Proximal methods in vector optimization

Henri Bonnel[*], Alfredo Noel Iusem[†], Benar Fux Svaiter[‡]

June 12, 2003

Abstract

We consider the vector optimization problem of finding weakly efficient points for maps from a Hilbert space X to a Banach space Y, with respect to the partial order induced by a closed, convex and pointed cone $C \subset Y$ with nonempty interior. We develop for this problem an extension of the proximal point method for scalar-valued convex optimization. In this extension, the subproblems consist of finding weakly efficient points for suitable regularization of the original map. We present both an exact version an an inexact one, in which the subproblems are solved only approximately, within a constant relative tolerance. In both cases, we prove weak convergence of the generated sequence to a weakly efficient point, assuming convexity of the map with respect to C and C-completeness of the initial section. In cases where this last assumption fails, we still establish that the generating sequence is a minimizing one, in a suitable sense.

1 Introduction

We discuss in this paper methods for vector-valued optimization, in the following setting. We consider maps $F : X \to Y$, where X is a Hilbert space, and Y is a Banach space containing a closed, convex and pointed cone C, which defines a partial order $\preceq_C$ in Y, given by $y \preceq_C y'$ if and only if $y' - y$ belongs to C, with its associate relation $\prec_C$, given by $y \prec_C y'$ if and only if $y' - y$ belongs to the interior $\mathrm{int}(C)$ of C. Actually, we admit the possibility that F takes the value $+\infty$, in a sense made precise in Section 2. Our goal is to analyze methods for finding a weakly efficient minimizer of F with respect to $\preceq_C$, meaning a point $a \in X$ such that there exists no $x \in X$ satisfying $F(x) \prec_C F(a)$.

This paper is part of a wider research program consisting of the extension to vector-valued optimization of several iterative methods for scalar-valued ones. In these extensions, we attempt to define the iterates in the vector-valued case by considering the order $\preceq_C$ in Y, mimicking,

[*]Université de la Nouvelle-Calédonie, e-mail bonnel@univ-nc.nc (former name : Serban Bolintineanu)

[†]Instituto de Matématica Pura e Aplicada, e-mail iusp@impa.br

[‡]Instituto de Matématica Pura e Aplicada, e-mail benar@impa.br

whenever possible, the role of the usual order in $\mathbb{R}$ in the corresponding algorithm for scalar-valued optimization.

Several methods have been already extended in this fashion in a finite dimensional setting. The case of the steepest descent method for multi-objective optimization (i.e., when C is the nonnegative orthant of $\mathbb{R}^n$) was dealt with in [5]; the same method for general finite dimensional vector optimization (i.e., for partial orders given by rather general cones in $\mathbb{R}^n$) is analyzed in [7]. An extension of the projected gradient method to the case of convexly constrained vector optimization, for the order given by a rather general cone in $\mathbb{R}^n$, can be found in [6].

In this paper, we perform a similar extension for the case of the proximal point method for scalar-valued convex optimization. Thus, we give next a brief description of this method. Given a Hilbert space X and a point-to-set operator $T : X \to \mathcal{P}(X)$, the proximal point method, in its so called exact version, is an iterative procedure for finding a zero of T, i.e., a point $x^* \in X$ such that $0 \in T(x^*)$. The method generates a sequence $\{x_n\} \subset X$, starting from an arbitrary $x_0 \in X$, through the following iteration: given a bounded exogenous sequence of positive real numbers $\{\alpha_n\}$ (called regularization parameters) and the current iterate x_n, x_{n+1} is the unique vector in X such that

$$0 \in T_n(x_{n+1}) \tag{1}$$

where $T_n : X \to \mathcal{P}(X)$ is defined as $T_n(x) = T(x) + \alpha_n(x - x_n)$.

The origins of the proximal method can be traced back to the 60's (see, e.g. [11], [9] and [10]); it attained the formulation given above in Rockafellar's works in the 70's (see [12] and [13]).

It has been proved in [13] that whenever T is maximal monotone and has zeroes, the sequence $\{x_n\}$ generated by the proximal point method converges to a zero of T in the weak topology of X.

Take now a convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ and a closed and convex subset $K \subset X$. The solutions of the convex optimization problem min $f(x)$ subject to $x \in K$ are precisely the zeroes of the maximal monotone operator $T = \partial(f + I_K)$, where ∂ denotes the subdifferential of a convex function, and I_K is the indicator function of K, defined as $I_K(x) = 0$ if x belongs to K and $I_K(x) = +\infty$ otherwise. Thus, the proximal point method can be used to solve the above mentioned optimization problem; in this case, it is easy to check that the iteration (1) takes the form

$$x_{n+1} \in \operatorname{argmin}_{x \in K} \left\{ f(x) + \frac{\alpha_n}{2} \|x - x_n\|^2 \right\}. \tag{2}$$

We discuss next inexact versions of the method, in which case x_{n+1} needs not be the exact solution of the optimization subproblem in (2), but just an approximate solution of it. Clearly the analysis of an inexact version is essential if one wants the convergence results to hold for actual implementations of the method. Inexact versions were proposed as early as in [13], in which the n-th subproblem was allowed to be solved within a prescribed tolerance τ_n, and it was demanded that $\sum_{n=1}^{\infty} \tau_n < \infty$. Similar error criteria, requesting always summability of the tolerances, appeared in several papers later on; all these criteria are quite demanding and difficult to check in actual

implementations. More recently, less stringent error criteria, based on relative rather than absolute errors, were introduced by Solodov and Svaiter. In one of these approaches, the following iteration is considered for the problem min $f(x)$ subject to $x \in X$, with a convex $f : X \to \mathbb{R} \cup \{+\infty\}$: given x_n, take as x_{n+1} any vector $x \in X$ such that there exists $\varepsilon_n \in \mathbb{R}_+$ satisfying:

$$0 \in \partial_{\varepsilon_n} f(x) + \alpha_n (x - x_n), \tag{3}$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \|x - x_n\|^2 , \tag{4}$$

where $\sigma \in [0, 1)$ is a relative error constant.

We recall that, for $\varepsilon \geq 0$,

$$\partial_\varepsilon f(x) = \{u \in X : f(y) - f(x) - \langle u, x - y\rangle \geq -\varepsilon \ \ \forall\, y \in X\}.$$

The conditions in (3)-(4) just mean that $\alpha_n(x_n - x_{n+1})$ must belong to $\partial_{\varepsilon_n} f(x_{n+1})$, with the definition of ∂_{ε_n} given above, and with ε_n satisfying the inequality in (4) for $x = x_{n+1}$.

The algorithm given by (3)–(4) reduces to the exact version for the optimization case as in (2), when $\sigma = 0$ (or, equivalently, $\varepsilon_n = 0$ for all n).

This algorithm was developed in [14] and [4] for the more general problem of finding zeroes of operators (the exact version of this method is the one presented in (1)), where, instead of $\partial_{\varepsilon_n} f_n$, an adequate enlargement of the operator T_n, related to ε_n, is used. This type of enlargements was introduced in [3].

