Noname manuscript No.
(will be inserted by the editor)

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

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

Received: date / Accepted: date

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

Keywords Multiobjective optimization · Pareto-Clarke optimality · Nonconvex optimization

Mathematics Subject Classification (2000) 90C30 · 90C29 · 90C26

G. C. Bento

Universidade Federal de Goiás, Goiânia, GO, Brazil.

E-mail: glaydston@ufg.br

O. P. Ferreira

Universidade Federal Goiás, Goiânia, GO, Brazil.

E-mail: orizon@ufg.br

V. L. Sousa Junior

Universidade Federal de Goiás, Goiânia, GO, Brasil.

E-mail: valdinesldsj@gmail.com

1 Introduction

Multiobjective optimization is the process of simultaneously optimizing two or more real-valued objective functions. Usually, no single point will minimize all 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-Clare critical*; see [14]. These types of problems have applications in the economy, industry, agriculture, and the others filds; see [22]. Bonnel et al. [8] considered extensions of the proximal point method to the multiobjectivel setting, see also, Bento et al. [4], Ceng and Yao [10], Ceng et al. [9], Choung et al. [11], Villacorta and Oliveira [33] and references therein. We point out that other methods, associated to scalar-valued optimization, have already been extended to multiobjective optimization; see, for instance, Fliege and Svaiter [18], Graña Drummond and Iusem [16], Graña Drummond and Svaiter [24], Fliege et al. [17], Fukuda and Graña Drummond [19, 20], Bello Cruz et al. [1], Bento et al. [3, 5], Cruz Neto et al. [15], Bento and Cruz Neto [2].

Our goal is to study the proximal method for multiobjective problems, introduced by Bonnel et al. [8], where the coordinate functions are like Lower-C^1, which include a special class of nonconvex functions. In the last four decades, several authors have proposed generalized proximal point methods for certain nonconvex minimization problems as well as problems of finding singularities of nonmonotone operators. As far as we know the first generalization has been performed by Fukushima and Mine [21], see also Kaplan and Tichatschke [26] for a review. For the problem of finding singularities of nonmonotone operators (e.g. where the operator is hypomonotone), see e.g., Spingarn and Jonathan [32], Pennanen [29], Iusem et al. [25], Combettes and Pennanen [13] and Garciga and Iusem [23].

Our approach extend to the multiobjective context the results of Kaplan and Tichatschke [26] (see also, [6]). More precisely, we show that the method is well defined and the accumulation points of any generated sequence, if any, are Pareto-Clarke critical points for the multiobjective function. Moreover, under some additional assumptions, we show full convergence of the generated sequence.

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

2 Basic results and definitions

Let $\Omega \subset \mathbb{R}^n$ be a convex set. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said *strongly convex* on Ω with constant $L > 0$ if,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) - \frac{1}{2} L\alpha(1-\alpha)\|x-y\|^2,$$

for all $x, y \in \Omega$ with $\alpha \in [0,1]$. If $L = 0$ then f is said *convex* on Ω. It is possible to show that f is strongly convex with constant L iff

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

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

Remark 2.1 The subdifferential of a convex functions is always nonempty, convex and compact . Furthermore, if f_1, f_2 are two convex functions on Ω and $\lambda_1, \lambda_2 \geq 0$, then $\partial(\lambda_1 f_1 + \lambda_2 f_2)(x) = \lambda_1 \partial f_1(x) + \lambda_2 \partial f_2(x)$, for all $x \in \Omega$; see [30, Theorem 23.8].

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

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

Remark 2.2 Note that, if $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function then the first-order optimality condition for the optimization problem $\min_{x\in C} g(x)$ takes the following form $0 \in \partial g(x) + N_C(x)$.

Proposition 2.1 *Let $C \subset \mathbb{R}^n$ be a closed and convex set and $\{y^k\} \subset C$ a bounded sequence. If $\{z^k\}$ is a sequence such that $z^k \in N_C(y^k)$, then $\{z^k\}$ is bounded.*

Proof The proof follows by combining [31, Theorem 9.13 and Proposition 5.15] for $f = I_C$, $S = \partial I_C = N_C$ and $B = \{y^k\}$, where I_C denotes the indicator function of C . □

Proposition 2.2 *Let $h_i : \mathbb{R}^n \to \mathbb{R}$ be a strongly convex function with constant L_i, for $i \in I = \{1,...,p\}$. Then $h : \mathbb{R}^n \to \mathbb{R}$ defined by $h(x) = \max_{i \in I} h_i(x)$ is strongly convex function with constant $L := \min_{i \in I} L_i$. In particular, if h_i is convex for all $i \in I$, h is convex on Ω.*

Proof It follows from the definition of strongly convex function.

The proof of next result follows directly from [7, Proposition 4.5.1].

