CrossMark

Proximal point method for a special class of nonconvex multiobjective optimization functions

G. C. Bento[1] · O. P. Ferreira[1] · V. L. Sousa Junior[1]

Abstract   The proximal point method for a special class of nonconvex multiobjective functions is studied in this paper. We show that the method is well defined and that the accumulation points of any generated sequence, if any, are Pareto–Clarke critical points. Moreover, under additional assumptions, we show the full convergence of the generated sequence.

1 Introduction

Multiobjective optimization is the process of simultaneously optimizing two or more real-valued objective functions. Usually, no single point will minimize all the given objective functions at once (i.e., there is no ideal minimizer), and so the concept of optimality has to be replaced by the concept of *Pareto optimality* or as we will see, *Pareto–Clarke critical*; see [11]. These types of problems have applications in the economy, industry, agriculture, and other fields; see [13]. Bonnel et al. [6] considered extensions of the proximal point method to the multiobjective setting, see also [1–4,7–9,20] and references therein.

✉ V. L. Sousa Junior
valdinesldsj@gmail.com

G. C. Bento
glaydston@ufg.br

O. P. Ferreira
orizon@ufg.br

[1]   Universidade Federal de Goiás, Goiânia, GO, Brazil

Our goal is to study the proximal point method introduced in [6] for the multi-objective problems, where each component function is lower-C^1, a special class of nonconvex functions. Over the last four decades, several authors have proposed generalized proximal point methods for certain nonconvex minimization problems. As far as we know, the first generalization was performed in [12], see also [15] for a review. Our approach extends to the multiobjective context the results of [15]. More precisely, we show that the method is well defined and that the accumulation points of any generated sequence, if any, are Pareto–Clarke critical for the multiobjective function. Moreover, under some additional assumptions, we show the full convergence of the generated sequence.

The organization of the paper is as follows. In Sect. 2, some notation and basic results used throughout the paper are presented. In Sect. 3, the main results are stated and proved. Some final remarks are made in Sect. 4.

2 Preliminaries

In this section, we present some basic results and definitions.

We denote $I := \{1, \ldots, m\}$, $\mathbb{R}^m_+ := \{x \in \mathbb{R}^m : x_j \geq 0, j \in I\}$, and $\mathbb{R}^m_{++} := \{x \in \mathbb{R}^m : x_j > 0, j \in I\}$. For $y, z \in \mathbb{R}^m$, $z \geq y$ (or $y \leq z$) means that $z - y \in \mathbb{R}^m_+$ and $z > y$ (or $y \prec z$) means that $z - y \in \mathbb{R}^m_{++}$. We consider the *unconstrained multiobjective problem*: $\min_{x \in \mathbb{R}^n} F(x)$, where $F : \mathbb{R}^n \to \mathbb{R}^m$, with $F(x) = (f_1(x), \ldots, f_m(x))$. Given a nonempty set $C \subset \mathbb{R}^n$, a point $x^* \in C$ is said to be a weak Pareto solution of the problem $\min_w\{F(x) : x \in C\}$ if, and only if, there is no $x \in C$ with $F(x) \prec F(x^*)$. We denote as $\operatorname{argmin}_w\{F(x) : x \in C\}$ the weak Pareto solutions set. In particular, when $C = \mathbb{R}^n$, we denote this set as U^*. Assume that C is convex. F is called ν-*strongly convex* (or simply strongly convex) on C, $\nu \in \mathbb{R}^m_{++}$, if, and only if, for every $x, y \in C$,

$$F((1 - t)x + ty) \preceq (1 - t)F(x) + tF(y) - \nu t(1 - t)\|x - y\|^2, \quad t \in [0, 1].$$

F is said to be convex when $\nu = 0$ in the above inequality. Note that F is convex (resp. strongly convex) if, and only if, F is component-wise convex (resp. strongly convex). Moreover, this definition generalizes the definition of a convex function in the scalar case. The proof of the next proposition can be found in [16, p. 95].

Proposition 1 *If C is a convex set and F is a convex function, then*

$$\bigcup_{z \in \mathbb{R}^m_+ \setminus \{0\}} \operatorname*{argmin}_{x \in C} \langle F(x), z \rangle = \operatorname*{argmin}_w \{F(x) : x \in C\}.$$

If $m = 1$, f is L-strongly convex on $\Omega \subset \mathbb{R}^n$ with constant L if, and only if,

$$\langle u - v, x - y \rangle \geq L\|u - v\|^2, \quad u \in \partial f(x), \quad v \in \partial f(y), \tag{1}$$

whenever $x, y \in \Omega$, where ∂f denotes the subdifferential.

Remark 1 Let $f_1, f_2: \mathbb{R}^n \rightarrow \mathbb{R}$ be convex on Ω. Thus, $\partial f_1(x)$ and $\partial f_2(x)$ are nonempty, convex, and compact for $x \in \Omega$. Moreover, if $\lambda_1, \lambda_2 \geq 0$ then $\partial(\lambda_1 f_1 + \lambda_2 f_2)(x) = \lambda_1 \partial f_1(x) + \lambda_2 \partial f_2(x)$, for $x \in \Omega$; see [19, Theorem 23.8].