The results in [14] and [4] establish that the sequence generated by (3)–(4) converges to a minimizer of f in the weak topology of X under the same assumptions required in the exact case, namely convexity of f and existence of solutions of the optimization problem.

All the above discussion refers, of course, to the proximal method for scalar-valued convex optimization. The essence of this paper consists of the extension of both the exact proximal method (2) and its inexact counterpart (3)–(4) to the vector-valued optimization problem introduced at the beginning of this section. Basically, in the exact case the n-th subproblem will consist of finding weakly efficient minimizers of $F_n : X \to Y$, with

$$F_n(x) = F(x) + \alpha_n \|x - x_n\|^2 e_n, \tag{5}$$

restricted to the set $\Omega_n \subset X$ defined as $\Omega_n = \{x \in X : F(x) \preceq_C F(x_n)\}$, where e_n is an exogenously selected vector belonging to $\mathrm{int}(C)$ and such that $\|e_n\| = 1$.

For our inexact version, we consider the topological dual space Y^* of Y, the positive polar cone $C^+ \subset Y^*$, given by $C^+ = \{z \in Y^* : \langle y, z\rangle \geq 0 \ \forall\ y \in C\}$ (where $\langle \cdot, \cdot \rangle : Y \times Y^* \to \mathbb{R}$ is the duality pairing), and the indicator function I_{Ω_n}, of the set Ω_n, defined as above. We take an exogenous sequence $\{z_n\} \subset C^+$, with $\|z_n\| = 1$ for all $n \in \mathbb{N}$, and define, at iteration n, the function $f_n : X \to \mathbb{R} \cup \{+\infty\}$ as

$$f_n(x) = \langle F(x), z_n\rangle + I_{\Omega_n}(x). \tag{6}$$

Then we take as x_{n+1} any vector $x \in X$ such that there exists $\varepsilon_n \in \mathbb{R}_+$ satisfying:

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, z_n \rangle (x - x_n), \tag{7}$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \langle e_n, z_n \rangle \left\| x_n - x \right\|^2, \tag{8}$$

where $\sigma \in [0, 1)$ is again a measure of the relative error.

We will establish that any sequence generated by either our exact version or our inexact one converges, in the weak topology of X, to a weakly efficient minimizer of F under the following two assumptions:

i) F is C-convex, i.e., $F(\lambda x + (1-\lambda)x') \preceq_C \lambda F(x) + (1-\lambda)F(x')$ for all $x, x' \in X$ and all $\lambda \in [0, 1]$.

ii) The set $(F(x_0) - C) \cap F(X)$ is C-complete, i.e., for every sequence $\{a_n\} \subset X$, with $a_1 = x_0$, such that $F(a_{n+1}) \preceq_C F(a_n)$ for all $n \in \mathbb{N}$, there exists $a \in X$ such that $F(a) \preceq_C F(a_n)$ for all $n \in \mathbb{N}$.

In the absence of assumption (ii), we establish weaker convergence results, namely that the generated sequence is a minimizing one for our problem, meaning that $F\{(x_n)\}$ approaches the set of infimal values of F, in a sense made precise in Propositions 2 and 3 of Section 3.

We mention that the exact form of the algorithm is indeed a particular case of the inexact one, corresponding to choosing $\sigma = 0$, or, equivalently $\varepsilon_n = 0$ for all n, in the sense that any vector x_{n+1} satisfying (6)–(8) with $\sigma = 0$ is a weakly efficient minimizer of F_n as defined in (4). Thus, a separate analysis of the exact version might seem superfluous. We have included it because the analysis of the inexact version is somewhat obscured by the introduction of the sequences ε_n and z_n; the exact version is "cleaner" and it allows an easier visualization of the essential arguments in the convergence proofs. Also, both versions are presented in a somewhat different dressing: the subproblems of the exact one are (vector-valued) optimization problems; in each subproblem of the inexact version we look for zeroes of approximate subdifferentials of scalar-valued convex functions. This rather inessential difference is elaborated upon in Remark 2 of Section 4.

We emphasize that all our analysis holds without requiring reflexivity of the Banach space Y. This detail is not negligible, because we do need the cone C to have nonempty interior. The prototypical infinite dimensional Banach spaces are the $\mathcal{L}^p$ spaces ($1 \leq p \leq \infty$), and the most relevant cones in them are the so called positive cones, consisting of all p-integrable functions which are nonnegative almost everywhere. As it is well known, these cones have empty interior, excepting for the case of $\mathcal{L}^\infty$, which happens to be nonreflexive. Thus, our analysis covers at least one meaningful example where the order is induced by a cone in an infinite dimensional space.

The paper is organized as follows. Section 2 states formally the problem and introduces some required preliminary material. The exact version of the algorithm is analyzed in Section 3 and the inexact one in Section 4. Section 5 contains several final remarks.

2 Problem statement and notation

Throughout this paper X is a real Hilbert space, Y is a real Banach space, $\langle \cdot, \cdot \rangle$ denotes the scalar product in X as well as the duality scalar product between Y^* (the topological dual of Y) and Y. Any norm will be denoted by $\| \cdot \|$. $C \subset Y$ is a pointed closed convex cone whose topological interior $\mathrm{int}(C)$ is nonempty. F is an extended-valued map from X to $Y \cup \{+\infty_C\}$. The extended space $\bar{Y} = Y \cup \{-\infty, +\infty\}$ was introduced in [1], where a neighborhood of $+\infty_C$ is defined as a set $N \subset \bar{Y}$ containing $a + C \cup \{+\infty_C\}$ for some $a \in Y$, and its opposite $-N$ is a neighborhood of $-\infty_C$. The partial order relation of Y given by

$$\forall y, y' \in Y \qquad y \preceq_C y' \Longleftrightarrow y' - y \in C,$$

and the transitive relation

$$\forall y, y' \in Y \qquad y \prec_C y' \Longleftrightarrow y' - y \in \mathrm{int}(C),$$

are extended to $\bar{Y}$ by

$$\forall y \in Y \qquad -\infty_C \prec_C y \prec_C +\infty_C, \qquad -\infty_C \preceq_C y \preceq_C +\infty_C.$$

Note that the embedding $Y \subset \bar{Y}$ is continuous and dense. The map F is assumed to be proper, i.e. not identically equal to $+\infty_C$.

