Hindawi
Complexity
Volume 2024, Article ID 5720769, 11 pages
https://doi.org/10.1155/2024/5720769

Research Article

A Nonconvex Proximal Bundle Method for Nonsmooth Constrained Optimization

Jie Shen,[1] **Fang-Fang Guo**,[2] **and Na Xu**[1]

[1]*School of Mathematics, Liaoning Normal University, Dalian 116029, China*
[2]*School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China*

Correspondence should be addressed to Jie Shen; tt010725@163.com

Received 15 February 2023; Revised 16 November 2023; Accepted 17 January 2024; Published 6 February 2024

Academic Editor: Sigurdur F. Hafstein

An implementable algorithm for solving nonsmooth nonconvex constrained optimization is proposed by combining bundle ideas, proximity control, and the exact penalty function. We construct two kinds of approximations to nonconvex objective function; these two approximations correspond to the convex and concave behaviors of the objective function at the current point, which captures precisely the characteristic of the objective function. The penalty coefficients are increased only a finite number of times under the conditions of Slater constraint qualification and the boundedness of the constrained set, which limit the unnecessary penalty growth. The given algorithm converges to an approximate stationary point of the exact penalty function for constrained nonconvex optimization with weakly semismooth objective function. We also provide the results of some preliminary numerical testing to show the validity and efficiency of the proposed method.

1. Introduction and Motivation

Nonsmooth optimization problems (NSO) arise from many fields of applications in engineering [1], economics [2], mechanics [3], and optimal control [4]. For example, multiobjective nonsmooth optimization has also been applied in many fields of engineering where optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives [5]. There exist several approaches to solving NSO, see [6–10]. Bundle methods are currently among the most efficient optimization methods; they can be used to study the engineering problem of the safe evaluation technology for concrete dams by applying nonsmooth bundle ideas to hydrostructure antiseismic fields [11–14]. These methods are based on the cutting plane method [15, 16], where the convexity of the objective function is the fundamental assumption. If the objective function f is convex, the model functions are lower approximations to the objective function. This feature is crucial to prove the convergence of most bundle methods. There exist lots of bundle methods [7, 17–19] for solving convex constrained optimization problems. In [7], the author presents a version of proximal bundle method for convex constrained optimization, where l_1 and l_∞ exact penalty functions are employed and a new penalty update is used to limit unnecessary penalty growth; the global convergence of the method is established. In [17], an infeasible bundle method for convex constrained optimization is proposed, which does not use either a penalty function or a filter, and in fact the method can be viewed as an unconstrained proximal bundle method applied directly to the improvement function, and it should be noted that the serious steps need neither be monotone nor feasible. The algorithm presented in [18] inherits attractive features from the proximal bundle methods and the filter strategy, which makes the criterion for accepting a candidate point as a serious step easier to satisfy.

However, for nonconvex cases, the corresponding model function does not stay below the objective function f and may even cut off a region containing a minimizer. There are few systematic studies for extending convex bundle methods to nonconvex cases. Most authors have considered forcing linearization errors to be positive by replacing negative values with a quadratic term or with the absolute value of the linearization errors; the piecewise affine models embedding

possible downward shifting of the affine pieces are also considered [20, 21]. In [22], the author presents a substitute for the cutting plane without convexity assumption and proves that every accumulation point of the sequence of serious steps is critical. Based on cutting plane models, a local convexification model of the objective function is constructed in [23]; it opens a new way to create nonconvex algorithms; Fuduli et al. [24] partitioned the bundle information into two subsets to capture convex and concave behaviors of the objective function around the current point. Other literatures about bundle methods for nonsmooth nonconvex optimization can be found in [25–28].

In this paper, we propose a new algorithm for constrained nonconvex optimization. The algorithm is based on the construction of two kinds of approximations to the objective function, and these two kinds of approximations correspond to the convex and concave behaviors of the objective function at the current point. If the linearization error is positive, we build a local lower approximation to the objective function, otherwise a local upper approximation is constructed. Besides that, the method employs l_1 exact penalty functions with a new penalty update rule that limits unnecessary penalty growth. Our method extends the exact penalty function algorithms for constrained convex minimization to nonconvex optimization. The following is the main difference of our paper in many respects from the existing ones [6, 7, 19, 29]. The proposed algorithm in this paper is quite different from the ones in [7, 17–19] since the objective function in our paper is nonconvex although the constraint function is convex, while both the objective function and the constraint function are convex in [7, 17–19]. In [7, 19], the l_1 and l_∞ exact penalty functions are employed to consider convex constraint optimization problems by combining with proximal bundle method. In this paper, we also use the similar exact penalty functions for solving nonconvex optimization problems, but we have to adjust suitably the construction of quadratic programming subproblems since the presence of nonconvexity can make the linearization errors negative, which enhances the difficulty to solve the problem. To solve this problem, we divide the bundle index set into two sets according to the signs of linearization errors and use the partitioned bundles during the process of the construction of the objective function model. Therefore, the direction finding subproblem is quite different from the existing ones, which leads to the overall changes for the design of the algorithm. The algorithms in [17, 18] use neither penalty functions nor relatively complex filters; they build on the theory of the well-developed unconstrained bundle methods by introducing the improvement function, which is essential for the convergence of the proposed algorithm. It is another approach proposed in recent years for solving nonsmooth convex constrained optimization problems. It should be noted that the descent condition in our method used to decide when the candidate point can be accepted as the next serious step is different from the ones in [17, 18], where the improvement function involving the objective function is employed to form the descent criterion, but in our method, the penalty function is used to serve the same role.

This paper is organized as follows: in Section 2, the model for constrained nonconvex optimization is established by using the exact penalty function. The nonconvex bundle algorithm is presented in Section 3. In Section 4, we prove that the sequence of serious steps generated by the proposed algorithm converges to an approximate stationary point of exact penalty function with weakly semismooth objective functions. Preliminary numerical experiments are provided in Section 5. Finally, some conclusions are given in Section 6.

2. Derivation of the Model

Consider the following constrained nonconvex optimization problem:

$$\begin{aligned} &\min_{x\in\mathbf{R}^n} \ f(x), \\ &\text{s.t.} \quad F(x)\le 0, \end{aligned} \tag{1}$$

where $f\colon R^n \longrightarrow R$ is a real-valued locally Lipschitz function and $F\colon R^n \longrightarrow R$ is a real-valued convex function. It is well known that for a locally Lipschitz function f, the generalized subdifferential (Clarke's subdifferential) at each point x is defined by

$$\partial_C f(x) = \operatorname{conv}\left\{g \mid g\in R^n, \nabla f\left(x^k\right) \longrightarrow g, x^k \longrightarrow x, x^k \notin \Omega_f\right\}, \tag{2}$$

where "conv" denotes the convex hull of a set and Ω_f is the set where f is not differentiable. The set $\partial_C f(x)$ is locally bounded [30]. An extension of the generalized subdifferential is the Goldstein ε-subdifferential $\partial_\varepsilon^G f(x)$ defined as

$$\partial_\varepsilon^G f(x) = \operatorname{conv}\{\partial_C f(y) \mid \|y-x\|\le\varepsilon\}. \tag{3}$$

For convex function F, the subdifferential of F at x is defined by

$$\partial F(x) = \{g \mid g\in R^n, F(y)\ge F(x) + \langle g, y-x\rangle, \forall y\in R^n\}, \tag{4}$$

which is locally bounded. Assume that we are able to compute at each point x both the function values $f(x), F(x)$ and subgradients $g_f(x)\in\partial_C f(x), g_F(x)\in\partial F(x)$. We denote the current iteration point (stability center) and the trial point by x^i and y^j, respectively. The bundles of available information are the sets B_{fi} and B_{Fi} of elements

$$\left(y^j, f\left(y^j\right), g_f^j, \alpha_f^j, a_f^j\right), \quad j \in B_{fi}, \left(y^j, F\left(y^j\right), g_F^j, \alpha_F^j\right), \quad j \in B_{Fi}, \tag{5}$$

where $g_f^j \in \partial_C f(y^j)$, $g_F^j \in \partial F(y^j)$, $\alpha_f^j = f(x^i) - f(y^j) - \langle g_f^j, x^i - y^j\rangle$, $a_f^j = \|x^i - y^j\|$, $\alpha_F^j = F(x^i)_+ - F(y^j) - \langle g_F^j, x^i - y^j\rangle$, and $F(x^i)_+ = \max\{F(x^i), 0\}$.