Let $C \subset \mathbb{R}^n$ be nonempty, closed, and convex. The normal cone is defined by

$$N_C(x) := \left\{ v \in \mathbb{R}^n : \langle v, y - x \rangle \leq 0, \ y \in C \right\}. \tag{2}$$

Remark 2 If $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, then the first-order optimality condition for $\min_{x \in C} g(x)$ is $0 \in \partial g(x) + N_C(x)$. If g is the maximum of a finite collection of continuously differentiable functions, then this constraint qualification is satisfied.

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz at $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$. The Clarke directional derivative [10, p. 25] of f at x in the direction d is defined as

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

and the Clarke subdifferential of f at x, denoted by $\partial^\circ f(x)$, is defined as

$$\partial^\circ f(x) := \left\{ w \in \mathbb{R}^n : \langle w, d \rangle \leq f^\circ(x, d), \ \forall \, d \in \mathbb{R}^n \right\}.$$

The previous definition can be found in [10, p. 27]. If f is convex, $f^\circ(x, d) = f'(x, d)$, where $f'(x, d)$ is the usual directional derivative. Moreover, $\partial^\circ f(x) = \partial f(x)$ for all $x \in \mathbb{R}^n$; see [10, Proposition 2.2.7]. The next lemmas can be found in [10, p. 39]

Lemma 1 *Let $\Omega \subset \mathbb{R}^n$ be open and convex. If $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz on Ω and $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is convex on Ω, then $(f + g)^\circ(x, d) = f^\circ(x, d) + g'(x, d)$ for each $x \in \Omega$ and $d \in \mathbb{R}^n$. Consequently, if $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable on Ω, $\partial^\circ(f + g)(x) = \partial^\circ f(x) + \mathrm{grad}\, g(x)$ for each $x \in \Omega$.*

Lemma 2 *Let $\Omega \subset \mathbb{R}^n$ be open and convex. Let $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable function on Ω, $i \in I$. Define $f(x) := \max_{i \in I} f_i(x)$, and $I(x) := \{i \in I : f_i(x) = f(x)\}$. Then, (a) f is locally Lipschitz on Ω and $conv\{\mathrm{grad}\, f_i(x) : i \in I(x)\} \subset \partial^\circ f(x)$, $x \in \Omega$; (b) if $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable and convex on Ω, $i \in I$, then $\partial f(x) = conv\{\mathrm{grad}\, f_i(x) : i \in I(x)\}$. In particular, x minimizes f on Ω if, and only if, there exists $\alpha_i \geq 0$, $i \in I(x)$, such that $0 = \sum_{i \in I(x)} \alpha_i \, \mathrm{grad}\, f_i(x)$ and $\sum_{i \in I(x)} \alpha_i = 1$; (c) if $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is L_i-strongly convex, for $i \in I$, f is $\min_{i \in I} L_i$ strongly convex.*

Proof The proofs of items (a) and (b) can be found in [5, Proposition 4.5.1] and [17, p. 49], respectively. The proof of item (c) follows from the definition of a strongly convex function. □

Definition 1 Let $F = (f_1, \ldots, f_m)^T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be locally Lipschitz on $\mathbb{R}^n$. We say that $x^* \in \mathbb{R}^n$ is a Pareto–Clarke critical point of F if, for all directions $d \in \mathbb{R}^n$, there exists $i_0 = i_0(d) \in \{1, \ldots, m\}$, such that $f_{i_0}^\circ(x^*, d) \geq 0$.

Remark 3 The previous definition can be found in [11]. When $m = 1$, the last definition becomes the classic definition of the critical point for the nonsmooth convex function. The last definition generalizes, for nonsmooth multiobjective optimization, the condition Im $(JF(x^*)) \cap (-\mathbb{R}^m_{++}) = \emptyset$, which characterizes a Pareto critical point when F is continuously differentiable.

3 Proximal algorithm for multiobjective optimization

In this section, we present a proximal point method to minimize a nonconvex function F, where its component is given by the maximum of continuously differentiable functions. Our goal is to prove the following theorem:

Theorem 1 *Let $\Omega \subset \mathbb{R}^n$ be open and convex and $I_j := \{1, \ldots, \ell_j\}$, with $\ell_j \in \mathbb{Z}_+$. Let $F(x) := (f_1(x), \ldots, f_m(x))$, where $f_j(x) := \max_{i \in I_j} f_{ij}(x)$, $j \in I$, and $f_{ij} : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function on Ω and continuous on $\bar{\Omega}$, for all $i \in I_j$. Assume that for all $j \in I$, $-\infty < \inf_{x \in \mathbb{R}^n} f_j(x)$, grad f_{ij} is Lipschitz on Ω with constant L_{ij} for each $i \in I_j$ and $S_F(F(\bar{y})) := \{x \in \mathbb{R}^n : F(x) \preceq F(\bar{y})\} \subset \Omega$, for some $\bar{y} \in \mathbb{R}^n$. Let $\bar{\lambda} > 0$ and $\bar{\mu} > 0$, such that $\bar{\mu} < 1$. Take $\{e^k\} \subset \mathbb{R}^m_{++}$ and $\{\lambda_k\} \subset \mathbb{R}_{++}$ satisfying*

$$\|e^k\| = 1, \quad \bar{\mu} < e^k_j, \quad \frac{1}{\bar{\mu}} \max_{i \in I_j} L_{ij} < \lambda_k \leq \bar{\lambda}, \quad j \in I, \quad k = 0, 1, \ldots. \tag{3}$$

