



A class of infeasible proximal bundle methods for nonsmooth nonconvex multi-objective optimization problems

Li-Ping Pang^{1,2} · Fan-Yun Meng³ · Jian-Song Yang¹

Received: 10 November 2019 / Accepted: 10 September 2022 / Published online: 13 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

We propose a class of infeasible proximal bundle methods for solving nonsmooth nonconvex multi-objective optimization problems. The proposed algorithms have no requirements on the feasibility of the initial points. In the algorithms, the multi-objective functions are handled directly without any scalarization procedure. To speed up the convergence of the infeasible algorithm, an acceleration technique, i.e., the penalty skill, is applied into the algorithm. The strategies are introduced to adjust the proximal parameters and penalty parameters. Under some wild assumptions, the sequence generated by infeasible proximal bundle methods converges to the globally Pareto solution of multi-objective optimization problems. Numerical results shows the good performance of the proposed algorithms.

Keywords Multi-objective optimization · Bundle method · Nonsmooth optimization

Mathematics Subject Classification 90C25 · 49J40 · 90C20

1 Introduction

In this paper we seek to solve the nonsmooth nonconvex multi-objective optimization problem:

$$\begin{aligned} \text{(MOP)} \quad & \min F(x) := (f_1(x), \dots, f_p(x)) \\ & \text{s. t. } c(x) \leq 0, \end{aligned}$$

The research are partially supported by the Natural Science Foundation of Shandong Province, Grant ZR2019BA014, the Natural Science Foundation of Zhejiang Province, Grant LY20A010025.

✉ Fan-Yun Meng
mengfanyundlut@163.com

¹ School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

² Key Laboratory for Computational Mathematics and Data Intelligence of Liaoning Province, Dalian 116024, China

³ School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China

where the objective functions $f_i : R^n \rightarrow R, i = 1, 2, \dots, p$ and the constraint function $c : R^n \rightarrow R$ are locally Lipschitz continuous. Without loss of generality, we consider only one constraint function in (MOP). When dealing with multiple inequality constraints, one can utilize the maximum function to transform these finite constraints into one constraint.

The multi-objective optimization problem is to optimize several conflicting objectives at the same time, which has arisen in many real-life applications of finance, engineering, transportation, or mechanics, see [1, 4, 18, 21, 23] etc. Since these multiple objectives are conflict, the optimal solution in single objective optimization problems is not suitable for multi-objective optimization problems. The Pareto (efficient) solution of multi-objective optimization problems is proposed. The Pareto solution means that there does not exist another solution whose objective function value is the same or smaller than that of the Pareto solution, and at least one objective function value is strictly smaller. There are many Pareto solutions for multi-objective optimization problems. The set of all Pareto solutions is usually called the Pareto set and its corresponding multi-objective function value is called the Pareto frontier.

The theoretical research on multi-objective programming has achieved abundant results [24–26, 32], including various optimality conditions and stability results. The research on the numerical algorithms of multi-objective programming is relatively little. There are two main types of algorithms for solving multi-objective programming. One classical algorithm is based on the scalarization strategy [2, 3], which is to transform the multi-objective programming into its equivalent single-objective programming. This scalarization strategy has great influence on the efficiency of the algorithm and the convergence of the algorithm depends on the scalarization skill. The other kind of algorithms is to solve the multi-objective programming directly, without any scalarizations, see [6, 14, 19, 29]. This kind of multi-objective algorithms has its own special advantages and it can use all the information of the functions, which does not depend on any scalarization technologies.

It is universally acknowledged that the bundle method is a highly reliable and effective method for solving nonsmooth nonconvex optimization problems [7–10, 31]. In recent years, bundle methods are applied into various types of optimization problems, such as infinite programs [12, 27], control optimization problems [30], equilibrium problems [17, 22]. With the development of multi-objective optimization, the bundle methods are applied into solving nonsmooth multi-objective programs, refer to [13, 20]. However, the existing algorithms for solving nonsmooth multiobjective programming still have some limitations. The proximal bundle method proposed in [14] has the requirements on the feasibility of initial points, which makes great impact on the convergence of the algorithm. Obtaining a feasible initial point is very difficult because that finding a feasible initial point of multi-objective programming is almost the same as that of solving the whole problem. The paper [29] used the cutting-plane method to solve the multi-objective programming. The proposed auxiliary problem in [29], which adds the multi-objective functions to the constraint. The increasing number of constraints brings the result that the constraint qualification of the original problem may be hardly to satisfy, and it increases the difficulty of solving the subproblem of the cutting-plane model.

The aim of this paper is to utilize the proximal bundle method to seek Pareto solutions of multiobjective optimization problems. The multiple objective functions are handled individually without any scalarization procedure. By virtue of improvement functions and the bundle of information, we construct the subproblem of the multi-objective optimization problem. What difference from other existed multi-objective algorithms is that, in this paper, the proposed algorithm has no requirement on the feasibility of initial points, which means that the starting point is arbitrary and it may be infeasible. To further increase the convergent speed of the algorithm, an acceleration procedure is proposed, which is the penalty procedure. Some

strategies are also put forward to adjust the penalty parameters. Under the wild assumptions, the sequence generated by penalized infeasible proximal bundle methods globally converges to the Pareto solution of multi-objective optimization problems.

The remainder of the paper is organized as follows. Section 2 gives some preliminaries about nonsmooth optimization and multiobjective optimization. Section 3 gives the infeasible proximal bundle method and the penalized infeasible proximal bundle method, moreover, the corresponding convergence results are provided. Section 4 is devoted to the numerical experiments illustrating the efficiency of the methods. Section 5 is the conclusion.

2 Preliminaries

In this section we give some preliminaries from the convex analysis and multi-objective optimization problems, which are needed in the following discussion.

Definition 2.1 [14] A function $f : R^n \rightarrow R$ is f° -pseudoconvex, if it is locally Lipschitz continuous, and for all $x, y \in R^n$,

$$f(y) < f(x) \text{ implies } f^\circ(x; y - x) < 0.$$

Definition 2.2 [14] For a locally Lipschitz continuous function $f : R^n \rightarrow R$ the Clarke generalized directional derivative at x in the direction $d \in R^n$ is defined by

$$f^\circ(x; d) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y + td) - f(y)}{t}$$

and the Clarke subdifferential of f at x is

$$\partial f(x) = \{\xi \in R^n \mid f^\circ(x; d) \geq \xi^T d, \forall d \in R^n\}.$$

From the property of the Clarke subdifferential, it has $f^\circ(x; d) = \max_{g \in \partial f(x)} g^T d$.

Next we present two assumptions which are used in the later theorems.

Assumption 2.1 For the given sequence $\{x^k\} \in R^n, \{y^l\} \in R^n$, the constant $M_0 \geq 0$, the level set satisfies

$$L_0 = \{x \in R^n \mid \max\{f_i(y^{l+1}) - f_i(x^k), c(y^{l+1}) - c(x^k) \mid i = 1, 2, \dots, p\} < M_0\} \subseteq \mathcal{O},$$

where \mathcal{O} is an open and bounded set.

Definition 2.3 [15] The functions $f : R^n \rightarrow R$ is said to be subdifferentially regular at $x \in R^n$ if it is locally Lipschitz continuous at x and for all $d \in R^n$ the classical directional derivative

$$f'(x; d) = \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

exists and $f'(x; d) = f^\circ(x; d)$.

The following theorems give the important properties of f° -pseudoconvex function.

Theorem 2.1 [15] A f° -pseudoconvex function $f(x)$ attains its global minimum at x^* , if and only if $0 \in \partial f(x^*)$.

Theorem 2.2 [15] *Let $f_i : R^n \rightarrow R$ be f° -pseudoconvex for all $i = 1, 2, \dots, p$. Then the function*

$$f(x) = \max\{f_i(x) \mid i = 1, 2, \dots, p\}$$

is also f° -pseudoconvex, and

$$\partial f(x) \subseteq \text{conv}\{\partial f_i(x) \mid f_i(x) = f(x), i = 1, \dots, p\}.$$

Furthermore, if f_i satisfies Definition 2.3, then

$$\partial f(x) = \text{conv}\{\partial f_i(x) \mid f_i(x) = f(x), i = 1, \dots, p\}.$$

The ideal solution of (MOP) is:

$$x^* \in C, \quad f_i(x^*) \leq f_i(x), \forall i = 1, 2, \dots, p,$$

where C is the feasible set of (MOP). The ideal solution is hard to obtain because that these objectives are often balanced against each other. In the following we give the Pareto solution of (MOP).

Definition 2.4 [14] A vector x^* is said to be a global Pareto solution of (MOP), if there does not exist $x \in C$ such that

$$f_i(x) \leq f_i(x^*), \text{ for all } i = 1, 2, \dots, p \text{ and } f_j(x) < f_j(x^*) \text{ for some } j.$$

Vector x^* is said to be a global weak Pareto solution of (MOP), if there does not exist $x \in C$ such that

$$f_i(x) < f_i(x^*), \text{ for all } i = 1, 2, \dots, p.$$

Definition 2.5 [14] A vector x^* is called the local (weakly) Pareto solution of (MOP), if there exists $\delta > 0$ such that x^* is the (weakly) Pareto solution on $B(x^*, \delta) \cap C$. It can be seen that the Pareto solution must be the weakly Pareto solution.

The tangent cone and polar cone of set $K \subseteq R^n$ at x are defined as

$$T_K(x) = \{d \in R^n \mid \text{there exist } t_i \rightarrow 0 \text{ and } d_i \rightarrow d \text{ with } x + t_i d_i \in K\},$$

$$K^\leq = \{d \in R^n \mid w^T d \leq 0, \forall w \in K\}.$$

The convex hull and closure hull of K are denoted by $\text{conv} K$ and $\text{cl} K$. Denote

$$\text{ray} K = \{\lambda v \mid \lambda \geq 0, v \in K\} \text{ and } \text{cone} K = \text{ray conv} K.$$

Let $C := \{x \in R^n \mid c(x) \leq 0\}$, $K(x) := \partial c(x)$, and the constraint qualification

$$K^\leq(x) \subseteq T_C(x) \tag{2.1}$$

is posed to obtain optimality conditions.

The following theorem gives the optimality condition of (MOP).

Theorem 2.3 [15] *x^* is the local weakly Pareto solution of (MOP), and the constraint qualification (2.1) is satisfied, then*

$$0 \in \text{conv} \cup_{i=1}^p \partial f_i(x^*) + \text{cl cone } \partial c(x^*). \tag{2.2}$$

Moreover, if $f_i, i = 1, 2, \dots, p, c$ are f° -pseudoconvex, the condition (2.2) is the sufficient condition of that x^ is the global weakly Pareto solution.*

3 Infeasible multi-objective proximal bundle methods

In this section, firstly we give the infeasible multi-objective proximal bundle method (IMPB), the proposed method has an advantage that it has no requirement on the feasibility of initial points, which means that the algorithm can begin from any starting points.

To improve the speed of the IMPB method, we propose an acceleration procedure, which is the penalty skill, and, the strategy to adjust the penalty parameters is also given.

For both two algorithms, under wild assumptions, the sequences generated by the algorithms converge to the globally Pareto solutions of (MOP).