If f and F are real-valued convex functions on R^n, under Slater constraint qualification, problem (1) can be solved by minimizing the following l_1 exact penalty function

$$e(x, c) = f(x) + cF(x)_+, \tag{6}$$

where c is the penalty parameter which is greater than the Lagrange multiplier of problem (1), see [7]. For real-valued locally Lipschitz function $f: R^n \longrightarrow R$ and real-valued convex function $F: R^n \longrightarrow R$, we try to combine the ideas of bundle methods, proximity control, and exact penalty functions to solve problem (1). We define the cutting plane approximations of f, F, and $e(x, c)$ by their linearizations:

$$\begin{aligned} \hat{f}^i(x) &= \max_{j\in B_{fi}}\left\{f\left(y^j\right) + \langle g_f^j, x - y^j\rangle\right\}, \\ \hat{F}^i(x) &= \max_{j\in B_{Fi}}\left\{F\left(y^j\right) + \langle g_F^j, x - y^j\rangle\right\}, \\ \hat{e}^i(x, c_i) &= \hat{f}^i(x) + c_i\widehat{F}^i(x)_+, \end{aligned} \tag{7}$$

where $B_{fi}, B_{Fi} \subset \{1, 2, \ldots, i\}$, c_i is the corresponding penalty parameter. The next trial point y^{j+1} is obtained by solving the following problem:

$$\min_{x\in R^n}\left\{\hat{e}^i(x, c_i) + \frac{u_i}{2}\left\|x - x^i\right\|^2\right\}, \tag{8}$$

where $u_i > 0$ is the proximal parameter. Note that if x is feasible, $F(x)_+ = 0$. Therefore, we introduce additionally a trial point y^0, the subgradient $g_F^0 = 0$, and the linearization error $\alpha_F^0 = F(x^i)_+ (=0)$ with respect to y^0, and then we define $B_{Fi0} = B_{Fi} \cup \{0\}$. Hence, problem (8) can be written in equivalent form

$$\begin{cases} \min v_f + c_i v_F + \dfrac{u_i}{2}\|d\|^2, \\ \text{s.t.} \quad -\alpha_f^j + \left(g_f^j\right)^\top d \le v_f, \quad j \in B_{fi}, \\ -\alpha_F^j + \left(g_F^j\right)^\top d \le v_F, \quad j \in B_{Fi0}, \end{cases} \tag{9}$$

where $d = x - x^i$. The introduction of index set $\{0\}$ into $B_{F_{i0}}$ can make sure $v_F \ge 0$ in (9) for $j = 0$.

It should be noted that α_f^j may be negative since f is nonconvex; we divide B_{fi} into two sets B_{fi}^+ and B_{fi}^- defined by

$$B_{fi}^+ = \left\{j \in B_{fi} \mid \alpha_f^j \ge 0\right\}, B_{fi}^- = \left\{j \in B_{fi} \mid \alpha_f^j < 0\right\}, \tag{10}$$

where B_{fi}^+ is nonempty since $i \in B_{fi}^+$. We define the following two piecewise affine functions:

$$H^+(d) = \max_{j\in B_{fi}^+}\left\{\left(g_f^j\right)^\top d - \alpha_f^j\right\}, H^-(d) = \min_{j\in B_{fi}^-}\left\{\left(g_f^j\right)^\top d - \alpha_f^j\right\}. \tag{11}$$

Let $h(d) = f(x^i + d) - f(x^i)$, the affine function $H^+(d)$, be considered as the approximation of $h(d)$ since $h(d) = f(x^i + d) - f(x^i) = f(x) - f(x^i)$ and $H^+(d) = \max_{j\in B_{fi}^+}\left\{(g_f^j)^\top d - \alpha_f^j\right\} = \max_{j\in B_{fi}^+}\left\{f(y^j) + (g_f^j)^\top(x - y^j)\right\} - f(x^i)$. Because $H^+(0) < 0$, $H^-(0) > 0$ if $B_{fi}^- \ne \varnothing$, therefore $H^+(0) < H^-(0)$. Summing up, around $d = 0$ (around the stability center x^i), it appears that the set $S_f = \{d \in R^n | H^+(d) \le H^-(d)\}$ is more important and reliable.

Let $(v_{f_{u_i}}, v_{F_{u_i}}, d_{u_i})$ be the optimal solution to the following problem:

$$QP_{(u_i)}\begin{cases} z_{u_i} = \min v_f + c_i v_F + \dfrac{u_i}{2}\|d\|^2, \\ \text{s.t.} \quad -\alpha_f^j + \left(g_f^j\right)^\top d \le v_f, \quad j \in B_{fi}^+, \\ -\alpha_f^j + \left(g_f^j\right)^\top d \ge v_f, \quad j \in B_{fi}^-, \\ -\alpha_F^j + \left(g_F^j\right)^\top d \le v_F, \quad j \in B_{Fi0}. \end{cases} \tag{12}$$

Let $y^{j+1} = x^i + d_{u_i}$. Since v_{Fu_i} is the optimal solution of problem (12), $v_{Fu_i} = \widehat{F}^i(y^{j+1})_+ - F(x^i)_+$, similarly, $v_{fu_i} = \hat{f}^i(y^{j+1}) - f(x^i)$. We define the predicted descent $v_{u_i} = \hat{e}^i(y^{j+1}, c_i) - e(x^i, c_i)$, it is not difficult to find that $v_{u_i} = v_{fu_i} + c_i v_{Fu_i}$. We notice that $z_{u_i} \le 0$ (therefore, $v_{u_i} \le 0$) since $(v_f, v_F, d) = (0,0,0)$ is a feasible point of problem (12). Set

$$\begin{aligned} L\left(v_f, v_F, d, \lambda, \mu, \gamma\right) = {} & v_f + c_i v_F + \frac{u_i}{2}\|d\|^2 \\ & - \sum_{j\in B_{fi}^+}\lambda_j\left(v_f - \left(g_f^j\right)^\top d + \alpha_f^j\right) \\ & - \sum_{j\in B_{fi}^-}\mu_j\left(\left(g_f^j\right)^\top d - \alpha_f^j - v_f\right) \\ & - \sum_{j\in B_{Fi0}}\gamma_j\left(v_F - \left(g_F^j\right)^\top d + \alpha_F^j\right). \end{aligned} \tag{13}$$