Let $\hat{x} \in S_F(F(\bar{y}))$. If $\Omega_k := \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$, then

$$x^{k+1} \in \underset{w}{\operatorname{argmin}} \left\{ F(x) + \frac{\lambda_k}{2} \|x - x^k\|^2 e^k : x \in \Omega_k \right\}, \quad k = 0, 1, \ldots, \tag{4}$$

starting with $x^0 = \hat{x}$ is well defined, the generated sequence $\{x^k\}$ rests in $S_F(F(\bar{y}))$ and any accumulation point of $\{x^k\}$ is a Pareto–Clarke critical point of F, as long as Ω_k is convex, for each k.

In order to prove the above theorem we need some preliminaries. Hereafter, we assume that all the assumptions of Theorem 1 hold. We start proving the well-definedness of the sequence in (4).

Proposition 2 *The proximal point method (4) applied to F with starting point $x^0 = \hat{x}$ is well defined.*

Proof The proof will be made by induction on k. Let $\{x^k\}$ be as in (4). By assumption, $\hat{x} \in S_F(F(\bar{y}))$. Thus, we assume that $x^k \in S_F(F(\bar{y}))$ for some k. Take $z \in \mathbb{R}^m_+ \setminus \{0\}$ and define $\varphi_k(x) := \langle F(x), z \rangle + (\lambda_k/2)\langle e^k, z \rangle \|x - x^k\|^2$. As $-\infty < \inf_{x \in \mathbb{R}^n} f_j(x)$ for all $j \in I$, the function $\langle F(\cdot), z \rangle$ is bounded below and, taking into account that $\langle e^k, z \rangle > 0$, it follows that φ_k is coercive. Then, as Ω_k is closed, there exists $\tilde{x} \in \Omega_k$, such that $\tilde{x} = \operatorname{argmin}_{x \in \Omega_k} \varphi_k(x)$. Therefore, from Proposition 1 we can take $x^{k+1} := \tilde{x}$ and the induction is done, proving the proposition. □

Lemma 3 *For all $\tilde{x} \in \mathbb{R}^n$, $v := (v_1, \ldots, v_m) \in \mathbb{R}_{++}^m$, $j \in I$ and λ satisfying $\sup_{i \in I_j} L_{ij} < \lambda v_j$, the functions $f_{ij} + \lambda v_j \| \cdot - \tilde{x} \|^2 / 2$, $f_j + \lambda v_j \| \cdot - \tilde{x} \|^2 / 2$ and $F + (\lambda/2) \| \cdot - \tilde{x} \|^2 v$ are strongly convex on Ω. Moreover, $\langle F(\cdot), z \rangle + \lambda \langle v, z \rangle \| \cdot - \tilde{x} \|^2 / 2$ is strongly convex on Ω for each $z \in \mathbb{R}_+^m \backslash \{0\}$.*

Proof Take $j \in I$, $i \in I_j$, $\tilde{x} \in \mathbb{R}^n$, $v_j \in \mathbb{R}_{++}$ and define $h_{ij} = f_{ij} + \lambda v_j \| \cdot - \tilde{x} \|^2 / 2$. Since $\operatorname{grad} h_{ij}(x) = \operatorname{grad} f_{ij}(x) + \lambda v_j(x - \tilde{x})$, we have $\langle \operatorname{grad} h_{ij}(x) - \operatorname{grad} h_{ij}(y), x - y \rangle = \langle \operatorname{grad} f_{ij}(x) - \operatorname{grad} f_{ij}(y), x - y \rangle + \lambda v_j \| x - y \|^2$. Using the Cauchy inequality, last equality becomes

$$\langle \operatorname{grad} h_{ij}(x) - \operatorname{grad} h_{ij}(y), x - y \rangle \geq -\| \operatorname{grad} f_{ij}(x) - \operatorname{grad} f_{ij}(y) \| \| x - y \|$$
$$+ \lambda v_j \| x - y \|^2.$$

As $\operatorname{grad} f_{ij}$ is Lipschitz on Ω with constant L_{ij}, $\langle \operatorname{grad} h_{ij}(x) - \operatorname{grad} h_{ij}(y), x - y \rangle \geq (\lambda v_j - L_{ij}) \| x - y \|^2$ holds. Hence, the last inequality along with the assumption $\lambda v_j > \sup_{i \in I_j} L_{ij}$ implies that $\operatorname{grad} h_{ij}$ is strongly monotone. Therefore, (1) implies that h_{ij} is strongly convex, proving the first part of the lemma. The second and third parts of the lemma follow from the first one. $\square$

Hereafter, $\{x^k\}$ is generated by (4). Note that Proposition 1 implies that there exists a sequence $\{z^k\} \subset \mathbb{R}_+^m \backslash \{0\}$, such that

$$x^{k+1} = \operatorname*{argmin}_{x \in \Omega_k} \psi_k(x), \tag{5}$$

where the function $\psi_k : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\psi_k(x) := \left\langle F(x), z^k \right\rangle + \frac{\lambda_k}{2} \left\langle e^k, z^k \right\rangle \left\| x - x^k \right\|^2. \tag{6}$$

The solution of the problem in (5) is not altered through the multiplication of z^k by positive scalars. Thus, we can suppose $\| z^k \| = 1$ for $k = 0, 1, \ldots$.