We denote by $\mathrm{dom}\,(F) = \{x \in X |\ F(x) \neq +\infty_C\}$ the effective domain of F, and by $C^+ = \{z \in Y^* |\ \langle y, z \rangle \geq 0\ \forall\ y \in Y\}$ the positive polar cone of C. We extend by continuity every $z \in C^+ \setminus \{0\}$ to $\bar{Y}$, by putting $\langle \pm\infty_C, z \rangle = \pm\infty$ (see [1] and [2] for more details). For a set $T \subset \bar{Y}$, we denote its topological closure in the topological space $\bar{Y}$ by $\bar{T}$. The *infimal set* (resp. *weakly infimal set*) of a subset $T \subset \bar{Y}$, is the set

$$C\text{-INF}\,(T) = \{y \in \bar{T} |\ \nexists z \in T \setminus \{y\},\ z \preceq_C y\}\ (\text{resp. } C\text{-INF}_w\,(T) = \{y \in \bar{T} |\ \nexists z \in T,\ z \prec_C y\}).$$

A point $a \in X$ is called an *efficient (or Pareto)* (resp. *weakly efficient (or weakly Pareto)) solution* for a vector optimization problem of the form $C - \mathrm{INF}_w\,\{G(x) |\ x \in S\}$, where $G : S \to Y \cup \{+\infty_C\}$ and $S \subset X$, if $a \in S$, and there exists no $x \in S$ verifying $G(x) \preceq_C G(a)$, $G(x) \neq G(a)$ (resp. if $a \in S$, and there exists no $x \in S$ verifying $G(x) \prec_C G(a)$); in other words, if $a \in S$, $G(a) \in \mathrm{INF}\,(G(S))$ (resp. $a \in S$, $G(a) \in \mathrm{INF}_w\,(G(S))$). The set of efficient (resp. weakly efficient) solutions will be denoted $C - \mathrm{ARGINF}\,\{G(x) \mid x \in S\}$ (resp. $C - \mathrm{ARGINF}_w\,\{G(x) \mid x \in S\}$).

For $y \in Y$, $T \subset Y \cup \{+\infty_C\}$, $T \neq \{+\infty_C\}$ we denote $\mathrm{d}(y, T) = \inf\{\|y - z\| \mid z \in T \cap Y\}$.

We deal with the following vector optimization problem (VOP):

$$C - \mathrm{INF}_w\,\{F(x) |\ x \in X\}.$$

Notice that any *constrained* vector optimization problem (CVOP)

$$C - \mathrm{INF}_w\, F_0(x) \ \text{ subject to } x \in S$$

where $S \subset X$ is the feasible set, and F_0 is a map from S to Y, is equivalent to the unconstrained extended-valued vector optimization problem (VOP) with

$$F(x) = \begin{cases} F_0(x) & \text{if } x \in S \\ +\infty_C & \text{if } x \in X \setminus S. \end{cases}$$

The problems (CVOP) and (VOP) are equivalent in the sense that they have the same weakly efficient solutions, and the same weakly infimal set.

A map $G : S \subset X \to Y \cup \{+\infty_C\}$, where S is a convex set, is called $C-$*convex* if :

$$\forall x,\, x' \in S,\ \lambda \in [0,1] \qquad G((1-\lambda)x + \lambda x') \preceq_C (1-\lambda)G(x) + \lambda G(x')$$

(with the conventions $\infty_C + \infty_C = \infty_C$, $\lambda \cdot (+\infty_C) = +\infty_C$, for each positive number λ).

A map $G : X \to Y \cup \{+\infty_C\}$ will be called *positively lower semicontinuous*, if for every $z \in C^+$, the extended valued scalar function $x \mapsto \langle G(x), z\rangle$ is lower semicontinuous.

Throughout this paper we consider a $C-$convex vector optimization problem (VOP), which means that the map F is C-convex.

We recall the following scalarization result (see e.g. [8]) known for finite vector-valued maps, but which is immediately generalized to extended valued vector ones.

Theorem 1. *If $S \subset X$ is a convex set and $G : S \to Y \cup \{+\infty_C\}$ is a $C-$convex proper map, then*

$$C - \mathrm{ARGINF}_w \{G(x) \mid \ x \in S\} = \bigcup_{z \in C^+ \setminus \{0\}} \mathrm{argmin}\{\langle G(x), z\rangle \mid x \in S\}.$$

3 An exact proximal algorithm

The exact version of our proximal algorithm (to be called **ALG1**) requires two exogenous sequences: a bounded sequence of positive real numbers $\{\alpha_n\}$, and a sequence $\{e_n\} \subset \mathrm{int}(C)$ such that $\|e_n\| = 1$ for all n. The method generates a sequence $\{x_n\} \subset X$ in the following way:

Initialization: Choose $x_0 \in \mathrm{dom}\,(F)$.
Stopping rule: Given x_n, if $x_n \in C - \mathrm{ARGINF}_w \{F(x) |\ x \in X\}$, then $x_{n+p} = x_n$ for all $p \geqslant 1$.
Iterative step: Given x_n, if $x_n \notin C - \mathrm{ARGINF}_w \{F(x) |\ x \in X\}$, then take as the next iterate any $x_{n+1} \in X$ such that

$$x_{n+1} \in C - \mathrm{ARGINF}_w \left\{ F(x) + \frac{\alpha_n}{2}\|x - x_n\|^2 e_n \,\middle|\, x \in \Omega_n \right\}, \tag{9}$$

with $\Omega_n := \{x \in X |\ F(x) \preceq_C F(x_n)\}$.

We make now the following assumption on the map F and the initial iterate x_0:
A) The set $(F(x_0) - C) \cap F(X)$ is *C-complete*, meaning that for all sequence $\{a_n\} \subset X$, with $a_1 = x_0$, such that $F(a_{n+1}) \preceq_C F(a_n)$ for all $n \in \mathbb{N}$, there exists $a \in X$ such that $F(a) \preceq_C F(a_n)$ for all $n \in \mathbb{N}$.

Next, we state and prove our main convergence result.

Theorem 2. *Let $F : X \to Y \cup \{+\infty_C\}$ be a proper $C-$convex map, positively lower semicontinuous. Under assumption (A) above, any sequence $\{x_n\}$ generated by algorithm* **ALG1** *is convergent, with respect to the weak topology of X, to a weakly efficient solution of problem (VOP).*

Proof. We divide the proof in five steps.
Step 1: existence of iterates. x_0 is chosen in the initialization step. Assuming that the algorithm reached iteration n, we show next that an appropriate x_{n+1} exists. In view of the stopping rule, this is certainly the case if x_n is weakly efficient for (VOP). Otherwise, take any $z \in C^+ \setminus \{0\}$. Since e_n belongs to $\text{int}(C)$, it follows from the definition of C^+ that $\langle e_n, z\rangle > 0$. Define $\varphi_n : X \to \mathbb{R} \cup \{+\infty\}$ as

$$\varphi_n(x) = \langle F(x), z\rangle + I_{\Omega_n}(x) + \frac{\alpha_n}{2}\langle e_n, z\rangle \|x - x_n\|^2 . \tag{10}$$

C-convexity of F implies convexity of $\langle F(\cdot), z\rangle$ and convexity of Ω_n. Lower semicontinuity of F implies closedness of Ω_n, so that $\langle F(\cdot), z\rangle + I_{\Omega_n}$ is convex and lower semicontinuous. Since $\langle e_n, z\rangle > 0$, φ_n is strongly convex, and the existence of minimizers of φ_n results from the standard argument for existence of iterates of the scalar-valued proximal method (see [13]): the subdifferential of φ_n is maximal monotone and strongly monotone, and hence onto, by virtue of Minty's theorem. Thus, the subdifferential has some zero, which is a minimizer of φ_n. In view of Theorem 1, such minimizer satisfies (9) and can be taken as x_{n+1}.