Let $\nabla_d L = 0$; we obtain

$$d = \frac{1}{u_i}\left[-G_+^f\lambda + G_-^f\mu - G^F\gamma\right], \tag{14}$$

where G_+^f, G_-^f and G^F are matrices whose columns are the vectors $g_f^j, j \in B_{fi}^+, g_f^j, j \in B_{fi}^-$ and $g_F^j, j \in B_{Fi0}$. λ, μ and γ are the vectors with components $\lambda_j, j \in B_{fi}^+, \mu_j, j \in B_{fi}^-$ and $\gamma_j, j \in B_{Fi0}$, respectively. Let $\nabla_{v_f} L = 0$; we obtain

$$\sum_{j\in B_{fi}^+} \lambda_j - \sum_{j\in B_{fi}^-} \mu_j = 1, \tag{15}$$

i.e., $\tilde{e}^\top\lambda - \tilde{e}^\top\mu = 1$, where $\tilde{e} = (1,1,\ldots,1)^T$. Let $\nabla_{v_F} L = 0$; we obtain

$$c_i = \tilde{c}_i + \gamma_0, \tilde{c}_i = \sum_{j\in B_{Fi}} \gamma_j \geq 0. \tag{16}$$

Substituting (14)–(16) into L, we have

$$L\left(v_f, v_F, d, \lambda, \mu, \gamma\right) = -\frac{1}{2u_i}\left\|G_+^f\lambda - G_-^f\mu + G^F\gamma\right\|^2 - \alpha_{f+}^\top\lambda + \alpha_{f-}^\top\mu - \alpha_F^\top\gamma, \tag{17}$$

where α_{f+}, α_{f-} and α_F are vectors whose components are $\alpha_f^j, j \in B_{fi}^+; \alpha_f^j, j \in B_{fi}^-$ and $\alpha_F^j, j \in B_{Fi0}$, respectively. Then, the duality problem of $QP_{(u_i)}$ is the following minimization problem:

$$DP_{(u_i)}\begin{cases} \min \dfrac{1}{2u_i}\left\|G_+^f\lambda - G_-^f\mu + G^F\gamma\right\|^2 + \alpha_{f+}^\top\lambda - \alpha_{f-}^\top\mu + \alpha_F^\top\gamma, \\ \text{s.t.} \quad \lambda, \mu, \gamma \geq 0, \\ \tilde{e}^\top\lambda - \tilde{e}^\top\mu = 1, \end{cases} \tag{18}$$

and the primal optimal solution $(v_{f_{u_i}}, v_{F_{u_i}}, d_{u_i})$ is related to the dual optimal solution (λ, μ, γ) by the following formulae:

$$d_{u_i} = \frac{1}{u_i}\left[-G_+^f\lambda + G_-^f\mu - G^F\gamma\right], \tag{19}$$

$$v_{u_i} = v_{fu_i} + c_i v_{Fu_i} = -u_i\left\|d_{u_i}\right\|^2 - \alpha_{f+}^\top\lambda_{u_i} + \alpha_{f-}^\top\mu_{u_i} - \alpha_F^\top\gamma_{u_i}. \tag{20}$$

3. Algorithm

In this section, we present a nonconvex bundle algorithm for constrained nonconvex optimization problem (1). Our method is based on repeatedly solving problem (12).

A few comments on Algorithm 1 are in order.

The stopping criterion in Step 1 is used to assess the stationarity of current stability center. If it is satisfied, the approximate stationarity of exact penalty function is achieved, Algorithm 1 stops, and the approximate solution is obtained.

The solution $(d_{\hat{u}}^j, v_{f\hat{u}}^j, v_{F\hat{u}}^j)$ of $QP_{(\hat{u})}$ at Step 2 may be obtained by using the dual quadratic programming method of [29] or [31], which can efficiently solve sequences of subproblems $QP_{(u_i)}$ with varying u_i and c_i.

The result $d_{\hat{u}}^j \leq \theta$ is never a consequence of the choice of too big $\hat{u}$. In fact, we note that if $\|g_f^i + c_i g_F^i\| > \delta$, it holds that

$$\left\|d_{u_{\max}}\right\| \leq \frac{2}{u_{\max}}\left\|g_f^i + c_i g_F^i\right\| = \frac{2\theta\left\|g_f^i + c_i g_F^i\right\|}{r\delta}. \tag{21}$$

The right-hand side of the above inequality is more than θ, so too big $\hat{u}$ may not lead to $d_{\hat{u}}^j \leq \theta$, therefore we intend to decrease the value of u in Step 2.

Large c_i may force the iterate points generated by Algorithm 1 to approach closely to the boundary of the feasible set $S_F = \{x \in R^n | F(x) \leq 0\}$ and may damage the fast convergence of Algorithm 1. Our rule increases c_i (from c_{i-1}) only to ensure a significant predicted decrease in constraint violation of the form $\widehat{F}^i(y^{i+1}) \leq \kappa_i F(x^i)_+$, where $\kappa_i \in [0,1)$ is a contraction factor, see Step 3.

Finally, note that the insertion of a bundle index into B_{fi}^+ or B_{fi}^- at Step 6 is not simply based on the sign of α_f^j, see [15].

4. Convergence Results

The presented work in this section follows a line of investigation initiated in [6, 7], where nonconvex bundle algorithm is used to solve unconstrained minimization problem

Step 0 (Initialization):

Choose x^1 such that $F(x^1) \le 0$. Choose the stationarity tolerance $\delta > 0$, the proximity measure $\varepsilon \ge 0$, the improvement parameter $m \in (0,1)$, the cut parameter $\rho \in (m,1)$, the reduction parameter $r \in (0,1)$, the increase parameter $R > 1$, the infeasibility contraction bound $k_{\max} \in [0,1)$, the initial penalty coefficient $c_1 > 0$, and the maximal number of stored subgradients $M \ge n+2$. Set $y^1 = x^1, B_{f1}^+ = \{1\}, B_{f1}^- = \varnothing, B_{F1} = \{1\}$. Set the outer iteration counter $i = 1$, the inner iteration counter $j = 1$.

Step 1 (Safeguard parameters setting):

If $\|g_f^i + c_i g_F^i\| \le \delta$, terminate; otherwise set $u_{\max} = (\|g_f^i + c_i g_F^i\|/r\varepsilon)$, $u_{\min} = (u_{\max}/R)$, $\theta = (r\delta/u_{\max})$.

Step 2 (Direction finding):

Solve $QP_{(u_i)}$ repeatedly by choosing decreasing value of $u_i \in [u_{\min}, u_{\max}]$ (for the first time choose $u_i = u_{\max}$) and find the solution $(v_{f_{u_i}}^j, v_{F_{u_i}}^j, d_{u_i}^j)$ of problem (12) for $u = u_i$ until $e(x^i + d_{u_i}^j, c_i) > e(x^i, c_i) + m v_{u_i}^j$,

where $v_{u_i}^j = v_{f_{u_i}}^j + c_i v_{F_{u_i}}^j$. If such u_i does exist, let $\hat{u}$ equals the maximum value of $u_i \in [u_{\min}, u_{\max}]$; otherwise set $\hat{u} = u_{\min}$. Denote the optimal solution of $QP_{(\hat{u})}$ by $(v_{f\hat{u}}^j, v_{F\hat{u}}^j, d_{\hat{u}}^j)$, and set $y^{j+1} = x^i + d_{\hat{u}}^j$. If $\|d_{\hat{u}}^j\| > \theta$, go to Step 5.

Step 3 (Penalty updating):

If $\widehat{F}^i(y^{j+1}) > \kappa_i F(x^i)_+$, choose $\kappa_i \in [0, \kappa_{\max}]$ and replace c_i by $2c_i$, go to Step 2.