Proof of Theorem 1 The well-definedness of (4) follows from Proposition 2. As $x_0 = \hat{x} \in S_F(F(\bar{y})) \subset \Omega$, (4) implies $\{x^k\} \subset S_F(F(\bar{y}))$. Let $\bar{x}$ be an accumulation point of $\{x^k\}$. Assume that Ω_k is convex and, by contradiction, that $\bar{x}$ is not Pareto–Clarke critical of F. Then, there exists $d \in \mathbb{R}^n$, such that

$$f_i^\circ(\bar{x}, d) < 0, \quad i \in I. \tag{7}$$

Thus, d is a descent direction for F in $\bar{x}$ and there exists $\delta > 0$, such that $F(\bar{x} + td) \prec F(\bar{x})$ for all $t \in (0, \delta]$. Hence, $\bar{x} + td \in \Omega_k$, for $k = 0, 1, \ldots$.

Let $\{z_k\}$ be a sequence satisfying (5). Hence, we can combine Lemma 3 and Remark 2 to obtain

$$0 \in \partial \left(\left\langle F(\cdot), z^k \right\rangle + \frac{\lambda_k}{2} \left\langle e^k, z^k \right\rangle \| \cdot - x^k \|^2 \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1}), \quad k = 0, 1, \ldots.$$

Letting $z^k = (z_1^k, \ldots, z_m^k)$ and $e^k = (e_1^k, \ldots, e_m^k)$, Remark 1 gives us,

$$0 \in \sum_{j=1}^m z_j^k \partial \left(f_j + \frac{\lambda_k}{2} e_j^k \left\| \cdot - x^k \right\|^2 \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1}), \quad k = 0, 1, \ldots.$$

The last inclusion implies that there exists $v^{k+1} \in N_{\Omega_k}(x^{k+1})$, such that

$$0 \in \sum_{j=1}^m z_j^k \partial \left(f_j + \frac{\lambda_k}{2} e_j^k \left\| \cdot - x^k \right\|^2 \right) (x^{k+1}) + v^{k+1}, \quad k = 0, 1, \ldots.$$

Since $\max_{i \in I_j} L_{ij} < \lambda_k e_j^k$, Lemma 3 implies that $f_{ij} + \lambda_k e_j^k \| \cdot - x^k \|^2 / 2$ and $f_j + \lambda_k e_j^k \| \cdot -x^k \|^2 / 2$ are strongly convex for all $j \in I$, $k = 0, 1, \ldots$. Applying Lemma 2(b), for $I = I_j$ and for the functions $f_{ij} + \lambda_k e_j^k \| \cdot - x^k \|^2 / 2$ and $f_j + \lambda_k e_j^k \| \cdot -x^k \|^2 / 2$, for each $j \in I$, we obtain

$$0 = \sum_{j=1}^m z_j^k \left(\sum_{i \in I_j(x^{k+1})} \alpha_{ij}^{k+1} \mathrm{grad} \left(f_{ij} + \frac{\lambda_k e_j^k}{2} \left\| \cdot - x^k \right\|^2 \right) (x^{k+1}) \right) + v^{k+1},$$

$$\times \sum_{i \in I_j(x^{k+1})} \alpha_{ij}^{k+1} = 1,$$

which holds for all $k = 0, 1, \ldots$, with $\alpha_{ij}^{k+1} \geq 0$, $i \in I_j(x^{k+1})$. This tells us that

$$0 = \sum_{j=1}^m z_j^k \left(\sum_{i \in I_j(x^{k+1})} \alpha_{ij}^{k+1} \left(\mathrm{grad}\, f_{ij}(x^{k+1}) + \lambda_k e_j^k (x^{k+1} - x^k) \right) \right) + v^{k+1},$$

$$\times \sum_{i \in I_j(x^{k+1})} \alpha_{ij}^{k+1} = 1, \tag{8}$$

for all $k = 0, 1, \ldots$. For all $j \in I$, let $\{\alpha_{ij}^{k+1}\} \subset \mathbb{R}^m$ be the sequence defined by $\alpha_j^{k+1} = (\alpha_{1j}^{k+1}, \alpha_{2j}^{k+1}, \ldots, \alpha_{mj}^{k+1})$, $\alpha_{ij}^{k+1} = 0$, $i \in I_j \backslash I_j(x^{k+1})$, for all $k = 0, 1, \ldots$. Since $\sum_{i \in I_j(x^{k+1})} \alpha_{ij}^{k+1} = 1$, $\|\alpha_j^{k+1}\|_1 = 1$ for all k, where $\| \cdot \|_1$ is the sum norm in $\mathbb{R}^n$. Thus, $\{\alpha_j^{k+1}\}$ is bounded. As $\{x^k\} \subset S_F(F(\bar{y}))$ and F is continuous on Ω, we have $\bar{x} \in S_F(F(\bar{y}))$. Since I_j is finite we can assume without loss of generality that $I_j(x^{k_1+1}) = I_j(x^{k_2+1}) = \cdots =: \tilde{I}_J$, and (8) becomes