3.1 The infeasible multi-objective proximal bundle method

The IMPB method is the extension of proximal bundle methods into the multi-objective case, which is usually used to handle the nonconvex constrained problems [6, 16]. The idea of the IMPB method is to find the direction from any starting point, in which the values of all objective functions improve simultaneously. For the IMPB method, it still needs the scalarization to obtain the minimization function in the bundle. In this paper, we use the improvement function $H : R^n \times R^n \rightarrow R$:

$$H_{x^k}(y) = \max\{f_i(y) - f_i(x^k), c(y) \mid i = 1, 2, \dots, p\}. \tag{3.1}$$

The following theorem gives the connection between the problem (MOP) and the improvement function.

Theorem 3.1 [14] *A necessary condition for $x^* \in R^n$ to be a global weak Pareto solution of (MOP) is that*

$$x^* = \arg \min_{y \in R^n} H_{x^*}(y). \tag{3.2}$$

Moreover, if $f_i, i = 1, \dots, p, c$ are f° -pseudoconvex and the constraint qualification (2.1) is valid, then the condition (3.2) is sufficient for x^ to be a global weak Pareto solution of (MOP).*

Suppose that there are some reference points $x^j \in R^n$ from the past iterations and subgradients $g_{f_i}^j \in \partial f_i(x^j), j \in L_l^{f_i}, L_l^{f_i} \subseteq \{1, 2, \dots, l\}$ and $g_c^j \in \partial c(x^j), j \in L_l^c, L_l^c \subseteq \{1, 2, \dots, l\}$. The bundle of information are given as follows:

$$\begin{aligned} B_k^{f_i} &= \{(x^j, f_i(x^j), g_{f_i}^j) : j \in L_l^{f_i}\}, L_l^{f_i} \subseteq \{1, 2, \dots, l\}; \\ B_k^c &= \{(x^j, c(y^j), g_c^j) : j \in L_l^c\}, L_l^c \subseteq \{1, 2, \dots, l\}. \end{aligned}$$

Let x^k be the current approximation point of (MOP) at the iteration k . The linearization errors of $f_i(x)$ and $c(x)$ at y^j are defined as

$$\begin{aligned} e_{f_{ij}}^k &= f_i(x^k) - f_i(y^j) - \langle g_{f_i}^j, x^k - y^j \rangle, j \in L_l^{f_i}, \\ e_{c_j}^k &= c(x^k) - c(y^j) - \langle g_c^j, x^k - y^j \rangle, j \in L_l^c. \end{aligned} \tag{3.3}$$

Because $f_i(x)$ and $c(x)$ are nonsmooth nonconvex functions, the corresponding linearization errors $e_{f_{ij}}^k$ and $e_{c_j}^k$ may be negative. In order to make the linearization errors nonnegative, we consider the generalized linearization errors $\hat{e}_{f_{ij}}^k = e_{f_{ij}}^k + \eta_l d_j^k, \hat{e}_{c_j}^k + \eta_l d_j^k$, where $d_j^k =$

$\frac{1}{2}\|y^j - x^k\|^2$, η_l is the convexificaton parameter [5]. In this paper, the minimal value of convexificaton parameter η_l is chosen as

$$\eta_l^{\min} = \max_{1 \leq i \leq p} \left\{ \max_{j \in L_i^{f_i} \setminus \{j_k\}} \frac{-2e_{f_{ij}}^k}{\|y^j - x^k\|^2}, \max_{j \in L_i^{c_j} \setminus \{j_k\}} \frac{-2e_{c_j}^k}{\|y^j - x^k\|^2}, 0 \right\}. \tag{3.4}$$

When $\eta_l \geq \eta_l^{\min}$, it has $\hat{e}_{f_{ij}}^k \geq 0, \hat{e}_{c_j}^k \geq 0$. For simplicity, denote $\Delta_j^k = y^j - x^k, c^+(x^k) = \max\{c(x^k), 0\}$.

The piecewise approximation model of the objective function $f_i(x)$ at y^j is

$$\begin{aligned} \check{f}_{il}(y) &= \max_{j \in L_i^{f_i}} \{f_i(y^j) + \eta_l d_j^k + \eta_l \langle g_{f_i}^j + \eta_l \Delta_j^k, y - y^j \rangle\} \\ &= f_i(x^k) + \max_{j \in L_i^{f_i}} \{-\hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle\}. \end{aligned}$$

The piecewise approximation model of the constraint function $c(x)$ at y^j is

$$\begin{aligned} \check{c}_l(y) &= \max_{j \in L_l^c} \{c(y^j) + \eta_l d_j^k + \eta_l \langle g_c^j + \eta_l \Delta_j^k, y - y^j \rangle\} \\ &= c(x^k) + \max_{j \in L_l^c} \{-\hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle\}. \end{aligned}$$

Then the approximation of the improvement function (3.1) is

$$\check{H}_l(y) = \max\{\check{f}_{il}(y) - f_i(x^k), c_l(y) \mid i = 1, 2, \dots, p\}. \tag{3.5}$$

Solve the following quadratic problem

$$y^{l+1} = \arg \min_{y \in \mathbb{R}^n} \{\check{H}_l(y) + \frac{u_l}{2} \|y - x^k\|^2\} \tag{3.6}$$

to get the candidate point y^{l+1} , where x^k is called the proximal center and u_l is the proximal parameter.

The quadratic problem (3.6) is equivalent to the problem

$$\begin{aligned} \min r + \frac{u_l}{2} \|y - x^k\|^2 \\ \text{s. t. } f_i(x^k) - \hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle \leq r, j \in L_i^{f_i}, i = 1, 2, \dots, p, \\ c(x^k) - \hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle \leq r, j \in L_l^c. \end{aligned} \tag{3.7}$$

The Lagrange function of (3.7) is

$$\begin{aligned} L(y, r, \lambda, u) &= r + \frac{u_l}{2} \|y - x^k\|^2 + \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} [f_i(x^k) - \hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle - r] \\ &\quad + \sum_{j \in L_l^c} u_j [c(x^k) - \hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle - r]. \end{aligned}$$

The dual problem of (3.7) is

$$\begin{aligned} \min_{\lambda, u \geq 0} & \frac{\|G^l\|^2}{2} - \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} [f_i(x^k) - \hat{e}_{f_{ij}}^k] - \sum_{j \in L_i^c} u_j [c(x^k) - \hat{e}_{c_j}^k] \\ \text{s. t.} & \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} + \sum_{j \in L_i^c} u_j = 1, \lambda_{ij} \geq 0, u_j \geq 0, \end{aligned} \tag{3.8}$$

where $G^l = \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} (g_{f_i}^j + \eta_l \Delta_j^k) + \sum_{j \in L_i^c} u_j (g_c^j + \eta_l \Delta_j^k)$.

Solve the dual problem (3.8) and obtain the solution y^{l+1} , which is

$$y^{l+1} = x^k - \frac{1}{u_l} G^l. \tag{3.9}$$

The aggregate linearization function at y^{l+1} is $\psi_l(y) = \check{H}_l(y^{l+1}) + \langle G^l, y - y^{l+1} \rangle$. The function $\psi_l(y)$ is affine and $\psi_l(y^{l+1}) = \check{H}_l(y^{l+1})$, $G^l = \nabla \psi_l(y)$, hence, we obtain $\psi_l(y) \leq \check{H}_l(y), \forall y \in R^n$. The aggregate linearization error at y^{l+1} is

$$\varepsilon_l^k = H_k(x^k) - \psi_l(y^{l+1}) = c^+(x^k) - \check{H}_l(y^{l+1}) - u_l \|y^{l+1} - x^k\|^2.$$

Furthermore, the predicted decrease is

$$\delta_l = c^+(x^k) + \frac{\eta_l}{2} \|y^{l+1} - x^k\|^2 - \check{H}_l(y^{l+1}) = \varepsilon_l^k + R_l \|y^{l+1} - x^k\|^2, \tag{3.10}$$

where $R_l = u_l + \frac{\eta_l}{2}$.

Now we propose the infeasible multi-objective proximal bundle method.

The IMPB method

Step 0 (Initialization) Choose a infeasible starting point $x^0 \in C$, and select a unacceptable functional increasing parameter M_0 , a convex parameter Γ_0 , a proximal parameter Γ_1 , a stopping tolerance τ_{stop} . Set $y_0^0 = x^0$, compute $g_{f_i}^0 \in \partial f_i(y^0), i = 1, 2, \dots, p, g_c^0 \in \partial c(y^0)$. Initialize $u_1, \eta_1, e_{f_{i_0}}^0 = e_{c_0}^0 = d_0^0 = \Delta_0^0 = 0$.

Step 1 (Computing the candidat point) Solve the quadratic programming (3.6) and obtain the solution y^{l+1} . Compute the predicted decrease δ_l by (3.10).

Step 2 (Stopping test) If $\delta_l \leq \tau_{stop}$, then stop.

Step 3 (Updating the bundle) Update the index set $L_{l+1}^{f_i} \supseteq \{l+1, i_k\}, L_{l+1}^c \supseteq \{l+1, i_k\}$. Compute $g_{f_i}^{l+1} \in \partial f_i(y^{l+1}), i = 1, 2, \dots, p, g_c^{l+1} \in \partial c(y^{l+1})$. Compute the bundle of information:

$$\begin{aligned} e_{f_{i,l+1}}^k &= f_i(x^k) - f_i(y^{l+1}) - \langle g_{f_i}^{l+1}, x^k - y^{l+1} \rangle, j \in L_i^{f_i}, \\ e_{c_{l+1}}^k &= c(x^k) - c(y^{l+1}) - \langle g_c^{l+1}, x^k - y^{l+1} \rangle, j \in L_l^c. \end{aligned}$$

Step 4 (Testing the descent criterion) If the candidat point y^{l+1} is good enough, i.e.,

$$\begin{cases} \max_{1 \leq i \leq p} (f_i(y^{l+1}) - f_i(x^k)) \leq -m\delta_l \text{ and } c(y^{l+1}) \leq 0, & \text{if } c(x^k) \leq 0; \\ c(y^{l+1}) \leq c(x^k) - m\delta_l, & \text{if } c(x^k) > 0. \end{cases} \tag{3.11}$$

Update the stability center, and set $x^{k+1} = y^{l+1}$; otherwise, set $x^{k+1} = x^k$.

Step 5 (Updating the convexification parameter) Compute η_{l+1}^{min} by (3.4), and set

$$\begin{cases} \eta_{l+1} = \eta_l, & \text{if } \eta_{l+1}^{min} \leq \eta_l; \\ \eta_{l+1} = \Gamma_0 \eta_{l+1}^{min}, & \text{if } \eta_{l+1}^{min} > \eta_l. \end{cases} \tag{3.12}$$

Step 6 (Updating the proximal parameter) If $\max\{f_i(y^{l+1}) - f_i(x^k), c(y^{l+1}) - c(x^k) \mid i = 1, 2, \dots, p\} > M_0$, set $u_{l+1} = \Gamma_1 u_l$; otherwise, set $u_{l+1} = u_l$, and go to Step 1.

There are some explanations about the IMPB method.

- (i) The descent criterion (3.11) has two distinct advantages: firstly, when the last stability center is not feasible, the second formula of (3.11) is used to reduce the infeasibility; secondly, when the last stability center is feasible, on the basis of maintaining the feasibility of candidate points, the value of objective functions is reduced.
- (ii) In Step 3, in order to find the new searching direction, it needs to reserve the information of new iteration points into the bundle. The size of the bundle increases with the number of iterations. When the number of iterations is large, due to the limited computer memory space, it is necessary to compress the bundle, which is called the aggregate technique. The readers can refer to [28] for more details about these aggregate technique.
- (iii) In Step 5, the updating manner of η_l will finally make $\eta_l \geq \eta_{l+1}^{min}$, $\hat{e}_{f_{ij}}^k \geq 0$, $\hat{e}_c^k \geq 0$ after many iterations. In Step 6, in order to obtain $\{y^l\} \subseteq L_0$, it needs to increase the proximal parameter u_l and adjust the piecewise approximation model, however, this may lead the infinite loop between Step 1 and Step 6. The following lemma proves that this loop will stop in finite steps.