Step 4 (Stationarity test):

Set

$$B_{fi}^+ = B_{fi}^+ \backslash \left\{ j \in B_{fi}^+ | a_f^j > \varepsilon \right\}, B_{fi}^- = B_{fi}^- \backslash \left\{ j \in B_{fi}^- | a_f^j > \varepsilon \right\}.$$

Calculate

$$g^* = \min\left\{ \|g\| | g \in \operatorname{conv}\left\{ g_f^j + c_i g_F^j \right\}, g_f^j, j \in B_{fi}^+; g_F^j, j \in B_{Fi} \right\}$$

If $\|g^*\| \le \delta$, terminate; else set $u_{\min} := u_{\min} u_{\max} / ((1-r)u_{\max} + r u_{\min})$, go to Step 2.

Step 5 (Trial point calculating):

Compute $g_f^{j+1} \in \partial_C f(y^{j+1})$, $g_F^{j+1} \in \partial F(y^{j+1})$ and set $\alpha_f^{j+1} = f(x^i) - f(y^{j+1}) + (g_f^{j+1})^T d_{\hat{u}}^j$, $\alpha_F^{j+1} = F(x^i)_+ - F(y^{j+1}) + (g_F^{j+1})^T d_{\hat{u}}^j$.

Step 6 (Insertion of index):

(a) If $\alpha_f^{j+1} < 0$ and $\|d_{\hat{u}}^j\| > \varepsilon$, insert the element $(y^{j+1}, f(y^{j+1}), g_f^{j+1}, \alpha_f^{j+1}, \|d_{\hat{u}}^j\|)$ into the bundle for an appropriate value of $j \in B_{fi}^-$ and set $\hat{u} := \hat{u} u_{\max} / ((1-r)u_{\max} + r\hat{u})$.

(b) Else, if $(g_f^{j+1} + c_i g_F^{j+1})^T d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$, insert the element $(y^{j+1}, f(y^{j+1}), g_f^{j+1}, \max\left\{0, \alpha_f^{j+1}\right\}, \|d_{\hat{u}}^j\|)$ into the bundle for an appropriate value of $j \in B_{fi}^+$.

(c) Else find a scalar $t \in (0,1)$ such that $(g_1(t) + c_i g_2(t)) \in \partial e(x^i + t d_{\hat{u}}^j, c_i)$ satisfies the condition $(g_1(t) + c_i g_2(t))^T d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$, where $g_1(t) \in \partial_C f(x^i + t d_{\hat{u}}^j)$, $g_2(t) \in \partial F(x^i + t d_{\hat{u}}^j)$, and insert the element $(x^i + t d_{\hat{u}}^j, f(x^i + t d_{\hat{u}}^j), g_1(t), \max\{0, \alpha_t\}, t\|d_{\hat{u}}^j\|)$ into the bundle for an appropriate value of $j \in B_{fi}^+$, where $\alpha_t = f(x^i) - f(x^i + t d_{\hat{u}}^j) + t g_1(t)^T d_{\hat{u}}^j$.

(d) Insert the element $(y^{j+1}, F(y^{j+1}), g_F^{j+1}, \alpha_F^{j+1})$ into the bundle B_{Fi}.

Step 7 (Descent test):

If

$$e(x^i + d_{\hat{u}}^j, c_i) \le e(x^i, c_i) + m v_{\hat{u}}^j,$$

set $u_i = \hat{u}$, $d_{u_i}^i = d_{\hat{u}}^j$, the new stability center $x^{i+1} = x^i + d_{u_i}^i$ (serious step); otherwise set $x^{i+1} = x^i$ (null step).

Step 8 (Bundle updating):

Select sets $\widetilde{B}_{fi}, \widetilde{B}_{Fi}$ such that $\widehat{B}_{fi} \subset \widetilde{B}_{fi} \subset B_{fi}$, $\widehat{B}_{Fi} \subset \widetilde{B}_{Fi} \subset B_{Fi}$ and $|\widetilde{B}_{fi}| + |\widetilde{B}_{Fi}| \le M - 2$, where $\widehat{B}_{fi} \subset B_{fi}$, $\widehat{B}_{Fi} \subset B_{Fi}$ corresponding to nonzero multipliers satisfy $|\widehat{B}_{fi}| + |\widehat{B}_{Fi}| \le n+2$. Set $B_{f(i+1)} = \widetilde{B}_{fi} \cup \{i+1\}$, $B_{F(i+1)} = \widetilde{B}_{Fi} \cup \{i+1\}$. Set $c_{i+1} = c_i$. If a serious step is taken, increase i by 1 and go to Step 1, otherwise increase j by 1 and go to Step 2.

End of Algorithm 1

Algorithm 1: A proximal bundle algorithm for nonconvex functions.

and the idea of exact penalty function is employed in proximal bundle method for constrained convex minimization problem. Here, we expand and generalize the central idea [6] to constrained nonconvex minimization problems; some techniques have to be adjusted to the new situations for the presence of constraints and nonconvexity.

Throughout the section, we make the following assumptions:

(A1) f and F are weakly semismooth (f is said to be weakly semismooth if the directional derivative $f'(x;d) = \lim_{t\downarrow 0} t^{-1}[f(x+td) - f(x)]$ exists for all x and d, and $f'(x;d) = \lim_{t\downarrow 0} g(x+td)^T d$ where $g(x+td) \in \partial f(x+td)$).

(A2) The set $S_1 = \{x \in R^n | f(x) \le f(x^1)\}$ is compact, where x^1 is the initial point provided by the user in Algorithm 1.

(A3) The feasible set $S_F = \{x \in R^n | F(x) \le 0\}$ is bounded.

(A4) The Slater constraint qualification holds, i.e., there exists $\tilde{x} \in R^n$ such that $F(\tilde{x}) < 0$.

The assumption that the feasible set of problem (1) is bounded is usual and reasonable; it was also assumed in [7, 32–34]. In [27], the boundedness of the feasible set was assumed in order to guarantee the existence of the supremum of the range of a set-valued mapping on the feasible set. In [32], the authors assumed the feasible sets were bounded closed convex for finding the saddle point of the objective function.

Lemma 1. *Let $(d_{\hat{u}}^j, v_{\hat{u}}^j)$ be the sequence generated within an inner iteration such that $\|d_{\hat{u}}^j\| > \theta$ and*

$$e\left(x^i + d_{\hat{u}}^j, c_i\right) > e\left(x^i, c_i\right) + m v_{\hat{u}}^j, \tag{22}$$

with Algorithm 1 looping between Step 2 and Step 8. Then, the following conclusions hold:

(i) There is an index $\hat{j}$ such that for each $j \ge \hat{j}$, every new bundle index with respect to f is inserted into B_{fi}^+ and $\hat{u}$ remains unchanged.

(ii) Step 6(c) is appropriate, feasible, and not difficult to realize.

(iii) Whenever a new bundle index is inserted into B_{fi}^+, the condition $(g_f^{j+1} + c_i g_F^{j+1})^\top d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$ holds, where g_f^{j+1}, g_F^{j+1} are the subgradients of f and F at y^{j+1}, respectively.

Proof

(i) Since $\hat{u}$ increases at Step 6(a) of Algorithm 1, the situation that infinite bundle indices are inserted into B_{fi}^- can not happen. Hence, once $\hat{u}$ exceeds $((\varepsilon/2\|g_f^i + c_i g_F^i)\|)^{-1}$, no bundle index with respect to f can be inserted into B_{fi}^+.