Proposition 2.3 *Let $\Omega \subset \mathbb{R}^n$ be a convex set and $h_i : \mathbb{R}^n \to \mathbb{R}$ a differentiable convex function on Ω for $i \in I = \{1,...,p\}$. If $h : \mathbb{R}^n \to \mathbb{R}$ is defined by $h(x) = \max_{i \in I} h_i(x)$, then:*

$$\partial h(x) = \left\{ \sum_{i \in I(x)} \alpha_i \text{grad } h_i(x) \ : \ \sum_{i \in I(x)} \alpha_i = 1, \ \alpha_i \geq 0 \right\},$$

where $I(x) = \{i \in I \ : \ h(x) = h_i(x)\}$. In particular, x minimizes h on Ω, if and only if, there exist $\alpha_i \geq 0$, $i \in I(x)$, such that:

$$0 = \sum_{i \in I(x)} \alpha_i \, \text{grad} \, h_i(x), \qquad \sum_{i \in I(x)} \alpha_i = 1.$$

Next definition can be found in [12, page 10].

Definition 2.1 Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function at $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$. The Clarke directional derivative of f at x in the direction d, denoted by $f^\circ(x,d)$, is defined as

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

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

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

Remark 2.3 If f is a convex function, $f^\circ(x,d) = f'(x,d)$, where $f'(x,d)$ is the usual directional derivative of f at x in the direction d. Moreover $\partial^\circ f(x) = \partial f(x)$ for all $x \in \mathbb{R}^n$, i. e., the Clarke's subdifferential generalize of the usual subdifferential for convex functions; see [12, Proposition 2.2.7].

The following two Lemmas, are results well known from Clarke's subdiffential theory. They can be found in [12, page 39] and [28, page 49] respectively.

Lemma 2.1 *Let $\Omega \subset \mathbb{R}^n$ be an open convex set. If $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz on Ω and $g : \mathbb{R}^n \to \mathbb{R}$ is convex on Ω, then $(f+g)^\circ(x,d) = f^\circ(x,d) + g'(x,d)$, for each $x \in \Omega$ and $d \in \mathbb{R}^n$. As a consequence, if $g : \mathbb{R}^n \to \mathbb{R}$ is continously differentiable on Ω, $\partial^\circ(f+g)(x) = \partial^\circ f(x) + \operatorname{grad} g(x)$, for each $x \in \Omega$.*

Lemma 2.2 *Let $\Omega \subset \mathbb{R}^n$ be an open convex set and $I = \{1,\ldots,m\}$. Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be a continuously diffentiable function on Ω for all $i \in I$. Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) := \max_{i\in I} f_i(x)$, and $I(x) := \{i \in : f_i(x) = f(x)\}$. Then, f is locally Lipschitz on Ω and conv$\{grad f_i(x) : i \in I(x)\} \subset \partial^\circ f(x)$, for each $x \in \Omega$.*

Next definition can be found in [14].

Definition 2.2 Let $F = (f_1,\ldots,f_m)^T : \mathbb{R}^n \to \mathbb{R}^m$ 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 2.4 We point out that a similiar definition, in the Riemannian context, has also appeared in [2]. Note that, when $m = 1$, last definition becomes the classic definition of critical point for nonsmooth convex function. Note also that the last definition generalizes, for nonsmooth multiobjective optimization, the condition $\operatorname{Im}(JF(x^*)) \cap \left(-\mathbb{R}^m_{++}\right) = \emptyset$, (where $\operatorname{Im}(JF(x^*))$ denotes the image of the jacobian of F at point $x^* \in \mathbb{R}^n$), which characterizes a Pareto critical point when F is a continuously differentiable vector function.

3 Multiobjective Optimization

In this section, we present the Multiobjective problem. First we need some basic definitions.

$$\mathbb{R}^m_+ := \left\{x \in \mathbb{R}^m \ : \ x_j \geq 0, \ j \in I\right\}, \qquad \mathbb{R}^m_{++} := \left\{x \in \mathbb{R}^m \ : \ x_j > 0 \ j \in I\right\}, \qquad I := \{1,\ldots,m\}.$$

For $y,z \in \mathbb{R}^m$, $z \succeq y$ (or $y \preceq z$) means that $z-y \in \mathbb{R}^m_+$ and $z \succ y$ (or $y \prec z$) means that $z-y \in \mathbb{R}^m_{++}$. We consider the *unconstrained multiobjective optimization problem*:

$$\min_{x\in\mathbb{R}^n} F(x), \tag{3}$$

where $F:\mathbb{R}^n\to\mathbb{R}^m$, with $F(x)=(f_1(x),\ldots,f_m(x))$ and $f_i:\mathbb{R}^n\to\mathbb{R}$ for $i=1,\ldots m$.