Step 2: Fejér convergence to the set of lower bounds of the initial section. If the stopping rule applies at some iteration, then the sequence remains constant thereafter, and thus it is strongly convergent to the stopping iterate, which is a weakly efficient solution of (VOP). We assume from now on that the stopping rule never applies. Define $E \subset X$ as:

$$E = \{x \in X : F(x) \preceq_C F(x_n)\ \forall\ n \in \mathbb{N}\}.$$

By assumption (A), E is nonempty. Now, since x_n solves the vector optimization problem in (9), by Theorem 1 there exists $z_n \in C^+ \setminus \{0\}$ such than x_{n+1} solves the problem

$$\min \varphi_n(x) \tag{11}$$

$$x \in \Omega_n, \tag{12}$$

where φ_n is as in (10), with z_n substituting for z. Since the solution of (11)–(12) is not altered through multiplication of z_n by positive scalars, we can assume without loss of generality that $\|z_n\| = 1$ for all $n \in \mathbb{N}$. Since x_n is not weakly efficient, there exists some $x \in X$ such that $F(x) \prec_C F(x_n)$ so that Ω_n has nonempty relative interior. In view of this constraint qualification, it follows that x_{n+1} satisfies the first order optimality conditions for problem (11)–(12), i.e., there exists $u_n \in X$ such that

$$u_n \in \partial\varphi_n(x_{n+1}) \tag{13}$$

and

$$0 \leq \langle u_n, x - x_{n+1}\rangle \tag{14}$$

for all $x \in \Omega_n$. Define now $\psi_n : X \to \mathbb{R} \cup \{+\infty\}$ as

$$\psi_n(x) = \langle F(x), z_n\rangle. \tag{15}$$

In view of (10) and (13), we have

$$u_n = v_n + \alpha_n\langle e_n, z_n\rangle(x_{n+1} - x_n), \tag{16}$$

for some

$$v_n \in \partial\psi_n(x_{n+1}). \tag{17}$$

Take now $x^* \in E$. By definition of E, x^* belongs to Ω_n for all $n \in \mathbb{N}$. Combining (14) with $x = x^*$ and (16), we have

$$0 \leq \langle v_n, x^* - x_{n+1}\rangle + \alpha_n\langle e_n, z_n\rangle\langle x_{n+1} - x_n, x^* - x_{n+1}\rangle \leq$$

$$\langle F(x^*) - F(x_{n+1}), z_n\rangle + \alpha_n\langle e_n, z_n\rangle\langle x_{n+1} - x_n, x^* - x_{n+1}\rangle \leq \alpha_n\langle e_n, z_n\rangle\langle x_{n+1} - x_n, x^* - x_{n+1}\rangle, \tag{18}$$

using (15) and (17) in the second inequality, and the facts that x^* belongs to E and z_n belongs to $C^+\backslash\{0\}$ in the third one: it is clear that $F(x^*) - F(x_{n+1}) \preceq_C 0$ and therefore $\langle F(x^*) - F(x_{n+1}), z_n\rangle \leq 0$.

Define now $\gamma_n = \alpha_n\langle e_n, z_n\rangle$. Note that $\gamma_n > 0$, because $\alpha_n > 0$, e_n belongs to $\mathrm{int}(C)$ and z_n belongs to $C^+ \setminus \{0\}$. By (18),

$$0 \leq \gamma_n\langle x_{n+1} - x_n, x^* - x_{n+1}\rangle = \gamma_n\left(\|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 - \|x_n - x_{n+1}\|^2\right).$$

implying, since $\gamma_n > 0$,

$$\|x_{n+1} - x^*\|^2 \leq \|x_n - x^*\|^2 - \|x_n - x_{n+1}\|^2. \tag{19}$$

Step 3: boundedness of the sequence and proximity of consecutive iterates. This step is identical to the corresponding one in the convergence proof of the scalar-valued method. We present it for the sake of completeness. It follows from (19) that, for all $x^* \in E$, $\{\|x_n - x^*\|^2\}$ is a nonnegative and nonincreasing sequence, hence convergent. Rewriting (19) as

$$\|x_n - x_{n+1}\|^2 \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 , \tag{20}$$

and observing that the right hand side of (20) converges to 0 as $n \to \infty$ because $\{\|x_n - x^*\|^2\}$ is convergent, we conclude that

$$\lim_{n\to\infty} (x_n - x_{n+1}) = 0. \tag{21}$$

It follows from (19) that

$$\|x_n - x^*\| \leq \|x_0 - x^*\| \tag{22}$$

for all $x^* \in E$. Since E is nonempty, we conclude that $\{x_n\}$ is bounded.

Step 4: optimality of the weak cluster points of $\{x_n\}$. Since $\{x_n\}$ is bounded, it has weak cluster points. We will prove next that all of them are weakly efficient solutions of (VOP). Let $\hat{x}$ be a weak cluster point of $\{x_n\}$ and $\{x_{k_n}\}$ a sequence weakly convergent to it.

Define $\psi_z : X \to \mathbb{R}$ as $\psi_z(x) = \langle F(x), z\rangle$. We claim that

$$\psi_z(\hat{x}) \leq \psi_z(x_n) \tag{23}$$

for all $z \in C^+$ and all $n \in \mathbb{N}$. Since F is positively lower semicontinuous and C-convex, ψ_z is lower semicontinuous and convex, so that $\psi_z(\hat{x}) \leq \lim_{n\to\infty} \psi_z(x_{k_n})$. Since $F(x_{n+1}) \preceq_C F(x_n)$ for all n, we have $\psi_z(x_{n+1}) \leq \psi_z(x_n)$ for all n, so that $\lim_{n\to\infty} \psi_z(x_{k_n}) = \inf\{\psi_z(x_n)\}$. The claim follows from the last two statements. It follows easily from (23) that

$$F(\hat{x}) \preceq_C F(x_n) \tag{24}$$

for all $n \in \mathbb{N}$.