$$0 = \sum_{j=1}^m z_j^{k_s} \left(\sum_{i \in \tilde{I}_J} \alpha_{ij}^{k_s+1} \mathrm{grad}\, f_{ij}(x^{k_s+1}) + \lambda_{k_s} e_j^{k_s}(x^{k_s+1} - x^{k_s}) \right) + v^{k_s+1},$$

$$\times \sum_{i \in \tilde{I}_J} \alpha_{ij}^{k_s+1} = 1, \quad s = 0, 1, \ldots. \tag{9}$$

From the continuity of F we obtain that Ω_k is closed. Considering that $x^{k_s} \in \Omega_{k_s}$, Ω_{k_s} is convex and $\Omega_{k_s+1} \subset \Omega_{k_s}$, for $s = 0, 1, \ldots$, we obtain that

$$\tilde{\Omega} := \cap_{s=0}^{+\infty} \Omega_{k_s}, \tag{10}$$

is nonempty, closed, and convex. As $v^{k_s+1} \in N_{\Omega_{k_s}}(x^{k_s+1})$ and $\tilde{\Omega} \subset \Omega_{k_s}$, (2) implies

$$\left\langle v^{k_s+1}, \ x - x^{k_s+1} \right\rangle \leq 0, \quad x \in \tilde{\Omega}, \quad s = 0, 1, \ldots. \tag{11}$$

On the other hand, let $\{z^{k_s+1}\}$, $\{x^{k_s+1}\}$, $\{e_j^{k_s+1}\}$, $\{\lambda_{k_s+1}\}$, and $\{\alpha_j^{k_s+1}\}$ be the subsequences of $\{z^{k+1}\}$, $\{x^{k+1}\}$, $\{e_j^{k+1}\}$, $\{\lambda_{k+1}\}$, and $\{\alpha_j^{k+1}\}$, respectively, such that $\lim_{s\to+\infty}(z^{k_s+1}, x^{k_s+1}, e_j^{k_s+1}, \lambda_{k_s+1}, \alpha_j^{k_s+1}) = (\bar{z}, \bar{x}, \bar{e}_j, \hat{\lambda}, \bar{\alpha}_j)$. This fact along with (9), implies that $\lim_{s\to+\infty} v^{k_s+1} = \bar{v}$. From (11), $\bar{v} \in N_{\tilde{\Omega}}(\bar{x})$. Hence, in view of (9), we have $0 = \sum_{j=1}^m \bar{z}_j \sum_{i\in\tilde{I}_J} \bar{\alpha}_{ij} \operatorname{grad} f_{ij}(\bar{x}) + \bar{v}$ and $\sum_{i\in\tilde{I}_J} \bar{\alpha}_{ij} = 1$. Let $x \in \tilde{\Omega}$. Taking $u_j = \sum_{i\in\tilde{I}_J} \bar{\alpha}_{ij} \operatorname{grad} f_{ij}(\bar{x})$, we have

$$0 = \sum_{j=1}^m \bar{z}_j \langle u_j, x - \bar{x} \rangle + \langle \bar{v}, x - \bar{x} \rangle. \tag{12}$$

As $\bar{x} + td \in \Omega_k$, for all $k = 0, 1, \ldots$, the definition of $\tilde{\Omega}$ in (10) implies that $\bar{x} + td \in \tilde{\Omega}$, $t \in (0, \delta]$. Since $u_j = \sum_{i\in\tilde{I}_J} \bar{\alpha}_{ij} \operatorname{grad} f_{ij}(\bar{x})$ and $\sum_{i\in\tilde{I}_J} \bar{\alpha}_{ij} = 1$, Lemma 2 (a) and (b) implies that $u_j \in \partial° f_j(\bar{x})$. Hence, using that $\bar{v} \in N_{\tilde{\Omega}}(\bar{x})$ and definition of $f_j°(\bar{x}, d)$, equality (12) with $x = \bar{x} + td$ yields $0 \leq \sum_{j=1}^m \bar{z}_j \langle u_j, d \rangle \leq \sum_{j=1}^m \bar{z}_j f_j°(\bar{x}, d)$. Thus, there exists $j \in I$ such that $f_j°(\bar{x}, d) \geq 0$, which contradicts (7). Therefore, $\bar{x}$ is Pareto–Clarke critical point of F.

Now let us introduce some conditions that will guarantee that $\{x^k\}$ converges to a point $x^* \in U^*$. Suppose that

(H1) $U = \{y \in \mathbb{R}^n : F(y) \preceq F(x^k), \ k = 0, 1, \ldots\} \neq \varnothing$;

(H2) there exists $c \in \mathbb{R}$ such that the following conditions hold:

(a) $S_F(ce) := \{x \in \mathbb{R}^n : F(x) \preceq ce\} \neq \varnothing$ and $S_F(ce) \subsetneq S_F(F(\bar{y}))$;

(b) $S_F(ce)$ is convex and F is convex on $S_F(ce)$, where $e := (1, \ldots, 1) \in \mathbb{R}^m$;

(H3) there exists $\delta > 0$ such that for all $z \in \mathbb{R}_+^m \setminus \{0\}$, $x \in S_F(F(\bar{y})) \setminus S_F(ce)$ and $w_z(x) \in \partial° (\langle F(\cdot), z \rangle)(x) + N_{\Omega_k}(x)$, it holds that $\|w_z(x)\| > \delta > 0$.