Definition 3.1 A point $x^*\in\mathbb{R}^n$ is a weak Pareto solution of the problem (3) iff there is no $x\in\mathbb{R}^n$ with $F(x)\prec F(x^*)$. Denotes by $U^*:=\operatorname{argmin}_w\{F(x)\ :\ x\in C\}$, the weak Pareto solutions set.

The proof of the next proposition can be found in [8].

Proposition 3.1 *Let $H:\mathbb{R}^n\to\mathbb{R}^m$ be a vectorial function and $C\subset\mathbb{R}^n$ a convex set. Then,*

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

Definition 3.2 Let $H:\mathbb{R}^n\to\mathbb{R}^m$ be a multiobjective function.

i) H is called *convex* iff for every $x,y\in\mathbb{R}^n$, the following holds:

$$H((1-t)x+ty)\preceq(1-t)H(x)+tH(y),\qquad t\in[0,1];$$

ii) H is called *strongly convex* with constant $\nu\in\mathbb{R}^m_{++}$ iff for every $x,y\in\mathbb{R}^n$, the following holds:

$$H((1-t)x+ty)\preceq(1-t)H(x)+tH(y)-t(1-t)\|x-y\|^2\nu,\qquad t\in[0,1].$$

Remark 3.1 For the first definition above, see [27, Definition 6.2]. Note that H is convex (resp. strongly convex) iff H is component-wise convex (resp. strongly convex). From the above definitions, it is easy to see that if H is strongly convex, in particular, it is also convex (the reciprocal is clearly false).

3.1 Convergence Analysis of a Proximal Algorithm for Multiobjective Optimization

In this section, we present an application of the proximal point method to minimize a multiobjective function $F:\mathbb{R}^n\to\mathbb{R}^m$ (not necessarily convex) where each component is given by the maximum of a certain class of continuously differentiable functions. Our goal is to prove the following theorem:

Theorem 3.1 *Let $\Omega\subset\mathbb{R}^n$ be an open convex set, $\bar{y}\in\Omega$, $\hat{I}:=\{1,\ldots,m\}$ and $I_j:=\{1,\ldots,\ell_j\}$, with $\ell_j\in\mathbb{Z}_+$. Let $F:\mathbb{R}^n\to\mathbb{R}^m$ defined by $F(x):=(f_1(x),\ldots,f_m(x))$, where*

$$f_j(x):=\max_{i\in I_j}f_{ij}(x),\quad j\in\hat{I},$$

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* $j \in \hat{I}$*,* $-\infty < \inf_{x\in\mathbb{R}^n} f_j(x)$*,* $\mathrm{grad} f_{ij}$ *is Lipschitz on* Ω *with constant* L_{ij} *for each* $i \in I_j$ *and*

$$S_F(F(\bar{y})) := \{x \in \mathbb{R}^n \,:\, F(x) \preceq F(\bar{y})\} \subset \Omega, \qquad \inf_{x\in\mathbb{R}^n} f_j(x) < \min_{s\in\hat{I}} f_s(\bar{y}), \qquad j \in \hat{I}.$$

Let $\bar{\lambda} > 0$ *and* $\bar{\mu} > 0$*. Take sequences* $\{e^k\} \subset \mathbb{R}^m_{++}$ *and* $\{\lambda_k\} \subset \mathbb{R}_{++}$ *satisfying*

$$\|e^k\| = 1, \qquad \bar{\mu} < e^k_j, \qquad \frac{1}{\bar{\mu}} \max_{i\in I_j} L_{ij} < \lambda_k \le \bar{\lambda}, \qquad j \in \hat{I}, \quad k = 0,1,\ldots. \tag{4}$$

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

$$x^{k+1} \in \operatorname*{argmin}_{w} \left\{ F(x) + \frac{\lambda_k}{2}\|x - x^k\|^2 e^k \,:\, x \in \Omega_k \right\}, \qquad k = 0,1,\ldots, \tag{5}$$

with starting ponit $x^0 = \hat{x}$ *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 a convex set, for each* k*. Moreover, assuming that*

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

Let $c \in \mathbb{R}$ *such that* $\inf_{x\in\mathbb{R}^n} f_j(x) < c < \min_{s\in\bar{I}} f_s(\bar{y})$*, for all* $j \in \hat{I}$*. If, in addition, hold:*

(H2) $S_F(c) := \{x \in \mathbb{R}^n \,:\, F(x) \preceq ce\}$ *is convex and* F *is convex on* $S_F(c)$*, where* $e := (1,\ldots,1) \in \mathbb{R}^m$*;*

(H3) *For each point* $z \in \mathbb{R}^m_+ \backslash \{0\}$*,* $x \in S_F(F(\bar{y})) \setminus S_F(c)$ *and* $w_z(x) \in \partial^\circ(\langle F(\cdot), z\rangle)(x) + N_{\Omega_k}(x)$ *we have* $\|w_z(x)\| > \delta_z > 0$*,*

then the sequence $\{x^k\}$ *generated by proximal point method converges to a point* $x^* \in U^*$*.*

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 3.2 If the sequence $\{x^k\}$ has an accumulation point, then U is nonempty, 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.