Assume that $\hat{x}$ is not weakly efficient for (VOP), i.e., that there exists $\bar{x} \in X$ such that $F(\bar{x}) \prec_C F(\hat{x})$. Take z_n as chosen before equation (11). Since $\|z_n\| = 1$ for all n, there exists, by virtue of Bourbaki-Alaoglu's Theorem, a weak* cluster point of $\{z_n\}$, say $\bar{z}$, weak* limit of the subsequence $\{z_{j_n}\}$. We claim now that C^+ is weakly* closed. Observe that $C^+ = \cap_{y\in C}\{z \in Y^* : \langle y, z\rangle \geq 0\}$. Since the linear forms $z \mapsto \langle y, z\rangle$ are weakly* continuous for all $y \in Y$, we have written C^+ as an intersection of weakly* closed sets, establishing the claim. It follows that $\bar{z}$ belongs to C^+. Thus

$$\langle F(\bar{x}) - F(\hat{x}), z_{j_n}\rangle \geq \langle F(\bar{x}) - F\left(x_{j_n+1}\right), z_{j_n}\rangle = \psi_n(\bar{x}) - \psi_n\left(x_{j_n+1}\right) \geq$$

$$\langle v_{j_n}, \bar{x} - x_{j_n+1}\rangle = \langle u_{j_n}, \bar{x} - x_{j_n+1}\rangle - \gamma_n\langle, x_{j_n+1} - x_{j_n}, \bar{x} - x_{j_n+1}\rangle \geq$$

$$\gamma_n \langle , x_{j_n+1} - x_{j_n}, \bar{x} - x_{j_n+1} \rangle \geq -\gamma_n \left\| x_{j_n+1} - x_{j_n} \right\| \ \left\| \bar{x} - x_{j_n+1} \right\|. \tag{25}$$

using (23) in the first inequality, (17) in the second one, (16) in the second equality, and (14) in the third inequality, together with the fact that $\bar{x}$ belongs to Ω_n for all $n \in \mathbb{N}$, because $F(\bar{x}) \prec_C F(\hat{x}) \preceq_C (x_n)$, in view of (24).

We will take limits now in the extreme expressions of (25). Regarding the rightmost one, since $\{\alpha_n\}$ is bounded, and $\|z_n\| = \|e_n\| = 1$, we have that $\{\gamma_n\}$ is bounded. Note that $\{x_n\}$ is bounded, so that $\{\|\bar{x} - x_{j_n+1}\|\}$ is also bounded. Finally, $\lim_{n\to\infty} \|x_{j_n} - x_{j_n+1}\| = 0$, by (21). We conclude that the limit of the rightmost expression in (25) as $n \to \infty$ vanishes, and therefore, since $\bar{z}$ is the weak* limit of $\{z_{j_n}\}$, we conclude from (25) that

$$\langle \bar{F}(\bar{x}) - F(\hat{x}), \bar{z} \rangle \geq 0. \tag{26}$$

Now we claim that $\bar{z} \neq 0$. Take $e \in \text{int}(C)$. It follows from Lemma 2.2 of [2] that $\langle e, z_n \rangle \geq d(e, Y \setminus C) > 0$ for all $n \in \mathbb{N}$. Since $\bar{z}$ is the weak* limit of$\{z_{j_n}\}$, we conclude that $\langle e, \bar{z} \rangle > 0$, establishing the claim. Since $\bar{z} \neq 0$, it is clear that (26) contradicts the fact that $\bar{z}$ belongs to C^+ and the assumption that $F(\bar{x}) \prec_C F(\hat{x})$. Thus, such assumption is false, and $\hat{x}$ is indeed weakly efficient for problem (VOP).

Step 5: uniqueness of the weak cluster point of $\{x_n\}$. This part of the proof, presented also for the sake of completeness, is again identical to the scalar-valued case, as in [13], and it uses Brézis' uniqueness argument. Suppose that both $\hat{x}$ and $\tilde{x}$ are weak cluster points of $\{x_n\}$. With the argument as in(24), both $\hat{x}$ and $\tilde{x}$ belong to E, so that both $\{\|\hat{x} - x_n\|\}$ and $\{\|\tilde{x} - x_n\|\}$ converge, as shown at the beginning of Step 3, i.e., there exist $\hat{\beta}, \tilde{\beta} \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \|x_n - \hat{x}\| = \hat{\beta}, \qquad \lim_{n\to\infty} \|x_n - \tilde{x}\| = \tilde{\beta}. \tag{27}$$

Since

$$\|x_n - \hat{x}\|^2 = \|x_n - \tilde{x}\|^2 + 2\langle x_n - \tilde{x}, \tilde{x} - \hat{x} \rangle + \|\tilde{x} - \hat{x}\|^2$$

we conclude from (27) that

$$\lim_{n\to\infty} \langle x_n - \tilde{x}, \tilde{x} - \hat{x} \rangle = \frac{1}{2} \left(\hat{\beta}^2 - \tilde{\beta}^2 - \|\tilde{x} - \hat{x}\|^2 \right). \tag{28}$$

The left hand side of (28) vanishes, because $\tilde{x}$ is a weak cluster point of $\{x_n\}$, so that

$$\hat{\beta}^2 - \tilde{\beta}^2 = \|\tilde{x} - \hat{x}\|^2. \tag{29}$$

Reversing the roles of $\hat{x}$ and $\tilde{x}$, a similar reasoning leads to $\tilde{\beta}^2 - \hat{\beta}^2 = \|\tilde{x} - \hat{x}\|^2$, which, combined with (29), gives $\|\tilde{x} - \hat{x}\| = 0$, i.e., $\tilde{x} = \hat{x}$, establishing uniqueness of the weak cluster points of $\{x_n\}$. It follows that $\{x_n\}$ is weakly convergent to a weakly efficient solution of (VOP) and the proof is complete. □

The next result is connected to the possibility of replacing the stopping rule of **ALG1** by the following one, which is easier to check: after computing x_{n+1}, the algorithm stops if $x_{n+1} = x_n$ (i.e. we set $x_{n+p} = x_n$ for all $p \geq 0$ in such a case). Under this new rule, if the algorithm stops at iteration n then x_n is a weakly Pareto solution of the problem. On the other hand, this rule might fail to recognize weakly Pareto solutions: the sequence might hit such a solution at iteration n and proceed with the iterative process, eventually generating a whole sequence of different weakly Pareto solutions. However, this situation is not possible if x_n is a Pareto solution (not just a weakly Pareto one), in which case, as shown in the next proposition, the following iterate x_{n+1} coincides with x_n.

Proposition 1. *Let x_n be the $n-th$ iterate of a sequence generated by Algorithm* **ALG1**.

1. *If* $x_n \in C-\mathrm{ARGINF}_w \left\{F(x)+\frac{\alpha_n}{2}\|x-x_n\|^2 e_n \middle|\ x \in \Omega_n\right\}$, *then* $x_n \in C-\mathrm{ARGINF}_w \{F(x)|\ x \in X\}$.

2. *If* $x_n \in C - \mathrm{ARGINF}\{F(x)|\ x \in X\}$, *then* $C - \mathrm{ARGINF}_w \left\{F(x) + \frac{\alpha_n}{2}\|x - x_n\|^2 e_n \middle|\ x \in \Omega_n\right\} = \{x_n\}$.