Lemma 3.1 *Suppose that the sequence of the candidate points generated by the IMPB method is bounded, i.e., $\{y^l\} \subseteq L_0$. Then there exists the iteration index $l' > 0$ such that $u_l = \bar{u}$, $\eta_l = \bar{\eta}$ for all $l > l'$.*

Proof Because $L_0 \subseteq O$, where O is an open bounded set, L_0 is a closed bounded set. The functions $f_i, i = 1, 2, \dots, p, c$ are Lipschitz continuous and there are Lipschitzian constants $L_{il}, i = 1, 2, \dots, p, L_c$. Take $L_f = \max_{1 \leq i \leq p} \{L_{il}\}, L = \max\{L_f, L_c\}, r = \frac{M_0}{L}$, for all $x \in B_r(x^k)$, we have

$$\begin{aligned} \|f_i(x) - f_i(x^k)\| &\leq L\|x - x^k\| \leq M_0, \\ \|c(x) - c(x^k)\| &\leq L\|x - x^k\| \leq M_0. \end{aligned}$$

Due to $g_{f_i}^j \in \partial f_i(y^j), g_c^j \in \partial c(y^j), y^j \in L_0$, it has $\|g_{f_i}^j\| \leq L, \|g_c^j\| \leq L$. Take $i = i_k$, then $g_{h_k}^{i_k} \in \partial \check{H}_l(x^k)$, which is $\check{H}_l(y) \geq \check{H}_l(x^k) + \langle g_{h_k}^{i_k}, y - x^k \rangle$. Moreover, it holds that

$$\begin{aligned} &\arg \min_y \{ \check{H}_l(y) + \frac{u_l}{2} \|y - x^k\|^2 \} \\ &\subseteq \{ y : \check{H}_l(x^k) + \langle g_{h_k}^{i_k}, y - x^k \rangle + \frac{u_l}{2} \|y - x^k\|^2 \leq \check{H}_l(x^k) \} \\ &\subseteq \{ y : -\|g_{h_k}^{i_k}\| \|y - x^k\| + \frac{u_l}{2} \|y - x^k\|^2 \leq 0 \} \\ &\subseteq \{ y : \|y - x^k\| \leq \frac{2\|g_{h_k}^{i_k}\|}{u_l} \} \\ &\subseteq \{ y : \|y - x^k\| \leq \frac{2T}{u_l} \}, \text{ where } T = L_f + L_c. \end{aligned}$$

When $l \rightarrow \infty, u_l$ becomes very large, then $\frac{2T}{u_l}$ becomes small. Moreover, the functions $f_i, i = 1, 2, \dots, p, c$ are Lipschitz, then $\max_{1 \leq i \leq p} \{f_i(y^{l+1}) - f_i(x^k), c(y^{l+1}) - c(x^k)\} > M_0$ does not hold, hence, there exists $l' > 0$, when $l > l'$, it has $u_l = \bar{u}$.

The convexification parameter η_l in Step 5 is nondecreasing during the iteration. In contrast, assume that η_l does not converge to a stable value. After the iteration l' , the generalized linearization errors are nonnegative, i.e., $\hat{e}_{f_l}^k \geq 0, \hat{e}_{c_l}^k \geq 0$. When $l > l'$, it has that $\eta_l > \eta_k^{min}$. Hence, the parameter η_l no longer updates, and $\eta_l = \bar{\eta}$. \square

Next we will give the convergence of the IMPB method. The set C is the feasible set of (MOP), i.e., $C = \{x \in R^n \mid c(x) \leq 0\}$. k is the index of serious steps, and L_s is the index set of all serious steps.

$l(k)$ is the index when the candidate point $y^{l(k)}$ becomes the stability center x^k , i.e., $x^k = y^{l(k)}$. When the number of iterations becomes very large, the convexification parameter η_l and the proximal parameter u_l will finally remain unchanged. Let the stopping parameter $\tau_{stop} = 0$ and the predicted decreasing $\delta_l > 0$. Suppose that the algorithm generates two kinds of infinite steps, which are: (1) there are infinite serious steps; (2) there are many finite serious steps and then there are many infinite null steps. Before obtaining the convergence of the algorithm, a lemma is given.

Lemma 3.2 *For any decreasing index $k_0 > 0$, the following conclusions are satisfied:*

- (i) $x^k \in \{x \in R^n \mid c(x) \leq c^+(x^{k_0})\}, \forall k \geq k_0$;
- (ii) *if there exists k_1 satisfying $x^{k_1} \in C$, then $x^k \in C, \forall k \geq k_1$.*

Proof Fix the last decreasing index by k_0 , i.e., all the iterate steps are null steps after k_0 , the conclusion (i) holds clearly. Assume that there are $k_0 + 1$ serious steps, if $x^{k_0} \notin C$, by (3.11), we have $c(x^{k_0+1}) < c^+(x^{k_0}) = c(x^{k_0})$. Moreover, if $x^k \notin C, \forall k \geq k_0$, repeating the previous step, then $\{c(x^k)\}$ is nonincreasing, therefore, $c(x^k) < c^+(x^{k_0}) = c(x^{k_0}), k \geq k_0$.

Assume that it exists k_1 satisfying $x^{k_1} \in C$. If k_1 is the last serious step, the conclusion (ii) holds. Suppose that there are $k_1 + 1$ serious steps, by (3.11), it has $c(x^{k_1+1}) \leq 0 = c^+(x^{k_1})$. Repeating the previous step, we obtain $c(x^k) \leq 0 = c^+(x^{k_1}), \forall k \geq k_1$, which means $x^k \in C$. The conclusion holds. \square

Lemma 3.2 shows that, no matter the initial point is feasible or infeasible, the proposed algorithm can be applied into both two cases. Once the iteration point is feasible, the algorithm becomes feasible and the speed of the algorithm is increased. It is of great practical significance for applying the algorithm into more general multi-objective optimization fields.

Lemma 3.3 *Suppose that the IMPB method generates a infinite sequence of serious steps, and L_s is the index set of all serious steps, then we have*

$$\lim_{k \rightarrow \infty} \delta_k = 0, \lim_{k \rightarrow \infty} \varepsilon_l^k = 0, \lim_{k \rightarrow \infty} \|G^k\| = 0, k \in L_s.$$

Proof We prove the conclusion from the following two cases. The first case is that the serious step satisfying the first condition of (3.11), i.e., $\exists \hat{k}$ s.t. $c(x^{\hat{k}}) \leq 0, c(x^{k+1}) \leq 0, \forall k \geq \hat{k}, k \in L_s$, moreover, $\max_{1 \leq i \leq p} \{f_i(x^{k+1}) - f_i(x^k)\} \leq -m\delta_k$. Then we obtain $f_i(x^{k+1}) - f_i(x^k) \leq -m\delta_k, \forall i = 1, 2, \dots, p$, and $\{f_i(x^k)\}$ is nonincreasing. Because f_i is bounded and f° -pseudoconvex, there exists the limit of the function. Denoting $\{f_i(x^k)\} \rightarrow \hat{f}_i, i = 1, 2, \dots, p$, it has that

$$m \sum_{k \in L_s} \delta_k \leq \sum_{k \in L_s} (f_1(x^k) - f_1(x^{k+1})), \forall k \geq \hat{k},$$

i.e.,

$$\sum_{k \in L_s} \delta_k \leq \frac{1}{m} \sum_{k \in L_s} (f_1(x^k) - f_1(x^{k+1})), k \geq \hat{k}.$$

Similarly,

$$\sum_{k \in L_s} \delta_k \leq \frac{1}{m} \sum_{k \in L_s} (f_i(x^k) - f_i(x^{k+1})), i = 1, 2, \dots, p.$$

After many infinite iterations, we obtain that

$$\sum_{k \in L_s} \delta_k \leq \frac{1}{m} \sum_{k \in L_s} (f_i(x^{\hat{k}}) - \hat{f}_i), i = 1, 2, \dots, p.$$

Therefore,

$$\sum_{k \in L_s} \delta_k \leq \frac{1}{mp} (f_1(x^{\hat{k}}) - \hat{f}_1 + f_2(x^{\hat{k}}) - \hat{f}_2 + \dots + f_p(x^{\hat{k}}) - \hat{f}_p),$$

hence, we have $\lim_{k \rightarrow \infty} \delta_k = 0$.

The second case is that the serious step satisfying the second condition of (3.11), which is $c(x^k) > 0$. From the second formula of (3.11), it has $c(x^{k+1}) - c(x^k) \leq -m\delta_k$, which means $0 < \delta_k \leq \frac{1}{m}(c(x^k) - c(x^{k+1}))$. Because $c(x)$ is bounded and f° -pseudoconvex, the limit of $c(x)$ exists. Denoting $\{c(x^k)\} \rightarrow \hat{c}$, it holds that

$$\sum_{k \in L_s} \delta_k \leq \frac{1}{m} \sum_{k \in L_s} (c(x^k) - c(x^{k+1})) = \frac{1}{m} (c(x^{\hat{k}}) - \hat{c}).$$

Therefore, we have $\lim_{k \rightarrow \infty} \delta_k = 0$. It follows from (3.10) that $0 \leq \varepsilon_l^k \leq \delta_k$ and $\lim_{k \rightarrow \infty} \varepsilon_l^k = 0$.

Due to $y^{l+1} = x^k - \frac{1}{u_l} G^k$, $\frac{1}{u_k} \geq \frac{1}{\bar{u}}$, it holds $\delta_k \geq \varepsilon_l^k + \frac{1}{u_k} \|G^k\|^2$. Take the limit and we obtain $\lim_{k \rightarrow \infty} \|G^k\| = 0$. □

Now we present the convergence of the IMPB method.