Besides, using definition of $\{x^k\}$ in (5), we conclude that $\{F(x^k)\}$ is a decreasing sequence. Hence, usual arguments easily show that the whole sequence $\{F(x^k)\}$ converges to $F(\bar{x})$ and $\bar{x} \in U$, i.e., $U \neq \varnothing$.

In order to prove the above theorem we need some preliminary results. From now on we assume that every assumptions on Theorem 3.1 hold, with the exception of (H1), (H2) and (H3), which will be considered to hold only when explicitly stated. First of all we show, in the next remark, that if Ω is bounded and F is convex on Ω then F satisfies the assumptions (H1), (H2) and (H3).

Remark 3.3 Assume that Ω is bounded and F is convex on Ω. Since Ω is bounded, Remark 3.2 implies that assumptions (H1) holds. As F is convex on Ω, assumptions (H2) holds. For each $z \in \mathbb{R}^m_+ \backslash \{0\}$, define $F_z(x) = \langle F(x), z\rangle$. If F is convex on Ω, then F_z is convex on Ω and from Remark 2.3,

$$\partial^\circ F_z(x) = \partial F_z(x), \qquad x \in \Omega. \tag{6}$$

Take $c \in \mathbb{R}$ such that $\inf_{x\in\mathbb{R}^n} f_j(x) < c < \min_{s\in\bar{I}} f_s(\bar{y})$, for all $j \in \hat{I}$, $x^* \in U$ and set

$$0 < \varepsilon = \sup\{\|x - x^*\| \ : \ x \in S_F(F(\bar{y})) \setminus S_F(c)\} < +\infty. \tag{7}$$

Let $x \in S_F(F(\bar{y})) \setminus S_F(c)$ and $w_z(x) = u_z(x) + v(x)$ with $u_z(x) \in \partial F_z(x)$ and $v(x) \in N_{\Omega_k}(x)$. The convexity of F_z on Ω implies $F_z(x^*) \geq F_z(x) + \langle u_z(x), x^* - x\rangle$. As $v(x) \in N_{\Omega_k}(x)$, we obtain

$$F_z(x^*) \geq F_z(x) + \langle w_z(x), x^* - x\rangle - \langle v(x), x^* - x\rangle \geq F_z(x) + \langle w_z(x), x^* - x\rangle. \tag{8}$$

Since $\|w_z(x)\|\|x^* - x\| \geq F_z(x) - F_z(x^*)$, the set U is a proper subset of $S_F(c)$ and $x \in S_F(F(\bar{y})) \setminus S_F(c)$, we conclude that $\|w_z(x)\|\|x^* - x\| > c\langle e, z\rangle - F_z(x^*) > 0$. Thus, from (7) and the latter inequality, we have

$$\|w_z(x)\|\varepsilon > c\langle e, z\rangle - F_z(x^*) > 0.$$

Therefore, choosing $\delta_z = (c\langle e, z\rangle - F_z(x^*))/\varepsilon$, we obtain $\|w_z(x)\| > \delta_z > 0$ which, combined with (6), shows that F_z satisfies (H3).

Proposition 3.2 *The proximal point method* (5) *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 a sequence defined in (5). 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 $z \in \mathbb{R}^m_+ \backslash \{0\}$ and define $\varphi : \mathbb{R}^n \to \mathbb{R}$ by

$$\varphi_k(x) = \langle F(x), z \rangle + \frac{\lambda_k}{2} \langle e^k, z \rangle \|x - x^k\|^2. \tag{9}$$

As $-\infty < \inf_{x \in \mathbb{R}^n} f_j(x)$ for all $j \in \hat{I}$, the function $\langle F(\cdot), z \rangle$ is bounded below and, tanking into account that $\langle e^k, z \rangle > 0$, it is also coercive. Then, as Ω_k is closed, there exists $\tilde{x} \in \Omega_k$ such that

$$\tilde{x} = \arg\min_{x \in \Omega_k} \varphi_k(x).$$

Therefore, from Proposition 3.1 we can take $x^{k+1} := \tilde{x}$ and the induction is done, which conclude the proof. □

From now on, $\{x^k\}$ is the sequence generated by (5). Note that Proposition 3.1 implies that there exists a sequence $\{z^k\} \subset \mathbb{R}^m_+ \backslash \{0\}$ such that

$$x^{k+1} = \arg\min_{x \in \Omega_k} \psi_k(x), \tag{10}$$

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

$$\psi_k(x) := \left\langle F(x), z^k \right\rangle + \frac{\lambda_k}{2} \left\langle e^k, z^k \right\rangle \left\| x - x^k \right\|^2. \tag{11}$$

Note that the solution of problem in (10) is not altered through multiplication of z^k by positive scalars. Hence, from now on we assume, without loss of generality, that $\|z^k\| = 1$ for $k = 0, 1, \ldots$.