Proof. 1. Suppose that $x_n \in C - \mathrm{ARGINF}_w \left\{F(x) + \frac{\alpha_n}{2}\|x - x_n\|^2 e_n \middle|\ x \in \Omega_n\right\}$, but $x_n \notin C - \mathrm{ARGINF}_w \{F(x)|\ x \in X\}$. Then, there exists some $\bar{x} \in X$ such that $F(\bar{x}) \prec_C F(x_n)$, i.e., there exists some $c \in \mathrm{int}(C)$ such that $F(\bar{x}) = F(x_n) - c$. For $t \in [0,1)$, define $x_t = (1-t)\bar{x} + tx_n$. By C-convexity of F,

$$F(x_t) \preceq_C (1-t)F(\bar{x}) + tF(x_n) = F(x_n) - (1-t)c. \tag{30}$$

Since x_n solves the vector optimization problem in (9) (with x_n instead of x_{n+1}), by Theorem 1 there exists $z_n \in C^+ \setminus \{0\}$ such that

$$\varphi_n(x_n) = \min_{x \in \Omega_n} \varphi_n(x) \leq \varphi_n(x_t), \tag{31}$$

where φ_n is as in (10), with z_n substituting for z. In view of (30) and (31), and calling as before $\gamma_n = \alpha_n \langle e_n, z_n \rangle$, we have

$$\langle F(x_n), z_n \rangle \leq \langle F(x_t), z_n \rangle + \frac{\gamma_n}{2}\|x_t - x_n\|^2 = \langle F(x_t), z_n \rangle + (1-t)^2 \frac{\gamma_n}{2}\|\bar{x} - x_n\|^2 \leq$$

$$\langle F(x_n), z_n \rangle - (1-t)\langle c, z_n \rangle + (1-t)^2 \frac{\gamma_n}{2}\|\bar{x} - x_n\|^2. \tag{32}$$

It follows from (32) that

$$\langle c, z_n \rangle \leq (1-t)\frac{\gamma_n}{2}\|\bar{x} - x_n\|^2 \tag{33}$$

for all $t \in [0,1)$. Letting $t \to 1^-$, we obtain from (33) that $\langle c, z_n \rangle \leq 0$, which is a contradiction.

2. Suppose now that $x_n \in C - \text{ARGINF}\{F(x) |\ x \in X\}$, i.e., that x_n is a Pareto solution of the original problem. It is obvious that $x_n \in C - \text{ARGINF}_w \{F(x) + \frac{\alpha_n}{2}\|x - x_n\|^2 e_n \Big|\ x \in \Omega_n\}$. Let $\bar{x} \in C - \text{ARGINF}_w \{F(x) + \frac{\alpha_n}{2}\|x - x_n\|^2 e_n \Big|\ x \in \Omega_n\}$. Since $\bar{x} \in \Omega_n$ and x_n is a Pareto solution, we must have $F(\bar{x}) = F(x_n)$. On the other hand, by Theorem 1, there exists $z \in C^+ \setminus \{0\}$ such that

$$\langle F(\bar{x}), z \rangle + \frac{\alpha_n}{2}\langle e_n, z \rangle \|\bar{x} - x_n\|^2 = \min_{x \in \Omega_n} (\langle F(x), z \rangle + \frac{\alpha_n}{2}\langle e_n, z \rangle \|x - x_n\|^2) \leq \langle F(x_n), z \rangle = \langle F(\bar{x}), z \rangle.$$

It follows immediately that $\bar{x} = x_n$. □

We discuss next a couple of issues connected to assumption (A), i.e., the C-completeness hypothesis. First, note that in the scalar minimization problems ($Y = \mathbb{R}$, $C = [0, +\infty)$), when the set of minimizers is not empty, every section $(F(x_0) - C) \cap F(X)$ is $C-$complete. We also the have the following observation.

Remark 1. If $\dim Y < +\infty$, $F(X)$ and $F(X) + C$ are closed, then according to [8, Corollary 4.5, p. 54, and Theorem 3.3 p. 46], it holds that $(F(x_0) - C) \cap F(X)$ is $C-$complete for all $x_0 \in \text{dom}\,(F)$ if and only if there exist efficient solutions for (VOP).

In practical situations, assumption (A) could be difficult to verify. Thus, it is interesting to establish some convergence results, albeit weaker than Theorem 2, for situations in which this assumption fails to hold. This is the content of the following two propositions.

Proposition 2. *Let $F : X \to Y \cup \{+\infty_C\}$ be a proper $C-$convex map, positively lower semicontinuous. If the sequence $\{x_n\}$ generated by algorithm* **ALG1** *has a weak cluster point, (without assuming that (A) holds) then it is weakly convergent to a weakly efficient solution of the problem (VOP).*

Proof. By assumption, there exists a weakly convergent subsequence $\{x_{k_n}\}$ of $\{x_n\}$ whose weak limit is some $\hat{x} \in X$. Since the sequence $\{F(x_n)\}$ is $C-$nonincreasing, we have that for all $z \in C^+$, the sequence $\{\langle F(x_{k_n}), z \rangle\}$ is nonincreasing, and since the scalar convex function $\langle F(\cdot), z \rangle$ is weakly lower semicontinuous, we obtain that $\langle F(\hat{x}), z \rangle$ is a lower bound for $\{\langle F(x_{k_n}), z \rangle\}$. Since this result holds for all $z \in C^+$, we get that $F(\hat{x}) \preceq_C F(x_{k_n})$ for all $n \in \mathbb{N}$, implying that $F(\hat{x}) \preceq_C F(x_n)$. From this point on, the argument continues as in the proof of Theorem 2 after (24). □

Proposition 3. *Let $F : X \to Y \cup \{+\infty_C\}$ be a proper $C-$convex map, positively lower semicontinuous. Assume that* $\dim Y < \infty$, *and that int$(C^+) \neq \emptyset$. Let $\{x_n\}$ be a sequence generated by Algorithm* **ALG1**. *Suppose that $\{x_n\}$ is unbounded, and that the set of cluster points of $\{F(x_n)\}$ is finite and nonempty. Then the sequence $\{F(x_n)\}$ converges to an element of the weakly-infimal set* $\text{INF}_w\,(F(X))$ *of the problem (VOP).*

Proof. Note first that, in view of Proposition 2, our assumption implies that $\{x_n\}$ is not weakly convergent.

Consider a subsequence $\{F(x_{k_n})\}$ convergent to some $\bar{y} \in Y$. We want to prove that the whole sequence $\{F(x_n)\}$ converges to $\bar{y}$. Note that $\bar{y}$ belongs to $\overline{F(X)}$. For every $z \in C^+$, the sequence $\{\langle F(x_n), z\rangle\}$ is convergent to $\langle \bar{y}, z\rangle$, because it is nonincreasing and admits a subsequence converging to $\langle \bar{y}, z\rangle$. On the other hand, since $\{F(x_n)\}$ is C-nonincreasing, and C is closed, it follows that $\bar{y}$ is a $C-$lower bound for this sequence. Next we take some $z \in \text{int}(C^+)$, and, using Lemma 2.4 of [2], we get

$$\langle F(x_n) - \bar{y}, z\rangle \geqslant \ \mathrm{d}(z, Y^* \setminus C^+)\, \|F(x_n) - \bar{y}\| \tag{34}$$

for all $n \in \mathbb{N}$. Since the left hand side of (34) goes to zero as $n \to \infty$, and $\mathrm{d}(z, Y^* \setminus C^+) > 0$, we obtain that $\{F(x_n)\}$ is convergent to $\bar{y}$.