(ii) According to Assumption (A1), $e(x,c)$ is weakly semismooth, the directional derivative $e'(x^i + t_i d_{\hat{u}}^j, c_i; d_{\hat{u}}^j)$ exists for any $t_i > 0$. It follows from the mean value theorem that $e(x^i + d_{\hat{u}}^j, c_i) - e(x^i, c_i) = c$ for some $c \in (e'_{\inf}, e'_{\sup})$, where

$$e'_{\inf} = \inf_{0 \le t_i \le 1} e'\left(x^i + t_i d_{\hat{u}}^j, c_i; d_{\hat{u}}^j\right), e'_{\sup} = \sup_{0 \le t_i \le 1} e'\left(x^i + t_i d_{\hat{u}}^j, c_i; d_{\hat{u}}^j\right). \tag{23}$$

Since the sufficient decrease condition Algorithm 1 is not satisfied, we have

$$\rho v_{\hat{u}}^j < m v_{\hat{u}}^j < e\left(x^i + d_{\hat{u}}^j, c_i\right) - e\left(x^i, c_i\right), \tag{24}$$

there exists a scalar $\overline{t_i} \in (0,1)$ such that $\rho v_{\hat{u}}^j < e'(x^i + \overline{t_i} d_{\hat{u}}^j, c_i; d_{\hat{u}}^j)$. By weak semismoothness of $e(x,c)$, it is not difficult to find a scalar $t \in (0,1)$ such that $(g_1(t) + c_i g_2(t)) \in \partial e(x^i + t d_{\hat{u}}^j, c_i)$ satisfies the condition $(g_1(t) + c_i g_2(t))^T d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$, where $g_1(t) \in \partial_C f(x^i + t d_{\hat{u}}^j)$, $g_2(t) \in \partial F(x^i + t d_{\hat{u}}^j)$.

(iii) By construction of Algorithm 1, we have $(g_f^{j+1} + c_i g_F^{j+1})^\top d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$. If $\alpha_f^{j+1} - c_i \alpha_F^{j+1} \ge 0$ (the next Lemma 4 shows that c_i can not be increased for infinitely many times, therefore, $\alpha_f^{j+1} - c_i \alpha_F^{j+1} \ge 0$ is possible), we also have

$$\left(g_f^{j+1} + c_i g_F^{j+1}\right)^\top d_{\hat{u}}^j \ge \left(g_f^{j+1} + c_i g_F^{j+1}\right)^\top d_{\hat{u}}^j - \alpha_f^{j+1} + c_i \alpha_F^{j+1} = e\left(x^i + d_{\hat{u}}^j, c_i\right) - e\left(x^i, c_i\right) > m v_{\hat{u}}^j > \rho v_{\hat{u}}^j, \tag{25}$$

the condition $(g_f^{j+1} + c_i g_F^{j+1})^\top d_{\hat{u}}^j \ge \rho v_{\hat{u}}^j$ also holds.

The next lemma shows the finite termination of the inner iteration. □

Lemma 2. *The inner iteration terminates after a finite number of steps.*

Proof. It is enough to demonstrate that, in a finite number of steps, either the condition of the stop at Step 1 or the exit at Step 4 is satisfied. Firstly, we prove Algorithm 1 cannot pass through Step 4 infinitely many times. Assume that such a case occurs, since at each iteration, the algorithm enters Step 4, then we have $\|d_{\hat{u}}^j\| \le \theta$ and $\|g^*\| > \delta$. Observe that $\hat{u} \ge u_{\min}$ and $u_{\min}$ will exceed the threshold $(2\|g_f^i + c_i g_F^i\|/\varepsilon)$

in a finite number of steps. It follows that $\|d^j_{u_i}\| \le (2/u_i)\|g^j_f + c_i g f^j_F\|$, therefore we obtain $\|d^j_{\hat{u}}\| \le \varepsilon$, which means that the indices of the new bundle elements are inserted into B^+_{fi} and are never removed.

According to Step 6, we insert an index into B^-_{fi} only if $\|d^j_{\hat{u}}\| > \varepsilon$, which implies that whenever entering Step 4, all the elements in $i \in B^-_{fi}$ are removed. Taking into account (15), (16), (19), and (20), there is an index $\tilde{j}$ such that for all $j \ge \tilde{j}$,

$$d^j_{\hat{u}} = -\frac{1}{\hat{u}}\left[G^f_+\lambda + G^F\gamma\right] = -\frac{1}{\hat{u}}\left[g^i_{cf} + c_i g^i_{cF}\right], \tag{26}$$

where $g^i_{cf} = \mathrm{conv}\left\{g^j_f | j \in B^+_{fi}\right\}$, $g^i_{cF} = \mathrm{conv}\left\{g^j_F | j \in B_{Fi}\right\}$. But since $\|d^j_{\hat{u}}\| \le \theta$ and $\|g^*\| > \delta$, we have

$$\theta \ge \left\|d^j_{\hat{u}}\right\| = \frac{1}{\hat{u}}\left\|g^i_{cf} + c_i g^i_{cF}\right\| \ge \frac{1}{u_{\max}}\left\|g^*\right\| > \frac{\theta}{r\delta}\delta > \theta, \tag{27}$$

which leads to a contradiction.

Next, we show that it is impossible to have $d^j_{\hat{u}} > \theta$ for infinitely many times and the descent condition Algorithm 1 is not satisfied with the algorithm looping between Step 4 and Step 8. Indexing by $j \in B_{Fi} \cup B_{fi}$, the jth passage through such a loop, we observe that, by Lemma 1(i), there exists an index $\bar{j}$ such that for every $j \ge \bar{j}$, the index of each new bundle element is put into B^+_{fi} with $\hat{u}$ remaining unchanged. Therefore, for $j \ge \bar{j}$, the sequence $\left\{z^j_{\hat{u}}\right\}$ is nondecreasing and bounded and hence convergent. Since $\left\{d^j_{\hat{u}}\right\}$ is bounded, suppose $\left\{d^j_{\hat{u}}\right\}_{j \in B_{Fi} \cup B_{fi}}$ is its convergent subsequence. The sequence $\left\{v^j_{\hat{u}}\right\}_{j \in B_{Fi} \cup B_{fi}}$ also converges to a nonpositive limit $\bar{v}$. Now assume that $\bar{v} < 0$. Let s and t be two successive indices in $B_{Fi} \cup B_{fi}$ and let $\beta^s_f = \max\left\{0, \alpha^s_f\right\}$ with $\alpha^s_f = e\ (x^i, c_i) - e(x^i + d^s_{\hat{u}}, c_i) + (g^s_f + c_i g^s_F)^T d^s_{\hat{u}}$ and $g^s_f \in \partial_C f\quad (x^i + d^s_{\hat{u}})$, $g^s_F \in \partial F(x^i + d^s_{\hat{u}})$; we have

$$v^t_{\hat{u}} \ge \left(g^s_f + c_i g^s_F\right)^\top d^t_{\hat{u}} - \beta^s_f, \tag{28}$$

$e(x^i + d^s_{\hat{u}}, c_i) - e(x^i, c_i) > m v^s_{\hat{u}}$, and $(g^s_f + c_i g^s_F)^T d^s_{\hat{u}} \ge \rho v^s_{\hat{u}}$. If $\beta^s_f = 0$; we have

$$\left(g^s_f + c_i g^s_F\right)^\top d^t_{\hat{u}} - \beta^s_f \ge \rho v^s_{\hat{u}}, \tag{29}$$

if $\beta^s_f = \alpha^s_f$, it holds that

$$\left(g^s_f + c_i g^s_F\right)^\top d^s_{\hat{u}} - \beta^s_f = e\left(x^i + d^s_{\hat{u}}, c_i\right) - e\left(x^i, c_i\right) > m v^s_{\hat{u}} > \rho v^s_{\hat{u}}. \tag{30}$$

Combing (28) and (29), we obtain $v^t_{\hat{u}} - \rho v^s_{\hat{u}} \ge (g^s_f + c_i g^s_F)^\top (d^t_{\hat{u}} - d^s_{\hat{u}})$, hence by taking the limits $(1-\rho)\bar{v} \ge 0$, it contradicts $\bar{v} < 0$, hence $\bar{v} = 0$. It follows from $z^j_{\hat{u}} \le 0$ that $|v^j_{\hat{u}}| \ge (\hat{u}/2)\|d^j_{\hat{u}}\|^2$, which contradicts the fact that $\|d^j_{\hat{u}}\| > \theta$.