Lemma 3.1 *Assume that* (H1) *and* (H2) *hold and* λ_k *satisfies* (4). *If* $x^k \in S_F(c)$ *for some* k *then* $\{x^k\}$ *converges to a point* $x^* \in U^* \subset \Omega$.

Proof By hypothesis, $x^k \in S_F(c)$ for some k, i. e., there exists k_0 such that $F(x^{k_0}) \preceq ce$. Hence, definition of $\{x^k\}$ in (5) implies that $\{x^k\} \subset S_F(c)$, for all $k \geq k_0$. Therefore, using (4), (H1) and (H2) the result follows by applying [8, Theorem 3.1] with $F_0 = F$, $S = S_F(c)$, $X = \mathbb{R}^n$, $C = \mathbb{R}^m_+$ and by taking into account that (H1) replaced assumption (A). □

Lemma 3.2 *For all $\tilde{x} \in \mathbb{R}^n$, $v := (v_1, \ldots, v_m) \in \mathbb{R}^m_{++}$, $j \in \hat{I}$ and λ satisfying $\sup_{i \in I_j} L_{ij} < \lambda v_j$, the function $f_{ij} + \lambda v_j \| \cdot - \tilde{x}\|^2/2$ is strongly convex in Ω with constant $\lambda v_j - \sup_{i \in I_j} L_{ij}$. Consequently, the function $f_j + \lambda v_j \| \cdot - \tilde{x}\|^2/2$ is strongly convex in Ω with constant $\lambda v_j - \sup_{i \in I_j} L_{ij}$ and $F + (\lambda/2)\| \cdot - \tilde{x}\|^2 v$ is strongly convex in Ω. Moreover, $\langle F(\cdot), z\rangle + \lambda \langle v, z\rangle \| \cdot - \tilde{x}\|^2/2$ is strongly convex in Ω for each $z \in \mathbb{R}^m_+ \backslash \{0\}$.*

Proof Take $j \in \hat{I}$, $i \in I_j$, $\tilde{x} \in \mathbb{R}^n$, $v_j \in \mathbb{R}_{++}$ 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 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 in Ω with constant L_{ij}, last inequality give us

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

Hence, combining latter inequality with assumption $\lambda v_j > \sup_{i \in I_j} L_{ij}$ we conclude that $\operatorname{grad} h_{ij}$ is strongly monotone with constant $\lambda v_j - \sup_{i \in I_j} L_{ij}$. Therefore, from (1) we conclude that h_{ij} is strongly convex with constant $\lambda v_j - \sup_{i \in I_j} L_{ij}$, which proof the first part of the lemma.

The second and third part of the lemma follows easily by combination of Proposition 2.2 and Remark 3.1. □

Lemma 3.3 *Assume that* (H1) *and* (H3) *hold and λ_k satisfies* (4). *Then after a finite number of steps the proximal iterates go into the set $S_F(c)$, i. e., there exists k_0 such that $\{x^k\} \subset S_F(c)$, for all $k \geq k_0$.*

Proof Suppose, by contradiction, that $x^k \in S_F(F(\bar{y})) \backslash S_F(c)$ for all k. Let $\{z_k\}$ be a sequence satisfying (10). Hence, we can combine Lemma 3.2, Remark 2.2 and Lemma 2.1 to obtain

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

Then,

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

As $x^{k+1} \in S_F(F(\bar{y}))\backslash S_F(c)$, the assumption (H3) and last inclusion give us

$$\frac{\lambda_k}{2}\left\langle e^k, z^k\right\rangle \left\|x^{k+1}-x^k\right\| > \delta_z, \qquad k=0,1,\ldots. \tag{12}$$

On the other hand, since that $x^{k+1} = \operatorname{argmin}_{x\in\mathbb{R}^n} \psi_k(x)$, see (10) and (11), and $\|z^k\| = 1$, we conclude that

$$\frac{\lambda_k}{2}\left\langle e^k, z^k\right\rangle \left\|x^{k+1}-x^k\right\|^2 \leq \left\langle F(x^k), z^k\right\rangle - \left\langle F(x^{k+1}), z^k\right\rangle \leq \|F(x^k)-F(x^{k+1})\|.$$