If $\bar{y} \notin \text{INF}_w\,(F(X))$, then there exists $\bar{x} \in X$ with $F(\bar{x}) \prec_C \bar{y}$. Using the same argument as in the proof of Proposition 2, we get that $\bar{y} \preceq_C F(x_n)$ for all $n \in \mathbb{N}$. Then, we can establish weak convergence of $\{x_n\}$ to a weakly efficient solution of (VOP), with the same argument as in the proof of Theorem 2 from (24) on. This contradicts the assumption that $\{x_n\}$ is unbounded, establishing the result. □

4 An inexact proximal algorithm

We present now the inexact version of our algorithm, to be called **ALG2**. It requires an exogenous relative error tolerance $\sigma \in [0, 1)$ and three exogenous sequences: the ones required by **ALG1**, i.e., a bounded sequence of positive real numbers $\{\alpha_n\}$ and a sequence $\{e_n\} \subset \text{int}(C)$ such that $\|e_n\| = 1$ for all n, and now also a sequence $\{z_n\} \subset C^+$ such that $\|z_n\| = 1$ for all $n \in \mathbb{N}$. The method generates a sequence $\{x_n\} \subset X$ in the following way:

Initialization: Choose $x_0 \in \text{dom}\,(F)$.
Stopping rule: Given x_n, if $x_n \in C - \text{ARGINF}_w\,\{F(x)|\ x \in X\}$, then $x_{n+p} = x_n$ for all $p \geqslant 1$.
Iterative step: Given x_n, if $x_n \notin C - \text{ARGINF}_w\,\{F(x)|\ x \in X\}$, then let $\Omega_n = \{x \in X : F(x) \preceq_C F(x_n)\}$ and define $f_n(x) = \langle F(x), z_n\rangle + I_{\Omega_n}(x)$. Take as x_{n+1} any vector $x \in X$ such that there exists $\varepsilon_n \in \mathbb{R}_+$ satisfying

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, z_n\rangle (x - x_n), \tag{35}$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \langle e_n, z_n\rangle \, \|x - x_n\|^2 . \tag{36}$$

Remark 2. We mention now that the difference between the presentation of the iterative steps in **ALG1** and **ALG2** is not substantial, indeed. The subproblems of **ALG1** require finding weakly

efficient points for a regularized vector-valued function; the subproblems of **ALG2** demand ε-subgradients of a scalar-valued one, which could be seen, in the light of Theorem 1, as approximate weakly efficient points of the vector-valued one. We chose the presentation above in order to avoid the possibly cumbersome tasks of defining, for vector-valued maps, either some kind of approximate weakly efficient points or some kind of ε-subgradients. As discussed in Remark 5 of Section 5, the subproblems of **ALG2**, despite its scalar-valued presentation, in some cases are more suitable for being solved with algorithms specifically devised for vector optimization.

The convergence result for **ALG2** is the following.

Theorem 3. *Let $F : X \to Y \cup \{+\infty_C\}$ be a proper $C-$convex map, positively lower semicontinuous. Under assumption (A), introduced just before the statement of Theorem 2, any sequence $\{x_n\}$ generated by algorithm* **ALG2** *is convergent, with respect to the weak topology of X, to a weakly efficient point for problem (VOP).*

Proof. Step 1: existence of the iterates. Recall that $\partial f(x) = \partial_0 f(x) \subset \partial f_\varepsilon(x)$ for all convex $f : X \to \mathbb{R} \cup \{+\infty\}$, all $x \in X$ and all $\varepsilon \in \mathbb{R}_+$. Thus, assuming some x_n exists, the strongly convex function

$$\bar{f}_n(x) := f_n(x) + \frac{\alpha_n}{2}\langle e_n, z_n\rangle \|x - x_n\|^2$$

has a strongly convex subdifferential, which has a zero by virtue of Minty's Theorem. Such zero satisfies the inclusion in (35), with $\varepsilon_n = 0$, which trivially satisfy (36).

Step 2: Fejér convergence to the set of lower bounds of the initial section. Let again $\gamma_n = \alpha_n\langle e_n, z_n\rangle$. In view of (35),

$$\gamma_n(x_n - x_{n+1}) \in \partial_{\varepsilon_n} f_n(x_{n+1}). \tag{37}$$

Take any $x^* \in E$. Then, in view of (37) and the definition of ∂_{ε_n}, we have

$$-\varepsilon_n + \gamma\langle x_n - x_{n+1}, x^* - x_{n+1}\rangle \leq f_n(x^*) - f_n(x_{n+1}). \tag{38}$$

Note that, since x^* belongs to $E \subset \Omega_n$, we have that $x^* \in \Omega_n$, so that

$$f_n(x^*) = \langle F(x^*), z_n\rangle, \tag{39}$$

and also

$$\langle F(x^*) - F(x_{n+1}), z_n\rangle \leq 0, \tag{40}$$

because $F(x^*) \preceq_C F(x_{n+1})$ and $z_n \in C^+$. Note also that, in view of the definitions of f_n and I_{Ω_n}, we have $\partial_{\varepsilon_n} f_n(x) = \emptyset$ for all $x \notin \Omega_n$. Thus x_{n+1} belongs to Ω_n, i.e.,

$$f_n(x_{n+1}) = \langle F(x_{n+1}), z_n\rangle. \tag{41}$$

Combining (38)–(41), we get

$$-\varepsilon_n + \gamma_n \langle x_n - x_{n+1}, x^* - x_{n+1} \rangle \leq \langle F(x^*) - F(x_{n+1}), z_n \rangle \leq 0. \tag{42}$$

It follows from (42) that

$$0 \leq \varepsilon_n + \gamma_n \langle x_{n+1} - x_n, x^* - x_{n+1} \rangle = \varepsilon_n + \gamma_n \left(\|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 - \|x_n - x_{n+1}\|^2 \right). \tag{43}$$

Combining now (43) and (36), we get

$$0 \leq \gamma_n \left(\|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 - (1-\sigma) \|x_n - x_{n+1}\|^2 \right). \tag{44}$$

Since $0 \leq \sigma < 1$ and $\gamma_n > 0$, it follows from (44) that $\|x_{n+1} - x^*\| \leq \|x_n - x^*\|$, so that $\{\|x_n - x^*\|\}$ is nonincreasing and nonnegative, hence convergent, and also that

$$(1-\sigma) \|x_n - x_{n+1}\|^2 \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2. \tag{45}$$

Since the right hand side of (45) converges to 0 as $n \to \infty$ and $1 - \sigma > 0$, we conclude from (45) that $\lim_{n\to\infty}(x_n - x_{n+1}) = 0$. Note that we have obtained in this step exactly the same results as in Step 2 in the proof of Theorem 2.