Theorem 3.2 *Suppose that the IMPB method generates a infinite sequence of serious steps, then the accumulation point of $\{x^k\}$ is the weakly Pareto solution of (MOP).*

Proof For every $y \in R^n$, it has $H_{x^k}(y) \geq c^+(x^k) + \langle G^k, y - x^k \rangle - \varepsilon_l^k, l = l(k) \in L_s$. Due to $u_l \leq \bar{u}$, by Lemma 3.2, it holds $\lim_{k \rightarrow \infty} \varepsilon_l^k = 0, \lim_{l \rightarrow \infty} \|G^l\| = 0$. The sequence $\{x^k\}$ is bounded and denote the subsequence by $\{x^{k_i}\}$. When $k \rightarrow \infty$, it has $x^{k_i} \rightarrow \bar{x}$. Moreover, there is $H_{\bar{x}}(y) \geq c^+(\bar{x}) + \langle 0, y - \bar{x} \rangle - 0 = c^+(\bar{x}) = H_{\bar{x}}(\bar{x})$. From the arbitrariness of y , it yields that $\min\{H_{\bar{x}}(y) \mid y \in R^n\} = H_{\bar{x}}(\bar{x}) = c^+(\bar{x})$, i.e., $\bar{x} = \arg \min_{y \in R^n} H_{\bar{x}}(y)$. In what follows,

we prove that $H_{\bar{x}}(\bar{x}) = c^+(\bar{x}) = 0$ by contradiction. Suppose that $c^+(\bar{x}) > 0$, by the local Lipschitz continuity of c , it holds that $c^+(y) = c(y) > f(y) - f(x), \forall y \in B(\bar{x})$. Because \bar{x} is the local optimality solution, it has $c^+(\bar{x}) > 0$, which contradicts with the solution set being nonempty, hence, the relation $H_{\bar{x}}(\bar{x}) = c^+(\bar{x}) = 0, 0 \in \partial H_{\bar{x}}(\bar{x})$ is satisfied. By Theorem 3.1, we obtain that \bar{x} is the weakly Pareto solution of (MOP). □

Theorem 3.3 *Suppose that the IMPB method generates an infinite sequence of null steps and $x^{k_{last}}$ is the last stability center, then $x^{k_{last}}$ is the weakly Pareto solution of (MOP).*

Proof Fix $H(\cdot) = H_{k_{last}}(\cdot)$, by Lemmas 3.2 and 3.3, it holds that $\lim_{l \rightarrow \infty} y^l = \bar{y}$ and $\lim_{l \rightarrow \infty} \check{H}_l(y^{l+1}) = H(\bar{y}) + \frac{\eta}{2} \|\bar{y} - x^{k_{last}}\|^2$. Hence, the equality $H(\bar{y}) + \frac{\eta}{2} \|\bar{y} - x^{k_{last}}\|^2 \leq c(x^{k_{last}}) = H(x^{k_{last}})$ is satisfied. Because that k_{last} is the index of the last stability center, when $l > k_{last}$, the decreasing criterion (3.11) is not satisfied, the following relations are satisfied.

(i) If $c(x^{k_{last}}) \leq 0$, it has $\max_{1 \leq i \leq p} (f_i(y^{l+1}) - f_i(x^{k_{last}})) > -m\delta_l$ and $c(y^{l+1}) > 0$. Hence,

$$\begin{aligned} H(y^{l+1}) &\geq \max_{1 \leq i \leq p} (f_i(y^{l+1}) - f_i(x^{k_{last}})) > -m\delta_l = c^+(x^{k_{last}}) - m\delta_l; \\ H(y^{l+1}) &\geq c(y^{l+1}) > c^+(x^{k_{last}}) - m\delta_l. \end{aligned}$$

(ii) If $c(x^{k_{last}}) > 0$, it has $c(y^{l+1}) > c(x^{k_{last}}) - m\delta_l$, $c^+(x^{k_{last}}) = c(x^{k_{last}})$, $\lim_{l \rightarrow \infty} \delta_l = 0$.

Then $H(y^{l+1}) \geq c(y^{l+1}) > c^+(x^{k_{last}}) - m\delta_l$ holds. Passing the limit to the above inequality and we get $H(\bar{y}) \geq c^+(x^{k_{last}})$. Moreover, from $H(\bar{y}) + \frac{\eta}{2} \|\bar{y} - x^{k_{last}}\|^2 \leq c^+(x^{k_{last}})$, we obtain that $\bar{y} = x^{k_{last}}$, i.e., $0 \in \partial H(x^{k_{last}})$. According to Theorem 3.1, $x^{k_{last}}$ is the weakly Pareto solution of (MOP). The proof of the theorem is finished. \square

3.2 The penalized infeasible multi-objective proximal bundle method

In the improvement function of the IMPB algorithm, we use the functional value of testing points to approximate the optimal value of the objective function. This approach of the approximation is slow. The reason is that, the feasibility of the testing point in current iteration is not considered, and the number of iterations of the algorithm is increased. On the other hand, this approximation may make the algorithm fall into the infeasible local optimum, hence, the feasible optimality solution of the problem is hardly reached and the algorithm is invalid. For the above reasons, we make an improvement of the IMPB method. In the new algorithm, a penalty item is added into the improvement function, which considers the feasibility of testing points. The new algorithm is called the penalized infeasible multi-objective proximal bundle (PIMPB) method. In PIMPB algorithm, the feasibility of the current iteration point is considered, which not only reduces the risk of iteration points falling into the infeasible local optimal point, but also speeds up the process of approximating the optimal value of objective functions.

The new improvement function is $H : R^n \times R^n \rightarrow R$,

$$H_{x^k}(y) = \max\{f_i(y) - \theta_{i1}, c(y) - \theta_2 \mid i = 1, 2, \dots, p\}, \tag{3.13}$$

where $\theta_{i1} = f_i(x^k) + s_k \max\{0, c(x^k)\}$, $\theta_2 = t_k \max\{0, c(x^k)\}$. The PIMPB algorithm is divided into two phases. The first phase is the infeasible iteration, in this phase, the iteration point is infeasible. The second phase is the feasible iteration, in this phase, the iteration point is feasible. Under this feasibility, the function value is decreased until the stopping criterion is satisfied.

Compared with the improvement function in the IMPB algorithm, the improvement function in the PIMPB algorithm is to subtract two penalty terms from the former. The penalty term only works in the first phase. In this phase, when the infeasibility of the constraint function is reduced, the objective function is allowed to have a slight increment, which can make the iteration point avoid becoming the infeasible local optimality solution. When the algorithm enters the second phase, the penalty term is zero, and the improvement function

is the same as that in the IMPB algorithm. In addition, in numerical experiments, it is found that adding a penalty term can speed up the convergence of the algorithm.

In the following, we present the model of the PIMPB algorithm. The piecewise approximation models of objective functions and constraint function are the same as that in the IMPB algorithm. The new improvement function is

$$\begin{aligned} \check{H}_l(y) &= \max\{\check{f}_{il}(y) - \theta_{il}, \check{c}_l(y) - \theta_2 \mid i = 1, 2, \dots, p\} \\ &= \max\{\check{f}_{il}(y) - f_i(x^k) - s_k c^+(x^k), \check{c}_l(y) - t_k c^+(x^k) \mid i = 1, 2, \dots, p\}. \end{aligned} \tag{3.14}$$

Solve the following quadric programming problem

$$\min_{y \in R^n} \check{H}_l(y) + \frac{u_l}{2} \|y - x^k\|^2 \tag{3.15}$$

to obtain the candidate point y^{l+1} . The equivalent problem of (3.15) is

$$\begin{aligned} \min \quad & r + \frac{u_l}{2} \|y - x^k\|^2 \\ \text{s. t.} \quad & -s_k c^+(x^k) - \hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle \leq r, \quad j \in L_l^{f_i}, i = 1, 2, \dots, p, \\ & -t_k c^+(x^k) - \hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle \leq r, \quad j \in L_l^c. \end{aligned} \tag{3.16}$$

The following theorem gives the formula of the candidate point.

Theorem 3.4 *Suppose that y^{l+1} is the solution of (3.16), then*

$$y^{l+1} = x^k - \frac{1}{u_l} G^l, \tag{3.17}$$

where $G^l = \sum_{i=1}^p \sum_{j \in L_l^{f_i}} \lambda_{ij} (g_{f_i}^j + \eta_l \Delta_j^k) + \sum_{j \in L_l^c} u_j (g_c^j + \eta_l \Delta_j^k)$. Moreover, the following conclusions hold.

- (i) $G^l \in \partial \check{H}_l(y^{l+1})$;
- (ii) $\varepsilon_l^k = \sum_{i=1}^p \sum_{j \in L_l^{f_i}} \lambda_{ij} \hat{e}_{f_{ij}}^k + \sum_{j \in L_l^c} u_j \hat{e}_{c_j}^k$, and $\varepsilon_l^k \geq 0$.

Proof First we prove that (3.17) holds. The Lagrange function of (3.16) is

$$\begin{aligned} L(y, r, \lambda, u) &= r + \frac{u_l}{2} \|y - x^k\|^2 + \sum_{i=1}^p \sum_{j \in L_l^{f_i}} \lambda_{ij} [-s_k c^+(x^k) - \hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle - r] \\ &\quad + \sum_{j \in L_l^c} u_j [c(x^k) - t_k c^+(x^k) - \hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle - r]. \end{aligned}$$

Since the problem (3.16) is strongly convex, the optimal solution is unique and the dual gap is zero, then

$$\min_{y,r} \max_{\lambda,u} L(y, r, \lambda, u) = \min_{\lambda,u} \max_{y,r} L(y, r, \lambda, u).$$

From $\frac{\partial L}{\partial r} = 0$, it has $\sum_{i=1}^p \sum_{j \in L_l^{f_i}} \lambda_{ij} + \sum_{j \in L_l^c} u_j = 1$. Solve the following problem

$$\begin{aligned} \min_y \quad & r + \frac{u_l}{2} \|y - x^k\|^2 + \sum_{i=1}^p \sum_{j \in L_l^{f_i}} \lambda_{ij} [-s_k c^+(x^k) - \hat{e}_{f_{ij}}^k + \langle g_{f_i}^j + \eta_l \Delta_j^k, y - x^k \rangle] \\ & + \sum_{j \in L_l^c} u_j [c(x^k) - t_k c^+(x^k) - \hat{e}_{c_j}^k + \langle g_c^j + \eta_l \Delta_j^k, y - x^k \rangle] \end{aligned}$$

to obtain y^{l+1} , taking the gradient on y , we obtain that $0 = u_l(y^{l+1} - x^k) + G^l$, which is $y^{l+1} = x^k - \frac{1}{u_l}G^l$.

Next we prove the assertions (i) and (ii).

(i) Since y^{l+1} is the optimal solution of (3.15), it has $0 \in \partial \check{H}_l(y^{l+1}) + u_l(y^{l+1} - x^k)$, which is $G^l \in \partial \check{H}_l(y^{l+1})$, and the first assertion holds.

(ii) Bring y^{l+1} in the Lagrange function $L(y, r, \lambda, u)$ and obtain

$$L(y^{l+1}, r, \lambda, u) = \frac{u_l}{2} \|y^{l+1} - x^k\|^2 + \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} [-s_k c^+(x^k) - \hat{e}_{f_{ij}^k}] + \sum_{j \in L_i^c} u_j [c(x^k) - t_k c^+(x^k) - \hat{e}_{c_j^k}] - \frac{1}{u_l} \|G^l\|^2.$$

On the other hand, $L(y^{l+1}, \lambda, u) = \check{H}_l(y^{l+1}) + \frac{u_l}{2} \|y^{l+1} - x^k\|^2$, from the definition of (3.4), we have

$$\sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} \hat{e}_{f_{ij}^k} + \sum_{j \in L_i^c} u_j \hat{e}_{c_j^k} = H_{x^k}(x^k) - \check{H}_l(y^{l+1}) - u_l \|y^{l+1} - x^k\|^2.$$

Due to the aggregate linearization errors and $\hat{e}_{f_{ij}^k} \geq 0, \hat{e}_{c_j^k} \geq 0$, it holds that $\varepsilon_l^k = \sum_{i=1}^p \sum_{j \in L_i^{f_i}} \lambda_{ij} \hat{e}_{f_{ij}^k} + \sum_{j \in L_i^c} u_j \hat{e}_{c_j^k} \geq 0$. The second assertion is proven. \square

The aggregate linearization function is $\psi_l(y) = \check{H}_l(y^{l+1}) + \langle G^l, y - y^{l+1} \rangle$. The function $\psi_l(y)$ is affine and $\psi_l(y^{l+1}) = \check{H}_l(y^{l+1})$, $G^l = \nabla \psi_l(y)$, hence, we obtain $\psi_l(y) \leq \check{H}_l(y), \forall y \in R^n$. The aggregate linearization error is

$$\varepsilon_l^k = H_l(x^k) - \psi_l(y^{l+1}) = H_k(x^k) - \check{H}_l(y^{l+1}) - u_l \|y^{l+1} - x^k\|^2.$$

By the definition of subdifferentials, (3.17) and $\psi_l(y) \leq \check{H}_l(y)$, we have

$$H_{x^k}(x^k) - \varepsilon_l^k \leq \check{H}_l(y) + \langle G^k, x^k - y \rangle. \tag{3.18}$$

The predicted decrease is

$$\delta_l = H_{x^k}(x^k) - \check{H}_l(y^{l+1}) = \varepsilon_l^k + u_l \|y^{l+1} - x^k\|^2. \tag{3.19}$$

Now we propose the PIMPB method.