Using assumption (H1) we conclude that the sequence $\{\|F(x^k)-F(x^{k+1})\|\}$ goes to zero. Hence, combining assumption (4) with last inequality, we obtain that $\{x^{k+1}-x^k\}$ converges to zero, which contradicts (12). Therefore, the proof is done. □

Proof of Theorem 3.1. The well definedness of the proximal point method follows from Proposition 3.2. Let $\{x^k\}$ be the sequence generated by the proximal point method. As $x_0 = \hat{x} \in S_F(F(\bar{y})) \subset \Omega$, definition of $\{x^k\}$ in (5) implies that $\{x^k\} \subset S_F(F(\bar{y}))$. Let $\bar{x}$ be an accumulation point of $\{x^k\}$, assume that Ω_k is a convex set and, by contradiction, that $\bar{x}$ is not Pareto-Clarke critical point in $\mathbb{R}^n$. Then, there exists a direction $d \in \mathbb{R}^n$ such that

$$f_i^\circ(\bar{x}, d) < 0, \quad i \in \{1,\ldots,m\}. \tag{13}$$

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

Let $\{z_k\}$ be a sequence satisfying (10). Hence, we can combine Lemma 3.2 and Remark 2.2 to obtain

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

Letting $z^k = (z_1^k,\ldots,z_m^k)$ and $e^k = (e_1^k,\ldots,e_m^k)$, Remark 2.1 give us,

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

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

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

Since $\max_{i\in I_j} L_{ij} < \lambda_k e_j^k$, Lemma 3.2 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 functions, for all $j \in \hat{I}$ and $k = 0, 1, \ldots$. For each $j \in \hat{I}$, we can apply Proposition 2.3 with $I = I_j$, $h_i = f_{ij} + \lambda_k e_j^k \|\cdot - x^k\|^2/2$ and $h = f_j + \lambda_k e_j^k \|\cdot - x^k\|^2/2$, to conclude that there exists constant $\alpha_{ij}^{k+1} \geq 0$ with $i \in I_j(x^{k+1})$ such that

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

for all $k = 0, 1, \ldots$. This tells us that

$$0 = \sum_{j=1}^{m} z_j^k \left(\sum_{i\in I_j(x^{k+1})} \alpha_{ij}^{k+1} \left(\mathrm{grad} f_{ij}(x^{k+1}) + \lambda_k e_j^k (x^{k+1} - x^k) \right) \right) + v^{k+1}, \qquad \sum_{i\in I_j(x^{k+1})} \alpha_{ij}^{k+1} = 1, \tag{14}$$

for all $k = 0, 1, \ldots$. Now, for each $j \in \hat{I}$, let $\{\alpha_{ij}^{k+1}\} \subset \mathbb{R}^m$ be the sequence defined by

$$\alpha_j^{k+1} = (\alpha_{1j}^{k+1},\ \alpha_{2j}^{k+1},\ \ldots\ ,\ \alpha_{mj}^{k+1}), \qquad \alpha_{ij}^{k+1} = 0, \qquad i \in I_j \backslash I_j(x^{k+1}), \quad k = 0, 1, \ldots.$$

Since $\sum_{i\in I_j(x^{k+1})} \alpha_{ij}^{k+1} = 1$, we have $\|\alpha_j^{k+1}\|_1 = 1$ for all k, where $\|\cdot\|_1$ denotes the sum norm in $\mathbb{R}^n$. Thus $\{\alpha_j^{k+1}\}$ is bounded. 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 $\lim_{s\to+\infty} z^{k_s+1} = \bar{z}$, $\lim_{s\to+\infty} x^{k_s+1} = \bar{x}$, $\lim_{s\to+\infty} e_j^{k_s+1} = \bar{e}_j$, $\lim_{s\to+\infty} \lambda_{k_s+1} = \hat{\lambda}$ and $\lim_{s\to+\infty} \alpha_j^{k_s+1} = \bar{\alpha}_j$. As $\{x^k\} \subset S_F(F(\bar{y}))$, continuity of F on Ω allow us to conclude that $\bar{x} \in S_F(F(\bar{y}))$. Since I_j is finite we can assume without loss of generality that

$$I_j(x^{k_1+1}) = I_j(x^{k_2+1}) = \cdots =: \tilde{I}_J, \tag{15}$$

and (14) becomes

$$0 = \sum_{j=1}^{m} z_j^{k_s} \left(\sum_{i\in \tilde{I}_J} \alpha_{ij}^{k_s+1} \mathrm{grad} f_{ij}(x^{k_s+1}) + \lambda_{k_s} e_j^{k_s} (x^{k_s+1} - x^{k_s}) \right) + v^{k_s+1}, \quad \sum_{i\in \tilde{I}_J} \alpha_{ij}^{k_s+1} = 1, \quad s = 0, 1, \ldots. \tag{16}$$