Steps 3, 4 and 5: boundedness of $\{x_n\}$, optimality and uniqueness of its weak cluster point. The argument in Steps 3, 4 and 5 in the proof of Theorem 2 hold "verbatim" for this case. □

The results in Propositions 2 and 3 can be extended, "mutatis mutandis", to the sequences generated by **ALG2**. We refrain to do so for the sake of conciseness.

5 Final remarks

Remark 3. The proximal point method is indeed a conceptual scheme rather than an implementable algorithm, in the sense that it merely transforms a given problem into a sequence of better behaved subproblems. Thus, the performance of the method depends essentially on the algorithm used to solve the subproblems. In this situation, it makes little sense to compare the proximal method with other methods for finding efficient points (weakly or not) in terms of computational efficiency, unless a specific procedure is specified for solving the subproblems. In this paper, we refrain from discussing algorithms for the solution of the subproblems, and hence we skip discussions of implementation issues or comparisons with alternative approaches. Nevertheless, some hints in this direction are commented upon in Remark 5 below.

Remark 4. One could fix from the onset some $z \in C^+$, define $\psi : X \to \mathbb{R} \cup \{+\infty\}$ as $\psi(x) = \langle F(x), z \rangle$ and proceed to apply the proximal point method for scalar-valued functions

to the minimization of ψ. Under C-convexity of F and assumption (A), ψ has minimizers and, taking into account the well known convergence results for the scalar proximal point method (e.g. [13] for the exact method as in (1), [14] for the inexact method in (3)–(4)) the sequence generated by the proximal method would be weakly convergent to a minimizer of ψ, which would be weakly efficient for (VOP), in view of Theorem 1. This is the well known *scalarization* approach for solving (VOP). On the other hand, as can be observed in the proof of Theorem 2 and in the statement of Algorithm **ALG2**, our approach can be seen as choosing in each step a different scalarization through the selection of some $z_n \in C^+$. Thus, the novelty of the method in this paper with respect to the above mentioned approach lies in the additional flexibility, allowing a different scalarization in each iteration. As mentioned in Remark 3, this paper does not deal with the computational performance of the method, but we intend to study, in the future, good strategies for the choice of the scalarization vectors z_n, from the point of view of actual implementations. In view of Remark 3 above, such choice will very likely be linked to the method chosen for solving the subproblems.

Remark 5. We mentioned in Remark 4 that our method can be seen as a procedure which chooses at each iteration a certain scalarization and minimizes a regularized scalar-valued function. We emphasize now that this is just a way to visualize the procedure, but it does not entail, even in the case of the inexact version **ALG2**, that the subproblems should actually be treated as scalar minimization problems. In fact, it could be inconvenient to do so, as we discuss next. The presence of the term I_{Ω_n} in the definition of f_n demands that $F(x_{n+1}) \preceq_C F(x_n)$. This requirement could be difficult to comply with if we intend to solve the subproblem by minimizing $\bar{f}(x) = f_n(x) + (\gamma_n/2)\|x - x_n\|^2$. On the other hand, there exist methods for vector minimization which work directly with the order $\preceq_C$, and which are thus able to deal without trouble with constraints of the type $F(x) \preceq_C F(x_n)$. For instance, we mention the method introduced in [7] for finding weakly efficient points of differentiable maps $G : \mathbb{R}^p \to \mathbb{R}^m$ with respect to the partial order induced by a cone $C \subset \mathbb{R}^m$. This method, which extends the steepest descent method for scalar-valued minimization, generates a sequence $\{x_j\}$ which converges to a weakly efficient point for G, with the following decreasing property: $G(x_{j+1}) \preceq_C G(x_j)$ for all j. It is easy to check that if we are in the n-th subproblem of either **ALG1** or **ALG2**, and we apply a method with these properties to $G(x) = F(x) + (\alpha_n/2)\|x - x^n\|^2 e_n$, then the limit x^* of the generated sequence will satisfy $F(x^*) \preceq_C F(x_n)$, because of the decreasing property, and also (36) with $\varepsilon_n = 0$, because x^* is weakly efficient for G. Thus, such x^* can be taken as the next iterate x_{n+1} of our methods in both cases. In fact, for Algorithm **ALG2** in the finite dimensional and differentiable setting, the method in [7] has an even better feature: applied to the same G as above, after a finite number of iterations it will generate an iterate $\bar{y}$ with the property that $\bar{y} - \theta J_G(\bar{y})^t z_n$ (where $J_G(x)$ is the Jacobian matrix of G at x, and θ is some positive number) satisfies (35)-(36) for some $\varepsilon_n > 0$, and thus can be taken as the next iterate x_{n+1}. Since the method in [7] is an explicit gradient-type scheme, this approach provides a truly finite procedure for the subproblems in **ALG2**, in the smooth and finite dimensional case.

References

[1] Bolintinéanu, S. Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems. *Journal of Optimization Theory and Applcations* **106** (2000) 265-296.

[2] Bolintinéanu, S. Vector variational principles; $\varepsilon-$efficiency and scalar stationarity. *Journal of Convex Analysis* **8** (2001) 71-85.

[3] Burachik, R.S., Iusem, A.N., Svaiter, B.F. Enlargement of monotone operators with applications to variational inequalities *Set-Valued Analysis* **5** (1997) 159-180.

[4] Burachik, R.S., Scheimberg, S., Svaiter, B.F. Robustness of the hybrid extragradient proximal point algorithm. *Journal of Optimization Theory and Applications* **111** (2001) 117-136.

[5] Fliege, J., Svaiter, B.F. Steepest descent methods for multicriteria optimimization. *Mathematical Methods of Operations Research* **51** (2000) 479-494.

[6] Graña Drummond, L.M., Iusem, A.N. A projected gradient method for vector optimization problems (to be published in *Computational Optimization and Applications*).

[7] Graña Drummond, L.M., Svaiter, B.F. A steepest descent method for vector optimization (to be published).

[8] Luc, T.D. *Theory of Vector Optimization*. Lecture Notes in Economics and Mathematical Systems **319**, Springer (1989).

[9] Martinet, B. Régularisation d'inéquations variationelles par approximations succesives. *Revue Française de Informatique et Recherche Opérationelle* **2** (1970) 154-159.

[10] Martinet, B. *Algorithmes pour la résolution de problèmes d'optimisation et minimax*. Thèse d'état, Université de Grenoble (1972).

[11] Moreau, J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Societé Mathématique de France* **93** (1965) 273-299.

[12] Rockafellar, R.T. A dual approach to solving nonlinear programming problems by unconstrained optimization. *Mathematical Programming* **5** (1973) 354-373.

[13] Rockafellar, R.T. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization* **14** (1976) 877-898.

[14] Solodov, M.V., Svaiter, B.F. An inexact hybrid extragardient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis* **7** (1999) 323-345.