The next lemma shows that the penalty coefficients are increased finitely many times under the conditions of Slater constraint qualification and the boundedness of S_F. □

Lemma 3. *There exists $\bar{c} < \infty$ such that $\hat{F}^i(y^{j+1}) \le 0$ if $c_i \ge \bar{c}$.*

Proof. Denote the Lagrangian of $QP_{(u_i)}$ by

$$\begin{aligned} L\left(v_f, v_F, d, \lambda, \mu, \gamma\right) = v_f + c_i v_F + \frac{u_i}{2}\|d\|^2 - \sum_{j \in B^+_{fi}} \lambda_j\left(v_f - \left(g^j_f\right)^\top d + \alpha^j_f\right) \\ - \sum_{j \in B^-_{fi}} \mu_j\left(\left(g^j_f\right)^\top d - \alpha^j_f - v_f\right) - \sum_{j \in B_{Fi0}} \gamma_j\left(v_F - \left(g^j_F\right)^\top d + \alpha^j_F\right). \end{aligned} \tag{31}$$

The Lagrange multipliers satisfy the usual saddle-point condition:

$$L\left(v_{fu_i}, v_{Fu_i}, d_{u_i}, \lambda_i, \mu_i, \gamma_i\right) \le L\left(v_f, v_F, d, \lambda_i, \mu_i, \gamma_i\right), \tag{32}$$

for all d, v_f, v_F. For the above inequality, we take $d = \tilde{d}_{u_i} = \tilde{x} - x^i$, where $\tilde{x}$ is the one in Assumption (A4), and then by using (15) and (16), we obtain

$$\begin{aligned} \frac{u_i}{2}\left\|d_{u_i}\right\|^2 + \left(p^i_f\right)^T d_{u_i} - \sum_{j \in B^+_{f_i}} \lambda_j \alpha^j_f + \sum_{j \in B^-_{f_i}} \mu_j \alpha^j_f + c_i\left(\left(g^j_F\right)^T d_{u_i} - \alpha^j_F\right) \\ \le \frac{u_i}{2}\left\|\tilde{d}_{u_i}\right\|^2 + \left(p^i_f\right)^T \tilde{d}_{u_i} - \sum_{j \in B^+_{f_i}} \lambda_j \alpha^j_f + \sum_{j \in B^-_{f_i}} \mu_j \alpha^j_f + c_i\left(\left(g^j_F\right)^T \tilde{d}_{u_i} - \alpha^j_F\right), \end{aligned} \tag{33}$$

where $p_f^i = \sum_{j\in B_{fi}^+}\lambda_j g_f^j - \sum_{j\in B_{fi}^-}\mu_j g_f^j$. Remove the same terms from both sides of (33) and note that $(g_F^j)^T d_{u_i} - \alpha_F^j = v_{F_{u_i}} = \widehat{F}^i(y^{j+1})_+ - F(x^i)_+$, $(g_F^j)^T\widetilde{d}_{u_i} - \alpha_F^j = (g_F^j)^T\widetilde{d}_{u_i} - F(x^i)_+ + F(y^j) + \langle g_F^j, x^i - y^j\rangle$, and the subgradient inequality of convex function $F(y^j) + \langle g_F^j, x^i - y^j\rangle + \langle g_F^j, \widetilde{x} - x^i\rangle \le F(\widetilde{x})$; we have

$$\frac{1}{2}u_i\left\|d_{u_i}\right\|^2 + \left(p_f^i\right)^T d_{u_i} + c_i\widehat{F}^i\left(y^{j+1}\right)_+ \le \frac{1}{2}u_i\left\|\widetilde{d}_{u_i}\right\|^2 + \left(p_f^i\right)^T\widetilde{d}_{u_i} + c_iF(\widetilde{x}). \tag{34}$$

Since $c_i = \widetilde{c}_i + \gamma_0$, $\gamma_0 \ge 0$, and Assumption (A4) tells us that $F(\widetilde{x}) < 0$, the following inequality holds:

$$\frac{1}{2}u_i\left\|d_{u_i}\right\|^2 + \left(p_f^i\right)^T d_{u_i} + c_i\widehat{F}^i\left(y^{j+1}\right)_+ \le \frac{1}{2}u_i\left\|\widetilde{d}_{u_i}\right\|^2 + \left(p_f^i\right)^T\widetilde{d}_{u_i} + \widetilde{c}_iF(\widetilde{x}). \tag{35}$$

Recalling that g_f^j is locally bounded, let $D = \sup\{\|x - y\| | x, y \in S_F\}$ and $C_g = \sup\left\{\|g_f^j\| | j \in I_{fi}\right\}$. The existence of C_g can be guaranteed by the aggregate technique of bundle methods. Then,

$$\widetilde{c}_i(-F(\widetilde{x})) + c_i\widehat{F}^i\left(y^{j+1}\right)_+ \le \frac{1}{2}\bar{u}D^2 + DC_g =: L, \tag{36}$$

where $\bar{u}$ is the upper bound of u_i. Now, if we take $\bar{c} = (-L)/(F(\widetilde{x}))$, the conclusion can be obtained. □

Lemma 4. *There exist an index $\bar{i}$ and $\bar{c} > 0$ such that $c_i = \bar{c}$ for all $i \ge \bar{i}$.*

Proof. According to the result of Lemma 3, there exists one $\bar{c} > 0$ such that $\widehat{F}^i(y^{j+1}) > \kappa_i F(x^i)$ does not hold for $c_i \ge \bar{c}$. Therefore, Step 3 in Algorithm 1 will not be executed once $c_i \ge \bar{c}$. The penalty coefficient c_i remains constant $\bar{c}$.

Note that Lemmas 3 and 4 ensure the number of iterations between Step 2 and Step 3 is finite, and the penalty coefficients stay unchanged after finitely many iterations.

Now, we are ready to prove the overall finiteness of Algorithm 1. □

Theorem 5. *Suppose Assumptions (A1)–(A4) hold, then for any $\varepsilon > 0$ and $\delta > 0$, Algorithm 1 stops at a point satisfying the approximate stationarity condition*

$$\left\|g^*\right\| \le \delta, \tag{37}$$

with $g^* \in \partial_\varepsilon^G f(x^i) + c_i\partial_\varepsilon^G F(x^i)$.