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

$$\tilde{\Omega} := \cap_{s=0}^{+\infty} \Omega_{k_s}, \tag{17}$$

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

$$\langle v^{k_s+1},\, x - x^{k_s+1}\rangle \leq 0, \qquad x \in \tilde{\Omega}, \qquad s = 0,1,\ldots. \tag{18}$$

From Proposition 2.1 we conclude that $\{v^{k_s}\}$ is bounded. Thus, we suppose that $\lim_{s\to+\infty} v^{k_s+1} = \bar{v}$. By (18), we have $\bar{v} \in N_{\tilde{\Omega}}(\bar{x})$. Letting s goes to infinity in (16), we obtain

$$0 = \sum_{j=1}^{m} \bar{z}_j \sum_{i\in\bar{I}_J} \bar{\alpha}_{ij} \mathrm{grad} f_{ij}(\bar{x}) + \bar{v}, \qquad \sum_{i\in\bar{I}_J} \bar{\alpha}_{ij} = 1.$$

Let $x \in \tilde{\Omega}$. Taking $u_j = \sum_{i\in\bar{I}_J} \bar{\alpha}_{ij} \mathrm{grad} f_{ij}(\bar{x})$ it follows from above equality that

$$0 = \sum_{j=1}^{m} \bar{z}_j \langle u_j, x - \bar{x}\rangle + \langle \bar{v}, x - \bar{x}\rangle. \tag{19}$$

As $\bar{x} + td \in \Omega_k$, for all $k = 0,1,\ldots$, definition of $\tilde{\Omega}$ in (17) implies that $\bar{x} + td \in \tilde{\Omega}$, $t \in (0,\delta]$. Since $u_j = \sum_{i\in\bar{I}_J} \bar{\alpha}_{ij} \mathrm{grad} f_{ij}(\bar{x})$ and $\sum_{i\in\bar{I}_J} \bar{\alpha}_{ij} = 1$, combining Proposition 2.3 with Lemma 2.2 we conclude that $u_j \in \partial^\circ f_j(\bar{x})$. Hence, using that $\bar{v} \in N_{\tilde{\Omega}}(\bar{x})$ and definition of $f_j^\circ(\bar{x},d)$, equality (19) with $x = \bar{x} + td$ yields

$$0 \leq \sum_{j=1}^{m} \bar{z}_j \langle u_j, d\rangle \leq \sum_{j=1}^{m} \bar{z}_j f_j^\circ(\bar{x},d).$$

So, there exists $j \in \{1,\ldots,m\}$ such that $f_j^\circ(\bar{x},d) \geq 0$, which contradicts (13). Therefore, $\bar{x}$ is Pareto-Clarke critical point and the first part is proved.

The second part follows from Lemma 3.3 and Lemma 3.1, which conclude the proof of the theorem. □

4 Conclusions

In this paper a convergence analysis of proximal point method for special class of nonconvex multiobjective functions is studied. In one sense, it is a continuation of [6], where the proximal point method, for special class of nonconvex functions, has been studied in the Riemannian context. We expect that the results of the present paper become a further step towards solving general multiobjective optimization. We foresee further progress in this topic in the nearby future.

Acknowledgements First author was partially supported by CAPES-MES-CUBA 226/2012, FAPEG 201210267000909 - 05/2012 and CNPq Grants 458479/2014-4, 471815/2012-8, 303732/2011-3, 236938/2012-6, 312077/2014-9. The second author was partially supported by FAPEG 201210267000909 - 05/2012, PRONEX–Optimization(FAPERJ/CNPq), CNPq Grants 4471815/2012-8, 305158/2014-7. The third author was partially supported by CAPES.