The PIMPB method

Step 0 (Initialization) Choose a infeasible starting point $x^0 \in C$ and a stopping tolerance τ_{stop} . Select a convex parameter Γ_0 , a proximal parameter Γ_1 , a penalty parameter Γ_2 and an improvement parameter m . Set $y_0^0 = x^0$ and compute $g_{f_i}^0 \in \partial f_i(y_0^0), i = 1, 2, \dots, p, g_c^0 \in \partial c(y_0^0)$. Initialize $u_1, \eta_1, e_{f_{i_0}}^0 = e_{c_0}^0 = d_0^0 = \Delta_0^0 = 0$.

Step 1 (Computing candidate points) Establish the piecewise linearization approximation model (3.15), and solve its dual problem (3.16) to obtain the dual parameters λ, u . Compute y^{l+1} by (3.17) and δ_l by (3.19).

Step 2 (Stopping test) If $\delta_l \leq \tau_{stop}$, then stop; otherwise, go to the next step.

Step 3 (Updating the bundle) Update the index set $L_{l+1}^{f_i} \supseteq \{l + 1, i_k\}$, $L_{l+1}^c \supseteq \{l + 1, i_k\}$. Compute $g_{f_i}^{l+1} \in \partial f_i(y^{l+1})$, $i = 1, 2, \dots, p$, $g_c^{l+1} \in \partial c(y^{l+1})$. Compute the bundle of information:

$$\begin{aligned} e_{f_i, l+1}^k &= f_i(x^k) - f_i(y^{l+1}) - \langle g_{f_i}^{l+1}, x^k - y^{l+1} \rangle, j \in L_{l+1}^{f_i}, \\ e_{c, l+1}^k &= c(x^k) - c(y^{l+1}) - \langle g_c^{l+1}, x^k - y^{l+1} \rangle, j \in L_{l+1}^c. \end{aligned}$$

Step 4 (Testing the descent criterion) If the candidate point y_k^i is good enough, i.e.,

$$\begin{cases} \max_{1 \leq i \leq p} (f_i(y^{l+1}) - f_i(x^k)) \leq -m\delta_l \text{ and } c(y^{l+1}) \leq 0, & \text{if } c(x^k) \leq 0; \\ c(y^{l+1}) \leq c(x^k) - m\delta_l, & \text{if } c(x^k) > 0. \end{cases} \tag{3.20}$$

then update the stability center, and set $x^{k+1} = y^{l+1}$; otherwise, set $x^{k+1} = x^k$.

Step 5 (Updating the convexification parameter) Compute η_{l+1}^{min} by (3.11), and set

$$\begin{cases} \eta_{l+1} = \eta_l, & \text{if } \eta_{l+1}^{min} \leq \eta_l; \\ \eta_{l+1} = \Gamma_0 \eta_{l+1}^{min}, & \eta_{l+1}^{min} > \eta_l. \end{cases} \tag{3.21}$$

Step 6 (Updating the proximal parameter) If y^{l+1} is the descent point, then set $u_{k+1} \leq u_{max} < +\infty$; otherwise, set $u_{l+1} \in [u_l, u_{max}]$.

Step 7 (Updating the penalty parameter) Choose the penalty parameter Γ_2 satisfying

$$0 \leq t_k \leq 1, s_k \geq 0, s_k - t_k \geq \Gamma_2 > 0, \tag{3.22}$$

and set $k = k + 1$, and go to Step 1.

Now we make the comparison between IMPB, PIMPB algorithms and other algorithms given in [14] and [19]. The main differences among them are as follows:

- (i) The first difference is the choice of the initial solution. The initial point of [14] and [19] is feasible. The initial point of this paper is arbitrary and it can be infeasible.
- (ii) The second difference is the descent criterion. Since the initial point of the algorithm is arbitrary, it needs to consider the feasibility of the testing points in the descent criterion. In the descent criterion, when the current stability center is feasible, the decrease of objective functions and the feasibility of the candidate point are considered. When the current stability center is infeasible, it needs to have a sufficient decrease for the constraint function.
- (iii) The third difference is the improvement function. To further guarantee the feasibility of the testing points, the penalty procedure is added to the improvement function. It not only reduces the risk of iteration points falling into the infeasible local optimum, but also speeds up the process of approximating the optimal value of objective functions.

Next we present several lemmas before giving the convergence of the PIMPB algorithm.

Lemma 3.4 *The penalty parameter is updated according to (3.22), and the following conclusions hold.*

- (i) *If $c(x^k) \leq 0$, it has $H_{x^k}(x^k) = 0$. If $c(x^k) > 0$, it has $H_{x^k}(x^k) = c(x^k)(1 - t_k)$.*
- (ii) *When k is the index of null steps, it holds $H_{x^k}(y^{l+1}) > H_{x^k}(x^k) - m\delta_k$.*

Proof (i) If $c(x^k) \leq 0$, it has $H_{x^k}(x^k) = 0$. If $c(x^k) > 0$, it follows from (3.22) that $H_{x^k}(x^k) = c(x^k)(1 - t_k) \geq 0$ and (i) is proved.

(ii) When k is the index of null steps, the descent criterion (3.19) doesn't hold. We prove the conclusion from the following two cases. The first case is $c(x^k) > 0$, and it has

$$\begin{aligned} H_{x^k}(y^{k+1}) &= \max\{f_i(y^{k+1}) - f_i(x^k) - s_k c(x^k), c(y^{k+1}) - t_k c(x^k)\} \\ &\geq c(y^{k+1}) - t_k c(x^k) \\ &\geq c(x^k) - m\delta_k - t_k c(x^k) \\ &= (1 - t_k)c(x^k) - m\delta_k \\ &= H_{x^k}(x^k) - m\delta_k. \end{aligned}$$

The second case is $c(x^k) \leq 0$, then $\max_{1 \leq i \leq p} \{f_i(y^{k+1}) - f_i(x^k)\} \geq -m\delta_k$ and $c(y^{k+1}) > 0$, hence,

$$\begin{aligned} H_{x^k}(y^{k+1}) &= \max_{1 \leq i \leq p} \{f_i(y^{k+1}) - f_i(x^k), c(y^{k+1})\} \\ &\geq \max_{1 \leq i \leq p} \{f_i(y^{k+1}) - f_i(x^k)\} \\ &\geq -m\delta_k \\ &= H_{x^k}(x^k) - m\delta_k. \end{aligned}$$

So we obtain $H_{x^k}(y^{k+1}) \geq c(y^{k+1}) > 0 \geq H_{x^k}(x^k) - m\delta_k$, and (ii) is proved. □

Lemma 3.5 *Suppose that the PIMPB algorithm generates an infinite sequence of serious steps, and L_s is the index set of serious steps, then*

- (i) $\lim_{k \rightarrow \infty} \delta_k = 0, \lim_{k \rightarrow \infty} \varepsilon_l^k = 0, \lim_{k \rightarrow \infty} \|G^k\| = 0, k \in L_s.$
- (ii) $H_{x^k}(x^k) \leq \check{H}_k(y) + o(1/k), k \in L_s$ when k is large enough.

Proof (i) The proof is similar with Lemma 3.3 and we omit the proof.

- (ii) Since $\lim_{k \rightarrow \infty} \varepsilon_l^k = 0$, from (3.18), we obtain that $H_{x^k}(x^k) \leq \check{H}_k(y) + o(1/k), k \in L_s$, when k is large enough. □

Lemma 3.6 *Suppose that the PIMPB algorithm generates an infinite sequence of null steps, and L'_s is the index set of null steps. $\hat{x} = x^{\hat{k}}$ is the last serious step, and there are all null steps after \hat{k} steps. Then*

- (i) $\lim_{k \rightarrow \infty} \delta_k = 0, \lim_{k \rightarrow \infty} \varepsilon_l^k = 0, \lim_{k \rightarrow \infty} \|G^k\| = 0, k \in L'_s;$
- (ii) when $k \in L'_s$ is large enough, it has $y^k \rightarrow \hat{x};$
- (iii) $H_{x^k}(x^k) \leq \check{H}_k(y) + o(1/k), k \in L'_s$ and k is large enough.

Proof (i) From (3.17), it has $y^{k+1} = \hat{x} - \frac{1}{u_k} G^k, G^k = u_k(\hat{x} - y^{k+1})$, then

$$\begin{aligned} 2\langle G^k, y - y^{k+1} \rangle &= 2u_k \langle \hat{x} - y^{k+1}, y - y^{k+1} \rangle \\ &= u_k \langle \hat{x} - y^{k+1}, \hat{x} - y^{k+1} + y - \hat{x} \rangle + u_k \langle y - y^{k+1} + \hat{x} - y, y - y^{k+1} \rangle \\ &= u_k \|y^{k+1} - \hat{x}\|^2 + u_k \|y - y^{k+1}\|^2 - u_k \|y - \hat{x}\|^2. \end{aligned}$$

By the aggregate linearization function, it yields

$$\begin{aligned} \psi_k(y) &= \check{H}_k(y^{k+1}) + \langle G^k, y - y^{k+1} \rangle \\ &= \check{H}_k(y^{k+1}) + \frac{u_k}{2} \|y^{k+1} - \hat{x}\|^2 + \frac{u_k}{2} \|y - y^{k+1}\|^2 - \frac{u_k}{2} \|y - \hat{x}\|^2. \end{aligned}$$

Denote $V_k = \check{H}_k(y^{k+1}) + \frac{u_k}{2} \|y^{k+1} - \hat{x}\|^2$, where V_k is the optimal value of (3.15), then

$$\psi_k(y) + \frac{u_k}{2} \|y - \hat{x}\|^2 = V_k + \frac{u_k}{2} \|y - y^{k+1}\|^2.$$

When $y = \hat{x}$, there exists $\hat{M} > 0$ satisfying $V_k + \frac{u_k}{2} \|\hat{x} - y^{k+1}\|^2 = \psi_k(\hat{x}) \leq \check{H}_k(\hat{x}) \leq \hat{M}$, then V_k has the upper bound. Due to $\check{H}_{k+1}(y) \geq \psi_k(y)$, for $\forall y \in R^n$, it holds $\check{H}_{k+1}(y) + \frac{u_k}{2} \|y - \hat{x}\|^2 \geq V_k + \frac{u_k}{2} \|y - y^{k+1}\|^2$. In a null step, from Step 6 of the PIMPB algorithm, it has $u_{k+1} \geq u_k$. Taking $y = y^{k+2}$, it holds $V_{k+1} \geq V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2$, and $\{V_k\}$ is nondecreasing, then $\{V_k\}$ is convergent and $V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2 \rightarrow 0$. From Lemma 3.4, it has $H_{x^k}(y^{k+1}) > H_{x^k}(x^k) - m\delta_k$. Adding δ_k to the above inequality, we obtain

$$\begin{aligned} 0 &\leq (1 - m)\delta_k < \delta_k + H_{x^k}(y^{k+1}) - H_{x^k}(x^k) \\ &= H_{x^k}(y^{k+1}) - \hat{H}_k(y^{k+1}) \\ &= H_{x^k}(y^{k+1}) - \hat{H}_k(y^{k+2}) + \hat{H}_k(y^{k+2}) - \hat{H}_k(y^{k+1}). \end{aligned}$$