Proof. For contradiction, assume that the approximate stationarity condition (37) cannot be satisfied for an infinite number of iterations. In other words, the termination condition $\|g^*\| \le \delta$ in Step 4 is not satisfied for each iteration. Therefore, Algorithm 1 is executed for infinitely many times. It follows from Lemma 2 that the descent condition Algorithm 1 is satisfied for each iteration. Let x^i be the stability center at the jth iteration through inner iteration, then $\|d_{\hat{u}}^j\| > \theta$ and $e(x^{i+1}, c_{i+1}) \le e(x^i, c_i) + mv_{\hat{u}}^j$, hence

$$e\left(x^{i+1}, c_{i+1}\right) - e\left(x^1, c_1\right) \le m\sum_{s=1}^{j} v_{\hat{u}}^s. \tag{38}$$

Since $u_{\max} = (2\|g_f^i + c_ig_F^i\|/r\varepsilon)$ and $\theta = r\delta(1/u_{\max})$, it follows that, if i is big enough,

$$\theta = \frac{r^2\varepsilon\delta}{2\left\|g_f^i + c_ig_F^i\right\|} \ge \frac{r^2\varepsilon\delta}{2\left\|L_f + c_iL_F\right\|} \ge \frac{r^2\varepsilon\delta}{2\left\|L_f + \bar{c}L_F\right\|}, \tag{39}$$

where $\bar{c}$ is the constant appeared in Lemma 3 and L_f and L_F are the locally Lipschitz constants of f and F on S_1. Therefore, $\|d_{\hat{u}}^j\|$ is bounded away from zero. It follows from $z_{\hat{u}}^j \le 0$ that $|v_{\hat{u}}^j| \ge (\hat{u}/2)\|d_{\hat{u}}^j\|^2$. Hence, $v_{\hat{u}}^j$ is bounded away from zero as well. Therefore, by passing to the limit, we obtain

$$\lim_{i\longrightarrow\infty} e\left(x^{i+1}, c_{i+1}\right) - e\left(x^1, c_1\right) \le -\infty. \tag{40}$$

Note that f is bounded from below as a consequence of the semismoothness of f and the compactness of the level set S_1 of f. By combining the fact $f(x^{i+1}) \le e(x^{i+1}, c_{i+1})$ and (40), we obtain $\lim_{i\longrightarrow\infty} f(x^{i+1}) = -\infty$, and f is unbounded, which leads to a contradiction. Hence, Algorithm 1 cannot be executed for infinitely many times; it stops at a point satisfying condition (37). □

5. Numerical Experiments

To assess practical performance of the presented method, we coded Algorithm 1 in MATLAB and ran it on a PC with 1.80 GHz CPU.

5.1. Examples for Nonconvex Optimization Problems. In this subsection, we first introduce the nonconvex test problems. We prefer a series of polynomial functions developed in [35], also see [23, 36]. For each $i = 1, 2, \ldots, n$, the function h_i: $R^n \longrightarrow R$ is defined by

$$h_i(x) = \sum_{j=1}^{n} x_j + \left(ix_i^2 - 2x_i\right). \tag{41}$$

There are four classes of test functions defined by h_i in [36] as objective functions. It has been proved in [23, 36] that they are nonconvex, globally lower $-C^1$, level coercive, and bounded on compact $\mathcal{X} = B_{15}(0)$. We use one of these test functions to verify the validity and efficiency of the proposed method.

Example 1. Consider problem (1):

$$\begin{aligned} &\min_{x \in \mathbf{R}^n} \ f(x), \\ &\text{s.t.} \quad F(x) \le 0. \end{aligned} \tag{42}$$

For objective function, we define the nonconvex function

$$f(x) := \sum_{i=1}^{n} \left|h_i(x)\right| + \frac{1}{2}\|x\|_2, \quad n = 4. \tag{43}$$

For constraint function, we consider the pointwise maximum of a finite collection of quadratic functions:

$$\begin{aligned} F(x) &= \max\{F_1(x), F_2(x), F_3(x), F_4(x)\}, \\ F_1(x) &= -x_1^2 - 27 * x_1 - 2 * x_4^2 - 22 * x_4 - 23 * x_2 - 21 * x_3 - x_2 * \left(x_2 + x_3\right) - 9; \\ F_2(x) &= -x_1^2 - 28 * x_1 - 2 * x_2^2 - 29 * x_2 - x_4^2 - 21 * x_4 - 21 * x_3 - 3; \\ F_3(x) &= -27 * x_1 - 22 * x_2 - 21 * x_3 - 24 * x_4 - x_3 * \left(x_2 + 2 * x_3\right) - x_2^2 - 5; \\ F_4(x) &= -22 * x_1 - 23 * x_2 - 31 * x_3 - 22 * x_4 - x_1 * \left(x_1 + x_3\right) - x_1 * x_2 - x_3^2 - x_4^2 - 3. \end{aligned} \tag{44}$$

For problem (1) with (43) and (44), we can obtain that $0 = \min_{F(x) \le 0} f(x)$ and $\{0\} \subseteq \operatorname{argmin}_{F(x) \le 0} f(x)$. The results of numerical experiment are as follows:

The initial point: $x_0 = (3,3,3,3)$;

Optimal solution: $x_{\text{final}} = 1.0e - 03 \times (0.63, 0.50, 0.74, 0.36)$;

The final objective function value: $f_{\text{final}} = 0.0056$;

The final constraint function value: $F_{\text{final}} = -3.0051$;

The CPU time: 0.25 seconds.

Example 2. Consider problem (1):

$$\begin{aligned} &\min_{x \in \mathbf{R}^n} \ f(x), \\ &\text{s.t.} \quad F(x) \le 0. \end{aligned} \tag{45}$$

For objective function, we define the nonconvex function

$$f(x) := \sum_{i=1}^{n} \left|h_i(x)\right| + \frac{1}{2}\|x\|_2, \quad n = 6. \tag{46}$$

For constraint function, we consider the pointwise maximum of a finite collection of quadratic functions:

$$\begin{aligned} F(x) &= \max\{F_1(x), F_2(x)\}, \\ F_1(x) &= -37 * x_1 - 33 * x_2 - 41 * x_3 - 32 * x_4 - 33 * x_5 - 36 * x_6 - x_2 * x_4 - x_5 * x_6 \\ &\quad - x_1 * \left(x_1 + 2 * x_5\right) - x_6 * \left(3 * x_2 + x_5\right) - x_4 * \left(x_1 + x_3 + x_4\right) \\ &\quad - x_3 * \left(x_2 + x_5 + x_6\right) - 19; \\ F_2(x) &= -39 * x_1 - 52 * x_2 - 27 * x_3 - 32 * x_4 - 26 * x_5 - 32 * x_6 - x_3 * \left(x_4 + x_6\right) \\ &\quad - x_2 * \left(x_2 - x_3 + 3 * x_5\right) - x_4 * \left(2 * x_2 + 2 * x_5 + x_6\right) - x_1 * \left(x_1 - x_6\right) \\ &\quad - x_5 * \left(2 * x_1 + x_3 + x_4\right) - 11. \end{aligned} \tag{47}$$

For problem (1) with (46) and (47), we can obtain that $0 = \min_{F(x)\leq 0} f(x)$ and $\{0\} \subseteq \operatorname{argmin}_{F(x)\leq 0} f(x)$. The results of the numerical experiment are as follows:

The initial point: $x_0 = (1, 1, 1, 1, 1, 1)$;

Optimal solution: $x_{\text{final}} = 1.0e-05 \times (0.12, -0.07, 0.09, -0.02, 0.00, -0.02)$;

The final objective function value: $f_{\text{final}} = 9.78e-06$;

The final constraint function value: $F_{\text{final}} = -11.00$;

The CPU time: 15.47 seconds.

The above two examples show that the proposed Algorithm 1 does perform not badly since the optimal solutions $1.0e-03 \times (0.63, 0.50, 0.74, 0.36)$ and $1.0e-05 \times (0.12, -0.07, 0.09, -0.02, 0.00, -0.02)$ computed, respectively, by Algorithm 1 are not far away from the true optimal solution to problem (1).