In general, the set U defined in assumption (H1) may be an empty set. To guarantee that U is nonempty, an additional assumption on the sequence $\{x^k\}$ is needed. In the next remark we give such a condition.

Remark 4 If the sequence $\{x^k\}$ has an accumulation point, then $U \neq \varnothing$, i.e., assumption (H1) holds. Indeed, let $\bar{x}$ be an accumulation point of the sequence $\{x^k\}$. Then, there exists a subsequence $\{x^{k_j}\}$ of $\{x^k\}$ which converges to $\bar{x}$. Since F is continuous, $\{F(x^k)\}$ has $F(\bar{x})$ as an accumulation point. Using the definition of $\{x^k\}$ in (4), we conclude that $\{F(x^k)\}$ is a decreasing sequence. Hence, the usual arguments easily show that the whole sequence $\{F(x^k)\}$ converges to $F(\bar{x})$ and $\bar{x} \in U$, i.e., $U \neq \varnothing$.

Next, we present a function satisfying the assumptions of Theorem 1, as well as (H1), (H2) and (H3), when $m = 2$, $j = 2$ and $I = I_j = 2$.

Example 1 Take $0 < \epsilon < 0.4$, $\Omega = (\epsilon, +\infty)$, $I := \{1, 2\}$, and $\bar{y} = 2.718\ldots$ with $\ln \bar{y} = 1$. Let $F : \mathbb{R} \to \mathbb{R}^2$ be defined by $F(x) = (0, 0)$ for $x \in \mathbb{R}\backslash\mathbb{R}_{++}$ and $F(x) := (f_1(x), f_2(x))$, where $f_j(x) := \max_{i \in I} f_{ij}(x)$ for $j \in I$ and $f_{11}(x) = \ln x + 1/x$, $f_{21}(x) = \ln x - 1/x$, $f_{12}(x) = 2\sqrt{x} + 1/x$, $f_{22}(x) = 2\sqrt{x} - 1/x$, for all $x \in \mathbb{R}_{++}$. Note that f_{1j}, f_{2j} are continuously differentiable on Ω and continuous on $\bar{\Omega}$, for all $j \in I$. Since f''_{1j}, f''_{2j} are bounded on Ω, we conclude that f'_{1j}, f'_{2j} are Lipschitz on Ω, for $j \in I$. Since $\max\{a, b\} = (a + b)/2 + |a - b|/2$, for all $a, b \in \mathbb{R}$, we conclude $f_1(x) = \ln x + 1/x$ and $f_2(x) = 2\sqrt{x} + 1/x$. Therefore, $F(x) = \left(\ln x + 1/x, 2\sqrt{x} + 1/x\right)$, $x \in \mathbb{R}_{++}$. It is easy to see that F is nonconvex and $\{x \in \Omega : F(x) \preceq (\zeta, \zeta)\}$ is convex and nonempty, for all $\zeta \geq 1$. Consider the following multiobjective optimization problem $\min_w \{F(x) : x \in \Omega\}$, which has $x^* = 1$ as the unique solution. In fact, $F(1) \prec F(x)$, for all $x \in \mathbb{R}_{++}$. Hence, we obtain that $-\infty < \inf_{x \in \mathbb{R}} f_j(x)$, for $j \in I$. Since $0 < \epsilon < \bar{y}$, we conclude that $S_F(F(\bar{y})) \subset \Omega$ and $S_F(F(\bar{y})) \neq \varnothing$. Therefore, taking into account that Ω_k is convex, F satisfies all the assumptions of Theorem 1. We are going to prove that F also satisfies (H1), (H2), and (H3). Since $F(1) \prec F(x)$, for all $x \in \mathbb{R}_{++}$, we conclude that F satisfies (H1). Let $c = f_2(2)$ and note that $(0.6, 2] \subset S_F(ce) \subsetneq [0.5, 2.7] \subset S_F(F(\bar{y}))$. But this tells us that, in particular, F satisfies (H2). Finally, we are going to prove that F satisfies (H3). First, note that $S_F(F(\bar{y}))\backslash S_F(ce) \subset [0.47, 0.57) \cup (2, 2.72]$. For each point $z = (z_1, z_2) \in \mathbb{R}_+^2\backslash\{0\}$ with $\|z\|_1 := z_1 + z_2 = 1$, take $x \in S_F(F(\bar{y}))\backslash S_F(ce)$ and $w_z(x) \in \partial° (\langle F(\cdot), z \rangle)(x) + N_{\Omega_k}(x)$. Hence, there exists $v \in N_{\Omega_k}(x)$, such that