Furthermore,

$$\begin{aligned} \check{f}_{i,k+1}(y^{k+2}) &\geq f_i(\hat{x}) - \hat{e}_{f_{i,k+1}}^{k+1} + \langle g_{f_{i,k+1}} + \eta_{k+1}(y^{k+1} - \hat{x}), y^{k+2} - \hat{x} \rangle \\ &= f_i(x^{k+1}) + \langle g_{f_{i,k+1}}, \hat{x} - y^{k+1} \rangle - \frac{\eta_{k+1}}{2} \|\hat{x} - y^{k+1}\|^2 \\ &\quad + \langle g_{f_{i,k+1}} + \eta_{k+1}(y^{k+1} - \hat{x}), y^{k+2} - \hat{x} \rangle \\ &\geq f_i(x^{k+1}) + \langle g_{f_{i,k+1}}, y^{k+2} - y^{k+1} \rangle + \eta_{k+1} \langle y^{k+1} - \hat{x}, y^{k+2} - y^{k+1} \rangle \\ &= f_i(x^{k+1}) + \langle g_{f_{i,k+1}} + \eta_{k+1} \Delta_{k+1}^k, y^{k+2} - y^{k+1} \rangle. \end{aligned}$$

Similarly, we have $\check{c}_{k+1}(y^{k+2}) \geq c(x^{k+1}) + \langle g_{c,k+1} + \eta_{k+1} \Delta_{k+1}^k, y^{k+2} - y^{k+1} \rangle$. Hence, it holds that

$$\begin{aligned} \check{H}_{k+1}(y^{k+2}) &= \max\{\check{f}_{i,k+1}(y^{k+2}) - \theta_{i1}, \check{c}_{k+1}(y^{k+2}) - \theta_{2}\} \\ &\geq \max\{f_i(x^{k+1}) + \langle g_{f_{i,k+1}} + \eta_{k+1} \Delta_{k+1}^k, y^{k+2} - y^{k+1} \rangle - \theta_{i1}, c(x^{k+1}) \\ &\quad + \langle g_{c,k+1} + \eta_{k+1} \Delta_{k+1}^k, y^{k+2} - y^{k+1} \rangle - \theta_{2}\}. \end{aligned}$$

Because $\{y^k\}$ is bounded, there exists $N > 0, \forall k, \|\Delta_{k+1}^k\| \leq N$. $f_i, i = 1, 2, \dots, p, c$ are Lipschitz functions, then it exists $L > 0$ satisfying $\|g_{f_{i,k+1}}\| \leq L, \|g_{c,k+1}\| \leq L$. According to the Cauchy–Schwarz inequality, it has

$$\check{H}_{k+1}(y^{k+2}) \geq H_k(y^{k+1}) - (L + \eta_{k+1}N) \|y^{k+2} - y^{k+1}\|,$$

then

$$\begin{aligned} 0 &\leq (1 - m)\delta_k < \check{H}_{k+1}(y^{k+2}) - H_k(y^{k+1}) + (L + \eta_{k+1}N) \|y^{k+2} - y^{k+1}\| \\ &= V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - \hat{x}\|^2 + \frac{u_k}{2} \|y^{k+1} - \hat{x}\|^2 \\ &\quad + (L + \eta_{k+1}N) \|y^{k+2} - y^{k+1}\| \\ &= \langle y^{k+2} - y^{k+1}, G^k \rangle + (L + \eta_{k+1}N) \|y^{k+2} - y^{k+1}\| \\ &\quad + V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2 \\ &\leq (L + \eta_{k+1}N + \|G^k\|) \|y^{k+2} - y^{k+1}\| + V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2 \\ &\leq (2L + \eta_{k+1}N) \|y^{k+2} - y^{k+1}\| + V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2. \end{aligned}$$

Due to $V_{k+1} - V_k - \frac{u_k}{2} \|y^{k+2} - y^{k+1}\|^2 \rightarrow 0, \|y^{k+2} - y^{k+1}\| \rightarrow 0$, it has $\lim_{k \rightarrow \infty} \delta_k = 0$.

Because $0 \leq \varepsilon_l^k \leq \delta_k$, it has that $\lim_{k \rightarrow \infty} \varepsilon_l^k = 0$.

Furthermore, $\delta_k \geq \varepsilon_l^k + \frac{1}{u_{max}} \|G^k\|^2 \geq 0$, it holds that $\lim_{k \rightarrow \infty} \|G^k\| = 0$, from (3.18), we obtain $H_{x^k}(x^k) \leq \check{H}_k(y) + o(1/k), k \in L'_s$ when k is large enough. □

Now we give the convergence result of the PIMPB algorithm.

Theorem 3.5 *Let $f_i, i = 1, 2, \dots, p$ and g be f° -pseudoconvex and the constraint qualification (2.1) be valid.*

The penalty parameter of the PIMPB algorithm is updated according to (3.22). The PIMPB algorithm generates the sequence $\{x^k\}$, denoting $\{x^k\} \rightarrow \bar{x}, \{s_k\} \rightarrow \bar{s}, \{t_k\} \rightarrow \bar{t}$, then the following conclusions hold.

- (i) $\forall y \in R^n$, it has $\max\{\bar{c}, 0\}(1 - \bar{t}) \leq \max_{1 \leq i \leq p} \{f_i(y) - \bar{f}_i - \bar{s} \max\{\bar{c}, 0\}, c(y) - \bar{t} \max\{\bar{c}, 0\}\}$.
- (ii) If $\bar{c} > 0$, it exists the constant R_0 satisfying $\bar{x} = \arg \min_{y \in R^n} c(y)$ when $\bar{s} \geq R_0$.
- (iii) If $\bar{c} \leq 0$, \bar{x} is the weakly Pareto solution of (MOP).

Proof Suppose that the PIMPB algorithm generates the infinite loops, denoting by L_s, L'_s the infinite serious steps and null steps, from Lemmas 3.4–3.6, there exists the subsequence K satisfying $\{x^k\}_{k \in K} \rightarrow \bar{x}$. When $K \subseteq L'_s$, it has $\{y^k\} \rightarrow \bar{x}$, and the stability point $x^k = \bar{x}$ is unchanged. Moreover, it holds that

$$\forall y \in R^n, k \in K, H_{x^k}(x^k) = \max\{c(x^k), 0\}(1 - t_k) \leq \check{H}_k(y) + o(1/k).$$

Because $\check{H}_k(y) \leq H_k(y)$, we have

$$\max\{c(x^k), 0\}(1 - t_k) \leq \max_{1 \leq i \leq p} \{f_i(y) - \theta_{i1}, c(y) - \theta_2\} + o(1/k), \tag{3.23}$$

where $\theta_{i1} = f_i(x^k) + s_k \max\{c(x^k), 0\}, \theta_2 = t_k \max\{c(x^k), 0\}$. When $\{x^k\}_{k \in K} \rightarrow \bar{x}$, denoting $\bar{f}_i = \lim_{k \in K} f_i(x^k), \bar{c} = \lim_{k \in K} c(x^k), \bar{s} = \lim_{k \in K} s_k, \bar{t} = \lim_{k \in K} t_k$, pass the limit to (3.23) and obtain

$$\forall y \in R^n, \max\{\bar{c}, 0\}(1 - \bar{t}) \leq \max_{1 \leq i \leq p} \{f_i(y) - \bar{f}_i - \bar{s} \max\{\bar{c}, 0\}, c(y) - \bar{t} \max\{\bar{c}, 0\}\}.$$

The proof of (i) is finished.

(ii) When $\bar{c} > 0$, take $R_0 = \frac{1}{\bar{c}} \{ \max_{1 \leq i \leq p} \{f_i(y) - \bar{f}_i\} - c(y) \} + \bar{t}$. If $\bar{s} \geq R_0$, it has

$$(\bar{s} - \bar{t})\bar{c} > \max_{1 \leq i \leq p} \{f_i(y) - \bar{f}_i\} - c(y) \Leftrightarrow f_i(y) - \bar{f}_i - \bar{s} \max\{\bar{c}, 0\} < c(y) - \bar{t} \max\{\bar{c}, 0\}.$$

From (i), it holds $\bar{c} \leq c(y), \forall y \in R^n$, which is $\bar{x} = \arg \min_{y \in R^n} c(y)$, and (ii) holds. (iii) When $\bar{c} \leq 0$, we have $H_{\bar{x}}(\bar{x}) = 0 \leq \max_{1 \leq i \leq p} \{f_i(y) - \bar{f}_i, c(y)\}, \forall y \in R^n$, so it has $0 \in \partial H_{\bar{x}}(\bar{x})$, by Lemma 3.4, we know that \bar{x} is the weakly Pareto solution of (MOP). □

4 Numerical Results

In this section we give some numerical experiments to test the IMPB algorithm and the PIMPB algorithm for (MOP). The algorithms are completed in MATLAB 9.2. The subproblem is solved by quadratic-programming QuadProg.m which is available in the MATLAB optimization toolbox.

Example 1 The multiobjective optimization problem is defined by

$$\begin{aligned} \min \quad & f_1(x) = \sqrt{\|x\| + 2} \\ & f_2(x) = \max\{-x_1 - x_2, -x_1 - x_2 + x_1^2 + x_2^2 - 1\} \\ \text{s.t.} \quad & g(x) = \max\{x_1^2 + x_2^2 - 10, 3x_1 + x_2 + 1.5\} \leq 0, \end{aligned}$$

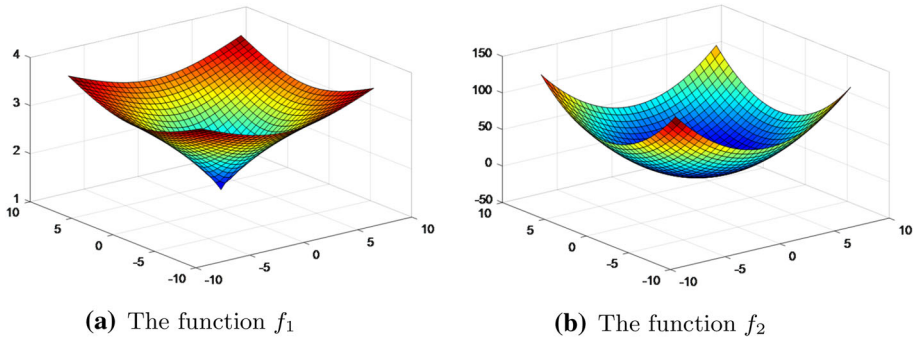


Fig. 1 The mulobjective function

where $x = (x_1, x_2)^T \in R^n$. The function f_1 is f° -pseudoconvex, f_2 is convex and the constraint function g is also convex.

Figure 1 gives the images of objectives functions f_1 and f_2 .

Since the IMPB algorithm has no requirement on the feasibility of the starting point, we choose the infeasible starting point $x^0 = (1, 1)^T$.

The parameters of the algorithm are given as follows: the increasing parameter $M_0 = 8$, the convex parameter $\Gamma_0 = 1.009$, the proximal parameter $\Gamma_1 = 1.005$, the stopping tolerance $\tau_{stop} = 10^{-4}$, the improvement parameter $m = 0.1$, the convex parameter $\eta_1 = 1$, the proximal parameter $\mu_1 = 1$, and the constraint tolerance is 10^{-3} .

