^k$, $0 < k < 1$, and nonsmooth $\max\{x, 0\}$ functions by controlling the parameter k.

According to the numerical results given in Section 5, we show that the Algorithm 4.2 and Algorithm 4.5 are effective for both medium scale and large constrained optimization problems. Moreover, these algorithms have a good convergence for a global solution or an approximate global solution of the original constrained optimization problem.

Acknowledgement

The authors would like to thank referees for their constructive comments and suggestions, which significantly improved the presentation of the paper. This work is supported by grants from the National Natural Science Foundation of China (No. 11371242 and 61572099).

References

[1] M. S. Bazaraa and J. J. Goode, *Sufficient conditions for a globally exact penalty function without convexity*, Mathematical Programming Study, *19* (1982), 1–15.

[2] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, Nonlinear Programming Theory and Algorithms, John Wiley and Sons, Inc., New York, second editon, 1993.

[3] N. T. Binh, *Smoothing approximation to l_1 exact penalty function for constrained optimization problems*, Journal of Applied Mathematics and Informatics, **33** (2015), 387–399.

[4] N. T. Binh, *Smoothed lower order penalty function for constrained optimization problems*, IAENG International Journal of Applied Mathematics, **46** (2016), 76–81.

[5] C. H. Chen and O. L. Mangasarian, *Smoothing methods for convex inequalities and linear complementarity problems*, Mathematical Programming, **71** (1995), 51–69.

[6] G. Di Pillo and L. Grippo, *An exact penalty function method with global conergence properties for nonlinear programming problems*, Mathematical Programming, **36** (1986), 1–18.

[7] S. P. Han and O. L. Mangasrian, *Exact penalty function in nonlinear programming*, Mathematical Programming, **17** (1979), 257–269.

[8] J. B. Lasserre, *A globally convergent algorithm for exact penalty functions*, European Journal of Opterational Research, **7** (1981), 389–395.

[9] Z. Q. Meng, C. Y. Dang and X. Q. Yang, *On the smoothing of the square-root exact penalty function for inequality constrained optimization*, Computational Optimization and Applications, **35** (2006), 375–398.

[10] J. Nocedal and S. T. Wright, Numerical Optimization, Springer-Verlag, New York, 1999.

[11] M. C. Pinar and S. A. Zenios, *On smoothing exact penalty function for convex constrained optimization*, SIAM Journal on Optimization, **4** (1994), 486–511.

[12] E. Rosenberg, *Exact penalty functions and stability in locally Lipschitz programming*, Mathematical Programming, **30** (1984), 340–356.

[13] Z. Y. Wu, F. S. Bai, X. Q. Yang and L. S. Zhang, *An exact lower order penalty function and its smoothing in nonlinear programming*, Optimization, **53** (2004), 51–68.

[14] Z. Y. Wu, H. W. J. Lee, F. S. Bai and L. S. Zhang, *Quadratic smoothing approximation to l_1 exact penalty function in global optimization*, Journal of Industrial and Management Optimization, **1** (2005), 533–547.

[15] T. Yang, N. T. Binh, T. M. Thang and D. T. Hoa, *A new smoothing nonlinear penalty function for constrained optimization*, Mathematical and Computational Applications, **22** (2017), 2–31.

[16] C. J. Yu, K. L. Teo, L. S. Zhang and Y. Q. Bai, *A new exact penalty function method for continuous inequality constrained optimization problems*, Journal of Industrial and Management Optimization, **6** (2010), 895–910.

[17] C. J. Yu, K. L. Teo and Y. Q. Bai, *An exact penalty function method for nonlinear mixed discrete programming problems*, Optimization Letters, **7** (2013), 23–38.

[18] C. J. Yu, B. Li, R. Loxton and K. L. Teo, *Optimal discrete-valued control computation*, Journal of Global Optimization, **56** (2013), 503–518.

[19] W. I. Zangwill, *Nonlinear programming via penalty function*, Management Science, **13** (1967), 334–358.

[20] S. A. Zenios, M. C. Pinar and R. S. Dembo, *A smooth penalty function algorithm for network-structured problems*, European Journal of Opterational Research, **83** (1995), 220–236.

Department of Mathematics, Shanghai University, Shanghai 200444, China. and YenBai Teacher's Training College, Yen Bai city, Vietnam.

E-mail: bingbongyb@gmail.com

Department of Mathematics, Shanghai University, Shanghai 200444, China.

E-mail: yqbai@shu.edu.cn

School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, China.

E-mail: shirleyyx_1989@126.com

School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China.

E-mail: touna1988@yahoo.com