$$w_z(x) = z_1\varphi_1(x) + z_2\varphi_2(x) + v, \tag{13}$$

where $\varphi_1(x) := (1/x - 1/x^2)$ and $\varphi_2(x) := (1/\sqrt{x} - 1/x^2)$. First, we assume that $x \in [0.47, 0.57)$. In this case, $N_{\Omega_k}(x) \subset \mathbb{R}_-$ and, using the above equality, we obtain $w_z \leq z_1\varphi_1(x) + z_2\varphi_2(x)$. Since $(x - 1)/x^2 < -0, 4/(0, 47)^2$ and $(x^2 - \sqrt{x})/x^{3/2} < -0, 2/(0, 47)^{3/2}$, we have $w_z(x) \leq z_1\varphi_1(x) + z_2\varphi_2(x) < -0, 4/(0, 47)^2 z_1 - 0, 2/(0, 47)^{3/2} z_2$. Then, for all $w_z(x) \in \partial° (\langle F(\cdot), z \rangle)(x) + N_{\Omega_k}(x)$,

$$|w_z(x)| > \frac{0, 4}{(0, 47)^2} z_1 + \frac{0, 27}{(0, 47)^{3/2}} z_2 > \|z\|_1 \frac{0, 27}{(0, 47)^{3/2}} = \frac{0, 27}{(0, 47)^{3/2}}, \tag{14}$$

for $x \in [0.47, 0.57)$. Assuming that $x \in (2, 2.72]$, it follows that $N_{\Omega_k}(x) \subset \mathbb{R}_+$. Hence, it follows from (13) that $w_z(x) \geq z_1\varphi_1(x) + z_2\varphi_2(x)$. From $(x - 1)/x^2 > 1/(2, 72)^2$ and $\left(1/\sqrt{x} - 1/x^2\right) > 2, 3/(2, 72)^{3/2}$ we obtain $w_z(x) \geq z_1\varphi_1(x) + z_2\varphi_2(x) > 1/(2, 72)^2 z_1 + 2, 3/(2, 72)^{3/2} z_2$. Thus, for all $w_z(x) \in \partial° (\langle F(\cdot), z \rangle)(x) + N_{\Omega_k}(x)$,

$$|w_z(x)| > \frac{1}{(2, 72)^2} z_1 + \frac{2, 3}{(2, 72)^{3/2}} z_2 > \|z\|_1 \frac{1}{(2, 72)^2} = \frac{1}{(2, 72)^2}, \quad x \in (2, 2.72].$$

Since $S_F(F(\bar{y}))\backslash S_F(ce) \subset [0.47, 0.57) \cup (2, 2.72]$, combining (14) with the last inequality, we conclude that F satisfies (H3) with $\delta = 1/(2, 72)^2$.

Lemma 4 *Assume that* (H1)*, (a) in* (H2)*, and* (H3) *hold and λ_k satisfies* (3)*. Then, after a finite number of steps the proximal iterates lies in $S_F(ce)$, i.e., there exists k_0 such that $\{x^k\} \subset S_F(ce)$, for all $k \geq k_0$.*

Proof Condition (a) in (H2) implies that $S_F(ce) \neq \varnothing$. Suppose, by contradiction, that $x^k \in S_F(F(\bar{y})) \backslash S_F(ce)$ for all k. Let $\{z_k\}$ be a sequence satisfying (5). Hence, we can combine Lemma 3, Remark 2 and Lemma 1 to obtain $0 \in \partial^\circ \left(\langle F(\cdot), z^k \rangle\right)(x^{k+1}) + (\lambda_k/2)\langle e^k, z^k \rangle (x^{k+1} - x^k) + N_{\Omega_k}(x^{k+1})$, $k \geq 0$. Then, $-(\lambda_k/2)\langle e^k, z^k \rangle (x^{k+1} - x^k) \in \partial^\circ \left(\langle F(\cdot), z^k \rangle\right)(x^{k+1}) + N_{\Omega_k}(x^{k+1})$, for $k \geq 0$. As $x^{k+1} \in S_F(F(\bar{y})) \backslash S_F(ce)$, (H3) along with the last inclusion gives us

$$\frac{\lambda_k}{2}\left\langle e^k, z^k \right\rangle \left\| x^{k+1} - x^k \right\| > \delta, \quad k \geq 0. \tag{15}$$

Hence, $(\lambda_k/2)\langle e^k, z^k \rangle \left\| x^{k+1} - x^k \right\|^2 \leq \| F(x^k) - F(x^{k+1})\|$ holds, in view of (5) and (6), and $\|z^k\| = 1$, for all $k \geq 0$. Thus, from (H1), $\| F(x^k) - F(x^{k+1})\| \to 0$, contradicting (15). $\qquad\square$

Lemma 5 *Assume that* (H1) *and* (H2) *hold and λ_k satisfies* (3)*. If $x^k \in S_F(ce)$ for some k, then $\{x^k\}$ converges to a point $x^* \in U^*$.*