Figure 2 presents the variation of the constrained function during the iteration of the algorithm. From Fig. 2, it can be seen that the initial point is infeasible, and it leads to some initial iteration being infeasible.

When the iteration points are infeasible, the amplitude of the value of the constraint function is relatively large.

After nine iterations, the algorithm achieves the feasible point and executes the feasible solution mode.

The amplitude of the constraint function value is relatively small, and the Pareto optimal solution is found on the basis of keeping the feasibility.

Figure 3 describes the changes of two objective functions during the iteration of the algorithm.

From Fig. 3, it can be seen that the two objective functions are non-monotonic. It is reasonable because the initial point is not feasible, and later some initial iteration are also not feasible. At this time, the values of functions are not necessarily reduced. The IMPB algorithm pays more attention to increasing the functional value to reduce the infeasibility, once the feasible point is reached, the values of two objective functions are reduced while the feasibility is maintained.

Table 1 shows that, at the same stopping criterion, after six iterations, the multi-objective proximal bundle(MPB) method provided by Karmitsa [14] reaches the Pareto optimal solution $x^* = (-0.4620, -0.1138)^T$, and the corresponding optimal value is $f^* = (1.5735, 0.5759)^T$. The IMPB algorithm reaches the Pareto optimal solution $x^* = (-0.4486, -0.1544)^T$, the optimal value $f^* = (-0.4486, -0.1544)^T$ and the constraint function $g^* = -3.4931 \times 10^{-4}$ after 33 iterations.

The IMPB algorithm is an infeasible algorithm and the initial point can be not feasible. Karmitsa’s MPB algorithm is a feasible algorithm, and it requires the feasibility of initial

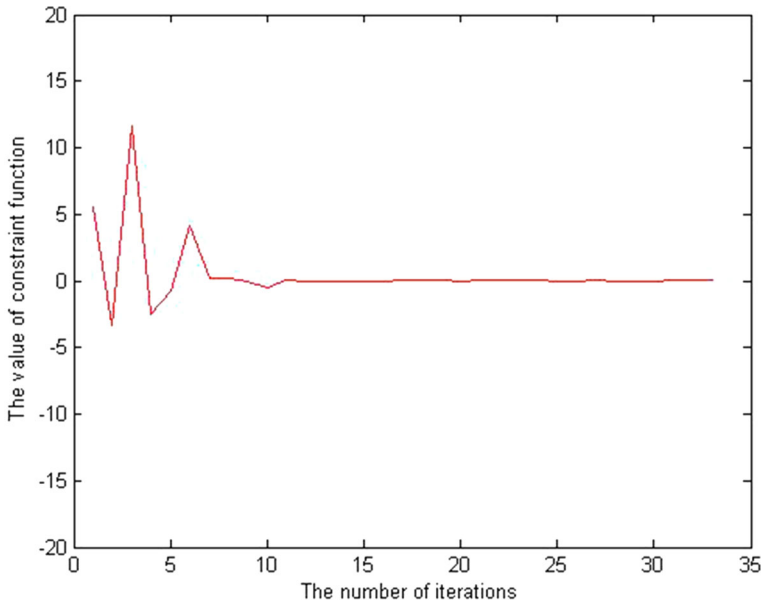


Fig. 2 The constraint function

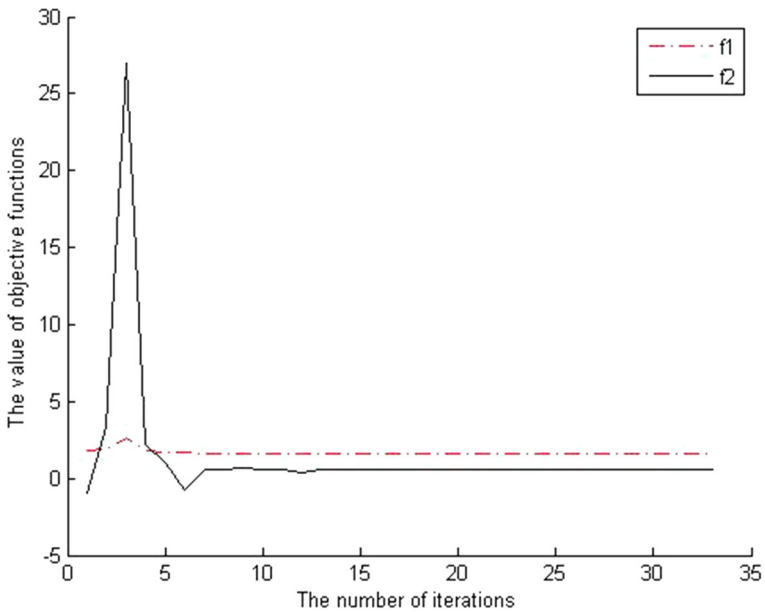


Fig. 3 The variation of the objective function

points and all the iterations need to be feasible. From this aspect, the number of iterations of the IMPB algorithm is more than that of Karmitsa’s MPB algorithm.

To increase the speed of the IMPB algorithm, an acceleration procedure, i.e., the penalty procedure is added to the IMPB algorithm, which is called the PIMPB algorithm. In the

Table 1 The comparison of MPB and IMPB

	Iterations	Optimal solution	Optimal value
MPB	6	(− 0.4620, − 0.1138)	(1.5735, 0.5759)
IMPB	33	(− 0.4486, − 0.1544)	(1.5730, 0.6031)

Table 2 The comparison of IMPB and PIMPB for Example 1

	Iterations	Optimal solution	Optimal value
IMPB	33	(− 0.4486, − 0.1544)	(1.5730, 0.6031)
PIMPB	20	(− 0.4508, − 0.1482)	(1.5731, 0.5990)

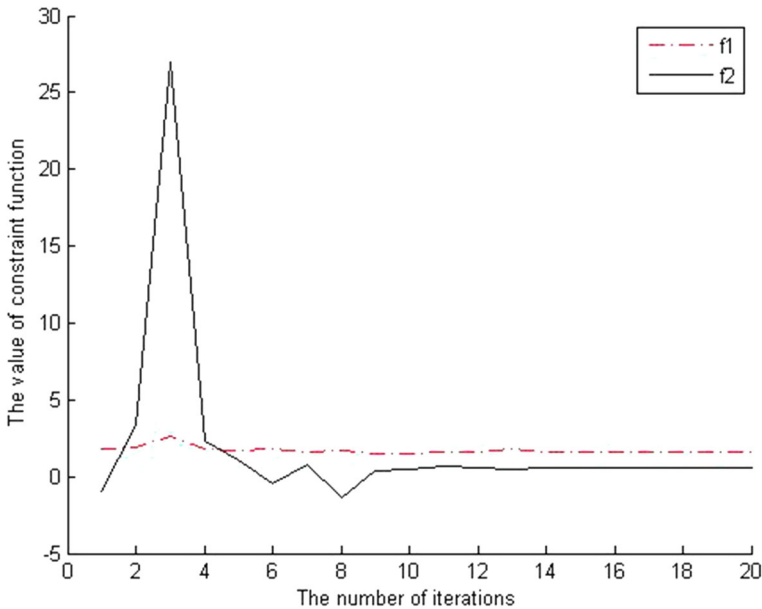


Fig. 4 The variation of the objective function

PIMPB algorithm, set the penalty parameter $s_0 = 1$, the increasing penalty parameter $\Gamma_2 = 1.8$, and the other parameters are set the same with that in the IMPB algorithm.

Figures 4 and 5 describe the changes of the values of objective functions and the constraint function by using the PIMPB algorithm. Because that the penalty term of constraint function is added to the improvement function, the convergent speed of the PIMPB algorithm is obviously faster than that of the IMPB algorithm. It is follows from Table 2 that the IMPB algorithm performs 33 steps to achieve the stopping accuracy, while the PIMPB algorithm only needs 20 steps to achieve the same precision.

Example 2 In this example we consider the Pentagon optimization problem [11]. There are three objective functions of 6-dimensions with 15 constraint functions in this nonsmooth nonconvex problem, which is

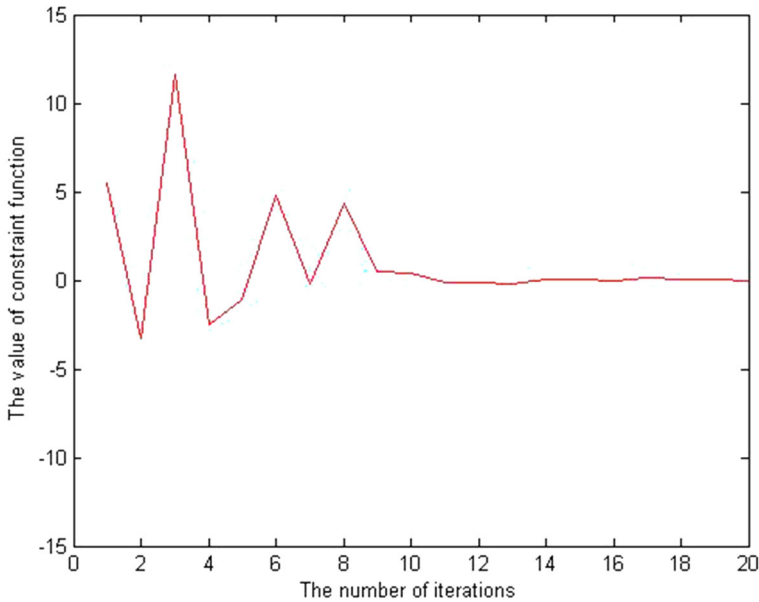


Fig. 5 The variation of the constraint function

$$\begin{aligned}
 \min f_1(x) &= -\sqrt{(x_1 - x_3)^2 + (x_2 - x_4)^2} \\
 f_2(x) &= -\sqrt{(x_1 - x_3)^2 + (x_4 - x_6)^2} \\
 f_3(x) &= -\sqrt{(x_1 - x_5)^2 + (x_2 - x_6)^2} \\
 \text{s.t. } g_{ij}(x) &= x_i \cos\left(\frac{2\pi j}{5}\right) + x_{i+1} \sin\left(\frac{2\pi j}{5}\right) - 1 \leq 0,
 \end{aligned}$$

where $x = (x_1, x_2, \dots, x_6)^T, i = 1, 2, 3; j = 0, 1, \dots, 4$. Figure 6 gives the image of the objective function f_1 .

We use the PIMPB method to solve this example and change the setting of parameters as follows: the proximal parameter $\Gamma_1 = 1.2$, the increasing penalty parameter $\Gamma_2 = 1.8$, the improvement parameter $m = 0.08$, the stopping criterion $\tau_{stop} = 10^{-3}$. Because the PIMPB method has no feasibility requirement on the starting point, we set the infeasible starting point $x^0 = (-1, 0.5, 0.6, -1, 0.9, 1.8)^T$. At this starting point, the two constraint functions $g_{12}(x^0) = 0.1029, g_{34}(x^0) = 0.1365$ are infeasible, and the other 13 constraint functions are feasible .

Figure 7 shows the change of the maximum constraint function during iterations. From Fig. 7, it can be seen that the initial implementation of the PIMPB algorithm is infeasible. When the iteration point is infeasible, the value of the constraint function is larger. Fifteen times later, the feasible point is reached, and then the feasible solution mode is executed. At this time, the amplitude of the constraint function value is relatively small, and the Pareto optimal solution is found on the basis of keeping the feasibility. Figure 8 describes the changes of three objective functions during iterations.