6. Conclusion

For constrained nonconvex optimization problem, we propose an implementable algorithm by combining bundle ideas, proximal control, and exact penalty functions. The results extend the ideas of cutting plane and proximity control to the constrained nonconvex case. We present some techniques for choosing penalty coefficients which ensures the limitation of penalty growth. The penalty parameters are increased only a finite number of times which prevents the algorithm from following closely the curvature boundary of the constrained set. For weakly semismooth functions, the convergence of the presented algorithm to an approximate stationary point of the exact penalty function is proved without any additional assumptions except for the conditions of Slater constraint qualification and the boundedness of the constrained set.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (grant numbers 61877032 and 11601061) and Basic Scientific Research Project of Educational Department of Liaoning Province (grant number JYTMS20231042).

References

[1] E. S. Mistakidis and G. E. Stavroulakis, *Nonconvex Optimization in Mechanics. Smooth and Nonsmooth Algorithms, Heuristics and Engineering Applications*, Kluwer Academic Publisher, Dordrecht, Netherlands, 1998.

[2] J. Outrata, M. Kocvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1998.

[3] J. J. Moreau, P. D. Panagiotopoulos, and G. Strang, *Topics in Nonsmooth Mechanics*, Birkhuser Verlag, Basel, Switzerland, 1988.

[4] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, Berlin, Germany, 1998.

[5] K. Miettinen, *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1999.

[6] A. Fuduli, M. Gaudioso, and G. Giallombardo, "Minimizing nonconvex nonsmooth functions via cutting planes and proximity control," *SIAM Journal on Optimization*, vol. 14, no. 3, pp. 743–756, 2004.

[7] K. C. Kiwiel, "Exact penalty functions in proximal bundle methods for constrained convex nondifferentiable minimization," *Mathematical Programming*, vol. 52, no. 1-3, pp. 285–302, 1991.

[8] L. P. Pang and Z. Q. Xia, "A PVT-type algorithm for minimizing a nonsmooth convex function," *Serdica Mathematical Journal*, vol. 29, no. 1, pp. 11–32, 2003.

[9] L. P. Pang, Z. Q. Xia, and L. W. Zhang, "On a second order parallel variable transformation approach," *Journal of Applied Mathematics and Computing*, vol. 11, no. 1-2, pp. 201–213, 2003.

[10] M. Gaudioso and E. Gorgone, "Gradient set splitting in nonconvex nonsmooth numerical optimization," *Optimization Methods and Software*, vol. 25, no. 1, pp. 59–74, 2010.

[11] C. Lemarechal, "Lagrangian relaxation," in *Computational combinatorial optimization: optimal or provably near-optimal solutions*, pp. 112–156, Springer, Berlin, Germany, 2001.

[12] D. Noll, "Bundle method for non-convex minimization with inexact subgradients and function values," in *Computational and Analytical Mathematics, Volume 50 of Springer Proceedings in Mathematics and Statistics*, D. H. Bailey, H. H. Bauschke, P. Borwein et al., Eds., pp. 555–592, Springer, Berlin, Germany, 2013.

[13] W. Hare, C. Sagastizabal, and M. Solodov, "A proximal bundle method for nonsmooth nonconvex functions with inexact information," *Computational Optimization and Applications*, vol. 63, no. 1, pp. 1–28, 2016.

[14] Y. Yang, L. Pang, X. Ma, and J. Shen, "Constrained nonconvex nonsmooth optimization via proximal bundle method," *Journal of Optimization Theory and Applications*, vol. 163, no. 3, pp. 900–925, 2014.

[15] E. W. Cheney and A. A. Goldstein, "Newton's method for convex programming and Tchebycheff approximation," *Numerical Mathematics*, vol. 1, pp. 253–268, 1959.

[16] J. E. Kelley Jr., "The cutting-plane method for solving convex programs," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 4, pp. 703–712, 1960.

[17] C. A. Sagastizabal and M. V. Solodov, "An infeasible bundle method for nonsmooth convex constrained optimization without a penalty function or a filter," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 146–169, 2005.

[18] E. Karas, A. Ribeiro, C. A. Sagastizabal, and M. V. Solodov, "A bundle-filter method for nonsmooth convex constrained optimization," *Mathematical Programming*, vol. 116, no. 1-2, pp. 297–320, 2009.

[19] K. C. Kiwiel, "An exact penalty function algorithm for nonsmooth convex constrained minimization problems,"

IMA Journal of Numerical Analysis, vol. 5, no. 1, pp. 111–119, 1985.

[20] H. Schramm and J. Zowe, "A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 121–152, 1992.

[21] K. C. Kiwiel, "Restricted step and Levenberg-Marquardt techniques in proximal bundle methods for nonconvex nondifferentiable optimization," *SIAM Journal on Optimization*, vol. 6, no. 1, pp. 227–249, 1996.

[22] D. Noll, "Cutting plane oracles to minimize non-smooth nonconvex functions," *Set-Valued and Variational Analysis*, vol. 18, no. 3-4, pp. 531–568, 2010.

[23] W. Hare and C. Sagastizabal, "A redistributed proximal bundle method for nonconvex optimization," *SIAM Journal on Optimization*, vol. 20, no. 5, pp. 2442–2473, 2010.

[24] A. Fuduli, M. Gaudioso, and E. A. Nurminski, "A splitting bundle approach for non-smooth non-convex minimization," *Optimization*, vol. 64, no. 5, pp. 1131–1151, 2015.

[25] A. Fuduli, M. Gaudioso, and G. Giallombardo, "A DC piecewise affine model and a bundling technique in nonconvex nonsmooth minimization," *Optimization Methods and Software*, vol. 19, no. 1, pp. 89–102, 2004.

[26] J. V. Burke, A. S. Lewis, and M. L. Overton, "A robust gradient sampling algorithm for nonsmooth, nonconvex optimization," *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 751–779, 2005.

[27] K. C. Kiwiel, "Convergence of the gradient sampling algorithm for nonsmooth nonconvex optimization," *SIAM Journal on Optimization*, vol. 18, no. 2, pp. 379–388, 2007.

[28] W. Hare and C. Sagastizabal, "Computing proximal points of nonconvex functions," *Mathematical Programming*, vol. 116, no. 1-2, pp. 221–258, 2009.

[29] K. C. Kiwiel, "A dual method for certain positive semidefinite quadratic programming problems," *SIAM Journal on Scientific and Statistical Computing*, vol. 10, no. 1, pp. 175–186, 1989.

[30] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, Hoboken, NJ, USA, 1983.

[31] A. Fuduli and M. Gaudioso, "The proximal trajectory algorithm for convex minimization," Technical Report 7/98, Universita della Calabria, Arcavacata, Italy, 1998.

[32] C. Lemarechal, A. Nemirovskii, and Y. Nesterov, "New variants of bundle methods," *Mathematical Programming*, vol. 69, no. 1-3, pp. 111–147, 1995.

[33] R. Mifflin, "An algorithm for constrained optimization with semismooth functions," *Mathematics of Operations Research*, vol. 2, pp. 191–207, 1977.

[34] R. Fletcher and S. Leyffer, "A bundle filter method for nonsmooth nonlinear optimization," Numerical Analysis Report NA/195, The University of Dundee, Dundee, Scotland, 1999.

[35] C. Ferrier, *Bornes Duales de Problemes d'Optimisation Polynomiaux*, Ph.D. thesis, Universite Paul Sabatier, Toulouse, France, 1997.

[36] C. Ferrier, "Computation of the distance to semi-algebraic sets," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 5, pp. 139–156, 2000.