Proof By hypothesis, $x^k \in S_F(ce)$ for some k, i.e., there exists k_0 such that $F(x^{k_0}) \preceq ce$. Hence, the definition of $\{x^k\}$ in (4) implies that $\{x^k\} \subset S_F(ce)$, for all $k \geq k_0$. Therefore, using (3), (H1), and (H2), the result follows by applying [6, Theorem 3.1] with $F_0 = F$, $S = S_F(ce)$, $X = \mathbb{R}^n$, $C = \mathbb{R}^m_+$, and using (H1) instead of (A). $\qquad\square$

Theorem 2 *Under the conditions* $(H1)$*,* $(H2)$*, and* $(H3)$*, the sequence $\{x^k\}$ generated by* (4) *converges to a point $x^* \in U^*$.*

Proof It follows by combination of Lemma 4 with Lemma 5. $\qquad\square$

Remark 5 As the function in Example 3.1 is not convex, the analysis in [6] does not allow us to conclude that $\{x^k\}$ converges to a minimizer. However, as the function satisfies (H1) to (H3), Theorem 2 guarantees that $\{x^k\}$ converges.

4 Conclusions

The main contribution of this paper is the extension of the convergence analysis of the proximal method (4), which has been studied in [6], in order to increase the range of its applications; see Example 1. The proximal point method is indeed a conceptual scheme that transforms a given problem into a sequence of better behaved subproblems, which has been proven to be an efficient tool in several instances through the methods that can be derived from it (e.g., Augmented Lagrangians, both classical or generalized). In this sense, the proximal point method is basic and we expect that the results of the present paper will become an additional step toward solving general multiobjective optimization problems. Finally, it is worth noting that in our analysis of the preference relation was induced by a cone with a nonempty interior. Various vector optimization problems can be formalized by using convex ordering cones with empty interiors, this

restriction could open some new perspectives in the point of view of the numerical methods, see [2,14,18]. We foresee new developments in this direction in the near future.

References

1. Apolinário, H.C.F., Papa Quiroz, E.A., Oliveira, P.R.: A scalarization proximal point method for quasiconvex multiobjective minimization. J. Glob. Optim. **64**(1), 79–96 (2016)
2. Bao, T.Q., Mordukhovich, B.S.: Relative Pareto minimizers in multiobjective optimization: existence and optimality conditions. Math. Program. **122**, 301–347 (2010)
3. Bao, T.Q., Mordukhovich, B.S.: Necessary conditions for super minimizers in constrained multiobjective optimization. J. Glob. Optim. **43**, 533–552 (2009)
4. Bento, G.C., Cruz Neto, J.X., Soubeyran, A.: A proximal point-type method for multicriteria optimization. Set-Valued Var. Anal. **22**(3), 557–573 (2014)
5. Bertsekas, D.P.: Convex Analysis and Optimization. Athena Scientific, Belmont (2003)
6. Bonnel, H., Iusem, A.N., Svaiter, B.F.: Proximal methods in vector optimization. SIAM J. Optim. **15**(4), 953–970 (2005)
7. Ceng, L.C., Mordukhovich, B.S., Yao, J.C.: Hybrid approximate proximal method with auxiliary variational inequality for vector optimization. J. Optim. Theory Appl. **146**(2), 267–303 (2010)
8. Ceng, L.C., Yao, J.C.: Approximate proximal methods in vector optimization. Eur. J. Oper. Res. **183**(1), 1–19 (2007)
9. Chuong, T.D., Mordukhovich, B.S., Yao, J.C.: Hybrid approximate proximal algorithms for efficient solutions in vector optimization. J. Nonlinear Convex Anal. **12**(2), 257–286 (2011)
10. Clarke, F.H.: Optimization and Nonsmooth Analysis, Volume 5 of Classics in Applied Mathematics, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1990)
11. Custódio, A.L., Madeira, J.F.A., Vaz, A.I.F., Vicente, L.N.: Direct multisearch for multiobjective optimization. SIAM J. Optim. **21**(3), 1109–1140 (2011)
12. Fukushima, M., Mine, H.: A generalized proximal point algorithm for certain nonconvex minimization problems. Int. J. Syst. Sci. **12**(8), 989–1000 (1981)
13. Gal, T., Hanne, T.: On the development and future aspects of vector optimization and MCDM. A tutorial. In: Multicriteria Analysis (Coimbra, 1994), pp. 130–145. Springer, Berlin (1997)
14. Grad, S.M., Pop, E.L.: Vector duality for convex vector optimization problems by means of the quasi-interior of the ordering cone. Optimization **63**(1), 21–37 (2014)
15. Kaplan, A., Tichatschke, R.: Proximal point methods and nonconvex optimization. J. Glob. Optim. **13**(4), 389–406 (1998)
16. Luc, D.T.: Theory of Vector Optimization, volume 319 of Lecture Notes in Economics and Mathematical Systems. Springer, Berlin (1989)
17. Mäkelä, M.M., Neittaanmäki, P.: Nonsmooth Optimization. World Scientific Publishing Co., Inc., River Edge (1992)
18. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, II: Applications. Grundlehren Series in Fundamental Principles of Mathematical Sciences, vol. 331. Springer, Berlin (2006)
19. Rockafellar, R.T.: Convex Analysis. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton (1970)
20. Villacorta, K.D.V., Oliveira, P.R.: An interior proximal method in vector optimization. Eur. J. Oper. Res. **214**(3), 485–492 (2011)