Table 3 provides the comparison results of IMPB and PIMPB for this problem. The number of iterations by the PIMPB method is 55, about almost half of the IMPB method with 98 iterations. The CPU time of the PIMPB method is much less than that of the IMPB method.

To conclude, since the IMPB and PIMPB methods are infeasible algorithms, they are performed well for these two examples. The proposed algorithms reduces the requirements

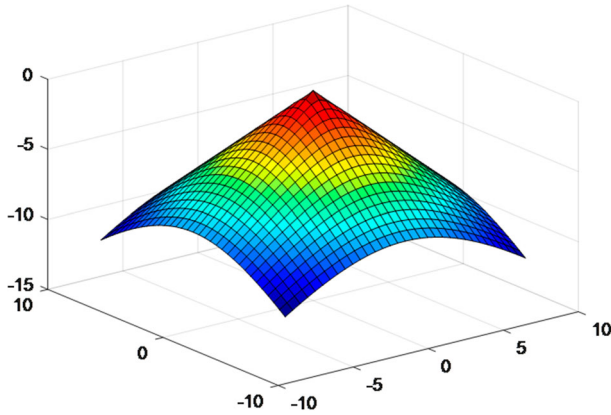


Fig. 6 The objective function of f_1

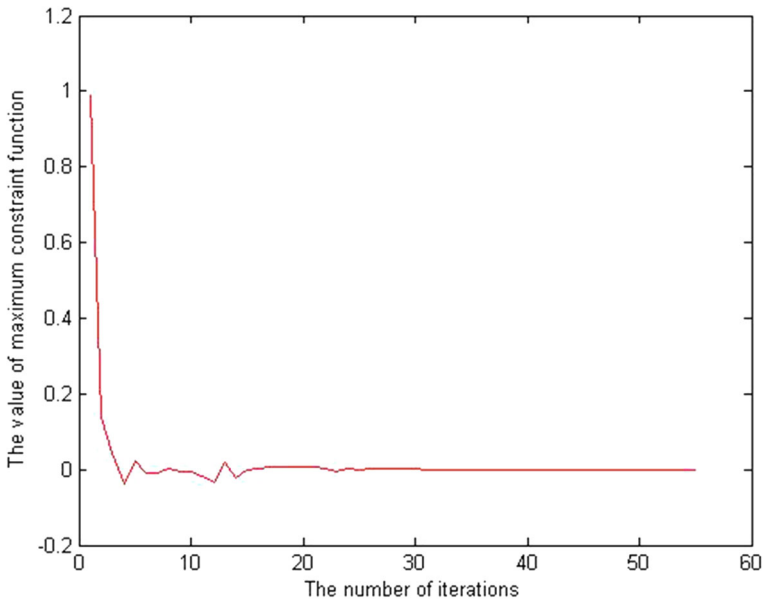


Fig. 7 The maximum constraint function

on the initial points of the problems, which decreases the computation of the algorithms. The overall behavior of the PIMPB algorithm is better than the IMPB method due to its distinctive penalty procedure, which is of great efficiency to solve constrained optimization problems.

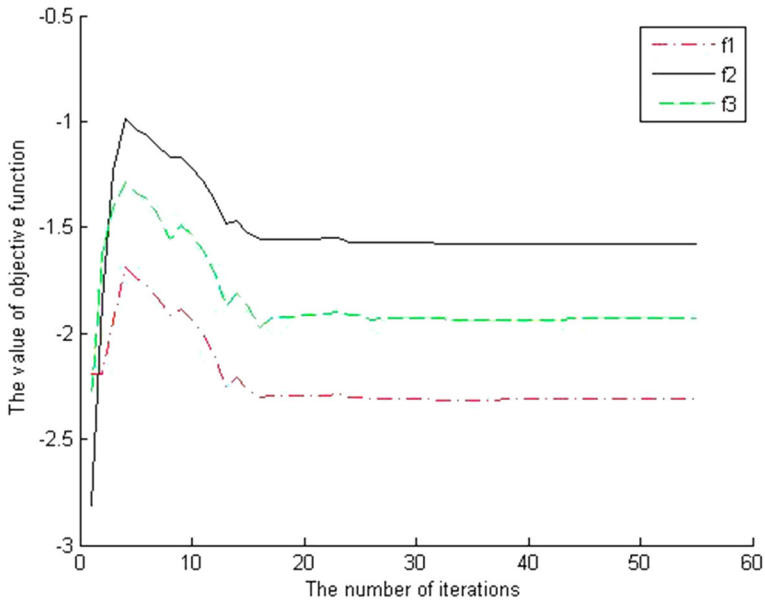


Fig. 8 The three objective functions

Table 3 The comparison of IMPB and PIMPB for Example 2

	Iterations	Optimal value	CPU time
IMPB	98	(− 2.2998, − 1.8858, − 1.7402)	157.87
PIMPB	55	(− 2.3138, − 1.5799, − 1.9387)	16.83

5 Conclusion

We have presented two infeasible proximal bundle methods for nonsmooth nonconvex multi-objective optimization problems. In the algorithms, the multiobjective functions are handled individually without employing any scalarization. The proposed algorithm has no requirement on the feasibility of initial points. The penalty procedure is given to increase the speed of the convergence of the infeasible algorithm. Under some generalized convexity assumptions, we prove that the algorithm finds the global weakly Pareto optimal solutions of the problem.

Acknowledgements The authors thank two anonymous referees for a number of valuable and helpful suggestions.

References

1. Banichuk, N.V., Neittaanmaki, P.J.: Structural Optimization with Uncertainties. Solid Mechanics and Its Applications. Springer, Berlin (2010)
2. Eichfelder, G.: Adaptive Scalarization Methods in Multiobjective Optimization. Springer, Berlin (2009)
3. Fliege, J., Ismael, A.: A method for constrained multiobjective optimization based on SQP techniques. SIAM **26**, 2091–2119 (2016)

4. Haslinger, J., Neittaanmaki, P.: *Finite Element Approximation for Optimal Shape, Material and Topology Design*. Wiley, Chichester (1996)
5. Hare, W., Sagastizbal, C., Solodov, M.: A proximal bundle method for nonsmooth nonconvex functions with inexact information. *Comput. Optim. Appl.* **63**(1), 1–28 (2016)
6. Kiwiel, K.C.: A descent method for nonsmooth convex multiobjective minimization. *Large Scale Syst.* **8**(2), 119–129 (1985)
7. Kiwiel, K.C.: An algorithm for nonsmooth convex minimization with errors. *Math. Comput.* **45**(171), 173–180 (1985)
8. Kiwiel, K.C.: A proximal bundle method with approximate subgradient linearizations. *SIAM J. Optim.* **16**(4), 1007–1023 (2006)
9. Kiwiel, K.C.: Bundle methods for convex minimization with partially inexact oracles. Technical Report, Systems Research Institute, Polish Academy of Sciences (2010)
10. Kiwiel, K.C.: Proximity control in bundle methods for convex nondifferentiable Optimization. *Math. Program.* **46**, 105–122 (1990)
11. Luksan, L., Viecek, J.: Test problems for nonsmooth unconstrained and linearly constrained optimization. Institute of Computer Science (2000)
12. Lv, J., Pang, L.P., Xu, N., Xiao, Z.H.: An infeasible bundle method for nonconvex constrained optimization with application to semi-infinite programming problems. *Numer. Algorithms* **80**, 397–427 (2019)
13. Montonen, O., Joki, K.: Bundle-based descent method for nonsmooth multiobjective DC optimization with inequality constraints. *J. Glob. Optim.* **72**, 403–429 (2018)
14. Mäkelä, M., Karitsa, N., Wilppu, O.: Proximal bundle method for nonsmooth and nonconvex multiobjective optimization. *Math. Model. Optim. Mech.* (2014). <https://doi.org/10.1007/978-3-319-23564-612>
15. Mäkelä, M.M., Eronen, V.P., Karitsa, N.: On nonsmooth optimality conditions with generalized convexities. TUCS Technical Reports 1056, Turku Centre for Computer Science, Turku (2012)
16. Mäkelä, M.M., Neittaanmki, P.: *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co., Singapore (1992)
17. Meng, F.Y., Pang, L.P., Lv, J., Wang, J.H.: An approximate bundle method for solving nonsmooth equilibrium problems. *J. Glob. Optim.* **68**, 537–562 (2017)
18. Mistakidis, E.S., Stavroulakis, G.E.: *Nonconvex Optimization in Mechanics. Smooth and Nonsmooth Algorithms, Heuristics and Engineering Applications*. Kluwer Academic Publisher, Dordrecht (1998)
19. Miettinen, K., Makela, M.M.: Interactive bundle-based method for nondifferentiable multiobjective optimization: NIMBUS. *Optimization* **34**, 231–246 (1995)
20. Miettinen, K., Makela, M.: Interactive bundle based method for nondifferentiable multiobjective optimization: nimbus. *Multiobject. Program. Goal Program.* **3**, 231–246 (1995)
21. Mordukhovich, B.S.: Multiobjective optimization problems with equilibrium constraints. *Math. Program.* **117**(1), 331–354 (2008)
22. Nguyen, T.T., Strodiot, J.J., Nguyen, V.H.: A bundle method for solving equilibrium problems. *Math. Program. Ser. B* **116**(1), 529–552 (2009)
23. Outrata, J., Kocvara, M., Zowe, J.: *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Theory, Applications and Numerical Results*. Kluwer Academic Publishers, Dordrecht (1998)
24. Pang, L.P., Meng, F.Y., Chen, S., Li, D.: Optimality conditions for multiobjective optimization problem constrained by parametric variational inequalities. *Set Valued Var Anal.* **22**, 285–298 (2014)
25. Pang, L.P., Meng, F.Y., Wang, J.H.: Asymptotic convergence of stationary points of stochastic multiobjective programs with parametric variational inequality constraint via SAA approach. *J. Ind. Manag. Optim.* **15**(4), 1653–1675 (2019)
26. Pang, L.P., Meng, F.Y., Xiao, Z.H., Xu, N.: Optimality conditions for vector optimization problem governed by the cone constrained generalized equations. *Optimization* **68**(5), 921–954 (2019)
27. Pang, L.P., Lv, J., Wang, J.H.: Constrained incremental bundle method with partial inexact oracle for nonsmooth convex semi-infinite programming problems. *Comput. Optim. Appl.* **64**, 433–465 (2016)
28. Sagastizabal, C., Solodov, M.: An infeasible bundle method for nonsmooth convex constrained optimization without a penalty function or a filter. *SIAM* **16**, 146–169 (2005)
29. Vieira, D.A.G., Lisboa, A.C.: A cutting-plane method to nonsmooth multiobjective optimization problems. *Eur. J. Oper. Res.* **275**, 822–829 (2019)
30. Wu, Q., Wang, J.H., Zhang, H.W., Wang, S., Pang, L.P.: Nonsmooth optimization method for H^∞ output feedback control. *Asia Pac. J. Oper. Res.* **36**(03), 1–23 (2019). <https://doi.org/10.1142/S0217595919500155>
31. Yang, Y., Pang, L.P., Ma, X.F., Shen, J.: Constrained nonsmooth nonsmooth optimization via proximal bundle method. *J. Optim. Theory Appl.* **163**, 900–925 (2014)

32. Ye, J.J., Zhu, Q.J.: Multiobjective optimization problem with variational inequality constraints. *Math. Program.* **96**(1), 139–160 (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.