References

1. J. Y. Bello Cruz, L. R. Lucambio Pérez, and J. G. Melo. Convergence of the projected gradient method for quasiconvex multiobjective optimization. *Nonlinear Anal.*, 74(16):5268–5273, 2011.
2. G. C. Bento and J. X. Cruz Neto. A subgradient method for multiobjective optimization on Riemannian manifolds. *J. Optim. Theory Appl.*, 159(1):125–137, 2013.
3. G. C. Bento, J. X. Cruz Neto, P. R. Oliveira, and A. Soubeyran. The self regulation problem as an inexact steepest descent method for multicriteria optimization. *European J. Oper. Res.*, 235(3):494–502, 2014.
4. G. C. Bento, J. X. Cruz Neto, and A. Soubeyran. A proximal point-type method for multicriteria optimization. *Set-Valued Var. Anal.*, 22(3):557–573, 2014.
5. G. C. Bento, O. P. Ferreira, and P. R. Oliveira. Unconstrained steepest descent method for multicriteria optimization on Riemannian manifolds. *J. Optim. Theory Appl.*, 154(1):88–107, 2012.
6. G. C. Bento, O. P. Ferreira, and P. R. Oliveira. Proximal point method for a special class of nonconvex functions on Hadamard manifolds. *Optimization*, 64(2):289–319, 2015.
7. D. P. Bertsekas. *Convex analysis and optimization*. Athena Scientific, Belmont, MA, 2003. With Angelia Nedić and Asuman E. Ozdaglar.
8. H. Bonnel, A. N. Iusem, and B. F. Svaiter. Proximal methods in vector optimization. *SIAM J. Optim.*, 15(4):953–970 (electronic), 2005.
9. L. C. Ceng, B. S. Mordukhovich, and J. C. Yao. Hybrid approximate proximal method with auxiliary variational inequality for vector optimization. *J. Optim. Theory Appl.*, 146(2):267–303, 2010.
10. L.-C. Ceng and J.-C. Yao. Approximate proximal methods in vector optimization. *European J. Oper. Res.*, 183(1):1–19, 2007.
11. T. D. Chuong, B. S. Mordukhovich, and J.-C. Yao. Hybrid approximate proximal algorithms for efficient solutions in vector optimization. *J. Nonlinear Convex Anal.*, 12(2):257–286, 2011.
12. F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.
13. P. L. Combettes and T. Pennanen. Proximal methods for cohypomonotone operators. *SIAM J. Control Optim.*, 43(2):731–742 (electronic), 2004.

14. A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz, and L. N. Vicente. Direct multisearch for multiobjective optimization. *SIAM J. Optim.*, 21(3):1109–1140, 2011.
15. J. X. Da Cruz Neto, G. J. P. Da Silva, O. P. Ferreira, and J. O. Lopes. A subgradient method for multiobjective optimization. *Comput. Optim. Appl.*, 54(3):461–472, 2013.
16. L. M. G. Drummond and A. N. Iusem. A projected gradient method for vector optimization problems. *Comput. Optim. Appl.*, 28(1):5–29, 2004.
17. J. Fliege, L. M. Graña Drummond, and B. F. Svaiter. Newton's method for multiobjective optimization. *SIAM J. Optim.*, 20(2):602–626, 2009.
18. J. Fliege and B. F. Svaiter. Steepest descent methods for multicriteria optimization. *Math. Methods Oper. Res.*, 51(3):479–494, 2000.
19. E. H. Fukuda and L. M. Graña Drummond. On the convergence of the projected gradient method for vector optimization. *Optimization*, 60(8-9):1009–1021, 2011.
20. E. H. Fukuda and L. M. Graña Drummond. Inexact projected gradient method for vector optimization. *Comput. Optim. Appl.*, 54(3):473–493, 2013.
21. M. Fukushima and H. Mine. A generalized proximal point algorithm for certain nonconvex minimization problems. *Internat. J. Systems Sci.*, 12(8):989–1000, 1981.
22. T. Gal and T. Hanne. On the development and future aspects of vector optimization and MCDM. A tutorial. In *Multicriteria analysis (Coimbra, 1994)*, pages 130–145. Springer, Berlin, 1997.
23. R. Gárciga Otero and A. N. Iusem. Proximal methods in reflexive Banach spaces without monotonicity. *J. Math. Anal. Appl.*, 330(1):433–450, 2007.
24. L. M. Graña Drummond and B. F. Svaiter. A steepest descent method for vector optimization. *J. Comput. Appl. Math.*, 175(2):395–414, 2005.
25. A. N. Iusem, T. Pennanen, and B. F. Svaiter. Inexact variants of the proximal point algorithm without monotonicity. *SIAM J. Optim.*, 13(4):1080–1097, 2003.
26. A. Kaplan and R. Tichatschke. Proximal point methods and nonconvex optimization. *J. Global Optim.*, 13(4):389–406, 1998. Workshop on Global Optimization (Trier, 1997).
27. D. T. Luc. *Theory of vector optimization*, volume 319 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1989.
28. M. M. Mäkelä and P. Neittaanmäki. *Nonsmooth optimization*. World Scientific Publishing Co., Inc., River Edge, NJ, 1992. Analysis and algorithms with applications to optimal control.
29. T. Pennanen. Local convergence of the proximal point algorithm and multiplier methods without monotonicity. *Math. Oper. Res.*, 27(1):170–191, 2002.

30. R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
31. R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
32. J. E. Spingarn. Submonotone mappings and the proximal point algorithm. *Numer. Funct. Anal. Optim.*, 4(2):123–150, 1981/82.
33. K. D. V. Villacorta and P. R. Oliveira. An interior proximal method in vector optimization. *European J. Oper. Res.*, 214(3):485–492, 2011.